6-2016

An Averaging Method for Advection-Diffusion Equations

Nicholas Spizzirri

Graduate Center, City University of New York

Recommended Citation
https://academicworks.cuny.edu/gc_etds/1369

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
An Averaging Method for Advection-Diffusion Equations

by

Nicholas Spizzirri

A dissertation submitted to the Graduate Faculty in Mathematics in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2016
The manuscript has been read and accepted for the
Graduate Faculty in Mathematics in satisfaction of the
Dissertation requirements for the degree of Doctor of Philosophy

April 19, 2016
Date

(Print) Tobias Schaefer
(Signature) Chair of Examining Committee

April 19, 2016
Date

(Print) Ara Basmajian
(Signature) Executive Officer

(Print) Jesenko Vukadinovic
(Print) Jozef Dodziuk
(Print) Olympia Hadjiladis

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract

AN AVERAGING METHOD FOR ADVECTION DIFFUSION EQUATIONS

by

NICHOLAS SPIZZIRRI

Adviser: Tobias Schaefer

Many models for physical systems have dynamics that happen over various different time scales. For example, contrast the everyday waves in the ocean with the larger, slowly moving global currents. The method of multiple scales is an approach for approximating the solutions of differential equations by separating out the dynamics at slower and faster time scales. In this work, we apply the method of multiple scales to generic advection-diffusion equations (both linear and non-linear, and in arbitrary spatial dimensions) and develop a method for averaging over the faster time scales, giving us an ‘effective’ solution governing the dynamics on the slower time scales. Numerical results are then obtained to confirm the effectiveness of this technique.
## CONTENTS

1 CHAPTER 1 1  
1.1 Introduction 1  
1.2 Diffusion 5  
1.3 Advection 8  
1.4 The Advection-Diffusion Equation 11  
1.5 The Method of Characteristics 13  
1.6 Inhomogeneous Problems 19  

2 CHAPTER 2 21  
2.1 Asymptotic Expansions 21  
2.2 The Method of Multiple Scales 24  
2.3 Multi-Scale Analysis on a PDE 32  
2.4 A not-quite-as-trivial example 37  
2.5 Discussion 43  

3 CHAPTER 3 45  
3.1 One More Example of the Multi-Scale Method on Advection-Diffusion Equations 45  
3.2 General Formula 50  
3.3 Non-Periodic Advection 57  
3.4 One More Example 61  

4 CHAPTER 4 64  
4.1 The Advection-Diffusion Equation in Arbitrary Dimensions 64
List of Figures

Figure 1.1 Intuition behind the heat equation 6
Figure 1.2 Diffusion of an initial Gaussian at various time slices 8
Figure 1.3 Intuition behind the Transport equation 9
Figure 1.4 Advection of an initial Gaussian at various time slices 10
Figure 1.5 A visual example of a path for which $u$ is constant 13
Figure 1.6 Two ways characteristics could cross 14
Figure 2.1 Approximate and exact solutions for (2.1.1) for $\epsilon = 0.1$ 23
Figure 2.2 Approximate and exact solutions for (2.1.1) for $\epsilon = 0.1$ 29
Figure 6.1 Initial condition 82
Figure 6.2 Simulation of (6.1.1) at $t = 2\pi$. 82
Figure 6.3 Simulation of (6.1.1) at $t = 2\pi$ 83
Figure 6.4 Simulation of (6.1.1) at $t = 3\pi$ and $t = 4\pi$. 83
Figure 6.5 Decay of $L^2$ norms over two periods. 85
Figure 6.6 Characteristic curves to equation (6.1.5) 86
Figure 6.7 Simulation of (6.1.5) at $t = \pi/2$. 86
Figure 6.8 Simulation of (6.1.5) at $t = \pi/2$, $\pi$, $3\pi/2$, $2\pi$. 87
Figure 6.9 Characteristics for $a(x,t) = 0.3 \cos(2x + t)$ 88
Figure 6.10 Above: Profiles at $t = 6\pi$, Below: Decay of $L^2$ norms 89
Figure 6.11 Simulation of (6.2.2) at time $t = 0$ and $\pi/2$. 92
Figure 6.12 Simulation of (6.2.2) at time $t = \pi$, $t = 3\pi/2$, and $t = 2\pi$. 93
Figure 6.13 $L^2$ decay of burgers equation 94
Figure 6.14 Simulation of (6.2.3) at times $t = 0$, $\pi/2$, $\pi$, $3\pi/2$, $2\pi$. 95
Figure 6.15 Close up view of solution to (6.2.3) at $t = 2\pi$. 96
Figure 6.16 Initial Condition 101
Figure 6.17 Above: $L^2$ decay, Below: $L^2$ decay zoomed in 101
Figure 6.18 Left: $u_{\text{true}} - u_{\text{diff}}$, Right: $u_{\text{ave}} - u_{\text{diff}}$, at $t = 10\pi$. 102

Figure 6.19 Initial Condition 103

Figure 6.20 4 time/space slices for dissipative ABC flow (true solution) 104

Figure 6.21 4 time/space slices for dissipative ABC flow (ave. solution) 104

Figure 6.22 Left: $u_{\text{true}}$, Right: $u_{\text{ave}}$, at $t = 10\pi$. 105

Figure 6.23 Left: $u_{\text{true}}$, Right: $u_{\text{ave}}$, at $t = 10\pi$, zoomed in. 105

Figure 6.24 Left: $u_{\text{true}} - u_{\text{diff}}$, Right: $u_{\text{ave}} - u_{\text{diff}}$, at $t = 10\pi$. 105

Figure 6.25 Decay of $L^2$ norm for ABC flow. 106
Overview and Introduction to the Advection-Diffusion Equation.

1.1 INTRODUCTION

The advection-diffusion equation,

$$u_t + a \cdot \nabla u = \epsilon \Delta u,$$

is a partial differential equation that is ubiquitous in pure and applied mathematics as a model for many physical systems. Though its applications are wide and varied, it is perhaps easiest to understand in the context of fluids. The word ‘advection’ often refers to the transport of material through a fluid by the motion of the fluid itself, while the word ‘diffusion’ refers to the fact that material tends to spread out and diffuse as time goes by. Thus, solutions of the advection-diffusion equation are functions that simultaneously display two very different types of phenomena. How the advection and diffusion interact with and affect each other is a deep and rich question, and it is one that we explore in this work.

It is often the case that the advection and diffusion phenomena transpire on very different time and length scales. An oil spill in the ocean, for example, might slowly diffuse over the course of weeks or months, whereas steady advective currents might transport the oil around rather quickly. When such disparities between time and/or length scales are present, we may consider employing the method of multiple scales. This is a
technique for approximating solutions to PDEs by introducing new, auxiliary parameters that help articulate what is happening on the different time/length scales. Sometimes it is even possible to effectively ‘average out’ the dynamics on smaller/faster scales to determine what is the ‘net’ effect on the larger/slower scales. In this work, we derive and implement such a technique.

Typically, the diffusion process happens on a much slower time scale than the advection process (as suggested by the oil spill example). This situation corresponds to the diffusive parameter $\epsilon$ being significantly smaller than the other relevant parameters in the problem. Such a scenario has been called the ‘Batchelor Regime’ due to the work of G. K. Batchelor in the 1950’s who, among others, was one of the first to highlight the relevancy of this setting. Because it is typically diffusion which governs the long term limiting behavior of a dissipative system, then it is often of primary interest to describe how the advection alters the rate of diffusion, in particular.

The primary inspiration for the work presented here is the 1991 paper by M. S. Krol. In this paper, Krol approximates the solution to the advection-diffusion equation on an unbounded domain by (approximately) decomposing it in terms of the faster and slower time scales. Assuming that the advection is nearly periodic in time, this then sets the stage for an averaging process by which the faster scale dependence of the solution can be averaged out, leaving us with an effective equation governing the evolution of the solution over the slower time scales. Moreover, he shows that the error of the effective solution is of the order $O(\epsilon)$ on the time scale of $[0, k/\epsilon)$, for some constant $k$.

Shortly thereafter, Heijnekamp et al. considered the 2-D advection diffusion equation on a bounded domain, again with time-periodic advection, and subject to Dirichlet boundary conditions. Obtaining similar results as Krol, they in fact show that the $O(\epsilon)$ convergence is valid for all time in this case, not just on the $1/\epsilon$ scale, even providing some explicit examples on the disk. We will derive similar techniques here to
that of Krol’s, but by a different approach, while expanding and generalizing the scope and applicability.

The general program of finding an effective equation governing the slower time scales has been taken up by many. Krol’s approach was one of the earliest, but it is by no means the only avenue of pursuit. Homogenization techniques are another popular approach. The idea here is to find an effective diffusion equation (specifically, an effective diffusion operator)

$$u_t = \Delta_{\text{eff}} u,$$

such that $u \to \overline{u}$ in the limit as $t \to \infty$. This is often found by rescaling the variables, leading to an auxiliary PDE, from which $\Delta_{\text{eff}}$ can be computed via derivatives of the solution to the auxiliary problem $[20]$. In the work of Freidlin and Wentzell $[11]$, they too define an effective equation which is valid only as $t \to \infty$, although their approach is entirely different, arriving at their equation by randomly perturbing Hamiltonian systems. In a slightly more abstract setting, Buitelaar $[4]$ looks at initial value problems posed on a Banach space with operators possessing a discrete spectrum. Transforming the evolution equation into an integral equation of the form $z(t) = u + \int_t^t F(z(s), t, ) ds$, he computes a ‘long-time’ average for the integrand $F$ by $F^0(z) = \lim_{T \to \infty} \int_0^T F(z, s)ds$. Replacing $F$ with $F^0$ gives us a corresponding ‘averaged’ solution $\overline{z}$. One unifying property to note from the previous three examples is that the effective solution is only relevant in the long-time limit, whereas the averaged solution proposed by Krol is accurate on the $[0, k/\epsilon)$ scale (accurate in a sense made precise in section 2.5 and 4.2). For a good survey of these and other averaging techniques, see $[29]$.

Another drawback of the abstract methods of Buitelaar and others, is the fact that it is often difficult to compute explicit examples and/or techniques. Rather than find an effective equation, many have instead sought to analyze the long term behavior via the spectrum of the advection-diffusion operator. This approach typically takes place on a bounded domain, where the spectrum is usually easier to analyze. Of particular
interest is the case where the advection tends to ‘enhance’ the rate of dissipation. Finding sharp conditions for when exactly we should expect dissipation enhancement has been studied, for example by Berestycki et al. \(3\) in the case of Dirichlet boundary conditions, and Constantin et al. \(6\) in the case of Neumann boundary conditions. All of this work however, takes place on a bounded domain, and there is little by way of analogous results for unbounded domains.

In the last two decades, the idea of dissipation enhancement due to chaotic flow has grown in popularity, aided in no small part by the increase in computing power. As Popovich describes it, complex chaotic advection, with its “stretching and folding, [...] produces intricate patterns down to ever finer detail, until finally the diffusion takes over and completes the mixing...” \(24\). The existence of long term, nearly stable states exhibiting complex, self-similar behavior (dubbed ‘strange eigenmodes’ by Pierre-rehumbert\(22\)) has attracted attention in particular. Once more, the decay rate of these eigenmodes can be studied via the spectrum of the differential operator. In this context, Schaefer et al. \(23\) consider a linear advection-diffusion operator on the 2-D annulus. Inspired by Krol, they perform a change of variables to action-angle coordinates which facilitates averaging. The corresponding spectral analysis is done in a follow up paper \(30\), and their numerical results will be used for comparison in chapter 6 of this work.

It is the interest of this author to investigate the problem on an unbounded domain, and moreover, not to prove convergence theorems or obtain abstract conditions, but rather, in the spirit of Krol, to develop concrete techniques and applicable algorithms for determining effective equations. Furthermore, whereas much of the previous work has been concentrated on bounded domains, incompressible flows, and linear operators, it is the goal in this work to provide a more general framework, using the method of multiple scales. We will seek an algorithm (both theoretically and numerically) to find the effective equation, which can be applied to arbitrary advection fields and in arbitrary dimensions,
even treating, in one case, the nonlinear Burgers equation (with the setting being strictly on unbounded domains).

Chapter 1 is an introduction to the advection-diffusion equation as well as a review of some basic techniques that will be needed throughout the rest of the work. Chapter 2 is an introduction to the method of multiple scales. By chapter 3, we will have all the prerequisite tools we need to start building some theory for the multiple scale analysis on linear advection-diffusion equations in one spatial dimension and we will see some examples of its application. In chapter 4, we extend the results from chapter 3 to higher spatial dimensions. In chapter 5 we apply this method to a non-linear Burgers style advection-diffusion equations and finally, in chapter 6, we put the theorems to the test by performing numerical simulations.

1.2 DIFFUSION

The Diffusion Equation (or heat equation) is the partial differential equation

\[
\frac{\partial u}{\partial t}(x,t) = \kappa \frac{\partial^2 u}{\partial x^2}(x,t)
\]

\[u(x,0) = f(x)\]

for some prescribed boundary conditions and \(\kappa > 0\). Since solutions of this equation will be functions of space and time which depend on our choice of initial condition \(f\), we can think of the PDE as a machine which takes initial conditions (as functions of space) and evolves them forward in time. It is from this perspective, then, that we call it the diffusion equation, for initial conditions tend to diffuse and smoothen out as they evolve according to the PDE. Why this is so can be summarized quite succinctly by the following picture:

This can be seen more rigorously by considering specific solutions. For example, if we take our spatial domain to be the interval \([0, L]\) with Dirichlet boundary conditions,
Figure 1.1   Intuition behind the heat equation

\[ u(0) = u(L) = 0, \text{ and initial condition } f(x), \text{ then it can easily be verified that the solution to the diffusion equation is } \]

\[ u(x, t) = \sum_n C_n e^{-\kappa \frac{n^2 \pi^2}{L^2} t} \sin \left( \frac{n\pi}{L} x \right), \tag{1.2.2} \]

where the coefficients \( C_n \) are the Fourier coefficients of the function \( f(x) \). Again, if we think of \( u(x, t) \) as a sequence of spatial profiles that evolve with time, then we can ask questions like “towards what function does \( u \) approach as time goes on?” With the boundary conditions pinned down at zero, and the diffusion equation seeking out profiles that are featureless and diffuse, we might expect \( u \) to tend towards being identically zero. Sure enough, it is clear from \ref{1.2.2} that \( u \to 0 \) as \( t \to \infty \). In fact, if we look at the \( L^2 \) norm (a measure analogous to ‘mass’) of the spatial profile as it evolves, we see

\[ ||u||_{L^2}(t) = \int_0^L \left( \sum_n C_n \exp \left( -\kappa \frac{n^2 \pi^2}{L^2} t \right) \sin \left( \frac{n\pi}{L} x \right) \right)^2 \, dx \]

\[ = \sum_n \frac{L}{2} C_n \exp \left( -2\kappa \frac{n^2 \pi^2}{L^2} t \right), \]
which, in the long run decays like

\[ \|u\|_{L^2}(t) \sim \exp \left( -\frac{2\pi^2}{L^2\kappa t} \right). \]  

(1.2.3)

On the other hand, we may consider a diffusion equation problem on an unbounded domain, like \( \mathbb{R} \), with the boundary conditions simply being that we demand the function decay to zero at infinity. In this case, one can verify that the solution to (1.2.1) with the same initial value will be given by the formula \[ \text{(8)} \]

\[ u(x,t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{\mathbb{R}} f(y) \exp \left( -\frac{(x-y)^2}{4\kappa t} \right) dy. \] 

(1.2.4)

Looking at the decay of the solution over time is not so trivial in this case, now that the solution is represented as an integral. If we take a simple example, however, such as a Gaussian for the initial condition, \( f(x) = \exp(-bx^2) \), then the solution according to (1.2.4), is

\[ u(x,t) = \frac{1}{\sqrt{1+4b\kappa t}} \exp \left( -\frac{bx^2}{1+4b\kappa t} \right). \] 

(1.2.5)

By squaring and integrating over \( \mathbb{R} \), one can compute the \( L^2 \) norm to be

\[ \|u\|_{L^2}(t) = \frac{\pi/2b}{1+4b\kappa t}, \]

which clearly goes to zeros as \( t \to \infty \) at the rate of

\[ \|u\|_{L^2}(t) \sim \frac{\pi}{8b^2\kappa t}. \] 

(1.2.6)

In Figure 2 we see some spatial profiles at various slices in time, illustrating the diffusive nature of its evolution.

An important thing to note in both of these examples is how the coefficient \( \kappa \) affects the decay rate. In both (1.2.3) and (1.2.6), we see that the decay rate is faster if \( \kappa \) is
Figure 1.2  Diffusion of an initial Gaussian at various time slices

increased. For this reason, \( \kappa \) is called the ‘diffusivity’ of the system, a measure of how quickly the system will diffuse. In fact, since \( \kappa \) always appears as a coefficient of \( t \), we might think of it as ‘rescaling’ time.

We also mention here for future reference, that for \( x \in \mathbb{R}^n \), the diffusion equation takes the form

\[
\frac{\partial u}{\partial t}(x,t) = \kappa \Delta u(x,t),
\]

where \( \Delta \) is the Laplacian operator

\[
\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \ldots
\]

1.3 \textsc{advection}

Next, we turn to the PDE

\[
\frac{\partial u}{\partial t}(x,t) + a \frac{\partial u}{\partial x}(x,t) = 0 \tag{1.3.1}
\]
\[ u(x,0) = f(x). \]

Let us assume again, that this equation is posed on the whole real line, with \( a \in \mathbb{R} \), and the boundary conditions again being decay at infinity. One can check that the solution to this equation is given by

\[ u(x,t) = f(x-at). \]  

We can surmise from this formula that the evolution of an initial condition will simply be given by the initial profile rigidly translating along the \( x \) axis at a speed \( a \) (to the right if \( a > 0 \) and to the left if \( a < 0 \)). For this reason, this equation is often called the transport equation. Though it is easy to verify that the expression above is indeed the solution (simply substitute it back into the PDE), it is also somewhat intuitive, considering the picture in Figure 3.

Figure 1.3  Intuition behind the Transport equation
It is reasonable to guess therefore that the solution does not decay at all. In fact, this is true, as we can check. First, by \((1.3.2)\), we have
\[
 u_t = -a u_x.
\]

Now we integrate both sides:
\[
 \int_R u_t = -a \int_R u_x \quad \Rightarrow \quad \frac{d}{dt} \int_R u = -au \bigg|_\infty^\infty = 0.
\]
Thus, \(\int_R u\) is constant.

We now consider a variant of the transport equation with non-constant coefficients of \(u_x\):
\[
 \frac{\partial u}{\partial t}(x,t) + a(x,t) \frac{\partial u}{\partial x}(x,t) = 0. \tag{1.3.3}
\]

If we recall from before that \(a\) was a measure of the speed at which the profile was transported, then we can interpret \((1.3.3)\) as a transport equation, where each point in space and time has its own speed of transport. Thus, for example, in the solution to the equation
\[
 u_t + \sin(t)u_x = 0,
\]
we will see an initial profile transported back and forth, harmonically in time. Whereas, in the solution to the equation
\[
 u_t + xu_x = 0,
\]
we will see an initial profile get ‘stretched apart’, because points farther from the origin are more inclined to transport away faster. This is illustrated by the solution \(u(x,t) = \exp(-5(xe^{-t} - 2)^2)\), in Figure 4:

In general, as is indicated in \((1.3.3)\), we can have both temporal and spatial dependence in \(a\). A PDE of this kind is often called an Advection Equation. If we begin to stretch the
Figure 1.4 Advection of an initial Gaussian at various time slices

initial profile by employing a non-uniform ‘advection field’, then we would no longer expect the $L^p$ norm to remain constant. However, we would not necessarily expect decay either, like we see in the diffusion equation.

