Snapshots of modern mathematics
from Oberwolfach

№ xx/xxxx

Snapshot on regularization by noise

Ana Djurdjevac [1] • Henri Elad Altman
Tommaso Rosati

Regularization by noise is a term which refers to different phenomena, appearing in areas ranging from fluid dynamics to artificial intelligence. The common thread is that the analysis of certain deterministic systems is mathematically extremely challenging or computationally very expensive. Sometimes a perturbation of the system through the action of a noise simplifies the mathematical analysis or diminishes the computational cost of the problem. In many examples this simplification is related to a regularizing effect of the noise. In a nutshell, if we do not take into account the effect of noise, our deterministic prediction of a future state can vary dramatically based on our current knowledge. Instead, in the presence of noise, our probabilistic prediction of future states is stable with respect to changes in the surrounding environment.

[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.

1 A toy example

Newton's laws of motion. Imagine a perfectly spherical ball, positioned at the exact top of a symmetric hill (see Figure 1).

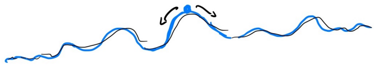

Figure 1: A ball on a hill.

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 neighboring valleys. How well can we predict where the ball will end up?

Following Newton's laws, we can approximate the finial position of the ball within the limits of computational power, provided we are given the exact initial state (including the mass of the ball, the strength of friction and so on). But this approximation 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 microscopic impurity or force (here microscopic could simply mean so small that our measurement devices cannot perceive it) could be enough to set it rolling down either of the two slopes.

The prediction that our precariously positioned ball will never leave its location, which follows from a naïve application of Newton's laws, is now contrary to our intuition. 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.

Does noise help? One way to motivate our probabilistic estimate is to assume that there is a random force, or noise, acting on our system, capturing small unexpected changes in the surrounding environment. The correct language to formulate this assumption is that of stochastic differential equations:

$$\underbrace{\mathrm{d}x_t}_{velocity} = -\underbrace{\partial_x V(x_t)}_{landscape}\,\mathrm{d}t + \varepsilon \underbrace{\mathrm{d}B_t}_{\color{red}{noise}}\ . \tag{1}$$

This is an infinitesimal notation, which indicates that the instantaneous variation of the position x_t of the ball (its velocity) follows downwards the slope $\partial_x V(x_t)$ of the landscape – here $V(x)$ is the height of the hill at position x – and is perturbed in that instant by a random force $\mathrm{d}B_t$ (the B stands for Brownian motion, typically the most natural choice of noise: see the discussion in later sections). The effect of the noise is multiplied with a parameter $0 < \varepsilon \ll 1$, which indicates that the effect is *small*, in our running example smaller than what can possibly be perceived by our measuring device.

Then, after a time of order $\log(\varepsilon^{-1})$ one can hope to find the ball in either one of the two bottom valleys with roughly 50% chance, as should be expected from the intuition we described. The time is of logarithmic order because the ball will leave its unstable position at an *exponential* speed: note that while ε^{-1} might be too long a time for us to wait, $\log(\varepsilon^{-1})$ is enormously smaller! This is one of the first and simplest instances of regularization by noise: a result of this kind has been obtained by Bafico and Baldi [1] in the study of Peano phenomena, which we describe in the upcoming sections.

2 Differential equations

Let us take a step back and start from the analysis of differential equations. Typically ordinary differential equations (ODEs) describe the velocity – that is the time derivative – of the position $x_t \in \mathbb{R}^d$ of a particle at time t:

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

where $\varphi \colon \mathbb{R}^d \to \mathbb{R}^d$ is a function representing the velocity field or drift in which a particle moves. A convenient way to rewrite the above equation is by using differentials:

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

This notation should be understood in its integral form

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

The study of 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. 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. The study of regularity aims at quantifying how rapidly a function changes its value (perhaps oscillating) over time, or more generally over its domain. A first simple distinction between different levels of regularity is to separate continuous functions (whose graph is a line without interruption) from functions that have jumps, such as the one given by $\varphi(x) = 1$ for $x > 0$ and $\varphi(x) = 0$ for $x \leq 0$. In order to measure regularity, we need to choose some notion of distance on the domain of the function.

A function $\varphi : \mathbb{R} \to \mathbb{R}$ is said to be Lipschitz continuous (or just Lipschitz for short) if for some constant $L > 0$

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

In the above, $|x - y|$ measures the distance between x and y. Hence, for any Lipschitz function φ there is a limitation on how fast it can change, the Lipschitz constant L describing the most dramatic change that the function φ may undergo. If φ is differentiable, then $L = \max_{x\in\mathbb{R}} |\varphi'(x)|$. An example of a Lipschitz function is $\sin x$, while $\varphi(x) = \sqrt{|x|}$ is not Lipschitz.

