This article was downloaded by:[Picinbono, B.]
On: 2 April 2008
Access Details: [subscription number 791840432]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

International Journal of Control

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713393989>

Singularity in signal theory

B. Picinbono^a

^a Laboratoire des Signaux et Systèmes Supélec, Plateau de Moulon, 91192 Gif sur Yvette, France

Online Publication Date: 01 March 2008

To cite this Article: Picinbono, B. (2008) 'Singularity in signal theory', International Journal of Control, 81:3, 495 - 506

To link to this article: DOI: 10.1080/00207170701556906

URL: <http://dx.doi.org/10.1080/00207170701556906>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Singularity in signal theory

B. PICINBONO*

Laboratoire des Signaux et Systèmes Supélec, Plateau de Moulon, 91192 Gif sur Yvette, France

(Received 5 July 2007; in final form 6 July 2007)

A discrete-time random signal is singular if its values are singular random variables defined by a distribution function continuous but with a derivative equal to zero almost everywhere. Singular random signals can be obtained at the output of some linear filters when the input is a discrete-valued white noise. Sufficient conditions for singularity are established. In particular it is shown that if the poles of the filter are inside a circle called the circle of singularity and if the input is white and discrete-valued the output is singular. Computer experiments using histograms at different scales exhibit the structure of singular signals. The influence of input correlation is also analysed. It is shown that when the input is not white, but has a specific Markovian structure, the output can be singular. This is also verified by computer experiments. Finally, mixtures of singular and discrete-valued random signals are analysed.

1. Introduction

Singularity is a concept introduced in probability theory and related to properties of distribution functions of random variables. Usually only two kinds of random variables are considered: those that are continuous and having a distribution function (DF) with a derivative (probability density function) and those that are discrete with a DF varying by steps defining possible values and their corresponding probabilities. However there is a third kind of random variables: those that are singular.

A random variable (RV) is said to be singular if its DF is continuous but with a derivative equal to zero almost everywhere. Then such a RV is neither continuous (no probability density function) nor discrete because its DF is continuous and not a stepwise function.

Singular RVs are often considered as mathematical curiosities without interest in signal theory and engineering sciences. They are rarely introduced in standard textbooks (Papoulis 1984, Ochi 1990, Pfeifer 1990, Helstrom 1991, Picinbono 1993) and appear only in mathematically oriented books (Loève 1977, Wong and Hajek 1985). Very simple RVs however can be singular. The best example comes from an old result known for more than sixty years. It says that if w_k is a set of

independent and identically distributed (IID) RVs taking the values ± 1 with the same probability, then the sum of the series $\sum_{k=0}^{\infty} a^k w_k$ is singular as soon as $|a| < 1/2$. So one of the simplest RV that can be considered is singular. This result is often omitted because its usual proof requires very abstract reasoning. One of the first tasks of this paper is to show its origin and to introduce an elementary proof that can afterwards be used in signal theory. Indeed the series considered above is similar to the output of an exponential discrete-time filter whose input is a discrete valued white noise. This is the case of autoregressive signals widely used in signal processing. Then the question of knowing whether or not this can be extended to other class of signals and systems appears immediately.

The main result of this paper is that singularity of the output of a linear filter depends on two facts: the discrete character of the input, which is a common situation in communication theory, and some specific properties of the filter such as the location of its poles in a circle called the circle of singularity. In order to provide a better understanding of a problem usually widely ignored, and to visualize how singularity can appear for rather simple signals, number of computers experiments are presented.

In the last part of the paper it is also shown that the assumption of whiteness, fundamental in the proof of

*Email: bernard.picinbono@lss.supelec.fr

the basic result, can be partially deleted and we present some examples of coloured input signals leading also to singular outputs. This is especially the case of some Markov processes. Finally mixtures of discrete and singular random signals can be obtained depending on the properties of the correlation function of the input, and theoretical and experimental examples of such signals are presented.

2. Statement of the problem and review of known results

Let $X(\omega)$ be a RV defined on some probability space. Let $F(x)$ be its distribution function (DF). In all of what follows we do not make any difference between two RVs distinct but equal with probability 1. This means that $X(\omega)$ is entirely defined by its DF $F(x)$. The RV $X(\omega)$ is said to be singular if its DF $F(x)$ is continuous but with a derivative equal to zero almost everywhere.

At the beginning let us remind some fundamental results of probability theory used in the discussion that follows and especially the Lebesgue decomposition theorem (Lukacs 1970). For this let us recall that a RV is said to be continuous if its DF is continuous and has a derivative which is its probability density function (PDF). On the other hand a RV is said to be discrete if its DF varies only by steps of amplitudes p_i at points x_i . This means that the RV takes only the values x_i with the probabilities p_i . The Lebesgue theorem says that any DF $F(x)$ can be decomposed in a unique way as a sum of three terms, or that

$$F(x) = a_1 F_c(x) + a_2 F_d(x) + a_3 F_s(x);$$

$$a_1 + a_2 + a_3 = 1; \quad a_i \geq 0. \quad (1)$$

In this equation the three functions $F_i(x)$ are DFs and $F_c(x)$, $F_d(x)$, and $F_s(x)$ are the continuous, discrete and singular components of $F(x)$ respectively. If $a_1 = 1$, the RV is continuous and its PDF is the derivative of $F_c(x)$. If $a_2 = 1$, X is a discrete RV. If, finally, $a_3 = 1$, the RV X is singular. If one of the coefficients a_i is equal to 1, the DF is said to be pure. The spectrum S_F of a RV is the set of the points of increase of its DF $F(x)$ and the spectral measure (SM) is the Lebesgue measure $L(S_F)$ of this set. It is clear that if the SM is zero the continuous part in (1) is zero, or $a_1 = 0$. Then the RV can be discrete ($a_3 = 0$), singular ($a_2 = 0$), or a mixture of a discrete and a singular parts. Then to show that a RV is singular it suffices to show that its SM is zero and that there is no discrete part in the decomposition of its DF.

