







REVIEW ARTICLE | JULY 30 2024

The challenge of non-Markovian energy balance models in climate

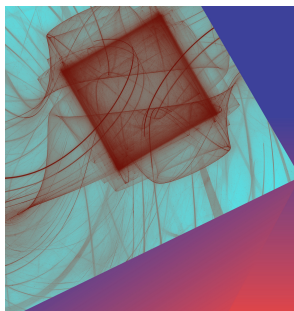
Special Collection: [Nonlinear dynamics, synchronization and networks: Dedicated to Jürgen Kurths' 70th birthday](#)

Nicholas W. Watkins   ; Raphael Calel  ; Sandra C. Chapman  ; Aleksei Chechkin  ; Rainer Klages  ; David A. Stainforth



Chaos 34, 072105 (2024)

<https://doi.org/10.1063/5.0187815>



Chaos: An Interdisciplinary Journal of Nonlinear Science

Focus Issue:

From Sand to Shrimps: In honor of Jason A.C. Gallas

Guest Editors: Marcus W. Beims, Thorsten Pöschel, Pedro G. Lind

[Submit Today!](#)

The challenge of non-Markovian energy balance models in climate

Cite as: Chaos 34, 072105 (2024); doi: 10.1063/5.0187815

Submitted: 16 November 2023 · Accepted: 16 June 2024 ·

Published Online: 30 July 2024



View Online



Export Citation



CrossMark

Nicholas W. Watkins,^{1,2,a}  Raphael Calel,^{1,3}  Sandra C. Chapman,^{2,4,5}  Aleksei Chechkin,^{6,7,8} 
Rainer Klages,^{9,10}  and David A. Stainforth^{1,2}

AFFILIATIONS

¹Grantham Research Institute on Climate Change and the Environment, The London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom

²Centre for Fusion Space and Astrophysics, University of Warwick, Gibbett Hill Road, Coventry CV4 7AL, United Kingdom

³McCourt School of Public Policy, Georgetown University, 125 E St. NW, Washington DC 20001, USA

⁴Department of Mathematics and Statistics, Faculty of Science and Technology, UiT—The Arctic University of Norway, N-9037 Tromsø, Norway

⁵International Space Science Institute, Hallerstrasse 6, 3012 Bern, Switzerland

⁶Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam-Golm, Germany

⁷Faculty of Pure and Applied Mathematics, Hugo Steinhaus Center, Wrocław University of Science and Technology, Wyspińskiego 27, 50-370 Wrocław, Poland

⁸Akhiezer Institute for Theoretical Physics National Science Center “Kharkiv Institute of Physics and Technology,” 61108 Kharkiv, Ukraine

⁹Centre for Complex Systems, School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom

¹⁰London Mathematical Laboratory, 8 Margravine Gardens, London W6 8RH, United Kingdom

Note: This paper is part of the Focus Issue on Nonlinear dynamics, synchronization and networks: Dedicated to Juergen Kurths' 70th birthday.

^a**Author to whom correspondence should be addressed:** nickwatkins62@fastmail.com

ABSTRACT

We first review the way in which Hasselmann's paradigm, introduced in 1976 and recently honored with the Nobel Prize, can, like many key innovations in complexity science, be understood on several different levels. It can be seen as a way to add variability into the pioneering energy balance models (EBMs) of Budyko and Sellers. On a more abstract level, however, it used the original stochastic mathematical model of Brownian motion to provide a conceptual superstructure to link slow climate variability to fast weather fluctuations, in a context broader than EBMs, and led Hasselmann to posit a need for negative feedback in climate modeling. Hasselmann's paradigm has still much to offer us, but naturally, since the 1970s, a number of newer developments have built on his pioneering ideas. One important one has been the development of a rigorous mathematical hierarchy that embeds Hasselmann-type models in the more comprehensive Mori–Zwanzig generalized Langevin equation (GLE) framework. Another has been the interest in stochastic EBMs with a memory that has slower decay and, thus, longer range than the exponential form seen in his EBMs. In this paper, we argue that the Mori–Kubo overdamped GLE, as widely used in statistical mechanics, suggests the form of a relatively simple stochastic EBM with memory for the global temperature anomaly. We also explore how this EBM relates to Lovejoy *et al.*'s fractional energy balance equation.

© 2024 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0187815>

Hasselmann's stochastic energy balance model embodies a long-standing paradigm in climate science. Recent evidence that climate may show long-range memory has, thus, motivated newer stochastic models such as Lovejoy *et al.*'s fractional energy

balance equation (FEBE). Physical arguments for extending the Hasselmann formalism are more general than this, however, and are motivated by many other well-known features of the system, including periodicity. We, thus, need a model that allows a

more general range of dependency structures than either just the shortest possible or the longest possible ranged behaviors while allowing both as limiting cases. In this article, we propose the use of such a formalism to extend Hasselmann's EBM, the stochastic Mori-Kubo generalized Langevin equation (GLE).

I. INTRODUCTION

The value of climate science would be substantial even without anthropogenic interference in the climate system. The presence of such an interference, however, raises a number of categorically new challenges. The context of human-induced climate change has made it important to quantify the observed changes, understand their origin, and, as far as possible,^{1,2} predict their future scope, in the presence of the nontrivial intrinsic variability that is already present (see, e.g., Refs. 3 and 4). The development of this need has coincided with the rise of the science of complexity, which has itself both pioneered and promoted new approaches to similar problems across a very wide range of topical application areas.⁵

Several key achievements in the study of fluctuations, at the interface^{6–8} of climate modeling and the fast-developing sciences of complexity, have recently been recognized by the award of the 2021 Nobel Prize for Physics to Hasselmann, Manabe, and Parisi. The Nobel Committee have given a very useful and accessible short introduction⁹ to these branches of science, but two more recent articles have demonstrated between them an emerging dichotomy in responses to the Prize. Our paper seeks to reconcile them.

In Ref. 10, Franzke *et al.* have shown how “the Hasselmann program,” first enunciated in Ref. 11, of explaining how slow climate arises from fast weather using stochastic EBMs, has given rise to decades of productive and insightful research. The original stochastic EBMs were Markovian, depending only on the current values of their (stochastic) variables, and by construction assumed strongly separated fast and slow time scales. A reader might easily infer from this excellent review, however, that the broader “Hasselmann program” built around such stochastic EBMs, their developments, and his subsequent spatiotemporal fingerprinting methods has fully solved the problems of realistic climate modeling. This impression is reinforced by the fact that the paper does not delve into more recent, and maximally non-Markovian, EBMs that exhibit “long-range memory” (LRM; though see a complementary review of this by Franzke *et al.*,¹² and recent work by him with his colleagues¹³).

In contrast, Lovejoy¹⁴ has stressed the differences between Hasselmann's Markovian paradigm and more recent developments in complexity science, especially those related to LRM and fractional responses. He has articulated^{15–17} the view that only models incorporating LRM can fully describe the already known properties of climate data. Arguments for the relevance of LRM have at times been received with skepticism by some climate scientists,¹⁸ while others (e.g., Ref. 19), who do acknowledge its importance, have nevertheless emphasized the deterministic behavior present in the climate system because of the predictability that it affords. Recently, however, the Nobel Committee has remarked⁹ on an increasing awareness that “on decadal timescales of relevance to humans, evidence is consistent with the memory becoming effectively infinite . . .”

An AGU editor highlight (Ana Baros, “Eminently Complex – Climate Science and the 2021 Nobel Prize,” January 1, 2023) described Ref. 14 as making “*the case that for the last 100 years there has been little communication between atmospheric sciences and nonlinear geophysics, and this must change to improve useful predictability (i.e., lead time) of climate models.*” In this paper, we address the need for communication between these two fields in a number of ways.

(i) First, by reviewing the topic, we will show that the situation is in some ways already better than it may appear, because, in fact, there has been work, notably that of Ref. 20, reviewed in Ref. 8, which, while proceeding directly from the Hasselmann paradigm, has allowed the construction of models with a non-Markovian structure, via the use of memory kernels. It seems not to have always been very accessible to a wider audience, perhaps because it exploits the Mori-Zwanzig formalism,²¹ well known in applied mathematics, data assimilation, and statistical mechanics, to decompose the effect of fast degrees of freedom into a memory term and a noise-like term. The fundamental (and nonlinear) equation that expresses the Mori-Zwanzig decomposition is an identity, and deterministic, but confusingly, to those already familiar with the linear stochastic equation first studied by Langevin in 1908; it is sometimes^{8,10,20–23} referred to as the generalized Langevin equation (GLE). This Mori-Zwanzig GLE is intractable without further approximations²¹ or other modifications.⁸

(ii) In view of this intractability, our paper takes an alternative, more direct stochastic approach to an EBM with memory, based on an already existing extension of Langevin's original equation²⁴ of 1908. We advocate for the use of Mori's and Kubo's stochastic GLE, widely known and applied since the 1960s in condensed matter theory, quantum optics, and physical chemistry.^{25–29} We show how the form of the Mori-Kubo GLE, occasionally known as the non-Markovian Langevin equation,³⁰ suggests an analogous generalized Hasselmann stochastic EBM for the global temperature anomaly. Our paper expands on and revises our contribution to the 2020 International Conference on Complex Systems,³¹ which was itself expanded as Ref. 32.

(iii) We go on to show that Hasselmann's model is one limiting case of our proposed EBM, that of short-ranged memory, while Lovejoy's fractional model, FEBE,^{16,33} is closely related to the opposite special case, long-ranged memory. FEBE differs only by using a Riemann-Liouville fractional derivative, instead of a Caputo. We, thus, bridge these two seemingly opposed perspectives. We further note that an EBM based on the Mori-Kubo GLE can handle not just the two extremes of delta-correlated (shortest ranged) or power law decaying (longest ranged) memory kernels but also those with intermediate-ranged memory such as that which arises from a few widely separated timescales. In addition, modifications of such an EBM could, in principle, accommodate many types of periodic (or quasiperiodic) determinism that are present, such as ENSO. As early as 1976, Mitchell³⁴ asserted that “*there are no known deterministic mechanisms of climatic variability that are internal to the atmosphere . . . or any other part of the climatic system.*” From this viewpoint, the emergence of any deterministic mode(s), even if nonlinear and only quasiperiodic, is, thus, crucial to predictability³⁵ and its inclusion essential in any compact description.

(iv) We conclude by noting that an EBM based on the Mori–Kubo GLE would have the option of embodying a fluctuation–dissipation relation (FDR) *by construction*, because this was the main original purpose of their stochastic GLE.

We hope our contribution will open a new channel of communication between two disparate worlds and will facilitate the construction of a new class of EBMs, which will, in turn, improve the study of climate predictability.

In Sec. II, we will first briefly recap the basic notions of EBMs and then discuss some ways in which nonlinearity and stochasticity have been incorporated into them.

II. ENERGY BALANCE MODELS

A. Nonlinearity in EBMs

The first energy balance models^{36–38} of Budyko and Sellers were deterministic. As the name implies, they captured how the earth’s surface temperature T will, after a perturbation, evolve in time t to restore the balance between the incoming $R \downarrow$ and outgoing $R \uparrow$ heat fluxes at the earth’s surface, expressed by the following equation:

$$C_p \frac{dT}{dt} = R \downarrow - R \uparrow. \quad (1)$$

Here, C_p is the effective heat capacity per unit area.

The Budyko–Sellers model in its simplest form then makes the outgoing and incoming radiation terms in (1) explicit. The former is modeled with the Stefan–Boltzmann law and the latter includes an albedo feedback function $a(T)$,

$$C_p \frac{dT}{dt} = S_0(1 - a(T)) - \varepsilon \sigma_{SB} T^4, \quad (2)$$

where S_0 is the mean solar radiation incident per meter squared at the earth’s surface,³⁶ ε is the earth’s emissivity, and σ_{SB} is the Stefan–Boltzmann constant.

Since the 1960s, EBMs have been explored and extended in many directions. In particular, much work has been done on nonlinear and stochastic climate dynamics in this framework, one topic of the recently edited book by Franzke and O’Kane.⁶

Considerable attention has been focused on the nonlinearity of the complicated potential created by the interplay of the T^4 term and the albedo feedback, for example. Taking a quadratic choice for the latter gives the model studied by Fraedrich,³⁹

$$C_p \frac{dT}{dt} = \frac{1}{4} \mu I_0 (1 - a_2 + b_2 T^2) - \varepsilon \sigma_{SB} T^4, \quad (3)$$

where I_0 is the solar constant, μ is a fractional premultiplier³⁶ that gives the radiation incident at the earth’s surface, and the quarter results in an average over the planet’s surface.

