Interacting Particle Systems

Subtitle

Thomas M. Hodgson
Maxwell Institute

July 30, 2019

Contents

List of Figures

Abstract

+++ +++

Acknowledgements

1   Introduction

+++Motivation, particle systems and flocking in biology+++

2   The Model

We here elucidate the model of Buttà, Flandoli, Ottobre and Zegarlinski [1]. This is a Vicsek-type continuum model for an interacting particle system. There are three key parts to this model that together create interaction and ordered motion: a herding function, denoted G; an averaging function, denoted M; and an interaction function, φ. Each will be described before the full model is introduced. Throughout, we will be considering the phase space distribution function $f_t(x, v)$, where $(x, v) \in \mathbb{T} \times \mathbb{R}$ and $t \in \mathbb{R}$ with $\mathbb{T} \cong \mathbb{R}\backslash\mathbb{Z} \cong S^1$ denoting the unit one-dimensional torus. The distribution function represents the density of particles at time t at a given point in phase space (x, v).

The interaction function, $\varphi : \mathbb{T} \to \mathbb{R}$ acts on the smallest scale, between individual points. It is smooth and satisfies the following:

$$\varphi(x) \geq \varepsilon > 0, \qquad\qquad \varphi(x) = \varphi(-x), \qquad\qquad \int_{\mathbb{T}} \varphi(x)\, \mathrm{d}x = 1. \tag{1}$$

These are quite natural ways to describe an interaction. The first means that all points interact with each other to some degree. +++ Positivity requires that all interactions are attractive, in contrast to many real-world phenomena, give examples of only attractive systems +++. The second requires symmetry when interacting, and finally the function must have unit integral – this is only for simplicity and could in fact be any value, so long as it is finite. For example, we may take $\varphi \equiv 1$. This will satisfy the above requirements and corresponds to all points interacting an equal amount, regardless of position on the torus. This will create space-homogeneity, which we will return to later.

The function $M : \mathbb{R} \to \mathbb{R}$ takes a weighted average of the interactions across the space.

$$M(t, x) = \frac{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y, w)\varphi(x - y)w\, \mathrm{d}w\, \mathrm{d}y}{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y, w)\varphi(x - y)\, \mathrm{d}w\, \mathrm{d}y}$$

The interaction function provides the weighting. Note the argument of φ is the distance between two points in space and also that M depends on f_t. +++The denominator is required in the corresponding particle model to prevent the size of the group of particles from having an effect on the dynamics. It can be thought of as a normaliser. – Poor explanation.+++

Finally, the herding function $G : \mathbb{R} \to \mathbb{R}$ is chosen such that

$$G(u) = -G(-u), \qquad\qquad \begin{cases} G(u) > u & \text{if } 0 < u < 1 \\ G(u) < u & \text{if } u > 1 \end{cases}.$$

We shall consider three forms of G.

i) Step:

$$G(u) = \frac{1 + \beta\,\mathrm{sgn}(u)}{1 + \beta}, \qquad \beta > 0$$

This is the usual choice in the biological literature, with $\beta = \frac{1}{2}$. It is differentiable everywhere with $G'(u) = \frac{1}{1+\beta}$, except at $u = 0$. This discontinuity makes analysis more difficult.

ii) Smooth:

$$G(u) = \frac{\arctan(u)}{\arctan(1)}$$

Here we get a much smoother curve that is everywhere differentiable. However the gradient isn't very steep and there is no way to control the shape of the curve.

iii) Sigmoid:

$$G(u) = \frac{2}{1 + \mathrm{e}^{-\alpha u}} - 1, \qquad \alpha > 0$$

The sigmoid herding function provides a middle ground between the previous two. It is smooth, but can be made as sharp as required by adjusting the parameter α. However note that it only satisfies second requirement above in the limit as $\alpha \to \infty$ as the function is bounded above by 1.

During the analysis, we will use the smooth herding function for its amenability. The role of G is to herd the mean velocity towards a fixed value – in this case plus or minus one. How it does so will become clearer when we discuss the dynamics of the model.

We now have all the necessary ingredients to give the full continuum model. The evolution of the distribution is described by

$$\partial_t f_t(x,v) = -v\partial_x f_t(x,v) + \partial_v v f_t(x,v) - \partial_v\left[G(M(t,x))f_t(x,v)\right] + \sigma\partial_{vv}f_t(x,v), \qquad (2) \quad \boxed{\texttt{eq:fullPDE}}$$

where $\sigma > 0$ is a constant describing the rate of diffusion. At its heart, this is just an advection-diffusion equation. The first term illustrates the movement around the torus, the second describes a damping effect and the fourth gives a diffusive effect in velocity. The third term is where the interest lies as it contains all the information on interaction and herding.

3 Dynamics

$\boxed{\texttt{ec:dynamics}}$

To better understand the dynamics of ($\overset{\texttt{eq:fullPDE}}{2}$), first let $G \equiv 0$. This removes any interaction between particles and the equation becomes

$$\partial_t f_t(x,v) = -v\partial_x f_t(x,v) + \partial_v v f_t(x,v) + \sigma\partial_{vv}f_t(x,v). \qquad (3) \quad \boxed{\texttt{eq:OUFPE}}$$

This is the Fokker-Planck equation for the following system:

$$\begin{cases} \mathrm{d}x_t = v_t\,\mathrm{d}t \\ \mathrm{d}v_t = -v_t\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t. \end{cases} \qquad (4) \quad \boxed{\texttt{eq:2dOU}}$$

6

Here, W_t is a standard one-dimensional Wiener process. We recognise this as overdamped Langevin dynamics, and so we can exactly calculate the unique invariant measure. To do so it is necessary to solve $\partial_t \rho_t = 0$, that is, when is the density not changing in time. The evolution in v is described by an Ornstein-Uhlenbeck process so it is known that the unique stationary distribution is Gaussian with mean 1 and variance σ. The position x depends only on v with no noise. There is no reason to assume any sort of mean reverting behaviour and expect the density to spread out uniformly on the torus. A guess at the solution would then be $\rho(x, v) = \frac{1}{2\pi} \exp(\frac{-v^2}{2})$, with some normalising constant. Notice that this equation does not depend on x. By substitution into the stationary Fokker-Planck equation, this is indeed a solution.

Similarly, when $G \not\equiv 0$ the density $f_t(x, v)$ of (2) is the Fokker-Planck equation of the two dimensional SDE,

$$\begin{cases} \mathrm{d}x_t = v_t \, \mathrm{d}t \\ \mathrm{d}v_t = [G(M(t, x_t)) - v_t] \, \mathrm{d}t + \sqrt{2\sigma} \, \mathrm{d}W_t. \end{cases} \qquad (x_t, v_t) \in \mathbb{T} \times \mathbb{R} \qquad (5)$$

This is a McKean-Vlasov equation as the function $M(t, x_t)$ conceals a dependence on the law of the process.

$$M(t, x) = \frac{\mathbb{E}\left[v'_t \varphi(x - x'_t)\right]}{\mathbb{E}\left[\varphi(x - x'_t)\right]}$$

The pair (x'_t, v'_t) is a random variable equal in distribution to (x_t, v_t). This representation of $M(t, x)$ is the same as the one given previously, however in this setting it is natural to use expectation notation over integrals. To illustrate the necessity of introducing another random variable, albeit one that is equal in law, consider the case when the interaction function is the indicator on the interval $[-b, b]$, $\varphi(x) = \mathbb{1}_{[-b,b]}(x)$. This does not meet the requirements in Section 2, in particular it is neither strictly positive nor does it have unit integral.

$$\begin{aligned} M(t, x) &= \frac{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y, w) \varphi(x - y) w \, \mathrm{d}w \, \mathrm{d}y}{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y, w) \varphi(x - y) \, \mathrm{d}w \, \mathrm{d}y} \\ &= \frac{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y, w) \mathbb{1}_{[-b,b]} w \, \mathrm{d}w \, \mathrm{d}y}{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y, w) \mathbb{1}_{[-b,b]} \, \mathrm{d}w \, \mathrm{d}y} \\ &= \frac{\mathbb{E}\left[v'_t \mathbb{1}_{[-b,b]}\right]}{\mathbb{P}(x' \in [x - b, x + b])} \\ &= \mathbb{E}\left[v'_t | x' \in [x - b, x + b]\right] \end{aligned}$$

It makes no sense to consider $\mathbb{P}(x \in [x - b, x + b])$ as it is always equal to one so the auxiliary variable (x'_t, v'_t) is used. +++More detail?+++ This calculation also illustrates how φ describes the interaction or 'where to look' while M acts as a average velocity of the points 'within range'. The transport term, $\partial_v ([v - G(M(t, x))] f_t(x, v))$ then acts to change the velocity towards the local average given by M. It is in this sense that G is the herding function 'herding' the system towards a common velocity.

The McKean-Vlasov representation of the system is also the weak+++not weak limit, mention convergence of empirical measure to solution f +++ limit of an N-body particle system described by

$$\begin{cases} \mathrm{d}x_t^{i,N} = v_t^{i,N} \, \mathrm{d}t \\ \mathrm{d}v_t^{i,N} = -v_t^{i,N} \, \mathrm{d}t + G\left(\frac{\frac{1}{N}\sum_{j=1}^{N} \varphi(x_t^{i,N} - x_t^{j,N}) v_t^{j,N}}{\frac{1}{N}\sum_{j=1}^{N} \varphi(x_t^{i,N} - x_t^{j,N})}\right) \, \mathrm{d}t + \sqrt{2\sigma} \, \mathrm{d}W_t^i \end{cases} \quad , \qquad i = 1, \dots, N. \qquad (6)$$

This gives another view of the full system (2), and one that is particularly amenable to analysis. It also allows a more intuitive view of the system – rather than thinking of how a density changes over time, we can think of particles moving around a torus and interacting according to φ.