The fundamental theorem opening all this discussion is the following. Consider a set of RVs w_k independent and identically distributed (IID) with a symmetric Bernoulli distribution. This means that the w_k s take

only two values with the same probabilities $1/2$. When this is not otherwise explicitly indicated the two possible values of each RV w_k are ± 1 . Consider the RV X defined by the series

$$X = \sum_{k=0}^{\infty} a^k w_k, \quad (2)$$

which is convergent provided that $|a| < 1$. Note that the possible values of w_k imply that X is symmetric, which means that X and $-X$ have the same DF. Furthermore, for the same reasons, changing a in $-a$ does not change the DF of X . Then we can assume that $a \geq 0$. The following results holds:

- (1) if $a < 1/2$, X is singular;
- (2) if $a = 1/2$, X is uniformly distributed in $[-2, +2]$;
- (3) if $1/2 < a < 1$, X is in general continuous, but can also be singular for values of a belonging to a set of zero measure.

Point 2 can be shown directly by a simple calculation. Point 3 is without interest for this paper and the complete set of points of singularity is still a subject of research. Point 1 was shown long time ago in the framework of infinite products of convolutions (Kershner and Wintner 1935) and discussed more recently (Peres and Solomyak 1998, Solomyak *et al.* 2000). Because of its importance for all what follows, we shall now present a direct proof.

For this purpose consider the partial sum X_N and the rest R_N defined by

$$X_N = \sum_{k=0}^N a^k w_k; \quad R_N = \sum_{k=N+1}^{\infty} a^k w_k = a^{N+1} \sum_{k=0}^{\infty} a^k w_{N+k+1}. \quad (3)$$

The RV X_N takes 2^{N+1} distinct values v_i^N with the same probability $1/2^{N+1}$. The rest R_N satisfies $|R_N| \leq a^{N+1}/(1-a)$, this limit being obtained if all the w_{N+k+1} are equal to $+1$ or -1 . As a consequence the SM of R_N is smaller than $a^{N+1} [2/(1-a)]$. Since there are 2^{N+1} distinct values v_i^N of X_N , the SM of X is smaller than $(2a)^{N+1} [2/(1-a)]$. As this is valid for all N , this SM is zero whenever $a < 1/2$.

Note that this property of the SM is due to the value of a and to the fact that the input w_k has only two possible values. On the other hand the probabilities of these outcomes do not play any role. We shall see later that this is general.

This result means that there is no continuous part in the decomposition of the DF of X , or that $a_1 = 0$ in (1). Let us see now that there is also no discrete component, or $a_2 = 0$. Indeed suppose that this is not the case. This would mean that there is a value x_0 such that the probability that $X = x_0$ is positive. But as x_0 is a value of

X , there exists a set of numbers η_k taking only the values ± 1 such that $x_0 = \sum_{k=0}^{\infty} a^k \eta_k$. Furthermore, as the values v_i^N are distinct, this set is unique. Since the RVs w_k are IID and $P[w_k = \eta_k] = 1/2$, we deduce

$$P[X = x_0] = \prod_{k=0}^{\infty} P[w_k = \eta_k] = 0. \quad (4)$$

As a consequence the RV X is singular, or $a_3 = 1$. On the contrary to the property of the SM, it is clear that this proof uses the fact that the w_k s are independent and that their two values have the same probability $1/2$. It is simple to see however that the result remains valid if one of the values has the probability p , except when $p = 0$ or $p = 1$, which corresponds to a situation where the input signal w_k is no longer random.

For the discussion that follows it is important to understand that the singularity comes from two completely distinct properties. The first one depends only on a and on the fact that w_k takes only two values ± 1 . But it is insufficient to imply singularity and we shall see later that for some particular probability distributions of the w_k the sum X can be discrete or a mixture of discrete and singular parts. The second properties arises from the whiteness of w_k . It is the combination of these two properties which ensures that X is singular.

Let now present an interpretation of X_N by its tree of construction presented in figure 1. For each value v_i^N of X_N we can associate the value $-v_i^N$. It is obtained simply by changing the signs of the w_k s appearing in (3). This means, as noted above, that the the RV X is symmetric. As a consequence we can consider only positive v_i^N s. For the same reason it is always possible to assume that $a > 0$. With these assumptions we have $v_0^0 = 1$. The two positive values of v_i^1 are $1 - a$ and $1 + a$. The construction of the 8 positive values of X_3 appears in the tree of figure 1. Let us now see that the assumption $a < 1/2$ means that there is no crossing of the branches of the tree. Indeed consider the two branches of the tree starting from a point v_i^N . There is no crossing between all the branches starting from this point if $v_i^N - a^{N+1} + a^{N+2}/(1 - a) < v_i^N + a^{N+1} - a^{N+2}/(1 - a)$. This yields $a < 1/2$. It is clear that the construction of this

Figure 1. Tree of construction of the successive values v_i^N .

tree is similar to the one of Cantor sets. This is why it is sometimes said that the RV X has a Cantor-type distribution (Wittke *et al.* 1988).

Singularity is not limited to random geometric series like (2) but can appear with RVs such as

$$X = \sum_{k=0}^{\infty} h_k w_k, \quad (5)$$

where $h_k > 0$ and the w_k s have the same properties as in (2). In this case X_N and R_N of (3) are written simply by replacing a^k by h_k . The last equality of (3) does not hold.

It is shown without complete proof in p. 66 of (Lukacs 1970) that if

$$\rho_n = \sum_{k=n+1}^{\infty} h_k < h_n, \quad \forall n, \quad (6)$$

then X is singular. In the case where $h_k = a^k$, this yields $a < 1/2$. It is possible to construct a tree as in figure 1 with the h_k s instead of the a^k s. One can then see that the condition (6) implies that there is no crossing between the branches of the tree, which introduces again a Cantor structure (Picinbono and Tournet 2005).

