Towards a definition of Climate and Weather

Franco Flandoli and Elisa Tonello

October 2, 2019

1 Introduction

In the study of climate we are faced with several very different time-scales (Figure 1). Some of the variabilities are directly associated to external inputs, others presumably have an internal, dynamical character. We focus on a subcase of this wide field: a case where, with good approximation, we may think to have two external time-varying inputs, one slowly varying, the other fluctuating at short time scale. An example could be the *dichotomy Climate-Weather* (with climate at a time-scale of decades): Climate is influenced by slowly varying inputs like annual solar radiation of human production of CO_2; but at a faster time-scale there are processes like the daily solar radiation with its random impact on the ground due to clouds, which may be modeled as random external inputs, fast varying, having certainly great influence on weather, and maybe having a less direct but still relevant influence on climate.

Systems with non autonomous inputs and different time scales can be found in many other fields. In Economy, financial assets fluctuate at short time scale but we are also interested in long-term average quantities like gdp. Examples in biology may found in [37]. In a sense, it is natural to introduce the dichotomy Climate-Weather also in other sciences.

Accurate geophysical models are of course very complicated. We make the following abstraction, that they are described by a stochastic non-autonomous differential equation

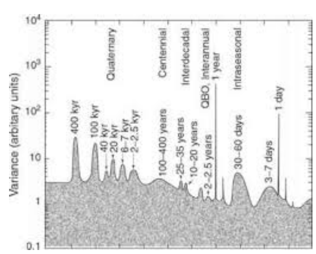


Figure 1: Spectrum of climate variability (from M. Ghil [28])

of the following form, in $\mathbb{R}^d$ (hence with a finite number of degrees of freedom):

$$\begin{aligned} dX_t &= b(q(t), X_t)\, dt + dW_t \qquad t \geq t_0 \\ X_{t_0} &= x_0 \end{aligned}$$

where $X_t \in \mathbb{R}^d$ is the state of the system, the collection of relevant variable of the model; W_t is a Brownian Motion (BM) in $\mathbb{R}^d$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ modeling the fast-varying external input; $q(t)$ is the slowly varying external input, a function

$$q : \mathbb{R} \to \mathbb{R}^k$$

and $b : \mathbb{R}^k \times \mathbb{R}^d \to \mathbb{R}^d$ is subject to assumptions that guarantee solvability in a suitable sense. Except for the necessary technical additional difficulties, what we are going to describe may be adapted to more general cases, like a noise term of the form

$$\sigma(q(t), X_t)\, dW_t$$

instead of dW_t alone, or the infinite dimensional case described by stochastic partial differential equations. The general framework we are going to describe accepts in principle these cases and others, although it may be difficult to check assumptions depending on the example.

Especially in climate studies, it is fundamental to consider the evolution of initial conditions starting from a certain time in the past, or more generally a time different from zero. This is why we have taken a generic t_0 as initial time for the Cauchy problem above.

In the sequel we thus have to introduce a concept of non autonomous dynamics depending on two external inputs q and ω: the *solution map*

$$U_{q,\omega}(t_0, t)$$

from time t_0 to time t. In Physics it is often called the *propagator*. Then we introduce associated relevant invariant objects (sets and measures) and *we give a definition of Climate and Weather*, analyzing some of its foundational properties, like existence and uniqueness. For the definition of Weather, see Section 3; for the definition of Climate, see Section 5. The notions we are going to introduce are related to the concept of pull-back attractor introduced in [17], [16], [52] and developed and applied to several examples by many authors; notions that have found particular interest in Climate Sciences. see for instance [5], [11], [12], [29], [30], [40], [43], [49], [50] and many others. Although depending on the time-scale and simplifications of the models it is possible to associate the Climate to different concepts in this theory, we try here to give a unified and critical view when both weather and climate coexist, hence two different time-scales are taken into account. A summary of the ideas is given in the final section.

2 Non autonomous dynamical system

By *Non Autonomous Dynamical System (NADS) with two external inputs* we mean the following structure:

1. a set $\mathcal{Q}$
2. a probability space $(\Omega, \mathcal{F}, \mathbb{P})$
3. a metric space (X, d)
4. a family of continuous maps $U_{q,\omega}(t_0, t) : X \to X$, indexed by $t_0 \leq t \in \mathbb{R}$ and $(q, \omega) \in \mathcal{Q} \times \Omega$, such that for all (q, ω) we have

$$U_{q,\omega}(t_0, t_0) = Id \qquad \text{for all } t_0 \in \mathbb{R}$$
$$U_{q,\omega}(t_0, t) = U_{q,\omega}(s, t) \circ U_{q,\omega}(t_0, s) \qquad \text{for all } t_0 \leq s \leq t \in \mathbb{R}$$

   and such that for every $q \in \mathcal{Q}$ and $x \in X$ the map

$$(\omega, t_0, t) \mapsto U_{q,\omega}(t_0, t)(x)$$

   is measurable (with respect to the Borel σ-algebra on the codomain and the product σ-algebra between $\mathcal{F}$ and Borel σ-algebra on $\mathbb{R} \times \mathbb{R}$ on the domain).

If we want to summarize the notations, we shall speak of a NADS

$$(\mathcal{Q}, (\Omega, \mathcal{F}, \mathbb{P}), (X, d), U_{q,\omega}(t_0, t))$$

with implicit understanding of the range of variations of the parameters in the maps $U_{q,\omega}(t_0, t)$.

Sometimes it is possible to require less from the probabilistic side, like that for every $q \in \mathcal{Q}$ and $t_0 \leq s \leq t \in \mathbb{R}$ one has

$$\mathbb{P}\left(U_{q,\cdot}(t_0, t) = U_{q,\cdot}(s, t) \circ U_{q,\cdot}(t_0, s)\right) = 1$$

with the possibility that the exceptional zero measure depends on q, t_0, s, t. But we leave such generalizations for specialized literature. Here we assume uniformity in all parameters.

The interpretation is:

$$q = q(t) = \text{slow external input}$$
$$\omega = \omega(t) = \text{fast external input}$$

$$\mathcal{Q} = \text{space of slow external inputs } q(\cdot)$$
$$\Omega = \text{space of fast external inputs } \omega(\cdot)$$

$$X = \text{space of configurations (states)}$$

$U_{q,\omega}(t_0, t)$ map that associates the state at time t to the state at time t_0 corresponding to a given pair of inputs (q, ω).

Compared to the SDE example above, if the SDE is uniquely globally solvable and we call X_t^{q,ω,t_0,x_0} its solution, we set (see also Proposition 3 below)

$$U_{q,\omega}(t_0, t)(x_0) = X_t^{q,\omega,t_0,x_0}.$$

In the case of additive noise and suitable assumptions on the drift b, it is a relatively simple exercise to check that the family $U_{q,\omega}(t_0, t)$, so defined, satisfies the required properties. For more general models it may be very technical (requiring the theory of stochastic flows and suitable perfection properties) or even an open question (for SPDEs).

2.1 Cocycle

A more advanced structure, that we call a cocycle, is made of:

1. a set $\mathcal{Q}$ and a family of transformations $\vartheta_t : \mathcal{Q} \to \mathcal{Q}$, $t \in \mathbb{R}$, satisfying (group property)
$$\vartheta_0 = Id, \qquad \vartheta_{t+s} = \vartheta_t \circ \vartheta_s, \qquad s, t \in \mathbb{R}$$

2. a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a family of $\mathcal{F}$-measurable transformations $\theta_t : \Omega \to \Omega$, $t \in \mathbb{R}$, satisfying (group property)
$$\theta_0 = Id, \qquad \theta_{t+s} = \theta_t \circ \theta_s, \qquad s, t \in \mathbb{R}$$
and stationarity (or invariance of $\mathbb{P}$):
$$\theta_t \mathbb{P} = \mathbb{P}.$$

3. a metric space (X, d)

4. a family of continuous maps $\varphi_t(q, \omega) : X \to X$, indexed by $t \geq 0$ and $(q, \omega) \in \mathcal{Q} \times \Omega$, such that for all (q, ω) we have
$$\begin{aligned} \varphi_0(q, \omega) &= Id \\ \varphi_{t+s}(q, \omega) &= \varphi_s(\Phi_t(q, \omega)) \circ \varphi_t(q, \omega) \qquad \text{for all } t, s \geq 0 \end{aligned}$$
where $\Phi_t : \mathcal{Q} \times \Omega \to \mathcal{Q} \times \Omega$ is defined as
$$\Phi_t(q, \omega) = (\vartheta_t q, \theta_t \omega)$$
and such that for every $q \in \mathcal{Q}$ and $x \in X$ the map
$$(\omega, t) \mapsto \varphi_t(q, \omega)\, x$$
is measurable.

Here and below, when we have a map $T : (F, \mathcal{F}) \to (G, \mathcal{G})$ between two measurable spaces and a probability measure μ on $(F, \mathcal{F})$, we write $T\mu$ for the image probability measure on $(G, \mathcal{G})$, defined by

$$(T\mu)(B) = \mu\left(T^{-1}(B)\right)$$

for all $B \in \mathcal{G}$. It is also called the push-forward, usually denoted by $T_\sharp\mu$ (but we simplify this notation).

If we want to summarize the notations, we shall speak of a cocycle

$$(\mathcal{Q}, \vartheta_t, (\Omega, \mathcal{F}, \mathbb{P}), \theta_t, (X, d), \varphi_t(q, \omega))$$

with implicit understanding of the range of variations of the parameters in the maps ϑ_t, θ_t, $U_{q,\omega}(t_0, t)$.

Compared to the SDE example above, if the SDE is uniquely globally solvable and we call X_t^{q,ω,t_0,x_0} its solution, we set (as in Proposition 3 below)

$$\varphi_t(q, \omega) x_0 = X_t^{q,\omega,0,x_0}.$$

In other words, now we solve the SDE only from time $t_0 = 0$; but we need two shifts on (q, ω) to describe the evolution property. The shift ϑ_t is simply defined as

$$(\vartheta_t q)(s) = q(t + s)$$

but the shift θ_t, on the so called two-sided Wiener space, is defined as

$$(\theta_t \omega)(s) = \omega(t + s) - \omega(t)$$

corresponding to the fact that we have to shift the time derivative of ω. More details later. See [3] for an extended exposition of cocycles and several foundational results on them.

2.2 Cocycle implies NADS

If we have a NADS $U_{q,\omega}(t_0, t)$, the cocycle is obviously defined as

$$\varphi_t(q, \omega) = U_{q,\omega}(0, t). \tag{1}$$

However, nothing in the abstract definition of NADS allows one to introduce the groups ϑ_t and θ_t and thus we cannot reconstruct, at an abstract level, a cocycle from a NADS (in SDE examples we construct directly the shift and the cocycle, without passing through a NADS). On the constrary, if we have a cocycle $\varphi_t(q, \omega)$, we define

$$U_{q,\omega}(s, t) := \varphi_{t-s}(\Phi_s(q, \omega)). \tag{2}$$

Let us formalize this fact (that cocycle implies NADS, while the converse requires more).

Proposition 1 *Given a cocycle* $(\mathcal{Q}, \vartheta_t, (\Omega, \mathcal{F}, \mathbb{P}), \theta_t, (X, d), \varphi_t(q, \omega))$*, the family of maps*

$$U_{q,\omega}(s,t) : X \to X \qquad s \leq t$$

defined by (2) is a NADS.

Proof. Clearly continuity holds as well as $U_{q,\omega}(s,s) = Id$. For $s \leq r \leq t$ we have

$$\begin{aligned} U_{q,\omega}(r,t) \circ U_{q,\omega}(s,r) &= \varphi_{t-r}(\Phi_r(q,\omega)) \circ \varphi_{r-s}(\Phi_s(q,\omega)) \\ &= \varphi_{t-r}(\Phi_{r-s}\Phi_s(q,\omega)) \circ \varphi_{r-s}(\Phi_s(q,\omega)) \\ &= \varphi_{t-s}(\Phi_s(q,\omega)). \end{aligned}$$

■

Proposition 2 *Let* $U_{q,\omega}(s,t)$ *be a NADS. Assume* (q,ω) *belong to a structure* $(\mathcal{Q}, \vartheta_t, (\Omega, \mathcal{F}, \mathbb{P}), \theta_t)$*. Assume*

$$U_{\Phi_t(q,\omega)}(0,s) = U_{q,\omega}(t, t+s).$$

Then equation (1) defines a cocycle. The condition is also necessary; more generally it holds

$$U_{\Phi_r(q,\omega)}(s,t) = U_{q,\omega}(s+r, t+r). \tag{3}$$

Proof. It follows from the identities

$$\begin{aligned} \varphi_s(\Phi_t(q,\omega)) \circ \varphi_t(q,\omega) &= U_{\Phi_t(q,\omega)}(0,s) \circ U_{q,\omega}(0,t) \\ \varphi_{t+s}(q,\omega) &= U_{q,\omega}(0, t+s). \end{aligned}$$

Formula (3) follows from

$$U_{\Phi_r(q,\omega)}(s,t) = \varphi_{t-s}(\Phi_s\Phi_r(q,\omega)) = \varphi_{t+r-(s+r)}(\Phi_{s+r}(q,\omega)) = U_{q,\omega}(s+r, t+r).$$

■

2.3 Examples

Let $C_0(\mathbb{R}; \mathbb{R}^d)$ be the space of continuous functions null at $t = 0$. Let us define the two-sided Wiener measure. Take, on some probability space, two independent copies of the BM W, say $W_t^{(i)}$, $i = 1, 2$; define the two-sided BM:

$$W_t = W_t^{(1)} \text{ for } t \geq 0,\ W_t = W_{-t}^{(2)} \text{ for } t \leq 0;$$

and call $\mathbb{P}$ its law, on Borel sets of $C_0(\mathbb{R}; \mathbb{R}^d)$; this is the two-sided Wiener measure. On $C_0(\mathbb{R}; \mathbb{R}^d)$ consider the "shift"

$$\theta_t(\omega) = \omega(t + \cdot) - \omega(t)$$

for every $\omega \in C_0\left(\mathbb{R};\mathbb{R}^d\right)$. By known properties of Brownian motion it follows $(\theta_t)_\sharp \mathbb{P} = \mathbb{P}$. Indeed, if $(X_t)_{t\in\mathbb{R}}$ is a two-sided BM with law $\mathbb{P}$, we have

$$\theta_t\left(X_\cdot\right)(r) = X_{t+r} - X_t$$

which is a new two-sided Brownian motion.

