

# Lewis Fry Richardson: scientist, visionary and pacifist

Angelo Vulpiani

© Centro P.RI.ST.EM, Università Commerciale Luigi Bocconi 2014

**Abstract** The aim of this paper is to present the main aspects of the life of Lewis Fry Richardson and his most important scientific contributions. Of particular importance are the seminal concepts of turbulent diffusion, turbulent cascade and self-similar processes, which led to a profound transformation of the way weather forecasts, turbulent flows and, more generally, complex systems are viewed.

**Keywords** Lewis Fry Richardson · Weather forecasting · Turbulent diffusion · Turbulent cascade · Numerical methods · Fractals · Self-similarity

Lewis Fry Richardson (1881–1953), while undeservedly little known, had a fundamental (often posthumous) role in twentieth-century science. Even though his name is linked to several important results in fluid dynamics, in meteorology, and in numerical analysis (let us just recall his stability criterion for fluids, his idea of a scale-dependent diffusion coefficient, and the algorithm that bears his name, which is still used to integrate differential equations), physicists and mathematicians themselves often don't know who he was. His originality was too often mistaken for eccentricity, but some of the ideas and methods that he conceived would be rediscovered only decades later. G.I. Taylor, a great expert in fluid dynamics, wrote that Richardson “seldom thought on the same lines as his contemporaries and often was not understood by them” [5].

The last of seven children in a thriving English Quaker family, in 1898 Richardson enrolled in Durham College of Science, and two years later was awarded a grant to study at King's College in Cambridge, where he received a diverse education: he read physics, mathematics, chemistry, meteorology, botany, and psychology. At the beginning of his career he was quite uncertain about which road to follow, but he later resolved to work in several areas, just like the great German scientist Helmholtz, who was a physician and then a physicist, but following a different order. In his words, he decided “to spend the first half of my life under the strict discipline of physics, and afterwards to apply that training to researches on living things”.

In addition to contributing important results to meteorology, numerical analysis and fluid dynamics, Richardson was the first to attempt a mathematical description of conflicts, and was a pioneer in the study of self-similar systems as well as one of the fathers of fractals.

In one of his first jobs, he worked for the Meteorological Office; during World War I the Office was amalgamated into the Air Ministry, and Richardson, as a Quaker and a conscientious objector, lost the job. Nevertheless, he took part (unarmed) in war operations, in the Friends' Ambulance Unit, on the French front. Apparently, he was a mediocre driver but an excellent mechanic.

It was during the war that Richardson conceived his great, visionary notion for weather forecasting: using the fundamental equations of fluid dynamics and thermodynamics to determine the future state of the atmosphere. At the time, the procedure suggested by Richardson was nearly impossible to perform, due to the lack of suitable computing tools. All the same, he managed to outline the problem in a correct way and also to define numerical algorithms (still used today) to integrate differential equations. The manuscript of the book he wrote when off

---

Translation by Daniele A. Gewurz.

---

A. Vulpiani (✉)  
Dipartimento di Fisica, Università degli Studi di Roma “La Sapienza”, Piazzale Aldo Moro 2, 00185 Rome, Italy  
e-mail: angelo.vulpiani@roma1.infn.it

duty while on the front line, *Weather Prediction by Numerical Process*, was lost during the battle of Champagne in April 1917, to be fortuitously (and fortunately for the development of meteorology) found, months later, under a pile of coal.

At the age of 47, Richardson got a Ph.D. in psychology. In 1939, thanks to a small inheritance, he retired early to devote himself to his studies in mathematical psychology: he was one of the first to attempt a study of conflicts in mathematical terms.

He died in his sleep on 30 September 1953. The year after, BBC would broadcast the first TV programme of weather forecast.

Before Richardson, in order to forecast the weather a semi-empirical method was used, based on the idea that several phenomena show some regularity, not so different from what is said in the Bible: “The thing that hath been, it is that which shall be; and that which is done is that which shall be done: and there is no new thing under the sun” (Ecclesiastes, 1:9; KJV). Basically, one looks for a past state of things “near” to that of the present: if it can be found at day  $k$ , then it makes sense to assume that tomorrow the system will be “near” to day  $k + 1$ . In more formal terms, given the series  $(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M)$ , where  $\mathbf{x}_j = \mathbf{x}(j\Delta t)$  is the vector describing the state of the system at time  $j\Delta t$ , we look in the past for an analogous state, that is a vector  $\mathbf{x}_k$  with  $k < M$ , “near enough” (that, is such that  $|\mathbf{x}_k - \mathbf{x}_M| < \epsilon$ , where  $\epsilon$  denotes the desired degree of accuracy). When we find such a vector, we “predict” the future at times  $M + n > M$ , by simply assuming for  $\mathbf{x}_{M+n}$  the state  $\mathbf{x}_{k+n}$ .

