Lecture 1 From random walks to fractional diffusion equations

Dr Qianqian Yang

AMSI Winter School 2019

Fractional calculus is a rapidly growing field of research, at the interface between probability, differential equations, and mathematical physics. Fractional calculus is used to model anomalous diffusion, in which a cloud of particles spreads in a different manner than traditional diffusion. In our first lecture, we will derive the space fractional diffusion equations from the point of view of probability, using random walks.

Case study 1: Random walks with constant jump sizes

The traditional diffusion model in one dimension space

$$\partial u(x,t)/\partial t = -D(-\Delta)u(x,t)$$

can be derived in many ways. We will be interested in its derivation using random walks, since that approach generalises nicely to the fractional diffusion case. Notice that we have written the equation using the operator $-\Delta$ (which can also be written as $-\nabla^2$). The reason for the minus sign is that $-\Delta$ is then a *positive definite* operator, so its eigenvalues are positive. In one dimension this operator is just $-\Delta = -\frac{\partial^2}{\partial x^2}$. We will derive the equation in one dimension only, but the idea generalises to any number of space dimensions.

Consider a large number of particles all undergoing independent random walks. For simplicity, we suppose that they all begin at the origin $x = 0$ at time $t = 0$. Then at each discrete time step Δt, they randomly jump one unit left, $-\Delta x$ or one unit right, Δx.

Let $u(x,t)$ denote the *density* of particles at position x at some time t: the number of particles per unit distance.

At the next time $t + \Delta t$, *all* of these particles at position x will jump away. And since we're dealing with a large number of particles, we can say that *half* of the particles from the left and *half* of the particles from the right will each jump here. So the new *number* of particles will be

$$u(x, t+\Delta t)\Delta x = \frac{1}{2}u(x-\Delta x, t)\Delta x + \frac{1}{2}u(x+\Delta x, t)\Delta x$$

Clearly we can cancel the factors of Δx and just talk about particle *density:*

$$u(x, t+\Delta t) = \frac{1}{2}u(x-\Delta x, t) + \frac{1}{2}u(x+\Delta x, t)$$

Taking Taylor expansions we find that

$$\frac{\partial u(x,t)}{\partial t} = \frac{\Delta x^2}{2\Delta t}\frac{\partial^2 u(x,t)}{\partial x^2} + O(\Delta t) + \frac{O(\Delta x^3)}{\Delta t}.$$

Now we take the limit as the stepsize Δx and time between jumps Δt go to zero, keeping the factor $D = \frac{\Delta x^2}{2\Delta t}$ fixed. We derive the diffusion equation

$$\frac{\partial u(x,t)}{\partial t} = -D\left(-\frac{\partial^2}{\partial x^2}\right)u(x,t) = -D(-\Delta)u(x,t)$$

where D is called the diffusivity, or diffusion coefficient. In nature, there is never really the *limit* as Δx and Δt go to zero. Instead, the diffusion coefficient is whatever it is, for the small but finite values of Δx and Δt that govern the diffusion (e.g. molecular scale).

The initial condition that all the particles begin at the origin can be written as $u(x,0) = \delta(x)$ (Dirac delta function), where for convenience we can measure the density in units such as "billions of particles per millimetre" (or whatever), so that that the initial *number* of particles is just

$$\int_{-\infty}^{\infty} u(x,0)dx = \int_{-\infty}^{\infty} \delta(x)dx = 1 \text{ (billion particles)}$$