Proposition 3 *Let $b : \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^d$ continuous and satisfying*

$$|b(a,x) - b(a,y)| \le L\,|x-y| \qquad \text{for all } a \in \mathbb{R}^m \text{ and } x,y \in \mathbb{R}^d.$$

Let $q \in C\left(\mathbb{R};\mathbb{R}^m\right)$. Consider the non-autonomous stochastic equation (SDE)

$$dX_t = b\left(q(t), X(t)\right)dt + dW_t$$

where W_t is a two-sided Brownian motion in $\mathbb{R}^d$. It generates a cocycle and thus a NADS.

If we solve the SDE from time $t_0 = 0$ and initial condition x_0 and call X^{q,ω,x_0} its solution, we have $\mathbb{P}$-a.s.

$$\varphi_t(q,\omega)(x_0) = X^{q,\omega,x_0}(t) \qquad \text{for all } t \ge 0.$$

Similarly, if we solve the SDE from a generic time t_0 and initial condition x_0 and call X^{q,ω,t_0,x_0} its solution, we have $\mathbb{P}$-a.s.

$$U_{q,\omega}(t_0,t)(x_0) = X^{q,\omega,t_0,x_0}(t) \qquad \text{for all } t \ge t_0.$$

Proof. This equation can be easily solved by usual stochastic methods. For the purpose of introducing the associated cocycle we interpret this equation *pathwise*: for every $\omega \in C_0\left(\mathbb{R};\mathbb{R}^d\right)$ we consider the integral equation

$$x(t) = x_0 + \int_0^t b\left(q(s), x(s)\right)ds + \omega(t).$$

By classical contraction principle, iterated on finite time intervals (where all coefficients satisfy uniformly the necessry continuity and Lipschitz properties), one can easily prove it has a unique solution

$$x^{q,\omega,x_0} \in C\left(\mathbb{R};\mathbb{R}^d\right).$$

Then we set

$$\mathcal{Q} = C\left(\mathbb{R};\mathbb{R}^m\right) \qquad \vartheta_t q = q(t+\cdot)$$

$$\Omega = C_0\left(\mathbb{R};\mathbb{R}^d\right) \qquad \theta_t\omega = \omega(t+\cdot) - \omega(t)$$

$\mathcal{F}$ = Borel σ field of Ω, $\mathbb{P}$ = two-sided Wiener measure;

$$\varphi_t(q,\omega)(x_0) = x^{q,\omega,x_0}(t).$$

In other words, with the notation $\gamma = (q, \omega)$,

$$\varphi_t(\gamma)(x_0) = x_0 + \int_0^t b(q(s), \varphi_s(\gamma)(x_0))\, ds + \omega(t).$$

In order to say that this example satisfies the abstract properties of a cocycle we have to check a few conditions. Group properties of shifts are obvious; property $(\theta_t)_\sharp \mathbb{P} = \mathbb{P}$ was already discussed above. We have to prove the cocycle property

$$\varphi_{t+s}(\gamma) = \varphi_s(\Theta_t \gamma) \circ \varphi_t(\gamma).$$

We have

$$\begin{aligned}\varphi_{t+s}(\gamma)(x_0) &= x_0 + \int_0^{t+s} b(q(r), \varphi_r(\gamma)(x_0))\, dr + \omega(t+s) \\ &= x_0 + \int_0^t b(q(r), \varphi_r(\gamma)(x_0))\, dr + \omega(t) \\ &+ \int_t^{t+s} b(q(r), \varphi_r(\gamma)(x_0))\, dr + \omega(t+s) - \omega(t)\end{aligned}$$

$$\begin{aligned}&= \varphi_t(\gamma) + \int_t^{t+s} b(q(r), \varphi_r(\gamma)(x_0))\, dr + (\theta_t \omega)(s) \\ &\overset{r=t+u}{=} \varphi_t(\gamma) + \int_0^s b((\vartheta_t q)(u), \varphi_{t+u}(\gamma)(x_0))\, du + (\theta_t \omega)(s)\end{aligned}$$

hence we see that

$$z(s) := \varphi_{t+s}(\gamma)(x_0) \qquad s \geq 0$$

satisfies

$$z(s) = z_0 + \int_0^s b(q^*(u), z(u))\, du + \omega^*(s)$$

where $z_0 = \varphi_t(\gamma)$, $q^*(u) = (\vartheta_t q)(u)$, $\omega^*(s) = (\theta_t \omega)(s)$. By uniqueness of solutions to this equation,

$$z(s) = \varphi_s(q^*, \omega^*)(z_0).$$

Collecting all identities and definitions,

$$\begin{aligned}\varphi_{t+s}(\gamma)(x_0) &= \varphi_s(q^*, \omega^*)(z_0) \\ &= \varphi_s(\vartheta_t q, \theta_t \omega)(\varphi_t(\gamma))\end{aligned}$$

which means precisely $\varphi_{t+s}(\gamma) = \varphi_s(\Theta_t \gamma) \circ \varphi_t(\gamma)$. $\blacksquare$

Remark 4 *At the structural level (group and cocycle properties) there is nothing special in the additive noise. What is special is the possibility to solve, uniquely, the equation for all elements* $\omega \in C\left(\mathbb{R};\mathbb{R}^d\right)$ *of the Wiener space, namely the pathwise solvability. Similar tricks hold only for few equations, like*

$$dX_t = b\left(q\left(t\right), X\left(t\right)\right)dt + X_t dW_t.$$

However, in general there are two possibilities. One is the theory of rough paths, which at the price of some additional regularity of coefficients and a certain high level background in stochastic analysis, enables to solve "any" stochastic equation pathwise. One has to replace the Wiener space $C\left(\mathbb{R};\mathbb{R}^d\right)$ *by the more structured space of (geometric) rough paths. This non trivial theory has the advantage to be fully pathwise and thus entirely analog to the previous additive noise example. The other possibility is to use the theory of stochastic flows. That theory claims that, given the stochastic-process solutions* $X_t^{x_0}(\omega)$*, which at time t are equivalence classes (hence not pointwise uniquely defined in* ω*), there is a version of the family such that we can talk of the space-time trajectories*

$$(t, x_0) \to X_t^{x_0}(\omega)$$

and, even more, we can do the same starting from an arbitrary time t_0*. Elaborating this concept with non trivial further elements, one can construct a naRDS (the technical difficulty is only in the stochastic part, not the non-autonomous one). See [4] for a relevant theorem on this topic.*

The second example we consider is a finite dimensional model inspired to the Navier-Stokes equations. We consider the stochastic equation in $\mathbb{R}^d$

$$dX_t + (AX_t + B(X_t, X_t))dt = q\left(t\right)dt + \sqrt{Q}dW_t, \qquad t \geq t_0 \tag{4}$$

where A is an $d \times d$ symmetric matrix satisfyng $\langle Ax, x\rangle \geq \lambda |x|^2$ for each $x \in \mathbb{R}^d$, B is a bilinear continuous mapping with the property

$$\langle B(y, x), x\rangle = 0 \qquad \forall x, y \in \mathbb{R}^d \tag{5}$$

Q is a positive semidefinite symmetric matrix, q is a continuous function and W_t is a two-sided Brownian motion. Except for the minor generalization due to $\sqrt{Q}$, the main point is the fact that B is not globally Lipschitz continuous, only locally Lipschitz, whch a priori may cause problems of global existence. Thus let us discuss some details. As in the previous example (with $m = d$) we take $\mathcal{Q} = C\left(\mathbb{R};\mathbb{R}^d\right)$, $\vartheta_t q = q\left(t + \cdot\right)$, $\Omega = C_0\left(\mathbb{R};\mathbb{R}^d\right)$, $\theta_t \omega = \omega\left(t + \cdot\right) - \omega\left(t\right)$, $\mathcal{F} =$ Borel σ field of Ω, $\mathbb{P} =$ two-sided Wiener measure. Given $q \in \mathcal{Q}$ and $\omega \in \Omega$ we may consider the equation on a time interval $[t_0, \infty)$ pathwise

$$x_t = x_0 - \int_{t_0}^{t} (Ax_s + B(x_s, x_s) - q\left(s\right))ds + \sqrt{Q}\omega_t - \sqrt{Q}\omega_{t_0}.$$

By classical contraction principle arguments, since B is locally Lipschitz continuous, we have local-in-time existence, uniqueness and continuous dependence on the initial condition x_0. If we prove an a priori bound, the solution is global (and unique). Since we work pathwise, we cannot use stochastic calculus to prove the a priori bound. Thus we introduce the function z_t, globally defined and continuous, satisfying

$$z_t = -\int_{t_0}^{t} A z_s ds + \sqrt{Q}\omega_t - \sqrt{Q}\omega_{t_0}$$

(it is given by the previous proposition, for instance). Set $v_t = x_t - z_t$. We have

$$v_t = x_0 - \int_{t_0}^{t} (Av_s + B(v_s + z_s, v_s + z_s) - q(s))ds.$$

Hence v_t is differentiable and

$$\frac{d}{dt} v_t + Av_t + B(v_t + z_t, v_t + z_t) = q(t).$$

We deduce

$$\frac{d}{dt} |v_t|^2 + 2\langle Av_t, v_t\rangle + 2\langle B(v_t + z_t, v_t + z_t), v_t\rangle = 2\langle q(t), v_t\rangle$$

where $|\cdot|$ is the Euclidean norm in $\mathbb{R}^d$ and $\langle\cdot,\cdot\rangle$ the scalar product. Using the assumptions on A and B we get

$$\frac{d}{dt} |v_t|^2 + 2\lambda |v_t|^2 \leq 2\langle q(t), v_t\rangle - 2\langle B(v_t + z_t, z_t), v_t\rangle.$$

We have

$$2\langle q(t), v_t\rangle \leq |q(t)|^2 + |v_t|^2$$

$$\begin{aligned} -2\langle B(v_t + z_t, z_t), v_t\rangle &\leq C|v_t|^2 |z_t| + C|v_t| |z_t|^2 \\ &\leq C|v_t|^2 (|z_t| + 1) + C|z_t|^4 \end{aligned}$$

and thus, summarizing,

$$\frac{d}{dt} |v_t|^2 \leq |v_t|^2 (C|z_t| + C + 1) + |q(t)|^2 + C|z_t|^4$$

which implies, by Gronwall lemma, an a priori bound on any finite time interval. The a priori bound allows to repeat the application of contraction arguments on any finite time interval in a finite number of steps, getting also global continuous dependence on initial conditions as a consequence.

Whence global existence, uniqueness and continuous dependence is assured, the cocicle property follow as in the proof of the previous proposition.

2.4 White noise NADS

The previous example of SDE gives rise to what we call a *white noise* NADS. There is an abstract definition for this concept.

Definition 5 *A white noise NADS is made of*

1. a NADS $\left(\mathcal{Q}, (\Omega, \mathcal{F}, \mathbb{P}), (X, d), U_{q,\omega}(t_0, t)\right)$
2. a family of σ-algebras $\mathcal{F}_{t_0,t}$ indexed by $t_0 \leq t \in \mathbb{R}$ with the property that for every $q \in \mathcal{Q}$ and $x_0 \in X$

$$\omega \mapsto U_{q,\omega}(t_0, t)(x_0) \text{ is } \mathcal{F}_{t_0,t}\text{-measurable}$$

$$\omega \mapsto U_{q,\omega}(t_0, t)(x_0) \text{ is } \mathcal{F}_{s,u}\text{-independent}$$
$$\text{for every } s \leq u \text{ in } (-\infty, t_0] \text{ or in } [t, +\infty).$$

The interpretation of $\mathcal{F}_{t_0,t}$ is of the σ-algebra of events associated to a white noise on the time interval $[t_0, t]$. White noise is independent on disjoint intervals. The interpretation of (1-2) is that the solution map on $[t_0, t]$ depends only on the white noise on that interval and it is independent on the noise on other disjoint intervals. All these properties are satisfied in the example of the SDE above, with

$$\mathcal{F}_{t_0,t} = \sigma\left\{W_s - W_u; u \leq s \in [t_0, t]\right\}.$$

3 The Weather

The weather is what we observe, at the short time scale. Let us divide the concept in two definitions: one is what we could observe, the set of all possibilities; the other is what we statistically observe. For shortness we could call "potential Weather" what we could observe, "Weather" what we observe statistically.

