Snapshots of modern mathematics
from Oberwolfach

№ 2/2024

Randomness is natural - an introduction to regularisation by noise

Ana Djurdjevac [1] • Henri Elad Altman [2]
Tommaso Rosati

Differential equations make predictions on the future state of a system given the present. In order to get a sensible prediction, sometimes it is necessary to include randomness in differential equations, taking microscopic effects into account. Surprisingly, despite the presence of randomness, our probabilistic prediction of future states is stable with respect to changes in the surrounding environment, even if the original prediction was unstable. This snapshot will unveil the core mathematical mechanism underlying this "regularisation by noise" phenomenon.

[1] The research of the first author was funded by Deutsche Forschungsgemeinschaft (DFG) through the grant CRC 1114: Scaling Cascades in Complex Systems, Project Number 235221301.
[2] Work on this Snpashot started when the second author was Dirichlet Postdoctoral Fellow at Freie Universität Berlin. His research was funded by the BMS, Forschungsprojekt des Exzellenzclusters MATH+ EXC 2046.

1 A first example

Newton's laws of motion. Imagine a perfectly spherical ball, positioned at the exact top of a symmetric hill, like in Figure 1. Under the action of gravity, no matter how slightly we move the ball, it will fall down one of the two sides of the hill, and – depending on the strength of friction – settle in one of the neighbouring valleys. How well can we predict where the ball will end up?

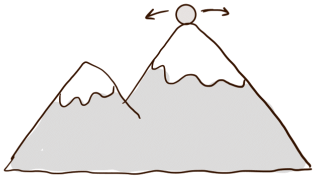

Figure 1: A ball on a hill.

A first, naïve idea would be to apply Newton's laws of motion, so to approximate the final position of the ball, given the exact initial state, which includes the mass of the ball, the strength of friction, and much more. This would bring us to the conclusion that the ball will never leave its location.

But this approximation is contrary to our intuition, and it does not account for effects such as a small whiff of wind: We intuitively understand that the ball at the top of the hill finds itself in a precarious state and any ever-so-small impurity or force could be enough to set it rolling down either of the two slopes.

A more reasonable expectation on the final state is that with a 50% chance the ball will end up at the bottom of the valley on its left and with the remaining 50% chance it will end up at the bottom of the valley on its right, always assuming that friction is sufficiently strong, so that the ball does not fall into an even further valley. We have passed from a *deterministic* prediction on the location or our ball, to a *probabilistic* one, which better fits our expectations.

We encode those small, unexpected changes in the surrounding environment as a random force, called *noise* or *stochastic term*. In the next section, we will gradually build up a mathematical formulation for this, in the language of *stochastic differential equations*.

2 Differential equations

We start by exploring the movement of the ball in the language of *ordinary differential equations*. The study of ordinary differential equations (ODEs) dates back several centuries, and such systems find an enormity of applications, from the description of trajectories of planets to the evolution of a population in a Petri dish. In the language of ODEs, we describe the velocity of the little ball as the time derivative of its position $x(t) \equiv x_t \in \mathbb{R}^d$ at time t. Denoting by $V(x)$ the height of the hill at position x, we can write the basic movement of the ball, which rolls downwards the slope $\partial_x V(x)$ of the hill, as

$$\frac{\mathrm{d}x_t}{\mathrm{d}t} = -\partial_x V(x_t) \; . \tag{1}$$

Let us rewrite this equation of motion in a slightly more abstract way, and let us speak of general particles $x_t \in \mathbb{R}^d$ moving in a velocity field $\varphi\colon \mathbb{R}^d \to \mathbb{R}^d$, often called *drift*, so $\frac{\mathrm{d}x_t}{\mathrm{d}t} = \varphi(x_t)$. Usually, especially in physics, a different way of writing this equation, using *differentials*,

$$\mathrm{d}x_t = \varphi(x_t)\,\mathrm{d}t \tag{2}$$

is common, and we adopt it here to align with its usage in the stochastic case that we will describe in Section 3.

Existence and uniqueness of solutions

A fundamental mathematical question is whether an ODE admits solutions and, if so, whether the solution is unique. The answer to this question depends heavily on the *regularity* of the drift φ, i.e. how rapidly φ changes its value over time, or more generally over its domain. A first simple distinction between different levels of regularity is to separate functions whose graph is a line without interruption, the *continuous functions*, from functions that have jumps. But also continuous functions can be quite wild, for example when they are *oscillating*, in the sense that they change values abruptly.