I will also mention here for future reference, that for $x \in \mathbb{R}^n$, the advection equation takes the natural form of

$$\frac{\partial u}{\partial t}(x,t) + \vec{a}(x,t) \cdot \nabla u(x,t) = 0,$$

which, in the context of fluid dynamics, is related to the material derivative of $u$.

1.4 THE ADVECTION-DIFFUSION EQUATION

Now we consider the PDE which has both a diffusive term \textit{and} an advective term:

$$\frac{\partial u}{\partial t}(x,t) + a(x,t) \frac{\partial u}{\partial x}(x,t) = \kappa \frac{\partial^2 u}{\partial x^2}(x,t).$$

(1.4.1)
This is called, naturally, an *Advection-Diffusion Equation*. This equation, which, in general, is significantly harder to solve than the advection or diffusion equations separately. In fact, it is almost never possible to find a closed-form of the solution. Can we, however, make a qualitative guess about its behavior? Given the previous discussion, we would expect that the evolution of an initial condition according to this PDE would be that the initial profile is advected around according to \( a(x, t) \), while it simultaneously diffuses overall, according to \( \kappa \). This is ‘somewhat’ accurate. Consider, for example, the simple case of \( a(x, t) = a \), a constant, and let this be posed on the whole real line, with decay at infinity for the boundary conditions. Then, as one can check, the solution, for a given initial condition, will be

\[
u(x, t) = \frac{1}{\sqrt{4\pi \kappa t}} \int_{\mathbb{R}} f(y) \exp \left( \frac{-(x - at - y)^2}{4\kappa t} \right) . \tag{1.4.2}
\]

For the particular choice of \( f(x) = e^{-bx^2} \), we have

\[
u(x, t) = \frac{1}{\sqrt{1 + 4b \kappa t}} \exp \left( \frac{-b(x - at)^2}{1 + 4b \kappa t} \right) . \tag{1.4.3}
\]

It is clear in this particular case that we literally just have the solution to the diffusion equation \([1.2.5]\) being transported around at the speed \( a \). What is important to notice here, is that one can view this solution as depending on two different but relevant time parameters: the advection happens through the parameter \( t \), while the diffusion happens through the ‘parameter’ \( \kappa t \). In section 2.2 we will make this idea concrete.

In more general cases, however, the behavior of a solution to the advection-diffusion equation will not just be a superposition of the advection and diffusion alone. As we will see, the advection is often responsible for altering the rate of diffusion. In fact, it will be one of our primary goals to understand and approximate the net effect of the advection on the rate of diffusion.
To put this idea in physical terms, we can imagine putting a concentration of dye in a solution. If the solution is still and placid, then as time goes on, we expect the dye to slowly diffuse - due to thermal properties of the solution - governed by a pure diffusion equation. However, we can instead introduce some advection by stirring up the solution. Now we can ask how the effective rate at which the dye diffuses throughout the solution has been altered. The methods described in the following chapters will answer this question quantitatively.

1.5 The Method of Characteristics

Because it will be of significant importance later on, we quickly review the method of characteristics [8]. This is the method used (implicitly and sometimes explicitly) when solving purely advective equations. Consider the advection equation

\[ \frac{\partial u}{\partial t}(x,t) + a(x,t) \frac{\partial u}{\partial x}(x,t) = 0 \]  

\[ u(x,0) = f(x). \]  

Imagine a path in the \((x,t)\)-plane, parameterized by \(s\): \((x(s),t(s))\). In particular, we will seek out paths for which \(u\) is constant, \(u(x(s),t(s)) = C\). These can be thought of as the level sets of the surface \(u(x,t)\).

The form of equation (1.5.1) actually makes finding such paths very straightforward. Notice that if we impose the conditions

\[ \dot{s} = 1 \quad \dot{x}(s) = a(x(s),t(s)), \]  

(where \(\cdot = \frac{d}{ds}\), then equation (1.5.1) reduces to \(\dot{u}(x(s),t(s)) = 0\), in other words, \(u\) is constant along the path. Thus, paths satisfying (1.5.2) are our desired paths. Such paths
are often called *characteristic* curves, or simply ‘characteristics’. Let us, then, solve these ODEs. The first equation results in \( t = s + c \), and if we parametrize the path such that \( t = 0 \) when \( s = 0 \), then we have \( t = s \). The second ODE then reduces to \( \dot{x}(t) = a(x, t) \).

Denote the solution to this equation by \( \chi \), which will depend on one arbitrary integration constant:

\[
x = \chi_C(t)
\]

Clearly each choice of \( C \) corresponds to a different path (and therefore a potentially different constant value of \( u \)). We can parametrize the space of paths by where they intersect the \( x \) axis. Call this parameter \( x_0 \). Then, choose the integration constant for each path such that \( x_0 = \chi_C(0) \). Solving for \( C \) and plugging it back into \( \chi \), we will have \( x \) as a function of \( t \) and \( x_0 \):

\[
x = x(t, x_0).
\]

We will want to invert this map to obtain

\[
x_0 = x_0(x, t).
\]
But how do we know it is invertible? If it were *not* invertible, that would mean that there were two initial points \( x_{0_1} \) and \( x_{0_2} \) that produce characteristic paths which intersect at the point \((x, t)\). Either they cross and keep going, or they simply run into each other, as in Figure 5.

Figure 1.6  Two ways characteristics could cross

In the first case, the ODE \( \dot{x}(t) = a(x, t) \) would not be well-posed, for it gives two solutions starting from the point \((x, t)\). In the second case, we can see that at the point where the two paths meet, we have \( \dot{x}(t) = \infty \), which means that \( a(x, t) \) is singular there. More concretely, we can invoke existence and uniqueness theorems for first order ODEs (applied to \( \dot{x}(t) = a(x, t) \)) to guarantee that the characteristic paths emanating from a point exist and are unique \[8\].

Now that we are confident that we *can* invert the map \( x = x(t, x_0) \), let us consider the inverse map

\[
x_0 = x_0(x, t).
\]

I will refer to this map often as *the* characteristic map. The interpretation of this map is simple; if we choose an arbitrary point \((x, t)\), it will exist on some characteristic curve. Trace this curve back to where it crosses the \( x \)-axis. This value will be \( x_0 \). What makes this
so useful is that we know the value of $u$ at the point $(x_0, 0)$; it is given by $u(x_0, 0) = f(x_0)$. Since $u$ is constant along the curve, then it will have the same value at $(x, t)$:

$$u(x, t) = f(x_0(x, t)).$$

This is easiest to understand through examples. Consider the transport equation

$$u_t + au_x = 0.$$ 

In this case we would have

$$\dot{x}(t) = a,$$

which means $x = at + x_0$. Inverting this gives us $x_0 = x - at$. Thus, the solution will be given by

$$u(x, t) = f(x - at),$$

as we have seen before. A less trivial example is

$$u_t + \cos(t)xu_x = 0.$$ 

In this case we have

$$\dot{x}(t) = \cos(t)x.$$ 

Separating variables gives us

$$\frac{dx}{x} = \cos(t)dt,$$

which, after integrating, becomes

$$\log|x| = \sin(t) + C.$$
Plugging in $t = 0$ gives us
\[ \log |x| = \sin(t) + \log |x_0|. \]

Exponentiating both sides yields
\[ x = x_0e^{\sin(t)}. \]

Inverting gives us
\[ x_0(x, t) = xe^{-\sin(t)}, \tag{1.5.3} \]

and we finally arrive at a solution
\[ u(x, t) = f(xe^{-\sin(t)}). \]

We will use this particular example at the beginning of Chapter 3. Consider one more example, whose solution we will also use in Chapter 3:
\[ u_t + \cos(t)x^2u_x = 0. \]

(Aprically the same as the previous example, only $x$ has been promoted to $x^2$.) We must solve the equation
\[ \dot{x}(t) = \cos(t)x^2. \]

By separating variables and integrating, one arrives at
\[ x(t) = \frac{1}{\frac{1}{x_0} - \sin(t)}. \]

The inverse of this is
\[ x_0(t) = \frac{1}{\frac{1}{x} + \sin(t)}, \tag{1.5.4} \]
and so the solution will be given by

$$u(x,t) = f \left( \frac{1}{\frac{1}{x} + \sin(t)} \right).$$

A property of the characteristic map $x_o(x,t)$ which will be useful later on, is the fact that this map itself solves the associated advection equation. To see this, we simply choose the initial condition to be $f(x) = x$. Then the solution $u$ is just $u(x,t) = x_o(x,t)$, which means that the map $x_o(x,t)$ must satisfy the original advection equation. This is true even in higher dimensions, where the advection equation takes the form

$$\frac{\partial u}{\partial t}(\vec{x},t) + \vec{a}(\vec{x},t) \cdot \nabla u(\vec{x},t) = 0.$$

The characteristic curves are obtained by solving the system of equations

$$\dot{x}^i(t) = a^i(\vec{x},t).$$

Solving these ODEs yields $\vec{x} = \vec{x}(\vec{x}_o,t)$, which we invert to obtain $\vec{x}_o = \vec{x}_o(\vec{x},t)$. The solution will then be given by

$$u(\vec{x},t) = f(\vec{x}_o(\vec{x},t)).$$

But the fact remains, each component of $\vec{x}_o$, i.e. $x^i_o(\vec{x},t)$, will satisfy the advection equation, for we could always simply choose the initial condition to be $f(\vec{x}) = x^i$. In this case, it would have to be true that

$$\frac{\partial x^i_o}{\partial t} + \vec{a} \cdot \nabla x^i_o = 0. \quad (1.5.5)$$

This fact will be useful in section 4.1.
1.6 INHOMOGENEOUS PROBLEMS

An advection equation whose right hand side is non-zero,

\[ \frac{\partial u}{\partial t}(x,t) + a(x,t) \frac{\partial u}{\partial x}(x,t) = g(x,t) \tag{1.6.1} \]

\[ u(x,0) = f(x), \]

is said to be inhomogeneous (as opposed to homogeneous, for which \( g \equiv 0 \)). If, in the spirit of the characteristic method, we solve the ODE

\[ \dot{x}(t) = a(x,t), \quad x(0) = x_0, \]

to obtain \( x = x(x_0, t) \), then this time the PDE (1.6.1) tells us that, along the characteristic path, \( u \) is no longer constant, but in fact

\[ \dot{u}(t) = g(x(x_0, t), t). \]

On a particular characteristic, identified by a choice of \( x_0 \), we can integrate this equation to compute \( u \) at any time:

\[ u(x_0, t) = u(x_0, 0) + \int_0^t g(x(x_0, s), s) \, ds. \]

Now, write \( x_0 \) in terms of \((x, t)\):

\[ u(x, t) = u(x_0(x(t), 0)) + \int_0^t g(x_0(x(t), s), s) \, ds. \]
But the term $u(x_0(x,t),0)$ is just the value of the initial condition at the point $x_0(x,t)$. Thus, we have

$$u(x,t) = f(x_0(x,t)) + \int_0^t g(x(x_0(x,t),s),s)ds,$$

which completely solves the initial value problem. The first term is often called the \textit{homogeneous solution} while the second term is often called the \textit{particular solution}. This formula is a particular example of what’s known as \textit{Duhamel’s Principle} \cite{12}. 
CHAPTER 2

The Method of Multiple Scales

2.1 ASYMPTOTIC EXPANSIONS.

Many if not most problems in partial differential equations are unsolvable. While the solution may be guaranteed to exist, finding a closed form expression for the solution is often difficult or impossible. The typical response to this is either to obtain an approximate solution by numerical brute force, or to obtain qualitative information about the solutions, such as long term behavior and stability. An alternative approach is to obtain approximate solutions by asymptotic expansions. Often, this occurs if there is a small parameter in the differential equation (typically denoted by \( \epsilon \)) which, if omitted makes the problem significantly easier. Our problem is then viewed as a perturbation to the easier problem, and approximate solutions are sought after as perturbations to the solutions of the corresponding easier problem. These approximate solutions are usually obtained by asymptotic expansions in \( \epsilon \).

An asymptotic expansion is an expansion of an expression with respect to a parameter \( \epsilon \), around a given value \( a \), with the requirement that the expansion approaches the original expression in the limit as \( \epsilon \to a \). For example, the Taylor series of a function around a point \( a \) is an asymptotic expansion (where the parameter is the variable of the function itself). However, whereas the Taylor series is unique, asymptotic expansions in general are not unique. By our definition, this should be clear; for the Taylor series, if we
simply change the coefficient of a single term, then the expansion is no longer the Taylor series, but clearly the series still approaches the value of the function as \((x - a)\) goes to zero. With all this freedom comes the responsibility to choose the expansion wisely and practically. Many authors go into more detail about additional criteria that an asymptotic expansion should satisfy, and what exactly makes a ‘good’ asymptotic expansion \([14],[21]\). But for present purposes, we will take a simpler, if naive approach.

This is best understood through the following example, which follows closely an example worked out in the text by Holmes \([14]\). Consider the ordinary partial differential equation

\[
  u'' + \epsilon u' + u = 0, \quad u(0) = 0, \quad u'(0) = 1. \tag{2.1.1}
\]

If \(\epsilon\) were zero, this would be a simple model for periodic harmonic motion. However, as any skydiver (who has studied physics) can attest, the presence of the first derivative causes damping in the dynamics of the solution. If \(\epsilon\) is large, we expect the solution to decay quickly. However, if \(\epsilon\) is small, then we might expect slow decay as the solution oscillates harmonically. In fact, as one can check, for \(\epsilon < 2\), the solution to this equation will be

\[
  u(t) = \frac{1}{\sqrt{1 - \epsilon^2/4}} \exp \left( -\frac{\epsilon}{2} t \right) \sin \left( \sqrt{1 - \epsilon^2/4} t \right), \tag{2.1.2}
\]

which is exactly harmonic motion modulated by a decaying exponential. Now suppose we were not able to solve this equation exactly. We might try to approximate the solution by some, as of yet unknown asymptotic expansion:

\[
  u(t) = u_0(t) + \epsilon u_1(t) + \epsilon^2 u_2(t) + \ldots \tag{2.1.3}
\]

Taking the first two terms of this expansion and plugging into the differential equation \((2.1.1)\) gives us

\[
  (u_0'' + \epsilon u_1'' + \ldots) + \epsilon (u_0' + \epsilon u_1' + \ldots) + (u_0 + \epsilon u_1 + \ldots) = 0 \tag{2.1.4}
\]
Now, let’s rearrange this by collecting similar powers of $\epsilon$:

\[(u''_0 + u_0) + \epsilon(u''_1 + u'_0 + u_1) + ... = 0.\]

If we have a power series set equal to zero, then each term (each coefficient of a given power of $\epsilon$) must itself be zero. So, for example, by looking at the $O(\epsilon^0)$ terms, we conclude

\[u''_0 + u_0 = 0.\]

The solution to this is

\[u_0(t) = A \cos(t) + B \sin(t).\]

How do we apply the boundary conditions? If the boundary conditions were expanded asymptotically in powers of $\epsilon$, each term would apply to the corresponding $u_i$. Since our current boundary conditions do not depend on $\epsilon$, then they apply only to the zeroth order ODE above (all the other $u_i$’s will then have boundary conditions equal to zero). Imposing this boundary condition gives us $u_0(t) = \sin(t)$. Next, the terms up to $O(\epsilon)$ give us

\[u''_1 + u'_0 + u_1 = 0,\]

or rather

\[u''_1 + u_1 = -\cos(t).\]

Imposing the conditions that $u_1(0) = u'(0) = 0$, we get $u_1(t) = -t/2 \sin(t)$. So, for example, if we take only the first two terms of the asymptotic expansion, we have

\[u(t) \approx u_0(t) + \epsilon u_1(t) = \left(1 - \frac{\epsilon}{2}t\right) \sin(t).\]

What should we notice about this approximation? Comparing (2.1.5) to (2.1.2) we see that indeed, the approximation approaches the true solution in the limit for small $\epsilon$. The
factor $\sqrt{1 - \epsilon^2/4}$ is approximately equal to one, while the term $(1 - \epsilon/2t)$ approximates the decaying exponential for small $\epsilon$ and small $t$. For a small, but fixed $\epsilon$, this might seem like a good approximation, and indeed it would be. However, as $t$ increases, the linear factor of $t$ in (2.1.5), which was initially responsible for decay in amplitude, begins to grow, once $t > 2/\epsilon$. This can be seen in the Figure 2.1, where $\epsilon$ is taken to be 1/10. The approximation is good for small $t$, but near the turning point $2/\epsilon \sim 20$, we see the approximate solution begins to grow, which is exactly counter to what we wanted. Let us turn, then, to a more sophisticated approach.

2.2 THE METHOD OF MULTIPLE SCALES

Similar to the solution of the advection-diffusion equation given in Section 1.4 , we see that the exact solution to our current problem (2.1.1) which is given by

$$u(t) = \frac{1}{\sqrt{1 - \epsilon^2/4}} \exp \left( -\frac{\epsilon}{2} t \right) \sin \left( \sqrt{1 - \epsilon^2/4} t \right),$$

(2.2.1)
effectively has two relevant time parameters: the harmonic oscillator sees the variable \( t \), whereas the decaying exponential sees the variable \( \epsilon t \). The idea of the method of multiple scales is the following \([14]\). Let us pretend that there are, in fact, two time variables,

\[
t_0 := t, \quad t_1 := \epsilon t,
\]

which we will treat as independent variables. Furthermore, if we have an expression \( f(t) \), where \( t \) may appear naked or as \( \epsilon t \), then we could think of this function instead as a function of two variables \( f(t, \epsilon t) \), or rather, \( f(t_0, t_1) \). The derivative with respect to \( t \) would then be

\[
\partial_t f(t) = \partial_t f(t, \epsilon t) = \partial_{t_0} f + \epsilon \partial_{t_1} f.
\]

Hence, we will associate the operator \( \partial_t \) with the operator \( \partial_{t_0} + \epsilon \partial_{t_1} \). This implies further that

\[
\partial_t^2 = \partial_{t_0}^2 + 2\epsilon \partial_{t_0} t_1 + \epsilon^2 \partial_{t_1}^2.
\]

Moreover, we are still going to search for an asymptotic expansion, only now it will depend on both of our time parameters:

\[
u(t) = u_0(t_0, t_1) + \epsilon u_1(t_0, t_1) + \ldots.
\]

Let us see what effect this has on the original differential equation \((2.1.1)\).

\[
\left( \partial_{t_0}^2 + 2\epsilon \partial_{t_0} t_1 + \epsilon^2 \partial_{t_1}^2 \right) \left( u_0 + \epsilon u_1 + \ldots \right) + \epsilon \left( \partial_{t_0} + \epsilon \partial_{t_1} \right) \left( u_0 + \epsilon u_1 + \ldots \right)
\]

\[
+ \left( u_0 + \epsilon u_1 + \ldots \right) = 0, \quad (2.2.2)
\]
keeping in mind that $u_0, u_1, \ldots$ depend on both $t_0$ and $t_1$. As before, we will collect all the terms for each power of $\epsilon$ and demand that each must independently vanish. To order $O(\epsilon^0)$ we have

$$\partial_{t_0}^2 u_0 + u_0 = 0.$$  

The solution to this is typically

$$u_0 = A \cos(t_0) + B \sin(t_0),$$

but keeping in mind that $u_0$ could now depend also on $t_1$, we modify:

$$u_0(t_0, t_1) = A(t_1) \cos(t_0) + B(t_1) \sin(t_0). \quad (2.2.3)$$

Notice how any potential $t_1$ dependence is contained in the degrees of freedom that belong to the solution of the $O(\epsilon^0)$ differential equation. Now, we apply the initial conditions exclusively to $u_0$, as before, since they do not depend on $\epsilon$. First we have

$$u_0(0, 0) = A(0) = 0. \quad (2.2.4)$$

Secondly, we have

$$u_0'(0, 0) = (\partial_{t_0} + \epsilon \partial_{t_1}) u_0|_{(0,0)} = B(0) + \epsilon A'(0) = 0.$$

Since the coefficients of various powers of $\epsilon$ must independently vanish, then we have

$$B(0) = 0 \quad A'(0) = 0. \quad (2.2.5)$$

Evidently, we can’t fully solve for $u_0$ yet, there is still ambiguity in the functions $A(t_1)$ and $B(t_1)$. However, it turns out that if we investigate the term $u_1$, we can clear up this
ambiguity without ever solving for $u_1$. Going back to (2.2.2), we look at the $O(\epsilon)$ terms. We have

$$2\partial_t \partial_t u_0 + \partial_t^2 u_1 + \partial_t u_0 + u_1 = 0.$$ 

Substituting our explicit expression for $u_0$ and rearranging gives us

$$\left(\partial_t^2 + 1\right) u_1 = A(t_1) \sin(t_o) - B(t_1) \cos(t_o) - 2A'(t_1) \sin(t_o) - 2B'(t_1) \cos(t_o)$$

$$= \left(A(t_1) + 2A'(t_1)\right) \sin(t_o) - \left(B(t_1) + 2B'(t_1)\right) \cos(t_o)$$

(2.2.6)

Solving this equation would be lengthy and tedious. Instead, we will see what conditions this equation imposes on the solution for $u_0$ (2.2.3). To do this we will take a quick detour into functional analysis.