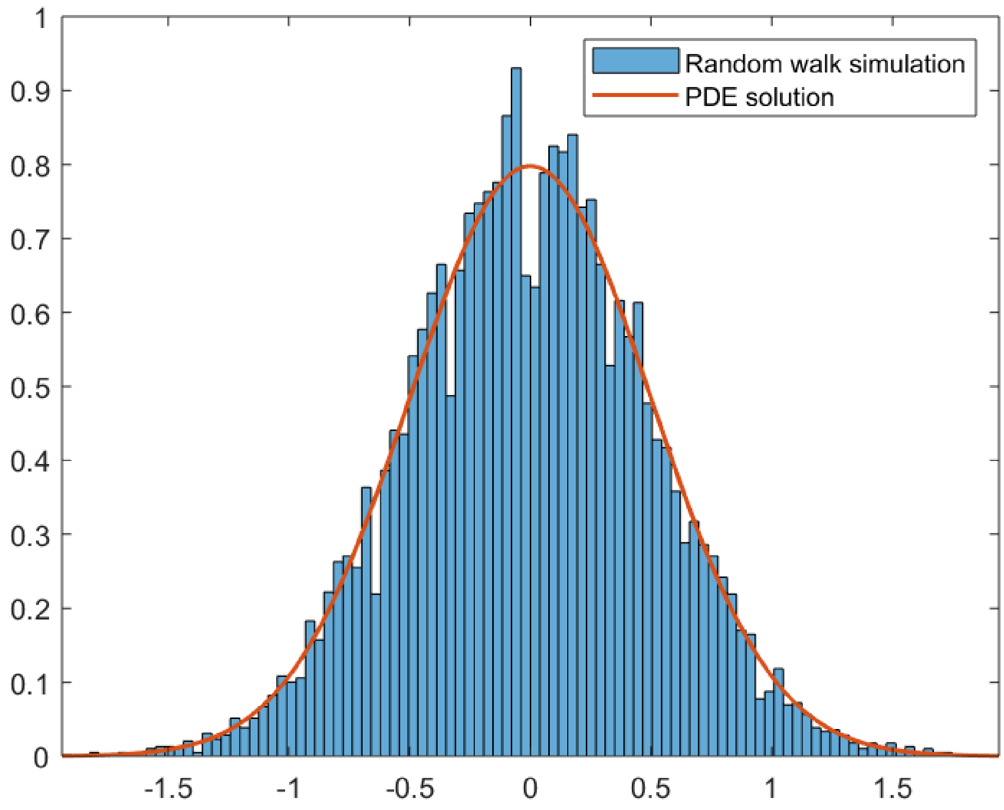
With this initial condition, the diffusion equation has the solution

$$u(x,t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right)$$

You might recognise this as the equation for the *normal distribution*, also called *Gaussian distribution* probability density function. That's no coincidence! The particles' positions at time t are the sum of all the independent random jumps up to that time. From statistics, we know that the sum of many independent random variables approaches the Gaussian distribution. This is called the Central Limit Theorem.

Let's do a simulation! The code is provided in the file random_walk_simulation.m. This simulation uses 10^4 particles each taking 10^4 steps of size $\Delta x = 5 \times 10^{-3}$. The final time t is arbitrarily set to $t = 1$, meaning the effective timestep is $\Delta t = 10^{-4}$ and so the diffusivity for the PDE model is $D = \Delta x^2/(2\Delta t) = 0.125$. Let's compare the histogram of particle locations after 10^4 steps with the prediction of the continuum PDE model.

```
random_walk_simulation
```




```
D = 0.1250
```

The solutions are in good agreement. Better agreement would be obtained with more particles (thereby smoothing over the random bumps) and more steps (meaning the random walk was closer to the continuum limit with Δx and Δt going to zero).

Case study 2: Random walks with jump sizes drawn from a symmetric distribution with *finite* variance

We might wonder how general this diffusion equation is. Instead of assuming the jumps are all constant size Δx, could we have drawn each jump size from a probability distribution? Let the random variable X denote the jump size now, and suppose it comes from a probability distribution with probability density function $q(X)$. We suppose that this is a *symmetric* distribution around zero, so the jump size can be positive or negative with equal probability. Thus, the mean μ of the distribution is

$$E[X] = \int_{-\infty}^{\infty} x\, q(x) dx = \mu = 0$$

and we suppose its variance is some finite value σ^2

$$var[X] = E[(X-\mu)^2] = E[X^2] = \int_{-\infty}^{\infty} x^2 q(x) dx = \sigma^2.$$

Now, previously we had the equation

$$u(x,t+\Delta t) = \frac{1}{2}u(x-\Delta x,t) + \frac{1}{2}u(x+\Delta x,t)$$

to update the density at time $t+\Delta t$. But that assumed the only possible jump sizes were $-\Delta x$ and Δx. Now, in theory *any jump size is possible* (though, very large jumps are less likely because of finite variance). So we have to consider all the ways that particles could jump to location x at time $t+\Delta t$. We generalise the two terms in the sum (each with probability $1/2$) to a continuous range of possibilities, each with their own probability given by q. Note that to jump from some arbitrary position x' to position x requires a jump size of $x-x'$. So we have,

$$u(x,t+\Delta t) = \int_{-\infty}^{\infty} u(x',t)\,q(x-x')dx'$$

How to proceed from here? We recognise the integral on the right hand side as the *convolution of* u and q. Any time you see convolutions, there's a good chance that Fourier transforms can help, because the Fourier transform of a convolution is the *product* of the individual Fourier transforms.

