Mandelbrot's fractional renewal model for 1/f noise: weak ergodicity breaking in the 1960s ?

Nick Watkins

NickWatkins@mykolab.com

The Open University

CATS CENTRE FOR THE ANALYSIS OF TIME SERIES

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

WARWICK THE UNIVERSITY OF WARWICK

Thanks

- Rainer for inviting me to participate in the ASG
- ONR for funding my participation in the ASG
- PKS for supporting my visiting position in 2013-2014, and visits to ASG in Dresden
- Holger for hosting me in his group at PKS in 2013-14
- Ralf Metzler for hosting me at Potsdam during ASG

Discussions with many people including the above, Sandra Chapman, Eli Barkai, Daniela Froemberg, Igor Sokolov, Igor Goychuk, Aleksei Chechkin, Martin Gerlach, the PKS complexity reading group 2013-15, and all of the 2015 ASG team and its visitors.

ROADMAP

Storytelling

- This story mixes physics and history
- Can lead with physics (my PKS talk in summer 2014-see my Slideshare page)
- Can lead with history (my talk from January this year-ditto)
- Today will try and balance the two
- Note: In interests of time have dropped the “panorama” mentioned in abstract

1/f noise : why it matters,
Why it's puzzling

$$S(f) \sim f^{-\beta}$$

Many solutions (or one) ?

Ergodic route

- Non-ergodic route

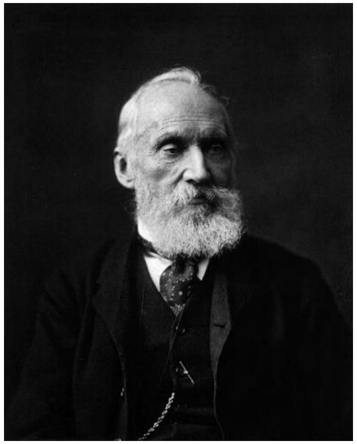
Another: Long tailed waiting times between switching-fractional renewal process/continuous time random walk. Motivated by physics of weak ergodicity breaking.

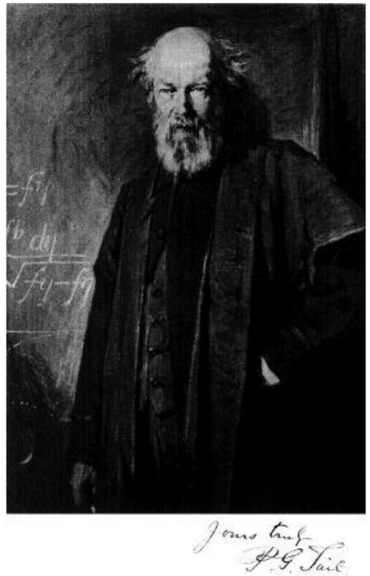
One solution: Long range dependent kernel, non-Markovian, fractional Gaussian noise. Physics <u>now</u> known to be the generalised Langevin equation, fGn is its noise term

$$S(f) \sim Q(T)\, f^{-\beta}$$

$$S(f) \sim f^{-\beta}$$

Formula versus fact





*"Nothing can be more fatal to progress than a too confident reliance on mathematical symbols; for the student is only too apt to … **consider the formula and not the fact as the physical reality".*** ***Thomson (Kelvin) & Tait, 1890 edition.***

"*Like the ear,* ***the eye is very sensitive to features that the spectrum does not reflect****. Seen side by side, different 1/f noises, Gaussian [i.e. fGn], dustborne [i.e. fractional renewal] and multifractal,* ***obviously*** *differ from one another"*- **Mandelbrot, Selecta N, 1999*.***

THE ENIGMA OF “1/F”

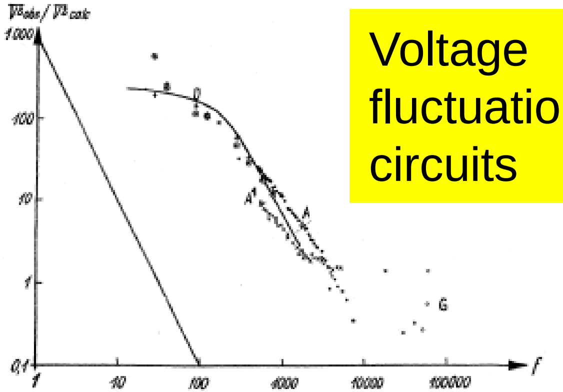
“1/f” spectra

- 1/f noise-fractals in time
- One explanation (c.f. Bak’s SOC) or several ?
- SOC links 1/f in time to spatial correlation functions in “avalanching” critical systems
- But spectra and correlation functions based on ergodicity-“misbehave” on non ergodic 1/f signals
- Will later show you that that Mandelbrot made this last point in 1965-67, but mainly to engineers !

1/f noise – fractals in time

"There is …[a] … ubiquitous phenomenon which has defied explanation for decades. … a power spectrum decaying with an exponent near unity at low frequencies .... This type of behavior is known as "1/ f" noise, or flicker noise."

It was the dependence of the observed effect upon frequency which suggested to Johnson more definite conclusions regarding the origin of the spontaneous perturbation in the emission from the surface. This frequency variation is shown, to the same logarithmic scale as

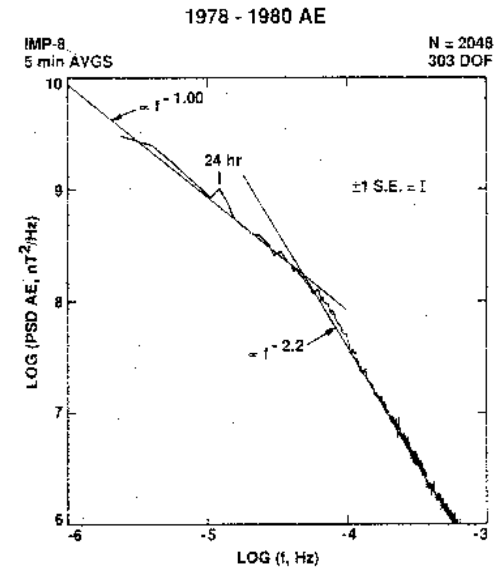


Voltage fluctuations in circuits

Fig. 4.

before, in Figs. 3 and 4, taken from Johnson's paper. Fig. 3 corresponds

Schottky, 1926



Tsurutani et al, GRL, 1991

Fluctuating magnetic field in ionosphere

Why is 1/f puzzling ?

If spectral density $S'(f) \sim f^{-\beta}$ then i) it is singular as $f \to 0$

and ii) if we define an acf via time average $\rho(\tau) = \overline{x(t)x(t+\tau}$

and use Wiener-Khinchine theorem to get ρ from Fourier transform of $S'(f)$

then ρ falls off as power law, and its summed lags "blow up" $\sum_{\tau} \rho(\tau) \to \infty$

Why is 1/f puzzling ?

If spectral density $S'(f) \sim f^{-\beta}$ then i) it is singular as $f \to 0$

- So two odd features
- Infrared catastrophe in the psd
- Highly non-Markovian autocorrelation behaviour

The cutoffs that would be needed to tame these are often not seen

1/f: a single origin ?

"The importance of Mandelbrot's discovery that fractals occur widespread in nature can hardly be exaggerated. Many things which we used to think of as messy and structureless are in fact characterized by well-defined power-law spatial correlation functions. … is it possible to construct a dynamical theory of the physics of fractals?"

"There is another ubiquitous phenomenon which has defied explanation for decades. … a power spectrum decaying with an exponent near unity at low frequencies .... This type of behavior is known as "1/ f" noise, or flicker noise."

*"We believe that those two phenomena are **often** two sides of the same coin: they are the spatial and temporal manifestations of a self-organized critical state."*

-Bak and Chen, "The physics of fractals" Physica D, 38(1-3), 1989.

Or several … ?

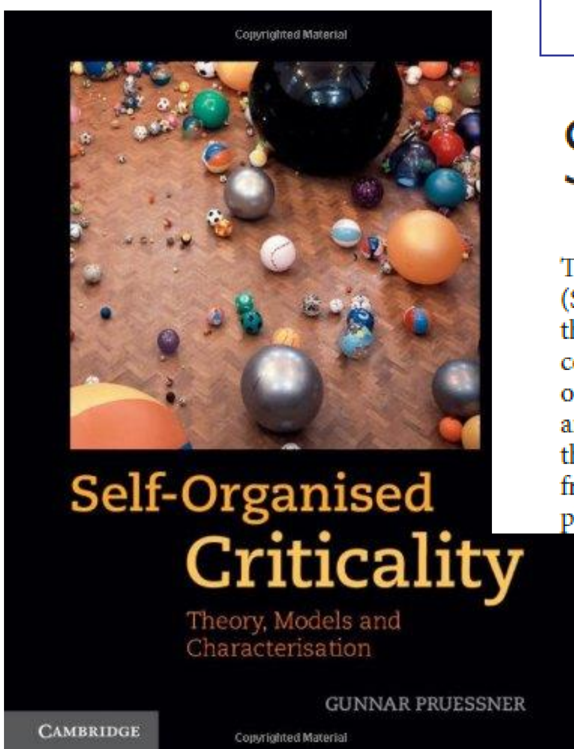
As a matter of fact, I started with one, a very simplified view of the 1/f noise, which was lucky, because had I seen the whole monster I would have been totally overwhelmed … But very soon I realised that … in fact you could have a spectrum, 1/f , while being of very many different kinds, … a discovery that at the same time is ... a shallow observation and at the same time also very profound, because 1/f is a formula. <u>The same formula can be used as caption to all kinds of different phenomena</u> – Benoit Mandelbrot, Web of Stories, 1998.

Space Sci Rev
DOI 10.1007/s11214-015-0155-x

25 Years of Self-organized Criticality: Concepts and Controversies

Nicholas W. Watkins[1,2,3,4] · Gunnar Pruessner[5] · Sandra C. Chapman[4,6] · Norma B. Crosby[7] · Henrik J. Jensen[5]

Received: 24 January 2015 / Accepted: 4 May 2015



to any of the simple models known to exhibit SOC. For example, schematic models for the geometric spread of forest fires are essentially isomorphic to some of the archetypal SOC models. The SOC interpretation of power law distributions for such fires therefore carries more weight.

thesis

SOC revisited

The idea of self-organized criticality (SOC) has inspired physicists for more than 25 years. It was born in 1987 as a conjecture about the dynamics of a system of coupled pendulums. Per Bak, Chao Tang and Kurt Wiesenfeld were motivated by the widespread existence of scale-invariant fractal structures, both in space and time, in physical and biological phenomena. Their

Watkins and colleagues quote a wonderful sentence from Philip Anderson, which perhaps best captures how the SOC idea should be seen. It's not quite a specific theory, as one might consider general relativity or the kinetic theory of gases. Rather, it's a more generally scheme for thinking about many systems out of equilibrium, which may come close to capturing many of their essential elements. "Self-organized criticality seems to me to be," as Anderson put it, "not the right and unique solution to these and other similar problems, but to have paradigmatic value, as the kind of generalization which will characterize the next stage of physics."

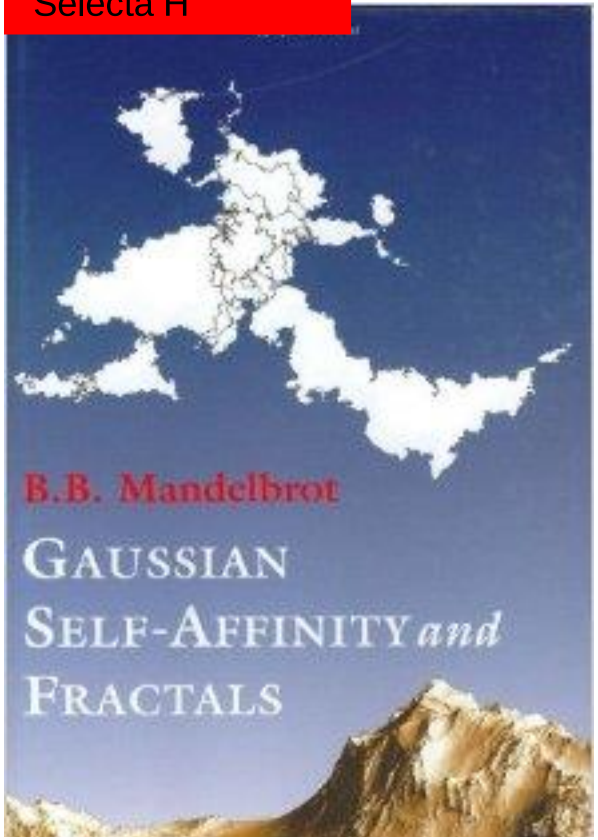
That seems about right. At times, no doubt, the importance of the idea has been overstated. Much of the controversy and confusion has been engendered by some sloppy thinking and over-enthusiasm by SOC supporters. But it remains an idea of enduring value.

MARK BUCHANAN

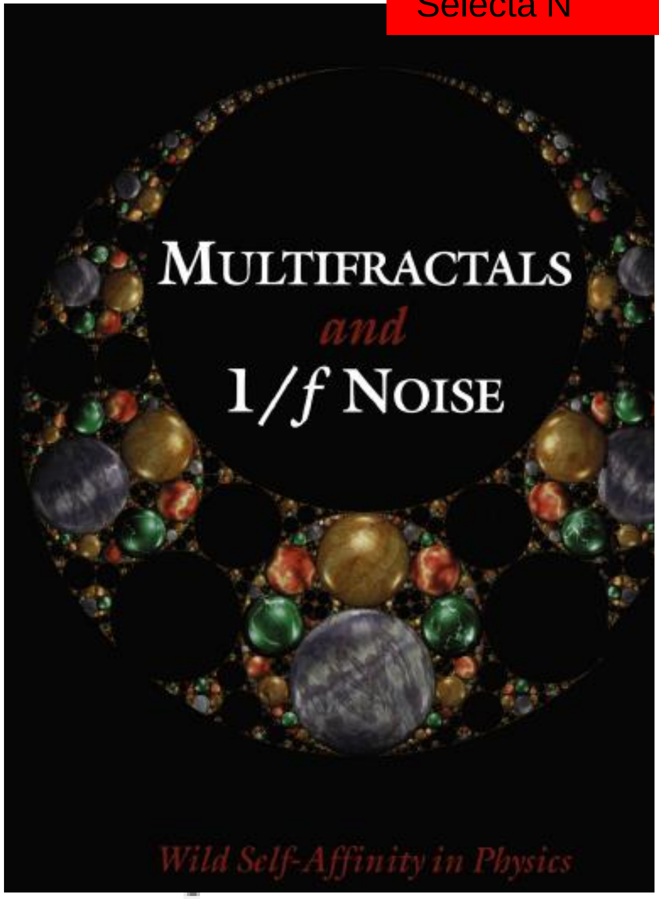
NATURE PHYSICS | VOL 11 | JUNE 2015 | www.nature.com/naturephysics

Models for "1/f"

Selecta H



Selecta N



My models of both telephone errors and Nile floods involved spectra of the form f^{-B}. Despite this common property, those processes were of totally different character. That is, a common spectrum did not imply any deeper commonality.

1/f noise : why it matters,
Why it's puzzling

$$S(f) \sim f^{-\beta}$$

Many solutions (or one) ?

