Nonlinear Optimal Perturbations

Daniel Lecoanet

October 1, 2013

1 Introduction

In this report, I will describe a series of calculations employing nonlinear optimal perturbations. This technique was used by [4] to study transition to turbulence in shear flows. They searched for the lowest energy perturbation to the laminar state that would yield turbulence at late time. Since the kinetic energy of the perturbation is higher in the turbulent state than in the laminar state (where it is zero), they tried to maximize the perturbation kinetic energy at some late time T. This maximization was subject to the constraint that the initial condition has a given perturbation kinetic energy E_0. For E_0 lower than some threshold energy, E_c, they found that the optimization procedure was not able to find a turbulent state. However, for E_0 greater than E_c, the optimization procedure was able to find initial perturbations which evolved into turbulence. This suggests that E_c is the minimum energy required to trigger turbulence, and the perturbation with this energy that yields turbulence is referred to as the minimum seed.

To illustrate this technique, I will describe the nonlinear optimization procedure for a much simpler system: a system of 2 ODEs. The ODEs are

$$\partial_t x_1 = -x_1 + 10x_2 \tag{1}$$
$$\partial_t x_2 = x_2(x_2 - 1)\left(10\exp(-x_1^2/100) - x_2\right). \tag{2}$$

This system has two stable fixed points, one at $\boldsymbol{x} = \boldsymbol{x}_l \equiv \boldsymbol{0}$, and the other at $\boldsymbol{x} = \boldsymbol{x}_t \approx (14.0174, 1.40174)$. To make an analogy to the transition to turbulence problem, the former can be thought of as the laminar state, and the latter as the turbulent state. It is straightforward to check that the basin of attraction of the laminar state is the region with $x_2 < 1$, and the basin of attraction of the turbulent state is the region with $x_2 > 1$. The analogous problem to finding the minimum seed is then to maximize x_2 at some late time T, subject to the constraint that the initial perturbation from $\boldsymbol{0}$ has norm $|\boldsymbol{x}|^2 = E_0$. I will now describe how to perform this optimization.

Consider the functional $\mathcal{L}$ given by

$$\mathcal{L} = x_2(T) + \alpha\left(|\boldsymbol{x}(0)|^2 - E_0\right) + \int_0^T dt\ \boldsymbol{\nu}(t)\cdot\left(\partial_t\boldsymbol{x}(t) - \boldsymbol{f}(\boldsymbol{x}(t))\right), \tag{3}$$

where α and $\boldsymbol{\nu}(t)$ are Lagrange multipliers imposing the constraints that $\boldsymbol{x}(0)^2 = E_0$ and that $\boldsymbol{x}(t)$ satisfies the system of ODEs (1 & 2). $\boldsymbol{\nu}(t)$ are referred to as the adjoint variables.

Varying $\mathcal{L}$ yields

$$\begin{aligned}\delta\mathcal{L} = \quad & \delta x_2(T) + \delta\alpha\left(|\boldsymbol{x}(0)|^2 - E_0\right) + 2\alpha\boldsymbol{x}(0)\cdot\delta\boldsymbol{x}(0) \\ & + \int_0^T dt\ \delta\boldsymbol{\nu}(t)\cdot\left(\partial_t\boldsymbol{x}(t) - \boldsymbol{f}(\boldsymbol{x}(t))\right) \\ & + \int_0^T dt\ \boldsymbol{\nu}(t)\cdot\left(\partial_t\delta\boldsymbol{x}(t) - \frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}}\cdot\delta\boldsymbol{x}(t)\right). \end{aligned} \tag{4}$$

The partial derivatives of $\mathcal{L}$ are thus

$$\frac{\delta\mathcal{L}}{\delta\alpha} = |\boldsymbol{x}(0)|^2 - E_0, \tag{5}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{\nu}(t)} = \partial_t\boldsymbol{x}(t) - \boldsymbol{f}(\boldsymbol{x}(t)), \tag{6}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}(T)} = (0,1) + \boldsymbol{\nu}(T), \tag{7}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}(t)} = -\partial_t\boldsymbol{\nu}(t) - \boldsymbol{\nu}\cdot\frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}}, \tag{8}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}(0)} = 2\alpha\boldsymbol{x}(0) - \boldsymbol{\nu}(0). \tag{9}$$

If all these conditions are satisfied, then $\boldsymbol{x}(0)$ maximizes $x_2(T)$ subject to $|\boldsymbol{x}(0)|^2 = E_0$. (8) is an evolution equation (backward in time) for the adjoint variables, and is referred to as the adjoint equation.

The optimization procedure is an iterative algorithm which uses the above expressions to update a guess for $\boldsymbol{x}(0)$, call it $\boldsymbol{x}^{(0)}(0)$, to another initial condition, $\boldsymbol{x}^{(1)}(0)$, which has a larger $x_2(T)$ than the first guess.

First, pick $\boldsymbol{x}^{(0)}(0)$ which satisfies $|\boldsymbol{x}^{(0)}(0)| = E_0$ (5). Then use (6) to integrate $\boldsymbol{x}(t)$ from $t = 0$ to $t = T$. Then use (7) to set $\boldsymbol{\nu}(T)$, which in this case is always equal to $(0,-1)$, but in general will depend on $\boldsymbol{x}(T)$. Next, (8) is integrated backward in time to get $\boldsymbol{\nu}(0)$. Finally, if $\boldsymbol{\nu}(0)$ is parallel to $\boldsymbol{x}^{(0)}(0)$, (9) shows that $\boldsymbol{x}^{(0)}(0)$ maximizes $x_2(T)$. If $\boldsymbol{\nu}(0)$ is not parallel to $\boldsymbol{x}^{(0)}(0)$, then use (9) to update to a new initial perturbation using steepest ascent,

$$\boldsymbol{x}^{(1)}(0) = \boldsymbol{x}^{(0)}(0) + \epsilon\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}(0)}, \tag{10}$$

where ϵ sets the size of the update. For sufficiently small ϵ, the initial condition $\boldsymbol{x}^{(1)}(0)$ will lead to a larger $x_2(T)$ than $\boldsymbol{x}^{(0)}(0)$. Note that for arbitrary ϵ and α, $\boldsymbol{x}^{(1)}(0)$ does not satisfy the norm constraint (5) $|\boldsymbol{x}^{(1)}(0)|^2 = E_0$. Rather, α must be chosen to satisfy this constraint.

In the remainder of this report, I will apply this technique to two new problems. The first is an extension of the optimization procedure to allow for multiple perturbations, instead

of just a single perturbation at $t = 0$. The second is an application of the optimization to the Vlasov-Poisson equations.

2 Multiple Perturbations

For simplicity, I will consider the multiple perturbation problem only in the context of 2D ODE systems. However, the approach is easily generalized to more complicated ODE or PDE systems.

To be general, assume that I will maximize the quantity $\phi(\boldsymbol{x}(T))$, for some function ϕ. I want to derive a procedure for finding the set of perturbations $\boldsymbol{\xi}_i$, $0 \leq i \leq n-1$, which act on $\boldsymbol{x}$ at

$$0 = T_0 < T_1 < \cdots < T_{n-1}, \tag{11}$$

with $T_{n-1} < T = T_n$, which maximize $\phi(\boldsymbol{x}(T))$. I will also limit the size of the perturbations $\boldsymbol{\xi}_i$ using some constraint $N(\boldsymbol{\xi}_i) = 0$, e.g., $N(\boldsymbol{\xi}_i) = \sum |\boldsymbol{\xi}_i|^2 - E_0$. This is a generalization of the single perturbation case, in which $n = 1$.

To perform the optimization analysis, I will also split the dependent variable $\boldsymbol{x}$ into n parts, $\boldsymbol{x}_i(t)$ with $0 \leq i \leq n-1$, where $\boldsymbol{x}_i(t)$ is defined between $T_i \leq t \leq T_{i+1}$. I take $\boldsymbol{x}_0(T_0) = \boldsymbol{\xi}_0$, and for $1 \leq i \leq n-1$, $\boldsymbol{x}_i(T_i) = \boldsymbol{x}_{i-1}(T_i) + \boldsymbol{\xi}_i$. Thus, I am maximizing $\phi(\boldsymbol{x}_{n-1}(T)) = \phi(\boldsymbol{x}_{n-1}(T_n))$. To simplify the equations, I will define $\boldsymbol{x}_{-1}(T_0) = \boldsymbol{x}_s$.