Before leaving this section let us present a short review of some papers from the engineering literature where the problem of singularity is discussed. The treatment of sequences of Bernoulli RVs appear frequently in the context of digital communications. The first discussion concerning consequences of singularity was presented in (Hill and Blanco 1973). The discussion was limited to geometric series like (2) and the purpose was to obtain an approximation of the DF for the calculation of the error probability or the performance of communication systems. Extensions of the same problem to Cantor-type distribution was presented in (Wittke *et al.* 1988). In this paper the condition (6) is explicitly used and various examples of filters satisfying this condition are introduced. However, as indicated by the authors, the condition $h_k > 0$ used in (5) and introducing the Cantor structure is very restrictive and it is not satisfied by a large class of filters containing for example terms like $a^k \cos(\omega k)$. The principle of calculation of the distribution function and of some expectations is then presented and used for the evaluation of the error probability. A rather more theoretical approach of the same problem is presented in (Smith *et al.* 1993). Finally other calculations of expectations of singular RVs are discussed in (Campbell *et al.* 1995). In this paper the singularity is introduced from some properties of the entropy of the RVs by using an approach introduced in (Garsia 1962). Similar discussions appear in (Naraghi-Pour *et al.* 1990, Tournet *et al.* 1994).

The first purpose of the present paper is to show that singularity can be introduced from considerations of

properties of the poles of a linear filter. Furthermore in all these papers it is assumed that the RVs w_k are independent, and this assumption is a corner stone for the introduction of singularity. Then it arises immediately the question of knowing whether or not it can be relaxed without suppressing singularity. This question is discussed in the second part of the paper.

3. Singularity of autoregressive of order 1 signals

The previous discussion has an immediate application in the case of autoregressive signals. A signal x_k is said to be autoregressive of order 1 [AR(1)] if it is deduced from a white noise w_k by the first order recursion $x_k = ax_{k-1} + w_k$. It is then defined by the parameter a called the regression coefficient and by the DF of w_k . This recursion introduces a linear filter and its input-output relationship is

$$x_n = \sum_{k=0}^{\infty} a^k w_{n-k}. \quad (7)$$

It is then a discrete time causal filter with the impulse response a^k . If the input signal is a white symmetric Bernoulli signal, it results from the previous discussion that the RVs x_k are singular as soon that $|a| < 1/2$, and we say that the signal x_k is singular. Indeed in order to come at (2) it suffices to introduce $\hat{w}_k = w_{n-k}$ and it is clear that the RVs $\hat{w}_k$ are still IID Bernoulli. The simplicity of this signal explains why it was said previously that singularity is a common phenomenon.

Five trajectories of the signal x_k are represented in figure 2. The corresponding values of a are : 0; (1); 0.2; (2); 1/3; (3); 0.4; (4); 0.5; (5). Trajectory 1 corresponds to the white noise w_k . The only specific property apparent on these trajectories is the damping phenomenon characterized by the fact that the speed of variation of x_k decreases when a increases. There is however no element which can illustrate the singularity.

For this purpose it is necessary to study the DF of the signals and the appropriate mean for this purpose is to realize histograms at various scales. A histogram is related to the DF according to the following property. Let $n(x, \Delta x)$ be the number of samples of the signal x_k recorded in an interval $]x, x + \Delta x]$. Let S be the total number of samples of the signal analysed in the histogram. According to the law of large numbers, the ratio $n(x, \Delta x)/S$ tends to the increment $F(x + \Delta x) - F(x)$ of the DF when $S \rightarrow \infty$. If the RV x_k is continuous ($a_1 = 1$) the histogram yields an estimation of the PDF, and this is the usual way for this purpose. On the other hand if x_k is singular the PDF is zero and this is why an analysis at different scales is necessary. For this let us remind that a histogram is characterized by

Figure 2. Trajectories of x_k for various values of the regression coefficient a : 1. $a=0$; 2. $a=0.2$; 3. $a=1/3$; 4. $a=0.4$; 5. $a=0.5$.

three parameters. The first is the total number S of samples analysed. The second is the interval of analysis $[\alpha, \beta]$. This means that only the samples x_k satisfying $\alpha < x_k < \beta$ are taking into account in the histogram. The third is the number B of adjacent bins in this interval. The bin i of this histogram corresponds to the interval $[\alpha + i(\beta - \alpha)/B, \alpha + (i + 1)(\beta - \alpha)/B]$, $0 \leq i \leq B - 1$.

In order to analyse singularity by histograms we consider the successive intervals centered at some nodes v_i^N of the tree of figure 1 and defined by $[v_i^N - \rho_N, v_i^N + \rho_N]$, where ρ_N is defined by (6). For $h_k = a^k$, it satisfies $\rho_N = a^{N+1}/(1 - a)$.

In the computer experiments presented in what follows S is of the order of 10^7 and $B=400$. In figure 3 four histograms corresponding to various values of the regression coefficient a are presented. In these figures the centres of the histograms are the origin and we verify that X is symmetric. We observe clearly an evolution from the case where there are only two lines ($a=0$), which is not represented, to a uniform distribution between $[-2, +2]$ obtained for $a=1/2$, as predicted by the theory.

A more detailed analysis of the histogram 2 of this figure corresponding to $a=1/3$ is presented in figure 4. In this case the extreme values of X are $\pm 3/2$.

Figure 3. Histograms of x_k for various values of the regression coefficient a : 1. $a=0.2$; 2. $a=1/3$; 3. $a=0.4$; 4. $a=0.5$.

The number of samples of the signal in the experiment is $S=2 \cdot 10^6$. The centres v_i^N of the histograms are obtained by taking $i=2^N-1$. These successive centres of the seven histograms are 0, 1, 1.133, 1.444, 1.4815, 1.4938, 1.4979. The ratio of the two extreme window widths is 729, and even with this very great diminution of the interval of analysis the structure of the histograms remains the same. Of course it would be possible to continue the experiment, but this would require a much greater value of S . This figure exhibits a fractal property explaining why the DF has no derivative (Falconer 1990).