So take Fourier transforms to obtain

$$\hat{u}(k,t+\Delta t) = \hat{u}(k,t)\,\hat{q}(k)$$

and we will use the convention from statistics in defining the Fourier transform to be

$$\hat{f}(k) = \int_{-\infty}^{\infty} e^{ikx} f(x)dx.$$

(In statistics, if f is a probability density function, then $\hat{f}$ is known as the characteristic function*.)*

We can now expand q in terms of its *moments*:

$$\hat{u}(k,t+\Delta t) = \hat{u}(k,t)\left[\int_{-\infty}^{\infty} q(x)dx + ik\int_{-\infty}^{\infty} x\,q(x)dx - \frac{k^2}{2}\int_{-\infty}^{\infty} x^2 q(x)dx + \ldots\right]$$

Expanding in a Taylor series on the left as before, we find that

$$\frac{d\hat{u}(k,t)}{dt} = -\frac{\sigma^2}{2\Delta t}k^2\hat{u}(k,t) + O(\Delta t) + \frac{o(\sigma^2)}{\Delta t}.$$

Remember the big O notation $O(\Delta t)$ means that term goes to zero at least as fast as Δt, while the 'little o notation' $o(\sigma^2)$ means that term goes to zero *faster* than σ^2.

So, letting $D = \frac{\sigma^2}{2\Delta t}$ and taking limits as Δt and σ go to zero, keeping D fixed, we have

$$\frac{d\hat{u}(k,t)}{dt} = -Dk^2\hat{u}(k,t)$$

With initial condition $\hat{u}(k,0) = 1$ (corresponding to $u(x,0) = \delta(x)$) we obtain

$$\hat{u}(k,t) = \exp(-Dk^2 t)$$

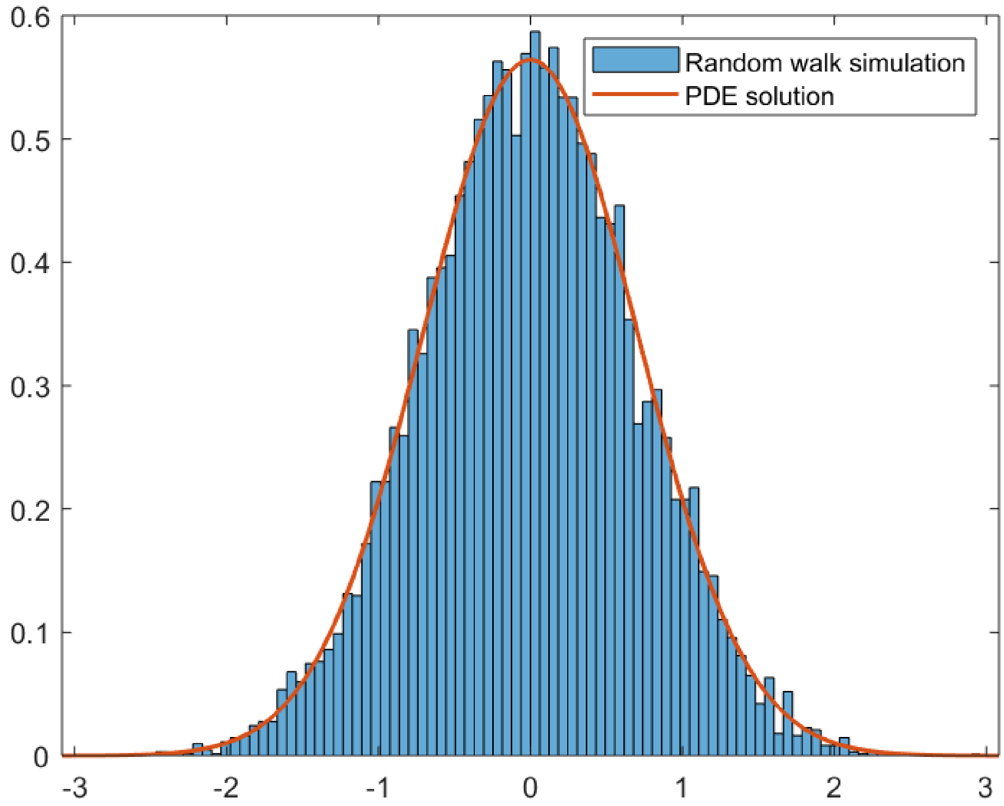
and by Fourier inversion

$$u(x,t) = \frac{1}{\sqrt{4\pi Dt}}\exp\left(-\frac{x^2}{4Dt}\right)$$

as before. But now we see that the result is true much more generally than for constant-size jumps, which was just a special case with $\sigma^2 = \Delta x^2$. Instead, we can choose any symmetric distribution of jump sizes with variance σ^2. This includes distributions for which there is considerable probability for taking very long jumps.