The optimization will follow from extremizing

$$\begin{aligned}
\mathcal{L} &= \phi(\boldsymbol{x}_{n-1}(T_n)) + \alpha N(\boldsymbol{\xi}_i) \\
&+ \sum_{i=0}^{n-1} \boldsymbol{\beta}_i \cdot (\boldsymbol{x}_i(T_i) - \boldsymbol{x}_{i-1}(T_i) - \boldsymbol{\xi}_i) \\
&+ \sum_{i=0}^{n-1} \int_{T_i}^{T_{i+1}} dt\ \boldsymbol{\nu}_i(t) \cdot (\partial_t \boldsymbol{x}_i - \boldsymbol{f}(\boldsymbol{x}_i))\,.
\end{aligned} \tag{12}$$

The variation of $\mathcal{L}$ is

$$\begin{aligned}\delta\mathcal{L} = \quad & \frac{\partial\phi}{\partial \boldsymbol{x}} \cdot \delta\boldsymbol{x}_{n-1}(T) + \delta\alpha N(\boldsymbol{\xi}_i) + \alpha \sum_{i=0}^{n-1} \frac{\partial N}{\partial \boldsymbol{\xi_i}} \cdot \delta\boldsymbol{\xi}_i \\ & + \sum_{i=0}^{n-1} \delta\boldsymbol{\beta}_i \cdot (\boldsymbol{x}_i(T_i) - \boldsymbol{x}_{i-1}(T_i) - \boldsymbol{\xi}_i) \\ & + \sum_{i=0}^{n-1} \boldsymbol{\beta}_i \cdot (\delta\boldsymbol{x}_i(T_i) - \delta\boldsymbol{x}_{i-1}(T_i) - \delta\boldsymbol{\xi}_i) \\ & + \sum_{i=0}^{n-1} \int_{T_i}^{T_{i+1}} dt\ \delta\boldsymbol{\nu}_i(t) \cdot (\partial_t \boldsymbol{x}_i - \boldsymbol{f}(\boldsymbol{x}_i)) \\ & + \sum_{i=0}^{n-1} \int_{T_i}^{T_{i+1}} dt\ \boldsymbol{\nu}_i(t) \cdot \left(\partial_t \delta\boldsymbol{x}_i - \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}_i} \cdot \delta\boldsymbol{x}_i\right). \end{aligned} \tag{13}$$

Thus, the partial derivatives are

$$\frac{\delta\mathcal{L}}{\delta\alpha} = N(\boldsymbol{\xi}_i), \tag{14}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{\nu}_i(t)} = \partial_t \boldsymbol{x}_i - \boldsymbol{f}(\boldsymbol{x}_i), \tag{15}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{\beta}_i} = \boldsymbol{x}_i(T_i) - \boldsymbol{x}_{i-1}(T_i) - \boldsymbol{\xi}_i, \tag{16}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}_{n-1}(T)} = \frac{\partial\phi(\boldsymbol{x}_{n-1}(T))}{\partial\boldsymbol{x}_{n-1}(T)} + \boldsymbol{\nu}_{n-1}(T), \tag{17}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}_i(t)} = -\partial_t\boldsymbol{\nu}_i(t) - \boldsymbol{\nu}_i \cdot \frac{\partial\boldsymbol{f}}{\partial\boldsymbol{x}_i}, \tag{18}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}_{i-1}(T_i)} = -\boldsymbol{\beta}_i + \boldsymbol{\nu}_{i-1}(T_i), \tag{19}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{x}_i(T_i)} = \boldsymbol{\beta}_i - \boldsymbol{\nu}_i(T_i), \tag{20}$$

$$\frac{\delta\mathcal{L}}{\delta\boldsymbol{\xi}_i} = \alpha\frac{\partial N}{\partial\boldsymbol{\xi}_i} + \boldsymbol{\beta}_i. \tag{21}$$

The optimization algorithm now follows from these partial derivatives. As before, start with an initial set of perturbations $\boldsymbol{\xi}_i^{(0)}$ satisfying the norm condition (14). Integrate $\boldsymbol{x}(t)$ (for simplicity, I drop the subscripts) forward in time using (15), adding in the perturbations at the times T_i according to (16). Set $\boldsymbol{\nu}(T)$ using (17) and $\boldsymbol{x}(T)$. Then integrate $\boldsymbol{\nu}$ back in time to $t = 0$ using (18). (19) and (20) imply that $\boldsymbol{\nu}(t)$ is continuous at T_i. To update $\boldsymbol{\xi}_i$, I use steepest ascent along with (21), identifying $\boldsymbol{\beta}_i$ with $\boldsymbol{\nu}(T_i)$ (20).

2.1 Changing the times of the perturbations, T_i

I also change T_i for $1 \leq i \leq n-1$ to better optimize the functional. The variation of $\mathcal{L}$ with T_i is

$$\frac{\delta \mathcal{L}}{\delta T_i} = \boldsymbol{\beta}_i \cdot \left(\left. \frac{\partial \boldsymbol{x}_i}{\partial t} \right|_{T_i} - \left. \frac{\partial \boldsymbol{x}_{i-1}}{\partial t} \right|_{T_i} \right), \tag{22}$$

or,

$$\frac{\delta \mathcal{L}}{\delta T_i} = \boldsymbol{\beta}_i \cdot \left(\boldsymbol{f}(\boldsymbol{x}_i) - \boldsymbol{f}(\boldsymbol{x}_{i-1}) \right). \tag{23}$$

T_i is updated using steepest ascent.

2.2 Numerical Experiments

I found nonlinear optimal sets of perturbations for two systems of second order ODEs. The first set are (1 & 2). The optimal perturbation for these equations is $\boldsymbol{x} = (0, 1+\epsilon)$. The optimization algorithm has no difficulty in finding this optimal perturbation. When considering multiple perturbations, it is important to pick the current norm for describing the size of the set of perturbations. A naive choice would be to use the sum of magnitudes, i.e.,

$$N(\boldsymbol{\xi}_i) = \sum_i |\boldsymbol{\xi}_i|^2. \tag{24}$$

However, this norm has the flaw that $N(\{(0,1)\}) = 1$, but $N(\{(0,\frac{1}{2}),(0,\frac{1}{2})\}) = \frac{1}{2}$, where $\{\boldsymbol{\xi}_i\}$ denotes the set of perturbations. That is to say, having multiple perturbations that act at the same time and in the same direction is more efficient than a single perturbation. To avoid this issue, I use the magnitude of sum norm,

$$N(\boldsymbol{\xi}_i) = \left| \sum_i \boldsymbol{\xi}_i \right|^2. \tag{25}$$

Using this norm, I find that having more than one perturbation does not change the minimum norm required to perturb the system into the attractor of the second fixed point $\boldsymbol{x}_t$. The minimum seed is the single perturbation $(0, 1+\epsilon)$, which has energy $1+2\epsilon$.

The reason that multiple perturbations do not help the system transition to the second fixed point $\boldsymbol{x}_t$ is that $\partial_t x_2 < 0$ between $0 < x_2 < 1$ near $\mathbf{0}$ – $\partial_t x_2$ only becomes positive for $0 < x_2 < 1$ when $|x_1| \gtrsim 15.17$. This implies that the ODE flows downwards near $\mathbf{0}$. Having multiple perturbations is thus not efficient, because the perturbations have to fight against the ODE flow. Thus, putting all the energy into one big perturbation is the most efficient way of pushing the system into the attractor of the second fixed point $\boldsymbol{x}_t$.

With this in mind, I have also investigated the related ODE system,

$$\partial_t x_1 = -x_1 + 10x_2 \tag{26}$$

$$\partial_t x_2 = x_2(x_2-1)\left(20\exp(-(x_1-7)^2/14) - 10x_2\right). \tag{27}$$

As for (1 & 2), these ODEs also possess two linearly stable fixed points, one at $\boldsymbol{x} = \boldsymbol{x}_l \equiv \mathbf{0}$, and the other at $\boldsymbol{x} = \boldsymbol{x}_t \approx (10.094, 1.0094)$. As before, the boundary between the basins of attraction of the two fixed points is $x_2 = 1$. This ODE system was chosen such that there is a region with $x_2 > 0$ and $|\boldsymbol{x}|^2 < 1$ which has $\partial_t x_2 > 0$, i.e., the ODE can flow the system toward the boundary at $x_2 = 1$.