For our ODE problem, a good class is provided by *Lipschitz continuous* functions: a function $\varphi : \mathbb{R} \to \mathbb{R}$ is said to be Lipschitz continuous (or just Lipschitz for short) if for some constant $L > 0$

$$\frac{|\varphi(x) - \varphi(y)|}{|x - y|} \leq L \; , \qquad \forall x, y \in \mathbb{R}, \quad x \neq y \; .$$

Hence, for any Lipschitz function φ, there is a limitation on how fast it can change. The limitation is given by the Lipschitz constant L which describes the most dramatic change that the function φ may undergo. If φ is differentiable, then we can determine the Lipschitz constant easily as $L = \max_{x\in\mathbb{R}} |\varphi'(x)|$. The function $\varphi(x) = \sqrt{|x|}$ is not Lipschitz, as it becomes steeper and steeper the closer $|x|$ is to zero, see Figure 2.

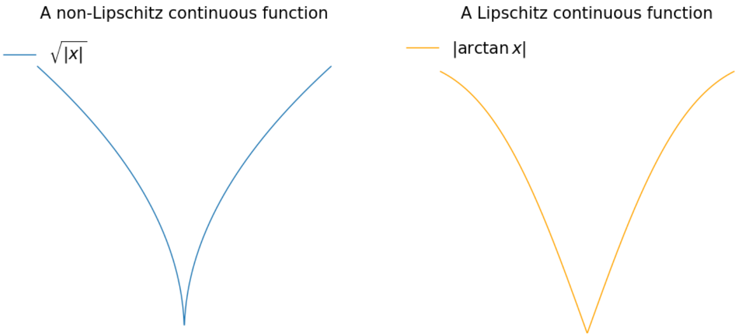

Figure 2: Examples of a non-Lipschitz and a Lipschitz function

Equipped with Lipschitz continuous functions, we can go back to the question of the existence and uniqueness of the solution of Equation (2), the Picard–Lindelöf Theorem. This classical result discusses the *Cauchy problem*, namely the solvability of a differential equation for a prescribed *initial condition*, such as the evolution of a ball starting from a specified point on the top of a hill, with a specified velocity.

Theorem 2.1 (Picard–Lindelöf) *If φ is Lipschitz continuous, then, for any initial condition $\bar{x} \in \mathbb{R}^d$, the Cauchy problem*

$$\mathrm{d}x_t = \varphi(x_t)\,\mathrm{d}t, \qquad x_0 = \bar{x}$$

admits a unique solution, defined over all $t \geq 0$. The solution depends continuously on the initial datum $\bar{x}$.

The Picard–Lindelöf theorem tells us that the initial value problem is *well-posed*, namely, for every initial datum there exists a unique solution to the differential equation and this solution is stable, i.e. it depends continuously on the initial datum. Let us investigate now if we can allow also less regular functions φ, namely those which are only continuous. This is the setting of Peano's theorem:

Theorem 2.2 (Peano) *If φ is continuous and bounded, then the Cauchy problem*

$$\mathrm{d}x_t = \varphi(x_t)\,\mathrm{d}t, \qquad x_0 = \bar{x}$$

admits a solution.

The assumption of a continuous drift φ in Peano's theorem ensures the existence of a solution, but not its uniqueness, as was the case with Lipschitz continuity in the Picard–Lindelöf theorem. We can see this at the example of the ODE

$$\mathrm{d}x_t = \operatorname{sign}(x_t)\,|x_t|^\alpha\,\mathrm{d}t, \qquad x_0 = 0$$

where α is a parameter in $(0,1)$ and $\operatorname{sign}(x)$ denotes the sign of x, which is the function that is 1 when x is positive, -1 when x is negative, and taken to be equal to 0 if $x = 0$. Clearly, $x_t = 0$ is a solution to the equation. But one can check that $x_t := ((1-\alpha)t)^{\frac{1}{1-\alpha}}$ is a solution as well, and so is $x_t := -\left((1-\alpha)t\right)^{\frac{1}{1-\alpha}}$.