B. Stochasticity in EBMs

Stochasticity can be introduced into nonlinear EBMs in a number of ways, such as white noise $\xi(t)$ with a controllable variance σ^2 added as an extra term to Eq. (3) by Sutera.⁴⁰ This changes it from an Ordinary Differential Equation (ODE) to a stochastic differential equation (SDE), an important subtlety.

The Fraedrich–Sutera model is Markovian because its deterministic part depends only on the current value of its single variable, the temperature $T(t)$, while its stochastic part, Gaussian white noise, is delta-correlated and, thus, also has no memory of its previous state.

The marginal distributions of such a Markovian climate model need not be Gaussian, however.⁴¹ Reference 42 is a comprehensive study of α -stable Lévy noise of the more recent Ghil–Sellers extension to the original EBMs. The PDF of the noise here has an infinite variance and is heavy-tailed.

C. Memory in EBMs

The Markovian restriction on the earliest EBMs has since been relaxed in several different ways. A pioneering approach was that of Bhattacharya *et al.*,⁴³ who introduced memory into the deterministic part via a modification of the albedo term. The resulting work has been based on delay differential equations.⁴⁴

Multi-box (and, thus, multivariate) stochastic models have incorporated memory effects through the presence of a range of feedback time scales in their different “boxes.”^{45–47}

Most of the remainder of this paper concerns memory in stochastic EBMs, but in Sec. III, we first examine the simplest linear Markovian stochastic models in more detail. In particular, we consider the univariate autonomous and non-autonomous versions in which they have been used to model the earth’s global temperature anomaly.

III. THE HASSELMANN PROGRAM I: LINEARIZED MARKOVIAN STOCHASTIC EBMs

Linear Markovian stochastic EBMs were explored in detail in the second of Hasselmann’s classic mid-1970s papers,⁴⁸ complementing the treatment in the first paper,¹¹ which was largely in PDE (Fokker–Planck) rather than SDE (Langevin) terms. They can be motivated in several ways.

The first is an ansatz and was essentially the approach followed by Mitchell in the 1960s (Sec. III). It recognizes that the discrete time autoregressive first-order Markov process $AR(1)$, and the continuous time Langevin Eq. (9) whose solution is called the Ornstein–Uhlenbeck (O–U) process, are the simplest possible linear stochastic models of the response to perturbation.⁴⁹

A more physically intuitive argument, summarized in Sec. III B, and given more fully by Lemke in Ref. 50, comes from linearizing deterministic EBMs of the Budyko–Sellers type and adding white noise. At this level of abstraction, it is already apparent that the result is what we will refer to as Hasselmann’s autonomous SDE. However, Hasselmann’s SDE is isomorphic to not one but two such equations arising in Brownian motion (BM); so Sec. III C describes the dilemma posed by the choice of possible mappings.

Hasselmann’s mathematical analogy with Brownian motion was not limited simply to mapping an EBM to a Langevin equation. The most sophisticated and abstract level of his program, proposed first in Ref. 11, replaced Mitchell’s ansatz, which has sometimes been seen as being purely statistical, with a postulate, inspired by statistical physics: that of the separation of fast and slow timescales, and the identification of the former with weather and the latter with climate.

It is very important to realize that Hasselmann's postulate is potentially quite generally applicable to climate variables and not just to the global mean temperature (GMT) anomaly—it was not simply a way to make a stochastic EBM, but to kick off a research program.

A. Mitchell's stochastic models of 1966

Mitchell's pioneering chapter,⁵¹ in a RAND Corporation conference proceedings volume from December 1966, proposed two stochastic climate models.

The first, as he put it, bore “*primarily on the origins of relatively rapid fluctuations, as, for example, those reflected in the interannual variability of temperature observed at a wide variety of locations during the past century.*” He postulated that the air temperature anomaly in a marine environment was composed of two additive parts, one proportional to the sea surface temperature (SST) anomaly and one independent of it. He then gave arguments suggesting that each component's data, comprising monthly values, would follow first-order Markov or “AR(1),” stochastic processes. An AR(1) process is defined at discrete times; in this case, the set $t, t + dt, t + 2dt, \dots$, and takes the following form:

$$x(t + dt) = \phi x(t) + \xi(t), \quad (4)$$

where $0 < \phi < 1$ and $\xi(t)$ are an iid normal random variable of zero mean and unit variance, respectively.

Mitchell derived the standard results for the autocorrelation function and power spectrum of the resulting compound first-order linear Markov process and pointed out that the measurements of SST implied that its ϕ would be in the range 0.5–0.9, giving a red noise behavior. He also fitted the compound model to the famous 269-year (1689–1957) Manley series of monthly temperatures for central England, illustrating how the two component model's Autocorrelation Function (ACF) gave a good description of the observed one. Hasselmann was made aware of Mitchell's work by a referee and took care to credit it, both in Ref. 11 and subsequently, e.g., the oral history of Ref. 52.

Mitchell's second model operated on longer time scales than his first and is less directly relevant to Hasselmann's work. He viewed it as bearing “*contrastingly on the origin of much longer fluctuations, and*” thought that it might “*have some unexpected relevance to the origins of the glacial-interglacial succession of the Pleistocene ice age.*” In a sense, it anticipated the previous section's Fraedrich–Sutera model and is further evidence of how far Mitchell was ahead of his time.

B. Adding noise to a linearized Budyko–Sellars model

Following Ref. 9, we start with Budyko–Sellars model (2) and expand T about the steady state average surface temperature T_S ,

$$T = T_S + \Delta T.$$

So,

$$C_p \frac{d(T_S + \Delta T)}{dt} = S_0(1 - a(T_S + \Delta T)) - \varepsilon \sigma_{SB}(T_S + \Delta T)^4,$$

from which we can subtract the corresponding steady state values to get

$$C_p \frac{d\Delta T}{dt} = -S_0 \frac{\partial a}{\partial T} \Delta T - 4\varepsilon \sigma_{SB} T_S^3 \Delta T + O(\Delta T^3),$$

and, thus, dropping terms in ΔT of third order and above, we obtain

$$C_p \frac{d\Delta T}{dt} = - \left[4\varepsilon \sigma_{SB} T_S^3 - S_0 \left| \frac{\partial a}{\partial T} \right| \right] \Delta T = -\Lambda \Delta T, \quad (5)$$

where we have additionally used the fact that albedo sensitivity, $\partial a / \partial T$, is negative and have defined a feedback parameter Λ .

As in the Fraedrich–Sutera model, high frequency fluctuations are then introduced as additive white noise $\xi(t)$, a standard Gaussian random variable (with zero mean and unit variance) to give Hasselmann's stochastic EBM in its SDE form,

$$C_p \frac{d\Delta T}{dt} = -\Lambda \Delta T + \sigma \xi(t). \quad (6)$$

This SDE is Markovian in the sense that it depends only on the current value of its variable, which is now the global mean temperature (GMT) anomaly ΔT . This SDE is a Langevin equation governing the OU process $\Delta T(t)$. It expresses the competition between the Gaussian fluctuations and the mean-reverting $-\Lambda \Delta T$, which produces a dynamic steady state.

There is also a non-autonomous version of Hasselmann's model, in which $F(t)$ represents the deterministic part of forcing, in particular, that arising from anthropogenic carbon emissions,

$$C_p \frac{d\Delta T}{dt} = F(t) - \Lambda \Delta T + \sigma \xi(t). \quad (7)$$

Equations (6) and (7) are formal and, like all SDEs, contain objects whose meanings must be made explicit and rigorous. For white noise-driven SDEs, this process has taken decades of mathematical work in developing the now mature field of stochastic calculus.^{21,53} In that Markovian case, there is a widely used convention to write SDEs as differentials rather than derivatives. We do not use this notation here because many of the equations we study have colored or even fractional noise driving terms for which there is less of a notational consensus.

C. Relation of EBM to the equations of Brownian motion

Hasselmann was directly inspired by a conceptual and mathematical analogy with Brownian motion. As he observed in an AIP oral history interview,⁵² “my stochastic model is an application of the concept of Brownian motion . . . One of the simplest stochastic processes. The idea that one could explain long-term climate variability very simply by the short term fluctuations of the atmosphere in analogy with Brownian motion, is really rather obvious and I thought I would write it up somewhere in a little note.”

His paper explored an important distinction between the mathematical and physical usages of the term Brownian motion (BM). For mathematicians, the Wiener process defines BM, while for physicists, the underdamped LE defines BM. Hasselmann showed that the former was already insightful in addition to the then-current

paradigms for fluctuations in climate but that the latter was necessary to create a power spectrum that more closely resembled those already being observed.

Interestingly, however, there are not one but two possible SDEs in physical BM to choose from, the underdamped and overdamped Langevin equations. If Hasselmann had proceeded purely by mapping one of these onto an EBM, he could ostensibly have used either, as we now show. While this ambiguity was relatively unimportant to his work, we will then argue that it becomes significant when thinking about extensions to the Hasselmann EBM.

1. Underdamped Langevin equation

The fundamental SDE used in Brownian motion is the underdamped Langevin equation,

$$M \frac{dv}{dt} = -V'(y) - \gamma v + \sigma \xi(t). \quad (8)$$

It describes the acceleration dv/dt , where velocity $v = dy/dt$, of a particle of mass M , which experiences a deterministic force due to a potential $V(y)$, a stochastic force $\sigma \xi(t)$, and a linear damping force $-\gamma v$.

If we (i) map velocity into GMT anomaly, and mass into heat capacity C_p , (ii) ignore the question of mapping and meaning of the position variable y , and (iii) replace the derivative of a potential by deterministic forcing we obtain the non-autonomous version (7) of the Hasselmann EBM.

Autonomous version (6), with $F(t) = 0$, is also frequently studied, e.g., Ref. 54, where an attempt has been made to remove the prompt effects of the deterministic forcing from the time series.

2. Overdamped Langevin equation

We now consider forced Brownian motion in a potential. If the mass is very large and the particle motion is close to a minimum of the potential, we can neglect the acceleration term in (8) and set $V'(y) \approx -ky$. This gives the overdamped Langevin equation,

$$\gamma \frac{dy}{dt} = -ky + \sigma \xi(t). \quad (9)$$

The solution of this for the position of the overdamped particle is the O-U process.

Equation (9), thus, offers an alternative conceptual starting point for a stochastic EBM, where instead of velocity, it is the position that is mapped to ΔT . The result is an autonomous equation,

$$C_p \frac{d\Delta T}{dt} = -\Lambda \Delta T + \sigma \xi(t). \quad (10)$$

Although mathematically identical to the unforced version of (7), rather than the mass M being mapped to the heat capacity C_p , it is now the friction constant γ , which is so mapped. This is why although γ is usually absorbed into \bar{k} and $\bar{\sigma}$, did not do so in Eq. (9). Another conceptual difference is that the linear term in ΔT of Eq. (10) arises from mapping a linearized potential rather than a friction coefficient.

The overdamped Langevin equation can be modified to study problems with additional deterministic forcing of the velocity, such

as the ion trap modeled by Ref. 55, and so its analogous Hasselmann EBM could also have an additional deterministic forcing term.

3. Which of the underdamped or overdamped Langevin equations is closer to Hasselmann's EBM ?

The conceptual differences between the underdamped and overdamped Langevin equations start to become important when we try to derive EBMs from first principles. This is even more true if we want to extend them to include effects such as long-range memory and fluctuation-dissipation relations. We illustrate this in Sec. IV by considering Hasselmann's own motivation¹¹ of his stochastic model embodied in Eq. (10), and the outline derivation of it offered by the Nobel Committee.⁹ We will find that the latter, in particular, has more in common with the overdamped than the underdamped Langevin equation.

IV. THE HASSELMANN PROGRAM II: FLUCTUATIONS, FAST AND SLOW

A. Assumptions of the Hasselmann program

A modernized version of Hasselmann's argument is sketched by the Nobel Committee,⁹ who describe it as an "interpretative and notationally uncluttered outline." We follow it and its notation here, supplemented by reference to Hasselmann.¹¹

1. Fast and slow variables

Hasselmann assumes that the instantaneous state of the atmosphere-ocean-cryosphere-land system is described by a finite set of discrete variables $\mathbf{z} = (z_1, z_2, \dots)$. He then assumes that its evolution is described by a series of prognostic equations like

$$\frac{dz_i}{dt} = w_i(\mathbf{z}) \quad (11)$$

and then divides the system into two subsystems $\mathbf{z} = (\mathbf{x}, \mathbf{y})$.