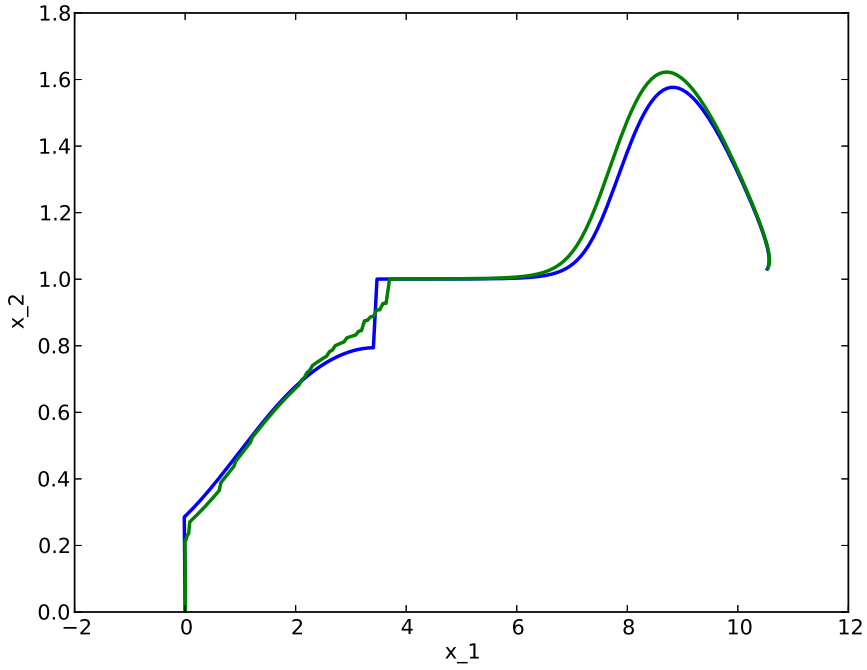

Figure 1: Trajectories for optimal sets of perturbations for ODEs (26 & 27), for two perturbations (blue) and fifty perturbations (green).

As before, optimizing over one perturbation yields a minimum seed of $(0, 1+\epsilon)$, having energy of $1 + 2\epsilon$. However, optimizing over two perturbations allows the system to take advantage of the upwards flow of the ODE. In Figure 1, I have plotted the trajectories of the optimal set of perturbations, for both two and fifty perturbations. The (mostly vertical) discontinuities are the perturbations. Looking at the blue curve, there is an initial perturbation upwards from $\boldsymbol{x} = \mathbf{0}$. This is followed by allowing the ODE flow to the point where $\partial_t x_2 = 0$, at which point a second perturbation is used to push the system above $x_2 = 1$. Once $x_2 > 0$, the system is in the basin of attraction of the second fixed point $\boldsymbol{x}_t$, and flows into this fixed point. This set of perturbations has a "magnitude of sum" norm of 0.485, about a factor of two smaller than the single perturbation optimal.

Interestingly, the optimal set of perturbations for fifty perturbations is similar to the optimal set for two perturbations. As before, there is a large perturbation near $\boldsymbol{x} = \mathbf{0}$ to push the system into the region of phase space where $\partial_t x_2 > 0$. This is followed by perturbations near the point at which $\partial_t x_2 = 0$ that push the system above $x_2 = 1$. Furthermore, the "magnitude of sum" norm of this set of perturbations is also 0.485, i.e.,

it does no better than two perturbations. For this system of ODEs, two perturbations is enough to realize the minimum seed. Presumably, one could cook up a slightly more complicated ODE system which has two disconnected regions where $\partial_t x_2 > 0$ between $x_2 = 0$ and $x_2 = 1$ – in this case, optimal sets of perturbations would require at least three perturbations: the first perturbing the system from $\boldsymbol{x} = \boldsymbol{0}$ to the first region where $\partial_t x_2 > 0$, the second moving the system from the first region where $\partial_t x_2 > 0$ to the second such region, and then the third perturbation pushing the system above $x_2 = 1$. In a real fluids system, this would correspond to using several distinct mechanisms to amplify the energy of a perturbation sufficiently to access a new nonlinear state.

2.3 Stochastic Forcing

In experiments, systems are rarely perturbed by the optimal perturbation. A more relevant perturbation might be small amplitude random noise. I will analyze how easily an ODE system can switch between stable equilibria under small amplitude random noise using the mean exit time from the attractor of the fixed point at $\boldsymbol{x} = \boldsymbol{0}$. It turns out that the mean exit time is related to the energy of the minimum seed (allowing for multiple perturbations). This is because, in the limit of low amplitude noise, the randomly perturbed system leaves the attractor of the fixed point at $\boldsymbol{x} = \boldsymbol{0}$ only when the random perturbations almost coincide with the minimum seed. This is related to Large something-something theory [1].

I determine the mean exit time using two approaches. First, I integrate the stochastic differential equations (SDEs)

$$\partial_t x_i = f_i(\boldsymbol{x}) + \sqrt{2\sigma_i}\partial_t W_{ti}, \tag{28}$$

where W_{ti} are Weiner processes, and $\boldsymbol{\sigma}$ is a constant. To integrate this SDE system, I integrate the ODE system (ignoring the noise), applying a forcing $\sqrt{2\sigma_i}\Delta W_i$ at each time step, where ΔW_i are iid normally distributed random variables with expected value zero and variance Δt, the time step.

As an example, consider the SDE where $\boldsymbol{f}(\boldsymbol{x})$ given by (1 & 2), and with $\sigma_1 = 0.1$, $\sigma_2 = 0.4$. In Figure 2, I plot three trajectories generated from this SDE, starting at $\boldsymbol{x} = \boldsymbol{0}$ at $t = 0$ and integrated to $t = 20$. The green trajectory never leaves the basin of attraction of the fixed point at $\boldsymbol{x} = \boldsymbol{0}$. The red trajectory makes it to the basin of attraction of the second fixed point, $\boldsymbol{x}_t$. However, the most interesting trajectory is the blue trajectory, which makes it to the basin of attraction second fixed point at $\boldsymbol{x}_t$, but eventually is perturbed back into the basin of attraction of the fixed point at $\boldsymbol{0}$.

Second, I evolve the probability distribution function (pdf), $F(\boldsymbol{x}, t)$, for the state of the system as a function of time. Integrating $F(\boldsymbol{x}, t)$ in the neighborhood of $\boldsymbol{x}_0$ gives the probability that $\boldsymbol{x} = \boldsymbol{x}_0$ at the time t. $F(\boldsymbol{x}, t)$ satisfies a modified diffusion equation known as a Fokker-Planck equation,

$$\partial_t F = -\boldsymbol{\nabla} \cdot (F\boldsymbol{f}(\boldsymbol{x})) + \partial^2_{x_i}(\sigma_i F). \tag{29}$$

I integrate the Fokker-Planck equation forward in time, using a sharply peaked Gaussian centered at $\boldsymbol{x} = \boldsymbol{0}$ as the initial condition. I set $F(\boldsymbol{x}, t) = 0$ on the outer edge of the domain in $\boldsymbol{x}$-space – this is a good approximation provided that the domain is large enough (I test