Ergodic route

- Non-ergodic route

Another: Long tailed waiting times between switching-fractional renewal process/continuous time random walk. Motivated by physics of weak ergodicity breaking.

$$S(f) \sim \mathrm{Q(T)}\, f^{-\beta}$$

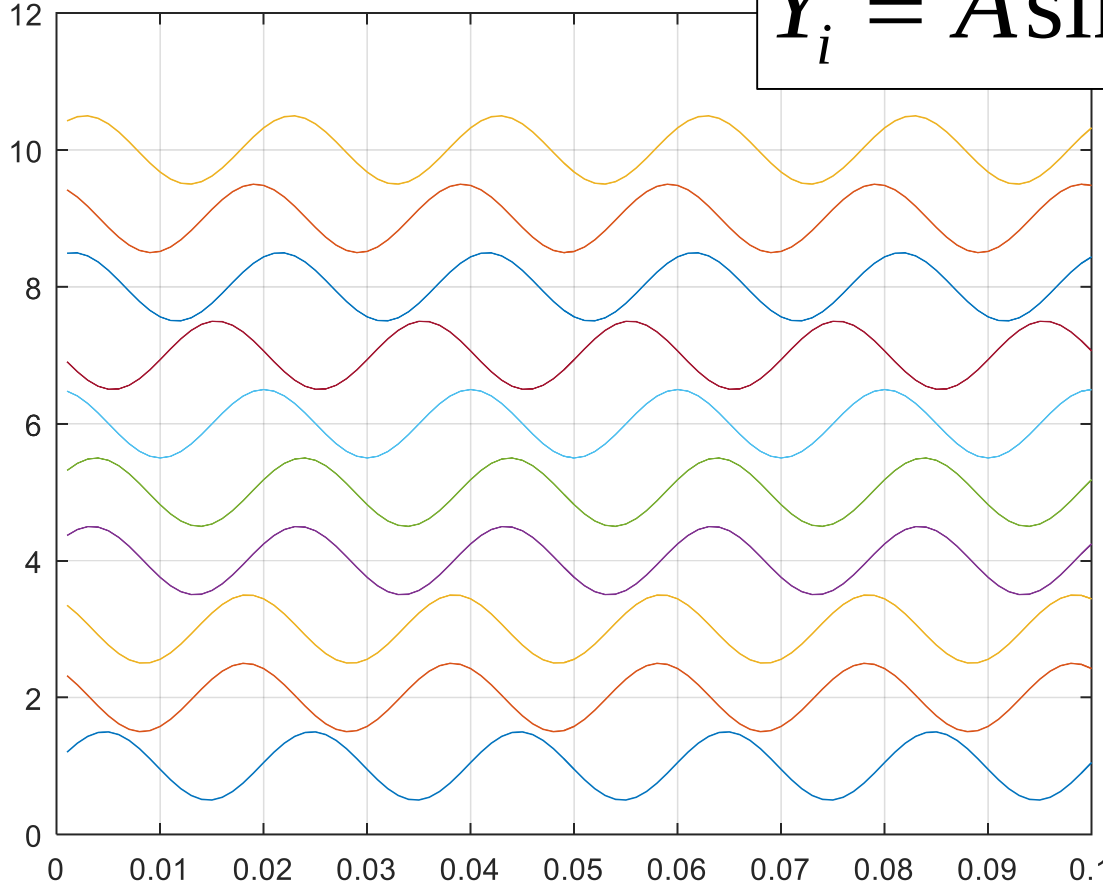
One solution: Long range dependent kernel, non-Markovian, fractional Gaussian noise. Physics <u>now</u> known to be the generalised Langevin equation, fGn is its noise term

$$S(f) \sim f^{-\beta}$$

ERGODICITY BREAKING

Ergodic

$$Y_i = A\sin(\omega t + \phi_i)$$



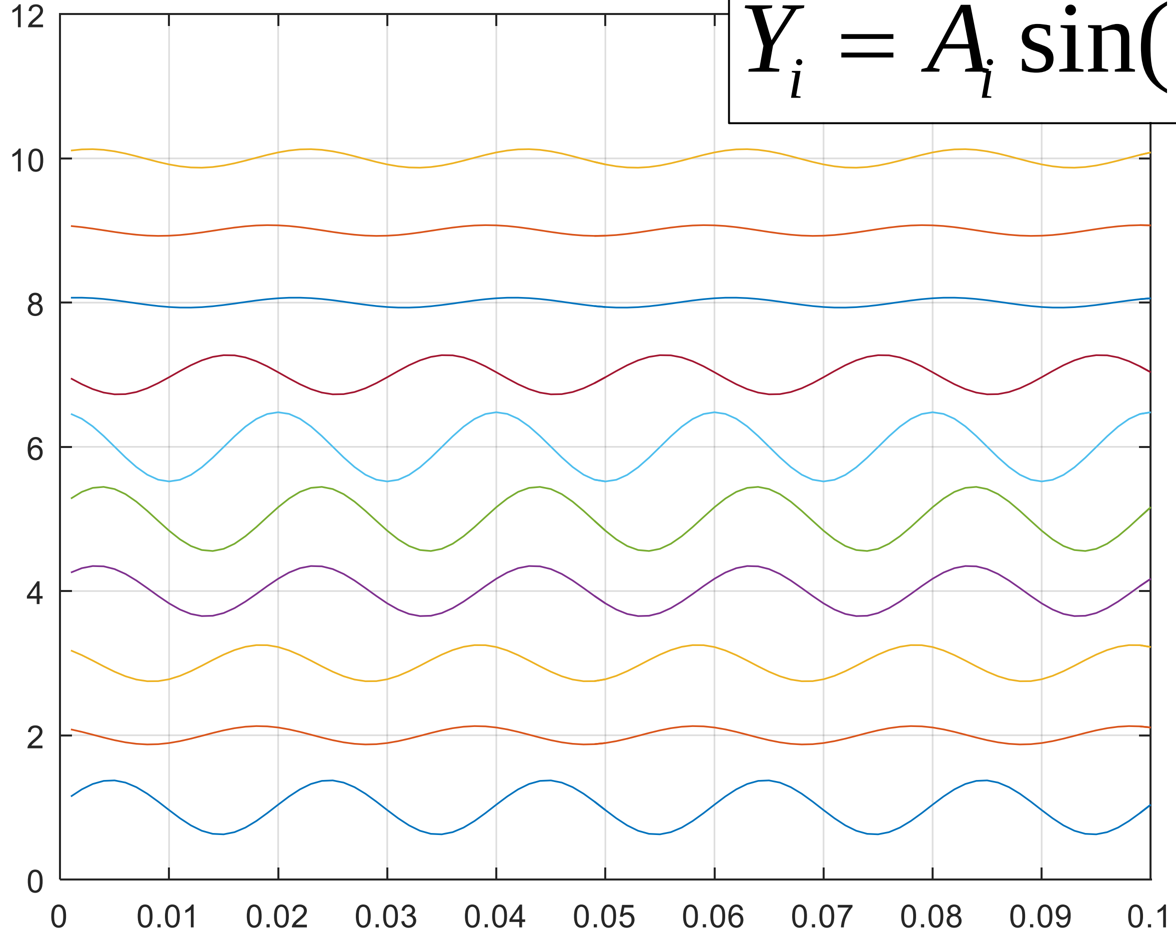
$$\overline{Y_i^2} \propto A^2$$

$$< Y_i^2 > \propto A^2$$

After **Bendat & Piersol, "Random data"**

Non-Ergodic

$$Y_i = A_i \sin(\omega t + \phi_i)$$



$$\overline{Y_i^2} \propto A_i^2$$

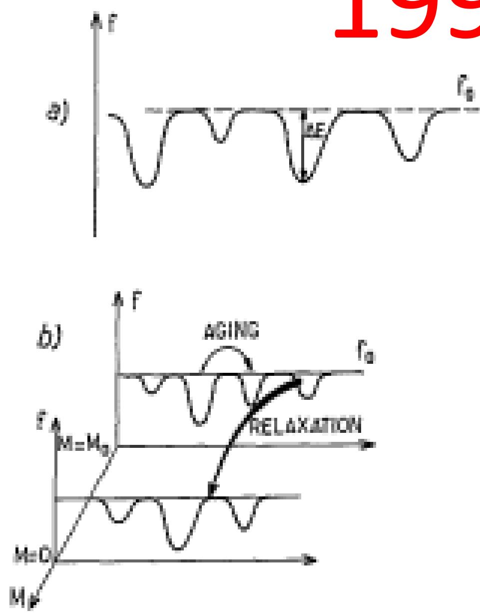
$$< Y_i^2 > \propto \sum_i A_i^2$$

After **Bendat & Piersol, “Random data”**

Non-ergodicity matters

- Physics can be non-ergodic
- Weak ergodicity breaking [Bouchaud,1992]
- Single particle tracking
- Blinking quantum dots
- Also economics ?
- St Petersburg paradox [Bernoulli, 1713]

(Weak) ergodicity breaking [Bouchaud 1992]



The most interesting property of these distribution laws is the fact that $\langle \tau \rangle = \int_0^\infty \mathrm{d}\tau \, \tau \psi(\tau)$ diverges, when $x \leqslant 1$ for equation (2) and for all finite ζ for equation (4).

Hence, the time needed to explore an infinite system is infinite. This is however a non-conventional scenario for ergodicity breaking, since the phase space is *a priori* not broken into mutually inaccessible regions in which local equilibrium may be achieved. We shall call this situation « weak » ergodicity breaking. It may be that « true » ergodicity breaking also occurs, i.e. the existence of many « pure states » between which infinite barrier stand. We shall carefully avoid this issue (the reader is referred to [19] for a recent experimental discussion), since the dynamics is by definition restricted to only one pure state, and for which the concept of « weak » ergodicity breaking is relevant.

Single particle tracking

Physics

Physics **1**, 8 (2008)

Viewpoint

Statistics and the single molecule

By Igor M. Sokolov
Institut für Physik, Humboldt-Universität zu Berlin, Newtonstrasse 15, D-12489 Berlin, Germany

Published July 28, 2008

One hundred years ago, the atomic-molecular theory of matter was having a hard time, and many physicists considered it merely a kind of convenient shorthand rather than a real description of nature; after all, nobody had really seen a molecule, let alone an atom. Today, developments in micromanipulation and in single-molecule tracking have not only made individual molecules visible, but have led to real breakthroughs in understanding of the molecular basis of life. This ability to follow and to manipulate single molecules has opened new perspectives in nanoscience and nanotechnology. Experts in single-molecule tracking often say that observation of individual trajectories gives more information about the system than only looking at ensemble averages, which is the approach taken in statistical thermodynamics. The idea is that the closer one looks, the more information one can get.

tems that rapidly relax to equilibrium or to a stationary state, implying the system is ergodic. Systems far from equilibrium or showing very slow relaxation may be nonergodic, and subdiffusion as modeled by CTRW may be one of the simplest theoretical examples. One has to be cautious when applying our intuition gained for the close-to-equilibrium cases to such processes: the information contained in the time-averaged and ensemble-averaged results is different and is pertinent to different aspects of the system's behavior. Understanding this fact is necessary when interpreting the results of existing experiments and when planning future studies.

[Physics Today,2008]

feature article

Beyond quantum jumps: Blinking nano-scale light emitters

Fernando D. Stefani, Jacob P. Hoogenboom, and Eli Barkai

On the nanoscale, almost all light sources blink. Surprisingly, such blinking occurs on time scales much larger than predicted by quantum mechanics and has statistics governed by nonergodicity.

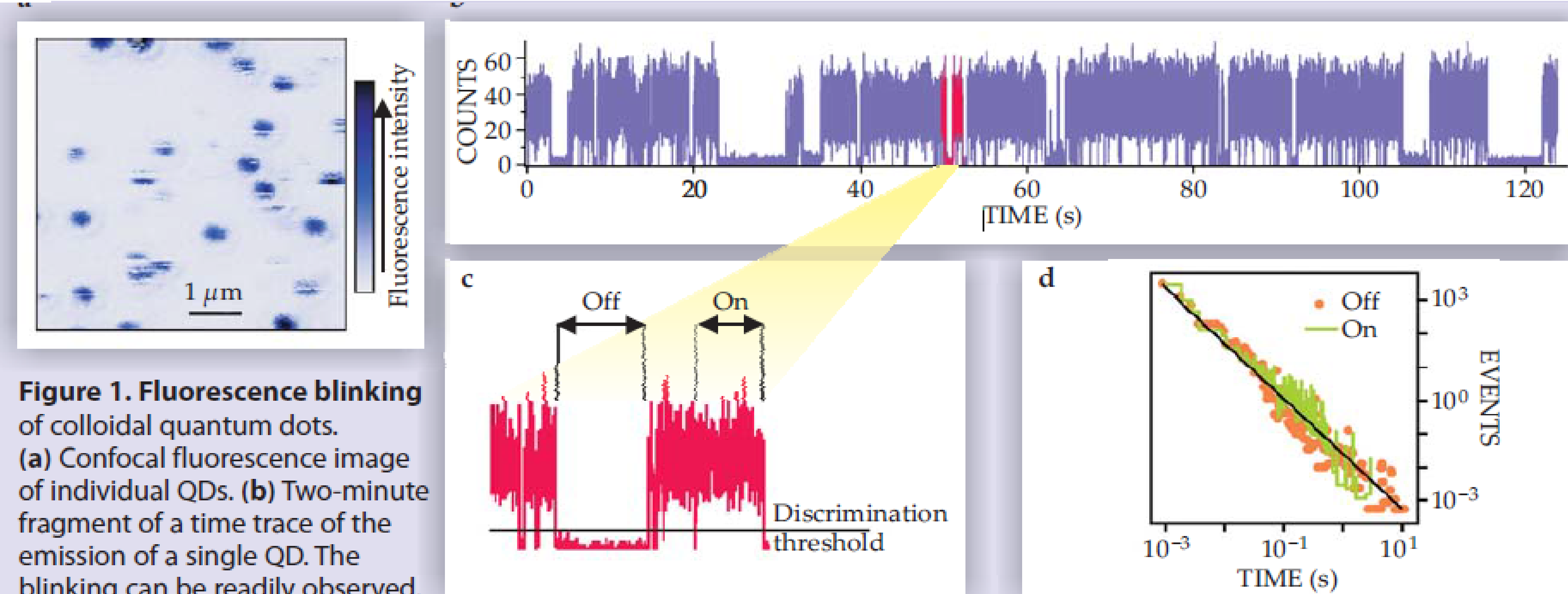


Figure 1. Fluorescence blinking of colloidal quantum dots. (**a**) Confocal fluorescence image of individual QDs. (**b**) Two-minute fragment of a time trace of the emission of a single QD. The blinking can be readily observed. (**c**) A two-second zoom-in. The horizontal line marks the threshold used to discern the off state from the on state. (**d**) The distribution of times in the on and off states. The line is a power law (equation 1) with an exponent $\alpha = 1.65$.

St Petersburg Paradox (Peters,2011)

The St Petersburg paradox was first put forward by Nicolaus Bernoulli in 1713 [13, p. 402]. He considered lotteries of the following type:

A fair coin is tossed.

(1) On heads, the lottery pays \$1, and the game ends. On tails, the coin is tossed again.
(2) On heads, the lottery pays \$2, and the game ends. On tails, the coin is tossed again.

⋮

(n) On heads, the lottery pays $\$2^{n-1}$, and the game ends. On tails, the coin is tossed again.

⋮

In other words, the random number of coin tosses, n, follows a geometric distribution with parameter 1/2, and the payouts increase exponentially with n.