In addition, for all $t_0 \geq 0$, $t \geq 0$ defined by

$$x_t := ((1-\alpha)(t-t_0)_+)^{\frac{1}{1-\alpha}} = \begin{cases} 0 & \text{if} \quad t < t_0 \\ ((1-\alpha)(t-t_0))^{\frac{1}{1-\alpha}} & \text{if} \quad t \geq t_0 \end{cases}$$

is also a solution. We see therefore that there are infinitely many different solutions to the same Cauchy problem, see Figure 3: we call this a *Peano phenomenon*.

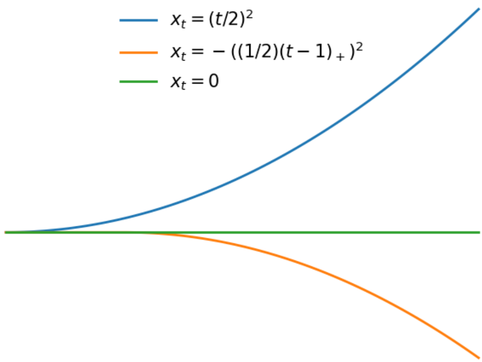

Figure 3: Several solutions to the same Cauchy problem $\mathrm{d}x_t = \operatorname{sign}(x_t)\sqrt{|x_t|}\,\mathrm{d}t$ with $x_0 = 0$: a *Peano phenomenon*.

Before we continue with the stochastic case, let us discuss the assumptions of Peano's theorem more closely: The boundedness assumption may be further weakened, at the expense of having solutions which may only exist for a finite time. But the continuity of φ is a necessary assumption, we cannot ask for less. To see this, we integrate both sides of Equation (2) to avoid the usage of derivatives. This gives us the integral formulation of the ODE

$$x_t = x_0 + \int_0^t \varphi(x_s)\,\mathrm{d}s\,. \tag{3}$$

If φ is not continuous, it is not clear how to define the integral $\int_0^t \varphi(x_s)\, ds$ in Equation (3), and existence may actually not hold.

3 Stochastic differential equations

Let us now study what happens if we perturb our original ODE by adding a forcing term, also called a *stochastic process* or *noise*, denoted by $(w_t)_{t\geq 0}$.

A common choice of noise is *Brownian motion*. Brownian motion was first discovered in 1827 by the botanist Robert Brown, who was observing particles within a grain of pollen suspended in water. We omit the formal definition of Brownian motion, referring the interested reader to the book [5], but it can be thought of as the zoomed-out trajectory of a particle moving right-up or right-down with probability 1/2 at each step: as we zoom out more and more, this saw-tooth-like trajectory takes the form of a wiggly function, which, albeit continuous, is not Lipschitz continuous, see Figure 4.

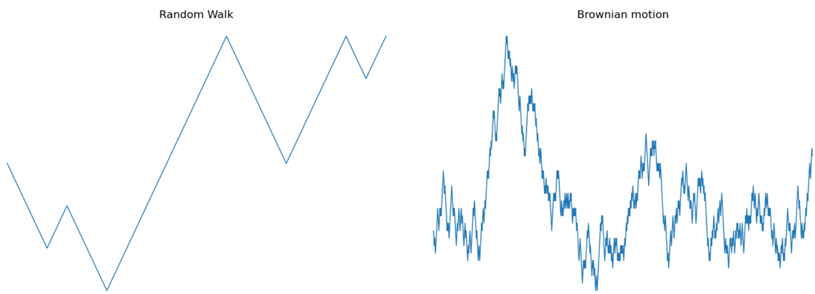

Figure 4: Brownian motion is the limit of a random walk as we “zoom out”, see Section 5.

We now have all ingredients to formulate our initial example of the ball on the top of the hill in the language of stochastic differential equations. Using the infinitesimal notation from Equation (2), we can write a more accurate version of Equation (1) of the movement of the little ball as

$$\underbrace{\mathrm{d}x_t}_{velocity} = -\underbrace{\partial_x V(x_t)}_{landscape}\,\mathrm{d}t + \varepsilon\,\underbrace{\mathrm{d}B_t}_{noise}\ . \tag{4}$$

In this new equation, the velocity of the ball, or, more precisely, the instantaneous variation of the position x_t of the ball, follows the slope $\partial_x V(x_t)$ downwards and is perturbed in that instant by a random force $\mathrm{d}B_t$, where B stands for Brownian motion. The effect of the noise is multiplied with a parameter

$0 < \varepsilon \ll 1$, which indicates that the effect is very small. Such a very small and maybe undetectable perturbation is often called *microscopic* in the literature.