It all seems quite easy, but it is not at all obvious that an analogue can be found. Richardson understood that the method could not work; there is no particular reason why a previous analogous state should exist, or at least that we can find it:

“... the *Nautical Almanac*, that marvel of accurate forecasting, is not based on the principle that astronomical history repeats itself in the aggregate. It would be safe to say that a particular disposition of stars, planets and satellites never occurs twice. Why then should we expect a present weather map to be exactly represented in a catalogue of past weather?” [13, p. vii]

The problem of finding an analogue is strictly linked to the Poincaré recurrence theorem: after a suitable time, a deterministic system with a bounded phase space returns to a state near to its initial condition. Thus an analogue is sure to exist, but how long have we to go back to find it? The answer, first basically understood by Boltzmann during his heated debate with Zermelo about irreversibility, was given

by the Polish mathematician Mark Kac: the average return time in a region  $A$  is proportional to the inverse of the probability  $P(A)$  that the system is in  $A$ :

$$\langle T_R \rangle = \frac{\tau}{P(A)},$$

where  $\tau$  is a characteristic time. To understand how hard it is to observe a recurrence, and hence to find an analogue, consider in a system of dimension  $D$  (to be precise, if the system is dissipative,  $D$  is the fractal dimension of the attractor) the probability  $P(A)$  of being in a region  $A$  that extends in every direction by a fraction  $\epsilon$  is proportional to  $\epsilon^D$ , and so we have  $\langle T_R \rangle \sim \epsilon^{-D}$ . If  $D$  is large (say, larger than 10), even for not very high levels of precision (for instance, 5 %, that is,  $\epsilon = 0.05$ ), the return time is so large that in practice a recurrence is never observed (or, equivalently, an analogue cannot be found).

In his attempts to forecast weather, Richardson introduced several of the ideas on which modern meteorology is based. In his book, he suggested the use of the equations regulating the evolution of the atmosphere: given a certain initial condition (that is, today’s weather), we can determine the future state (that is, tomorrow’s or next week’s weather), by numerically integrating the equations.

By now, his approach looks obvious (and is indeed the one presently used in weather forecasting): we know that the atmosphere evolves according to the equations of hydrodynamics (for the fields describing velocity  $\mathbf{u}$ , density  $\rho$ , pressure  $p$ , water percentage  $s$ , and temperature  $T$ ) and the thermodynamics giving the relation (equation of state) among  $\rho$ ,  $T$ ,  $s$ , and  $p$ .

So, by knowing the present state of the atmosphere, we can solve seven partial differential equations (three for velocity  $\mathbf{u}$ , and then those for  $\rho$ ,  $p$ ,  $s$ , and  $T$ ) to obtain—at least in principle—a weather forecast. Of course, these equations cannot be solved by pen and paper, so a numerical solution is the only option.

The basic idea is the right one, but the paradox lies in the fact that the equations proposed by Richardson are too accurate, and this has adverse consequences (see Appendix 1); moreover, in those times the practical obstacles were to all intents and purposes unsurmountable. The initial conditions used by Richardson consisted of a record of the weather conditions observed in Northern Europe at 4 A.M. on 20 May 1910 during an international balloon day. The numerical work by Richardson was long, taxing and wearisome: it has been estimated that, finding the time around his other duties, in the course of two years he worked for at least one thousand hours, computing by hand and with some rudimentary computing machine. The result, giving a six-hour forecast, was quite disappointing. Richardson correctly understood that “the scheme is complicated

because the atmosphere is complicated” [13, p. xi]. Nevertheless, he was moderately optimistic in his conclusive remarks: “perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances. ... But that is a dream” [13, p. vii].

In Richardson’s time the term “computer” used to refer to the person actually doing the computations, by hand and using the first calculators; anticipating parallel computers, he had the idea of using several (human) calculators simultaneously:

“Imagine a large hall like a theatre ... A myriad computers are at work upon the weather of the part of the map where each sits, but each computer attends only to one equation or part of an equation. The work of each region is coordinated by an official of higher rank” [13, p. 219].

The realisation of Richardson’s dream had to wait until the 1950s, with the development of three “ingredients”, all far from trivial (see Appendix 1):

1. finding effective equations;
2. fast numerical algorithms;
3. computers suitable for numerical calculations.