Let's do another simulation! This time we'll use a two-sided exponential distribution for our step sizes, with *mean absolute* stepsize 5×10^{-3}. Then the *variance* for the two-sided distribution is $\sigma^2 = 5 \times 10^{-5}$. The other parameters are as before, so the diffusivity for the PDE model is now $D = \sigma^2/(2\Delta t) = 0.25$. Although the mean absolute step size is the same as previously, it has more variance (large steps are possible), so the overall diffusivity is twice as large. Let's again compare the histogram of particle locations after 10^4 steps with the prediction of the continuum PDE model.

```
random_walk_simulation('exponential')
```




```
D = 0.2500
```

The solutions are again in good agreement, with more spreading of the particles compared to previously.

Case study 3: Random walks with jump sizes drawn from a symmetric distribution with *infinite* variance

You might wonder then, how would we ever obtain *anomalous diffusion*, where the particle density doesn't follow a Gaussian distribution? Doesn't the Central Limit Theorem guarantee that the sum of many jumps from any distribution will approach a Gaussian? The catch is in the fine print of the theorem. It only applies to jump size distributions with *finite variance* σ^2.

If the particle jump sizes are drawn from a probability distribution with *infinite variance*, then the Central Limit Theorem doesn't apply. The prototype example of such a distribution is the Pareto distribution which has probability density function

$$q(x) = \frac{\alpha x_m{}^{\alpha}}{2|x|^{\alpha+1}}, \qquad |x| > x_m,$$

(This is the two-sided form of the Pareto distribution, which allows for positive and negative jumps with equal probaiblity.)

The parameter α is the key here. It determines the "heaviness" of the tails. We will be concerned with the range $1 < \alpha < 2$. For this range, the mean of this distribution is equal to zero, but the variance is infinite.

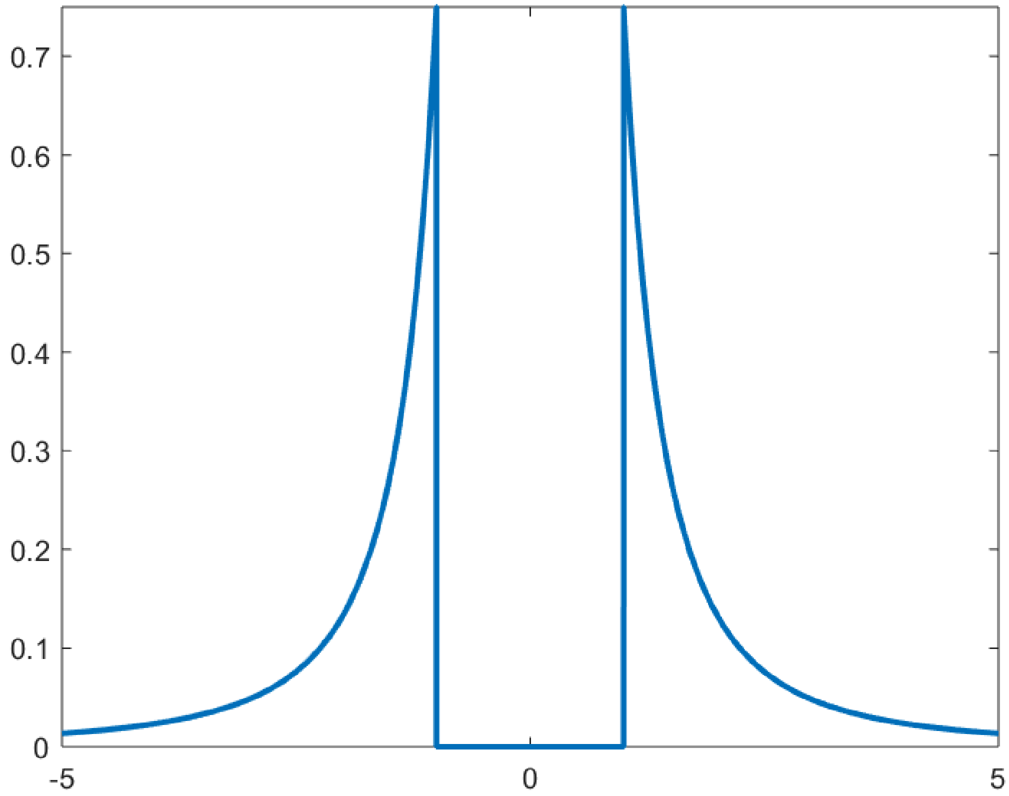
Let's use MATLAB to check these properties of the Pareto distribution.

```
syms x alpha x_m
assume(alpha > 1 & alpha < 2)
assume(x_m > 0)
pareto = piecewise(abs(x)>=x_m, alpha*x_m^alpha / 2 / abs(x)^(alpha+1), 0)
```