The effect of noise

As suggested by the intuition we described at the beginning of this exposition, if a ball is positioned at the top of a hill and is perturbed by some random velocity, then we expect to find the ball in either one of the two bottom valleys next to the peak from which it started, with roughly 50% chance. This mechanism appears also in the study of differential equations. Indeed, the infinitely many solutions depicted in Figure 3 and the preceding discussion describe the position of a particle x_t which sits at level $x = 0$ for an arbitrary amount of time t_0, before escaping either to the upper or to the lower region. If we add noise to the differential equation describing x_t, the particle will not be able to sit indefinitely at zero. Instead, it will be instantaneously kicked out, either towards the top or towards the bottom: we have ruled out all non-physical solutions. Bafico and Baldi [1] were among the first to observe this remarkable effect, and they were able to prove that if the noise is small, then the probability with which the particle immediately moves up or down is exactly 50%, as described by our original intuition. This is one of the first, simplest, and most striking instances of regularisation by noise.

4 Restoring well-posedness via additive noise

In this section, we will unveil how adding a stochastic term as irregular as Brownian motion can actually make a solution to a differential equation well-defined, in the sense that solutions exist and are unique, even when the deterministic counterpart is not: this is what we call a *regularisation by noise* phenomenon.

Let us consider a stochastic differential equation with a continuous drift function φ and very fluctuating noise term w_t, for example Brownian motion.

$$\mathrm{d}x_t = \varphi(x_t)\,\mathrm{d}t + \mathrm{d}w_t, \qquad x_0 = x, \tag{5}$$

At first, adding noise does not seem to make our problem any simpler: if φ is not a Lipschitz function, so not regular enough to ensure well-posedness of the problem, adding another term just as bad will not make the situation better. However, we will see that a careful choice of noise may induce an *averaging effect* that may restore well-posedness of the equation.

How does noise regularise?

To explain the kind of averaging that will smoothen (or regularise) the drift φ, let us assume that φ is the discontinuous function defined by $\varphi(x) = 1$ if $x \geq 0$

and $\varphi(x) = 0$ if $x < 0$. We can rewrite Equation (5) in integral form as

$$x_t = x_0 + \int_0^t \varphi(x_s)\,\mathrm{d}s + w_t \ .$$

Setting $y_t := x_t - w_t$, we may in turn rewrite this as

$$y_t = x_0 + \int_0^t \varphi(y_s + w_s)\,\mathrm{d}s \ .$$

Let us now study the function

$$x \mapsto \int_0^t \varphi(x + w_s)\,\mathrm{d}s \ . \tag{6}$$

A priori, because of the discontinuity of φ, we would guess that this function would be discontinuous. However, for sufficiently "fluctuating" noise w, such as Brownian motion, it turns out that $x \mapsto \int_0^t \varphi(x + w_s)\,\mathrm{d}s$ will be quite regular, actually even Lipschitz continuous, see Figure 5. This is because, almost

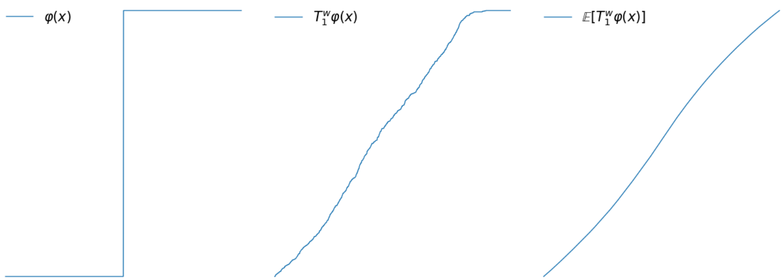

Figure 5: From left to right, a deterministic function $\varphi(x)$, its randomly regularised version $\int_0^1 \varphi(x+B_s)\,\mathrm{d}s$ and averaged version $\mathbb{E}\int_0^1 \varphi(x+B_s)\,\mathrm{d}s$.

instantaneously, w will explore the entire area surrounding its starting point, thereby "averaging out" the discontinuity of φ. It is like a staircase whose steps are being levelled off under a snow storm: when the storm is over, the discontinuities are covered with snow, and you will be able to slide down the slope without any problem.

