Transport Phenomena & Random Walks

One of the unifying themes of this course is to see how **macrophysical** processes (large-scale fluid averages) are connected to **microphysical** processes (particle-by-particle motions; *random?*).

The bridge that joins them: **probability distributions** (i.e., integrals over large numbers of independent random trials) of finding particles in a particular "macro" state.

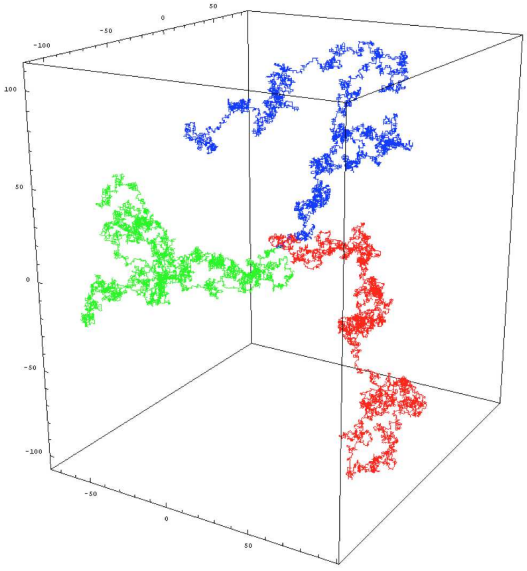
We'll develop these ideas in layers... hopefully each time we re-visit, we'll have extra physics insight. Summary of the layers:

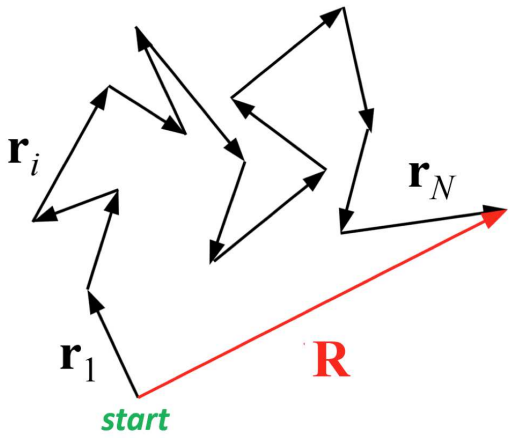
(1) Basic statistics of random walks (3D)
(2) Einstein-Smoluchowski theory: express as probabilities (1D)
(3) Langevin equation: incorporate more real physics (3D) **(optional)**

. . . . . .

On the micro scale, we'll often think about **random walks.**

Consider a point-like particle moving in 3D space, undergoing "interactions" every so often that randomly reorient its direction.





Initial position: $\mathbf{r} = 0$. Final position after N scatterings:

$$\mathbf{R} = \sum_{i=1}^{N} \mathbf{r}_i \qquad \text{(net displacement)} \; .$$

Consider random steps (i.e., random path lengths $|\mathbf{r}_i|$ and random directions in space θ_i, ϕ_i).

What's the particle's *averaged* behavior? Let's compute expectation values (e.g., $\langle \mathbf{R} \rangle$, $\langle R^2 \rangle$), wherein we average over an ensemble of $\gg 1$ independently realized trials.

Because motion in any direction is equally likely, intuition tells us that $\langle \mathbf{R} \rangle$ should be zero.

Prove it: Look at just one Cartesian coordinate:

$$\mathbf{R} \;=\; X\hat{\mathbf{e}}_x \;+\; Y\hat{\mathbf{e}}_y \;+\; Z\hat{\mathbf{e}}_z \;\;, \qquad \text{where}$$

$$X \;=\; \sum_{i=1}^{N} x_i \;=\; \sum_{i=1}^{N} r_i \,\sin\theta_i \,\cos\phi_i \qquad \text{and so on.}$$

The ensemble average is

$$\begin{aligned} \langle X \rangle \;&=\; \langle x_1 + x_2 + x_3 + \cdots \rangle \\ &=\; \langle x_1 \rangle + \langle x_2 \rangle + \langle x_3 \rangle + \cdots \end{aligned}$$

since the average of a sum = the sum of the averages.

$$\text{Thus, each piece is:} \qquad \langle x_i \rangle \;=\; \langle r_i \,\sin\theta_i \,\cos\phi_i \rangle$$

and if r_i, θ_i, and ϕ_i are all sampled from *independent* random distributions (i.e., they're not correlated with one another in any way), then we can make use of the basic law of joint probabilities...

$$\boxed{P(A \text{ and } B) \;=\; P(A)\,P(B)} \qquad \text{(for } A \text{ and } B \text{ being uncorrelated)}$$

so thus,

$$\langle x_i \rangle \;=\; \langle r_i \rangle \,\langle \sin\theta_i \rangle \,\langle \cos\phi_i \rangle \;\;.$$

There's another way to think about these ensemble averages. They're kind of equivalent to weighted averages over a probability distribution:

$$\langle f(x) \rangle \;=\; \frac{\int dx' \;P(x') \;f(x')}{\int dx' \;P(x')}$$

but for now, the probabilities are simple—i.e., for angles sampled uniformly in space (i.e., θ_i sampled between 0 and π, and ϕ_i sampled between 0 and 2π), we can easily compute, say,

$$\langle \cos\phi_i \rangle \;=\; \frac{\int_0^{2\pi} d\phi \;\cos\phi}{\int_0^{2\pi} d\phi} \;=\; \rightsquigarrow \;=\; 0 \qquad \text{for} \quad P = \frac{1}{2\pi} \;.$$

Thus, each $\langle x_i \rangle = 0$ (similarly for $\langle y_i \rangle$ & $\langle z_i \rangle$), and we get $\boxed{\langle \mathbf{R} \rangle \;=\; 0}$.

. . . . . .