pareto =

$$\begin{cases} \frac{\alpha x_m{}^{\alpha}}{2\,|x|^{\alpha+1}} & \text{if } x_m \le |x| \\ 0 & \text{otherwise} \end{cases}$$

```
figure;
fplot(subs(pareto, {alpha, x_m}, {1.5, 1}), [-5,5], 'LineWidth', 2)
```




```
int(pareto, x, -inf, inf)
```

Warning: Unable to check whether the integrand exists everywhere on the integration interval.
ans = 1

```
int(x*pareto, x, -inf, inf)
```

Warning: Unable to check whether the integrand exists everywhere on the integration interval.
ans = 0

```
int(x^2*pareto, x, -inf, inf)
```

Warning: Unable to check whether the integrand exists everywhere on the integration interval.
ans = ∞

The equation for updating the density is still the same as before:

$$u(x, t + \Delta t) = \int_{-\infty}^{\infty} u(x', t)\, q(x - x')dx'$$

and its Fourier transform is

$$\widehat{u}(k, t + \Delta t) = \widehat{u}(k, t)\, \widehat{q}(k)\,.$$

But the variance is infinite, we can't apply our moment expansion like we did before. We need a new technique to tackle this equation. It will be great if we know the formula for $\hat{q}(k)$.

MATLAB is too stupid to calculate the Fourier transform

```
fourier(pareto)
```

```
ans =
```

$$\mathrm{fourier}\left(\begin{cases} \frac{\alpha\, {x_m}^{\alpha}}{2\,|x|^{\alpha+1}} & \text{if } x_m < |x| \\ 0 & \text{otherwise} \end{cases}, x, w\right)$$

but from Mathematica we find that

$$\hat{q}(k) = 1 + \alpha \cos(\pi\alpha/2)\,\Gamma(-\alpha)\, {x_m}^{\alpha}|k|^{\alpha} + \frac{\alpha\, {x_m}^2 k^2}{2(2-\alpha)} + O({x_m}^4)$$

or in simpler terms,

$$\hat{q}(k) = 1 - C_\alpha {x_m}^{\alpha}|k|^{\alpha} + o({x_m}^{\alpha})$$

where $C_\alpha = -\alpha\cos(\pi\alpha/2)\,\Gamma(-\alpha)$ (which is positive), and the term $\frac{\alpha\, {x_m}^2 k^2}{2(2-\alpha)}$ goes to zero faster than ${x_m}^{\alpha}$ in the limit as x_m goes to zero (remember that $1 < \alpha < 2$).

Now, substitute $\hat{q}(k)$ into the Fourier transformed equation

$$\hat{u}(k, t+\Delta t) = \hat{u}(k, t)\,\hat{q}(k)$$

to get

$$\hat{u}(k, t+\Delta t) = \hat{u}(k, t)\,(1 - C_\alpha {x_m}^{\alpha}|k|^{\alpha} + o({x_m}^{\alpha})).$$

Taylor series on the left as usual, and we find

$$\hat{u}'(k, t) = -\frac{C_\alpha {x_m}^{\alpha}}{\Delta t}|k|^{\alpha}\hat{u}(k, t) + O(\Delta t) + \frac{o({x_m}^{\alpha})}{\Delta t}$$

So, letting $D = \frac{C_\alpha {x_m}^{\alpha}}{\Delta t}$ and taking limits as Δt and x_m go to zero, keeping D fixed, we have