Suppose we have a linear operator $L$ in a linear vector space $V$ with an inner product $\langle f, g \rangle$. It is natural to ask whether, for a given $v \in V$, there exists a solution $u$ to the equation $Lu = v$. It turns out (not surprisingly) that this is guaranteed if and only if the kernel of $L$ is empty. This is known as the Fredholm alternative [10] (when it is phrased as an either/or statement). One (more surprising) manifestation of this is that there exists a solution $u$ if and only if $v$ is orthogonal to every element of the kernel of the adjoint of $L$, denoted by $L^\dagger$.

$$\left\{\text{solution of } Lu = v \text{ exists} \right\} \Leftrightarrow \left\{ \langle v, \varphi \rangle = 0, \forall \varphi \in \ker(L^\dagger) \right\}.$$
Proving the ‘only if’ part of this statement is much easier than proving the ‘if’ part, and fortunately, that is all we will need. Specifically, assume that for a given \( v \), there exists a \( u \) such that \( L u = v \). Let \( \varphi \in \ker(L^+) \). Then

\[
\langle v, \varphi \rangle = \langle L u, \varphi \rangle = \langle u, L^+ \varphi \rangle = \langle u, 0 \rangle = 0,
\]

and hence, \( v \) is orthogonal to \( \ker(L^+) \).

Now, returning to equation (2.2.6), we can cast this in the form \( L u = v \) if we identify

\[
L := \partial_{t_0}^2 + 1
\]

and

\[
v := A(t_1) \sin(t_0) - B(t_1) \cos(t_0) - 2A'(t_1) \sin(t_0) - 2B'(t_1) \cos(t_0).
\]

We are going to apply the Fredholm Alternative to (2.2.6) but this requires knowing the adjoint of \( L \), which in turn, requires that we pick an inner product. Since all the \( t_0 \) dependence so far has been through periodic functions on \([0,2\pi]\), let us choose our vector space to be functions that are periodic on \([0,2\pi]\) with respect to the \( t_0 \) variable, which gives us a natural inner product

\[
\langle f, g \rangle = \frac{1}{2\pi} \int_{0}^{2\pi} f g \, dt_0.
\]

28
Then we can compute the adjoint of $L$ using integration by parts:

\[
\langle f, Lg \rangle = \frac{1}{2\pi} \int_0^{2\pi} f \left( \partial^2_{t_o} + 1 \right) g \, dt_o
\]

\[
= \frac{1}{2\pi} \int_0^{2\pi} f \left[ \partial^2_{t_o} g \right] \, dt_o + \frac{1}{2\pi} \int_0^{2\pi} f g \, dt_o
\]

\[
= \frac{1}{2\pi} \int_0^{2\pi} \left( \partial^2_{t_o} + 1 \right) f g \, dt_o
\]

\[
= \langle L^\dagger f, g \rangle
\]

where we see that $L^\dagger = L$. Going from the second to third line above is justified by integrating by parts twice on the first term and noticing that the boundary terms cancel out each time. Now, what is the kernel of $L^\dagger$? As was stated before when solving (2.2.3), we have

\[
\ker(L^\dagger) = \{ C \cos(t_o) + D \sin(t_o) \}.
\]

The Fredholm Alternative says that if we want a solution $u_1$ of (2.2.6) to exist, then we must have $v$ be orthogonal to a member of $\ker(L^\dagger)$. In other words,

\[
0 = \frac{1}{2\pi} \int_0^{2\pi} \left( C \cos(t_o) + D \sin(t_o) \right) \left[ A(t_1) \sin(t_o) - B(t_1) \cos(t_o) \right.
\]

\[
- 2A'(t_1) \sin(t_o) - 2B'(t_1) \cos(t_o) \] \, dt_o.
\]

Using the fact that $\int_0^{2\pi} \cos(t) \sin(t) \, dt = 0$ and that $\int_0^{2\pi} \cos^2(t) \, dt = \int_0^{2\pi} \sin^2(t) \, dt = \pi$, the previous line reduces to

\[
0 = D \left[ A(t_1) - 2A'(t_1) \right] + C \left[ B(t_1) - 2B'(t_1) \right].
\]
But since the choice of $D$ and $C$ were arbitrary, then we must have

\[ A(t_1) - 2A'(t_1) = 0 \quad B(t_1) - 2B'(t_1) = 0. \]

Combining this with the initial conditions (2.2.4), (2.2.5), we have

\[ A(t_1) = 0 \quad B(t_1) = \exp \left( -\frac{1}{2} t_1 \right). \]

Though we haven’t yet solved for $u_1$ (and don’t really care to, anyway) we have actually cleared up any ambiguity from (2.2.3). Putting it all together we have

\[ u_0(t_o, t_1) = \exp \left( -\frac{1}{2} t_1 \right) \sin(t_o), \]

or, in terms of the original variable $t$, we have

\[ u_o(t) = \exp \left( -\frac{\epsilon}{2} t \right) \sin(t). \quad (2.2.8) \]

We could go on and try to explicitly solve for $u_1$, but let us stop here and consider how much of an improvement (2.2.8) already is over the previous approximation (2.1.5). This approximation has the same desired oscillatory behavior, namely $\sin(t)$, which closely mimics (2.2.1). However, the unbounded growth that plagued our last approximation (2.1.5) is no longer present. In fact, (2.2.8) decays at exactly the same rate as the true solution. Compare them visually in Figure 2.2. They are indistinguishable to the eye.
Now, a lot has happened in the last few pages, so let us recap the important steps that happened here, and in a sense outline the procedure that we are going to see over and over again:

1. We were confronted with a differential equation which had a small parameter $\epsilon$. Suspecting that the solution would display two types of behavior, one depending on $t$ and the other on $\epsilon t$, we introduce two new time variables $t_0$ and $t_1$.

2. Then, writing out the solution as an asymptotic expansion in $\epsilon$ which depends on the two new time parameters, we collect the $O(\epsilon^0)$ terms and solve the resulting differential equation corresponding to $u_0$.

3. At this stage, however, $u_0$ was not completely determined, as there was still $t_1$ ambiguity contained in the degrees of freedom.

4. Recognizing the $O(\epsilon)$ differential equation for $u_1$ as a linear operator equation, we set up the Fredholm Alternative integral, which will henceforth be referred to as the FAI.
5. This integral forced conditions that gave us an evolution equation, determining how $u_0$ depends on $t_1$.

6. Having obtained a well defined $u_0$ we stopped there, rather satisfied with how accurate it appeared to be.

This is going to be, more or less, the procedure we follow for the rest of this work, so it would be behoove the reader to internalize this procedure at this point.

2.3 MULTI-SCALE ANALYSIS ON A PDE

Now let us apply this method to a partial differential equation. We return to the advection-diffusion equation from section 1.4:

$$\frac{\partial u}{\partial t}(x, t) + \cos(t)\frac{\partial u}{\partial x}(x, t) = \epsilon\frac{\partial^2 u}{\partial x^2}(x, t) \quad (2.3.1)$$

$$u(x, 0) = f(x)$$

where $x \in \mathbb{R}$, $t \in [0, \infty)$, and the only ‘boundary conditions’ are that we demand that the function $u$ decay to zero at infinity. As before, we will switch to the two time parameters $t_0 = t$ and $t_1 = \epsilon t$. Also as before, we expand $u$ asymptotically in $\epsilon$. Then (2.3.1) becomes

$$\left( \partial_{t_0} + \epsilon \partial_{t_1} \right) \left( u_0 + \epsilon u_1 + \ldots \right) + \cos(t_0)\partial_x \left( u_0 + \epsilon u_1 + \ldots \right)$$

$$= \epsilon \partial_{xx} \left( u_0 + \epsilon u_1 + \ldots \right). \quad (2.3.2)$$

Now, collecting the $O(\epsilon^0)$ terms, we have

$$\partial_{t_0} u_0 + \cos(t_0)\partial_x u_0 = 0. \quad (2.3.3)$$
Computing the characteristic map, we see that solutions to this equation are given by 
\[ u_0 = f(x - \sin(t_0)), \] as one can easily check. Unfortunately, this does not include any \( t_1 \) dependence. However, just as in (2.2.3), where we snuck the \( t_1 \) dependence into the underdetermined constants, similarly, we will use the envelope \( f \) to sneak in the \( t_1 \) dependence:

\[ u_0(t_0, t_1) = F\left(x - \sin(t_0), t_1\right), \tag{2.3.4} \]

where \( F(x, 0) = f(x) \). It is simple matter to substitute this back into (2.3.3) and check that this indeed solves the equation. Again, however, we are left with \( t_1 \) ambiguity. We have no idea, as of yet, how \( F(t_0, t_1) \) depends on \( t_1 \), only that it *does* depend on \( t_1 \). However, as we saw before, it will be in our investigation of the \( O(\epsilon) \) terms that we will clear up this ambiguity.

The \( O(\epsilon) \) term in (2.3.2) gives us the equation

\[ \left( \partial_{t_0} - \cos(t_0)\partial_x \right) u_1 = \partial_{xx} u_0 - \partial_{t_1} u_0. \tag{2.3.5} \]

If we define

\[ L := \partial_{t_0} - \cos(t_0)\partial_x \quad v := \partial_{xx} u_0 - \partial_{t_1} u_0, \]

then this can again be cast as an equation \( Lu = v \). To apply the Fredholm Alternative to this, we need to compute \( L^\dagger \). Again, we must define a vector space with a suitable inner product. Since our \( t_0 \) dependence seems to be periodic, let us define our vector space to be the space of functions that are periodic with period \( 2\pi \) and integrable in \( x \). Then a suitable inner product would be

\[ \langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} f g \, dx \, dt_0. \tag{2.3.6} \]
Now, given this inner product, what is the adjoint of $L$? We compute, using integration by parts:

$$\langle f, Lg \rangle = \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} f \left( \partial_{t_0} - \cos(t_0) \partial_x \right) g \, dx \, dt_0$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} f \left[ \partial_{t_0} g \right] \, dx \, dt_0 - \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} f \left[ \cos(t_0) \partial_x g \right] \, dx \, dt_0$$

$$= -\frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} \left[ \partial_{t_0} f \right] g \, dx \, dt_0 + \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} \left[ \cos(t_0) \partial_x f \right] g \, dx \, dt_0$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} \left( -\partial_{t_0} + \cos(t_0) \partial_x \right) f g \, dx \, dt_0$$

$$= \langle L^+ f, g \rangle,$$

which tells us that $L^+ = -L$. (The boundary terms from integration by parts in the third line vanished for the time integral because of periodicity, and vanished for the space integral because of decay at infinity). Now we need the kernel of $L^+$. As was we saw when solving (2.3.3), we have

$$\ker(L^+) = \{ w(x - \sin(t_0)) \},$$

where $w$ can be any function of one variable. Thus, to set up the FAI, we need the inner product of an element of $\ker(L^+)$ with the right hand side of (2.3.5):

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} w \left( x - \sin(t_0) \right) \left[ \partial_{xx} u_0 - \partial_t u_0 \right] \, dx \, dt_0.$$

Now, remembering our expression for $u_0$ given by (2.3.4), we have

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{-\infty}^{\infty} w \left( x - \sin(t_0) \right) \left[ \partial_{xx} F \left( x - \sin(t_0), t_1 \right) - \partial_t F \left( x - \sin(t_0), t_1 \right) \right] \, dx \, dt_0.$$
Notice how often the expression $x - \sin(t_o)$ occurs in this integral, almost as if it’s asking for a substitution. Let us oblige. Define $\bar{x} := x - \sin(t_o)$. Then

$$d\bar{x} = dx \quad \partial_{xx} = \partial_{\bar{x}\bar{x}},$$

and so our integral becomes

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_R w(\bar{x}) \left[ \partial_{\bar{x}\bar{x}} F(\bar{x}, t_1) - \partial_{t_1} F(\bar{x}, t_1) \right] d\bar{x} dt_0.$$  \hspace{1cm} (2.3.7)

Now the $t_o$ dependence has been completely masked in this expression, so we may as well integrate in the $t_o$ variable:

$$0 = \int_R w(\bar{x}) \left[ \partial_{\bar{x}\bar{x}} F(\bar{x}, t_1) - \partial_{t_1} F(\bar{x}, t_1) \right] d\bar{x}. \hspace{1cm} (2.3.8)$$

But, remember that $w$ was a completely arbitrary function. The only way this integral can be guaranteed to vanish, for any $w$, is if the expression inside the braces vanishes:

$$\partial_{\bar{x}\bar{x}} F(\bar{x}, t_1) - \partial_{t_1} F(\bar{x}, t_1) = 0.$$

A more familiar way of writing this is

$$\partial_{t_1} F(\bar{x}, t_1) = \partial_{\bar{x}\bar{x}} F(\bar{x}, t_1),$$

that is, the diffusion equation.

Recall that the ambiguity in (2.3.4) was that we didn’t know how $F$ depended on $t_1$. Well, now we do; it apparently evolves according to a diffusion equation. Specifically, as we saw in Chapter 1, the solution is given by

$$F(\bar{x}, t_1) = \frac{1}{\sqrt{4\pi t_1}} \int_R f(y) \exp \left( -\frac{(\bar{x} - y)^2}{4t_1} \right).$$
Thus, we have

$$u_0(x, t_0, t_1) = \frac{1}{\sqrt{4\pi t_1}} \int_{\mathbb{R}} f(y) \exp \left( \frac{-(x - \sin(t_0) - y)^2}{4t_1} \right),$$

or returning to the original time variable

$$u_0(x, t) = \frac{1}{\sqrt{4\pi \epsilon t}} \int_{\mathbb{R}} f(y) \exp \left( \frac{-(x - \sin(t) - y)^2}{4\epsilon t} \right).$$

Recall that this was only the first term in the asymptotic expansion. We can ask how accurate is this approximation? Well, in this particular case, it’s exact! This is just a matter of luck, as we’ll see later. In fact, as long as the advection coefficient (in this case $\cos(t)$) is independent of space, this procedure will be exact. Recall the transport-diffusion equation from Section 1.4, where the advection coefficient was just a constant $a$. Repeating this whole process, we would end up with

$$u_0 = \frac{1}{\sqrt{4\pi \epsilon t}} \int_{\mathbb{R}} f(y) \exp \left( \frac{-(x - at - y)^2}{4\epsilon t} \right).$$

If, once again, we take the initial condition $f(x) = \exp(-bt^2)$, then we can compute this integral explicitly and we end up with

$$u_0(x, t) = \frac{1}{\sqrt{1 + 4\epsilon t}} \exp \left( \frac{-b(x + at)^2}{1 + 4\epsilon t} \right),$$

exactly as before.

Recall the step that got us from (2.3.7) to (2.3.8), integration with respect to $t_0$. At first, this seems like a fairly innocuous step. However, we will in see in less trivial examples that this step is actually quite meaningful. It is here that we are effectively ‘averaging out’ the faster scale phenomena (that which depends on $t_0$) to uncover what kind of net effect
it has on the slower scale phenomena, i.e., the evolution of \( F(\cdot, t_1) \). We will discuss this further in section 2.5.

### 2.4 A Not-Quite-As-Trivial Example

For the next example, we will look at another advection-diffusion equation, similar to the equation (2.3.1), but with a slightly more interesting advection coefficient. In this and all the following examples, the procedure will be the same as before, outlined in section 2.2. It may help the reader to review these steps before we proceed again.

