# Fractional phenomenology of cosmic ray anomalous diffusion 

V V Uchaikin

## Contents

1. Introduction ..... 1074
2. Classical sources ..... 10762.1 Brownian motion; 2.2 Some remarks on Einstein's derivation; 2.3 Turbulent diffusion; 2.4 Percolation;2.5 Combined equation and its interpretation; 2.6 Fractionally stable statistics
3. Isotropic anomalous diffusion1084
3.1 Traps and voids; 3.2 First studies with a fractional differential model; 3.3 ' Knee' in the spectrum and the model parameterization; 3.4 Equations of restricted-speed anomalous diffusion; 3.5 Anomalous diffusion distributions; 3.6 Range-time correlations; 3.7 Anisotropy; 3.8 Reaching the galactic disc boundary
4. Anisotropic anomalous diffusion
4.1 Compound diffusion model; 4.2 Fractional Brownian model of field lines; 4.3 Transverse diffusion in fractional- order operators; 4.4 Longitudinal finite-velocity random walk1100
5. Diffusion acceleration models
5.1 Classical Fermi model; 5.2 Fractional differential kinetic equation; 5.3 Fractional differential Fokker-Planck equation; 5.4 Integro-fractional differential model1108
6. Problems and outlook1113
6.1 Fractional time derivative; 6.2 Fractional Laplacian; 6.3 Gradient and material derivative
7. Conclusions ..... 1116
References ..... 1118


#### Abstract

We review the evolution of the cosmic ray diffusion concept from the ordinary (Einstein) model of Brownian motion to the fractional models that appeared in the last decade. The mathematical and physical foundations of these models are discussed, as are their consequences, related problems, and prospects for further development.


## 1. Introduction

The physical foundations for the applicability of the classical isotropic diffusion model to the description of the propagation of cosmic rays in the Galaxy can already be found in the pioneering papers by Fermi and Ginzburg. In his first paper on the nature of cosmic rays [1], Fermi proposed a hypothesis that "cosmic rays originate and are accelerated primarily in the interstellar space, although they are assumed to be prevented by magnetic fields from leaving the boundaries of the galaxy.... Such fields have a remarkably great stability because of their large dimensions (of the order of magnitude of light years), and of the relatively high electrical conductivity of the interstellar space. Indeed, the conductivity is so high that one might describe the magnetic lines of force as attached

V V Uchaikin Ul'yanovsk State University,
ul. L. Tolstogo 42, 432917 Ul'yanovsk, Russian Federation
E-mail: vuchaikin@gmail.com
Received 3 December 2012, revised 12 April 2013
Uspekhi Fizicheskikh Nauk 183 (11) 1175-1223 (2013)
DOI: 10.3367/UFNr.0183.201311b. 1175
Translated by M Sapozhnikov; edited by A M Semikhatov
to the matter and partaking in its streaming motions... . The evidence indicates, however, that this matter is not uniformly spread, but that there are condensations where the density may be as much as ten or a hundred times as large and which extend to average dimensions of the order of 10 parsec.... Such relatively dense clouds occupy approximately 5 percent of the interstellar space" [2]. Fermi argued that the acceleration of a particle moving in the interstellar space was the result of its scatterings in collisions with magnetized clouds.

Five years later, Ginzburg wrote: "The motion of charged particles in the interstellar space resembles Brownian motion or motion of molecules in a gas. Indeed, due to the presence of the interstellar magnetic field, in the region where this field is quasihomogeneous, the trajectory of a particle winds around a magnetic field line and, upon averaging over the rotation period, is close to a straight line. However, on passing to a region with a different field direction, the trajectory changes and becomes a broken line as a whole. If the size of regions where the field direction noticeably changes is small compared to that of regions with a quasihomogeneous field, the particle motion can be treated as the motion of a molecule in a gas: the motion is free in the homogeneous field, and a change in the velocity direction at a boundary is similar to a collision with another molecule and can be usually assumed instantaneous. Hence, the size of the region with a quasihomogeneous field plays the role of the mean free path $l$. The mean free time is $\tau=l / v_{0}$, where $v_{0}$ is the translational velocity along the trajectory, which is by an order of magnitude equal to the usual velocity of the particle itself (and we therefore assume below that $\tau \sim l / v$, where $v$ is the particle velocity). If magnetic fields do not change in time, this collision process
leads only to the diffusion of particles and the 'mixing' of their velocities over directions but not to a change in the energy of the particles. It is known from the diffusion theory that the mean square distance $L$ propagated by a particle in a time $t$ is

$$
L=\sqrt{6 D t} \sim \sqrt{l v t},
$$

where $D \sim l v / 3$ is the diffusion coefficient. According to astronomical data, $l>10^{19} \mathrm{~cm}$ in the interstellar medium, and for $v \sim c$ and $t \sim T \sim 10^{16} \mathrm{~s}$ (the proton lifetime), we obtain $L \sim 3 \times 10^{22} \mathrm{~cm}$, which is of the order of the size of the Galaxy. Therefore, for $l<10^{19} \mathrm{~cm}$, protons, and all the more so nuclei, have no time to escape in great numbers from the Galaxy" [3, pp. 368, 369].

Of course, it was clear from the very beginning that the intricate cosmic-ray transfer process cannot be fully described by the classical diffusion model. Ginzburg and Syrovatskii write in the first monograph on cosmic rays [4]: "The high degree of isotropy of cosmic rays was one of the first indications that cosmic rays fall on Earth not directly from sources but after complicated motion and scattering in interstellar magnetic fields. This motion can be considered a 'diffusion' of cosmic rays in the interstellar space during which particles 'forget' about their initial direction of motion. However, the determination of the real nature of this diffusion is a quite challenging problem." Giving the model of the adiabatic motion of particles along field lines the due credit, the authors point out that the necessary condition for this motion (the radius of curvature being much smaller than the size of inhomogeneities of the magnetic field) is not satisfied everywhere, and its violation (in shock waves with a small front width or in regions with a zero magnetic field strength) eventually leads to the diffusion process.

Despite such apparently intuitive and phenomenological, ${ }^{1}$ rather than physico-mathematical, foundations, the diffusion direction in cosmic-ray physics exists and is still being developed. To account for a change in the energy spectrum of particles with the distance from their source due to ionization, synchrotron radiation, and additional acceleration by fluctuations of magnetic fields, shock waves, and supernova remnants, the energy dependence was introduced for the only material parameter in the isotropic diffusion model, the diffusion coefficient. In view of the observed decrease in the fraction of secondary nuclei with energy, the diffusion coefficient was approximated by a power-law function

$$
\begin{equation*}
D(E)=D_{0} E^{\delta} \tag{1}
\end{equation*}
$$

where $E$ is the particle energy in GeV , with the exponent $\delta \simeq 0.3-0.7$, which was consistent, in particular, with the data on the cosmic-ray anisotropy [6].

Along with the isotropic model, the anisotropic diffusion model is widely used in local problems of galactic cosmic-ray
${ }^{1}$ In this connection, we quote a remarkable note by Heisenberg [5]: "A 'phenomenological' theory is understood as the formulation of regularities in the field of observed physical phenomena that does not attempt to reduce the described relations to the underlying general laws of nature through which they could be understood. Such phenomenological theories have always played a considerable role in the development of physics... . Of course, phenomenological theories are always developed where the observed phenomena cannot yet be reduced to the general laws of nature. The reason for this can be either an extreme complexity of these phenomena, which makes such a reduction impossible because of mathematical difficulties, or the lack of knowledge about these laws."
transfer. This model was initially developed in theoretical studies of the motion of charged particles in quasihomogeneous regions with a fluctuating magnetic field slightly different from a constant homogeneous field. The development of this model led to the separation of the diffusion of charged particles into the longitudinal and transverse components, each of which was described by a diffusion equation of the corresponding dimension with its own diffusion coefficient [7-10]. The transverse diffusion was the first example of anomalous diffusion. The transverse diffusion anomaly was manifested not only in its slowness compared to the normal diffusion (which could be achieved by simply introducing a smaller diffusion coefficient) but also in a different expansion law for a diffusion packet and its different shape. Some authors believe that the local interpretation of such a composite model of anomalous diffusion (compound diffusion) can be extended to the entire galactic disc. For example, Hayakawa writes: "In this model, interstellar magnetic fields are assumed almost homogeneous along spiral sleeves. Particles are drifting along field lines and are reflected at mirror points... . Particles captured and kept on a field line continue to diffuse... in accordance with the chaotic motion of the field line.... Because the magnetic field is homogeneous only at the distance of a few kiloparsecs, we can assume that particles have escaped from the Galaxy if they have propagated a path longer than the field homogeneity length" [10].

Papers by Urch [11, 12] devoted to the study of the motion of charged particles in random magnetic fields can probably be considered forerunners of the appearance of fractional derivatives in the cosmic-ray phenomenology. Discussing the applicability of the Fokker-Planck equation under the assumption that the trajectories of particles propagating along the unperturbed trajectory over distances many times exceeding the correlation length are only slightly perturbed and a number of other conditions are satisfied (the gyroradius $r_{\mathrm{g}}^{\prime}$ of particles due to the field perturbation is negligibly small compared to the correlation length $L_{c}$ of the field, the stochastic magnetic field consists of unpolarized Alfvén waves propagating along the main field $\mathbf{H}_{0}$ directed along the $z$ axis, and the velocity $v$ of particles greatly exceeds the Alfén wave velocity), Urch reaches the conclusion that for

$$
\frac{v_{z}}{v}<\sqrt{\frac{L_{\mathrm{c}}}{r_{\mathrm{g}}^{\prime}}}
$$

the Fokker-Planck equation leading to Fick's law $J_{x}=-D \partial f / \partial x$ of normal diffusion is inapplicable. Urch performed calculations not related to the Fokker-Planck equation and found that the relation between the transverse component $J_{x}$ of the particle current density and their concentration $N$ in the given problem under consideration differs from the usual Fick law by the presence of the third derivative instead of the first one:

$$
J_{x}=-D_{\|} D_{\mathrm{L}}^{2} \frac{\partial^{3} f}{\partial x^{3}}
$$

In [11, 12], fractional derivatives were not mentioned, but Webb and coauthors [13] later noticed that the Urch formula in conjunction with the continuity equation

$$
\frac{\partial f}{\partial t}+\frac{\partial J_{x}}{\partial x}=0
$$

leads to an equation with the fourth derivative with respect to the coordinate,

$$
\left(\frac{\partial}{\partial t}-D_{\|} D_{\mathrm{L}}^{2} \frac{\partial^{4}}{\partial x^{4}}\right) f(x, t)=0 .
$$

Factoring the operator in the left-hand side of this equation,

$$
\begin{aligned}
\frac{\partial}{\partial t}-D_{\|} D_{\mathrm{L}}^{2} \frac{\partial^{4}}{\partial x^{4}} & =\left(\sqrt{\frac{\partial}{\partial t}}+\sqrt{D_{\|}} D_{\mathrm{L}} \frac{\partial^{2}}{\partial x^{2}}\right) \\
& \times\left(\sqrt{\frac{\partial}{\partial t}}-\sqrt{D_{\|}} D_{\mathrm{L}} \frac{\partial^{2}}{\partial x^{2}}\right)
\end{aligned}
$$

formally leads to a fractional differentiation operator (of the order $1 / 2$ ).

The next paper "in the vicinity of a fractional derivative" was preprint [14], where various compound-diffusion regimes were investigated in more detail. The description of one of these regimes resulted in an equation close to the one with a fractional derivative (see the details in [15]). The equation involved an integral of the solution over the time interval preceding the observation instant, which can be interpreted as a peculiar effect of 'magnetic traps', characteristic of plasma dynamics [16]. At the same time, estimates in [6, p. 90] based on a comparison of the anisotropy $\delta \sim 10^{-3}$ and the mean free path $L \lesssim 10^{21} \mathrm{~cm}$ in the disc with the size of the disc itself confirmed the old assumption in [17] that "cosmic rays cannot freely propagate along the disc but should also efficiently shift in the transverse direction. Such a motion can be caused by the mixing and entanglement of lines of force themselves, which carry the frozen relativistic cosmic-ray gas away to the disc boundaries. Thus, something similar to diffusion should obviously take place."

However, it is reasonable to assume that calculations of the transverse diffusion performed in the perturbation theory approximation should not be extended to longer times, because this method exhausts its possibilities as the cumulative effect of the perturbations builds up. For example, the calculations in [18] showed that the direct solution of the compound diffusion problem, which can be obtained in a simple model, also gives the normal diffusion in transverse directions in the long-time asymptotics. The authors of [18] also showed that this conclusion can be obtained in the perturbation theory by changing the sequence of averaging procedures. This remarkable fact emphasizes that calculations based on the perturbation theory should be treated with caution.

Another important aspect of the propagation of galactic cosmic rays is the convection transfer mechanism caused by large-scale motions of a medium as a whole with a convection velocity $\mathbf{u}(\mathbf{r}, t)$. "The large-scale motions of a medium can be random, and then on the average at scales greatly exceeding the main turbulence scale $L$, the motion of particles in some cases (for $D \ll u L / 3$ ) is reduced to diffusion with the effective turbulent diffusion coefficient of the order of $u L / 3$ " [6]. The most important difference between turbulent diffusion and molecular diffusion is the nonlocal (specifically, spatially nonlocal) character of the former: the presence of vortex formations at different scales gives rise to long-range spatial correlations of the velocity field.

Following the evolution of the diffusion model as additional information is being gradually included, we infer that imitation possibilities of the model are already nearly
exhausted. The reason for this is clear: the diffusion process is determined by the only parameter (except the space-time scales), the diffusion coefficient, and this single parameter (even if split into components, as in the case of compound diffusion) is insufficient. The natural way out, by replacing the diffusion coefficient with its random analog and subsequently averaging the equation and obtaining its averaged solution, was mathematically found only in the case of small fluctuations, which are of minor interest in our problems: in the turbulent interstellar space, a major role is played by large fluctuations alternating with different-scale voids and characterized by long-range power-law correlations. To describe the transfer of such a 'ragged' ( fractal) medium, a special mathematical apparatus was required, and it was developed in the framework of fractional calculus.

The introduction of the method of fractional derivatives for solving a number of relevant problems [19, 20] was stimulated by the use of the fractional differential technique in [21,22] and our experience, described in [23], of working with stable non-Gaussian distributions. The appearance of two new parameters, the spatial $(\alpha)$ and temporal $(\beta)$ fractional-order derivatives, remarkably extended the family of solutions of the diffusion equation, formally preserving its form. The most important feature of new solutions is the power law of their asymptotic behavior, which is in excellent agreement with the known properties of the turbulent interstellar medium, the Fermi acceleration mechanism, and other processes affecting cosmic rays. At the same time, the fractional differential approach, unlike other nonlocal approaches, demonstrated a peculiar 'correspondence principle', incorporating normal diffusion as a particular case corresponding to $\alpha=2$ and $\beta=1$.

Beginning from the abovementioned works, a series of papers developing this area were published over more than a decade. But in hindsight, reviewing these papers have shown the scarcity of a clear physical motivation for introducing fractional derivatives. The treatment was dry and laconic, and the papers (although most of them were reports presented at conferences on cosmic rays) mainly contained formulas that the readers were not familiar with, plus traditional numerical materials: plots, tables, and comparisons with experimental data, which apparently made any additional explanations unnecessary. Other recent papers using the fractional differential approach are not free of this drawback either.

In this review, we try to fill this gap as much as possible and to show that the fractional differential phenomenology naturally appears in cosmic-ray physics as a direct logical development of the concepts proposed by its founders.

## 2. Classical sources

### 2.1 Brownian motion

To clarify the probabilistic nature of the diffusion process, we consider Einstein's paper [24] of 1905, where the diffusion equation was derived based on the probabilistic concept of Brownian motion. Here is an excerpt from this paper: "Obviously, it must be assumed that each individual particle moves independently of other particles; in addition, motions of the same particle at different time intervals should be regarded as independent of each other until these intervals become too small."

By introducing the probability distribution density $f(x, t)$ of the coordinate of a particle experiencing a random walk
along the $x$ axis and letting $\varphi(\Delta)$ denote the symmetric distribution density of a random displacement of the particle in the time $\tau$, Einstein writes the equation

$$
\begin{equation*}
f(x, t+\tau)=\int_{-\infty}^{\infty} f(x+\Delta, t) \varphi(\Delta) \mathrm{d} \Delta \tag{2}
\end{equation*}
$$

[Instead of $x+\Delta$ in the argument of $f$, it should be $x-\Delta$, but the symmetry condition for $\varphi(\Delta)$ neutralizes Einstein's slip of the pen.] We next quote the Russian translation in [24, p. 115]: "Because $\tau$ is very small, we can write

$$
f(x, t+\tau)=f(x, t)+\tau \frac{\partial f}{\partial t} .
$$

Now we expand $f(x+\Delta, t)$ in a power series in $\Delta$ :

$$
\begin{aligned}
f(x+\Delta, t) & =f(x, t)+\Delta \frac{\partial f(x, t)}{\partial x} \\
& +\frac{\Delta^{2}}{2} \frac{\partial^{2} f(x, t)}{\partial x^{2}}+\ldots \text { to infinity } .
\end{aligned}
$$

This expansion can be introduced in the integrand because only very small values of $\Delta$ are important here. We obtain

$$
\begin{aligned}
f+\frac{\partial f}{\partial t} \tau= & f \int_{-\infty}^{\infty} \varphi(\Delta) \mathrm{d} \Delta+\frac{\partial f}{\partial x} \int_{-\infty}^{\infty} \Delta \varphi(\Delta) \mathrm{d} \Delta \\
& +\frac{\partial^{2} f}{\partial x^{2}} \int_{-\infty}^{\infty} \frac{\Delta^{2}}{2} \varphi(\Delta) \mathrm{d} \Delta+\ldots
\end{aligned}
$$

Because $\varphi(\Delta)=\varphi(-\Delta)$, the second, fourth, etc. terms in the right-hand side vanish, whereas among the first, third, fifth, etc. terms, each next term is very small compared with the preceding one. Taking into account that

$$
\int_{-\infty}^{\infty} \varphi(\Delta) \mathrm{d} \Delta=1
$$

and setting

$$
\frac{1}{\tau} \int_{-\infty}^{\infty} \frac{\Delta^{2}}{2} \varphi(\Delta) \mathrm{d} \Delta=D
$$

we restrict ourselves to the first and third term in the righthand side of the equation to obtain

$$
\begin{equation*}
\frac{\partial f(x, t)}{\partial t}=D \frac{\partial^{2} f(x, t)}{\partial x^{2}} . \tag{3}
\end{equation*}
$$

This is the known differential equation for diffusion, and $D$ is the diffusion coefficient."

We note that in introducing the time interval $\tau$, Einstein defined it as "very small compared to the observed time intervals, but large enough to allow considering the motions of a particle during two successive time intervals as independent events" [24, p. 114]. During time intervals much shorter than $\tau$, the particle can move without collisions with atoms of the medium; correlations are then strictly determined by Newton's law of motion. It is for this reason that Einstein defines $\varphi(\Delta)$ in Eqn (2) separately, not identifying this density with $f(x, \tau)$. Such a step would have led him to the ChapmanKolmogorov equation for Markov processes,

$$
\begin{equation*}
f(x, t+\tau)=\int_{-\infty}^{\infty} f(x+\Delta, t) f(\Delta, \tau) \mathrm{d} \Delta, \quad f(x, 0)=\delta(x), \tag{4}
\end{equation*}
$$

which underlies the theory of random processes with independent increments (see review [25]). It is to this class of processes that the Brownian motion belongs: to separate it from the entire class, it suffices to require two conditions be satisfied:
(i) The motion must be self-similar,

$$
\begin{equation*}
f(x, t)=t^{-\gamma} f\left(x t^{-\gamma}, 1\right) ; \tag{5a}
\end{equation*}
$$

(ii) the motion must have a finite dispersion,

$$
\begin{equation*}
\int_{-\infty}^{\infty} x^{2} f(x, t) \mathrm{d} x<\infty \tag{5b}
\end{equation*}
$$

As a result, we obtain Eq. (3) with $\gamma=1 / 2$. The solution of this equation with the initial condition $f(x, 0)=\delta(x)$ has form (5a), where $f(x, t)$ is the Gaussian density,

$$
f(x ; 1)=\frac{1}{2 \sqrt{\pi D}} \exp \left(-\frac{x^{2}}{4 D}\right)
$$

with the dispersion equal to $2 D$, and the diffusion coefficient is

$$
D=\frac{\left\langle[X(t)]^{2}\right\rangle}{2 t}
$$

We note that Einstein derived diffusion equation (3) as the asymptotic form of integral equation (2) as $t \rightarrow \infty$, whereas conditions (5) imply the equivalence of Eqns (3) and (4) for all times. Such a process is called the Wiener process (Brownian motion is often understood just as a Wiener process).

Trajectories in a Wiener process are continuous, nowhere differentiable lines. Such a trajectory observed at any magnification always represents an infinitely broken line without any single smooth interval. The length of a part of such a trajectory, even between closely spaced points, is infinite, and therefore the velocity of such a particle is infinite. Neither a magnetic field line with such properties nor the trajectory of a real physical particle with a charge and a mass can be imagined. But hardly anyone ponders this: the 'good old Gaussian', familiar since the student desk, inspires confidence....

### 2.2 Some remarks on Einstein's derivation

To understand Einstein's random walk model more clearly, a few remarks are in order. First, Einstein did not identify $\varphi(\Delta)$ with $f(\Delta, \tau)$, thereby providing the possibility of introducing additional information into his model (for example, the velocity of the free motion of particles). But if we set $\varphi(\Delta)=f(\Delta, \tau)$, as in Eqn (4), and at the same time relax the conditions, retaining only (5a) and removing constraint (5b), we obtain a broader family of random processes, which are called Lévi motion (like Brownian motion). This family plays a key role in probability theory, because it includes exactly the limit distributions in the scheme of summation of independent identically distributed random quantities. This explains the wide applicability of the diffusion model in physics, from atomic to cosmic scales, and attracts our attention to other members of the family generated by Eqn (4) under condition (5a). It was shown in [25] that the problem in (4), (5a) can be identically transformed into a fractional-derivative equation without any additional assumptions or simplifications.

Second, only the first term of the expansion is kept in the right-hand side of the final equation, although other terms could also be kept. Einstein wrote 'to infinity', assuming that the function is infinitely differentiable, which is of course unnecessary. In the given case, it is sufficient to assume only the existence of the second derivative and to use the Taylor formula containing the function itself at the point $x$, the terms with the first and second derivatives, and the residual term, say, in the integral form:

$$
\begin{aligned}
f(x+\Delta) & =f(x)+\Delta f^{\prime}(x)+\frac{\Delta^{2}}{2!} f^{\prime \prime}(x) \\
& +\frac{1}{2!} \int_{0}^{\Delta}(\Delta-\xi)^{2} f^{(3)}(x+\xi) \mathrm{d} \xi
\end{aligned}
$$

This is an exact formula, whereas the Einstein equation is derived from its truncated version (with the residual term omitted). In principle, nothing prevents the continuation of this expansion, somewhat strengthening the conditions on the function, for example, by requiring the existence of the fourth-order derivative. Such an equation,

$$
\frac{\partial f}{\partial t}=D_{2} \frac{\partial^{2} f}{\partial x^{2}}+D_{4} \frac{\partial^{4} f}{\partial x^{4}},
$$

was derived by Burnett [26] based on physical considerations, which initiated a series of papers in the area that was later called 'generalized hydrodynamics'.

Because we are always considering symmetric diffusion, we can pass in these equations to derivatives with respect to the absolute coordinate:

$$
\frac{\partial f}{\partial t}=D_{2} \frac{\partial^{2} f}{\partial|x|^{2}}+\ldots+D_{2 n} \frac{\partial^{2 n} f}{\partial|x|^{2 n}}, \quad n=1,2, \ldots
$$

Einstein delicately bypassed the problem of the convergence of improper integrals,

$$
\int_{A}^{B} \Delta^{2 n} \varphi(\Delta) \mathrm{d} \Delta \rightarrow \int_{-\infty}^{\infty} \Delta^{2 n} \varphi(\Delta) \mathrm{d} \Delta, \quad A \rightarrow-\infty, \quad B \rightarrow \infty
$$

assuming that $\varphi(\Delta)=0$ outside a narrow symmetric interval near zero; all the moments then converge. This cannot be applied to Eqn (4): we cannot impose such a condition on the solution of the problem, which is now the function $\varphi(\Delta) \equiv f(\Delta, \tau)$. However, supplementing the condition that the process be Markovian with the requirement that the process be self-similar and its dispersion be finite, we necessarily arrive at normal diffusion equation (3) [25]. The next (fourth) moment can be infinite. However, this is unimportant because we do not use an infinite series, but take only two of its first terms with finite coefficients, while all the rest is included in the residual term, which we do not expand, and therefore higher-order coefficients simply do not appear, whereas the residual term itself is finite.

It now becomes clear what to do if the coefficient at the second derivative (i.e., the diffusion coefficient) already diverges: we must then terminate the series at the derivative of a fractional order $\alpha<2$ such that the corresponding moment converges. In this case, Einstein would have arrived at the equation

$$
\frac{\partial f(x, t)}{\partial t}=D_{\alpha} \frac{\partial^{\alpha} f(x, t)}{\partial|x|^{\alpha}}, \quad 0<\alpha<2 .
$$

The fractional operator appearing in the right-hand side can be considered a one-dimensional version of the fractional Laplace operator

$$
\frac{\partial^{\alpha}}{\partial|x|^{\alpha}}=\frac{\partial^{2 \alpha / 2}}{\partial\left(x^{2}\right)^{\alpha / 2}}=\left(\frac{\partial^{2}}{\partial x^{2}}\right)^{\alpha / 2},
$$

which we discuss below.
And the last remark: Einstein was not fully satisfied by the result he obtained, which clearly contradicted the most important principle of his theory of relativity: a diffusion packet, being concentrated at the initial instant at the coordinate origin, the next instant fills the entire space, including its remotest regions. Practically, this did not cause any inconveniences due to the vanishingly small probability of the residence of the diffusion packet there, but conceptually Einstein could not help feeling some discomfort, as has now become fashionable to say. However, he did not develop this topic any further.

### 2.3 Turbulent diffusion

As we saw in Sections 2.1 and 2.2, the fractional differential concepts of anomalous diffusion are already contained in the equation for Brownian motion derived by Einstein. But this fact was realized only after half a century, and it was related to the very popular topic in the 1950s, the nature of turbulence, or more exactly, turbulent diffusion (TD). The specificity of TD is determined by the action of vortices of different sizes on a particle in a turbulent medium. The distance between two test particles can considerably change in a short time only under the action of a vortex whose size is comparable to this distance. It is this condition that is satisfied in a turbulent medium filled with vortices of various sizes. The larger the separation between particles is, the larger the size of vortices carrying them from each other and the faster the distance $l$ between them increases. In the framework of the classical diffusion theory, this effect can be achieved by introducing the dependence of the diffusion coefficient $D$ on relative coordinates, i.e., on the distance $D=D(l)$. This approach was used by Richardson, who wrote the TD equation for the distribution density $p(l, t)$ of a random distance between two advected particles located at the instant $t=0$ at one point in the form

$$
\frac{\partial p}{\partial t}=\frac{\partial}{\partial l}\left(D(l) \frac{\partial p}{\partial l}\right)
$$

with the diffusion coefficient

$$
D(l) \propto l^{4 / 3}
$$

corresponding to the increase in the diffusion-packet width by the law $\propto t^{3 / 2}$, which considerably differs from the normal diffusion law $t^{1 / 2}$.

The Richardson $4 / 3$ law was theoretically substantiated by Kolmogorov [27, 28] and Obukhov [29, 30] based on the hypothesis of the self-similarity of locally isotropic turbulence determined by a single dimensional parameter, the dissipation rate of the turbulent energy $\varepsilon$. It follows from the dimensional considerations that

$$
D(l)=c \varepsilon^{2 / 3} l^{4 / 3}
$$

These results were in qualitative agreement with experiments. But the fact that the diffusion coefficient in a
homogeneous (on average) medium should depend on a spatial variable gave rise to some inconveniences. The method for combining the accelerated character of TD with a constant coefficient characterizing a medium was proposed by Monin in [31]. That paper can be viewed as the beginning of non-Gaussian unstable distributions and fractional-derivative equations penetrating into the TD theory (this terminology itself was introduced into the TD theory a decade later [32]). Monin considered the diffusion of a cloud of advected particles in a coordinate system associated with the cloud center by expressing the concentration distribution $f(\mathbf{r}, t)$ at an instant $t$ in terms of the initial distribution $f(\mathbf{r}, 0)$ using a time-dependent linear operator $A(t)$ :

$$
f(\mathbf{r}, t)=A(t) f(\mathbf{r}, 0), \quad t>0
$$

In the model of stationary homogeneous locally isotropic turbulence, the operator $A(t)$ can be assumed to be invariant under displacements and rotations of the coordinate system and dependent on the only dimensional parameter $\varepsilon$ (if molecular diffusion is neglected). After the Fourier transformation, the operator $A(t)$ becomes a function $a(k, t, \varepsilon)$ of the modulus $k \equiv|\mathbf{k}|$ of the wave vector $\mathbf{k}$. Based on the dimensional considerations, Monin represents this quantity as a function of $\varepsilon^{1 / 3} k^{2 / 3} t$, such that

$$
\tilde{f}(\mathbf{k}, t)=a\left(\varepsilon^{1 / 3} k^{2 / 3} t\right) \tilde{f}(\mathbf{k}, 0) .
$$

Monin's hypothesis, as Monin and Yaglom write in the second volume of Statistical Fluid Mechanics [32], is based on the assumption that the operators $A(t)$ form a semigroup,

$$
A\left(t_{1}\right) A\left(t_{2}\right)=A\left(t_{1}+t_{2}\right),
$$

and hence

$$
\begin{equation*}
a\left(\varepsilon^{1 / 3} k^{2 / 3} t_{1}\right) a\left(\varepsilon^{1 / 3} k^{2 / 3} t_{2}\right)=a\left(\varepsilon^{1 / 3} k^{2 / 3}\left(t_{1}+t_{2}\right)\right) . \tag{6}
\end{equation*}
$$

The solution of Eqn (6) has the form

$$
a\left(\varepsilon^{1 / 3} k^{2 / 3} t\right)=\exp \left(-c \varepsilon^{1 / 3} k^{2 / 3} t\right),
$$

and the Fourier transformation of the required configuration

$$
\tilde{f}(\mathbf{k}, t)=\exp \left(-c \varepsilon^{1 / 3} k^{2 / 3} t\right) \tilde{f}(\mathbf{k}, 0)
$$

satisfies the differential equation

$$
\frac{\mathrm{d} \tilde{f}(\mathbf{k}, t)}{\mathrm{d} t}=-c \varepsilon^{1 / 3} k^{2 / 3} \tilde{f}(\mathbf{k}, t) .
$$

Using the interpretation of the factor $k^{2 / 3}$ given by Monin and Yaglom ("the Fourier transform of the Laplace operator to the power $1 / 3$ " $[32$, p. 510]), we obtain the equation

$$
\frac{\partial f(\mathbf{r}, t)}{\partial t}=-D_{2 / 3}(-\Delta)^{1 / 3} f(\mathbf{r}, t), \quad D_{2 / 3}=c \varepsilon^{1 / 3}
$$

which belongs to the family of equations

$$
\frac{\partial f(\mathbf{r}, t)}{\partial t}=-D_{\alpha}(-\Delta)^{\alpha / 2} f(\mathbf{r}, t), \quad 0<\alpha<2
$$

Of course, these equations acquire a meaning only after the operators with a fractional exponent $\alpha$ are defined. This
can be done in several ways; the simplest is to define their action through the Fourier transformation of differentiable functions:

$$
\begin{aligned}
& \int \exp (\mathrm{i} \mathbf{k r})(-\Delta)^{\alpha / 2} f(\mathbf{r}, t) \mathrm{d} \mathbf{r}=|\mathbf{k}|^{\alpha} \int \exp (\mathrm{i} \mathbf{k r}) f(\mathbf{r}, t) \mathrm{d} \mathbf{r} \\
& \quad=|\mathbf{k}|^{\alpha} \tilde{f}(\mathbf{k}, t)
\end{aligned}
$$

The inverse transformation (in the $d$-dimensional case) leads to the prescription

$$
\begin{aligned}
& (-\Delta)^{\alpha / 2} f(\mathbf{x}, t)=\frac{2^{\alpha} \Gamma((\alpha+d) / 2)}{\pi^{d / 2} \Gamma(-\alpha / 2)} \int_{\mathrm{R}^{d}} \frac{\left[f(\mathbf{x})-f\left(\mathbf{x}^{\prime}\right)\right] \mathrm{d} \mathbf{x}^{\prime}}{\left|\mathbf{x}-\mathbf{x}^{\prime}\right|^{d+\alpha}} \\
& \mathbf{x} \in \mathrm{R}^{d}, \quad \alpha \in(0,1)
\end{aligned}
$$

where $\Gamma(\ldots)$ is the gamma function. For $\alpha \in(1,2)$, the difference in the numerator is replaced with the secondorder difference. The solutions of these equations are expressed in terms of isotropic stable densities (ISDs):

$$
\begin{aligned}
& \Psi_{3}^{(\alpha)}(r)=\Psi_{3}^{(\alpha)}(\mathbf{r})=\frac{1}{(2 \pi)^{3}} \int \exp (-\mathrm{i} \mathbf{k r}) \exp \left(-|\mathbf{k}|^{\alpha}\right) \mathrm{d} \mathbf{k} \\
& 0<\alpha \leqslant 2: \\
& f(\mathbf{r}, t)=\left[D_{\alpha} t\right]^{-3 / \alpha} \Psi_{3}^{(\alpha)}\left(r\left[D_{\alpha} t\right]^{-1 / \alpha}\right)
\end{aligned}
$$

For $\alpha=2$, this expression represents the normal distribution density corresponding to Brownian motion, and the other values of the parameter $\alpha$ correspond to Lévi motions. Fixing the argument value $r\left[D_{\alpha} t\right]^{-1 / \alpha}=\xi$, we see that the width of this distribution increases with time proportionally to $t^{1 / \alpha}$, i.e., for $\alpha<2$, faster than in the case of normal diffusion, which justifies the introduction of the term 'superdiffusion'.

The further development of the fractional differential approach for describing turbulence is presented in monographs [33-35] and reviews [36, 37].

### 2.4 Percolation

In inhomogeneous media with irregular, porous, large-grain, and winding structures, the slowed down diffusion (subdiffusion) of a liquid can be observed (percolation). There are several models of such a process leading to fractional differential equations [33]. One of them involves a periodic lattice each of whose sites can be either occupied or free with a certain probability. A set of neighboring free points forms a cluster. Lines (paths) connecting these points can be either conducting (with their ends coming out to the cluster surface) or blind. We take one conducting line and stretch it along a straight line, accurately directing other lines coming out of its sites perpendicular to it. Those lines can, in turn, branch or, on the contrary, break; without going into further details, we assume for simplicity that they are unbranching infinite lines. In this way, we obtain a 'comb' (Fig. 1).

We now trace a liquid particle (a point mass) entering the main line. For this, we specify the law of motion of the particle along the line; let it be the ordinary diffusion. Having arrived at the first site, the particle passes to a side branch. According to the diffusion laws, even in the case of an infinite (more exactly, a semi-infinite) branch, the particle will necessarily (i.e., with probability 1 ) return, but the distribution of the return time has a power tail $P(T>t) \propto\left(t / \tau_{0}\right)^{-\beta}$ with the exponent $\beta=1 / 2$ and the characteristic time $\tau_{0}$. Following only the coordinate $x$, we can say that the particle stopped at

b


Figure 1. Percolation process (a) and its comb model (b).
this point for some time (captured in a trap), then continued to diffuse along the $x$ axis with the coefficient $D_{x}$, then was again captured in the same or neighboring trap and remained in it for a different time, etc. In the limit of small distances between traps, such that

$$
\frac{\tau_{1} D_{x}}{\tau_{0}^{\beta}} \rightarrow D>0
$$

where $\tau_{1}$ is the mean diffusion time between hitting the traps, this process is described by the integral equation

$$
f(x, t)=\delta(x)+\frac{1}{\Gamma(\beta)} \int_{0}^{t}(t-\tau)^{\beta-1} D \Delta f(x, \tau) \mathrm{d} \tau, \quad \beta=\frac{1}{2}
$$

whose kernel reflects the delay in the diffusion of the coordinate $x$ caused by the residence of the particle outside this axis.

