

Hasselmann's Program Revisited: The Analysis of Stochasticity in Deterministic Climate Models

Ludwig Arnold
Institut für Dynamische Systeme
Fachbereich 3, Universität, Postfach 33 04 40
28334 Bremen, Germany
email: arnold@math.uni-bremen.de

Abstract

In his seminal 1976 paper on “Stochastic Climate Models”, K. Hasselmann proposed to improve deterministic models for the “climate” (slow variables) by incorporating the influence of the “weather” (fast variables) in the form of random noise.

We will recast this program in the language of modern probability theory as follows: While the transition from a GCM (general circulation model) to an SDM (statistical dynamical model) (both deterministic) is facilitated by the *method of averaging*, stochasticity comes into the picture when studying the error made in the averaging procedure, provided that the fast variables are sufficiently “chaotic”.

The study of *normal deviations* from the averaged system is described by the central limit theorem, while the study of *large deviations* from the average (events happening on an exponential time scale) is done by the theory of large deviations. We feel that the latter should be particularly appealing to meteorologists, as one can, for example, describe the “hopping” of the climate between its various local attractors due to the forcing by chaotic weather.

Key words and phrases: Method of averaging, central limit theorem, large deviations, stochastic climate models.

AMS 1991 subject classification: primary 34C29, 60F05, 60F10, 86A10; secondary 34F05, 60H10.

1 Introduction

Twentyfour years ago, K. Hasselmann in his fundamental and visionary paper [13] entitled “Stochastic Climate Models. Part I. Theory” laid out a new

approach for explaining and understanding climate variability and stochasticity.

The aim of the present paper is to describe Hasselmann's approach (which we call *Hasselmann's Program*) in modern mathematical language and add some new points of view to it.

Hasselmann's paper (Part I was followed by Parts II [7] and III [22] in which applications are given) aroused intensive activities by meteorologists (among them several contributors to these Proceedings) for about 10 years. For a good review see Olbers, Lemke and Wolf-Gladrow [30, Chap. 4] and the contribution of Olbers [29] to these Proceedings. Later the interest in stochastic climate models diminished.

An exception was the area of "stochastic resonance" which was offered to explain the glacial cycles as the combined effect of external astronomical periodic forcing and internal climatic noise, see the contributions of Imkeller [15] and of Freund, Neiman and Schimansky-Geier [12] to these Proceedings.

We believe that the dwindling interest in Hasselmann's program was caused, on the one hand, by disillusionment and frustration as the program did not seem to live up to its original expectations, and, on the other hand, by the availability of computing power that made qualitative research less pressing.

We claim that Hasselmann's program has never really been implemented. We believe that this is also caused by the fact that, to our knowledge, no mathematician has ever really worked on it. This paper is an attempt to change this situation as we are convinced that the program has not lost its importance and significance, and that there are today new promising ways and means which are worth being tried.

Hasselmann starts with the basic observation that a characteristic feature of climatic records is their pronounced *variability*, meaning that climatic time series have a positive continuous power spectrum (spectral density) over the whole range of frequencies (with an increase in spectral energy with decreasing frequency).

We can add today that this enormous variability is also shown by computer data produced by means of the most advanced deterministic (!) climate models, such as the ECHAM 3 and LSG General Circulation Models [5, 6]. In short, most output data look like (colored) random noise.

"An understanding of the origin of climatic variability, in the entire spectral range from extreme ice age changes to seasonal anomalies, is a primary goal of climate research" [13, p. 473] – to which we have nothing to add, as there is still no generally accepted explanation for the above phenomenon.

Hasselmann's *Ansatz* is to attribute climate variability to the existence of various time scales, more precisely to internal random forcing by the short time scale "weather" components (all variables with period up to 15 days, say, mainly related to atmospheric processes) of the system. Slowly responding components of the system (such as the ocean, the cryosphere and the biosphere) act as integrators of this random input in much the same way as a pollen grain in a liquid integrates the short time impact of the molecules to yield Brownian

motion. In short, the “weather” drives the “climate” through internal random forcing.

In fact, although most of the simple deterministic models fail to reproduce the observed as well as computed variability, it is a striking feature of even the simplest possible stochastic models that they produce the correct power spectrum.

Consider, for example, the scalar linear equation with additive white noise

$$\dot{x}_t = -\alpha x_t + \sigma \xi_t, \quad \alpha, \sigma > 0, \quad \xi_t \text{ white noise,}$$

correctly written as an Itô stochastic differential equation

$$dx_t = -\alpha x_t dt + \sigma dW_t,$$

where W_t is Brownian motion (also called a **Wiener process** by mathematicians) and white noise $\xi_t = \dot{W}_t$ is the derivative of W_t in the sense of generalized functions. The unique stationary solution

$$x_t = \sigma e^{-\alpha t} \int_{-\infty}^t e^{\alpha s} dW_s,$$

known as the Ornstein-Uhlenbeck process, has spectral density

$$f(\lambda) = \frac{\sigma^2}{2\pi\alpha} \frac{1}{\lambda^2 + \alpha^2}.$$

Our key problem in substantiating Hasselmann’s idea is now how to turn fast, but deterministic motion into random forcing in a mathematically rigorous manner. How can randomness emerge in a deterministic system? This question has been investigated for a long time and is discussed using terms like “mixing”, “decay of correlation” and “hyperbolicity”. For a recent treatment see Viana [34] and the references therein. See also the discussions in the contributions of Kifer [18] and of Rödenbeck, Beck and Kantz [31] to these Proceedings.

