10

A simple introduction to the KLT (Karhunen–Loève Transform)

10.1 INTRODUCTION

This chapter is a simple introduction about using the Karhunen–Loève Transform (KLT) to extract weak signals from noise of any kind. In general, the noise may be colored and over wide bandwidths, and not just white and over narrow bandwidths. We show that the signal extraction can be achieved by the KLT more accurately than by the Fast Fourier Transform (FFT), especially if the signals buried into the noise are very weak, in which case the FFT fails. This superior performance of the KLT happens because the KLT of any stochastic process (both stationary and non-stationary) is defined from the start over a *finite* time span ranging between 0 and a final and *finite* instant T (contrary to the FFT, which is defined over an infinite time span). We then show mathematically that the series of all the eigenvalues of the autocorrelation of the (noise + signal) may be differentiated with respect to T yielding the "Final Variance" of the stochastic process $X(t)$ in terms of a sum of the first-order derivatives of the eigenvalues with respect to T. Finally, we prove that this new result leads to the immediate reconstruction of a signal buried into the thick noise. We have thus put on a strong mathematical foundation a set of very important practical formulae that can be applied to improve SETI, the detection of exoplanets, asteroidal radar, and also other fields of knowledge like economics, genetics, biomedicine, etc. to which the KLT can be equally well applied with success.

10.2 A BIT OF HISTORY

The Karhunen–Loève Transform (KLT) is the most advanced mathematical algorithm available in the year 2008 to achieve both noise filtering and data compression in processing signals of any kind.

It took about two centuries (~1800–2000) for mathematicians to create such a jewel of thought little by little, piece after piece, paper after paper. It is thus difficult to recognize who did what in building up the KLT and at the same time be fair in attributing each individual advance to the appropriate author. In addition, mathematicians, both pure and applied, often speak such a "clumsy" language of their own that even learned scientists sometimes find it hard to understand them. This unfortunate situation hides the esthetic beauty of many mathematical discoveries that were often historically made by their authors more for the joy of opening new lines of thought than for the sake of any immediate application to science and engineering.

In essence, the KLT is a rather new mathematical tool used to improve our understanding of physical phoenomena, far superior to the classical Fourier Transform (FT). The KLT is named for two mathematicians—the Finnish actuary Kari Karhunen (1915–1992) [1] and the French American Michel Loève (1907–1979) [2, 3]—who proved, independently and about the same time (1946), that the series (2) hereafter is *convergent*. Put this way, the KLT looks like a purely mathematical topic, but really this is hardly the case. As early as 1933 the American statistician and economist Harold Hotelling (1895–1973) used the KLT (for discrete time, rather than for continuous time), so that the KLT is sometimes called the "Hotelling Transform". Even much earlier than these three authors the Italian geometer Eugenio Beltrami (1835–1899) discovered as early as 1873 the SVD (Singular Value Decomposition), which is closely related to the KLT in that area of applied mathematics nowadays called Principal Components Analysis (PCA). Unfortunately, a complete historical account about how these contributions developed since 1865—when the English mathematician Arthur Cayley (1821–1895) "invented" matrices—simply does not exist. We only know about "fragments of thought" that impair an overall vision of both the PCA and the KLT.

In Sections 10.3–10.5, we'll derive *heuristically and step-by-step* the many equations that make up for the KLT. We think that this approach is much easier to understand for beginners than what is found in most "pure" mathematical textbooks, and hope that the readers will appreciate our effort to explain the KLT as easily as possible to non-mathematically trained people.

10.3 A HEURISTIC DERIVATION OF THE KL EXPANSION

We start by saying that the KLT was born during the years of World War Two out of the need to merge two different areas of classical mathematics.

(1) The expansion of a deterministic periodic signal $x(t)$ into a basis of orthonormal functions (sines and cosines, in this case), typified by the classical Fourier series—first put forward by the French mathematician Jean Baptiste Joseph Fourier (1768–1830) around 1807,

$$x(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty}[a_n \cos(nt) + b_n \sin(nt)] \qquad (-\pi \leq t \leq \pi). \tag{10.1}$$

(2) The need to extend this too narrow and deterministic view to probability and statistics. The much larger variety of phenomena called "noise" by physicists and engineers will thus be encompassed by the new transform. This enlarged view means considering a random function $X(t)$ (notice that we denote random quantities by capitals, and that $X(t)$ is also called a "stochastic process of the time"). We now seek to expand this stochastic process onto a set of orthonormal functions $\phi_n(t)$ according to the starting formula

$$X(t) = \sum_{n=1}^{\infty} Z_n \, \phi_n(t) \tag{10.2}$$

which is called the *Karhunen–Loève (KL) expansion of $X(t)$ over the finite time interval* $0 \le t \le T$.

What are then the Z_n and the $\phi_n(t)$ in (10.2)? To find out, let us start by recalling what "orthonormality" means for the Fourier series (10.1). Leonhard Euler (1707–1783) had already laid the first stone towards the Fourier series (10.1) by proving that, if $x(t)$ is assumed to be periodic over the time interval $-\pi \le t \le \pi$, then the coefficients a_n and b_n in (10.1) are obtained from the known function (or "periodic signal") $x(t)$ by virtue of the equations ("Euler formulae"):

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\cos(nt)\,dt \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} x(t)\sin(nt)\,dt. \tag{10.3}$$

If the same result is going to be true for the Karhunen–Loève expansion, the functions of the time, $\phi_n(t)$ in (10.2) must be orthornormal (i.e., both orthogonal and normalized to 1). That is,

$$\int_0^T \phi_m(t)\,\phi_n(t)\,dt = \delta_{mn} \tag{10.4}$$

where the δ_{mn} are the Kronecker symbols, defined by $\delta_{mn} = 0$ for $m \neq n$ and $\delta_{nn} = 1$.

But what then are the Z_n appearing in (10.2)? Well, a random function $X(t)$ can be thought of as something made up of two parts: its behavior in time, represented by the functions $\phi_n(t)$, and its behavior with respect to probability and statistics, which must therefore be represented by the Z_n. In other words, the Z_n must be random variables not changing in time (i.e., "just" random variables and not stochastic processes). By doing so we have actually made one basic, new step ahead: we have found that the KLT separates the probabilistic behavior of the random function $X(t)$ from its behavior in time, a kind of "untypical" separation that is achieved nowhere else in mathematics!

Having discovered that the Z_n are random variables, some trivial consequences follow at once. Let us denote by $E\{\ \}$ the linear operator yielding the average of a random variable or stochastic process. If one takes the average of both sides of the KL expansion (10.2), one then gets (we "freely" interchange here the average operator

$E\{\ \}$ with the infinite summation sign, bypassing the complaints of "subtle" mathematicians!)

$$E\{X(t)\} = \sum_{n=1}^{\infty} E\{Z_n\}\phi_n(t). \tag{10.5}$$

Now, it is not restrictive to suppose that the random function $X(t)$ has a zero mean value in time—namely, that the following equation is identically true for all values of the time t within the interval $0 \le t \le T$:

$$E\{X(t)\} \equiv 0. \tag{10.6}$$

In fact, were this not the case, one could replace $X(t)$ by the new random function $X(t) - E\{X(t)\}$ in all the above calculations, thus reverting to the case of a new random function with zero mean value. Thus, in conclusion, the random variables Z_n too must have a zero mean value

$$E\{Z_n\} \equiv 0. \tag{10.7}$$

This equation has a simple consequence: since the variance $\sigma^2_{Z_n}$ of the random variables Z_n is given by

$$\sigma^2_{Z_n} = E\{Z_n^2\} - E^2\{Z_n\} \tag{10.8}$$

by inserting (10.7) into (10.8) we get

$$\sigma^2_{Z_n} = E\{Z_n^2\}. \tag{10.9}$$

At this point, we can make a further step ahead, that has no counterpart in the classical Fourier series: we wish to introduce a new sequence of positive numbers λ_n such that every λ_n is the variance of the corresponding random variable Z_n, that is

$$\sigma^2_{Z_n} = \lambda_n = E\{Z_n^2\} > 0. \tag{10.10}$$

This equation provides the "answer" to the next "natural" question: Do the random variables Z_n fulfill a new type of "orthonormality" somehow similar to what the classical orthonormality (10.4) is for the $\phi_n(t)$? Since we are talking about random variables, the "orthogonality operator" can only be understood in the sense of *statistical independence*. The integral in (10.4) must then be replaced by the average operator $E\{\ \}$ for the random variables Z_n. In conclusion, we found that the random variables Z_n must obey the important equation

$$E\{Z_m Z_n\} = \lambda_n\, \delta_{mn}. \tag{10.11}$$

In this equation, we were forced to introduce the positive λ_n in the right-hand side in order to let (10.11) reduce to (10.10) in the special case $m = n$.

As for the KL equivalent of the Euler formulae (10.3) of the Fourier series, from the KL series (10.2) and the orthonormality (10.4) of the $\phi_n(t)$ one immediately finds that

$$Z_n = \int_0^T X(t)\, \phi_n(t)\, dt. \tag{10.12}$$

In other words: the random variables Z_n are obtained from the given stochastic process $X(t)$ by "projecting" this $X(t)$ over the correspoding eigenvector $\phi_n(t)$. If one likes the language of mathematicians and of quantum physics, then one may say that this projection of $X(t)$ onto $\phi_n(t)$ occurs in the "Hilbert space", which is the infinitely dimensional Euclidean space spanned by the eigenvectors $\phi_n(t)$ so that the square of $\phi_n(t)$ is integrable over the finite time span $0 \leq t \leq T$.