Claim 6 *The "potential Weather" is the minimal compact global attractor $A_{q,\omega}(t)$.*

Claim 7 *The "Weather" is the statistical equilibrium $\mu_{q,\omega}(t)$.*

Let us define the two rigorous concepts of minimal compact global attractor and statistical equilibrium.

3.1 Minimal compact global attractor

In this section we drop the notation (q,ω). The theory is entirely deterministic and applies to each value of the pair (q,ω). At the end, one has to check measurability in ω in case probabilistic arguments are required in the sequel. The results are taken from [17] and [16].

Definition 8 *A family of sets $\{A(t) \subset X, t \in \mathbb{R}\}$ is called a compact invariant set for $U(t_0,t)$ if $A(t)$ is compact for every t and*

$$U(t_0,t)\,A(t_0) = A(t) \qquad \text{for all } t_0 \leq t.$$

A family of Borel probability measures $\{\mu(t), t \in \mathbb{R}\}$ is invariant for $U(t_0,t)$ if

$$U(t_0,t)\,\mu(t_0) = \mu(t) \qquad \text{for all } t_0 \leq t.$$

These invariance properties are very poor compared to the analogous ones of the deterministic case. One way to reinforce them is to ask a variant of them for the cocycle. We shall develop this argument later.

Definition 9 *We say that $A(t)$ attracts B at time t if for every $\epsilon > 0$ there exists $s_0 < 0$ such that for all $s < s_0$ we have*

$$U(s,t)(B) \subset \mathcal{U}_\epsilon(A(t)).$$

We say that a family of sets $\{A(t), t \in \mathbb{R}\}$ attracts B if the set $A(t)$ attracts B at time t, for every $t \in \mathbb{R}$.

This definition can be formulated by means of the non-symmetric distance between sets. Given $A, B \subset X$ define

$$d(B,A) = \sup_{x\in B} d(x,A)$$
$$\text{where } d(x,A) = \inf_{y\in A} d(x,y).$$

Then $A(t)$ attracts B if

$$\lim_{s\to-\infty} d(U(s,t)(B), A(t)) = 0.$$

Definition 10 *We call pull-back omega-limit set of B at time t the set*

$$\begin{aligned}\Omega(B,t) &= \overline{\bigcap_{s_0\leq 0}\bigcup_{s\leq s_0} U(s,t)(B)} \\ &= \{y \in X : \exists x_n \subset B, s_n \to -\infty, U(s_n,t)(x_n) \to y\}.\end{aligned}$$

Notice that obviously it can be an empty set.

Proposition 11 *Assume $A(t)$ is compact. Then $A(t)$ attracts B if and only if $\Omega(B,t) \subset A(t)$.*

Proof. Let us prove that if $A(t)$ attracts B then $\Omega(B,t) \subset A(t)$. Take $y \in \Omega(B,t)$ and $x_n \in B$, $s_n \to -\infty$ such that

$$U(s_n,t)(x_n) \to y.$$

We have $U(s_n,t)(x_n) \in \mathcal{U}_\epsilon(A(t))$ eventually, hence y is in the closure of $\mathcal{U}_\epsilon(A(t))$. By arbitrariety of ϵ, $y \in A(t)$.

Let us prove the converse statement by contradiction. Assuming that $A(t)$ does not attract B means that there exists $\epsilon > 0$ such that, for every $s_0 < 0$ there exists $s < s_0$ and $x \in B$ such that $U(s,t)(x) \notin \mathcal{U}_\epsilon(A(t))$. We can thus construct a sequence with this property so that a point $y \in \Omega(B,t)$ does not belong to $\mathcal{U}_\epsilon(A(t))$. This contradicts the assumption. ■

Below we give the definition of compact absorbing family; we anticipate here for compactness of exposition a criterium partially based on such a concept.

Proposition 12 *In general,*

$$U(s,t)\,\Omega(B,s) \subset \Omega(B,t).$$

If there is a compact absorbing family, then

$$U(s,t)\,\Omega(B,s) = \Omega(B,t).$$

Proof. Take $y \in \Omega(B,t)$ and $x_n \in B$, $s_n \to -\infty$ such that

$$U(s_n,s)(x_n) \to y.$$

Then

$$U(s_n,t)(x_n) = U(s,t)\,U(s_n,s)(x_n) \to U(s,t)\,y$$

namely $U(s,t)\,y \in \Omega(B,t)$, i.e. $U(s,t)\,\Omega(B,s) \subset \Omega(B,t)$. Conversely, take $z \in \Omega(B,t)$ and $x_n \in B$, $s_n \to -\infty$ such that

$$U(s_n,t)(x_n) \to z.$$

Then

$$U(s,t)\,U(s_n,s)(x_n) \to z.$$

The existence of a compact absorbing set implies that $U(s_n,s)(x_n)$ is included, eventually, in a compact set, hence there is a convergent subsequence $U(s_{n_k},s)(x_{n_k}) \to y$, hence $y \in \Omega(B,s)$ and $z = U(s,t)\,y$, therefore $\Omega(B,t) \subset U(s,t)\,\Omega(B,s)$. ■

Definition 13 *Given the NADS*

$$U(s,t) : X \to X \qquad s \leq t, s, t \in \mathbb{R}$$

we say that a family if sets

$$A(t) \qquad t \in \mathbb{R}$$

is a pull-back global compact attractor (PBCGA) if:

i) $A(t)$ is compact for every $t \in \mathbb{R}$

ii) $A(\cdot)$ is invariant: $U(s,t)(A(s)) = A(t)$ for every $s \leq t, s, t \in \mathbb{R}$

iii) $A(\cdot)$ pull-back attracts bounded sets:

$$\Omega(B,t) \subset A(t)$$

for all bounded set $B \subset X$.

Definition 14 *A family of sets $D(t)$, $t \in \mathbb{R}$ is called a bounded (resp. compact) pull-back absorbing family if:*

i) $D(t)$ is bounded (resp. compact) for every $t \in \mathbb{R}$

ii) for every $t \in \mathbb{R}$ and every bounded set $B \subset X$ there exists $t_B < t$ such that

$$U(s,t)(B) \subset D(t) \qquad \text{for every } s < t_B.$$

Theorem 15 *Let $U(s,t)$ be a continuous NADS. Assume that there exists a compact pull-back absorbing family. Then a PBCGA exists.*

In infinite dimensional examples of parabolic type the existence of a compact absorbing family is usually proved by means of the following lemma; hence what we usually apply in those examples is the next corollary.

Lemma 16 *Let $U(s,t)$ be a continuous NADS. Assume that:*

i) (compact NADS) for every t and bounded set $B \subset X$, the set $\overline{U(t,t+1)(B)}$ is compact

ii) there exists a pull-back absorbing family.

Then there exists a compact pull-back absorbing family.

Proof. Call $D(t)$ the (bounded) absorbing family. Set $D'(t) = \overline{U(t-1,t)(D(t-1))}$. It is easy to check that this is a compact absorbing family. ■

Remark 17 *In the compactness assumption for $U(t,t+1)$ the time lag 1 is obviously arbitrary. Over this remark and the peculiarities of certain examples, the literature developed a more general criterium based on asymptotic compactness.*

Corollary 18 *Let $U(s,t)$ be a continuous NADS. Assume that:*

i) (compact NADS) for every t and bounded set $B \subset X$, the set $\overline{U(t,t+1)(B)}$ is compact

ii) there exists a pull-back absorbing family.

Then a PBCGA exists.

Proof. (Proof of Theorem 15) Set

$$A(t) = \overline{\bigcup_{B \text{ bounded}} \Omega(B,t)}.$$

Let us prove it fulfills all properties of a PBGCA. By definition, $\Omega(B,t) \subset A(t)$, hence we have pull-back attraction.

From property (ii) of absorbing set,

$$\Omega(B,t) \subset D(t)$$

for every t and every bounded set B. In particular,

$$\overline{\bigcup_{B \text{ bounded}} \Omega(B,t)} \subset \overline{D(t)} = D(t).$$

Hence $A(t) \subset D(t)$ namely it is compact (being closed subset of a compact set).

Let us prove forward invariance. It will be an easy consequence of the forward invariance of omega-limit sets: $U(s,t)\,\Omega(B,s) \subset \Omega(B,t)$. It implies

$$\overline{\bigcup_{B \text{ bounded}} U(s,t)\,\Omega(B,s)} \subset \overline{\bigcup_{B \text{ bounded}} \Omega(B,t)} = A(t).$$

We always have the set-theoretical property

$$U(s,t)\left(\bigcup_{B \text{ bounded}} \Omega(B,s)\right) \subset \bigcup_{B \text{ bounded}} U(s,t)\,\Omega(B,s)$$

hence we have proved that

$$\overline{U(s,t)\left(\bigcup_{B \text{ bounded}} \Omega(B,s)\right)} \subset A(t).$$

But, by continuity of $U(s,t)$ and definition of $A(s)$ we have

$$U(s,t)\,A((s)) = U(s,t)\,\overline{\left(\bigcup_{B \text{ bounded}} \Omega(B,s)\right)} \subset \overline{U(s,t)\left(\bigcup_{B \text{ bounded}} \Omega(B,s)\right)}$$

which implies

$$U(s,t)\,A((s)) \subset A(t).$$

The opposite inclusion

$$A(t) \subset U(s,t)\,A((s))$$

is the most tricky step of the proof. Let $z \in A(t)$. There exists a sequence of bounded sets B_n and points $z_n \in \Omega(B_n,t)$ such that

$$z = \lim_{n\to\infty} z_n.$$

Each z_n is equal to

$$z_n = \lim_{k\to\infty} U(s_k^n,t)(x_k^n)$$

where $\lim_{k\to\infty} s_k^n = -\infty$ and $(x_k^n)_{k\in\mathbb{N}} \subset B_n$. Hence

$$\begin{aligned} z_n &= \lim_{k\to\infty} U(s,t)\left(U(s_k^n,s)(x_k^n)\right) = \lim_{k\to\infty} U(s,t)(y_k^n) \\ y_k^n &:= U(s_k^n,s)(x_k^n). \end{aligned}$$

The existence of a compact absorbing set $D(s)$ implies that there exists a subsequence $\left(y_{k_m^n}^n\right)_{m\in\mathbb{N}}$ with a limit

$$\begin{aligned} y^n &= \lim_{m\to\infty} y_{k_m^n}^n \\ y^n &\in D(s). \end{aligned}$$

Then, by continuity of $U(s,t)$,

$$z_n = U(s,t)\,y^n.$$

Again by compactness of $D(s)$ there is a subsequence $(y^{n_j})_{j\in\mathbb{N}}$ with a limit $y \in D(s)$. By continuity of $U(s,t)$,

$$z = U(s,t)\,y.$$

It remains to check that $y \in A(s)$. Due to the closure in the definition of $A(s)$, it is sufficient to prove that $y^{n_j} \in A(s)$. Looking at the definition of $y_{k_m^n}^n$, we see that its limit (in m) y^n is in $\Omega(B_n,s)$, hence in $A(s)$. ■

The attraction property has a strong power of identification, with respect to the poor property of invariance. However, given a global attractor $A(t)$ we may always add to it trajectories $\{x(t)\}$ and still have all properties, because compactness and invariance are satisfied and attraction continue to hold when we enlarge the attracting sets. One way to escape this artificial non-uniqueness is by asking a property of minimality. We call it so because of the use in the literature of attractors, although minimal in set theory is a different notion.

Definition 19 *We say that a global compact attractor $A(t)$ is minimal if any other global compact attractor $A'(t)$ satisfies*

$$A(t) \subset A'(t)$$

for all $t \in \mathbb{R}$.

When it exists, the minimal global compact attractor is obviously unique.

Proposition 20 *Under the assumptions of Theorem 15, the attractor*

$$A(t) = \overline{\bigcup_{B \text{ bounded}} \Omega(B,t)}$$

is minimal.

Proof. If $A'(t)$ is a global attractor, it includes all omega-limit sets, hence it includes (being closed) $A(t)$. ■

The previous solution of the uniqueness problem is very simple and efficient under the assumptions of Theorem 15, which are those we verify in examples. However, at a more conceptual level one could ask whether there are cases when we can establish uniqueness from the definition itself. Indeed, in the autonomous case (see for instance the first chapter of [56]) it is known that the analogous concept is unique as an immediate consequence of the definition, because the attractor is itself a bounded set and the attraction property of bounded sets easily implies uniqueness.

We thus mimic the autonomous case by a definition and a simple criterium.