Warning: The technical level of our subject matter is very high so that a mathematically rigorous treatment would be beyond the limits of these Proceedings. We hence adopt a style that is conceptual and narrative, and most statements are true only “modulo technical details”, some of them are even mere speculations – thus our statements should not be quoted as a mathematical reference.

2 Stochasticity in Deterministic Climate Models With Two Separate Time Scales

2.1 Hasselmann’s Approach

Suppose we have a deterministic **General Circulation Model (GMC)** that describes the climate-weather dynamics by an ordinary differential equation

(ODE) (also called “prognostic equation”) of the form

$$\dot{z} = h(z)$$

in a finite-dimensional space, like the ECHAM models [5, 6].

The **basic assumption** for Hasselmann (as well as in this paper) is that we can separate the components of z as

$$z = (x, y)$$

with strongly differing “response times” τ_x and τ_y , where y is the vector of *fast* variables (with response times of “a few days”, typically from the atmosphere), and x is the vector of *slow* variables (with response times of “several months, years and longer”, typically related to the ocean, the cryosphere and the biosphere), so that $\tau_y \ll \tau_x$. We henceforth call y the weather variables and x the climate variables (omitting quotation marks).

Mathematically, strongly differing response times amount to the fact that we can introduce a small scaling parameter ε such that the GCM $\dot{z} = h(z)$ is equivalent to the coupled system of the two ODE

$$\dot{x} = f(x, y), \quad x = \text{climate, slow variables}, \quad (2.1)$$

$$\dot{y} = \frac{1}{\varepsilon} g(x, y), \quad y = \text{weather, fast variables}, \quad (2.2)$$

so that $\tau_y \approx \varepsilon \ll \tau_x \approx 1$. In this paper, we remain on this formal mathematical level and do not try to further specify x , y , f and g on the grounds of actual physical climate models.

Statistical Dynamical Models (SDM) are derived from the scaled GCM (2.1,2.2) by averaging the fast variables out of eq. (2.1), arriving at an averaged ODE for $u := \langle x \rangle$,

$$\dot{u} = F(u), \quad (2.3)$$

where $F(x) := \langle f(x, y) \rangle$ is the average over the fast variables y for “frozen” slow variables x . In spite of their name, SDM are in fact deterministic rather than statistical.

Hasselmann’s basic idea was to establish models called *stochastic models* that are “in between” the GCM (2.1,2.2) and the SDM (2.3). These are equations for x only, thus simpler than the GCM, but more precise than the SDM as the non-averaged weather components are retained and appear formally as random forcing terms.

We will try to make this procedure, described by Hasselmann in “physicist’s style”, mathematically precise.

2.2 Hasselmann's Program in Mathematical Language

We claim that the basic procedure of passing from a GCM to a SDM is the well-known and rather classical *Method of Averaging* (also called the *Averaging Principle*). This method has long been used throughout the 18th and 19th century in celestial mechanics (e.g. by Lagrange, Laplace and others). The first mathematically rigorous justification was provided by Bogolyubov and Mitropolskii [2] only about 40 years ago, establishing that, modulo assumptions, $\lim_{\varepsilon \rightarrow 0} x_t^\varepsilon = u_t$ on the time interval $[0, T]$, where x_t^ε is the x component of the solution of (2.1,2.2), while u_t is the solution of (2.3), both starting at the same initial value.

We are also convinced that stochastic models like Hasselmann's mathematically amount to studying the error in the Method of Averaging, i.e. the deviation of x_t^ε from u_t in the course of time.

Working on a fixed time interval $[0, T]$ it was discovered by Khasminskii [17] in 1966 that if the fast motion is a stochastic process then

$$\zeta_t^\varepsilon := \frac{1}{\sqrt{\varepsilon}}(x_t^\varepsilon - u_t)$$

has a limiting Gaussian (or normal) distribution as $\varepsilon \rightarrow 0$, a fact that is known as the *Central Limit Theorem*.

However, there are many qualitative phenomena of x_t^ε which are captured *neither* by the Method of Averaging *nor* by the Central Limit Theorem as they happen on longer time scales (of order $\varepsilon^{1/\varepsilon}$). These include: x_t^ε leaving a fixed neighborhood of a stable steady state of $\dot{u} = F(u)$, the "hopping" of x_t^ε between various local attractors of $\dot{u} = F(u)$, or the building-up of invariant measures μ^ε for x_t^ε . The study of those so-called *Large Deviations* from the averaged system was initiated by Freidlin and Wentzell in 1979 and has since become one of the pillars of modern probability theory.

We hence believe that Hasselmann's program can be spelled-out today in the following form.

Mathematical Synopsis of Hasselmann's Program

Given a dynamical system described by the coupled system of ODE (2.1,2.2), with $\varepsilon > 0$ a smallness parameter, and solutions $(x_t^\varepsilon, y_t^\varepsilon)$. Let u_t denote the solution of the averaged ODE (2.3), with $x_0^\varepsilon = u_0$.

1. **Method of Averaging:** Study conditions under which $x_t^\varepsilon \rightarrow u_t$ on a fixed time interval $[0, T]$.
2. **Normal Deviations from the Averaged System, Central Limit Theorem:** Study the error $x_t^\varepsilon - u_t$ made in the Method of Averaging on a fixed time interval $[0, T]$ and prove the Central Limit Theorem for it, i.e. prove that

$$\zeta_t^\varepsilon := \frac{1}{\sqrt{\varepsilon}}(x_t^\varepsilon - u_t)$$

is asymptotically normally distributed as $\varepsilon \rightarrow 0$.