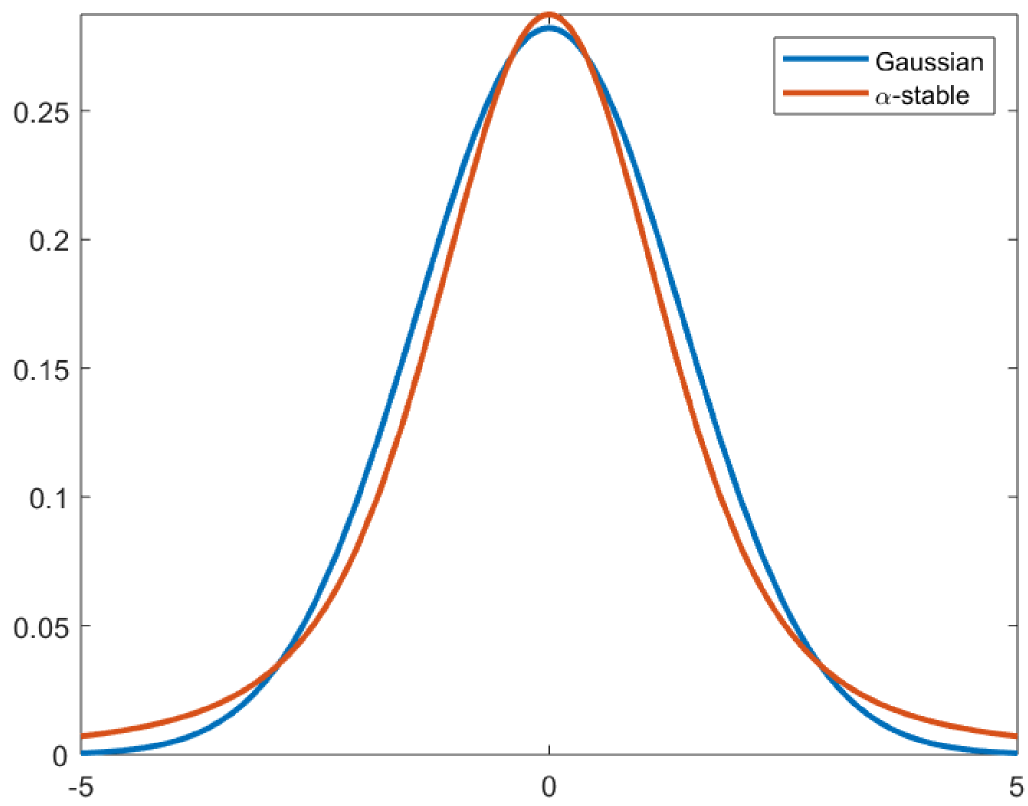
$$\hat{u}'(k, t) = -D\,|k|^{\alpha}\hat{u}(k, t)$$

and so

$$\hat{u}(k, t) = \exp(-D\,|k|^{\alpha} t).$$

But how do we invert this Fourier transform? Equivalently, for what probability distribution is this $\hat{u}$ the *characteristic function*? Statisticians will recognise the characteristic function as that of the so-called α-stable distribution. . Its density function u has no analytic expression, so this formula for $\hat{u}$ is as far as we can write down. But its properties as a probability distribution are very well known. It looks "Gaussian-like", but with heavier tails.

```
pd = makedist('Stable', 2, 0, 1, 0);  % Gaussian is a special case of alpha-stable (with alpha = 2)
figure, fplot(@pd.pdf, [-5,5], 'LineWidth', 2, 'DisplayName', 'Gaussian')
pd = makedist('Stable', 1.5, 0, 1, 0);  % alpha = 1.5
hold on, fplot(@pd.pdf, [-5,5], 'LineWidth', 2, 'DisplayName', '\alpha-stable')
legend
```