Hold on... It may not be clear why the law of joint probabilities can actually *apply* to our case of expectation values. Let's take a minute to explore it for two abstract independent quantities A and B, for which we want to know the expectation value of their product:

$$\langle AB \rangle \;=\; \iint dA\; dB\; AB\; P(A,B)$$

where $P(A,B)$ is the joint probability of A and B happening together. Using the law of joint probabilities, we see that

$$\langle AB \rangle \;=\; \iint dA\; dB\; AB\; P(A)\, P(B)$$

and we can rearrange the integrals

$$\langle AB \rangle \;=\; \int dA\; A\; P(A) \int dB\; B\; P(B) \quad = \quad \langle A \rangle \langle B \rangle$$

which is the result we used above.

. . . . . .

We showed that $\langle \mathbf{R} \rangle = 0$. What can we say about $\langle R^2 \rangle \;=\; \langle |\mathbf{R}|^2 \rangle \;=\; \langle \mathbf{R} \cdot \mathbf{R} \rangle$?

This is the variance; i.e., the average of the distribution of squared distances (from the origin) after the particle random-walks N steps.

We can't use probability theory to split it up into a product of two averages, since $\mathbf{R}$ isn't independent of itself! It's clear that $\langle R^2 \rangle \neq \langle \mathbf{R} \rangle \cdot \langle \mathbf{R} \rangle$. Thus,

$$\begin{aligned} \langle \mathbf{R} \cdot \mathbf{R} \rangle \;&=\; \langle (\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 + \cdots) \cdot (\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3 + \cdots) \rangle \\ &=\; \langle |\mathbf{r}_1|^2 \rangle + \langle |\mathbf{r}_2|^2 \rangle + \langle |\mathbf{r}_3|^2 \rangle + \cdots \\ &\quad +\; 2\langle \mathbf{r}_1 \cdot \mathbf{r}_2 \rangle + 2\langle \mathbf{r}_1 \cdot \mathbf{r}_3 \rangle + \cdots \end{aligned}$$

i.e., it's a sum of lots of terms that look like

$$\begin{aligned} \langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle &= \langle |\mathbf{r}_i|\, |\mathbf{r}_j| \cos\alpha \rangle \\ &= \langle |\mathbf{r}_i|\, |\mathbf{r}_j| \rangle \langle \cos\alpha \rangle \qquad \text{(some with } i = j\text{, some with } i \neq j\text{)} \end{aligned}$$

where α is the angle between $\mathbf{r}_i$ and $\mathbf{r}_j$.

Of course, for $i \neq j$, the angle α is sampled randomly between 0 and π, and we find that

$$\langle \cos\alpha \rangle = 0$$

and thus all of the "cross-terms" in $\langle \mathbf{R} \cdot \mathbf{R} \rangle$ are zero.

We may have also reasoned this from the fact that (for $i \neq j$) $\mathbf{r}_i$ is completely uncorrelated with $\mathbf{r}_j$ (i.e., they're independent), so

$$\langle \mathbf{r}_i \cdot \mathbf{r}_j \rangle = \langle \mathbf{r}_i \rangle \cdot \langle \mathbf{r}_j \rangle = 0 \cdot 0 = 0 \ .$$

In any case, we keep only the $i = j$ terms,

$$\langle R^2 \rangle = \sum_{i=1}^{N} \langle |\mathbf{r}_i|^2 \rangle$$

and they keep adding to one another "in quadrature" (like in the Pythagorean theorem).

If we define the **mean free path** (i.e., r.m.s. step size)

$$\ell_{\text{mfp}} \equiv \sqrt{\frac{1}{N} \sum_{i=1}^{N} \langle |\mathbf{r}_i|^2 \rangle} \qquad \text{then} \qquad \langle R^2 \rangle = \ell_{\text{mfp}}^2 \, N$$

(i.e., if all steps were the same length, that length would be ℓ_{mfp}).

The r.m.s. distance from the origin σ (after taking N steps) is

$$\sigma = \langle R^2 \rangle^{1/2} = \ell_{\text{mfp}} \sqrt{N} \ .$$

Not to put too fine a point on it, but consider a particle moving at a constant speed v. The total path length traveled is $d = N\ell_{\text{mfp}}$, over total time t (i.e., $v = d/t$). Thus, we get $N = vt/\ell_{\text{mfp}}$, and we plug into the above to get $\sigma = \sqrt{\ell_{\text{mfp}} vt} \propto \sqrt{t}$, just like in the diffusion equation, where $D \sim \ell_{\text{mfp}} v$.

This is *suggestive* of a deep similarity between random walks and diffusion... but we're still not quite to the point of mathematically demonstrating that they're the same thing. To do that, let's look at...

Einstein–Smoluchowski theory of Brownian motion

In 1827, botanist Robert Brown looked at the motion of pollen grains in water, through a microscope. These grains seemed to undergo random-walk-like jittery motion.

In 1905, Einstein and Smoluchowski (independently) showed that Brownian motion can occur as the result of random **collisions** between the "large" grain and $\gg 1$ smaller, more rapidly moving molecules.

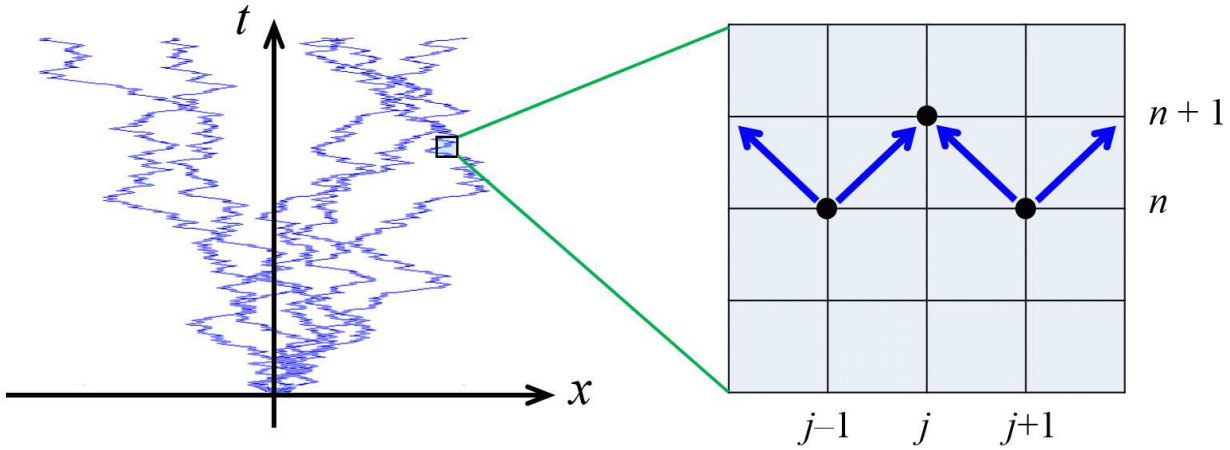
(This was a big deal: First real evidence that atoms and molecules actually exist, and aren't just abstractions that make chemical reactions tractable!)