To sum up, we have actually achieved a remarkable generalization of the Fourier series by defining the Karhunen–Loève expansion (10.2) as the only possible statistical expansion in which all the expansion terms are *uncorrelated* from each other. This word "uncorrelated" comes from the fact that the *autocorrelation* of a random function of the time, $X(t)$, is defined as the mean value of the product of $X(t)$ at two different instants t_1 and t_2:

$$R_{XX}(t_1, t_2) \equiv R_X(t_1, t_2) = E\{X(t_1)X(t_2)\}. \tag{10.13}$$

If we assume, according to (10.6), that the mean value of $X(t)$ vanishes identically in the interval $0 \leq t \leq T$, the autocorrelation (10.13) reduces to the variance of $X(t)$ when the two instants are the same

$$\sigma^2_{X(t)} = E\{X^2(t)\} = E\{X(t)X(t)\} = R_X(t, t). \tag{10.14}$$

Let us add one final remark about the basic notion of statistical independence of the random viariables Z_n. It can be proven that, while the Z_n in (10.2) always are uncorrelated (by construction), they also are statistically independent if they are Gaussian-distributed random variables. This is fortunately the case for the Brownian motion and for the background noise we face in SETI. So we are not concerned about this subtle mathematical distinction between uncorrelated and statistically independent random variables.

10.4 THE KLT FINDS THE BEST BASIS (EIGEN-BASIS) IN THE HILBERT SPACE SPANNED BY THE EIGENFUNCTIONS OF THE AUTOCORRELATION OF *X(t)*

Up to this point, we have not given any hint about how to find the orthonormal functions of the time, $\phi_n(t)$, and positive numbers λ_n (i.e., the variances of the corresponding uncorrelated random variables Z_n). In this section, we solve this problem by showing that the $\phi_n(t)$ are the eigenfunctions of the autocorrelation $R_X(t_1, t_2) = E\{X(t_1)X(t_2)\}$ and that the λ_n are the corresponding eigenvalues. This is the correct mathematical phrasing of what we are going to prove. However, in order to ease the understanding of the further maths involved hereafter, a "translation" into the language of "common words" is now provided. Consider an object—for instance, a book—and a three-axes rectangular reference frame, oriented in an arbitrary fashion with respect to the book. Then, the classical Newtonian mechanics shows that all the mechanical properties of the book are described by a 3×3 symmetric matrix called the "inertia matrix" (or, more correctly, "inertia tensor") whose elements are, in general, all different from zero. Handling a matrix whose elements are all nonzero is obviously more complicated than handling a matrix where all entries are zeros except for those on the main diagonal (i.e., a *diagonal matrix*). Thus, one may be led to wonder whether a certain transformation of axes exists that changes the inertia matrix of the book into a diagonal matrix. Newtonian mechanics shows then that only *one* privileged orientation of the reference frame with respect to the book exists yielding a *diagonal* inertia matrix: the three axes must then coincide with a set of three axes (parallel to the book edges) called "principal axes" of the book, or "eigenvectors" or "proper vectors" of the inertia matrix of the book. In other words, each body posesses an intrinsic set of three rectangular axes that describes at best its dynamics (i.e., in the most concise form). This was proven again by Euler, and one can always compute the position of the eigenvectors with respect to a generic reference frame by means of a certain mathematical procedure called "finding the eigenvectors of a square matrix".

In a similar fashion, one can describe any stochastic process $X(t)$ by virtue of the statistical quantity called the autocorrelation (or simply the correlation), defined as the mean value of the product of the values of $X(t)$ at two different instants t_1 and t_2, and formally written $E\{X(t_1)X(t_2)\}$. The autocorrelation, obviously symmetric in t_1 and t_2, plays for the stochastic process $X(t)$ just the same role as the inertia matrix for the book example above. Thus, if one first seeks the eigenvectors of the correlation, and then changes the reference frame over to this new set of vectors, one achieves the simplest possible description of the whole (signal + noise) set.

Let us now translate the whole above description into equations. First of all, we must express the autocorrelation $E\{X(t_1)X(t_2)\}$ by virtue of the KL expansion (10.2). This goal is achieved by writing down (10.2) for two different instants, t_1 and t_2, taking the average of their product, and then (freely) interchanging the average and the summations in the right-hand side. The result is

$$E\{X(t_1)X(t_2)\} = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \phi_m(t_1)\phi_n(t_2)\, E\{Z_m Z_n\}. \tag{10.15}$$

Taking advantage of the statistical orthogonality of the Z_n, given by (10.11), (10.15) simplifies to

$$E\{X(t_1)X(t_2)\} = \sum_{m=1}^{\infty} \lambda_m \phi_m(t_1)\phi_m(t_2). \tag{10.16}$$

Finally, we now want to let the $\phi_n(t)$ "disappear" from the right-hand side of (10.16) by taking advantage of their orthonormality (10.4). To do so, we multiply both sides of (10.16) by $\phi_n(t_1)$ and then take the integral with respect to t_1 between 0 and T. One then gets:

$$\begin{aligned}\int_0^T E\{X(t_1)X(t_2)\}\phi_n(t_1)\,dt_1 &= \sum_{m=1}^{\infty} \lambda_m \phi_m(t_2) \int_0^T \phi_m(t_1)\phi_n(t_1)\,dt_1 \\ &= \sum_{m=1}^{\infty} \lambda_m \phi_m(t_2)\delta_{mn} = \lambda_n \phi_n(t_2),\end{aligned} \tag{10.17}$$

that is

$$\int_0^T E\{X(t_1)X(t_2)\}\,\phi_n(t_1)\,dt_1 = \lambda_n\,\phi_n(t_2). \tag{10.18}$$

This basic result is an *integral equation*, called by mathematicians "of Fredholm type". Once the correlation $E\{X(t_1)X(t_2)\}$ of $X(t)$ is known, the integral equation (10.18) yields (upon its solution, which may not be easy at all to find analytically!) both the Karhunen–Loève eigenvalues λ_n and the corresponding eigenfunctions $\phi_n(t)$. Readers familiar with quantum mechanics will also recognize in (10.18) a typical "eigenvalue equation" having the kernel $E\{X(t_1)X(t_2)\}$.

Let us finally summarize what we have proven so far in Sections 10.3 and 10.4, and let us use the language of signal processing, which will lead us directly to SETI, the main theme of this chapter.

By adding random noise to a deterministic signal one obtains what is called a "noisy signal" or, in case the signal power is much lower than the noise power, "a signal buried into the noise". The noise + signal is a random function of the time, denoted hereafter by $X(t)$. Karhunen and Loève proved that it is possible to represent $X(t)$ as the infinite series (called the KL expansion) given by (10.2), and this series is convergent. Assuming that the (signal + noise) correlation $E\{X(t_1)X(t_2)\}$ is a known function of t_1 and t_2, then the orthonormal functions $\phi_n(t)$ $(n = 1, 2, \ldots)$ turn out to be just the eigenfunctions of the correlation. These eigenfunctions $\phi_n(t)$ form an orthonormal basis in what physicists and mathematicians call the space of square-integrable functions, also called the Hilbert space. The eigenfunctions $\phi_n(t)$ actually are *the best* possible basis to describe the (signal + noise), much better than any classical Fourier basis made up by sines and cosines only. One can conclude that *the KLT automatically adapts itself to the shape of the (signal + noise), whatever behavior in time it may have, by adopting as a new reference frame in the Hilbert space the basis spanned by the eigenfunctions, $\phi_n(t)$, of the autocorrelation of the (signal + noise), $X(t)$.*

10.5 CONTINUOUS TIME VS. DISCRETE TIME IN THE KLT

The KL expansion in continuous time, *t*, is what we have described so far. This may be more "palatable" to theoretical physicists and mathematicians inasmuch as it may be related to other branches of physics, or of science in general, in which time obviously must be a continuous variable. For instance, this author spent 15 years of his life (1980–1994) in investigating mathematically the connection between Special Relativity and KLT. The result was the mathematical theory of optimal telecommunications between the Earth and a relativistic spaceship either receding from the Earth or approaching it. Although this may sound like "mathematical science fiction" to some folks (who we would call "short sighted"), the possibility that, in the future, humankind will send out relativistic automatic probes or even manned spaceships, is not unrealistic. Nor is it science fiction to imagine that an *alien* spaceship might approach the Earth slowing down from relativistic speeds to zero speed. So, a mathematical physics book like [4] can make sense. There, the KLT is obtained for any acceleration profile of the relativistic probe or spaceship. The result is that the KL eigenfunctions are Bessel functions of the first kind (suitably modified) and the eigenvalues are determined by the zeros of linear combinations of these Bessel functions and their derivatives, as we shall prove in Appendices F through K of this book, and especially in Appendix G.

Other continuous-time applications of the KLT are to be found in other branches of science, ranging, for instance, from genetics to economics. But, whatever the application may be, if time is a continuous variable, then one must solve the integral equation (10.18), and this may require considerable mathematical skills. In fact, (10.18) is, in general, an integral equation of the Fredholm type, and the usual "iterated nuclei" procedure used to solve Fredholm integral equations may be particularly painful to achieve. The task may be much easier if one is able to reduce the Fredholm integral equation to a Volterra integral equation, in just the way shown in the book [4] for the time-rescaled Brownian motion in relation to Special Relativity.

But let us go back to the time variable t in the KL expansion (10.2). If this variable is *discrete*, rather than continuous, then the picture changes completely. In fact, the integral equation (10.2) now becomes ... a system of simultaneous algebraic equations of the first degree, that can *always* be solved! The difficulty here is that this system of linear equations is *huge*, because the autocorrelation matrix is huge (hundreds or thousands of elements are the rule for autocorrelation matrices in SETI and in other applications, like image processing and the like). Also huge are the eigenvalues of the characteristic equation (i.e., the algebraic equations whose roots are the KL). Can you imagine solving *directly* an algebraic equation of degree 10,000?