We recall Cauchy's integral formula representing the $n$-fold integral $I^{n}[f(t)]$ in terms of the ordinary integral,

$$
I^{n}[f(t)]=\frac{1}{(n-1)!} \int_{0}^{t}(t-\tau)^{n-1} f(\tau) \mathrm{d} \tau, \quad n=1,2,3, \ldots
$$

and allowing an analytic continuation to fractional (and even complex) values of the exponent $n$ :

$$
I^{\beta}[f(t)]=\frac{1}{\Gamma(\beta)} \int_{0}^{t}(t-\tau)^{\beta-1} f(\tau) \mathrm{d} \tau, \quad \beta>0 .
$$

Introducing the Riemann-Liouville fractional derivatives ${ }^{2}$

$$
\frac{\partial^{\beta}}{\partial t^{\beta}} f(t)=\frac{\partial}{\partial t} I^{1-\beta}(t)=\frac{1}{\Gamma(1-\beta)} \frac{\partial}{\partial t} \int_{0}^{t} \frac{f(\tau) \mathrm{d} \tau}{(t-\tau)^{\beta}}, \quad 0<\beta<1,
$$

and applying them to the integral equation written above, we obtain the fractional differential subdiffusion equation

$$
\frac{\partial^{\beta} f(x, t)}{\partial t^{\beta}}=D \Delta f(x, t)+\delta(x) \delta_{\beta}(t)
$$

where $\delta_{\beta}(t)=t^{-\beta} / \Gamma(1-\beta)$.
We make some remarks. First, the 'fractional delta function' is called so because it is a continuation of the known definition of the delta function as the derivative (in

[^0]the generalized sense) of the Heaviside function
\[

1_{+}(t)= $$
\begin{cases}0, & t<0, \\ 1, & t>0,\end{cases}
$$
\]

namely,

$$
\delta_{\beta}(t)=\frac{\mathrm{d}^{\beta} 1_{+}(t)}{\mathrm{d} t^{\beta}}
$$

Second, a solution of the subdiffusion equation has the meaning of a probability density, and $\delta_{\beta}(t)$ ensures that the integral of the solution is constant in $x$ (normalization). Finally, the fractional character of the derivative is here a consequence of the infinite length of side branches. If their length is limited, we again have the first-order time derivative instead of a fractional derivative in the $x$-diffusion equation, i.e., the usual diffusion equation (albeit with the diminished diffusion coefficient).

The problem of the percolation of a liquid through a porous medium (which might look like a very particular problem) attracted the attention of researchers in different fields (including cosmic-ray physics) because the percolation process turned out to be critical. A signature of this important property is the existence of a number - a percolation threshold, the minimum density of free sites of the lattice above which the liquid percolates over the entire infinite lattice and below which the liquid occupies only a finite region of the medium. For densities close to the threshold density, percolation occurs over a fractal set, and this process is governed exclusively by the laws of criticality, irrespective of the macroscopic properties of the medium [34, 38]. A deep connection between the percolation model and cosmic electrodynamics (the multiscale interaction of fields and currents in the distant Earth's magnetotail, self-organization processes in magnetized plasmas, the evolution of large-scale magnetic fields in the solar photosphere and interstellar space, and the construction of a self-consistent model of the turbulent current sheet) is demonstrated in the remarkable review by Zelenyi and Milovanov [39].

### 2.5 Combined equation and its interpretation

We now consider a combined equation containing fractional derivatives with respect to coordinates (a fractional Laplacian) and time:
$\frac{\partial^{\beta} G}{\partial t^{\beta}}=-D_{\alpha}(-\Delta)^{\alpha / 2} G(\mathbf{r}, t)+\delta(\mathbf{r}) \delta_{\beta}(t), \quad \alpha \in(0,2], \quad \beta \in(0,1]$.

We note that the dimension of each terms in this equation is $L^{-3} T^{-\beta}$, and the coefficient $D_{\alpha}$ replacing the diffusion coefficient and having the dimension $L^{\alpha} T^{-\beta}$ is strictly speaking the diffusion coefficient only for $\alpha=2$ and $\beta=1$. Typically, Eqn (7) is derived from the random walk pattern (see, e.g., [40]). To 'read' the physical content of this equation, we should go in the reverse direction. For this, we represent (7) in the Fourier-Laplace variables:

$$
\begin{equation*}
\left(\lambda^{\beta}+D_{\alpha}|\mathbf{k}|^{\alpha}\right) \tilde{G}(\mathbf{k}, \lambda)=\lambda^{\beta-1} . \tag{8}
\end{equation*}
$$

The right-hand side of Eqn (8) can be easily understood without additional calculations. Indeed, for $\mathbf{k}=0$, this equation takes the form

$$
\lambda^{\beta} \tilde{G}(0, \lambda)=\lambda^{\beta-1}
$$

where

$$
\begin{aligned}
\tilde{G}(0, \lambda) & =\int_{0}^{\infty} \exp (-\lambda t)\left[\int G(\mathbf{r}, t) \mathrm{d} \mathbf{r}\right] \mathrm{d} t \\
& =\int_{0}^{\infty} \exp (-\lambda t) \mathrm{d} t=\frac{1}{\lambda}
\end{aligned}
$$

(the spatial integral of the spatial distribution density in square brackets is unity, according to the normalization conditions). Multiplying both sides of Eqn (8) by a positive constant $B$ and setting $C=B D_{\alpha}$, we transform the operator in the left-hand side:

$$
\begin{aligned}
B \lambda^{\beta}+C|\mathbf{k}|^{\alpha} & =1-\left(1-B \lambda^{\beta}-C|\mathbf{k}|^{\alpha}\right) \\
& =1-\left(1-B \lambda^{\beta}\right)\left(1-C|\mathbf{k}|^{\alpha}\right)+B C \lambda^{\beta}|\mathbf{k}|^{\alpha} .
\end{aligned}
$$

As $|\mathbf{k}| \rightarrow 0$ and $\lambda \rightarrow 0$, we have

$$
B \lambda^{\beta}+C|\mathbf{k}|^{\alpha} \sim 1-\left(1-B \lambda^{\beta}\right)\left(1-C|\mathbf{k}|^{\alpha}\right),
$$

and the substitution of the last expression in Eqn (8),

$$
\left[1-\left(1-B \lambda^{\beta}\right)\left(1-C|\mathbf{k}|^{\alpha}\right)\right] \tilde{G}(\mathbf{k}, \lambda)=B \lambda^{\beta-1},
$$

brings it to the form

$$
\begin{equation*}
\tilde{G}(\mathbf{k}, \lambda)=\left(1-B \lambda^{\beta}\right)\left(1-C|\mathbf{k}|^{\alpha}\right) \tilde{G}(\mathbf{k}, \lambda)+B \lambda^{\beta-1} . \tag{9}
\end{equation*}
$$

The content of the two parentheses in the first term in the right-hand side can be treated as the asymptotic expression for the characteristic functions of the temporal and spatial probability densities $q(t)$ and $p(\mathbf{r})$ :

$$
\begin{equation*}
\hat{q}(\lambda) \equiv \int_{0}^{\infty} \exp (-\lambda t) q(t) \mathrm{d} t \sim 1-B \lambda^{\beta}, \quad \lambda \rightarrow 0, \quad \beta \leqslant 1, \tag{10}
\end{equation*}
$$

and, with an isotropic distribution $p(\mathbf{r}) \mathrm{d} \mathbf{r}=p_{R}(r) \mathrm{d} r \mathrm{~d} \boldsymbol{\Omega} / 4 \pi$,

$$
\begin{equation*}
\tilde{p}(\mathbf{k}) \equiv \int \exp (-\mathrm{i} \mathbf{k r}) p(\mathbf{r}) \mathrm{d} \mathbf{r} \sim 1-C|\mathbf{k}|^{\alpha}, \quad \mathbf{k} \rightarrow 0, \quad \alpha \leqslant 2 . \tag{11}
\end{equation*}
$$

We also note that
$B \lambda^{\beta-1} \sim \tilde{Q}(\lambda)=\int_{0}^{\infty} \exp (-\lambda t) Q(t) \mathrm{d} t, \quad Q(t)=\int_{t}^{\infty} q(t) \mathrm{d} t$.
Replacing the expressions in the parentheses and the free term in Eqn (9) with the left-hand sides of expressions (10)(12),

$$
\tilde{G}(\mathbf{k}, \lambda)=\tilde{p}(\mathbf{k}) \hat{q}(\lambda) \tilde{G}(\mathbf{k}, \lambda)+\hat{Q}(\lambda),
$$

and performing the inverse Fourier-Lorentz transformation, we obtain the integral equation with a factored kernel:
$G(\mathbf{r}, t)=\int \mathrm{d} \mathbf{r}^{\prime} \int_{0}^{\infty} \mathrm{d} t^{\prime} p\left(\mathbf{r}^{\prime}\right) q\left(t^{\prime}\right) G\left(\mathbf{r}-\mathbf{r}^{\prime}, t-t^{\prime}\right)+Q(t) \delta(\mathbf{r})$.

Of course, the densities $p(\mathbf{r})$ and $q(t)$ are not uniquely determined by the asymptotic form of their transforms (10) and (11), but, being probability densities, they are nonnegative and satisfy the normalization conditions $\int_{R^{3}} p(\mathbf{r}) \mathrm{d} \mathbf{r}=1$ and $\int_{0}^{\infty} q(t) \mathrm{d} t=1$ and, in addition, Tauberian theorems relating the asymptotic behavior of transforms
in the vicinity of zero to the long-distance asymptotic form of the originals,

$$
\begin{aligned}
& 1-\tilde{p}(\mathbf{k}) \propto|\mathbf{k}|^{\alpha}, \quad|\mathbf{k}| \rightarrow 0 \Leftrightarrow p(\mathbf{r}) \propto|\mathbf{r}|^{-\alpha-3}, \quad|\mathbf{r}| \rightarrow \infty, \\
& 1-\hat{q}(\lambda) \propto \lambda^{\beta}, \quad \lambda \rightarrow 0 \Leftrightarrow q(t) \propto t^{-\beta-1}, \quad t \rightarrow \infty,
\end{aligned}
$$

determine a power-law behavior of the intermediate densities for large spatial and temporal arguments. ${ }^{3}$

The physical meaning of prelimit Eqn (13) (with respect to fractional differential) is already clear: representing the solution of this integral equation in the form of a Neumann series,

$$
G(\mathbf{r}, t)=Q(t) \delta(\mathbf{r})+\int_{0}^{t} \mathrm{~d} t^{\prime} Q\left(t-t^{\prime}\right) p(\mathbf{r}) q\left(t^{\prime}\right)+\ldots
$$

we see that the probability of finding the particle at an instant $t$ at a point $\mathbf{r}$ is the sum of the probability of the permanent residence of the particle at the initial point without displacements (the first term) and the probability that the particle performs an instant hop at one of the intermediate instants $t^{\prime} \in(0, t)$ from the creation point $\mathbf{r}_{0}=0$ to the observation point $\mathbf{r}$ and remains there until the observation instant $t$; the next term would give the probability of finding a particle that has performed two instant hops separated by a random time interval with the density $q(t)$, etc. The lengths of these hops are random, mutually independent, and independent of the residence times of the particle at the traps. This hopping process is called the continuous-time random walk (CTRW). Applying the Fourier-Laplace transformation to Eqn (13), without imposing any conditions on the intermediate probabilities $p$ and $q$, we obtain the known expression [41]

$$
\begin{equation*}
\tilde{G}(\mathbf{k}, \lambda)=\frac{\hat{Q}(\lambda)}{1-\tilde{p}(\mathbf{k}) \hat{q}(\lambda)}=\frac{1-\hat{q}(\lambda)}{\lambda(1-\tilde{p}(\mathbf{k}) \hat{q}(\lambda))} . \tag{14}
\end{equation*}
$$

For $\alpha=2$ and $\beta=1$, Eqn (7) describes Brownian motion (Fig. 2a), and for $\alpha<2$ and $\beta=1$, it describes Levi motion, which differs from Brownian motion by numerous discontinuities of the trajectory along the $x$ axis (which is indicated by horizontal segments corresponding to instant hops by finite distances). As long as the derivative with respect to $x$ has an integer (second) order, discontinuities are absent (only kinks exist), and when the order of the derivative becomes fractional, discontinuities appear. The distribution of the coordinate jumps at these discontinuities is described by a power law, and therefore the discontinuities are observed at different scales. A process in which a diffusion packet expands faster than in the normal case, i.e., proportionally to $t^{\gamma}$ with $\gamma>1 / 2$, belongs to the class of superdiffusion processes.

Replacing the first time derivative $(\beta=1)$ with a fractional derivative $(\beta<1)$ gives rise to discontinuities of the trajectory along the time axis (the particle stops; time goes on while the particle does not move). For $\alpha<1$ and $\beta<1$, instead of a trajectory, we see point clusters separated by hopping regions (horizontal segments) and residences at traps (vertical segments) (Figs 2b, c) [25]).

### 2.6 Fractionally stable statistics

The self-similar solution of Eqn (7) is expressed in terms of special functions, which we called isotropic fractionally stable

[^1]

Figure 2. Typical realizations of the $x-t$ trajectories of particles in the three models under study.


Figure 3. One-dimensional (a) symmetric and (b) one-sided stable densities.
densities (IFSDs) and studied in a number of papers (see [40] and the references therein). We let the IFSD be denoted by $\Psi_{d}^{(\alpha, \beta)}(r)$ (where $d$ is the space dimension), such that

$$
\begin{align*}
& G(\mathbf{x}, t)=\left[D t^{\beta}\right]^{-3 / \alpha} \Psi_{d}^{(\alpha, \beta)}\left(|\mathbf{x}|\left[D t^{\beta}\right]^{-1 / \alpha}\right),  \tag{15}\\
& 0<\alpha \leqslant 2, \quad 0<\beta \leqslant 1
\end{align*}
$$

Isotropic fractionally stable densities have no general explicit expression in terms of elementary functions but are defined by the characteristic function

$$
\begin{equation*}
\tilde{\Psi}_{d}^{(\alpha, \beta)}(\mathbf{k})=\int_{0}^{\infty} \tilde{\Psi}_{d}^{(\alpha)}\left(|\mathbf{k}| \tau^{-\beta / \alpha}\right) g_{+}(\tau ; \beta) \mathrm{d} \tau, \tag{16}
\end{equation*}
$$

where $g_{+}(\tau ; \beta)$ is the one-sided stable density defined by the Laplace transform:

$$
\int_{0}^{\infty} \exp (-\lambda \tau) g_{+}(\tau ; \beta) \mathrm{d} \tau=\exp \left(-\lambda^{\beta}\right)
$$

Because $g_{+}(\tau ; 1)=\delta(\tau-1)$, the IFSD class includes the family

$$
\begin{equation*}
\tilde{\Psi}_{d}^{(\alpha)}(\mathbf{k}) \equiv \tilde{\Psi}_{d}^{(\alpha, 1)}(\mathbf{k})=\exp \left(-|\mathbf{k}|^{\alpha}\right) \tag{17}
\end{equation*}
$$

as a subset, in particular,

$$
\begin{equation*}
\tilde{\Psi}_{d}^{(2)}(\mathbf{k}) \equiv \tilde{\Psi}_{d}^{(2,1)}(\mathbf{k})=\exp \left(-|\mathbf{k}|^{2}\right) \tag{18}
\end{equation*}
$$

To clarify the role of fractionally stable laws in the hierarchy of probability distributions, we consider expressions (16)-(18). Expression (18), which is the characteristic function of the normal (Gaussian) distribution with dispersion 2, characterizes the limit distribution of the properly normalized sums of a fixed (nonrandom) number of independent, identically distributed random vectors with a finite second moment. Expression (17) is related to a similar sum of vectors with an infinite second moment, but with a power-law distribution tail, $P(|\mathbf{R}|>r) \propto r^{-\alpha}$; the corresponding limit distributions are expressed in terms of isotropic stable densities, which are characterized by the dispersion divergence. Finally, expression (16) characterizes the asymptotic behavior (as $t \rightarrow \infty$ ) of the distribution of the sum of a random number of such terms. This extension of the family of stable laws to fractionally stable ones provides a representation of solutions of fractional differential equations.

One-dimensional symmetric and one-sided stable densities are represented in Fig. 3. For $\alpha=2$, the functions $\Psi_{1}^{(2)}(\xi)$ and $\Psi_{3}^{(2)}(\xi)$ represent the one-dimensional and three-dimen-


Figure 4. Three-dimensional isotropic stable densities. Curves $1-8$ (from top down) correspond to $\alpha=0.3,0.4,0.6,0.8,1.0,1.2,1.7$, and 2.0 .
sional normal (Gaussian) densities with the dispersion equal to 2 :

$$
\begin{aligned}
& \Psi_{1}^{(2)}(\xi)=\frac{1}{2 \sqrt{\pi}} \exp \left(-\frac{\xi^{2}}{4}\right), \quad-\infty<\xi<\infty, \\
& \Psi_{3}^{(2)}(\xi)=\frac{1}{(2 \sqrt{\pi})^{3}} \exp \left(-\frac{\xi^{2}}{4}\right), \quad \xi>0 .
\end{aligned}
$$

As $\alpha$ decreases, the distributions become narrower and reach higher values in the central part, and, at the same time, an increasing part of the probability goes from the intermediate region to the tails. For $\alpha=1$, we have the Cauchy distribution densities

$$
\begin{aligned}
\Psi_{1}^{(1)}(\xi) & =\frac{1}{\pi\left(1+\xi^{2}\right)}, \quad-\infty<\xi<\infty \\
\Psi_{3}^{(1)}(\xi) & =\frac{1}{\left[\pi\left(1+\xi^{2}\right)\right]^{2}}, \quad \xi>0
\end{aligned}
$$

For other values of $\alpha$, stable densities cannot be expressed in terms of elementary functions. At large distances, they decrease in accordance with a power law,

$$
\Psi_{3}^{(\alpha)}(\xi) \sim \frac{\alpha 2^{\alpha-1}}{\pi^{3 / 2}} \frac{\Gamma((\alpha+3) / 2)}{\Gamma(1-\alpha / 2)} \xi^{-(\alpha+3)}, \quad \xi \rightarrow \infty
$$

and are described by a bell-shaped contour centered at the coordinate origin with the maximum

$$
\Psi_{3}^{(\alpha)}(0)=\frac{\Gamma(1+3 / \alpha)}{(4 \pi)^{3 / 2} \Gamma(1+3 / 2)}
$$

In the double logarithmic scale, ranging many orders of magnitude, we observe a horizontal plateau changing into a sloping part on the right (Fig. 4). Fractionally stable densities $\Psi_{d}^{(\alpha, \beta)}(\xi)$ retain the same asymmetry type as $r \rightarrow \infty$, but, unlike stable densities, have a power-law singularity at zero:

$$
\Psi_{3}^{(\alpha, \beta)}(\xi) \propto \begin{cases}\xi^{-(3-\alpha)}, & \xi \rightarrow 0 \\ \xi^{-(3+\alpha)}, & \xi \rightarrow \infty\end{cases}
$$

Correspondingly, the plot in the double logarithmic scale is given by two half-lines with different continuously joining slopes (Fig. 5).


Figure 5. Three-dimensional isotropic fractionally stable densities, $\beta=0.8$. Curves $1-5$ correspond to $\alpha=0.3,0.5,1.0,1.4$, and 1.7.

To date, a large set of physical phenomena demonstrating power-law distributions has been collected [33]. We here consider a phenomenon involving a fractionally stable distribution of particles diffusing in a turbulent medium: the results of numerical simulation of the motion of charged particles in an infinitely long cylinder filled with plasma in a constant homogeneous magnetic field directed along the cylinder axis [43]. The plasma turbulence excited by the noise component of pressure was described by a system of magnetohydrodynamic equations for fluctuating components of the pressure and electrostatic potential. The magnetic field was assumed fixed. The nondiffusion transport of charged particles in such a medium is caused only by random fluctuations of the electrostatic potential produced by the generator of the noise component of pressure. Thus, the problem was reduced to the calculation of the diffusion of particles in the field of random velocities $\mathbf{V}$, which determined the statistical properties of an ensemble of random trajectories by the equation

$$
\frac{\mathrm{d} \mathbf{r}}{\mathrm{~d} t}=\mathbf{V}(\mathbf{r}, t)
$$

Particular details of the model and calculation procedure can be found in [43] and the references therein. Here, we are only interested in the conclusion made by the authors of the paper. They observed the behavior of $25 \times 10^{3}$ particles in a numerical experiment, which were initially homogeneously distributed over the side surface $r=a / 2$ of a cylinder with the radius $r=a$; having obtained the particle distributions over the variable $x=(r-a / 2) / a$ for several successive instants of time, the authors found that:

- The particle distribution density over the variable $x$ at different instants is well approximated by the self-similar function $t^{-v} f\left(x t^{-v}\right), v=0.66 \pm 0.20$.
- The function $f(\xi)$ is symmetric with respect to $\xi=0$ and its tails decay as $|\xi|^{-\alpha-1}, \alpha \approx 3 / 4$.
- The density at a fixed point $x$ increases at small times as $\sim t^{\beta}$ and decreases at large times as $t^{-\beta}$, where $\beta \approx 1 / 2$.

A comparison of the results of these calculations (Fig. 6) with our representation of this distribution in terms of the fractionally stable density,

$$
P(x, t)=\left[D_{\alpha} t^{\beta}\right]^{-1 / \alpha} \Psi_{1}^{(\alpha, \beta)}\left(x\left[D_{\alpha} t^{\beta}\right]^{-1 / \alpha}\right), \quad \alpha=\frac{3}{4}, \quad \beta=\frac{1}{2},
$$

demonstrates good agreement. The authors of [43] interpret the meaning of the fractional differential equation in the


Figure 6. Fractionally stable distribution in plasma. Triangles show the results of numerical simulation [43]. The curve is the solution of the onedimensional fractional differential equation expressed in terms of the $\operatorname{IFSD} \Psi_{1}^{(3 / 4,1 / 2)}(x)$.
framework of a CTRW process assuming that particles entering a vortex can remain inside it for a long time until they are ejected from it and pass by several vortices during one flight, until their capture in a next trap.

The impressive coincidence of the simulation results with the solution of a fractional differential equation does not eliminate all doubts, however. For $\alpha<1$, the mean free path is infinite, and the authors of [43] report that the displacement momenta behave in the 'superdiffusion way', $\left\langle X^{n}(t)\right\rangle \propto t^{n \beta / \alpha}$. The spatial fractional operator is written for an infinite medium, although a finite size of the cylinder cannot be ignored for such $\alpha$. The assumption about the infinitely large velocity of flights is also incompatible with this value of $\alpha$. Finally, turbulence must have some special structure for successive passages to be considered statistically independent.

Another example of a non-Gaussian statistics in a plasma, involving a velocity distribution, is presented in Section 5.3.

## 3. Isotropic anomalous diffusion

### 3.1 Traps and voids

Fermi's picture of magnetic clouds, which was mentioned in the Introduction, was later supplemented by the concept of magnetic traps, where charged particles can stay for a long time. Dorman writes in [44, p. 53]:
"Cosmic rays in the cosmos are confined, in fact, in magnetic traps of one scale or another, not propagating freely in space (except for cosmic-ray gamma quanta and neutrinos, for which traps are absent). Giant traps for cosmic rays exhibit a wide variety of properties, and the behavior of charged particles in them essentially depends on the particle energy. A trap in the vicinity of Earth formed by a magnetic field close to a dipole field is highly stable, and the lifetime of particles in it is long. At the same time, traps in the vicinity of chromospheric flares or in solar corpuscular fluxes of magnetized plasmas are much more transparent to particles and the particles escape from them similarly to how diffusion in irregular magnetic fields occurs. Traps of various types are also formed in the vicinity of usual stars, in particular, in the solar system and in supernova shells. On the other hand, the


Figure 7. Passing from (a) Brownian diffusion in a continuous homogeneous medium to (b) random walks in the medium alternating with voids and an example of the Levi random walk trajectory obtained by the Monte Carlo method, with $\alpha=1.67$ (c).
galaxy (the galactic disc together with the halo) also forms a certain type of trap a few thousand parsecs in size, which well confines moderate- and high-energy particles (with a lifetime of $\sim 10^{7}$ years) and is quite transparent to ultra-high-energy particles. It is quite possible that galactic clusters form even more gigantic traps for ultra-high-energy particles.
... it seems reasonable to treat any magnetic regions where the motion and lifetime of charged particles considerably differ from those in free space of the same volume as cosmic magnetic traps."

For such a large spread in the size of objects, it is impossible to imagine their uniform or even independent position in space. The distribution of visible matter in space (star clusters, galaxies, and galactic clusters) produces examples of hierarchic structures, which approximately preserve their inhomogeneity type upon varying scales in a broad range. The mathematical model of such inhomogeneities, which cannot be smoothed by scale transformations, uses fractals characterized by power-law correlations of spatial structures. At the end of the 20th century, a new avenue analyzing the structure of magnetic fields based on the fractal concept emerged in the astrophysics of the interstellar medium [45].

To elucidate the relation between the CTRW model and the real cosmic-ray transfer process in a galactic magnetic field, we consider a homogeneous (diffusion-wise) medium and divide it into cubic volumes (cells) (Fig. 7). We assign the coordinate $\mathbf{r}_{i}$ to each particle entering the $i$ th cell centered at the point $\mathbf{r}_{i}$ at an instant $t$ and leaving it at an instant $t+T$, neglecting the motion of the particle inside the cell itself. After some (random) time $T$, the particle moves to one of the six neighboring cells, and the vector assigned to the particle moves from the center of the previous cell to the center of this new cell at the instant of the intersection of the face separating the cells. After a random time $T^{\prime}$, the particle moves to another neighboring cell, and so on. Thus, in a coarse-grain description, the particle coordinate moves jumpwise over three-dimensional lattice sites, staying in them for a random time. At a large scale, we then see a random walk, which is very close to Brownian motion (which it is, in fact).

However, a strong turbulent magnetic field does not necessarily exist in each cell. A considerable part of space between magnetic clouds is filled with weaker and quieter fields with magnetic field lines running smoothly over large distances. Charged particles in cosmic rays move along these lines and are sometimes captured in cloud traps, where they can be confined for a long time, forgetting their initial direction. To construct such a model, we remove part of the elements (cells), keeping others in their places. The passage
from one cell to another is then no longer instantaneous, as was the case with crossing the face between neighboring cells, and the particle passes through 'almost empty' cells, propagating over random distances $R$. The time spent for these passages is proportional to the distance propagated in free space (the mean free path). The character of the entire process depends on the mutual arrangement of the remaining parts. If they are scattered homogeneously (in the statistical sense) and are independent of each other, as in the Bershadskii model [46], we obtain a normal process with a larger diffusion coefficient. If they are arranged in a fractal manner, such that clusters and voids are observed at different scales (the method for constructing such distributions is shown in [47]), the particles can move over long distances in one run. In that case, the resulting process is determined by the order of the converging moments of a single run. We say that an anomalous diffusion process is a process of the first kind if the mean free path $\langle R\rangle=\infty$, and a process of the second kind if $\langle R\rangle<\infty$ and $\left\langle R^{2}\right\rangle=\infty$. Below, we see a significant difference between these two types of anomalous diffusion.

We must note, however, that these considerations, despite their intentionally schematic and illustrative character, are in qualitative agreement with the picture of magnetic fields randomly spread in space, which was already proposed by Fermi. Quantitative agreement can probably be achieved by specifying (i) the appropriate distributions of clouds producing the corresponding distributions of mean free paths of particles between them and (ii) the laws of interaction of particles with individual clouds. In such a formulation, the problem can be considered in the framework of the standard multiple scattering theory, which, in particular, allows investigating the interaction of a particle with a cloud (the scattering cross section) and the transfer itself of particles in space (solving the kinetic equation) separately. In passing to the diffusion limit, the integro-differential kinetic equation is transformed into a differential diffusion equation, while the interaction cross section containing information on the mechanisms of this interaction (resonance interaction, scattering by Alfvén and magnetosonic waves, etc.) is transformed into the diffusion coefficient, which then accumulates this information.

As regards the kinetics themselves, there are no grounds to assume that magnetic clouds are arranged homogeneously and move independently of each other. In that case, it would be reasonable to adopt an exponential distribution law for transits between scattering events. An example is given by an ideal gas, whose molecules do not interact with each other, which leads to the independence and exponential range distribution. But measurements of the electromagnetic radiation of the charged component of cosmic rays show that the interstellar region is not an ideal gas and is characterized by long-range power-law correlations, and this can be manifested in the higher probability of the long transits of particles propagating through voids. The mathematical model of such a process is already developed. This is Lévy motion: the random walk of a particle with an asymptotically power-law range distribution. An example of the trajectory of such a process is presented in Fig. 7c. We can see that the increase in the fraction of long free paths is accompanied by the increase in the fraction of short ones. This occurs due to the decrease in the probability of intermediate-length free paths. Now we can say that a trajectory consists of clusters of short free runs separated by long ones. The clusters of short free runs localized in space are capable of simulating the behavior of a
particle in cells, which we mentioned above. At the same time, stable laws governing Lévy motion are directly and rigorously related to fractional derivatives [23, 25].

All this appeared quite intriguing and stimulated us 13 years ago to develop a fractional differential model of the transfer of particles in the Galaxy.

### 3.2 First studies with a fractional differential model

Among the first studies on galactic cosmic-ray transfer, whose results can already be explained in the framework of the fractional differential approach, were preprint [14] (although the authors did not use the relevant term) and paper [48]. We discuss preprint [14] in Section 4.4, and here consider the results in [48].

As pointed out by the authors of [48], although galactic magnetic fields are located predominantly in the galactic plane (see [49]) and charged particles with small Larmor radii (compared to mean free paths) move along these lines, only slightly diffusing in transverse directions, the observed angular distribution of particles has a surprisingly high isotropy degree. The authors explain this by large-scale fluctuations of the interstellar fields and by exponential divergence of magnetic field lines. An important mechanism of active mixing of cosmic rays is also their acceleration at the leading edges of shock waves from supernovae - cosmic tsunamis that shake the particles off their field lines and bring them to a chaotic regime [50-52]. Under the action of these and other factors, particles that initially moved along 'their' field lines lose the connection with them after some correlation time and enter the isotropic diffusion regime. The turbulent character of the interstellar medium, which is manifested in the alternation of weakly irregular magnetic fields with randomly dispersed islands (clouds, regions) of strong fluctuations, affects only the distributions of free paths and residence times of particles in different regions. The authors of [48] generalized the standard diffusion model to the CTRW model with power-law distributions of free paths and lifetimes in traps corresponding to (10) and (11) and wrote the mean square of the resulting distribution at the observation instant $t$ in the form
$\int|\mathbf{r}|^{2} G(\mathbf{r}, t) \mathrm{d} \mathbf{r}=\frac{1}{2 \pi \mathrm{i}} \frac{1}{2 \pi} \int_{\Gamma} \mathrm{d} \lambda \int \mathrm{d} \mathbf{k} \int \mathrm{d} \mathbf{r}|\mathbf{r}|^{2} \exp (\lambda t-\mathrm{i} \mathbf{k r}) \tilde{G}(\mathbf{k}, \lambda)$,
where

$$
\begin{equation*}
\tilde{G}(\mathbf{k}, \lambda)=\frac{B \lambda^{\beta-1}}{B \lambda^{\beta}+C|\mathbf{k}|^{\alpha}} \tag{20}
\end{equation*}
$$

in our notation. The last expression is consistent with expression (8) representing fractional differential equation (7) in Fourier-Laplace variables. The authors justify the choice of representations (10) and (11) by the existence of 'stability islets' (traps) in the plasma capturing particles for a long time, thereby slowing down the diffusion process [52], and by the property of field lines to perform rapid (compared to diffusion) and long-range flights, preserving their direction [8,53]. The competition between these two processes is reflected in the distribution obtained.

Two remarks concerning paper [48] are in order. ${ }^{4}$ First, expression (19) makes no sense for $\alpha \neq 2$, because the integral

[^2]diverges and the packet width should be described differently. However, the authors consider only the case $\alpha=2$, and no objections can be raised here. The seconds remark concerns the range of the parameter $\beta$ (denoted by $\alpha$ in their paper). The authors solved the problem for synchrotron radiation of electrons with $\beta=0.5,1.0$, and 1.5 , and pointed out that in contrast to papers [14-16], they extended the range of $\beta$ to the region $\beta>1$, and this extension can be regarded as the passage from subdiffusion to superdiffusion. At first glance, such an interpretation is quite reasonable: $\beta<1$ corresponds to subdiffusion and $\beta=1$ to normal diffusion, and therefore $\beta>1$ should correspond to superdiffusion, and the authors had no doubts about it. However, a reason to doubt existed: it suffices to inspect expression (14) in [48], which has the form
\[

$$
\begin{equation*}
q(t) \sim \frac{1}{\tau} \frac{\beta}{\Gamma(1-\beta)}\left(\frac{t}{\tau}\right)^{-\beta-1}, \quad t \rightarrow \infty \tag{21}
\end{equation*}
$$

\]

Obviously, a function with such an asymptotic behavior can never be a probability density for $\beta>1$, simply because it is negative $(\Gamma(1-\beta)<0)$. However, the issue is not that $q(t)$ has a power-law asymptotic behavior, because nothing prevents us from taking it, say, in the form

$$
\begin{equation*}
q(t) \sim \frac{\beta^{\prime}}{\tau}\left(\frac{t}{\tau}\right)^{-\beta^{\prime}-1}, \quad t \rightarrow \infty \tag{22}
\end{equation*}
$$

where $\beta^{\prime}$ is any positive number, but that expression (2) is obtained from expansion (10), which is valid only for $\beta^{\prime}=\beta \leqslant 1$. If $1<\beta^{\prime} \leqslant 2$, then the expansion

$$
\tilde{q}(\lambda) \sim 1-\langle T\rangle \lambda+B^{\prime} \lambda^{\beta^{\prime}}, \quad \lambda \rightarrow \infty
$$

holds instead of (10). For small $\lambda$ corresponding to long times, the term with $\lambda^{\beta^{\prime}}$ is asymptotically small compared to the linear term and can be omitted. Hence, for $\beta^{\prime}>1$ and $\alpha=2$, we obtain (and the authors of [48] should have obtained) normal diffusion, rather than superdiffusion. ${ }^{5}$ The latter would only follow for $\alpha<2 \beta$. However, the question can arise: what about the case $\alpha>1$ ? Should a term proportional to a gradient appear in this case, in addition to the Laplacian? Indeed, this term should appear, but $\langle\mathbf{R}\rangle=0$ for isotropic diffusion, and this term safely disappears. The duration of the interval $T$ cannot be negative; therefore, assuming $\langle T\rangle=0$, we automatically assume $T=0$, i.e., we eliminate the traps. Because the particle velocity is infinite, the particle immediately escapes to infinity in the absence of traps, and we do not observe the time sweep of the process.

If for some reason we wish to see a fractional time derivative of an order $\beta \in(1,2)$ in the equation, then, as follows from the previous expansion, this derivative should be introduced together with the first derivative:

$$
\begin{aligned}
& \langle T\rangle \frac{\partial G(x, t)}{\partial t}+B \frac{\partial^{\beta} G(x, t)}{\partial t^{\beta}}+C D_{x}^{\alpha / 2} G(x, t) \\
& \quad=\langle T\rangle \delta(t) \delta(x)+B \delta_{\beta}(t) \delta(x), \quad 1<\beta<2 .
\end{aligned}
$$

[^3]Now everything is in place. Only we should bear in mind that the long-term asymptotic form of the solution already determines an equation with a lower-order time derivative, i.e., the asymptotic solution is still characterized by $\beta=1$.

In our first paper written in the framework of the fractional differential model [19], we discussed the nature of a 'knee' in the energy spectrum of primary cosmic radiation at $E \sim 3 \times 10^{15} \mathrm{eV}$. The fractional differential character of the diffusion equation was substantiated in the following way.

Based on the standard diffusion equation

$$
\frac{\partial N}{\partial t}=D \Delta N(\mathbf{r}, t)+S(\mathbf{r}, t)
$$

this knee is related to the decrease in the confinement efficiency of high-energy particles in the Galaxy, which in turn requires a long lifetime ( $10^{7}-10^{8}$ years) of the protonnuclear component in the system and the presence of remote sources [6]. But replacing the diffusion equation with the superdiffusion equation involving the fractional Laplacian,

$$
\begin{equation*}
\frac{\partial N}{\partial t}=-D_{\alpha}(-\Delta)^{\alpha / 2} N(\mathbf{r}, t)+S(\mathbf{r}, t) \tag{23}
\end{equation*}
$$

showed the steepening of the spectrum even without a special assumption about the leakage of particles, if the sources of particles are explosions of supernovae nearest to the Solar System. In this case, the knee appears due to the fractional Laplacian, caused by a power-law distribution of free paths, which can in turn be interpreted as a result of averaging the exponential distribution of free paths over different-scale (fractal) fluctuations of the interstellar magnetic field (see, e.g., $[54,55])$.