Definition 21 *A family of compact sets $A(t)$, $t \in \mathbb{R}$, is called backward frequently bounded if there is a bounded set B and a sequence $s_n \to -\infty$ such that $A(s_n) \subset B$ for every $n \in \mathbb{N}$.*

Proposition 22 *In the class of backward frequently bounded families, the PBCGA is unique, when it exists.*

Proof. Assume $A(\cdot)$ and $\widetilde{A}(\cdot)$ are two PBCGA, both backward frequently bounded and let B be a bounded set and $s_n \to -\infty$ be a sequence such that $A(s_n) \subset B$ for every $n \in \mathbb{N}$. Take $z \in A(t)$ and, thanks to property $U(s_n,t)A(s_n) = A(t)$, let $x_n \in A(s_n)$ be such that $U(s_n,t)x_n = z$. We have $x_n \in B$. Hence we have

$$z \in \{y \in X : \exists x_n \subset B, s_n \to -\infty, U(s_n,t)(x_n) \to y\}.$$

This proves

$$A(t) \subset \Omega(B,t).$$

But

$$\Omega(B,t) \subset \widetilde{A}(t)$$

hence

$$A(t) \subset \widetilde{A}(t).$$

The converse is also true, hence the two families coincide. ■

3.2 Examples

We give two examples. The first one is related to pitchfork bifurcation (not discussed here). The second one is the finite dmensional Navier-Stokes type example given above.

Consider, over all $t \in \mathbb{R}$, the equation

$$X'(t) = q(t) X(t) - X^3(t)$$

where $q(t)$ is a given bounded function. In the following arguments we can think that we work on a "pull-back" interval $[s, t]$ or equivalently on a standard forward interval $[0, T]$: the bounds are the same, depending only on $\|q\|_\infty$.

We prove now that a pull-back absorbing family exists, made of a single bounded set (hence backward frequently bounded); in finite dimensions the closure of a bounded set is compact, hence a PBCGA exists.

Let us investigate the time-evolution of the "energy" of a solution:

$$\frac{d}{dt} X^2 = 2XX' = -2q(t) X^2 - 2X^4 \leq 2\|q\|_\infty X^2 - 2X^4.$$

What counts is the strong dissipation for large values of $|X|$ provided by the term $-2X^4$; the term $2\|q\|_\infty X^2$ may complicate the dynamics in a bounded region around the origin but not "at infinity". One way to capture rigorously these features is to estimate

$$2\|q\|_\infty X^2 - 2X^4 \leq C - X^4$$

for some $C = C(\|q\|_\infty) > 0$, then reducing the inequality to

$$\frac{d}{dt} X^2 \leq C - X^4$$

(this step is not really necessary, one can work with the original inequality). Let $y(t)$ (for us $= X^2(t)$) be a non negative differentiable functions which satisfies, for every $t \geq 0$,

$$y'(t) \leq C - y^2(t).$$

A simple picture immediately clarifies the result. Since the function $y \mapsto C - y^2(t)$, for $y \geq 0$, is positive in the interval $\left(0, \sqrt{C}\right)$, negative for $y > \sqrt{C}$, the same happens to $y'(t)$. Therefore, if $y(0) \in \left[0, \sqrt{C}\right]$, we can show that $y(t) \in \left[0, \sqrt{C}\right]$ for every $t \geq 0$. If $y(0) > \sqrt{C}$, $y(t)$ decreases until $y(t) > \sqrt{C}$. More precisely, if $y(0) \in \left[0, \sqrt{C}+1\right]$, then $y(0) \in \left[0, \sqrt{C}+1\right]$ for every $t \geq 0$; if $y(0) > \sqrt{C}+1$, in a finite time depending on $y(0)$ the function $y(t)$ enters $\left[0, \sqrt{C}+1\right]$ and then (for what already said) it never leaves it. Translated to $X(t)$: the ball $B\left(0, \sqrt{C}+1\right)$ is an absorbing set.

Now we consider equation (4) of Section 2.3 that we rewrite here:

$$x_t = x_0 - \int_{t_0}^{t} (Ax_s + B(x_s, x_s) - q(s))ds + \sqrt{Q}\omega_t - \sqrt{Q}\omega_{t_0}$$

where we now assume that q is a bounded continuous function. We give a proof of existence of absorbing set inspired to the proof given for the 2D stochastic Navier-Stokes equations in [17]. It requires a small but essential modification of the idea already explained in Section 2.3: we introduce here the auxiliary equation

$$z_t = z_{t_0} - \int_{t_0}^{t} (A + \alpha) z_s ds + \sqrt{Q}\omega_t - \sqrt{Q}\omega_{t_0}$$

with an extra damping factor $\alpha > 0$ which will play a major role. The value z_{t_0} will be properly chosen below. Set again $v_t = x_t - z_t$ to have

$$v_t = x_0 - z_{t_0} - \int_{t_0}^{t} (Av_s + B(v_s + z_s, v_s + z_s) - q(s) - \alpha z_s)ds$$

and then, as in Section 2.3,

$$\frac{d}{dt}v_t + Av_t + B(v_t + z_t, v_t + z_t) = q(t) + \alpha z_t.$$

$$\frac{d}{dt}|v_t|^2 + 2\langle Av_t, v_t\rangle + 2\langle B(v_t + z_t, v_t + z_t), v_t\rangle = 2\langle q(t) + \alpha z_t, v_t\rangle$$

$$\begin{aligned}\frac{d}{dt}|v_t|^2 + 2\lambda |v_t|^2 &\le 2\langle q(t) + \alpha z_t, v_t\rangle - 2\langle B(v_t + z_t, z_t), v_t\rangle \\ &\le 2|q(t)|^2 + 2\alpha^2 |z_t|^2 + C|v_t|^2(|z_t| + 1) + C|z_t|^4\end{aligned}$$

and thus

$$\frac{d}{dt}|v_t|^2 \le (-2\lambda + C(|z_t| + 1))|v_t|^2 + 2|q(t)|^2 + 2\alpha^2|z_t|^2 + C|z_t|^4.$$

It follows

$$\begin{aligned}|v_t|^2 &\le e^{\int_{t_0}^{t}(-2\lambda + C(|z_s|+1))ds}(x_0 - z_{t_0}) \\ &\quad + \int_{t_0}^{t} e^{\int_r^t(-2\lambda + C(|z_s|+1))ds}\left(2|q(r)|^2 + 2\alpha^2|z_r|^2 + C|z_r|^4\right)dr.\end{aligned}$$

The problem, a priori, is that $-2\lambda + C(|z_s| + 1)$ could be non-dissipative, the term $C(|z_s| + 1)$ could spoil the dissipativity of the term -2λ. Only if $|z_s|$ is very small, we still have dissipativity. But we can get $|z_s|$ very small by a proper choice of the auxiliary parameter α.

Understood this idea, we only describe the strategy, without necessary details that can be found in [17]. Firse, one takes as $z_{t_0} = z_{t_0}(\omega)$ the value which makes $z_t(\omega)$ a stationary stochastic process. This process is ergodic, also in the negative direction of time and we have, $\mathbb{P}$-a.s.,

$$\lim_{T\to+\infty} \frac{1}{T}\int_{-T}^{0} |z_s|\, ds = \mathbb{E}\left[|z_0|\right].$$

The number $\mathbb{E}[|z_0|]$ can be made arbitrarily small by chosing α large. This way it is possible to have

$$\lim_{T\to+\infty} \frac{1}{T}\int_{-T}^{0} C\left(|z_s|+1\right) ds \leq \lambda$$

and thus

$$\lim_{T\to+\infty} \frac{1}{T}\int_{-T}^{0} \left(-2\lambda + C\left(|z_s|+1\right)\right) ds \leq -\lambda. \tag{6}$$

Moreover, $|z_s|$ increases less than linearly as $s \to -\infty$. These facts allow to prove that, given any x_0,

$$e^{\int_{t_0}^{t}(-2\lambda+C(|z_s|+1))ds} \left|x_0 - z_{t_0}\right| \leq 1$$

for all $t_0 < 0$ with $|t_0|$ large enough. And similarly we can prove a (random) bound, uniformly in $t_0 < 0$ with $|t_0|$ large, for the integral term

$$\int_{t_0}^{t} e^{\int_r^t(-2\lambda+C(|z_s|+1))ds} \left(2\left|q(r)\right|^2 + 2\alpha^2\left|z_r\right|^2 + C\left|z_r\right|^4\right) dr$$

exploiting the sub polynomial growth of $\left(2\left|q(r)\right|^2 + 2\alpha^2\left|z_r\right|^2 + C\left|z_r\right|^4\right)$ as $r \to -\infty$ and the exponential bound for $e^{\int_r^t(-2\lambda+C(|z_s|+1))ds}$ coming from (6).

3.3 About invariance

In the autonomous case we say that a set $A \subset X$ is invariant if

$$\varphi_t(A) = A \qquad \text{for all } t$$

(also the simpler property of positive invariance, $\varphi_t(A) \subset A$, sometimes is already quite useful). To avoid trivialities (like $\varphi_t(X) = X$), one usually add the requirement of boundedness or compactness of invariance sets. This identifies a notion, for instance compact invariant set, which is usually quite useful and interesting. A probability measure μ is invariant if

$$\varphi_t\mu = \mu \qquad \text{for all } t$$

where $\varphi_t\mu$ is the probability measure defined by $(\varphi_t\mu)(B) = \mu\left(\varphi_t^{-1}(B)\right)$ for all Borel sets B. Being a probability measure implies it is almost supported on compact sets (in the sense of tightness), hence it incorporates a requirement similar to the boundedness or compacteness of invariant sets.

In the non-autonomous case the analogous definitions are:

Definition 23 *A family of sets* $\{A_{q,\omega}(t) \subset X, t \in \mathbb{R}\}$ *is called a compact invariant set for* $U_{q,\omega}(s,t)$ *if* $A_{q,\omega}(t)$ *is compact for every* t *and*

$$U_{q,\omega}(s,t)\, A_{q,\omega}(s) = A_{q,\omega}(t) \qquad \text{for all } s \leq t.$$

A family of Borel probability measures $\{\mu_{q,\omega}(t), t \in \mathbb{R}\}$ *is invariant for* $U_{q,\omega}(s,t)$ *if*

$$U_{q,\omega}(s,t)\, \mu_{q,\omega}(s) = \mu_{q,\omega}(t) \qquad \text{for all } s \leq t.$$

The main remark of this section is that these concepts are almost empty, opposite to the autonomous case. indeed, they are satisfied by trivial object of no long-time interest. For instance, assume the dynamics can be run both forward and backward globally in time - as it is for the examples of Section 2.3) and let

$$x_{q,\omega}(t), \qquad t \in \mathbb{R}$$

be a trajectory. It satisfies

$$U_{q,\omega}(s,t)\,(x_{q,\omega}(s)) = x_{q,\omega}(t)$$

and thus the singleton $\{x_{q,\omega}(t)\}$ is a compact invariant set. The same fact holds for any family of trajectories which is compact at some time t. Let $\{\mu_{q,\omega}(t), t \in \mathbb{R}\}$ be simply defined by

$$\mu_{q,\omega}(t) = \delta_{x_{q,\omega}(t)}.$$

It is an invariant measure. This genericity is very far from the specificity of invariant sets and measures of the autonomous case!

There are two ways to excape this triviality. One is to ask that invariant sets and measures come, in a sutable pull-back sense, from $-\infty$. This idea for sets has been developed above by means the definition of attractor. For measures, it is less obvious, see Section 3.4 below, keeping also in mind the followng remark.

Remark 24 *Also in the autonomous case it happens for many examples that invariant sets and invariant measures are non unique but this corresponds to interesting different long time objects of the dynamics. In the example above the non-uniqueness is simply related to different initial conditions, in spite of the fact that all trajectories tend to approach each other when time increases. Thus it is an artificial non-uniqueness due to a drawback of the concept.*

Another way to excape the triviality described above is to define *cocycle invariance*. If $A_{q,\omega}(t)$ is invariant for the NADS in the sense of Definition 23, and the NADS comes from a cocycle $\varphi_t(q,\omega)$, then from (2) we deduce

$$\varphi_{t-s}(\Phi_s(q,\omega))\, A_{q,\omega}(s) = A_{q,\omega}(t) \qquad \text{for all } s \leq t.$$

Namely, changing times,

$$\varphi_t(\Phi_s(q,\omega))\, A_{q,\omega}(s) = A_{q,\omega}(t+s) \qquad \text{for all } t \geq 0 \text{ and all } s;$$

in particular,

$$\varphi_t(q,\omega)\, A_{q,\omega}(0) = A_{q,\omega}(t) \qquad \text{for all } t \geq 0.$$

Until now nothing has changed. But now we can ask more: that

$$A_{q,\omega}(t) = A_{\Phi_t(q,\omega)}(0) \tag{7}$$

in which case the condition becomes

$$\varphi_t(q,\omega)\, A_{q,\omega}(0) = A_{\Phi_t(q,\omega)}(0) \qquad \text{for all } t \geq 0.$$

The new condition (7) is not automatically satisfied. *It is a sort of Φ_t-stationarity of the random set $A_0(\omega)$.*

Definition 25 *Given a cocycle $(\mathcal{Q}, \vartheta_t, (\Omega, \mathcal{F}, \mathbb{P}), \theta_t, (X, d), \varphi_t(q,\omega))$, we say that a set $A(q,\omega)$ parametrized by $q \in \mathcal{Q}$ and $\omega \in \Omega$ is cocycle invariant if*

$$\varphi_t(q,\omega)\, A(q,\omega) = A(\Phi_t(q,\omega)) \qquad \textit{for all } t \geq 0.$$

Moreover, we say that a Borel probability measure $\mu(q,\omega)$ parametrized by $q \in \mathcal{Q}$ and $\omega \in \Omega$ is cocycle invariant if

$$\varphi_t(q,\omega)\, \mu(q,\omega) = \mu(\Phi_t(q,\omega)) \qquad \textit{for all } t \geq 0.$$

Proposition 26 *If $A(q,\omega)$ (resp. $\mu(q,\omega)$) is cocycle invariant then*

$$A_t(q,\omega) := A(\Phi_t(q,\omega))$$

(resp. $\mu_t(q,\omega) := \mu(\Phi_t(q,\omega))$) is invariant for $U_{q,\omega}(s,t)$.