3. **Large Deviations from the Averaged System:** Study the long-term phenomena on a time scale $\varepsilon^{1/\varepsilon}$ not captured by the Method of Averaging and the Central Limit Theorem.

Since the Method of Averaging can be considered a *Law of Large Numbers* in case when the fast motion is a stochastic process, Hasselmann’s program contains indeed all three basic asymptotic methods of probability theory: the Law of Large Numbers, the Central Limit Theorem, and the Theory of Large Deviations.

We will now briefly describe the essence of all three methods. We first present what we call the “classical” case in which the fast motion is not back-coupled to the slow motion and has nice statistical properties. Then we discuss the more complicated case where $(x_t^\varepsilon, y_t^\varepsilon)$ is the solution of the cross-coupled system (2.1,2.2) which, for short, we call “Hasselmann’s case”.

We refrain, however, from presenting complete mathematical theorems with a full list of all technical assumptions, but rather concentrate on the conceptual side. We also do not give proofs, and some statements do not have proofs (yet). Hence, we repeat, our statements here should not be cited to justify any analytic derivation.

3 The Method of Averaging

We refer to Sanders and Verhulst [32] for more details and the history of the method.

3.1 Method of Averaging: Classical Case

Consider the ODE in \mathbb{R}^d

$$\dot{z}_\tau = \varepsilon f(z_\tau, \xi_\tau), \quad z_0 = z, \tag{3.1}$$

where ξ_τ is some forcing function. By the continuous dependence of the solution $z_\tau^\varepsilon = z_\tau^\varepsilon(z)$ of the initial value problem on the right-hand side of (3.1) clearly

$$\lim_{\varepsilon \rightarrow 0} z_\tau^\varepsilon(z) = z \quad \text{uniformly on } [0, T]$$

for any fixed finite time $T > 0$. In other words, the solution of (3.1) will for small ε not develop any interesting features on a finite time interval $[0, T]$.

The situation changes if we look at time intervals $[0, \frac{T}{\varepsilon}]$ or longer whose length increases as $\varepsilon \rightarrow 0$. Then $z_\tau^\varepsilon(z)$ can indeed develop significant excursions from the initial value.

It is hence convenient to speed up time by putting

$$t = \varepsilon\tau, \quad x_t = x_t^\varepsilon := z_{t/\varepsilon}^\varepsilon, \quad \tau \in [0, \frac{T}{\varepsilon}] \iff t \in [0, T]$$

to arrive at our standard form of (3.1), namely

$$\dot{x}_t = f(x_t, \xi_{t/\varepsilon}), \quad x_0 = x, \quad t \in [0, T], \quad (3.2)$$

where $\xi_{t/\varepsilon}$ is now obviously a *fast* forcing function and the time interval $[0, T]$ is once again independent of ε .

We now average out ξ from the right-hand side of (3.2) at a “frozen” value of x , i.e. we form

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x, \xi_t) dt = \int f(x, y) \mu(dy) =: F(x), \quad (3.3)$$

where the time average exists for all x provided e.g. ξ_t is periodic, quasiperiodic, almost-periodic, stationary stochastic, . . . , and where we have used the ergodic theorem in (3.3) to replace the time average with the phase average with respect to the invariant measure μ of ξ_t (uniqueness assumed).

Assuming the knowledge of $\mu(dy)$ and calculating the average $F(x)$ of $f(x, y)$ under $\mu(dy)$ via (3.3) amounts to solving what is called the *closure problem* for the slow variables x .

The statement of the Method of Averaging for our case reads

Method of Averaging (Classical Case):

Let $x_t^\varepsilon(x)$ be the solution of (3.2). Then we have

$$\lim_{\varepsilon \rightarrow 0} x_t^\varepsilon(x) = u_t(x) \quad \text{uniformly in } [0, T] \quad \text{for all initial values } x, \quad (3.4)$$

where $u_t(x)$ is the solution of the **averaged ODE**

$$\dot{u}_t = F(u_t), \quad u_0 = x, \quad (3.5)$$

whose right-hand side is defined by (3.3).

In case ξ_t is a random process the convergence in (3.4) is “in probability”, i.e.

$$\lim_{\varepsilon \rightarrow 0} \mathbb{P} \left\{ \max_{0 \leq t \leq T} \|x_t^\varepsilon(x) - u_t(x)\| > \delta \right\} = 0 \quad \text{for all } \delta > 0 \quad \text{and all } x.$$

We stress that both the non-averaged motion $x_t^\varepsilon(x)$ and the averaged motion $u_t(x)$ start at the *same* initial value x .

Example: Let ξ be a zero mean ergodic stationary stochastic process and consider

$$\dot{x}_t = b(x_t) + \sigma(x_t)\xi_{t/\varepsilon}.$$

Then $\dot{u}_t = b(u_t)$.

3.2 Method of Averaging: Hasselmann's Case

We now consider the system

$$\dot{x}_t = f(x_t, y_t), \quad x_0 = x \in \mathbb{R}^d \text{ (climate, slow variables)}, \quad (3.6)$$

$$\dot{y}_t = \frac{1}{\varepsilon} g(x_t, y_t), \quad y_0 = y \in \mathbb{R}^m \text{ (weather, fast variables)}. \quad (3.7)$$

Denote by $(x_t^\varepsilon, y_t^\varepsilon) = (x_t^\varepsilon(x, y), y_t^\varepsilon(x, y))$ its solution with the initial value (x, y) . Now the slow and fast variables are cross-coupled to each other.