In $[19,20]$, we introduced Eqn (23) to describe the cosmicray transfer as the limit of a jumpwise process in the threedimensional space, described by the integro-differential equation

$$
\frac{\partial f_{\epsilon}}{\partial s}+\sigma f_{\epsilon}(\mathbf{r}, s)=\sigma \int f_{\epsilon}\left(\mathbf{r}^{\prime}, s\right) W\left(\frac{\mathbf{r}-\mathbf{r}^{\prime}}{\epsilon}\right) \frac{\mathrm{d} \mathbf{r}^{\prime}}{\epsilon^{3}}+S_{\epsilon}(\mathbf{r}, t)
$$

where $\epsilon$ is an auxiliary parameter, sending which to zero provides (in the normal case) the passage from the CTRW scheme to the diffusion approximation. This equation can also be written in the form

$$
\frac{\partial f_{\epsilon}}{\partial s}=\sigma \int\left[f_{\epsilon}\left(\mathbf{r}^{\prime}, s\right)-f_{\epsilon}(\mathbf{r}, s)\right] W\left(\frac{\mathbf{r}-\mathbf{r}^{\prime}}{\epsilon}\right) \frac{\mathrm{d} \mathbf{r}^{\prime}}{\epsilon^{3}}+S_{\epsilon}(\mathbf{r}, t)
$$

The equation in the first form describes the process of independent instantaneous jumps separated by random time intervals and distributed in accordance with the exponential law with the mean $1 / \sigma$. The kernel $W$ of the integral operator plays the role of the distribution density of the displacement vector in such a jump, and should therefore be integrable. For the equation represented in the second form, requirements imposed on the kernel $W$ are relaxed, because the factor given by the difference of the solution values at nearby points serves as a regulator, and the integral can converge even if the kernel diverges. Assuming that the asymptotic form of the kernel for large ranges ('Lévy flights') is described by a power-law function, as is typical for fractal structures,

$$
W(\mathbf{r}) \sim A r^{-3-\alpha}, \quad r \rightarrow \infty
$$

and introducing the notation $A^{\prime}=A \sigma, t=\epsilon^{\alpha} s, N(\mathbf{r}, t)=$ $\lim _{\epsilon \rightarrow 0} f_{\epsilon}\left(\mathbf{r}, t \epsilon^{-\alpha}\right)$, and $S(\mathbf{r}, t)=\lim _{\epsilon \rightarrow 0} \epsilon^{\alpha} S_{\epsilon}(\mathbf{r}, t)$, we obtain the equation for the three-dimensional isotropic Lévy motion:

$$
\frac{\partial N(\mathbf{r}, t)}{\partial t}=A^{\prime} \int \frac{N\left(\mathbf{r}^{\prime}, t\right)-N(\mathbf{r}, t)}{\left|\mathbf{r}^{\prime}-\mathbf{r}\right|^{3+\alpha}} \mathrm{d} \mathbf{r}^{\prime}+S(\mathbf{r}, t), \quad 0<\alpha<1 .
$$

The range of $\alpha$ indicated here is determined by the convergence condition for the integral at $\mathbf{r}^{\prime} \sim \mathbf{r}$. We assume that $N(\mathbf{r}, t)$ is a differentiable function of the coordinates, and therefore, as $\left|\mathbf{r}-\mathbf{r}^{\prime}\right| \rightarrow 0$, we have

$$
\left|N\left(\mathbf{r}^{\prime}, t\right)-N(\mathbf{r}, t)\right| \propto\left|\mathbf{r}^{\prime}-\mathbf{r}\right|,
$$

and the integral converges for $\alpha<1$. This integral operator can be continued to the domain of larger $\alpha$ by several methods [33]. The regularization by calculating the finite (Hadamard's) part of the integral brings it to the form

$$
J=\int \frac{\left[N\left(\mathbf{r}^{\prime}, t\right)-N(\mathbf{r}, t)\right]_{2}}{\left|\mathbf{r}^{\prime}-\mathbf{r}\right|^{3+\alpha}} \mathrm{d} \mathbf{r}^{\prime}, \quad 1<\alpha<2
$$

where $[\ldots]_{2}$ denotes a second-order difference.
In all these cases, the equation for the Fourier transform with respect to spatial variables,

$$
\frac{\partial \tilde{N}(\mathbf{k}, t)}{\partial t}=-A_{\alpha}|\mathbf{k}|^{\alpha} \tilde{N}(\mathbf{k}, t)+\tilde{S}(\mathbf{k}, t)
$$

contains a term proportional to the fractional power $\alpha$ of the wave vector $|\mathbf{k}|$. Because $-|\mathbf{k}|^{2}$ is the Laplace operator transform, this term can be represented as

$$
|\mathbf{k}|^{\alpha} \tilde{N}(\mathbf{k}, t) \equiv\left(|\mathbf{k}|^{2}\right)^{\alpha / 2} \tilde{N}(\mathbf{k}, t) \Leftrightarrow(-\Delta)^{\alpha / 2} N(\mathbf{r}, t),
$$

and the equation is then written in form (23).
From the physical standpoint, the fractional Laplacian can be regarded as the result of some averaging of the diffusion operator with a random diffusion coefficient $\tilde{D}$ :

$$
\langle\nabla(\tilde{D} \nabla N(\mathbf{r}, t))\rangle \mapsto-D_{\alpha}(-\Delta)^{\alpha / 2}\langle N(\mathbf{r}, t)\rangle .
$$

It seems that the general derivation of this relation from the fractal structure of a medium where diffusion occurs does not exist, but there is a good example of a particular process of the propagation of excitations in plasma by resonance radiation. When averaging the transfer equation with an exponential range distribution (asymptotically equivalent to the standard diffusion equation) over the Lorentzian frequency distribution, the integral transfer operator is indeed transformed into the fractional Laplacian [33], and the equation exactly coincides with the one written above.

The fractional-differential diffusion equation mainly differs from the standard diffusion equation in the range distribution: a power-like distribution in the first case and an exponential distribution in the second case. To elucidate the problem of cosmic-ray propagation in random magnetic fields, we refer again to the book by Ginzburg and Syrovatskii [4, p. 181]: "Assume that the motion occurs only along tubes of force, but these tubes themselves are randomly entangled, for example, consist of rectilinear segments with the mean length $l$, any angle between the directions of neighboring segments being equally probable. Then the
diffusion approximation can be fully applied to analyze the problem of averaging the spatial distribution of particles (cosmic rays) over large enough regions." What do we know about the lengths of these segments? Of course, it is natural to assume that they are random. As a continuous random quantity, this length is characterized by a distribution density $p(\xi), \xi>0$. The mention of the mean length of a rectilinear segment can be taken as an implicit assumption of the existence (i.e., convergence) of the improper integral

$$
\int_{0}^{\infty} p(\xi) \xi \mathrm{d} \xi=\langle R\rangle=l
$$

But this is not sufficient for providing the diffusive random walk of particles. The second moment of this distribution must also exist,

$$
\int_{0}^{\infty} p(\xi) \xi^{2} \mathrm{~d} \xi=\left\langle R^{2}\right\rangle
$$

The standard diffusion coefficient is expressed in terms of this moment. These questions do not occur in the standard kinetics because the initial range distribution is assumed exponential, and all the moments of an exponential distribution are finite. We can say, of course, that this distribution is 'derived', i.e., obtained as a solution of the first-order differential equation $\mathrm{d} p / \mathrm{d} \xi=-\sigma p(\xi)$, but a close inspection of this derivation shows that it is based on the assumption that the random numbers of atoms on the segment $[0, \xi]$ and the adjacent element d $\xi$ are mutually independent. In classical kinetics, a particle moves in an ideal gas of noninteracting and unrelated atoms. In the case under study, the role of such atoms, in 'collisions' with which the cosmic-ray particles change the direction of their motion, is played by the ends (or, it is better to say, breaking points) of the rectilinear parts of magnetic field lines. Can we agree without any doubt that two such points lying at the ends of the same segment are mutually independent? Apparently not. But rejecting this hypothesis of the independence, we thereby cast doubt on the validity of the exponential distribution. This is not yet problematic because many other distributions exist with two finite first moments; however, the beginning has already been made. Any such distribution in the long-time limit brings us again to classical diffusion, but now the question already arises: What is there behind the convergence of the second moment? What do we sacrifice when we accept the assumption about the convergence of the second moment? And it turns out that we sacrifice the entire class of stable laws with a power-law asymptotic behavior, in problems with plasmas, turbulence, and random kinetics, which we are dealing with in studying cosmic rays. The difference between transfer processes with exponential and power-law range distributions is the same as that between molecular and turbulent diffusion, and cosmic-ray transfer in the Galaxy is the turbulent diffusion.

## 3.3 'Knee' in the spectrum and the model parameterization

 Of all the measurements performed in cosmic ray physics, only data on the energy spectra cover more than ten orders of magnitude, whereas the range of variations of other parameters is considerably smaller. This suggests that the effect of replacing the normal spatial distribution by distributions with power-law asymptotic forms is manifested, first of all, in energy spectra. However, it is obvious that a simple introduction of the energy spectrum $S(E)$ of a source into the transferequation,

$$
\frac{\partial N}{\partial t}=-D_{\alpha}(-\Delta)^{\alpha / 2} N(\mathbf{r}, t, E)+\delta(\mathbf{r}) \delta(t) S(E)
$$

is not sufficient for estimating this effect: the source spectrum translates into the observed spectrum

$$
N(\mathbf{r}, t, E)=\left[D_{\alpha} t\right]^{-3 / \alpha} \Psi_{3}^{(\alpha)}\left(\left[D_{\alpha} t\right]^{-1 / \alpha} r\right) S(E)
$$

without changes, because particles with different energies move in space with the same 'diffusion coefficient'. The situation changes if we introduce the energy dependence of this coefficient: in this case, different energies correspond to different values of the dimensional variable $\xi=$ $\left[D_{\alpha}(E) t\right]^{-1 / \alpha} r$. The choice of a power-law dependence, as in (1), means that the parameter $\xi=\left[D_{0 \alpha} t E^{\delta}\right]^{-1 / \alpha} r \equiv \xi_{1} E^{-\delta / \alpha}$ of a high-energy particle lies near the top of the stable density (the left asymptotic region, for which we conventionally assume that $\xi<\xi_{0}$ ), whereas low-energy particles correspond to large values of $\xi$ lying at the periphery of the spatial distribution $\Psi_{3}^{(\alpha)}(\xi)$ (the right asymptotic region, $\xi>\xi_{0}$ ). In the first case,

$$
N(\mathbf{r}, t, E) \sim S_{0} r^{-3} E^{-p}\left[\xi^{3} \Psi_{3}^{(\alpha)}(0)\right] \propto E^{-p-3 \delta / \alpha}, \quad E \rightarrow \infty
$$

for any admissible $\alpha \in(0,2]$. In the second case $(E \rightarrow 0)$, a power-law asymptotic behavior is observed only for anomalous diffusion $(\alpha<2)$ :

$$
N(\mathbf{r}, t, E) \propto \begin{cases}E^{-(p+3 \delta / 2)} \exp \left(-\frac{\xi_{1}^{2}}{4 E^{\delta}}\right), & \alpha=2 \\ E^{-(p-\delta)}, & \alpha<2\end{cases}
$$

This was the reason in $[19,20]$ to pass from the usual diffusion equation to the equation with fractional Laplacian (23), which allows relating this Laplacian to the fractal properties of the medium due to which the free paths of particles acquire the power-law form. The difference between the exponents of power-law asymptotic forms for low and high energies manifested for $\alpha<2$ was interpreted as an indication of a 'knee'.

Using experimental data on the position of the intermediate knee region $\left(E_{<}, E_{>}\right)$and the exponent of the spectrum for $E<E_{<}$and $E>E_{>}$, the main parameters $\left(D_{0 \alpha}, \delta\right)$ of the model and the exponent $p$ of particle generation in a source as a function of the exponent $\alpha$ were found. In [19], calculations were originally performed for a source located at a distance $r$ from the observation point and acting with a constant intensity for a time interval $\tau_{\mathrm{S}}$ preceding the observation. In this case,

$$
\begin{aligned}
N(\mathbf{r}, t, E) & =S_{0} E^{-p}\left[D_{0 \alpha} E^{\delta}\right]^{-3 / \alpha} \\
& \times \int_{\max \left\{0, t-\tau_{\mathrm{s}}\right\}}^{t} \Psi_{3}^{(\alpha)}\left(r\left[D_{0 \alpha} E^{\delta} \tau\right]^{-1 / \alpha}\right) \tau^{-3 / \alpha} \mathrm{d} \tau
\end{aligned}
$$

and the exponent of the observed spectrum changes from $p-\delta$ for $E \leqslant E_{<}$to $p+\delta$ for $E \geqslant E_{>}$. The results of calculations showed that the best agreement with experimental data on the spectra of protons and nuclei and the total spectrum of all particles was achieved for $\alpha \approx 5 / 3$, $E_{\mathrm{rad}}=\left(E_{>}+E_{<}\right) / 2=3 \times 10^{4} \mathrm{GeV}$ per nucleon, $\delta=0.25$, and the injection exponent for all nuclei in the source $p \approx 2.9$ ( $r \sim 200 \mathrm{pc}, \tau_{0} \sim 10^{5}$ years). The exponent of the spectrum observed in the kink region then changes from 2.65 to 3.15 (Fig. 8), which does not contradict the hypothesis that the


Figure 8. Results of the first calculations in a model with the fractional Laplacian $(\alpha=5 / 3)$ [19, 20]: comparison of the proton spectrum (curve) calculated in this model with different experimental data (see the details in [56]).
sources of cosmic rays could be the explosions of the nearest supernovae during the last 100,000 years.

In [57] (see also [58]), we passed to the equation of a more general type, containing, along with the fractional Laplacian, the fractional time derivative to take the influence of magnetic traps into account (which, without a doubt, exist in the galactic medium):

$$
\begin{equation*}
\frac{\partial N}{\partial t}=-D_{\alpha}(E) \frac{\partial^{1-\beta}}{\partial t^{1-\beta}}(-\Delta)^{\alpha / 2} N(\mathbf{r}, t, E)+S(\mathbf{r}, t, E) \tag{24}
\end{equation*}
$$

Because the family of these equations includes Eqn (23) as a particular case $(\beta=1)$, we sacrificed nothing, but simply extended the class of solutions: instead of the one-parameter family of solutions expressed in terms of stable densities $\Psi_{3}^{(\alpha)}$, we obtained the two-parameter family of solutions determined by the fractionally stable densities $\Psi_{3}^{(\alpha, \beta)}$ [40]. For a point-like instantaneous source, the solution has the form

$$
N(\mathbf{r}, t, E)=S_{0} E^{-p}\left[D_{0 \alpha} E^{\delta} t^{\beta}\right]^{-3 / \alpha} \Psi_{3}^{(\alpha, \beta)}\left(r\left[D_{0 \alpha} E^{\delta} t^{\beta}\right]^{-1 / \alpha}\right),
$$

whence it follows (see Section 2.6) that

$$
N(\mathbf{r}, t, E) \approx \begin{cases}S_{0} D_{0 \alpha} t^{\beta} r^{-3-\alpha} E^{-p+\delta}, & E<E_{<} \\ S_{0}\left[D_{0 \alpha} t^{\beta}\right]^{-1} r^{-3+\alpha} E^{-p-\delta}, & E>E_{>}\end{cases}
$$

Thus, as the energy increases, the spectrum steepness increases, which is manifested in the increase in the absolute value of the exponent of the spectrum by $2 \delta$ after passing through the interval $\left(E_{<}, E_{>}\right)$. In the case of a source acting with a constant intensity for a finite time, the spectrum steepness also increases, but the exponent changes by $(1+1 / \beta) \delta$. We note that in both cases, the exponents characterizing the exponential behavior of the spectrum outside the knee region are independent of $\alpha$, but a noticeable effect is observed in the intermediate region: the smaller $\alpha$ is, the broader the transition region, and the transition occurs more smoothly (as shown in Fig. 5).

The choice of numerical values of these parameters was discussed in [59]: "To estimate the parameter $\beta$, we used results from paper [60], where the anomalous diffusion of solar magnetic elements was studied. The authors showed that the distribution of the lifetime in a trap in asymptotics takes the form of the Lévy distribution with the spectral
exponent $\beta \approx 0.8$. Assuming that the capture mechanism is characterized by a certain self-similarity, we can expect the same value of $\beta$ at all scales under study. For this reason, we used the value $\beta=0.8$ in our calculations. Assuming then that $\eta_{E<E_{0}} \sim 2.63$ and $\eta_{E \gtrdot E_{0}} \sim 3.24$, we finally obtain $p \approx 2.9$ and $\delta \sim 0.27$.

To determine the next important parameter, the anomalous diffusion coefficient $D_{0 \alpha}$, we used experimental data on the anisotropy of $10^{3}-10^{4}-\mathrm{GeV}$ particle fluxes within the framework of the scheme proposed by Osborne and coauthors in 1976 [61] and Dorman and coauthors in 1985 [62]. In particular, we found that $D_{0 \alpha} \approx(1-4) \times 10^{-3} \mathrm{pc}^{1.7}$ year $^{-0.8}$ for $\alpha=1.7$ and $\beta=0.8$ for the three nearest sources.

In the model considered here, only one parameter $\alpha$ $(1<\alpha<2)$, related to the fractal structure of the medium, was found by fitting. Test calculations of cosmic-ray spectra showed that the best fit of experimental data was achieved for $\alpha \approx 1.7$."

In conclusion, the authors of [59] note that for the parameters presented above, the results of calculations are in agreement not only with the experimental increase in the spectrum steepness but also with the mass content in the energy region $10^{2}-10^{5} \mathrm{GeV}$ per nucleon if the source content is $p=72 \%, \mathrm{He}=18 \%, \mathrm{CNO}=5 \%, \mathrm{Ne}-\mathrm{Si}=3 \%$, and $\mathrm{Fe}=2 \%$.

For $\alpha=1$, the stable density can be written in a simple analytic form (the three-dimensional Cauchy distribution; see Section 2.6), which was used in [59] in solving Eqn (23) with the source

$$
S(\mathbf{r}, t, E)=S_{0} E^{-p} \delta(\mathbf{r}) 1_{+}(t)
$$

As a result, the simple expression
$N(\mathbf{r}, t, E)=\frac{S_{0} E^{-p-\delta}}{2 \pi D_{0.1} r^{2} t} \frac{\left(E / E_{\mathrm{rad}}\right)^{2 \delta}}{\left(E / E_{\mathrm{rad}}\right)^{2 \delta}+1} 1_{+}(t), \quad \alpha=1, \quad \beta=1$
was obtained with $E_{\mathrm{rad}}=\left[r /\left(D_{0.1} t\right)\right]^{1 / \delta}$, which allows passing from one set of asymptotic expressions to another. In particular, we can estimate the energy gap $\Delta E=E_{>}-E_{<}$ separating these two asymptotic forms. For a $20 \%$ agreement margin (the difference between the exact spectrum and its power-law asymptotic form), we obtain the energy gap somewhat smaller than two and a half orders of magnitude, which does not contradict the experimental data in general.

In [63], a problem with a constant (in time) point-like source was considered. The solution of the stationary equation

$$
D_{\alpha}(E)(-\Delta)^{\alpha / 2} N(\mathbf{r}, E)=S_{0} E^{-p} \delta(\mathbf{r}),
$$

following from (23) and found by using the known Mellin transformation of the three-dimensional stable density

$$
\int_{0}^{\infty} \Psi_{3}^{(\alpha)}(r) r^{s-1} \mathrm{~d} r=\frac{2^{s} \Gamma(s / 2) \Gamma((3-s) / \alpha)}{\alpha(4 \pi)^{3 / 2} \Gamma((3-s) / 2)}
$$

under the same assumption about the energy dependence of the diffusion coefficient, has the form

$$
\begin{aligned}
N(\mathbf{r}, E) & =S_{0} E^{-p}\left[D_{\alpha}(E)\right]^{-3 / \alpha} \int_{0}^{\infty} \Psi_{3}^{(\alpha)}\left(r\left[D_{\alpha}(E) t\right]^{-1 / \alpha}\right) \mathrm{d} t \\
& =\frac{2^{-\alpha} S_{0}}{\pi^{3 / 2} D_{0 \alpha} r^{3-\alpha}} \frac{\Gamma((3-\alpha / 2) / 2)}{\Gamma(\alpha / 2)} E^{-p-\delta} .
\end{aligned}
$$

We emphasize that this expression is an exact solution of the stationary equation, although this diffusion equation itself is approximate, representing the asymptotic regime of the process at large distances, and therefore the use of its solution at small distances is risky. In conclusion, the authors of [63] made an attempt to obtain the two-parameter (with $\beta \neq 1$ ) stationary solution as the limit (for $t \rightarrow \infty$ ) of the distribution of particles from a source switched on at the instant $t=0$ :

$$
\begin{aligned}
& N_{\mathrm{st}}(\mathbf{r}, E)=\lim _{t \rightarrow \infty} N(\mathbf{r}, t, E) \\
& \quad=\frac{S_{0} E^{-p}}{\left[D_{\alpha}(E)\right]^{3 / 2}} \int_{0}^{\infty} \Psi_{3}^{(\alpha, \beta)}\left(r\left[D_{\alpha}(E) \tau^{\beta}\right]^{-1 / \alpha}\right) \tau^{-3 \beta / \alpha} \mathrm{d} \tau
\end{aligned}
$$

However, this integral diverges, and they were forced to cut off the time integral at a large upper limit ( $10^{10}$ years), which would make sense if the limit existed. The divergence is explained by the fact that a fraction of the particles confined in traps escape from them after some time and continuously make up the total flux; however, due to the infinite mean lifetime in a trap, the equilibrium between particles captured in traps and those leaving them is not established. Mathematically, the matter is that the Riemann-Liouville fractional derivative used in this model vanishes when applied to a constant only for an integer order $\beta$, whereas a fractionalorder derivative of a constant is not zero [33]:

$$
\frac{\partial^{v} C}{\partial t^{v}}=\frac{C t^{-v}}{\Gamma(1-v)}
$$

Therefore, for a stationary (time-independent) distribution of particles in a medium to exist, i.e., for the condition $\partial N_{\mathrm{st}} / \partial t=0$ to be satisfied, as follows from Eqn (24), the source density must satisfy the equation

$$
\begin{aligned}
S(\mathbf{r}, t, E) & =D_{\alpha}(E)(-\Delta)^{\alpha / 2} \frac{\partial^{1-\beta}}{\partial t^{1-\beta}} N_{\mathrm{st}}(\mathbf{r}, E) \\
& \left.=\left[D_{\alpha}(E)(-\Delta)^{\alpha / 2} N_{\mathrm{st}} \mathbf{r}, E\right)\right] \frac{t^{\beta-1}}{\Gamma(\beta)} .
\end{aligned}
$$

Such a behavior of sources can be explained by the fact that they themselves are traps of the same type as others, emitting particles at a rate decreasing in accordance with a power law, which seems more natural than a source perpetually emitting particles at a constant rate.

Using the results in [59], researchers in Lagutin's group calculated the energy spectra and mass composition by separating the particle flux into contributions from particles in a direct (unscattered) flux from near ( $r<1 \mathrm{kpc}$ ) and distant $(r>1 \mathrm{kpc})$ sources, represented by three terms in the expression

$$
J_{i}=\frac{v_{i}}{4 \pi}\left[C_{0 i} E^{-p+\delta / \beta}+\sum_{j: r_{j}<1 \mathrm{kpc}} N_{i}\left(r_{j}, t_{j}, E\right)+C_{1 i} E^{-p-\delta / \beta}\right],
$$

$i=\mathrm{p}, \mathrm{He}, \mathrm{CNO}, \mathrm{Ne}-\mathrm{Si}, \mathrm{Fe}$.
The constants $C_{0 i}$ and $C_{1 i}$ were determined from the same experimental data with which the results of calculations were compared. By introducing the correction for the solar modulation of galactic rays, the authors of [63, 64] satisfactorily reproduced the energy spectra (the total spectrum is shown in Fig. 9) and the mass composition of components with different energies.


Figure 9. Spectrum of all particles. The curve is calculated in the fractional differential model with the parameters $\alpha=1.7$ and $\beta=0.8$. Symbols are different experimental results obtained in 1970-2001 (see the details in [64]).

The same model was used in [65, 66] for calculating the electron and positron spectra. The sources of high-energy ( $E \geqslant 100 \mathrm{GeV}$ ) electrons and positrons observed in the Solar System were shown to be comparatively young local sources (the distance is no more than 200 pc and the age is $10^{5}$ years).

The fraction of positrons obtained from these calculations was in agreement with experimental data (Fig. 10) and the exponent $p_{\mathrm{e}}=2.95$ of the spectrum of the source for electrons and positrons proved to be close to the exponent $p_{\mathrm{p}}=2.9$ found previously for protons. The authors believe that this suggests that the acceleration mechanism of these particles is the same. The energy losses of relativistic electrons were taken in the form

$$
-\frac{\mathrm{d} E}{\mathrm{~d} t}=b(E)=b_{0}+b_{1} E+b_{2} E^{2}
$$

(corresponding to a homogeneous medium, however), whereas the inhomogeneities of the magnetic field were assumed fractal. Such an approach assumes the absence of a correlation between the magnetic field and matter in the interstellar region, which contradicts the conventional concept of magnetic field lines 'frozen' into the material medium. In the case of complete correlation, the term with energy losses should enter the equation as a part of the material derivative operator raised to a fractional power (see Section 3.4).

The calculations by Lagutin et al. with the parameters $\alpha=1.7$ and $\beta=0.8$ (which we refer to as the LagutinUchaikin (LU) model for convenience) were continued in [67, 68], and it seems that no serious disagreements with experiments were observed (Figs 10 and 11).

However, the authors of [69] performed calculations with $\alpha=0.5,1.0$, and 1.5 and concluded that the value $\alpha=1.0$ provides the best fit of the experimental data and at the same time is consistent with the Kraichnan spectrum $F(k) \propto k^{-\chi}$ known from turbulence theory (the parameter $\alpha$ is related to the index $\chi$ as $\alpha=(3-\chi) / 2$; we have $\chi=3 / 2$ and $\alpha=3 / 4$ for the magnetic energy). ${ }^{6}$ Moving further in this direction,

[^4]

Figure 10. Fraction of positrons in the electron-positron component of primary cosmic rays. Symbols are different experimental results obtained in 1987-2000 (see the details in [66]); the curve is calculated in the LU $\operatorname{model}(\alpha=1.7, \beta=0.8)$.


Figure 11. Mean depth of the maximum of broad atmospheric showers as a function of the primary energy. Symbols show different experimental data, curves (dotted for primary protons, dashed for primary iron nuclei, and solid for the mixed composition) are calculated with the parameters $\alpha=1.7$ and $\beta=0.8$ [67].

Lagutin and Tyumentsev [73] took a new value of the key parameter $\alpha=0.3$ (we call this variant the $L T$ version of the fractional differential model of galactic cosmic ray transfer) and performed extensive calculations with this parameter in 2004-2010 (which are available in proceedings of cosmic ray conferences). The motivation for choosing this value of $\alpha$ was the finding in [74] that the fractional exponent $\alpha$ in the range distribution does not coincide with the fractal dimension $d_{\mathrm{F}}$ of the medium. In addition, it was desirable to match mean free paths with the known parameters of the real interstellar medium. The first calculations of this type were performed in our paper [75] (see also [76]). Lagutin et al. simulated the free paths of particles in a medium with randomly distributed spherical targets and wrote in [77] that because "for media with a fractal dimension $1<d_{\mathrm{F}}<2$,

$$
\begin{equation*}
\alpha \approx 2-d_{\mathrm{F}} \tag{25}
\end{equation*}
$$

we find $\alpha=0.3$ for the galactic medium with $d_{\mathrm{F}}=1.7$ [51]. We set the exponent $\beta$ equal to 0.8 , as in [58]. Another important parameter of the model, the anomalous diffusion coefficient $D_{0 \alpha}$, can be estimated by comparing the position of the knee in


Figure 12. Difference in the range distribution for the same distribution of centers and different radii of spheres simulating magnetic clouds: (a) LT model, (b) LU model.
the observed cosmic-ray spectrum with the position of the 'breaking point' in fractionally stable distributions $\Psi_{3}^{(\alpha, \beta)}(\xi)$. Because the breaking point of $\Psi_{3}^{(0.3 ; 0.8)}(\xi)$ is observed at $\xi \approx 2.3$, we obtain

$$
r\left(D_{0 \alpha} E_{\text {knet }}^{\delta} t^{\beta}\right)^{-1 / \alpha} \approx 2.3 .
$$

Assuming that near sources are also involved in the formation of the knee in the energy spectrum, we find for $r \approx 10^{2} \mathrm{pc}$ and $t \approx 10^{5}$ years that

$$
D_{0 \alpha} \approx(3-5) \times 10^{-6} \mathrm{pc}^{0.3} / \mathrm{y}^{0.8} .
$$

For such $D_{0 \alpha}$ and the parameters $\delta, \alpha$, and $\beta$ chosen above, a unique relation exists between $r$ and $t$ for the sources providing the knee in the spectrum at $E_{\mathrm{fr}} \approx 3 \times 10^{6} \mathrm{GeV}$."

This conclusion is incorrect, because expression (25) is valid only for extremely small elements forming a fractal. The size of magnetic inhomogeneities of the galactic medium does not belong to this type: they occupy $5 \%$ of the volume (according to old estimates made in [2], which, however, were confirmed by subsequent astronomical measurements: for example, the volume fraction of the interstellar space not filled with hydrogen was estimated as $95 \%$ in [57]), which becomes more than $30 \%$ on passing to the linear scale. That formula (25) is invalid for such scales can be clearly seen, in particular, from [55], where, based on the experimental studies of the cloudy structure of interstellar hydrogen, the value $d_{\mathrm{F}}=2.3$ was obtained, which is obviously incompatible with expression (25). We also recall that the fractal dimension alone does not characterize a fractal-like structure completely: the size of its elements and their concentration in space are determined not by the exponent $d_{\mathrm{F}}$ in the fractal formula $V_{\mathrm{F}}(R)=C R^{d_{\mathrm{F}}}$ but by the coefficient $C=V_{\mathrm{F}}(1)$ determining the volume fraction of a unit-radius ball filled with a fractal. An increase in the size of magnetic clouds, preserving the same fractal dimension $d_{\mathrm{F}}$, obviously reduces the range distribution, thereby increasing the exponent $\alpha$ (Fig. 12). This effect is also clearly seen in Fig. 13 taken from [77].

The original LU model has been used by other authors. The authors of [78] used the stationary solution for $\alpha=1.8$ in Monte Carlo simulations of cosmic-ray diffusion from a supernova (in the Erlykin-Wolfendale model) and obtained good agreement with the observed characteristics. Based on these results, they concluded that the source of cosmic rays is


Figure 13. Exponent $\alpha$ of the range distribution as a function of the fractal dimension of the distribution of the centers of spheres and their radii $R=10.0,2.0,1.0,0.5,0.25$, and 0.01 rel. units for respective curves $1-6$ [77].
a supernova. The authors of [79] in fact repeated our first calculations and concluded again that the value $\alpha=1.65$ provides the best fit of the energy spectrum and radial gradient in the vicinity of the Solar System, and the admissible range of $\alpha$ values is $1.6-1.9$, whereas the value $\alpha=2$, corresponding to classical diffusion, is unacceptable. Paper [80] is entirely devoted, in fact, to estimating the exponent $\alpha$. The authors studied the transfer of cosmic rays with the energy in the range $10^{12}-10^{19} \mathrm{eV}$ from a supernova with the energy fraction converted to cosmic rays from 0.01 to 0.1 and the supernova age from $10^{4}$ to $10^{7}$ years for different $\alpha \in[0.5,2.0]$ and concluded that the propagation of cosmic rays in the Galaxy is governed by anomalous diffusion of the second kind (with $\alpha=1.7$ ) and is not described by the normal diffusion model $(\alpha=2)$. Therefore, this conclusion also rejects the LT hypothesis about anomalous diffusion with $\alpha=0.3$ (diffusion of the first kind).

After our critical paper [81] in 2010, the authors of the LT version decided to return to the original model. In 2011, they published calculations with $\alpha=1.1$ [82] and presented the results of Monte Carlo simulations with $\alpha=1.7$ at the European Cosmic Ray Symposium in 2012 [83].

The coefficient $D_{0 \alpha}$ suffers the greatest changes under these variations of $\alpha$ (Table 1). However, the final results (the spectrum and composition) were practically unchanged (Table 2). Obviously, a change in the parameter $\alpha$ was compensated by the corresponding change in the diffusion coefficient. However, as $\alpha$ changes, the dimension of the diffusion coefficient also changes, and we cannot quantitatively estimate these changes or estimate the difference between velocity and acceleration values. It is possible that the coefficients $C_{i}$ and $C_{0 i}$ play an important role here, and the results can be changed by varying these coefficients. At the

Table 1. Parameters of the fractional differential model in different papers.

| $\alpha$ | $\beta$ | $p$ | $\delta$ | $D_{0}, \mathrm{pc}^{\alpha} \mathrm{year}^{-\beta}$ | References (year) |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1.7 | 0.8 | 2.90 | 0.27 | $(1-4) \times 10^{-3}$ | $[64](2001)$ |
| 0.3 | 0.8 | 2.86 | 0.27 | $(3-5) \times 10^{-6}$ | $[73](2004)$ |
| 0.7 | 1.0 | 2.60 | 0.27 | $2 \times 10^{-5}$ | $[84](2008)$ |
| 1.1 | 0.8 | 2.85 | 0.27 | $1 \times 10^{-4}$ | $[82](2011)$ |
| 1.7 | 0.8 | 2.90 | 0.27 | $2.4 \times 10^{-3}$ | $[83](2012)$ |

Table 2. Mass composition in the LU model and its LT version.

| Model | $\alpha$ | $\beta$ | p, <br> $\%$ | He, <br> $\%$ | CNO, <br> $\%$ | $\mathrm{Ne}-\mathrm{Si}$, <br> $\%$ | Fe, <br> $\%$ | References <br> (year) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| LU | 1.7 | 0.8 | 72 | 18 | 5 | 3 | 2 | $[64](2001)$ |
| LT | 0.3 | 0.8 | 77 | 16 | 4 | 2 | 1 | $[83](2012)$ |
| $[73](2004)$ |  |  |  |  |  |  |  |  |

same time, as shown in Sections 3.7 and 3.8, these manipulations drastically change the space-time shapes of particle trajectories, which was neglected by the authors of the LT version. If the dynamics of the process described by a fractional differential equation are not consistent with the reality, simulations become a multiparametric approximation of the known experimental results.

We also note that the parameter $\gamma=\beta / \alpha$ characterizing the expansion law of a diffusion packet in the LU model turns out to be suspiciously close to the classical value $1 / 2$. Hence, if we initially attempt to construct a model in which the expansion of a diffusion packet is consistent with the standard theory, but the shape is self-similar and has a power-law asymptotic behavior required for the description of the observed knee, then, for $\beta=0.8$, we arrive precisely at the LU model $(\beta / \alpha \approx 0.5)$ rather than at the LT version $(\beta / \alpha \approx 2.66)$. Obviously, the diffusion packet dynamics in the LT version in the large-time asymptotic regime contradicts he physical reality, because the packet expansion velocity increases infinitely, despite the natural restriction imposed on the velocities of particles in the packet.

### 3.4 Equations of restricted-speed anomalous diffusion

In the process considered above, flights were assumed to be instantaneous [we called this process unrestricted anomalous diffusion (UAD)]. The solutions of random-walk problems for cosmic rays with a finite velocity of free motion were considered in our papers [81, 85-88] [restricted anomalous diffusion (RAD)].