So, the KLT is practically impossible to find numerically, unless we resort to *simplifying tricks* of some kind. This is precisely what was done for the SETI-Italia program by this author and his students, strongly supported by Ing. Stelio Montebugnoli and his team [5].

10.6 THE KLT: JUST A LINEAR TRANSFORMATION IN THE HILBERT SPACE

Although we have explained the KL expansion (10.2), we have yet to explain what the *KLT* is! We do so in this section.

The next step towards the KLT proper is the rearrangement of the eigenvalues λ_n in decreasing order of magnitude. Suppose we have done this. Consequently, we also rearrange the eigenfunctions $\phi_n(t)$ so that each eigenfunction keeps corresponding to its own eigenvalue. It can be proved that no mismatch can possibly arise in doing so, inasmuch as each eigenfunction corresponds to one eigenvalue only—namely, it can be proved that there is no degeneracy (contrary to what happens in quantum physics, where, for instance, there is a lot of degeneracy in the eigenfunctions of even the simplest atom of all, the hydrogen atom!). Furthermore, all eigenvalues are positive, and so, once rearranged in decreasing order of magnitude, they form a decreasing sequence where the first eigenvalue is the largest, and is called the "dominant" eigenvalue by mathematicians.

We are now ready to compute the *Direct KLT* of the (signal + noise). Let us use the new set of eigen-axes to describe the (signal + noise). Then, in the new representation, the (signal + noise) is just the Direct KLT of the old (signal + noise). In other words, the KLT is properly called just a *linear* trasformation of axes, and nothing is easier than that! (Incidentally, this accounts for the title of Karhunen's first paper "Über Lineare Methoden in der Wahrscheinlichkeitsrechnung" = "On linear methods in the calculus of probabilities", [1], which obviously refers to the linear character of the transformation of axes in the Hilbert space.)

10.7 A BREAKTHROUGH ABOUT THE KLT: MACCONE'S "FINAL VARIANCE" THEOREM

The importance of the KLT as a mathematical tool superior to the FFT has already been pointed out. However, the implementation of the KLT by a numerical code running on computers has always been a difficult problem. Both François Biraud in France [6] and Bob Dixon in the USA [16] failed to do so in the 1980s because all computers then available could not make the N^2 calculations required to solve the huge system of simultaneous algebraic equations of the first degree corresponding (in the discrete case) to the integral equation (10.18). At the SETI-Italia facilities at Medicina we faced the same problem, of course. But we did better than our predecessors because we discovered the new theorem about the KLT that we demonstrate in this section and call "the Final Variance theorem". This new theorem seems to be even more important than the rest of research work about the KLT since it solves directly the problem of extracting a weak sinusoidal carrier (a tone) from noise of whatever kind (both colored and white).

The key idea of the Final Variance theorem is to differentiate the first eigenvalue (briefly called the "dominant eigenvalue") of the autocorrelation of the (noise + signal) with respect to the final instant T of the general KLT theory. Remember here that this final instant T simply does not exist in the ordinary Fourier theory, because this T equals infinity according to the Fourier theory. Therefore, the final instant T in itself is possibly the most important "novelty" introduced by the KLT regarding the classical FFT. With respect to T, we may take derivatives (called "final derivatives" in the remainder of this book because they are time derivatives taken with respect to the final instant T) and integrals that have no analogs in the ordinary Fourier theory. The "error" that was made in the past—even by many KLT scholars—was to set $T = 1$, thus obscuring the fundamental novelty represented by the finite, real positive T as a new continuous variable playing in the game! This error made by other scholars clearly appears, for instance, in the Wikipedia site about the "Karhunen–Loève Theorem", *http://en.wikipedia.org/wiki/Karhunen-Loève_* *theorem*. So, by removing this silly $T = 1$ convention we opened up new prospects for KLT theory, as we now show by proving our "Final Variance theorem".

Consider the eigenfunction expansion of the autocorrelation again—Equation (10.16)—with the traditional dummy index n rewritten instead of m. Upon replacing $t_1 = t_2 = t$, this equation becomes

$$E\{X^2(t)\} = \sum_{n=1}^{\infty} \lambda_n \phi_n^2(t). \tag{10.19}$$

Since the eigenfunctions $\phi_n(t)$ are normalized to 1, we are prompted to integrate both sides of (10.19) with respect to t between 0 and T, so that the integral of the square of the $\phi_n(t)$ becomes just 1:

$$\int_0^T E\{X^2(t)\}dt = \sum_{n=1}^{\infty} \lambda_n \int_0^T \phi_n^2(t)\, dt = \sum_{n=1}^{\infty} \lambda_n. \tag{10.20}$$

On the other hand, since the mean value of $X(t)$ is identically equal to 0, one may now introduce the *variance* $\sigma^2_{X(t)}$ of the stochastic process $X(t)$ defined by

$$\sigma^2_{X(t)} = E\{X^2(t)\} - E^2\{X(t)\} = E\{X^2(t)\}. \tag{10.21}$$

Replacing (10.21) into (10.20), one gets

$$\int_0^T \sigma^2_{X(t)}\, dt = \sum_{n=1}^{\infty} \lambda_n. \tag{10.22}$$

This formula was first given by this author in his 1994 book [4, eq. (1.13), p. 12]. At that time, however, (10.22) was regarded as interesting inasmuch as (upon interchanging the two sides) it proves that the series of all the eigenvalues λ_n is indeed convergent (as one would intuitively expect) and its sum is given by the integral of the variance between 0 and T.

Back in 1994, however, the author did not understand that (10.22) had a more profound meaning: since the final instant T is the upper limit of the time integral on the left-hand side, the right-hand side also must depend on T. In other words, all the eigenvalues λ_n must be some functions of the final instant T:

$$\lambda_n \equiv \lambda_n(T). \tag{10.23}$$

This new remark is vital in order to make further progress. In fact, one is now prompted to let the integral on the left-hand side of (10.22) disappear by differentiating both sides with respect to the final instant T. One thus gets:

$$\sigma^2_{X(T)} = \sum_{n=1}^{\infty} \frac{\partial \lambda_n(T)}{\partial T}. \tag{10.24}$$

This result we call the *Final Variance theorem*. It was discovered by this author in May 2007 and is the key new result put forward in this chapter. It states that for any (either non-stationary or stationary) stochastic process $X(t)$, *the Final Variance* $\sigma^2_{X(T)}$

is the sum of the series of the first-order partial derivatives of the eigenvalues $\lambda_n(T)$ with respect to the final instant T.

Let us now consider a few particular cases of this theorem that are especially interesting.

(1) In general, only the first N terms of the decreasing sequence of eigenvalues will be retained as "significant" by the user, and all the other terms, from the $(N+1)$th term onward, will be declared to be "just noise". Therefore, the infinite series in (10.24) becomes in practice the finite sum

$$\sigma^2_{X(T)} \approx \sum_{n=1}^{N} \frac{\partial \lambda_n(T)}{\partial T}. \tag{10.25}$$

In numerical simulations, however, one always wants to make computation time as short as possible! Therefore, one might be led to consider the *first* (or *dominant*) eigenvalue only in (10.25); that is

$$\sigma^2_{X(T)} \approx \frac{\partial \lambda_1(T)}{\partial T}. \tag{10.26}$$

This clearly is "the roughest possible" approximation to the full $X(t)$ process since we are actually replacing the full $X(t)$ by its first KLT term $Z_1\phi_1(t)$ only. However, using (10.26) instead of the N-term sum (10.25) is indeed a good shortcut for application of the KLT to the extraction of very weak signals from noise, as we now stress in the very important practical case of stationary processes.

(2) If we restrict our considerations to *stationary* stochastic processes only (i.e., processes for which both the mean value and the variance are constant in time), then (10.25) simplifies even further. In fact, by definition, the stationary processes have the *same final variance at any time (i.e., for stationary processes σ^2_X is a constant)*. Then (10.22) immediately shows that, for stationary processes only, all the KLT eigenvalues are *linear* functions of the final instant T:

$$\lambda_n(T) \propto T \quad \textit{for stationary processes only.} \tag{10.27}$$

As a consequence, the first-order partial derivatives of all the λ_n with respect to T for stationary processes are just constants. In yet other words, for stationary processes only, (10.25) becomes

$$\sum_{n=1}^{N} \frac{\partial \lambda_n(T)}{\partial T} \approx \text{a constant with respect to } T. \tag{10.28}$$

In particular, if one sticks again to the first, dominant eigenvalue only (i.e., to the roughest possible approximation), then (10.28) reduces to

$$\frac{\partial \lambda_1(T)}{\partial T} \approx \text{a constant with respect to } T. \tag{10.29}$$

In Section 10.8 we will discuss the deep, practical implications of this result for SETI, extrasolar planet detection, asteroidal radar, and other KLT applications.

(3) Please notice that, for non-stationary processes, the dependence of the eigenvalues on T certainly is non-linear. For instance, for the well-known Brownian motion (i.e., "the easiest of the non-stationary processes"), one has

$$\lambda_n(T) = \frac{4T^2}{\pi^2(2n-1)^2} \quad (n = 1, 2, \dots) \tag{10.30}$$

and so the dependence on T is quadratic. For the proof, just place the Brownian motion variance $\sigma^2_{B(t)} = t$ into (10.22) and perform the integration, yielding the T^2 directly. Of course, this is in agreement with (10.30), which will be proven in Appendix F when we search for the KLT of the standard Brownian motion—see, in particular, (F.21).

(4) Even higher than quadratic is the dependence on T for the eigenvalues of other highly non-stationary processes. For instance, for the zero-mean square of the Brownian motion, the KLT eigenvalues depend cubically on the final instant T, as will be proven in Appendix I by Equation (I.60). And so on for more complicated processes, like the time-rescaled squared Brownian motions whose KLT will found in Appendix I.