Einstein and Smoluchowski also showed how these random jitters produce classical **diffusion.**

Let's simplify the problem into one spatial dimension (1D) & discrete steps:

$x(t)$ = the position of the Brownian grain at time t.

Assume each molecular impact (occurring at constant time intervals τ) causes the grain to jump *either* $+\ell$ or $-\ell$ in the x direction.



Animated version:
https://twitter.com/SarlCagan93/status/1332124614391173123

We assume successive jumps are **uncorrelated** with previous ones. (This is called a "memoryless" Markov process.) Thus:

$$\begin{aligned} p &= \text{probability of a step to the right } (+\ell) \\ q &= \text{probability of a step to the left } (-\ell) \\ p+q=1 &\quad \text{(i.e., it always has to move; it never stays still).} \end{aligned}$$

For now, let's just assume $p = q = 1/2$.

Quantify the steps as occurring on an integer grid in space (j) & time (n),

$$t = n\tau \quad , \qquad x = j\ell \quad , \qquad P_j^n \equiv \text{probability of the grain being at } (n, j)$$

At $t > 0$, the probability of ending up at a given place depends on the **history.** Just going back one step in time, we see it's possible to reach an arbitrary point in 2 different ways:

$$P_j^{n+1} \;=\; p\,P_{j-1}^n \,+\, q\,P_{j+1}^n \;=\; \frac{1}{2}\left(P_{j-1}^n \,+\, P_{j+1}^n\right)$$

We're going to just look at this single "step back" from P_j^{n+1}, but Einstein and Smoluchowski applied this *iteratively* over many steps... and eventually got a binomial distribution.

Trick: It's okay to do the same thing to both sides. Subtracting P_j^n from both sides, we get

$$P_j^{n+1} - P_j^n \;=\; \frac{1}{2}\left(P_{j-1}^n - 2P_j^n + P_{j+1}^n\right) \;.$$

It's sometimes necessary to make a problem more complicated (in the short term) in order to get to an answer that's simpler and more understandable (in the long term)!

This now looks like a finite-difference equation. For small ℓ and τ,

$$\frac{P_j^{n+1} - P_j^n}{\tau} \approx \frac{\partial P}{\partial t} \qquad \text{(forward difference in time)}$$

$$\frac{P_{j-1}^n - 2P_j^n + P_{j+1}^n}{\ell^2} = \frac{1}{\ell}\left(\frac{P_{j+1}^n - P_j^n}{\ell} - \frac{P_j^n - P_{j-1}^n}{\ell}\right) \approx \frac{\partial^2 P}{\partial x^2} \quad \text{(2 space diff's)}$$

Thus, if we transition to thinking about $P(x,t)$ as a continuous variable,

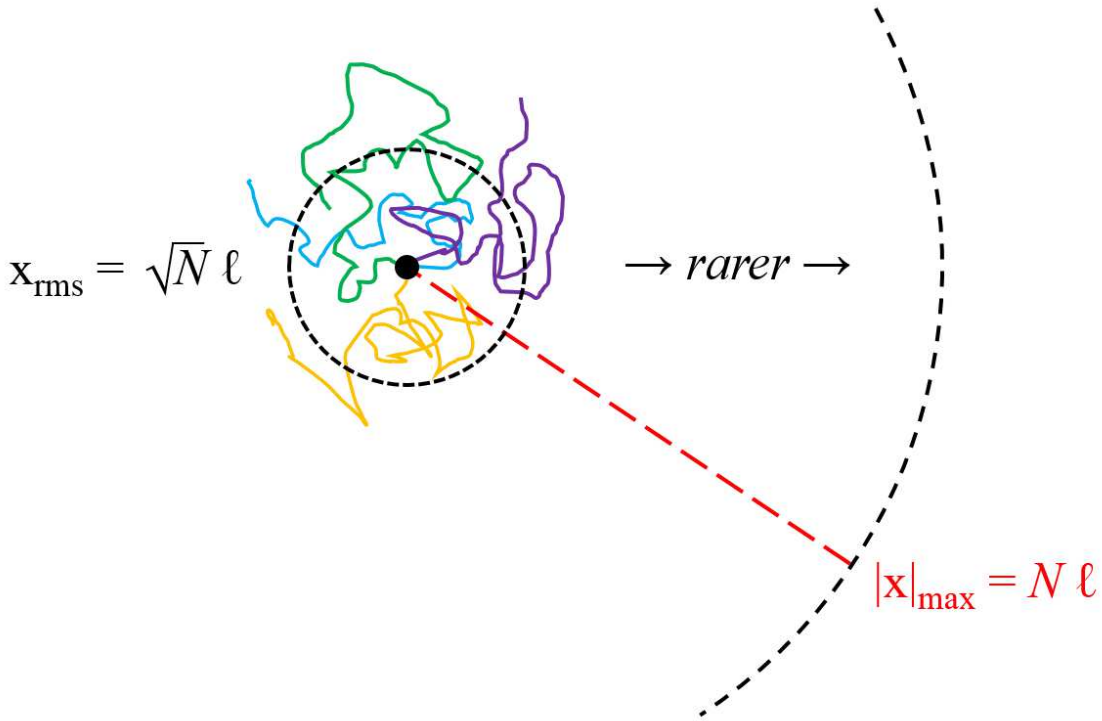
$$\boxed{\frac{\partial P}{\partial t} \approx D\frac{\partial^2 P}{\partial x^2} \qquad \text{where} \qquad D \equiv \frac{\ell^2}{2\tau}}$$

i.e., the probability of ending up at a particular place (in a random walk) obeys a diffusion equation. **Probability is a quantity that 'diffuses.'**

Problem: Hold on! If we start out with a particle at an exactly known position x_0 (i.e., delta function initial condition), our Fourier transform solution says the probability of being at an arbitrarily large $x \neq x_0$ is **finite** for any $t > 0$.