In order to have a more precise idea of the phenomenon we present the result of the same experiment realized with $a=1/2$ where the theory predicts a uniform distribution in $[-2, +2]$. This appears in figure 5 which indicates clearly this uniform distribution. The fluctuations appearing in the last histogram are simply due to too small number of samples appearing in the interval of analysis.

4. Extensions to other filters

It is intuitive that the previous results are not restricted to AR(1) signals with symmetric Bernoulli inputs. In this

Figure 4. Histograms with scaling effect for an AR (1) signal with $a=1/3$.

Figure 5. Histograms with scaling effect for an AR (1) signal with $a=1/2$.

section we shall effectively show that they can be extended to a large number of other signals. This can be done either by changing the filter that yields x_k from w_k or by changing the statistical properties of the input w_k .

Let us first present some preliminary considerations. We restrict our discussion to dynamical discrete time filters. These filters are characterized by the facts that they are causal and that their transfer function $H(z)$ is a rational function of z . Such filters are defined by their poles and their zeros and almost all the filters used in signal processing are dynamical. We consider further only IIF filters, because RIF filters cannot introduce singularity. Indeed this property, as seen previously, is due to a series and the input-output relationships in RIF filters is a simple sum. This means that we exclude from our analysis transfer functions with only one pole at the origin.

We consider also white discrete-valued inputs signals w_k , or sequences of IID random variables taking only a finite number q of possible values.

Let finally introduce a circle called circle of singularity with center O and with the radius equal to $1/q$. This allows to introduce the following theorem.

Theorem 1: *Let x_k be the output of a dynamical non-RIF filter generated by the input w_k . If w_k is a white signal taking only q values and if the poles of the filter lie inside the circle of singularity, then the output x_k is singular.*

Proof: As in the previous section it is made in two steps: 1. Proof that the SM is zero, 2. Proof that there is no discrete component in the DF.

Let $\mathcal{F}$ be a dynamical filter defined by its transfer function $H(z)$ or its impulse response h_k . Let $\mathcal{F}'$ be the filter with the impulse response $g_k = q^k h_k$. It is obvious that its transfer function is $G(z) = H(z/q)$. This implies that if the poles of $\mathcal{F}$ are z_i , those of $\mathcal{F}'$ are qz_i . The assumption that the poles of $\mathcal{F}$ are inside the circle of singularity implies that the poles of $\mathcal{F}'$ are inside the unit circle, or that $\mathcal{F}'$ is a dynamical filter. As such a filter is stable we deduce that $\sum_{k=0}^{\infty} |g_k| < \infty$. The signal x_p is defined by $x_p = \sum_{k=0}^{\infty} h_k w_{p-k}$ and by introducing $x_p = X$ and $\hat{w}_k = w_{p-k}$ we have $X = \sum_{k=0}^{\infty} h_k \hat{w}_k$, where the $\hat{w}_k$ s have the same properties as the w_k s or are IID and take only q values. As indicated above the finite sum X_N and the rest R_N are defined in (3) where a^k is replaced by h_k and w_k by $\hat{w}_k$. Let A be the greatest possible value of w_k . We have then $|R_N| \leq A \rho_N$ with $\rho_N = \sum_{k=N+1}^{\infty} |h_k|$. As X_N can take only q^N values, the SM $\mathcal{S}$ of X satisfies $\mathcal{S} \leq 2Aq^N \rho_N$. This is valid for all N . Then $\mathcal{S} \leq 2A \lim_{N \rightarrow \infty} (q^N \rho_N)$, but we have

$$q^N \rho_N = q^N \sum_{k=N+1}^{\infty} |h_k| < \sum_{k=N+1}^{\infty} q^k |h_k| = \sum_{k=N+1}^{\infty} |g_k|, \quad (8)$$

and the limit is 0 because the filter $\mathcal{F}'$ is stable or $\sum_{k=0}^{\infty} |g_k| < +\infty$. This implies that $\mathcal{S} = 0$, or $a_1 = 0$. The proof that $a_2 = 0$ is exactly the same as previously and comes only from the whiteness of the w_k s or from their independence. Another proof of the fact that there is non discrete component is given by Lukacs (1970, p. 65).

It is important to note that this theorem yields only a sufficient condition of singularity. However the condition that the poles lie in the singularity circle is not at all necessary. Let us discuss this point on a particular example.

Consider the filter defined by the transfer function $H(z) = z^4/(z^4 - a^4)$. It has four poles located on the circle of radius a . Its impulse response is such that $h_{4k} = a^{4k}$ and the other h_k are zero. It introduces the RV $X = \sum_{k=0}^{\infty} a^{4k} w_{4k}$. It results from the previous result that X is singular if $a^4 < 1/2$, or $a < 2^{-1/4} = 0.8409$ and uniformly distributed if $a = 2^{-1/4}$. Then singularity appears for poles outside the singularity circle. This is presented in figure 6 which is very similar to figure 4. The histograms corresponding to the values of a equal to 0.5, 0.75, and 0.8 exhibit singular behaviour.

Let us now present some comments on this theorem. The first point to note is that it implies the result analysed in the previous section. Indeed the causal filter

Figure 6. Four poles $z_i = a^{1/4}$: 1. $a = 0.5$; 2. $a = 0.75$; 3. $a = 0.8$; 4. $a = 2^{-1/4} = 0.8409$.

with impulse response $h_k = a^k$ is obviously dynamical because its transfer function is simply $H(z) = z / (z - a)$. Then if $q=2$, or if the input signal w_k takes only two distinct values, the radius of the circle of singularity is $1/2$ and we find again the condition $a < 1/2$.

The second point to note is that the possible values of w_k do not play any role in the result. However for a given set of values it is sometimes possible to extend the domain of singularity by more abstract methods (Peres and Solomyak 1998). But if we impose the singularity for all the possible sets of values, we come back to the conditions of Theorem 1.

