

From Scatter-Free to Diffusive Propagation of Energetic Particles with Exact Solution of Fokker-Planck Equation

M.A. Malkov*

University of California, San Diego

E-mail: mmalkov@ucsd.edu

Propagation of energetic particles through magnetized turbulent media is reconsidered using the exact solution of Fokker-Planck equation. This solution reveals that our understanding of cosmic ray (CR) transport is inaccurate when it relies on a diffusive approximation for weakly scattered energetic particles. The poor understanding of CR transport obscures their sources and acceleration mechanisms. After the Fermi, PAMELA, and AMS-02 have discovered the electron/positron and p/He spectral anomalies, it becomes crucial to improve transport models for improving our understanding of the anomalies. We discuss and simplify the exact solution of Fokker-Planck equation [1] which accurately describes a ballistic, diffusive and transdiffusive (intermediate between the first two) propagation regimes. It is found that the transdiffusive phase lasts for a (surprisingly) long time, about five scattering times ($5t_c$) while starting as early as at $0.5t_c$. Since the scattering time of CRs is energy dependent, $t_c = t_c(E)$, a significant part of their spectrum propagates transdiffusively, thus requiring the exact solution of Fokker-Planck equation. This limitation is particularly relevant to the heliospheric modulation of galactic CRs. We present a new, simplified version of an exact Fokker-Planck propagator. It can easily be employed in place of the Gaussian propagator, currently used in major solar modulation and other CR transport models. Flaws in using the telegraph equation to model the CR propagation before the onset of diffusive regime are also briefly discussed.

35th International Cosmic Ray Conference — ICRC2017

10–20 July, 2017

Bexco, Busan, Korea

*Speaker.

1. Lacuna in CR Transport Model

The cosmic ray (CR) propagation in turbulent environments, such as the interstellar medium (ISM) or Heliosphere, has been actively researched for a long time [2]. Time asymptotically, CRs propagate diffusively; after several collisions, they “forget” their initial velocities and enter a random walk process. However, in astrophysical objects with infrequent particle collisions, there may not be enough time or room for even a few collisions. In such systems, the focus shifts to earlier propagation phases, which are better described as ballistic rather than diffusive propagation. The question is, what is in between these two regimes and how long it lasts?

The transition from ballistic to diffusive transport regime has always been a challenge for the theory. At the same time, it is often the key to understanding the CR sources. Since the particle mean free path (m.f.p) usually grows with energy, some part of their spectrum almost inescapably falls into a transient category where neither ballistic nor diffusive approximation applies. We will call this regime transdiffusive and demonstrate that it lasts for long enough to compromise both the ballistic and diffusive model predictions. During this propagation phase, CR protons accelerated in supernova remnants (SNR), for example, may reach a nearby molecular cloud, making themselves visible by interacting with its dense gas. The CR protons of lower energies would instead be diffusively confined to the SNR shell and evade detection. Another example is the propagation of solar energetic particles to 1 AU. Also, in this case, the m.f.p. is comparable to, or even exceeds, 1 AU, so neither the diffusive nor ballistic approximation applies.

Galactic CRs ultimately propagating through the Heliosphere to the observer cannot always be propagated back to their source within simple diffusion or ballistic paradigms, so their spectra cannot be fully understood. This problem is particularly relevant to striking anomalies in the CR spectra and composition, which are becoming a general trend in the CR observations. Besides the e^+/e^- anomaly, there is a ~ 0.1 difference in rigidity indices of proton and He. Although the explanations are available (see, e.g., [3, 4], and the paper by Hanusch et al. in the Proceedings), the low-energy parts of these spectra are strongly affected by the solar modulation. Curvature and gradient drifts in the Heliospheric magnetic field are mostly treated by considering particle propagation along the field line as diffusive which we will show to be inaccurate for sufficiently energetic particles with long m.f.p.