We may call n a 'waiting time', although in this study it is assumed that the lottery is performed instantaneously, i.e. a geometric random variable is drawn and no significant physical time elapses. The expected payout from this game is

$$\$\sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^{n}2^{n-1}=\$\left(\frac{1}{2}+\frac{1}{2}+\cdots\right), \quad (3.1)$$

which is a diverging sum. A rational person, N. Bernoulli argued, should therefore be willing to pay any price for a ticket in this lottery. In reality, however, people are rarely willing to pay more than \$10, which constitutes the paradox.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A

Phil. Trans. R. Soc. A (2011) **369**, 4913–4931
doi:10.1098/rsta.2011.0065

The time resolution of the St Petersburg paradox

BY OLE PETERS[1,2,3,*]

[1]*Department of Mathematics and Grantham Institute for Climate Change, Imperial College London, London SW7 2AZ, UK*
[2]*Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, 7127 Math Science Building, 405 Hilgard Avenue, Los Angeles, CA 90095-1565, USA*
[3]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

A resolution of the St Petersburg paradox is presented. In contrast to the standard resolution, utility is not required. Instead, the time-average performance of the lottery is computed. The final result can be phrased mathematically identically to Daniel Bernoulli's resolution, which uses logarithmic utility, but is derived using a conceptually different argument. The advantage of the time resolution is the elimination of arbitrary utility functions.

Thus, the St Petersburg paradox relies for its existence on the assumption that the expected gain (or growth factor or exponential growth rate) is the relevant quantity for an individual deciding whether to take part in the lottery. This assumption can be shown to be implausible by carefully analysing the physical meaning of the ensemble average. A quantity that is more directly relevant to the financial well-being of an individual is the growth of an investment over time. Utility, which can obscure risks, is not necessary to evaluate the situation and resolve the paradox. It is the actual wealth, in \$, of a player, not the utility, that grows with $\bar{g}$ (equation (6.10)). It is manifestly not true that the commonly used ensemble-average performance of the lottery equals the time-average performance. In this sense, the system is not ergodic, and statements

Candidate models

- Fractional renewal process (e.g. Lowen & Teich,1993, who always use cutoffs !)
- Continuous Time Random Walk (Montroll &Weiss,1965)
- Renewal reward process
- Heavy tailed random telegraph (Niemann et al, 2013)
- Common features
- Discrete states
- Heavy tailed switching time distribution

Alternating Fractional Renewal Process (AFRP)

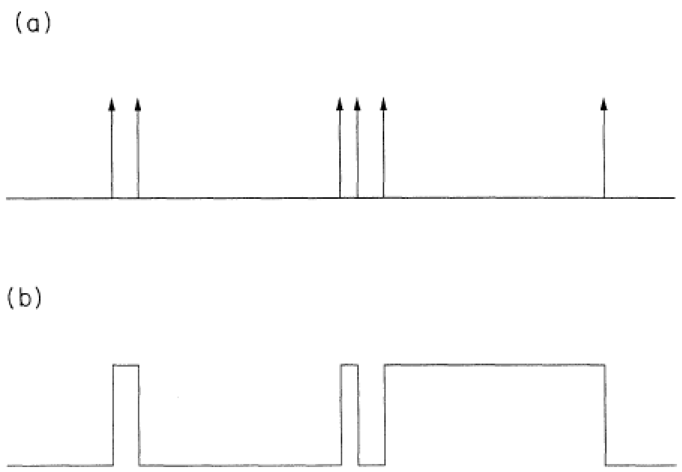


FIG. 1. Sample functions of fractal renewal processes. Interevent times are power-law distributed. (a) The standard fractal renewal process (SFRP) consists of Dirac δ functions and is zero-valued elsewhere. (b) The alternating fractal renewal process (AFRP) switches between values of zero and unity. The symmetric case is shown here.

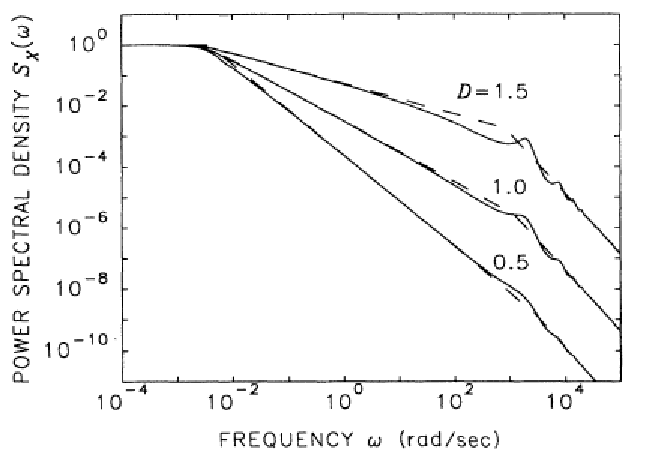
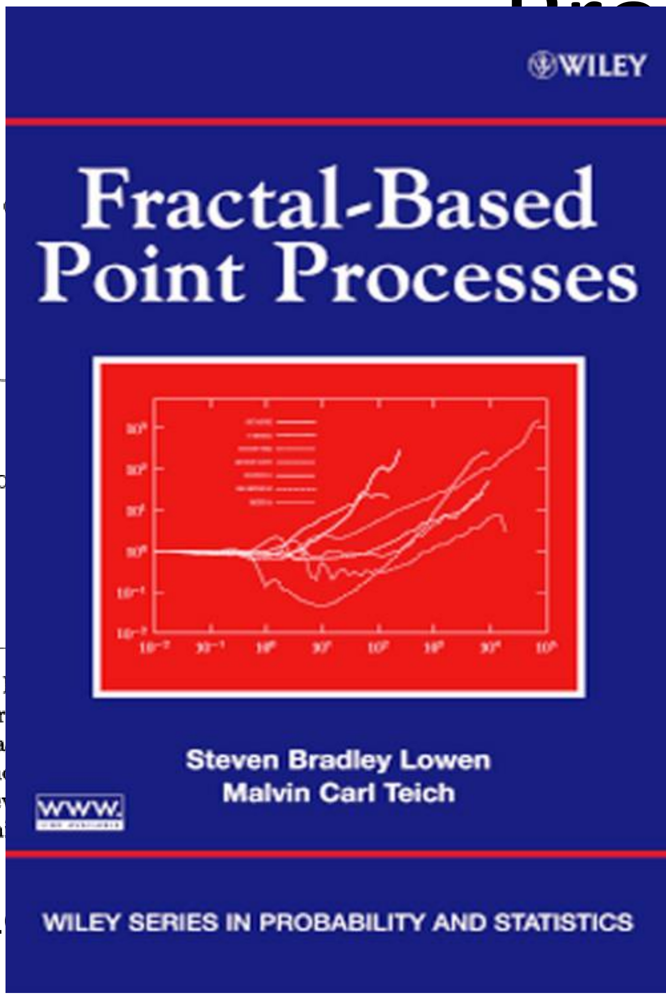


FIG. 6. Double logarithmic plot of the normalized power spectral density for the AFRP, with an abrupt-cutoff power-law probability density. Three values of the exponent D are shown: 0.5, 1.0, and 1.5 ($A = 10^{-3}$, $B = 10^{3}$). Asymptotic forms from Eqs. (8), (9), and (29) are included for comparison. The abrupt cutoff in the interevent-time probability density function gives rise to oscillations in the frequency domain, especially for larger values of D.

Lowen and Teich, PRE, 1993

Note cutoffs on inter event time

Alternating Fractional Renewal Process (AFRP)



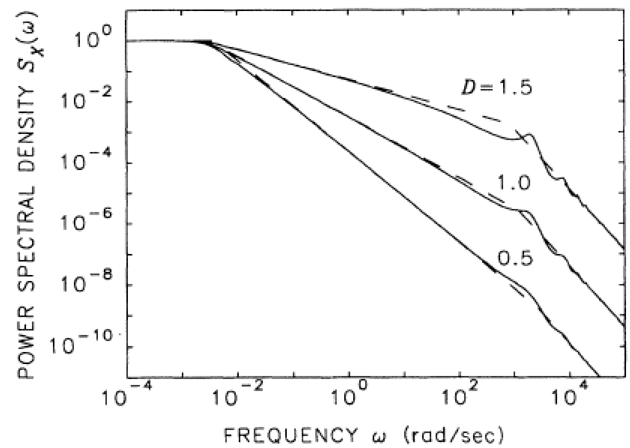
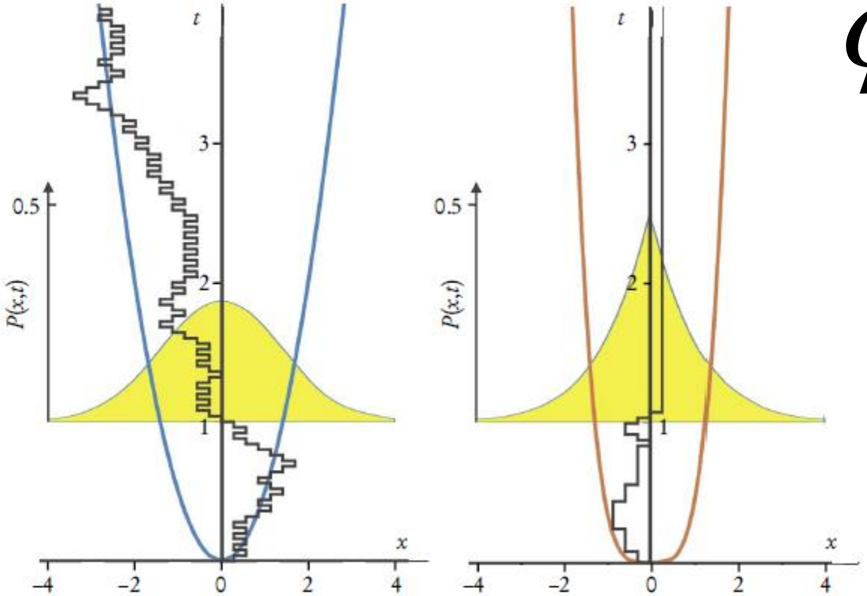


FIG. 6. Double logarithmic plot of the normalized power spectral density for the AFRP, with an abrupt-cutoff power-law probability density. Three values of the exponent D are shown: 0.5, 1.0, and 1.5 ($A = 10^{-3}$, $B = 10^{3}$). Asymptotic forms from Eqs. (8), (9), and (29) are included for comparison. The abrupt cutoff in the interevent-time probability density function gives rise to oscillations in the frequency domain, especially for larger values of D.

Note cutoffs on inter event time

Continuous Time Random Walk (CTRW)



$$\varphi(\xi,\tau)=\lambda(\xi)\psi(\tau)$$

$$X(t)=\sum_{i=1}^{n}\xi_i$$

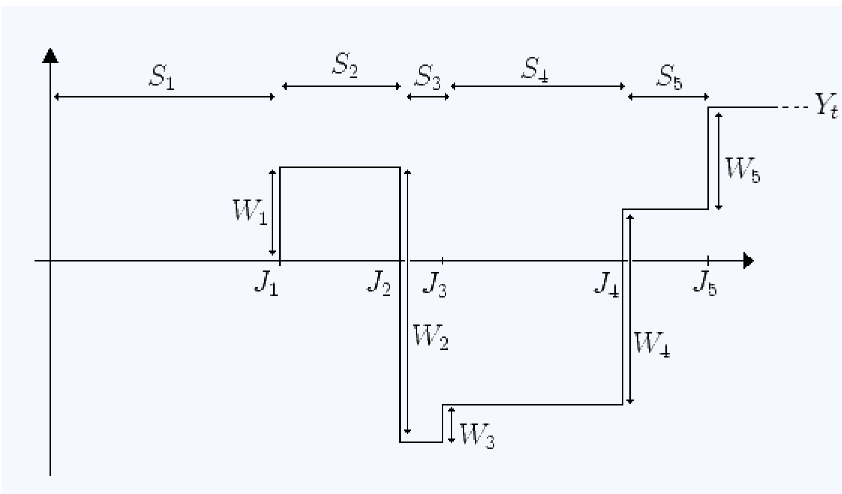
$$t_n=\sum_{i=1}^{n}\tau_i$$

Montroll and Weiss 1965

Figure from Sokolov et al, Physics Today, 2002

Renewal reward process

Y_t W_i



"An … analogy is that we have a magic goose which lays eggs at intervals (holding times). Sometimes it lays golden eggs of random weight, and sometimes it lays toxic eggs (also of random weight) which require responsible (and costly) disposal. The "rewards" are the successive (random) financial losses/gains resulting from successive eggs (i = 1,2,3,...) ."-Wikipedia

Here jumps at {J} become rewards {W} and waiting times become holding times {S}

Heavy tailed random telegraph (Niemann et al, 2013)

Model.—For simplicity, we consider a two-state model, with a state *up* where $I(t) = I_0$ and *down* with $I(t) = -I_0$. The sojourn times in these states are independently identically distributed random variables with PDFs $\psi(\tau)$. Thus, after waiting a random time in any state (called an *epoch*), the particle chooses the next state to be *up* or *down* with equal probability. The waiting time PDFs have long tails $\psi(\tau) \propto \tau^{-(1+\alpha)}$ with $0 < \alpha < 1$; hence, the averages of the *up* and *down* times are infinite. The Laplace $t \to \lambda$ transform of these PDFs is for small λ: $\hat{\psi}(\lambda) \simeq 1 - (\overline{\tau}\lambda)^{\alpha}$, where $\overline{\tau}$ is a scaling constant. This is a simple stochastic model of a blinking quantum dot, for which typically $\alpha = 1/2$, although $1/2 < \alpha < 1$ was also reported.

- Like Lowen & Teich AFRP but, crucially, without cutoffs

FIG. 1. Sample functions of fractal renewal processes. Interevent times are power-law distributed. (a) The standard fractal renewal process (SFRP) consists of Dirac δ functions and is zero-valued elsewhere. (b) The alternating fractal renewal process (AFRP) switches between values of zero and unity. The symmetric case is shown here.

Nonergodic spectrum

Statement of the main results.—For $\alpha < 1$, the expectation value of the spectrum is not constant but decreases with measurement time $\langle S_t(\omega)\rangle \simeq t^{\alpha-1}\sigma_\alpha(\omega)$. Expanding the t-independent function $\sigma_\alpha(\omega)$ for small frequencies ω, one finds a typical nonintegrable $1/f$ noise

$$\langle S_t(\omega)\rangle \simeq C\frac{t^{\alpha-1}}{\omega^{2-\alpha}}. \tag{4}$$

Spectra depend on observed length of time series as well as waiting time exponent alpha

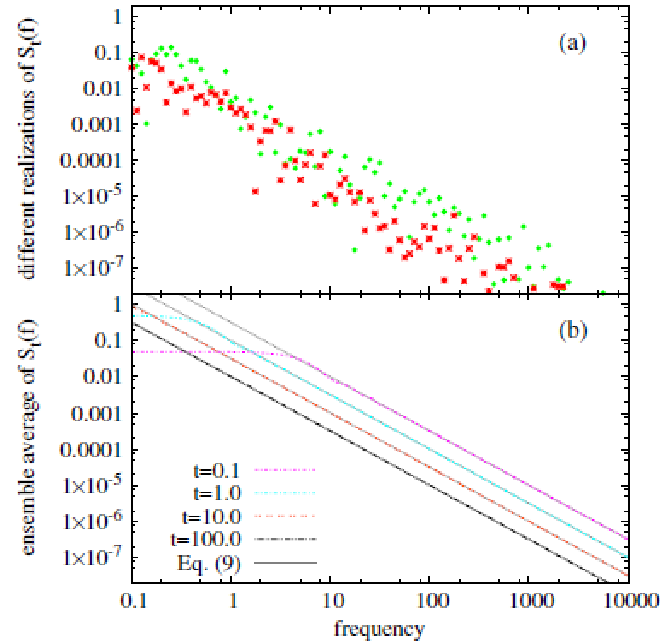


FIG. 1 (color online). (a) $S_t(f)$ plotted for different realizations ($\alpha = 0.5$, $t = 10^2$). Inside each realization one has fluctuations following exponential distributions. Different realizations are shifted with respect to each other due to the random prefactor Y_α. (b) Ensemble average of $S_t(f)$ plotted for different lengths t of the time series. One sees the decay of the spectrum $\langle S_t(f)\rangle \simeq 0.101t^{-1/2}f^{-3/2}$ [Eq. (9)] both in time and frequency. The crossover frequency is around $f_c \simeq 0.51/t$ [Eq. (14)]. The simulations perfectly match the theory [Eqs. (9) and (13)].