In the case of a finite flight velocity, a particle can be in one of two states at an arbitrary observation instant: rest (0) or motion (1) (Fig. 14b-d). We let the $1 \rightarrow 0$ and $0 \rightarrow 1$ transition rates per unit volume be denoted by $F_{1 \rightarrow 0}(\mathbf{r}, t)$ and $F_{0 \rightarrow 1}(\mathbf{r}, t)$, in the vicinity of a point $\mathbf{r}$ for a particle located at the coordinate origin at the initial instant. The unit velocity vector $\boldsymbol{\Omega}$ of a particle leaving a source or a trap has an isotropic distribution. A particle that has undergone a transition to the rest state at the instant $t-t^{\prime}$ remains there until the observation instant $t$ with the probability

$$
Q\left(t^{\prime}\right)=\int_{0}^{\infty} q\left(t^{\prime}+\tau\right) \mathrm{d} \tau
$$

and a particle that has left a trap at the point $\mathbf{r}-\mathbf{r}^{\prime}$ intersects, without interaction, a unit area at the point $\mathbf{r}$ with the probability

$$
P\left(\mathbf{r}^{\prime}\right)=\int_{0}^{\infty} p\left(\mathbf{r}^{\prime}+\xi \boldsymbol{\Omega}\right) \mathrm{d} \xi
$$

Because the particle should spend $r^{\prime} / v$ seconds for this transition, we obtain the total probability density of finding the particle at the point $\mathbf{r}$ at the instant $t$ :

$$
\begin{align*}
G(\mathbf{r}, t) & =\int_{0}^{\infty} \mathrm{d} t^{\prime} Q\left(t^{\prime}\right) F_{1 \rightarrow 0}\left(\mathbf{r}, t-t^{\prime}\right) \\
& +\frac{1}{v} \int \mathrm{~d} \mathbf{r}^{\prime} P\left(\mathbf{r}^{\prime}\right) F_{0 \rightarrow 1}\left(\mathbf{r}-\mathbf{r}^{\prime}, t-\frac{r^{\prime}}{v}\right) \tag{26a}
\end{align*}
$$

The transition velocities (under the condition that the particle history begins with the particle capture in a trap located at the coordinate origin) are related as
$F_{1 \rightarrow 0}(\mathbf{r}, t)=\int \mathrm{d} \mathbf{r}^{\prime} p\left(\mathbf{r}^{\prime}\right) F_{0 \rightarrow 1}\left(\mathbf{r}-\mathbf{r}^{\prime}, t-\frac{r^{\prime}}{v}\right)+\delta(\mathbf{r}) \delta(t)$,
$F_{0 \rightarrow 1}(\mathbf{r}, t)=\int_{0}^{t} \mathrm{~d} \tau q(\tau) F_{1 \rightarrow 0}(\mathbf{r}, t-\tau)$.
Applying the Fourier transformation in the spatial coordinate and the Laplace transformation in the time coordinate to system of integral equations (26), we obtain the system of algebraic equations

$$
\begin{aligned}
& \tilde{G}(\mathbf{k}, \lambda)=\hat{Q}(\lambda) \tilde{F}_{1 \rightarrow 0}(\mathbf{k}, \lambda)+\frac{1}{v} \tilde{P}\left(\mathbf{k}, \frac{\lambda}{v}\right) \tilde{F}_{0 \rightarrow 1}(\mathbf{k}, \lambda) \\
& \tilde{F}_{1 \rightarrow 0}(\mathbf{k}, \lambda)=\tilde{p}\left(\mathbf{k}, \frac{\lambda}{v}\right) \tilde{F}_{0 \rightarrow 1}(\mathbf{k}, \lambda)+1 \\
& \tilde{F}_{0 \rightarrow 1}(\mathbf{k}, \lambda)=\hat{q}(\lambda) \tilde{F}_{1 \rightarrow 0}(\mathbf{k}, \lambda)
\end{aligned}
$$

where

$$
\begin{aligned}
& \hat{Q}(\lambda)=\frac{1-\hat{q}(\lambda)}{\lambda} \\
& \tilde{P}\left(\mathbf{k}, \frac{\lambda}{v}\right)=\int P(\mathbf{r}) \exp \left(-\frac{\lambda}{v} r\right) \exp (\mathbf{i k r}) \mathrm{d} \mathbf{r} \\
& \tilde{p}\left(\mathbf{k}, \frac{\lambda}{v}\right)=\int p(\mathbf{r}) \exp \left(-\frac{\lambda}{v} r\right) \exp (\mathbf{i k r}) \mathrm{d} \mathbf{r}
\end{aligned}
$$

The solution of this system has the form

$$
\begin{equation*}
\tilde{G}(\mathbf{k}, \lambda)=\frac{\hat{Q}(\lambda)+(1 / v) \tilde{P}(\mathbf{k}, \lambda / v) \hat{q}(\lambda)}{1-\tilde{p}(\mathbf{k}, \lambda / v) \hat{q}(\lambda)} . \tag{27}
\end{equation*}
$$



Figure 14. Initial segment of a trajectory in the ( $x, t$ ) coordinates with (a) instant flights, ( $\mathrm{b}, \mathrm{c}$ ) a finite velocity of flights in a medium with traps (two possible states at the observation instant $t$ ), and (d) continuous motion (traps are absent).

For $v=\infty$, expression (27) takes the form of Montroll-Weiss formula (14) describing the motion of a particle as a sequence of instant hops from one point in space to another: irrespective of the distance between these two points, the particle arrives at one of them at the same instant that it leaves another point. In this case, the delay time (the lifetime in a trap) is connected neither with this distance nor with the motion as a whole. If traps are removed from this model (which amounts to setting $\hat{Q}(\lambda)=0$ ), then the model has no meaning at all: the particle instantly flies away to infinity, disappearing from the system. However, the introduction of a finite velocity of free motion brings everything back to normal even in the absence of traps: the particle moves continuously with time, always remaining within a sphere with the radius $v t$ centered at the initial point of the trajectory. In this case, solution (27) becomes

$$
\begin{equation*}
\tilde{G}(\mathbf{k}, \lambda)=\frac{(1 / v) \tilde{P}(\mathbf{k}, \lambda / v)}{1-\tilde{p}(\mathbf{k}, \lambda / v)} . \tag{28}
\end{equation*}
$$

Delays in the motion of particles are now not caused by independent traps and become closely related (directly proportional) to particle ranges, and this considerably changes the situation, especially for $\alpha<1$ (in the original paper [81], the RAD equations were derived under the assumption of independent traps).

We now pass from Eqn (28) in the Fourier-Laplace variables to the asymptotic equation in natural variables. For this, we represent (28) in the form

$$
\begin{equation*}
\tilde{L}(\mathbf{k}, \lambda) \tilde{N}(\mathbf{k}, \lambda) \equiv\left[1-\tilde{p}\left(\mathbf{k}, \frac{\lambda}{v}\right)\right] \tilde{N}(\mathbf{k}, \lambda)=v^{-1} \tilde{S}(\mathbf{k}, \lambda) \tag{29}
\end{equation*}
$$

where the function $\tilde{S}(\mathbf{k}, \lambda)$ is not necessarily equal to the function $P(\mathbf{k}, \lambda / v)$ characterizing an instantaneous point-like source, but is related to an arbitrary source (therefore, the solution $N$ is not necessarily the Green's function). We assume, as previously, that

$$
P(R>r) \sim A r^{-\alpha}, \quad r \rightarrow \infty,
$$

whence

$$
p(\mathbf{r}) \mathrm{d} \mathbf{r}=p_{R}(r) \mathrm{d} r \frac{\mathrm{~d} \boldsymbol{\Omega}}{4 \pi},
$$

where $\mathrm{d} \boldsymbol{\Omega}$ is a solid angle element and $p_{R}(r)$ is the range distribution density,

$$
\begin{equation*}
p_{R}(r)=-\frac{\mathrm{d} P(R>r)}{\mathrm{d} r} \sim \alpha A r^{-\alpha-1} \mathrm{~d} r, \quad r \rightarrow \infty . \tag{30}
\end{equation*}
$$

The Laplace transform of this distribution behaves differently in the vicinity of $s=0$, depending on the convergence or divergence of the mean and dispersion:

$$
\begin{align*}
\hat{p}_{R}(s) & =\int_{0}^{\infty} p_{R}(r) \exp (-s r) \mathrm{d} r \\
& \sim \begin{cases}1-A_{\alpha} s^{\alpha}, & 0<\alpha<1, \\
1-\langle R\rangle s+A_{\alpha} s^{\alpha}, & 1<\alpha<2, \\
1-\langle R\rangle s+\left\langle\frac{R^{2}}{2}\right\rangle s^{2}, & \alpha>2,\end{cases} \tag{31}
\end{align*}
$$

and we therefore obtain

$$
\begin{aligned}
& 1-\tilde{p}\left(\mathbf{k}, \frac{\lambda}{v}\right)=\int\left\{1-\exp \left[-\left(\frac{\lambda}{v}-\mathrm{i} \mathbf{k} \boldsymbol{\Omega}\right) r\right]\right\} p(\mathbf{r}) \mathrm{d} \mathbf{r} \\
& \sim \begin{cases}A_{\alpha}\left\langle\left(\frac{\lambda}{v}-\mathrm{i} \mathbf{k} \boldsymbol{\Omega}\right)^{\alpha}\right\rangle, & 0<\alpha<1, \\
\langle R\rangle\left\langle\left(\frac{\lambda}{v}-\mathbf{i} \mathbf{k} \boldsymbol{\Omega}\right)\right\rangle-A_{\alpha}\left\langle\left(\frac{\lambda}{v}-\mathrm{i} \mathbf{k} \boldsymbol{\Omega}\right)^{\alpha}\right\rangle, & 1<\alpha<2, \\
\langle R\rangle\left\langle\left(\frac{\lambda}{v}-\mathbf{i} \mathbf{k} \boldsymbol{\Omega}\right)\right\rangle-\left\langle\frac{R^{2}}{2}\right\rangle\left\langle\left(\frac{\lambda}{v}-\mathrm{i} \mathbf{k} \boldsymbol{\Omega}\right)^{2}\right\rangle, & \alpha>2,\end{cases}
\end{aligned}
$$

where the angular brackets denote averaging over the isotropically distributed vector $\boldsymbol{\Omega}$ and the power-law functions are treated, as usual, in the sense of the principal branch of the analytic function $z^{\alpha}$ in the plane with a cut along the positive semiaxis [89]:

$$
z^{\alpha}=|z|^{\alpha} \exp (\mathrm{i} \alpha \arg z),\left.\quad \lim _{\varepsilon \rightarrow 0} \arg z\right|_{z=s+\mathrm{i}, s>0}=0 .
$$

In accordance with this choice,

$$
\left(\frac{\lambda}{v}-i \mathbf{i} \boldsymbol{\Omega}\right)^{\alpha}=\left[\left(\frac{\lambda}{v}\right)^{2}+(\mathbf{k} \boldsymbol{\Omega})^{2}\right]^{\alpha / 2} \exp (-\mathrm{i} \phi \alpha)
$$

where

$$
\tan \phi=\frac{\mathbf{k} \boldsymbol{\Omega}}{\lambda / v} .
$$

The functions $P(\mathbf{k}, \lambda)$ corresponding to each of these intervals are found similarly.

Because the distribution of the random directions of $\boldsymbol{\Omega}$ is isotropic, the mean of $\boldsymbol{\Omega}$ is zero and the fractional diffusion asymptotic form of the operator $\tilde{L}$ as $k \rightarrow 0$ and $\lambda \rightarrow 0$ is

$$
\tilde{L}(\mathbf{k}, \lambda) \sim \begin{cases}A_{\alpha}\left\langle\left(\frac{\lambda}{v}-\mathbf{i} \mathbf{k} \boldsymbol{\Omega}\right)^{\alpha}\right\rangle, & 0<\alpha<1 \\ \langle R\rangle \frac{\lambda}{v}-A_{\alpha}\left\langle\left(\frac{\lambda}{v}-\mathbf{i} \mathbf{k} \boldsymbol{\Omega}\right)^{\alpha}\right\rangle, & 1<\alpha<2 \\ \langle R\rangle \frac{\lambda}{v}+k^{2}\left\langle\frac{R^{2}}{6}\right\rangle, & \alpha>2 .\end{cases}
$$

The corresponding equations in natural variables take the form

$$
\begin{array}{ll}
\left\langle\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right)^{\alpha}\right\rangle N(\mathbf{r}, t)=S_{\alpha}(\mathbf{r}, t), & 0<\alpha<1 \\
{\left[\frac{\partial}{\partial t}-\frac{A_{\alpha}}{v^{\alpha-1}\langle R\rangle}\left\langle\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right)^{\alpha}\right\rangle\right] N(\mathbf{r}, t)=\frac{1}{\langle R\rangle} S_{\alpha}(\mathbf{r}, t)} \\
& 1<\alpha<2 \\
\left(\frac{\partial}{\partial t}-\frac{v\left\langle R^{2}\right\rangle}{6\langle R\rangle} \Delta\right) N(\mathbf{r}, t)=\frac{1}{\langle R\rangle} S_{2}(\mathbf{r}, t), & \alpha>2 \tag{34}
\end{array}
$$

The pseudo-differential operator in Eqns (32) and (33) can be considered a fractional power of the material (total) derivative operator:

$$
\begin{gathered}
\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right)^{\alpha} N(\mathbf{r}, t)=\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right)\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right)^{\alpha-1} N(\mathbf{r}, t) \\
=\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right) \int_{0}^{t} \frac{N(\mathbf{r}-\mathbf{v}(t-\tau), \tau)}{\Gamma(1-\alpha)(t-\tau)^{\alpha}} \mathrm{d} \tau .
\end{gathered}
$$

In an homogeneous distribution of particles, the fractional material derivative of a function is equal to the fractional Riemann-Liouville time derivative, and in the stationary case, it is equal to the fractional directional derivative.

We recall that Eqns (32) and (33) respectively describe superdiffusion of the first and second kinds, and Eqn (33) can be simplified. Taking into account that its first term in the Fourier-Laplace representation contains $\lambda$ to the first power, and the second term contains $\lambda$ to a higher power, we can conclude that as $\lambda \rightarrow 0$, the term containing $\lambda$ in the second term can be neglected and the operator can be written in the form

$$
\tilde{L}(\mathbf{k}, \lambda) \sim\langle R\rangle \frac{\lambda}{v}-A_{\alpha}\left\langle(-\mathbf{i} \mathbf{k} \boldsymbol{\Omega})^{\alpha}\right\rangle, \quad 1<\alpha<2
$$

After averaging over directions in the last term,

$$
\begin{aligned}
& \left\langle(-\mathrm{i} \mathbf{k} \boldsymbol{\Omega})^{\alpha}\right\rangle=\frac{1}{2} k^{\alpha} \int_{-1}^{1}(-\mathrm{i} \mu)^{\alpha} \mathrm{d} \mu=\frac{\mathrm{i}}{2} k^{\alpha} \int_{\mathrm{i}}^{-\mathrm{i}} z^{\alpha} \mathrm{d} z \\
& \quad=\frac{\mathrm{i}}{2(\alpha+1)} k^{\alpha}\left\{\exp \left[-\mathrm{i}(\alpha+1) \frac{\pi}{2}\right]-\exp \left[\mathrm{i}(\alpha+1) \frac{\pi}{2}\right]\right\} \\
& \quad=\frac{A_{\alpha} \cos (\alpha \pi / 2)}{\alpha+1} k^{\alpha}=-\frac{A_{\alpha}|\cos (\alpha \pi / 2)|}{\alpha+1} k^{\alpha}, \quad 1<\alpha<2
\end{aligned}
$$

we arrive at the approximate representation of RAD equation (32) in the form of the UAD equation

$$
\begin{equation*}
\left(\frac{\partial}{\partial t}+D_{\alpha}(-\Delta)^{\alpha / 2}\right) N(\mathbf{r}, t)=\frac{1}{\langle R\rangle} S_{\alpha}(\mathbf{r}, t), \quad 1<\alpha<2, \tag{35}
\end{equation*}
$$

where

$$
D_{\alpha}=\frac{A_{\alpha}|\cos (\alpha \pi / 2)|}{v^{\alpha-1}(\alpha+1)\langle R\rangle} .
$$

Finally, the third equation describes normal diffusion with the diffusion coefficient

$$
D_{2}=\frac{v\left\langle R^{2}\right\rangle}{6\langle R\rangle}
$$

In the exponential distribution of free paths,

$$
p_{R}(r)=\frac{1}{\langle R\rangle} \exp \left(-\frac{r}{\langle R\rangle}\right)
$$

the mean square $\left\langle R^{2}\right\rangle=2\langle R\rangle^{2}$, and the diffusion coefficient takes the conventional form

$$
D_{2}=\frac{v\langle R\rangle}{3}
$$

The inclusion of independent power-law traps would result in adding a fractional partial time derivative to the left-hand sides of these equations [81].

In conclusion, we note that the solutions of the equations presented above are not continuous functions of the exponent $\alpha$ in the range of its values, changing jumpwise at $\alpha=1$ and 2 . For $\alpha<2$, the coefficient $D_{\alpha}$ is not the diffusion coefficient and does not transform into it even as $\alpha \uparrow 2$, because $D_{\alpha}$ is determined by the asymptotic behavior of the range distribution at large arguments. But the classical diffusion coefficient is determined by dispersion, being finite for $\alpha=2$, but infinite
for any $\alpha<2$, which means that as $\alpha$ approaches 2 from below, the dispersion limit is also infinite and cannot coincide with the classical value. This leads to a discontinuity.

### 3.5 Anomalous diffusion distributions

We see from Section 3.4 that the RAD equations are more complicated than the UAD equations, and they are more difficult to solve. However, the RAD equations have a certain advantage because, due to the boundedness of the spatial region of the RAD model, all moments of the RAD equations are finite. This opens up the possibility of using the method of moments, which has been tested in the classical theory, and also allows obtaining asymptotic analytic estimates of the moments.

In [85, 87], the time evolution of the root mean square

$$
\sqrt{\left\langle[R(t)]^{2}\right\rangle}=\left[\int r^{2} G(\mathbf{r}, t) \mathrm{d} \mathbf{r}\right]^{1 / 2}
$$

of a diffusion packet was considered. The Laplace transform of the root mean square,

$$
\int_{0}^{\infty} \exp (-\lambda t)\left\langle[R(t)]^{2}\right\rangle \mathrm{d} t=-\left.\frac{\partial^{2} \tilde{G}(\mathbf{k}, \lambda)}{\partial k^{2}}\right|_{\mathbf{k}=0} \equiv-\tilde{G}^{\prime \prime}(0, \lambda)
$$

was found from Eqn (28). Using the asymptotic form of the transforms $p_{R}(\lambda)$ and $q_{T}(\lambda)$ as $\lambda \rightarrow 0$,

$$
\begin{aligned}
& p_{R}(\lambda) \sim \begin{cases}1-A \lambda^{\alpha}, & 0<\alpha<1 \\
1-\langle R\rangle \lambda+C \lambda^{\alpha}, & 1<\alpha \leqslant 2\end{cases} \\
& q_{T}(\lambda) \sim 1-B \lambda^{\beta}, \quad 0<\beta \leqslant 1
\end{aligned}
$$

and Tauber's theorems [42], the inverse transformation was performed, with the result that, as $t \rightarrow \infty$,

$$
\left\langle[R(t)]^{2}\right\rangle \sim \begin{cases}(1-\alpha) v^{2} t^{2}, & \alpha<\beta<1, \\ \frac{A(1-\alpha) v^{2}}{A+B v^{\alpha}} t^{2}, & \alpha=\beta<1, \\ \frac{2(1-\alpha) A v^{2-\alpha}}{B \Gamma(3-\alpha+\beta)} t^{2-\alpha+\beta}, & \beta<\alpha<1, \\ \frac{2(\alpha-1) C v^{2-\alpha}}{B \Gamma(3-\alpha+\beta)} t^{2-\alpha+\beta}, & \alpha>1 .\end{cases}
$$

We see from the last line that for $\alpha>1$ and $\beta=\alpha-1$, we have the normal expansion law for the diffusion packet:

$$
\left\langle[R(t)]^{2}\right\rangle \sim 2(\alpha-1) \frac{C}{B} v^{2-\alpha} t
$$

However, the packet shape differs from normal, coinciding with it only for $\alpha=2$ and $\beta=1$ (a quasi-normal diffusion becomes normal). For $\beta<\alpha-1$, the packet spreads more slowly than the normal packet (the subdiffusion regime), while for $\beta>\alpha-1$, it spreads faster (superdiffusion). In all other cases, the condition $\alpha<1$ gives rise to superdiffusion.

The conclusions made in [85] are as follows. For $v=\infty$, the distribution mean square diverges and cannot be used to characterize the distribution width. Thus, the assumption of a finite velocity of motion between collisions in anomalous diffusion drastically changes the asymptotic behavior of the diffusion packet width as $t \rightarrow \infty$. In this case, subdiffusion appears only when the free path is finite (i.e., $\alpha>1$ ), the waiting time in a trap is distributed by a power law, and the condition $\beta<\alpha-1$ is satisfied. For $\beta>\alpha-1$ and any $\alpha<1$,
$\beta<1$ (the mean free path and mean waiting time are infinite), superdiffusion occurs. The linear time dependence of the root mean square for $\alpha \leqslant \beta$ means that the ballistic regime plays the main role in the asymptotic behavior; as $\alpha \rightarrow 0$, we obtain the free motion of a particle in a pure form:

$$
\sqrt{\left\langle[R(t)]^{2}\right\rangle} \sim v t
$$

We next discuss the distributions. We noted above that superdiffusion in the RAD model has two qualitatively different regimes: superdiffusion of the first kind $(\alpha \in(0,1))$ and superdiffusion of the second kind $(\alpha \in(1,2))$. The asymptotic analysis of the system of integral equations for the process of the second kind leads to the same fractional differential equation as for the UAD model, but with a corrected diffusion coefficient. ${ }^{7}$

To avoid cumbersome calculations, we present only qualitative considerations leading to the same result [21]. Let $T_{i}$ be the confinement time in a trap before the $i$ th jump and $R_{i}$ be its random length. Then a random instant of time

$$
\Theta_{n}=\sum_{i=1}^{n}\left(T_{i}+\frac{R_{i}}{v}\right)
$$

corresponds to the $n$th jump of a particle. According to the law of large numbers, we can write

$$
\Theta_{n} \approx t \sim\left(\langle T\rangle+\frac{\langle R\rangle}{v}\right) n .
$$

Hence,

$$
n=\frac{t}{\langle T\rangle+\langle R\rangle / v} .
$$

Introducing the notation

$$
t_{v}=\frac{t}{1+\langle R\rangle /\langle v T\rangle}
$$

we obtain the same distribution as in the case of normal jumps, but at a shifted instant of time:

$$
G(\mathbf{r}, t)=\left[D t_{v}^{\beta}\right]^{-3 / \alpha} \Psi_{3}^{(\alpha, \beta)}\left(r\left[D t_{v}^{\beta}\right]^{-1 / \alpha}\right), \quad 1<\alpha \leqslant 2 .
$$

The result is clear: the finite velocity of free motion slows the expansion of the diffusion packet compared to the case of unrestricted diffusion, where $v=\infty$. The replacement $t \rightarrow t_{v}$ takes this slowing down into account (in the asymptotic sense). However, it is convenient to add a correcting factor to the diffusion coefficient by introducing the notation

$$
D_{v}=\frac{D}{(1+\langle R\rangle /\langle v T\rangle)^{\beta}}
$$

and writing the result in the form
$G(\mathbf{r}, t)=\left[D_{v} t^{\beta}\right]^{-3 / \alpha} \Psi_{3}^{(\alpha, \beta)}\left(r\left[D_{v} t^{\beta}\right]^{-1 / \alpha}\right), \quad 1<\alpha \leqslant 2$.
Figure 15 demonstrates the applicability of this approximation. At first, of course, ballistic restrictions play their role,

[^5]

Figure 15. Influence of a finite velocity on the one-dimensional distribution in superdiffusion of the second kind $\left(\alpha=3 / 2, \beta=1, \mu=1, t=10^{3}\right)$. The dashed curve is obtained for $v=\infty$, the solid curve is calculated by expression (36) with $v=5$, the histogram is calculated by the Monte Carlo method for the same velocity $v=5$ over $2 \times 10^{5}$ trajectories [21].


Figure 16. Evolution of a superdiffusion packet in the process of the second kind. The influence of the velocity on the distribution shape is not observed, the distribution itself rapidly escapes from the 'custody' of ballistic bounds. The probability concentration is maximal at the center and virtually vanishes long before reaching the boundary.
constraining the diffusion packet, but with time the ballistic bound moves away from the packet, which expands with a smaller velocity, and the effect of the bound on the distribution shape decreases with time, disappearing completely in the asymptotic regime (Fig. 16). Hence, in superdiffusion of the second kind, which is described by the LU model, the assumption of a finite velocity in the asymptotic regime leads only to a change (decrease) in the diffusion coefficient in the equation, preserving the form of the solution itself expressed in terms of a spherically symmetric threedimensional stable density and satisfying the same fractional differential equation with the corrected coefficient.

However, this conclusion is valid if two conditions are satisfied. First, the modified process is constructed based on a jumpwise process with instantaneous flights alternating with the rest states of a particle. Therefore, the rest states remain in


Figure 17. Evolution of a superdiffusion packet in the process of the first kind [21]. A finite velocity drastically affects the distribution, which becomes U-shaped (histogram). At large times, the probability concentration is maximal at ballistic bounds and minimal at the center. In the case $v=\infty$, the distribution is described by a rapidly spreading stable density (solid curves near the horizontal axes).
the modified process, and alternate with flights with a finite velocity. As a result, the time interval between the beginning of the one flight and the beginning of the next flight consists of two parts, one of which is the confinement time of a particle in a trap, which is independent of other variables of the process, and the other is proportional to the flight length, which produces partial correlations of temporal and spatial intervals absent in the initial process. We emphasize once more that expression (36) is approximate and is derived assuming the dominant role of traps. The second restriction is the condition $\alpha>1$ (anomalous diffusion of the second kind), which was used in the replacement $\Theta_{n} \rightarrow t$.

For $\alpha<1$ (the process of the first kind), the situation drastically changes, as can be seen from the following considerations. The size ('width') of a superdiffusion packet increases with time proportionally to $t^{1 / \alpha}$. Due to a finite velocity, the distribution density outside a sphere with the radius $v t$ vanishes. For $\alpha<1$, a superdiffusion packet spreads faster than $v t$ (faster than for free motion) and the kinematic constraint becomes the main factor determining the asymptotic distribution of the diffusion packet. When restricted by a sphere of the radius $v t$, this distribution has a completely different form: it is W -shaped for $\alpha>1 / 2$ and U -shaped for $\alpha<1 / 2$ (see [90], Fig. 17 and also Figs 18 and 25 below). However, the authors of the LT version, setting $\alpha=0.3$, still use the unrestricted fractionally stable distribution, which is no longer related to the process under study. The applicability of this approach could be substantiated if the coefficient in the distribution of the residence time in traps was so large that all the times considered would be related to the beginning of the process, when the diffusion packet has not yet reached
ballistic bounds. Leaving aside the physical interpretation of such an unconventional assumption, we note that it is also beyond reason in this case to use the solutions of fractional differential equations; these equations describe the asymptotic behavior of a process caused by a large number of transitions, while the introduction of long residence times in traps reduces this number and pushes the random walk process away from the asymptotic region described by the fractional differential equation.

### 3.6 Range-time correlations

The trajectory of a charged particle in the interstellar magnetic field, together with its derivative, is an extremely complex continuous curve in the phase space, which can be described in detail by an integer-order differential equation if the magnetic field and other characteristics of the medium affecting the particle motion are also specified in detail. Needless to say, we have no such information.

We return to the idea of passing from a continuous description to a coarse-grain description (see Section 3.1), this time focusing our attention on the trajectory shape. If we were dealing with a homogeneous medium, the partition elements would be only slightly different from each other, and, in the finite-element representation, we would again obtain an analog of the classical differential equation with the diffusion coefficient somewhat corrected due to the coarsegraining. But the interstellar medium is not simply inhomogeneous in reality: its inhomogeneities have a multiscale character, as mentioned in the Introduction. This multiscale (fractal) property of the structure does not allow choosing the partition size for which the characteristics of elements could be assumed approximately identical. Selecting one such element, we find that the magnetic field in neighboring elements is an order of magnitude lower (and a particle freely goes away to a remote region) or higher (and a particle remains confined in one of the elements for a long time). The fractal structure assumes that this holds in a broad range of partition sizes. But while crossing the interface between neighboring elements occurs instantaneously, crossing a number of large-scale elements almost transparent to particles can no longer be assumed instantaneous, and the time spent for the passage should be taken into account. This is the difference between the RAD model, taking the time spent by a particle to move from one trap to another into account, and the UAD model, in which such a transition is assumed instantaneous irrespective of its scale. If $T$ is the time interval separating the instants of a particle arriving at some partition element and the next nonempty element, then in the RAD model $T$ consists of two terms: the residence time $T_{0}$ of a particle in the first element and the time $T_{1}$ spent to pass to the next element, whereas in the UAD model, this time is the residence time of a particle in a trap. Ranges and waiting times in both models are independent of each other, but the flight time in the RAD model is proportional to the range (for a constant velocity) and gives rise to a $T-R$ correlation,

$$
T=T_{0}+\frac{R}{v}
$$

As shown in Section 2.5, a fractional order $\beta<1$ of the time derivative means that residence times in traps are distributed with a density proportional to $t^{-\beta-1}, t \rightarrow \infty$, and the fractional order of the Laplacian suggests that ranges are distributed with a density proportional to $r^{-\alpha-1}, r \rightarrow \infty$. The


Figure 18. Influence of range-time correlations on the form of (a) one-dimensional and (b) three-dimensional propagators. Filled symbols correspond to the process with independent $R$ and $T$, and open symbols to the process with the linear dependence $R=v t$. The vertical straight lines $x= \pm v t$ show the boundaries of distributions.
unrelated terms $(\partial / \partial t)^{\beta}$ and $(-\Delta)^{\alpha / 2}$ in the UAD equation operator mean the absence of correlations between $T$ and $R$, and, as a consequence, a jumpwise (discontinuous) form of the trajectories, whereas the presence of the composite operator $\left\langle(\partial / \partial t+\mathbf{v} \nabla)^{\alpha}\right\rangle$ in the UAD equation indicates the presence of $R-T$ correlations.

Including $R-T$ correlations in the model
(i) transforms discontinuous trajectories into continuous ones, thereby eliminating nonphysical phantoms such as long instant flights and long rest times followed by instant gains of an infinite velocity;
(ii) restricts the spatial position of a particle by a spherical region of the radius $v t$ centered at the initial point, which allows reconciling the process with the relativistic concept, restricts the expansion law of the diffusion packet by a linear velocity, and returns the method of moments to the toolbox of computational techniques in the transfer theory [88] (which is inapplicable in the UAD model due to the divergence of moments);
(iii) for $\alpha<1$, considerably changes the shape of the spatial distribution, giving rise to splashes near the ballistic boundary $r=v t$ and thereby transforming the usual bellshaped diffusion packet to a W-shaped packet, and, for $\alpha<1 / 2$, to the U-shaped packet.

The differences between the UAD and RAD models are most strongly manifested for $\alpha, \beta<1$, when mathematical expectations of random variables are infinite and jumps of space-time trajectories can be seen at any scale. For $\alpha>1$ and $\beta=1$, the situation is different: as the size of the chosen spacetime region increases, the relative role of jumps becomes less noticeable and becomes insignificant when the expectation values are greatly exceeded. This is well seen for a usual Brownian trajectory: despite the independence of the spatial and temporal parts of the differential operator, the trajectory is everywhere continuous (although not differentiable).

To elucidate the role of $R-T$ correlations in the model under study, we compare propagators in the RAD and UAD models. In the RAD model, we assume that independent traps are absent $\left(T_{0}=0\right)$, and therefore $T$ and $R=v t$ completely correlate, such that their distributions coincide up to the scale factor $v$ (which we set equal to 1 ). In the UAD model, $v=\infty$ and random variables $T \equiv T_{0}$ and $R$ are completely independent. We take the same distribution $P(R>x)=P(T>x) \propto x^{-v}, v>0$, for them and note that in this case,

$$
\alpha=\left\{\begin{array}{ll}
v, & v \leqslant 2, \\
2, & v>2,
\end{array} \quad \beta= \begin{cases}v, & v \leqslant 1 \\
1, & v>1\end{cases}\right.
$$

The principal difference between the propagators disappears only for $v>2$. We can see from Fig. 18, representing calculations of one-dimensional and three-dimensional random walks in both models, that the $R-T$ correlations drastically change the process for $v \leqslant 1$ : the propagators differ in their form, expansion law, and behavior near ballistic boundaries and near the radiation source (we return to these distributions in the study of one-dimensional random walks along field lines in Section 4.5). For $1<v<2$ (anomalous diffusion of the second kind), the differences are weaker, although quite noticeable: in one case, distributions are bounded, while in the other, they are not bounded; in the RAD model, a front appears near the region $|\mathbf{r}|=v t$ and the densities are quantitatively different near the source.

### 3.7 Anisotropy

The anisotropy coefficient of a flux in the standard diffusion theory, determined by the ratio of the current density $J_{r}=J$ to the particle flux density $v N$,

$$
\begin{equation*}
\delta(r, t)=3 \frac{J(r, t)}{v N(r, t)}, \tag{37}
\end{equation*}
$$

is described at a distance $r$ from a point-like instantaneous source in a homogeneous infinite medium by the known expression [10]

$$
\begin{equation*}
\delta(r, t)=\frac{3 r}{2 v t} . \tag{38}
\end{equation*}
$$

We note that (i) the derivation of expressions (37) and (38) a priori assumes a weak anisotropy of the flux at the point under study and (ii) the diffusion approximation itself is applicable to random walks at rather large times, when the number of flights performed by a particle is large enough for the total displacement to reach the asymptotic regime.

In the case of anomalous diffusion based on the usual Fick law, this derivation is of course invalid; but the fact that anomalous diffusion equations preserve the self-similarity of their solutions is sufficient for finding the general expression relating the current and concentration, even without specifying the form of a self-similar solution. According to the physical meaning of $J$, we have

$$
\begin{equation*}
J(r, t)=\frac{1}{r^{2}} \frac{\mathrm{~d}}{\mathrm{~d} t} \int_{r}^{\infty} N(r, t) r^{2} \mathrm{~d} r \tag{39}
\end{equation*}
$$

in a spherically symmetric problem. Substituting $N(r, t)=G(\mathbf{r}, t)$ from (36) in (39) and changing the variable $r\left(D t^{\beta}\right)^{-1 / \alpha} \rightarrow \xi$ in the integrand, we find

$$
\begin{aligned}
& \int_{r}^{\infty} N(r, t) r^{2} \mathrm{~d} r=\int_{r}^{\infty}\left[D t^{\beta}\right]^{-3 / \alpha} \Psi_{3}^{(\alpha, \beta)}\left(r\left[D t^{\beta}\right]^{-1 / \alpha}\right) r^{2} \mathrm{~d} r \\
& \quad=\int_{r\left(D t^{\beta}\right)^{-1 / \alpha}}^{\infty} \Psi_{3}^{(\alpha, \beta)}(\xi) \xi^{2} \mathrm{~d} \xi
\end{aligned}
$$

Substituting this result in (39) and differentiating, we obtain a simple expression for the current density:

$$
J(r, t)=N(r, t) \frac{\beta}{\alpha} \frac{r}{t} .
$$

Therefore, the ratio of the current density $J$ to the concentration $N$ is given by a simple general relation for any (not necessarily fractionally stable) self-similar concentration
$N(r, t)=t^{-3 \beta / \alpha} \Phi\left(r t^{-\beta / \alpha}\right):$

$$
\frac{J(r, t)}{N(r, t)}=\frac{\beta}{\alpha} \frac{r}{t}
$$

If all the particles continuously moved with the same constant velocity $v$, the anisotropy coefficient would also have the model-independent form

$$
\begin{equation*}
\delta(r, t)=3 \frac{J(r, t)}{v N(r, t)}=3 \frac{\beta}{\alpha} \frac{r}{v t} \tag{40}
\end{equation*}
$$

This was the case in paper [73] describing the LT version of the fractional differential model: the anisotropy in this model was described by classical formula (37), in which the concentration $N$ was replaced with the product $k N_{\text {near }}$, where $k$ is the number of nearest sources and $N_{\text {near }}$ is the concentration of one of them. However, it is obvious that in this case, the motion of particles never satisfies the continuity condition mentioned above. On the contrary, the concentration of particles moving at a given instant in the LT model is zero, while the velocity of particles performing instant flights is infinite. This gives rise to an uncertainty in the product $v N$, which can lead to any result. Another consideration requiring a careful treatment of general expression (40) is that this expression is exact with respect to the fractional differential equation, but this equation itself is only an asymptotic form of the system of integral equations presented in Section 3.4. Integral equations include the distributions of ranges and time intervals between flights, and therefore correctly describe random walks with specified characteristics, whereas the fractional differential equation contains some part of this information in the diffusion coefficient, and the characteristics of flight lengths no longer can be separated from waiting times. In considering this model less formally, we must admit that the concept of particles at rest in traps cannot correspond to reality. Simply put, we are dealing with regions which a particle can leave only with difficulty because of the small diffusion coefficient strongly entangling the trajectory along which the particle can continue its motion with the same velocity $v$ (which for light particles (electrons and positrons) gives enhanced synchrotron radiation). In this case, we should consider real anisotropy at a point located in a trap (with $N$ meaning the total concentration) and anisotropy at a point located outside the trap (with $N$ being the concentration of particles not affected by traps).