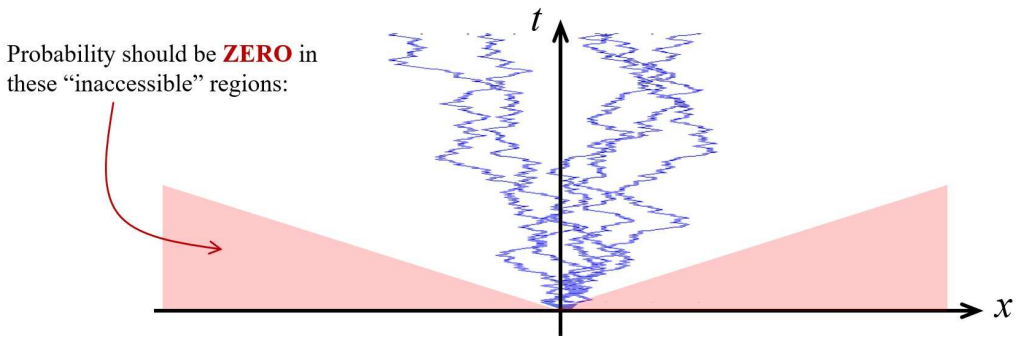
That doesn't make physical sense. We know that **IF** all steps happened to line up in the same direction $\longrightarrow \longrightarrow \longrightarrow \longrightarrow \longrightarrow \longrightarrow \longrightarrow$ then there's a finite maximum distance that could have been traveled.

If $t = N\tau$, then the probability ought to be exactly 0 for reaching any $|x| > |x|_{\max} = N\ell$.



Most trajectories will remain within $x_{\rm rms} = \sqrt{2Dt} = \sqrt{\ell^2 t/\tau} = \ell\sqrt{N}$.

Another way to view what's going on:



Solution: We weren't careful enough in our finite differencing!

Let's more carefully rewrite the original probability equation using *continous* variable language... $P_j^n \longrightarrow P(x,t)$.

$$P_j^{n+1} \;=\; \frac{1}{2}\left(P_{j-1}^n \,+\, P_{j+1}^n\right) \qquad \text{becomes...}$$

$$P(x,t+\tau) \;=\; \frac{1}{2}\left[P(x-\ell,t) + P(x+\ell,t)\right] \;.$$

Expand these terms as Taylor series...

$$P(x,t+\tau) \;=\; P(x,t) \,+\, \frac{\tau}{1!}\left(\frac{\partial P}{\partial t}\right)_{x,t} + \frac{\tau^2}{2!}\left(\frac{\partial^2 P}{\partial t^2}\right)_{x,t} + \,\cdots \mathcal{O}(\tau^3)\,\cdots$$

$$P(x\pm\ell,t) \;=\; P(x,t) \,\pm\, \frac{\ell}{1!}\left(\frac{\partial P}{\partial x}\right)_{x,t} + \frac{\ell^2}{2!}\left(\frac{\partial^2 P}{\partial x^2}\right)_{x,t} \pm \,\cdots \mathcal{O}(\ell^3)\,\cdots$$

Above, we were thinking about diffusion with $D \sim \ell^2/\tau$, so we essentially kept everything up to $\mathcal{O}(\tau^1)$ in time, and $\mathcal{O}(\ell^2)$ in space.

In 1934, Boris Davydov realized we should be more consistent, and keep every term up to second-order in time and space...

$\longrightarrow$ $P(x,t)$ cancels out on both sides.
$\longrightarrow$ $\ell(\partial P/\partial x)$ terms cancel each other out on RHS.

and we get... $$\boxed{\frac{\partial P}{\partial t} + \frac{\tau}{2}\frac{\partial^2 P}{\partial t^2} \;=\; D\,\frac{\partial^2 P}{\partial x^2}}$$

where we still define $D = \ell^2/(2\tau)$, a finite ratio of tiny quantities.

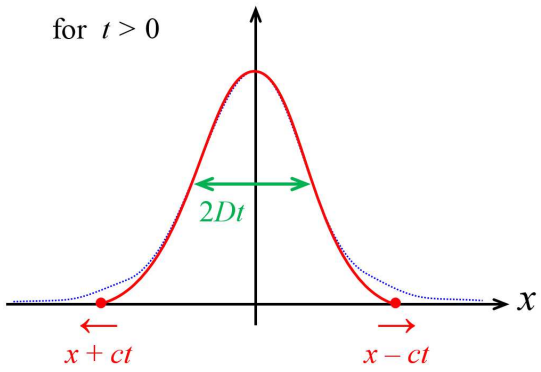
This equation is sometimes called the **telegraph equation** (or telegrapher's equation), since it's used to model lossy propagation of signals along wires.

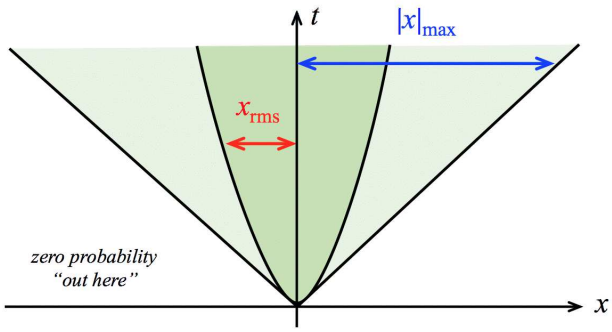
In the limit of $t \gg \tau$ (long time), it's a diffusion equation like before.