Strong fluctuations

In general, the value $S_t(\omega)$ of the spectrum is a fluctuating quantity even in the $t \to \infty$ limit. The statistical behavior of the general class of processes for large t (for pairwise disjoint $\omega_i \neq 0$) is fully described by the convergence in distribution of

$$\left(\frac{S_t(\omega_1)}{\langle S_t(\omega_1)\rangle}, \ldots, \frac{S_t(\omega_n)}{\langle S_t(\omega_n)\rangle}\right) \to Y_\alpha(\xi_1, \ldots, \xi_n), \qquad (5)$$

where Y_α is a random variable of normalized Mittag-Leffler distribution with exponent α whose moments are $\langle Y_\alpha^n \rangle = n!\Gamma(1+\alpha)^n/\Gamma(1+n\alpha)$ [30]. The ξ_i are independent exponential random variables with a unit mean. For $\alpha = 1$, the Mittag-Leffler random variable becomes $Y_1 = 1$, so that the powers $S_t(\omega_i)$ of different frequencies become independent exponentially distributed random variables—a result known for several ergodic random processes [31]. In the case of weak ergodicity breaking ($\alpha < 1$), the whole spectrum has a common random prefactor Y_α which shifts the complete observed spectrum.

Many procedures for the estimation of the spectrum from one finite time realization are designed to suppress the statistical fluctuations due to the uncorrelated random variables ξ_i [27,31]. These cannot account for the fluctuations of Y_α, common to all estimators of a given realization. For these procedures, the prefactor affects all estimated values for the spectrum. However, being a common prefactor, it does not affect the shape of the estimated spectrum so that such features as $1/f$ noise can be detected independently of the realization.

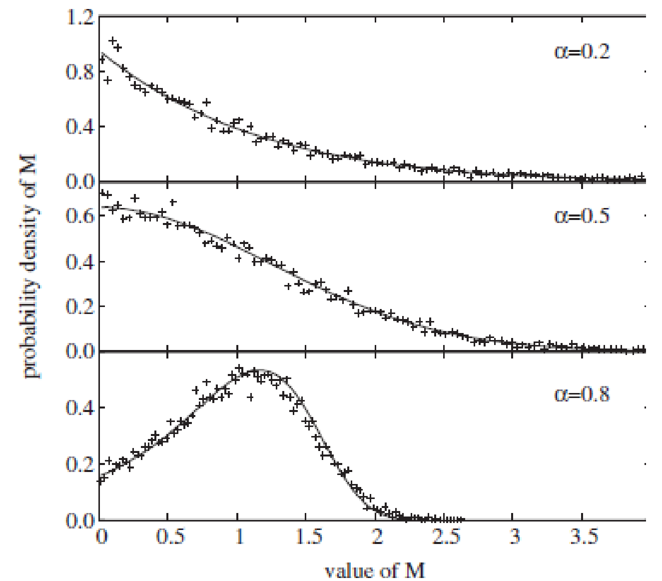


FIG. 2. Distributions of the frequency averaged spectra [see Eq. (12)]. The lines are the analytic probability densi of the Mittag-Leffler distributions ($t = 10^4$).

Strongly fluctuating estimates of average quantities such as power spectra.

$$M = \frac{1}{N}\sum_{i=1}^{N} \frac{S_t(\omega_i)}{\langle S_t(\omega_i)\rangle}.$$

Differs from fGn

- Two state random telegraph clearly non-ergodic
- Strong contrast to long range dependent, fractional Gaussian noise model (fGn) (Kolmogorov, 1940; Mandelbrot, 1965) which is

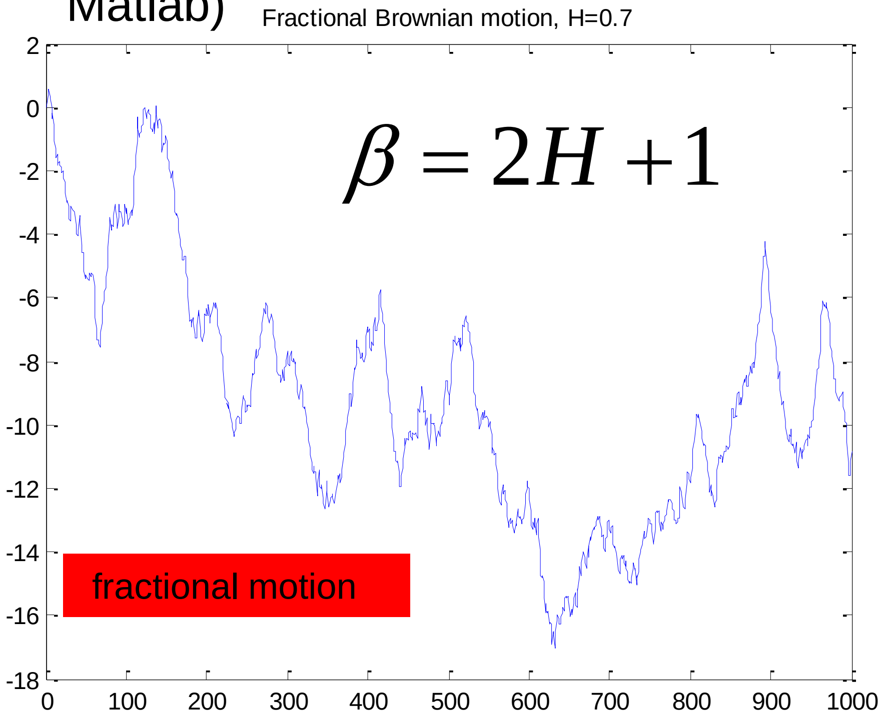
- ergodic [Mandelbrot & Van Ness, 1968]
- "better" behaved [M 1967; M and van Ness, 1968]
- at cost of assuming “unphysical”, fully non-Markovian, memory kernel

$$\text{fBm: } X_{H,2}(t) \sim \int_R \left((t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right) dL_2(s)$$

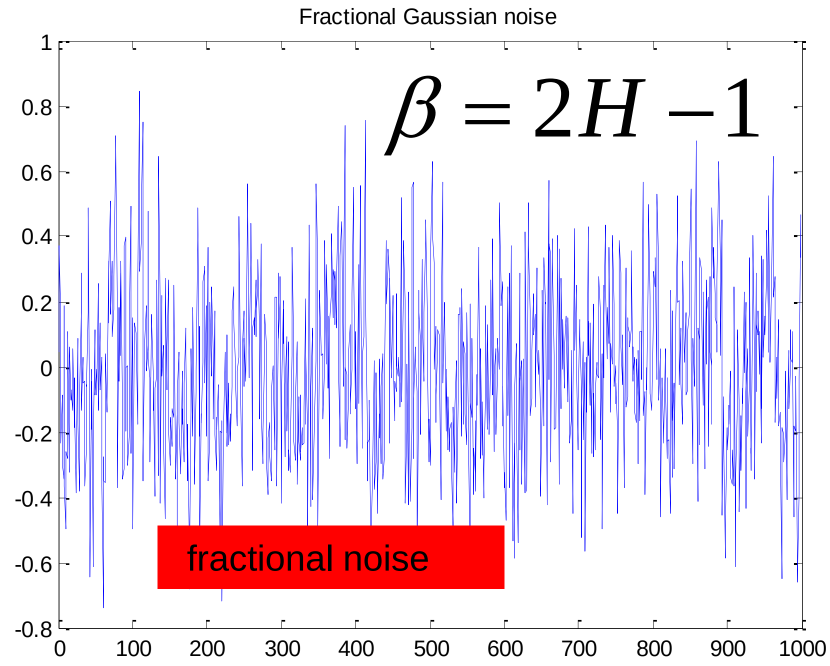
Fractional motions and noises

$$\text{fBm: } X_{H,2}(t) \sim \int_R \left((t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right) dL_2(s)$$

Build a nonstationary, self similar walk … (used wfbm in Matlab)



Then differentiate to give a stationary LRD noise



1/f noise : why it matters,
Why it's puzzling

$$S(f) \sim f^{-\beta}$$

Many solutions (or one) ?

Ergodic route

- Non-ergodic route

Another: Long tailed waiting times between switching-fractional renewal process/continuous time random walk. Motivated by physics of weak ergodicity breaking.

$$S(f) \sim \mathrm{Q(T)}\, f^{-\beta}$$

One solution: Long range dependent kernel, non-Markovian, fractional Gaussian noise. Physics now known to be the generalised Langevin equation, fGn is its noise term

$$S(f) \sim f^{-\beta}$$

BACK TO THE 60S

1967 …

- Surprisingly, weak ergodicity foreshadowed by Mandelbrot !
- Will summarise his fractional renewal models inspired by telephone errors …
- .. & results, especially those in remarkably far-sighted paper [IEEE Trans 1967].
- Why has it been almost totally overlooked in stats and physics literatures since 60s ?
- Has this affected history of complexity science and way the subject is taught ? [Graves et al, 2014; Watkins et al, 2015]
- Epilogue: Where do fGn and fractional renewal classes sit in wider modern context, as part "the panorama of grid bound variability".

Famously hard to read …

- *At once a compendium of Mandelbrot's pioneering work and a sampling of new results, the presentation seems modeled on the brilliant avant-garde movie "Last Year in Marienbad", in which the usual flow of time is suspended, and the plot is gradually revealed by numerous but slightly different repetitions of a few underlying events. As Mandelbrot himself admits in the Preface, the presentation allows the reader unusual freedom of choice in the order in which the book is read.*   - Nigel Goldenfeld, "Last Year in Mandelbrot "

… but worthwhile ?

In fact, I enjoyed this work most when I read it in random order, juxtaposing viewpoints and analyses separated in time by 3 decades, and making clear the progression of ideas that Mandelbrot has generated. These include:

- *the classification of different forms of randomness,*
- *their manifestation in terms of distribution theory,*
- *their ability to be represented compactly,*
- *the notion of trading time,*
- *the importance of discontinuities,*
- *the relationship between financial time series and turbulent time series,*
- *the pathologies of commonly abused distributions, particularly the log-normal, and*
- *a catalogue of the methods used to derive scaling distributions*

Berger & Mandelbrot, 1963

A New Model for Error Clustering in Telephone Circuits

Abstract: This paper proposes a new mathematical model to describe the distribution of the occurrence of errors in data transmission on telephone lines. We suggest: a) that the statistics of telephone errors can be described in terms of an error probability depending solely on the time elapsed since the last occurrence of an error; b) that the distribution of inter-error intervals can be well approximated by a law of Pareto of exponent less than one; the relative number of errors and the equivocation tend, therefore, to zero as the length of the message is increased. The validity of those concepts is demonstrated with the aid of experimental data obtained from the German telephone network. Further consequences, refinements, and uses of the model are described in the body of the paper.

- IBM J. Research & Development

Mandelbrot 1965

Self-Similar Error Clusters in Communication Systems and the Concept of Conditional Stationarity

BENOIT MANDELBROT, SENIOR MEMBER, IEEE

Abstract—The purpose of this paper is twofold. From the viewpoint of communications engineering, it presents a model of certain random perturbations that appear to come in clusters, or bursts. This will be achieved by introducing the concept of "self-similar stochastic point process in continuous time." The resulting mechanism presents fascinating peculiarities from the mathematical viewpoint. In order to make them more palatable as well as to help in the search for further developments, the basic concept of "conditional stationarity" will be discussed in greater detail than would be strictly necessary from the viewpoint of the immediate engineering problem of errors of transmission.

This work is related to an earlier "new model" due to J. M. Berger and the author; the logical structure of the theory has been further streamlined and a number of fresh consequences have been derived; the empirical fit has been further improved, while recourse to *ad hoc* corrective terms was made unnecessary.

I. Introduction

B. Prospect

The now classical technique of spectral analysis is inapplicable to the processes examined in this paper but it is sometimes unavoidable that otherwise excellent spectral estimates be applied in this context. Another publication of the author [18] is devoted to an examination of the expected outcomes of such operations. This will lead to fresh concepts that appear most promising indeed in the context of a statistical study of turbulence, excess noise, and other phenomena when interesting events are intermittent and bunched together (see also [19]).

[18] Mandelbrot, B., Time-varying channels, 1/*f* noises, and the infrared catastrophy. *Conv. Rec., 1st IEEE Communication Convention* (in press).

[19] ——, Self-similar turbulence and non-Wienerian conditioned spectra, to be published.

[18] Mandelbrot, B., Time-varying channels, $1/f$ noises, and the infrared catastrophy. *Conv. Rec., 1st IEEE Communication Convention* (in press).
[19] ——, Self-similar turbulence and non-Wienerian conditioned spectra, to be published.

Mandelbrot 1967

IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. IT-13, NO. 2, APRIL 1967 289

Some Noises with $1/f$ Spectrum, a Bridge Between Direct Current and White Noise

BENOIT MANDELBROT, SENIOR MEMBER, IEEE

[18] became 1965 IEEE Boulder conference paper **[N8 in Selecta 1999]** & [19] the 1967 journal paper **[N9].**

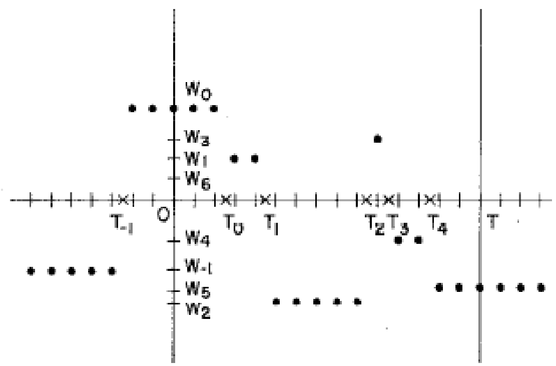


Fig. 2.