2. Governing Equation

The Fokker-Planck (FP) equation is a minimalist model suitable for the CR transport. An ambient magnetic field justifies a 1D treatment, while its fluctuating part supports the particle scattering in pitch angle. The simplest FP equation for the CR distribution f reads:

$$\frac{\partial f}{\partial t} + v\mu \frac{\partial f}{\partial x} = \frac{\partial}{\partial \mu} (1 - \mu^2) D(\mu, E) \frac{\partial f}{\partial \mu}. \quad (2.1)$$

Here x is directed along the local magnetic field, μ is the cosine of the particle pitch angle, v, E are the particle velocity and energy, conserved in interactions with quasi-static magnetic turbulence. D is the scattering rate (collision frequency).

One propagation scenario that Eq.(2.1) describes very well comes about through an instant release of a cloud of particles into a scattering medium. Again, Galactic SNRs, widely believed

to generate CRs with energies up to $\sim 10^{15}$ eV, must accelerate them in SNR shock waves with a subsequent release into a turbulent ISM. The question then is how exactly the particle density (the isotropic component of f) propagates along a magnetic flux tube that intersects the SNR shell. Our goal is to achieve the simplicity of diffusive description [2] which is a well-known derivative of Eq.(2.1). As emphasized earlier, the diffusive treatment is inadequate in the preceding ballistic and transdiffusive propagation phases, while the latter is often the key for probing into the source.

2.1 Restricting Propagation Models by Limiting Cases

Because of the difficulties in reducing the FP equation to a manageable isotropic form, it is useful to set a framework for such reduction by the two limiting cases of ballistic and diffusive propagation. We derive both regimes directly from Eq.(2.1), by eliminating angular dynamics.

In the ballistic case, which strictly applies to times shorter than the collision time $t \ll t_c \sim 1/D$, one can neglect the r.h.s. altogether. The solution then follows from an integration along the particle trajectories, $x - \mu vt = \text{const}$ (Liouville's theorem), with a conserved pitch angle, $\mu = \text{const}$. The solution is simply $f(x, \mu, t) = f(x - v\mu t, \mu, 0)$. Now, consider an isotropic point source: $f(x, \mu, 0) = \delta(x) \Theta(1 - \mu^2)/2$, where δ and Θ denote the Dirac's delta and Heaviside unit step functions, respectively. From the above solution for $f(x, \mu, t)$, one obtains the ballistic expansion in form of the second moment, $\langle x^2 \rangle = v^2 t^2/3$ by integrating $x^2 f = x^2 \delta(x - v\mu t) \Theta(1 - \mu^2)/2$ over x and μ . The result describes a free escape with the mean square velocity $v/\sqrt{3}$, while the maximum particle velocity (along x) is v . The pitch angle averaged particle distribution, $f_0(x, t) = (2vt)^{-1} \Theta(1 - x^2/v^2 t^2)$, is best described as an expanding 'box' of decreasing height.

Among earlier attempts to reduce f to its pitch angle-averaged part, $f(t, x, \mu) \rightarrow f_0(t, x)$, an approach leading to a "telegraph" equation, can be readily tested using the above box solution. We will briefly discuss this approach later and show that it is inconsistent with the ballistic limit of f_0 obtained directly from the FP equation. Needless to say that the exact solution of Eq.(2.1), presented further in this paper, converges to the above-described box distribution at $t \ll t_c$.

The second, well studied propagation regime is diffusive. It dominates at $t \gg t_c \sim 1/D$ and is treated in a way opposite to the above-described ballistic regime, [2]. The r.h.s. of Eq.(2.1) is now the leading term, thus implying that the particle distribution is close to isotropy, $\partial f/\partial \mu \rightarrow 0$. Working to higher orders in anisotropic corrections $\sim 1/D$, and averaging the equation over μ , one obtains the following infinite hierarchy of equations for $f_0(x, t)$ [5]

$$\frac{\partial f_0}{\partial t} - \kappa_2 \frac{\partial^2 f_0}{\partial x^2} = -\kappa_4 \frac{\partial^4 f_0}{\partial x^4} + \kappa_6 \frac{\partial^6 f_0}{\partial x^6} - \dots, \quad (2.2)$$