Proof. We have, for all $s \leq t$,

$$\begin{aligned} U_{q,\omega}(s,t)\, A_s(q,\omega) &= \varphi_{t-s}(\Phi_s(q,\omega))\, A(\Phi_s(q,\omega)) \\ &= A(\Phi_{t-s}\Phi_s(q,\omega)) \\ &= A_t(q,\omega). \end{aligned}$$

■

Example 27 *Consider the cocycle $\varphi_t(\omega)$ (let us drop q for simplicity) associated to the stochastic equation*

$$dX_t = -\alpha X_t dt + \sigma dW_t.$$

For almost every ω, the integral

$$x_0(\omega) = -\int_{-\infty}^{0} \alpha e^{\alpha s} \sigma \omega(s) ds$$

is convergent, as a consequence of the fact that, for almost every ω, $\frac{W_t}{t}$ converges to 0 as t goes to infinity (see for example [35], problem 9.3). We show that $\{x_0(\omega)\}$ is an invariant set, and then the random measure $\delta_{x_0(\omega)}$ is invariant for φ. We have

$$\begin{aligned}
\varphi(t,\omega)x_0(\omega) &= e^{-\alpha t}x_0(\omega) + \sigma\omega(t) - \int_0^t \alpha e^{-\alpha(t-s)}\sigma\omega(s)ds \\
&= -\int_{-\infty}^{0} \alpha e^{-\alpha(t-s)}\sigma\omega(s)ds + \sigma\omega(t) - \int_0^t \alpha e^{-\alpha(t-s)}\sigma\omega(s)ds \\
&= -\int_{-\infty}^{t} \alpha e^{-\alpha(t-s)}\sigma\omega(s)ds + \sigma\omega(t) \\
&= -\int_{-\infty}^{0} \alpha e^{\alpha s}\sigma\omega(t+s)ds + \sigma\omega(t)\left(\int_{-\infty}^{0} \alpha e^{\alpha s} ds\right) \\
&= -\int_{-\infty}^{0} \alpha e^{\alpha s}\sigma(\omega(t+s) - \omega(t))ds = x_0(\theta_t\omega).
\end{aligned}$$

The compact global attractor constructed by Theorem 15 is invariant in this new sense, when the NADS comes from a cocycle.

Theorem 28 *Let $(\mathcal{Q}, \vartheta_t, (\Omega, \mathcal{F}, \mathbb{P}), \theta_t, (X,d), \varphi_t(q,\omega))$ be a cocycle and let $U_{q,\omega}(s,t)$ be the associated NADS. Let $A_t(q,\omega)$ be the minimal global compact attractor associated to $U_{q,\omega}(s,t)$. Then it holds (7), which implies that $A_0(q,\omega)$ is invariant for the cocycle, in the sense of Definition 25.*

Proof. We prove the analog of (7) for the omega-limit sets:

$$\Omega_t(B; q,\omega) = \Omega_0(B; \Phi_t(q,\omega)).$$

By definition of $A_t(q,\omega)$, it follows (7). The identity of omega-limit sets is due to the following fact. The set $\Omega_t(B;q,\omega)$ is made of points y such that there exist $x_n \in B$, $s_n \to -\infty$ such that $U_{q,\omega}(s_n,t)\, x_n \to y$. The set $\Omega_0(B;\Phi_t(q,\omega))$ is made of points y' such that there exist $x'_n \in B$, $s'_n \to -\infty$ such that $U_{\Phi_t(q,\omega)}(s'_n, 0)\, x'_n \to y'$. But the latter formula, from (3), can be rewritten as $U_{q,\omega}(s'_n + t, t)\, x'_n \to y'$. Thus (renaming $s_n = s'_n + t$) we see that the two conditions are equivalent, they define the same set. ■

3.4 Statistical equilibrium

Consider a NADS $(\mathcal{Q}, (\Omega, \mathcal{F}, \mathbb{P}), (X, d), U_{q,\omega}(t_0, t))$. We have defined above invariance of a family of measures $\mu_{\omega,q}(t)$ but, as in the case of the attractor, we have to add some property to avoid triviality. One possibility is simply the concept of cocycle invariance.

Let us also discuss a second possibility, based on the intuitive idea of coming from $-\infty$.

Definition 29 *Let $\mu_q(t)$, $q \in \mathcal{Q}$, $t \in \mathbb{R}$, be a family of Borel measures on X. If, given any $q \in \mathcal{Q}$, $t \in \mathbb{R}$, the limit*

$$\mu_{\omega,q}(t) = \lim_{s \to -\infty} U_{q,\omega}(s, t)_\sharp \mu_q(s)$$

exists $\mathbb{P}$-a.s., we call $\mu_{\omega,q}(t)$ the statistical equilibrium associated to the family $\mu_q(t)$.

The most interesting case from the viewpoint of interpretation is when

$$\mu_q(t) = \lambda$$

a given measure λ with geometrical or physical meaning, like normalized volume measure on a Riemannian manifold. In this case we say that $\mu_{\omega,q}(t)$ *is the physical statistical equilibrium relative to* λ.

Quite interesting is also the case when $\mu_q(t)$ is the Climate, following the definition below. In such case, the associated $\mu_{\omega,q}(t)$ (if it exists) will be called *Climate statistical equilibrium*.

There are general results stating that these statistical equilibria are supported on the random time-dependent attractor. Hence we may see them as statitical specifications, on the attractor, of what we really observe with relevant frequency.

Opposite to the case of the attractor, which exists under very general conditions satisfied in a myriad of examples, the existence of a statistical equilibrium, especially a physical one, is very difficult. We shall see below a result. For general results on this topic see [3], [6], [21], [36], [39], [58].

4 Markov and Kolmogorov operators

Consider a NADS

$$(\mathcal{Q}, (\Omega, \mathcal{F}, \mathbb{P}), (X, d), U_{q,\omega}(t_0, t)).$$

On the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ denote by $\mathbb{E}$ the mathematical expectation. We introduce a family of linear operators $P^q_{t_0,t}$ acting on functions, indexed by $q \in \mathcal{Q}$. Denote by $B(X)$ the space of bounded measurable functions on X, endowed with the Borel σ-algebra. Given $f \in B(X)$ we call $P^q_{t_0,t} f$ the function on X, element of $B(X)$ again, defined as

$$P^q_{t_0,t} f(x) = \mathbb{E}[f(U_{q,\cdot}(t_0, t)(x))] \qquad x \in X.$$

It is called, sometimes *Kolmogorov evolution operator* (Kolmogorov semigroup when q is constant), and satisfies the evolution property

$$P^q_{t_0,t_0} = Id$$
$$P^q_{t_0,t} = P^q_{s,t} \circ P^q_{t_0,s}$$

for every $t_0 \leq s \leq t$ in $\mathbb{R}$. Denote by $C_b(X)$ the space of continuous bounded functions on X; obviously $C_b(X) \subset B(X)$ hence we can compute $P^q_{t_0,t}f$ for $f \in C_b(X)$. When it happens that $P^q_{t_0,t}f \in C_b(X)$ for every $f \in C_b(X)$, we say that $P^q_{t_0,t}$ is *Feller.* This property is very common, satisfied for instance when $U_{q,\cdot}(t_0,t)$ is a continuous map on X, for $\mathbb{P}$-a.e. ω, as we have assumed in our NADS. Thus in our set-up Feller property holds.

In the special case when $P^q_{t_0,t}$ even maps $B(X)$ into $C_b(X)$ we say that $P^q_{t_0,t}$ is *Strong Feller.* This is verified in special but important cases, for many SDEs and SPDEs; it is a strong regularizing property.

Let μ be a Borel probability measure on X. We call $\mu P^q_{t_0,t}$ a new Borel probability measure on X defined for all $f \in C_b(X)$ by the relation

$$\int_X f(x) \left(\mu P^q_{t_0,t}\right)(dx) = \int_X \left(P^q_{t_0,t}f\right)(x)\,\mu(dx)$$

shortly written as a duality as

$$\left\langle \mu P^q_{t_0,t}, f\right\rangle = \left\langle \mu, P^q_{t_0,t}f\right\rangle.$$

We have implicitly used the Feller property (but for systems without it, it is sufficient to replace $C_b(X)$ by $B(X)$). The action of $P^q_{t_0,t}$ on the left on measures is dual to the action on the right on functions, although duality here is understood precisely in the sense of the formula just given and not as a duality of operators between dual spaces (because, even extending the previous definitions to bounded signed Borel measures, they are not dual to $C_b(X)$ in general; they are when X is compact). We shall call $P^q_{t_0,t}$, when acting on the left on measures, *Markov evolution operator* (semigroup when q is constant); it satisfies the evolution properties above.

Notice that, by definition,

$$\mu P^q_{t_0,t} = \mathbb{E}\left[U_{q,\cdot}(t_0,t)\,\mu\right]$$

where $U_{q,\cdot}(t_0,t)\,\mu$ is the image of μ under the map $U_{q,\cdot}(t_0,t)$ and the expectation is a meaningful operation in the convex set of Borel probability measures. The formula holds because, by the very definition of such expectation, plus the rule of change of measure and

Fubini-Tonelli theorem,

$$\begin{aligned}\int_X f(y)\,\mathbb{E}\left[U_{q,\cdot}(t_0,t)\,\mu\right](dy) &:= \mathbb{E}\left[\int_X f(y)\left(U_{q,\cdot}(t_0,t)\,\mu\right)(dy)\right] \\ &= \mathbb{E}\left[\int_X f\left(U_{q,\cdot}(t_0,t)(x)\right)\mu(dx)\right] \\ &= \int_X \mathbb{E}\left[f\left(U_{q,\cdot}(t_0,t)(x)\right)\right]\mu(dx).\end{aligned}$$

Definition 30 *We say that a family $\mu_q(t)$ of Borel probability measures is invariant for $P^q_{t_0,t}$ if*

$$\mu_q(t) = \mu_q(t_0)\,P^q_{t_0,t}$$

for every $q \in \mathcal{Q}$, $t_0 \leq t$ in $\mathbb{R}$.

4.1 The Fokker-Planck equation

The density $p_{t_0,t}(x)$ (assuming for simplicity of notations that the density exists) of the law on Borel sets of $\mathbb{R}^d$ of a solution $X_t^{t_0}$ of the equation

$$dX_t = b(q(t), X_t)\,dt + dW_t \qquad t \geq t_0$$

with a (potentially random) initial condition X_{t_0} at time t_0 is a (weak, in the sense of distributions) solution of the PDE, called Fokker-Planck equation,

$$\partial_t p = \frac{1}{2}\Delta p - \operatorname{div}\left(p \cdot b(q(t), x)\right)$$

with initial condition at time t_0 given by the density of X_{t_0}. If we call μ_{t_0} the law of X_{t_0}, then the law of $X_t^{t_0}$ is $\mu_{t_0}P^q_{t_0,t}$, which is also $p^q_{t_0,t}(x)\,dx$ (emphasizing its dependence on q), hence

$$\left(\mu_{t_0}P^q_{t_0,t}\right)(dx) = p^q_{t_0,t}(x)\,dx.$$

A family $\mu_q(t)$ with densities $p^q_t(x)$ is invariant if $p^q_t(x)$ is a solution for all times of the Fokker-Planck equation, without initial condition.

5 The Climate

Claim 31 *Climate is a particular family of Borel probability measures*

$$\mu_q(t)$$

on X, invariant for the Markov evolution operator in the sense that

$$\mu_q(t) = \mu_q(t_0)\,P^q_{t_0,t}$$

for every $q \in \mathcal{Q}$, $t_0 \leq t$ in $\mathbb{R}$.

Opposite to the attractor $A_{q,\omega}(t)$, which is unique by definition of minimality and just needs to exist, Climate in principle is less easy to identify, among the potential several invariant measures. However, there are easy conditions for uniqueness that we recall below and thus, for later use, we shall assume it is unique.