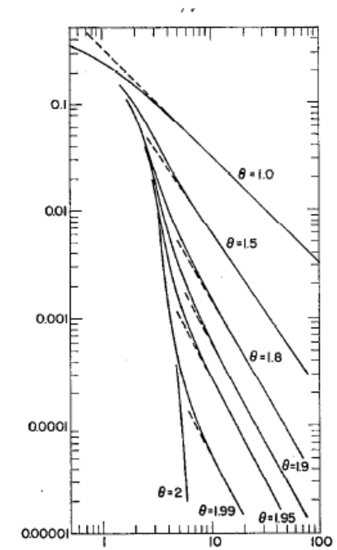
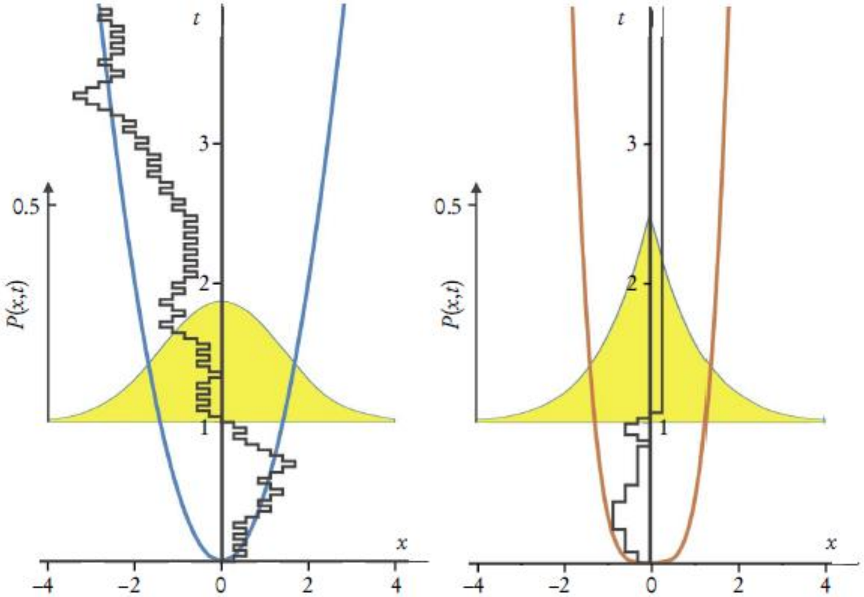


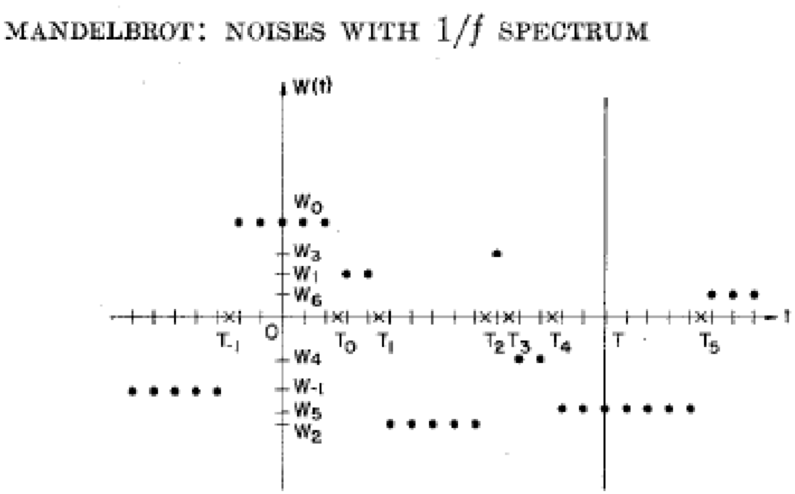
Fig. 3. This is a representation of the positive portions of Paul Lévy's symmetric stable distributions, with $1 \leq \theta \leq 2$. Abscissa: $\log u$; ordinate: $\log \Pr(U > u)$. The limit cases $\theta = 2$ and $\theta = 1$ are the classical laws of Gauss and of Cauchy. Note that, if u is large enough, one has $\Pr(U > u) \sim (u/u_\theta^*)^{-\theta}$.

- Abrupt state changes "conditionally stationary"
- Fat tailed distributions of switching times: "Levy" (E[t^2] = ∞) case.

Fig. 3. This is a representation of the positive portions of Paul Lévy's symmetric stable distributions, with $1 \leq \theta \leq 2$. Abscissa: $\log u$; ordinate: $\log \Pr(U > u)$. The limit cases $\theta = 2$ and $\theta = 1$ are the classical laws of Gauss and of Cauchy. Note that, if u is large enough, one has $\Pr(U \geq u) \sim (u/u_\theta^*)^{-\theta}$.

Continuous Time Random Walk (CTRW)

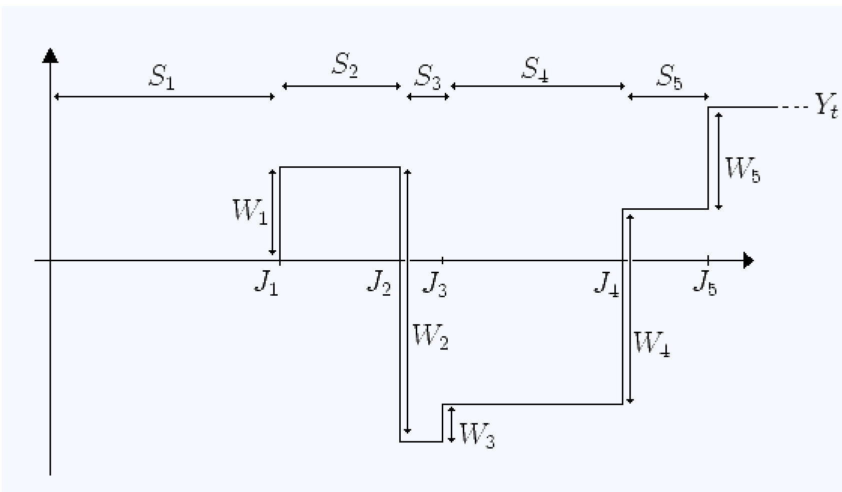




Montroll and Weiss 1965,
Figure from Sokolov et al, Physics Today, 2002

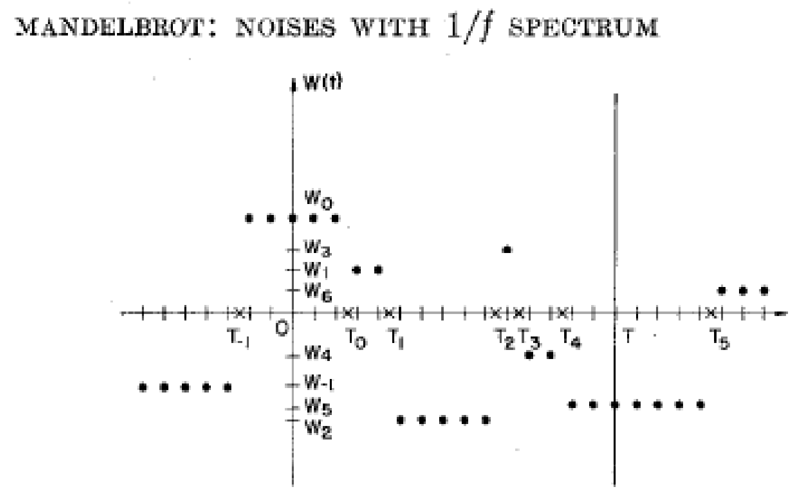
Y_t Renewal reward process

W_i



Here jumps at {J} become rewards {W} and waiting times become holding times {S}

“An alternative analogy is that we have a magic goose which lays eggs at intervals (holding times). Sometimes it lays golden eggs of random weight, and sometimes it lays toxic eggs (also of random weight) which require responsible (and costly) disposal. The "rewards" are the successive (random) financial losses/gains resulting from successive eggs (i = 1,2,3,...) .”-Wikipedia



Abstract—Noises in thin metallic films, semiconductors, nerve tissues, and many other media, have measured spectral densities proportional to $f^{\theta-2}$, with f the frequency and θ a constant $0 \leq \theta < 2$. The energy of these "$f^{\theta-2}$ noises" behaves more "erratically" in time, than expected from functions subject to the Wiener–Khinchin spectral theory. Moreover, blind extrapolation of the "$f^{\theta-2}$ law" to $f = 0$ incorrectly suggests, when $\theta \leq 1$, that the total energy is infinite ("infrared catastrophe"). The problems thus raised are of the greatest theoretical interest, and of the greatest practical importance in the design of electronic devices.

The present paper reinterprets these spectral measurements without paradox, by introducing a concept to be called "conditional spectrum." Examples are given of functions ruled by chance, that have the observed "erratic" behavior and conditional spectral density.

A conditional spectrum is obtained when a procedure, meant to measure a sample Wiener–Khinchin spectrum, is applied to a sample conditioned to be nonconstant. The conditional spectrum is defined, not only for nonconstant samples from all random functions of the Wiener–Khinchin theory, but also for nonconstant samples from certain nonstationary random functions, and for nonconstant samples from a new generalization of random functions, called "sporadic functions."

The conditional spectrum:

- First key finding was that Wiener-Khinchine inspired measures like periodogram would return a ``1/f'' shape for such models, but that a more useful object was conditioned also on series length T. This conditional spectrum S(f,T) would factor into two parts, one dependent on f and one on T.
- In the 1967 paper he describes it thus (theta is determined by exponent of power law waiting time pdf = usual Levy alpha):

The properties of the corresponding functions $V(t)$ $W(t)$, and $X(t)$ can be summarized as follows:

If $1 < \theta \leq 2$, $V(t)$ or $W(t)$ have *Wiener–Khinchir spectral densities* $S'(f)$ of the form $f^{\theta-2}L(f)$, where $L(f$ varies slowly near $f = 0$.

$W(t)$ has a *conditional spectrum* $S(f, T)$, such that $S(f, T) = f^{\theta-1}L(f)Q(T)$, where $5/T \ll f \ll \frac{1}{2}$. If $1 < \theta \leq 2$ $0 < \lim_{T\to\infty} Q(T) < \infty$. If $0 \leq \theta < 1$, $Q(T)T^{1-\theta}$ varie slowly for $T \to \infty$.

The conditional spectrum:

Fourier psd $S^{'} \sim f^{\theta-2} L(f)$

Cond. Spectrum $S(f, T) \sim f^{\theta-1} L(f) Q(T)$

$Q(T) T^{1-\theta}$ varies slowly

So $\dfrac{d}{d\theta} S(f, T) = S'(\mathrm{f, T}) \sim f^{\theta-2} T^{\theta-1} L(f)$

i.e. $\dfrac{T^{\theta-1}}{f^{2-\theta}} L(f)$

The conditional spectrum:

- First key finding was that Wiener-Khinchine inspired measures like periodogram would return a ``1/f'' shape for such models, but that a more useful object was conditioned also on series length T. This <u>conditional</u> spectrum S(f,T) would factor into two parts, one dependent on f and one on T.
- In the 1967 paper he describes it thus (theta is determined by exponent of power law waiting time distribution = usual Levy alpha):

  - "*Numerical … 1/f … spectrum … need not … estimate … Wiener-Khinchine spectrum*". **M67 reviewed in N2, Selecta, 1999**

    Instead "*depends on conditioning length T*".

    Unlike the stationary LRD model, singularity is an artefact.

The infrared catastrophe as mirage:

- Rather than representing a true singularity in power at the lowest frequencies, in this model he described the infrared catastrophe in the power spectral density as a ``mirage“:

The infrared mirage. In a finite sample, the accessible frequencies do not range down tO, only down to $f = 1/T$. Therefore, an effect of letting $T \to \infty$, is to modify this the accessible range, and force an unchanging total energy to "flow along" toward increasingly low frequencies.

As a result, the threatened low frequency divergence or infrared catastrophe *never materializes* and the self-consistency of nature is preserved. However, the interpretation of spectra is deeply affected. The fact that the additional prefactor T^{1-B} is not numerical but a function of T expresses that the measured square Fourier moduli do *not* estimate a Wiener-Khinchin spectral density, but something *different*. Thus, the differences in geometry have obvious practical consequences that one could not deduce from the form of the spectrum alone.

M67 revisited in N2, Selecta, 1999

Distinct from fBm and fGN:

Mandelbrot 1967 was prepared in the same period as Mandelbrot and van Ness on fBm and fGn, which it cites as ``to be published". In it contrast is clearly drawn between the sampling behaviour of conditionally stationary, non-Gaussian renewal process as a 1/f model and his stationary, Gaussian (fGn) model:

Section VI showed that *some* $f^{\theta-2}L(f)$ noises have a very erratic sampling behavior. Some *other* $f^{\theta-2}$ noises, however, are Gaussian and, therefore, perfectly "well-behaved;" an example is provided by "fractional white noise" which is the formal derivative of the process of Mandelbrot and Van Ness.[10] This difference in behavior

1 jump is extreme case: not Wiener-Khinchine

- **M 1967 first** illustrated non ergodicity with single jump in an infinite interval

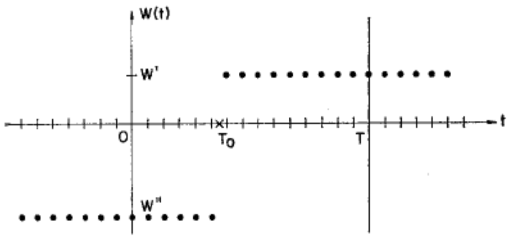


Fig. 1.

The Nonergodic Character of the Function $W(t, T_0)$

A function $X(t)$ is ergodic if $E(X)$ is defined and if $\lim_{|t''-t'|\to\infty} (t'' - t')^{-1} \sum_{s=t'+1}^{t''} X(s)$ is nonrandom and equals $E(X)$ almost surely. When T_0 is a proper random variable, the function $W(t, T_0)$ is nonergodic, being highly nonstationary. When T_0 is uniformly distributed over $(-\infty, \infty)$, $W(t, T_0)$ is also nonergodic: it is "almost surely" constant over any prescribed finite time span, and its sample mean equals $W(t')$, which is random and independent of T.

Sampling Distribution of the Conditional Covariance

A corollary of nonergodicity is that, if $T \to \infty$ while n/T remains constant, the expression

$$R = [(1/T) \sum_{t=1}^{T} W(t, T_0)W(t + n, T_0) \mid T_0 \varepsilon (1, T)]$$

does *not* tend to the limit $1 - |n|/T$ of its expectation. For example, when $0 < n < T/2$, the span of variation of $H = T_0/T$ is to be divided into three portions, 0 to n/T, n/T to $1 - n/T$, and $1 - n/T$ to 1. In the first span, $R = (1 - h - n/T)W''^2 + hW'W''$; in the second span, $R = (1 - n/T)W'^2 + (1 - h - n/T)W''^2 + hW'W''$; in the third span, $R = (h - n/T)W'^2 + (1 - h)W'W''$. Thus, for every fixed couple (W', W''), R is the mixture of three uniformly distributed random variables.

Also explicitly discussed non-ergodicity of divergent mean waiting time case

B. The Case $0 < \theta < 1$

The ergodic theorem fails to apply to $W(t)$, for reasons similar to those encountered for the function $W(t, T_0)$ of Section II: As $t \rightarrow \infty$, $t^{-1}W^0(t)$ tends in distribution to a *random* limit. The ergodic theorem is trivial for $V(t)$, because the sample mean almost surely tends to zero as $T \rightarrow \infty$.

Thus, one cannot speak of the "dc component" of processes such as $W(t)$, with $0 < \theta < 1$. For processes

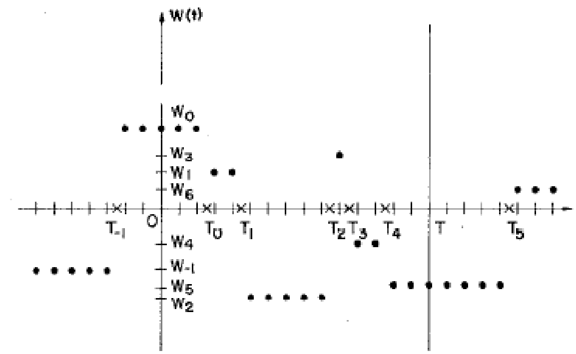


Fig. 2.

So was it W.E.B ?