The task of the Method of Averaging is again to average out y from f such that $x_t^\varepsilon(x, y)$ converges to $u_t(x)$, the solution of the averaged equation. This will, of course, only work if y is sufficiently "chaotic".

The situation now becomes less nice and much more complicated than in the classical case. We refer the reader to Kifer [20, Remark 2.5] and to Sect. 3 of Kifer's contribution [18] to these Proceedings for the only presently available mathematically rigorous results. See also Chap. 7 of Freidlin and Wentzell [11] for a general introduction.

To prepare the averaging procedure for (3.6, 3.7) we consider the auxiliary weather ODE

$$\dot{y}_t = g(x, y_t) \quad y_0 = y, \quad (3.8)$$

with frozen climate variables x , and denote by $(t, y) \mapsto \varphi_t^x(y)$ the dynamical system (flow) generated by (3.8) indexed by the fixed parameter x .

The first basic difficulty is that the dynamical system φ^x typically has many invariant measures $\mu_x(dy)$ (except in the rare case of unique ergodicity) which clearly also depend on x .

The average of $f(x, \cdot)$ over y now becomes

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x, \varphi_t^x(y)) dt = \int_{\mathbb{R}^m} f(x, y) \mu_x(dy) =: F_{\mu_x}(x), \quad (3.9)$$

where equality in (3.9) only holds for μ_x -almost all y , and the (non-uniform) limit depends on our choice of μ_x and its respective domain of attraction, hence is different for different y 's.

Example: Let $\dot{y}_t = g(x, y_t) = xy_t - y_t^3$ for $x > 0$. Then φ^x has the three Dirac measures $\mu_x^1 = \delta_0$ (unstable) and $\mu_x^{1,2} = \delta_{\pm\sqrt{x}}$ (stable) as invariant measures, so that the limits in (3.9) are

$$F_{\mu_x^i}(x) = \begin{cases} F_{\mu_x^0}(x) = f(x, 0), & y = 0, \\ F_{\mu_x^1}(x) = f(x, \sqrt{x}), & y > 0, \\ F_{\mu_x^2}(x) = f(x, -\sqrt{x}), & y < 0. \end{cases}$$

Here is a form of the Method of Averaging for the above situation we wish to have implemented.

Method of Averaging (Hasselmann's Case):

Let (μ_x) be a family of invariant measures for the auxiliary weather dynamical system generated by (3.8). Then

$$\lim_{\varepsilon \rightarrow 0} x_t^\varepsilon(x, y) = u_t(x) \quad (3.10)$$

on $[0, T]$ for all $x \in \mathbb{R}^d$ and μ_x -almost all $y \in \mathbb{R}^m$, where $u_t(x)$ is the solution of the averaged equation

$$\dot{u}_t = F_{\mu_{u_t}}(u_t), \quad u_0 = x, \quad (3.11)$$

and $F_{\mu_x}(x)$ is defined by (3.9).

If φ^x is nicely hyperbolic then it has a particular invariant measure μ_x^{SRB} called the “physical” or SRB (Sinai-Ruelle-Bowen) measure (see Viana [34]) which builds-up as the occupation measure of orbits $\varphi_t^x(y)$ for Lebesgue-almost all initial values y . In this case we would like (3.10) to hold for all climate initial conditions x and Lebesgue-almost all weather initial conditions y .

4 Normal Deviations from the Averaged System: The Central Limit Theorem

4.1 Central Limit Theorem: Classical Case

We now “improve” the averaged motion u_t by taking the error made in the Method of Averaging into account.

We first consider the case

$$\dot{x}_t^\varepsilon = f(x_t^\varepsilon, \xi_{t/\varepsilon}), \quad x_0 = x, \quad (4.1)$$

where $\xi_{t/\varepsilon}$ is a fast motion that does not depend on the variables x .

In case ξ_t is T -periodic, the averaged equation is

$$\dot{u} = F(u) := \frac{1}{T} \int_0^T f(u, \xi_t) dt,$$

and the error $x_t^\varepsilon(x) - u_t(x)$ can be asymptotically expanded as

$$x_t^\varepsilon(x) - u_t(x) = \varepsilon u_1(t) + \varepsilon^2 u_2(t) + \dots \quad (4.2)$$

(see Sanders and Verhulst [32] for a precise statement).

Thus the first error term is of order ε , the second of order ε^2 etc, and no stochasticity appears.

The reason is that a periodic motion is not sufficiently “mixing” or “chaotic”.

In case the fast motion ξ happens to be sufficiently “mixing” (typically measured by conditions on the sufficiently rapid decay of the correlation of ξ_0

and ξ_t as $t \rightarrow \infty$), then $x_t^\varepsilon(x) - u_t(x)$ is of a character completely different from (4.2). The error $x_t^\varepsilon(x) - u_t(x)$ is then of order $\sqrt{\varepsilon}$, but **no** further term of the asymptotic expansion can be written down.

To be more precise, $\frac{1}{\sqrt{\varepsilon}}(x_t^\varepsilon(x) - u_t(x))$ does not converge, but has a limiting probability distribution which is Gaussian.

Here is a more detailed mathematical formulation of this phenomenon which was discovered by Khasminskii [17].