10.8 BAM ("BORDERED AUTOCORRELATION METHOD") TO FIND THE NUMERIC KLT OF *STATIONARY* PROCESSES ONLY

The BAM (an acronym for "Bordered Autocorrelation Method") is an alternative numerical technique to evaluate the KLT of stationary processes (only) that may run faster on computers than the traditional full-solving KLT technique described in Section 10.5. The BAM has its mathematical foundation in our Final Variance theorem already proved in Section 10.7. In this section we describe the BAM in detail and provide the results of numerical simulations showing that, by virtue of the BAM, the KLT succeeds in extracting a sinusoidal carrier embedded in a lot of noise when the FFT utterly fails.

Let us start by recalling that the standard, traditional technique to find the KLT of any stochastic process (whether stationary or not) numerically amounts to solving N simultaneous linear algebraic equations whose coefficient matrix is the (huge) autocorrelation matrix. This N^2 amount of calculations is much larger than the $N * \ln(N)$ amount of calculations required by the FFT and that's precisely the reason the FFT has been preferred to the KLT in the last 50 years!

Because of the Final Variance theorem proved in the previous section, however, one is tempted to confine oneself to the study of the dominant eigenvalue, only by virtue of just using (10.29). This means studying (10.29) for different values of the final instant T (i.e., as a function of the final instant T).

Also, we now confine ourselves to a *stationary* $X(t)$ over a *discrete* set of instants $t = 0, \dots, N$.

In this case, the autocorrelation of $X(t)$ becomes the Toeplitz matrix (for an introduction to the research field of Toeplitz matrices, see the Wikipedia site, *http://en.wikipedia.org/wiki/Toeplitz_matrix*) which we denote by $R_{Toeplitz}$.

$$R_{Toeplitz} = \begin{bmatrix} R_{XX}(0) & R_{XX}(1) & R_{XX}(2) & \cdots & \cdots & R_{XX}(N) \\ R_{XX}(1) & R_{XX}(0) & R_{XX}(1) & \cdots & \cdots & R_{XX}(N-1) \\ R_{XX}(2) & R_{XX}(1) & R_{XX}(0) & \cdots & \cdots & R_{XX}(N-2) \\ \cdots & \cdots & \cdots & R_{XX}(0) & \cdots & \cdots \\ R_{XX}(N) & R_{XX}(N-1) & \cdots & \cdots & R_{XX}(1) & R_{XX}(0) \end{bmatrix}. \tag{10.31}$$

This theorem had already been proven by Bob Dixon and Mike Kline back in 1991 [16], and will not be proven here again. We may choose N at will, but clearly the higher the N, the more accurate the KLT of $X(t)$. On the other hand, the final instant T in the KLT can be chosen at will and now is $T = N$. So, we can regard $T = N$ as a sort of "new time variable" and even take derivatives with respect to it, as we'll do in a moment.

But let us now go back to the Toeplitz autocorrelation (10.31). If we let N vary as a new free variable, that amounts to *bordering it* (i.e., adding one (last) column and one (last) row to the previous correlation). This means solving yet again the system of linear algebraic equations of the KLT for $N + 1$, rather than for N. So, *for each different value of N, we get a new value of the first eigenvalue* λ_1 *now regarded as a function of N (i.e.,* $\lambda_1(N)$*). Doing this over and over again, for as many values as we wish (or, more correctly, for how many values of N our computer can still handle!) constitutes our BAM, the Bordered Autocorrelation Method.*

But then we know from the Final Variance theorem that $\lambda_1(N)$ *is proportional to N. And such a function* $\lambda_1(N)$ *of course has a derivative,* $\partial\lambda_1(N)/\partial N$*, that can be computed numerically as a new function of N. And this derivative turns out to be a constant with respect to N. This fact paves the way for a new set of applications of the KLT to all fields of science!*

In fact, numeric simulations lead to the results shown in the four plots in Figures 10.1–10.4. The first plot is the ordinary Fourier spectrum of a pure tone at 300 Hz buried in noise with a signal-to-noise ratio of 0.5, abbreviated hereafter as SNR = 0.5. For a definition of the SNR see the Wikipedia site, *http://en.wikipedia.org/wiki/Signal-to-noise_ratio* Please note the following two facts:

(1) This is about as low an SNR can be before the FFT starts failing to denoise a signal, as is well known by electrical and electronic engineers.
(2) This Fourier spectrum is obviously computed by taking the Fourier Transform of the *stationary* autocorrelation of $X(t)$, as is well known from the Wiener–Khinchin theorem (for a concise description of this theorem, see *http://en.wikipedia.org/wiki/Wiener-Khinchin_theorem*).

Notice, however, that this procedure would *not* work for non-stationary $X(t)$ because the Wiener–Khinchin theorem does not apply to non-stationary processes. For

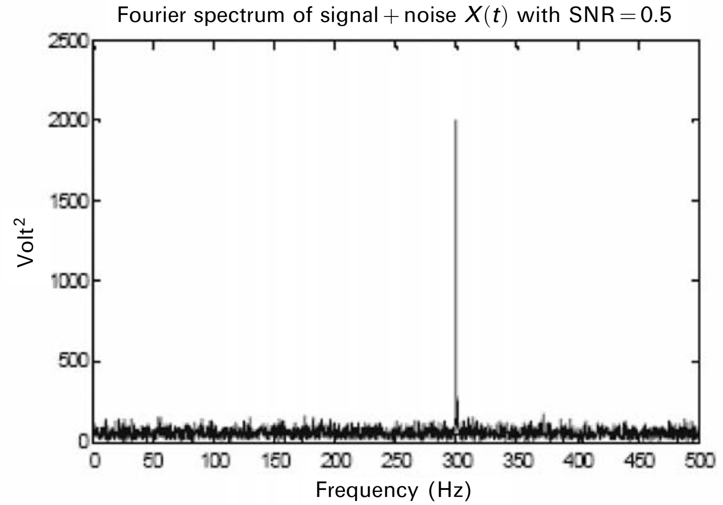

Figure 10.1. Fourier spectrum of a pure tone (i.e., just a sinusoidal carrier) with frequency at 300 Hz buried in stationary noise with a signal-to-noise ratio of 0.5.

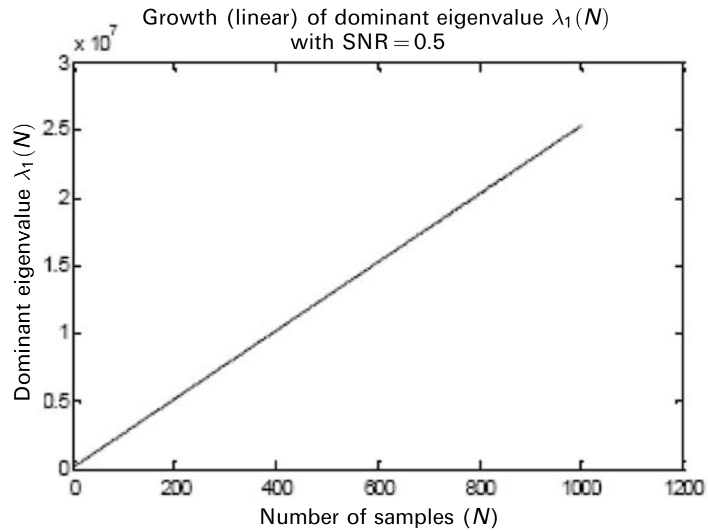

Figure 10.2. The KLT dominant eigenvalue $\lambda_1(N)$ over $N = 1{,}200$ time samples, computed by virtue of the BAM, the Bordered Autocorrelation Method.

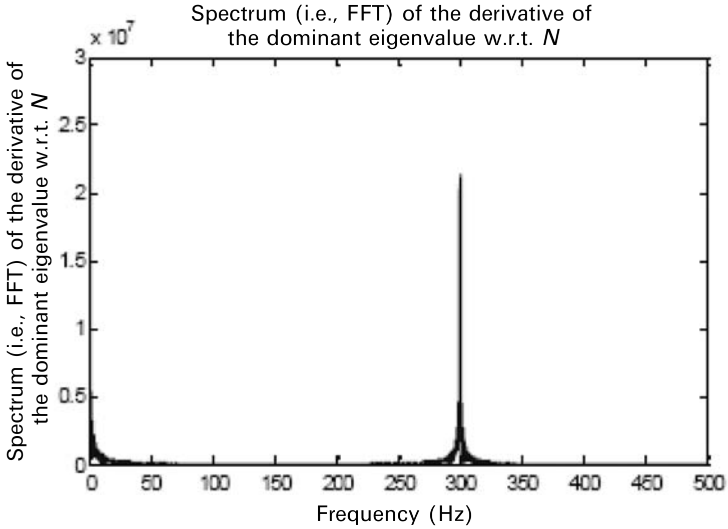

Figure 10.3. The spectrum (i.e., the Fourier Transform) of the *constant* derivative of the KLT dominant eigenvalue $\lambda_1(N)$ with respect to N as given by the BAM. This is clearly a Dirac delta function (i.e., a peak, at 300 Hz), as expected.