Yes and No.

- Noted the consequences of naively interpreting a periodogram as a spectrum even when Wiener-Khinchine not satisfied, and the solution to “1/f catastrophe”
- Derived correct expression for periodogram and dependence on T that Niemann et al confirmed
- Noted that fluctuations in averages were “wild” compared to those of fGn and had some results that I have yet to decode
- But saw it in a much more abstract context than the modern physical one-but also more generic
- Focus was signal processing rather than physics

A neglected paper … ?

- **Mandelbrot 1967** received far less attention than either papers on heavy tails in finance in early 1960s or the series with **van Ness and Wallis in 1968-69** on stationary fractional Gaussian models for LRD. **M & van Ness 1968 SIAM Review is alone now ~2600 citations.**
- **M 67 got about 20 citations in its first 20 years (c.f. M & van Ness ~100) despite fact that the 1967 IEEE paper did cite the 1968 fGn paper and vice versa.**
- Was apparently unknown to **Vit Klemes [Water Resources Research, 1974],** who essentially reinvented it to criticise fBm. Still seems relatively little known. Citations in the 100s … last time I looked … after 50 years.
- Not cited by **Beran et al [2013],** and while listed in the citations of **Beran [1994]** I haven't found it in the text.
- Some exceptions, e.g. **Lenoir, Fluctuation and Noise Letters, 2013**

… Whose time has come ?

- Should we pay more attention to this class of models ?

- Fortunately, as I showed you, we now are.

- Of course, the above workers going much further than Mandelbrot, in a physics-driven context

- I think most direct value of looking back nearly 50 years to how Mandelbrot saw these models is to see how they fit into *"the panorama of grid-bound self-affine variability"* as he later put it **[Selecta, 1999, N1].**

- Helps link maths and physics, the formula & the phenomenon, and inform future work.

Why neglected ?

Although he revisited the paper with new commentary in **Selecta Volume N [1999]** dealing with multifractals and 1/f noise, **Mandelbrot** neglected to mention it explicitly in his popular and historical accounts of the genesis of LRD such as **Mandelbrot and Hudson [2008].**

Why ?

- Because it wasn’t as popular as fBm/fGn ?
- Because it wasn’t as “beautiful” as the self-similar LRD kernel?
- Because he wanted to keep it for his “day job” at IBM ?
- Or because it complicated the story of how he got from heavy tails in finance to fGn too much ?

Why does this matter ?

While it is common that ideas don't flourish if they are far ahead of their time-think of **Ada Lovelace** for just one of many examples-it is perhaps not so common for such an crucial step in a very famous person's output to be ignored. I think it does matter, because

- Almost all discussion of 1/f in stats tends to be either in terms of fGn (or its relative ARFIMA) or conceptually framed as breakpoints etc
- Almost all physics books or review papers on 1/f use one of the CTRW or fGn as a paradigm, and few compare them.
- This has affected geophysics, economics, neurology, …

1/f noise : why it matters,
Why it's puzzling

$$S(f) \sim f^{-\beta}$$

Many solutions (or one) ?

Ergodic route

- Non-ergodic route

Another: Long tailed waiting times between switching-fractional renewal process/continuous time random walk. Motivated by physics of weak ergodicity breaking.

One solution: Long range dependent kernel, non-Markovian, fractional Gaussian noise. Physics <u>now</u> known to be the generalised Langevin equation, fGn is its noise term

$$S(f) \sim Q(T) f^{-\beta}$$

$$S(f) \sim f^{-\beta}$$

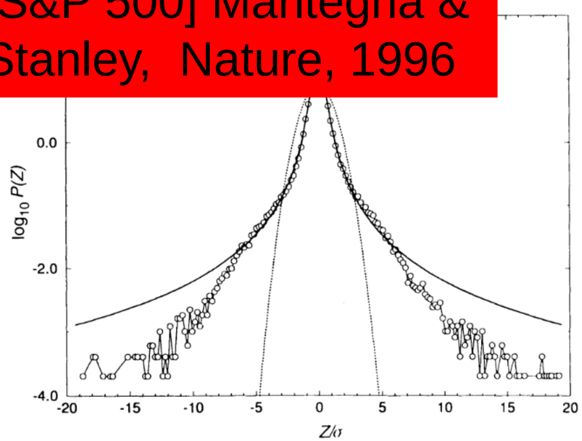
EPILOGUE: MANDELBROT’S PANORAMA

Fact: Wild Fluctuations

THE VARIATION OF CERTAIN SPECULATIVE PRICES*

BENOIT MANDELBROT†

[S&P 500] Mantegna & Stanley, Nature, 1996



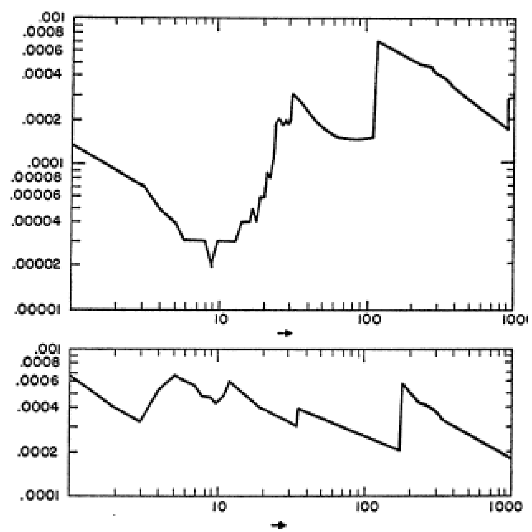


FIG. 2.—Both graphs are relative to the sequential sample second moment of cotton price changes.

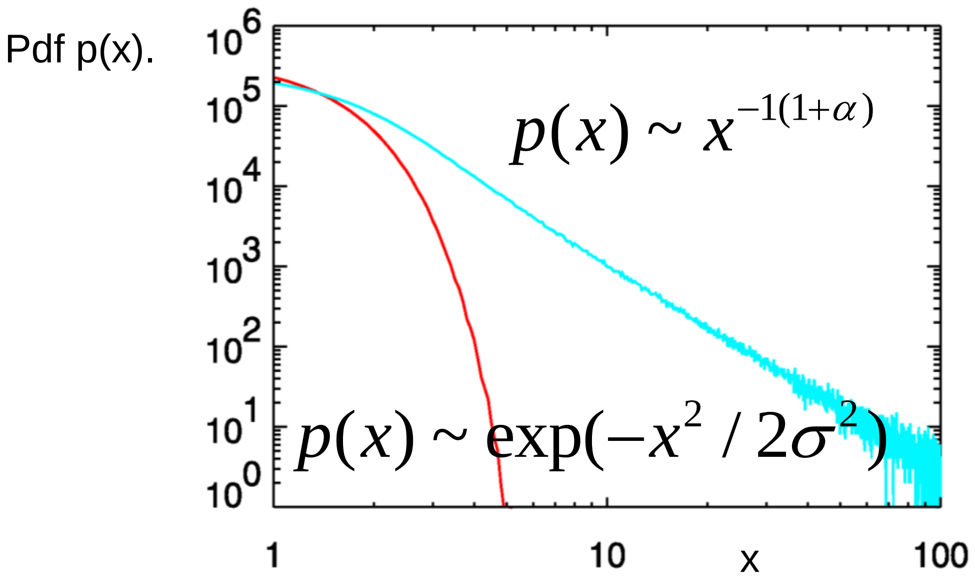
Horizontal scale represents time in days, with two different origins T°: on the upper graph, T° was September 21, 1900; on the lower graph T° was August 1, 1900. Vertical lines represent the value of the function

$$(T - T^\circ)^{-1} \sum_{t=T^\circ}^{t=T} [L(t,1)]^2 ,$$

where $L(t, 1) = \log_e Z(t+1) - \log_e Z(t)$ and $Z(t)$ is the closing spot price of cotton on day t, as privately reported by the United States Department of Agriculture.

J. Business, 1963

Formula: Heavy tails

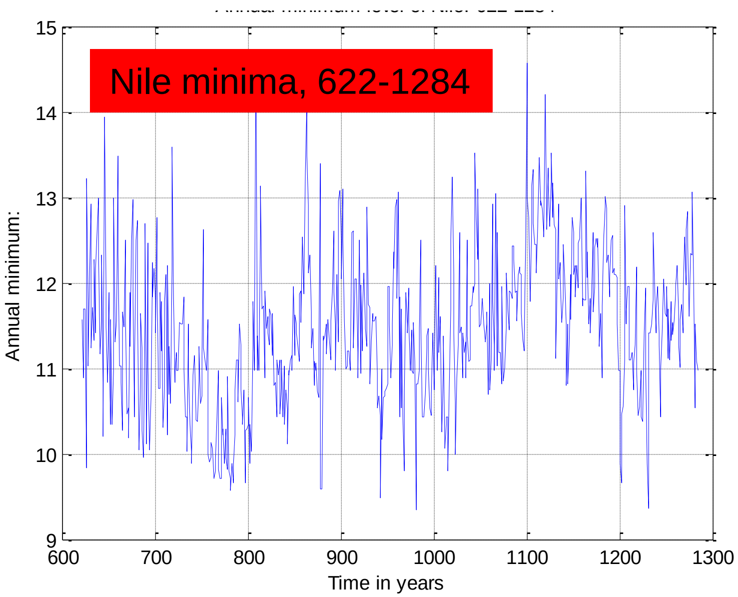


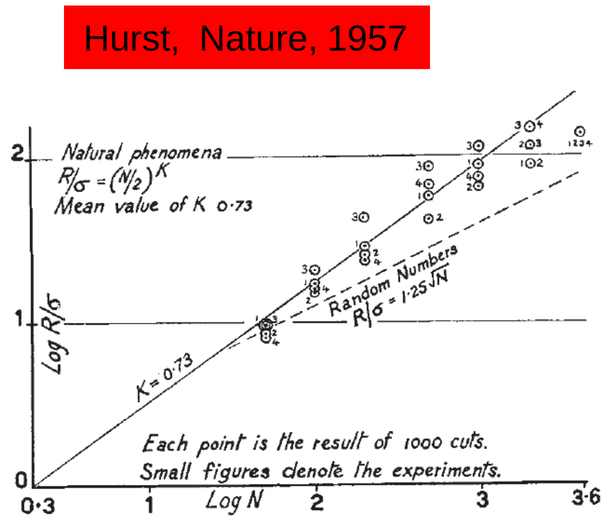
Heavy tailed example, alpha stable distribution which has a power law tail.

Light tailed example Gaussian

$$p(x) \sim x^{-(1+\alpha)}, \; H = 1/\alpha$$

Another fact: Hurst's growth of range





"I heard about the … Nile … in '64, ... the variance doesn't draw like time span as you take bigger and bigger integration intervals; it goes like time to a certain power different from one. … Hurst … was getting results that were incomprehensible".

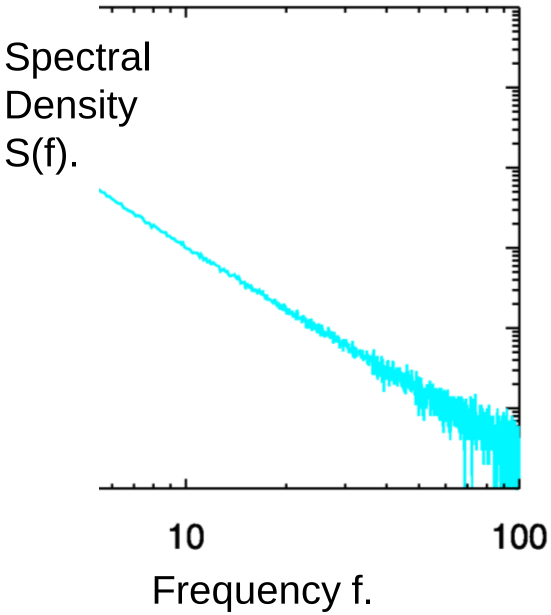
– Mandelbrot, 1998, interviewed by Bernard Sapoval for Web of Stories

But what's the formula ?

"This was very much noticed and the literature grew about it … it was viewed as a major puzzle,
this thing which didn't work out" -Mandelbrot, 1998.

In collected papers (Selecta) said he initially thought could explain Hurst's observations with heavy tailed model like 1963 financial one.

But when saw data realised wasn't heavy tailed in amplitude ! Instead abstracted out property of self similarity, but in spectral rather than amplitude domain --- i.e. proposed a model with a heavy tailed power spectrum ~ 1/f, **even down to lowest frequency.** Advocated idea to mathematicians with van Ness (68), and hydrologists with Wallis (68-69)**.**

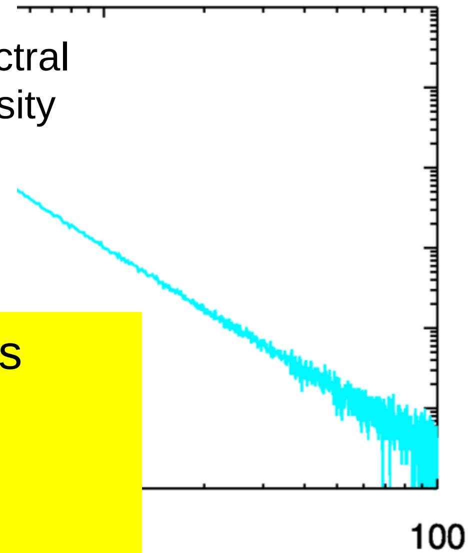


$$S(f) \sim f^{-\beta}$$

But what's the formula ?

"This was very much noticed and the literature grew about it … it was viewed as a major puzzle, this thing which didn't work out" -Mandelbrot, 1998.

Spectral Density S(f).



quency f.

- The leap of imagination and abstraction from long tails in a pdf to long tails in an acf and thus a singular power spectrum may seem very large …

perhaps to large to make in one go …

Hold that thought …

$$S(f) \sim f^{-\beta}$$

amplitude domain --- i.e. proposed a model with a heavy tailed power spectrum ~ 1/f, **even down to lowest frequency.** Advocated idea to mathematicians with van Ness (68), and hydrologists with Wallis (68-69).

Formula: fBm & fGn, 1965-68

SIAM Review
Vol. 10, No. 4, October 1968

FRACTIONAL BROWNIAN MOTIONS, FRACTIONAL NOISES AND APPLICATIONS*

BENOIT B. MANDELBROT† AND JOHN W. VAN NESS‡

1. Introduction. By "fractional Brownian motions" (fBm's), we propose to designate a family of Gaussian random functions defined as follows:[1] $B(t)$ being ordinary Brownian motion, and H a parameter satisfying $0 < H < 1$, fBm of exponent H is a moving average of $dB(t)$, in which past increments of $B(t)$ are weighted by the kernel $(t - s)^{H-1/2}$. We believe fBm's do provide useful models for a host of natural time series and wish therefore to present their curious properties to scientists, engineers and statisticians.