Normal Deviations from the Averaged System: The Central Limit Theorem (Classical Case):

Assume that the Method of Averaging as formulated in Subsection 3.1 applies to (4.1). Then for all $x \in \mathbb{R}^d$

$$\zeta_t^{\varepsilon, x} := \frac{1}{\sqrt{\varepsilon}}(x_t^\varepsilon(x) - u_t(x)) \Rightarrow \zeta_t^x \text{ as } \varepsilon \rightarrow 0 \text{ on } [0, T], \quad (4.3)$$

where ζ_t^x is a Gauss-Markov process on $[0, T]$ (also called Ornstein-Uhlenbeck process) given by the solution of the linear stochastic differential equation

$$d\zeta_t^x = DF(u_t(x)) \zeta_t^x dt + \sqrt{\sigma(u_t(x))} dW_t, \quad \zeta_0^x = 0, \quad (4.4)$$

written in form of a Langevin equation as

$$\dot{\zeta}_t^x = DF(u_t(x)) \zeta_t^x + \sqrt{\sigma(u_t(x))} \dot{W}_t,$$

where \dot{W} is white noise. Here \Rightarrow means “weak convergence” (i.e. the probability distribution of the left-hand side converges to the probability distribution of the right-hand side as $\varepsilon \rightarrow 0$), W_t is a standard Brownian motion in \mathbb{R}^d , the $d \times d$ matrix

$$DF(u) = \left(\frac{\partial^i F}{\partial u_j} \right) (u)$$

is the Jacobian of the averaged vector field F , and the positive-definite $d \times d$ matrix σ is defined as the “nonlinear average”

$$\sigma(x) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \int_0^T \text{cov}(f(x, \xi_t), f(x, \xi_s)) ds dt = 2\pi S_x(0), \quad (4.5)$$

where $S_x(\omega)$ is the spectral density matrix of $\eta_t := f(x, \xi_t)$ for frozen x at frequency ω .

In (4.5), $\text{cov}(X, Y)$ denotes the covariance matrix of the vectors X and Y .

The **result** is that the averaged equation $\dot{u}_t = F(u_t)$ is replaced by the more informative pair of equations

$$\dot{u}_t = F(u_t), \quad u_0 = x, \quad (4.6)$$

$$\dot{\zeta}_t = DF(u_t) \zeta_t + \sqrt{\sigma(u_t)} \dot{W}_t, \quad \zeta_0 = 0. \quad (4.7)$$

The pair (4.6, 4.7) has a triangular or skew product structure, i.e. equation (4.6) is decoupled from (4.7), while the solution u_t of (4.6) forces the stochastic differential equation (4.7).

Comparing the original process x_t^ε with the averaged process u_t we can write

$$x_t^\varepsilon = u_t + \sqrt{\varepsilon} \zeta_t^\varepsilon, \quad (4.8)$$

where $\zeta_t^\varepsilon \Rightarrow \zeta_t$ (ζ_t the solution of (4.7)) as $\varepsilon \rightarrow 0$, or, sloppily,

$$x_t^\varepsilon \stackrel{\mathcal{D}}{\approx} u_t + \sqrt{\varepsilon} \zeta_t,$$

($\stackrel{\mathcal{D}}{\approx}$ meaning approximate equality of probability distributions), which is also known in physics as *Van Kampen's approximation*.

We are convinced that (4.6, 4.7), resp. (4.8), is the appropriate mathematically rigorous form of Hasselmann's stochastic model, formulated for the classical case and where the weather is a sufficiently mixing process.

4.2 Central Limit Theorem: Hasselmann's Case

We also believe that there are suitable versions of the Central Limit Theorem based on the Method of Averaging for the coupled system (3.6, 3.7). There are results by Kifer [18, 20, 21] for the case where g on the right-hand side of (3.7) is independent of x .

Here again, the weather dynamical system φ^x at frozen climate x solving (3.8) has to be sufficiently "mixing" to ensure that for each fixed x

$$\zeta_t^{\varepsilon, x}(y) := \frac{1}{\sqrt{\varepsilon}} (x_t^\varepsilon(x, y) - u_t(x)), \quad 0 \leq t \leq T,$$

converges weakly (based on the measure $\mu_x(dy)$ in y space \mathbb{R}^m chosen to perform the Method of Averaging) to the Ornstein-Uhlenbeck process ζ_t^x solving the stochastic differential equation (4.4), where $u_t(x)$ is the solution of the averaged equation $\dot{u} = F_{\mu_u}(u)$ based on the chosen reference measure μ_x .

If φ^x is nicely hyperbolic (chaotic) to ensure the existence of a unique SRB measure μ_x^{SRB} , then we expect situations in which μ_x^{SRB} can be replaced in the convergence statement by some (normalized) form of Lebesgue measure in y space.

Particular Case: The case

$$\dot{x}_t = f(y_t), \quad \dot{y}_t = \frac{1}{\varepsilon} g(y_t), \quad x_0 = x, \quad y_0 = y,$$

where $\dot{y}_t = g(y_t)$ generates a nice hyperbolic dynamical system $(t, y) \mapsto \varphi_t(y)$, has been investigated by many authors. For recent reviews see Denker [4] or Viana [34] and the references therein.

Assume that φ has a unique SRB measure μ . We average out y , i.e. consider

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(\varphi_t(y)) dt = \int f(y) \mu(dy) =: \mu(f)$$

for Lebesgue-almost all y , hence $\dot{u}_t = \mu(f)$ and $u_t(x) = \mu(f)t + x$.