The set of fast "weather" variables \mathbf{x} have a response time scale τ_x of the order of a few days and evolve under one set of ODEs,

$$\dot{x}_i = f_i(x, y), \quad (12)$$

while the slow "climate" variables \mathbf{y} , such as sea surface temperature, ice coverage, and land foliage, have a longer response time scale τ_y , of the order of months, years, or more:

$$\dot{y}_i = g_i(x, y). \quad (13)$$

Here, f, g correspond to u, v in Hasselmann's original notation.¹¹

2. Conditional average over fast dynamics

In Ref. 11, Hasselmann then discusses the key idea of averaging over the fast dynamics. The Nobel Committee's argument⁹ writes the dynamics of the fast variables in terms of the slow ones and is simplified without the loss of generality by considering scalar x and y ; hence,

$$x = \langle x|y \rangle + x^*, \quad (14)$$

where following Hasselmann, the first term is an ensemble average of a set of values of x for a given y , i.e., a conditional average, and the second term defines a fluctuation with respect to this average.

B. The Nobel Committee's derivation

1. Drift force and multiplicative noise

At this stage, the Nobel Committee⁹ gives a simpler but more explicit argument for a Langevin-based EBM than Hasselmann's own. His original paper, in contrast, dealt with both Langevin and Fokker–Planck approaches.

Their approach is to substitute for x using Eq. (14) in the equation for the slowly varying dynamics \dot{y} , and then, Taylor expand in the relatively small but fast fluctuations x^* ,

$$\begin{aligned}\dot{y} &= g(x, y) \\ &= g(x|y) + x^*, y \approx g(x|y, y) + \partial_x g(x|y, y)x^*.\end{aligned}$$

One can then model the slowly varying first term as a drift force arising from a potential $U(y)$, while the fast fluctuations are modeled as white noise $\xi(t)$ whose strength $\bar{\sigma}$ is dependent on the value of y , i.e., this noise is not additive but multiplicative,

$$\dot{y} = -\frac{dU(y)}{dy} + \bar{\sigma}(y)\xi(t). \quad (15)$$

2. Linearizing the drift force and assuming additive noise

To transform the more general model Eq. (15) into the specific Hasselmann form with linear damping and additive noise, the Nobel Committee then argues that noise intensity is small, so its amplitude can be taken to be approximately constant,

$$\bar{\sigma}(y) \approx \bar{\sigma} \quad (16)$$

and that in the case of interest, the system loiters near a dynamical fixed point y_E around which the potential and, thus, the drift force can be expanded,

$$U(y) = U(y_E) + (y - y_E)\frac{dU(y_E)}{dy} + \frac{1}{2!}(y - y_E)^2\frac{d^2U(y_E)}{dy^2} + \dots$$

and linearized by dropping terms above second order in $y - y_E$.

The SDE for the resulting dynamics then has the well-known O–U solution, but we can see from the above arguments that the SDE must be in the variable $y^* = y - y_E$, which tracks the relative distance from the fixed point, rather than y ,

$$\frac{dy^*}{dt} = -\bar{\gamma}y^* + \bar{\sigma}\xi(t). \quad (17)$$

The result can be compared to the autonomous Hasselmann model (10), which is in ΔT , not T .

V. STOCHASTIC EBMs AND IAMs

Economic assessments of climate change typically rely on a suite of integrated assessment models (IAMs) descended from the work of another Nobel laureate, Nordhaus.^{56,57} These models have now been refined and extended in many directions—from increasing their spatial and sectoral resolutions^{58–60} to explicitly modeling technological choices⁶¹ to incorporating agent-based and non-equilibrium models of the economy.^{62–64} Yet, although Hasselmann himself contributed to the early development of IAMs,^{65,66}

the climate modules of IAMs are generally based on deterministic EBMs.⁶⁷

Climate stochasticity, when it is acknowledged at all, most commonly enters in the form of uncertain temperature thresholds (effectively “tipping points”) at which the model's parameter values suddenly change.⁶⁸

One strand of economic research has added noise to the temperature dynamics directly.^{69–71} The main objective has been to study how fast we can learn the model's parameter values—the equilibrium climate sensitivity, in particular—when we observe a noisy temperature time series. Anticipated learning tends to reduce the effect of parameter uncertainty on the optimal mitigation policy, but there is no agreement on whether this anticipated learning-channel is quantitatively meaningful.⁷²

One possible reason that the research on IAMs has paid so little attention to temperature stochasticity is that this source of uncertainty does not affect the *marginal* warming resulting from an additional tonne of carbon emissions. As a consequence, temperature stochasticity itself has very little effect on the optimal mitigation policy.^{73,74} However, a recent paper by Calel *et al.*,⁷³ which perhaps offers the first example of a Hasselmann EBM coupled to an economic damage function, points out that temperature variability could still greatly affect optimal investment in adaptation. They conclude that a forward-looking social planner would be willing to make substantial investments to avoid or reduce the effects of temperature fluctuations. Evidently, Hasselmann's “obvious” idea still has much to teach us.

VI. BUT SHOULD STOCHASTIC CLIMATE MODELS REALLY BE MARKOVIAN?

A. Leith's “infrared climate problem”

In addition to Mitchell and Hasselmann, a third great climate pioneer, Leith, argued for the use of Langevin models in the mid-1970s. In his case, this was explicitly because of their connection with fluctuation–dissipation relations. As he put it at an NCAR conference on statistics and climate in 1994,⁷⁵ “a stochastic climate model that incorporates the fluctuation–dissipation relation can be devised and would provide a crude estimate of climate sensitivity. This model would be of the Langevin type.”

However, by then, he had identified a problem of enduring relevance that “there is evidence that this kind of model, would not on its own be satisfactory to capture some of the low frequency phenomena observed in the atmosphere.” He referred to this, in a phrase that has sadly not gained currency, as the “infrared climate problem.” This was not a literal reference to thermal radiation but rather to an “infrared catastrophe” in fluctuations in temperature, i.e., the “piling up of extra variance at low frequencies.”

The Nobel Committee⁹ illustrates and acknowledges the continuing relevance of Leith's worry. It points out that the Markovian nature of Hasselmann-type models results in an exponential decay of their autocorrelation functions,

$$R(\tau) = \langle y^*(t)y^*(t + \tau) \rangle. \quad (18)$$

Such a model predicts that the typical power spectrum of a climate variable would be Lorentzian,

$$P(\omega) = \bar{\sigma}^2 / (\bar{\gamma}^2 + \omega^2) \quad (19)$$

and that the ratio of strength $\bar{\sigma}$ of the rescaled noise term and, $\bar{\gamma}$, the rescaled damping term would control the spectrum's shape. The power spectral density would, thus, be flat at low frequencies, a property that was reported in early comparisons of the Hasselmann paradigm with data from weather ships.

B. Phenomenological evidence for LRM in climate

The simplest stochastic process with memory, AR(1) Eq. (4), is Markovian, being formulated only in instantaneous variables. It can additionally be given random initial conditions, although this is not always the relevant case for climate, where, for example, the initial temperature anomaly may be set to zero.

The emergent physics of the earth system on larger scales need not be similarly memoryless, though, and in fact, the Nobel Committee's summary states that in contrast to the atmosphere "*the ocean has a very long memory of events, at least in the hundreds . . . 1000 year range.*" As we noted in our introduction, the committee then remarks on the increasing awareness of effectively infinite memory and that "*this suggests a potential self-similarity or fractal character,*" citing Moon *et al.*'s study⁷⁶ in support of their statement. Moon *et al.* in their turn refer to the by now quite extensive literature on long-range dependence in climate, from several research teams. We note, in particular, the groups at Tromsø (Refs. 47, 77, and 78) and Montreal,^{16,33,79} and the recent PAGES CVAS group.¹²

In Ref. 14, Lovejoy addressed the issue of LRM directly. He compared Frankignoul and Hasselmann's original Markovian power spectral fit⁴⁸ to 16 years of sea surface temperature data from the Ocean Weather Ship India, with the data now available from 452 stations over about 100 years. He argued that a compound power law fit $S(f) \sim f^{-0.6}$ at low frequencies and $\sim f^{-1.8}$ at higher frequencies is much more satisfactory.

C. Theoretical approaches to LRM in climate

In other papers,^{16,33} Lovejoy and colleagues have argued for a fractional energy balance equation (FEBE) as a method for incorporating long-range memory into EBMs.

In Sec. VII, we will emphasize that there was already a well-studied and standard way, the Mori–Kubo GLE,^{80–83} to incorporate memory in a Langevin equation. The Mori–Kubo GLE is suitable for use as a univariate stochastic equation, and thus, to add memory to EBMs of the Hasselmann type. We show that FEBE, but with a Caputo fractional derivative rather than the Riemann–Liouville type used by Lovejoy *et al.*, corresponds to one particular limit of this equation.

VII. WHAT IS A GENERALIZED LANGEVIN EQUATION?

As frequently happens when mathematics and physics intersect, the terminology used in this subject varies, which has sometimes caused confusion.

An example is the otherwise excellent Nobel Committee article,⁹ which describes the SDE corresponding to the O–U process as a "generalized Langevin equation." It is literally so, but only in the

relatively limited sense that both the underdamped and overdamped Langevin equations have the same mathematical form, and that any other linearly damped model with additive white noise will also obey an SDE of this same form, whether physical in origin or not.

Two other equations which much more substantially merit the description "generalized" are

- the equation which has been widely studied in applied mathematics as a concise statement of the Mori–Zwanzig decomposition, and which is occasionally referred to as the GLE^{8,22} and
- the stochastic Mori–Kubo GLE which has found widespread application in statistical physics.

The Mori–Zwanzig decomposition has played an important conceptual role in formalizing an approach to modeling the dynamics of a few preferred degrees of freedom in the presence of many other unobserved ones in many areas of science, notably quantum mechanics. The resulting equations are sometimes directly derived from Hamiltonian equations of motion for the whole system, for example, in a reservoir plus subsystem model.²² The technique is not restricted to such cases, though, and it has also become arguably the most influential paradigm for thinking about how to make the Hasselmann program mathematically rigorous and has been employed in climate science to order empirical modes²⁰ (see also Ref. 8).

In contrast, throughout the statistical mechanics community, the more standard usage of the term GLE^{25–29} is to refer to the linear integrodifferential Langevin equation of Mori and Kubo with a memory kernel on its velocity term,

$$M \frac{d^2 y}{dt^2} = -V'(y) - \int_{t_m}^t dt' \gamma(t-t') \frac{dy(t')}{dt'} + v(t), \quad (20)$$

where the Markovian friction term $\gamma dy/dt$ is replaced by a non-Markovian integral over past values of the velocity starting at a "memory time" t_m (cf. Ref. 55).

In the Mori–Kubo formalism, the autocorrelation function of noise $v(t)$ is proportional to $\gamma(t-t')$ with the prefactor depending on the temperature of the medium

$$\langle v(t)v(t') \rangle = k_B T \gamma(t-t'),$$

where T is the temperature of the medium and k_B is the Boltzmann constant. It is, thus, non-white noise.

VIII. THE MORI-KUBO GENERALIZED LANGEVIN EQUATION: A NEW APPROACH TO THE INFRARED CLIMATE PROBLEM?

A. The overdamped Mori–Kubo GLE

Just as there is an overdamped Langevin equation whose solution is the O–U process (9), there is a corresponding overdamped generalized Langevin equation, shown here for motion near the minimum of the potential $V(y)$,

$$\int_{t_m}^t dt' \gamma(t-t') \frac{dy(t')}{dt'} = -ky + v(t). \quad (21)$$

We note that the memory kernel is on the left hand side of this equation, on the term in dy/dt . We also remark that the same FDR as that for the underdamped Mori–Kubo GLE also applies here. The

power spectrum of the noise $v(t)$ is matched to that of the kernel $\gamma(t-t')$, and so as above, $v(t)$ is no longer white noise.

It is important to realize that while mathematically this is adding memory to the Langevin model, physically it is simply making explicit something that was already known, the presence of finite collision times in physical Brownian motion.

B. An EBM with memory

The significance of the Mori–Kubo GLE for our purposes is that it allows us to suggest a Mori–Kubo GLE-based extension of the Hasselman EBM. We show our suggested EBM below in its autonomous version,

$$\int_{t_m}^t dt' C_p(t-t') \frac{d\Delta T(t')}{dt'} = -\Lambda \Delta T + v(t). \quad (22)$$

We remark that an integral over a time dependent heat capacity $C_p(t-t')$ has replaced the previous constant C . We do not prescribe the form of the noise, so the notation $v(t)$ does not presume that a fluctuation–dissipation relation exists.