Let us begin. Consider the equation

\[
\frac{\partial u}{\partial t} (x, t) + \cos(t) x \frac{\partial u}{\partial x} (x, t) = \epsilon \frac{\partial^2 u}{\partial x^2} (x, t) \tag{2.4.1}
\]

\[ u(x, 0) = f(x). \]

This is almost the same as the previous example, except for the additional \( x \) in the advection coefficient. As before, we both (a) introduce two time parameters \( t_0 \) and \( t_1 \), and (b) expand in powers of \( \epsilon \):

\[ u(x, t_0, t_1) = u_0(x, t_0, t_1) + \epsilon u_1(x, t_0, t_1) + \ldots \]

Plugging this into equation (2.4.1) gives us

\[
\left( \partial_{t_0} + \epsilon \partial_{t_1} \right) (u_0 + \epsilon u_1 + \ldots) + \cos(t_0) x \partial_{x} (u_0 + \epsilon u_1 + \ldots) \\
= \epsilon \partial^2_{x} (u_0 + \epsilon u_1 + \ldots). \tag{2.4.2}
\]
We proceed as before. Collecting the $O(\epsilon^0)$ terms, we have

$$\partial_{t_0} u_0 + \cos(t_0)x \partial_x u_0 = 0.$$  \hspace{1cm} (2.4.3)

In section 1.4, we computed the characteristic map of this equation \[\text{(1.5.3)}, \text{namely } x_0(x,t_0) = xe^{-\sin(t)}. \text{Thus, the general solution to equation (3.1.2) will be}

$$u_0 = F(xe^{-\sin(t)}, t_1).$$  \hspace{1cm} (2.4.4)

with $F(\cdot,0) = f(\cdot)$. As expected, there is ambiguity present in $u_0$. Like before, we will investigate the $O(\epsilon^1)$ terms, set up the FAI, and clear up this ambiguity. Collecting the $O(\epsilon^1)$ terms from \[\text{(2.4.2)}\] gives us

$$(\partial_{t_0} + \cos(t_0)x \partial_x) u_1 = \partial_x^2 u_0 - \partial_{t_1} u_0.$$  

Once again, define

$$L := \partial_{t_0} + \cos(t_0)x \partial_x, \quad v := \partial_x^2 u_0 - \partial_{t_1} u_0.$$ 

The Fredholm alternative tells us that if we wish for a solution $u_1$ to exist, we must have $v$ orthogonal the kernel of $L^\dagger$, which of course means we must choose an inner product. Since our advection is periodic on $[0,2\pi]$, we will once again choose the inner product from \[\text{(2.3.6)}\]. First, we use this to identify the adjoint of $L$. (Note, once again, that when
integrating by parts, all boundary terms will vanish; for the spatial integrals, this is because of decay at infinity, for the time integrals, this is because of periodicity).

\[
\langle f, Lg \rangle = \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} f(\partial_{t_0} + \cos(t_0)x\partial_x)g\, dx\, dt_0
\]

\[
= \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} f[\partial_{t_0}g]\, dx\, dt_0 - \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} f[\cos(t_0)x\partial_x g]\, dx\, dt_0
\]

\[
= -\frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} [\partial_{t_0}f]g\, dx\, dt_0 - \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} [\cos(t_0)x\partial_x f]g\, dx\, dt_0
\]

\[
= \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} \left( -\partial_{t_0} + \cos(t_0)x\partial_x - \cos(t)f \right) g\, dx\, dt_0
\]

\[
= \langle L^+ f, g \rangle
\]

and so we see that

\[ L^+ = -L - \cos(t_0). \]

Next, we must find the kernel of $L^+$. Suppose $\varphi \in \ker(L^+)$. Then

\[ \varphi_{t_0} + \cos(t_0)x\varphi_x = -\cos(t_0)\varphi. \tag{2.4.5} \]

A trick for solving this equation is to decompose $\varphi$ as the product of two functions, one that depends exclusively on $x_0(x, t_0)$ and one that doesn’t:

\[ \varphi(x, t_0, t_1) = W(x_0(x, t_0), t_1)Y(x, t_0, t_1). \]

Furthermore, let us write the function $Y$ as an exponential:

\[ \varphi(x, t_0, t_1) = W(x_0(x, t_0), t_1)e^{p(x, t_0, t_1)}. \]
(There is no loss of generality here because we could allow \( p \) to take on complex or singular values). Then, plugging this into equation (2.4.5), we have

\[
W'(x_o)\partial_t x_o e^p + W(x_o)e^p p_{t_o} + \cos(t_o)xW'(x_o)\partial_x x_o e^p + \cos(t_o)xW(x_o)e^p p_x = -\cos(t_o)W(x_o)e^p,
\]

or upon rearranging,

\[
W'(x_o)e^p \left[ \partial_t x_o + \cos(t_o)x\partial_x x_o \right] + W(x_o)e^p \left[ p_{t_o} + \cos(t)p_x + \cos(t) \right] = 0.
\]

From our discussion of characteristics, we know that the characteristic map \( x_o(x, t_o) \) will satisfy \( Lx_o = 0 \), and so \( \partial_t x_o + \cos(t_o)x\partial_x x_o = 0 \), leaving us with simply

\[
p_{t_o} + \cos(t)x p_x = -\cos(t_o).
\]

This equation is now the same original advection equation but with an additional inhomogeneous term on right hand side. As we saw in section 1.5, the solution will be something of the form

\[
p(x, t_o, t_1) = p_{h}(x_o(x, t_o), t_1) + \tilde{p}(x, t_o, t_1),
\]

where \( \tilde{p} \) is the particular solution given by the integral (1.6.2). In this case, one can easily check that the particular solution will be \( \tilde{p}(x, t_o, t_1) = -\sin(t_o) \). The term \( p_{h} \) can be absorbed into \( W \) since it depends only on \( x_o(x, t_o) \) and \( t_1 \). Thus, we’re left with

\[
\varphi(x, t_o, t_1) = W(x_o(x, t_o), t_1)e^{-\sin(t_o)}.
\]
The arbitrariness of $W$ is what allows us to span the entire space of $\ker(L^\dagger)$:

$$\ker(L^\dagger) = \left\{ W(x_0(x, t_0), t_1)e^{-\sin(t_0)} \parallel W \text{ is arbitrary} \right\}.$$ 

Now we set up the FAI. Recall we want $\langle \phi, v \rangle = 0$. This takes the form

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0(x, t_0), t_1)e^{-\sin(t_0)} \left[ \partial_x^2 u_0 - \partial_{t_1} u_0 \right] dx dt_0.$$  \hspace{1cm} (2.4.6)

Now, if we write out explicitly what $u_0$ is, we have

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0(x, t_0), t_1)e^{-\sin(t_0)} \left[ \partial_x^2 F(x_0(x, t_0), t_1) - \partial_{t_1} F(x_0(x, t_0), t_1) \right] dx dt_0.$$  \hspace{1cm} (2.4.7)

As in the previous example, this integral looks ripe for a substitution. In particular, let us integrate with respect to $dx_0$ instead of $dx$, since no naked $x$’s appear anywhere. To do this, we need to compute

$$dx = \frac{\partial x}{\partial x_0} dx_0,$$

which, by (2.4.4) is

$$dx = e^{\sin(t)} dx_0.$$

Putting this back into (2.4.7) gives us

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0, t_1)e^{-\sin(t_0)} \left[ \partial_x^2 F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] e^{\sin(t_0)} dx_0 dt_0.$$  \hspace{1cm} (2.4.7)

Fortunately, the two exponential terms cancel and we are left with

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0, t_1) \left[ \partial_x^2 F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] dx_0 dt_0.$$
Now, we need to convert the $\partial_{xx}$ operator to be in terms of $x_0$. Since

$$\partial_x = \frac{\partial x_0}{\partial x} \partial_{x_0},$$

then, as we have already computed,

$$\partial_x = e^{-\sin(t_0)} \partial_{x_0}.$$  

This implies further that

$$\partial^2_x = \partial_x \partial_x = e^{-2\sin(t_0)} \partial^2_{x_0}.$$  

Plugging this back into our integral, we have

$$0 = \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} W(x_0, t_1) \left[ e^{-2\sin(t_0)} \partial_{x_0} x_0 F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] dx_0 dt_0. \quad (2.4.8)$$

Now we have completely converted to the $x_0$ variable. Notice that the $t_0$ dependence only occurs in the coefficient of $\partial_{x_0} x_0$. Let us now perform the $t_0$ integral, (denoting $\langle \cdot \rangle = \frac{1}{2\pi} \int_0^{2\pi} \cdot dt_0$),

$$0 = \int_\mathbb{R} W(x_0, t_1) \left[ \langle e^{-2\sin(t_0)} \rangle \partial^2_{x_0} F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] dx_0 \quad (2.4.9)$$

Since $W$ was a completely arbitrary function, then the only way to guarantee that this integral vanishes is if the expression in the brackets vanishes:

$$\langle e^{-2\sin(t_0)} \rangle \partial^2_{x_0} F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) = 0.$$  

Once again, this is just a diffusion equation with a diffusivity of $\langle e^{-2\sin(t)} \rangle \approx 2.279$:

$$\partial_{t_1} F(x_0, t_1) = 2.279 \partial^2_{x_0} F(x_0, t_1). \quad (2.4.10)$$
Let us pause here to articulate exactly what this is telling us. Our approximate solution

\[ u_0(x, t_0, t_1) = F\left(x_0(x, t_0), t_1\right), \]

will be given by looking at the spatial profile \( F(\cdot, t_1) \) which evolves according to \((2.4.10)\), and interpolating this profile over the characteristic map. There are two main points to be made here. The first is that, in effect, we have (approximately) decomposed the dynamics of the solution into a purely advective part (the characteristic map) and a purely diffusive part (equation \((2.4.10)\)). The second point is that this was possible because of the step going from \((2.4.8)\) to \((2.4.9)\), where we integrated in \( t_0 \). The \( t_0 \) dependence was in the exponential term, but by integrating and computing it’s average value, we obtained the number 2.279, which is the effective diffusion rate of the envelope \( F \). This is precisely the idea hinted at earlier of ‘averaging out’ the fast scale dynamics in order to see what effect it has on the slow scale dynamics. The biggest take away from this particular example is the fact that the advection tends to effectively double the diffusion rate.

### 2.5 Discussion

An important point which is illustrated in the last example is that we essentially “averaged out” the faster scale dynamics. After all, by introducing two time parameters, \( t_0 \) and \( t_1 \), we don’t want to simply increase the number of parameters describing the system. Rather, the real advantage of doing this is that we are decomposing the system on two different time scales, and thus we can decide which information we do and don’t care about. By averaging in \( t_0 \), we are basically deciding that the fast scale dynamics are not relevant to us as long as we can compute their effect on the slow scale dynamics. This idea of averaging a time-dependent differential equation has been around for some time, in various incarnations. Interest in systems that are best described with multiple time
scales can be traced back at least to the work of Kuzmak [18], Cole and Kevorkian [5]. Specifically, advection dominated systems with weak diffusion (sometimes called the ‘Batchelor regime’ [2]) have been popular since the midcentury. However, using multiple scales on such regimes to actually ‘average out’ the faster time scales didn’t really get it’s start until the late 80’s with the work of Krol [16], [17]. Krol not only develops explicit techniques and formulas for a certain class of problems, he also proves a theorem on the accuracy of this method (see Section 4.2 for his theorem and further discussion). There have been many subsequent developments on this idea, as described in section 1.1, but many of these developments have been limited in scope and context. For example, as mentioned in the introduction, in the work of Schaefer et al. [23], they consider a two dimensional vorticial system (which we will see in Section 6.4) on an annular domain. In order to average the equation, they perform a change of variables (to action-angle coordinates) which facilitates a further transformation of coordinates to the characteristic map, analogous to what happened in our previous example. From there, they are able to average out the faster time dependence. Providing both a convergence theorem and numerical confirmation, their technique works wonderfully. However, it is limited to 2-D, separable, incompressible, linear advection on a bounded domain with Dirichlet boundary conditions. In fact, it is common that most averaging techniques are tailor-made for particular settings. In the next two chapters, we will develop an averaging technique which is significantly more general. In particular, we will arrive at an algorithm which is capable of handling any linear advection operator in any dimension on an unbounded domain. To be explicit, some of the major restrictions that will be cast aside are that of incompressibility, separability, and time-periodicity, to name a few. Moreover, we will even treat (although not generalize) a non-linear case.

It should be mentioned here that the program of averaging out the faster scales via the multiple scale approach has also been applied to stochastic differential equations [25],[26], but this will not be pursued here.
Applications to Advection-Diffusion Equation

3.1 ONE MORE EXAMPLE OF THE MULTI-SCALE METHOD ON ADVECTION-DIFFUSION EQUATIONS

For the next example, consider the PDE

\[
\frac{\partial u}{\partial t}(x,t) + \cos(t) x^2 \frac{\partial u}{\partial x}(x,t) = \epsilon \frac{\partial^2 u}{\partial x^2}(x,t)
\] (3.1.1)

\[u(x,0) = f(x).\]

The only change from the previous example is that the advective coefficient \(x\) is now \(x^2\). The multi-scale expansion will be almost identical, and we therefore just repeat our work from the previous example, replacing \(x\) with \(x^2\). First, the \(O(\epsilon^0)\) terms give us

\[
\partial_{t_0} u_0 + \cos(t_0) x^2 \partial_x u_0 = 0.
\] (3.1.2)

From Chapter 1, we know that the solution to this advection equation is given by equation (1.5.4), namely

\[
x_0(x,t_0) = \frac{1}{\frac{1}{x} + \sin(t_0)}.
\] (3.1.3)

Thus, \(u_0\) becomes

\[u_0(x,t_0,t_1) = F(x_0(x,t_0),t_1).\]
with \( F(\cdot, 0) = f(\cdot) \). To see how \( F \) evolves, we set up the FAI for \( u_1 \). First, the \( O(\epsilon) \) terms in the multi-scale expansion give us

\[
(\partial_{t_0} + \cos(t_0)x^2\partial_x)u_1 = \partial_x^2 u_0 - \partial_{t_1}u_0.
\]

Once again, we call

\[
L := \partial_{t_0} + \cos(t_0)x^2\partial_x, \quad \nu := \partial_x^2 u_0 - \partial_{t_1}u_0.
\]

Using the same inner product as before, we can compute the adjoint of \( L \). After integrating by parts, we find

\[
L^\dagger = -L - \cos(t_0)2x.
\]

Thus, if \( \varphi \in \ker(L^\dagger) \), it will satisfy

\[
\varphi_{t_0} - \cos(t_0)x^2 \varphi_x = -\cos(t_0)2x\varphi.
\]

Like before, we decompose \( \varphi \) into

\[
\varphi = W(x_0(x, t_0), t_1)e^{p(x, t_0, t_1)},
\]

from which we find

\[
W'(x_0)\partial_{t_0}x_0e^p + W(x_0)e^p p_{t_0} + \cos(t_0)x^2W'(x_0)\partial_x x_0e^p + \cos(t_0)x^2W(x_0)e^p p_x = -\cos(t_0)2xW(x_0)e^p.
\]

Once more, this can be collected into

\[
W'(x_0)e^p \left[\partial_{t_0}x_0 + \cos(t_0)x^2\partial_x x_0\right] + W(x_0)e^p \left[p_{t_0} + \cos(t)xp_x + \cos(t_0)2x\right] = 0.
\]
from which we conclude that

\[ p_{t_0} + \cos(t)x^2p_x = -\cos(t_0)2x. \]

Again, \( p \) will have a homogeneous solution and an inhomogeneous particular solution \( p = p_h + \tilde{p} \). To find the particular solution, we use the integral in (1.6.2),

\[ \tilde{p} = -2 \log |x \sin(t_0) + 1|. \]

Putting it all together then,

\[ \ker(L^\dagger) = \left\{ W\left(x_o(x, t_0), t_1\right) \exp \left( -2 \log |x \sin(t_0) + 1| \right) \right\}, \]

which simplifies to

\[ \ker(L^\dagger) = \begin{cases} \frac{W\left(x_o(x, t_0), t_1\right)}{(x \sin(t_0) + 1)^2} \end{cases}, \]

Now we set up the FAI:

\[ 0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W\left(x_o(x, t_0), t_1\right) \left[ \partial_x^2 F(x_o(x, t), t_1) - \partial_{t_1} F(x_o(x, t_0), t_1) \right] dx dt_0. \quad (3.1.4) \]

Once again, due to the presence of so many \( x_o \)'s, we would like to switch variables from \( x \) to \( x_o \). However, there is an uncooperative factor below \( W \) which still contains a naked \( x \). Let us put that on hold, and see what happens when we exchange the differentials \( dx \) for \( dx_o \). Using (3.1.3), we see that

\[ x = \frac{1}{\frac{1}{x_o} - \sin(t_0)}. \]
Thus,

\[ dx = \frac{\partial x}{\partial x_0} dx_0 = \frac{1}{(1 - x_0 \sin(t_0))^2} dx_0. \]

Once again, using (3.1.3), this becomes

\[ dx = (1 + x \sin(t_0))^2 dx_0. \]

By an apparent miracle, this is precisely the factor we need to cancel with the denominator of (3.1.4). Putting this together we have

\[ 0 = \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} W(x_0, t_1) \left[ \partial_x^2 F(x_0, t_1) - \partial_t F(x_0, t_1) \right] dx dt. \]

All that’s left is to express \( \partial_x^2 \) in terms of \( x_0 \). This is tedious but straightforward, using (3.1.3) and its inverse. First,

\[ \partial_x = \frac{\partial x_0}{\partial x} \partial x_0 = \frac{1}{(1 + x \sin(t_0))^2} \partial x_0. \]

\[
\partial_x^2 = \frac{1}{(1 + x \sin(t_0))^4} \partial x_0 \left( \frac{1}{(1 + x \sin(t_0))^2} \partial x_0 \right)
= \frac{1}{(1 + x \sin(t_0))^4} \partial x_0^2 + \frac{1}{(1 + x \sin(t_0))^2} \left( \frac{\partial x}{\partial x_0} \frac{1}{(1 + x \sin(t_0))^2} \right) \partial x_0
= \frac{1}{(1 + x \sin(t_0))^4} \partial x_0^2 + \frac{1}{(1 + x \sin(t_0))^2} \left( \frac{\partial x}{\partial x_0} \frac{1}{(1 + x \sin(t_0))^2} \right) \partial x_0
= \frac{1}{(1 + x \sin(t_0))^4} \partial x_0^2 + \left( \frac{-2 \sin(t_0)}{(1 + x \sin(t_0))^3} \right) \partial x_0.
\]
Now we would like to put this back into our FAI (3.1.4), but then we would be introducing $x$ dependence again. Before we do that, then, let us substitute $x(x_0, t_0)$ into these expressions:

\[
\frac{\partial^2}{\partial x^2} = (1 - x_0 \sin(t_0))^4 \frac{\partial^2}{\partial x_0^2} - 2 \sin(t_0) (1 - x_0 \sin(t_0))^3 \frac{\partial}{\partial x_0}
\]

\[
= \alpha \frac{\partial^2}{\partial x_0^2} + \beta \frac{\partial}{\partial x_0},
\]

where we have introduced the names $\alpha$ and $\beta$ for the coefficients. Now this is something we can be proud to put back into (3.1.4):

\[
0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0, t_1) \left[ \alpha \frac{\partial^2}{\partial x_0^2} F(x_0, t_1) + \beta \frac{\partial}{\partial x_0} F(x_0, t_1) - \frac{\partial}{\partial t_1} F(x_0, t_1) \right] dx_0 dt_0
\]  

(3.1.5)

Now that we’ve successfully hidden all the $x$ dependence, let us integrate in the $t_0$ variable.

\[
0 = \int_{\mathbb{R}} W(x_0, t_1) \left[ \langle \alpha \rangle \frac{\partial^2}{\partial x_0^2} F(x_0, t_1) + \langle \beta \rangle \frac{\partial}{\partial x_0} F(x_0, t_1) - \frac{\partial}{\partial t_1} F(x_0, t_1) \right] dx_0,
\]

where

\[
\langle \alpha \rangle = \frac{1}{2\pi} \int_0^{2\pi} (1 - x_0 \sin(t))^4 dt = 1 + 3x_0^3 + \frac{3}{8}x_0^4
\]

\[
\langle \beta \rangle = \frac{1}{2\pi} \int_0^{2\pi} -2 \sin(t) (1 - x_0 \sin(t))^3 dt = 3x_0 + \frac{3}{4}x_0^3.
\]

If this integral is to be identically zero for any function $W$, then we must have the expression in the brackets vanish:

\[
\frac{\partial}{\partial t_1} F(x_0, t_1) = \langle \alpha \rangle \frac{\partial^2}{\partial x_0^2} F(x_0, t_1) + \langle \beta \rangle \frac{\partial}{\partial x_0} F(x_0, t_1),
\]

or

\[
\frac{\partial}{\partial t_1} F(x_0, t_1) = \left[ 1 + 3x_0^3 + \frac{3}{8}x_0^4 \right] \frac{\partial^2}{\partial x_0^2} F(x_0, t_1) + \left[ 3x_0 + \frac{3}{4}x_0^3 \right] \frac{\partial}{\partial x_0} F(x_0, t_1).
\]

(3.1.6)
Again, this is the equation which tells us how $F$ evolves. Though we can’t solve it explicitly right now (it is, after all another advection-diffusion equation), it tells us the slow scale evolution of the spatial profile which we then interpolate over the characteristic map.