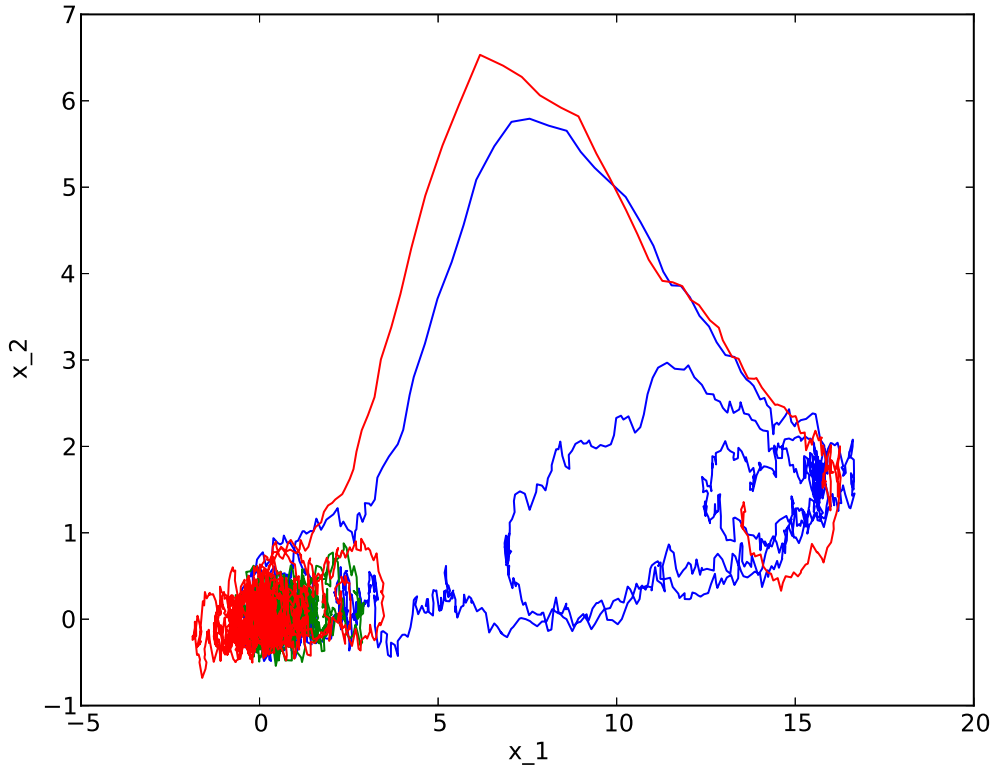

Figure 2: Three trajectories from the SDEs corresponding to (1 & 2), for $\sigma_1 = 0.1$ and $\sigma_2 = 0.4$, integrated to $t = 20$.

this by confirming that the solution is insensitive to the domain size). I found that explicit integration schemes have stability problems, so I integrate the Fokker-Planck equation using the forward Euler scheme.

For example, Figure 3 shows $F(\boldsymbol{x}, 10)$ for when $\boldsymbol{f}(\boldsymbol{x})$ is given by (1 & 2), and $\sigma_1 = 0.1$, $\sigma_2 = 0.4$. Note the similarities between the pdf and the trajectories in Figure 2.

The outer-most contour in Figure 3 shows the typical trajectory between the two fixed points. Going from the fixed point at $\boldsymbol{x} = \boldsymbol{0}$ to the fixed point at $\boldsymbol{x}_t$, the system typically crosses $x_2 = 1$ at around $x_1 \approx 2 - 3$, goes up to $\boldsymbol{x} \approx (6, 7)$, and then approaches $\boldsymbol{x}_t$, staying near $\boldsymbol{x} = (16, 2)$. This has larger x_1 and x_2 than $\boldsymbol{x}_t \approx (14, 4)$. I assume this is because the pull of the ODE back to the fixed point is stronger in the southwest direction than in the northeast direction. However, the system is often kicked out of the attractor of this fixed point. These features occur in both the trajectories and in the pdf.

As t increases, the pdf $F(\boldsymbol{x}, t)$ approaches a steady distribution. One could in principle find this distribution by setting the $\partial_t F$ term in the Fokker-Planck equation equal to zero, and then solving the corresponding elliptic equation. I have not done this, because one would presumably need to be careful about the boundary conditions. However, if one integrates the Fokker-Planck equation in time long enough, one can verify that the distribution function

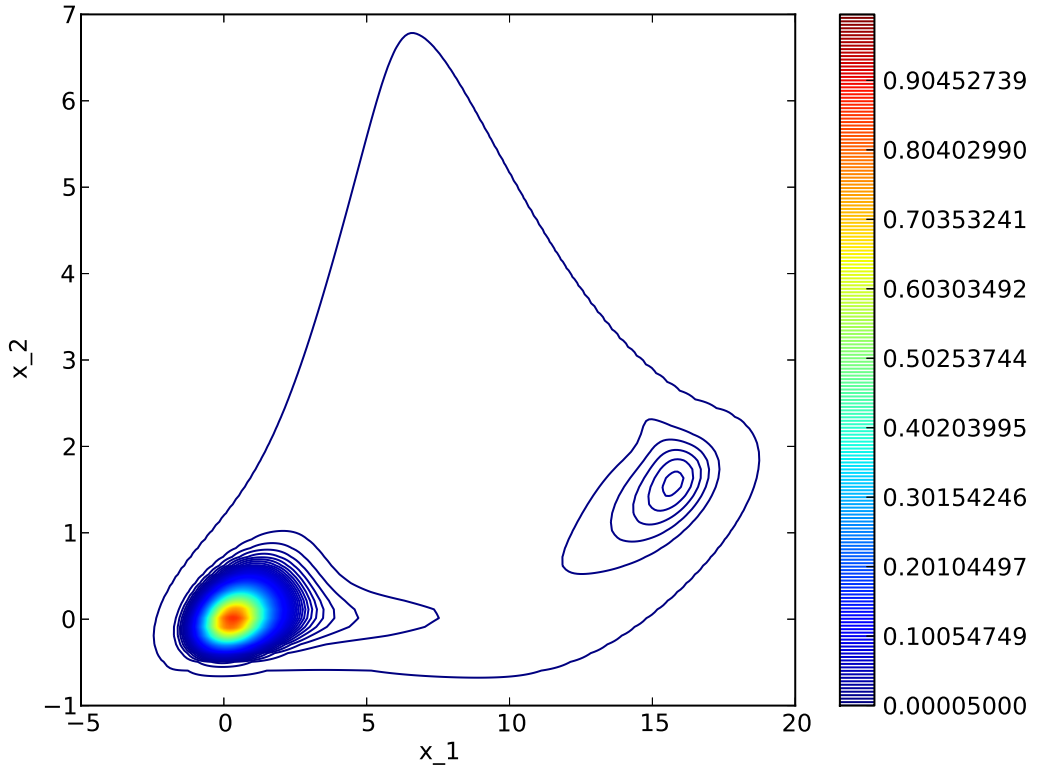

Figure 3: Contours of the probability density, as derived from the Fokker-Planck equation, for the ODEs (1 & 2), using $\sigma_1 = 0.1$, $\sigma_2 = 0.4$, and integrating to $t = 10$.

converges. For instance, in this example,

$$\frac{\int d^2\boldsymbol{x}\,|F(\boldsymbol{x},15) - F(\boldsymbol{x},10)|}{\int d^2\boldsymbol{x} F(\boldsymbol{x},15)} \approx 0.01, \tag{30}$$

and the analogous change between $t = 15$ and $t = 20$ is 0.002. Thus, it seems that $F(\boldsymbol{x}, t)$ has reached the a steady state distribution.

Another comparison is to consider an ensemble average of many integrations of the SDEs. I have integrated 20,000 instances of the SDEs to $t = 10$, and calculated a pdf of the states (Figure 4). In the limit of the SDEs being integrated infinitely many times, the pdf of the system states should converge to the result of the pdf given by the Fokker-Planck equation. Indeed, Figure 4 agrees well with Figure 3. Presumably, including even more integrations of the SDEs would improve the agreement.