The basic feature of fBm's is that the "span of interdependence" between their increments can be said to be infinite. By way of contrast, the study of

$$\text{fBm: } X_{H,2}(t) \sim \int_R \left((t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right) dL_2(s)$$

Memory kernel

Gaussian

See also **Kolmogorov's "Wiener Spirals and some other interesting curves in a Hilbert space" (1940).**

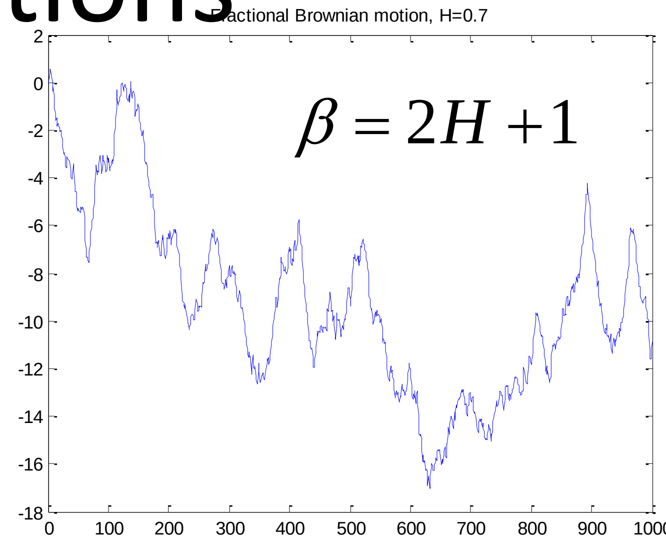
And **Mandelbrot, Comptes Rendus, 1965**

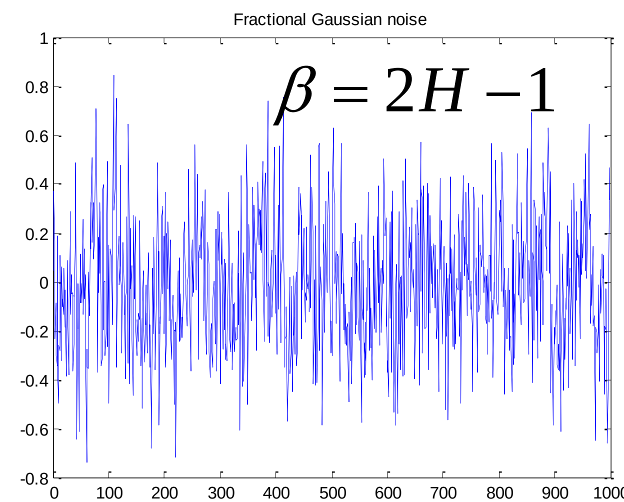
Noises vs. motions

fBm: $X_{H,2}(t) \sim \int_R \left((t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right) dL_2(s)$

- **Mandelbrot and van Ness [1968]** proposed use of fractional Brownian motion. Non stationary, H-self similar model. Generalises Wiener process, has spectral index between -1 and -3.

- … and its derivative, fractional Gaussian noise, which is stationary, and long range dependent.





So what did BBM think it meant ?

*[…], **if infinite dependence is necessary it does not mean that IBM's details of ten years ago influence IBM today, because there's no mechanism within IBM for this dependence.** However, IBM is not alone. The River Nile is [not] alone. **They're just one-dimensional corners of immensely big systems**. The behaviour of IBM stock ten years ago does not influence its stock today through IBM, but IBM the enormous corporation has changed the environment very strongly. The way its price varied, went up or went up and fluctuated, had discontinuities, had effects upon all kinds of other quantities, and they in turn affect us. –*
Mandelbrot, interviewed in 1998 by B. Sapoval for Web of Stories

In modern fractional Langevin models fGn is noise term e.g. Metzler et al, PCCP, 2014; Watkins GRL, 2013; Taloni et al, 2010; Kupferman, 2004; Lutz, 2001.

So what did BBM think it meant ?

*[…], **if infinite dependence is necessary it does not mean that IBM's details of ten years ago influence IBM today, because there's no mechanism within IBM for this dependence.** However, IBM is not alone. The River Nile is [not] alone. **They're just one-dimensional corners of***

- Resolution of apparent paradox is that physical laws are Markovian, the infinite memory is consequence of looking at a piece of world ? In spirit of Mori-Zwanzig.

very strongly. The way its price varied, went up or went up and fluctuated, had discontinuities, had effects upon all kinds of other quantities, and they in turn affect us. – **Mandelbrot, interviewed in 1998 by B. Sapoval for Web of Stories**

In modern fractional Langevin models fGn is noise term e.g. Metzler et al, PCCP, 2014; Watkins GRL, 2013; Taloni et al, 2010; Kupferman, 2004; Lutz, 2001.

(Ohmic) Langevin equation

$$M\ddot{q} = -V'(q) - \eta\dot{q} + f(t)$$

Beyond the Ohmic case

In deriving Langevin equation can consider other types of reservoir oscillator spectral function including but not limited to power laws :

$$J(\omega) \propto \omega^s$$

where $s > 1$ is super-Ohmic

and $s < 1$ is sub-Ohmic

In the presence of a memory in the heat bath we then have generalised Langevin equation of the form:

$$M\ddot{q} = -V'(q) - M\int_0^t dt' \rho(t-t')\dot{q}(t') + f(t)$$

Where memory kernel rho replaces constant eta **[e.g. Kupferman, 2004; Caldeira, 2010]**

Fractional Langevin equation

If memory kernel has slowest decay $\rho(\tau) \sim \tau^{-(1+2d)}$

then GLE: $M\ddot{q} = -V'(q) - M\int_0^t dt' \rho(t-t')\dot{q}(t') + f(t)$

becomes FLE: $M\ddot{q} = -\mathbf{V}'(q) - M\rho_{(1+2d)} \frac{\partial^{2d}}{\partial t^{2d}} \dot{q}(t') + f(t)$

where frac. derivative is $\frac{\partial^{\lambda} F(t)}{\partial t^{\lambda}} = \frac{1}{\Gamma(-\lambda)} \int_0^t d\tau F(\tau)(t-\tau)^{-(1+\lambda)}$

The many faces of "1/f"

- Late in his life, Mandelbrot re-emphasised that the formula **wasn't** the fact, and the property of self-similarity seen in his most famous models wasn't the whole story. *"Reducing the notion of "1/f noise" to self-affinity ... shows it to be very severely underspecified"*- Selecta volume N, 1999*.*

- Why was he saying this ? Because his eyes told him to: "*Like the ear, the eye is very sensitive to features that the spectrum does not reflect. Seen side by side, different 1/f noises, Gaussian, dustborne and multifractal, **obviously** differ from one another"*- Selecta, op cit.

So what were these models ?

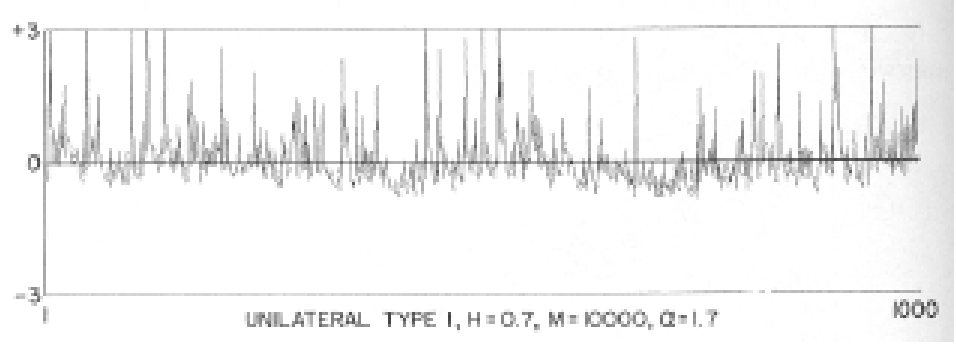
- Additive, stable, models extending fGn like fractional hyperbolic model of Mandelbrot & Wallis [1969].
- Multiplicative, multifractal models exhibiting volatility bunching as well as 1/f spectra and fat tails-1972 (turbulence), 1990s (finance).
- And the class he referred to as "dustborne": the least known of his papers, from 1963-67, though closely related to the Alternating Fractional Renewal Process, the CTRW and modern work on weak ergodicity breaking.
- We'll very briefly recap first two, then dwell on last one.

Additive fractional stable class

Water Resources Research: 5, 1969, 967-988

H25

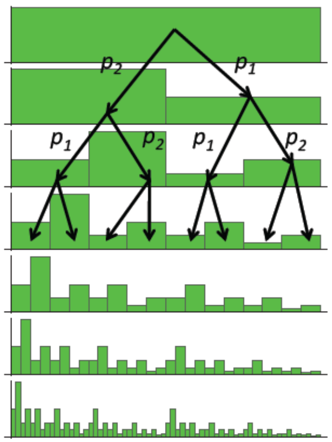
Robustness of *R/S* in measuring noncyclic global statistical dependence (M & Wallis 1969c)



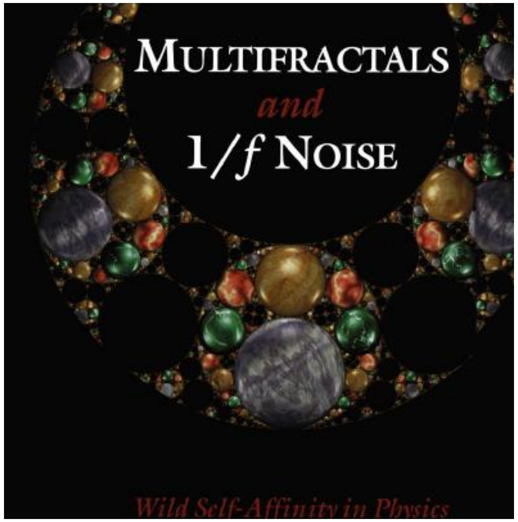
Mandelbrot & Wallis [1969] first attempt to unify long range memory kernel of fGn with heavy tailed amplitude fluctuations - called it "fractional hyperbolic" model because of its power law tails.

Anticipated today's versatile linear fractional stable noises, but it didn't satisfy him completely for problems he was looking at.

Multiplicative multifractal cascades



Selecta Volume N 1999

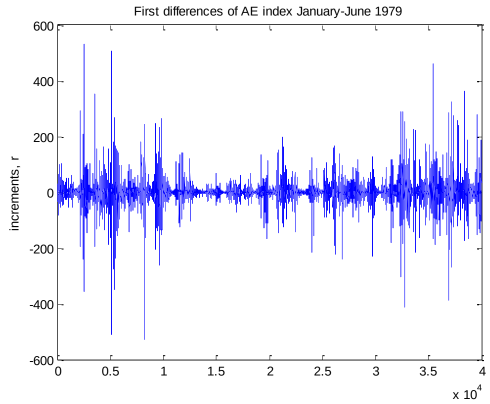


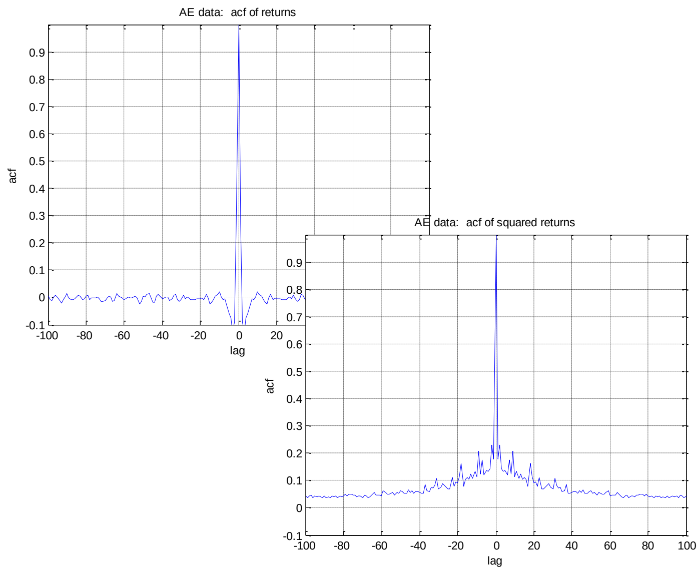
Many systems have aggregation, but not by an additive route. Classic example is turbulence.

One indicator is a lognormal or stretched exponential pdf

Multifractals and volatility clustering

another is correlations between the **absolute values** of the time series- or here, in ionospheric data, the first differences.

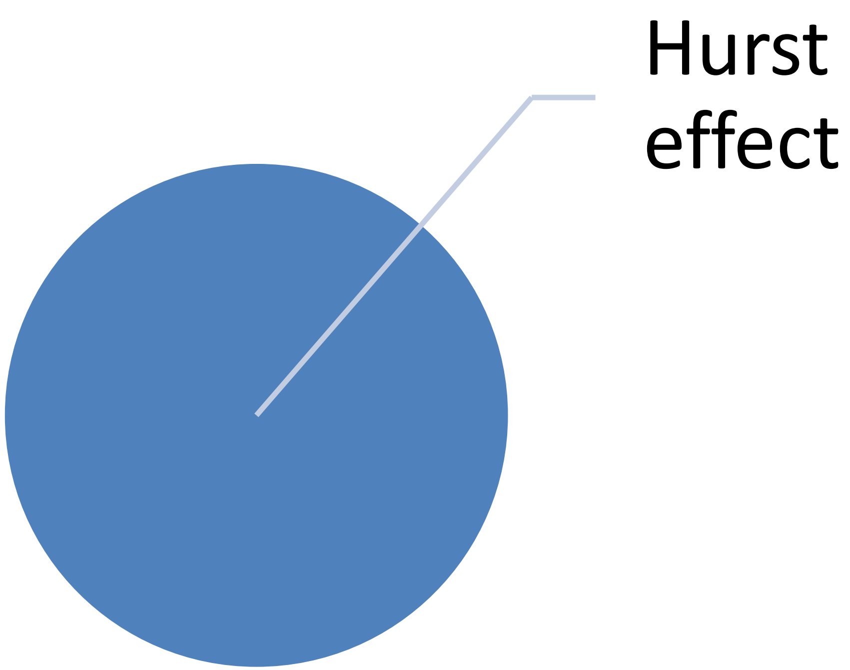




Watkins et al, in “Extreme Events and Natural Hazards”, 2012

SPARES

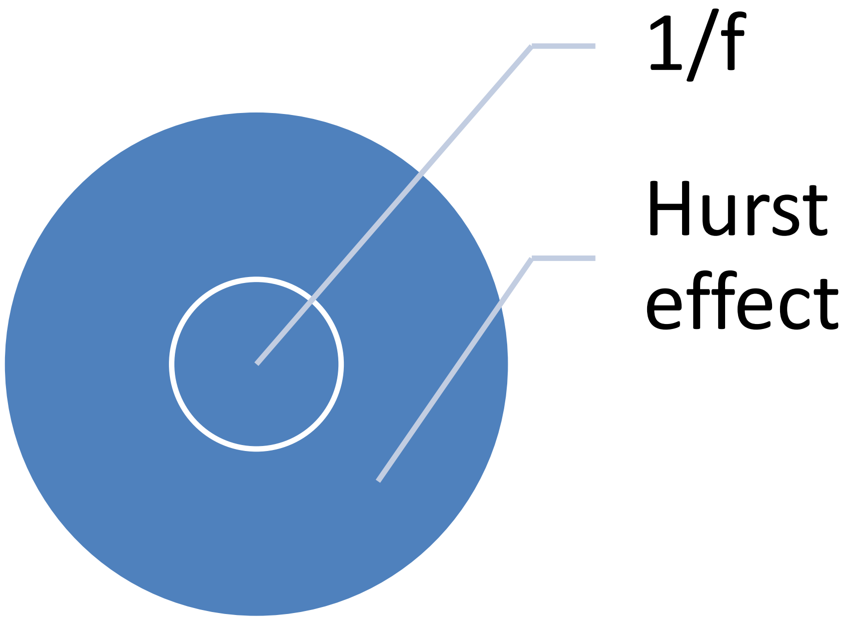
Theme



Will today distinguish three things often taken as same

- Observed growth of range in time series: “Hurst effect”