Another of Richardson’s many contributions to science is his study of turbulent diffusion: he showed that the distance between pairs of particles carried by a velocity field (for instance, the wind) is a non-Gaussian variable with large deviations from the average value. This is a key problem in many applications, for instance the spreading of a polluting agent in the sea, or of a cloud of dust in the atmosphere (see Appendix 2). As a coherent pacifist, Richardson refused for his research about diffusion to be used for war purposes. For several years, he left the study of turbulence and is supposed to have destroyed some of his unpublished works to prevent military uses.

In his studies on fluid dynamics, Richardson realised that not all natural phenomena can be described by regular functions. For instance, he remarked that, in turbulence, rather than a typical scenario with small variations around a mean value and some rare fluctuation (within, say, one standard deviation), we have long intervals of quiescence, when the signal has a regular behaviour, close to the mean value, alternating with short, irregular periods of activity with huge fluctuations. With Gaussian variables, on the contrary, such large fluctuations are not possible.

From these remarks, Richardson seriously asked himself the (apparently nonsensical) question “Does wind have a speed?” Starting from just a few empirical data, he guessed the self-similar structure of turbulence; here is how he summarised his insight in a verse (adapted from a satirical one by Swift):

Big whirls have little whirls  
that feed on their velocity,  
and little whirls have lesser whirls  
and so on to viscosity  
– in the molecular sense. [13, p. 66]

The mathematical formalisation of this idea arrived in the 1940s, thanks to Andrey Nikolaevich Kolmogorov, who showed how, in the so-called inertial interval, the velocity field is very “wrinkled”, quite different from the usual functions we are used to: the speed difference  $\delta v(\ell)$  between two points at a distance  $\ell$  is not proportional to  $\ell$ ; rather, we have a non-analytical behaviour  $\delta v(\ell) \sim \ell^{1/3}$ , with enormous speed gradients (which become infinite when Reynolds numbers tend to infinity).

The first to ask “How long is the coast of Britain?” was not Mandelbrot (who is regarded as the father of fractals), but Richardson himself. Among the papers found after his death, there are graphs in bilogarithmic coordinates where he plotted, as a function of the resolution  $\ell$ , the length  $L(\ell)$  of the coastline of Great Britain, of the land border of Germany, of the Spain-Portugal border, and of the coastlines of Australia and South Africa. Rather than a convergence to a constant value, Richardson observed a behaviour of the form  $L(\ell) \sim \ell^{-\alpha}$ , where  $\alpha$  is about zero for the coastline of South Africa, while in the other cases it is positive, and increases with the “wrinkleness” of the curve. In modern terms,  $\alpha = D_F - 1$  (where  $D_F$  is the fractal dimension).

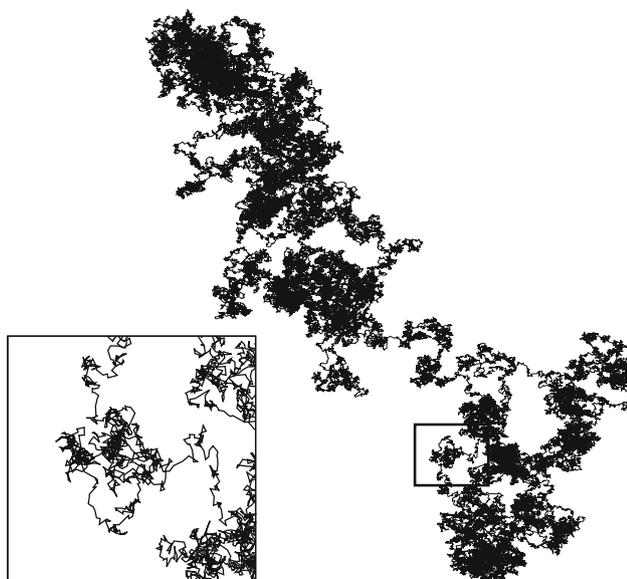
We are witnessing the beginning of fractal objects and self-similar structures which, at least in some particular cases, had already been studied between nineteenth and twentieth centuries by Cantor, Weierstrass, Hausdorff, and Julia. For instance, the function

$$f(x) = \sum_{n=1}^{\infty} A^{-n} \cos(2\pi B^{n-1}x) \quad 0 \leq x \leq 1$$

where  $A$  and  $B$  are integer and  $A < B$ , studied by Weierstrass within the theory of Fourier series, is extremely irregular, is not differentiable, and its graph has an infinite length, with fractal dimension (see Box C)  $D_F = 2 - \ln A / \ln B$  between 1 and 2.

Perrin was among the first to appreciate the importance for physics of self-similar systems, that is, systems with the property that an enlarged part looks similar to the whole system.