3 Vlasov-Poisson

The Vlasov-Poisson equations describe a collision less electrostatic plasma. I will consider the case of an electron plasma, with a uniform and stationary background of ions. Then

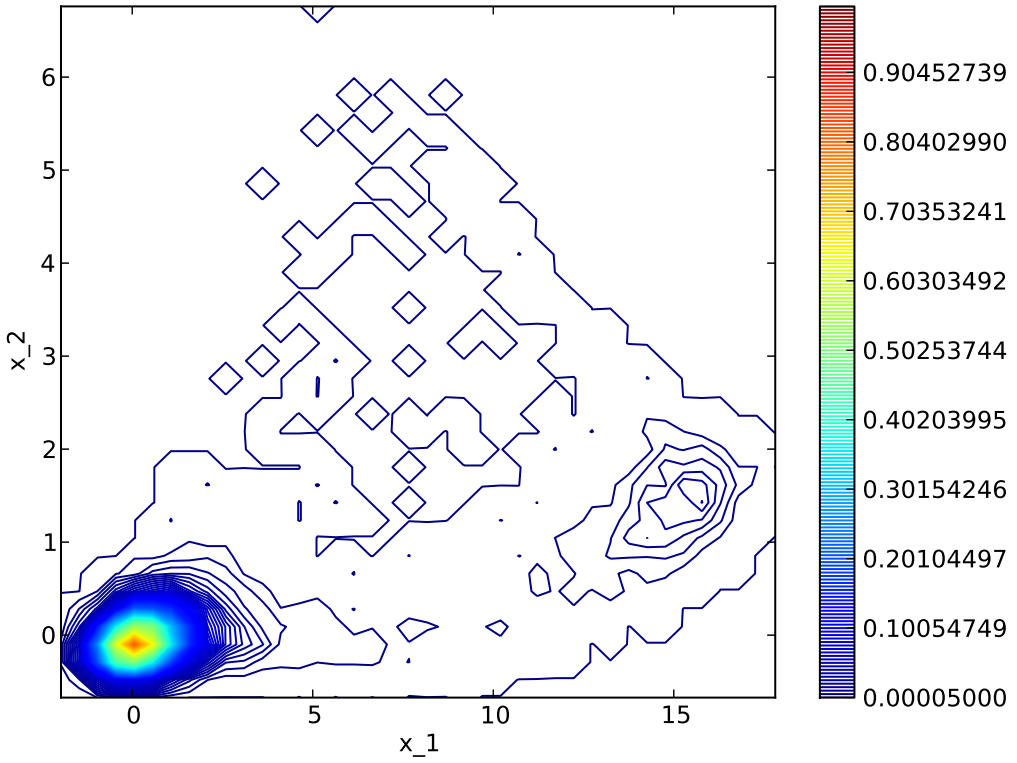

Figure 4: Contours of the probability density from 20,000 integrations of the SDEs corresponding to (1 & 2), using $\sigma_1 = 0.1$, $\sigma_2 = 0.4$, and integrating to $t = 10$.

the electron distribution function, $f(x, v, t)$, satisfies

$$\partial_t f(x, v, t) + v\partial_x f(x, v, t) - \frac{e}{m_e} E(x, t)\partial_v f(x, v, t) = 0, \tag{31}$$

$$\epsilon_0 \partial_x E(x, t) = e n_e - e \int_{-\infty}^{+\infty} dv \; f(x, v, t), \tag{32}$$

where e is the electron charge, m_e is the electron mass, n_e is proton number density (which equals the average electron number density), and ϵ_0 is the vacuum permittivity. The natural frequency scale is the electron plasma frequency,

$$\omega_{pe} = \sqrt{\frac{n_e e^2}{m_e \epsilon_0}}. \tag{33}$$

For simplicity, I assume that the plasma is periodic in the x direction with periodicity L. Non-dimensionalizing according to

$$\begin{aligned} x &\rightarrow xL, \\ t &\rightarrow t/\omega_{pe}, \\ v &\rightarrow vL\omega_{pe}, \\ f &\rightarrow fn_e/(L\omega_{pe}), \\ E &\rightarrow ELen_e/\epsilon_0, \end{aligned} \tag{34}$$

the Vlasov-Poisson equations become

$$\partial_t f(x,v,t) + v\partial_x f(x,v,t) - E(x,t)\partial_v f(x,v,t) = 0, \tag{35}$$

$$\partial_x E(x,t) = 1 - \int_{-\infty}^{+\infty} dv\ f(x,v,t). \tag{36}$$

An alternative form of (36) is

$$\partial_t \partial_x E(x,t) = \int_{-\infty}^{+\infty} dv\ v\partial_x f(x,v,t) \tag{37}$$

It has been observed through experiments and numerical simulation (e.g., [2] and references therein) that some perturbations to a spatially uniform electron plasma decay to zero in time, whereas others lead to non-linear states analogous to BGK waves, in which the electric field stays finite as $t \rightarrow \infty$. I am interested in finding the smallest perturbation to the distribution function which will generate an electric field which stays finite as $t \rightarrow \infty$. There are many ways to perturb the distribution function, but I will limit myself to perturbations which change the electron density as a function of x, but which do not change the velocity dependence of the distribution function.

The variational problem is as follows. The optimization is over the electric field at $t = 0$. I want to maximize

$$\int_0^L dx\ E(x,T)^2. \tag{38}$$

To limit the size of the perturbation, I fix

$$\int_0^L dx\ |\partial_x E(x,0)|^2 = N_0. \tag{39}$$

Recall that the velocity-averaged distribution function is given by $\partial_x E$. Furthermore, for $E(x,0)$ to be representable as the derivative of a periodic potential, it must have zero mean. Thus,

$$\int_0^L dx\ E(x,0) = 0. \tag{40}$$

The initial distribution function must be consistent with $E(x,0)$, so I impose

$$f(x,v,0) = F(v)(1 - \partial_x E(x,0)), \tag{41}$$

where $F(v)$ is the initial velocity-space structure of the distribution function, which is assumed to integrate to one. For instance, one could take,

$$F(v) = \frac{1}{\sqrt{2\pi}} \exp(-v^2/2). \tag{42}$$

The optimization is based on the initial electric field, rather than the full distribution function, because it is one dimensional. It would be possible to introduce additional degrees of freedom regarding the velocity distribution. For instance, one could take

$$f(x,v,0) = F(v) - \partial_x E(x,0) G(v), \tag{43}$$

where $G(v)$ integrates to one. If $G(v)$ is assumed to be a Gaussian with width σ centered at v_0, then it would be possible to optimize over σ and v_0 as well as $E(x,0)$.

3.1 Variational Problem

The functional I want to maximize is

$$\begin{aligned}
\mathcal{L} = \quad & \int_0^L dx\ E(x,T)^2 \\
& + \lambda \left(\int_0^L dx\ |\partial_x E(x,0)|^2 - N_0 \right) + \gamma \left(\int_0^L dx E(x,0) \right) \\
& + \iint dx dv\ \beta(x,v) \left[f(x,v,0) - F(v)(1 - \partial_x E(x,0)) \right] \\
& + \int dt \int_0^L dx\ \nu(x,t) \left(\partial_t \partial_x E(x,t) - \int_{-\infty}^{+\infty} dv\ v \partial_x f(x,v,t) \right) \\
& + \int dt \iint dx dv\ \mu(x,v,t) \left(\partial_t f(x,v,t) + v \partial_x f(x,v,t) \right. \\
& \qquad \left. - E(x,t) \partial_v f(x,v,t) \right).
\end{aligned} \tag{44}$$

Varying $\mathcal{L}$ yields

$$\begin{aligned}
\delta\mathcal{L} = \quad & \int_0^L dx\ 2\delta E(x,T)E(x,T) + \delta\gamma \int_0^L dx E(x,0) + \gamma \int_0^L dx \delta E(x,0) \\
& + \delta\lambda \left(\int_0^L dx\ |\partial_x E(x,0)|^2 - N_0 \right) + \lambda \int_0^L dx\ 2\partial_x \delta E(x,0) \partial_x E(x,0) \\
& + \iint dx dv\ \delta\beta(x,v) \left[f(x,v,0) - F(v)(1 - \partial_x E(x,0)) \right] \\
& + \iint dx dv\ \beta(x,v) \left[\delta f(x,v,0) + F(v) \partial_x \delta E(x,0) \right] \\
& + \int dt \int_0^L dx\ \delta\nu(x,t) \left(\partial_t \partial_x E(x,t) - \int_{-\infty}^{+\infty} dv\ v \partial_x f(x,v,t) \right) \\
& + \int dt \int_0^L dx\ \nu(x,t) \left(\partial_t \partial_x \delta E(x,t) - \int_{-\infty}^{+\infty} dv\ v \partial_x \delta f(x,v,t) \right) \\
& + \int dt \iint dx dv\ \delta\mu(x,v,t) \left(\partial_t f(x,v,t) + v \partial_x f(x,v,t) \right. \\
& \qquad \left. - E(x,t) \partial_v f(x,v,t) \right) \\
& + \int dt \iint dx dv\ \mu(x,v,t) \left(\partial_t \delta f(x,v,t) + v \partial_x \delta f(x,v,t) \right. \\
& \qquad \left. - \delta E(x,t) \partial_v f(x,v,t) - E(x,t) \partial_v \delta f(x,v,t) \right) . \qquad (45)
\end{aligned}$$