Getting back to Equation (5), we can use this enhanced continuity induced by the noise to prove well-posedness of the problem at hand. The rule of thumb is: provided that the noise w "fluctuates" enough, so that the regularity of $x \mapsto \int_0^t \varphi(x + w_s)\,\mathrm{d}s$ is sufficiently enhanced, well-posedness of Equation (5) can be restored.

But let's not deceive ourselves by the word "regularising effect": while we gain well-posedness of an ODE for which we may previously have had multiple solutions (no uniqueness) or no solution at all, it does by no means indicate that sample paths of solutions are smoother as functions of time: in fact, the exact opposite is true.

How do we measure the regularising effect of noise?

What do we mean when we say that the noise $(w_t)_{t\geq 0}$ has to "fluctuate sufficiently"? One way to quantify fluctuations, or irregularity, is to consider the *time spent* by the process $(w_t)_{t\geq 0}$ at a given location. We can prove that for a wide class of random noises w such as Brownian motion, it is possible to define, for all $a \in \mathbb{R}$, a number $L_t^w(a)$ quantifying the amount of time spent by w at a, before time t, called a *local time*. It satisfies, for all functions φ, the relation

$$\int_0^T \varphi(w_s + x)\, ds = \int_{\mathbb{R}} \varphi(a + x) L_T^w(a)\, \mathrm{d}a \ . \tag{7}$$

Intuitively, this means that we can assign to each location $a \in \mathbb{R}$ a clock that ticks every time the process $(w_t)_{t\geq 0}$ visits a. If the process $(w_t)_{t\geq 0}$ fluctuates much, so that it explores the space evenly, the times $L_t(a)$ and $L_t(b)$ displayed at two nearby points a and b will be close. This heuristics can be made quantitative as follows. Let C^r denote the space of r times differentiable functions. The noise w is said to be r-regularising if, for all $t > 0$, $a \mapsto L_t^w(a)$ is of class C^r. If w satisfies this property with $r = \infty$, we say that it is infinitely regularising. Let us explain how the regularising noise can be of any help to our problem. We first recall the notion of convolution, which has the property of smoothing out irregular functions.

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a function and w be an r-regularising noise, for some $r \geq 0$. Then the *convolution* of L_T^w and φ is the function $L_T^w * \varphi$ defined by

$$(L_T^w * \varphi)(x) := \int_{\mathbb{R}} L_T^w(a)\, \varphi(x + a)\, \mathrm{d}a \ , \tag{8}$$

whenever the integral is defined.[3] In the right-hand side above, we recognise the quantity of Equation (7). From the properties of the convolution, as soon as $a \mapsto L_t^w(a)$ is of class C^r, we deduce that the function $x \mapsto \int_0^T \varphi(w_s + x)\, \mathrm{d}s$ is of class C^r as well, even if φ is a very wild function. In particular, if w is infinitely regularising, the function $x \mapsto \int_0^T \varphi(w_s + x)\, \mathrm{d}s$ is smooth. We stress that this

[3] For simplicity, the definition of convolution used here differs slightly from the usual one which, if adopted, would require to replace the left-hand side of Equation (8) by $L_T^w * \check{\varphi}$, where $\check{\varphi}(x) = \varphi(-x)$.

will be independent of the continuity of φ and indeed this approach allows to choose highly irregular, so called *generalised* functions, for example the Dirac function which may be loosely defined as a function vanishing away from 0 and integrating to 1. The fact that $x \mapsto \int_0^T \varphi(w_s + x)\,\mathrm{d}s$ is more regular than the original function φ is a key observation to prove that adding regularising noise to our original ODE restores well-posedness.

5 Examples and outlook

Where do we find regularising paths?

Constructing a path with the mentioned regularising properties with bare hands is quite challenging. Recent works have shown that regularising paths must be rough, to an amount that can be quantified precisely, see [3, Corollary 74]. Conversely, identifying paths having the required regularising property has been a long-standing open problem. One way to circumvent this difficulty consists, instead of trying to construct a fixed regularising path, in using probability theory to obtain such a path by picking it at *random*.