The final point is that it is in general no longer possible to interpret the result by a tree of construction. This is especially the case when the dynamical filter has complex poles as for example when the impulse response is of the kind $a^k \cos(k\phi)$.

In order to illustrate the results of this section let us consider the filter defined by $H(z) = z^3 / (z - a)^3$, $0 < a < 1$. Its impulse response is $h_k = (1/2)k(k+1)a^k$. It has a triple pole at $z = a$. Suppose that the input is a Bernoulli white noise as in the previous sections. In this case if $a < 1/2$ the pole is inside the circle of singularity and the output is singular. This is illustrated in figure 7 calculated for $a = 1/3$ as in figure 4. The results are quite different and the fractal structure appears only in the last three histograms. It is interesting to note that the

Figure 7. Histograms with scaling effect for an AR (3) signal with $a = 1/3$.

condition (6) requires that $a < 0.2063$, which is not satisfied in this example. This means that the DF is not a Cantor-type distribution, and this appears in the figure. This is a good example among many others that the condition (6) is not required to introduce singularity.

The results of various other computer experiments are presented in (Picinbono and Tourneret 2005) and the fractal behaviour of the histograms appears clearly, but in a less regular form than in figure 4 valid in the case of the exponential filter. All these results show the great variety of possible singular signals.

5. The influence of the correlation

The assumption of independence, or of whiteness of the input, plays a fundamental role in the previous results. It is introduced in all the papers indicated in the list of references. This assumption allows us to show that, even if the SM is zero, there is no discrete component in the DF, of that $a_2 = 0$. Thus appears immediately the question of knowing whether it is still possible to meet singularity in the case of coloured inputs.

Consider a filter $\mathcal{F}$ defined by its impulse response h_k and satisfying the conditions of Theorem 1. The problem of singularity of the output depends on the properties of the RV X given by (5) where the condition $h_k > 0$ is relaxed. We assume that the w_k s are symmetric Bernoulli, but not necessarily independent. The maximum value of X is $X_m = \sum_{k=0}^{\infty} |h_k|$ which is finite because $\mathcal{F}$ is stable. Let X_N be the partial sum analogue to (3) and defined by

$$X_N = \sum_{k=0}^N h_k w_k. \quad (9)$$

It takes at the maximum 2^{N+1} distinct values v_i^N and we assume that this maximum is reached. This assumption of distinct values is obviously satisfied when there is no crossing of the branches of the tree constructed with the h_k . This appears with filters satisfying (6), or for the Cantor-type structure. It is clear that this assumption depends only of the impulse response h_k of the filter. It is however satisfied by a large class of filters which is not discussed here. This means that for any N there is no pair (i, j) , $i \neq j$, such that $v_i^N = v_j^N$. Because of the symmetry of the w_k s, the RVs X and X_N are also symmetric and this implies that $v_i^N \neq 0$. Indeed the symmetry and the existence of a zero value would imply that the number of distinct values is odd, which is not the case.

To each value v_i^N we associate a node V_i^N in the tree of construction similar to the one appearing in figure 1. The assumption of distinct values v_i^N means that the nodes of the tree are single, which means that each node

V_i^N is reached by only one path coming from only one node V_j^{N-1} at the step $N-1$. This is obviously satisfied when there is no crossing of the branches of the tree constructed with the h_k s. This especially appears if condition (6) is satisfied. However, since h_k is not necessarily positive as in figure 1, it is not possible to restrict this tree to the nodes V_i^N corresponding to positive values v_i^N . Then we assume that there are 2^{N+1} distinct nodes satisfying $v_0^N < v_1^N < \dots < v_{2^{N+1}-1}^N$. Finally we assume that the nodes V_0^0 and V_1^0 defined by $-v_0^0 = v_1^0 = |h_0|$ are generated from an origin node V which does not correspond to a value of X_N .

The fundamental consequence of the assumption of distinct values v_i^N is that for any N and i there is a unique path going from V to V_i^N . Let $i_k^N(i)$, $0 \leq k \leq N-1$, be the indices j characterizing the nodes V_j^k of this path. These nodes can then be written $V_{i_k^N(i)}^k$.

The problem is to calculate the probabilities

$$p_N(i) \triangleq P[X_N = v_i^N], \quad 0 \leq i \leq 2^{N+1} - 1. \quad (10)$$

When the w_k s are IID this probability is $1/2^{N+1}$. When they are no longer independent, its calculation is much more complicated.

For this we introduce the conditional probability

$$p_N(i, j) \triangleq p_N[X_N = v_i^N | X_{N-1} = v_j^{N-1}], \quad (11)$$

called also transition probability. It has two fundamental properties for the discussion that follows.

The first comes from the fact that, as any probability, it is normalized or satisfies for all j the relation $\sum_{i=0}^{2^{N+1}-1} p(i, j) = 1$. However a node V_j^{N-1} of the tree of construction generates only two nodes V_i^N characterized by the indices $i_+(j)$ and $i_-(j)$ and, according to (9), corresponding to the values $v_{i_\pm}^N(j) = v_j^{N-1} \pm h_N w_N$. As a consequence for a given j there is only two terms in the previous sum and we have

$$p_N[i_+(j), j] + p_N[i_-(j), j] = 1. \quad (12)$$

The second starts from the fact that there is only one V_j^{N-1} at the step $N-1$ of the tree generating V_i^N and called $j(i)$. Thus $p_N(i, j)$ is zero except when $j=j(i)$, and the only non-zero values of $p_N(i, j)$ are

$$q_N(i) = p_N[i, j(i)] \quad (13)$$

for $N > 0$ and $q_0(i) = 1/2$.