“Consider, for instance, one of the white flakes that are obtained by salting a soap solution. At a distance its contour may appear sharply defined, but as soon as we draw nearer its sharpness disappears.... The use of magnifying glass or microscope leaves us just as



**Fig. 1** An example of Brownian motion; in the inset, an enlargement of a detail

uncertain, for every time we increase the magnification we find fresh irregularities appearing, and we never succeed in getting a sharp, smooth impression, such as that given, for example, by a steel ball.” [27, p. ix]

Figure 1 shows an example of a Brownian motion: it is clear how an enlarged part has the same properties as the whole.

However, these studies remained on the fringe: within mathematics, they were regarded as sophisticated constructions of curious monsters; within physics, as pathological situations that were not representative of those phenomena (erroneously) considered to be the important. The person who recognised the widespread occurrence and the importance of such behaviours in natural sciences, and who coined the term “fractal” itself was Benoît Mandelbrot who, in his books, correctly acknowledges the credit due to Richardson, who is to be considered as the grandfather of fractals and self-similarity.

### Appendix 1: Weather forecasting: from Richardson to today

The key idea by Richardson to forecast the weather was correct, but in order to put it in practice it was necessary to introduce one further ingredient that he could not possibly have known. Even using modern computers, integrating numerically the equations of hydrodynamics is not easily done, one of the reasons being the so-called numerical instability, which forces us to use very small integration

steps  $\Delta t$ . Indeed, in the numerical treatment of partial differential equations it is necessary to introduce a discretisation of space, by means of a grid whose intervals measure  $\Delta x$ , and of time (with integration step  $\Delta t$ ). For the numerical algorithm to converge to the solution of the partial differential equation,  $\Delta t$  and  $\Delta x$  cannot be assigned arbitrary values, but have to satisfy constraints that depend on the equation under exam and, in general, on the initial conditions. Just to give an idea of the problem, consider the equation (far simpler than the one used by Richardson)

$$\frac{\partial u(x, t)}{\partial t} = -v \frac{\partial u(x, t)}{\partial x} \quad (1)$$

where  $v$  is a constant and  $-\infty < x < \infty$ . A simple, but efficient, approximation of this equation is

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = -v \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta x} \quad (2)$$

where  $u_j^n = u(j\Delta x, n\Delta t)$ . In this case, in order to get stability for algorithm (1.2), “Courant condition” (not yet known in Richardson’s times) has to hold:

$$\frac{\Delta t}{\Delta x} \leq C = \frac{1}{|v|}; \quad (3)$$

since the equation is linear,  $C$  does not depend on the initial condition.

In the problem studied by Richardson the negative result (the necessity of using a very small  $\Delta t$ ) is not just a mathematical detail, but has a precise physical origin. It is a consequence of the presence in the atmosphere of phenomena, such as sound waves and gravity waves, that have very short characteristic times. Directly using the equations of fluid dynamics and thermodynamics is not a promising approach; so, hoping for significant progress, it is necessary to understand which aspects of the problem have to be taken into account and which ones can be ignored.

We have to remark that fast phenomena are not especially interesting for weather forecasting, but they influence the slow variables, so they have to be somehow accounted for. The way to solve the problem was found by Charney and colleagues in the 1940s–1950s, within the Meteorological Project at the Institute for Advanced Study, in Princeton. The project involved scientists from different fields: mathematicians (such as the great John von Neumann), experts in meteorology, engineering, and computer science. We might say that his project marked the beginning of making Richardson’s dream come true.

In this important scientific effort, a fundamental role was played by the achievement of technological advances (such as the construction of the first modern computer, ENIAC), of numerical methods (algorithms for the integration of partial differential equations), and of modelling (that is, the development of effective equations for meteorology).

Charney and his colleagues noticed that the equations originally proposed by Richardson, even though correct, are not suitable for weather forecasting. The apparently paradoxical reason is that they are too accurate: they also describe high-frequency wave motions that are irrelevant for meteorology. So it is necessary to construct effective equations that get rid of the fast variables. The introduction of a sifting procedure that sets the meteorologically significant part apart from the irrelevant one has a clear practical advantage: the numerical instabilities are less severe, and so we can use a relatively large integration step  $\Delta t$ , thus obtaining more efficient numerical computations.

In addition to the computational point of view, it is important that the effective equations for the slow dynamics make it possible to detect the most important factors, which on the contrary remain hidden in the detailed description given by the original equations. The equations now used are called quasi-geostrophic; the simplest case is the barotropic one, where pressure only depends on the horizontal coordinates.