A typical example is obtained by sampling the path w as a Brownian motion, see Figure 4. Then one can prove that w possesses the nice property of being regularising with $r = 1/2 - \varepsilon$ for any $\varepsilon > 0$, where the meaning of fractional regularity is that a function is "half-way" between merely continuous and differentiable, such as for example $\varphi(x) = \sqrt{|x|}$.

Universality of Brownian motion

Besides having the desired regularising effect, Brownian motion turns out to be commonly observed in real-life problems, because of its so-called *universality*. In a nutshell, coming back to our description in terms of the zoomed-out trajectory of a microscopic particle performing a random walk, it does not matter how precisely our particle moves: as long as on average it jumps as much up as it does down, and as long as the choice of the first jump is chosen independent of the jump at a distant, later time, we will always see a Brownian motion when we zoom out. Of course, we have to zoom out at an appropriate speed. If X_t is the position of the particle at time $t \geqslant 0$ (say started at $X_0 = 0$), then after a time $n \gg 1$ the particle will have reached a distance of order $\sqrt{n}$: it is the same scaling as for the central limit theorem[4], which is closely linked to this derivation of Brownian motion. We obtain that, as $n \to \infty$, $\frac{1}{\sqrt{n}} X_{nt} \to B_t$, where B_t is a Brownian motion.

[4] See [2, Chapter 3] for a detailed exposition.

Infinitely regularising paths

Can one construct a noise that regularises better than Brownian motion? For example, if we take any $r \geq 0$, potentially very large, can we construct a noise that is r-regularising? The answer is positive, and one can even construct a noise that is *infinitely* regularising. We do not describe how such a noise is constructed, but just mention that it can be obtained as some "rougher version" of a Brownian motion, and refer the interested reader to [4] for more details.

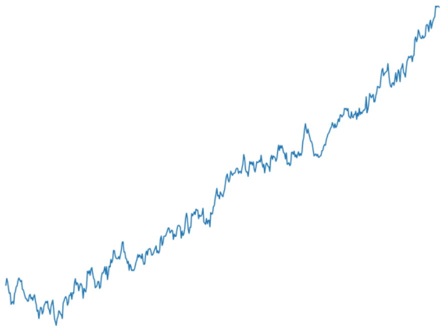

Figure 6: An infinitely regularising noise.

References

[1] R. Bafico and P. Baldi, *Small random perturbations of Peano phenomena*, Stochastics **6** (1981/82), no. 3-4, 279–292.

[2] R. Durrett, *Probability theory and examples*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49.

[3] L. Galeati and M. Gubinelli, *Prevalence of ρ-irregularity and related properties*, Arxiv Preprint, arXiv:2004.00872v2 (2020).

[4] F. A. Harang and N. Perkowski, *C^∞-regularization of ODEs perturbed by noise*, Stoch. Dyn. **21** (2021), no. 8.

[5] P. Mörters and Y. Peres, *Brownian motion*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 30.

Ana Djurdjevac *is a junior professor at Freie Universität Berlin.*

Henri Elad Altman *is maître de conférence at Université Sorbonne Paris Nord.*

Tommaso Rosati *is a Harrison early career assistant professor at the University of Warwick.*

Mathematical subjects
Probability Theory and Statistics

Connections to other fields
Chemistry and Earth Science, Physics

DOI
10.14760/SNAP-2024-002-EN

Snapshots of modern mathematics from Oberwolfach provide exciting insights into current mathematical research. They are written by participants in the scientific program of the Mathematisches Forschungsinstitut Oberwolfach (MFO). The snapshot project is designed to promote the understanding and appreciation of modern mathematics and mathematical research in the interested public worldwide. All snapshots are published in cooperation with the IMAGINARY platform and can be found on www.imaginary.org/snapshots and on www.mfo.de/snapshots.

ISSN 2626-1995

Mathematisches Forschungsinstitut
Oberwolfach gGmbH
Schwarzwaldstr. 9 – 11
77709 Oberwolfach
Germany

Director
Gerhard Huisken

Mathematisches Forschungsinstitut Oberwolfach

Member of Leibniz Association

IMAGINARY open mathematics