Since

$$x_t^\varepsilon(x, y) = x + \int_0^t f(\varphi_{\frac{s}{\varepsilon}}(y)) ds,$$

we have (modulo conditions)

$$\begin{aligned} \frac{1}{\sqrt{\varepsilon}} (x_t^\varepsilon(x, y) - u_t(x)) &= \frac{1}{\sqrt{\varepsilon}} \int_0^t (f(\varphi_{\frac{s}{\varepsilon}}(y)) - \mu(f)) ds \\ &= \sqrt{\varepsilon} \int_0^{t/\varepsilon} (f(\varphi_s(y)) - \mu(f)) ds \\ &\Rightarrow \zeta_t, \end{aligned}$$

where $\zeta_t = \sqrt{\sigma} W_t$, since $DF \equiv 0$ here, and σ is a constant matrix defined as in (4.5).

5 Large Deviations from the Averaged System

5.1 Deficiencies of the Method of Averaging and the Central Limit Theorem. Large Deviations

Consider the coupled climate-weather system (3.6, 3.7).

While the weather system $\dot{y}_t = \frac{1}{\varepsilon} g(x, y_t)$ undergoes significant changes (for constant climate x) already in a time interval $[0, \varepsilon]$, the climate system $\dot{x}_t = f(x_t, y_t)$ needs a time interval $[0, T]$ to undergo observable deviations from the initial condition $x_0 = x$. How climate actually evolves in the time interval $[0, T]$ for generic weather conditions is determined by the Method of Averaging in the zero'th approximation, and by the Central Limit Theorem in the first (and last) approximation.

The Method of Averaging and the Central Limit Theorem become typically invalid for time intervals growing with $1/\varepsilon$, and can thus in principle not describe phenomena which happen at a time scale much longer than 1.

There are indeed many of such long-term phenomena that are of crucial importance for the climate and that (as it will turn out) evolve on an exponential time scale. Here are some examples.

1. Exit from Neighborhood of Attractor: Suppose x_0 is an asymptotically stable fixed point of the averaged system $\dot{u}_t = F(u_t)$, i.e. a locally attracting stable climatic state if the weather is averaged out. Then there exists a neighborhood D of x_0 which is positively invariant, i.e. the solution of $\dot{u}_t = F(u_t)$ with $u_0 = x \in D$ cannot leave D .

However, the non-averaged climate x_t^ε is typically able to leave D in finite time, and Large Deviations Theory determines the order of magnitude of the first exit time τ^ε (namely $\varepsilon^{1/\varepsilon}$) as well as the exit position.

2. “Hopping” Between Attractors: Suppose $\dot{u}_t = F(u_t)$ has several local attractors. Then x_t^ε can “hop” between them in the following sense: When x_t^ε exits a neighborhood of the first attractor after having spent an exponentially long time there, it can be attracted by another local attractor of the averaged system, stay exponentially long in a neighborhood of this new attractor, then leave this neighborhood etc. Large deviations theory can determine the dynamics of this “hopping” rather precisely, in particular the order in which the local attractors are visited.

This could describe the switching between different locally attracting climatic states described by SDM (in particular, by an EBM) if we take the forcing of the climate by the weather into account. We believe that the recent paper [14] by Hasselmann can be interpreted in this direction.

3. Invariant Measures: The invariant measures μ^ε of x_t^ε build-up as occupation measures of orbits only during very long times, as the value $\mu^\varepsilon(B)$ for some set $B \subset \mathbb{R}^d$ is the proportion of time spent by a typical trajectory of x_t^ε in the set B .

Hence invariant measures μ^ε cannot be found from the Method of Averaging or the Central Limit Theorem, nor can we determine to which one of the typically many invariant measures of $\dot{u}_t = F(u_t)$ the sequence μ^ε converges as $\varepsilon \rightarrow 0$. For example, often “non-uniqueness breaking” is observed, i.e. the sequence (μ^ε) singles-out a particular measure μ^0 of $\dot{u}_t = F(u_t)$ to which it converges.

It thus appears to us that the Theory of Large Deviations should be most interesting and useful to meteorologists.

5.2 Large Deviations: Classical Case

We begin by looking again at

$$\dot{x}_t = f(x_t, \xi_{t/\varepsilon}), \quad x_0 = x \in \mathbb{R}^d, \quad (5.1)$$

where ξ_t is a nice ergodic stationary stochastic process. We assume that the Method of Averaging applies, where the averaged equation is

$$\dot{u}_t = F(u_t), \quad u_0 = x, \quad (5.2)$$

with

$$F(x) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(x, \xi_t) dt = \int_{\mathbb{R}^m} f(x, y) \mu(dy). \quad (5.3)$$

It was Freidlin [8, 9] who first (in fact, in 1976, the same year Hasselmann wrote his paper) studied large deviations of $x_t^\varepsilon(x)$ from the averaged process $u_t(x)$, i.e. the type of problems we sketched in Subsection 5.1.

A canonical reference for the Theory of Large Deviations is the classical book by Freidlin and Wentzell [11], in particular Chap. 7 for Large Deviations Theory in the context of the Method of Averaging.

We just make some very brief motivational remarks here, e.g. on the study of the first exit time of the solution $x_t^\varepsilon(x)$ of (5.1) from a neighborhood D of an asymptotically stable steady state x_0 of the averaged equation (5.2).