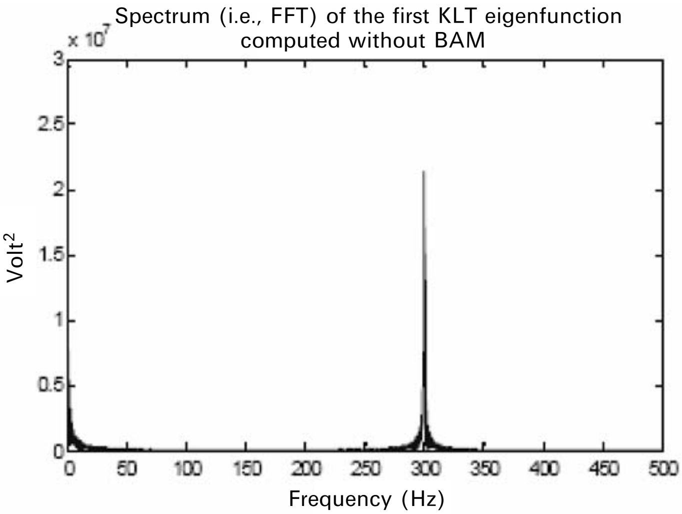

Figure 10.4. The spectrum (i.e., the Fourier Transform) of the first KLT eigenfunction *not* obtained by the BAM, but rather by the very long procedure of solving N linear algebraic equations corresponding, in discrete time, to the integral equation (10.18). Clearly, the result is the same as obtained in Figure 10.3 by the much less time-consuming BAM. So, one can say that adoption of the BAM actually made the KLT "feasible" on small computers by circumventing the difficulty of the N^2 calculations requested by the "straight" KLT theory.

non-stationary processes there are other "tricks" to compute the spectrum from the autocorrelation, like the Wigner–Ville Transform, but we shall not consider them here.

The second plot (Figure 10.2) shows the first (i.e., the dominant) KLT eigenvalue $\lambda_1(N)$ over $N = 1{,}200$ time samples. Clearly, this $\lambda_1(N)$ is proportional to N, as predicted by our Final Variance theorem (10.27).

So, its derivative, $\partial\lambda_1(N)/\partial N$, is a constant with respect to N. But we may then take the Fourier Transform of such a constant and get a Dirac delta function (i.e., a peak just at 300 Hz). In other words, we have KLT-reconstructed the original tone by virtue of the BAM. The third plot (Figure 10.3) shows such a BAM-reconstructed peak.

Finally, this plot is of course identical to the fourth plot (Figure 10.4), showing the ordinary FFT of the first KLT eigenfuction as obtained, not by the BAM, but by solving the full and long system of N algebraic first-degree equations.

Let us now do the same again … but with an incredibly low SNR of 0.005.

Poor Fourier here is in a mess! Just look at the plot in Figure 10.5! No classical FFT spectrum can be identified at all for such a terribly low SNR!

But for the KLT no problem!

The next plot (Figure 10.6) shows that $\lambda_1(N) \propto N$, as predicted by our Final Variance theorem (10.27).

The third plot (Figure 10.7, *KLT fast* way via the BAM) is the *neat KLT spectrum* of the 300 Hz tone obtained by computing the FFT of the *constant* $\partial\lambda_1(N)/\partial N$.

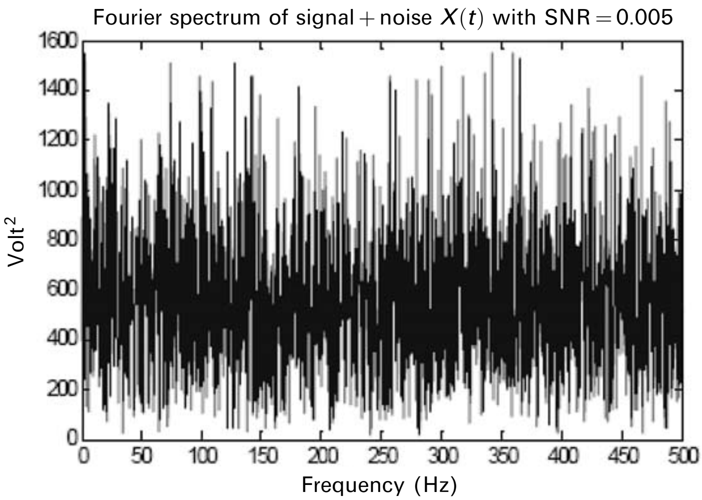

Figure 10.5. Fourier spectrum of a pure tone (i.e., just a sinusoidal carrier) with frequency at 300 Hz buried in stationary noise with the terribly low signal-to-noise ratio of 0.005. This is clearly beyond the reach of the FFT, since we know there should just be one peak only at 300 Hz. Fourier *fails* at such a low SNR.

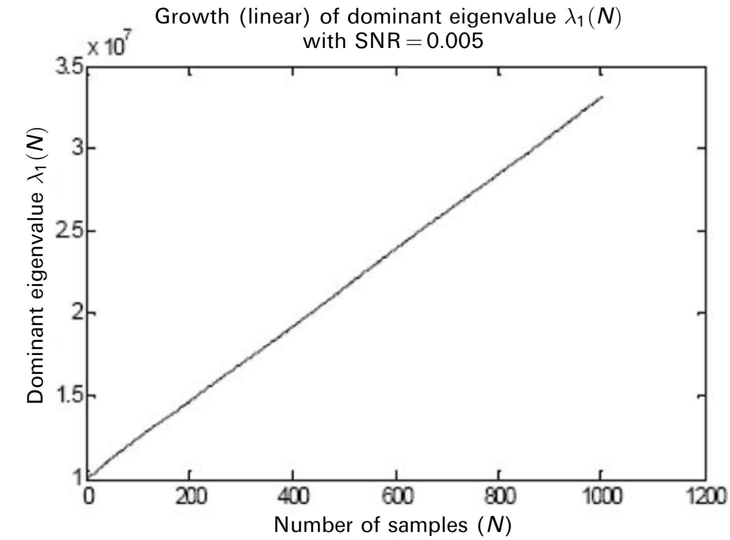

Figure 10.6. The KLT dominant eigenvalue $\lambda_1(N)$ for $N = 1{,}200$ time samples, computed by virtue of the BAM, for the very low SNR = 0.005.

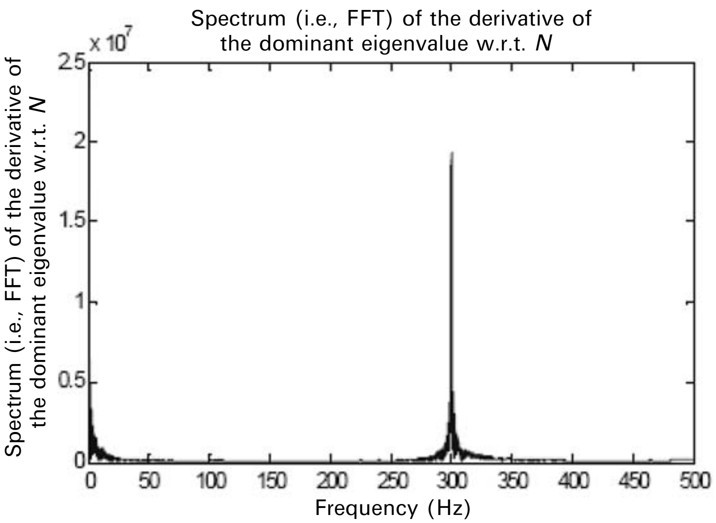

Figure 10.7. The spectrum (i.e., the Fourier Transform) of the *constant* derivative of the KLT dominant eigenvalue $\lambda_1(N)$ with respect to N as given by the BAM. This is a neat Dirac delta function (i.e., it has a peak at 300 Hz, as expected).

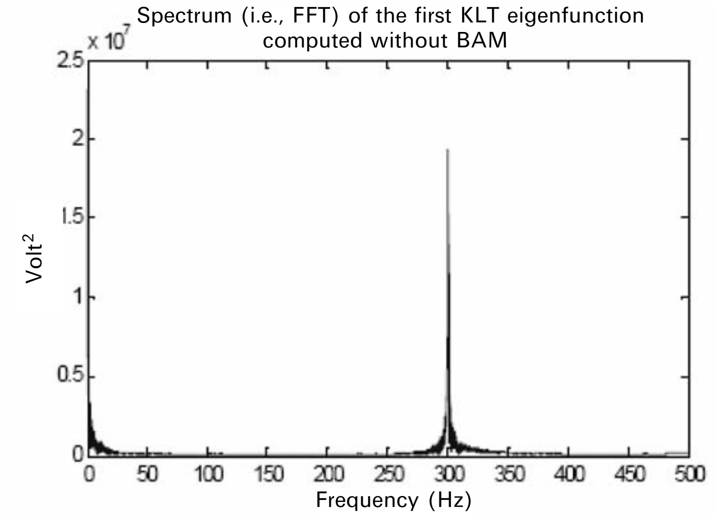

Figure 10.8. The spectrum (i.e., the Fourier Transform) of the first KLT eigenfunction, *not* obtained by the BAM but rather by the very long procedure of solving N linear algebraic equations corresponding, in discrete time, to the integral equation (10.18). Clearly, the result is the same as obtained in Figure 10.7, but this time by the much less time-consuming BAM. So, one can say that the adoption of the BAM actually made the KLT "feasible" on small computers by circumventing the difficulty of N^2 calculations requested by the "straight" KLT theory.

And this is just the same as the last plot (Figure 10.4) of the dominant KLT eigenfunction obtained by the KLT *slow* way of doing N^2 calculations. This proves the superior behavior of the KLT.

10.9 DEVELOPMENTS IN 2007 AND 2008

The numerical simulations described in the previous section were performed at Medicina during the winter 2006–2007 by Francesco Schillirò and Salvatore "Salvo" Pluchino [22]. These simulations suggested in a purely numerical fashion (i.e., without any analytic proof) that the BAM leads to the following result for stationary processes: the ordinary Fourier transform (i.e., "the spectrum" in the common sense, since the processes are supposed to be stationary) of the first-order partial derivative with respect to the final instant T of the dominant eigenvalue, $\dfrac{\partial \lambda_1(T)}{\partial T}$, is just the frequency of the feeble sinusoidal carrier buried in the mountain of noise. In SETI language, if we are looking for a simple sinusoidal carrier sent by ET and buried in a lot of cosmic noise, then the frequency we are looking for is given by the FFT of $\dfrac{\partial\lambda_1(T)}{\partial T}$.