C. A Mori–Kubo GLE with power law kernel: The fractional Langevin equation

An interesting special case of the overdamped Mori–Kubo GLE arises when we take a decaying power law form for its memory kernel

$$\gamma(t-t') = \frac{(t-t')^{-\alpha}}{\Gamma(1-\alpha)} \quad (23)$$

instead of an arbitrary memory function.

The resulting integrodifferential SDE is written as (for noninteger α)

$$\frac{1}{\Gamma(1-\alpha)} \int_{t_m}^t dt' (t-t')^{-\alpha} \frac{dy(t')}{dt'} = -ky + v_\alpha(t), \quad (24)$$

and is known in the physics literature as the overdamped fractional Langevin equation (FLE). Both overdamped and underdamped cases of the FLE have been extensively studied^{84–87} following the pioneering investigation by Mainardi and Pironi⁸⁸ of the $\alpha = 3/2$ case.

The notation v_α here acknowledges the presence of a relationship, the FDR,^{89,90} between the noise and the damping kernel.

D. Fractional derivatives

The integrodifferential form of the FLE is sometimes simplified by the use of a fractional derivative. Fractional calculus forms a large and growing field of research in itself (see, e.g., Refs. 91–93), so to aid the reader, we have included the Appendix to summarize those aspects we have used here. We again emphasize that the FLE is only one special case of the more general Mori–GLE, and that we remain “agnostic” about its relevance to the climate system, preferring to suggest a model with a more flexible memory kernel.

The (left-handed) Caputo derivative of order α for a suitable function $f(x)$ is defined^{94,95} for $0 < \alpha < 1$ as

$${}_a^C D_x^\alpha f(x) = \frac{1}{\Gamma(1-\alpha)} \int_a^x du f'(u) (x-u)^{-\alpha}, \quad (25)$$

where $f'(u) = df/du$.

As well as its order α , the Caputo derivative has a second parameter, the lower limit a of the integral in its definition. The Caputo fractional derivative has physically more intuitive initial conditions than the most widely used alternative to it, the Riemann–Liouville derivative. This is because they are specified as integers rather fractional powers of physical quantities. However, as noted by Wei *et al.*,⁹⁵ the problem can be circumvented by the Schneider–Weiss formulation of the Riemann–Liouville derivative.

The Caputo derivative also has a nice property that while of fractional order for $\alpha < 1$ it becomes the familiar integer derivative in the $\alpha = 1$ case, i.e., ${}_a^C D_x^1 f(x) = df/du$, as is shown in Sec. 2.4.1 of Ref. 92. A Caputo FLE can, thus, be defined as

$${}_m^C D_t^\alpha y = -ky + v_\alpha(t) \quad (26)$$

that straightforwardly encompasses the familiar Langevin equation as its $\alpha = 1$ case. This equation and its close relatives are increasingly being studied in several disciplines,^{96–100} typically under the name “Caputo fractional SDE,” though typically specialized to the case of white noise driving.

E. The LRM EBM special case and its relation to FEBE

In the same way, we can postulate a decaying power law form for the kernel in the EBM above

$$C(t-t') = \frac{(t-t')^{-\alpha}}{\Gamma(1-\alpha)} \quad (27)$$

instead of an arbitrary memory function and can, thus, write the integral over the kernel in our EBM as a Caputo fractional derivative, giving an autonomous version as follows:

$${}_m^C D_t^\alpha \Delta T = -\Lambda \Delta T + v_\alpha(t). \quad (28)$$

In this picture, the external climate forcing can then be incorporated as an additional term $F_2(t)$; so, we obtain

$${}_m^C D_t^\alpha \Delta T = -\Lambda \Delta T + v_\alpha(t) + F_2(t), \quad (29)$$

which can be seen to be closely related to Lovejoy and co-workers’ FEBE by comparison with, for example, Eqs. (13) and (14) of Ref. 16. The key difference is that we have a Caputo derivative rather than their Riemann–Liouville one. We postpone to future work the exploration of the implications of this choice in the model’s design. Importantly, the $\alpha = 1, H = 1/2$ limit of both derivatives is of course the same and gives the Hasselman model as required. The determination of the order of the Caputo derivative empirically is an area of current research, both generally and in the specific climate context. It is directly related to the form of the response function of the system under study, if the Caputo model is itself applicable, and indirectly related to the power spectral density of the main observable, here ΔT . The specific relationship, if any, between the response function and the noise driving depends on the presence or absence of a fluctuation–dissipation relation.

We note that although the integral in the Caputo derivative is frequently started at $t_m = 0$ (including, for example, in its current implementation in `Mathematica`), its definition allows it to start anywhere on the real line.¹⁰¹ It, thus, naturally encompasses the $t_m = -\infty$, “Weyl” case to which Lovejoy³³ has drawn attention in the climate context, without additional effort.

IX. SOLVING THE MORI-KUBO GLE AND THE FLE ANALYTICALLY

In this section and the following one, we have used the words “analytic,” “numerical,” and “solution” as is commonly done in the physics literature. The mathematics literature has different conventions and the reader is advised to beware.

A. General solution of the overdamped Mori-Kubo GLE

There are by now numerous sources in the literature for analytic solutions of the Mori-Kubo GLE, starting from near-contemporary work such as that of Fox.⁸³ They typically use Laplace transform methods, e.g., those of Ref. 86.

In our present application to climate, it is important to note that solutions of the GLE have not been limited to assuming a starting time t_m of zero. A notable recent example is where t_m has been more general is Ref. 55 whose authors studied an overdamped Mori-Kubo GLE,

$$\int_{t_m}^t dt' K(t-t')\dot{x}(t') + kx(t) = F_2(t) + \xi(t),$$

which described a particle of mass M moving in a harmonic confining potential $V(x) = (1/2)kx^2$ (in their case, an optical trap) and experiencing an additional deterministic forcing (from dragging of the trap) $F_2(t)$ and a stochastic forcing which they assumed to be white noise $\xi(t)$.

The authors obtained the general solution of their Mori-Kubo GLE as

$$x(t) = x_{t_m}(1 - k\chi(t - t_m)) + \int_{t_m}^t dt' \chi_x(t - t')[F_2(t') + \xi(t')], \quad (30)$$

where x_{t_m} denoted the initial condition for position $x(t = t_m)$. The position susceptibility $\chi_x(t)$ here was defined through its Laplace transform

$$\hat{\chi}_x(s) = [Ms^2 + s\hat{K}(s) + k]^{-1}$$

and the integral of it appearing in their solution was defined as

$$\chi(t) = \int_0^t dt' \chi_x(t').$$

The cases treated in Ref. 55 were relatively tractable, so they should, in principle, also allow some special cases of our proposed stochastic EBM to be solved analytically—a problem which we expect to return to in future work.

B. Solving the FLE by Green's functions

1. The overdamped FLE

For now, though, we will specialize to the fractional special case of the overdamped Mori-Kubo GLE, i.e., the overdamped FLE. We recall that this will specialize further for $\alpha = 1$ to the familiar exponential solutions of the Hasselmann model.

The overdamped FLE, both with and without deterministic forcing, has been studied in the mathematics and statistical physics literature, e.g., in 2017 by Li *et al.* in Ref. 102. With a Riemann-Liouville derivative rather than the Caputo type, it has been studied by Lovejoy and co-workers as a stochastic climate model since 2019, and dubbed “the fractional energy balance equation.”

The overdamped forced FLE,

$$[{}_{t_m}^C D_t^\alpha + k]x(t) = F_2(t) + v_\alpha(t), \quad (31)$$

can be seen to have an entirely deterministic fractional operator on its left hand side, while the right hand side incorporates both the deterministic forcing term $F_2(t)$ and a stochastic term v_α .

2. The delta function-forced Caputo FDE and Green's function of the FLE

One way to solve the overdamped forced FLE is to first obtain the solution of the deterministic relaxation and oscillation fractional differential equation, i.e., the Caputo FDE, when forced by a δ -function,

$$[{}_{t_m}^C D_t^\alpha + k]x(t) = \delta(x), \quad (32)$$

which can be done using Laplace transforms, e.g., Ref. 103.

With an appropriate initial condition, this gives the impulse response of the fractional operator, which can then be used to construct Green's function to solve either the deterministic or stochastic versions of the inhomogenous equation.

The Laplace transform solution of (32) for $t_m = 0$ and $x(0) = 0$ is

$$x(t) = t^{\alpha-1} E_{\alpha,\alpha}(-kt^\alpha),$$

as can be checked in version 13.1 or later of Mathematica using its `CaputoD` and `DiracDelta` functions.

Figure 1 plots $x(t)$ for values of α from 1 to 1/2, and $k = 1$. On the semilog axes chosen, we can see that the $\alpha = 1$ case is exponential while smaller values of α correspond to more slowly decaying tails in the Mittag-Leffler function $E_{\alpha,\alpha}(t)$ and to an increasingly singular “spike” at zero from the power law prefactor. It may be compared with Figs. 2 and 3 of Ref. 16, which used a Riemann-Liouville derivative.

3. Step function forcing of the Caputo FDE and FLE

If we then include deterministic forcing $F_2(t)$ and take $t_m = 0$, $x(t_m) = x_0$, the solution becomes

$$x(t) = x_0 E_\alpha(-kt^\alpha) + \int_0^t F_2(\tau)(t - \tau)^{\alpha-1} E_{\alpha,\alpha}[-k(t - \tau)^\alpha] d\tau.$$

In the well-studied case most relevant to climate science, when forcing becomes a Heaviside step function applied from $t = 0$ onward,

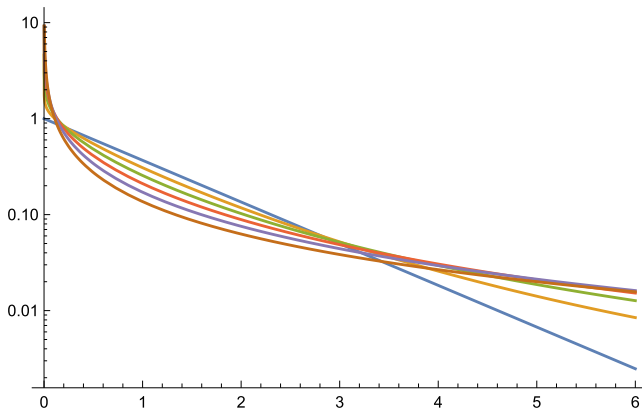


FIG. 1. Impulse response of the fractional relaxation-oscillation equation for $\alpha = 1$ to 0.5.

i.e., $F(t) = \theta(t)$, the integral can be evaluated and the solution becomes

$$x(t) = x_0 E_\alpha(-kt^\alpha) + t^\alpha E_{\alpha,1+\alpha}[-kt^\alpha],$$

which again can be confirmed by Mathematica.

This function is illustrated in Fig. 2, which plots $x(t)$, again for values of α from 1 to 0.5, but now for $k = 0.1$. The topmost, blue, line is the $\alpha = 1$, exponential, case, and $x_0 = 0$.

The effect of memory for the cases where $\alpha < 1$ is clear in lengthening the time it takes for the solution to take its asymptotic value. In a simplest forcing experiment, such as doubling CO_2 , for a stochastic EBM, this would correspond to the approach to a new equilibrium value of the GMT anomaly.

4. The overdamped FLE with arbitrary deterministic and stochastic forcing

We finally come to the full overdamped forced FLE (31), again with $t_m = 0$, when the stochastic term has also been included. This is

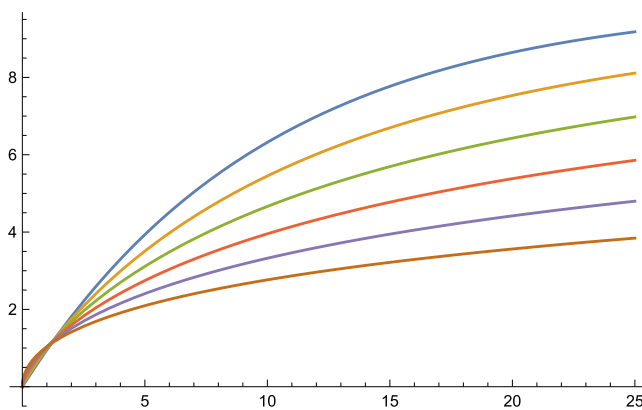


FIG. 2. Step response of the fractional relaxation-oscillation equation for $k = 0.1$, $\alpha = 1$ to 0.5.

too general to have a useful analytic solution, but we can make some comments to help elucidate its meaning.