To measure the cost for going from a point x in the interior of D to a point z on the boundary ∂D of D , a *quasipotential*

$$V(x, z) := \inf\{S_T(h) : h(0) = x, h(T) = z, T > 0 \text{ arbitrary}\} \quad (5.4)$$

is introduced, where $S_T(h) \geq 0$ is the *action functional* of the problem which assigns to any continuous function $h : [0, T] \rightarrow \mathbb{R}^d$ a non-negative number, and the infimum in (5.4) is taken over all functions which start at x and reach z in some arbitrary, but finite time $T > 0$. We have $S_T(h) = 0$ if and only if h is a solution of the averaged equation (5.2).

In the simplest case where

$$\dot{x}_t = f(x_t) + \sqrt{\varepsilon} \dot{W}_t$$

we have $\dot{u}_t = f(u_t)$ and the action functional is

$$S_T(h) = \int_0^T \left(\dot{h}(t) - f(h(t)) \right)^2 dt.$$

It turns out that $V(x, z) = V(x_0, z)$ in (5.4). The lowest cost of ever reaching the boundary ∂D from inside D is

$$V_0 := \inf\{V(x_0, z) : z \in \partial D\} > 0. \quad (5.5)$$

One can prove that the first exit time $\tau^\varepsilon(x)$ of $x_t^\varepsilon(x)$ from D satisfies for all x in the interior of D

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \log \tau^\varepsilon(x) = V_0 \quad \text{in probability,} \quad (5.6)$$

i.e.

$$\tau^\varepsilon(x) \approx e^{V_0/\varepsilon} \quad \text{with high probability.}$$

Further, $x_{\tau^\varepsilon(x)}^\varepsilon(x) = \bar{x}$ with high probability, where \bar{x} is the point at which the infimum in (5.5) is attained.

The action functional $S_T(h)$ is also the appropriate instrument for describing the time scale (again $e^{c/\varepsilon}$) of the “hopping” of x_t^ε between the various local attractors of $\dot{u}_t = F(u_t)$, see [11].

Freidlin [10] has recently shown that the so-called *stochastic resonance* is also a large deviations phenomenon. This is explained and made more precise in the contributions of Imkeller [15] to these Proceedings.

5.3 Large Deviations: Hasselmann’s Case

Unfortunately, there are no proven results yet for this case, but Remark 2.5 on page 348 of [20] gives some hope. There are, however, results by Kifer [18, 19, 20] for the case where (3.7) is decoupled from (3.6).

We speculate that there is a Theory of Large Deviations from the averaged motion also for Hasselmann’s case (3.6, 3.7), provided that the weather is sufficiently “chaotic”.

More specifically, let the dynamical system φ^x generated by (3.8) with frozen x stay in a compact part K of its state space (with normalized Lebesgue measure m on K) and be nicely hyperbolic so that it has a unique SRB measure $\mu_x^{\text{SRB}}(dy)$ for all x .

Assume that the Method of Averaging holds for $\dot{u}_t = F(u_t)$, $u_0 = x$, where

$$F(x) := \int_K f(x, y) \mu_x^{\text{SRB}}(dy).$$

Then we expect Large Deviations statements to hold, now based on the reference measure m .

For example, the first exit time $\tau^\varepsilon(x, y)$ of $x_t^\varepsilon(x, y)$ from a neighborhood D of an asymptotically stable steady state x_0 of the averaged motion is expected to satisfy

$$\int_K \tau^\varepsilon(x, y) m(dy) \sim e^{V_0/\varepsilon} \quad \text{as } \varepsilon \rightarrow 0$$

for all x in the interior of D . Furthermore, the “hopping” of $x_t^\varepsilon(x, y)$ between the various local attractors of $\dot{u}_t = F(u_t)$ on a time scale $\varepsilon^{1/\varepsilon}$ could then also be described.

If these Large Deviations results come into existence, they should be most appealing and interesting to meteorologists, as they would capture the crucial influence of the forcing of the weather on the long-term evolution of the climate.

6 Extensions of Hasselmann’s Program. Comments

6.1 Extensions

Hasselmann’s program, by which we understand the study of the coupled system (3.6, 3.7) by means of the Method of Averaging, the Central Limit Theorem and the Theory of Large Deviations, should be implemented in the first place, but should also be extended to

(i) the infinite-dimensional case, i.e. to coupled PDE’s, where f and g in (3.6, 3.7) are now partial differential operators containing in particular derivatives of x and y with respect to spatial variables (see the work of Watanabe [35])

for averaging in infinite dimensions, and of Da Prato and Zabczyk [3] for the general background),

(ii) other variants of fast/slow systems, e.g. to a situation described by Lions, Temam and Wang [23, 24, 25] (see also the contribution of Temam [33] to these Proceedings), namely an atmosphere-ocean model, where the (otherwise uncoupled) (fast) atmosphere is coupled to the (slow) ocean via interface conditions on the surface of the ocean. The models of Müller [27, 28] (see also his Proceedings contribution [26]) for the ocean circulation forced by fast random wind belong to this context.

6.2 Comments

(i) Hasselmann's nonlinear stochastic differential equation

$$dx_t = F(x_t) dt + \sqrt{\sigma(x_t)} dW_t,$$

where $F(x)$ is the averaged vector field and $\sigma(x)$ is the nonlinear average defined by (4.5) [13, Sect. 4]), still defies our attempts of interpretation. In what sense can it be considered as an approximation of x_t^e , or as an extension of $\dot{u}_t = F(u_t)$?