with $\kappa_{2n} \sim 1/D^n$. This particular hierarchy results from an asymptotic (Chapman-Enskog) expansion of the problem in $1/D$ under the scattering symmetry: $D(-\mu) = D(\mu)$. It is valid only for $t \gg t_c \sim 1/D$, and all the short-time-scale, ballistic propagation effects are intentionally eliminated (cf. elimination of secular terms in perturbative treatments). A failure to do so results in a second order time derivative in Eq.(2.2) (already mentioned telegraph term) which is illegitimate unless $t \gg t_c$. Nevertheless, the telegraph equation has been put forward over since the 1960s as a viable tool for describing the CR propagation from the ballistic to diffusive phases.

Meanwhile, the r.h.s. of eq.(2.2) provides small hyperdiffusive corrections negligible for $t > t_c$, as the higher spatial derivatives quickly decay because of the smoothing effect from the diffusive term on its l.h.s. They do not shed much light on the ballistic and transdiffusive propagation regimes, unless the series is summed up with no truncation. The latter requirement follows from an exact solution of the parent FP equation [1]. It contains an *infinite* series of moments $\langle x^{2n} f_0(x,t) \rangle$ that are evidently connected with the infinite series of coefficients $\{\kappa_{2n}\}$ in eq.(2.2). Conversely, by including only a few hyperdiffusive correction terms outside of their validity range, $t \gg t_c$, one may worsen the diffusive approximation. It can also be shown [5] that within its validity range, a truncated version of eq.(2.2), with $\kappa_{2n} = 0$ for $n > 2$, can be mapped onto the telegraph equation. It follows that neither a *truncated* hyperdiffusive approach nor the telegraph equation (a subset of the former) cannot adequately reproduce the FP solution at times shorter than $t \gg t_c$. This was recently demonstrated in [6], by a numerical integration of Eq.(2.1). We will quantify the constraint $t \gg t_c$, repeatedly stressed above, by comparing the full FP solution with its diffusive limit (see [1] for more details).

The primary failure of the diffusive approach is an unrealistically fast (acausal) propagation, which is especially pronounced during the ballistic and transdiffusive phases. Mathematically, the approximation violates an upper bound $|x| \leq vt$ that immediately follows from Eq.(2.1) for a point source solution, discussed above. There have been attempts to overcome this problem, but no adequate *ab initio* description of particle spreading that would cover ballistic and diffusive phases was elaborated. The most persistent such attempt is based on the telegraph equation discussed above. It has a misleading impact on the field of CR propagation for that simple reason