Thus, the partial derivatives are

$$\frac{\delta \mathcal{L}}{\delta \lambda} = \int_0^L dx\ |\partial_x E(x,0)|^2 - N_0, \tag{46}$$

$$\frac{\delta \mathcal{L}}{\delta \beta(x,v)} = f(x,v,0) - F(v)(1-\partial_x E(x,0)), \tag{47}$$

$$\frac{\delta \mathcal{L}}{\delta \gamma} = \int_0^L dx\ E(x,0), \tag{48}$$

$$\frac{\delta \mathcal{L}}{\delta \nu(x,t)} = \partial_t \partial_x E(x,t) - \int_{-\infty}^{+\infty} dv\ v\partial_x f(x,v,t), \tag{49}$$

$$\frac{\delta \mathcal{L}}{\delta \mu(x,v,t)} = \partial_t f(x,v,t) + v\partial_x f(x,v,t) - E(x,t)\partial_v f(x,v,t), \tag{50}$$

$$\frac{\delta \mathcal{L}}{\delta E(x,T)} = 2E(x,T) - \partial_x \nu(x,T), \tag{51}$$

$$\frac{\delta \mathcal{L}}{\delta E(x,0)} = -2\lambda \partial_x^2 E(x,0) + \partial_x \nu(x,0) + \gamma - \int dv\ \partial_x \beta(x,v) F(v), \tag{52}$$

$$\frac{\delta \mathcal{L}}{\delta f(x,v,T)} = \mu(x,v,T), \tag{53}$$

$$\frac{\delta \mathcal{L}}{\delta f(x,v,0)} = -\mu(x,v,0) + \beta(x,v), \tag{54}$$

$$\frac{\delta \mathcal{L}}{\delta f(x,v,t)} = v\partial_x \nu(x,t) - \partial_t \mu(x,v,t) - v\partial_x \mu(x,v,t) + E(x,t)\partial_v \mu(x,v,t), \tag{55}$$

$$\frac{\delta \mathcal{L}}{\delta E(x,t)} = \partial_t \partial_x \nu(x,t) - \int dv\ \mu(x,v,t)\partial_v f(x,v,t). \tag{56}$$

The nonlinear optimization algorithm is as follows: First, make a guess $E^{(0)}(x,0)$. Then integrate the Vlasov-Poisson equations to $t = T$. Then solve for the adjoint variables at $t = T$ using

$$\rho(x,T) = 2E(x,T), \tag{57}$$

$$\mu(x,v,T) = 0, \tag{58}$$

where I define $\rho(x,t) = \partial_x \nu(x,t)$. Integrate the adjoint equations back to $t = 0$ using

$$\partial_t \rho(x,t) = \int dv\ \mu(x,v,t)\partial_v f(x,v,t), \tag{59}$$

$$\partial_t \mu(x,v,t) + v\partial_x \mu(x,v,t) - E(x,t)\partial_v \mu(x,v,t) = v\rho(x,t). \tag{60}$$

Finally, update the guess to $E^{(1)}(x,0)$ by

$$E^{(1)}(x,0) = E^{(0)}(x,0) + \epsilon \frac{\delta \mathcal{L}}{\delta E(x,0)}, \tag{61}$$

where

$$\frac{\delta \mathcal{L}}{\delta E(x,0)} = -2\lambda \partial_x^2 E^{(0)}(x,0) + \rho(x,0) - \int dv\ \partial_x \beta(x,v) F(v) + \gamma. \tag{62}$$

At this stage λ, γ, and $\beta(x,v)$ are all still unknown. They must be chosen to satisfy

$$\int_0^L dx\ |\partial_x E(x,0)|^2 = N_0, \tag{63}$$

$$\int_0^L dx\ E(x,0) = 0, \tag{64}$$

$$\frac{\delta \mathcal{L}}{\delta f(x,v,0)} = \beta(x,v) - \mu(x,v,0) = F(v)\partial_x \left(\frac{\delta \mathcal{L}}{\delta E(x,0)} \right). \tag{65}$$

The last condition implies

$$\beta(x,v) = F(v)\partial_x \left(\frac{\delta \mathcal{L}}{\delta E(x,0)} \right) + \mu(x,v,0), \tag{66}$$

and thus

$$\begin{aligned} \frac{\delta \mathcal{L}}{\delta E(x,0)} = \quad & -2\lambda \partial_x^2 E^{(0)}(x,0) + \rho(x) + \gamma - \int dv\ \partial_x \mu(x,v,0) F(v) \\ & -\partial_x^2 \left(\frac{\delta \mathcal{L}}{\delta E(x,0)} \right) \int dv\ F(v)^2. \end{aligned} \tag{67}$$

Putting the $\delta \mathcal{L}/\delta E$ terms on the LHS yields

$$\left(1 + \int dv\ F(v)^2 \partial_x^2 \right) \frac{\delta \mathcal{L}}{\delta E(x,0)} = -2\lambda \partial_x^2 E^{(0)}(x,0) + \hat{\rho}(x), \tag{68}$$

where

$$\hat{\rho}(x) = \rho(x) - \int dv\ \partial_x \mu(x,v,0) F(v) + \gamma. \tag{69}$$

To solve for λ, first calculate $\hat{\rho}(x)$, picking γ such that $\hat{\rho}(x)$ has zero mean. Then Fourier transform the equations, and invert the derivative operator acting on $\delta \mathcal{L}/\delta E$,

$$\frac{\delta \mathcal{L}}{\delta E(k)} = \frac{2\lambda k^2 E^{(0)}(k) + \hat{\rho}(k)}{1 - k^2 \int dv\ F(v)^2}, \tag{70}$$

where k denotes the wavenumber. Then (61) becomes

$$E^{(1)}(k) = 2\lambda\epsilon \frac{k^2 E^{(0)}(k)}{1 - k^2 \int dv\ F(v)^2} + E^{(0)}(k) + \frac{\epsilon\hat{\rho}(k)}{1 - k^2 \int dv\ F(v)^2}. \tag{71}$$

Multiplying this equation by its complex conjugate times k^2 and then integrating over k yields

$$\begin{aligned} N_0 &= 4\lambda^2\epsilon^2 \int dk\ \frac{k^6 |E^{(0)}(k)|^2}{(1 - k^2 \int dv\ F(v)^2)^2} \\ &\quad + 4\lambda\epsilon \int dk\ \frac{k^4}{1 - k^2 \int dv\ F(v)^2} \times \\ &\qquad \Re\left[E^{(0)}(k) \left(E^{(0)}(k)^* + \frac{\epsilon\hat{\rho}(k)^*}{1 - k^2 \int dv\ F(v)^2} \right) \right] \\ &\quad + \int dk\ k^2 \left| E^{(0)}(k) + \frac{\epsilon\hat{\rho}(k)}{1 - k^2 \int dv\ F(v)^2} \right|^2. \end{aligned} \tag{72}$$

Solving this quadratic equation for λ specifies $\delta\mathcal{L}/\delta E$ (70), allowing $E^{(0)}$ to be updated to $E^{(1)}$. For sufficiently small ϵ, this update will increase $\int dx E(x,T)^2$. To find a NLOP, $E(x,0)$ is updated in this way until a local maximum is found.

3.2 Numerical Method

Stefan and Neil have provided a code that solves the Vlasov-Poisson system (35-36). The code uses operator splitting, separately solving

$$\partial_t f(x,v,t) + v\partial_x f(x,v,t) = 0, \tag{73}$$

and

$$\partial_t f(x,v,t) - E(x,t)\partial_v f(x,v,t) = 0. \tag{74}$$

These can be solved by shifting the x and v coordinates of $f(x,v,t)$ appropriately.