3.2 General Formula

The procedure was the same for the previous three examples, and some of the steps should be starting to feel familiar. It is therefore time to develop a general formula. Consider the general advection-diffusion equation:

$$\frac{\partial u}{\partial t}(x,t) + a(x,t)\frac{\partial u}{\partial x}(x,t) = \epsilon \frac{\partial^2 u}{\partial x^2}(x,t) \quad (3.2.1)$$

$$u(x,0) = f(x).$$

Let us assume that $a(x,t)$ is periodic in $t$ with period $2\pi$, so that we may use the same inner product as before. Now, setting up a multi-scale asymptotic expansion, we have

$$\left(\partial_{t_0} + \epsilon \partial_{t_1}\right)(u_0 + \epsilon u_1 + ...) + a(x,t_0)\partial_x(u_0 + \epsilon u_1 + ...)$$

$$= \epsilon \partial_x^2(u_0 + \epsilon u_1 + ...). \quad (3.2.2)$$

The $O(\epsilon^0)$ terms give us

$$\partial_{t_0}u_0 + a(x,t_0)\partial_xu_0 = 0.$$  

Without specifying $a(x,t)$, we can’t possibly solve this equation. But we know how we would solve it - by finding the characteristic map $x_0(x,t_0)$. Then the solution will be

$$u(x,t_0,t_1) = F\left(x_0(x,t_0),t_1\right).$$
The $O(\epsilon)$ terms give us
\[
\left(\partial_{t_0} + a(x, t_0)\partial_x\right)u_1 = \partial^2_x u_0 - \partial_t u_0.
\]

Defining
\[
L := \partial_{t_0} + a(x, t_0)\partial_x \quad \nu := \partial^2_x u_0 - \partial_t u_0,
\]
we compute $L^\dagger$:
\[
\langle f, Lg \rangle = \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} f(\partial_{t_0} + a(x, t_0)\partial_x) g \, dxdt_0
\]
\[
= \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} f[\partial_{t_0}g] \, dxdt_0 + \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} f[a(x, t_0)\partial_xg] \, dxdt_0
\]
\[
= -\frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} [\partial_{t_0}f] g \, dxdt_0 - \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} [a(x, t_0)\partial_xf] g \, dxdt_0
\]
\[
= \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} \left[ -\partial_{t_0} - a(x, t_0)\partial_x - a_x(x, t_0) \right] f \, dxdt_0
\]
\[
= \langle L^\dagger f, g \rangle
\]

Evidently, $L^\dagger = -L - a_x(x, t_0)$, which one can check, is in agreement with the previous examples. In computing the kernel of $L^\dagger$, we search for a function $\varphi$ satisfying
\[
L\varphi = -a_x\varphi.
\]

We decompose $\varphi = W(x_0(x, t_0), t_1) e^{p(x, t_0, t_1)}$, as before. Then
\[
W'(x_0)\partial_{t_0} x_0 e^{p} + W(x_0) e^{p} p_{t_0} + a(x, t_0) W'(x_0)\partial_x x_0 e^{p} + a(x, t_0) W(x_0) e^{p} p_x
\]
\[
= -a_x(x, t_0) W(x_0) e^p,
\]
Again, rearranging this gives us

\[ W'(x_0)e^p \left[ \partial_{t_0} x_0 + a(x, t_0) \partial_x x_0 \right] + W(x_0)e^p \left[ p_{t_0} + a(x, t_0) p_x + a_x(x, t_0) \right] = 0. \]

Recalling that the characteristic map itself satisfies the original advection equation, this then reduces to

\[ p_{t_0} + a(x, t_0) p_x = -a_x(x, t_0). \]

The solution to this equation will be a combination of a homogeneous solution \( p_h \) with the particular solution \( \tilde{p} \):

\[ p(x, t_0, t_1) = p_h(x_0(x, t_0), t_1) + \tilde{p}(x, t_0, t_1). \]

The homogeneous solution can be absorbed into \( W \), and we are left with

\[ \ker(L^+) = \left\{ W(x_0(x, t_0), t_1)e^{\tilde{p}(x, t_0, t_1)} \right\}. \]

Setting up the FAI, we have

\[ 0 = \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} W(x_0(x, t_0), t_1)e^{\tilde{p}(x, t_0, t_1)} \left[ \partial_x^2 u_0 - \partial_{t_1} u_0 \right] dx dt_0. \quad (3.2.3) \]

Now, at this point in the previous two examples, we computed the change of differentials, which miraculously introduced a term that canceled with the \( e^\theta \) term in the integral. But without knowing \( x_0(x, t_0) \) explicitly, we can’t actually do this computation now. However, there is a wonderful result that will take care of this for us.

**Theorem 1.** Let \( x_0(x, t) \) be the characteristic map for an advection equation \( Lu = 0 \), where \( L = \partial_t + a(x, t) \partial_x \). Then \( \frac{\partial x_0}{\partial x} \in \ker(L^+) \).
This hardly seems like a theorem at first, because the proof will be so simple. However, we will see in the next chapter how this is actually a particular case of a bigger, more legitimate theorem.

Proof. We already know that the characteristic map \( x_0(x, t) \) is in the kernel of \( L \):

\[
\partial_t x_0 + a(x, t) \partial_x x_0 = 0.
\]

Now take the derivative of both sides of this equation with respect to \( x \):

\[
\partial_t \frac{\partial x_0}{\partial x} + a(x, t) \partial_x x_0 + a(x, t) \partial_x \frac{\partial x_0}{\partial x} = 0.
\]

Or, slightly rearranging,

\[
0 = \partial_t \frac{\partial x_0}{\partial x} + a(x, t) \partial_x \frac{\partial x_0}{\partial x} + a(x, t) \partial_x \frac{\partial x_0}{\partial x} - a(x, t) \partial_x \frac{\partial x_0}{\partial x} = -L^\dagger \left( \frac{\partial x_0}{\partial x} \right)
\]

and hence \( \frac{\partial x_0}{\partial x} \in \ker(L^\dagger) \). □

Since we know that \( \frac{\partial x_0}{\partial x} \in \ker(L^\dagger) \), and we know what the kernel of \( L^\dagger \) looks like, then we conclude

\[
\frac{\partial x_0}{\partial x} = V(x_0(x, t_0), t_1) e^{\beta(x_0, t_0, t_1)}
\]

for some \( V \). Consequently, we have

\[
\frac{\partial x}{\partial x_0} = \frac{1}{V(x_0(x, t_0), t_1) e^{\beta(x_0, t_0, t_1)}}.
\]
Thus, the differentials convert via

\[
dx = \frac{\partial x}{\partial x_0} dx_0 = \frac{1}{V(x_0(x, t_0), t_1) e^{\tilde{p}(x, t_0, t_1)}} dx_0.
\]

Now, plugging this into equation (3.2.3), the Fredholm alternative, we have

\[
0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0, t_1) e^{\tilde{p}(x, t_0, t_1)} \left[ \frac{\partial^2 F(x, t_1) - \partial t_1 F(x, t_1)}{V(x_0, t_1) e^{\tilde{p}(x, t_0, t_1)}} \right] dx_0 dt_0.
\]

Now we can see explicitly from where came the fortuitous cancellations of the exponential term. The integrand reduces to (absorbing \( V \) into \( W \))

\[
0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0, t_1) \left[ \frac{\partial^2 F(x_0, t_1) - \partial t_1 F(x_0, t_1)}{V(x_0, t_1) e^{\tilde{p}(x_0, t_1)}} \right] dx_0 dt_0. \tag{3.2.4}
\]

Now all that’s left is the business of putting \( \partial^2 x \) into the language of \( x_0 \). Since

\[
\partial_x = \frac{\partial x_0}{\partial x} \partial_{x_0},
\]

then we can compute

\[
\partial_{xx} = \partial_x \partial_x = \partial_x \left( \frac{\partial x_0}{\partial x} \right) \left( \frac{\partial x_0}{\partial x} \right) = \left( \frac{\partial x_0}{\partial x} \right)^2 \partial_{x_0} \partial_{x_0} + \frac{\partial x_0}{\partial x} \left( \partial_{x_0} \frac{\partial x_0}{\partial x} \right) \partial_{x_0} = \left( \frac{\partial x_0}{\partial x} \right)^2 \partial_{x_0} \partial_{x_0} + \frac{\partial x_0}{\partial x} \frac{\partial x}{\partial x} \frac{\partial x_0}{\partial x} \left( \frac{\partial x_0}{\partial x} \right) \partial_{x_0} = \left( \frac{\partial x_0}{\partial x} \right)^2 \partial_{x_0} \partial_{x_0} + \frac{\partial^2 x_0}{\partial x^2} \partial_{x_0}.
\]

54
For the rest of this chapter, we will denote
\[ \alpha := \left( \frac{\partial x_0}{\partial x} \right)^2 \quad \beta := \frac{\partial^2 x_0}{\partial x^2}. \] (3.2.6)

However, \( \alpha \) and \( \beta \) will be functions of \( x \) and \( t_0 \) in general. Since we are switching to the \( x_0 \) variable inside the integral, let us substitute \( x = x(x_0, t_0) \) into the expression for \( \frac{\partial^2 x_0}{\partial x} \) and \( \left( \frac{\partial x_0}{\partial x} \right) \), as in the previous example, to obtain \( \alpha \) and \( \beta \) as functions of \( x_0 \), and \( t_0 \):
\[
\alpha(x_0, t_0) := \left( \frac{\partial x_0}{\partial x} \right)^2 \quad \beta(x_0, t_0) := \frac{\partial^2 x_0}{\partial x^2}.
\]

Going back to our integral (3.2.4),
\[
0 = \frac{1}{2\pi} \int_0^{2\pi} \int_{\mathbb{R}} W(x_0, t_1) \left[ \alpha \partial^2_{x_0} F(x_0, t_1) + \beta \partial_{x_0} F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] dx_0 dt_0.
\]

Next, we perform the integration in the \( t_0 \) variable:
\[
0 = \int_{\mathbb{R}} W(x_0, t_1) \left[ \langle \alpha \rangle \partial^2_{x_0} F(x_0, t_1) + \langle \beta \rangle \partial_{x_0} F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] dx_0.
\]

For this integral to vanish regardless of choice of \( W \), then we need the expression in the brackets to vanish:
\[
\langle \alpha \rangle \partial^2_{x_0} F(x_0, t_1) + \langle \beta \rangle \partial_{x_0} F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) = 0,
\]
which gives us our evolution equation of \( F \) with respect to \( t_1 \):
\[
\partial_{t_1} F(x_0, t_1) = \langle \alpha \rangle \partial^2_{x_0} F(x_0, t_1) + \langle \beta \rangle \partial_{x_0} F(x_0, t_1).
\]

Thus, we can summarize these results in the following theorem.
**Theorem 2.** For \((x, t) \in \mathbb{R} \times [0, \infty)\), let \(u(x, t)\) be the solution of the advection-diffusion equation

\[
\frac{\partial u}{\partial t}(x, t) + a(x, t) \frac{\partial u}{\partial x}(x, t) = \epsilon \frac{\partial^2 u}{\partial x^2}(x, t)
\]  

(3.2.7)

\[u(x, 0) = f(x),\]

where \(a(x, t)\) is periodic in time with period \(2\pi\). Then the zeroth order term of the multi-scale asymptotic expansion of \(u\) will be given by

\[u_o(x, t) = F\left(x_o(x, t), \epsilon t\right),\]  

(3.2.8)

where \(x_o(x, t)\) is the characteristic map to the advection equation \(u_t + a(x, t)u_x = 0\), and the evolution of the envelope \(F(\cdot, t_1)\) is given by the PDE

\[
\partial_t F(x_o, t_1) = \langle \alpha \rangle \frac{\partial^2}{\partial x_o^2} F(x_o, t_1) + \langle \beta \rangle \partial_{x_o} F(x_o, t_1),
\]  

(3.2.9)

\[F(x, 0) = f(x),\]

where

\[
\langle \alpha \rangle = \frac{1}{2\pi} \int_0^{2\pi} \left( \frac{\partial x_o}{\partial x} \right)^2 dt, \quad \langle \beta \rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial^2 x_o}{\partial x^2} dt.
\]  

(3.2.10)

but written in terms of \(x_o\).

**Proof.** The entirety of section 3.2 is the proof of this theorem. \(\Box\)
3.3 Non-periodic Advection

In the previous examples, we had advection operators that were periodic in time, with period $2\pi$. This motivated an inner product

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \int_\mathbb{R} f(x,t) g(x,t) \, dx \, dt.$$  

This choice of this inner product then made computing $L^\dagger$ significantly easier, as the temporal boundary terms vanished. But now suppose that our advection is not periodic, or that even if it is, suppose we wish to average over some other interval. Let’s choose an arbitrary advection $a(x,t)$ and an arbitrary interval $[\tau_1, \tau_2]$ over which to define our inner product:

$$\langle f, g \rangle = \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} f(x,t) g(x,t) \, dx \, dt,$$

where $T := \tau_2 - \tau_1$. How will this affect our formulas (3.3.1) and (3.3.2)? Let us compute the adjoint of $L$:

$$\langle f, Lg \rangle = \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} f(\partial_{t_0} - a(x,t_0) \partial_x) g(x,t_0) \, dx \, dt_0$$

$$= \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} f(\partial_{t_0}g) + fa(x,t_0) [\partial_x g] \, dx \, dt_0$$

$$= \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} f g \bigg|_{\tau_1}^{\tau_2} \partial_{t_0} g \, dx \, dt_0 - \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} \partial_x \left([fa(x,t_0)] g\right) \, dx \, dt_0$$

$$= \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \left[\left(\delta(t_0 - \tau_2) - \delta(t_0 - \tau_1)\right)fg - \partial_{t_0} f g\right] \, dx \, dt_0$$

$$- \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} \left(\partial_x f a(x,t_0) + fa(x,t_0) \right) g \, dx \, dt_0$$

57
Let us then define \( \tilde{\delta} := \delta(t_o - \tau_2) - \delta(t_o - \tau_1) \). Carrying on,

\[
\frac{1}{T} \int_{R}^{\tau_2} \left( [\tilde{\delta}(t_o) f g - [\partial_t f] g] \right) dt_o dx
- \frac{1}{T} \int_{\tau_1}^{\tau_2} \left( a(x, t_o) [\partial_x f] g + a_x(x, t_o) f g \right) dx dt_o
\]

\[
= \frac{1}{T} \int_{R}^{\tau_2} \left( \tilde{\delta}(t_o) f g - \partial_t o - a(x, t_o) \partial_x - \partial_x a(x, t_o) \right) f g dx dt_o
\]

\[
= \langle L^+ f, g \rangle.
\]

Thus, our adjoint is

\[
L^+ = -L - a_x(x, t_o) + \tilde{\delta}(t_o).
\]

This is the same result we had in the previous section, except now with the additional \( \tilde{\delta} \) term. This will have a minimal effect on the kernel. Suppose \( \varphi \) is an element of the kernel of \( L^+ \) from the previous example, that is

\[
L \varphi = -a_x(x, t_o) \varphi.
\]

Then the kernel of \( L^+ \) will now be simply

\[
\ker(L^+) = \left\{ \varphi e^{\Lambda(t_o)} \mid L \varphi = -a \varphi \right\},
\]
where $\Delta(t)$ is a function such that $\Delta' = \bar{\delta}$. (Formally at least, the antiderivative of the delta function is thought of as a step function. We will look at this momentarily). To verify this claim, simply apply $L^+$ to $\varphi \Delta(t_o)$:

$$L^+(\varphi \Delta) = \left(-L - a \varphi + \overline{\delta}\right) \varphi \Delta$$

$$= \left(-\partial_{t_o} - a \partial_x - a \varphi + \overline{\delta}\right) \varphi \Delta$$

$$= \left[-\partial_{t_o} \varphi \Delta\right] + \left[-a \partial_x \varphi \Delta\right] + \overline{\delta} \varphi \Delta$$

$$= -[\partial_{t_o} \varphi] e^{\Delta} - \varphi e^{\Delta} \bar{\delta} + \left[-a \partial_x \varphi \right] e^{\Delta} + \overline{\delta} \varphi e^{\Delta}$$

$$= -[\partial_{t_o} \varphi] e^{\Delta} + \left[-a \partial_x \varphi \right] e^{\Delta}$$

$$= \left[-\left(L - a \varphi\right) \varphi \right] e^{\Delta}$$

$$= 0$$

Thus, our FAI will take the form

$$0 = \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} W(x_o(x,t_o),t_1)e^{\tilde{h}(x,t_o,t_1)}e^{\Delta(t_o)} \left[\partial^2_x u_0 - \partial_t u_0\right] dx dt_0.$$

However, what does the function $\Delta$ look like on the interval $[\tau_1, \tau_2]$? Formally, the anti-derivative of the delta function is the step function:

$$\int_{-\infty}^{t} \delta(t) dt = \begin{cases} 
0 & \text{if } t < 0 \\
1 & \text{if } t \geq 0
\end{cases}$$

Therefore, if our $\Delta$ is supposed to be the anti-derivative of $\delta(t - \tau_2) - \delta(t - \tau_1)$, then we want

$$\Delta(t) := \begin{cases} 
0 & \text{if } t < \tau_2 \\
1 & \text{if } t \geq \tau_2
\end{cases} - \begin{cases} 
0 & \text{if } t < \tau_1 \\
1 & \text{if } t \geq \tau_1
\end{cases} = \begin{cases} 
0 & t < \tau_1 \\
-1 & \tau_1 \leq t < \tau_2 \\
0 & t \geq \tau_2
\end{cases}$$

59
Thus, on our interval of interest, \([\tau_1, \tau_2], \Delta(t_0)\) is simply equal to \(-1\), and so we can pull the factor of \(e^{-1}\) outside of the integral,

\[
0 = \frac{1}{e} \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} W(x_0(x, t_0), t_1) e^{\beta(x, t_0, t_1)} \left[ \frac{\partial^2}{\partial x^2} u_0 - \partial_{t_1} u_0 \right] dx dt.
\]

Hence, there is no effect on the integral. What do we conclude from this? Changing the interval that we average over does not change our formulas, even if our advection is not periodic anymore. However, this tells us nothing about whether \(u_0\) will still be a good approximation. We can now slightly generalize theorem 2:

**Theorem 3.** Let \(u\) be the solution of the advection-diffusion equation \(u_t + a(x, t)u_x = \epsilon u_{xx}\) defined on \(\mathbb{R} \times [0, \infty)\) with \(u(x, 0) = f(x)\), and \(a(x, t)\) is arbitrary. Then the zeroth order term of the multi-scale asymptotic expansion of \(u\) will be given by

\[
u_0(x, t) = F(x_0(x, t), \epsilon t),
\]

where \(x_0(x, t)\) is the characteristic map of the advection operator and the evolution of the envelope \(F(\cdot, t_1)\) is given by the PDE

\[
\partial_{t_1} F(x_0, t_1) = \langle \alpha \rangle \partial^2_{x_0} F(x_0, t_1) + \langle \beta \rangle \partial_{x_0} F(x_0, t_1),
\]

\[(3.3.1)\]

\[
F(x, 0) = f(x),
\]

where

\[
\langle \alpha \rangle = \frac{1}{2\pi} \int_0^{2\pi} \left( \frac{\partial x_0}{\partial x} \right)^2 dt, \quad \langle \beta \rangle = \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial^2 x_0}{\partial x^2} dt.
\]

\[(3.3.2)\]

**Proof.** Once again, sections 3.2 and 3.3 are the proof of this theorem. \(\square\)
For the final example in this section, consider the equation

\[
\frac{\partial u}{\partial t}(x,t) + a(t) \sin(x) \frac{\partial u}{\partial x}(x,t) = \epsilon \frac{\partial^2 u}{\partial x^2}(x,t) \tag{3.4.1}
\]

\[ u(x,0) = f(x). \]