As an example of a construction of effective equations for large-scale behaviour, consider the diffusion equation in a single spatial dimension:

$$\frac{\partial}{\partial t} \theta = \frac{\partial}{\partial x} \left( D(x, x/\epsilon) \frac{\partial \theta}{\partial x} \right) \tag{4}$$

where in the diffusion coefficient  $D(x, x/\epsilon)$  there are two spatial scales: a scale  $O(\epsilon)$  and another  $O(1)$ . For instance, we may interpret  $D(x, y)$  as periodic in  $y$  with period  $L$ . The system (1.4) describes such physical processes as heat conduction in composite materials. We aim at finding an effective equation that holds for longer times on scales much larger than  $\epsilon$ , that is, an equation for the function  $\Theta(x, t)$  obtained by “filtering”  $\theta(x, t)$ ; in practice, we want a moving average  $\Theta(x, t) = \int_0^L \theta(x + z, t) dz / L$ . Homogenisation techniques (the term used in physics is *multiscale*) show that:

$$\frac{\partial}{\partial t} \Theta = \frac{\partial}{\partial x} \left( D^E(x) \frac{\partial \Theta}{\partial x} \right) \tag{5}$$

where

$$D^E(x) = \frac{1}{\left\langle \frac{1}{D(x, y)} \right\rangle} = \left[ \frac{1}{L} \int_0^L \frac{dy}{D(x, y)} \right]^{-1}$$

**Appendix 2: Diffusion is not always Gaussian**

In the simplest case of diffusion, where only molecular diffusion is at work, the probability density  $P(x, t)$  evolves according to the well-known Fick’s equation (for the sake of simplicity, we consider the one-dimensional case on an infinite line):

$$\frac{\partial}{\partial t} P(x, t) = D \frac{\partial^2}{\partial x^2} P(x, t) \tag{6}$$

where  $D$  is the diffusion coefficient. If we consider initial conditions  $P(x, 0)$  centred around the origin, the asymptotic solution is a Gaussian curve:

$$P(x, t) \simeq \frac{1}{\sqrt{4\pi Dt}} \exp -\frac{x^2}{4Dt},$$

hence  $\langle x^2(t) \rangle \simeq 2Dt$ .

Consider now the problem of *relative diffusion*, which involves two particles with positions  $x_1$  and  $x_2$ ; in the case where no mutual interaction is present and the probability density of each particle evolves according to Fick’s law, it is easy to see that for the variable  $\ell = x_2 - x_1$  the probability density follows Fick’s law (2.2) with the only difference that its diffusion coefficient is doubled:

$$\frac{\partial}{\partial t} P(\ell, t) = 2D \frac{\partial^2}{\partial \ell^2} P(\ell, t),$$

so  $\langle \ell^2(t) \rangle \simeq 4Dt$ .

Now let us ask a more ambitious question: we have a patch of some substance, transported by a velocity field, and to a first approximation let us assume that the substance is “passive”, that is, it does not influence the velocity field. This is not a mere theoretical problem; it has important real-world applications. Just think of a polluting substance in the sea or in the atmosphere: it is very important to understand how the patch gets larger in time. This is exactly the problem of relative diffusion. Of course, Fick’s law in its simpler form (2.2) does not hold anymore: indeed, the motion of the two particles is influenced by the presence of the velocity field of the fluid (air or water). Richardson, with a deeply insightful understanding of turbulence-related transport phenomena, proposed the following equation:

$$\frac{\partial}{\partial t} P(\ell, t) = \frac{\partial}{\partial \ell} \left( D(\ell) \frac{\partial}{\partial \ell} P(\ell, t) \right), \tag{7}$$

that is, a Fick’s law in which the diffusion coefficient depends on the distance, in order to keep the turbulent velocity field into account. How do we determine  $D(\ell)$ ? From data about diffusion in the atmosphere, he guessed (somewhat daringly) the law  $D(\ell) \sim \ell^{4/3}$ , from which  $\langle \ell^2(t) \rangle \sim t^3$ .

All of this took place in the 1920s, about two decades before the great Soviet mathematician Andrey Nikolaevich Kolmogorov devised (in part, following Richardson’s ideas about a turbulent cascade) the first modern theory of turbulence: in this context it is easy to prove the correctness of Richardson’s conjecture  $D(\ell) \sim \ell^{4/3}$ . Of course, the importance of his work does not lie in his “lucky guess”  $D(\ell) \sim \ell^{4/3}$ , but in his understanding of the physical processes leading to a diffusion equation of the form (7).

It is not hard to find an asymptotic solution to (7):

$$P(\ell, t) \sim \frac{1}{t^{3/2}} \exp - C_R \frac{|\ell|^{2/3}}{t},$$

a function quite different from a Gaussian.