5.1 Uniqueness of the Climate

The family $\mu_q(t)$, also called evolution system of measures in part of the literature, is unique under assumptions satisfied by several systems, including 2D Navier-Stokes equations in certain regimes. Here we develop some elements of that theory in the simplest possible case; see [19], [13], [14] for other results.

Example 32 *Consider the example of Proposition 3 and call X_t^{q,t_0,x_0} the solution for $t \geq t_0$ corresponding to q, with $X_{t_0}^{q,t_0,x_0} = x_0$; let $P_{t_0,t}^q$ be the associated Kolmogorov operator, $\left(P_{t_0,t}^q \phi\right)(x) = \mathbb{E}\left[\phi\left(X_t^{q,t_0,x}\right)\right]$. Assume that $q(t)$ is given by the solution of a differential equation*

$$\frac{d}{dt} q(t) = f(q(t)).$$

Call $q^{t_0,q_0}(t)$ the solution for $t \geq t_0$ with $q^{t_0,q_0}(t_0) = q_0$. Then the pair $(X_t, q(t))$ satisfies the autonomous stochastic differential system

$$\begin{aligned} dX_t &= b(q(t), X(t))\,dt + dW_t \\ dq(t) &= f(q(t))\,dt. \end{aligned}$$

Call S_t its Kolmogorov operator,

$$(S_t \Phi)(x_0, q_0) = \mathbb{E}\left[\Phi\left(X_t^{q^{0,q_0},0,x_0}, q^{0,q_0}(t)\right)\right]$$

Then

$$\begin{aligned} (S_t \Phi)(x_0, q_0) &= \left(P_{0,t}^{q^{0,q_0}} \Phi\left(\cdot, q^{0,q_0}(t)\right)\right)(x_0) \\ \left(P_{t_0,t}^q \phi\right)(x_0) &= \left(S_{t-t_0} \phi\right)(x_0, q(t_0)). \end{aligned}$$

The previous example motivates the following structure. Assume to have an ergodic metric dynamical system $\left(Y, \mathcal{G}, \nu, (T_t)_{t\in\mathbb{R}}\right)$, namely a probability space $(Y, \mathcal{G}, \nu)$ and a group $(T_t)_{t\in\mathbb{R}}$ of measurable transformations of Y that leaves ν invariant and such that ν is ergodic for $(T_t)_{t\in\mathbb{R}}$. Assume that

$$q(t) = T_t q_0$$

for some $q_0 \in Y$. Let $P_{t_0,t}^q$ be the Kolmogorov operator associated to a NADS. Assume that the equation

$$(S_t \Phi)(x_0, q_0) = \left(P_{0,t}^{T_\cdot q_0} \Phi(\cdot, T_t q_0)\right)(x_0)$$

defines a semigroup $(S_t)_{t\geq 0}$ on $C_b(X \times Y)$. Following [19] we may translate ergodic informations from S_t to $P^q_{t_0,t}$. Under this additional structure, the following results are proved in [19], Section 4.

We say that $P^q_{t_0,t}$ is *regular* if, for every $t_0 \leq t$ and $x, y \in X$ the following two measures are equivalent:

$$\delta_x P_{t_0,t} \sim \delta_y P^q_{t_0,t}.$$

Here $\delta_x P^q_{t_0,t}$ can be interpreted as the transition probability starting from x. The main result is:

Theorem 33 *If $P^q_{t_0,t}$ is regular, then there exists at most one invariant family $\mu(t)$.*

Let us also discuss classical sufficient conditions for regularity. We say that $P^q_{t_0,t}$ is *irreducible* if, for every $t_0 \leq t$ and $x \in X$ and open set $A \subset X$ we have

$$\delta_x P^q_{t_0,t}(A) > 0.$$

With positive probability the system goes from any point to any open set, in arbitrary time. The following two lemmata do not require the additional structure above.

Lemma 34 *If $P^q_{t_0,t}$ is irreducible and Strong Feller, then it is regular.*

In SDE examples with additive noise, Strong Feller property is related to the regularizing properties of the Laplacian in the corresponding Fokker-Planck equation. In the case of degenerate multiplicative noise it is more difficult and depends on examples. In any case, conceptually, it is a property of regularity of the PDEs associated to the SDE.

Irreducibility is a more geometric property, related to control theory. The question is whether controls replacing noise may lead from any point to any open set in arbitrary time. In the case of full additive noise this is always true.

Lemma 35 *If $P^q_{t_0,t}$ is regular and $\mu(t)$ is invariant, then*

$$\mu(t) \sim \delta_x P^q_{t_0,t}$$

for every $t_0 \leq t$ and $x \in X$.

The existence of Climate can be proved by Krylov-Bogoliubov argument as soon as some tightness property of the system, very close to the existence of a compact absorbing set (in fact easier), holds.

The verification of the assumptions for the 2D Navier-Stokes equations with noise and non-autonomous input, are given in [19].

5.2 Weather dynamics and Climate dynamics

The family $U_{q,\omega}(s,t)$ is the *weather dynamics*.

The evolution operators $P_{s,t}^{q}$ are the *climate dynamics*. This is related to Kolmogorov and Fokker-Planck non autonomous equations, which then becomes central object of investigation in climate research.

5.3 Examples

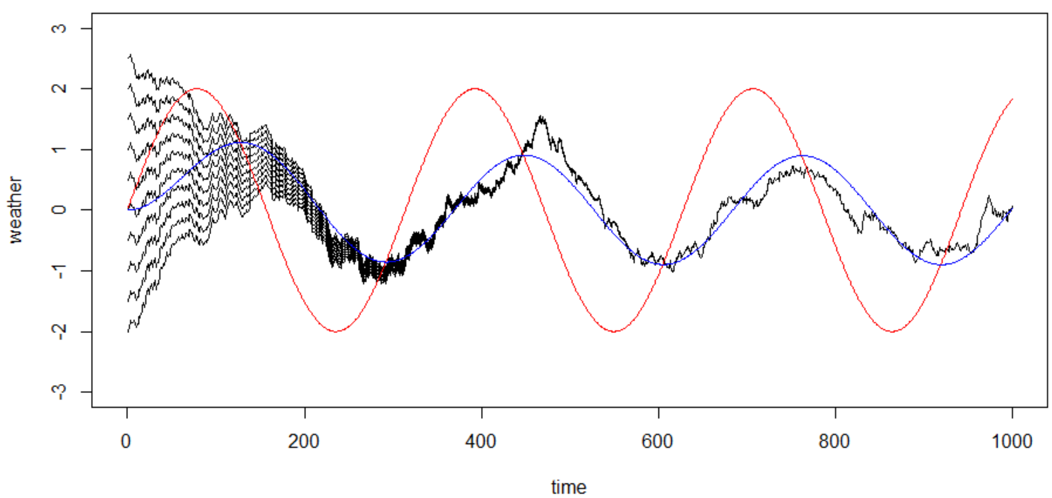
We consider here two quite artificial examples from the viewpoint of climate but useful to see the objects we have introduced.

In the first example, consider the equation

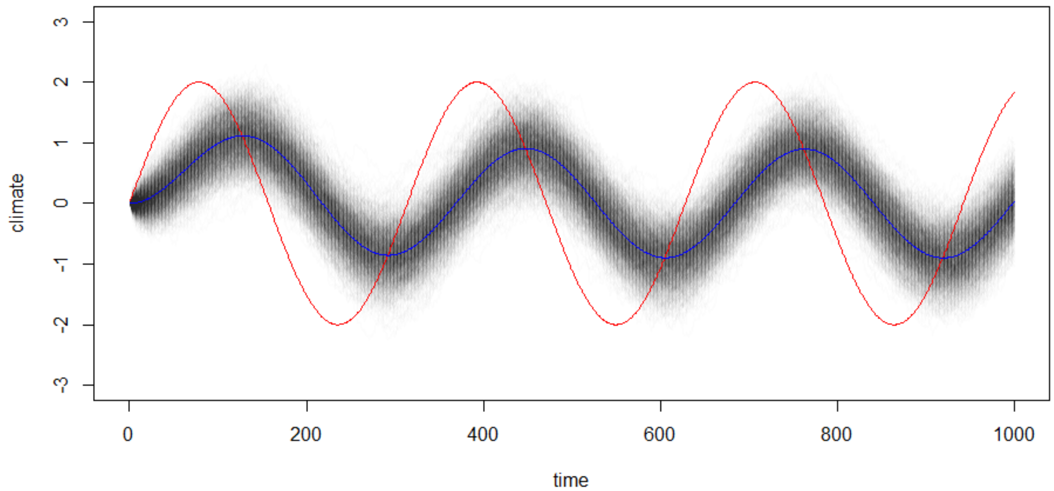
$$dX_t = -X_t dt + q(t)\, dt + \sigma dW_t$$

with

$$q(t) = 2 \cdot \sin t.$$



Weather: ten different initial conditions with noise intensity $\sigma = 0.5$. The blue curve is the deterministic trajectory ($\sigma = 0$) and the red one the signal $q(t)$.



Climate: at every time t, the grey zone gives a visual idea of the probability distribution $\mu_q(t)$. The picture is obtained running Monte Carlo with sample size 100.

Remark 36 *A natural alternative to our definitions would be to assume that t varies so slowly that, at any time t_0, the relevant invariant measure $\widetilde{\mu}_q(t_0)$ is the invariant measure of the system*

$$dX_t = -X_t dt + q(t_0)\, dt + \sigma dW_t$$

namely freezing $q(t_0)$. The invariant measure is a Gaussian with average

$$m(t_0) = q(t_0).$$

But we see from the plot that this is completely false. It is like to conjecture that the deterministic periodic trajectory ($\sigma = 0$) is obtained at time t_0 as the fixed point of the equation

$$x_t' = -x_t + q(t_0)$$

which is $q(t_0)$; but the blue and red lines are completely different. It is easy to see that the difference has to do with the link between the frction parameter, here equal to 1, and the period of the signal $q(t)$.

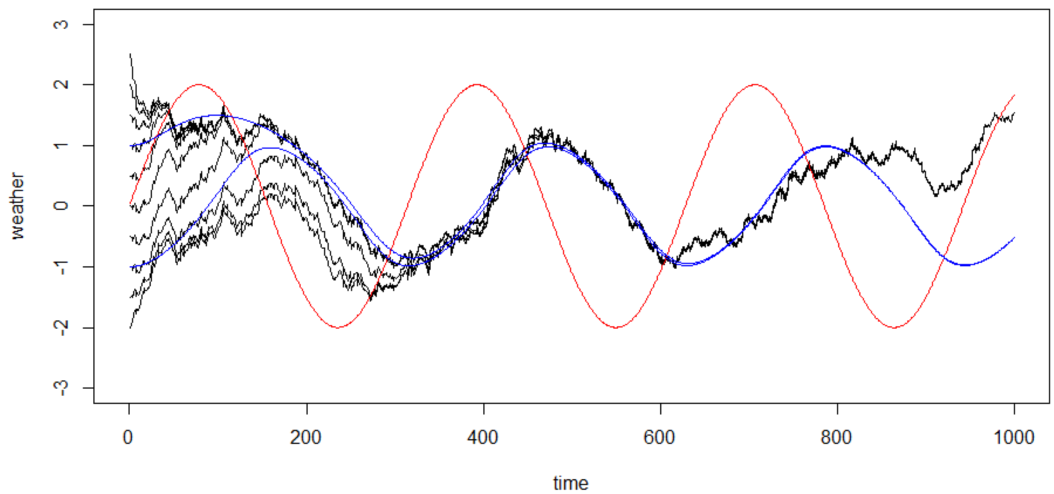
In the second example, consider the equation with double well drift:

$$dX_t = \left(X_t - X_t^3\right) dt + q(t)\, dt + \sigma dW_t$$

with

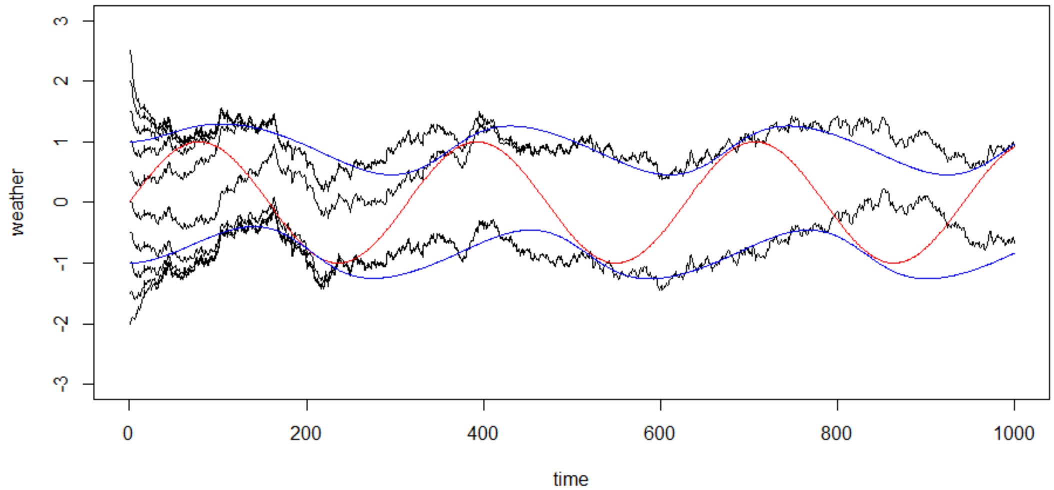
$$q(t) = \alpha \cdot \sin t.$$

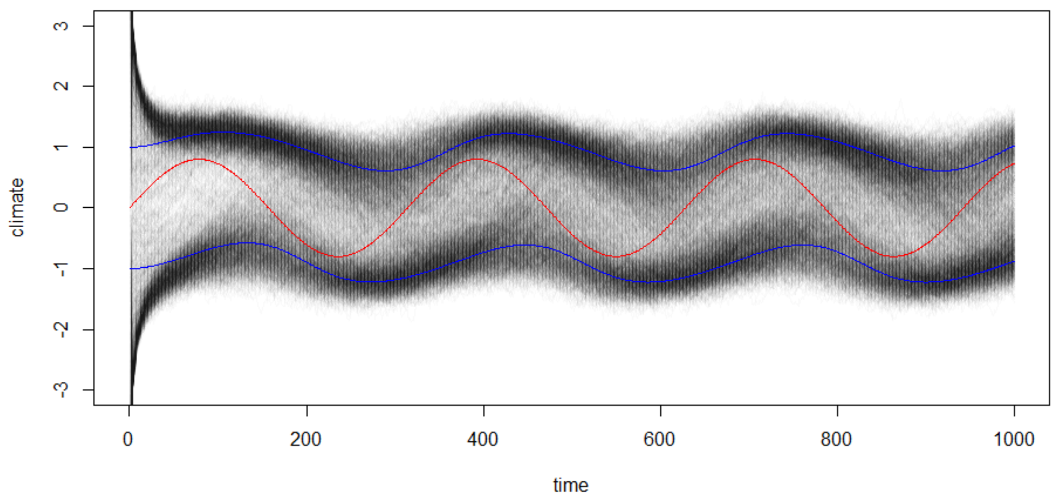
First of all, a specal phenomenon may occur depending on parameters: synchronization. If we take $\alpha = 2$ as above, already in the deterministic case all initial conditions synchronize to a single trajectory, in spite of the double well.



Double well example with too strong periodic force: synchronization.

Let us then take a smaller value, $\alpha = 1$ and see weather and climate.





6 Relations between Weather and Climate

The climate is what we expect; weather is what we get. A mathematical translation of the first part of this sentence is given by Theorem 37 below.

Consider a NADS

$$(\mathcal{Q}, (\Omega, \mathcal{F}, \mathbb{P}), (X, d), U_{q,\omega}(t_0, t)).$$

Recall that *weather* is either $A_{q,\omega}(t)$ or, more precisely, a certain family of measures $\mu_{q,\omega}(t)$, invariant for the dynamics

$$U_{q,\omega}(s,t)\,\mu_{q,\omega}(s) = \mu_{q,\omega}(t) \tag{8}$$

and with suitable properties of emergence from past times. Recall also that *climate* is a certain family of measures $\mu_q(t)$, invariant for the Markov evolution operator

$$\mu_q(s)\,P^q_{s,t} = \mu_q(t)\,. \tag{9}$$

Also in this case we should think to suitable properties of emergence from past but the uniqueness result of Section 5.1 help us to avoid difficult arguments.

In this section we relate the two concepts, assuming to have a white noise NADS.

Theorem 37 *If $\mu_{q,\omega}(t)$ satisfies (8) for a white noise NADS and*

$$\mu_{q,\cdot}(s) \text{ is } \mathcal{F}_{s,t}\text{-independent}$$

then

$$\mu_q(t) := \mathbb{E}\left[\mu_{q,\cdot}(t)\right]$$

satisfies (9).

Proof. Using well known properties of conditional expectation,

$$\begin{aligned}\mu_q(s)\,P^q_{s,t} &= \mathbb{E}\left[U_{q,\cdot}(s,t)\,\mu_q(s)\right] = \mathbb{E}\left[U_{q,\cdot}(s,t)\,\mathbb{E}\left[\mu_{q,\cdot}(s)\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[U_{q,\cdot}(s,t)\,\mu_{q,\cdot}(s)\,|\mathcal{F}_{s,t}\right]\right] \\ &= \mathbb{E}\left[U_{q,\cdot}(s,t)\,\mu_{q,\cdot}(s)\right] = \mathbb{E}\left[\mu_{q,\cdot}(t)\right] = \mu_q(t)\,.\end{aligned}$$

■

Theorem 38 *If $\mu_q(t)$ satisfies 9) for a white noise NADS, then the measure-valued process*

$$\left(\mu_{q,\omega}(s,t)\right)_{s\in(-\infty,t]}$$

defined as

$$\mu_{q,\omega}(s,t) := U_{q,\omega}(s,t)\,\mu_q(s)$$

is a $\mathcal{F}_{s,t}$ martingale in the s variable in the decreasing order: if $s < s'$, in the sense of measures applied to test functions of class $C_b(X)$ we have the identity

$$\mathbb{E}\left[\mu_{q,\omega}(s,t)\,|\mathcal{F}_{s',t}\right] = \mu_{q,\omega}\left(s',t\right).$$

Its limit as $s \to -\infty$ exists a.s. and defines a measure

$$\mu_{q,\omega}(t) = \lim_{s\to-\infty} \mu_{q,\omega}(s,t)$$

which is, then, the statistical equilibrium associated to the family $\mu_q(t)$. It holds that $\mu_{q,\cdot}(s)$ is $\mathcal{F}_{s,t}$-independent and thus, by Theorem 37,

$$\mu_q(t) := \mathbb{E}\left[\mu_{q,\cdot}(t)\right].$$

Moreover, if the global compact attractor $A_{q,\omega}(t)$ exists, then the support of $\mu_{q,\omega}(t)$ is in $A_{q,\omega}(t)$.

Proof. The martingale property, which should be written against test functions of class $C_b(X)$, is

$$\begin{aligned}
\mathbb{E}\left[U_{q,\cdot}(s,t)\,\mu_q(s)\,|\mathcal{F}_{s',t}\right] &= \mathbb{E}\left[U_{q,\cdot}\left(s',t\right)U_{q,\cdot}\left(s,s'\right)\mu_q(s)\,|\mathcal{F}_{s',t}\right] \\
&= U_{q,\cdot}\left(s',t\right)\mathbb{E}\left[U_{q,\cdot}\left(s,s'\right)\mu_q(s)\,|\mathcal{F}_{s',t}\right] \\
&= U_{q,\cdot}\left(s',t\right)_\sharp\,\mathbb{E}\left[U_{q,\cdot}\left(s,s'\right)\mu_q(s)\right] \\
&= U_{q,\cdot}\left(s',t\right)\left(\mu_q(s)\,P^q_{s,s'}\right) \\
&= U_{q,\cdot}\left(s',t\right)\mu_q\left(s'\right).
\end{aligned}$$

We omit the verification of the other properties. ■

The interpretation of Theorem 37 is that *climate is the average of weather* (under suitable assumptions). The interpretation of Theorem 38 is that behind a climate there is a weather, the average of which is the climate; moreover, the weather is a statistical subset of possibilities of the attractor, the potential weather.

More important, Theorem 38 provides the *existence of the weather*, question which was left open until now. However, this theory of weather is only partially satisfactory: it exists, its average is the climate, but it comes from the past from a time-dependent measure of moderate interpretation; it would be better to know it comes from a basic measure like Lebesgue, what we have called a physical measure. We add some remarks below on this issue.

6.1 Is Climate a time-average of Weather?

Theorem 37 states that climate is the average of weather, but average understood as a probabilistic average over possible noise realizations. For physical interpretation it would be more important to establish that Climate a time-average of weather.

If we take literally this sentence, being weather a random time-dependent measure $\mu_{q,\omega}(t)$, what we would like to prove is that

$$\frac{1}{\Delta T}\int_{I(t,\Delta T)}\mu_{q,\omega}(s)\,ds \sim \mu_q(t) \qquad (10)$$

where $I(t,\Delta T)$ is an interval around t of size ΔT.

Remark 39 *To be even closer to reality, we should take the time average of the empirical measure of a trajectory started very far in the past:*

$$\frac{1}{\Delta T}\int_{I(t,\Delta T)}\delta_{U_{q,\omega}(t_0,s)(x_0)}ds$$

for some very large and negative t_0 and some x_0. Investigation of such quantity is even more difficult because of the ambiguity of t_0 and x_0 and the lack of stationarity of such object even in the autonomous case; thus for the time being we restrict ourselves to the time average of the statistical equilibrium.

The size ΔT should be "small" compared to the typical time-length of variation of $\mu_q(t)$. For instance, if we discuss climate at the scale of years, namely we aim to appreciate variations of climate $\mu_q(t)$ at times of order of 300 days, for the time-average of weather we have to take time intervals of the order of 30 days. If we want to appreciate variations at times of the order of 10 years (like in the investigation of the effect of CO_2 emissions) we may take averages over one year.

Thus it is not meaningful to take a limit as $\Delta T \to +\infty$; on the contrary we should take $\Delta T \to 0$, in a *suitable scaling limit*, namely after the system is parametrized by $\epsilon > 0$ and we take the limit as $\epsilon \to 0$.

The problem resembles hydrodynamic limits, with time replacing space: the slow time variations correspond to the fact that the system is not at equilibrium at the macroscopic level. The time average in (10) correspond to the idea of *local equilibrium*.

Let us take as weather $\mu_{q,\omega}(s)$ the statistical equilibrium associated to the climate $\mu_q(t)$, under the assumption that the NADS is white noise and that $\mu_q(t)$ exists and it is unique. Let $f \in C_b(X)$ be an observable. Let us investigate the approximation (10) in mean square:

$$\mathbb{E}\left[\left|\frac{1}{\Delta T}\int_{I(t,\Delta T)}\langle\mu_{q,\omega}(s), f\rangle ds - \langle\mu_q(t), f\rangle\right|^2\right].$$

If ΔT is small and the function $s \mapsto \langle\mu_q(s), f\rangle$ is continuous, the error

$$\delta_1 := \left|\langle\mu_q(t), f\rangle - \frac{1}{\Delta T}\int_{I(t,\Delta T)}\langle\mu_q(s), f\rangle ds\right|$$

is small. Hence, up to an error related to δ_1, we have to investigate the quantity

$$\mathbb{E}\left[\left|\frac{1}{\Delta T}\int_{I(t,\Delta T)}\left(\langle\mu_{q,\omega}(s), f\rangle - \langle\mu_q(s), f\rangle\right) ds\right|^2\right].$$

Let us introduce the covariance function

$$\begin{aligned} C_{f,g}(s,s') &:= \mathbb{E}\left[\left(\langle\mu_{q,\omega}(s), f\rangle - \langle\mu_q(s), f\rangle\right)\left(\langle\mu_{q,\omega}(s'), g\rangle - \langle\mu_q(s'), g\rangle\right)\right] \\ &= \mathbb{E}\left[\langle\mu_{q,\omega}(s), f\rangle\langle\mu_{q,\omega}(s'), g\rangle\right] - \langle\mu_q(s), f\rangle\langle\mu_q(s'), g\rangle \end{aligned}$$

for every pair $f, g \in C_b(X)$. We have

$$\begin{aligned}
&\mathbb{E}\left[\left|\frac{1}{\Delta T}\int_{I(t,\Delta T)}\left(\langle\mu_{q,\omega}(s), f\rangle - \langle\mu_q(s), f\rangle\right) ds\right|^2\right] \\
&= \frac{1}{(\Delta T)^2}\int_{I(t,\Delta T)}\int_{I(t,\Delta T)} C_{f,f}\left(s, s'\right) ds ds' \\
&= \frac{2}{(\Delta T)^2}\int\int_{\substack{I(t,\Delta T)^2 \\ s\leq s'}} C_{f,f}\left(s, s'\right) ds ds'.
\end{aligned}$$

Assume there exists $\alpha > 0$ such that for all $f \in C_b(X)$ there is a constant $k_f > 0$ of order one in magnitude with the property

$$\left|C_{f,f}\left(s, s'\right)\right| \leq k_f e^{-\alpha|s'-s|}.$$

In such case we get

$$\begin{aligned}
\mathbb{E}\left[\left|\frac{1}{\Delta T}\int_{I(t,\Delta T)}\left(\langle\mu_{q,\omega}(s), f\rangle - \langle\mu_q(s), f\rangle\right) ds\right|^2\right] &\leq \frac{2k_f}{\alpha(\Delta T)^2}\int_{I(t,\Delta T)}\left(1 - e^{-\alpha(s'-s_0)}\right) ds' \\
&\leq \frac{2k_f}{\alpha\Delta T}
\end{aligned}$$

where $s_0 = \inf I(t, \Delta T)$. Therefore the mean square error is small if α is very high.