When written in integral form, the SDE has three parts. The first and second are those seen in the previous sections, while the third must be interpreted as a stochastic integral. We will not deal with this term, corresponding to $v_\alpha(t)$ in (31), in full generality but will instead consider the case of fractional Gaussian driving,

$$x(t) = x_0 E_\alpha(-kt^\alpha) + \int_0^t (t-\tau)^{\alpha-1} F_2(\tau) E_{\alpha,\alpha}[-k(t-\tau)^\alpha] d\tau + C_H \int_0^t (t-\tau)^{\alpha-1} E_{\alpha,\alpha}[-k(t-\tau)^\alpha] dB_H(\tau). \quad (33)$$

The above stochastic integral is best understood in the $H = 1/2$ case of white noise when it can be described by Itô calculus, although there is by now also a substantial mathematical literature concerning integrals with respect to fractional Brownian motion^{94,104} (fBm).

Furthermore, when we then specialize to $F_2 = 0$ and $\alpha = 1$, the solution then corresponds to the very well studied O-U process,

$$x(t) = x_0 \exp(-kt) + \bar{\sigma} \int_0^t e^{-k(t-\tau)} dW(\tau).$$

The stochastic integral in this case can be defined as the limit of a sum and dW as a suitably rescaled increment of the Brownian motion,

$$\int_0^t e^{-k(t-\tau)} dW(\tau) = \lim_{\Delta \rightarrow 0} \sum_{j=0}^{n-1} e^{-k(j\Delta-\tau)} (W((j+1)\Delta) - W(j\Delta)),$$

as described on pages 90–91 of Ref. 21.

X. NUMERICALLY EVALUATING TRAJECTORIES OF THE MORI-KUBO GLE AND FLE

As is so often the case, analytical intractability of the most physically interesting and relevant SDEs means that we must evaluate them numerically. We discuss this below, first for the Mori-Kubo GLE and then for its fractional special case.

A. Numerics for the Mori-Kubo GLE

There are by now several modern sources of numerical algorithm for the stochastic Mori-Kubo GLE, e.g., Refs. 28, 29, 105, and 106.

The algorithm of Lim²⁸ requires that the system of SDEs to be solved must be “Bohl,” i.e., the matrix elements of the memory kernel $K(t)$ are finite linear combinations of the functions of the form $t^k e^{\beta t} \cos(\omega t)$ and $t^k e^{\beta t} \sin(\omega t)$, where k is an integer and β and ω are real numbers. As Lim²⁸ notes, when these functions decay as a power law, for example, the resulting fractional Mori-Kubo GLE cannot be studied as a finite-dimensional SDE system. Although one can work formally in an infinite-dimensional setting, numerics will typically resort (e.g., Ref. 107) to approximating the power law, frequently by using a weighted sum of exponentials.

B. Numerics for the overdamped FLE

The overdamped FLE can be solved in a number of ways. Lovejoy *et al.*'s work^{15,16} on FEBE has so far used Fourier methods similar

to those widely used to simulate fBm and its increment, fGn, and has focused on an initial condition specified at $t_m = -\infty$. In our paper, we will instead consider finite t_m , which allows the infinitely distant case as a limit, and wish to use a non-Fourier based algorithm for its flexibility in potential future developments.

Our approach is illustrative, rather than rigorous. We use the explicit solution of the overdamped FLE and evaluate it at each required value of t . We do this by discretizing t using the definitions of the stochastic and deterministic integrals as the limits of sums. However, we must stress that even in the O-U case, it is known (e.g., Ref. 24) that this algorithm will **not** ensure that a solution will have the required autocorrelation properties. This is why our results must be considered illustrative rather than definitive for now.

In our future work, in order to achieve the required ACF, a time evolution equation rather than an integral will be needed. A prototype of such algorithms was derived by Gillespie^{24,108} for the O-U case. Candidates include the spectral method of Ref. 98, and the recently developed fast Euler-Maruyama algorithm given in Refs. 109 and 110.

Here, we also restrict ourselves to the white noise-forced case, but with the additional deterministic forcing term, and intend to explore colored noise forcing in future work.

We first approximate the stochastic integral using

$$\begin{aligned} & \int_0^t (t-\tau)^{\alpha-1} E_{\alpha,\alpha}(-k(t-\tau)^\alpha) dW(\tau) \\ & \approx \sum_{j=0}^{n-1} (t-j\Delta)^{\alpha-1} E_{\alpha,\alpha}(-k(t-j\Delta)^\alpha) (W[(j+1)\Delta] - W(j\Delta)) \\ & = \sum_{j=0}^{n-1} \Delta^{\alpha-1} (n-j)^{\alpha-1} E_{\alpha,\alpha}(-k\Delta^\alpha (n-j)^\alpha) \Delta^{1/2} N_{j\Delta}^{(j+1)\Delta}, \end{aligned}$$

where following Ref. 24, each finite difference of W has been replaced by the scaled Gaussian random number $\Delta^{1/2}N$.

The deterministic integral is approximated by

$$\begin{aligned} & \int_0^t F_2(\tau) (t-\tau)^{\alpha-1} E_{\alpha,\alpha}(-k(t-\tau)^\alpha) d\tau \\ & \approx \sum_{j=0}^{n-1} \Delta^{\alpha-1} (n-j)^{\alpha-1} E_{\alpha,\alpha}(-k\Delta^\alpha (n-j)^\alpha) \Delta, \end{aligned} \quad (34)$$

where $F_2(\tau) = \theta(\tau)$ for the step function problem.

In Fig. 3, we illustrate the behavior of the FLE by plotting 100 sample paths obtained this way for forcing by a Heaviside step function $\Theta(t)$ and a stochastic term of amplitude $\bar{\sigma} = 0.16$. We used MATLAB's function `boxplot` where for each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the "+" marker symbol. Parameters were $t_m = 0, x(t_m) = 0, k = 0.1$, and $\alpha = 0.7$. The mean of the paths is overplotted and is seen to compare well with its known analytic result.

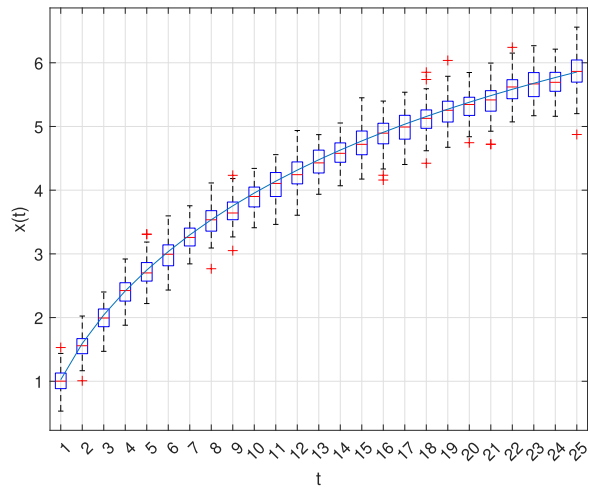


FIG. 3. Box plot of an ensemble of 100 sample paths of the overdamped fractional Langevin equation, forced by a Heaviside step function $\Theta(t)$ and a stochastic term of amplitude $\bar{\sigma} = 0.16$. The parameters were $t_m = 0, x(t_m) = 0, k = 0.1$, and $\alpha = 0.7$. Overplotted is the analytic solution, which can be seen to coincide well with the means of the ensemble.

C. Numerics for our proposed Mori-Kubo GLE-type EBM driven by RCP 8.5

We now briefly illustrate how our stochastic GLE-based EBM, in its fractional Caputo special case (29), can be used to study the global mean surface temperature. We do this in the white noise driven ($\nu_\alpha = \bar{\sigma}\xi$) case by repeating the previous subsection's approach, but this time taking the deterministic part of the driving from the RCP 8.5 Representative Concentration Pathway¹¹¹, a (worst case) model of anthropogenic forcing. We first write the fractional EBM, with its parameters and scalings made more explicit, as

$${}_0^C D_t^\alpha \Delta T + \tau_H^{-\alpha} \Delta T = \tau_H^{-\alpha} \frac{F}{\lambda} + \bar{\sigma} \xi,$$

where s is the climate sensitivity, $\lambda = 1/s$, $C^{(\alpha)}$ is an α -dependent scale factor for the effective heat capacity (see Refs. 16 33), $\tau_H = C^{(\alpha)}/\lambda$ is the Hasselmann time, and the noise term has variance $\bar{\sigma}^2$. We used MATLAB's quadrature function `trapz` to speed up the deterministic part of the integration, a step size of 0.1 years, and, as in section X.B used Podlubny and Kacena's function `mlf` from the MATLAB contributed toolbox to evaluate the Mittag-Leffler functions. Figure 4 is a box plot of 64 paths for ΔT , again with the, at present arbitrary, choice $\alpha = 0.7$, and with representative values $\lambda = 1.23$, $\tau_H = 25.78$ years and $\bar{\sigma} = 0.1$. The appropriate choice of parameter values for such models remains an area of active research¹⁶. Future more detailed investigations will, rather than quadrature, use a more accurate SDE solver suitable for all GLE-derived EBMs, not just fractional ones, as discussed above.

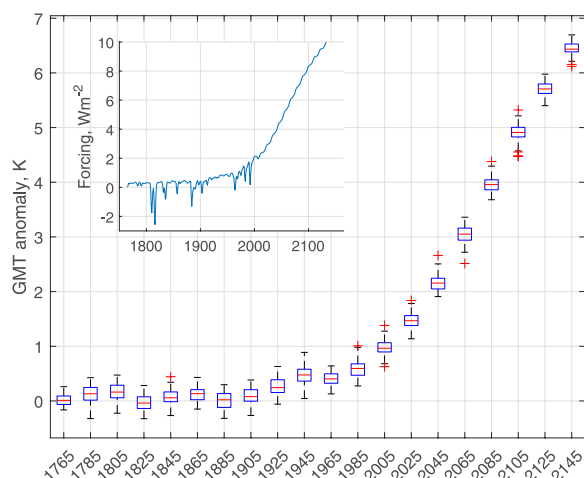


FIG. 4. Box plot of 64 paths for global mean temperature anomaly ΔT from a fractional Caputo EBM driven by RCP 8.5 (shown in the inset) with added white noise. $\alpha = 0.7$, $\lambda = 1.23$, $\tau_H = 25.78$ years and $\bar{\sigma} = 0.1$.

XI. ARE THERE ALTERNATIVE STOCHASTIC GLE-BASED EBMs?

Although this paper is focused on a stochastic EBM inspired by a mapping to the overdamped Mori–Kubo Langevin equation, other options exist—some of which have been explored so far in the literature.

A. Mori–Kubo GLE-based EBMs and other intrinsically stochastic approaches

In addition to Lovejoy *et al.*'s FEBE discussed in some detail above, and the overdamped GLE-based EBM we have proposed here, there have been several other proposals that extend the Hasselmann formalism to include multiscale memory. We enumerate some below. As noted in the introduction, Ghil *et al.*¹⁹ have argued that a key requirement of any such model is the ability to add quasiperiodic determinism as well as stochasticity.

1. Modified fractional Brownian motion

In several pioneering papers,^{77,78} Rypdal and colleagues discussed various extensions of the fractional Brownian motion as generalizations of the Hasselmann EBM. The main limitation of their idea was the absence of the damping term, as noted by Lovejoy.^{16,33}

2. Direct mapping to underdamped generalized Langevin equation rather than overdamped GLE

In our earlier conference proceedings chapter³¹ and subsequent preprint,³² we considered a different GLE-based EBM. Here, the mapping was made from the underdamped Langevin equation rather than the overdamped Langevin equation.

This results in a mapping from velocity onto GMT anomaly, rather than from position. We erroneously thought that in its fractional special case, our suggested stochastic EBM was isomorphic

to FEBE after suitable conversions between their parameters, but in fact, the comparisons made in Sec. VIII of the present paper show that FEBE is much closer to the overdamped Langevin equation and, thus, to our newer proposal.

3. Two box models

The two box EBM is an interesting alternative which may be useful in some cases. It replaces the single decay scale of Hasselmann's model with two. This allows, for example, the atmosphere and ocean timescales to be explicitly represented.

As shown in Ref. 45, the two box model results in a second order SDE with constant coefficients, isomorphic to the underdamped Langevin equation for position.

Interestingly, this second order equation is also an alternative representation for the Mori–Kubo GLE when its memory kernel is of exponential form, as shown in Ref. 112, a special case of a more general result given in Ref. 106.