3.1 Invariant Measures under Space Homogeneity

Thus far, we have found the invariant measure for only the rather uninteresting case when $G \equiv 0$. To progress further, we must simplify the system. One way to do this is to impose that the density does not depend on the position in space and set $\varphi \equiv 1$. The weighted local average velocity M becomes neither local nor weighted. In the particle system, this means that each particle interacts with every other particle, irrespective of the distance between them. Indeed,

$$
\begin{aligned}
M(t,x) &= \frac{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y,w)\varphi(x-y)w \, \mathrm{d}w \, \mathrm{d}y}{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y,w)\varphi(x-y) \, \mathrm{d}w \, \mathrm{d}y} \\
&= \frac{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y,w)w \, \mathrm{d}w \, \mathrm{d}y}{\int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y,w) \, \mathrm{d}w \, \mathrm{d}y} \\
&= \int_{\mathbb{T}} \int_{\mathbb{R}} f_t(y,w)w \, \mathrm{d}w \, \mathrm{d}y \\
&:= \langle w \rangle_{f_t}.
\end{aligned}
$$

So the weighted local average velocity is just the average of the distribution $f_t(v)$ or equivalently, the average velocity of all particles in the system. The evolution is then described by

$$
\partial_t f_t(v) = \partial_v v f_t(v) - \partial_v G(\langle w \rangle_{f_t}) f_t(v) + \sigma \partial_{vv} f_t(v). \tag{7}
$$

The above problem's well-posedness can be shown as follows. Let

$$
h_t(v) = f_t(u), \text{ where } u = v + \int_0^t G(\langle w \rangle_{f_t}) \, \mathrm{d}s.
$$

To find an equation that $h_t(v)$ solves, we use the chain rule.

$$
\begin{aligned}
\partial_t f_t(u) &= \partial_t f_t(u) + G(\langle w \rangle_{f_t}) \partial_v f_t(u) && \text{as } \frac{\partial g}{\partial v} = 1, \frac{\partial g}{\partial t} = G((\langle w \rangle_{f_t}) \\
&= \partial_v v f_t(u) - \partial_v G(\langle w \rangle_{f_t}) f_t(u) + \sigma \partial_{vv} f_t(u) + G(\langle w \rangle_{f_t}) \partial_v f_t(u) && \text{as } f_t \text{ solves (7)} \\
&= \partial_v v h_t(v) + \sigma \partial_{vv} h_t(v), \qquad h_0(v) = f_0(v).
\end{aligned}
$$

So h_t solves the Fokker-Planck equation for the one-dimensional Ornstein-Uhlenbeck process which is well posed +++cite+++. Therefore the solution $f_t(v)$ of (7) exists and is unique given initial data. To find invariant measures of this evolution, set the transient term to zero and solve the remaining PDE, which can be written in gradient form. First note that if $G \equiv 0$, this system for f_t is the same as that for h_t, and so has a unique invariant measure $\mathcal{N}(0,\sigma)$. When the herding function is non-zero, other measures are also stationary as we will now show. For brevity, we omit the dependence on v of f and write $\langle w \rangle_{f_t} = M_1$. The latter is natural as it denotes the first moment of the distribution.

$$
\begin{aligned}
\partial_v \left[(-G(M_1) + v)f + \sigma \partial_v f \right] &= 0 \\
\implies (-G(M_1) + v)f + \sigma \partial_v f &= C_1
\end{aligned}
$$

This ODE is solvable and gives

$$f(v) = C_2 \exp\left(\frac{2G(M_1)v - v^2}{2\sigma}\right).$$

Applying the constraint $\int_{\mathbb{R}} f(w)\,\mathrm{d}w = 1$ gives $C_2 = \exp(-G^2(M_1))/\sqrt{2\pi\sigma}$. +++note here this is a Gaussian with mean $G(M_1)$? Immediately implies that $M_1 = G(M_1)$ without need for integration +++ Furthermore, from the definition of space average,

$$\begin{aligned}
M_1 &:= \int_{\mathbb{R}} w f(w)\,\mathrm{d}w \\
&= \int_{\mathbb{R}} \frac{w \exp(-G^2(M_1))}{\sqrt{2\pi\sigma}} \exp\left(\frac{2G(M_1)v - v^2}{2\sigma}\right) \\
&= \int_{\mathbb{R}} \frac{w}{\sqrt{2\pi\sigma}} \exp\left(\frac{-(v - G(M_1))^2}{2\sigma}\right) \\
&= G(M_1)
\end{aligned}$$

Using the smoothed version of the herding function from Section 2 gives $G(M_1) = 0, \pm 1$. Substituting this into our expression for f gives three invariant measures:

$$\rho_0 = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{v^2}{2\sigma}\right), \qquad \rho_\pm = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(v \pm 1)^2}{2\sigma}\right).$$

These are Gaussian densities with mean $0, \pm 1$ and variance σ.

We have found three invariant measures for the evolution (7), however whether the system will converge to these distributions given initial data remains to be settled. In the space homogeneous case it is possible to close equations on the moments of the distribution. Let $M_n = \int_{\mathbb{R}} v^n f_t(v)\,\mathrm{d}v$ denote the n$^{\text{th}}$ moment of the distribution $f_t(v)$. Then

$$\begin{aligned}
\partial_t M_n(t) &= \int_{\mathbb{R}} v^n \partial_t f_t(v)\,\mathrm{d}v \\
&= -\int_{\mathbb{R}} v^n \partial_v G(M_1) f_t(v)\,\mathrm{d}v + \int_{\mathbb{R}} v^n \partial_v v f_t(v)\,\mathrm{d}v + \sigma \int_{\mathbb{R}} v^n \partial_v f_t(v)\,\mathrm{d}v,
\end{aligned}$$

as f_t solves the space-homogeneous PDE (7). Using integration by parts and simplifying gives, for the first three moments,

$$\begin{aligned}
\dot{M}_0 &= 0, \\
\dot{M}_1 &= G(M_1) - M_1, \\
\dot{M}_2 &= 2[M_1 G(M_1) - M_2 + \sigma].
\end{aligned} \tag{8}$$

The first tells us there is no change in the zeroth moment over time, that is probability is conserved. The equation for the first moment has equilibria at -1, 0 and 1 for the smoothed choice of G. Figure +++ref+++ shows the phase plane diagram for this ODE. It shows that given any initial data with positive average velocity, the first moment will converge to one. By symmetry, the same is true for negative initial data, the mean will converge towards negative one. It also shows that mean zero is an unstable equilibrium – a slight perturbation will move the system from random movement to ordered motion. To find the convergence of the variance, we must look at the difference between

Figure 1: Phase plane of the first +++(and second?)+++ moment ODE

the second moment and the square of the first.

$$\frac{\mathrm{d}}{\mathrm{d}t}[M_2 - M_1^2] = \dot{M_2} - 2M_1\dot{M_1}$$
$$= 2[G(M_1)M_1 - M_2 + \sigma] - 2M_1[G(M_1) - M_1]$$
$$= -2M_2 + 2\sigma + 2M_1^2$$
$$= 2[M_2 - M_1^2 - \sigma]$$

Solving this ODE gives $M_2 - M_1^2 = \sigma + (\sigma_0 - \sigma)\mathrm{e}^{-2t}$, where σ_0 is the variance of the initial data f_0. Then as $t \to \infty$, it can be seen that the system forgets its initial data and tends towards a distribution with variance σ.

The equations on the moments have shown that the system indeed converges to a measure with mean ± 1 and variance σ however this is not sufficient to conclude that they are Gaussian. To do so requires closing equations on all the cumulants of the distribution. It can be shown that these vanish for all cumulants of order greater than three and so the system indeed converges towards the invariant measures found by solving the space-homogeneous PDE (7). This means that irrespective of initial data, we expect all particles to move on average in the same direction with the same velocity – a phenomenon known as unconditional flocking. The only exception is when the initial data has mean zero. In this case the system will converge to the Gaussian with mean zero, corresponding to disordered motion. +++not a great explanation+++

The dynamics of the space-homogeneous system have thus been fully characterised analytically. We will now turn our attention to the particle model in this case.

3.1.1 Particle System Dynamics

As the particle system approximates the McKean-Vlasov SDE, whose density solves the PDE (7), one would expect that it behaves in entirely the same way. This is in fact not the case, as we will now show. Consider the particle system (6) in the space-homogeneous case,

$$\mathrm{d}v_t^{i,N} = G\left(\frac{1}{N}\sum_{j=1}^{N} v_t^{j,N}\right)\mathrm{d}t - v_t^{i,N}\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i \tag{9}$$

It has been shown for the kinetic model that the invariant measures have mean zero and ± 1 when we consider a smoothed herding function. Here we will show that the space-homogeneous particle model only admits invariant measures with mean zero for any herding function satisfying the earlier requirements in Section 2, before finding analytic forms for the functions we will be using. First note that in the case of no interaction, that is $G \equiv 0$, the dynamics are governed by an Ornstein-Uhlenbeck process leading to a unique stationary distribution with mean zero. This is in agreement with the solution to the PDE under no interaction in Section 3. When $G \not\equiv 0$, the invariant distribution still always has mean zero, unlike what we have seen for the space-homogeneous kinetic model. Given that we wish to find the behaviour of the mean velocity, it would be pertinent to find an equation

that it solves. To this end, let $m_t = \frac{1}{N}\sum_{j=1}^{N} v_t^{j,N}$. Summing the system across i gives an equation for m_t:

$$\mathrm{d}m_t = [-m_t + G(m_t)]\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t$$

Let $V(x) = \frac{x^2}{2} + \tilde{V}(x)$, where $-\tilde{V}'(x) = G(x)$ so that $-V'(x) = -x + G(x)$. The aim here is to find a function V such that the equation for m_t resembles the Langevin equation. The existence and uniqueness of the invariant measure will then follow from standard theory for the Langevin equation and its associated Fokker-Planck equation. Thus

$$\mathrm{d}m_t = -V'(m_t)\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t. \tag{10}$$

`eq:meanSDE`

By definition, $\tilde{V}(x) = -\int_{-\infty}^{x} G(u)\,\mathrm{d}u$. As G is bounded, $\tilde{V}(x)$ grows at most linearly. This means $V(x) \to \infty$ as $|x| \to \infty$, V is a well-defined potential well and $\mathrm{e}^{-V(x)}$ is integrable. Hence (10) admits a unique invariant measure, $\rho(x) = \mathcal{Z}^{-1}\mathrm{e}^{-V(x)}$ with $\mathcal{Z}$ a normalising constant.

What can we say about this invariant measure? Recall the only restrictions placed on G in Section 2 were firstly G must be odd, that is $G(u) = -G(-u)$, and secondly,

$$\begin{cases} G(u) > u & \text{if } 0 < u < 1, \\ G(u) < u & \text{if } u > 1. \end{cases}$$

If G is odd, $\tilde{V}(x)$ is an even function, so $V(x)$ is also an even function. Then the mean of the invariant distribution ρ is

$$\mathbb{E}_\rho[m] = \mathcal{Z}^{-1}\int_{\mathbb{R}} m\mathrm{e}^{-V(m)}\,\mathrm{d}m = 0,$$

as this is the integral of an odd function. The behaviour of the particle system is then fundamentally different to that of its corresponding kinetic model, despite weakly+++check+++ converging to the McKean-Vlasov SDE (5) whose density solves the PDE (7). This is in fact a well-known problem in the area of mean field equations +++cite+++.

The behaviour of the particle system here is not wholly unexpected – the change in velocity is inherently random. To see how this affects the invariant measure, first recall the case when $G \equiv 0$, in which the particle system is the Ornstein-Uhlenbeck process and the kinetic model is its corresponding Fokker-Planck equation (3). Here, the particle moves in a potential well with a parabolic shape with a minima at zero. When $G \not\equiv 0$, the particles move within a double-well potential. For the choices of the herding function described here, this well will have minima at -1, 0 and 1[a]. Moving deterministically, as the evolution described by the kinetic model is, the particles must settle in one of the two wells leading to a system with mean 1, 0 or -1. Under random movement however, there is a chance that the Wiener process forces the particle to jump over the potential barrier and into the other well. If this happens to enough particles, the remaining particles will also make the jump due to the effect of the herding function. We thus expect, for a finite number of particles, the motion to switch randomly between having average velocity positive one and negative one leading to the invariant measure having mean zero. This switching of mean velocity is known as metastability – the Gaussian invariant measures for the PDE are only metastable in the particle model. We emphasise again that the dynamics under the kinetic model are fully deterministic and as such no switch of stability is possible. +++Dynkin's formula? Calc on exp time to switch?+++

---

[a]This is not quite true for any sigmoid G in which α is not large enough.