Why?

No analytic proof of this numerical result was ever found at Medicina. But this author had made the first step towards the then missing analytic proof by proving the Final Variance Theorem in May 2007, and persisted in discussing this "frontier result" with other radioastronomers. One year later, in June 2008, he went to Dwingeloo, the Netherlands, and met with the ASTRON Team working on a possible implementation of SETI on the brand-new LOFAR radiotelescope. Dr. Sarod Yatawatta of ASTRON then made the next step toward the missing analytic proof: he derived an unknown analytic expression for the KLT eigenvalues of the ET sinusoidal carrier [24]. Unfortunately, Dr. Yatawatta made two analytical errors in his derivation (described hereafter), which this author discovered and corrected in September 2008.

In conclusion, the final, correct version of all these equations is explained in the next two sections, and it proves that the Fourier Transform of the first derivative of the KLT eigenvalues with respect to the final instant T is indeed the frequency of the "unknown" ET signal, but only for stationary processes, of course.

For non-stationary processes (i.e., for transient phoenomena as actually happens in practical SETI, since all celestial bodies move, rather than rest), the story is much more complicated, and this author is convinced that a much more refined mathematical investigation has to be made: but this will be our next step, not described in this book yet!

10.10 KLT OF STATIONARY WHITE NOISE

Before we give the analytic proof that the Fourier Transform of $\dfrac{\partial\lambda_1(T)}{\partial T}$ is the frequency of the unknown ET signal, we must understand what the KLT of stationary white noise is.

Stationary white noise is defined as the one "limit" stochastic process that is completely uncorrelated (i.e., the autocorrelation of which is the Dirac delta function). In other words, denoting the stationary white noise by $W(t)$, one has by definition

$$E\{W(t_1)W(t_2)\} = \delta(t_1 - t_2). \tag{10.32}$$

If one now seeks the KLT of stationary white noise, one must of course insert the autocorrelation (10.32) into the KLT integral equation (10.18), getting

$$\lambda_n\phi_n(t_2) = \int_0^T E\{W(t_1)W(t_2)\}\phi_n(t_1)\,dt_1 = \int_0^T \delta(t_1 - t_2)\phi_n(t_1)\,dt_1 = \phi_n(t_2). \tag{10.33}$$

This proves that:

(1) The KLT eigenvalues of stationary white noise are all equal to 1.
(2) *Any* set of orthonormal eigenfunctions $\phi_n(t)$ in the Hilbert space is a suitable basis to represent stationary white noise.

Since *any* set of orthonormal eigenfunctions $\phi_n(t)$ in the Hilbert space is a suitable basis to represent stationary white noise, from now one we shall adopt the easiest possible such basis; that is, the simple Fourier basis made up only by orthonormalized sines over the finite interval $0 \leq t \leq T$:

$$\phi_n(t) = \sqrt{\frac{2}{T}} \sin\left(\frac{2\pi n}{T} t\right) \equiv W_n(t). \tag{10.34}$$

Of course, this set of basis functions fulfills the orthonormality condition

$$\int_0^T W_m(t) W_n(t)\, dt = \int_0^T \sqrt{\frac{2}{T}} \sin\left(\frac{2\pi m}{T} t\right) \cdot \sqrt{\frac{2}{T}} \sin\left(\frac{2\pi n}{T} t\right) dt = \delta_{mn}. \tag{10.35}$$

This property will be used in the next section, where we give the proof that the Fourier Transform of $\dfrac{\partial \lambda_n(T)}{\partial T}$ is indeed (twice) the frequency of the unknown ET sinusoidal carrier buried in white, cosmic noise. We conclude this section by pointing out the first analytical error made by Dr. Yatawatta in his personal communication to this author [24]: he forgot to put the square root in (10.34). This of course means that his further results were flawed, even more so since he made a second analytical error later, which we shall not describe. But the key ideas behind his proof were perfectly correct, and we shall describe them in the next section.

10.11 KLT OF AN ET SINUSOIDAL CARRIER BURIED IN WHITE, COSMIC NOISE

Consider a new stochastic process $S(t)$ made up by the sum of stationary white noise $W(t)$ plus an alien ET sinusoidal carrier of amplitude a and frequency $\nu = \dfrac{\omega}{2\pi}$; that is,

$$S(t) = W(t) + a \sin(\omega t). \tag{10.36}$$

What is the KLT of such a (signal + noise) process? This is the central problem of SETI, of course.

To find the answer, first build up the autocorrelation of this process:

$$\begin{aligned} E\{S(t_1)S(t_2)\} = E\{W(t_1)W(t_2)\} + a^2 \sin(\omega t_1)\sin(\omega t_2) \\ + aE\{W(t_1)\sin(\omega t_2)\} + aE\{W(t_2)\sin(\omega t_1)\}. \end{aligned} \tag{10.37}$$

The last two terms in (10.37) represent the two cross-correlations between the white noise and the sinusoidal signal. It is reasonable to assume that the white noise and the signal are uncorrelated, and so we shall simply replace these two cross-correlations by zero. The autocorrelation (10.37) of the (signal + noise) stochastic process $S(t)$ thus becomes

$$E\{S(t_1)S(t_2)\} = E\{W(t_1)W(t_2)\} + a^2 \sin(\omega t_1)\sin(\omega t_2). \tag{10.38}$$

In order to proceed, we now make use of the eigenfunction expansion of the autocorrelation (10.16), which, replaced into (10.38), changes it into

$$\sum_{m=1}^{\infty} \lambda_{S_m} S_m(t_1) S_m(t_2) = \sum_{m=1}^{\infty} \lambda_{W_m} W_m(t_1) W_m(t_2) + a^2 \sin(\omega t_1)\sin(\omega t_2). \tag{10.39}$$

In the last equation, the $S_m(t)$ clearly are the (unknown) eigenfunctions of the (signal + noise) process $S(t)$, and the λ_{S_m} are (unknown) corresponding eigenvalues. In the right-hand side, the λ_{W_m} are the eigenvalues of the stationary white noise, which we know to be equal to 1, but, for the sake of clarity, let us keep the symbol λ_{W_m} rather than replacing it by 1.

To proceed further, we now must get rid of both t_1 and t_2 in (10.39), and there is only one way to do so: use the orthonormality of the eigenfuctions appearing in (10.39). We shall do so in a moment. Before, however, let us make the following practical consideration: since the signal is much waker than the noise (by assumption) (i.e., the signal-to-noise ratio is much smaller than 1, or SNR $\ll$ 1), then, numerically speaking, the (signal + noise) eigenfunctions $S_m(t)$ must not differ very much from the pure white noise eigenfunctions $W_m(t)$. And, similarly, the (signal + noise) eigenvalues λ_{S_m} must not differ very much from the corresponding pure white noise eigenvalues λ_{W_m}. In other words, the hypothesis that SNR $\ll$ 1 amounts to the two approximate equations

$$\left.\begin{aligned} S_m(t) &\approx W_m(t) \\ \lambda_{S_m} &\approx \lambda_{W_m} = 1. \end{aligned}\right\} \tag{10.40}$$

Of course, only the first of these two equations will play a role in the two integrations that we are now going to perform: once with respect to t_1 and once with respect to t_2, and both over the interval $0 \le t \le T$. As a consequence, the new orthonormality condition (nearly) holds:

$$\int_0^T S_m(t_1) W_n(t_1)\, dt_1 \approx \delta_{mn} \tag{10.41}$$

and, similarly,

$$\int_0^T S_k(t_2) W_n(t_2)\, dt_2 \approx \delta_{kn} \tag{10.42}$$

So, let us now multiply both sides of (10.39) by $W_n(t_1)$ and integrate with respect to t_1 between 0 and T. Because of (10.41) and (10.35) one has:

$$\sum_{n=1}^{\infty} \lambda_{S_n} S_n(t_2) \approx \sum_{n=1}^{\infty} \lambda_{W_n} W_n(t_2) + a^2 \sin(\omega t_2) \int_0^T W_n(t_1)\sin(\omega t_1)\, dt_1 \tag{10.43}$$

The good point is that the integral appearing in the right-hand side of this equation can be found. In fact, replacing $W_n(t_1)$ by virtue of (10.34) and integrating, one gets

$$\sum_{k=1}^{\infty} \lambda_{S_k} S_k(t_2) \approx \sum_{k=1}^{\infty} \lambda_{W_k} W_k(t_2) + a^2 \sin(\omega t_2) \cdot \frac{2\sqrt{2}\pi n \sqrt{T} \sin(\omega T)}{\omega^2 T^2 - 4\pi^2 n^2} \tag{10.44}$$

We next multiply this equation by $W_n(t_2)$ and integrate with respect to t_2 between 0 and T. Because of (10.42) and (10.35), (10.44) becomes:

$$\lambda_{S_n} \approx \lambda_{W_n} + a^2 \frac{2\sqrt{2}\pi n \sqrt{T} \sin(\omega T)}{\omega^2 T^2 - 4\pi^2 n^2} \int_0^T W_n(t_2) \sin(\omega t_2)\, dt_2. \tag{10.45}$$

Again, the integral in the last equation can be computed—it is actually the same integral as in (10.43)—and so the conclusion is

$$\lambda_{S_n} \approx \lambda_{W_n} + a^2 \frac{8\pi^2 n^2 T \sin^2(\omega T)}{(\omega^2 T^2 - 4\pi^2 n^2)^2}. \tag{10.46}$$

This is Yatawatta's main result (corrected by Maccone). Let us now point out clearly that the eigenvalues on the left are a function of the final instant T; that is,

$$\lambda_{S_n}(T) \approx \lambda_{W_n} + a^2 \frac{8\pi^2 n^2 T \sin^2(\omega T)}{(\omega^2 T^2 - 4\pi^2 n^2)^2}. \tag{10.47}$$

This equation clearly shows that

(1) For $T \to 0$, the fraction in the right-hand side approaches zero, and so the eigenvalues of the signal + noise approach the pure white noise eigenvalues (as is intuitively obvious).
(2) For $n \to \infty$, again the fraction in the right-hand side approaches zero, and so the eigenvalues of the signal + noise approach the pure white noise eigenvalues (as again is intuitively obvious). This result may justify numerically the practical approximation made by the Medicina engineers when they confined their simulations to the first eigenvalue only (roughest approximation). In other words, the dominant eigenvalue of the signal + noise is given by

$$\lambda_{S_1}(T) \approx \lambda_{W_1} + a^2 \frac{8\pi^2 T \sin^2(\omega T)}{(\omega^2 T^2 - 4\pi^2)^2} = 1 + a^2 \frac{8\pi^2 T \sin^2(\omega T)}{(\omega^2 T^2 - 4\pi^2)^2}. \tag{10.48}$$

This completes our analysis of the KLT of a sinusoidal carrier buried in white, cosmic noise.