But for $t \ll \tau$ (short time), it's a wave equation,

$$\frac{\partial^2 P}{\partial t^2} \approx c^2 \frac{\partial^2 P}{\partial x^2} \qquad \text{where} \quad c = \pm\frac{\ell}{\tau} \ .$$

It really *doesn't* communicate at infinite speed. It describes the "ballistic" propagation of a front, at phase speed c, and diffusion "fills in" behind it.





After $N \gg 1$ steps, $|x_{\max}| = \ell N$, but $x_{\rm rms} = \ell\sqrt{N}$.

It's an interesting mix of deterministic & stochastic motion. Random-walk irreversibility eventually ($t \gg \tau$) "wins" by filling in the slowly-expanding region with $|x| < x_{\rm rms}$.

The Fourier transform technique gives an analytic solution to the full telegraph equation (involving Bessel functions), but we won't really need it.

Fisk & Axford (1969, *Solar Phys.*, 7, 486) applied it to the quasi-diffusive transport of cosmic rays (solar energetic particles) in the heliosphere.

. . . . . .

We've learned a lot about how random micro-processes give rise to diffusive macro-behavior... but it's still been very much a "toy model." But we still **don't know** how to compute the actual values of ℓ, τ, D, and so on.

For a *physical* description of Brownian motion, we need more quantitative information about the forces that act on the "pollen grain."

The Langevin Equation (optional section)

Consider a grain (of mass m) bombarded by a force due to multiple impacts from randomly moving molecules.

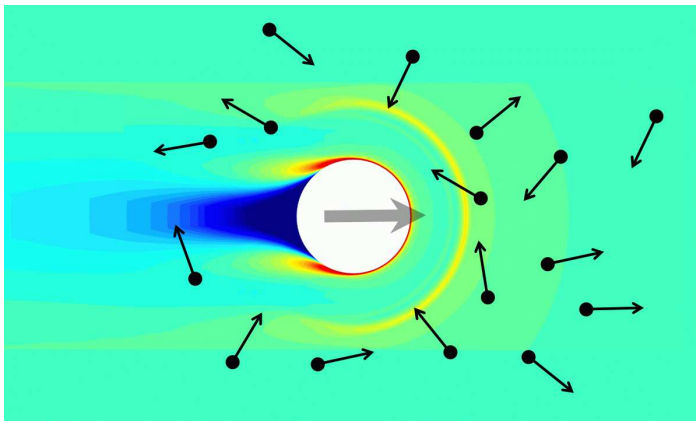
$$m\,\frac{d\mathbf{v}}{dt} \;=\; \mathbf{F}(t) \qquad \text{(a real equation of motion!)}$$

Paul Langevin (1908) postulated that the force can be decomposed into two parts: $\mathbf{F}(t) \;=\; \mathbf{F}_0 + \mathbf{f}(t)$.

Zero-order part: a temporal average over the multiple impacts, but still can be slowly varying in time.

First-order part: rapidly fluctuating part that averages to zero over long times.

Heuristically, we see that $\mathbf{F}_0$ should act like a **viscous drag** term; i.e., if the grain is moving rapidly in the $+x$ direction, it will collide with more particles coming at it from the right (with $-x$ speeds in the grain's frame) than particles from the left.



The net sense of momentum transfer (from the molecules to the big grain) **slows it down.** For now, let's postulate Stokes Law of viscous drag:

$$\mathbf{F}_0 \;\propto\; -\mathbf{v} \qquad \text{i.e.,} \quad \mathbf{F}_0 \;=\; -\left(\frac{1}{B}\right)\mathbf{v}$$

we'll see what happens. Here, B is called the "mobility." Also, define the relaxation time: $\boxed{\tau \;=\; m\,B}$.

B has strange units (velocity/force, or time/mass), but it's a traditional proportionality constant in fluid dynamics.

Later, we will **prove** that the zero-order force for Coulomb collisions in a plasma is indeed of this form (i.e., $F \propto -v$).

Thus, we get the Langevin equation:

$$\boxed{\frac{d\mathbf{v}}{dt} = -\frac{\mathbf{v}}{\tau} + \frac{\mathbf{f}(t)}{m}} \qquad \text{where} \quad \overline{\mathbf{f}(t)} = 0 \ ,$$

and the 'over-bar' notation denotes a long-time average (over many molecular collisions) in a single grain's evolution.

We also can say that $\langle \mathbf{f}(t) \rangle = 0$, which means that an ensemble-average over many random trials gives zero for the fluctuating part of the force.
Not the same as $\overline{\mathbf{f}(t)}$.

Is the Langevin equation directly applicable in astrophysics? Yes..

- Elongated dust grains in the ISM are sometimes partially magnetized; their alignment produces polarization, which traces **B**. Grain alignments evolve in time because they're embedded in gas. It's a kind of Brownian *torque* (e.g., Roberge et al. 1993, *ApJ,* 418, 287).
- Are pulsar magnetospheres responsible for NS spindown? They undergo random fluctuations (timing glitches, radio flares) that can be modeled with Langevin-like eqns (Ou et al. 2016, *MNRAS,* 457, 3922).
- MHD turbulence in the solar wind! (Bandyopadhyay et al. 2018, *Phys Rev E,* 97, 053211)

. . . . . .

If we take the ensemble average of each term in the Langevin equation, the **f** term drops out, and we just have

$$\left\langle \frac{d\mathbf{v}}{dt} \right\rangle = \left\langle -\frac{\mathbf{v}}{\tau} \right\rangle \qquad \rightsquigarrow \rightsquigarrow \qquad \frac{d\langle \mathbf{v} \rangle}{dt} = -\frac{\langle \mathbf{v} \rangle}{\tau}$$

where we can take out the d/dt because differentiation is a *linear* operator.