Here, \( a(t) \) could be any function, not necessarily periodic. Now that we have our formulas (3.2.8)-(3.3.2), we can skip most of the busy work. We need only find the characteristic map, and then compute \( \alpha \) and \( \beta \) from it. To find the characteristic map, we need to solve

\[ \dot{x}(t) = a(t) \sin(x). \]

Separating variables gives us

\[ -\log \left| \frac{1 + \cos(x)}{\sin(x)} \right| = A(t) + C, \]

where \( A(t) := \int_0^t a(\tau) d\tau \). We would like to solve this equation for \( x \). Denote

\[ w := \cos(x), \quad z := e^{-A(t) - C}. \]

Then we have

\[ \left| \frac{1 + w}{\sqrt{1 - w^2}} \right| = z. \]

We can drop the absolute values because both the numerator and denominator are always non-negative:

\[ \frac{1 + w}{\sqrt{1 - w^2}} = z. \]
Squaring both sides and using the quadratic formula, we obtain

\[ w = \frac{-1 + z^2}{1 + z^2}. \]

Putting this back in terms of our original variables, this becomes

\[ \cos(x) = \frac{-1 + e^{-2A(t)+2C}}{1 + e^{-2A(t)+2C}} = \tanh \left[ -A(t) - C \right]. \]

Hence,

\[ x = \cos^{-1} \left( \tanh \left[ -A(t) - C \right] \right). \]

Using the ambiguity in \( C \), we can demand that \( t = 0 \) produces the value \( x_0 \). This gives us

\[ x = \cos^{-1} \left( \tanh \left[ \tanh^{-1}[\cos(x_o)] - A(t) \right] \right). \]

Inverting this is straightforward:

\[ x_o = \cos^{-1} \left( \tanh \left[ \tanh^{-1}[\cos(x)] + A(t) \right] \right). \]

Hence, our characteristic map will be

\[ x_0(x, t_o) = \cos^{-1} \left( \tanh \left[ \tanh^{-1}[\cos(x)] + A(t_o) \right] \right). \]

We would like to now compute \( \alpha \) and \( \beta \), which involves computing the partial derivatives \( \frac{\partial x_o}{\partial x} \) and \( \frac{\partial^2 x_o}{\partial x^2} \) and then, as we have seen, substitute back in \( x(x_o, t_o) \) to get \( \alpha, \beta \) entirely in terms of \( x_o \) and \( t_o \). When the dust settles, we are left with

\[ a(x_o, t_o) = \sin(x_o)^2 \cosh^2 \left( \log \left| \frac{1 + \cos(x_o)}{\sin(x_o)} \right| - A(t_o) \right). \]  

(3.4.2)
and
\[
\beta(x_o, t_o) = \sin(x_o) \cosh^2 \left( \log \left| \frac{1 + \cos(x_o)}{\sin(x_o)} \right| - A(t_o) \right) \\
\times \left[ \cos(x_o) - \tanh \left( \log \left| \frac{1 + \cos(x_o)}{\sin(x_o)} \right| + A(t_o) \right) \right].
\] (3.4.3)

Finally, we would choose a function \(a(t)\), and then average \(\alpha\) and \(\beta\) over the specified time interval. These functions don’t look fun to integrate with respect to \(t\). Fortunately, when it comes to the numerics, we are actually going to bypass these formulas all together, as described in chapter 5.

Also note that, although it looks like these expression might have a singularity at \(x = n\pi\), they actually don’t. For a rough idea of why this is, consider the fact that near \(n\pi\), the sine function goes to zero almost linearly: \(\sin(x) \sim O(x)\). Thus, near \(n\pi\), \(\alpha\) will look like
\[
\alpha \sim x^2 \cosh^2 \left[ \log(1/x) \right] \sim x^2(1/x^2) \sim 1.
\]

A similar, but more complicated computation gives a similar result for \(\beta\).
CHAPTER 4

Generalizing to Higher Dimensions.

4.1 THE ADVECTION-DIFFUSION EQUATION IN ARBITRARY DIMENSIONS

The advection-diffusion equation over 2, 3 or an arbitrary number of dimensions is given by

\[ u_t(x,t) + \vec{a}(x,t) \cdot \nabla u(x,t) = \epsilon \Delta u(x,t), \]

with

\[ u(x,0) = f(x), \]

where now \( x = (x^1, ..., x^n) \in \mathbb{R}^n \). As before, we set up a multi-scale asymptotic expansion:

\[ (\partial_{\epsilon o} + \epsilon \partial_{t_1})(u_o + \epsilon u_1 + ...) + \vec{a} \cdot \nabla (u_o + \epsilon u_1 + ...) = \epsilon \Delta (u_o + \epsilon u_1 + ...). \]

Collecting the \( O(\epsilon^0) \) terms, we have

\[ \partial_{t_0} u_o + \vec{a} \cdot \nabla u_o = 0. \]  

(4.1.1)

As we discussed in Chapter 1, we must solve a system of characteristic equations:

\[ \dot{x}^i = a^i(x,t). \]
The solution to equation (4.1.1) will then be

\[ u_0(x, t_0, t_1) = F(x_0(x, t_0), t_1). \]

Collecting the terms of order \( O(\epsilon) \) yields

\[ (\partial_{t_0} + \vec{a} \cdot \nabla) u_1 = \Delta u_0 - \partial_{t_1} u_0. \]

As before, we will label

\[ L := \partial_{t_0} + \vec{a} \cdot \nabla \quad \quad \quad v := \Delta u_0 - \partial_{t_1} u_0. \]

Again, in our aim to set up the Fredholm alternative integral, we need to compute \( L^\dagger \).

Integrating by parts reveals

\[ L^\dagger = -\partial_{t_0} - \vec{a} \cdot \nabla - \nabla \cdot \vec{a} = -L - \nabla \cdot \vec{a}. \]

Now we seek the kernel of \( L^\dagger \), that is, functions \( \varphi \) that satisfy

\[ L \varphi = - (\nabla \cdot \vec{a}) \varphi. \quad (4.1.2) \]

Analogous to before, let us decompose such a function as a product of two functions, one which depends on \( x_0 \) and another which doesn’t:

\[ \varphi = W(x_0(x, t_0), t_1)e^{p(x, t_0, t_1)}. \]
Substituting this back into (4.1.2) gives us

\[
\sum_k \frac{\partial W}{\partial x^k_0} e^{p(x,t_0,t_1)} + W e^{p(x,t_0,t_1)} p_{t_0} + \sum_k \frac{\partial W}{\partial x^k_0} (\vec{a} \cdot \nabla x^k_0) e^{p(x,t_0,t_1)} + W e^{p(x,t_0,t_1)} (\vec{a} \cdot \nabla p \vec{x})
\]

\[
= - (\nabla \cdot \vec{a}) W e^{p(x,t_0,t_1)}.
\]

Dividing through by \( e^{p(x,t_0,t_1)} \) and collecting terms by derivatives of \( W \) we have

\[
\sum_k \frac{\partial W}{\partial x^k_0} (\frac{\partial x^k_0}{\partial t_0} + (\vec{a} \cdot \nabla x^k_0)) + W \left( p_{t_0} + (\nabla \cdot \vec{a}) \right) = -W (\nabla \cdot \vec{a}).
\]

But each component of the characteristic map solves equation (4.1.1) (see equation (1.5.5)) and thus the first summation vanishes. After dividing through by \( W \), we are left with

\[
p_{t_0} + (\nabla \cdot \vec{a}) = -\nabla \cdot \vec{a}.
\]

As before, the homogeneous part of PDE is simply the original advection equation, so it will have a homogeneous solution \( p_h(x_0(x,t_0),t_1) \) as well as a particular solution \( \tilde{p}(x,t_0,t_1) \). We absorb \( p_h \) into \( W \), leaving us with

\[
\ker(L^\dagger) = \left\{ W(x_0(x,t_0),t_1) e^{p(x,t_0,t_1)} \right\}.
\]

We can now set up the FAI:

\[
0 = \frac{1}{T} \int_{T_1}^{T_2} \int_{\mathbb{R}} W(x_0(x,t_0),t_1) e^{p(x,t_0,t_1)} \left[ \Delta u_0 - \partial_{t_0} u_0 \right] d\vec{x}^n dt_0.
\]

In the one dimensional case, it was through the marvelous fact that \( \frac{\partial x_0}{\partial x} \in \ker(L^\dagger) \) that an additional exponential term entered the integral and cancelled with the current one. This happened, of course, when we switched from \( dx \) to \( dx_0 \). In the multi-dimensional
setting we find ourselves in now, it is natural to ask, is the Jacobian \( \left| \frac{\partial x_o}{\partial x} \right| \), which arrives when switching coordinates, an element of the kernel of \( L^+ \)? The answer, is yes.

**Theorem 4.** Let \( x_o(x, t) \) be the characteristic map of the differential equation \( Lu = 0 \), where \( L \) is of the form \( L = \partial_t + \vec{a} \cdot \nabla \). Then \( \left| \frac{\partial x_o}{\partial x} \right| \in \ker(L^+) \).

This is the serious version of the theorem, of which the theorem from Chapter 3 is just a particular case.

**Proof.** This will be complicated, so let us sketch the idea before we begin. First, we write down an explicit expression for the Jacobian, which we will label \( J \), in terms of the determinant of the matrix of partial derivatives. Then, we want to see if \( L^+ J = 0 \), in other words,

\[
J_t + \vec{a} \cdot \nabla J = - \left( \nabla \cdot \vec{a} \right) J.
\]

Therefore, we will compute the expressions \( J_t \) and \( \nabla J \) and see if all the pieces fit together properly.

The Jacobian is defined by

\[
J := \left| \frac{\partial x_o}{\partial x} \right| = \sum_{\sigma} \epsilon_{\sigma} \prod_k \frac{\partial x_o^k}{\partial x^{\sigma(k)}},
\]

where \( \sigma \) is any permutation of the set \( \{1, 2, ..., n\} \) and

\[
\epsilon_{\sigma} = \begin{cases} 
+1 & \text{if } \sigma \text{ is an even permutation} \\
-1 & \text{if } \sigma \text{ is an odd permutation}
\end{cases}.
\]

To show that this quantity is in the kernel of \( L^+ \), we need to show that

\[
J_t + \vec{a} \cdot \nabla J = - \left( \nabla \cdot \vec{a} \right) J, \quad (4.1.3)
\]
or rather, to make things smoother later on, let us show

\[ J_t = -\vec{a} \cdot \nabla J - (\nabla \cdot \vec{a}) J. \]  

(4.1.4)

Let us compute. First, the time derivative of \( J \):

\[ J_t = \sum_\sigma \epsilon_\sigma \sum_j \partial_t \left( \frac{\partial x_\sigma^j}{\partial x^{\sigma(j)}} \right) \prod_{k \neq j} \frac{\partial x_\sigma^k}{\partial x^{\sigma(k)}}, \]  

(4.1.5)

which follows simply from the product rule. Now, we can write the term

\[ \partial_t \left( \frac{\partial x_\sigma^j}{\partial x^{\sigma(j)}} \right) \]

as

\[ \partial_{x^{\sigma(j)}} \left( \frac{\partial x_\sigma^j}{\partial t} \right) \]

by switching the order of derivatives. And, since each \( x_\sigma^j \) satisfies the equation \( L x_\sigma^j = 0 \) (as was discussed in Chapter 1), then we can replace \( \partial_t x_\sigma^j \) with \( (\vec{a} \cdot \nabla x_\sigma^j) \):

\[ \partial_{x^{\sigma(j)}} \left( \frac{\partial x_\sigma^j}{\partial t} \right) = \partial_{x^{\sigma(j)}} \left( -\vec{a} \cdot \nabla x_\sigma^j \right) = -\sum_l a^l \partial_{x^l} \left( \frac{\partial x_\sigma^j}{\partial x^{\sigma(j)}} \right) - \sum_l (a^l)_{x^{\sigma(j)}} \left( \frac{\partial x_\sigma^j}{\partial x^l} \right). \]

Putting this back into (4.1.5) gives us

\[ J_t = -\sum_\sigma \epsilon_\sigma \sum_j \left( \sum_l a^l \partial_{x^l} \left( \frac{\partial x_\sigma^j}{\partial x^{\sigma(j)}} \right) + \sum_l (a^l)_{x^{\sigma(j)}} \left( \frac{\partial x_\sigma^j}{\partial x^l} \right) \right) \prod_{k \neq j} \frac{\partial x_\sigma^k}{\partial x^{\sigma(k)}}. \]  

(4.1.6)

Now let us compute the spatial derivatives of \( J \):

\[ J_{x^l} = \sum_\sigma \epsilon_\sigma \sum_j \partial_{x^l} \left( \frac{\partial x_\sigma^j}{\partial x^{\sigma(j)}} \right) \prod_{k \neq j} \frac{\partial x_\sigma^k}{\partial x^{\sigma(k)}}. \]
Thus, the right hand side of equation (4.1.4) is

\[-\vec{a} \cdot \nabla J - (\nabla \cdot \vec{a}) J = \sum_l \left[ -d^l \sum_{\sigma} \epsilon_{\sigma} \sum_j \partial x^l_j \left( \frac{\partial x^l_0}{\partial x^{\sigma(j)}} \right) \prod_{k \neq j} \frac{\partial x^k_0}{\partial x^{\sigma(k)}} \right.\]

\[-\sum_{\sigma} \epsilon_{\sigma} \prod_k \frac{\partial x^k_0}{\partial x^{\sigma(k)}} (d^l)_x^l \]\n
or, upon rearranging the order or the sums, we have

\[-\vec{a} \cdot \nabla J - (\nabla \cdot \vec{a}) J = -\sum_{\sigma} \epsilon_{\sigma} \sum_l (a^l)_x^l \prod_k \frac{\partial x^k_0}{\partial x^{\sigma(k)}} - \sum_{\sigma} \epsilon_{\sigma} \sum_j \sum_l (a^j)_x^j \left( \frac{\partial x^j_0}{\partial x^{\sigma(j)}} \right) \prod_{k \neq j} \frac{\partial x^k_0}{\partial x^{\sigma(k)}}.\]

(4.1.7)

In order to prove the theorem, we must show that equations (4.1.6) and (4.1.7) are equal to each other. Notice that the first term in (4.1.6) is equal to the second term in (4.1.7).

Thus, what is left to show is

\[\sum_{\sigma} \epsilon_{\sigma} \sum_j \sum_l (a^l)_x^j \frac{\partial x^k_0}{\partial x^{\sigma(k)}} \prod_k \frac{\partial x^k_0}{\partial x^{\sigma(k)}} \]

\[= \sum_{\sigma} \epsilon_{\sigma} \sum_l (a^l)_x^l \prod_k \frac{\partial x^k_0}{\partial x^{\sigma(k)}}.\]

(4.1.8)

In the triple sum on the left hand side of the above equation

\[\sum_{\sigma} \epsilon_{\sigma} \sum_l \sum_j ,\]

we can split up the summation over \(j\) into two terms, that for which \(\sigma(j) = l\) and that for which \(\sigma(j) \neq l\). Explicitly, the left side of (4.1.8) can be written as

\[\sum_{\sigma} \epsilon_{\sigma} \sum_l (a^l)_x^l \prod_k \frac{\partial x^k_0}{\partial x^{\sigma(k)}} + \sum_{\sigma} \epsilon_{\sigma} \sum_l \sum_j (a^l)_x^j \frac{\partial x^j_0}{\partial x^{\sigma(j)}} \prod_{k \neq j} \frac{\partial x^k_0}{\partial x^{\sigma(k)}}.\]

(4.1.9)
Notice that the first term above is already equal to the right hand side of \((4.1.8)\). Thus, to show the equality of \((4.1.8)\) we need only show that the second term in \((4.1.9)\) vanishes:

\[
\sum_{\sigma} \varepsilon_{\sigma} \sum_{l, \sigma(j) \neq l} (a^l)_{x_{\sigma(j)}} \frac{\partial x_j^l}{\partial x_l} \prod_{k \neq j} \frac{\partial x_k^l}{\partial x_{\sigma(k)}} = 0. \tag{4.1.10}
\]

If we can show this holds, then the theorem is proved. Notice, however, that since \(\varepsilon_{\sigma}\) is +1 for even permutations and -1 for odd permutations, we could equivalently try to show that if we split this sum into a sum over even permutations and a sum over odd permutations, they cancel each other out:

\[
\sum_{\sigma \text{ even}} \varepsilon_{\sigma} \sum_{l, \sigma(j) \neq l} (a^l)_{x_{\sigma(j)}} \frac{\partial x_j^l}{\partial x_l} \prod_{k \neq j} \frac{\partial x_k^l}{\partial x_{\sigma(k)}} = - \sum_{\sigma \text{ odd}} \varepsilon_{\sigma} \sum_{l, \sigma(j) \neq l} (a^l)_{x_{\sigma(j)}} \frac{\partial x_j^l}{\partial x_l} \prod_{k \neq j} \frac{\partial x_k^l}{\partial x_{\sigma(k)}}. \tag{4.1.11}
\]

This can be shown by the following argument. Let \(\sigma\) be any even permutation of \(\{1, 2, ..., n\}\), and let \(l, j\) be such that \(\sigma(j) \neq l\). Define the quantity

\[
E_{l,j}(\sigma) := (a^l)_{x_{\sigma(j)}} \frac{\partial x_j^l}{\partial x_l} \prod_{k \neq j} \frac{\partial x_k^l}{\partial x_{\sigma(k)}} \\
= (a^l)_{x_{\sigma(j)}} \frac{\partial x_1^l}{\partial x_{\sigma(1)}} \frac{\partial x_2^l}{\partial x_{\sigma(2)}} \cdots \frac{\partial x_j^l}{\partial x_l} \cdots \frac{\partial x_n^l}{\partial x_{\sigma(n)}} \\
= (a^l)_{x_{\sigma(j)}} \frac{\partial x_1^l}{\partial x_{\sigma(1)}} \frac{\partial x_2^l}{\partial x_{\sigma(2)}} \cdots \frac{\partial x_j^l}{\partial x_l} \cdots \frac{\partial x_i^l}{\partial x_{\sigma(i)}} \cdots \frac{\partial x_n^l}{\partial x_{\sigma(n)}} \\
\text{(since } \sigma(j) \neq l \rightarrow \exists i \text{ s.t. } \sigma(i) = l, \text{ with } i \neq j).}
\]
Now consider an odd permutation $\tilde{\sigma}$ obtained by interchanging $\sigma(i) \leftrightarrow \sigma(j)$. Then the quantity $E_{l,j}(\tilde{\sigma})$ will be

$$E_{l,j}(\tilde{\sigma}) := (a^l)_{x^{(i)}} \frac{\partial x^j}{\partial x^l} \prod_{k \neq l} \frac{\partial x^k}{\partial x^{\sigma(k)}}$$

$$= (a^l)_{x^{\sigma(i)}} \frac{\partial x^1}{\partial x^{\sigma(1)}} \frac{\partial x^2}{\partial x^{\sigma(2)}} \cdots \frac{\partial x^i}{\partial x^l} \cdots \frac{\partial x^n}{\partial x^{\sigma(n)}}$$

$$(\text{now, by construction, } \tilde{\sigma}(j) = l)$$

$$= (a^l)_{x^{\sigma(i)}} \frac{\partial x^1}{\partial x^{\sigma(1)}} \frac{\partial x^2}{\partial x^{\sigma(2)}} \cdots \frac{\partial x^i}{\partial x^l} \cdots \frac{\partial x^n}{\partial x^{\sigma(n)}}$$

Thus, for each even $\sigma$, there exists an odd $\tilde{\sigma}$ for which $E_{l,j}(\tilde{\sigma}) = E_{l,j}(\sigma)$ and so the sums

$$\sum_{\sigma \text{ even}} \sum_l \sum_{[j, \sigma(j) \neq l]} E_{l,j}(\sigma) = \sum_{\sigma \text{ odd}} \sum_l \sum_{[j, \sigma(j) \neq l]} E_{l,j}(\sigma)$$

will be equal. If we throw in the $\epsilon_\sigma$'s, then we have

$$\sum_{\sigma \text{ even}} \epsilon_\sigma \sum_l \sum_{[j, \sigma(j) \neq l]} E_{l,j}(\sigma) = -\sum_{\sigma \text{ odd}} \epsilon_\sigma \sum_l \sum_{[j, \sigma(j) \neq l]} E_{l,j}(\sigma) \quad (4.1.12)$$

But this equation (4.1.12) is precisely identical to the equation (4.1.11). Thus, we have shown that equation (4.1.10) is valid, which means we are done. (because then equation (4.1.9) is valid, which means (4.1.8) is valid, which in turn shows that (4.1.6) and (4.1.7) are equal, which is the same as saying $J \in \ker(L^\dagger)$.)