(ii) Another version of letting (continuous time) randomness emerge from deterministic (discrete time) chaotic dynamics which is closely related to what we call the Central Limit Theorem has been studied by Beck [1], Kantz and Olbrich [16] and others and is one of the main themes of the contribution of Rödenbeck, Beck and Kantz [31] to these Proceedings.

Acknowledgement: The author is grateful to Yuri Kifer and Peter Kloeden for useful comments.

References

- [1] C. Beck. Dynamical systems of Langevin type. *Physica A*, 233:419–440, 1996.
- [2] N. Bogolyubov and Y. Mitropolskii. *Asymptotic methods in the theory of nonlinear oscillations*. Gordon and Breach, New York, 1961.
- [3] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge University Press, 1992.
- [4] M. Denker. The central limit theorem for dynamical systems. In *Dynamical Systems and Ergodic Theory*, pages 33–62. Banach Center Publications, Volume 23, 1989.
- [5] Deutsches Klimarechenzentrum. The Hamburg large scale geostrophic ocean general circulation model (cycle 1). Technical report, No. 2, 1992.

- [6] Deutsches Klimarechenzentrum. The ECHAM 3 atmospheric general circulation model. Technical report, No. 6, 1994.
- [7] C. Frankignoul and K. Hasselmann. Stochastic climate models, part II. Application to sea-surface temperature anomalies and thermocline variability. *Tellus*, 29:289–305, 1977.
- [8] M. Freidlin. Fluctuations in dynamical systems with averaging. *Soviet Math. Dokl.*, 17:104–108, 1976.
- [9] M. Freidlin. The averaging principle and theorems on large deviations. *Russian Math. Surveys*, 33:117–176, 1978.
- [10] M. Freidlin. Quasi-deterministic approximation, metastability and stochastic resonance. *Physica D*, 2000.
- [11] M. I. Freidlin and A. Wentzell. *Random perturbations of dynamical systems*. Springer-Verlag, Berlin Heidelberg New York, 1984.
- [12] J. Freund, A. Neiman, and L. Schimansky-Geier. Stochastic resonance and noise-induced phase coherence. These Proceedings.
- [13] K. Hasselmann. Stochastic climate models: Part I. Theory. *Tellus*, 28:473–485, 1976.
- [14] K. Hasselmann. Climate change: linear and nonlinear signature. *NATURE*, 398:755–756, 1999.
- [15] P. Imkeller. Energy balance models – viewed from stochastic dynamics. These Proceedings.
- [16] H. Kantz and E. Olbrich. The transition from deterministic chaos to a stochastic process. *Physica A*, 253:105–117, 1998.
- [17] R. Z. Khasminskii. A limit theorem for solutions of differential equations with random right-hand side. *Theory of Probability and its Applications*, 11:390–406, 1966.
- [18] Y. Kifer. Averaging and climate models. These Proceedings.
- [19] Y. Kifer. Large deviations in dynamical systems and stochastic processes. *Trans. Amer. Math. Soc.*, 321:505–524, 1990.
- [20] Y. Kifer. Averaging in dynamical systems and large deviations. *Invent. math.*, 110:337–370, 1992.
- [21] Y. Kifer. Limit theorems in averaging for dynamical systems. *Ergodic Theory and Dynamical Systems*, 15:1143–1172, 1995.

- [22] P. Lemke. Stochastic climate models, part 3. Application to zonally averaged energy models. *Tellus*, 29:385–392, 1977.
- [23] J.-L. Lions, R. Temam, and S. Wang. Models for the coupled atmosphere and ocean (CAO I). Part I. *Computational Mechanics Advances*, 1:5–54, 1993.
- [24] J.-L. Lions, R. Temam, and S. Wang. Numerical analysis of the coupled atmosphere-ocean models (CAO II). Part II. *Computational Mechanics Advances*, 1:55–119, 1993.
- [25] J.-L. Lions, R. Temam, and S. Wang. Mathematical theory for the coupled atmosphere-ocean models (CAO III). *J. Math. Pure Appl.*, 74:105–163, 1995.
- [26] P. Müller. Stochastic forcing of geostrophic oceanic motion. These Proceedings.
- [27] P. Müller. Stochastic forcing of quasi-geostrophic eddies. In R. Adler, P. Müller, and B. Rozovskii, editors, *Stochastic modelling in physical oceanography*, pages 381–395. Birkhäuser, Boston Basel Berlin, 1996.
- [28] P. Müller. Stochastic forcing of oceanic motion. In S. Molchanov and W. Woyczynski, editors, *Stochastic models in geosystems*, pages 219–237. Springer-Verlag, Berlin Heidelberg New York, 1997.
- [29] D. Olbers. A gallery of simple models from climate physics. These Proceedings.
- [30] D. Olbers, P. Lemke, and D. Wolf-Gladrow. Physik des Klimasystems. Lecture Notes, Bremen University, 1997.
- [31] C. Rödenbeck, C. Beck, and H. Kantz. Dynamical systems with time scale separation: averaging, stochastic modelling, and central limit theorem. These Proceedings.
- [32] J. A. Sanders and F. Verhulst. *Averaging methods in nonlinear dynamical systems*. Springer-Verlag, Berlin Heidelberg New York, 1985.
- [33] R. Temam. Mathematical problems in climate modeling. These Proceedings.
- [34] M. Viana. *Stochastic dynamics of deterministic systems*. IMPA, Rio de Janeiro, 1997.
- [35] H. Watanabe. Averaging and fluctuations for parabolic equations with rapidly oscillating random coefficients. *Probab. Th. Rel. Fields*, 77:359–378, 1988.