Is all of this consistent with reality? We might argue, as suggested for instance by George Batchelor, in a different way, by considering a Fick’s equation with a diffusion coefficient that is constant when  $\ell$  varies, but depends on time; that is,

$$\frac{\partial}{\partial t} P(\ell, t) = D(t) \frac{\partial^2}{\partial \ell^2} P(\ell, t)$$

with  $D(t) \sim t^2$  (when Kolmogorov’s theory is known, this become obvious); this way, we get, as in Richardson’s approach,  $\langle \ell^2(t) \rangle \sim t^3$ , but with a Gaussian probability density:

$$P(\ell, t) \sim \frac{1}{t^{3/2}} \exp - C_B \frac{\ell^2}{t^3}.$$

We therefore have two different approaches that yield the same result regarding  $\langle \ell^2(t) \rangle$ , but predict different probability densities. Now we know that Richardson’s approach (just like Kolmogorov’s in the 1940s) cannot be completely correct, since the velocity field is intermittent (has a multifractal structure). Only at the beginning of twenty-first century have accurate numerical simulations shown that  $P(\ell, t)$  is substantially in accordance with Richardson’s conclusions: the presence of intermittence only introduces small corrections.

### Appendix 3: Fractals, self-similarity and large fluctuations

Systematising the insights by Perrin and Richardson requires new mathematical objects (fractals), which at first sight might look like pathological monstrosities, but are on the contrary extremely common in nature.

Consider a regular curve and let us ask how to measure its length  $L$ . We might approximate the curve with a broken line consisting of line segments of length  $\ell$ , and then take  $\ell$  smaller and smaller, denoting by  $N(\ell)$  the number of segments. If  $\ell$  is small enough, the length  $L$  is about  $N(\ell)\ell$ , hence  $N(\ell)$  is proportional to  $1/\ell$ . In the case of a surface, we can “tile” it with  $N(\ell)$  small squares with side  $\ell$ ; the area  $A$  can be estimated by  $N(\ell)\ell^2$ , hence  $N(\ell)$  is proportional to  $1/\ell^2$ . Analogously, in order to fill a three-dimensional object, we will need a number of small cubes of side  $\ell$  proportional to  $1/\ell^3$ .

From these remarks, the idea of generalising the notion of dimension arises: an object has a fractal dimension  $D_F$  if

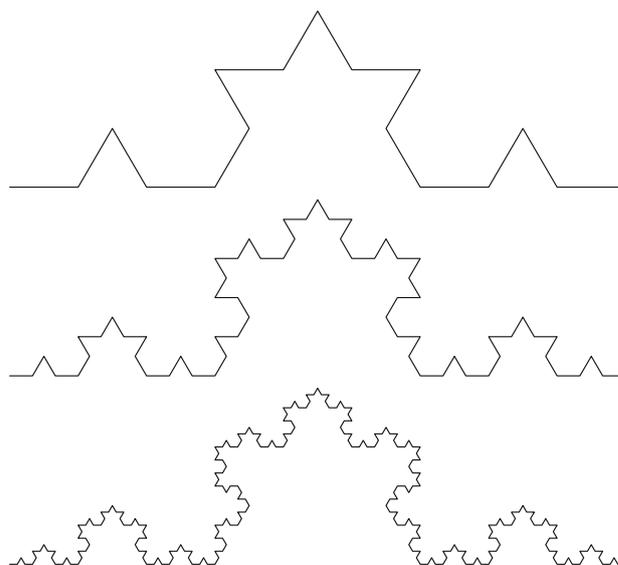


Fig. 2 Scheme for the construction of Koch snowflake

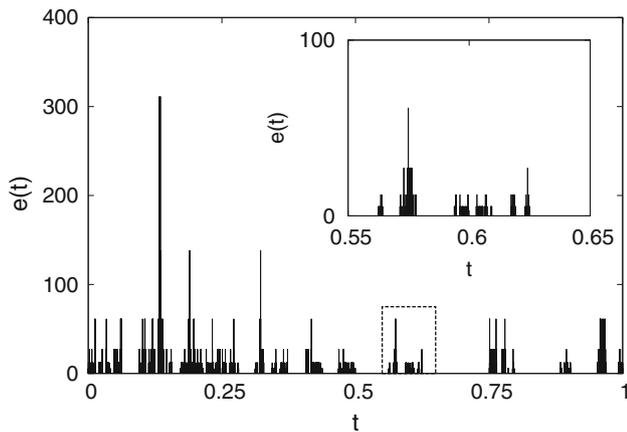
the number of “cubes” with side  $\ell$  that are necessary to cover the object behaves like

$$N(\ell) \sim \frac{1}{\ell^{D_F}}.$$

Clearly, for a regular object the fractal dimension is just the usual dimension. Are there objects for which the  $D_F$  is not an integer? The answer is in the affirmative. An example is given by Koch snowflake, whose construction is outlined in Fig. 2: take a line segment of length 1 and divide it into three equal parts. The middle part is taken out and substituted with two segments with the same length. Repeat the procedure on each of the four elements we have now obtained, then on the sixteen elements of the next generation, and so on, an infinite number of times. We get a “very angular” curve having fractal dimension  $\ln 4 / \ln 3 \simeq 1.2618$ .