Figure 2: Particles in wells. Parabolic (OU) and Double Well showing jump between

	Herding Function	Invariant Measure
Step	$\frac{1+\mathrm{sgn}(u)}{2}$	+++not correct, not double well or normalised+++ $\exp\left(\frac{|v|}{2}\right)$
Smooth	$\frac{\arctan(u)}{\arctan(1)}$	$\frac{1}{(v^2+1)^{\frac{2}{\pi}}}\exp\left(-\frac{v^2}{2}+\frac{4}{\pi}v\arctan(v)\right)$
Sigmoid	$\frac{2}{1+\mathrm{e}^{-\alpha u}}-1$	$\mathcal{Z}^{-1}(\mathrm{e}^{\alpha v}+1)^{\frac{2}{\alpha}}\exp\left(-\frac{v^2}{2}-v\right)$

Table 1: Normalised densities for the space-homogeneous particle model (9) with different herding functions

The method used above to characterise the mean of the distribution gives all that is required to calculate closed forms of the invariant measure. Table 1 gives expressions for the three herding functions given in Section 2.

+++Convergence to invariant measure? Depends on metastability, rate of switching between wells +++

Figure 3: Potential Wells for the invariant measure of the space-homogeneous particle system for various herding functions

+++Full particle model invariant measure?+++
In the space-homogeneous case, the analysis above gives a full picture of the dynamics of both the particle model and the kinetic model. However, the full system (2) is much harder to analyse. It can be shown that the invariant measures for (7) are also invariant, however it cannot be shown that they are the only ones. In particular, it can not be expressed in a gradient form like it could be in the space-homogeneous case. The aim now is to exploit our understanding of the space-homogeneous case to develop robust numerical techniques to learn more about the full system.

4 Numerical Methods

The analysis presented so far has been able to fully characterise the behaviour of the space-homogeneous kinetic model. However, it is unable to describe the dynamics of the full space-inhomogeneous model. It has been shown that the space-homogeneous model possesses only three invariant measures and converges to one of them dependent on the sign of the mean of the initial data. The full model also has these invariant measures, however it is not proven whether these are all of them or not. That is, the full model may have stationary distributions which are not homogeneous in space. By utilising numerical methods, we can relax some of the restrictions which the analysis requires. The goal is to show that for some choice of initial data and interaction function, the

particle system will have a stationary distribution that isn't uniform in space. At its most basic, this could be two clusters of particles moving on the same direction diametrically opposite each other. The system will still converge in velocity to the values predicted by the homogeneous model, but spatial heterogeneity will remain. While this may be the case for the particle system, we conjecture that such behaviour will not occur in the kinetic model.

The aim is to be able to accurately simulate the dynamics of both the space-inhomogeneous PDE (2) and the corresponding interacting particle system. The latter is relatively simple to simulate using techniques for SDEs. However, we have shown analytically that the particle system does not have the same invariant measures as the continuum model, or indeed its McKean-Vlasov equation. The kinetic model is more difficult as it contains both advective and diffusive terms. We begin this section with a presentation of numerical methods for SDEs before reviewing methods for advection-type and diffusion-type equations, following the treatment of Hundsdorfer and Verwer [3] as well as Morton and Meyers [4].

Mirroring the approach in the analysis, we will develop the techniques required first for the space-homogeneous system with no interaction, before adding further levels of complexity. Doing so allows the rigorous testing of any schemes developed as analytic solutions are available. Throughout this section it is recommended to follow along with the supplementary Jupyter notebook available at:

$$\texttt{https://github.com/Tom271/InteractingParticleSystems}$$

Every figure within this report was produced using the notebook and its accompanying source code.

4.1   Particle Systems

Simulating particle systems requires the numerical solution of coupled SDEs. The standard method for this is the Euler-Maruyama (EM) method, which can be seen as the stochastic analogue of Euler's method. Our motivating example here will be the Ornstein-Uhlenbeck process[b].

$$\mathrm{d}x_t = -x_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t \tag{11}$$

This is exactly the same evolution as that which drives the space-homogeneous particle system (9), with no interaction between particles, and so provides the ideal starting point. It describes the motion of a damped particle moving randomly within a potential well as illustrated in Section 3.1.1. Applying the EM method to the above equation gives:

$$x_{n+1} = x_n - x_n\Delta t + \sqrt{2\sigma\Delta t}Z_n,$$

where Z_n is a standard normal random variable. This discretisation is very quick to compute, and has strong order $\frac{1}{2}$ [2]. This gives us a method of finding the stationary distribution: simulate many particles for a long time. Eventually, the density of the particles will be approximately the stationary distribution of the system. In this case it is in fact not even necessary to simulate many particles – if the system is allowed to evolve for long enough, one particle will suffice as there is no interaction between particles and the scheme is ergodic. The particle system will be an excellent metric by
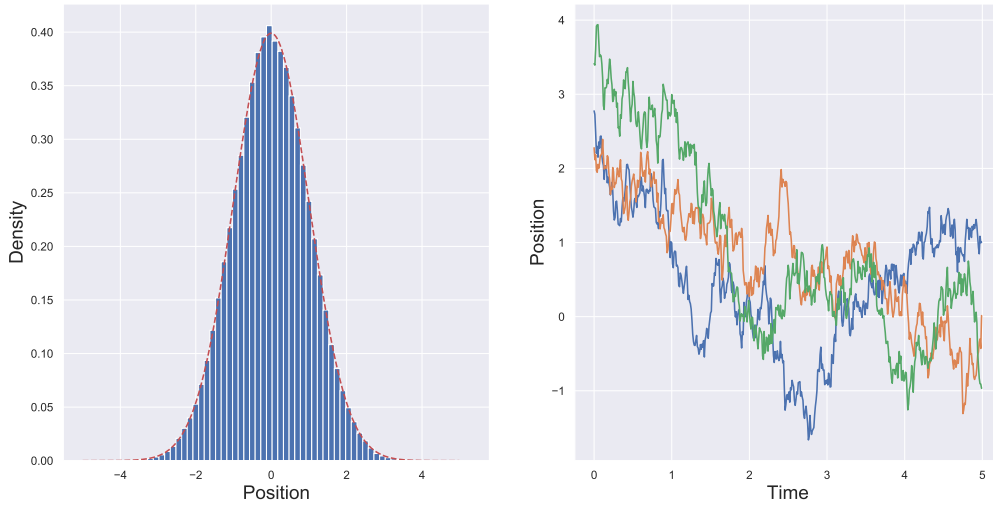
---

[b]This is also the overdamped Langevin equation

Figure 4: Histogram of positions of 1000 particles after 100s with $\sigma = 1$, and positions of 5 particles over time. +++conv in moments?+++

`articletraj`

which to test the schemes developed for the continuum models. One drawback however will be the presence of only one invariant measure, as was shown analytically in Section 3.1.1 and will be shown numerically in Section 5.1.

`sec:particledynamics`

`sec:homparticles`

4.2 Diffusion Equations

As a prototypical example of a diffusion equation, consider the heat equation in one dimension given by

$$\begin{cases} \partial_t u(t,x) = \sigma \partial_{xx} u(t,x), & \sigma > 0 \\ u(0,x) = u_0(x), & t \in \mathbb{R}^+, x \in \mathbb{R}, \end{cases} \tag{12}$$

`eq:heat`

To solve this numerically, we must truncate the domain in both time and space. To do this, a region is chosen where we expect most of the mass to be throughout the time of interest. Here, we truncate to $x \in [-L, L], t \in [0, T]$ for some $L > 0$ and a finite time horizon T. Doing so requires a boundary condition to be enforced. A natural choice for this equation is a zero Dirichlet condition, $u(t, L) = u(t, -L) = 0$. As long as L is chosen large enough, the heat which spreads beyond this limit is negligible – this will however be a source of error. To mitigate this, one could compare the difference between analytic solutions (where available) on the full real line and the truncated domain, then choosing L accordingly. In general, boundary conditions are determined by the phenomenon being modelled. The problem geometry may naturally suggest conditions, or they may be enforced by numerical constraints as in this case. We must also discretise in space and time, to create a mesh covering $[0, T] \times [-L, L]$. Let $\{x_j\}_{j=0}^{J}$ partition the space equally such that $x_0 = -L, x_J = L$ and $x_j - x_{j-1} = \Delta x$ for all j. Similarly, let $\{t_n\}_{n=0}^{N}$ partition $[0, T]$ in such a way that $t_0 = 0, t_N = T$ and $t_n - t_{n-1} = \Delta t$. The parameters $\Delta x, \Delta t$ are the space step and time step respectively. We thus have a discretised description of the continuum that we can implement.
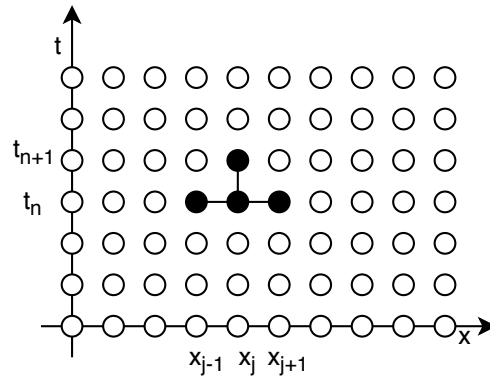
Figure 5: Points used in the FTCS method

Forward Time Centred Space

The aim is to solve the equation approximately on this grid. We shall denote approximate solutions by capital letters, that is $u(x_j, t_n) \approx U_j^n$. How can we approximate the solution? We must approximate each term in the equation separately. First the time derivative, or transient term can be approximated using the definition of the derivative. Recall

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \lim_{\Delta t \to 0} \frac{u(t + \Delta t) - u(t)}{\Delta t}.$$

We can then approximate the time derivative for fixed x by simply taking Δt small. Doing so leads to the forward difference approximation

$$\partial_t u(t_n, x_j) \approx \frac{U_j^{n+1} - U_j^n}{\Delta t}.$$

Now for the diffusion term, the most obvious approximation is to apply the forward difference scheme twice.

$$\partial_{xx} u(t_n, x_j) \approx \partial_x \left(\frac{U_{j+1}^n - U_j^n}{\Delta x} \right)$$

$$= \left(\frac{\partial_x U_{j+1}^n - \partial_x U_j^n}{\Delta x} \right)$$

$$\approx \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{(\Delta x)^2}$$

This is sufficient to give a first approximation to (12).