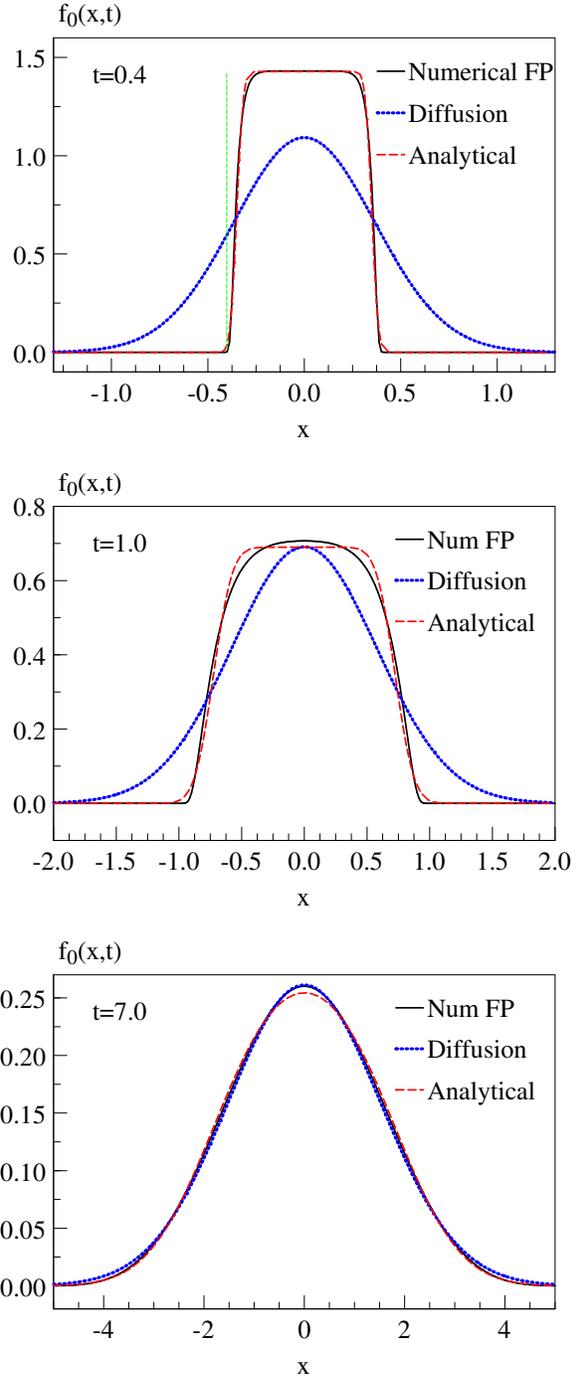


Figure 1: Fundamental solution of the Fokker-Planck equation shown for its isotropic component, $f_0(x,t) = \langle f(x,\mu,t) \rangle$ at $t = 0.4, 1.0, 7.0$. Analytic approximation is from Eq.(4.1), diffusive (Gaussian) solution from Eq.(4.3), numerical - from the FP eq.(3.1). Vertical green line in the upper panel shows the width of the front.

that the solution of this equation is inconsistent with its parent FP equation. We established this simple fact by considering the ballistic propagation phase directly from eq.(2.1) (see [5, 1] for more discussion).

It follows that there is no viable analytical tool to address the earlier phases of particle propagation, other than to either sum up the series of hyperdiffusive terms or to solve the FP equation directly. Below, we take the second option.

3. Exact Solution of FP equation

The energy dependence of the particle scattering frequency enters eq.(2.1) only as a *parameter*, i.e., $D(E)$. The possible pitch-angle dependence of D typically scales as $D(\mu) \propto |\mu|^{q-1}$ [2], thus being suppressed in an important case $q = 1$, where q is the power-law index of magnetic turbulence. Under these, quite realistic assumptions, the FP equation can be solved exactly [1]. To describe this solution, it is convenient to rewrite Eq.(2.1) using dimensionless time and length units according to the following transformations $D(E)t \rightarrow t$, $Dx/v \rightarrow x$. Instead of Eq.(2.1) we then have

$$\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} = \frac{\partial}{\partial \mu} (1 - \mu^2) \frac{\partial f}{\partial \mu} \quad (3.1)$$

This equation contains no parameters, thus precluding any direct, small-parameter expansion unless such a parameter implicitly enters the problem through an initial condition, $f(x, \mu, 0)$. In particular, if one is using Eq.(2.2) ($1/D$ - type expansion), not only should the initial distribution be close to isotropy, but it should also be spatially broad. The latter condition will prevent a high anisotropy from arising via the second term on the l.h.s. of Eq.(3.1). Hence, the problem of a point source spreading (Green's function, or fundamental solution) can not be treated using conventional $1/D$ expansion, until f becomes quasi-isotropic, that is broadened to $x \gtrsim 1$.

The exact solution of Eq.(3.1) can be obtained by resolving an infinite chain of equations for the moments of $f(\mu, x)$

$$M_{ij}(t) = \langle \mu^i x^j \rangle = \int_{-\infty}^{\infty} dx \int_{-1}^1 \mu^i x^j f d\mu / 2 \quad (3.2)$$

for all integer $i, j \geq 0$. The lowest moment M_{00} is automatically conserved by Eq.(3.1) (as being proportional to the number of particles) and we normalize it to unity, $M_{00} = 1$. All the higher moments can be explicitly obtained from the following recurrence relation [1]

$$M_{ij}(t) = M_{ij}(0) e^{-i(i+1)t} + \int_0^t e^{i(i+1)(t'-t)} [j M_{i+1, j-1}(t') + i(i-1) M_{i-2, j}(t')] dt' \quad (3.3)$$