This is indeed a self-similar structure: looking at such a figure under any given resolution we are not able to say the scale we are at. The truly important fact is that this kind of behaviour is not a mere artefact of pathological mathematical models. For instance, if we try to measure the length  $L(\ell)$  of a jagged coastline (as first suggested by Richardson), we find that it gets longer and longer the more the resolution  $\ell$  gets smaller: we have a fractal object with  $D_F$  between 1 and 2; for Koch snowflake,  $L(\ell) \sim \ell^{-(\ln 4 / \ln 3 - 1)}$ .

This kind of “wrinkledness” is very common: it appears in the attractors of chaotic dissipative dynamic systems and in many natural phenomena; for instance, in turbulence (as described by Richardson in his verse about whirls containing smaller whirls that in turn contain even smaller ones and so on) and in the large-scale structure of galaxies.



**Fig. 3** The energy dissipated in a turbulent fluid as a function of time; note that the enlargement of a part is similar to the whole

Two of the most important results in probability theory are the law of large numbers and the central limit theorem, which hold for independent (or weakly dependent) variables having finite variance. What happens in the case of strongly dependent variables or with infinite variance? As an example, consider a sequence of independent random variables  $\{x_n\}$ ,  $n = 1, 2, \dots, N$ , each of which has Cauchy probability density

$$p(x) = \frac{1}{\pi(1+x^2)};$$

for these variables (with infinite variance) it is possible to prove the following property: the “average”

$$Y_N = \frac{x_1 + x_2 + \dots + x_N}{N}$$

for each  $N$  (even an arbitrarily large one) has the same probability distribution as the single variable. We are witnessing here a situation which is very different from that in which the variance is finite, where by the law of large numbers, for  $N \rightarrow \infty$ ,  $(x_1 + \dots + x_N)/N$  converges, besides zero-probability cases, to the mean value.

In the case when the central limit theorem holds, we have that for large values of  $N$  the variable

$$z_N = \frac{x_1 + \dots + x_N - mN}{\sigma\sqrt{N}}$$

has a limit distribution (the Gaussian one). If the variance is infinite, in order to have a limit distribution we have to consider the variable

$$q_N = \frac{x_1 + \dots + x_N - mN}{C_\alpha N^\alpha}$$

for suitable values of  $\alpha$  and  $C_\alpha$ , which depend on the distribution of the variables (and especially on the behaviour

of the tails); for large values of  $N$  the probability distribution of  $q_N$  converges to a limit function (different from the Gaussian). For instance, in the case when for large values of  $|x|$  the probability density is of the form

$$p_x(x) \sim \frac{1}{x^2},$$

then  $\alpha = 1$  and the limit distribution is Cauchy; clearly, if the variance is finite we have  $\alpha = 1/2$  and the limit distribution is Gaussian.

Nowadays, the probability distributions with power-law tails and infinite variance are accepted as non-pathological, and widely used in modelling several physical and financial phenomena. It was not at all so in Richardson’s time; he was one of the first, in his studies about armed conflicts, to notice the non-pathological presence of this kind of probability density.

Even in the cases with a finite variance, when strong correlation between variables at different times is present, it is possible to get “wild fluctuations”; an example given by turbulence is shown in Fig. 3. We have there a non-Gaussian probability distribution; looking at the evolution of the variable over time, we note an alternation of long intervals with small fluctuations around its mean value with short, strong, intermittent departures.

Translated from the Italian by Daniele A. Gewurz.

**For further reading**

**Richardson’s life and work**

1. Richardson, S.A.: Lewis Fry Richardson (1881–1953): a personal biography. *Conflict Resolution* **1**, 300 (1957)
2. Ashford, O.M.: Prophet or Professor? The Life and Work of Lewis Fry Richardson. Adam Hilger (1985)
3. Hunt, J.C.R.: Lewis Fry Richardson and his contributions to mathematics, meteorology, and models of conflict. *Annu. Rev. Fluid Mech.* **30**, XIII (1998)

**Diffusion**

4. Richardson, L.F.: Atmospheric diffusion shown on a distance-neighbour graph. *Proc. R. Soc. Lond. Ser A* **110**, 709 (1926)
5. Taylor, G.I.: The present position in the theory of turbulent diffusion. *Adv. Geophys.* **6**, 101 (1959)
6. Majda, A.J., Kramer, P.R.: Simplified models for turbulent diffusion: theory, numerical modeling and physical phenomena. *Phys. Rep.* **314**, 237 (1999)
7. Boffetta, G., Sokolov, I.M.: Relative dispersion in fully developed turbulence: the Richardson’s law and intermittency corrections. *Phys. Rev. Lett.* **88**, 094501 (2002)