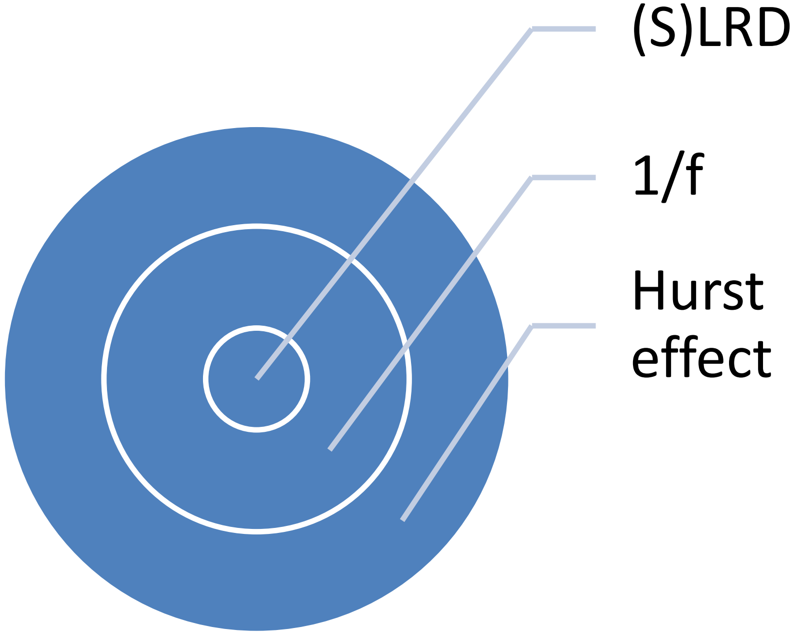
Theme



Will today distinguish three things often taken as same

- Observed growth of range in time series: “Hurst effect”
- Observation of a singularity at zero in Fourier spectra: “1/f”

Theme



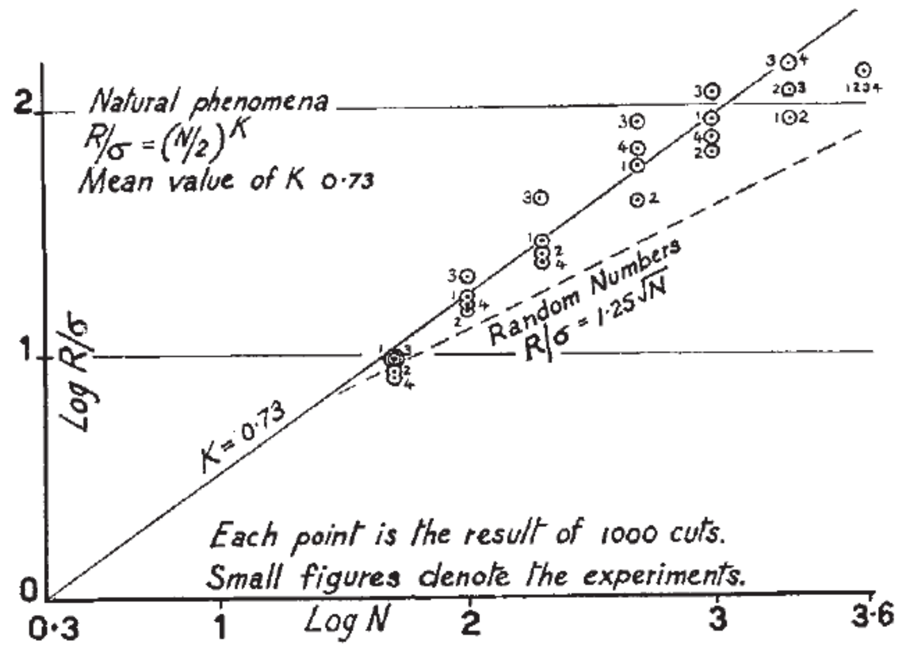
Will today distinguish three things often taken as same

- Observed growth of range in time series: "Hurst effect"
- Observation of a singularity at zero in Fourier spectra: "1/f"
- The long range dependence seen in stationary 1/f case: (S)LRD.
- Using 1/f as a diagnostic of LRD assumes stationarity

Fact: Anomalous growth of range

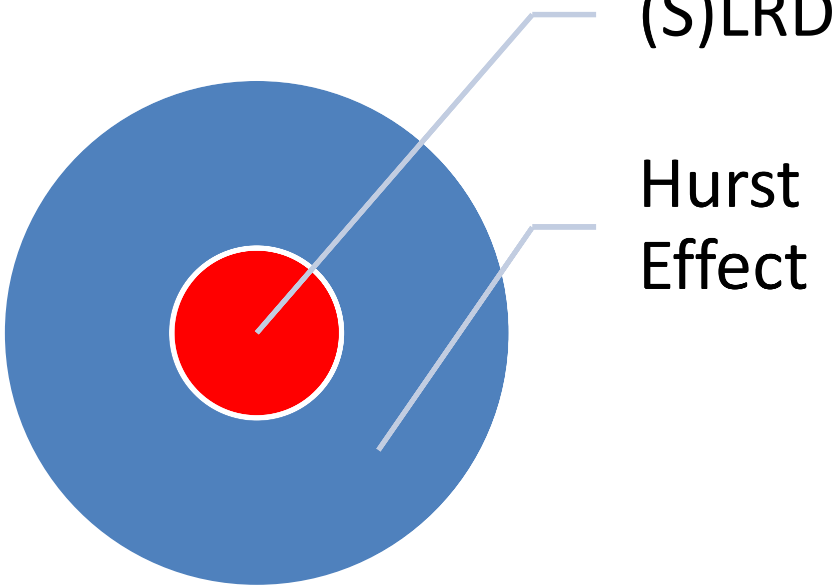
Hurst, Nature, 1957

Hurst Effect



"I heard about the … Nile … in '64, … the variance doesn't draw like time span as you take bigger and bigger integration intervals; it goes like time to a certain power different from one. … Hurst … was getting results that were incomprehensible". – ***Mandelbrot, '98***

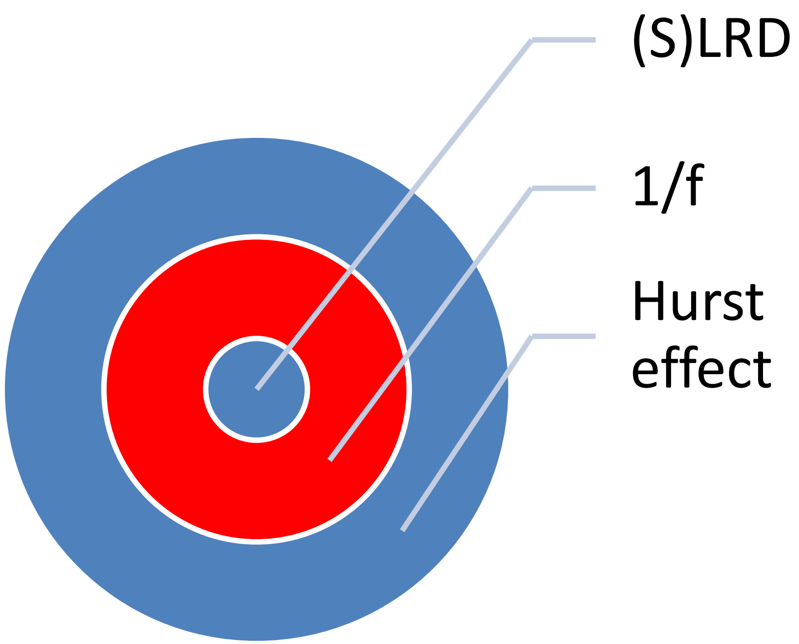
Formula: Long Range Dependence



- **Mandelbrot, van Ness, and Wallis, 1965-69**
- First **[history in Graves et al, arXiv, 2014a]** demonstration that Hurst effect could be explained by stationary long range dependent process
- Model, fractional Gaussian noise [see also **Kolmogorov's** "Wiener Spiral"], had singular spectral density at lowest frequencies.

$$S'(f) \sim f^{-\beta}$$

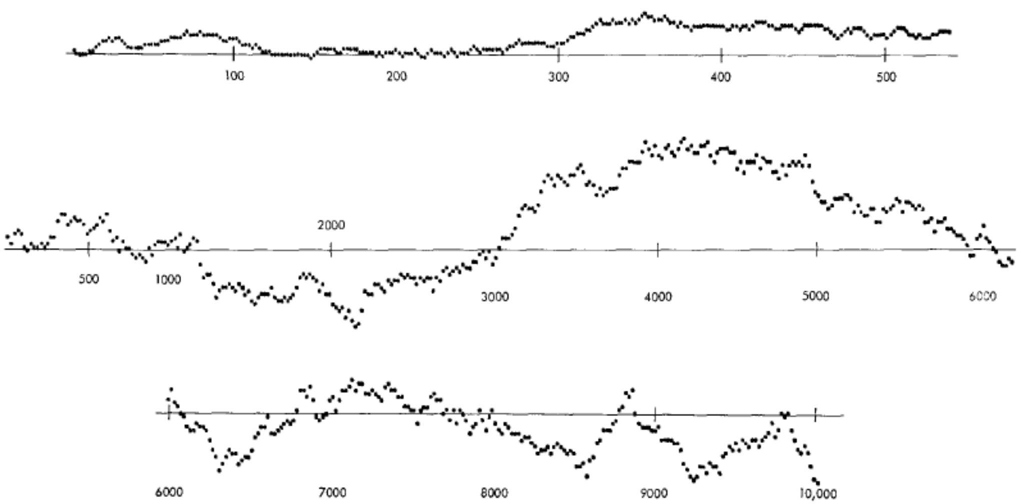
1/f without (S)LRD



- Before (S)LRD models, **Mandelbrot [1963-67]** had proposed other 1/f models which were not stationary LRD in same sense as fGn.
- Solved 1/f paradox by a different route. Still little known in the geosciences [but see **Klemes, WRR, 1974**].

Berger & Mandelbrot, 1963

Figure 1 **Record of Paul's winnings in a coin-tossing game, played with a fair coin.** *Zero-crossings appear strongly clustered, although the intervals between them are obviously statistically independent. In order to appreciate fully this Figure, one must note that the unit of time used on the second and third lines equals 20 plays. Hence, the second and third lines lack detail and each of the corresponding zero-crossings is actually a cluster or a cluster of clusters. For example, the details of the clusters around time 200 can be clearly read on line 1, which uses a unit of time equal 2. The present graph is reproduced from Fig. III.5, Ref. 13.*



225

• *Motivation of the fundamental postulates by the random walk process*

The intuitive reason that suggested the postulates (A) and (C′) is the striking qualitative resemblance which seemed to us to exist between the empirical records of inter-error intervals and the sequences of returns to the origin in the classical game of tossing a fair coin. Let us restate the rules of that game. The two celebrated old men Peter and Paul began to play (circa 1700 A.D.) with infinite fortunes; whenever their coin fell on "heads," Peter paid a cent to Paul, and whenever it fell on "tails," it was Paul who paid a cent to Peter. The behavior of $G(m)$, Peter's gain after m coin tosses, is well known to mathematicians and to some professional gamblers to be totally contrary to what is sometimes referred to as "intuition." Examine indeed our Fig. 1 (which is reproduced from Fig. III. 5 of Ref. 13). By definition, the intervals between successive roots of the equation $G(m) = 0$ are given by independent random variables; there is no question, however, that they appear to be grouped in clusters and that there are violent fluctuations in the intervals between such roots (Footnote 3). This suggests that error clustering and the violent fluctuations in the bit-error rate of telephone lines *need not* be due to dependence between the inter-error intervals; both perhaps may be described by peculiarities of the distribution of *independent* successive inter-error intervals.

Mandelbrot 1965

To try and broaden the existing theory, one may recall that, in its most general and loosest sense, a random function is simply an object chosen at random out of any prescribed collection of ordinary functions. This means that if the objects in that collection are labeled by an appropriate parameter λ, a random function is simply a function of two variables t and λ. The labeled collection is stationary if, whenever it contains a function, it also contains all its temporal translates, and if it attributes the same probability to two functions that differ by translation.

To this excessively loose description, certain measurability conditions are added solely in order to make it manageable, and in order to distract attention from irrelevant complications that can only present themselves on sets of λ's of measure zero.

This viewpoint cannot be made into a universal prescription. Indeed, for recurrent events with $E(U) = \infty$, the interesting cases are thus thrown into the dustbin. In these cases, to insist upon finite-dimensional distributions, and upon the usual measurability arguments, would have the same effect as if one insisted upon using a three-dimensional Euclidean measure to investigate problems in which all the probability is concentrated upon a one-dimensional line, or upon a surface. There is, however, a very simple way to avoid any such degenerate but meaningless answer, namely, to examine only the distributions conditioned by the assumption that one finds oneself upon the line or surface in question.

This suggests the following broader definition of *a stochastic process and the following conditional concept of stationarity*. Assume, not only that the value of the function $V(t'_0, t''_0)$ is known for some interval (t'_0, t''_0), but that it equals one rather than zero. The sequence $\{T_k\}$ will then be defined by the conditional distribution of the various finite sets of functions $V(t'_h, t''_h)$.

Examples: $t'_0 = t'_1 = t''_0 = t''_1$ means that the position of one of the events T_k is known; $t'_0 \leq \min t'_h$ and $t''_0 \geq \max t''_h$ means that at least one of the events t_k is located within the interval (t'_0, t''_0).

Moreover, $\{T_k\}$ will be called *conditionally stationary* if the conditioned distribution of the V remains unchanged when a nonvanishing increment is added to all the t'_h and t''_h, while V remains constrained by the condition $V(t'_0, t''_0) = 1$.

Innocuous as its definition may seem, this conditional concept may be the key to the necessary task of describing the structure of many empirical intermittent phenomena. It also suggests that various difficulties are likely to arise in applying other probabilistic and statistical arguments that were originally suggested by chance phenomena which are continuing rather than intermittent.

For sporadically varying g.r.f.'s, on the contrary, $\mu[B(u)]$ is unbounded and $G^*(0) = \infty$. In the asymptotically self-similar case,

$$G^*(\lambda) \sim \lambda^{1-D}, \qquad \text{as } \lambda \to 0. \tag{9.8}$$

In interpreting empirical spectral measurements, one may be tempted to handle G^* as if it were a Wiener-Khinchin spectrum. But $G^*(0) = \infty$ would then be interpreted as meaning that there is an infinite energy in low frequencies, which is impossible physically and therefore "catastrophic" for the identification of G^* to a Wiener-Khinchin spectrum. To distinguish this difficulty from high frequency divergences, it is called an "infrared catastrophe." As introduced in the theory of sporadically varying g.r.f.'s G^* is not a spectrum and its divergence is not impossible physically and hence not catastrophic for the theory.

More reasonable definitions of the spectrum will be proposed presently. They will show that, in order for $|G^*(\lambda'') - G^*(\lambda')| / \mu[B(\tilde{u})]$ to be a rough estimate of the energy in the frequency band (λ', λ''), one must assume that $1/\tilde{u} \leq \lambda' < \lambda'' \leq \infty$. In particular, the energy in the band $(1/\tilde{u}, \infty)$ is roughly $G^*(1/\tilde{u})/\mu[B(\tilde{u})]$. If $G^*(0) = \infty$, then both numerator and denominator increase as $\mu^* \to \infty$, but their ratio may well tend to a finite limit. The energy will seem to flow into ever lower frequencies, but the total *expected* energy will remain fixed.

Mandelbrot, Fifth Berkeley Symposium on Probability,1965.