Focusing on a point source (fundamental) solution, we assume the initial distribution $f(x, \mu, 0)$ to be symmetric in x . The initial spatial width must then be set to zero, $M_{02}(0) = \langle x^2 \rangle_0 = 0$. Assuming $f(x, \mu, 0)$ to be isotropic in μ for simplicity, nulls the odd moments, while the even moments become $M_{i0} = 1/(i+1)$. From the above solution with $j = 0$, for the leftmost column of the matrix $M_{ij}(t)$ we readily obtain $M_{i0}(t) = M_{i0}(0) = \delta_{i0}/(i+1)$. Starting from this column, eq.(3.3) can be recursively resolved for all $i, j \geq 0$ by using the chain rule for $M_{ij} : i \rightarrow i-1, j \rightarrow j+1$.

From the mathematical point of view, only a *full* set of moments in Eq.(3.3) provides a complete solution $f(x, \mu, t)$ of Eq.(3.1) given the initial value, $f(x, \mu, 0)$, that, in turn, determines the matrix $M_{ij}(0)$ in Eq.(3.3). To adequately reproduce the ballistic and transdiffusive phases the series of moments cannot be truncated. Considering the fundamental solution, we will focus on the isotropic part of particle distribution

$$f_0(x, t) = \int_{-1}^1 f(\mu, x, t) d\mu/2, \quad (3.4)$$

as only this part contributes to the particle number density. The fundamental solution to be obtained requires the initial condition $f_0(x, 0) = \delta(x)$. The matrix elements that represent f_0 are M_{0j} which we denote simply as $M_j \equiv M_{0j}$. Note, that $M_{ij}(i)$ with $i > 0$ are not small and absolutely essential for calculating the full set of the moments M_j . To link the set $M_j(t)$ to $f_0(x, t)$, we use the moment-generating function

$$f_\lambda(t) = \int_{-\infty}^{\infty} f_0(x, t) e^{\lambda x} dx = \sum_{n=0}^{\infty} \frac{\lambda^{2n}}{(2n)!} M_{2n}(t) \quad (3.5)$$

where we omitted the odd moments irrelevant to the fundamental (symmetric in x) solution. The above expansion may be cast in a familiar Fourier transform of $f_0(x, t)$ by setting $\lambda = -ik$.

Since expressions for the moments M_{2n} are becoming cumbersome with growing n , an exact form of the Green's function $f_0(x, t)$, which can be recovered from eq.(3.5) by inverting the Fourier integral, is also not simple, to say the least. Therefore, in the next section, we derive a new simplified version of the exact FP propagator.

4. Two-moment Fokker-Plank Propagator

The series in the moment generating function f_λ in eq.(3.5) can be evaluated for small and large values of t and λt . Although there are four limiting cases, all the expressions for the sum $f_\lambda(t)$ are surprisingly similar. They can be unified under a single *approximate* (but *valid* for all x and t) expression for f_0 :

$$f_0(x, t) \approx \frac{1}{4y} \left[\operatorname{erf} \left(\frac{x+y}{\Delta} \right) - \operatorname{erf} \left(\frac{x-y}{\Delta} \right) \right]. \quad (4.1)$$

It has been obtained in [1] from an inverse Fourier transform, $f_\lambda(t) \mapsto f_0(x, t)$, after summing up the series for f_λ in eq.(3.5). The two independent functions of time, $y(t)$ and $\Delta(t)$ can be expressed through the moment $M_2(t)$, which we calculate exactly from eq.(3.3). The solution f_0 with y and Δ so obtained compares very well with the numerical FP solution. The disadvantage of this single-moment representation of y and Δ is that it requires some (fairly minor, though) changes in $y(M_2)$ and $\Delta(M_2)$, between the cases $t \lesssim 1$ and $t \gtrsim 1$.