These problems are eliminated to a great extent in the kinetic CTRW model: the introduction of a finite flight velocity gives finite time intervals during which the particle moves. In the case of diffusion of the second kind $(1<\alpha<2)$, this procedure does not change the distribution shape but corrects only the diffusion coefficient. As regards anomalous diffusion of the first kind $(\alpha<1)$, the situation is different. The LT version is inapplicable, while calculating the propagator with a finite velocity taken into account for $\alpha<1$ is a more complicated problem. Based on qualitative considerations, we can conclude that for $\alpha<\beta$ (we recall that $\alpha=0.3$ and $\beta=0.8$ in the LT version), the diffusion packet is a thin spherical shell adjacent from the inner side to the front $r=v t$. Such a behavior is close to the ballistic regime (and passes into it as $\alpha \rightarrow 0$ ). This is explained by the presence of the leading range in the trajectory, which is considerably longer than other ranges and therefore has the greatest probability of crossing the observation sphere. Due to the


Figure 19. Anisotropy coefficient for a point-like instantaneous source as a function of time in three models: classical diffusion $(\alpha=2, \beta=1)$, superdiffusion of the second kind ( $\alpha=1.67, \beta=1$ ), and anomalous diffusion of the first kind (the LT version: $\alpha=0.3, \beta=0.8$ ). The tilted straight lines are calculated by expression (40), the dots are Monte Carlo simulation results, and $R$ is the distance from the source.
overall smallness of other ranges, the intersection virtually occurs along the radius, and the real anisotropy coefficient in the LT version is close to unity (Fig. 19, the upper straight line). We note that even if the LT version corresponded to reality, it would still be meaningless to consider random walks with such long ranges without the inclusion of processes at the Galaxy boundaries.

### 3.8 Reaching the galactic disc boundary

Because of certain difficulties encountered in the treatment of boundary conditions in nonlocal problems (see Section 6.2), we performed preliminary Monte Carlo calculations of anomalous diffusion in the galactic disc; the method simply reproduces the motion of particles in a bounded medium and therefore requires specifying only the characteristics of the boundaries themselves; no special problems with their inclusion appears. We considered a particular model problem of finding the path and time for a particle to reach the galactic disc $[91,92]$ and compared the results obtained in the


Figure 20. Galactic disc as an infinite planar layer.
framework of three models: the standard Ginzburg-Syrovartskii (GS) model, the $\operatorname{LU} \operatorname{model}(\alpha=1.67, \beta=1.00, v=c$, $D_{0,1.67}=2.4 \times 10^{-3} \mathrm{pc}^{1.67} \mathrm{year}^{-1}$ ), and the LT version $\left(\alpha=0.3, \beta=0.8, v=\infty, D_{0,0.3}=4 \times 10^{-6} \mathrm{pc}^{0.3}\right.$ year $\left.{ }^{-0.8}\right)$. It is known that a large fraction of particles leave the galactic disc through its bases [6, 93], and we therefore considered the isotropic random walk of particles in an infinite layer with the thickness $2 h=300 \mathrm{pc}$ and with semitransparent boundaries (Fig. 20).

Introducing the notation $t_{1}$ for the mean time to first reach the boundary, $t_{2}$ for the mean time interval between successive arrivals at the boundary (the same or opposite), and $\varepsilon$ for the transparency coefficient of a specularly reflecting boundary (a particle incident on a boundary is reflected from it with the probability $1-\varepsilon$, the angle of reflection being equal to the angle of incidence), we express the mean residence time $t_{\mathrm{G}}$ of a particle in the galactic disc as

$$
\begin{aligned}
t_{\mathrm{G}} & =\varepsilon t_{1}+\varepsilon(1-\varepsilon)\left(t_{1}+t_{2}\right)+\varepsilon(1-\varepsilon)^{2}\left(t_{1}+2 t_{2}\right)+\ldots \\
& =t_{1}+\left(\frac{1}{\varepsilon}-1\right) t_{2}
\end{aligned}
$$

Figure 21 shows the probability distribution densities for the time, $p\left(t_{1}\right)$, and path, $p\left(s_{1}\right)$, before the boundary is first


Figure 21. (a) Time and (b) path distributions until the boundary is first reached.
reached for an homogeneous distribution of creation points of a particle in a layer for the $\operatorname{GS}(\alpha=2), \operatorname{LU}(\alpha=1.7)$, and LT ( $\alpha=0.3$ ) models and for free ballistic motion (some authors consider the latter as an acceptable type of motion in the leaky box model; see, e.g., [94]). We can see that the distribution $s$ predicted in the LT version is a few times narrower than diffusion distributions and virtually coincides with the ballistic distribution (this is consistent with the nearly unit anisotropy in the LT version). At the same time, the distribution of $t_{1}$ is an order of magnitude broader than in the GS model, which is explained by the unjustifiably large confinement time of particles in the traps. The time until the complete escape increases by another order of magnitude. As a result, as follows from calculations, the distribution of particles from a point-like instantaneous source in the LT version does not become a homogeneous distribution (over the transverse coordinate) in the galactic disc even in $10^{7}$ years, whereas the process of relaxation to the homogeneous distribution in the LU model is completed in less than $10^{5}$ years.

## 4. Anisotropic anomalous diffusion

### 4.1 Compound diffusion model

By considering the motion of cosmic-ray particles in regions of limited sizes where fluctuations of the direction and strength of the interstellar magnetic field are relatively small, we imagine magnetic field lines on which the trajectories of charged particles are wound. The leading centers of particles move along these lines, slowing down their motion in front of 'plugs', reflect from them, and return. In this case, the field lines do not remain motionless: they are displaced and bent, carrying the nearby charged particles with them, and the distances between line condensations and switchings change, complicating the motion of particles along field lines randomly walking in space and time. Following a set of such lines passing through the vicinity of point O in Fig. 22, we see how they begin to diverge from each other, demonstrating a statistical ensemble, which has been described in many papers.

A purely phenomenological model was proposed in [7], with the ensemble of magnetic field lines represented as a family of independent three-dimensional trajectories consisting of successive independent segments with random lengths and random directions along which particles perform one-


Figure 22. Ensemble of magnetic field lines.
dimensional diffusive random walks. Such a compound diffusion causes the slowing down of diffusion in transverse directions:

$$
\left\langle R_{\perp}^{2}(t)\right\rangle \propto t^{1 / 2}
$$

However, this slowing down is caused by the assumption about the diffusion motion of particles along the field lines: if we assume that particles move freely along these lines, we return to the asymptotically normal diffusion (if the root-mean-square length of segments is finite).

Later, the compound diffusion model was used for solving the problem of the motion of particles in a weakly inhomogeneous magnetic field (see, e.g., [4, 10, 13, 95-97]). In the simplest formulation, a region was considered in a homogeneous stationary random field with the mean value

$$
\langle\mathbf{H}\rangle \equiv \mathbf{H}_{0}=H_{0} \mathbf{e}_{z}
$$

and the autocorrelation function

$$
\begin{align*}
& \left\langle H_{i}\left(z_{2}\right) H_{j}\left(z_{1}\right)\right\rangle=\left\langle\delta H_{i}\left(z_{2}\right) \delta H_{j}\left(z_{1}\right)\right\rangle=H_{0}^{2} C_{i j}(\zeta),  \tag{41}\\
& \zeta=\left|z_{2}-z_{1}\right|
\end{align*}
$$

The only random process completely determined by these two characteristics (the mean value and the correlation function) is a Gaussian process, whose particular case is the Brownian motion characterized by independent increments and the distribution density satisfying the standard diffusion equation

$$
\frac{\partial M}{\partial z}=D_{L} \Delta_{\perp} M(x, y, z)
$$

with the diffusion coefficient $D_{L}>0$. A particle diffuses (in the modified compound model $[10,13]$ ) with the longitudinal diffusion coefficient $D_{\|}$along one of the realizations of this ensemble symmetrically continued to the region $z<0$ (in the statistical sense). In this case, again,

$$
\left\langle R_{\perp}^{2}(t)\right\rangle \simeq D_{L}\left(D_{\|} t\right)^{1 / 2} .
$$

We note, however, that an ensemble of Brownian trajectories, which are nowhere differentiable fractal curves [98] with independent increments (Fig. 23), is difficult to reconcile with the concept of magnetic field lines. Returning


Figure 23. (a) Brownian field line according to [13] and (b) its actual realization.
to piecewise smooth lines in [7], but assuming that particles move in one direction at a constant velocity along these lines, we obtain the normal diffusion of transverse displacements of particles in the long-time asymptotic regime. An analysis based on the kinetic equation performed in [18] yields the same result.

To approach the real properties of the magnetic field of the interstellar medium, the authors of [99-101] represented autocorrelation function (41) in terms of the Fourier component of magnetic field fluctuations:

$$
\begin{aligned}
& C_{x x}(z)=\frac{1}{(2 \pi)^{6}} \\
& \quad \times \int \mathrm{d} \mathbf{k} \int \mathrm{~d} \mathbf{k}^{\prime}\left\langle\widetilde{\delta H}_{x}(\mathbf{k}) \widetilde{\delta H}_{x}^{*}\left(\mathbf{k}^{\prime}\right) \exp \left[-\mathrm{i}\left(\mathbf{k} \mathbf{R}(z)-\mathbf{k}^{\prime} \mathbf{R}(0)\right)\right]\right\rangle .
\end{aligned}
$$

Using the Corrsin independence hypothesis [102] and neglecting the correlations of field components with different wave vectors $\mathbf{k}$ and $\mathbf{k}^{\prime}$, the authors obtained

$$
\begin{align*}
& C_{x x}(z)=\int\langle\exp (-\mathrm{i} \mathbf{k} \Delta \mathbf{R}(z))\rangle P_{x x}(\mathbf{k}) \mathrm{d} \mathbf{k}  \tag{42}\\
& \left.P_{x x}(\mathbf{k})=\left.(2 \pi)^{-6}\langle |{\widetilde{\delta H_{x}}}_{x}(\mathbf{k})\right|^{2}\right\rangle
\end{align*}
$$

where $\Delta \mathbf{R}(z)=\mathbf{R}(z)-\mathbf{R}(0)$ is a random transverse displacement vector over which the exponential is averaged. As a result, they obtained the expression

$$
\begin{aligned}
& \left\langle(\Delta X(z))^{2}\right\rangle \equiv \frac{2}{H_{0}^{2}} \int_{0}^{z}(z-\zeta) C_{x x}(\zeta) \mathrm{d} \zeta=\frac{2}{H_{0}^{2}} \int \mathrm{~d} \mathbf{k} P_{x x}(\mathbf{k}) \\
& \quad \times \int_{0}^{z}(z-\zeta) \cos \left(k_{\|} \zeta\right) \exp \left[-\frac{1}{2}\left\langle(\Delta X(\zeta))^{2}\right\rangle k_{\perp}^{2}\right] \mathrm{d} \zeta
\end{aligned}
$$

providing a basis for a more natural model of randomly walking field lines. They continued these calculations using the hybrid approach, where a turbulent axially symmetric field is described by a combination of the one-dimensional planar $\left(\mathbf{k} \| \mathbf{H}_{0}\right)$ and two-dimensional $\left(\mathbf{k} \perp \mathbf{H}_{0}\right)$ components:

$$
P_{x x}(\mathbf{k})=g^{\prime}\left(k_{\|}\right) \frac{\delta\left(k_{\perp}\right)}{k_{\perp}}+g^{\prime \prime}\left(k_{\perp}\right) \frac{\delta\left(k_{\|}\right)}{k_{\perp}}\left(1-\frac{k_{x}^{2}}{k^{2}}\right)
$$

where

$$
\begin{aligned}
& g^{\prime}\left(k_{\|}\right)=\frac{c(v)}{2 \pi} \frac{l^{\prime} \delta H_{\mathrm{ID}}^{2}}{\left(1+k_{\|}^{2} l_{\|}^{2}\right)^{v}}, \quad g^{\prime \prime}\left(k_{\perp}\right)=\frac{2 c(v)}{\pi} \frac{l^{\prime \prime} \delta H_{2 \mathrm{D}}^{2}}{\left(1+k_{\perp}^{2} l_{\perp}^{2}\right)^{v}}, \\
& c(v)=\frac{1}{2 \sqrt{\pi}} \frac{\Gamma(v)}{\Gamma(1-v / 2)},
\end{aligned}
$$

with $2 v$ being the spectral index in the inertial interval. The combination of these components, with an asymptotically small term neglected (as $z \rightarrow \infty$ ), gave the equation

$$
\begin{aligned}
\frac{\mathrm{d}^{2}}{\mathrm{~d} z^{2}}\left\langle(\Delta X(z))^{2}\right\rangle & =\frac{2 \pi g^{\prime \prime}(0)}{H_{0}^{2}} \int_{0}^{\infty} \exp \left[-\frac{1}{2}\left\langle(\Delta X(z))^{2}\right\rangle k_{\perp}^{2}\right] \mathrm{d} k_{\perp} \\
& =\frac{2 \pi}{H_{0}^{2}} \sqrt{\frac{\pi}{2}} \frac{g^{\prime \prime}(0)}{\sqrt{\left\langle(\Delta X(z))^{2}\right\rangle}} .
\end{aligned}
$$

The solution of this equation

$$
\begin{equation*}
\left\langle(\Delta X(z))^{2}\right\rangle \sim\left(\sqrt{\frac{\pi}{2}} \frac{9 \pi g^{\prime \prime}(0)}{2 H_{0}^{2}}\right)^{2 / 3} z^{4 / 3}, \quad z \rightarrow \infty \tag{43}
\end{equation*}
$$

shows that taking the autocorrelations of magnetic field lines into account in the framework of the Corrsin hypothesis [102] leads to a superdiffusion behavior of transverse displacements. We note that the variable $z$ is not time but the longitudinal coordinate, and therefore the emerging 'superdiffusion' relates not to the development of the diffusion process in time but to the rapid divergence of a bunch of magnetic field lines in space. As a result, the decomposition of the displacement of a particle into independent longitudinal and transverse components becomes incompatible with the asymptotic regime of the sought solution. In this case, it is convenient to pass from the linear coordinate $z$ to the curvilinear coordinate $s$ measured along a field line, and finally to the model of isotropically randomly walking lines. The simulation of these lines by continuous broken lines with linear segments of random lengths ('ranges') allows representing the large-scale correlations of magnetic fields by using nonexponential range distributions. Asymptotic power-law distributions, which seem to be a natural continuation of the regularities of large-scale turbulence, are convenient for this purpose.

Another reason why reliable results at large scales cannot be expected from the compound model is the limited possibilities of analytic tools. The motion of particles in random inhomogeneous fields is analyzed using the perturbation theory, more exactly, the theory of small perturbations. Our experience with the use of this tool shows that it is reliable only when the calculated correction caused by perturbations does not exceed $10-15 \%$ of the initial (unperturbed, i.e., already known) solution. Everything that exceeds this value is a 'deception'. This method is very efficient in problems related to atomic and nuclear physics: optical measurements are performed with a very high accuracy and the perturbation theory corrections are under control. But in cosmic-ray physics, we are typically dealing with orders of magnitudes, and methods having a narrow range of applicability, such as the theory of small perturbations, are not particularly useful in global problems where perturbations become comparable with characteristics of an unperturbed medium.

### 4.2 Fractional Brownian model of field lines

Taking a model in which the root-mean-square displacement is finite but increases faster than in the normal diffusion model, we return to Richardson's idea of describing turbulent superdiffusion (now not of particles themselves but of randomly walking field lines) by using the traditional mathematical tools (without fractional derivatives), but with a variable diffusion coefficient. Surprisingly, however, when the diffusion coefficient increases in accordance with a power law, we again come to fractional differential operators; there are two complementary methods for describing the random motion of particles. One of them, which is the main tool in this review, is based on the distribution density method ${ }^{8}$ and is related to Maxwell, Boltzmann, Einstein, Fokker, Planck, Bogoliubov, and others.

Let $X(t)$ be a random coordinate of a moving particle at an instant $t$ and $1_{\Delta x}\left(x^{\prime}\right)$ be the indicator function of the element $\Delta x$, equal to 1 for $x^{\prime} \in \Delta x$ and zero otherwise. The simplest (single-time) density $f(x, t)$ is determined by averaging the indicator function over an ensemble of particle

[^6]trajectories,
$$
f(x, t) \Delta x \approx P(X(t) \in \Delta x)=\left\langle 1_{\Delta x}(X(t))\right\rangle
$$
and then the (Boltzmann, Einstein, Fokker-Planck, Bogoliubov, ...) kinetic equation is derived and solved for it.

In the alternative approach (Langevin, Stratonovich, Ito, and others), the Newtonian equation for a random coordinate (stochastic equation),

$$
m \frac{\mathrm{~d} X(t)}{\mathrm{d} t}=F(t)
$$

with a random force in the right-hand side is considered first. Specifying the required properties of the random process $F(t)$ and solving this equation, we can study the properties of the solution as an ensemble of random functions (correlations and higher-order moments of the functions, probabilities of various events, for example, the time the boundary is first reached, etc.). It is in this stochastic approach that fractional operators, which were eliminated from the kinetic approach, appear. We consider the stochastic approach in greater detail.

We somewhat generalize expression (43) by replacing the exponent $4 / 3$ with an arbitrary value $2 H$ and the coordinate $z$ with the curvilinear coordinate $s$ along a magnetic field line,

$$
\left\langle(\Delta X(s))^{2}\right\rangle=K s^{2 H}
$$

and write the distribution of the coordinate $x$ of a point $s$ of a planar field line in the form of the usual Gaussian density distribution

$$
\begin{equation*}
f^{H}(x, s)=\frac{1}{\sqrt{4 \pi K s^{2 H}}} \exp \left(-\frac{x^{2}}{4 K s^{2 H}}\right) \tag{44}
\end{equation*}
$$

This distribution satisfies the normal diffusion equation with the variable diffusion coefficient

$$
\frac{\partial f^{H}(x, s)}{\partial s}=2 H K s^{2 H-1} \frac{\partial^{2} f^{H}(x, s)}{\partial x^{2}}
$$

and the initial condition

$$
f^{H}(x, 0)=\delta(x)
$$

The distribution $f^{H}(x, s)$, which is not itself a solution of a fractional derivative equation, characterizes an ensemble of trajectories of fractional Brownian motion $B^{H}(s)$ - a nonMarkov Gaussian process with the zero mean and the correlation function that, unlike the correlation function of the usual Brownian motion,

$$
\left\langle B\left(s_{1}\right) B\left(s_{2}\right)\right\rangle=\frac{1}{2}\left(\left|s_{1}\right|+\left|s_{2}\right|-\left|s_{1}-s_{2}\right|\right)
$$

has the form ${ }^{9}$

$$
\left\langle B^{H}\left(s_{1}\right) B^{H}\left(s_{2}\right)\right\rangle=\frac{1}{2}\left(\left|s_{1}\right|^{2 H}+\left|s_{2}\right|^{2 H}-\left|s_{1}-s_{2}\right|^{2 H}\right)
$$

[^7](see [103-105]). Similarly to the usual Brownian motion, its fractional analog is a self-similar process, and the Hurst parameter value determines the order of the process selfsimilarity, $B^{H}(a s) \stackrel{d}{=} a^{H} B^{H}(s)(\stackrel{d}{=}$ means the equality of the distributions of random quantities rather than their values). The parameter $H$ also characterizes the type of the process memory: long-term memory (persistent motion or superdiffusion) for $H>1 / 2$ and short-term memory (anti-persistent motion, subdiffusion) for $H<1 / 2$. For $H=1 / 2$, the process is the usual Brownian motion (memory is absent: increments are independent of the prehistory).

To elucidate the specificity of this random process, we note that the classical Brownian motion can be represented for $t \geqslant 0$ as an integral of the white noise $\xi(s)$,

$$
B_{+}(s)=\int_{0}^{s} \xi\left(s^{\prime}\right) \mathrm{d} s^{\prime} \equiv{ }_{0} I_{s} \xi(s), \quad s>0,
$$

whose properties and simulation methods are well known. Fractional Brownian motion was introduced in [104] by replacing the usual integral operator with its fractional analog

$$
\begin{align*}
B_{+}^{H}(s) & ={ }_{0} I_{s}^{H+1 / 2} \xi(s) \equiv \frac{1}{\Gamma(H+1 / 2)} \int_{0}^{s}\left(s-s^{\prime}\right)^{H-1 / 2} \xi\left(s^{\prime}\right) \mathrm{d} s^{\prime} \\
& =\frac{1}{\Gamma(H+1 / 2)} \int_{0}^{s}\left(s-s^{\prime}\right)^{H-1 / 2} \mathrm{~d} B_{+}\left(s^{\prime}\right) \tag{45}
\end{align*}
$$

continued to the entire real axis. Process (45) can be called conditional fractional Brownian motion, defined such that $B_{+}(s)=0$ for $s \leqslant 0$. There is no problem to continue this process to the entire time axis: it suffices to join it with the independent time-reflected process $B_{+}(-s)$ :

$$
B(s) \stackrel{d}{=} \begin{cases}B_{+}(s), & s>0 \\ B_{+}(-s), & s<0\end{cases}
$$

Problems do not appear here because Brownian motions in adjacent intervals are independent, and it was only necessary to ensure the continuity, but this occurred automatically because Brownian motion satisfies the zero initial condition at one instant of time $t=0$. However, fractional Brownian motion gives rise to correlations of increments, and its behavior on the positive time semiaxis depends on its prehistory on the negative semiaxis, $t<0$, and not only on the value at the instant $t=0$. The accurate 'sewing' of this past part of the process $[104,105]$ gives

$$
B^{H}(s)=\frac{1}{C(H)} \int_{-\infty}^{\infty}\left[\left(s-s^{\prime}\right)_{+}^{H-1 / 2}-\left(-s^{\prime}\right)_{+}^{H-1 / 2}\right] \mathrm{d} B\left(s^{\prime}\right)
$$

where

$$
C(H)=\frac{\Gamma(H+1 / 2)}{\sqrt{\Gamma(2 H+1) \sin (\pi H)}}
$$

This process is also expressed in terms of a fractional integral of the white noise $\xi(s)=\dot{B}(s)$ in the form

$$
B^{H}(s)=\frac{\Gamma(H+1 / 2)}{C(H)} \int_{0}^{s}-\infty I_{s}^{H-1 / 2} \dot{B}\left(s^{\prime}\right) \mathrm{d} s^{\prime}
$$

which follows from the Lagrangian adjointness of fractional operators

$$
\left(\varphi,-\infty I_{s}^{\alpha} \psi\right)=\left({ }_{s} I_{\infty}^{\alpha} \varphi, \psi\right)
$$

and the readily verified equality

$$
\Gamma(\alpha+1)_{s} I_{\infty}^{\alpha} 1_{[a, b)}(s)=(b-s)_{+}^{\alpha}-(a-s)_{+}^{\alpha},
$$

where $(c-s)_{+}^{\alpha}=(c-s)^{\alpha} 1_{+}(c-s)$. Because passing to the fractional process involves integrating the trajectory over its prehistory (the inclusion of memory), Brownian lines, broken at all their points and containing no smooth segments, are locally smoothed and become suitable for the imitation of real field lines, while the variable $s$ acquires the meaning of the actual length. We note that this passage involves the Hurst parameter, which can be adjusted to efficiently control the properties of the entire ensemble.

In the model of turbulent divergence of magnetic field lines considered in (43), the role of time is played by the length $s$ of the line segment, which for small values is close to the length of a segment corresponding to the $z$ axis, and the Hurst exponent is $H=2 / 3$. This corresponds to the distribution

$$
f(x, s)=\frac{1}{\sqrt{4 \pi K s^{4 / 3}}} \exp \left(-\frac{x^{2}}{4 K s^{4 / 3}}\right), s>0
$$

which satisfies the equation

$$
\frac{\partial f(x, s)}{\partial s}=\frac{4}{3} K s^{1 / 3} \frac{\partial^{2} f(x, s)}{\partial x^{2}}
$$

with the initial condition

$$
f(x, 0)=\delta(x)
$$

### 4.3 Transverse diffusion in fractional-order operators

We now return to calculations performed in the approximation of small perturbations, not only to rehabilitate them as regards the description of the process of cosmic-ray galactic diffusion (the applicability of this approximation in local problems with a weakly inhomogeneous magnetic field raising no doubts), but also because cosmic-ray transfer equations with fractional-order derivatives appeared for the first time in these calculations.

Starting with the collisionless Vlasov equation for charged particles in an electromagnetic field

$$
\frac{\partial f}{\partial t}+\mathbf{v} \nabla f+\mathbf{F} \frac{\partial f}{\partial \mathbf{p}}=0, \quad \mathbf{F}=q \mathbf{E}+\frac{q}{c}[\mathbf{v} \mathbf{H}]
$$

and separating fluctuations from the means,

$$
\mathbf{H}=\langle\mathbf{H}\rangle+\mathbf{H}_{1}, \quad\left\langle\mathbf{H}_{1}\right\rangle=0,
$$

and then using the quasilinear approximation for a weakly turbulent plasma with small-scale fluctuations [106-109],

$$
f_{1} \ll\langle f\rangle,
$$

the authors of $[14,15]$ considered the equation

$$
\frac{\partial\langle f\rangle}{\partial t}+\mathbf{v} \nabla\langle f\rangle+\langle\mathbf{F}\rangle \frac{\partial\langle f\rangle}{\partial \mathbf{p}}=R,
$$

whose right-hand side, being the usual collisional term describing scattering of particles by small-scale and smallamplitude fluctuations, is expressed as a convolution of the scalar function $\mathbf{F}_{1} \partial\langle f\rangle / \partial \mathbf{p}$ and the Green's function of a linear equation (see [110, 111]). The approximation of this
term by the relaxational expression $-(f-\bar{f}) / \tau$, where

$$
\bar{f}=\bar{f}(\mathbf{r}, p, t)=\frac{1}{4 \pi} \int_{4 \pi} f(\mathbf{r}, p \mathbf{\Omega}, t) \mathrm{d} \boldsymbol{\Omega},
$$

leads to the equation

$$
\frac{\partial\langle f\rangle}{\partial t}+\mathbf{v} \nabla\langle f\rangle+\frac{q}{c}[\mathbf{v}\langle\mathbf{H}\rangle] \frac{\partial\langle f\rangle}{\partial \mathbf{p}}=-\frac{\langle f\rangle-\bar{f}}{\tau}
$$

where $\tau$ is the characteristic time of scattering by a small-scale inhomogeneity, the angular brackets denote averaging over small-scale fluctuations, and the bar over $f$ denotes time averaging. We let $\langle\bar{f}\rangle$ be denoted by $N$ and the anisotropic component by $N_{1}$ :

$$
\langle f\rangle=N+N_{1}, \quad \bar{N}_{1}=0 .
$$

In the limit of large distances $(R \gg v \tau)$ and large times $(t \gg R / v)$, the angular distribution is almost isotropic and averaging over directions leads to the equation

$$
\frac{\partial N}{\partial t}+\nabla\left(\overline{\mathbf{v}}_{1}\right)=0
$$

which in turn reduces to the standard diffusion equation

$$
\frac{\partial N}{\partial t}-\nabla_{i} \kappa_{i j} \nabla_{j} N=0
$$

with a local tensor diffusion coefficient $\kappa_{i j}$. The longitudinal, transverse, and asymmetric components of this equation are expressed in terms of the velocity $v$ and gyrofrequency $\omega_{H}=q\langle H\rangle /(p c)$ by approximate formulas

$$
\begin{aligned}
\kappa_{\|} & =\frac{v^{2} \tau}{3} \\
\kappa_{\perp} & =\frac{\kappa_{\|}}{1+\left(\omega_{H} \tau\right)^{2}} \\
\kappa_{\mathrm{A}} & =-\frac{\kappa_{\|} \omega_{H} \tau}{1+\left(\omega_{H} \tau\right)^{2}}
\end{aligned}
$$

(see book [108]). These expressions reflect the influence of weak long-wavelength perturbations of a magnetic field on the diffusion of charged particles. ${ }^{10}$ They show, in particular, that for $\omega_{H} \tau \gg 1$ (strongly magnetized plasma), we have $\kappa_{\perp} / \kappa_{\|} \simeq \simeq\left(\omega_{H} \tau\right)^{-2} \ll 1$ and $\kappa_{\mathrm{A}} / \kappa_{\|} \simeq\left(\omega_{H} \tau\right)^{-1} \ll 1$, and therefore transverse displacements of a particle are small.

Magnetic field fluctuations $\Delta \mathbf{H}=\mathbf{H}-\mathbf{H}_{0}$ give rise to fluctuations of $\kappa_{i j}$ and $N$. We let double angular brackets denote averaging over these large-scale fluctuations and $\Delta \kappa$ and $\Delta N$ deviations from these averaged values. Then

$$
\frac{\partial N}{\partial t}-\nabla_{i}\left\langle\left\langle\kappa_{i j}\right\rangle\right\rangle \nabla_{j} N-\nabla_{i}\left\langle\left\langle\Delta \kappa_{i j} \nabla_{j} \Delta N\right\rangle\right\rangle=0
$$

Fluctuations $\Delta N$, characterizing rapid random spatial changes, in turn experience slow variations in both longitudinal and transverse directions.

[^8]The authors of [14] obtained the equation (Eqn (B11) in [14])

$$
\begin{equation*}
\frac{\partial N_{\perp}\left(\mathbf{r}_{\perp}, t\right)}{\partial t}-D_{\perp} \Delta_{\perp}\left[N(\mathbf{r}, t)-\frac{1}{2} \int_{1}^{\infty} N\left(\mathbf{r}, t-\frac{y L^{2}}{\kappa_{\|}}\right) y^{-3 / 2} \mathrm{~d} y\right]=0, \tag{46}
\end{equation*}
$$

for the perpendicular diffusion of charged particles in a magnetic field with relatively small fluctuations $\left(\Delta \mathbf{H}(\mathbf{r})=\mathbf{A}(\mathbf{r}) H_{0}\right.$ with $\left.A \ll H_{0}, \mathbf{r}=\left(\mathbf{r}_{\perp}, z\right)\right)$, where $D_{\perp}=$ $\langle\langle A\rangle\rangle \kappa_{\|} / 2$ and $L$ is the correlation length of field fluctuations. Fractional derivatives are still absent here, but if we represent the expression in square brackets in the form

$$
\begin{aligned}
& N(\mathbf{r}, t)-\frac{1}{2} \int_{1}^{\infty} N\left(\mathbf{r}, t-\frac{y L^{2}}{\kappa_{\|}}\right) y^{-3 / 2} \mathrm{~d} y \\
& \quad=\frac{1}{2} \int_{1}^{\infty}\left[N(\mathbf{r}, t)-N\left(\mathbf{r}, t-\frac{y L^{2}}{\kappa_{\|}}\right)\right] y^{-3 / 2} \mathrm{~d} y \\
& \quad=L \sqrt{\frac{\pi}{\kappa_{\|}}}\left\{\frac{1}{\Gamma(-1 / 2)} \int_{L^{2} / \kappa_{\|}}^{\infty}(N(\mathbf{r}, t-\tau)-N(\mathbf{r}, t)) \tau^{-3 / 2} \mathrm{~d} \tau\right\}
\end{aligned}
$$

and, using the condition $t \gg t_{d}=L^{2} / \kappa_{\|}$, replace the lower integration limit by zero, the expression in braces becomes a fractional semi-derivative (in the Marchaut form, but identically equal to a fractional derivative in the RiemannLiouville form we used; see [33]). This was done in [97], where the transverse diffusion equation was represented in the fractional differential form

$$
\frac{\partial N_{\perp}}{\partial t}=D_{\perp} \frac{\partial^{1 / 2}}{\partial t^{1 / 2}} \Delta_{\perp} N_{\perp}\left(\mathbf{r}_{\perp}, t\right)+N_{\perp}\left(\mathbf{r}_{\perp}, 0\right) \delta(t)
$$

which, after applying the operator $(\partial / \partial t)^{-1 / 2}$, yields the equation

$$
\begin{equation*}
\frac{\partial^{1 / 2} N_{\perp}}{\partial t^{1 / 2}}=D_{\perp} \Delta_{\perp} N_{\perp}\left(\mathbf{r}_{\perp}, t\right)+N_{\perp}\left(\mathbf{r}_{\perp}, 0\right) \delta_{1 / 2}(t) \tag{47}
\end{equation*}
$$

The authors of [14] point out an important feature of their equation for the distribution of transverse displacements, preserved in its fractional differential version: "The presence of integrals in the equation means that particles have a good 'memory' during compound diffusion, the values of the function $N$ at the instant $t$ being dependent on the values of $N$ at preceding instants. Unlike the usual diffusion, compound diffusion is not a Markov process."

We note, however, that although these equations were mathematically obtained for the asymptotic behavior of the distribution (formally, for $t \rightarrow \infty$ ), their applicability is also restricted at long times, because, as mentioned above, a number of physical assumptions no longer correspond to the process. The authors of [14] begin the derivation of the corresponding equations with the comment: "We assume that a random magnetic field is small compared to the mean homogeneous magnetic field, i.e., $A \ll 1$. This means that the field lines weakly deviate from the direction of the mean field $\mathbf{H}_{0}=$ const. The compound diffusion proceeds in the plane perpendicular to $\mathbf{H}_{0}$, and the length parameter along a field line $S$ can be approximately replaced by displacement along the direction of $\mathbf{H}_{0}$." It is clear that the increase in the effect of strong large-scale inhomogeneities during the development of diffusion makes this process isotropic, and the correctness of
the assumptions used in its description and therefore of the equations derived is rapidly lost. For this reason, long-term correlations represented by the function $y^{-3 / 2}$ in the integral term in Eqn (46) are then shortened, the integral no longer balances the preceding term, its role decreases with time, and the transverse diffusion equation in the large-time limit takes the normal form: ${ }^{11}$

$$
\frac{\partial N\left(\mathbf{r}_{\perp}, t\right)}{\partial t}-D_{\perp} \Delta_{\perp} N\left(\mathbf{r}_{\perp}, t\right)=0
$$

This confirms the conclusion in [18] that the motion of charged particles in a random magnetic field eventually becomes normal diffusion (this conclusion was made assuming that field fluctuations are isotropic and have a small scale, the kinetic operator is Hermitian, and the first approximation of the perturbation theory with respect to the ratio of the particle free path to the field correlation length is valid).

In the conclusion in [18], the authors consider the simple problem of Brownian motion of a particle in the field of random velocities and again demonstrate the diffusive character of the asymptotic regime. This is quite consistent with the results obtained long ago in paper [112] devoted to the same problem, where it was also shown that this process at shorter times is described by an integro-differential equation like the one derived in [14]. That we are dealing here with different time regions is confirmed, in particular, in the second part of [113], where the Parker equation [114] describing the motion of cosmic rays in plasmas is analyzed taking convection, diffusion, and transfer into account. The authors of [113] also showed that the transverse displacement equation, having a fractional differential form at small times, asymptotically becomes the normal diffusion equation at large times. In other words, the solution of the fractional differential subdiffusion equation describes the pre-asymptotic behavior of the process, which is often called the intermediate asymptotic regime [115].