Now the question is: what partial differential equation does this characteristic function solve? We know for the *Gaussian* density it's the standard diffusion equation

$$\partial u(x,t)/\partial t = -D(-\Delta)u(x,t)$$

which in Fourier space is

$$\hat{u}'(k,t) = -Dk^2\hat{u}(k,t).$$

So what's the PDE that leads to the solution

$$\hat{u}'(k,t) = -D\,|k|^\alpha\hat{u}(k,t)$$

we found above? Simple! It's the *fractional* diffusion equation

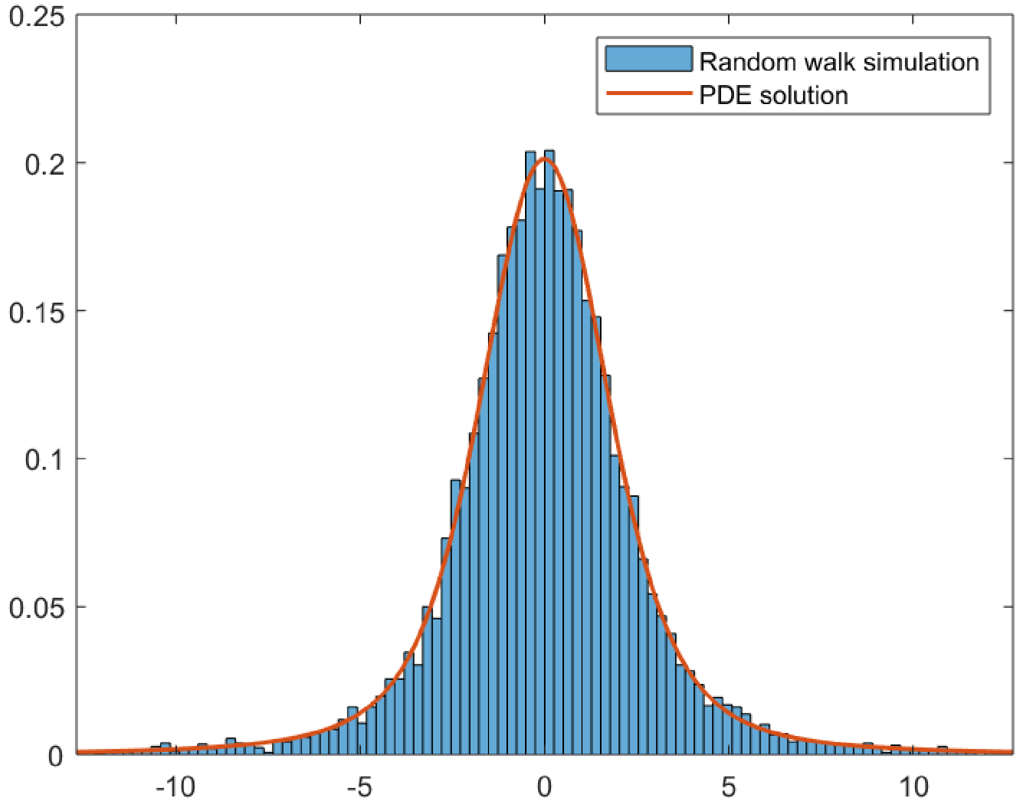
$$\partial u(x,t)/\partial t = -D(-\Delta)^{\alpha/2}u(x,t)$$

where the **fractional Laplacian** operator $(-\Delta)^{\alpha/2}$ is *defined* by its action in Fourier space:

$$\mathcal{F}\{(-\Delta)^{\alpha/2}f(x)\} = |k|^\alpha\mathcal{F}\{f(x)\}.$$

Since there is no analytical solution for the fractional diffusion equation on infinite domain, to compare the PDE solution to the random walk simulation we require a *numerical method* to solve the PDE. That will be the topic for the next two lessons. But for now, we'll trust that our code can do the trick.

```
random_walk_simulation('pareto',1.5)
```




```
D = 1.7055
```

We see the characteristic "heavy tails" associated with fractional diffusion.

Although we derived this fractional PDE using the Pareto distribution, the result is more general. We recognise that our two-sided Pareto distribution follows a form of two-sided power-law probability distribution

$$p(x) \sim A_\alpha \; \sigma^\alpha \, |x|^{-(1+\alpha)}, \text{ for large } |x|.$$

Here A_α is an arbitrary α-dependent constant, and σ is *playing the role of the variance* in determining how "spread out" the distribution is *(*remember though the the variance itself is infinite). Clearly our two-sided Pareto distribution meets this criteria, with $A_\alpha = \alpha/2$ and $\sigma = x_m$, but any other distribution with $1 < \alpha < 2$ meeting that asymptotic condition will lead to the same form of fractional PDE.

References:

R. Metzler, J. Klafter. The random walk's guide to anomalous diffusion: a fractional dynamics approach. Physics Reports 339 (2000) 1-77.