The code input is $f(x,v,t_n)$ in Fourier space (in x). To evolve to $f(x,v,t_{n+1})$, where $t_{n+1} = t_n + \Delta t$, the code does the following:

1. Shift x by an amount $-v\Delta t/2$:

$$f(x,v,t_{n+1/2}) = f(x - v\Delta t/2, v, t_n). \tag{75}$$

2. Solve the Poisson equation (36) for $E(x,t_{n+1/2})$ using $f(x,v,t_{n+1/2})$.

3. Transform to real space (in x).

4. Shift v by an amount $E(x,t_{n+1/2})\Delta t$:

$$f^*(x,v,t_{n+1/2}) = f(x, v + E(x,t_{n+1/2})\Delta t, t_{n+1/2}). \tag{76}$$

5. Transform to Fourier space (in x).

6. Shift x by an amount $-v\Delta t/2$:

$$f(x, v, t_{n+1}) = f^*(x - v\Delta t/2, v, t_{n+1/2}). \tag{77}$$

If this procedure is repeated many times, steps (6) and (1) can be done simultaneously (i.e., shift x by $-v\Delta t$).

The adjoint equations (59 & 60), have rather different structure. In particular, there is a source term for μ (the variable adjoint to f). This is important, given that μ is initialized to zero. Furthermore, ρ (adjoint to E) now satisfies a hyperbolic equation, rather than an elliptic equation.

My method for solving the adjoint equations is essentially a generalization of the operator splitting algorithm above. The adjoint equations evolve backward in time, so start with $\rho(x, t_{n+1/2})$ (in real space) and $\mu(x, v, t_n)$ (in Fourier space). In the following, $\Delta t > 0$.

1. Shift x by an amount $v\Delta t/2$ in μ:

$$\mu^*(x, v, t_{n-1/2}) = \mu(x + v\Delta t/2, v, t_n). \tag{78}$$

2. Transform $\mu^*(x, v, t_{n-1/2})$ to real space (in x).

3. Shift v by an amount $-E(x, t_{n-1/2})\Delta t/2$ in μ^*:

$$\mu^{**}(x, v, t_{n-1/2}) = \mu^*(x, v - E(x, t_{n-1/2})\Delta t/2, t_{n+1/2}). \tag{79}$$

4. Update ρ using the implicit step

$$\begin{aligned}\rho(x, t_{n-1/2}) \left(1 + \frac{\Delta t^2}{2} \int dv f(x, v, t_{n-1/2})\right) = \rho(x, t_{n+1/2}) \\ -\frac{\Delta t}{2} \int dv\ \mu(x, v, t_{n+1/2}) \partial_v f(x, v, t_{n+1/2}) \\ -\frac{\Delta t}{2} \int dv\ \mu^{**}(x, v, t_{n-1/2}) \partial_v f(x, v, t_{n-1/2}).\end{aligned} \tag{80}$$

5. Update μ^{**} using $\rho(x, t_{n-1/2})$:

$$\mu(x, v, t_{n-1/2}) = \mu^{**}(x, v, t_{n-1/2}) + v\Delta t \rho_{n-1/2}. \tag{81}$$

6. Shift v by an amount $-E(x, t_{n-1/2})\Delta t/2$ in μ

$$\tilde{\mu}(x, v, t_{n-1/2}) = \mu(x, v - E(x, t_{n-1/2})\Delta t/2, t_{n-1/2}). \tag{82}$$

7. Transform $\tilde{\mu}(x, v, t_{n-1/2})$ to Fourier space (in x).

8. Shift x by an amount $v\Delta t/2$ in $\tilde{\mu}$:

$$\mu(x, v, t_{n-1}) = \tilde{\mu}(x + v\Delta t/2, v, t_{n-1/2}). \tag{83}$$

When repeating the algorithm, steps steps (8) & (1) are combined.

3.3 Numerical Results

The goal for the optimization problem is to start with an initial distribution function for which the electric field decays to zero, and then find a different distribution function with an initial electric field of the same size for which the electric field does not decay to zero. I will start with initial distribution functions similar to those studied in [2, 3]. They considered

$$f(x, v, 0) = \frac{0.1}{\sqrt{2\pi}} \exp(-(v/0.1)^2/2) \times (1 + \epsilon \cos(8\pi x)), \tag{84}$$

i.e., a Gaussian with width 0.1, with a sinusoidal perturbation with a wavelength one fourth the box size, with strength ϵ.

It is nontrivial to determine the critical ϵ^* such that if $\epsilon > \epsilon^*$ the electric field will stay finite as $t \to \infty$, but if $\epsilon < \epsilon^*$, the electric field will decay to zero as $t \to \infty$. This is because numerically, the electric field can never decay to zero, only to the limits of the accuracy of the calculation (e.g., double point precision). For this problem, the typical evolution of the electric field is as follows (see Figure 5). The electric field strength is always oscillatory in time, but I will discuss the behavior of the envelop of this oscillation. First, linear Landau damping causes the electric field to decay. If $\epsilon > \epsilon^*$, then at a time $t_{\min}$, the electric field reaches a minimum value $E_{\min}$, and begins to grow. This sort of instability then saturates, and the electric field has a local maximum at $t_{\max}$, with field strength $E_{\max}$. After this point, the electric field oscillates near $E_{\max}$.

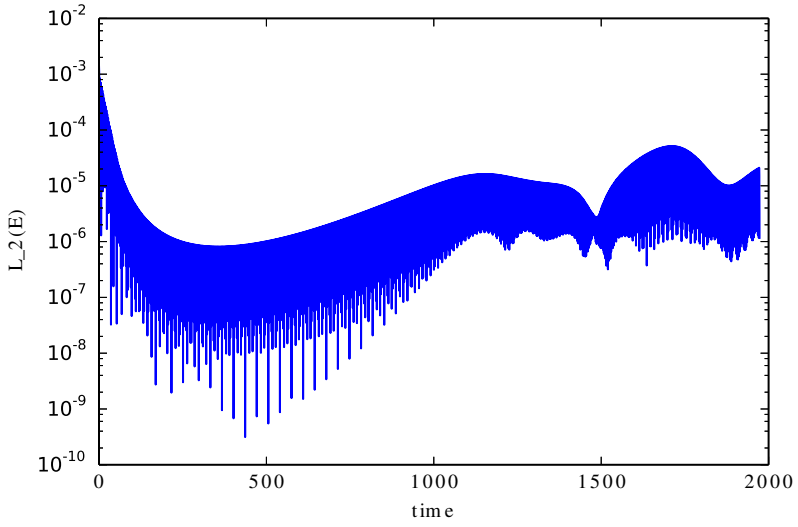

Figure 5: The L_2 norm of the electric field as a function of time, using the initial distribution function (84), for $\epsilon = 0.011$.

[3] found that as ϵ approaches ϵ^* from above, $t_{\min}$ and $t_{\max}$ go to infinity, and $E_{\min}$ and $E_{\max}$ go to zero as power-laws in $(\epsilon - \epsilon^*)$. This can be seen in the electric field traces in Figure 6. The five highest curves (blue, green, red, cyan, and purple) all have $\epsilon > \epsilon^*$, and have electric field minima which occur later and at lower electric field strengths as ϵ

decreases. The spikes occurring at e.g., $t \approx 1100$ and 1400 are numerical artifacts – they diminish in strength or disappear when the v resolution increases. Note that the electric field for $\epsilon < \epsilon^*$ does level off. This corresponds to reaching the numerical precision of the simulation.

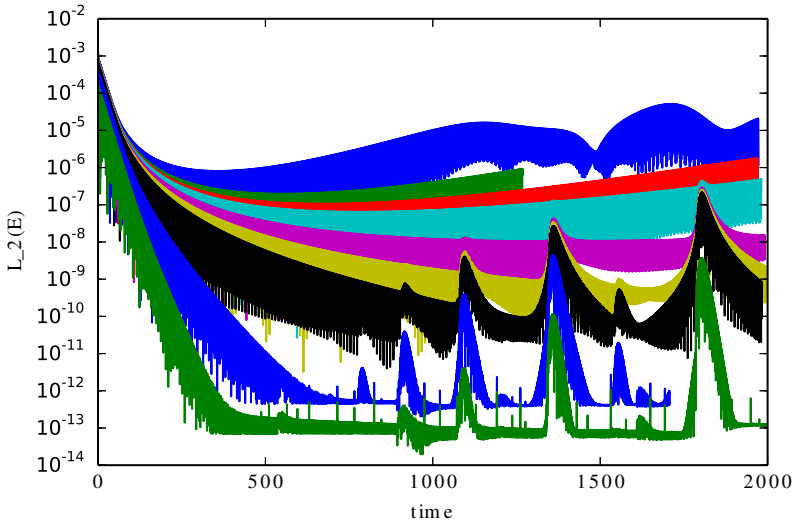

Figure 6: The L_2 norm of the electric field as a function of time, using the initial distribution function (84), for (from top to bottom) ϵ equal to 0.011, 0.01, 0.0097, 0.0095, 0.009, 0.0085, 0.008, 0.005, 0.002.