Now we can go back to the question of the existence and uniqueness of the solution of (2). One exceptionally important result is the following theorem.

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

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

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

If for every initial datum there exists a unique solution to the differential equation and this solution depends continuously on such initial datum (stability), then we say that the equation is **well-posed**. We may relax the assumptions of the above, at the cost of losing uniqueness and stability, and thus well-posedness, of solutions.

Theorem 2.2 (Peano) *If φ is continuous and bounded [2], then the Cauchy problem*

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

admits a solution.

There is a wide gap between the requirement of the two results above: let us see why this is the case. When the drift φ becomes irregular things can go wrong

[2] the boundedness assumption may actually be relaxed, at the expanse of having solutions which may blow up in finite time

in two different ways. First, if φ is not even continuous, it is not clear how to define the integral $\int_0^t \varphi(x_s)\, ds$ in the first place, and existence may actually not hold.

A second issue arises if φ is continuous but not Lipschitz, so that a solution exists by Peano's theorem, but uniqueness may fail to hold. Consider for example

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

where α is a parameter in $(0,1)$ and $\mathrm{sign}(x)$ denotes the sign of x (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 $x_t := -((1-\alpha)t)^{\frac{1}{1-\alpha}}$. 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 is an uncountable family of different solutions to the same Cauchy problem, see Figure 2.

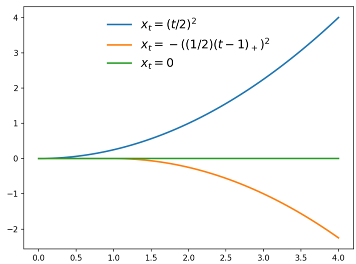

Figure 2: Several solutions to the same Cauchy problem $\mathrm{d}x_t = \mathrm{sign}(x_t)\sqrt{|x_t|}\,\mathrm{d}t$ with $x_0 = 0$.

3 Restoring well-posedness via additive noise

Let us now study what happens if we perturb our original ODE by adding a forcing term, or **noise** $(w_t)_{t\geq 0}$. In the spirit of (1), let us consider the equation

$$\mathrm{d}x_t = \varphi(x_t)\,\mathrm{d}t + \mathrm{d}w_t, \qquad x_0 = x, \tag{3}$$

where we have added a noise term that is very fluctuating. Adding a noise at first does not seem to make our problem any simpler. However, in cases in which

φ is not a Lipschitz function, so well-posedness of the original problem typically breaks down, adding such a noise induces an *averaging effect* that may restore well-posedness of the equation. This is an instance of *regularization by noise*. As we will attempt to explain, the more irregular the noise is, the stronger is its regularizing effect on the equation. To avoid confusion let us therefore stress once more that the regularising effect of the noise leads to 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.

3.1 Where do we find regularising paths?

Constructing a path with the mentioned regularising properties with bare hands is quite challenging. However, using probability theory, one can obtain such paths by choosing them at *random*.

One standard choice is Brownian motion. We omit the formal definition of this process (see [4]), 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 sawtooth-like trajectory takes the form of a wiggly function, which albeit continuous is not Lipschitz continuous. Brownian motion was first discovered in 1827 by the botanist Robert Brown who was observing particles within a grain of pollen suspended in water. Such a process is quite irregular, in particular, almost-surely, it is nowhere Lipschitz continuous, see Figure 3 and the upcoming discussion. The reason why Brownian motion is so commonly observed lies in its *universality*. In a nutshell, universality means that it does not matter how exactly our microscopic particle moves: as long as on average it jumps as much up as it does down, and as long as the jump between time n and time $n+1$ is roughly chosen independent of the jump between time 0 and time 1, for $n \gg 1$, 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 [2, Chapter 3], which is closely linked to this derivation of Brownian motion. We obtain that

$$\frac{1}{\sqrt{n}} X_{nt} \to B_t \ , \qquad \text{as } n \to \infty \ ,$$

where B_t is a Brownian motion. An interesting consequence of this derivation is that we should expect that $B_t - B_s \simeq \sqrt{t-s}$ for all $0 \leqslant t-s \ll 1$. We have thus constructed a function that is *nowhere* Lipschitz, in contrast to the square-root function, which is not Lipschitz only near the origin. It is exactly this kind of extreme irregularity (observe that it's not easy to build an example with the same property by hand, without choosing it randomly), which guarantees the regularization property of the noise.

What are the mathematical tools at our disposal to *quantify the irregularity* of a path? And are there other examples of regularizing paths? These are questions that we will attempt to answer over the next sections.

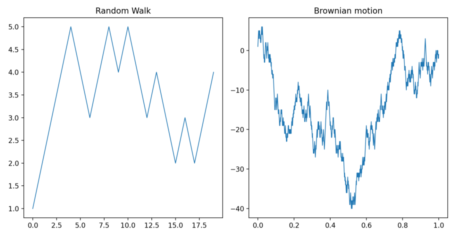

Figure 3: Brownian motion is the limit of a random walk as we "zoom out".