Here we obtain a uniformly valid expansion for f_0 as follows. As the general form of the approximate solution given in eq.(4.1) must arguably be the same for all t [1], we find the functions $y(t)$ and $\Delta(t)$ by requiring that $f_0(x, t)$ generates the moments M_2 and M_4 exactly:

$$M_2 = \int x^2 f_0(x, t) dx, \quad M_4 = \int x^4 f_0(x, t) dx \quad (4.2)$$

Note, that all odd moments vanish automatically for f_0 given by eq.(4.1). Recall, that we know exact values for all moments M_{2n} from eq.(3.3). Here, we will only use M_2 and M_4 , which satisfy the initial conditions, $M_2(0) = 0$ and $M_4(0) = 0$:

$$M_2 = \frac{t}{3} - \frac{1}{6}(1 - e^{-2t}), \quad M_4 = \frac{1}{270}e^{-6t} - \frac{t+2}{5}e^{-2t} + \frac{1}{3}t^2 - \frac{26}{45}t + \frac{107}{270}$$

Substituting f_0 from eq.(4.1) into eqs.(4.2), we find

$$y = \left[\frac{45}{2} \left(M_2^2 - \frac{1}{3}M_4 \right) \right]^{1/4}, \quad \Delta = \sqrt{2M_2 - \sqrt{10} \sqrt{M_2^2 - \frac{1}{3}M_4}}$$

The FP solution, eq.(4.1), is not more difficult than the familiar diffusive solution. Ignoring the time dependence of y in the error functions, the FP solution corresponds to a conventional diffusion problem with an initial particle density evenly distributed between $-y < x < y$, and zero otherwise. The difference is only in the form of $y(t)$ and $\Delta(t)$. At $t \ll 1$ this solution is the same as the ballistic one described in Sec. 2.1. Indeed, since $\Delta \propto t^2$ and $y \approx t$ for $t \ll 1$, the difference of the two error functions yields $2\Theta(1 - x^2/t^2)$ and f_0 in eq.(4.1) reproduces the expanding box solution. Note that the moments $M_n(t)$, being nonlocal, are not useful observationally but they define the above solution for $f(x, t)$. The telegraph solution, on the contrary, is inconsistent with this expansion regime. It contains two (nonexistent in the FP solution) singular components at the two propagating fronts, let alone inaccurate front positions and the overall profile [6].

The width of the propagating fronts at $x = \pm y$, determined by $\Delta(t)$, behaves as follows, Fig.1. At small $t \ll 1$, when the box is expanding ballistically, i.e. $y \approx t$, the wall thickness $\Delta \approx 2t^2/3\sqrt{5}$. After gradually proceeding through the transdiffusive phase, these quantities become $y \approx (11t/6)^{1/4}$ and $\Delta \approx (2t/3)^{1/2}$ for $t \gg 1$. Accordingly, the expression in eq.(4.1) converges (rather slowly, though) to:

$$f_0(x, t) = \sqrt{\frac{3}{2\pi t}} e^{-3x^2/2t} \quad (4.3)$$

which is the diffusive asymptotic solution of the pitch angle averaged FP equation, given by eq.(2.2) with $\kappa_2 = 1/6$ and all the hyperdiffusive coefficients $\kappa_{2n} = 0$ for $n > 1$.

Having at hand the Green's function for eq.(3.1), it is not difficult to obtain the solution of its inhomogeneous version. Adding to the r.h.s. of this equation the source function $S(x, t)$, the pitch-angle averaged solution can be expressed as follows

$$\langle f(x, t, \mu) \rangle = \int_0^t dt' \int_{-\infty}^{\infty} f_0(x - x', t - t') S(x', t') dt' \quad (4.4)$$

For the propagator f_0 here, either the simplified version, eq.(4.1), or the exact solution obtained by a Fourier inversion of eq.(3.5) [1] can be used. The source function is assumed to be isotropic here. This restriction can be relaxed by changing the initial conditions for the solution eq.(3.3) to $M_{i0}(t=0) = \mu^i/2$ and adding the required integral over μ' to eq.(4.4). The resulting propagator, however, will be no longer be simple.