Using the rules for dealing with fractional derivatives [33], we can represent the fractional differential subdiffusion equation in three equivalent forms: ${ }^{12}$

$$
\begin{align*}
& \frac{\partial^{\beta} G_{\perp}}{\partial t^{\beta}}=D_{\perp} \Delta_{\perp} G_{\perp}\left(\mathbf{r}_{\perp}, t\right)+\delta\left(\mathbf{r}_{\perp}\right) \delta_{\beta}(t)  \tag{48}\\
& \frac{\partial G_{\perp}}{\partial t}=D_{\perp} \Delta_{\perp} \frac{\partial^{1-\beta} G_{\perp}\left(\mathbf{r}_{\perp}, t\right)}{\partial t^{1-\beta}}+\delta\left(\mathbf{r}_{\perp}\right) \delta(t),  \tag{49}\\
& G_{\perp}\left(\mathbf{r}_{\perp}, t\right)=D_{\perp} \Delta_{\perp} \frac{\partial^{-\beta} G_{\perp}\left(\mathbf{r}_{\perp}, t\right)}{\partial t^{-\beta}}+\delta\left(\mathbf{r}_{\perp}\right) 1_{+}(t) . \tag{50}
\end{align*}
$$

Integral subdiffusion equation (50) in a multidimensional space was solved in [116]; the solution was expressed in terms of the Fox function. In the case of two-dimensional random walk, which is of interest to us, this solution has the form

$$
\begin{aligned}
& G_{\perp}(r, t)=\frac{1}{\beta \pi r^{2}} H_{12}^{20}\left(\left.\left(\frac{r}{2 \sqrt{D_{\perp}}}\right)^{2 / \beta} t^{-1}\right|_{(1,1 / \beta),(1,1 / \beta)} ^{(1,1)}\right), \\
& r \equiv\left|\mathbf{r}_{\perp}\right| .
\end{aligned}
$$

[^9]The Mellin transform of this distribution is expressed in the form

$$
\begin{aligned}
\bar{G}_{\perp}(s, t) & \equiv \int_{0}^{\infty} r^{s-1} G(r, t) \mathrm{d} r \\
& =\frac{\Gamma(s / 2) \Gamma(s / 2-1)}{2^{3-s} \pi \beta \Gamma(\beta(s / 2-1))}\left(D_{\perp} t^{\beta}\right)^{s / 2-1}
\end{aligned}
$$

In [22], where Eqn (48) was considered, another form of the solution was found, which relates it to fractionally stable distributions,

$$
\begin{equation*}
G_{\perp}(r, t)=\left(D_{\perp} t^{\beta}\right)^{-1} \Psi_{2}^{(2, \beta)}\left(\frac{r}{\sqrt{D_{\perp} t^{\beta}}}\right), \tag{51}
\end{equation*}
$$

where

$$
\begin{equation*}
\Psi_{2}^{(2, \beta)}(r)=\frac{1}{4 \pi} \int_{0}^{\infty} \mathrm{d} \tau \exp \left(-\frac{r^{2} \tau^{\beta}}{4}\right) \tau^{\beta} g_{+}(\tau ; \beta), \quad 0<\beta<1 \tag{52}
\end{equation*}
$$

It was shown in [22] in particular, that the necessary and sufficient condition for the appearance of the subdiffusion regime of the fractional differential type is the power-law asymptotic behavior of the waiting time in a trap with infinite dispersion,

$$
Q(t)=\int_{t}^{\infty} q(\tau) \mathrm{d} \tau \sim B_{0} t^{-\beta}, \quad t \rightarrow \infty,
$$

and Eqns (48) and (49) were derived based on simple probabilistic considerations. According to the generalized limit theorem, if independent positive random quantities $T_{i}$ are distributed with a density $q(t)$ satisfying this condition, then the normalized sum

$$
S_{n}=\sum_{i=1}^{n} \frac{T_{i}}{(n B)^{1 / \beta}}, \quad B=B_{0} \Gamma(1-\beta)
$$

is distributed at large $n$ with a stable one-sided density $g_{+}(t ; \beta)$, while the distribution density of the usual sum $\sum_{i=1}^{n} T_{i}$, characterized by the multiple convolution $q^{(n)}(t)$ $\left(q^{(1)}(t)=q(t)\right)$, has the large- $n$ asymptotic behavior

$$
q^{(n)}(t) \sim(n B)^{-1 / \beta} g_{+}\left((n B)^{-1 / \beta} t ; \beta\right) .
$$

To see more clearly how these statements relate to the transverse diffusion, we discuss the concept of a trap in this problem in view of the stochastic interpretation substantiated in Section 2.5. The position of a particle in a plane perpendicular to the $z$ axis, along which the main magnetic field is directed and the particle begins to move, is characterized by the two-dimensional vector $\mathbf{r}_{\perp}$. The statement "a particle is in a trap at the point $\mathbf{r}_{\perp}$ " means in this case that the particle moves along a line parallel to the $z$ axis, and its passage from one trap to another means its passage from one field line to another, also parallel to the $z$ axis. Neglecting the times spent for passages because of their smallness compared with the longitudinal motion time, we write the distribution of the number $v$ of captures of particles in traps for the time $t \rightarrow \infty$ in the form

$$
\begin{aligned}
P(v=n) & \approx Q^{(n)}(t)-Q^{(n+1)}(t)=\int_{0}^{t /(n B)^{1 / \beta}} g_{+}(\tau ; \beta) \mathrm{d} \tau \\
& -\int_{0}^{t /[(n+1) B]^{1 / \beta}} g_{+}(\tau ; \beta) \mathrm{d} \tau .
\end{aligned}
$$

Representing the upper limit of the second integral in the form

$$
\begin{aligned}
& {[(n+1) B]^{-1 / \beta} t=(n B)^{-1 / \beta} t-\varepsilon_{n}(t),} \\
& \varepsilon_{n}(t) \equiv(n B)^{-1 / \beta} t(n \beta)^{-1}
\end{aligned}
$$

and expanding this integral in a series in the small quantity $\varepsilon_{n}(t)$ (for $n \rightarrow \infty$ ), we find

$$
\begin{equation*}
P(v=n) \sim(n B)^{-1 / \beta} t(n \beta)^{-1} g_{+}\left((n B)^{-1 / \beta} t ; \beta\right), \quad t \rightarrow \infty \tag{53}
\end{equation*}
$$

The required propagator is expressed in terms of the conditional propagator $G_{\perp}\left(\mathbf{r}_{\perp}, t \mid n\right)$ by the formula for the total probability,

$$
\begin{equation*}
G_{\perp}\left(\mathbf{r}_{\perp}, t\right)=\sum_{n} G_{\perp}\left(\mathbf{r}_{\perp}, t \mid n\right) P(v=n), \tag{54}
\end{equation*}
$$

in which the conditional distribution

$$
\begin{equation*}
G_{\perp}\left(\mathbf{r}_{\perp}, t \mid n\right) \approx \frac{1}{4 \pi \sigma^{2} n} \exp \left(-\frac{r^{2}}{4 \sigma^{2} n}\right) \equiv\left(\sigma^{2} n\right)^{-1} g_{2}\left(\frac{r}{\sigma \sqrt{n}} ; 2\right) \tag{55}
\end{equation*}
$$

is a corollary of the central limit theorem. Substituting (53) and (55) in (54) and replacing the summation over $n$ with integration over the variable

$$
\tau=(n B)^{-1 / \alpha} t,
$$

we obtain the transverse propagator in (51).
For $\beta=1 / 2$, the one-sided stable density is expressed in terms of elementary functions,

$$
g_{+}\left(t ; \frac{1}{2}\right)=\frac{1}{2 \sqrt{\pi}} t^{-3 / 2} \exp \left(-\frac{1}{4 t}\right),
$$

and we obtain the transverse propagator

$$
G_{\perp}\left(\mathbf{r}_{\perp}, t\right)=\frac{1}{D_{\perp} \sqrt{t}} \Psi_{2}^{(2,1 / 2)}\left(\frac{r_{\perp}}{\sqrt{D_{\perp} \sqrt{t}}}\right), \quad r_{\perp}=\left|\mathbf{r}_{\perp}\right| .
$$

All moments of this propagator are finite and are given by

$$
\left\langle r_{\perp}^{2 n}\right\rangle=\frac{(n!)^{2}}{\Gamma(n / 2+1)}\left(4 D_{\perp} \sqrt{t}\right)^{n}, \quad n=1,2,3, \ldots
$$

The fractionally stable density

$$
\begin{gather*}
\Psi_{2}^{(2,1 / 2)}(\xi)=\frac{1}{4 \pi} \int_{0}^{\infty} \exp \left(-\frac{\xi^{2} \sqrt{\tau}}{4}\right) \sqrt{\tau} g_{+}\left(\tau ; \frac{1}{2}\right) \mathrm{d} \tau \\
\quad=\frac{1}{8 \pi^{3 / 2}} \int_{0}^{\infty} \exp \left(-\frac{\xi^{2} \sqrt{\tau}}{4}-\frac{1}{4 \tau}\right) \frac{\mathrm{d} \tau}{\tau}, \quad \xi>0 \tag{56}
\end{gather*}
$$

is presented in Fig. 24 together with the normal density with the same dispersion. We can see that at small and large distances, the subdiffusion density of transverse displacements exceeds the normal density, whereas for intermediate distances, the situation is the opposite. Unlike the normal density, the density $\Psi_{2}^{(2,1 / 2)}(\xi)$ at the coordinate origin has an integrable logarithmic singularity. This can be easily verified by dividing the integral in (52) into two parts and replacing the density $g_{+}(\tau ; 1 / 2)$ in the second part with its asymptotic


Figure 24. Fractional differential density (56) characterizing the distribution of the transverse displacement vector of a randomly walking particle (solid curve) and the normal density with the same dispersion (dashed curve).
expression,

$$
\begin{aligned}
\Psi_{2}^{(2,1 / 2)}(\xi) & \approx \frac{1}{4 \pi}\left[\int_{0}^{\theta} \exp \left(-\frac{\xi^{2} \sqrt{\tau}}{4}\right) \sqrt{\tau} g_{+}\left(\tau ; \frac{1}{2}\right) \mathrm{d} \tau\right. \\
& \left.+\frac{1}{2 \sqrt{\pi}} \int_{\theta}^{\infty} \exp \left(-\frac{\xi^{2} \sqrt{\tau}}{4}\right) \tau^{-1} \mathrm{~d} \tau\right]
\end{aligned}
$$

For small $\xi$, the second term dominates in this sum, which gives rise to a logarithmic singularity,

$$
\Psi_{2}^{(2,1 / 2)}(\xi) \sim \frac{1}{4 \pi^{3 / 2}} E_{1}\left(\xi^{2} \theta^{1 / 8}\right) \sim \frac{1}{2 \pi^{3 / 2}}|\ln \xi|, \quad \xi \rightarrow 0
$$

The process under study also differs from the normal diffusion in that the transverse coordinates $\mathrm{X}, Y$ of a particle are no longer independent, because their combined distribution

$$
P(X(t) \in \mathrm{d} x, Y(t) \in \mathrm{d} y)=\frac{1}{D_{\perp} \sqrt{t}} \Psi_{2}^{(2,1 / 2)}\left(\frac{\sqrt{x^{2}+y^{2}}}{\sqrt{D_{\perp} \sqrt{t}}}\right)
$$

does not reduce to the product $P(X(t) \in \mathrm{d} x) P(Y(t) \in \mathrm{d} y)$, although the coordinates remain uncorrelated because of the axial symmetry of the distribution.

### 4.4 Longitudinal finite-velocity random walk

We now discuss the random walk of particles along magnetic field lines. Because we neglect the time that a particle needs to pass between lines, the longitudinal motion of particles is apparently continuous and reverses the direction at random instants of time. The one-dimensional symmetric random walk of a particle with a constant velocity $v$ along the $z$ axis in the absence of traps is completely characterized by the distribution density of its free paths. In a medium consisting of independent atoms, free paths are distributed exponentially, and the process is described by a second-order partial differential equation (the telegraph equation) [117]. The solution of this equation is expressed in terms of the modified Bessel functions and transforms into the normal (Gaussian) distribution in the asymptotic regime of large times. In the case of an arbitrary density $p(z)$ with a finite second moment, the asymptotic part of the solution satisfies the telegraph
equation. The authors of [85-87, 118-120] considered onedimensional random walks with the asymptotically powerlaw range distribution $p(\xi) \propto \xi^{-\alpha-1}, 0<\alpha<2$, which are sometimes called fractal walks.

We now consider Eqn (28), which takes the form

$$
\tilde{G}_{\|}(k, \lambda)=\frac{\tilde{P}(k, \lambda)}{v[1-\tilde{p}(k, \lambda)]}
$$

in the case of one-dimensional walks. Here,

$$
\begin{aligned}
& \tilde{p}(k, \lambda)=\int_{0}^{\infty} \exp \left(-\frac{\lambda}{v} z\right) \cos (k z) p(z) \mathrm{d} z \\
& =\frac{1}{2}\left[\tilde{p}\left(\frac{\lambda}{v}-\mathrm{i} k\right)+\tilde{p}\left(\frac{\lambda}{v}+\mathrm{i} k\right)\right] \\
& \tilde{P}(k, \lambda)=\int_{0}^{\infty}\left[\exp \left(-\frac{\lambda}{v} z\right) \cos (k z) \int_{z}^{\infty} p\left(z^{\prime}\right) \mathrm{d} z^{\prime}\right] \mathrm{d} z \\
& = \\
& =\frac{2 \lambda / v-(\lambda / v-\mathrm{i} k) \tilde{p}(\lambda / v+\mathrm{i} k)-(\lambda / v+\mathrm{i} k) \tilde{p}(\lambda / v-\mathrm{i} k)}{2\left[(\lambda / v)^{2}+k^{2}\right]}
\end{aligned}
$$

where $\tilde{p}$ is the Laplace transform of the range distribution density:

$$
\tilde{p}(\lambda)=\int_{0}^{\infty} \exp (-\lambda \xi) p(\xi) \mathrm{d} \xi
$$

As a result, we have
$\tilde{G}_{\| \mid}(k, \lambda)$
$=\frac{2(\lambda / v)-(\lambda / v+\mathrm{i} k) \tilde{p}(\lambda / v-\mathrm{i} k)-(\lambda / v-\mathrm{i} k) \tilde{p}(\lambda / v+\mathrm{i} k)}{v\left[k^{2}+(\lambda / v)^{2}\right][2-\tilde{p}(\lambda / v-\mathrm{i} k)-\tilde{p}(\lambda / v+\mathrm{i} k)]}$.

We consider the three cases
$\tilde{p}(\lambda)=1-c \lambda^{\alpha}, \quad c=\frac{A}{\alpha} \Gamma(1-\alpha), \quad 0<\alpha<1$,
$\tilde{p}(\lambda)=1-m_{1} \lambda+c_{1} \lambda^{\alpha}, \quad c_{1}=\frac{(A / \alpha) \Gamma(2-\alpha)}{\alpha-1}, \quad 1<\alpha<2$,
$\tilde{p}(\lambda)=1-m_{1} \lambda+\frac{m_{2}}{2} \lambda^{2}-c_{2} \lambda^{\alpha}, \quad c_{2}=\frac{(A / \alpha) \Gamma(3-\alpha)}{(\alpha-1)(\alpha-2)}, \quad \alpha>2$,
where $m_{1}=\langle R\rangle$ and $m_{2}=\left\langle R^{2}\right\rangle$ are the moments of the range distribution.

In the first case (process of the first kind, $0<\alpha<1$ ), expression (57) takes the form

$$
\begin{equation*}
\tilde{G}_{\|}(k, \lambda)=\frac{(\lambda / v-\mathrm{i} k)^{\alpha-1}+(\lambda / v+\mathrm{i} k)^{\alpha-1}}{v\left[(\lambda / v-\mathrm{i} k)^{\alpha}+(\lambda / v+\mathrm{i} k)^{\alpha}\right]} \tag{58}
\end{equation*}
$$

For $\alpha=1 / 2$, the transformation

$$
\tilde{G}_{\|}(k, \lambda)=\frac{1}{v \sqrt{(\lambda / v)^{2}+k^{2}}}
$$

can easily be inverted to yield the symmetrized arcsine density [42]:

$$
\begin{equation*}
G_{\|}\left(z, t ; \frac{1}{2}\right)=\frac{1}{\pi \sqrt{(v t)^{2}-z^{2}}}, \quad-v t<z<v t . \tag{59}
\end{equation*}
$$

The explicit expression for the longitudinal propagator in terms of elementary functions for all $\alpha$ in the specified range is


Figure 25. Longitudinal RAD propagators of the first kind. (a) Calculation by expression (60) (curves) and Monte Carlo simulations (symbols). (b) The evolution of the RAD propagator with $\alpha=1 / 2$ compared with the evolution of the normal propagator $(\alpha=2)$. The dashed lines show the time dependence of the width of diffusion packets, and the solid lines show ballistic restrictions.
presented in our paper [120]:

$$
\begin{aligned}
& G_{\| \|}(z, t)=\frac{2 \sin \pi \alpha}{\pi v t} \\
& \quad \times \frac{\left(1-z^{2} / v^{2} t^{2}\right)^{\alpha-1}}{(1-z / v t)^{2 \alpha}+(1+z / v t)^{2 \alpha}+2\left(1-z^{2} / v^{2} t^{2}\right)^{\alpha} \cos \pi \alpha},
\end{aligned}
$$

$$
\begin{equation*}
\alpha \in(0,1) . \tag{60}
\end{equation*}
$$

Distributions of this type for several values of $\alpha$ are represented in Fig. 25.

For $1<\alpha<2$ (process of the second kind), the transform

$$
\tilde{G}_{\|}(k, \lambda)=\frac{2 m_{1}-c_{1}(\lambda / v-\mathrm{i} k)^{\alpha-1}-c_{1}(\lambda / v+\mathrm{i} k)^{\alpha-1}}{v\left[2 m_{1}(\lambda / v)-c_{1}(\lambda / v-\mathrm{i} k)^{\alpha}-c_{1}(\lambda / v+\mathrm{i} k)^{\alpha}\right]}
$$

in the asymptotic regime $\lambda \rightarrow 0, k \rightarrow 0,|\lambda / v k| \rightarrow 0$ takes the form

$$
\tilde{p}(k, \lambda)=\frac{1}{\lambda+D_{\|}|k|^{\alpha}},
$$

where

$$
D_{\|}=\frac{c_{1}}{m_{1}} v\left|\cos \frac{\alpha \pi}{2}\right|=v \frac{\Gamma(2-\alpha)}{\varepsilon^{1-\alpha}}\left|\cos \frac{\alpha \pi}{2}\right| .
$$

The inverse Laplace transformation leads to the characteristic function of a random coordinate of the walking particle,

$$
\tilde{G}_{\|}(k, t)=\exp \left(-D_{\|} t|k|^{\alpha}\right),
$$

and returning to the spatial variable gives the propagator

$$
G_{\|}(z, t)=\left(D_{\|} t\right)^{-1 / \alpha} \Psi_{1}^{(\alpha)}\left(z\left(D_{\|} t\right)^{-1 / \alpha}\right) .
$$

Finally, for $\alpha>2$, we obtain

$$
\begin{aligned}
\tilde{G}_{\|}(k, \lambda) & =\left[2 m_{1}-m_{2}\left(\frac{\lambda}{v}\right)+c_{2}\left(\frac{\lambda}{v}-\mathrm{i} k\right)^{\alpha-1}+c_{2}\left(\frac{\lambda}{v}+\mathrm{i} k\right)^{\alpha-1}\right] \\
& \times \frac{1}{v}\left\{2 m_{1}\left(\frac{\lambda}{v}\right)-m_{2}\left[\left(\frac{\lambda}{v}\right)^{2}-k^{2}\right]+c_{2}\left(\frac{\lambda}{v}-\mathrm{i} k\right)^{\alpha}\right. \\
& \left.+c_{2}\left(\frac{\lambda}{v}+\mathrm{i} k\right)^{\alpha}\right\}^{-1} \sim \frac{1}{\lambda+D k^{2}}
\end{aligned}
$$

where $D=m_{2} v / 2 m_{1}$. This gives

$$
G_{\|}(z, t)=\left(D_{\|} t\right)^{-1 / 2} \Psi_{1}^{(2)}\left(\frac{z}{\sqrt{D_{\|} t}}\right),
$$

where $\Psi_{1}^{(2)}(\xi)$ is the standard density of the normal distribution with dispersion 2.

Direct Monte Carlo simulation confirms the results of analytic calculations (see Fig. 25). The qualitative difference of distributions for $\alpha>1$ and $\alpha<1$ is explained by the competition between two processes: the diffusion process, expanding as $\propto t^{1 / \alpha}$ in the absence of restrictions, and the ballistic regime, restricting the position of a particle to the interval $[-v t, v t]$. For $\alpha>1$ and large times, the first process dominates (the rapidly extending interval [ $-v t, v t]$ no longer affects diffusion); for $\alpha<1$, the role of the kinematic constraint increases. Long ranges (with the infinite mean value) press the distribution against ballistic boundaries, producing characteristic peaks at its ends; for $\alpha<1 / 2$, specific $U$-shaped distributions are formed.

It may seem that we here avoid fractional derivatives, but the term $(\lambda \pm \mathrm{i} v k)^{\alpha}$ contained in the expressions presented above is the Fourier-Laplace transform of the material derivative operator to a fractional power,

$$
\begin{aligned}
& (\lambda \mp \mathrm{i} v k)^{\alpha} \tilde{f}(k, \lambda) \\
& \quad=\int_{0}^{\infty} \mathrm{d} t \int_{-\infty}^{\infty} \mathrm{d} z \exp (-\lambda t+\mathrm{i} k z)\left(\frac{\partial}{\partial t} \pm v \frac{\partial}{\partial z}\right)^{\alpha} f(z, t),
\end{aligned}
$$

as can be verified by direct calculation. For this, it suffices to represent this operator in the Riemann-Liouville form

$$
\begin{aligned}
& \left(\frac{\partial}{\partial t} \pm v \frac{\partial}{\partial z}\right)^{\alpha} f(z, t)=\frac{1}{\Gamma(1-\alpha)}\left(\frac{\partial}{\partial t} \pm v \frac{\partial}{\partial z}\right) \\
& \quad \times \int_{0}^{t} f(z-v(t-\tau), \tau)(t-\tau)^{-\alpha} \mathrm{d} \tau, \quad 0<\alpha<1
\end{aligned}
$$

and to apply the above transformation to it. Therefore, for example, Eqn (58) is equivalent to the equation with
fractional differential operators in the form

$$
\begin{aligned}
& {\left[\left(\frac{\partial}{\partial t}+v \frac{\partial}{\partial z}\right)^{\alpha}+\left(\frac{\partial}{\partial t}-v \frac{\partial}{\partial z}\right)^{\alpha}\right] G_{\|}(z, t)} \\
& \quad=v^{\alpha-1}\left[\left(\frac{\partial}{\partial t}+v \frac{\partial}{\partial z}\right)^{\alpha-1}+\left(\frac{\partial}{\partial t}-v \frac{\partial}{\partial z}\right)^{\alpha-1}\right] \delta(z) \delta(t) .
\end{aligned}
$$

Other equations, whose integral transforms are $(\lambda \pm \mathrm{i} v k)^{\alpha}$, can be represented similarly.

## 5. Diffusion acceleration models

### 5.1 Classical Fermi model

In this section, we consider the problem of acceleration (more exactly, additional or distributed acceleration) of cosmic rays, and we therefore study the behavior of particles in the momentum space. Successive interactions (collisions) of a charged particle with more or less localized inhomogeneities of the magnetic field-from magnetic clouds moving at comparatively small velocities, which were discussed by Fermi (see the Introduction), to strong shock waves in the remnants of supernovae, which were considered in review [121]-can be treated as instant flights from one point in the momentum space to another. Momentum increments $\Delta \mathbf{p}_{i}$ imparted to a particle in such collisions are random, and even in the case of their isotropic distribution, the point

$$
\mathbf{p}=\mathbf{p}_{0}+\Delta \mathbf{p}_{1}+\Delta \mathbf{p}_{2}+\Delta \mathbf{p}_{3}+\ldots,
$$

representing a particle in the momentum space, like a Brownian particle, moves away from the point (momentum) $\mathbf{p}_{0}$ of acceleration injection, which means the further acceleration (additional acceleration) of the particle. True, only a fraction of the particles moving away from the center are accelerated. It is this fluctuation component of the cosmic-ray acceleration mechanism that is of interest to us in this review.

From the statistical standpoint, the main result of the Fermi model is that for a power-law energy spectrum $N(E)$ to be formed, it is sufficient to have the exponential increase in the energy $E=E_{0} \exp (a t)$ of the particle being accelerated and the exponential distribution of the age $\mathrm{d} P=$ $\exp (-t / \tau) \mathrm{d} t / \tau$ of the detected particles:

$$
\begin{align*}
N(E) \mathrm{d} E & =\tau^{-1}\left[\int_{0}^{\infty} \delta\left(E-E_{0} \exp (a t)\right) \exp \left(-\frac{t}{\tau}\right) \mathrm{d} t\right] \mathrm{d} E \\
& =\frac{1}{a \tau}\left(\frac{E}{E_{0}}\right)^{-1-1 / a \tau} \frac{\mathrm{~d} E}{E_{0}}, \quad E>E_{0} \tag{61}
\end{align*}
$$

That is all, and no fluctuations are needed (except the abovementioned age fluctuations determined by the mean value $\tau$ of this age). In general, what fluctuations can be discussed here if, according to Fermi's estimates, the increment in the energy by a factor of $e$ can be obtained only after 100 million collisions?

To study the influence of other fluctuation sources on the energy spectrum of accelerated particles, it is necessary to consider processes with more diverse possibilities for acceleration in each of the collisions and the lower collision frequency. These processes include the above-mentioned interactions with strong shock waves in which the energy of a particle can be increased by a factor of 7 to 13 even in a
single collision with the shock wave front [6, p. 449]. For this, we pass from the degenerate spectral function $\delta\left(E-E_{0} \exp (a t)\right)$, characterizing the deterministic Fermi acceleration process, to the continuous function $N(E, t)$ related to the momentum distribution $f(\mathbf{p}, t)$ as

$$
N(E, t)=\int \delta(E-E(\mathbf{p})) f(\mathbf{p}, t) \mathrm{d} \mathbf{p}
$$

Assuming, as Fermi did, that the parameter $\tau$ is independent of energy, we represent the required spectrum as the spectrum $N(E, t)$ of observed particles averaged over the exponentially distributed age,

$$
\begin{align*}
N(E) & \equiv \bar{N}(E ; \tau) \equiv \tau^{-1}\left[\int_{0}^{\infty} N(E, t) \exp \left(-\frac{t}{\tau}\right) \mathrm{d} t\right] \\
& =\tau^{-1} \hat{N}\left(E, \tau^{-1}\right) \tag{62}
\end{align*}
$$

where

$$
\hat{N}(E, \lambda) \equiv \int_{0}^{\infty} \exp (-\lambda t) N(E, t) \mathrm{d} t
$$

is the Laplace transform of the spectral function over the time variable. These expressions take only the influence of age fluctuations on the energy spectrum into account. The influence of fluctuations caused by rare interactions followed by large changes in the particle momentum (energy) is taken into account at the stage of constructing the equations for $f(\mathbf{p}, t)$ or $N(E, t)$ by including additional terms containing differential and integral operators.

### 5.2 Fractional differential kinetic equation

The position of a point at rest in the momentum space, not coinciding with the coordinate origin, means that a particle moves with a constant momentum (velocity, energy). The exponent $\beta$ then characterizes the tail of the distribution of the random interval duration $T$ between successive collisions of the moving particle: $Q(t)=P(T>t), t \rightarrow \infty$. For an ultrarelativistic particle $(v \approx c)$, this interval is proportional to its free path between successive collisions, and therefore, along with the replacement of coordinates $\mathbf{r}$ by momenta $\mathbf{p}$, the exponent $\beta$ in the CTRW model is replaced by the exponent $\alpha$. The particle is assumed injected by a source at the instant $t=0$ with a distribution $f_{0}(\mathbf{p})$. Under these conditions, the process is described by an integral equation similar to (13):

$$
\begin{align*}
f(\mathbf{p}, t) & =\int \mathrm{d} \mathbf{p}^{\prime} \int_{0}^{t} \mathrm{~d} t^{\prime} w\left(\mathbf{p}-\mathbf{p}^{\prime} \rightarrow \mathbf{p}\right) q\left(t^{\prime}\right) \\
& \times f\left(\mathbf{p}-\mathbf{p}^{\prime}, t-t^{\prime}\right)+Q(t) f_{0}(\mathbf{p}) \tag{63}
\end{align*}
$$

The equation is next transformed by specifying the distributions of $q$ and $w$. We begin with the time distribution. Usually (virtually always), the exponential distribution $Q(t)=$ $\exp (-\mu t), q(t) \equiv-\mathrm{d} Q / \mathrm{d} t=\mu \exp (-\mu t)$ is selected, and this is typically done implicitly: the transfer equation is written beginning from the time derivative $\partial f(\mathbf{p}, t) / \partial t=\ldots$ assuming the process to be Markovian, which automatically leads to the exponential distribution. We have discussed the indications that this distribution is more likely to be power-like rather than exponential (we recall that here $q(t)$ is the range distribution up to a scale factor, which has a pronounced power-law form in a fractal medium, which is quite consistent


Figure 26. 'Fractional exponential density' $\psi_{\alpha}(t)$. Ten curves (from bottom up) correspond to $\alpha$ ranging from 0.1 to 1.0 with the step 0.1 .
with the self-similar picture of turbulent motions and its power-law regularities). The attractiveness of this idea becomes somewhat darkened by having to abandon the classical exponential, which at first glance produces an insurmountable barrier between these two distributions. The best compromise would be a family of distributions including both exponential and power-law ones. Fortunately, such a family exists. This is a set of functions

$$
Q_{\alpha}(t)=E_{\alpha}\left(-\mu t^{\alpha}\right), \quad \alpha \in(0,1], \quad \mu>0,
$$

where $E_{\alpha}(z)=\sum_{n=0}^{\infty} z^{n} / \Gamma(\alpha n+1)$ are Mittag-Leffler functions. For $\alpha=1$, the function $Q_{\alpha}(t)$ becomes the usual exponential, while for $\alpha<1$, this 'fractional exponential' has the asymptotic form $t^{-\alpha}, t \rightarrow \infty$. The corresponding
density (Fig. 26),

$$
\psi_{\alpha}(t)=\mu^{\alpha-1} E_{\alpha, \alpha}\left(-\mu t^{\alpha}\right), \quad E_{\alpha, \alpha}=\sum_{n=0}^{\infty} \frac{z^{n}}{\Gamma(\alpha n+\alpha)},
$$

satisfies a fractional differential equation [33], which brings integral equation (63) to the fractional differential form

$$
\begin{equation*}
\frac{\partial^{\alpha} f(\mathbf{p}, t)}{\partial t^{\alpha}}=\mu A f(\mathbf{p}, t)+f_{0}(\mathbf{p}) \delta_{\alpha}(t) \tag{64}
\end{equation*}
$$

Here, $A$ is the integral acceleration operator with the transition density $w\left(\mathbf{p}^{\prime} \rightarrow \mathbf{p}\right)$ :

$$
\begin{equation*}
A f(\mathbf{p}, t)=\int w\left(\mathbf{p}^{\prime} \rightarrow \mathbf{p}\right) f\left(\mathbf{p}^{\prime}, t\right) \mathrm{d} \mathbf{p}^{\prime}-f(\mathbf{p}, t) . \tag{65}
\end{equation*}
$$

The time sequence of collision instants forms the fractional Poisson process of the order $\alpha$ [122], which for $\alpha=1$ becomes the usual Poisson process underlying the classical kinetic equation (see Eqn (21.1) [123]). A new property of this process is that the mean number of collisions increases upon increasing the observation interval proportionally to $t^{\alpha}$, i.e., more slowly than in the case of the usual Poisson process $(\sim t)$, while the relative fluctuations of the number of collisions do not disappear as $t \rightarrow \infty$ but tend to a limit distribution depending on $\alpha$. Event flows obtained in Monte Carlo simulations are shown in Fig. 27. We can see that a change in the scale transforms the Poisson flow into the homogeneous flow, whereas the fractional Poisson flow remains inhomogeneous at all scales, which suggests that the point distribution is fractal.

As previously, expressing the solution $f_{\alpha}(\mathbf{p}, t)$ of fractional differential equation (64) in terms of the solution


Figure 27. Acceleration event fluxes in (a) Poisson and (b) fractional Poisson processes.
$f_{1}(\mathbf{p}, t)$ of the equation with the first-order derivative,

$$
\begin{equation*}
f_{\alpha}(\mathbf{p}, t)=\frac{t}{\alpha} \int_{0}^{\infty} f_{1}(\mathbf{p}, \tau) g_{+}\left(t \tau^{-1 / \alpha} ; \alpha\right) \tau^{-1 / \alpha-1} \mathrm{~d} \tau \tag{66}
\end{equation*}
$$

performing the time Laplace transformation of (66), and using the above expressions for the spectra, we obtain the relation

$$
\begin{equation*}
\bar{N}_{\alpha}(E ; \tau)=\bar{N}_{1}\left(E ; \tau^{\alpha}\right) . \tag{67}
\end{equation*}
$$

Expression (67) reflects the influence of the fractal dimension $\alpha \in(0,1]$ of the fractional Poisson collision process on the energy spectrum of cosmic rays: the spectrum $\bar{N}_{\alpha}(E ; \tau)$ formed by the ensemble of particles with a mean lifetime $\tau$ accelerated by the fractional Poisson law of an order $\alpha<1$ coincides with the spectrum of particles accelerated by the usual Poisson process $(\alpha=1)$, but with the mean lifetime $\tau^{\alpha}$ (we recall that the dimensionless lifetime is used here, and the injection spectra $f_{0}(E)$ in both problems coincide). It is easy to verify with the example of the Fermi spectrum that as the order of the process decreases, the acceleration efficiency decreases (the steepness of the spectrum increases): the fractal character of the distribution of accelerating regions in space reduces the acceleration efficiency (of course, for the same law of elementary accelerations).

### 5.3 Fractional differential Fokker-Planck equations

As in the classical case, passing from the kinetic equation to the Fokker-Planck equation involves the transformation of the collision integral to the differential form by expanding the integrand in a power series in the momentum increment through second-order terms. Two versions of this expansion exist, leading to somewhat different equations (see, e.g., [123]). The first assumes that the change in the absolute value of the momentum $|\Delta \mathbf{p}|=\left|\mathbf{p}-\mathbf{p}^{\prime}\right|$ is small, and therefore the momentum of the incident particle changes very weakly in magnitude and direction in one collision event (for example, as in the collision of a heavy particle with a light particle). The second version assumes that the change in the absolute value of the momentum $\Delta p=|\mathbf{p}|-\left|\mathbf{p}^{\prime}\right|$ is small, but the change in its direction can have an arbitrarily broad distribution, up to an isotropic one (as in the case of the collision of a light particle with a heavy particle). Assuming isotropic scattering, we obtain the fractional differential generalization of the Fokker-Planck equation in the form

$$
\begin{equation*}
\frac{\partial^{\alpha} f(\mathbf{p}, t)}{\partial t^{\alpha}}=\Delta_{\mathbf{p}}(K(p) f(\mathbf{p}, t))+f_{0}(\mathbf{p}) \delta_{\alpha}(t), \tag{68}
\end{equation*}
$$

where

$$
K(p)=\frac{\mu}{2} \int(\Delta \mathbf{p})^{2} w(\mathbf{p} \rightarrow \mathbf{p}+\Delta \mathbf{p}) \mathrm{d} \Delta \mathbf{p}
$$

is the diffusion coefficient in the momentum space. The energy analog of Eqn (68) (of course, for $\alpha=1$ ) in cosmicray physics has the form

$$
\begin{align*}
\frac{\partial^{\alpha} N(E, t)}{\partial t^{\alpha}}= & \frac{\partial\left[a_{1}(E) N(E, t)\right]}{\partial E} \\
& +\frac{\partial^{2}\left[a_{2}(E) N(E, t)\right]}{\partial E^{2}}+N_{0}(E) \delta_{\alpha}(t) \tag{69}
\end{align*}
$$

(see Eqn (14.2) in [4]). At the same time, Eqn (68), with the diffusion term

$$
\begin{aligned}
\Delta_{\mathbf{p}}(K(p) f(\mathbf{p}, t)) & =\left(\Delta_{\mathbf{p}} K(p)\right) f(\mathbf{p}, t)+2\left(\nabla_{\mathbf{p}} K(p)\right) \nabla_{\mathbf{p}} f(\mathbf{p}, t) \\
& +K(p) \Delta_{\mathbf{p}} f(\mathbf{p}, t)
\end{aligned}
$$

considerably differs from another diffusion equation (see Eqn (9.57) in [44]),

$$
\begin{equation*}
\frac{\partial^{\alpha} f(\mathbf{p}, t)}{\partial t^{\alpha}}=\nabla_{\mathbf{p}}\left(K(p) \nabla_{\mathbf{p}} f(\mathbf{p}, t)\right)+f_{0}(\mathbf{p}) \delta_{\alpha}(t) \tag{70}
\end{equation*}
$$

The reason for this difference is that Eqn (68) is derived in the collision model in which the point representing a particle instantly moves to another, possibly remote, geometrical point, violating the continuity of the trajectory in the momentum space, whereas the dynamic derivation of Eqn (70) assumes that the trajectory in the momentum space is continuous and even differentiable.

The classical versions ( $\alpha=1$ ) of Eqns (68)-(70) underlie the standard mathematical tools for describing fluctuation mechanisms of cosmic-ray acceleration, and their solutions are well known $[4,6,44]$. The diffusion model in the velocity space was used in $[124,125]$ to describe the interaction of particles with the waves of turbulent pulsations in the shortwavelength part of the spectrum that accelerate the particles. This acceleration was shown to be a universal property of turbulence in plasmas, caused by the most fundamental features of the corresponding processes, namely:
(i) The resonance character of scattering of particles by waves;
(ii) the continuity of the turbulent energy flux in the wave number space;
(iii) the increase in $K(v)$ proportionally to $v^{v-1}$, where $v$ is the exponent of the spectrum of turbulent pulsations.