In [3], ϵ^* is determined by finding the best fit of the different quantities (e.g., $t_{\min}$) to the curve $A(\epsilon - \epsilon^*)^\alpha$. They found that the best fits for all four quantities $t_{\min}$, $t_{\max}$, $E_{\min}$, and $E_{\max}$ had ϵ^* very close to 0.0084, although they each had different power law indices α. I have reproduced their calculations, and also find that assuming $\epsilon^* = 0.0084$, the $t_{\min}$ and $E_{\min}$ are power-laws in $(\epsilon - \epsilon^*)$. This is compelling evidence that $\epsilon^* = 0.0084$ for this problem.

Now that the threshold amplitude has been established, I can run the optimization procedure starting from an initial electric field corresponding to $\epsilon = 0.008$, which is less than the threshold amplitude. I did this by maximizing the L_2 norm of the electric field at $T = 50$, in the way described in section 3.1. I also limited the initial electric field to only consist of the first ten Fourier components. After many iterations of the algorithm, I find that the electric field in Figure 7 has a large electric field at $T = 50$. Note that the electric field remains dominated by the fourth Fourier mode, but also has a non-negligible component from the first Fourier mode.

In Figure 8, I plot the electric field strength versus time for the initial distribution function given in (84) for $\varepsilon = 0.008$, as well as for the initial distribution function corresponding to the electric field plotted in Figure 7. The electric field is much larger at late times for the electric field derived by the optimization procedure, indicating that the algorithm has worked. Furthermore, rather than decaying to zero, the electric field seems to have saturated at a relatively high amplitude. I also believe this to be a numerically converged result

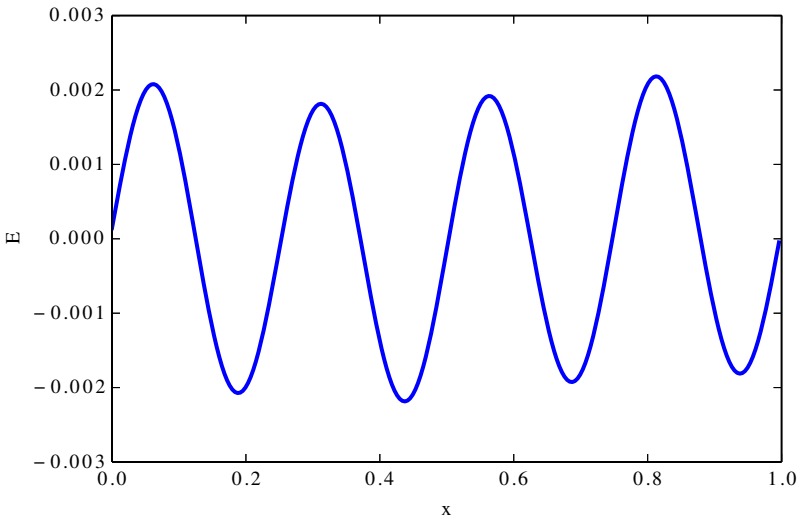

Figure 7: The initial electric field after many iteration of the optimization procedure for the L_2 norm of the electric field at $T = 50$.

– increasing the x, v, t resolution, as well as increasing $v_{\max}$, all have no effect on the electric field. Thus, the algorithm has successfully started from an electric field which decays to zero, and then find a different electric field of the same size which appears to remain finite for all time.

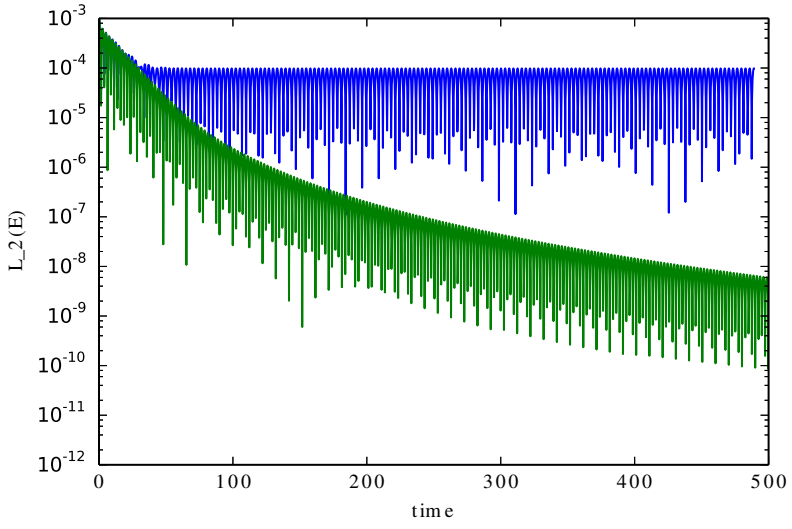

Figure 8: The L_2 norm of the electric field for the initial distribution function given in (84) for $\varepsilon = 0.008$ (green), and for the initial distribution function corresponding to the electric field plotted in Figure 7 (blue).

4 Further Work

In this report, I have described the application of nonlinear perturbation theory to two problems: low dimensional systems of ODEs, and the 1D Vlasov-Poisson equations. In both cases, I derived an algorithm for updating perturbations to increase a desirable quantity at late times. And in both cases, the numerical implementation of the algorithms seems to have been successful.

However, for the Vlasov-Poisson problem, there still remain many open questions. It is unclear what property of the initial electric field in Figure 7 allows the system to maintain a strong electric field. The distribution function at late times does not have any characteristics (e.g., a cat's eye pattern) that would suggest a mechanism for maintaining the electric field. Also, it is surprising that even when $\epsilon > \epsilon^*$ that the electric field can decay by many orders of magnitude before beginning to grow again and produce a cat's eye pattern. Both of these issues are likely amiable to analytic investigation.

These Vlasov-Poisson calculations were possible because I was able to parallelize the algorithm described above using MPI to efficiently run on about 32 processors. However, the algorithm uses global x and v interpolation techniques, which requires all-to-all communication calls. If a sufficiently accurate local interpolation scheme was implemented in the v direction, then the parallelization would be significantly more efficient.

5 Acknowledgements

First, I would like to thank my advisors, Rich Kerswell, Neil Balmforth, and Stefan Llewellyn Smith for proposing and shaping the project, and providing assistance and direction throughout the summer and beyond. Claudia Cenedese, Eric Chassignet, and Stefan Llewellyn Smith also deserve thanks for organizing a great summer. Most importantly, I would like to thank the other eight fellows for all the dinners at Quick's Hole (I got it; somehow?). I was partially supported over the summer by a Hertz Foundation Fellowship, and a National Science Foundation Graduate Research Fellowship under grant No. DGE 1106400.

References

[1] F. Bouchet, J. Laurie, and O. Zaboronski, *Control and instanton trajectories for random transitions in turbulent flows*, Journal of Physics Conference Series, 318 (2011), p. 022041.

[2] M. Brunetti, F. Califano, and F. Pegoraro, *Asymptotic evolution of nonlinear Landau damping*, Physical Review E, 62 (2000), pp. 4109–4114.

[3] A. V. Ivanov, I. H. Cairns, and P. A. Robinson, *Wave damping as a critical phenomenon*, Physics of Plasmas, 11 (2004), pp. 4649–4661.

[4] C. C. T. Pringle and R. R. Kerswell, *Using Nonlinear Transient Growth to Construct the Minimal Seed for Shear Flow Turbulence*, Physical Review Letters, 105 (2010), p. 154502.