Summarizing this section, the two-moment single formula representation of the FP solution in eq.(4.1) has correct asymptotic limits at $t \rightarrow 0, \infty$, both obtained independently of eq.(4.1). The deviations from the numerical solution at $t \sim 1$ are minor and more than compensated by the simplicity of eq.(4.1) and its validity for all $0 < t < \infty$.

5. Conclusions

The exact solution of FP equation obtained in [1] is simplified. The new propagator accurately evolves the isotropic part of the distribution $f_0(x, t)$, uniformly in $-\infty < x < \infty$ and $0 \leq t < \infty$.

The overall CR propagation categorizes into three phases: ballistic ($t \ll 1$), transdiffusive ($t \sim 1$) and diffusive ($t \gg 1$), (time in units of collision time t_c). In the ballistic phase, the source expands as a “box” of size $\Delta x \propto \sqrt{\langle x^2 \rangle} \propto t$ with thickening “walls” at $x = \pm y(t) \approx \pm t$ of the width $\Delta \propto t^2$. During the transdiffusive phase, the box’s walls thicken to a sizable fraction of the box $\Delta \sim \Delta x \sim y$ and the expansion slows down, Fig. 1. Finally, during the conventional diffusion phase, $\Delta x \sim \Delta \propto \sqrt{t}$, while the walls are completely smeared out, as $y \propto t^{1/4}$, so $y \ll \Delta$.

In constraining earlier FP-based models for the CR propagation, the exact FP solution reveals the following:

- the conventional diffusion approximation can be safely applied but, only after 5-7 collision times, depending on the accuracy requirements
- a popular telegraph approach, originally intended to cover also the earlier propagation phases at $t \lesssim 1$, is inconsistent with the exact FP solution (see also [1])
- no sub/super-diffusive propagation regime is present in the exact FP solution

The latter regimes are occasionally postulated, e.g., in studies of diffusive shock acceleration (DSA), in the form of a power-law dependence of particle dispersion $\sqrt{\langle x^2 \rangle} \propto t^\alpha$, with $1/2 < \alpha < 1$ (superdiffusion) or $0 < \alpha < 1/2$ (subdiffusion). The exact FP propagation leads to $\sqrt{\langle x^2 \rangle}$ that smoothly changes from the ballistic ($\alpha \rightarrow 1$) to diffusive ($\alpha \rightarrow 1/2$) propagation with no dwelling at any particular value of α between these limits.

Acknowledgments This work was supported by the NASA Astrophysics Theory Program under Grant No. NNX14AH36G.

References

- [1] M. Malkov, *Exact solution of the Fokker-Planck equation for isotropic scattering*, *Physical Review D* **95** (2017) 023007 [<https://arxiv.org/abs/1610.01584>].
- [2] J. R. Jokipii, *Propagation of cosmic rays in the solar wind.*, *Reviews of Geophysics and Space Physics* **9** (1971) 27–87.
- [3] M. A. Malkov, P. H. Diamond and R. Z. Sagdeev, *Proton-Helium Spectral Anomaly as a Signature of Cosmic Ray Accelerator*, *Physical Review Letters* **108** (#feb#, 2012) 081104, [<https://arxiv.org/abs/1110.5335>].
- [4] M. Malkov, P. Diamond and R. Sagdeev, *Positive charge prevalence in cosmic rays: Room for dark matter in the positron spectrum*, *Physical Review D* **94** (2016) 063006 [<https://arxiv.org/abs/1607.01820>].
- [5] M. A. Malkov and R. Z. Sagdeev, *Cosmic Ray Transport with Magnetic Focusing and the “Telegraph” Model*, *Astrophys. J.* **808** (#aug#, 2015) 157, [<https://arxiv.org/abs/1502.01799>].
- [6] Y. E. Litvinenko and P. L. Noble, *Comparison of the telegraph and hyperdiffusion approximations in cosmic-ray transport*, *Physics of Plasmas* **23** (June, 2016) 062901.