4. Multibox models

The above methods can be extended to the case of many relaxation time scales, an approach studied extensively for SDEs in the applied mathematics literature and also in climate science.^{22,46} Eigenvalue analysis can be used to predict some of the spectral properties of such multibox models, as done, for example, by Fredriksen and Rypdal.⁴⁷

5. The fractional Ornstein–Uhlenbeck equation

The Langevin equation can be modified to be driven by fractional Gaussian noise rather than white noise while still keeping the friction term constant. This has been studied in the mathematical finance literature extensively⁹⁴ and has been suggested in the climate context.¹¹³ These authors also develop a time series model^{113,114} based on the discrete time autoregressive fractionally integrated moving average ARFIMA process, which generalizes AR(1).

6. Yuan *et al.*'s generalized stochastic time series model

A fractional time series model, not simply isomorphic to ARFIMA or other existing models, has been proposed¹³ to systematically separate the “forcing-induced direct” and the “memory-induced indirect” trends in data.

B. Other stochastic Mori–Zwanzig GLE-based EBMs

As noted in our introduction, the Mori–Kubo GLE is in principle a subset of the more complete Mori–Zwanzig formalism, which has at its heart a deterministic identity expressing the division of a system into observed and unobserved degrees of freedom.²⁷ Space does not permit a full discussion of this here, but we refer the reader to a short introduction,²³ and reviews of the method²¹ and its applications to climate.⁸

XII. FLUCTUATION–DISSIPATION RELATIONS

Fluctuation–dissipation relations (FDRs) have been studied extensively in climate science (see, e.g., the reviews of Refs. 3 and 4

and recent examples like Ref. 115). We do not discuss them at length here, but simply note some aspects that relate to the stochastic EBMs we study.

A. Emergent parameters in the OU process

The meaning of the phrase “fluctuation–dissipation relation” can sometimes be as simple as its usage by the Nobel Committee,⁹ who noted that the variance $\langle y^2 \rangle$ of the OU process, which solves the overdamped Langevin equation (9) is proportional to the ratio of the noise amplitude $\bar{\sigma}^2$ to the mean reversion term $\bar{\gamma}$. The Nobel document appears to us to be stressing that **any** system to which the O–U process has been applied must in consequence obey this limited kind of FDR. This is because there will then always be a ratio of $\sigma^2/2\bar{\gamma}$ and it will empirically determine the value of $\langle y^2 \rangle$. This is not a trivial observation in the climate context because either the value of $\langle y^2 \rangle$ or the noise amplitude $\bar{\sigma}$ are in this sense “emergent” parameters whose values are unknown *a priori*, cf. the FDR-like relation found in Ref. 54.

B. The FDR in the Brownian motion

In the physical context (the motion of a tracer particle in a fluid) from which the underdamped and overdamped Langevin equations originated,^{4,24} however, this ratio was not a free or empirically determined parameter. It was instead constrained to be proportional to the temperature T of the surrounding heat bath. A proportionality to T is seen both in the Brownian motion special case of Kubo’s “FDR2,”^{89,90,116–118} i.e., the ratio of the noise amplitude to the mean reversion (friction) term in the underdamped Langevin equation, and in a corresponding expression arising from the overdamped Langevin equation. Importantly, these fluctuations are entirely “internal,” and the fast and slow terms in the Langevin equation arose from the same physics. This is not straightforwardly the case when an EBM is written down or derived for the climate problem, and extensive debates have occurred about the applicability or otherwise of the FDR in a climate context.^{3,4}

C. The FDR as seen in the Mori-Kubo GLE

As we noted above, the Mori–Kubo GLE (here in its overdamped form)

$$\int_{t_m}^t dt' \gamma(t-t') \frac{dy(t')}{dt'} = -ky + v(t) \quad (35)$$

was designed to embody an FDR which relates the covariant structure of the noise $v(t)$ to the auto-correlation structure of the kernel $\gamma(t-t')$, so that they are not free to vary independently. In the original context for the GLE, Brownian motion, this FDR is literally a fluctuation–dissipation relation because the $v(t)$ term is a fluctuating noise and the integral operator over $\gamma(t)$ is a friction term acting to reduce the magnitude of the velocity dy/dt .

If we naively take this functional form of the GLE over to the climate context, however, the integral operator is doing a physically different job. It is now an integral over a heat capacity $C_p(t-t')$, and so, an “FDR” of the classic GLE type would be relating the covariance structure of heat capacity to that of the noise $v(t)$.

This would still be doing essentially the same thing in the sense of balancing the noise and the kernel, scale by scale, in order to allow the system to relax to a steady state, but further investigation would be necessary to establish whether it made physical sense.

XIII. CONCLUSIONS

Hasselmann’s realization that the Brownian motion could be used as a mathematical superstructure to organize fast weather and slow climate fluctuations was a very powerful one, and his recent Nobel prize provided a timely context for our review. As we have shown, recent progress in going beyond his paradigm has generated, what at first sight seem to be divergent positions, exemplified by the recent reviews of Franzke *et al.* in Ref. 10 and Lovejoy in Ref. 14.

In this paper, we have sought to facilitate an improved dialogue between these views, and have found them to have more points of contact than was initially apparent. We first drew attention to the existing climate work using the Mori–Zwanzig approach and proposed using the stochastic GLE of Mori and Kubo as a framework to study memory in one-dimensional stochastic EBMs. We then related the stochastic GLE-based EBM to Lovejoy’s fractional energy balance equation and concluded with some speculations on fluctuation–dissipation relations.

ACKNOWLEDGMENTS

It is a pleasure to dedicate this paper to Professor Juergen Kurths on the occasion of his 70th birthday and to thank him for his interest in contribution to this topic. We are also particularly grateful to Ian Ford for many valuable discussions and to Nick Moloney for a thorough reading of the manuscript. We acknowledge valuable interactions with Lenin del Rio Amador, Peter Cox, Holger Kantz, Paul Kushner, Valerio Lembo, Shaun Lovejoy, Valerio Lucarini, Lauren Padilla, Cecile Penland, Sid Redner, Kira Rehfeld, Kristoffer Rypdal, Martin Rypdal, Jan Sieber, Geoff Vallis, and Dmitry Vyushin. N.W.W. and S.C.C. acknowledge AFOSR (Grant No. FA9550-17-1-0054) and STFC (Grant No. ST/T000252/1) and, with D.A.S. and R.C., the Warwick International Partnership Fund. N.W.W. and S.C.C. acknowledge the hospitality of the Max Planck Institute for the Physics of Complex Systems, the University of Potsdam, the Potsdam Institute (PIK), the Santa Fe Institute, the International Space Science Institute, and the CLIMA project at Caltech, during the development of this work. N.W.W. also acknowledges past support from the London Mathematical Laboratory, KLIMAFORSK (Project No. 229754) and ONR (Grant No. NICOP-N6209-15-1-N143). S.C.C. also acknowledges an ISSI Johannes Geiss Fellowship. A.V.C. acknowledges support of the Polish National Agency for Academic Exchange (NAWA). D.A.S. acknowledges support from the Grantham Research Institute on Climate Change and the Environment at the London School of Economics, the ESRC Centre for Climate Change Economics and Policy (CCCEP; ref. ES/R009708/1), and the Natural Environment Research Council through Optimising the Design of Ensembles to Support Science and Society (ODESSS; ref. NE/V011790/1). N.W.W. and R.K. thank the Isaac Newton Institute Cambridge for participation at the program on Mathematics for the Fluid Earth in 2013.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Nicholas W. Watkins: Conceptualization (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Raphael Calel:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Sandra C. Chapman:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Aleksei Chechkin:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Rainer Klages:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **David A. Stainforth:** Conceptualization (equal); Formal analysis (equal); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study

APPENDIX: SUMMARY OF FRACTIONAL CALCULUS

Because of their relationship to long-range dependence in time series and heavy-tailed amplitude distributions, fractional differential equations (FDEs) and fractional stochastic differential equations (FSDEs), such as those in the present paper, have also been extensively studied in complexity science. In order to make the paper more accessible to readers who have had little need to deal with fractional integration and differentiation, we give here a very brief summary of the results which we use. Fractional derivatives and the associated FDEs are now implemented directly in some numerical packages, like Mathematica whose demonstrations we encourage the reader to explore.

This Appendix follows the notational conventions of the physics literature. Readers seeking a more detailed but similarly physics-inspired treatment are referred to Ref. 119 and also the very useful general overview article of Ref. 120.

Readers who prefer a more mathematical development may consult Refs. 92, 93, 103, 121, and 122. In addition, two books which study SDEs driven by the fractional Brownian motion are Refs. 94 and 104.

1. Fractional integral

It is possible to begin by defining a fractional derivative but we follow most authors by starting with the inductive definition of integration, which the fractional integral generalizes. We are used to the idea that the taking of a derivative of a function $f(t)$ with respect to

time t is the inverse of integrating it with respect to t' up to time t ,

$$\int_a^t f(t') dt'. \quad (\text{A1})$$

The n th derivative is, thus, an inverse to the n -fold repeated integration,

$$\int_a^t dt'_1 \int_a^{t'_1} dt'_2 \dots \int_a^{t'_n} dt'_n f(t'_n), \quad (\text{A2})$$

and requires us to use the integral identity,

$$\int_a^t \int_a^{t'_1} \dots \int_a^{t'_{n-1}} f(t'_n) dt'_n \dots dt'_1 = \frac{1}{(n-1)!} \int_a^t f(t')(t-t')^{n-1} dt'. \quad (\text{A3})$$

The form of the right hand side of this expression shows why we can then allow n to be replaced by a non-integer value α , thus defining an integral of fractional order. In defining a fractional integral, we use the relationship between the factorial and the gamma function, $\Gamma(n) = (n-1)!$ the discovery of which was a consequence of the earliest work on fractional calculus.¹²³

The left-sided Riemann–Liouville fractional integral operator of order α on a finite interval (a, b) of the real line, acting on a function $f(t)$ is then

$$({}_a I_t^\alpha f)(t) = \frac{1}{\Gamma(\alpha)} \int_a^t f(t')(t-t')^{\alpha-1} dt'. \quad (\text{A4})$$

Note that this definition requires that $\alpha > 0$. Typically, $\alpha < 1$ is also assumed, but this is not obligatory.¹⁰¹ Other notations exist for this integral, e.g., $(I_{a+}^\alpha f)(x)$ used widely in the mathematics literature.

Taking $\alpha = 1$ recovers the usual integral of $f(t)$ from a to t ,

$$({}_a I_t^1 f)(t) = \int_a^t f(t') dt', \quad (\text{A5})$$

whereas $\alpha = 1/2$ gives

$$({}_a I_t^{1/2} f)(t) = \frac{1}{\Gamma(1/2)} \int_a^t \frac{f(t')}{\sqrt{t-t'}} dt', \quad (\text{A6})$$

which was one of the first fractional integrals to receive study in physics, chemistry, and engineering, notably in the heat transfer problem.⁹¹

Instead of being over a finite interval, we can define the left-sided Riemann–Liouville fractional integral of $f(t)$ on the complete real line,

$$({}_{-\infty} I_t^\alpha f)(t) = \frac{1}{\Gamma(\alpha)} \int_{-\infty}^t f(t')(t-t')^{\alpha-1} dt'. \quad (\text{A7})$$

2. Riemann–Liouville fractional derivative

We can now define fractional differentiation as the inverse of fractional integration. In practice, this is done by taking a fractional integral of order 1 higher than the desired order of the derivative, and then, using the usual integer derivative to lower its order by one. Because fractional integration and differentiation do not commute, two possible fractional derivatives exist, depending on whether the fractional integral of $f(t)$ is done first or the integer derivative. These are the Riemann–Liouville and Caputo cases, respectively.

Taking the Riemann–Liouville case first, and considering the finite time interval (a, b) , we define a derivative which is based on the inverse of the left-sided Riemann–Liouville integral. We use the fact that the Riemann–Liouville fractional derivative of the function $f(t)$ of order α , ${}_a D_t^\alpha f(t)$, is its Riemann–Liouville integral of order $-\alpha$. We calculate this by first writing the integral of order $1 - \alpha$, where $0 < \alpha < 1$,

$$({}_a I_t^{1-\alpha} f)(t) = \frac{1}{\Gamma(1-\alpha)} \int_a^t f(t') (t-t')^{-\alpha} dt', \quad (\text{A8})$$

and then differentiating

$$({}_a I_t^{1-\alpha} f)(t) = ({}_a D_t^\alpha f)(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_a^t f(t') (t-t')^{-\alpha} dt'. \quad (\text{A9})$$

The lower limit of this integral is usually taken to be 0, though this is not, in fact, essential.¹⁰¹ As in the case of the fractional integral, we can define the integral in the derivative over the whole real line, so a can be $-\infty$. This case is sometimes referred to as the Weyl or Riemann–Weyl derivative.