Referring to the Parker theory [126], the solution of the isotropic diffusion equation in the velocity space,

$$
\begin{equation*}
\frac{\partial f(\mathbf{v}, t)}{\partial t}=\frac{K_{1}}{v^{2}} \frac{\partial}{\partial v}\left(v^{v+1} \frac{\partial f(\mathbf{v}, t)}{\partial v}\right) \tag{71}
\end{equation*}
$$

with the initial condition $f(\mathbf{v}, 0) \propto \delta(v)$ was written in [124, 125] in the form

$$
\begin{equation*}
f(\mathbf{v}, t)=\text { const } t^{-3 /(3-v)} \exp \left[-\frac{v^{3-v}}{(3-v)^{2} K_{1} t}\right] \tag{72}
\end{equation*}
$$

which for $v=1$ reproduces the Maxwell spectrum and for $v=2$, the spectrum of protons from solar flares of the type $\exp \left(-v / v_{0}\right)$, while the values $v>2$ are of no interest because the acceleration is then determined not by the cyclotron resonance but by the Fermi mechanism.

The fractional differential version of this problem, in the form

$$
\begin{equation*}
\frac{\partial^{\gamma} \psi}{\partial t^{\gamma}}=\frac{K_{1}}{v^{2}} \frac{\partial}{\partial v}\left(v^{\gamma} \frac{\partial \psi(\mathbf{v}, t)}{\partial v}\right) \tag{73}
\end{equation*}
$$

with $\gamma \in[0,2]$ was considered in [127], where the problem setting was motivated, in particular, by a reference to the nonGaussian thermodynamic formalism developed in [128] for the description of strongly turbulent media exhibiting pronounced non-Markov behavior. The ideas and conclusions in this work are undoubtedly of interest, but two remarks should be made.

First, the three-dimensional isotropic density of the velocity distribution $\psi(\mathbf{v}, t)$ is normalized to the concentra-
tion $n=n(t)$ :

$$
\begin{equation*}
\int \psi(\mathbf{v}, t) \mathrm{d} \mathbf{v}=n . \tag{74}
\end{equation*}
$$

Multiplying both sides of Eqn (73) by $v^{4-\gamma}$ and integrating over the three-dimensional velocity space, the authors of [127] obtain

$$
\begin{equation*}
\frac{\partial^{\gamma}}{\partial t^{\gamma}} \int v^{4-\gamma} \psi \mathrm{d} \mathbf{v}=3(4-\gamma) n K \tag{75}
\end{equation*}
$$

and interpret the integral as the mean of the quantity $v^{4-\gamma}$ over the ensemble. But this is not the case: the mean over the ensemble is defined by the expression

$$
\left\langle v^{4-\gamma}\right\rangle=\frac{1}{n} \int v^{4-\gamma} \psi \mathrm{d} \mathbf{v}
$$

and Eqn (75) reduces to the form

$$
\frac{\partial^{\gamma}}{\partial t^{\gamma}}\left\langle v^{4-\gamma}\right\rangle=3(4-\gamma) K
$$

only if integral (74) remains constant. For a homogeneous equation, this condition is satisfied only for $\gamma=1$. For this condition to be also satisfied for $\gamma<1$, it is necessary to introduce the term with the fractional delta function in time,

$$
\frac{\partial^{\gamma} \psi}{\partial t^{\gamma}}=\frac{K_{1}}{v^{2}} \frac{\partial}{\partial v}\left(v^{\gamma} \frac{\partial \psi(\mathbf{v}, t)}{\partial v}\right)+\psi(\mathbf{v}, 0) \delta_{\gamma}(t), \quad 0<\gamma<1,
$$

into the right-hand side of Eqn (73). We note that in passing to the classical equation $(\beta \rightarrow 1)$, the fractional delta function transforms into the usual delta function, equal to zero everywhere except the only point $t=0$, which can be easily removed by restricting the consideration to the semiaxis $t>0$ and specifying the initial condition by the limit $\lim _{t \rightarrow 0+} \psi(v, t)$. Equation (71) then follows. Conversely, the fractional delta function

$$
\delta_{\gamma}(t)=\frac{t_{+}^{-\gamma}}{\Gamma(1-\gamma)},
$$

defined as the fractional derivative of the unit Heaviside function, is nonzero for all $t>0$; it describes particles delayed in their flight from the 'point of departure' and cannot be discarded in a similar manner.

The second remark is related to the applicability (more exactly, inapplicability) of the probabilistic interpretation of the equation with $\gamma>1$ in the absence of the first time derivative ${ }^{13}$ (see the discussion of this case in Section 3.2).

Returning to the problem considered in [124, 125], we formulate its fractional differential generalization in the form

$$
\frac{\partial^{\alpha} f_{\alpha}(\mathbf{v}, t)}{\partial t^{\alpha}}=\frac{K_{1}}{v^{2}} \frac{\partial}{\partial v}\left(v^{v+1} \frac{\partial f_{\alpha}(\mathbf{v}, t)}{\partial v}\right)+\delta(\mathbf{v}) \delta_{\alpha}(t) .
$$

The solution of this equation is expressed in terms of the solution of Eqn (71) as

$$
f_{\alpha}(\mathbf{v}, t)=\int_{0}^{\infty} f\left(\mathbf{v},\left(\frac{t}{\tau}\right)^{\alpha}\right) g_{+}(\tau ; \alpha) \mathrm{d} \tau, \quad 0<\alpha \leqslant 1
$$

Substituting $f(\mathbf{v}, t)$ from (72) in the last equation with $v=2$ and $\alpha=1 / 2$, we obtain the solution of the problem in the

[^10]integral form
\[

$$
\begin{aligned}
f_{1 / 2}(\mathbf{v}, t) & =\int_{0}^{\infty} f\left(\mathbf{v}, \sqrt{\frac{t}{\tau}}\right) g_{+}\left(\tau ; \frac{1}{2}\right) \mathrm{d} \tau \\
& =\frac{\text { const }}{2 \sqrt{\pi}} t^{-3 / 2} \int_{0}^{\infty} \exp \left(-\frac{v}{K_{1}} \sqrt{\frac{\tau}{t}}-\frac{1}{4 \tau}\right) \mathrm{d} \tau
\end{aligned}
$$
\]

The authors of [129] used a different fractional version of the Fokker-Planck equation (with the first time derivative and a fractional Laplacian) to describe the motion of a charge $e$ with mass $m$ in the magnetic field $\mathbf{H}=$ const under the action of the friction force $-\eta m \mathbf{v}$ and a random electric field $\mathbf{E}$. The latter was represented by homogeneous stationary white Lévy noise with the intensity $\mu$ and exponent $\alpha$. Lévy noise is a sequence of independent stationary increments of Lévy motion, just as white Gaussian noise is a sequence of increments of Brownian motion. The characteristic increment function is

$$
\tilde{p}(\mathbf{k}, \Delta t)=\exp \left(-\mu|\mathbf{k}|^{\alpha} \Delta t\right), \quad 0<\alpha \leqslant 2 .
$$

As $\alpha \rightarrow 2$, the Lévy noise transforms into Gaussian noise. The equation for the velocity distribution density $f(\mathbf{v}, t)$ in the magnetic field $\mathbf{H}=H \mathbf{e}_{z}$ corresponding to the given model has the form

$$
\frac{\partial f}{\partial t}+\omega_{H}\left[\mathbf{v e}_{z}\right] \nabla_{\mathbf{v}} f=\eta \nabla_{\mathbf{v}}(\mathbf{v} f)-\mu\left(-\Delta_{\mathbf{v}}\right)^{\alpha / 2} f, \quad \omega_{H}=\frac{e H}{m c}
$$

By the Fourier transformation

$$
\tilde{f}(\mathbf{k}, t)=\int \mathrm{d} \mathbf{v} \exp (\mathrm{i} \mathbf{k} \mathbf{v}) f(\mathbf{v}, t)
$$

it is reduced to the form

$$
\frac{\partial \tilde{f}}{\partial t}+\left(\omega_{H}\left[\mathbf{k} \mathbf{e}_{z}\right]+\eta \mathbf{k}\right) \nabla_{\mathbf{k}} \tilde{f}=-\mu|\mathbf{k}|^{\alpha} \tilde{f}
$$

The solution of this equation found in [129] for the zero initial condition $\mathbf{v}(0)=0$ has the form

$$
\tilde{f}(\mathbf{k}, t)=\exp \left\{-\frac{\mu}{\alpha \eta}[1-\exp (-\alpha \eta t)]|\mathbf{k}|^{\alpha}\right\}
$$

and hence the velocity distribution is expressed in terms of the isotropic stable density as

$$
\begin{aligned}
f(\mathbf{v}, t) & =\left\{\frac{\mu}{\alpha \eta}[1-\exp (-\alpha \eta t)]\right\}^{-3 / \alpha} \\
& \times \Psi_{3}^{(\alpha)}\left(\left\{\frac{\mu}{\alpha \eta}[1-\exp (-\alpha \eta t)]\right\}^{-1 / \alpha} \mathbf{v}\right) .
\end{aligned}
$$

For small times, we have

$$
f(\mathbf{v}, t) \sim(\mu t)^{-3 / \alpha} \Psi_{3}^{(\alpha)}\left((\mu t)^{-1 / \alpha} \mathbf{v}\right)
$$

and we are dealing with Lévy motion (the scaling factor is linear in time). In the large-time limit, the time dependence disappears, and we obtain the stationary velocity distribution

$$
f(\mathbf{v}, \infty)=\left(\frac{\mu}{\alpha \eta}\right)^{-3 / \alpha} \Psi_{3}^{(\alpha)}\left(\left(\frac{\mu}{\alpha \eta}\right)^{-1 / \alpha} \mathbf{v}\right)
$$

This solution coincides with the equilibrium Maxwell distribution only for $\alpha=2$, when $\Psi_{3}^{(\alpha)}$ is a three-dimensional Gaussian (with doubled dispersion). For other values of $\alpha$, it significantly differs from the equilibrium distribution by a redistribution of probability from the intermediate velocity
distribution to the regions of small and large velocities, forming asymptotic power-law tails:

$$
f(\mathbf{v}, \infty) \propto|\mathbf{v}|^{-\alpha-3}
$$

In the problem of energy losses by a fast particle in the absence of acceleration, its energy distribution is bounded by the initial energy, and all the moments of this distribution are finite. In the presence of acceleration, such a strict upper bound of the energy spectrum is absent, which gives additional grounds to examine the region with infinite dispersion, which is acquiring increasing popularity among researchers studying anomalous diffusion processes. Interesting results (in the asymptotic sense) are here obtained only when the infinite dispersion is caused by the power-law behavior of the distributions:

$$
\int_{|\Delta \mathbf{p}|>p} w\left(\mathbf{p}^{\prime} \rightarrow \mathbf{p}^{\prime}+\Delta \mathbf{p}\right) \mathrm{d} \Delta \mathbf{p} \propto p^{-\gamma}, \quad p \rightarrow \infty
$$

For the exponent $\gamma>2$, the second moment is finite, and we are in the 'classical' diffusion region. For $\gamma<2$, the second moment of the increment is infinite and we arrive at equations for the momentum distribution $f(\mathbf{p}, t)$,

$$
\begin{equation*}
\frac{\partial^{\alpha} f(\mathbf{p}, t)}{\partial t^{\alpha}}=-K\left(-\Delta_{\mathbf{p}}\right)^{v / 2} f(\mathbf{p}, t)+f_{0}(\mathbf{p}) \delta_{\alpha}(t) \tag{76}
\end{equation*}
$$

and the energy distribution $N(E, t)$,

$$
\begin{aligned}
& \frac{\partial^{\alpha} N(E, t)}{\partial t^{\alpha}} \\
& = \begin{cases}\frac{\partial^{v}\left[a_{v} N(E, t)\right]}{\partial E^{v}}+N_{0}(E) \delta_{\alpha}(t), & 0<v<1, \\
\frac{\partial\left[a_{1} N(E, t)\right]}{\partial E}+\frac{\partial^{v}\left[a_{v} N(E, t)\right]}{\partial E^{v}}+N_{0}(E) \delta_{\alpha}(t), & 1<v<2 .\end{cases}
\end{aligned}
$$

Here, $v=\gamma$ for $\gamma \leqslant 2$ and $v=2$ for $\gamma>2$, and it is important for the derivation of the equations themselves that the coefficients $K$, $a_{1}$, and $a_{2}$ are constant. Fractional differential equations are typically derived by integral transformations, which are efficient only if the coefficients are constant. Therefore, it would be incorrect, for example, to derive Eqn (76) with constant coefficients and then to place the variable diffusion coefficient $K(p)$ in front of the fractional Laplacian (as can already be easily verified in the example of Eqns (68)-(70) with an integer Laplacian).

A diffusion packet described by Eqn (76), propagating from the origin of momentum coordinates, has the form of the three-dimensional isotropic fractionally stable distribution $\Psi^{(v, \alpha)}(\xi), v \in(0,2], \alpha \in(0,1]$ 'spreading' by the $t^{\alpha / v}$ law:

$$
f(\mathbf{p}, t)=\left(K t^{\alpha}\right)^{-3 / v} \Psi_{3}^{(v, \alpha)}\left(\left(K t^{\alpha}\right)^{-1 / v} p\right) .
$$

The tails of this distribution are described by a power-law function with the exponent $v$. Physically, this means a peculiar leading-term effect: in the sum $\Delta \mathbf{p}_{1}+\Delta \mathbf{p}_{2}+\ldots+\Delta \mathbf{p}_{n}$ of a large number of independent terms, one of them always plays the leading role. This effect disappears for $v=2$, when the distribution becomes Gaussian (sub-Gaussian). As a result, for $v<2$, we obtain a spectrum of the form

$$
N_{1}(E) \mathrm{d} E \propto E^{-v-1} \mathrm{~d} E
$$

which is only superficially similar to the Fermi formula. The main difference is that the exponent $v$ here is not related to the age of the particles detected, being entirely determined only by the acceleration mechanism in an individual local event (collision). For this reason, the fractal distribution of accelerating regions in space also has no effect on the slope of the spectrum obtained.

### 5.4 Integro-fractional differential model

A disadvantage of the model presented in Section 5.3 is that momentum increments in the acceleration event are independent of the momentum of the particle involved in the interaction, whereas the energy increment (and therefore the momentum increment) in the Fermi model and its later versions is proportional, on average, to the particle energy (momentum) before the interaction. In this case, the energy of the accelerated particle is expressed not by the sum but by the product of independent random quantities. We call this model the multiplicative random walk to distinguish it from the additive random walk model considered above.

In the multiplicative model, the momentum increment is proportional (in the statistical sense) to the absolute value of the momentum $p^{\prime}$ of the interacting particle,

$$
\Delta \mathbf{p}=p^{\prime} \mathbf{q}, \quad \int_{|\Delta \mathbf{p}|>p} w\left(\Delta \mathbf{p} ; \mathbf{p}^{\prime}\right) \mathrm{d} \Delta \mathbf{p} \propto\left(\frac{p}{p^{\prime}}\right)^{-\gamma}, \quad p \rightarrow \infty
$$

Assuming that the distribution of the proportionality vector $\mathbf{q}$ is independent of $\mathbf{p}^{\prime}$ and is isotropic, $W\left(\mathbf{q} ; \mathbf{p}^{\prime}\right) \mathrm{d} \mathbf{q}=$ $(1 / 2) V(q) \mathrm{d} q \mathrm{~d} \xi, \xi=\cos (\mathbf{q}, \mathbf{p})$, we transform kinetic equation (64) into

$$
\begin{align*}
\frac{\partial^{\alpha} f(p, t)}{\partial t^{\alpha}} & =\mu\left\{\int_{-1}^{1} \frac{\mathrm{~d} \xi}{2} \int_{0}^{\infty} V(q) \frac{f\left(p / \sqrt{1+2 \xi q+q^{2}}, t\right)}{\left(\sqrt{1+2 \xi q+q^{2}}\right)^{3}} \mathrm{~d} q\right. \\
& -f(p, t)\}+f_{0}(p) \delta_{\alpha}(t) \tag{77}
\end{align*}
$$

This equation, derived in [130], is a modified version of the model in [131] (taking a change in the direction upon acceleration into account). To make it closer to real addi-tional-acceleration processes, for example, during the intersection of shock-wave fronts in supernova remnants, we assume, as in [132], that

$$
V(q)=\gamma q^{-\gamma-1}, \quad \gamma>1 .
$$

The obtained model can be called the multiplicative Lévy random walk model.

We consider equations for the spectral function in two cases.

Case 1. $\gamma>2$, there exists a second moment of the momentum increment, proportional to $E^{2}$, and we return to the region of classical diffusion with variable coefficients:

$$
\frac{\partial^{\alpha} n(E, t)}{\partial t^{\alpha}}=\frac{\partial\left[a_{1} E n(E, t)\right]}{\partial E}+\frac{\partial^{2}\left[a_{2} E^{2} n(E, t)\right]}{\partial E^{2}}+n_{0}(E) \delta_{\alpha}(t) .
$$

Case 2. $\gamma \ll 2$, and hence all the terms in the radicand in (77) except $q^{2}$ can be neglected,
$\frac{\partial^{\alpha} n(E, t)}{\partial t^{\alpha}}=\mu\left\{\int_{1}^{\infty} \gamma q^{-\gamma-1} n\left(\frac{E}{q}, t\right) \frac{\mathrm{d} q}{q}-n(E, t)\right\}+n_{0}(E) \delta_{\alpha}(t)$.
(We do not discuss the validity of this approximation here and point out that it is this acceleration operator that was used in [132] in specific calculations.)

Solving Eqn (78) with the use of the Mellin-Laplace transform and expressions (62) and (67), for a monoenergetic source $n(E)=\delta\left(E-E_{0}\right)$, we obtain

$$
\begin{equation*}
N_{\alpha}(E ; \tau)=\frac{\mu \tau^{\alpha} \gamma}{\left(1+\mu \tau^{\alpha}\right)^{2}}\left(\frac{E}{E_{0}}\right)^{-1-\gamma /\left(1+\mu \tau^{\alpha}\right)} \frac{1}{E_{0}} . \tag{79}
\end{equation*}
$$

Expression (79) was derived under the assumptions considerably simplifying the real situation, and it is largely a qualitative formula, but it nevertheless compactly reflects the influence of all three sources of the fluctuation acceleration on the shape of the energy spectrum of cosmic rays: fluctuations of the particle age (the parameter $\tau$ ), fluctuations of the number of acceleration events (the parameters $\alpha$ and $\mu$ ), and fluctuations of the energy imparted in a single event (the parameter $\gamma$ ). Representing the scale parameter $\mu$ in the form $\mu=\tau_{\mathrm{A}}^{\alpha}$, where $\tau_{\mathrm{A}}$ is the characteristic time interval between accelerations of particles in the remnants of different supernovae (we recall that $\tau$ is the mean lifetime in nuclear collisions), we can write the absolute value of the exponent of the integrated spectrum in a more transparent form: $\gamma^{\prime}=1+\gamma /\left[1+\left(\tau / \tau_{\mathrm{A}}\right)^{\alpha}\right]$. For $\alpha=1$ and $\mu \tau \gg 1$, we obtain Fermi formula (61) with $a=\mu / \gamma$.

## 6. Problems and outlook

### 6.1 Fractional time derivative

From the physical standpoint, cosmic rays are the highenergy component of the cosmic plasma, and it is not surprising that along with the development of fractional differential models describing cosmic-ray transfer considered in this review, these tools are also used in plasma physics. For example, Balescu [16] combined the collisionless plasma equation with the equations of motion for a charge in a homogeneous magnetic field and a fluctuating electric component with a random potential represented by homogeneous Gaussian noise, deriving a non-Markov equation for the mean concentration of particles, which he called the hybrid equation,

$$
\frac{\partial n(\mathbf{r}, t)}{\partial t}=\int_{0}^{t} \mathrm{~d} \tau \Lambda(\tau) \Delta n(\mathbf{r}, t-\tau),
$$

where $\Lambda(\tau)$ is the memory function related to the properties of the velocity field. In the case of weak turbulence, the delay can be neglected, $n(\mathbf{r}, t-\tau) \approx n(\mathbf{r}, t)$, and the $\tau$ integration can be extended to infinity. As a result, we obtain the usual diffusion equation. In the case of strong turbulence, the consideration of the self-similar (power) character of turbulence led to the integro-differential equation [16]

$$
\frac{\partial n(\mathbf{r}, t)}{\partial t}-\Delta\left[D n(\mathbf{r}, t)-H_{0} \int_{\theta}^{t} n(\mathbf{r}, t-\tau) \tau^{\beta-2} \mathrm{~d} \tau\right], \quad \beta=0.58
$$

which coincides with Eqn (46) and transforms into a differential equation as $\tau \rightarrow 0$.

Like any phenomenological derivation, this derivation is open to question. In most doubt here is not the nonMarkovian property or even the power-law reduction of correlations, but the assumed unboundedness of the power-
law distribution. The assumption that the power-law tail extends to infinity gives rise to a fractional RiemannLiouville derivative: if the power-law tail is cut off, this operator transforms at large times into a standard firstorder differential operator. Are there any physical reasons for restricting this distribution except for purely psychological ones (we do not like the mean confinement time in a trap to become infinite for $\beta<1$ )? This last possibility, incidentally, does not contradict anything: the mean (mathematical expectation value) is defined in terms of an improper integral (with an infinite upper limit), and the existence of such an artificial construction in Nature is unreasonable. Physically, the meaning of the retarded integral and its kernel is the retardation of particle motion caused not by the stopping of the particle but by its coming into a small turbulent region with a small diffusion coefficient.

Numerical simulation of the trajectories of particles in a turbulent plasma has shown [133] that charged particles are indeed captured by vortices ('traps') formed in the turbulent plasma and are confined in them for a long time. This time is random, and its probability distribution is characterized by a sufficiently long power-law region $t^{-\beta-1}, \beta=0.83 \pm 0.22$, followed by a rapid decay (Fig. 28a). A similar situation occurs in the dynamics of bright magnetic elements in the solar photosphere [60], with $\beta=0.83 \pm 0.05$ in the interval $0.3-22.0 \mathrm{~min}$, behind which the distribution also rapidly decays (Fig. 28b). In [134], we simulated the Brownian motion of a particle in a cube with a smooth surface and also found a power-law interval of the confinement time distribution (with the exponent $\beta=1 / 2$ ) terminated by a rapid drop. The difference between the last value of $\beta$ and the values indicated above can easily be explained by the influence of the boundary of the confining region: the natural porosity of the boundary surface reduces the confinement time in the trap, thereby increasing $\beta$. Although the value $\beta=0.8$ introduced in [20] is consistent with the results of plasma simulations, the boundedness of power-law intervals suggests that the results of these calculations should be considered with care. Obviously, such truncated distributions in the large-time limit yield first-order derivatives, but they can lead to fractional derivatives in the pre-asymptotic region. These two regions can be combined into a single equation in the case of soft decay by introducing an exponential factor in the kernel of the fractional differential operator [135]:

$$
\begin{aligned}
& \frac{\partial^{1-\beta} n(\mathbf{r}, t)}{\partial t^{1-\beta}}=\frac{1}{\Gamma(1-\beta)} \frac{\mathrm{d}}{\mathrm{~d} t} \int_{0}^{t}(t-\tau)^{-(1-\beta)} n(\mathbf{r}, \tau) \mathrm{d} \tau \\
& \quad \mapsto \frac{1}{\Gamma(1-\beta)} \frac{\mathrm{d}}{\mathrm{~d} t} \int_{0}^{t}(t-\tau)^{-(1-\beta)} \exp [-\gamma(t-\tau)] n(\mathbf{r}, \tau) \mathrm{d} \tau \\
& \quad=\exp (-\gamma t) \frac{\partial^{1-\beta}}{\partial t^{1-\beta}}[\exp (\gamma t) n(\mathbf{r}, t)] .
\end{aligned}
$$

The subdiffusion equation then takes the form

$$
\begin{aligned}
\frac{\partial n}{\partial t} & =D \Delta \exp (-\gamma t) \frac{\partial^{1-\beta}}{\partial t^{1-\beta}}[\exp (\gamma t) n(\mathbf{r}, t)]+\delta(\mathbf{r}) \delta(t) \\
& \Rightarrow \begin{cases}\frac{\partial n}{\partial t}=D \Delta n(\mathbf{r}, t)+\delta(\mathbf{r}) \delta(t), & t \rightarrow \infty \\
\frac{\partial^{\beta} n}{\partial t^{\beta}}=D \Delta n(\mathbf{r}, t)+\delta(\mathbf{r}) \delta_{\beta}(t), & t \rightarrow 0\end{cases}
\end{aligned}
$$



Figure 28. Confinement time distribution in a magnetic trap. (a) Results of the numerical simulation of plasma [133]; a segment of the straight line corresponds to $\beta=0.83$. (b) Data of observations of the motion of magnetic elements on the solar surface [60]; the straight line corresponds to $\beta=0.76$.
or, in terms of the Laplace transforms,

$$
\lambda \hat{n}(\mathbf{r}, \lambda)=(\lambda+\gamma)^{1-\beta} D \Delta \hat{n}(\mathbf{r}, \lambda)+\delta(\mathbf{r})
$$

For small times (i.e., for $\lambda \gg \gamma$ ), in the expression

$$
\lambda \hat{n}(\mathbf{r}, \lambda)=\lambda^{1-\beta} D \Delta \hat{n}(\mathbf{r}, \lambda)+\delta(\mathbf{r})
$$

we recognize the Laplace transform of the fractional differential subdiffusion equation, which we previously wrote in the form

$$
\lambda^{\beta} \hat{n}(\mathbf{r}, \lambda)=D \Delta \hat{n}(\mathbf{r}, \lambda)+\delta(\mathbf{r}) \lambda^{\beta-1}
$$

whereas in the large-time limit $(\lambda \ll \gamma)$, the same equation gives the transform of the parabolic equation of normal diffusion:

$$
\lambda \hat{n}(\mathbf{r}, \lambda)=D^{\prime} \Delta \hat{n}(\mathbf{r}, \lambda)+\delta(\mathbf{r}), \quad D^{\prime}=\gamma^{1-\beta} D
$$

It is clear that this effect can be obtained not only for the exponential but also for other factors providing a rapid decay (or indeed termination) of the asymptotic part of the powerlaw distribution, such that the mean time would become finite; the exponential factor can easily be incorporated into the result due to the use of the Laplace transform.

### 6.2 Fractional Laplacian

In light of the criticism of the LT version, the question can arise: How should we treat the Monin equation for turbulent diffusion, in which $\alpha=2 / 3$ and the velocity is not mentioned at all, and this is in relativistic hydrodynamics, in fact under laboratory conditions rather than in cosmic-ray physics with its astronomical scales and velocities close to the speed of light? In answering this question, we first note that the Monin equation has not found practical applications (incidentally, as many equations derived in theoretical physics and never used). Moreover, even in Statistical Hydromechanics [32] by Monin and Yaglom, the derivation of this equation is presented, but nothing is said about its applications. The
important reason behind this omission is that the power-law spectrum on which the derivation is based is valid only in a limited range of wave numbers $k$, whereas the fractional Laplacian corresponds to the unbounded power-law spectrum. There is an obvious discrepancy between expression (6) in [32], eventually leading to a fractional Laplacian, and the statement that the self-similar representation of the function

$$
a(k, t)=a_{0}\left(\varepsilon^{1 / 3} k^{2 / 3} t\right)
$$

used in this case is valid only in a limited interval of the arguments: "Restricting ourselves to the consideration of quasi-asymptotic diffusion times $t_{3}<t<t_{1}$, we should admit that the function $a(k, t)$, besides $k$ and $t$, depends on only one other dimensional parameter $\varepsilon$ and therefore has the form $a(k, t)=a_{0}\left(\varepsilon^{1 / 3} k^{2 / 3} t\right) \ldots$. Monin's hypothesis is that operators $A(t)$ form a semigroup, i.e., have the property $A\left(t_{1}\right) A\left(t_{2}\right)=A\left(t_{1}+t_{2}\right)$. This gives $a_{0}\left(x_{1}+x_{2}\right)=$ $a_{0}\left(x_{1}\right) a_{0}\left(x_{2}\right) \ldots$." [32, p. 510]. Obviously, the last expression is incompatible with the restriction presented above. This restriction itself follows from the fact that turbulence has a power-law spectrum only in the inertial interval of wave numbers $k \in\left[k_{\min }, k_{\max }\right]$. As shown in Section 2.5, the fractional Laplacian is determined by the behavior of the spectral function of the free path in the region $k \approx 0$, but it is this region that is excluded from the inertial interval. Therefore, the exponent $\alpha$ cannot be smaller than unity in a continuously turbulent medium, which is the case in the LT version. If $\alpha>1$, the mean free path becomes finite, producing a scale. The multiple excess of this scale allows approximating possible trajectories in a medium by such segments in the medium. At the same time, if particles are propagating in regions with relatively small localized turbulent zones with a fractal distribution, this objection can be ignored. An example is given by cosmic-ray transfer in the metagalaxy: here, a model with $\alpha<1$ and $\beta<1$, i.e., with long linear flights between galaxies and with confinement in the traps formed by galaxies, can be quite acceptable.

In the three-dimensional case, if the divergence of moments of an elementary displacement begins with the
next, the fourth, order, the fractional generalization of the Barnett equation can be introduced in the form
$\frac{\partial f(\mathbf{r}, t)}{\partial t}=D_{2} \Delta f(\mathbf{r}, t)-D_{\alpha}(-\Delta)^{\alpha / 2} f(\mathbf{r}, t), \quad 2<\alpha<4$.
The inclusion of the additional term in the equation violates the important property of the self-similarity of its solutions, but at the same time imparts an interesting feature to it. The Fourier transformation in $x$ gives terms with $k^{2}$ and $|k|^{\alpha} \equiv\left(k^{2}\right)^{\alpha / 2}$ in the right-hand side; as $k \rightarrow 0$, the first of them is leading, and as $k \rightarrow \infty$, the second one is (we recall that $\alpha>2$ ). Therefore, the process described by Eqn (80) looks like usual diffusion at large scales and like delayed diffusion at small scales. The same equation with $\alpha$ from the lower range,

$$
\frac{\partial f(\mathbf{r}, t)}{\partial t}=D_{2} \Delta f(\mathbf{r}, t)-D_{\alpha}(-\Delta)^{\alpha / 2} f(\mathbf{r}, t), \quad 0<\alpha<2,
$$

describes usual diffusion at smaller scales and superdiffusion at larger scales. The difference between these processes is explained, of course, by the different properties of the medium at different scales.

Another problematic aspect is the nonlocality of the fractional Laplace operator caused by its integral nature. Unlike the usual Laplacian

$$
\Delta=\frac{\partial^{2}}{\partial x^{2}}+\frac{\partial^{2}}{\partial y^{2}}+\frac{\partial^{2}}{\partial z^{2}},
$$

whose form is independent of the boundaries and boundary conditions, the fractional Laplacian depends on them. In the fractional Laplacian, it is necessary to specify not only the properties of the sought function at the domain boundaries, but also its values outside this domain. The popular Fourier transform $|\mathbf{k}|^{\alpha / 2}$ of the fractional Laplacian in a bounded medium is no longer applicable. In this case, the interpretation in terms of flights is very useful for determining the influence of the boundary quality (reflecting, transparent, semitransparent, diffusive) on the solution and for better understanding characteristics that are not so obvious, such as the times the boundary is first reached and first crossed. We note that the expansion of the Laplacian in Cartesian or other orthogonal coordinates, providing a theoretical basis for the method of separation of variables, is not applicable in the fractional differential case: the fractional three-dimensional Laplacian cannot be written as a sum of one-dimensional Laplacians along $x, y$, and $z$. This is obvious in the Fourier representation, where

$$
|\mathbf{k}|^{2}=k_{x}^{2}+k_{y}^{2}+k_{z}^{2}
$$

and

$$
|\mathbf{k}|^{\alpha}=\left(k_{x}^{2}+k_{y}^{2}+k_{z}^{2}\right)^{\alpha / 2} \neq\left|k_{x}\right|^{\alpha}+\left|k_{y}\right|^{\alpha}+\left|k_{z}\right|^{\alpha}, \quad \alpha \neq 2 .
$$

Similarly, it is impossible to separate the fractional Laplacian into the radial and angular components. The use of only a radial fractional Laplacian in any equation can only mean that the motion only along radial trajectories is considered, whereas in the case of the usual Laplacian, its radial component reflects the evolution of the radial coordinate of a complex spatial trajectory.

The authors of [136] introduced a matrix representation of the one-dimensional fractional Laplacian, which was used for
numerical solutions of problems with absorbing boundaries. The authors of [137] showed that in the presence of a reflecting wall, the fractional Laplacian for an infinite medium

$$
\begin{aligned}
-(-\Delta)^{\alpha / 2} f(x, t) & =-\frac{1}{2 \cos (\alpha \pi / 2) \Gamma(2-\alpha)} \\
& \times \frac{\partial^{2}}{\partial x^{2}} \int_{0}^{\infty}|x-\xi|^{1-\alpha} f(\xi, t) \mathrm{d} \xi, \quad 1<\alpha<2
\end{aligned}
$$

transforms into an integro-differential operator with a modified kernel:

$$
\begin{aligned}
& -(-\Delta)_{\text {refl }}^{\alpha / 2} f(x, t)=-\frac{1}{2 \cos (\alpha \pi / 2) \Gamma(2-\alpha)} \\
& \quad \times \frac{\partial^{2}}{\partial x^{2}} \int_{0}^{\infty}\left[|x-\xi|^{1-\alpha}+(x+\xi)^{1-\alpha}\right] f(\xi, t) \mathrm{d} \xi .
\end{aligned}
$$

In [138], the fractional Laplacian was introduced as the generalization of the one-dimensional expression for the fractional Marchaut derivative on a limited segment of the axis for a bounded domain of the $d$-dimensional space:

$$
\begin{aligned}
& \mathcal{D}_{\mathrm{G}}^{\alpha} f(\mathbf{x})=C(\alpha)\left[a_{\mathrm{G}}(\mathbf{x}) f(\mathbf{x})+\int_{\mathrm{G}} \frac{f(\mathbf{x})-f(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|^{d+\alpha}} \mathrm{d} \mathbf{y}\right], \\
& \mathbf{x} \in G \subset \mathrm{R}^{d}, \quad \alpha \in(0,1),
\end{aligned}
$$

where

$$
\begin{aligned}
& C(\alpha)=\frac{\alpha 2^{\alpha-1} \Gamma[(d+\alpha) / 2]}{\pi^{d / 2} \Gamma(1-\alpha / 2)}, \\
& a_{\mathrm{G}}(\mathbf{x})=\int_{\mathrm{R}^{d} \backslash \mathrm{G}} \frac{\mathrm{~d} \mathbf{y}}{|\mathbf{x}-\mathbf{y}|^{d+\alpha}} .
\end{aligned}
$$

The authors of [139], who studied reflected symmetric stable processes, called the limit

$$
-(-\Delta)_{\mathrm{G}}^{\alpha / 2} f(\mathbf{x}) \equiv \lim _{\varepsilon\lfloor 0} C(\alpha) \int_{\mathrm{G},|\mathbf{x}-\mathbf{y}|>\varepsilon} \frac{f(\mathbf{x})-f(\mathbf{y})}{|\mathbf{y}-\mathbf{x}|^{d+\alpha}} \mathrm{d} \mathbf{y}
$$

## the regional fractional Laplacian.

Because of the abovementioned difficulties encountered in the consideration of boundary conditions, the best method for formulating boundary value problems in the nonlocal theory is still based on the use of integral equations and Monte Carlo simulations.

### 6.3 Gradient and material derivative

Considering the problem of cosmic-ray transfer from the standpoint of the hypothesis of hidden variables, we see that the appearance of fractional time derivatives is quite expected because the motion of cosmic rays affects magnetic fields, and vice versa. The usual diffusion theory neglects this relation, whereas the fractional character of derivatives can suggest that this relation is partially taken into account in the remaining equations. However, there is a very serious question we must solve to acquire complete confidence in following the fractional differential way. Namely, as soon as we pass to the kinetic description ('include velocity'), the partial derivative $\partial / \partial t$ transforms into the material derivative $\mathrm{d} / \mathrm{d} t \equiv \partial / \partial t+\mathbf{v} \nabla$. In the fractional differential approach, this gives rise to the transformation $(\partial / \partial t)^{\alpha} \mapsto(\mathrm{d} / \mathrm{d} t)^{\alpha}$. The problem is that fractional operators are nonlocal, and it is
therefore necessary to indicate the domain in which the function involved in the calculation procedure is defined. Strictly speaking, the notation for the fractional time derivative we used is not quite satisfactory and should be replaced, for example, with the expression
${ }_{a} D_{t}^{\alpha} f(t) \equiv \frac{1}{\Gamma(1-\alpha)} \frac{\partial}{\partial t} \int_{a}^{t} \frac{f\left(t^{\prime}\right) \mathrm{d} t^{\prime}}{\left(t-t^{\prime}\right)^{\alpha}}, \quad-\infty \leqslant a<t, \alpha \in(0,1)$.
The lower limit is determined by physical conditions: if an object (a particle) was created at some instant, which we take as the onset of time counting, then $a$ should be set equal to zero, and if an object exists always (say, in the visible past), it is natural to set $a=-\infty$. By substituting different $a$ in the expression presented above, we obtain different values of the derivative of the same function $f(t)$ at the same instant of time $t$.