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \sigma \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{(\Delta x)^2} \tag{13}$$

$$U_j^{n+1} = U_j^n + \frac{\sigma \Delta t}{(\Delta x)^2}[U_{j+1}^n - 2U_j^n + U_{j-1}^n] \tag{14}$$

15

This is the Forward Time Centred Space (FTCS) method, a fully explicit scheme which means that we can write the next time step solely in terms of known values. We now have a method for approximating the true solution. When developing any numerical scheme, three properties need to be considered: stability, consistency and convergence. These concepts will be developed further below. Heuristically, a scheme is unstable if certain mesh spacings cause rapid growth in solutions – a blow-up. A method is consistent if the error it makes in each step tends to zero as the size of the step tends to zero. Finally, a scheme is convergent if the approximate solution tends towards the true solution for any point in space and time.

The heat equation ($\overset{\text{eq:heat}}{12}$) can be solved by separation of variables, giving the solution as a Fourier series. If the initial data is a single Fourier mode, the solution is

$$u(t,x) = \mathrm{e}^{-\sigma(2\pi k)^2 t}\varphi_k(x), \qquad\qquad \varphi_k(x) = \mathrm{e}^{2\pi i k x} \text{ for } k \in \mathbb{Z}.$$

If we assume our approximate solution is a Fourier mode, a criterion for stability will emerge.

$$U_j^n = \lambda^n \mathrm{e}^{ik(j\Delta x)}, \qquad k \in \mathbb{Z} \tag{15} \quad \boxed{\texttt{eq:fourierm}}$$

$$\implies U_j^{n+1} = \lambda U_j^n, \qquad U_{j\pm 1}^n = \mathrm{e}^{\pm ik\Delta x}U_j^n$$

$$\tag{16}$$

Clearly if $|\lambda| > 1$, the solution will be unstable. Substituting this into the finite difference scheme ($\overset{\text{eq:FTCSheat}}{13}$) gives

$$\lambda = 1 - 4\mu\sin^2\left(\frac{k\Delta x}{2}\right), \qquad \text{with } \mu = \frac{\sigma\Delta t}{(\Delta x)^2}, \tag{17} \quad \boxed{\texttt{eq:FTCSlamb}}$$

and so,

$$U_j^n = \left(1 - 4\mu\sin^2\left(\frac{k\Delta x}{2}\right)\right)^n \mathrm{e}^{ikj\Delta x}.$$

Looking at ($\overset{\text{eq:fouriermode}}{15}$), as $n \to \infty$, the numerical solution will grow unless $|\lambda| \leq 1$. If $|\lambda| \leq 1$, the solution will decay with higher modes being damped quicker. This is what we expect from our analytic knowledge of the heat equation. The restriction on the mesh spacing thus arises from ($\overset{\text{eq:FTCSlambda}}{17}$),

$$-1 \leq \lambda \leq 1 \implies \mu = \frac{\sigma\Delta t}{(\Delta x)^2} \leq \frac{1}{2}$$

This is not ideal: if more accuracy is required in the solution we can refine the space mesh, however in doing so the time step must also be shortened to maintain stability and the computational effort quickly increases.

The error made in each time step is the local truncation error, R, of the scheme. The local truncation error is the residual obtained by substituting the exact solution $u(t,x)$ into the discretisation. Performing a Taylor expansion around $t = n\Delta t$ gives

$$u_j^{n+1} = u_j^n + \Delta t\partial_t u_j^n + \frac{1}{2}(\Delta t)^2\partial_{tt}u_j^n + \mathcal{O}(\Delta t^3)$$

$$\implies \frac{u_j^{n+1} - u_j^n}{\Delta t} = \partial_t u_j^n + \frac{1}{2}\Delta t\partial_{tt}u_j^n + \mathcal{O}(\Delta t^2).$$

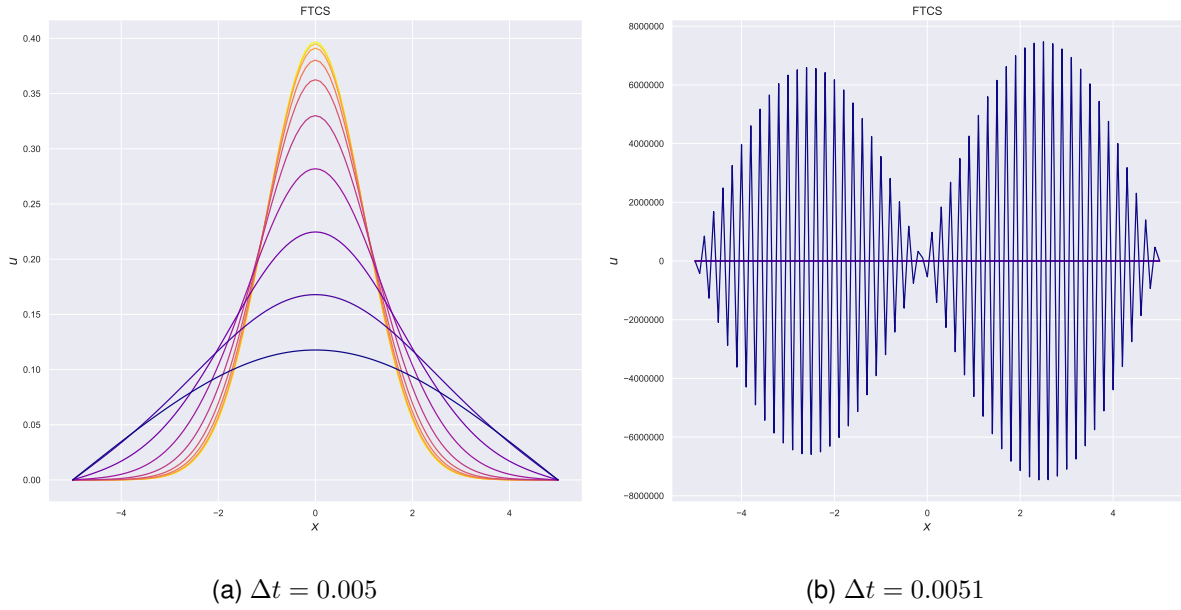(a) $\Delta t = 0.005$ (b) $\Delta t = 0.0051$

Figure 6: Using the FTCS scheme to solve the heat equation (12) on $[-5,5] \times [-1,1]$ with $\sigma = 1, \Delta x = 0.1$ with a Gaussian initial profile. When the stability condition is violated, $\lambda > 1$ and the solution grows exponentially.

And similarly around $x = j\Delta x$,

$$u_{j+1}^n = u_j^n + \Delta x \partial_x u_j^n + \frac{1}{2}(\Delta x)^2 \partial_{xx} u_j^n + \frac{1}{3!}(\Delta x)^3 \partial_{xxx} u_j^n + \frac{1}{4!}(\Delta x)^4 \partial_{xxxx} u_j^n + \mathcal{O}(\Delta x^5)$$

$$\implies \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{\Delta x^2} = \partial_{xx} u_j^n + (\Delta x)^2 u_{xxx} + \mathcal{O}(\Delta x^3).$$

The local truncation error is then

$$R_j^n = \underbrace{\partial_t u_j^n - \sigma \partial_{xx} u_j^n}_{=0} + \mathcal{O}(\Delta t) + \mathcal{O}(\Delta x^2)$$

$$\implies R_j^n = \mathcal{O}(\Delta t) + \mathcal{O}(\Delta x^2).$$

The first two terms are zero as u is solves the heat equation. The explicit scheme (13) is then first order in time and second order in space.

A numerical scheme is consistent if the local truncation error tends to zeros as $\Delta x, \Delta t \to 0$. The FTCS scheme is consistent of order one in time and two in space and stable so long as $\mu \leq \frac{1}{2}$. A scheme is convergent if for any fixed point $(x^*, t^*) \in [-L, L] \times [0, T]$,

$$x_j \to x^*, t_n \to t^* \implies U_j^n \to u(x^*, t^*).$$

The explicit scheme is convergent when it is stable. The stability condition is a great drawback of the FTCS method. In the next section we look to new methods to remove any restrictions on the mesh spacing, whilst maintaining (or improving) the truncation error.

17

Backward Time Centred Space

If instead of looking at the forward time difference we look at the backward time difference, the scheme becomes stable. This is surprising, very little appears to change but the scheme is fundamentally different. This method is implicit, meaning that its solution requires a more complex calculation at every time step. However, the lack of restriction on the size of the timestep means fewer steps (and therefore fewer calculations) need to be performed. The backward time centred space (BTCS) scheme for (12) is

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \sigma \frac{U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}}{(\Delta x)^2}. \tag{18}$$

The only difference here is the right hand side is calculated at the $n+1^{\text{th}}$ time step whereas in the FTCS algorithm it is taken at n. This scheme cannot be written explicitly in terms of known values, that is, the nth time step. Instead, we move all unknowns to the left hand side to give

$$-\mu U_{j-1}^{n+1} + (1 + 2\mu)U_j^{n+1} - \mu U_{j+1}^{n+1} = U_j^n$$

where, as before, $\mu = \frac{\Delta t}{(\Delta x)^2}$. This leads to a set of $J - 1$ simultaneous equations. As the system is tridiagonal, the Thomas algorithm can be used to efficiently solve the system. See Appendix B for a description.

The stability and consistency of the BTCS scheme can be found using the same techniques as was used for the explicit FTCS scheme. Assuming once again that the approximate solution as a Fourier mode, one obtains

$$\lambda = \frac{1}{1 + 4\mu \sin^2(\frac{1}{2}k\Delta x)}.$$

This is less than one for any positive μ, and so the BTCS method is unconditionally stable. The local truncation error of the scheme is the same as that of the FTCS method, namely the BTCS scheme is first order in time and second order in space. We now have a method with the same accuracy as the explicit scheme, but with no restriction on mesh size. The final method introduced will maintain this lack of constraint whilst increasing the accuracy, with little extra effort.