It is worth pointing out here, that while it was hoped that this was an original theorem, it can be seen that this theorem actually gives us an evolution equation for the so-called material derivative of the Jacobian, and is therefore actually a manifestation of the generalized Liouville theorem [19]. While the classical Liouville theorem tells us
that volume in phase space is preserved under Hamiltonian systems, the generalized Liouville theorem tells us how the volume (and hence the Jacobian) evolves under general advections [9].

Armed with this fact, we now know that \( \left| \frac{\partial x_0}{\partial x} \right| \), as a member of \( \text{ker}(L^+) \), must be of the form

\[
\left| \frac{\partial x_0}{\partial x} \right| = V(x_0(x, t_0), t_1) e^{\beta(x, t_0, t_1)}.
\]

And, since \( dx^n = \left| \frac{\partial x_0}{\partial x} \right|^{-1} dx_0^n \), then the FAI becomes

\[
0 = \frac{1}{T} \int_{T_1}^{T_2} \int_{\mathbb{R}} W(x_0, t_1) e^{\beta(x, t_0, t_1)} \left[ \Delta F(x_0, t_1) - \partial_t F(x_0, t_1) \right] \frac{dx_0^n}{W(x_0, t_1) e^{\beta(x, t_0, t_1)}} dt_0,
\]

or, after absorbing \( V \) into \( W \) and cancelling the exponential terms,

\[
0 = \frac{1}{T} \int_{T_1}^{T_2} \int_{\mathbb{R}} V(x_0, t_1) \left[ \Delta F(x_0, t_1) - \partial_t F(x_0, t_1) \right] dx_0^n dt_0. \quad (4.1.13)
\]

We are almost ready to switch entirely to the \( x_0 \) variable, but unfortunately, the derivatives in the Laplacian, \( \Delta \), are still with respect to the \( x \) variables. We now do a computation analogous to (3.2.5), only this time we start with

\[
\partial x^n = \sum_k \frac{\partial x^k}{\partial x^n} \partial x^k.
\]
Next, we compute the second derivative:

\[
\partial_{x^nx^n} = \left[ \sum_k \frac{\partial x^k}{\partial x^n} \partial x^k_n \right] \left[ \sum_k \frac{\partial x^k}{\partial x^n} \partial x^k_n \right] \]

\[
= \sum_{k,j} \left[ \frac{\partial x^k}{\partial x^n} \frac{\partial x^j}{\partial x^n} \partial x^k_n \partial x^j_n + \frac{\partial x^k}{\partial x^n} \left( \frac{\partial x^j}{\partial x^n} \frac{\partial x^k}{\partial x^n} \right) \partial x^j_n \right] \]

\[
= \sum_{k,j} \frac{\partial x^k}{\partial x^n} \frac{\partial x^j}{\partial x^n} \partial x^k_n \partial x^j_n + \sum_{k,j} \frac{\partial x^k}{\partial x^n} \left( \sum_m \frac{\partial x^m}{\partial x^k} \frac{\partial x^m}{\partial x^n} \right) \partial x^j_n \]

\[
= \sum_{k,j} \frac{\partial x^k}{\partial x^n} \frac{\partial x^j}{\partial x^n} \partial x^k_n \partial x^j_n + \sum_j \left[ \sum_m \left( \sum_k \frac{\partial x^k}{\partial x^n} \frac{\partial x^m}{\partial x^n} \right) \frac{\partial^2 x^j_n}{(\partial x^n)^2} \right] \partial x^j_n \]

\[
= \sum_{k,j} \frac{\partial x^k}{\partial x^n} \frac{\partial x^j}{\partial x^n} \partial x^k_n \partial x^j_n + \sum_j \left[ \frac{\partial^2 x^j_n}{(\partial x^n)^2} \right] \partial x^j_n \quad \text{(because } \sum_k \frac{\partial x^k}{\partial x^n} \frac{\partial x^m}{\partial x^n} = \delta_{nm} \text{)}
\]

Now, taking this expression for \( \partial_{x^nx^n} \) we can compute the Laplacian:

\[
\Delta x = \sum_n \partial x^n x^n 
\]

\[
= \sum_n \left[ \sum_{k,j} \frac{\partial x^k}{\partial x^n} \frac{\partial x^j}{\partial x^n} \partial x^k_n \partial x^j_n \right] + \sum_n \left[ \sum_j \frac{\partial^2 x^j_n}{(\partial x^n)^2} \partial x^j_n \right] \]

\[
= \sum_{k,j} \left[ \sum_n \frac{\partial x^k}{\partial x^n} \frac{\partial x^j}{\partial x^n} \partial x^k_n \partial x^j_n \right] + \sum_j \left[ \sum_n \frac{\partial^2 x^j_n}{(\partial x^n)^2} \right] \partial x^j_n \]

\[
= \sum_{k,j} (\nabla x^k \cdot \nabla x^j) \partial x^k_n \partial x^j_n + \sum_j \Delta x_j \partial x^j_n
\]
Compare this formula for (3.2.5). Now, let us give names to the coefficients of the differential operators above, analogous to what was done in the last chapter. Define

\[ \alpha_{k,j} := \nabla x_0^k \cdot \nabla x_0^j \quad \beta_j := \Delta x_0^j. \]

Although these would be functions of \((x, t_0)\), we can use the characteristic map to write them in terms of \((x_0, t_0)\). Then, putting these back into the FAI (4.1.13), we have

\[ 0 = \frac{1}{T} \int_{t_1}^{t_2} \int_{\mathbb{R}} V(x_0, t_1) \left[ \sum_{k,j} \alpha_{k,j} \partial_{x_0^k} F(x_0, t_1) + \sum_j \beta_j \partial_{x_0^j} F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] dx_0^n dt_0. \]

Now, after we perform the \(t_0\) integral, we will have

\[ 0 = \int_{\mathbb{R}} V(x_0, t_1) \left[ \sum_{k,j} \langle \alpha_{k,j} \rangle \partial_{x_0^k} F(x_0, t_1) + \sum_j \langle \beta_j \rangle \partial_{x_0^j} F(x_0, t_1) - \partial_{t_1} F(x_0, t_1) \right] dx_0^n. \]

But as always, if this integral is to vanish, for arbitrary choices of \(W\), then we must have the expression inside the brackets vanish:

\[ \partial_{t_1} F(x_0, t_1) = \sum_{k,j} \langle \alpha_{k,j} \rangle \partial_{x_0^k} F(x_0, t_1) + \sum_j \langle \beta_j \rangle \partial_{x_0^j} F(x_0, t_1). \]

Thus, we finally have our evolution equation for the profile of \(F\).

To sum up, then, we have the following theorem.

**Theorem 5.** For \((\bar{x}, t) \in \mathbb{R}^n \times [0, \infty)\), let \(u\) be the solution to the advection diffusion equation

\[ u_t(x, t) + \bar{a}(x, t) \cdot \nabla u(x, t) = \varepsilon \Delta u(x, t) \]

\[ u(x, 0) = f(x), \]
where \( \vec{a}(x, t) \) is arbitrary. Then the zeroth order term in the multi-scale asymptotic expansion for \( u \) is given by

\[
u_o(x, t_0, t_1) = F\left(x_o(x, t_0), t_1\right),
\]

where \( F(\cdot, 0) = f(\cdot) \), and \( x_o(x, t) \) is the characteristic map of the advection operator. Furthermore, the evolution of \( F(\cdot, t_1) \) with respect to \( t_1 \) is governed by the equation

\[
\partial_{t_1} F(x_o, t_1) = \sum_{k,j} \langle \alpha_{k,j} \rangle \partial_{x_o^{k} x_o^{j}} F(x_o, t_1) + \sum_{j} \langle \beta_{j} \rangle \partial_{x_o^{j}} F(x_o, t_1).
\]

where

\[
\langle \alpha_{k,j} \rangle := \frac{1}{T} \int_{t_1}^{t_2} \nabla x_o^{k} \cdot \nabla x_o^{j} dt_0 \quad \langle \beta_{j} \rangle := \frac{1}{T} \int_{t_1}^{t_2} \Delta x_o^{j} dt_0.
\]

Proof. Section 4.1 is the proof of the theorem.

4.2 DISCUSSION

It should be mentioned at this point that somewhat similar formulas were discovered by Krol and Steenstra over two decades ago [16], [27]. However, their formulas were much more limited in scope. Firstly, they are derived from a different (somewhat ad hoc) approach. Rather than setting up a multiple scale hypothesis and enforcing evolution conditions via the Fredholm Alternative, they use a more direct route, in which slow scale time dependence is manually inserted into the characteristic map, after which a time derivative is taken and a coordinate transformation (to the characteristic map) is performed. From this they arrive at an equation which is in good form to be averaged. The disadvantage of this approach is that it does not lend itself to as much generality as we have with our current formulas. Firstly, the Fredholm Alternative formalism provides us with a unifying framework for deriving and explaining several different questions. We see how the Jacobian conspires with us to facilitate a change of variables and from this
we can easily write down (explicit) formulas for the effective equations. Moreover, we can see directly how time-periodicity does not affect these formulas. Furthermore, there are no conditions imposed on $a(x,t)$ whatsoever, leaving it completely general. Finally, it isn’t clear how one could generalize to non-linear advection from Krol’s formulas, which were derived in the linear case. Comparatively, while it will be considerably more complicated, we will see in the next chapter how the method of multiple scales and the Fredholm Alternative integral allow us to generalize to a setting of non-linear advection. It should also be noted that in the one case where Krol does compute explicit formulas for the effective equation, they are, as far as this author can tell, incorrectly computed.

Perhaps the most crucial result to come from that paper, however, is a theorem on the accuracy of averaging a periodic advection-diffusion operator. Specifically, he proves that the following.

**Theorem 6.** Suppose $u$ satisfies $u_t = \epsilon L(t)u$, where $L$ is a time-periodic elliptic differential operator. Let $\overline{L}$ be the averaged operator and consider $\overline{u}$, the solution to $\overline{u}_t = \epsilon \overline{L}\overline{u}$. Then $\|u(t) - \overline{u}(t)\| = O(\epsilon)$ for $t \in [0, k/\epsilon]$, where $k$ depends on the operator $L$.

**Proof.** See [16] for the original proof, or [23] for a modern version. \qed

What this result essentially tells us is that on the slower time scales (the $1/\epsilon$ scale), the difference between the true solution and the averaged equation will be on the order of $\epsilon$. Since we are in the regime were $\epsilon$ is assumed to be small, then this theorem tells us that these methods of averaging can be quite accurate. Moreover, while his theorem requires $L$ to be periodic in time, it should be mentioned that numerical simulations for systems that are not periodic in time more or less conform to this theorem as well.

One final point to discuss is that of boundary conditions. It should be said that it appears that our current methods would be rather ineffective were we to impose boundary conditions (as opposed to working on an unbounded domain). The reason for this is simple. Characteristic maps are, in general, not compatible with boundary conditions.
There are specific criteria for when boundary conditions can be imposed on a first order advection equation, but roughly speaking, the idea is the following. Once we impose our initial condition, we will have exhausted all or our degrees of freedom and there will not be a characteristic map compatible with any additional conditions. In fact, were we to impose boundary conditions on our current advection-diffusion problem, it would be the second order diffusion operator which provides this extra ‘freedom’. Specifically, it allows for boundary layers to build up (a kind of compromise between the advection and the boundaries). The problem, then, would be that we could not write our $u_0$ as the elegant decomposition

$$u_0(x, t) = F(x_0(x, t), \epsilon t)$$

because the characteristic map needs to be in constant communication with the diffusion. In fact, the method of multiple scales is often used to deal with not only temporal scales, but spatial scales, and in particular, to describe boundary layers! Some work was done by this author early on towards this effort, but was abandoned to pursue more promising leads. Perhaps, this would be an area to pursue in the near future.
CHAPTER 5

Multi-Scale Analysis for a Non-linear (Burgers) Equation

5.1 A BURGERS’ STYLE ADVECTION-DIFFUSION EQUATION

Now we go back to the 1-D case \((x \in \mathbb{R})\) and add a little bit of non-linearity. Consider the Burgers-style PDE

\[
\frac{\partial u}{\partial t}(x,t) + a(x,t)u(x,t)\frac{\partial u}{\partial x}(x,t) = \epsilon \frac{\partial^2 u}{\partial x^2}
\]

\[
u(x,0) = f(x).
\]

One more time, we go through the motions as usual, but this time some extra inconveniences will arise. We make a multi-scale asymptotic expansion for \(u\) and substitute it into the PDE:

\[
(\partial_{t_0} + \epsilon \partial_{t_1})(u_0 + \epsilon u_1 + ...) + a(x,t_0)(u_0 + \epsilon u_1 + ...)\partial_x(u_0 + \epsilon u_1 + ...) = 
\]

\[
\epsilon \partial^2_x(u_0 + \epsilon u_1 + ...).
\]

The \(O(\epsilon^0)\) terms are

\[
\partial_{t_0}u_0 + a(x,t_0)u_0\partial_xu_0 = 0. \quad (5.1.1)
\]
To solve this equation, we begin by finding the characteristic map $x_0(x, t_0)$. If we blindly proceed in the usually way, we would assert

$$u_0 = F(x_0(x, t_0), t_1),$$

with $F(x,0) = f(x)$. However, this does not work in this case. The problem here is that the characteristic map of the Burgers equation depends on the initial condition. Contrast this with the linear advection equation, where the characteristic map is determined as soon as we write down the PDE, regardless of our choice of initial conditions. For the Burgers equation, the characteristic map isn’t determined until we specify the initial condition (which makes sense intuitively; since the direction of the ‘flow’ is dependent on the value of $u$ in this case, then the flow doesn’t know which direction to start out in until we specify the initial condition). To see explicitly what goes wrong, consider a simple Burgers style PDE with only one time variable

$$u_t + a(x, t)uu_x = 0. \quad (5.1.2)$$

With the characteristic map $x_0(x, t)$, the solution to the PDE will be given by $u = f(x_0(x, t))$. Substituting this back into the PDE gives

$$f'(x_0(x, t))\partial_t x_0(x, t) + a(x, t)f(x_0(x, t))f'(x_0(x, t))\partial_x x_0(x, t) = 0,$$

and after cancelling $f'$, we see that

$$\partial_t x_0(x, t) + a(x, t)f(x_0(x, t))\partial_x x_0(x, t) = 0. \quad (5.1.3)$$
Now, with this in mind, return to our ansatz \( u_0 = F(x_0(x,t_0), t_1) \), with \( F(x,0) = f(x) \), and substitute this back into (5.1.1) to find

\[
\partial_{x_0} F(x_0, t_1) \partial_{t_0} x_0(x, t_0) + a(x, t_0) F(x_0, t_1) \partial_{x_0} F(x_0, t_1) \partial_{x_0} x_0(x, t_0) = 0,
\]

or, after simplifying,

\[
\partial_{t_0} x_0(x, t_0) + a(x, t_0) F(x_0(x, t_0), t_1) \partial_{x_0} x_0(x, t_0) = 0. \tag{5.1.4}
\]

Comparing this to (5.1.3), we can see that there is something wrong; as \( t_1 \) moves forward, the envelope \( F(\cdot, t_1) \) will evolve away from \( f(\cdot) \), and hence the balance of (5.1.3) will not be held. Thus, our ansatz is wrong:

\[
u_0(x, t_0, t_1) \neq F(x_0(x, t_0), t_1).
\]

However, we are not far from the right expression. In fact, we might ask, for a fixed \( t_1 \), which function \textit{does} satisfy (5.1.4)? It would be the characteristic map from (5.1.2) with initial condition \( F(x, t_1) \). It is as if we want a different characteristic map for each time \( t_1 \). Let us try exactly that, then, and let our next guess be

\[
u_0(x, t_0, t_1) = F(x_0^t(x, t_0), t_1), \tag{5.1.5}
\]

where \( x_0^t(x, t_0) \) is the characteristic map to the Burgers equation (5.1.2) with initial condition \( F(x, t_1) \).

This is very strange. Our solution is implicitly defined in terms of our solution itself. But this is actually par for the course for the Burgers equation - anyone who has dealt with it before knows that one cannot avoid implicitly defined objects. At the very least,
we can check that this assertion is correct by direct substitution. Plugging \((5.1.5)\) into \((5.1.1)\), we have

\[
\partial_{x_0} F(x_{i_0}^t, t_1) \partial_{x_0} x_{i_0}^t (x, t_0) + a(x, t_0) F(x_{i_0}^t, t_1) \partial_{x_0} F(x_{i_0}^t, t_1) \partial_{x_0} x_{i_0}^t (x, t_0) = 0.
\]

Canceling the \(\partial_{x_0} F\) terms on both sides gives

\[
\partial_{t_0} x_{i_0}^t (x, t_0) + a(x, t_0) F(x_{i_0}^t (x, t_0), t_1) \partial_{x_0} x_{i_0}^t (x, t_0) = 0,
\]

which is exactly what we wanted (equation \((5.1.3)\)), considering which initial condition generated the map \(x_{i_0}^t\).