3.2 How does noise regularise?

When an irregular noise explores the area surrounding a point, it will induce a form of averaging that will smoothen (or regularise) the drift φ. To explain this phenomenon, let us for instance 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 (3) 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

$$x \mapsto \int_0^t \varphi(x + w_s)\,\mathrm{d}s\ . \tag{4}$$

A priori, because of the discontinuity of φ, we would guess that this function would be discontinuous. However, for a noise w that is sufficiently "fluctuating", 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 4. This is because, almost 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 our equation (3), 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 is "fluctuating" 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 (3) can be restored.

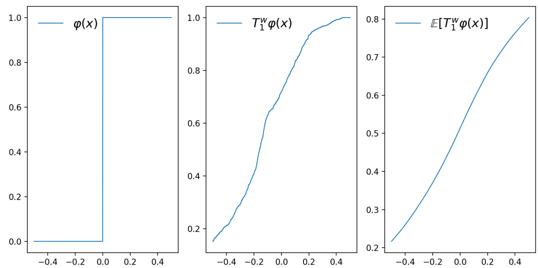

Figure 4: From left to right, a deterministic function $\varphi(x)$, its randomly regularized version $\int_0^1 \varphi(x+B_s)\,\mathrm{d}s$ and averaged version $\mathbb{E}\int_0^1 \varphi(x+B_s)\,\mathrm{d}s$.

3.3 How do we measure the regularising effect of a noise?

What do we mean when we say that a noise $(w_t)_{t\geq 0}$ has to be "sufficiently fluctuating"? 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. Such a *local time* satisfies, for all function φ, the relation

$$\int_0^T \varphi(w_s+x)\,ds = \int_{\mathbb{R}} \varphi(a+x) L_T^w(a)\,\mathrm{d}a\,. \tag{5}$$

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^w_T and φ is the function $L^w_T * \varphi$ defined by

$$(L^w_T * \varphi)(x) := \int_{\mathbb{R}} L^w_T(a)\, \varphi(x+a)\, \mathrm{d}a\ , \tag{6}$$

whenever the integral is defined [3]. In the right-hand side above we recognize the quantity of (5). From the properties of convolution, as soon as $a \mapsto L^w_t(a)$ 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 will be independent of the continuity of φ and indeed this approach allows to choose highly irregular, so called *generalized* functions, e.g. 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.

Let us consider for instance w to be a Brownian motion. 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 $\varphi(x) = \sqrt{|x|}$.

Can one construct a noise that further regularises? 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 "rough version" of a Brownian motion (see [3] for more details).

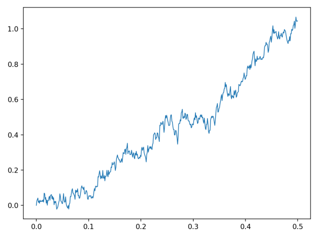

Figure 5: An infinitely regularizing noise.

[3] for simplicity, the definition of convolution used here differs slight from the usual one which, if adopted, would require to replace the left-hand side of (6) by $L^w_T * \check{\varphi}$, where $\check{\varphi}(x) = \varphi(-x)$.

References

[1] R. Bafico and P. Baldi, *Small random perturbations of Peano phenomena*, Stochastics **6** (1981/82), no. 3-4, 279–292, ISSN 0090-9491, https://doi.org/10.1080/17442508208833208. MR 665404

[2] Rick Durrett, *Probability theory and examples*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 49, Cambridge University Press, Cambridge, 2019, ISBN 978-1-108-47368-2, https://doi.org/10.1017/9781108591034, Fifth edition of [MR1068527]. MR 3930614

[3] F. A. Harang and N. Perkowski, *C^∞-regularization of ODEs perturbed by noise*, Stoch. Dyn. **21** (2021), no. 8, Paper No. 2140010, 29, ISSN 0219-4937, https://doi.org/10.1142/S0219493721400104. MR 4342752

[4] P. Mörters and Y. Peres, *Brownian motion*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 30, Cambridge University Press, Cambridge, 2010, ISBN 978-0-521-76018-8, https://doi.org/10.1017/CBO9780511750489, With an appendix by Oded Schramm and Wendelin Werner. MR 2604525

Ana Djurdjevac *is a Jun. Prof. at Freie Universität Berlin.*

Henri Elad Altman *is a Dirichlet postdoctoral fellow at Freie Universität Berlin.*

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

Mathematical subjects
will be filled out by the editors

Connections to other fields
will be filled out by the editors

DOI
will be filled out by the editors

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 Leibniz Association

IMAGINARY open mathematics