Theta Method

So far we have considered two methods, each using a set of three points to update. Given that we know how to solve implicit and explicit methods is there an optimal hybrid between the two? The theta method aims to find this optimum by taking a weighted average of the FTCS and BTCS schemes. Applied to our toy problem (12), the scheme is

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \sigma \left(\theta \frac{U_{j+1}^{n+1} - 2U_j^{n+1} + U_{j-1}^{n+1}}{(\Delta x)^2} + (1 - \theta) \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{(\Delta x)^2} \right),$$

where $0 \leq \theta 1$ is a weighting parameter. For $\theta = 0$, the FTCS scheme is recovered, and likewise when $\theta = 1$ the scheme is just the BTCS method. The aim of this section is to find an optimum weight for θ, if one exists. Optimum here means the value that decreases the truncation error
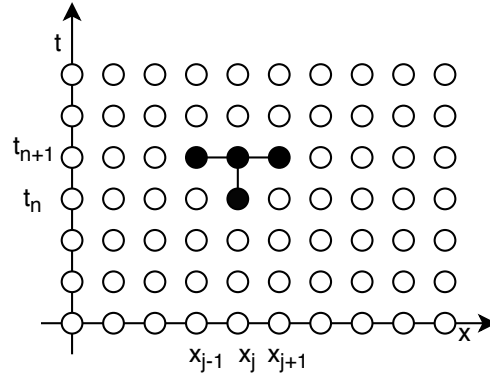
18

Figure 7: The points used in the implicit BTCS scheme

the most, whilst maintaining the unconditional stability of the BTCS method. Rearranging as a tridiagonal system gives

$$-\theta\mu U_{j-1}^{n+1} + (1 + 2\theta\mu)U_j^{n+1} - \theta\mu U_{j+1}^{n+1} = U_j^n + (1-\theta)\mu \left[U_{j+1}^n - 2U_j^n + U_{j-1}^n \right].$$

Using the stability analysis from the FTCS scheme, one obtains

$$\lambda = \frac{1 - 4(1-\theta)\mu \sin^2(\frac{1}{2}k\Delta x)}{1 + 4\theta\mu \sin^2(\frac{1}{2}k\Delta x)}.$$

Note again that if $\theta = 1$, λ is as it was for the BTCS scheme, and if $\theta = 0$, the value for the FTCS method is recovered. This is always less than one and is greater than minus one if

$$\mu(1 - 2\theta) \leq \frac{1}{2}.$$

Hence any choice of $\theta \geq \frac{1}{2}$ is sufficient for unconditional stability. This makes intuitive sense: the BTCS method is unconditionally stable, so any weighting preferring this scheme is also stable. If $\theta > \frac{1}{2}$, the mesh spacing can be adjusted to retain stability, as in the FTCS method. Thus far there is no reason to prefer the theta method over the simpler BTCS scheme. However, let's consider the truncation error of the scheme using similar analysis as above, expanding this time around the midpoint of the mesh, $(x_j, t_{n+\frac{1}{2}})$.

$$R_j^{n+\frac{1}{2}} = [\partial_t u - \partial_{xx} u] + \left[\left(\frac{1}{2} - \theta \right) \Delta t \partial_{xxt} u - \frac{1}{12}(\Delta x)^2 \partial_x^4 u \right]$$
$$+ \left[\frac{1}{24}(\Delta t)^2 \partial_t^3 u - \frac{1}{8}(\Delta t)^2 u_{xxtt} \right]$$
$$+ \left[\frac{1}{12} \left(\frac{1}{2} - \theta \right) \Delta t (\Delta x)^2 \partial_{xxxxt} u - \frac{2}{6!}(\Delta x)^4 \partial_x^6 u \right]$$

So for arbitrary choice of θ, the scheme remains first order in time and second order in space. However, if $\theta = \frac{1}{2}$, the first order in time term is zero, and the scheme is second order in both time and space. The optimal value of θ is the even weighting of the FTCS and BTCS scheme.
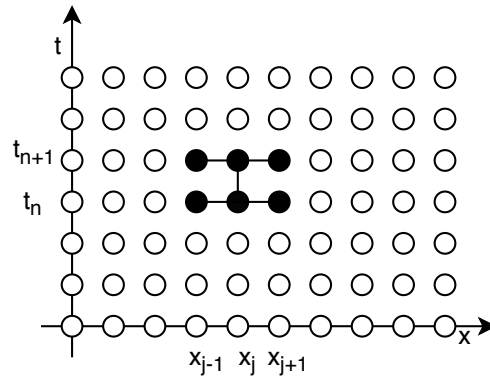
Figure 8: Points used in the Theta method

Method	FTCS ($\theta = 0$)	BTCS ($\theta = 1$)	Crank-Nicolson ($\theta = \frac{1}{2}$)
LTE	$\mathcal{O}(\Delta t), \mathcal{O}(\Delta x^2)$	$\mathcal{O}(\Delta t), \mathcal{O}(\Delta x^2)$	$\mathcal{O}(\Delta t^2), \mathcal{O}(\Delta x^2)$
Stablity Condition	$\mu \leq \frac{1}{2}$	None	None
Explicit/Implicit	Explicit	Implicit	Implicit

b:FDmethods

Table 2: Finite difference methods for diffusion equations

This method is known as the Crank-Nicolson (CN) method, and is the first choice when using finite difference methods – it is simple, quick and accurate to second order. Table 2 summarises the results of this section. The supplementary notebook contains all three solvers for the heat equation (12). For the model, a Crank-Nicolson scheme will be used.

4.3   Advection Equations

Our prototypical example in this case will be the one-way wave equation,

$$
\begin{cases}
\partial_t u(t,x) + a(x)\partial_x u(t,x) = 0, \\
u(0,x) = u_0(x), \qquad\qquad\qquad t \in \mathbb{R}^+, x \in \mathbb{R}.
\end{cases}
\tag{19}
$$

eq:wave

Advection equations are, in general, much more difficult to solve numerically than the diffusion equation. In this section we will consider a constant to illustrate the methods. When solving an advection equation numerically, any scheme must satisfy the CFL condition. To motivate the choice of scheme, we first introduce this condition.

The exact solution to an advection equation can be found by using the method of characteristics. Characteristics are lines on which the PDE (19) reduces to an ODE. Let

$$
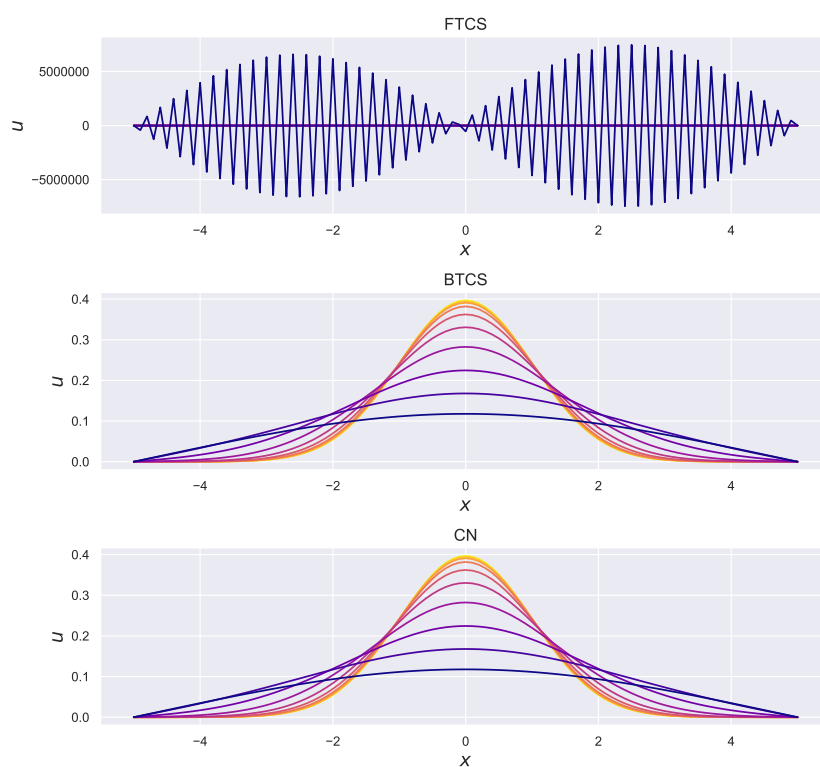\frac{\mathrm{d}x}{\mathrm{d}t} = a(x,t).
$$

fig:FTCSunstable

Figure 9: Solving the heat equation as in Figure 6 using FTCS, BTCS and CN solvers respectively. The implicit solvers remain stable regardless of step size, and return an accurate solution.

g:ftcsandcn

Then using the chain rule,

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \partial_t u + \frac{\mathrm{d}x}{\mathrm{d}t} \partial_x u = 0,$$

as u solves the PDE. Thus on a characteristic, the solution is constant. For the case when a is constant, the characteristics are the straight lines $x - at = const.$ and the solution is

$$u(x, t) = u_0(x - at).$$

The initial profile is simply transformed at a speed a. The solution at a point (x_j, t_n) is then obtained by finding the characteristic through this point and following it backwards in time to $t = 0$. If a discretisation is applied to equation (19), we obtain the explicit scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = -a \frac{U_{j+1}^n - U_j^n}{\Delta x}.$$

As in the FTCS method, to find the value of the solution at the next time point we evaluate

$$U_j^{n+1} = U_j^n - c(U_{j+1} - U_j),$$

where $c = \frac{a\Delta t}{\Delta x}$. Figure 10 shows the dependence of the current point on all previous points. Each point depends on two points from the previous time level, propagating all the way back to the initial line. This set of points, here a triangle, is the domain of dependence of the scheme. The CFL condition states that for a scheme to converge the characteristic curve passing through a point must stay within the domain of dependence for that point. In Figure 10, two example characteristics are plotted for $a > 0 (PQ)$ and for $a < 0 (PR)$. The explicit scheme outlined above is then only convergent for negative values of a, otherwise the characteristic lies outside of the domain of dependence. Intuitively, a scheme cannot be accurate if it cannot 'see' the wave arriving. If the wave is moving too fast, the CFL condition may be violated as the characteristic becomes much shallower and again lies outside the domain of dependence of the numerical scheme. This can be counteracted by reducing the timestep, effectively widening the base of the domain of dependence until it includes the characteristic. For this scheme this requires

$$c = \frac{a\Delta t}{\Delta x} < 1.$$

This condition ensures the solution does not travel further in one time step than one spacestep, thereby remaining in the domain of dependence. This condition is dependent on the geometry of the domain of dependence. The CFL condition is necessary but not sufficient for stability, as the following scheme shows. One way to satisfy the condition would be to make the triangle isosceles instead of right-angled. This can be achieved by taking a central difference in space:

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = -a \frac{U_{j+1}^n - U_{j-1}^n}{\Delta x}$$

The CFL condition here requires $|a|\frac{\Delta t}{\Delta x} < 1$. By applying the stability analysis used for the FTCS method, one obtains $\lambda = 1 - ci\sin(k\Delta x)$. This is always greater than one, and the scheme is thus unconditionally unstable, despite satisfying the CFL condition. The upwind scheme is the simplest stable method for advection equations. It switches the direction of the difference based on the direction of the wave, thereby always 'facing into the wave' (upwind).

$$U_j^{n+1} = \begin{cases} U_j^n - c(U_{j+1}^n - U_j^n) & \text{if } a < 0 \\ U_j^n - c(U_j^n - U_{j-1}^n) & \text{if } a > 0 \end{cases}$$
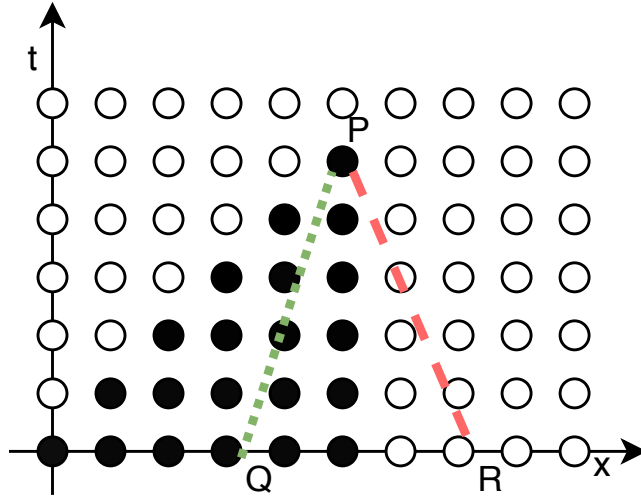
22