It results from (11) and from the unicity of the path between V and V_i^N that

$$p_N[(X_N = v_i^N) \cap (X_{N-1} = v_j^{N-1})] = p_{N-1}(j) \cdot p_N(i, j) \delta[j - j(i)], \quad (14)$$

where $\delta[\cdot]$ is the Kronecker delta symbol. By a summation on j , which contains only one term,

we obtain

$$p_N(i) = p_{N-1}[j(i)] q_N(i). \quad (15)$$

By repeating this at all the nodes of the unique path between V and V_i^N characterized by the indices $i_k^N(i)$ we obtain

$$p_N(i) = \prod_{k=0}^N q_k[i_k^N(i)]. \quad (16)$$

When the RVs w_k are IID we have of course $q_N[i_k^N(i)] = 1/2$, and we find again that the values v_i^N have equal probabilities $1/2^{N+1}$.

The probabilities $p_N(i)$ of (16) are normalized, or $\sum_i p_N(i) = 1$, where the sum is extended to all the indices i from 0 to $2^N - 1$. This property is valid for $N=0$ because $q_0(i) = 1/2$. Suppose that it is valid at the step $N-1$. Since each node V_j^{N-1} generates only two nodes $V_{i_-(j)}^N$ and $V_{i_+(j)}^N$ the result comes from (12).

The relation (16) is the basis for the discussion of the singularity. Indeed if all the $p_N(i)$ tend to 0 when $N \rightarrow \infty$, there is no value v_i^∞ with a finite probability, and this means that the RV cannot have a discrete component and then is singular. This can be specified by the following theorem.

Theorem 2: Let X be the RV $\sum_{k=0}^\infty h_k w_k$, where h_k is the impulse response of a dynamical non-RIF filter $\mathcal{F}$ and w_k a sequence of Bernoulli RVs. If the poles of $\mathcal{F}$ are inside the circle of singularity, if the possible values v_i^N of the partial sums X_N are distinct and if the transition probabilities $q_N(i)$ defined by (13) satisfy

$$0 < q_N(i) < B < 1, \quad (17)$$

then the RV X is singular, or $a_1 = a_2 = 0$.

Proof: If the poles are inside the circle of singularity, Theorem 1 shows that the SM is zero or $a_1 = 0$. It remains to show that $a_2 = 0$. This is a direct consequence of (16) and (17) because $p_N(i) < (1/2)B^N$, which tends to zero when $N \rightarrow \infty$.

Comments: It is clear that this situation appears for white input because in this case $q_N(i) = 1/2$. The question that remains concerns the conditions of the theorem on the filter. It is clear from the previous discussion and according to figure 1 that if $h_k = a^k$ with $|a| < 1/2$, these conditions are satisfied. There is a large class of filters satisfying also these conditions. However the question of characterizing all the dynamical filters with poles inside the circle of singularity and introducing distinct values v_i^N remains open. As a matter of fact it is possible to extend this theorem to the case where these values are not distinct, but this introduces other conditions that cannot be presented in this paper.

When the SM is zero the only consequence is that there is no continuous component, or $a_1 = 0$. It is now interesting to discuss conditions ensuring that there is only a discrete component, or that $a_1 = a_3 = 0$. This is partially solved by the following theorem.

Theorem 3: Let X be the RV $\sum_{k=0}^{\infty} h_k w_k$, where h_k is the impulse response of a dynamical non-RIF filter $\mathcal{F}$ and w_k a sequence of Bernoulli RVs. If the poles of $\mathcal{F}$ are inside the circle of singularity and if there exists an N_0 such that for $N > N_0$

$$p_N[i, j] = 0 \text{ or } 1, \quad \forall i, j, \quad (18)$$

then the RV X is discrete, or $a_1 = a_3 = 0$.

Proof: Let s be the number of values of j for which $p_N(j) \neq 0$ at the step N_0 . Let $v_j^{N_0}$ be one of these values and $V_j^{N_0}$ be the corresponding node of the tree. It generates at the step $N_0 + 1$ two nodes $V_i^{N_0+1}$ characterized by the indices $i_+(j)$ and $i_-(j)$. It results from (12) and (18) that one of the two $p_{N_0+1}(i_{\pm})$ is $p_{N_0}(j)$ and the other 0. Then at the step $N_0 + 1$ there are also only s non-zero probabilities $p_{N_0+1}(i)$ equal to those of the step N_0 .

By repeating this till infinity, we deduce that the RV X takes only s values with the same probabilities as at the step N_0 , which means that it is discrete.

Let us present a very simple example of this situation. Consider the signal $w_k = w_0(-1)^k$, where w_0 is a symmetric Bernoulli RV, and the filter $h_k = a^k$. It is clear that $X = w_0[1/(1+a)]$. Then X is a discrete RV taking only the values $\pm 1/(1+a)$ with equal probabilities. It is obvious that (18) is satisfied. In this case $N_0 = 1$ and $s = 1$ at step 1. The unique paths going from V to infinity are characterized at step N by $v_{\pm}^N = 1 \pm [-a + a^2 + \dots + (-a)^N]$.

Finally let us now present a relation between these considerations and the prediction of signals. This comes from the fact that the transition probability (11) can be expressed in terms of the RVs w_k . Indeed for a given v_j^{N-1} , or a given node V_j^{N-1} , there is only one path coming from V to V_j^{N-1} , or one sequence of values of w_k , $0 < k \leq N-1$. Let us note S_{N-1} this sequence. The transition probability $p_N(i, j)$ is simply related to the conditional probabilities of the w_k by

$$\begin{aligned} p_N[i_-(j), j] &= P[w_N = -1 | S_{N-1}]; \\ p_N[i_+(j), j] &= P[w_N = 1 | S_{N-1}], \end{aligned} \quad (19)$$

where $i_-(j)$ and $i_+(j)$ are the two indices i defining the two nodes V_i^N generated by the node V_j^{N-1} .

This is related with the time prediction of the signal w_k . Indeed the prediction of w_k in terms of its past values is entirely described by the conditional probabilities appearing in (19). In particular if the signal is

predictable with a finite past P we have for $N > P$ either $p_N[i_-(j)] = 1$ or $p_N[i_+(j)] = 1$, which implies $p_N[i_+(j)] = 0$ or $p_N[i_-(j)] = 0$ respectively.