10.12 ANALYTIC PROOF OF THE BAM–KLT

We are now ready for the analytic proof of the BAM-KLT method.

Let us first re-write (10.47) in a form in which the pure white noise eigenvalues are replaced by 1:

$$\lambda_{S_n}(T) \approx 1 + a^2 \frac{8\pi^2 n^2 T \sin^2(\omega T)}{(\omega^2 T^2 - 4\pi^2 n^2)^2}. \tag{10.49}$$

We then notice that the final instant T appears three times in the right-hand side of the last equation:

(1) once in the numerator outside the sine;
(2) once in the numerator inside the sine;
(3) once in the denominator.

Therefore, the partial derivative of (10.49) with respect to T will be made up by the sum of three terms:

(1) One term with the derivative of the T in the numerator (i.e., 1 times the sine square). This brings a term in the cosine of TWICE the sine argument, since one obviously has
$$\sin^2(\omega T) = \tfrac{1}{2} - \tfrac{1}{2}\cos(2\omega T). \tag{10.50}$$
(2) One term with the derivative of the T inside the sine. This brings a term in the sine of TWICE the sine argument, because one has
$$2\sin(\omega T)\cos(\omega T) = \sin(2\omega T). \tag{10.51}$$
(3) One term with the derivative of the T in the denominator. This does not bring any term in either the sine or the cosine, but just a rational function of T that we shall give in a moment. In fact, we now prefer to skip the lengthy and tedious steps leading to the derivative of (10.49) with respect to T and just give the final result.

In conclusion, the derivative of (10.49) with respect to T is given by the following sum of three terms:
$$\frac{\partial\lambda_{S_n}(T)}{\partial T} \approx \mathrm{Coeff}_1(T)\cdot\sin(2\omega T) + \mathrm{Coeff}_2(T)\cdot\cos(2\omega T) + \mathrm{Coeff}_3(T) \tag{10.52}$$
where the three coefficients turn out to be (after lengthy calculations)
$$\left.\begin{aligned}
\mathrm{Coeff}_1(T) &= a^2\,\frac{8\pi^2 n^2\omega T}{(\omega^2T^2 - 4\pi^2n^2)^2},\\
\mathrm{Coeff}_2(T) &= a^2\,\frac{4\pi^2n^2(3\omega^2T^2 + 4\pi^2n^2)}{(\omega^2T^2 - 4\pi^2n^2)^3},\\
\mathrm{Coeff}_3(T) &= -a^2\,\frac{4\pi^2n^2(3\omega^2T^2 + 4\pi^2n^2)}{(\omega^2T^2 - 4\pi^2n^2)^3}.
\end{aligned}\right\} \tag{10.53}$$

But the right-hand side of (10.52) is no more than ... the simple Fourier series expansion of $\frac{\partial\lambda_{S_n}(T)}{\partial T}$. Moreover, (10.52) shows that $\frac{\partial\lambda_{S_n}(T)}{\partial T}$ is a PERIODIC function of T with frequency $2\omega T$. We conclude that: ***The Fourier transform of*** $\frac{\partial\lambda_{S_n}(T)}{\partial T}$ ***equals TWICE the frequency of the buried alien sinusoidal carrier. In other words, the***

frequency of the alien signal is a HALF of the frequency found by taking the Fourier transform of $\dfrac{\partial \lambda_{S_n}(T)}{\partial T}$.

And the BAM–KLT method is thus proved analytically.

10.13 KLT SIGNAL-TO-NOISE (SNR) AS A FUNCTION OF THE FINAL T, EIGENVALUE INDEX n, AND ALIEN FREQUENCY ν

We now derive a consequence from the eigenvalue relationship (10.47) dealing with the signal-to-noise ratio (abbreviated SNR) in the KLT theory. We shall call it the "KLT–SNR Theorem". The proof is as follows.

Consider Equation (10.10), showing that the eigenvalues λ_n of any KL expansion are actually the variances of the zero-mean corresponding uncorrelated (i.e., orthogonal, in the probabilistic sense) random variables Z_n. If we apply this to the KLT of stationary unitary white noise, described in Section 10.10, the conclusion is that the λ_{W_m} are the mean values of the square of the corresponding orthogonal (i.e., uncorrelated random variables $Z^2_{W_n}$)

$$\lambda_{W_m} = E\{Z^2_{W_n}\}. \tag{10.54}$$

Now, the definition of the signal-to-noise ratio (which we prefer to denote SNR, rather than S/R) of a sinusoidal signal with amplitude a buried in the noise with amplitude Z_{W_n} is just:

$$\mathrm{SNR} = \frac{\text{power of the signal}}{\text{power of the noise}} = \frac{a^2}{E\{Z^2_{W_n}\}} = \frac{a^2}{\lambda_{W_m}}. \tag{10.55}$$

This definition can now be inserted into (10.47) divided by λ_{W_m}; that is,

$$\frac{\lambda_{S_n}(T)}{\lambda_{W_n}} \approx 1 + \frac{a^2}{\lambda_{W_n}} \cdot \frac{8\pi^2 n^2 T \sin^2(\omega T)}{(\omega^2 T^2 - 4\pi^2 n^2)^2}, \tag{10.56}$$

with the result that (10.56) is changed into

$$\frac{\lambda_{S_n}(T)}{\lambda_{W_n}} \approx 1 + \mathrm{SNR} \cdot \frac{8\pi^2 n^2 T \sin^2(\omega T)}{(\omega^2 T^2 - 4\pi^2 n^2)^2}. \tag{10.57}$$

Solving this for SNR yields

$$\mathrm{SNR}(T, n, \omega) \approx \left(\frac{\lambda_{S_n}(T)}{\lambda_{W_n}} - 1\right) \cdot \frac{(\omega^2 T^2 - 4\pi^2 n^2)^2}{8\pi^2 n^2 T \sin^2(\omega T)}. \tag{10.58}$$

For SETI applications, it may be preferable to re-express the last formula directly in terms of the "alien" frequency $\nu = \dfrac{\omega}{2\pi}$, instead of ω. Equation (10.58) is thus changed into

$$\mathrm{SNR}(T, n, \nu) \approx \left(\frac{\lambda_{S_n}(T)}{\lambda_{W_n}} - 1\right) \cdot \frac{2\pi^2 (\nu^2 T^2 - n^2)^2}{n^2 T \sin^2(2\pi\nu T)}. \tag{10.59}$$

This is our KLT–SNR Theorem. Since the quantity

$$\left(\frac{\lambda_{S_n}(T)}{\lambda_{W_n}} - 1\right) > 0 \tag{10.60}$$

has a positive numeric value just slighty above zero, from (10.59) we conclude that

$$\begin{cases} \mathrm{SNR}(T,n,\nu) = O(T^3) & \text{as } T \to \infty \\ \mathrm{SNR}(T,n,\nu) = O(n^2) & \text{as } n \to \infty \\ \mathrm{SNR}(T,n,\nu) = O(\nu^4) & \text{as } \nu \to \infty. \end{cases} \tag{10.61}$$

These equations yield the "pace of increase" of the KLT–SNR, and should be of importance in writing down the numeric codes for the actual implementation of the KLT.

10.14 HOW TO EAVESDROP ON ALIEN CHAT

Following the Paris *First IAA Workshop on Searching for Life Signatures* (held at UNESCO, Paris, September 22–26, 2008, and organized by this author), the British popular science magazine *New Scientist* published the following article on October 30, 2008, that well summarizes the key features of the present scientific discussion.

How to eavesdrop on alien chat

30 October 2008

From *New Scientist* Print Edition.

Jessica Griggs

ET, phone . . . each other? If aliens really are conversing, we are not picking up what they are saying. Now one researcher claims to have a way of tuning in to alien cellphone chatter.

On Earth, the signal used to send information via cellphones has evolved from a single carrier wave to a "spread spectrum" method of transmission. It's more efficient, because chunks of information are essentially carried on multiple low-powered carrier waves, and more secure because the waves continually change frequency so the signal is harder to intercept.

It follows that an advanced alien civilisation would have made this change too, but the search for extraterrestrial life (SETI) is not listening for such signals, says Claudio Maccone, co-chair of the SETI Permanent Study Group based in Paris, France.

An algorithm known as the Fast Fourier Transform (FFT) is the method of choice for extracting an alien signal from cosmic background noise. However, the technique cannot extract a spread spectrum signal. Maccone argues that SETI should use an algorithm known as the Karhunen–Loève Transform (KLT), which could find a buried conversation with a signal-to-noise ratio 1000 times lower than the FFT.