Figure 10

`fig:CFL`

Satisfying the CFL condition here requires $|a|\frac{\Delta t}{\Delta x} < 1$, and in fact stability analysis shows this is also sufficient for stability. Furthermore, the local truncation error can be calculated. Here we set $a > 0$.

$$R_j^n = \frac{u_j^{n+1} - u_j^n}{\Delta t} + a\frac{u_j^n - u_{j-1}^n}{\Delta x}$$
$$= \frac{1}{2}(\Delta t\partial_t^2 u - a\Delta x\partial_x^2 u) + \dots$$

So the upwind scheme is first order in both time and space. Higher order schemes are available for advection equations, however this will suffice for a first approximation to the model. There are problems, however, with artificial dissipation in this scheme, as Figure 11 shows. For a heuristic insight, we can consider the modified equation. By a Taylor expansion in the upwind scheme, we obtain

$$\frac{u_j^n - u_{j-1}^n}{\Delta x} = \partial_x u_j^n + \frac{1}{2}\Delta x\partial_x^2 u_j^n + \mathcal{O}(\Delta x^2).$$

This shows that the numerical solution will be more closely approximating an advection-diffusion equation:

$$\tilde{u}_t + a\tilde{u}_x = \frac{1}{2}a\Delta x\partial_x^2\tilde{u}.$$

The modified equation shows the source of the dissipative effects seen in the scheme.

Finite Volume Schemes

One aspect of the advection equation (19) the schemes thus far have not yet exploited is its conservation of mass. Finite volume methods ensure mass is conserved in the scheme, something not yet
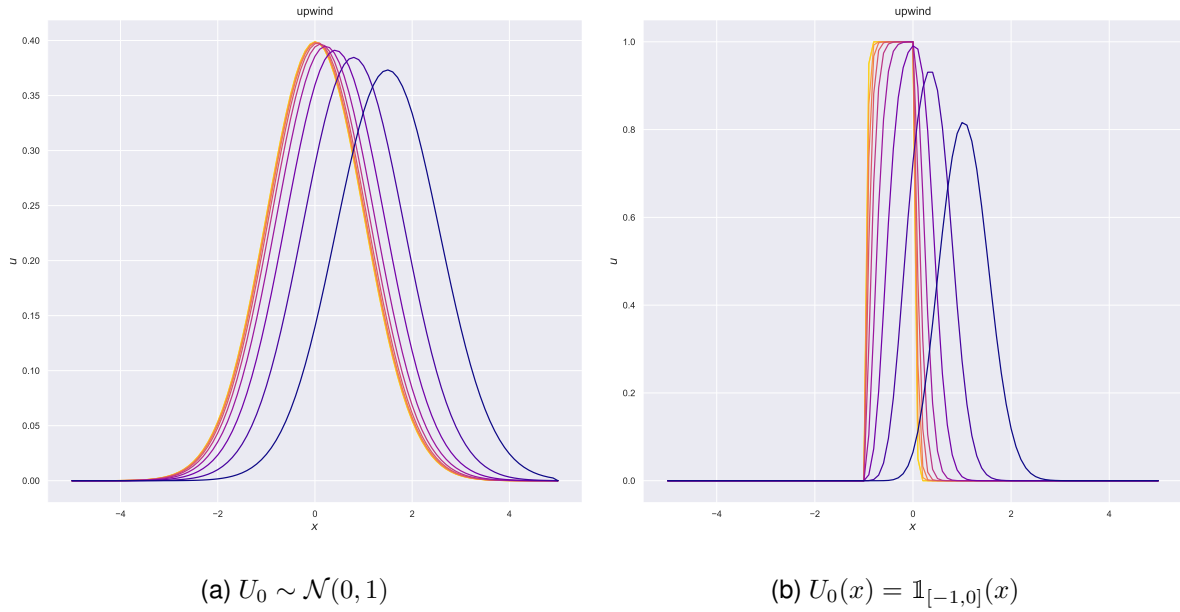
(a) $U_0 \sim \mathcal{N}(0,1)$ (b) $U_0(x) = \mathbb{1}_{[-1,0]}(x)$

Figure 11: Solving equation (19) with the upwind scheme. Even with very smooth initial data, dissipative effects can be severe.

g:advection

considered that can be a source of error. Finite volume schemes can be applied to any conservation (or continuity) equation. A conservative equation is one that can be written in the form

$$\partial_t u + \partial_x (a(x)u).$$

On the same mesh as previously, auxiliary points are introduced between points in space, namely $x_{j+\frac{1}{2}}$ for $0 \geq j \leq J - 1$. Within any interval (or volume), $\Omega_j = \left[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right]$, we can calculate the average density.

$$\bar{u}(x_j) = \frac{1}{\Delta x} \int_{\Omega_j} u(x) \, \mathrm{d}x = u(x_j) + \frac{(\Delta x)^2}{24} \partial_x^2 u(x_j) + \ldots$$

Furthermore, the change in the average density within the cell is equal to difference of the mass lost through the boundary of the volume, that is,

$$\Delta x \frac{\mathrm{d}\bar{u}(x_j)}{\mathrm{d}t} = a(x_{j-\frac{1}{2}})u(x_{j-\frac{1}{2}}) - a(x_{j-\frac{1}{2}})u(x_{j-\frac{1}{2}}),$$

This is also known as the flux across the boundary. A finite volume scheme balances fluxes across the domain, thereby maintaining total mass. Doing so leads to the first-order upwind scheme in flux form,

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = \frac{-1}{\Delta x} \left[a(x_{j+\frac{1}{2}})U_{j+\frac{1}{2}}^n - a(x_{j-\frac{1}{2}})U_{j-\frac{1}{2}}^n \right].$$

The value of a at the midpoint is chosen to be the average of its value at the cell centres, that is $a(x_{j+\frac{1}{2}}) = \frac{1}{2}\left(a(x_j) + a(x_{j+1})\right)$. The scheme as written will be unstable if $a(x) < 0$, for the same reason as in the previous section. The spatial bias needs to be changed depending on the direction of the wave.

$$a(x_{j+\frac{1}{2}})U_{j+\frac{1}{2}}^n = a^+(x_{j+\frac{1}{2}})U_j^n + a^-(x_{j+\frac{1}{2}})U_{j+1}^n,$$

24

where $a^+ = \max(a, 0), a^- = \min(a, 0)$. This scheme is still order 1, however it provides the ideal setting for developing higher order schemes.

5   Applying the Numerical Scheme

application

We now have enough methods to begin approximating both the particle system and the kinetic model. As in previous sections, we first focus on the particle system, before involving the continuum approach. Throughout this section the smooth G function has been used, although the other choices have also been implemented

5.1   Space-Homogeneous Particle Model

omparticles

Recall the homogeneous particle model (9),

eq:homparticle

$$\mathrm{d}v_t^{i,N} = G\left(\frac{1}{N}\sum_{j=1}^{n} v_t^{j,N}\right)\mathrm{d}t - v_t^{i,N}\,\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i. \tag{9}$$

eq:homparticle

Simulating this system is analagous to solving the Ornstein-Uhlenbeck process of Section ??. The
sec:particlemethods
interaction term is easily calculable using NumPy's `mean()` function as it is simply the average velocity of all the particles at each time step. The Euler-Maruyama scheme for this system is

$$v_{n+1}^{i,N} = v_n^{i,N} - v_n^{i,N}\Delta t + G\left(\frac{1}{N}\sum_{j=1}^{N} v_n^{j,N}\right)\Delta t + \sqrt{2\sigma\Delta t}Z_n^i.$$

As the system contains interactions, simulating one particle for a long time is no longer sufficient.
fig:homparticlemoments
Multiple particles must be simulated simultaneously. Figure 12 shows how the particle system behaves for a uniform initial distribution with negative mean. The mean velocity is herded towards -1, as predicted in the analysis. The variance behaves similarly, moving towards an average value of 1. Here the mean and variance of velocity is taken at each time step. If instead these statistics were taken across all time up to the point, by setting `timeavg = True`, the convergence is more clear. We can also plot an animation of the movement of the density of particles over time. Recall from
sec:dynamics
Section 3 that the empirical distribution of the particle system only approximates the true distribution as the number of particles tends to infinity. For any finite number of particles there is only one stationary distribution. When simulating the system, if the number of particles is high enough the probability that enough particles jump across the barrier is very small, and will only be seen if the simulation is ran for an extremely long time. Nevertheless, if the number of particles is low enough,
fig:switch
these switches happen often enough to observe, as Figure 13 shows. This something to be aware of when comparing the kinetic model with the particle model – a low number of particles produces a better approximation to the invariant distribution of the particle system, but a worse approximation of the kinetic model solution.

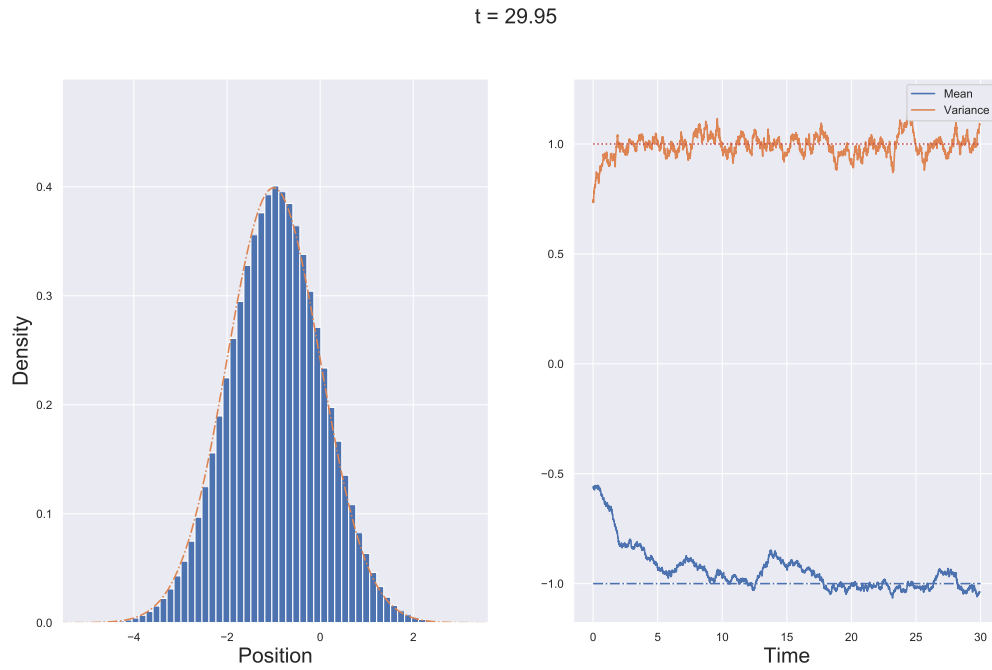+++List of things that can be done in code?+++