3. Caputo fractional derivative

If instead we take the integer derivative first, before doing the fractional integral, we can define the Caputo fractional derivative,

$$({}_a^C D_t^\alpha f)(t) = \frac{1}{\Gamma(1-\alpha)} \int_a^t dt' (t-t')^{-\alpha} \frac{df}{dt'}(t'). \quad (\text{A10})$$

The Caputo derivative can be defined via the Riemann–Liouville derivative,

$$({}_a^C D_t^\alpha f)(t) = ({}_a D_t^\alpha f)(t) - \frac{t^{-\alpha}}{\Gamma(-\alpha)} f(t), \quad (\text{A11})$$

so the Caputo and Riemann–Liouville fractional derivatives will coincide for negative orders α . One advantage of the Caputo definition over the Riemann–Liouville case is that it uses the values of $f(t)$ and its integer rather than fractional derivatives at 0 (or, in general, at any lower-limit point a). This is well suited to solving fractional-order initial-value problems using Laplace transforms, a popular approach for both FDEs and FSDEs.

REFERENCES

¹R. Saranavan, *The Climate Demon: Past, Present, and Future of Climate Prediction* (Cambridge University Press, Cambridge, 2022).
²D. A. Stainforth, *Predicting Our Climate Future: What We Know, What We Don't Know, and What We Can't Know* (Oxford University Press, Oxford, 2023).
³M. Ghil and V. Lucarini, "The physics of climate variability and climate change," *Rev. Mod. Phys.* **92**, 035002 (2020).
⁴M. Ghil, "A century of nonlinearity in the geosciences," *Earth Space Sci.* **6**, 1007–1042 (2019).
⁵N. Boccara, *Modelling Complex Systems* (Springer, London, 2010).
⁶*Nonlinear and Stochastic Climate Dynamics*, edited by C. Franzke and T. O'Kane (Cambridge University Press, Cambridge, 2017).
⁷*Stochastic Physics and Climate Modelling*, edited by T. Palmer and P. Williams (Cambridge University Press, Cambridge, 2010).
⁸V. Lucarini and M. Chekroun, "Theoretical tools for understanding the climate crisis from Hasselmann's program and beyond," *Nat. Rev. Phys.* **5**, 744 (2023).
⁹"Scientific background on the Nobel prize in physics 2021," The Nobel Committee (The Royal Swedish Academy of Sciences, Stockholm, 2021).

¹⁰C. L. E. Franzke, R. Blender, T. J. O'Kane, and V. Lembo, "Stochastic methods and complexity science in climate research and modelling," *Front. Phys.* **10**, 931596 (2022).
¹¹K. Hasselmann, "Stochastic climate models. Part I. Theory," *Tellus* **28**, 473–485 (1976).
¹²C. L. E. Franzke, S. Barbosa, R. Blender, H. Fredriksen, T. Laepple, F. Lambert, T. Nilsen, K. Rypdal, M. Rypdal, M. G. Scotto, S. Vannitsem, N. W. Watkins, L. Yang, and N. Yuan, "The structure of climate variability across scales," *Rev. Geophys.* **58**, e2019RG000657, (2020).
¹³N. Yuan, C. L. E. Franzke, F. Xiong, Z. Fu, and W. Dong, "The impact of long-term memory on the climate response to greenhouse gas emissions," *npj Clim. Atmos. Sci.* **5**, 70 (2022).
¹⁴S. Lovejoy, "The 2021 "complex systems" Nobel prize: The climate, with and without geocomplexity," *AGU Adv.* **3**, e2021AV000640 (2022).
¹⁵S. Lovejoy, *Weather, Macroweather, and the Climate: Our Random Yet Predictable Atmosphere* (Oxford University Press, Oxford, 2019).
¹⁶S. Lovejoy, R. Procyk, R. Hebert, and L. Del Rio Amador, "The fractional energy balance equation," *Q. J. R. Meteorol. Soc.* 1–25 (2021).
¹⁷S. Lovejoy, "Review article: Scaling, dynamical regimes, and stratification. How long does weather last? How big is a cloud?," *Nonlinear Process. Geophys.* **30**, 311–574 (2023).
¹⁸M. Mann, "On long range dependence in global surface temperature series," *Clim. Change* **7**, 267–276 (2011).
¹⁹M. Ghil, P. Yiou, S. Hallegatte, B. D. Malamud, P. Naveau, A. Soloviev, P. Friederichs, V. Keilis-Borok, D. Kondrashov, V. Kossobokov, O. Mestre, C. Nicolis, H. Rust, P. Shebalin, M. Vrac, A. Witt, and I. Zaliapin, "Extreme events: Dynamics, statics and prediction," *Nonlinear Processes in Geophysics* **18**, 295–350 (2011).
²⁰D. Kondrashov, M. D. Chekroun, and M. Ghil, "Data-driven non Markovian closure models," *Physica D* **297**, 33–55 (2015).
²¹A. J. Chorin and O. H. Hald, *Stochastic Tools in Mathematical Science*, 3rd ed. (Springer Verlag, Berlin, 2013).
²²G. A. Pavliotis, *Stochastic Processes and Applications: Diffusion Processes, The Fokker-Planck and Langevin Equations* (Springer, New York, 2014).
²³S.-H. Chung and M. Roper, "Generalized Langevin equation: An introductory review for biophysicists," *Biophys. Rev. Lett.* **14**, 171–196 (2019).
²⁴D. S. Lemons, *An Introduction to Stochastic Processes in Physics* (Johns Hopkins University Press, Baltimore, MD, 2002).
²⁵H. Mori, "Transport, collective motion and Brownian motion," *Progr. Theoret. Phys.* **33**, 423–455 (1965).
²⁶P. Pechukas, "Generalized Langevin equation of Mori and Kubo," *Phys. Rev.* **164**, 174–175 (1967).
²⁷R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, Oxford, 2001).
²⁸S. H. Lim, "Anomalous thermodynamics in homogenized generalized Langevin systems," *J. Phys. A: Math. Theor.* **54**, 155001 (2021).
²⁹I. Snook, *The Langevin and Generalised Langevin Approach to the Dynamics of Atomic, Polymeric and Colloidal Systems* (Elsevier, 2006).
³⁰V. Jaksic and C.-A. Pillet, "Ergodic properties of the non-Markovian Langevin equation," *Lett. Math. Phys.* **41**, 49–57 (1997).
³¹N. Watkins, S. Chapman, A. Chechkin, I. Ford, R. Klages, and D. Stainforth, "On generalised Langevin dynamics and the modelling of global mean temperature," in *Unifying Themes in Complex Systems X*, Springer Proceedings in Complexity, edited by D. Braha, M. A. M. de Aguiar, C. Gershenson, A. J. Morales, L. Kaufman, E. N. Naumova, A. A. Minai, and Y. Bar-Yam (Springer, 2021).
³²N. Watkins, S. Chapman, A. Chechkin, I. Ford, R. Klages, and D. Stainforth, "On generalized Langevin dynamics and the modelling of global mean temperature," [arxiv:2007.06464v2](https://arxiv.org/abs/2007.06464v2)[cond-mat.stat-mech] (2020).
³³S. Lovejoy, "Fractional relaxation noises, motions and the fractional energy balance equation," *Nonlinear Process. Geophys.* **29**, 93–121 (2022).
³⁴J. M. Mitchell, "An overview of climatic variability and its causal mechanisms," *Quat. Res.* **6**, 481–493 (1976).
³⁵M. Ghil, "Climate, zoonosis and interdisciplinarity. commentary," *Proc. Natl. Acad. Sci. U.S.A.* **120**(39), e2311575120 (2023).
³⁶K. McGuffie and A. Henderson-Sellers, *The Climate Modelling Primer*, 4th ed. (Wiley Blackwell, 2014).