Theorem 4: Let X be the RV $\sum_{k=0}^{\infty} h_k w_k$, where h_k is the impulse response of a dynamical non-RIF filter $\mathcal{F}$ and X_k a sequence of Bernoulli RVs. If the poles of $\mathcal{F}$ are inside the circle of singularity and if w_k is predictable with a finite past P , then X is discrete, or $a_1 = a_3 = 0$.

Proof: The assumption of predictability means that, for all $k > P$, w_k satisfies

$$w_k = f(w_1, w_2, \dots, w_P). \quad (20)$$

Since the RVs w_i take only 2 values, X of (5) takes at the maximum 2^P distinct positive values which means that it is a discrete RV.

6. Singularity and Markovian inputs

A symmetric Bernoulli Markovian signal of order P is a signal taking the values ± 1 with equal probabilities and which can be expressed as

$$w_k = f(w_{k-1}, w_{k-2}, \dots, w_{k-P}, \mathbf{b}_k), \quad (21)$$

where $\mathbf{b}_k$ is a vectorial white noise. Its value at time k conditionally to the whole past depends only on $w_{k-1}, w_{k-2}, \dots, w_{k-P}$. The function $f(\cdot)$ and the noise $\mathbf{b}_k$ must satisfy conditions ensuring that if the w_k s are signals taking the values ± 1 with equal probabilities, w_k has the same property.

The transition probabilities $p_N(i, j)$ are given by (19). The Markov assumption means that the sequence S_{N-1} can be replaced by a sequence using only the past of order P . This can be written

$$p_N[i_{\pm}(j), j] = P[w_N = \pm 1 | w_{N-1}, \dots, w_{N-P}]. \quad (22)$$

Since the RVs w_k take only the values ± 1 , there are only at most 2^P distinct values of $p_N[i_{\pm}(j), j]$. If we assume that the signal is not predictable there is no value of p_N equal to 1. Then the bound of (17) is simply the maximum value of the finite number of transition probabilities $p_N(i, j)$. Then Theorem 1 can be applied and the RV X is singular.

In order to simplify the discussion and the experiments we shall consider only linear Markovian signals of order 1 defined as follows. Let u_k be a strictly white noise taking only the values 0 or 1 with probabilities $1-p$ and p respectively. Similarly let v_k be a strictly white noise taking the values ± 1 with the same probabilities $1/2$ and independent of u_k . Consider now the signal

$$w_k = u_k w_{k-1} + \bar{u}_k v_k, \quad (23)$$

where $\bar{u}_k = 1 - u_k$. It has the form (21) where $P=1$ and $\mathbf{b}_k = [u_k, v_k]$.

It is obvious that if w_{k-1} takes only the values ± 1 with probabilities $1/2$, w_k has also the same property. It can be shown that, whatever the values of w_1 , w_k tends to have this property. It can also be shown that the covariance function γ_k of w_k is $p^{|k|}$, so it is an exponential covariance function. For $p=0$, $u_k=0$, and w_k is the white noise v_k with a zero covariance function. On the other hand if $p=1$, $u_k=1$ and $w_k=w_1$, which introduces constant covariance function or long range memory.

Let us now consider the signal x_k defined by the recursion $x_k = ax_{k-1} + w_k$, where w_k is given by (23). Some of its main properties appear in figures 8 and 9. The results of figure 8 are obtained by the same procedure as for figure 4 with the same value of the regression coefficient a and the same value of the number N of cells in each histogram. Note that histograms 1 and 2 of figure 8 are the same except a different scale of the y axis.

The first point to note is that there is no discrete component, at least within the precision of our measurements. This means that, as expected, x_k is singular.

The structures of histograms at various scales are the same as when w_k is white and they indicate a fractal

property which is the origin of the singularity of x_k . But there is a clear difference between these two series of histograms. In the case where the input signal w_k is white, the amplitudes, or numbers of samples recorded in each cell of the histograms, are constant. This is no longer true when w_k is a Markovian signal. Finally note that these amplitudes are maximum at the extreme values of the histograms. Even if the calculation of the DF of x_k is very complicated, the origin of this maximum can be explained (Picinbono 2006).

Note that when the input is white the histograms are symmetric with respect to their centres. This is no longer true when Markovian inputs are used, except for the histogram centered at 0 because X is symmetric and its histograms are symmetric with respect to 0. This explains that only histograms of positive values of x_k are represented.

The influence of the parameter p of the model is shown in figure 9 in the case where $a=1/2$. When $p=0$ the input w_k is white and we find again the results of histogram 4 of figure 3. The uniform distribution is strongly modified for non-zero values of p and this also can be explained (Picinbono 2006).

All the previous results can be extended to higher order linear Markovian inputs. For this purpose it suffices to replace $u_k w_{k-1}$ of (23) by a term depending on $w_{k-1}, w_{k-2}, \dots, w_{k-m}$. It is easy to

Figure 8. Histograms of x_k at different scales: $a=1/3$, $p=1/2$, $S=10^7$.

Figure 9. Influence of p for $a=1/2$: 1. $p=0$; 2. $p=0.1$; 3. $p=0.2$; 4. $p=0.4$; $S=10^7$.

construct models describing this situation. The calculations however are much more complicated and cannot be presented here.

7. Mixture of discrete and singular components

Consider the signal $u_k = u_0(-1)^k$ introduced above and the white signal v_k taking the values ± 1 with equal probabilities. Suppose now that w_k is equal either to u_k or to v_k with the probabilities α and $1 - \alpha$ respectively. This implies that its covariance function is $\gamma_k = \alpha(-1)^k + (1 - \alpha)\delta[k]$. The DF of the RVs w_k is $F(x) = \alpha F_d(x) + (1 - \alpha)F_s(x)$. Applying this signal again as the input to the exponential causal filter with impulse response $h_k = a^k$ with $a < 1/2$ yields an output signal taking the values $\pm 1/(1 + a)$ with probability α and being singular with probability $1 - \alpha$.