t = 29.95

Figure 12: Histogram of velocities of 1000 particles after 30s with $\sigma = 1$ against the analytic stationary distribution. A time step of $\Delta t = 0.01$ was used and the initial data was uniform on $[-2, 1]$. The right-hand plot shows the mean and variance moving towards the predicted values and oscillating about them.
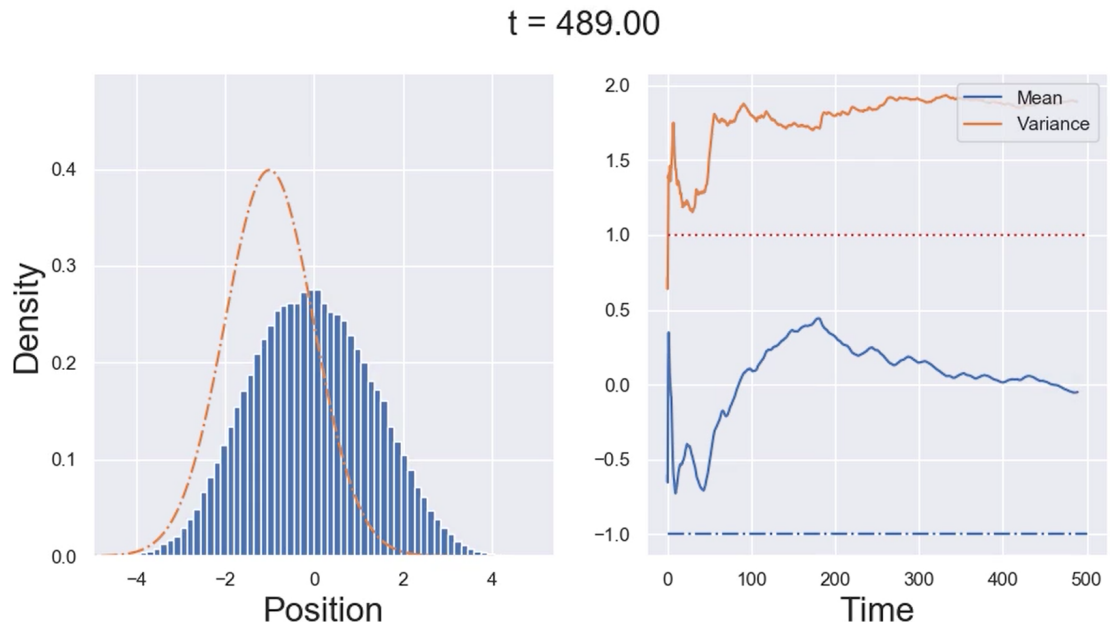
iclemoments

t = 489.00

Figure 13: When simulated for a long time with a low number of particles (here there are 6, and $T = 500$) the particles do not behave as the kinetic model predicts.

fig:switch

5.2 Space-Homogeneous Kinetic Model

sec:homkin

Consider the space-homogeneous evolution given by (7), that is

eq:spacehomPDE

$$\partial_t f_t(v) = \partial_v v f_t(v) - \partial_v G(\langle w \rangle_{f_t}) f_t(v) + \sigma \partial_{vv} f_t(v). \tag{7}$$

eq:spacehomPDE

Using the methods developed in Section 4, we are now able to approximate this system numerically. As when solving the heat equation, a zero boundary condition shall be enforced. This is valid as we know that the stationary distributions are Gaussian and centred at $-1, 0, +1$. The boundary L can then be chosen depending on the the diffusion so that almost no mass is contained beyond the boundary. For example if $\sigma = 1$, the mass contained beyond $L = 5$ is of the order 10^{-5}.

Simpson's rule will be used as a quick way to approximate the integral within the herding coefficient. To differentiate between the integral and its approximation, we write $\langle w \rangle_{F^n}$. Below is the scheme when both the herding coefficient and the velocity are positive, using a finite difference scheme with a CN discretisation for the diffusive term and an upwind method for the damping and herding terms.

$$\frac{F_j^{n+1} - F_j^n}{\Delta t} = -G(\langle w \rangle_{F^n}) \left[\frac{F_{j+1}^n - F_j^n}{\Delta v} \right] + \left[\frac{v_j F_j^n - v_{j-1} F_{j-1}^n}{\Delta v} \right]$$
$$+ \frac{\sigma}{2} \left[\frac{F_{j+1}^{n+1} - 2F_j^{n+1} + F_{j-1}^{n+1}}{(\Delta v)^2} + \frac{F_{j+1}^n - 2F_j^n + F_{j-1}^n}{(\Delta v)^2} \right]$$

This scheme will be first order in both time and space due to the use of the first-order upwinding. The stability will be restricted by the CFL condition, in particular, $\frac{|v_j - G(\langle w \rangle_{F^n})| \Delta t}{\Delta x} \leq 1$. A finite volume scheme has also been implemented. In this scheme the flux out of the right boundary is given by

$$\max(0, v_{j+\frac{1}{2}} - G(\langle w \rangle_{F^n})) F_{j+1}^n + \min(0, v_{j+\frac{1}{2}} - G(\langle w \rangle_{F^n})) F_j^n + \frac{\sigma(F_{j+1}^n - F_j^n)}{\Delta v}.$$

This is a simple FTCS method applied to the diffusion term and a first order upwind on the advective term. The stability here is limited by the mesh spacing, however it illustrates the use of a finite volume method without incorporating an implicit solver. Figure +++ref+++ is a frame from an animation of the PDE solution approximated using the finite difference scheme compared with the particle system. The finite difference method appears to closely follow the behaviour of the particle system, so long as the number of particles is high enough. Figure 15 shows the first moment and the variance contrasted with the analytic solution of (8). This clearly shows the almost immediate artificial dissipative effects of the upwind scheme in the finite volume case. The finite difference scheme appears to be better approximating the variance, however the spatial bias within this algorithm is not correct. Both schemes report no mass loss to 6 decimal places. This will be the case as long as the initial distribution does not have non-negligible mass at the boundary.

The error in the approximate stationary distribution can be approximated by simply finding the difference between the analytic invariant measure and the approximate solution. Figure 16 shows exactly this for $\sigma = 1, \Delta t = 0.001, T = 20, L = 6, \Delta v = 0.1$ and Gaussian initial data. The spatial bias in the finite difference method has introduced a bias in the error. No such bias exists for the upwinding in the finite volume solution, although the error is greater. This is very important for our aim in this project. If the initial data has mean zero, the scheme may always converge in the direction of the spatial bias, obscuring the true behaviour.
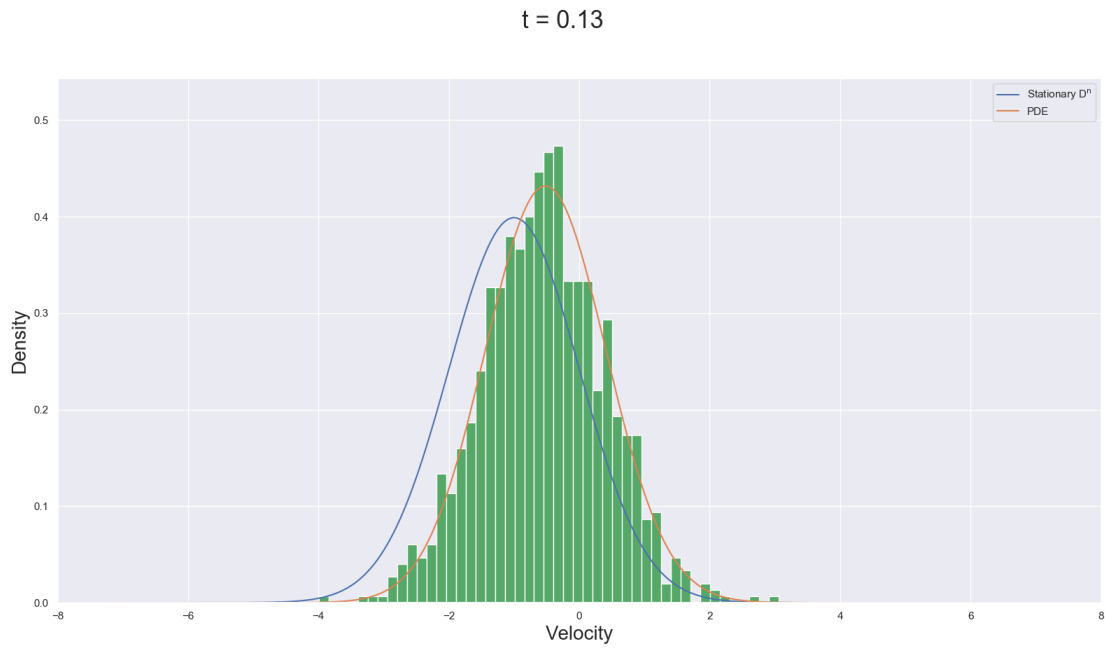
27

t = 0.13

PDEparticle

Figure 14: Frame from animation of finite difference model against the particle model histogram.
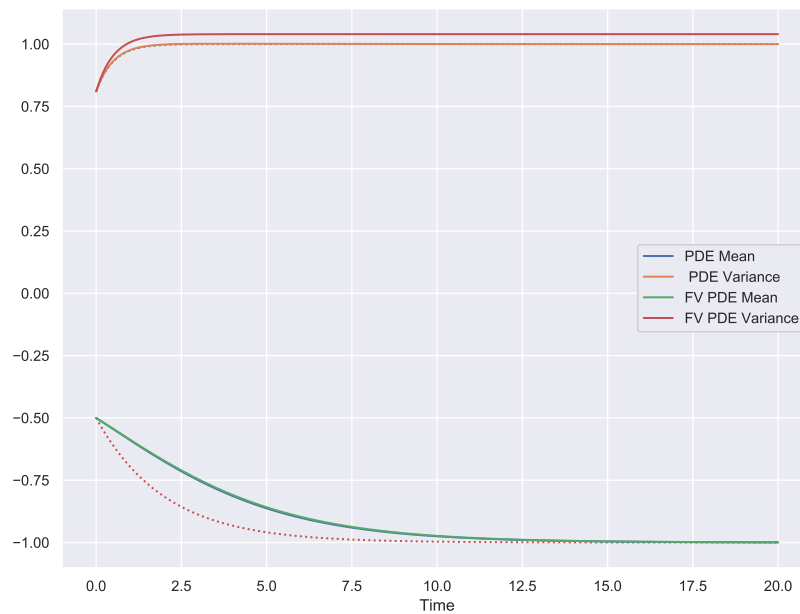

modelmomvar

Figure 15: For Gaussian initial data with mean -0.5 and variance 0.81, the schemes both converge with a time step of $\Delta t = 0.001$ and $\Delta v = 0.05$. For the mean, the solvers are almost indistinguishable. The time step must be small so the FTCS scheme within the finite volume solver is stable. Also note the overestimation of variance. Blue line shows the analytic stationary distribution.