- ³⁷G. R. North, R. F. Cahalan, and J. A. Coakley, Jr., “Energy balance climate models,” *Rev. Geophys. Space Phys.* **19**, 91–121 (1981).
- ³⁸G. R. North and K.-Y. Kim, *Energy Balance Climate Models* (Wiley-VCH, Weinheim, 2017).
- ³⁹K. Fraedrich, “Catastrophes and resilience of a zero-dimensional climate system with ice-albedo and greenhouse feedback,” *Quart. J. R. Met. Soc.* **105**, 147–167 (1979).
- ⁴⁰A. Sutera, “On stochastic perturbation and long-term climate behavior,” *Quart. J. R. Met. Soc.* **107**, 137–151 (1981).
- ⁴¹P. D. Ditlevsen, “Observation of α -stable noise induced millennial climate changes from an ice-core record,” *Geophys. Res. Lett.* **26**, 1441–1444, <https://doi.org/10.1029/1999GL900252> (1999).
- ⁴²V. Lucarini, L. Surdukova, and G. Margazoglou, “Lévy noise versus Gaussian-noise-induced transitions in the Ghil–Sellers energy balance model,” *Nonlinear Process. Geophys.* **29**, 183–205 (2022).
- ⁴³K. Bhattacharya, M. Ghil, and I. L. Vulis, “Internal variability of an energy-balance model with delayed albedo effects,” *J. Atmos. Sci.* **39**, 1747–1773 (1982).
- ⁴⁴M. Ghil, M. D. Chekroun, and G. Stepan, “A collection on ‘climate dynamics: Multiple scales and memory effects’, introduction,” *Proc. R. Soc. A* **471**, 20150097 (2015).
- ⁴⁵S. A. Soldatenko and R. A. Colman, “Power spectrum sensitivity analysis of the global mean surface temperature fluctuations simulated in a two-box stochastic energy balance model,” *Tellus A: Dyn. Meteorol. Oceanogr.* **74**, 68–84 (2022).
- ⁴⁶C. Penland, “Random forcing and forecasting using principal oscillation pattern analysis,” *Month. Weather Rev.* **117**, 2165–2185 (1989).
- ⁴⁷H.-B. Fredriksen and M. Rypdal, “Long-range persistence in global surface temperatures explained by linear multibox energy balance models,” *J. Clim.* **30**, 7157–7168 (2017).
- ⁴⁸C. Frankignoul and K. Hasselmann, “Stochastic climate models. Part II. Application to sea-surface temperature anomalies and thermocline variability,” *Tellus* **29**, 289–305 (1977).
- ⁴⁹C. W. Gardiner, *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences*, 3rd ed. (Springer Verlag, Berlin, 2004).
- ⁵⁰P. Lemke, “Stochastic climate models. Part III. Application to zonally averaged energy models,” *Tellus* **29**, 385–392 (1977).
- ⁵¹J. M. Mitchell, “Stochastic models of air-sea interaction and climatic fluctuation,” Tech. Rep. Mem. RM-5233-NSF (The RAND Corporation, Santa Monica, California, 1966).
- ⁵²“Oral history interview with Klaus Hasselmann,” <http://www.aip.org/history-programs/niels-bohr-library/oral-histories/33645> (AIP Publishing LLC, 2006).
- ⁵³F. C. Klebaner, *Introduction to Stochastic Calculus with Applications*, 3rd ed. (Imperial College Press, 2012).
- ⁵⁴P. M. Cox, C. Huntingford, and M. S. Williamson, “Emergent constraint on equilibrium climate sensitivity from global temperature variability,” *Nature* **553**, 319–322 (2018).
- ⁵⁵I. Di Terlizzi, F. Ritort, and M. Baiesi, “Explicit solution of the generalised Langevin equation,” *J. Stat. Phys.* **181**, 1609–1635 (2020).
- ⁵⁶W. D. Nordhaus, “An optimal transition path for controlling greenhouse gases,” *Science* **258**, 1315–1319 (1992).
- ⁵⁷L. Barrage, “The Nobel memorial prize for William D. Nordhaus,” *Scand. J. Econ.* **121**, 884–924 (2019).
- ⁵⁸W. D. Nordhaus and Z. Yang, “A regional dynamic general-equilibrium model of alternative climate-change strategies,” *Am. Econ. Rev.* **86**, 741–765 (1996).
- ⁵⁹C. Hope, “The marginal impact of CO₂ from PAGE2002: An integrated assessment model incorporating the IPCC’s five reasons for concern,” *Integr. Assess. J.* **6**, 19–56 (2006).
- ⁶⁰R. S. Tol, “On the optimal control of carbon dioxide emissions: An application of fund,” *Environ. Model. Assess.* **2**, 151–163 (1997).
- ⁶¹V. Bosetti, C. Carraro, M. Galeotti, E. Massetti, and M. Tavoni, “A world induced technical change hybrid model,” *Energy J.* **27**, 13–37 (2006).
- ⁶²F. Lamperti, G. Dosi, M. Napoletano, A. Roventini, and A. Sapio, “Faraway, so close: Coupled climate and economic dynamics in an agent-based integrated assessment model,” *Ecol. Econ.* **150**, 315–339 (2018).
- ⁶³M. Ghil, “Review article: Hilbert problems for the climate sciences in the 21st century—20 years later,” *Nonlinear Process. Geophys.* **27**, 429–451 (2020).
- ⁶⁴M. Ghil, “Foreword,” in *Handbook of Climate Change Mitigation and Adaptation*, 3rd ed., edited by M. Lackner, B. Sajjadi, and W. Y. Chen (Springer Nature, 2022), pp. vii–ix.
- ⁶⁵K. Hasselmann, S. Hasselmann, R. Giering, V. Ocana, and H. von Storch, “Sensitivity study of optimal CO₂ emission paths using a simplified structural integrated assessment model (SIAM),” *Clim. Change* **37**, 345–386 (1997).
- ⁶⁶K. Hasselmann, “Climate-change research after Kyoto,” *Nature* **390**, 225–226 (1997).
- ⁶⁷R. Calel and D. A. Stainforth, “On the physics of three integrated assessment models,” *Bull. Am. Meteorol. Soc.* **98**, 1199–1216 (2017).
- ⁶⁸S. Dietz, J. Rising, T. Stoerk, and G. Wagner, “Economic impacts of tipping points in the climate system,” *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2103081118 (2021).
- ⁶⁹D. L. Kelly and C. D. Kolstad, “Bayesian learning, growth, and pollution,” *J. Econ. Dyn. Control* **23**, 491–518 (1999).
- ⁷⁰A. J. Leach, “The climate change learning curve,” *J. Econ. Dyn. Control* **31**, 1728–1752 (2007).
- ⁷¹D. L. Kelly and Z. Tan, “Learning and climate feedbacks: Optimal climate insurance and fat tails,” *J. Environ. Econ. Manage.* **72**, 98–122 (2015).
- ⁷²D. Lemoine and I. Rudik, “Managing climate change under uncertainty: Recursive integrated assessment at an inflection point,” *Ann. Rev. Resour. Econ.* **9**, 117–142 (2017).
- ⁷³R. Calel, S. C. Chapman, D. A. Stainforth, and N. W. Watkins, “Temperature variability implies greater economic damages from climate change,” *Nat. Commun.* **11**, 5028 (2020).
- ⁷⁴S. Jensen and C. P. Traeger, “Pricing climate risk” (2021). CESifo Working Paper No. 9196, Available at SSRN: <https://ssrn.com/abstract=3892620> or <http://dx.doi.org/10.2139/ssrn.3892620>
- ⁷⁵C. Leith, “Climate signal and weather noise,” in *Applications of Statistics to Modeling the Earth’s Climate System*, edited by R. A. Madden, R. W. Katz (NCAR, Colorado, 1994), Technical Note NCAR/TN-409+PROC.
- ⁷⁶W. Moon, S. Agarwal, and J. S. Wettlaufer, “Intrinsic pink-noise multidecadal global climate dynamics mode,” *Phys. Rev. Lett.* **121**, 108701 (2018).
- ⁷⁷K. Rypdal, “Global temperature response to radiative forcing: Solar cycle versus volcanic eruptions,” *J. Geophys. Res. Atmos.* **117**, D06115 (2012).
- ⁷⁸M. Rypdal, H. Fredriksen, E. Myrsvoll-Nilsen, K. Rypdal, and S. H. Sorbye, “Emergent scale invariance and climate sensitivity,” *Climate* **6**, 93 (2018).
- ⁷⁹L. Del Rio Amador and S. Lovejoy, “Predicting the global temperature with the Stochastic Seasonal to Interannual Prediction System (StocSIPS),” *Clim. Dyn.* **53**, 4373 (2019).
- ⁸⁰S. A. McKinley and H. D. Nguyen, “Anomalous diffusion and the generalized Langevin equation,” *SIAM J. Math. Anal.* **50**, 5119–5160 (2018).
- ⁸¹N. Pottier, *Nonequilibrium Statistical Physics* (Oxford University Press, Oxford, 2010).
- ⁸²W. T. Coffey and Y. P. Kalmykov, *The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering*, 4th ed. (World Scientific, Singapore, 2017).
- ⁸³R. F. Fox, “The generalized Langevin equation with Gaussian fluctuations,” *J. Math. Phys.* **18**, 2331–2335 (1977).
- ⁸⁴E. Lutz, “Fractional Langevin equation,” *Phys. Rev. E* **64**, 051106 (2001).
- ⁸⁵R. Metzler, J. Jeon, A. G. Cherstvy, and E. Barkai, “Anomalous diffusion models and their properties: Non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking,” *Phys. Chem. Chem. Phys.* **16**, 24128 (2014).
- ⁸⁶J. M. Porrà, K. G. Wang, and J. Masoliver, “Generalized Langevin equations: Anomalous diffusion and probability distributions,” *Phys. Rev. E* **53**, 2044–2056 (1996).
- ⁸⁷V. Koblelev and E. Romanov, “Fractional Langevin equation to describe anomalous diffusion,” *Progr. Theor. Phys. Suppl.* **139**, 470–476 (2000).
- ⁸⁸F. Mainardi and P. Pironi, “The fractional Langevin equation: Brownian motion revisited,” *Extracta Math.* **11**, 140–154 (1996). <http://eudml.org/doc/38480>
- ⁸⁹A. V. Chechkin and R. Klages, “Fluctuation relations for anomalous dynamics,” *J. Stat. Mech.* **L03002**, 1–11 (2009).
- ⁹⁰A. V. Chechkin, F. Lenz, and R. Klages, “Normal and anomalous fluctuation relations for Gaussian stochastic dynamics,” *J. Stat. Mech.* **L11001**, 1–13 (2012).

- ⁹¹K. B. Oldham and J. Spanier, *The Fractional Calculus: Theory and Applications of Differentiation and Integration to Arbitrary Order* (Dover, 2006).
- ⁹²I. Podlubny, *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*, Mathematics in Science and Engineering No. 198 (Academic Press, 1999).
- ⁹³M. Meerschaert and A. Sikorskii, *Stochastic Models for Fractional Calculus*, 2nd ed. (de Gruyter, Berlin, 2019).
- ⁹⁴Y. Mishura, *Stochastic Calculus for Fractional Brownian Motion and Related Processes* (Springer, Berlin, 2008).
- ⁹⁵Q. Wei, W. Wang, H. Zhou, R. Metzler, and A. Chechkin, "Time-fractional Caputo derivative versus other integro-differential operators in generalized Fokker-Planck and generalized Langevin equations," *Phys. Rev. E* **108**, 024125 (2023).
- ⁹⁶T. S. Doan, P. T. Huong, P. E. Kloeden, and A. M. Vu, "Euler-Maruyama method for Caputo stochastic fractional differential equations," *J. Comput. Appl. Math.* **380**, 112989 (2020).
- ⁹⁷W. Wang, S. Cheng, Z. Guo, and X. Yan, "A note on the continuity for Caputo fractional stochastic differential equations," *Chaos* **30**, 073106 (2020).
- ⁹⁸K. Rybakov and A. Yushchenko, "Spectral method for solving linear Caputo fractional stochastic differential equations," *IOP Conf. Ser.: Mater. Sci. Eng.* **927**, 012077 (2020).
- ⁹⁹Z. Guo, X. Han, and J. Hu, "Averaging principle for stochastic Caputo fractional differential equations with non-Lipschitz condition," *Fract. Calc. Appl. Anal.* (2023).
- ¹⁰⁰P. T. Huong, P. Kloeden, and D. T. Son, "Well-posedness and regularity for solutions of Caputo stochastic fractional differential equations in L^p spaces," *Stochast. Anal. Appl.* **41**, 1–15 (2023).
- ¹⁰¹Y. Mishura, *Fractional Calculus and Fractional Stochastic Calculus, Including Rough-paths, with Applications* (Invited Lectures to London Mathematical Society, 2020).
- ¹⁰²L. Li, J.-G. Liu, and J. Lu, "Fractional stochastic differential equations satisfying fluctuation-dissipation theorem," *J. Stat. Phys.* **169**, 316–339 (2017).
- ¹⁰³O. C. Ibe, *Markov Processes for Stochastic Modelling*, 2nd ed. (Elsevier, London, 2013).
- ¹⁰⁴F. Biagini, Y. Hu, B. Oksendal, and T. Zhang, *Stochastic Calculus for Fractional Brownian Motion and Applications* (Springer, London, 2010).
- ¹⁰⁵C. Hohenegger and S. McKinley, "Fluid-particle dynamics for passive tracers advected by a thermally fluctuating viscous medium," *J. Comput. Phys.* **340**, 688–711 (2017).
- ¹⁰⁶G. A. Pavliotis, G. Stoltz, and U. Vaes, "Scaling limits for the generalized Langevin equation," *J. Nonlinear Sci.* **31**, 8 (2021).
- ¹⁰⁷R. Kupferman, "Fractional kinetics in Kac–Zwanzig heat bath models," *J. Stat. Phys.* **114**, 291–326 (2004).
- ¹⁰⁸D. T. Gillespie, "Exact numerical simulation of the Ornstein-Uhlenbeck process and its integral," *Phys. Rev. E* **54**, 2084–2091 (1996).
- ¹⁰⁹J. Zhang, Y. Tang, and J. Huang, "A fast Euler-Maruyama method for fractional stochastic differential equations," *J. Appl. Math. Comput.* **69**, 273–291 (2023).
- ¹¹⁰J. Zhang, J. Lv, J. Huang, and Y. Tang, "A fast Euler-Maruyama method for Riemann-Liouville stochastic fractional nonlinear differential equations," *Physica D* **446**, 133685 (2023).
- ¹¹¹M. Meinshausen, S. Smith, and K. Calvin, "The RCP greenhouse gas concentrations and their extensions from 1765 to 2300," *Clim. Change* **109**, 213 (2011).
- ¹¹²J. Slezak, "Langevin equation and fractional dynamics," Ph.D. thesis (2018).
- ¹¹³J. A. Kassel, "Nonlinear long-range correlated stochastic models of temperature time series: Inference and prediction," Ph.D. thesis (Technische Universitaet Dresden, Dresden, 2024).
- ¹¹⁴J. A. Kassel and H. Kantz, "Statistical inference of one-dimensional persistent nonlinear time series and application to predictions," *Phys. Rev. Res.* **4**, 013206 (2022).
- ¹¹⁵P. L. Langen and V. A. Alexeev, "Estimating $2 \times \text{CO}_2$ warming in an aquaplanet GCM using the fluctuation-dissipation theorem," *Geophys. Res. Lett.* **32**, L23708, <https://doi.org/10.1029/2005GL024136> (2005).
- ¹¹⁶R. Kubo, "The fluctuation-dissipation theorem," *Rep. Prog. Phys.* **29**, 255–284 (1966).
- ¹¹⁷R. Klages, A. V. Chechkin, and P. Dieterich, "Anomalous fluctuation relations," in *Nonequilibrium Statistical Physics of Small Systems*, edited by R. Klages, W. Just, and C. Jarzynski (Wiley-VCH, Weinheim, 2013), pp. 259–282.
- ¹¹⁸P. Dieterich, R. Klages, and A. V. Chechkin, "Fluctuation relations for anomalous dynamics generated by time-fractional Fokker-Planck equations," *New J. Phys.* **075004**, 1–14 (2015).
- ¹¹⁹I. M. Sokolov and Y. Klafter, *First Steps in Random Walks: From Tools to Applications* (Oxford University Press, Oxford, 2011).
- ¹²⁰I. M. Sokolov, J. Klafter, and A. Blumen, "Fractional kinetics," *Phys. Today* **55**, 48–54 (2002).
- ¹²¹S. Samko, A. Kilbas, and O. Marichev, *Fractional Integrals and Derivatives—Theory and Applications* (Gordon and Breach, 1993).
- ¹²²R. Gorenflo and A. Mainardi, "Fractional calculus," in *Fractals and Fractional Calculus in Continuum Mechanics*, edited by A. Carpentieri and F. Mainardi (Springer-Verlag Wien GmbH, 1997), pp. 223–276.
- ¹²³R. Hilfer, "Threefold introduction to fractional derivatives," in *Anomalous Transport: Foundations and Applications*, edited by R. Klages, G. Radons, and I. Sokolov (Wiley-VCH, Weinheim, 2008), pp. 17–XX.