Let us argue with time-unit of days. Weather loses memory after 10 days in the average, namely $e^{-\alpha|s'-s|}$ is very small when $|s' - s|$ is greater than 10 (days), say. This requires that α is of order $\frac{1}{2}$ (day^{-1}) (maybe smaller but not of order 0.1). If we choose $\Delta T = 40$ (days)

$$\frac{2k_f}{\alpha\Delta T} \sim \frac{1}{10}.$$

It is not a great result, but in the right direction; its poverty is due to the poor scale separation between weather and climate when we want to appreciate climate changes at the seasonal level. If we take several years, a decade say, as a time scale for climate variations (hence neglecting seasonal variations), the result is much improved.

Making rigorous this argument by a scaling limit in a parameter $\epsilon \to 0$ is an open interesting question.

7 Conclusions

Let us summarize here our main temptative definitions of what is climate and what is weather. We have introduced a concept of non autonomous dynamical system $U_{q,\omega}(s, t)$

depending on two external parameters, one deterministic, q, having the meaning of slowly varying external force; and the other random, ω, with the meaning of fast varying external input. We have introduced a concept of minimal time-dependent compact global attractor $A_{q,\omega}(t)$ and a notion of statistical equilibrium, a special class of invariant time dependent probability measures $\mu_{q,\omega}(t)$. By definition, when it exists, the minimal attractor $A_{q,\omega}(t)$ is unique. Not so for $\mu_{q,\omega}(t)$, which is related by pull-back to a given family $\mu_q(t)$ of probability measures, possibly a single given probability measure λ with geometrical or physical interest. Finally, we have introduced the associated Kolmogorov operators $P^q_{s,t}$, also non autonomous, but depending only on q. Correspondingly, we have defined a notion of Markov invariant family of probability measures $\mu_q(t)$. Often, $\mu_q(t)$ is unique. Finally, we have introduced the concept of white noise NADS and proved, in such a case, relations between $\mu_q(t)$ and $\mu_{q,\omega}(t)$, which state on one side that given a Markov invariant $\mu_q(t)$, the associated statistical equilibrium $\mu_{q,\omega}(t)$ exists; and on the other side, that $\mu_q(t)$ is the average of such $\mu_{q,\omega}(t)$.

Let us set ourselves in the following case, satisfied by many examples: that the NADS is white noise, that the compact global attractor $A_{q,\omega}(t)$ exists and that the Markov invariant measure $\mu_q(t)$ exists and is unique. In such case, we define:

Definition 40 *i) the "Weather dynamics" is* $U_{q,\omega}(s,t)$

ii) the "Climate dynamics" is $P^q_{s,t}$

iii) the "Climate" is $\mu_q(t)$

iv) the "potential Weather" is $A_{q,\omega}(t)$

v) the "Weather" is the statistical equilibrium $\mu_{q,\omega}(t)$ *associated to* $\mu_q(t)$*.*

As we have discussed, there are techniques to prove existence of $A_{q,\omega}(t)$ and existence and uniqueness of $\mu_q(t)$. We have however left open two main questions, very relevant for the interpretation of the concepts given in the definition. One is the interpretation of $\mu_q(t)$ as time average of $\mu_{q,\omega}(t)$, better corresponding to reality; we have shortly argued in Section 6.1 that a suitable scaling limit, with a property similar to local equilibrium, should be investigated. The other open question is under which conditions we may claim that the weather $\mu_{q,\omega}(t)$ is "physical", in the sense that comes as pull-back of a relevant probability measure λ. For autonomous stochastic systems on manifold, satisfying however very abstract ellipticity conditions, see [21]; it would be interesting to generalize this result to systems in $\mathbb{R}^d$ with additive noise - as a paradigm of simple and concrete noise - and then to non autonomous systems.

References

[1] V. Araujo and M. J. Pacifico and E. R. Pujals and M. Viana, Singular-hyperbolic attractors are chaotic,

[2] R. Z. Has'minsky, Stochastic stability of differential equations, Sijthoff & Noordhoff, Alphen aan den Rijn 1980.

[3] L. Arnold, Random dynamical systems, Springer-Verlag, Berlin 1998.

[4] L. Arnold, M. Scheutzow, Perfect cocycles through stochastic differential equations. Probab. Theory Related Fields 101 (1995), no. 1, 65–88.

[5] L. Arnold, Hasselmann's program revisited: the analysis of stochasticity in deterministic climate models. Stochastic Climate Models, P. Imkeller and J.-S.von Storch, Eds., Progress in Probability, Vol. 49, Springer 2001, 141–157.

[6] P. H. Baxendale, Statistical equilibrium and two-point motion for a stochastic flow of diffeomorphisms. Spatial stochastic processes, Progr. Probab. 19 (1991), 189-218.

[7] P. H. Baxendale and D. W. Stroock, Large deviations and stochastic flows of diffeomorphisms, Probability Theory and Related Fields 80 (1988), 169-215.

[8] P. Berti and L. Pratelli and P. Rigo, Almost sure weak convergence of random probability measures, Stochastics 78 (2006), 91-97.

[9] P. Billingsley, Convergence of probability measures, John Wiley & Sons, New York 1968.

[10] R. Buizza, The value of Probabilistic Prediction, Atmospheric Science Letters 9 (2008), 36-42.

[11] M. Chekroun, H. Liu, S. Wang, Stochastic Parameterizing Manifolds and Non-Markovian Reduced Equations: Stochastic Manifolds for Nonlinear SPDEs II. Springer, Berlin 2015.

[12] M. Chekroun, E. Simonnet, M. Ghil, Stochastic climate dynamics: random attractors and time-dependent invariant measures. Physica D 240 (2011), 1685-1700.

[13] X. Chen, J. Duan, M. Scheutzow, Evolution systems of measures for stochastic flows. Dyn. Syst. 26 (2011), no. 3, 323–334.

[14] L.-J. Cheng, A. Thalmaier, Evolution systems of measures and semigroup properties on evolving manifolds. Electron. J. Probab. 23 (2018), Paper No. 20, 27 pp.

[15] H. Crauel, Random probability measures on Polish spaces. Stochastics Monographs, 11, Taylor & Francis, London 2002.

[16] H. Crauel, A. Debussche, F. Flandoli, Random attractors. J. Dyn. Differ. Eq. 9 (1997), 307-341.

[17] H. Crauel, F. Flandoli, Attractors for random dynamical systems, Probab. Theory Related Fields 100 (1994), 365-393.

[18] H. Crauel, F. Flandoli, Additive noise destroys a pitchfork bifurcation, J. Dynam. Diff. Eq. 10 (1998), 259-274.

[19] G. Da Prato, A. Debussche, 2D stochastic Navier-Stokes equations with a time-periodic forcing term, J. Dyn. Diff. Eq. 20 (2008), 301-335.

[20] G. Da Prato, J. Zabczyk, Ergodicity for infinite-dimensional systems. London Mathematical Society Lecture Note Series, 229, Cambridge University Press, Cambridge 1996.

[21] D. Dolgopyat and V. Kaloshin and L. Koralov, Sample path properties of the stochastic flows, Ann. Probab. 32 (2004), 1-27.

[22] W. E and J. C. Mattingly, Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation, Comm. Pure Appl. Math. 54 (2001), 1386-1402.

[23] J.-P. Eckmann and D. Ruelle, Ergodic theory of chaos and strange attractors, Reviews of Modern Physics 57 (1985), 617-656.

[24] K. D. Elworthy, Stochastic differential equations on manifolds. London Mathematical Society Lecture Note Series, 70, Cambridge University Press, Cambridge-New York 1982.

[25] F. Flandoli, B. Gess, M. Scheutzow, Synchronization by noise, Probab. Theory Relat. Fields 168 (2017),511–556.

[26] F. Flandoli, B. Gess, M. Scheutzow, Synchronization by noise for order-preserving random dynamical systems, Ann. Probab. 45 (2017), 1325-1350.

[27] S. Galatolo, P. Giulietti, Linear response for dynamical systems with additive noise, Nonlinearity 32 (2019), 2269-2301.

[28] M. Ghil, Natural climate variability, Encyclopedia of Global Environmental Change, Vol. 1, M. C. MacCracken, J. S. Perry and T. Munn Eds., John Wiley and Sons, Chichester 2002.

[29] M. Ghil, The wind driven ocean circulation: applying dynamical system theory to a climate problem, Discr. Cont. Dyn. Syst. 37 (2017), 189-228.

[30] A. Gritsun, V. Lucarini, Fluctuations, response and resonances in a simple atmospheric model, Physica D: Nonlin. Phen. 349 (2017), 62–76.

[31] K. Hasselmann, Stochastic climate models. Part I. Theory. Tellus 28 (1976), 473–485.

[32] D. Holm, Variational principles for stochastic fluid dynamics, Proc. R. Soc. A 471 (2015), 20140963, 19 pp.

[33] Intergovernmental Panel on Climate Change (IPCC), Chair vision paper, AR6, 2017.

[34] B. Jakubczyk, Introduction to geometric nonlinear control; controllability and Lie bracket. Mathematical control theory, ICTP Lect. Notes, VIII, Abdus Salam Int. Cent. Theoret. Phys., Part 1, 2 (Trieste, 2001), pp. 107-168, 2002.

[35] I. Karatzas and S. Shreve, Brownian motion and stochastic calculus. Second edition. Graduate Texts in Mathematics, 113, Springer-Verlag, New York 1991.

[36] Y. Kifer, Ergodic theory of random transformations, Birkhäuser, Boston 1986.

[37] P. Kloeden, C. Pötzsche (Eds.), Nonautonomous Dynamical Systems in the Life Sciences, LNM 2102, Springer 2013.

[38] H. Kunita, Stochastic flows and stochastic differential equations, Cambridge University Press, Cambridge 1990.

[39] Y. Le Jan, Équilibre statistique pour les produits de diff\'eomorphismes aléatoires indépendants, Ann. Inst. H. Poincaré Probab. Statist. 23 (1987), 111-120.

[40] V. Lucarini, Stochastic perturbations to a dynamical system: a response theory approach. J. Stat. Phys., 14 (2012), 774-786.

[41] V. Lucarini, Response Operators for Markov Processes in a Finite State Space: Radius of Convergence and Link to the Response Theory for Axiom A Systems, J. Stat. Phys. 162 (2016), 312-333.

[42] V. Lucarini„ D. Faranda, A. Freitas, J. Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd, S. Vaienti, Extremes and Recurrence in Dynamical Systems, Wiley, New York 2016.

[43] V. Lucarini1, F. Ragone, F. Lunkeit, Predicting Climate Change Using Response Theory: Global Averages and Spatial Patterns, J. Stat. Phys. 166 (2017), 1036-1064.

[44] A. J., Majda, I. Timofeyev, E. Vanden-Eijnden, A mathematical framework for stochastic climate models, Comm. Pure Appl. Math. 54 (2001), 891-874.

[45] N. Masmoudi and L.-S. Young, Ergodic theory of infinite dimensional systems with applications to dissipative parabolic PDEs, 2001.

[46] J. C. Mattingly, A. M. Stuart, D. J. Higham, Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise, Stochastic Process. Appl. 101 (2002), 185-232.

[47] Neumann and Ulam, Random Ergodic Theorems, Bull. Amer. Math. Soc. 51 (1945).

[48] J. Norris, Simplified Malliavin calculus, Séminaire de probabilités (Strasbourg) 20 (1986),101-130.

[49] S. Pierini, M. Ghil, M. D. Chekroun, Exploring the pullback attractors of a low-order quasigeostrophic ocean model: The deterministic case, Journal of Climate, 2016

[50] S. Pierini, M. Ghil, M. D. Chekroun, The onset of chaos in nonautonomous dissipative dynamical systems: a low-order ocean-model case study, Nonlinear Processes Geophys. 25 (2018), 671-692.

[51] M. Romito, Ergodicity of the finite dimensional approximation of the 3D Navier-Stokes equations forced by a degenerate noise, J. Statist. Phys. 114 (2004), 155-177.

[52] B. Schmalfuss, Backward cocycles and attractors of stochastic differential equations, in V. Reitmann, T. Riedrich, N. Koksch Eds., Int. Sem. Appl. Math. 1992, 185-192.

[53] J. Seidler, Ergodic behaviour of stochastic parabolic equations, Czechoslovak Math. J. 47 (1997), 277-316.

[54] D. W. Stroock, Some applications of stochastic calculus to partial differential equations, Lecture Notes in Math. 976, Springer, Berlin 1983.

[55] D. W. Stroock and S. R. S. Varadhan, On the support of diffusion processes with applications to the strong maximum principle, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., Vol. III: Probability theory, 1970/1971, 333-359.

[56] R. Temam, Infinite-dimensional dynamical systems in mechanics and physics. Applied Mathematical Sciences 68, Springer-Verlag, New York 1997.

[57] R. Temam, Navier-Stokes equations and nonlinear functional analysis. Regional Conference Series in Applied Mathematics 66, Society for Industrial and Applied Mathematics (SIAM), Philadelphia 1995.

[58] L.-S. Young, What are SRB measures, and which dynamical systems have them? J. Statist. Phys. 108 (2002), 733-754.

[59] J. Zabczyk, Mathematical control theory: an introduction, Systems and Control: Foundations and Applications. Birkhäuser Boston, Boston 1992.