Figure 16: Difference between approximate solution at $t = 20$ and a Gaussian distribution with mean -1 and variance 1. Note the spatial bias in the 'static' finite difference method has introduced a bias in the error. No such bias exists for the upwinding in the finite volume case.

homkinerror

5.3 Space-Heterogeneous Particle Model

sec:hetkin

The next step is to reintroduce the spatial heterogeneity in the particle system. Recall the full particle model is given by

$$
\begin{cases}
\mathrm{d}x_t^{i,N} = v_t^{i,N}\,\mathrm{d}t \\
\mathrm{d}v_t^{i,N} = -v_t^{i,N}\,\mathrm{d}t + G\left(\dfrac{\frac{1}{N}\sum_{j=1}^{N} \varphi(x_t^{i,N}-x_t^{j,N})v_t^{j,N}}{\frac{1}{N}\sum_{j=1}^{N} \varphi(x_t^{i,N}-x_t^{j,N})} \right)\mathrm{d}t + \sqrt{2\sigma}\,\mathrm{d}W_t^i
\end{cases}
\quad , \qquad i = 1, \ldots, N.
\tag{6}
$$

eq:fullparticle

Letting $\varphi \equiv 1$ forces the interaction to be homogeneous in space, however there is now the movement of particles in space given by the first equation. The numerical simulations reflect what the analysis showed: particles will spread from their initial distribution until they are uniform in space, and average velocity converges to $-1, 0, +1$ depending on the sign of the mean of the initial data, see Figure 17. As in the space-homogeneous case, it is possible (and certain, if the system is allowed to run long enough) that the particles randomly switch from clockwise to counterclockwise movement. The animation `switch.mp4` shows exactly this behaviour.

fig:hetjointplot

6 Discussion

:discussion

So far we have discussed and implemented methods for the space-homogeneous kinetic and particle models, as well as the particle model with the space variable included. The particle model simulation has shown to closely reflect the behaviour predicted by the analysis, and is a good approximation to the kinetic model when the number of particles is high enough. It is also very efficient, many particles can be simulated very quickly. However, when spatial heterogeneity is introduced, complexity will increase hugely as the function phi must be evaluated $N-1$ times for each of the N
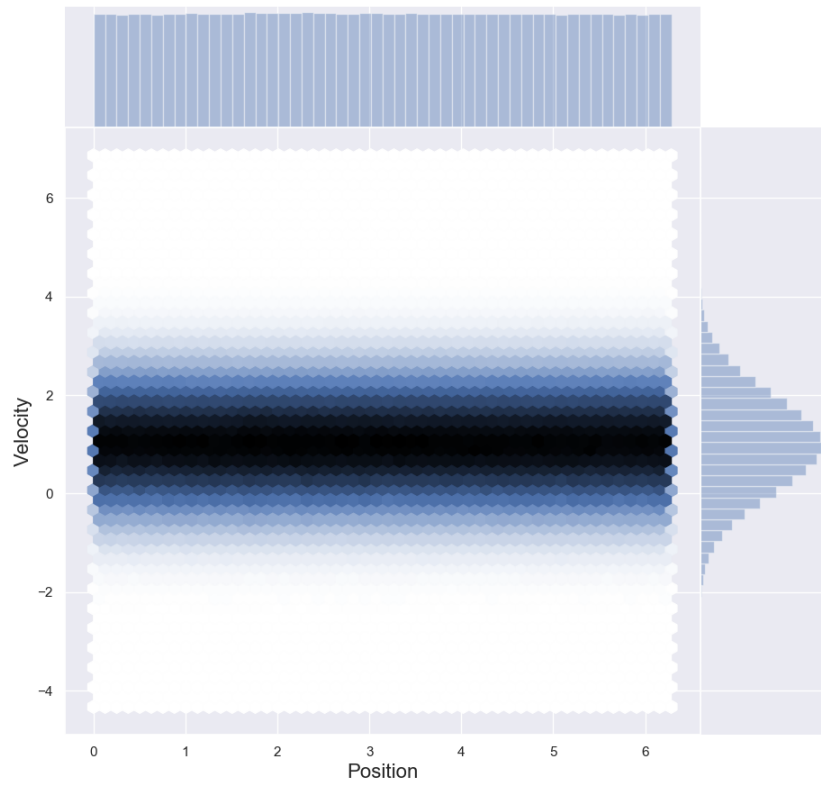
Figure 17: Hex plot of 100 particles' positions and velocities up to $T = 200$. This shows very clearly the uniformity in position and Gaussian distribution in velocity.
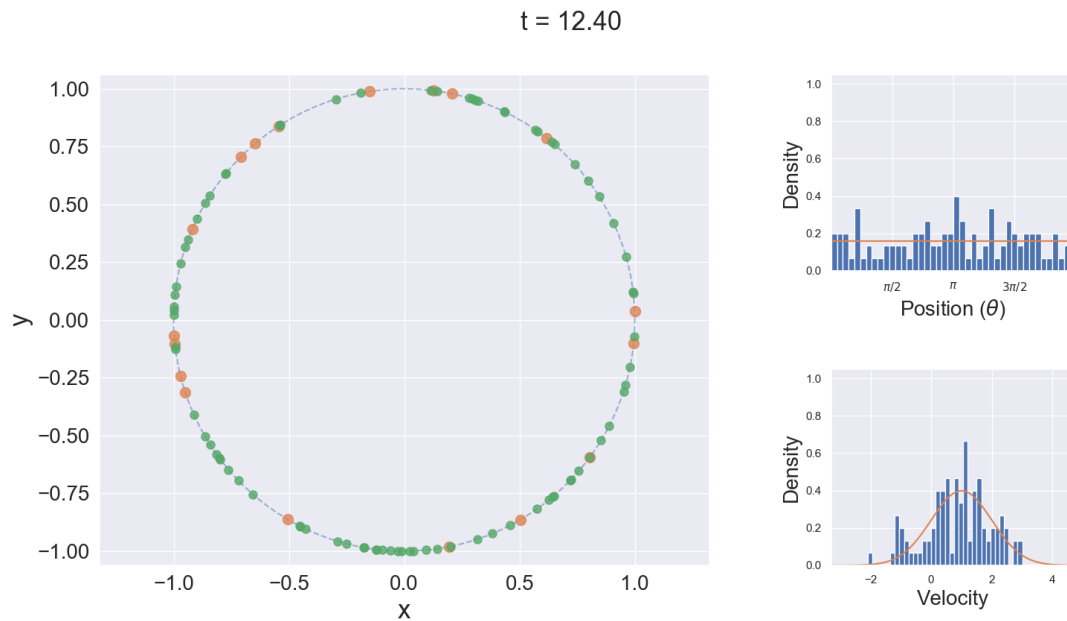
etjointplot

Figure 18: Particle model with position variable included. Green circles indicate those travelling with positive velocity, orange shows negative velocity. This is again visual confirmation of what the analysis predicted for $\varphi = 1$.

hetparticle

particles. In the kinetic model, φ appears only within the integral, which will be easier to numerically approximate. Therefore at this stage it is unwise to compare the two methods solely on computational speed. The next step in this direction is to include various choices of φ that do not satisfy the restrictions required for the analysis (1), to see if this has any effect on the convergence of the system.

Also, the error in the homogeneous kinetic approximation has not been rigorously checked. Theoretically the approximations will be first order in time and space, ideally one would have numeric confirmation of this. To this end, a higher order scheme will be implemented for the advection term in the finite difference method such as the third-order upwind scheme. Similarly, an implicit method used in the finite volume approach would be a simple way to increase accuracy. For example, a Crank-Nicolson scheme could be applied to the diffusive term. Finite volume methods provide many ways in which to solve the advection term whilst conserving mass, here we have only begun the exploration. The Lax-Wendroff scheme is second order in space and time, and would bring the truncation error of the advective term in line with that of the diffusion term. The change of variables described in Section 3.1 could also be utilised to assess the accuracy of the scheme. It gives another way of approximating the homogeneous system, only requiring the simulation of the Ornstein-Uhlenbeck process and the time integral of the herding function.

As well as consolidating the method for the space-homogeneous case, the method should be extended to solve the two-dimensional space-heterogeneous case. The one-dimensional finite volume scheme is the ideal base for this. This would allow us to ask and answer questions that the analysis cannot. For example, are their any combinations of interaction functions φ and initial distributions that mean the particle system will not homogenise in space? This would be in stark contrast to the behaviour of the kinetic model, and would suggest the existence of another invariant measure. We conjecture that a compactly supported interaction function, and an initial distribution in which two clusters of particles are placed diametrically opposite each other may produce such a measure. Figure 19 shows such a configuration.

References

[1] P. Buttà, F. Flandoli, M. Ottobre, and B. Zegarlinski. A non-linear kinetic model of self-propelled particles with multiple equilibria. *Kinetic & Related Models*, 12(4):791–827, 2019.

[2] D. Higham. An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, 43(3):525–546, 2001.

[3] W. Hundsdorfer and J. Verwer. *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Series in Computational Mathematics. Springer Berlin Heidelberg, 2007.

[4] K. Morton and D. Mayers. *Numerical Solution of Partial Differential Equations: An Introduction*. Cambridge University Press, 2005.
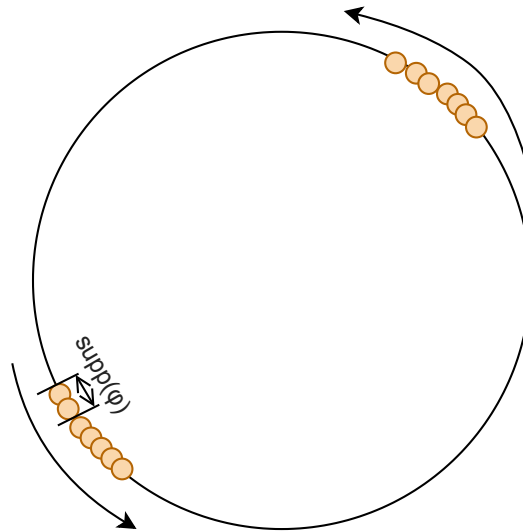
Figure 19: A possible inital configuration that may cause spatial heterogeneity.

A   Implementation

All the numerical schemes described above were implemented in Python 3. Before doing this project I had very little knowledge of Python beyond the basics, and a passing knowledge of MATLAB. As part of this project I have written a Python package to solve the toy problems, as well as the one-dimensional kinetic model and particle system. There is also a set of functions designed to produce animations and plots of the solutions. All of these are available and described further at

$$\texttt{https://github.com/Tom271/InteractingParticleSystems}$$

`Project Supplement.ipynb` shows the basic use of all the methods and how to produce plots and animations of the solution. The algorithms and functions themselves lie in the `src` folder as `.py` files, meaning they can be downloaded and used without Jupyter being installed.

B   Thomas Algorithm