A few people have been "preaching the KLT" since the early 1980s but until now it has been impractical as it involves computing millions of simultaneous equations, something even today's supercomputers would struggle with. At a recent meeting in Paris called Searching for Life Signatures, Maccone presented a mathematical method to get around this burden and suggested that the KLT should be programmed into computers at the new Low Frequency Array telescope in the Netherlands and the Square Kilometre Array telescope, due for completion in 2012.

Seth Shostak at the SETI Institute in California agrees that the KLT might be the way to go but thinks we shouldn't abandon existing efforts yet. "It is likely that for their own conversation they use a spread-spectrum method but it is not terribly crazy to assume that to get our attention they might use a 'ping' signal that has a lot of energy in a narrow band—the kind of thing the FFT could find."

"It is likely that aliens use the same spread-spectrum method of transmission as us on their cellphones."

From issue 2680 of *New Scientist* magazine,

30 October 2008, p. 14.

10.15 CONCLUSIONS

Let us summarize the main results of this chapter.

When the stochastic process $X(t)$ is stationary (i.e., it has both mean value and variance constant in time), then there are two alternative ways to compute the first KLT dominant eigenfunction (i.e., the roughest approximation to the full KLT expansion, which may be "enough" for practical applications!):

(1) (long way)—either you compute the first eigenvalue from the autocorrelation and then solve the huge (N^2) system of linear equations to get the first eigenfunction;
(2) (short way = BAM)—or you compute the derivative of the first eigenvalue with respect to $T = N$ and then Fourier-transform it to get the first eigenfunction.

In practical, numerical simulations of the KLT it may be much less time-consuming to choose option (2) rather than option (1).

In either case, the KLT of a given stationary process can retrieve a sinusoidal carrier out of the noise for values of the signal-to-noise ratio (SNR) that are three orders of magnitude lower than those that the FFT can still filter out. In other words, while the FFT (at best) can filter out signals buried in noise with an SNR of about 1 or so, the KLT can, say, filter out signals that have an SNR of, say, 0.001 or so.

This is the superior achievement of the KLT over the FFT.

The BAM (Bordered Autocorrelation Method) is an alternative numerical technique to evaluate the KLT of stationary processes (only) that may run faster on computers than the traditional full-solving KLT technique. In this chapter we have provided the results of numerical simulations that show, by virtue of the BAM, how the KLT succeeds in extracting a sinusoidal carrier embedded in a lot of noise when the FFT utterly fails.

10.16 ACKNOWLEDGMENTS

The author is indebted to many radioastronomers and scientists who helped him over the years to work out what is now the BAM–KLT method. Principal among them are Ing. Stelio Montebugnoli and his SETI-Italia Team, Dr. Mike Garrett and his ASTRON Team (in particular Dr. Sarod Yatawatta), Dr. Jill Tarter and the SETI Institute Team (in particular Drs. Seth Shostak and Doug Vakoch). Also, the Paris *SETI Conference of September 22–26, 2008*, organized by this author at UNESCO, was possible only through the full support of the Secretary General of the IAA, Dr. Jean-Michel Contant, and of the newly-born French SETI community. Finally, a number of other young and not-so-young folks continued to support this author in his efforts for SETI over the years, and their help is hereby gratefully acknowledged.

10.17 REFERENCES

[1] K. Karhunen, "Über lineare Methoden in der Wahrscheinlichkeitsrechnung," *Ann. Acad. Sci. Fennicae, Series A 1, Math. Phys.*, **37** (1946), 3–79.

[2] M. Loève, "Fonctions Aléatoires de Second Ordre," *Rev. Sci.*, **84**(4) (1946), 195–206.

[3] M. Loève, *Probability Theory: Foundations, Random Sequencies*, Van Nostrand, Princeton, NJ, 1955.

[4] C. Maccone, *Telecommunications, KLT and Relativity*, Volume 1, IPI Press, Colorado Springs, CO, 1994, ISBN # 1-880930-04-8. This book embodies the results of some 30 research papers published by the author about the KLT in the 15-year span 1980–1994 in peer-reviewed journals.

[5] S. Montebugnoli, C. Bortolotti, D. Caliendo, A. Cattani, N. D'Amico, A. Maccaferri, C. Maccone, J. Monari, A. Orlati, P. P. Pari *et al.*, "SETI-Italia 2003 Status Report and First Results of a KL Transform Algorithm for ETI Signal Detection," paper IAC-03-IAA.9.1.02 presented at the *2003 International Astronautical Congress* held in Bremen, Germany, September 29–October 3, 2003.

[6] F. Biraud, "SETI at the Nançay Radio-telescope," *Acta Astronautica*, **10** (1983), 759–760.

[7] C. Maccone, "Advantages of the Karhunen–Loève Transform over Fast Fourier Transform for Planetary Radar and Space Debris Detection," *Acta Astronautica*, **60** (2007), 775–779.

10.18 ANNOTATED BIBLIOGRAPHY

In addition to the above references, we would like to offer an "enlightened" list of a few key references about the KLT, subdivided according to the field of application.

Early papers by the author about the KLT in mathematics, physics, and the theory of relativistic interstellar flight, subdivided by journals

Il Nuovo Cimento

[8] C. Maccone, "Special Relativity and the Karhunen–Loève Expansion of Brownian Motion," *Nuovo Cimento, Series B*, **100** (1987), 329–342.

Bollettino dell'Unione Matematica Italiana

[9] C. Maccone, "Eigenfunctions and Energy for Time-Rescaled Gaussian Processes," *Bollettino dell'Unione Matematica Italiana, Series 6,* **3-A** (1984), 213–219;

[10] C. Maccone, "The Time-Rescaled Brownian Motion $B(t^{2H})$," *Bollettino dell'Unione Matematica Italiana, Series 6*, **4-C** (1985), 363–378; C. Maccone, "The Karhunen–Loève Expansion of the Zero-Mean Square Process of a Time-Rescaled Gaussian Process," *Bollettino dell'Unione Matematica Italiana, Series 7*, **2-A** (1988), 221–229.

Journal of the British Interplanetary Society

[12] C. Maccone, "Relativistic Interstellar Flight and Genetics," *Journal of the British Interplanetary Society*, **43** (1990), 569–572.

Acta Astronautica

[13] C. Maccone, "Relativistic Interstellar Flight and Gaussian Noise," *Acta Astronautica*, **17**(9) (1988), 1019–1027.

[14] C. Maccone, "Relativistic Interstellar Flight and Instantaneous Noise Energy," *Acta Astronautica*, **21**(3) (1990), 155–159.

KLT for data compression

[15] C. Maccone, "The Data Compression Problem for the 'Gaia' Astrometric Satellite of ESA," *Acta Astronautica*, **44**(7–12) (1999), 375–384.

Some important papers about the KLT for SETI

[16] R. S. Dixon, and M. Klein, "On the detection of unknown signals," *Proceedings of the Third Decennial US–USSR Conference on SETI* held at the University of California at Santa Cruz, August 5–9, 1991. Later published in the *Astronomical Society of the Pacific (ASP) Conference Series* (Seth Shostak, Ed.), **47** (1993), 128–140.

[17] C. Maccone, *Karhunen–Loève versus Fourier Transform for SETI*, Lecture Notes in Physics, Springer-Verlag, Vol. 390 (1990), pp. 247-253. These are the *Proceedings* (J. Heidmann and M. Klein, Eds.) *of the Third Bioastronomy Conference* held in Val Cenis, Savoie, France, June 18–23, 1990.

After these seminal works were published, the importance of the KLT for SETI was finally acknowledged by SETI Institute experts in

[18] R. Eckers, K. Cullers, J. Billingham, and L. Scheffer (Eds.), *SETI 2020*, SETI Institute, Mountain View, CA, 2002, p. 234, note 13. The authors say: "Currently (2002) only the Karhunen Loeve (KL) transform [Mac94] shows potential for recognizing the difference between incidental radiation technology and white noise. The KL transform is too

computationally intensive for the present generation of systems. The capability for using the KL transform should be added to future systems when computational requirements become affordable."

The paper [Mac94] referred to in the *SETI 2020* statement mentioned above is

[19] C. Maccone, "The Karhunen–Loève Transform: A Better Tool than the Fourier Transform for SETI and Relativity," *Journal of the British Interplanetary Society*, **47** (1994), 1.

An early paper about the KLT for SETI-Italia

[20] S. Montebugnoli, and C. Maccone, "SETI-Italia Status Report 2001", a paper presented at the 2001 *IAF Conference* held in Toulouse, France, October 1–5, 2001.

An early paper about the possibility of a "fast" KLT

[21] A. K. Jain, "A Fast Karhunen–Loève Transform for a Class of Random Processes," *IEEE Trans. Commun.*, **COM-24** (1976), 1023–1029.

Recent papers about the KLT and BAM–KLT

[22] F. Schillirò, S. Pluchino, C. Maccone, and S. Montebugnoli, *La KL Transform: considerazioni generali sulle metodologie di analisi ed impiego nel campo della Radioastronomia*, Istituto Nazionale di Astrofisica (INAF)/Istituto di Radioastronomia (IRA), Technical Report, January 2007 [in Italian].

[23] C. Maccone, "Innovative SETI by the KLT," *Proceedings of the Bursts, Pulses and Flickering Conference* held at Kerastari, Greece, June 13–18, 2007. Available at POS (Proceedings of Science) website *http://pos.sissa.it//archive/conferences/056/034/Dynamic2007_034.pdf*

[24] S. Yatawatta, pers. commun., June 17, 2008.

A recent paper about the KLT for relativistic interstellar flight

[25] C. Maccone, "Relativistic Optimized Link by KLT," *Journal of the British Interplanetary Society*, **59** (2006), 94–98.