This has a straightforward solution... $\quad \langle \mathbf{v} \rangle(t) \ = \ \langle \mathbf{v}_0 \rangle \, e^{-t/\tau} \ = \ \mathbf{v}_0 \, e^{-t/\tau} \quad .$

As expected, on average, the effective viscosity term damps out any initial motion $\mathbf{v}_0$, and slows the grain down to rest.

. . . . . .

Of course, we can't just ignore the random motions altogether.

We'd like to know more about their statistics. Do they give rise to *diffusion* like we expect from our earlier mathematical work?

Let's recall that the grain has a time-evolving position vector $\mathbf{r}$ in addition to a velocity $\mathbf{v} = d\mathbf{r}/dt$. We can derive an equation to see how r^2 evolves in time, then take the ensemble average to see how $\langle r^2 \rangle$ evolves in time.

To make progress, take $\mathbf{r} \cdot \{\text{each term in the full Langevin equation}\}$:

$$\mathbf{r} \cdot \frac{d\mathbf{v}}{dt} \;=\; -\frac{\mathbf{r} \cdot \mathbf{v}}{\tau} + \frac{\mathbf{r} \cdot \mathbf{f}}{m} \;\; .$$

What is $\mathbf{r} \cdot \mathbf{v}$? Note that

$$\frac{dr^2}{dt} \;=\; \frac{d}{dt}(\mathbf{r} \cdot \mathbf{r}) \;=\; 2\mathbf{r} \cdot \frac{d\mathbf{r}}{dt} \;=\; 2\mathbf{r} \cdot \mathbf{v} \;\; . \qquad \boxed{\mathbf{r} \cdot \mathbf{v} = \frac{1}{2}\frac{dr^2}{dt}}$$

Put that aside for a second. Also note from the chain rule,

$$\frac{d}{dt}(\mathbf{r} \cdot \mathbf{v}) \;=\; \mathbf{v} \cdot \frac{d\mathbf{r}}{dt} + \mathbf{r} \cdot \frac{d\mathbf{v}}{dt}$$

Rearranging,

$$\begin{aligned} \mathbf{r} \cdot \frac{d\mathbf{v}}{dt} \;&=\; \frac{d}{dt}(\mathbf{r} \cdot \mathbf{v}) - \mathbf{v} \cdot \frac{d\mathbf{r}}{dt} \\ &=\; \frac{d}{dt}\left(\frac{1}{2}\frac{dr^2}{dt}\right) - \mathbf{v} \cdot \mathbf{v} \\ &=\; \frac{1}{2}\frac{d^2}{dt^2}(r^2) \;-\; v^2 \;\; . \end{aligned}$$

That goes into the LHS of $\mathbf{r} \cdot \{\text{Langevin}\}$, and a bit more rearranging gives:

$$\frac{1}{2}\frac{d^2}{dt^2}(r^2) - v^2 \;=\; -\frac{1}{2\tau}\frac{dr^2}{dt} + \frac{\mathbf{r} \cdot \mathbf{f}}{m}$$

$$\frac{d^2}{dt^2}(r^2) \;+\; \frac{1}{\tau}\frac{d}{dt}(r^2) \;=\; \frac{2\,\mathbf{r} \cdot \mathbf{f}}{m} + 2v^2 \;\; .$$

What happens when we take the ensemble average of *this* equation?

Note that the fluctuating term ought to be

$$\langle \mathbf{r} \cdot \mathbf{f} \rangle \;=\; \langle \mathbf{r} \rangle \cdot \langle \mathbf{f} \rangle \;=\; 0$$

since $\langle \mathbf{f} \rangle = 0$. Also, there's no reason to expect a correlation to exist between the particle's absolute position $\mathbf{r}$ and the random molecular forces on it!

Thus, we get the EALE (ensemble averaged Langevin equation):

$$\frac{d^2}{dt^2}\langle r^2 \rangle \;+\; \frac{1}{\tau}\frac{d}{dt}\langle r^2 \rangle \;=\; 2\langle v^2 \rangle \;\; .$$

If we knew $\langle v^2 \rangle$, we could solve this 2nd order ODE for $\langle r^2 \rangle(t)$.
If, for example, we got $\langle r^2 \rangle \propto t$, then we know we're on the right (diffusive) track, and the constant of proportionality would give us D.

I know... I know... we used ensemble averaging to get rid of the fluctuations again. Or did we?

$\langle v^2 \rangle$ describes the ensemble variance of the distribution of speeds felt by the pollen grain. Thus, the effects of $\mathbf{f}$ are still, sort of, in there.

In reality, we'll see that $\langle v^2 \rangle$ varies in time like $\langle r^2 \rangle$ does. But **eventually** ($t \gg \tau$) the random fluctuations should be all that's left.

For a Maxwell-Boltzmann gas of particles in thermal equilibrium, let's express these things in "modern language,"