## Turbulence

8. Kolmogorov, A.N.: The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. Dokl. Akad. Nauk SSSR **30**, 301 (1941)
9. Kolmogorov, A.N.: A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. J. Fluid. Mech. **13**, 82 (1962)
10. Monin, A., Yaglom, A.: Statistical Fluid Dynamics, vols. I, II. MIT Press, Cambridge (1975)
11. Frisch, U.: Turbulence: The Legacy of A. N. Kolmogorov. Cambridge University Press, Cambridge (1995)
12. Bohr, T., Jensen, M.H., Paladin, G., Vulpiani, A.: Dynamical Systems Approach to Turbulence. Cambridge University Press, Cambridge (1998)

## Weather forecasting

13. Richardson, L.F.: Weather Prediction by Numerical Process. Cambridge University Press, Cambridge (1922)
14. Charney, J.G.: On a physical basis for numerical prediction of large-scale motions in the atmosphere. J. Meteor. **6**, 371 (1949)
15. Charney, J.G., Fjortoft, R., von Neumann, J.: Numerical integration of the barotropic vorticity equation. Tellus **2**, 237 (1950)
16. Dahan Dalmedico, A.: History and epistemology of models: meteorology as a case study. Arch. History Exact Sci. **55**, 395 (2001)
17. Lynch, P.: The Emergence of Numerical Weather Prediction: Richardson's Dream. Cambridge University Press, Cambridge (2006)
18. Lynch, P.: The origins of computer weather prediction and climate modeling. J. Comput. Phys. **227**, 3431 (2008)
19. Ceconi, F., Cencini, M., Falcioni, M., Vulpiani, A.: The prediction of future from the past: an old problem from a modern perspective. Am. J. Phys. **80**, 1001 (2012)
20. Tibaldi, S.: Che tempo farà. Le Scienze **538**, 42 (2013)

## Multiscale techniques

21. E, W., Engquist, B.: Multiscale modeling and computation. Notices Am. Math. Soc. **50**, 1062 (2003)
22. Mazzino, A., Musacchio, S., Vulpiani, A.: Multiple-scale analysis and renormalization for preasymptotic scalar transport. Phys. Rev. E **71**, 011113 (2004)

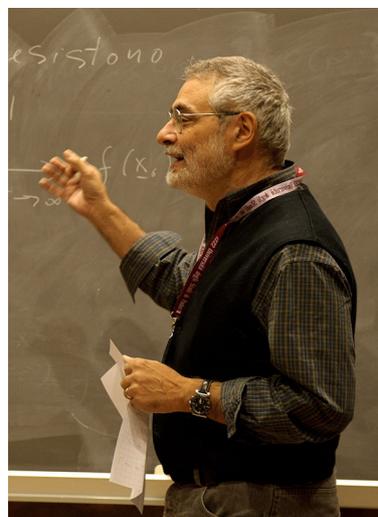
23. E, W.: Principles of Multiscale Modeling. Cambridge University Press, Cambridge (2011)

## Mathematics of conflicts

24. Richardson, L.F.: Could an arms race end without fighting? Nature **168**, 567 (1951)
25. Richardson, L.F.: The problem of contiguity: an appendix of statistics of deadly quarrels. Gen. Syst. Yearbook **6**, 139 (1961)
26. Hess, G.D.: An introduction to Lewis Fry Richardson and his mathematical theory of war and peace. Conflict Manage. Peace Sci. **14**, 77 (1995)

## Fractals and multifractals

27. Perrin, J.: Les atomes (1913)
28. Mandelbrot, B.B.: The Fractal Geometry of Nature. W.H. Freeman, New York (1982)
29. Paladin, G., Vulpiani, A.: Anomalous scaling laws in multifractal objects. Phys. Rep. **156**, 147 (1987)
30. Boffetta, G., Mazzino, A., Vulpiani, A.: Twenty-five years of multifractals in fully developed turbulence: a tribute to Giovanni Paladin. J. Phys. A Math. Theor. **41**, 363001 (2008)



**Angelo Vulpiani** is a professor of theoretical physics at the University of Rome “La Sapienza” and is a fellow of the Institute of Physics. He has been a visiting scientist in Belgium, Denmark, France, Sweden and the United States. His research interests include statistical mechanics, dynamical systems, turbulence, transport, and reaction-diffusion processes.