In some recent papers (see, e.g., [140, 141]), the concept of the fractional operator is extended to the gradient as

$$
\nabla^{\alpha} f(x, y)=f_{x}^{(\alpha)}(x, y) \mathbf{e}_{x}+f_{y}^{(\alpha)}(x, y) \mathbf{e}_{y}
$$

The reader may find this generalization natural (indeed, we are not aware of critical remarks in this respect), but this is only at first glance and as long as the picture of a particular process is not examined closely. The authors of [142] considered the motion of a point mass in a plane where the function $f(x, y)$ plays the role of the potential energy and $x$ and $y$ are the projections of velocities. For $\alpha \in(0,1)$, the system of corresponding equations has the form

$$
\begin{aligned}
& \frac{\mathrm{d} x}{\mathrm{~d} t}=-f_{x}^{(\alpha)}(x, y)=-\frac{1}{\Gamma(1-\alpha)} \frac{\partial}{\partial x} \int_{0}^{x} \frac{f\left(x^{\prime}, y\right) \mathrm{d} x^{\prime}}{\left(x-x^{\prime}\right)^{\alpha}} \\
& \frac{\mathrm{d} y}{\mathrm{~d} t}=-f_{y}^{(\alpha)}(x, y)=-\frac{1}{\Gamma(1-\alpha)} \frac{\partial}{\partial y} \int_{0}^{y} \frac{f\left(x, y^{\prime}\right) \mathrm{d} y^{\prime}}{\left(y-y^{\prime}\right)^{\alpha}}
\end{aligned}
$$

But what physical interpretation can be attached to these equations? Apparently, it is natural to integrate over the particle trajectory segment whereby the particle arrives at the given point P , but different components of the gradient at the specified point are calculated over different trajectories $\mathrm{O}^{\prime \prime} \mathrm{P}$ and $\mathrm{O}^{\prime} \mathrm{P}$ intersecting only at this point (Fig. 29). This operation denies physical interpretation.

Integrating along particle trajectory OP seems to be more natural,

$$
\begin{aligned}
& \frac{\mathrm{d} x}{\mathrm{~d} t}=-\frac{1}{\Gamma(1-\alpha)} \frac{\partial}{\partial x} \int_{0}^{x} \frac{f\left(x^{\prime}\left(t^{\prime}\right), y^{\prime}\left(t^{\prime}\right)\right) \mathrm{d} x^{\prime}\left(t^{\prime}\right)}{\left(x-x^{\prime}\left(t^{\prime}\right)\right)^{\alpha}} \\
& \frac{\mathrm{d} y}{\mathrm{~d} t}=-\frac{1}{\Gamma(1-\alpha)} \frac{\partial}{\partial y} \int_{0}^{y} \frac{f\left(x^{\prime}\left(t^{\prime}\right), y^{\prime}\left(t^{\prime}\right)\right) \mathrm{d} y^{\prime}\left(t^{\prime}\right)}{\left(y-y^{\prime}\left(t^{\prime}\right)\right)^{\alpha}}
\end{aligned}
$$

where $x^{\prime}\left(t^{\prime}\right)$ and $y^{\prime}\left(t^{\prime}\right)$ are coordinates of the particle that arrived to the point with coordinates $x$ and $y$ by the instant $t^{\prime}<t$.

Similarly, the definition of the fractional material derivative in the form

$$
\begin{aligned}
\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right)^{\alpha} f(\mathbf{r}, t) & =\frac{1}{\Gamma(1-\alpha)}\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right) \\
& \times \int_{-\infty}^{t} \frac{f(\mathbf{r}-(t-\tau) \mathbf{v}, \tau)}{(t-\tau)^{\alpha}} \mathrm{d} \tau
\end{aligned}
$$



Figure 29. The fractional gradient problem.
can be naturally replaced by the definition

$$
\begin{aligned}
\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right)^{\alpha} f(\mathbf{r}, t) & =\frac{1}{\Gamma(1-\alpha)}\left(\frac{\partial}{\partial t}+\mathbf{v} \nabla\right) \\
& \times \int_{-\infty}^{t} \frac{f(\mathbf{R}(\tau ; \mathbf{r}, t), \tau)}{(t-\tau)^{\alpha}} \mathrm{d} \tau
\end{aligned}
$$

where $\mathbf{R}(\tau ; \mathbf{r}, t)$ is the radius vector at the instant $\tau<t$ of a particle that is found at the observation point $\mathbf{r}$ at the measurement instant $t$, with

$$
\mathbf{R}(t ; \mathbf{r}, t)=\mathbf{r}
$$

For different trajectories leading the particle to the observation point, the gradients at this point are different. Interestingly, however, as $\alpha$ approaches 1 , the ever shortening segment of this curve adjacent to the point $\mathbf{r}$ affects the gradient and, finally, for $\alpha=1$, using the limit

$$
\lim _{\alpha \rightarrow 1} \frac{(t-\tau)^{-\alpha}}{\Gamma(1-\alpha)}=\delta(t-\tau)
$$

we obtain the usual independence of the gradient from the trajectory prehistory.

## 7. Conclusions

This review, the third in Physics-Uspekhi to reflect our experience in the application of the fractional differential apparatus to solution of physical problems, is devoted to the same goal as the two previous reviews [25, 143], the goal of rehabilitating this approach, removing some tinge of artificiality from it, and revealing its direct relation to the physics of natural processes. The problem is that fractional operators are nonlocal, and we cannot therefore derive master equations by the usual method, considering the relation between infinitesimal changes in a chosen quantity and infinitesimal increments in time and coordinates. Meanwhile, the number of publications devoted to the application of this apparatus to various physical problems is increasing, currently exceeding 3-4 thousand by our estimate. Many of these papers are constructed following a very primitive scheme: a known differential equation (for example, Newton's equation) is taken, a few magic words (like memory, fractality, complexity) are uttered, then integer-order (specifically, second-order) derivatives are replaced by fractional-order derivatives, some mathematical corollaries are discussed, and the paper can go to the printer! Some authors practice such a technique exclusively. Fractional differential analogs of the Lagrange, Hamilton,

Liouville, and Boltzmann equations, the chains of the Bogoliubov-Born-Green-Kirkwood-Yvon equations, diffusion equations, rheology equations of viscoelastic media, the Fokker-Planck, Navier-Stokes, Ginzburg-Landau, Schrödinger, Heisenberg, Klein-Gordon, and Dirac equations have been derived (although this procedure can hardly be called a derivation). It would appear that theoretical physics should receive a powerful impetus, if this opening up of 'new lands' can be compared to passing from integers to real numbers.

But the modesty of genuinely new results is discouraging. This is natural: it would be naive to assume that by simply replacing two in the derivative symbol with a fractional number $\alpha$ and solving the equation using the known scheme, we would at once arrive at 'new post-Newtonian physics'. Rather tedious work is required to provide a reliable understanding of natural reasons for the appearance (not the introduction) of fractional derivatives in a given problem, the specificity of reflection of the observed relations by them, a reliable interpretation of the results obtained, and an accurate establishment of the applicability limits of the relevant models.

More infallible are papers where fractional derivatives appear due to mathematical transformations of standard classical equations. An excellent example of this is the Basse or hereditary force acting on a sphere moving randomly in a viscous liquid (the last term in expression (4) in [144, p. 132]):

$$
F_{\mathrm{B}}=6 \pi \rho R^{2} \sqrt{\frac{v}{\pi}} \int_{-\infty}^{t} \frac{\mathrm{~d} u}{\mathrm{~d} t} \frac{\mathrm{~d} t}{\sqrt{t-\tau}} .
$$

We saw in Section 4.3 how the fractional derivative appears in discussing the transverse diffusion of cosmic rays. However, additional assumptions, which are inevitable in using the perturbation theory, significantly restrict the validity of the obtained results to the region of weak turbulence. The kinetic equation for particles in strongly turbulent plasmas derived in [145] contains an integral operator in the momentum logarithm (which is, generally speaking, not a fractional derivative, but which reflects the nonlocality of the process in the momentum space). The generalized hydrodynamics developed by Alekseev [146, 147] is also essentially nonlocal. In discussing its application to turbulent currents, the author points out that in the framework of the correlation approach using time-averaged products of the differences of velocities at two close points of the type

$$
B_{i k}=\overline{\left(v_{2 i}-v_{1 i}\right)\left(v_{2 k}-v_{1 k}\right)},
$$

it remains unclear what the term 'close points' means and how the time averaging is realized. In this connection, he quotes Hydrodynamics by Landau and Lifshitz [144, p. 200]: "One could think that there is the principal possibility of obtaining the universal formula (applicable to any turbulent motion) determining $B_{r r}$ and $B_{t t}$ for all distances that are small compared to $l_{\mathrm{L}}$. In reality, however, such a formula cannot exist at all, as is evident from the following considerations. The instant value of $\left(v_{2 i}-v_{1 i}\right)\left(v_{2 k}-v_{1 k}\right)$ could theoretically be expressed through the dissipation of energy $\varepsilon$ at the same instant of time $t$. However, upon averaging these expressions, the law of changing $\varepsilon$ during large-scale motions (of the order of $l_{\mathrm{L}}$ ), which is different for different particular motions, is important. Therefore, the result of averaging
cannot be universal." In other words, we cannot perform this averaging using only information about the infinitesimal vicinity of a given point: it is necessary to known the evolution of the field in a large region, including possible boundary conditions. However, the turbulent diffusion coefficient belongs to the same type of characteristics, and therefore the appearance of a nonlocal operator (fractional Laplacian) in the turbulent diffusion equation, which we observed, is a consequence of the fact pointed out in the above quotation.

It is likely that Ginzburg, who read Alekseev's manuscript and gave him advice, did not relate the potential possibilities of the nonlocal model of turbulent diffusion to the problem of cosmic-ray propagation in the Galaxy, where turbulence plays no less important role than in 'earth hydrodynamics'. Most likely, he no longer considered this problem among the urgent problems in cosmic-ray physics [148].

We can present other examples of the 'penetration' of nonlocal operators to plasma hydrodynamics and kinetics. Each time we meet an argument substantiating the introduction or explaining the appearance of nonlocal operators. Is there a general substantiation of the appearance of a nonlocal operator (e.g., in time), not related to a particular process and its approximate description? Yes, there is, and we find one, for example, in [149]. The authors consider a closed system of $N$ particles, whose motion in the $2 N$ dimensional phase space is governed by the Liouville differential equation of the Markov type, i.e., by the equation containing only the first time derivative. However, if we are interested only in the momentum distribution, then, using the technique of Zwan-zig-Mori projection operators, we obtain an equation for the momentum distribution without any simplifications and approximations, but that equation already contains a retarded integral, i.e., is not Markovian. The same is obtained for other variables. And we see the same if we observe a part of a conservative closed system, not seeing its other part. The invisible (hidden) part of the system affects the behavior of the visible part during the entire period before observation. By observing only the momentum distribution of particles, we ignore coordinates (assign them to the class of hidden variables). The state of the hidden part at the measurement instant is unknown, but it can, at least in principle, be reconstructed by analyzing the prehistory of the observed part. In attempting to describe the motion of cosmic rays without a detailed description of the evolution of the interstellar magnetic field, we assign the characteristics of the latter to the class of hidden variables and again arrive at a hereditary (non-Markov, retarded) equation for the cosmicray distribution. Physically, the information is delivered from different elements of the hidden part of the system to the observed characteristics at the observation point with a finite velocity, which in turn gives rise to a spatial nonlocality related to the appropriate structure of the medium. Fractional differential operators containing singular kernels of the power-law type are a consequence of the additional assumption about the self-similarity of the process. This is the 'philosophy' of the fractional differential phenomenology as we see it today.

## Acknowledgments

The author thanks the organizers of scientific seminars at the Skobeltsyn Research Institute for Nuclear Physics, Lomonosov Moscow State University, and the Lebedev Physical Institute, RAS, for the invitations to give talks on the topic
presented in the review, and the participants of these seminars for the useful comments and fruitful discussions. The author also highly appreciates the careful professional preparation of the manuscript by the editorial staff of Physics-Uspekhi and thanks his colleagues: Dr. R T Sibatov for numerous discussions of the problems considered here and some calculations and E V Kozhemyakina for her help in the preparation of the manuscript for publication.

The work was supported by the Ministry of Education and Science, Russian Federation, State Program 2.1894.2011 (agreement 14.B37.21.1296) and the Russian Foundation for Basic Research (project nos. 13-01-00585 and 11-01-00660).

## References

1. Fermi E Phys. Rev. 751169 (1949) [Translated into Russian: Nauchnye Trudy (Scientific Papers) Vol. 2 (Moscow: Nauka, 1972) p. 439]
2. Strömgren B Astrophys. J. 108242 (1948)
3. Ginzburg V L Usp. Fiz. Nauk 51343 (1953)
4. Ginzburg V L, Syrovatskii S I The Origin of Cosmic Rays (Oxford: Pergamon Press, 1964) [Translated from Russian: Proiskhozhdenie Kosmicheskikh Luchei (Moscow: Izd. AN SSSR, 1963) p. 173]
5. Heisenberg W, in Preludes in Theoretical Physics, in Honor of V.F. Weisskopf (Eds A de-Shalit, H Feshbach, L van Hove) (Amsterdam: North-Holland, 1966) p. 166 [Translated into Russian: Usp. Fiz. Nauk 91731 (1967)]
6. Berezinskii V Set al. Astrophysics of Cosmic Rays (Ed. V L Ginzburg) (Amsterdam: North-Holland, 1990) [Translated from Russian: Astrofiizika Kosmicheskikh Luchei (Ed. V L Ginzburg) (Moscow: Nauka, 1990)]
7. Getmantsev G G Sov. Astron. 6477 (1963) [Astron. Zh. 39607 (1962)]
8. Jokipii J R, Parker E N Astrophys. J. 155799 (1969)
9. Forman M A, Jokipii J R, Owens A J Astrophys. J. 192535 (1974)
10. Hayakawa S Cosmic Ray Physics; Nuclear and Astrophysical Aspects (New York: Wiley-Interscience, 1969) [Translated into Russian: Fizika Kosmicheskikh Luchei (Physics of Cosmic Rays) Pt. 2 (Moscow: Mir, 1974)]
11. Urch I H Astrophys. Space Sci. 48231 (1977)
12. Urch I H Astrophys. Space Sci. 49443 (1977)
13. Webb G M et al. Astrophys. J. 651211 (2006)
14. Ptuskin V S, Chuvil'gin L G, Preprint No. 49 (Troitsk: Institute of Terrestrial Magnetism, Ionosphere and Radio Wave Propagation, 1990)
15. Chuvilgin L G, Ptuskin V S Astron. Astrophys. 279278 (1993)
16. Balescu R Plasma Phys. Control. Fusion 42 B1 (2000)
17. Parker E N Space Sci. Rev. 9651 (1969)
18. Zybin K P, Istomin Ya N Sov. Phys. JETP 62479 (1985) [Zh. Eksp. Teor. Fiz. 89836 (1985)]
19. Lagutin A A, Nikulin Yu A, Uchaikin V V, Preprint No. 4 (Barnaul: Altai State Univ., 2000)
20. Lagutin A A, Nikulin Yu A, Uchaikin V V Nucl. Phys. B Proc. Suppl. 97267 (2001)
21. Zolotarev V M, Uchaikin V V, Saenko V V JETP 88780 (1999) [Zh. Eksp. Teor. Fiz. 1151411 (1999)]
22. Uchaikin V V JETP 881155 (1999) [Zh. Eksp. Teor. Fiz. 1152113 (1999)]
23. Uchaikin V V, Zolotarev V M Chance and Stability. Stable Distributions and Their Applications (Utrecht: VSP, 1999)
24. Einstein A Ann. Phys. 17549 (1905) [Translated into Russian: Sobranie Nauchnykh Trudov (Collection of Scientific Papers) Vol. 3 (Moscow: Nauka, 1966) p. 108]
25. Uchaikin V V Phys. Usp. 46821 (2003) [Usp. Fiz. Nauk 173847 (2003)]
26. Burnett D Proc. London Math. Soc. 39385 (1935)
27. Kolmogoroff A N C.R. (Dokl.) Acad. Sci. USSR 30301 (1941) [Dokl. Akad. Nauk SSSR 30299 (1941)]
28. Kolmogorov A N, in Mécanique de la turbulence: Colloq. intern. du Centre national de la recherche scientifique, Marseille, 28 Août2 Septembre 1961 (Paris: Éditions du Centre national de la recherche scientifique, 1962) p. 447 [Translated into English: The Mechanics of Turbulence. Intern. Symp. of the National Scientific Research Center, 1961 (New York: Gordon and Breach, 1964)]
29. Obukhov A M Izv. Akad. Nauk SSSR Ser. Geogr. Geofiz. 5453 (1941)
30. Oboukhov A M J. Fluid Mech. 1377 (1962)
31. Monin A S Dokl. Akad. Nauk SSSR 105256 (1955)
32. Monin A S, Yaglom A M Statistical Fluid Mechanics. Mechanics of Turbulence Vol. 2 (Cambridge, Mass.: MIT Press, 1975) [Translated from Russian: Statisticheskaya Gidromekhanika. Mekhanika Turbulentnosti Pt. 2 (Moscow: Nauka, 1967)]
33. Uchaikin V V Metod Drobnykh Proizvodnykh (Method of Fractional Derivatives) (Ul'yanovsk: Artishok, 2008)
34. Bakunin O G Turbulence and Diffusion: Scaling Versus Equations (New York: Springer, 2008)
35. Uchaikin V V Fractional Derivatives for Physicists and Engineers Vols 1, 2 (New York: Springer, 2012)
36. Bakunin O G Phys. Usp. 46733 (2003) [Usp. Fiz. Nauk 173757 (2003)]
37. Bakunin O G Phys. Usp. 56243 (2013) [Usp. Fiz. Nauk 183257 (2013)]
38. Isichenko M B Rev. Mod. Phys. 64961 (1992)
39. Zelenyi L M, Milovanov A V Phys. Usp. 47749 (2004) Usp. Fiz. Nauk 174809 (2004)]
40. Uchaikin V V JETP 97810 (2003) [Zh. Eksp. Teor. Fiz. 124903 (2003)]
41. Klafter J, Blumen A, Shlesinger M F Phys. Rev. A 353081 (1987)
42. Feller W An Introduction to Probability Theory and Its Applications Vol. 2 (New York: Wiley, 1971) [Translated into Russian: Vvedenie v Teoriyu Veroyatnostei i Ee Prilozheniya Vol. 2 (Moscow: Mir, 1984)]
43. Del-Castillo-Negrete D, Carreras B A, Lynch V E Phys. Rev. Lett. 94065003 (2005)
44. Dorman L I Eksperimental'nye i Teoreticheskie Osnovy Astrofiziki Kosmicheskikh Luchei (Experimental and Theoretical Foundations of Cosmic Ray Astrophysics) (Moscow: Nauka, 1975)
45. Combes F, in The Chaotic Universe (Advanced Series in Astrophysics and Cosmology, Vol. 10, Eds V G Gurzadyan, R Ruffini) (Singapore: World Scientific, 2000) p. 143
46. Bershadskii A G Sov. Phys. Usp. 331073 (1990) [Usp. Fiz. Nauk 160 (12) 189 (1990)]
47. Uchaikin V V, Korobko D A, Gismyatov I F Russ. Phys. J. 40711 (1997) [Izv. Vyssh. Uchebn. Zaved. Fiz. (8) 7 (1997)]
48. Ragot B R, Kirk J G Astron. Astrophys. 327432 (1997)
49. Zweibel E G, Heiles C Nature 385131 (1997)
50. Dendy R O et al., in Proc. of the 24th Intern. Cosmic Ray Conf., ICRC, Rome, 1995 Vol. 3 (Eds N Iucci, E Lamanna) (Urbino: Intern. Union of Pure and Applied Physics, 1995) p. 221
51. Duffy P et al. Astron. Astrophys. 302 L21 (1995)
52. White R B, Rax J M, Wu Y, in Statistical Description of Transport in Plasma, Astro- and Nuclear Physics (Eds J Misguich, G Pelletier, P Schuck) (New York: Nova Science Publ., 1993) p. 75
53. Kadomtsev B B, Pogutse O P, in Plasma Physics and Controlled Nuclear Fusion Research 1978. Proc. of the 7th Intern. Conf., IAEA, Innsbruck, Austria, August 23-30, 1978 Vol. 1 (Vienna: Intern. Atomic Energy Agency, 1978) p. 649
54. Kulakov A V, Rumyantsev A A Phys. Dokl. 39337 (1994) [Dokl. Ross. Akad. Nauk 336183 (1994)]
55. Elmegreen B G Publ. Astron. Soc. Australia 1574 (1998)
56. Shibata T Nucl. Phys. B Proc. Suppl. 7522 (1997)
57. Lagutin A A, Uchaikin V V, in Proc. of the 27th Intern. Cosmic Ray Conf., 7-15 August, 2001, Hamburg, Germany Vol. 5 (Hamburg: Copernicus Gesellschaft, 2001) p. 1896
58. Lagutin A A, Uchaikin V V Nucl. Instrum. Meth. Phys. Res. B 201 212 (2003)
59. Lagutin A A, Strelnikov D V, Tyumentsev A G, in Proc. of 27 th Intern. Cosmic Ray Conf., Hamburg, Germany, 2001 Vol. 5 (Hamburg: Copernicus Gesellschaft, 2001)
60. Cadavid A C, Lawrence J K, Ruzmaikin A A Astrophys. J. 521844 (1999)
61. Osborne J L,Wdowczyk J, Wolfendal A W J. Phys. A Math. Gen. 9 1399 (1976)
62. Dorman L I, Ghosh A, Ptuskin V S Astrophys. Space Sci. 10987 (1985)
63. Lagutin A A, Makarov V V, Tyumentsev A G, in Proc. of 27 th Intern. Cosmic Ray Conf., Hamburg, Germany, 2001 Vol. 5 (Hamburg: Copernicus Gesellschaft, 2001)
64. Lagutin A A Probl. Atom. Sci. Technol. (6) 214 (2001)
65. Lagutin A A, Osadchiy K I, Strelnikov D V, in Proc. of 27 th Intern. Cosmic Ray Conf., Hamburg, Germany, 2001 Vol. 5 (Hamburg: Copernicus Gesellschaft, 2001) p. 1852
66. Lagutin A A, Osadchiy K I, Gerasimov V V Probl. Atom. Sci. Technol. (6) 209 (2001)
67. Lagutin A A et al., in Proc. of the 28th Intern. Cosmic Ray Conf., July 31 -August 7, 2003, Tsukuba, Japan Vol. 2 (Eds T. Kajita et al.) (Tsukuba, 2003) p. 675; astro-ph/0306187
68. Lagutin A A, Tyumentsev A G Izv. Ross. Akad. Nauk Ser. Fiz. 67 439 (2003)
69. Erlykin A D, Lagutin A A, Wolfendale A W Astropart. Phys. 19351 (2003)
70. Kraichnan R H Phys. Fluids 81385 (1965)
71. Tchen C M, in Atmospheric Diffusion and Air Pollution (Advances in Geophysics, Vol. 6, Eds F N Frenkiel, P A Sheppard) (New York: Academic Press, 1959) p. 165 [Translated into Russian: Atmosfernaya Diffuziya i Zagryaznenie Vozdukha (Moscow: IL, 1962) p. 191]
72. Erlykin A D, Wolfendale A W Astropart. Phys. 4270 (2013); arXiv: 1212.2760
73. Lagutin A A, Tyumentsev A G Izv. Altaiskogo Gos. Univ. (Special issue) 4 (2004)
74. Uchaikin V V Gravitation Cosmology 105 (2004)
75. Uchaikin V V, Korobko D A Uchenye Zapiski Ul'yanovsk. Gos. Univ. Ser. Fiz. (1(4)) 3 (1998)
76. Korobko D A, Author's Abstract of Ph.D. Thesis (Candidate of Phys. Math. Sci.) (Ulyanovsk: Ulyanovsk State Univ., 1999)
77. Lagutin A A, Raikin R I, Tyumentsev A G Izv. Altaiskogo Gos. Univ. (Special issue) 27 (2004)
78. Ketabi N, Fatemi J Transaction B Mech. Eng. 16269 (2009)
79. Kermani H A, Fatemi J South African J. Sci. 1072 (2011)
80. Doostmohammadi S, Fatemi S J ISRN High Energy Phys. 2012 673250 (2012)
81. Uchaikin V V JETP Lett. 91105 (2010) [Pis'ma Zh. Eksp. Teor. Fiz. 91115 (2010)]
82. Lagutin A A et al., in Proc. of the 32nd Intern. Cosmic Ray Conf., ICRC2011, 11-18 August, 2011, Beijing, China
83. Lagutin A, Tyumentsev A J. Phys. Conf. Ser. 409012050 (2013)
84. Bugayov V V et al., in Proc. of the 30th Intern. Cosmic Ray Conf., July 3-11, 2007, Mérida, Yucatán, Mexico Vol. 2 (Eds R Caballero et al.) (Mexico City, Mexico: Univ. Nacional Autónoma de México, 2008) p. 179
85. Uchaikin V V Theor. Math. Phys. 115496 (1998) [Teor. Mat. Fiz. 115154 (1998)]
86. Uchaikin V V Tech. Phys. 43124 (1998) [Zh. Tekh. Fiz. 68 (1) 138 (1998)]
87. Uchaikin V V Physica A 25565 (1998)
88. Uchaikin V V, Yarovikova I V Comput. Math. Math. Phys. 431478 (2003) [Zh. Vychisl. Mat. Mat. Fiz. 431536 (2003)]
89. Samko S G, Kilbas A A, Marichev O I Fractional Integrals and Derivatives: Theory and Applications (Philadelphia, Pa.: Gordon and Breach Sci. Publ., 1993) [Translated from Russian: Integraly $i$ Proizvodnye Drobnogo Poryadka i Nekotorye Ikh Prilozheniya (Minsk: Nauka i Tekhnika, 1987)]
90. Uchaikin V V, Sibatov R T Gravitation Cosmology 18122 (2012)
91. Uchaikin V V, Sibatov R T, Saenko V V Bull. Russ. Acad. Sci. Phys. 77619 (2013) [Izv. Ross. Akad. Nauk Ser. Fiz. 77688 (2013)]
92. Uchaikin V V, Sibatov R T, Saenko V V, in 23rd European Cosmic Ray Symp., Moscow, Russia, July, 3-7, 2012, PCR1 ID 665
93. Taillet R, Maurin D Astron. Astrophys. 402971 (2003)
94. Skilling J Astrophys. J. 170265 (1971)
95. Krommes J A Prog. Theor. Phys. Suppl. 64137 (1978)
96. Rechester A B, Rosenbluth M N Phys. Rev. Lett. 4038 (1978)
97. Ramaty R, Lingenfelter R E, in Isotopic Composition of the Primary Cosmic Radiation. Proc. of a Symp., Lyngby, Denmark, March, 1971 (Ed. P M Dauber) (Lyngby: Danisch Space Res. Inst., 1971) p. 203
98. Zel'dovich Ya B, Sokolov D D Sov. Phys. Usp. 28608 (1985) [Usp. Fiz. Nauk 146493 (1985)]
99. Shalchi A, Kourakis I Phys. Plasmas 14092903 (2007)
100. Shalchi A, Kourakis I Phys. Plasmas 14112901 (2007)
101. Shalchi A et al. Phys. Rev. E 80066408 (2009)
102. Corrsin S, in Atmospheric Diffusion and Air Pollution (Advances in Geophysics, Vol. 6, Eds F N Frenkiel, P A Sheppard) (New York: Academic Press, 1959) p. 161 [Translated into Russian: Atmosfernaya Diffuziya i Zagryaznenie Vozdukha (Moscow: IL, 1962) p. 187]
103. Uchaikin V V, in Long-range Interactions, Stochasticity and Fractional Dynamics (Nonlinear Physical Science, Eds A C J Luo, V Afraimovich) (Beijing: Higher Education Press, Berlin: SpringerVerlag, 2010) p. 27
104. Mandelbrot B B, Van Ness J W SIAM Rev. 10422 (1968)
105. Wang K G, Lung C W Phys. Lett. A 151119 (1990)
106. Jokipii J R Astrophys. J. 1831029 (1973)
107. Schlickeiser R Cosmic Ray Astrophysics (Berlin: Springer, 2002)
108. Toptygin I N Cosmic Rays in Interplanetary Magnetic Fields (Dordrecht: D. Reidel, 1985) [Translated from Russian: Kosmicheskie Luchi v Mezhplanetnykh Magnitnykh Polyakh (Moscow: Nauka, 1983)]
109. Tsytovich V N Theory of Turbulent Plasma (New York: Consultants Bureau, 1977) [Translated from Russian: Teoriya Turbulentnoi Plazmy (Moscow: Atomizdat, 1971)]
110. Dorman L I, Ptuskin V S, Chuvil'gin L G Geomagn. Aeron. 29405 (1989) [Geomagn. Aeronom. 29529 (1989)]
111. Dorman L I, Ptuskin V S, Chuvil'gin L G Izv. Vyssh. Uchebn. Zaved. Radiofiz. 5172 (1962)
112. Roberts P H, Research Report No. HSN-2 (New York: New York Univ., 1960)
113. Webb G M et al. J. Phys. A Math. Theor. 42235502 (2009)
114. Parker E N Planet. Space Sci. 139 (1965)
115. Barenblatt G I, Zel'dovich Ya B Russ. Math. Surv. 26 (2) 45 (1971) [Usp. Mat. Nauk 26 (2) 115 (1971)]
116. Schneider W R, Wyss W J. Math. Phys. 30134 (1989)
117. Uchaikin V V, Saenko V V J. Phys. Studies 4371 (2000)
118. Uchaikin V V, Sibatov R T Tech. Phys. Lett. 30316 (2004) [Pis'ma Zh. Tekh. Fiz. 30 (8) 27 (2004)]
119. Uchaikin V V, Sibatov R T JETP 109537 (2009) [Zh. Eksp. Teor. Fiz. 136627 (2009)]
120. Uchaikin V V, Sibatov R T J. Phys. A Math. Theor. 44145501 (2011)
121. Berezhko E G, Krymskii G F Sov. Phys. Usp. 3127 (1988) [Usp. Fiz. Nauk 15449 (1988)]
122. Uchaikin V V, Cahoy D O, Sibatov R T Int. J. Bifurcat. Chaos 18 2717 (2008)
123. Lifshitz E M, Pitaevskii L P Physical Kinetics (Oxford: Pergamon Press, 1981) [Translated from Russian: Fizicheskaya Kinetika (Moscow: Nauka, 1979)]
124. Tverskoi B A Sov. Phys.JETP 25317 (1967) [Zh. Eksp. Teor. Fiz. 52 483 (1967)]
125. Tverskoi B A Sov. Phys. JETP 26821 (1968) [Zh. Eksp. Teor. Fiz. 53 1417 (1967)]
126. Parker E N Interplanetary Dynamical Processes (New York: Interscience Publ., 1963) [Translated into Russian: Dinamicheskie Protsessy v Mezhplanetnoi Srede (Moscow: Mir, 1965)]
127. Milovanov A V, Zelenyi L M Phys. Rev. E 64052101 (2001)
128. Treumann R A Phys. Scripta 59204 (1999)
129. Chechkin A V, Gonchar V Yu, Szydlowski M Phys. Plasmas 978 (2002)
130. Uchaikin V V JETP Lett. 92200 (2010) [Pis'ma Zh. Eksp. Teor. Fiz. 92226 (2010)]
131. Wandel A et al. Astrophys. J. 316676 (1987)
132. Blandford R D, Ostriker J R Astrophys. J. 221 L29 (1978)
133. Carreras B A, Lynch V E, Zaslavsky G M Phys. Plasmas 85096 (2001)
134. Uchaikin V V, Uchaikin D V Obozr. Prikl. Prom. Mat. 9477 (2002)
135. Sibatov R T, Uchaikin V V Commun. Nonlin. Sci. Numer. Simulat. 164564 (2011)
136. Zoia A, Rosso A, Kardar M Phys. Rev. E 76021116 (2007)
137. Krepysheva N, Di Pietro L, Néel M-C Phys. Rev. E 73021104 (2006)
138. Rafeiro H, Samko S Fract. Calc. Appl. Analys. 8393 (2005)
139. Guan Q-Y, Ma Z-M Stoch. Dyn. 5385 (2005)
140. Tarasov V E Ann. Physics 3232756 (2008)
141. Baleanu D et al. Int. J. Theor. Phys. 483114 (2009)
142. Tarasov V E J. Phys. A Math. Gen. 385929 (2005)
143. Sibatov R T, Uchaikin V V Phys. Usp. 521019 (2009) [Usp. Fiz. Nauk 1791079 (2009)]
144. Landau L D, Lifshitz E M Fluid Mechanics (Oxford: Pergamon Press, 1987) [Translated from Russian: Gidrodinamika (Moscow: Nauka, 1986)]
145. Bykov A M, Toptygin I N Phys. Usp. 50141 (2007) [Usp. Fiz. Nauk 177149 (2007)]
146. Alekseev B V Phys. Usp. 43601 (2000) [Usp. Fiz. Nauk 170649 (2000)]
147. Alekseev B V Phys. Usp. 46139 (2003) [Usp. Fiz. Nauk 173145 (2003)]
148. Ginzburg V L Phys. Usp. 42353 (1999) [Usp. Fiz. Nauk 169419 (1999)]
149. Résibois R, De Leener M Classical Kinetic Theory of Fluids (New York: Wiley, 1977) [Translated into Russian: Klassicheskaya Kineticheskaya Teoriya Zhidkostei i Gazov (Moscow: Mir, 1980)]

[^0]:    ${ }^{2}$ For $\beta=1$, an uncertainty occurs in the expressions for derivatives; however, refining their definition by continuity also allows including this boundary point in the considered interval of the fractional parameter. Transitions to the next unit-length intervals are performed by an additional integer-order differentiation [33].

[^1]:    ${ }^{3}$ Tauberian theorems are presented sufficiently well in vol. 2 of Feller's book [42] (see also review [25]).

[^2]:    ${ }^{4}$ That paper was written fairly long ago, and it would not be necessary to point out some of its inconsistencies if they were not repeated in other, later papers.

[^3]:    ${ }^{5}$ The type of anomalous diffusion is determined by the exponent $\gamma=\beta / \alpha$ in the expansion law of the diffusion packet $\Delta \propto t^{\gamma}$ : subdiffusion corresponds to $\gamma<1 / 2$ and superdiffusion, to $\gamma>1 / 2$. It is better to call the process with $\gamma=1 / 2, \alpha \neq 2$ a quasi-normal diffusion, because the shape of the diffusion packet in this case differs from the normal shape.

[^4]:    ${ }^{6}$ In [72], the value $\alpha=1$ is retained only for the periphery of the Galaxy, while $\alpha=0.8$ is used for the inner part. Such an approach seems dubious, because it uses propagators obtained for a homogeneous infinite medium with an unified exponent $\alpha$.

[^5]:    ${ }^{7}$ This concerns only the central part of distributions; the tails, of course, cannot be made coincident by any scale transformation.

[^6]:    ${ }^{8}$ In physics, the distribution density is often called the distribution function.

[^7]:    ${ }^{9}$ We emphasize, to avoid misunderstandings, that $B(s)$ here denotes the Brownian process rather than the magnetic induction, $H$ is the Hurst exponent rather than the field strength, and the subscript + at parentheses (see below) means that the negative values are replaced by zero inside the parentheses. The Hurst exponent, determining the law of the increase in the size of the diffusion packet $\propto s^{H}$, is simultaneously the fractal dimension of the trajectory of the anomalous process.

[^8]:    ${ }^{10}$ We note that the Boltzmann concept of collisions, according to which a trajectory is divided into ranges alternating with instant changes in the motion direction, has found its way into the result when transforming the collisionless Vlasov equation into the diffusion form, even though this concept was not postulated initially: the diffusion coefficient $\kappa_{\|}$presented above can well be interpreted as $v\langle R\rangle / 3$, where $\langle R\rangle$ is already the path length between collisions.

[^9]:    ${ }^{11}$ If the first moment (mean time) of the correlation function is finite [22]. ${ }^{12}$ We temporarily retain the parameter $\beta \in(0,1)$, although only $\beta=1 / 2$ is required for describing the transverse diffusion.

[^10]:    ${ }^{13}$ For this reason, we here retain the notation $\psi$ for the solution, used in [127]: the distribution density is denoted by $f$ in our review.