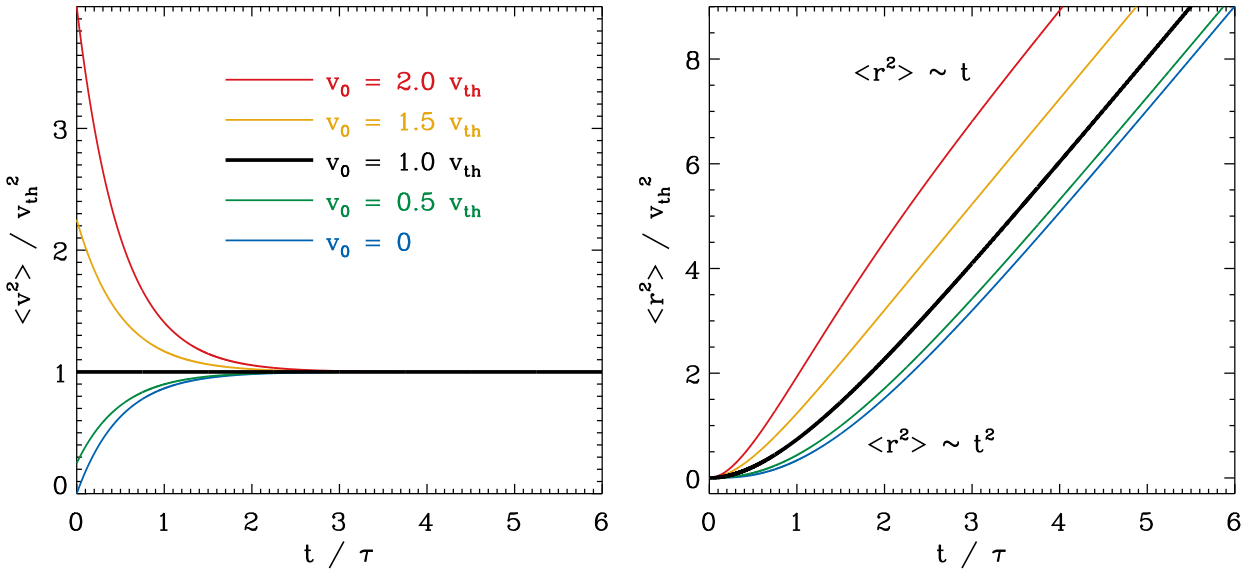
$$\left\langle \frac{1}{2} m v^2 \right\rangle \;=\; \frac{3}{2} k_{\mathrm{B}} T \qquad \text{i.e.,} \qquad \langle v^2 \rangle \;=\; \frac{3 k_{\mathrm{B}} T}{m} \;\equiv\; v_{\mathrm{th}}^2 \;\; .$$

If we assume that $\langle v^2 \rangle$ maintains this constant value, the EALE is just a 2nd order ODE with constant coefficients. It has a straightforward solution:

$$\langle r^2 \rangle \;=\; 2 v_{\mathrm{th}}^2 \tau^2 \left[\frac{t}{\tau} - \left(1 - e^{-t/\tau} \right) \right]$$

for which we used the initial condition $\langle r^2 \rangle = 0$ at $t = 0$.

This solution is the thick **black curve** (don't worry yet about the others):



Limiting values:

(1) For $t \ll \tau$ (i.e., times less than viscous relaxation time),

$$e^{-t/\tau} \approx 1 - \frac{t}{\tau} + \frac{t^2}{2\tau^2} - \cdots$$

$$\text{Thus,} \quad \langle r^2 \rangle = 2v_{\rm th}^2 \tau^2 \left[\frac{t}{\tau} - 1 + 1 - \frac{t}{\tau} + \frac{t^2}{2\tau^2} - \cdots \right]$$

$$\text{i.e.,} \quad \langle r^2 \rangle \approx v_{\rm th}^2 t^2 \approx \langle v^2 \rangle t^2$$

linear/ballistic short-term motion, like we saw in the telegraph equation.

(2) For $t \gg \tau$ (which is more appropriate when assuming $\langle v^2 \rangle = v_{\rm th}^2$),

$$\langle r^2 \rangle \approx 2v_{\rm th}^2 \tau\, t \equiv 6Dt \qquad \text{(let's just assign this definition)}$$

which is diffusive! This means there's an effective diffusion coefficient:

$$\boxed{D = \frac{k_{\rm B} T \tau}{m} = B k_{\rm B} T}$$

which is called the **Einstein relation** for Brownian motion.

It tells us there's a deep connection between the source of viscosity and the **random-walk** thermal diffusion. It's also testable: measure D from $\langle r^2 \rangle$, heat it up (i.e., change T), measure D again, and so on....

We're not quite done. We've included the random motions in an approximate way (i.e., assumed constant $\langle v^2 \rangle$), but we haven't really delved into the **statistical** properties of $\mathbf{f}(t)$.

(In other words, we neglected the real physics of the **friction.**)

We've made a lot of progress looking at the EALE, but let's take a step back and note there's actually a formal solution to the original Langevin equation:

$$\frac{d\mathbf{v}}{dt} = -\frac{\mathbf{v}}{\tau} + \frac{\mathbf{f}(t)}{m} \qquad \rightsquigarrow\rightsquigarrow \qquad \boxed{\mathbf{v}(t) = \mathbf{v}_0 e^{-t/\tau} + e^{-t/\tau} \int_0^t dt' \; e^{t'/\tau} \frac{\mathbf{f}(t')}{m}}$$

One can see that taking the ensemble average of this solution gives back our old ensemble solution

$$\langle \mathbf{v}(t) \rangle = \mathbf{v}_0 \, e^{-t/\tau} \qquad \text{since } \langle \mathbf{f} \rangle = 0 \text{ inside the integral.}$$

However, what we really want is

$$\begin{aligned} \langle v^2 \rangle(t) = \langle \mathbf{v}(t) \cdot \mathbf{v}(t) \rangle &= v_0^2 \, e^{-2t/\tau} + 2e^{-2t/\tau} \left\langle \mathbf{v}_0 \cdot \int_0^t dt' \; e^{t'/\tau} \, \frac{\mathbf{f}(t')}{m} \right\rangle \\ &+ e^{-2t/\tau} \left\langle \int_0^t dt_1 \int_0^t dt_2 \; e^{(t_1+t_2)/\tau} \, \frac{\mathbf{f}(t_1) \cdot \mathbf{f}(t_2)}{m^2} \right\rangle \end{aligned}$$

There are 3 terms on the RHS:

1st: already ensemble-averaged (i.e., it is its own ensemble average); no fluctuating components.

2nd: the only fluctuation component in the integral is $\mathbf{f}(t')$, and we know its ensemble average vanishes. $\boxed{\text{2nd term} = 0}$.