Results of computer experiments on this signal are presented in figures 10 and 11. In these experiments $a = 1/3$, as in figure 4, and $\alpha = 0.5$. This implies that x_k takes the values ± 0.75 with probability 0.25, which means that its DF is discontinuous for $x = \pm 0.75$ and the amplitude of the discontinuity is 0.25. The first histogram of figure 10 exhibits these two discontinuities characterized by two lines at the points ± 0.75 . One of these lines disappears in the second histogram of this figure and there is no longer a line in the last two histograms. The fractal structure introducing a singular component appears clearly in these histograms. In order to verify whether

Figure 10. Mixture discrete-singular, $\alpha = 0.5$. Fractal part of the histogram, $S = 10^7$.

or not the discontinuity of the DF is an experimental artefact, we present in figure 11 four histograms isolating the point 0.75 and with decreasing widths of the cells. We observe still a residual contribution of the singular part of the DF in the first histogram. There is however a line at 0.75 and the fact that its amplitude is constant indicates clearly that there is effectively a discrete component. As the experiment uses 10^6 samples, the discontinuity corresponding to the probability 0.25 must be equal to $2.5 \cdot 10^5$, which clearly appears in the four histograms of figure 11. This shows that the signal x_k is effectively a mixture of a discrete-valued signal and a singular signal.

8. Conclusion

When the input of a causal discrete-time exponential linear filter with impulse response a^k is a Bernoulli white noise, the output is singular for $a < 1/2$. A simple proof of this result was presented and it exhibits two steps. The condition on the parameter a and the fact that the input is a discrete-valued signal implies that the SM of the output is zero, which

Figure 11. Mixture discrete-singular, $\alpha = 0.5$. Analysis of the neighbourhood of 0.75, $S = 10^6$.

means that there is no continuous component in its DF. The fact that this DF does not contain a discrete component arises from the whiteness of the input. This result is not specific to exponential filters. We have shown that singularity can appear in many other situations and we have established a sufficient condition for singularity by using the positions of the poles of the filter with respect to the circle of singularity and the whiteness of the input.

This assumption of whiteness can however be partially relaxed without changing the singularity of the output. Some sufficient conditions ensuring the singularity with coloured inputs have been established. These conditions are obviously satisfied not only by white noise but also by a large class of correlated signals. It is especially the case of some Markovian signals of finite order. The theoretical analysis also shows that the output generated by coloured inputs can be a mixture of a discrete and a singular distribution. Computer experiments in order to verify the theoretical results have been realized and discussed. For this purpose a specific model of linear Markovian signal of order one was introduced and the experimental results are in complete agreement with the theory. Finally a model of Bernoulli input ensuring that the output contains a discrete and a singular part was introduced and here also the experimental results are in perfect agreement with the theory.

References

- L. Campbell, A. McKellips and P. Wittke, "Distributions and expectations of singular random variables", *J. Appl. Prob.*, 32, pp. 1032–1040, 1995.
- K. Falconer, *Fractal Geometry Mathematical Foundations and Applications*, Chichester: John Wiley and Sons, 1990.
- A. Garsia, "Entropy and singularity of infinite convolutions", *Pacific J. Math.*, 13, pp. 1159–1169, 1962.
- C.W. Helstrom, *Probability and Stochastic Processes for Engineers*, New-York: Macmillan, 1991.
- F.S. Hill and M.A. Blanco, "Random geometric series and intersymbol interference", *IEEE Trans. Inform. Theory*, IT 19, pp. 326–335, 1973.
- R. Kershner and A. Wintner, "On symmetric Bernoulli convolutions", *Amer. J. Math.*, 57, pp. 541–545, 1935.
- M. Loève, *Probability Theory*, New York: Springer Verlag, 1977.
- E. Lukacs, *Characteristic Functions*, London, UK: Griffin, 1970.
- M. Naraghi-Pour and D. Neuhoff, "On the continuity of the stationary state distribution of DPCM", *IEEE Trans. Inform. Theory*, IT-36, pp. 305–311, 1990.
- M.K. Ochi, *Applied Probability and Stochastic Processes*, New York: Wiley, 1990.
- A. Papoulis, *Probability, Random Variables and Stochastic Processes*, New York: McGraw Hill, 1984.
- Y. Peres and B. Solomyak, "Self similar measures and intersections of Cantor sets", *Trans. Amer. Math. Soci.*, 350, pp. 4065–4087, 1998.
- P.E. Pfeifer, *Probability for Applications*, New York: Springer Verlag, 1990.
- B. Picinbono, *Random Signals and Systems*, Englewood Cliffs: Prentice Hall, 1993.
- B. Picinbono, Correlation and singularity, L2S internal report, 2006.
- B. Picinbono and J.-Y. Tournet, "Singular random signals", *IEEE Trans. Sig. Proc.*, 53, pp. 499–504, 2005.
- W. Smith, P. Wittke and L. Campbell, "Error probabilities on fading channels with intersymbol interference and noise", *IEEE Trans. Inform. Theory*, 39, pp. 1598–1607, 1993.
- B. Solomyak, Y. Peres and W. Schlag, "Sixty years of Bernoulli convolutions, fractal and stochastics II", in *Progress in Probability*, C. Bandt, S. Graf and M. Zahle, Eds, Vol. 46. New York: Birkhäuser, 2000, pp. 39–65.
- J.-Y. Tournet and B. Lacaze, "Étude et simulation des lois de probabilité de sorties de modèles paramétriques", *Traitement du signal*, 11, pp. 271–281, 1994.
- P. Wittke, W. Smith and L. Campbell, "Infinite series of interference variables with Cantor-type distributions", *IEEE Trans. Inform. Theory*, 34, pp. 1428–1436, 1988.
- E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*, New York: Springer Verlag, 1985.