3rd: this involves the **autocorrelation** of **f.**

The 3rd term contains the integral

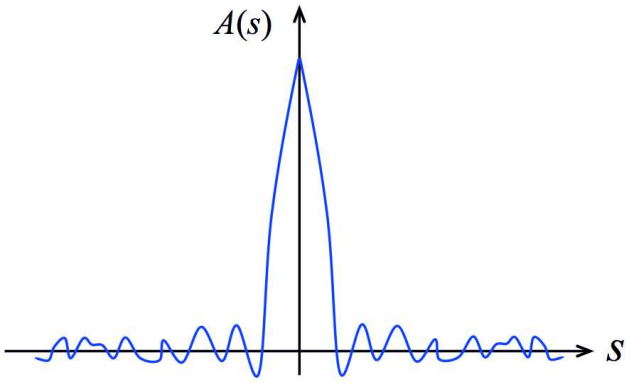
$$I = \frac{1}{m^2} \int_0^t dt_1 \int_0^t dt_2 \; e^{(t_1+t_2)/\tau} \; \langle \mathbf{f}(t_1) \cdot \mathbf{f}(t_2) \rangle \; .$$

Consider a change of variables: keep t_1, but define $s = t_2 - t_1$,

$$I = \frac{1}{m^2} \int_0^t dt_1 \int_{-t_1}^{t-t_1} ds \; e^{(2t_1+s)/\tau} \; \underbrace{\langle \mathbf{f}(t_1) \cdot \mathbf{f}(t_1+s) \rangle}_{\equiv A(s)}$$

$A(s)$ is a function of s, the *relative time lag* between t_1 and t_2.

However, because we assume $\mathbf{f}(t)$ is stationary (i.e., one absolute time is as good as any other, statistically), we can say that $A(s)$ does <u>not</u> depend on t_1.



Properties of the autocorrelation function for a stochastic process:

- maximum at $s = 0$.
- as $|s| \to \infty$, $A(s) \to 0$.
- symmetric: $A(s) = A(-s)$.

For a fluctuating force that varies **rapidly** compared to "usual" macroscopic timescales (like τ), why don't we just assume

$$A(s) \;=\; A_0\,\delta(s)$$

i.e., that there's NO correlation between subsequent "impact events" that produce the stochasticity.

Since we'll be integrating over s, maybe the actual width of $A(s)$ is unimportant.

Note: a more proper definition of the constant A_0 is

$$A_0 \;=\; \int_{-\infty}^{+\infty} ds\; A(s) \qquad \text{(i.e., the integral under the } A(s) \text{ curve).}$$

As long as $t > t_1$, we can write the inner integral (in I) as

$$\int_{-t_1}^{t-t_1} ds\; e^{(2t_1+s)/\tau}\, A(s) \;=\; A_0\, e^{2t_1/\tau}$$

$$\text{so that} \qquad I \;=\; \frac{A_0}{m^2}\int_0^t dt_1\; e^{2t_1/\tau} \;=\; \frac{\tau A_0}{2m^2}\left(e^{2t/\tau} - 1\right)\;.$$

Thus,

$$\langle v^2 \rangle \;=\; v_0^2\, e^{-2t/\tau} \;+\; \frac{\tau A_0}{2m^2}\left(1 - e^{-2t/\tau}\right)\;.$$

As $t \to \infty$, we still want to ensure that $\langle v^2 \rangle \to v_{\rm th}^2$. Thus,

$$\text{As } t \to \infty, \qquad \langle v^2 \rangle \;=\; \frac{\tau A_0}{2m^2} \;=\; \frac{3k_{\rm B}T}{m}$$

or in other words,

$$\boxed{A_0 = \int_{-\infty}^{+\infty} ds\; A(s) \;=\; \frac{6mk_{\rm B}T}{\tau}} \qquad \textbf{(the fluctuation-dissipation theorem).}$$

This is a kind of "strong version" of the Einstein relation.

We **had** been treating A_0 and τ as if they're independent of one another. They're not! For a given T, we're not free to specify both τ and A_0.

Also, it makes sense that $A_0 \propto T$, since stronger fluctuations ought to imply a higher temperature.

However, the remarkable thing is that if we specify the random micro–fluctuations completely (i.e., we specify A_0 and T), then that uniquely determines the macro–viscous (dissipative) timescale τ.

Note that if we set $v_0 = v_{\rm th}$, then $\langle v^2 \rangle = v_{\rm th}^2$ for all time.

Statistical equilibrium, once attained, has a natural tendency to persist!

. . . . . .

Lastly, one can substitute the full solution for $\langle v^2 \rangle$ into the EALE and solve for $\langle r^2 \rangle$ as a general function of t, τ, v_0, and $v_{\rm th}$. Too much detail for us here.

See multi-color curves in plot a few pages back. The behavior is similar to the simpler version: ballistic trajectories for $t \ll \tau$, and diffusion for $t \gg \tau$.

. . . . . .

Additional context from Pathria's *Statistical Mechanics,* 2nd ed.:

> Equation (1) establishes a fundamental relationship between the coefficient, $1/B$, of the "averaged-out" part of the total force $\mathscr{F}(t)$ experienced by the Brownian particle due to the impacts of the fluid molecules and the statistical character of the "fluctuating" part, $F(t)$, of that force; see Langevin's equation (14.4.2). In other words, it relates the coefficient of viscosity of the fluid, which represents *dissipative* forces operating in the system, with the temporal character of the molecular *fluctuations*; the content of eqn. (1) is, therefore, referred to as a *fluctuation–dissipation theorem*. The most striking feature of this theorem is that it relates, in a fundamental manner, the fluctuations of a physical quantity pertaining to the *equilibrium state* of the given system to a dissipative process which, in practice, is realized only when the system is subject to an external force that drives it *away from equilibrium*. Consequently, it enables us to determine the *non-equilibrium properties* of a given system on the basis of a knowledge of the thermal fluctuations occurring in the system when the system is in one of its *equilibrium states*!

In other words, the statistical properties of the equilibrium state (A_0 and $\langle v^2 \rangle$) are related to the external forces that exist only OUT of equilibrium (B, τ).

Ramshaw (2010, *Am. J. Phys.,* 78, 9) discusses the fact that if a system contained $\mathbf{f}(t)$ without any viscosity, there would be irreversible & continuous **heating.** The fluct-diss theorem says that, in order to find a steady state, viscosity must be there in order to counterbalance the "diverging" kinetic energy from the random force.