Proceeding now with the multi-scale analysis, the \(O(\epsilon)\) terms from the expansion give us

\[
\left( \partial_{t_0} + a(x, t_0) u_0 \partial_x + a(x, t_0) \partial_x u_0 \right) u_1 = \partial_x^2 u_0 - \partial_t u_0.
\]

As per usual, we call

\[
L := \partial_{t_0} + a(x, t_0) u_0 \partial_x + a(x, t_0) \partial_x u_0,
\]

\[
v := \partial_{xx} u_0 - \partial_t u_0,
\]

and seek to apply the FAI. This requires finding \(L^\dagger\) first:

\[
\langle f, Lg \rangle = \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} \left( f [\partial_{t_0}] g + f a u_0 [\partial_x g] + f a [\partial_x u_0] g \right) dx dt_0
\]

\[
= \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} \left( - [\partial_{t_0} f] g + [\partial_x f] a u_0 g - f a x u_0 g - f a [\partial_x u_0] g \right) dx dt_0
\]

\[
= \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} \left( - [\partial_{t_0} f] g - a u_0 [\partial_x f] g - a x u_0 f g \right) dx dt_0
\]

\[
= \frac{1}{T} \int_{\tau_1}^{\tau_2} \int_\mathbb{R} \left( - [\partial_{t_0} f] - a u_0 [\partial_x f] - a x u_0 f \right) g dx dt_0
\]

\[
= \langle L^\dagger f, g \rangle
\]
and so we have

$$L^\dagger = -\partial_{t_0} - a u_0 \partial_x - a x u_0.$$ 

Now we must determine the kernel of $L^\dagger$. Explicitly, this would be a function $\varphi$ such that

$$\partial_{t_0} \varphi + a(x, t_0) u_0(x, t_0, t_1) \partial_x \varphi = -a_x(x, t_0) u(x, t_0, t_1) \varphi. \quad (5.1.7)$$

As before, let us write the function $\varphi$ as a product two functions:

$$\varphi = W(x_0^{t_1} (x, t_0)) e^{p(x,t_0,t_1)}.$$

Substituting this into (5.1.7) gives us

$$W'[\partial_{t_0} x_0^{t_1}] e^p + W e^p p_{t_0} + a u_0 W' [\partial_x x_0^{t_1}] e^p + a u_0 W e^p p_x = -a_x u_0 W e^p.$$ 

Dividing through by $e^p$, writing $u_0 = F(x_0^{t_1}, t_1)$ and regrouping, we have

$$W' \left[ \partial_{t_0} x_0^{t_1} \right] + a F(x_0^{t_1}, t_1) \partial_x x_0^{t_1} + W [p_{t_0} + a F(x_0^{t_1}, t_1) p_x + a x u_0] = 0.$$ 

The first term (with $W'$) vanishes due to (5.1.6). Thus $p$ must solve the equation

$$p_{t_0} - a F(x_0^{t_1}, t_1) p_x = a x F(x_0^{t_1}, t_1).$$

Once again, the solution to this PDE will be a combination of solutions to the homogeneous part (which one can easily check will be of the form $W(x_0^{t_1}, t_1)$) as well as a particular solution $\tilde{p}(x, t_0, t_1)$. So, as before, we find

$$\ker(L^\dagger) = \left\{ W(x_0^{t_1} (x, t_0), t_1) e^{\tilde{p}(x,t_0,t_1)} \right\}.$$ 

82
Now as we set up the FAI,

$$0 = \frac{1}{T} \int_{t_1}^{t_2} \int_{\mathbb{R}} W(x^t_0(x, t_0), t_1) e^{\beta(x^t_0, t_1)} \left[ \partial_\tau^2 u_0 - \partial_t u_0 \right] d\tau dt_0,$$  \hspace{1cm} (5.1.8)

our past experience tells us to expect that the exponential factor will disappear when we switch integration variables. However, in this non-linear case, it is no longer true that $\frac{\partial x_0}{\partial x} \in \ker(L^\dagger)$. It is almost true, but not quite. Let us see.

Because it will be useful in a moment, consider without motivation the PDE

$$\partial_{t_1} q + a(x, t_0) u_0(x, t_0, t_1) \partial_x q = -a(x, t_0) \partial_x u_0(x, t_0, t_1).$$  \hspace{1cm} (5.1.9)

As usual, the solution $q$ will be a combination of a solution to the homogeneous equation and a particular solution:

$$q = q_h + \tilde{q}.$$  

With this, let’s define

$$\xi := \frac{\partial x^t_0}{\partial x} e^q(x, t_0, t_1).$$

Then it turns out that $\xi \in \ker(L^\dagger)$. To see why, first notice that if we take (5.1.4) (writing $u_0$ instead of $F(x^t_0, t_1)$), and take the derivative with respect to $x$, we have

$$\partial_{t_0} \left( \frac{\partial x^t_0}{\partial x} \right) + a u_0 \partial_x \left( \frac{\partial x^t_0}{\partial x} \right) + a_x u_0 \left( \frac{\partial x^t_0}{\partial x} \right) + a \partial_x u_0 \left( \frac{\partial x^t_0}{\partial x} \right) = 0.$$  \hspace{1cm} (5.1.10)
Now, hold on to this, and let us see what happens when we apply $L^\dagger$ to $\xi$: 

\[
L^\dagger \xi = -\partial_{t_o} \left( \frac{\partial x^{t_1}_o}{\partial x} \right) e^q - \left( \frac{\partial x^{t_1}_o}{\partial x} \right) e^q \partial_{t_o} q - au_o \partial_x \left( \frac{\partial x^{t_1}_o}{\partial x} \right) - au_o \left( \frac{\partial x^{t_1}_o}{\partial x} \right) e^q \partial_x q - a \partial_x u_o \left( \frac{\partial x^{t_1}_o}{\partial x} \right) e^q \\
= e^q \left[ -\partial_{t_o} \left( \frac{\partial x^{t_1}_o}{\partial x} \right) - au_o \partial_x \left( \frac{\partial x^{t_1}_o}{\partial x} \right) - a \partial_x u_o \left( \frac{\partial x^{t_1}_o}{\partial x} \right) \right] + e^q \left( \frac{\partial x^{t_1}_o}{\partial x} \right) \left[ \partial_{t_o} q + au_o \partial_x q + a \partial_x u_o \right] \\
= 0
\]

where the first term in brackets vanishes due to (5.1.10) and the second term vanishes due to (5.1.9). This confirms that $\xi \in \ker(L^\dagger)$ and so it can be written as

\[
\xi = \left( \frac{\partial x^{t_1}_o}{\partial x} \right) e^q = W(x^{t_1}_o, t_1) e^{\hat{\beta}}.
\]

Furthermore, since $q = q_h(x^{t_1}_o(x, t_0), t_1) + \hat{q}(x, t_0, t_1)$, we can absorb the $e^{\hat{\beta}}$ term into $W$, leaving us with

\[
\frac{\partial x^{t_1}_o}{\partial x} = W(x^{t_1}_o) e^{\hat{\beta} - \hat{q}}.
\]

From this we conclude

\[
dx = \frac{\partial x}{\partial x^{t_1}_o} dx^{t_1}_o = \frac{1}{W(x^{t_1}_o, t_1)} e^{\hat{\beta} - \hat{p}}.
\]

Returning to the FAI, we have

\[
0 = \frac{1}{T} \int_{t_1}^{T} \int_{\mathbb{R}} W(x^{t_1}_o(x, t_0), t_1) e^{\hat{\beta}(x, t_0, t_1)} \left[ \frac{\partial^2 u_0}{\partial x^2} - \partial_{t_1} u_0 \right] dx dt_0, \\
= \frac{1}{T} \int_{t_1}^{T} \int_{\mathbb{R}} W(x^{t_1}_o, t_1) e^{\hat{q}(x, t_0, t_1)} \left[ \frac{\partial^2 F(x^{t_1}_o, t_1)}{\partial x^2} - \partial_{t_1} F(x^{t_1}_o, t_1) \right] dx^{t_1}_o dt_0 \\
= \frac{1}{T} \int_{t_1}^{T} \int_{\mathbb{R}} W(x^{t_1}_o, t_1) e^{\hat{q}(x, t_0, t_1)} \left[ \beta \frac{\partial^2 F(x^{t_1}_o, t_1)}{\partial x^{t_1}_o^2} + \partial_{t_1} F(x^{t_1}_o, t_1) - \partial_{t_1} u_0 \right] dx^{t_1}_o dt_0,
\]

84
where $\alpha$ and $\beta$ are defined as usual by equations (3.2). Now we see the true cost of introducing non-linearity; the exponential factor $e^\tilde{q}$ does not disappear. This is not catastrophic, however - it just means that $e^\tilde{q}$ will be present as we integrate with respect to $t_0$:

$$0 = \int_{\mathbb{R}} W(x_0, t_1) \left[ \langle e^\tilde{q} \rangle \partial_{x_0}^2 F(x_0, t_1) + \langle e^\tilde{q} \rangle \partial_{x_0} F(x_0, t_1) - \langle e^\tilde{q} \rangle \partial_{t_1} u_0 \right] dx_0;$$

or, more compactly (suppressing the variable dependence), we have

$$0 = \int_{\mathbb{R}} W(x_0, t_1) \left[ \langle e^\tilde{q} \rangle \partial_{x_0}^2 F - \langle e^\tilde{q} \rangle \partial_{x_0} F - \langle e^\tilde{q} \rangle \partial_{t_1} F \right] dx_0.$$

Finally, this can only be guaranteed, for any function $W$, if the expression in the brackets identically vanishes:

$$\langle e^\tilde{q} \rangle \partial_{x_0}^2 F - \langle e^\tilde{q} \rangle \partial_{x_0} F - \langle e^\tilde{q} \rangle \partial_{t_1} F = 0.$$

In other words, we have our evolution equation for $F$:

$$\partial_{t_1} F(x_0, t_1) = \frac{\langle e^\tilde{q} \rangle \partial_{x_0}^2 F(x_0, t_1)}{\langle e^\tilde{q} \rangle} - \frac{\langle e^\tilde{q} \rangle \partial_{x_0} F(x_0, t_1)}{\langle e^\tilde{q} \rangle}.$$

where it’s understood that all terms are still potentially functions of $x_0^t$ and $t_1$. To summarize, then, we have the following theorem.

**Theorem 7.** For $(x, t) \in \mathbb{R} \times [0, \infty)$, let $u$ be the solution to the equation

$$\frac{\partial u}{\partial t} (x, t) + a(x, t) u(x, t) \frac{\partial u}{\partial x} (x, t) = \epsilon \frac{\partial^2 u}{\partial x^2},$$

$$u(x, 0) = f(x),$$

85
where \( a(x, t) \) is arbitrary. Then the zeroth order term of the multi-scale asymptotic expansion is given by

\[
u_0(x, t) = F \left( \chi_0^{\omega_1}(x, t), \epsilon t \right),
\]

where \( \chi_0^{\omega_1}(x, t) \) is the characteristic map to the equation \( u_t + a(x, t)uu_x = 0 \), subject to the initial condition \( u(x, 0) = F(x, t_1) \) and the evolution of the profile \( F(\cdot, t_1) \) is governed by the equation

\[
\partial_{t_1} F(x_0, t_1) = \frac{\langle e^{\tilde{q} \alpha} \rangle}{\langle e^{\tilde{q}} \rangle} \partial_{x_0}^2 F(x_0, t_1) - \frac{\langle e^{\tilde{q} \beta} \rangle}{\langle e^{\tilde{q}} \rangle} \partial_{x_0} F(x_0, t_1),
\]

provided

\[
\alpha = \left( \frac{\partial x_0}{\partial x} \right)^2, \quad \beta = \frac{\partial^2 x_0}{\partial x^2},
\]

and \( \tilde{q} \) is the particular solution to the equation

\[
\partial_{t_0} q + a(x, t_0)u_0(x, t_0, t_1) \partial_x q = -a(x, t_0)\partial_x u_0(x, t_0, t_1).
\]

**Proof.** The entirety of section 5.1 is the proof of the theorem. \( \Box \)
CHAPTER 6

Numerical Simulations

6.1 NUMERICS FOR 1-D LINEAR CASE

Now it is time to see what all this theory is worth by testing it numerically. Let us start with some examples in 1D. Our first example from Chapter 3 was (2.4.1),

$$\frac{\partial u}{\partial t}(x, t) + \cos(t)x \frac{\partial u}{\partial x}(x, t) = \epsilon \frac{\partial^2 u}{\partial x^2}(x, t)$$  (6.1.1)

for which we saw that the effective decay of the envelope function evolved according to (2.4.10):

$$\partial_{t_1} F(x_0, t_1) = 2.279 \partial^2_{x_0} F(x_0, t_1).$$  (6.1.2)

We can numerically simulate solutions for both of these equations and plot them together. As we have discussed, the averaged equation (6.1.2) is supposed to capture the longer term dynamics of the equation (6.1.1). Before our development of these averaging methods, a naive guess would have been that the long term behavior is governed by the pure diffusion equation

$$\frac{\partial u}{\partial t}(x, t) = \epsilon \frac{\partial^2 u}{\partial x^2}(x, t).$$  (6.1.3)

Thus, for each example that follows, we will simulate three things: the true (numerical) solution $u_{\text{true}}$, the averaged solution $u_{\text{ave}}$ and the solution of the pure diffusion equation
$u_{\text{diff}}$. The hope is that the solution to the averaged equation $u_{\text{ave}}$ captures the long term behavior of $u_{\text{true}}$ much better than $u_{\text{diff}}$ does.

For this and all the following one dimensional examples, our initial condition will be the Gaussian $u(x, 0) = e^{-2x^2}$.

![Figure 6.1. Initial Condition](image)

Given this choice of initial conditions, we expect that the evolution of the solution will take place near the center of our domain, away from the boundaries. Therefore, we will treat the problem as if it was periodic in space, allowing us to use a Fourier differentiation matrix to approximate spatial derivatives [28]. For evolving forward in time, we use a simple second order implicit-explicit Crank-Nicholson scheme [1]. Choosing $\epsilon = 0.1$, we begin to march forward in time.
Notice that after \( t = \pi \), the approximation doesn’t look as accurate as we might hope have h hoped. However, by \( t = 2\pi \), the approximation is, to the eye, indistinguishable from the true solution. This is because the averaged equation cares only about the entire interval, we don’t expect it to be accurate within the interval, only at the end of the interval. Because the advection is periodic, we can proceed in time:
We can quantify the accuracy further by comparing the decay of the $L^2$ norms:

Figure 6.4. Simulation of (6.1.1) at $t = 3\pi$ and $t = 4\pi$.

Figure 6.5. Decay of $L^2$ norms over two periods.
For this example, we computed analytically the coefficients $\alpha$ and $\beta$, thereby having an explicit expression for the effective equation (2.4.10). However, we would like to be able to handle problems where we can’t necessarily compute explicit formulas for $\alpha$ and $\beta$. Therefore, let us build a more robust algorithm, following in the spirit of section 3.2. Specifically, given a generic advection operator $\partial_t + a(x,t)\partial_x$, we will compute the characteristic paths by solving the ODE

$$\dot{x}(t) = a(x,t).$$

This can be done many ways, but a tried and true method for solving a system of ODEs is a fourth order Runge-Kutta scheme [15]. Once we have the characteristic map, we take the derivatives (numerically) according to the formulas (3.2.6):

$$\alpha := \left(\frac{\partial x_o}{\partial x}\right)^2, \quad \beta := \frac{\partial^2 x_o}{\partial x^2}.$$

However, there is a catch. In solving the characteristic equations numerically, we start with points $x_0$ for when $t = 0$. Solving the ODE forward, we arrive at points $x$, which yields (numerically) the function $x(x_0,t)$. To employ the formulae for $\alpha, \beta$, we’d like to have access to $x_0(x,t)$. One option is to compute the inverse function, but this can lead to numerical instability, especially in higher dimensions. Furthermore, we would encounter problems of domain of definition, since any numerically defined function is defined on a finite interval, and the inverse function may be defined outside this interval. Thus, if the function value strays too far away, it might be impossible to take the inverse. Instead, however, we can calculate the derivatives of the inverse function locally using the formulas from Calculus 1:

$$\alpha := \left(\frac{\partial x_o}{\partial x}\right)^2 = \left(\frac{1}{\frac{\partial x}{\partial x_o}}\right)^2, \quad \beta := \frac{\partial^2 x_o}{\partial x^2} = \frac{\frac{\partial^2 x}{\partial x^2}}{\left(\frac{\partial x}{\partial x_o}\right)^3}. \quad (6.1.4)$$
Once we have these, then we can integrate over time (numerically) to obtain the averaged coefficients $\langle \alpha \rangle$, $\langle \beta \rangle$, with which we can simulate the dynamics of the approximate solution. Note, however, that what we will then be simulating is the function

$$F(x, t_1),$$

as opposed to

$$F(x_0(x, t_0), t_1).$$

To do the latter, we would need to compute the inverse function $x_0(x, t_0)$ and interpolate $F$ over this new grid. This is possible to do, numerically, in one dimension, but not terribly interesting.

Proceeding, let us apply this approach to the last example in Chapter 3,

$$\frac{\partial u}{\partial t}(x, t) + a(t) \sin(x) \frac{\partial u}{\partial x}(x, t) = \epsilon \frac{\partial^2 u}{\partial x^2}(x, t). \quad (6.1.5)$$

The analysis of this problem in Chapter 3 led us to the coefficients $\alpha$ (3.4.2) and $\beta$ (3.4.3) of the effective equation that governs the slow scale dynamics. Using these formulas, it is possible to set up the averaging integral explicitly and, after some clever algebraic tricks, arrive at an explicit averaged equation, where the differential operator coefficients end up as cubic polynomials in $x_0$. However, as this is very time consuming and not generalizable, we will instead follow the approach described above. First, we compute the characteristics by solving the (uncoupled) system of ODE (see figure 6.6 for a plot of characteristic curves).
Using the same numerical schemes for spatial and temporal differentiation as before, as well as the same initial condition, we find the following:

![Figure 6.6. Characteristic curves to equation (6.1.5)](image1)

![Figure 6.7. Simulation of (6.1.5) at $t = \pi/2$.](image2)
Figure 6.8. Simulation of (6.1.5) at $t = \pi/2, \pi, 3\pi/2, 2\pi$. 
Once again, while we see interesting things happening throughout the interval, it is not until the end of the interval \((t = 2\pi)\) that we see how well \(u_{\text{ave}}\) approximates \(u_{\text{true}}\). In fact, in his paper [16], Krol discusses the numerical utility of averaging over a periodic advection. Specifically, he discusses how it is sometimes necessary to ignore the finer scales in some problems, due to memory constraints. The previous two examples suggest that if we first compute the averaged equation and simulate on the coarser time scale, then our averaged equation will be an effective model if we are only concerned with the longer time scales.

It is also important to note here that, as mentioned before, the main advantage of the techniques developed here is the generality. In many averaging techniques, the advection operator, which may depend on \(x\) and \(t\), is at least separable. To exemplify the generality of the present algorithm, let us consider a non-separable advection \(a(x, t) = 0.3 \cos(2x - t)\) (chosen for aesthetic value). It is worth noting also, that the characteristics are no longer periodic in the sense that they do not return to where they started:

![Figure 6.9: Characteristics for \(a(x, t) = 0.3 \cos(2x + t)\)](image-url)
Running a similar simulation, we find that the motion of the true solution is much more exotic, and in light of the fact that the characteristics do not return to their original point of departure, then we find that the averaged solution, after full periods, does not seem to accurately approximate the true solution. For that, we would need to interpolate the averaged solution over the characteristic map (i.e. plot $F(x_0(x,t), \epsilon t)$, rather than $F(x, \epsilon t)$). However, by comparing the decay of the $L^2$ norms, it appears that the averaged solution really is telling us something about the decay rate of the system (see Figure 6.9).

![Graph](image)

**Figure 6.10:** Above: Profiles at $t = 6\pi$, Below: Decay of $L^2$ norms
6.2 NUMERICS FOR THE NON-LINEAR CASE

It turns out (not surprisingly) that the non-linear case is much more complicated (even numerically) than the linear case. Recall that our approximating solution \( u_0 \) is given by

\[
u_0(x, t) = F(x^t_0(x, t), t_1),
\]

where \( x^t_0(x, t) \) is the characteristic map to the equation \( u_t + a(x, t)uu_x = 0 \), but with the initial condition \( F(x, t_1) \). Recall also that our effective equation over the slow time scales was (5.1.11)

\[
\partial_{t_1}F(x_o, t_1) = \frac{\langle e^{\tilde{q}_1} \alpha \rangle}{\langle e^{\tilde{q}_1} \rangle} \partial_x^2 F(x_o, t_1) - \frac{\langle e^{\tilde{q}_1} \beta \rangle}{\langle e^{\tilde{q}_1} \rangle} \partial_x F(x_o, t_1).
\]

where

\[
\alpha = \left( \frac{\partial x^t_0}{\partial x} \right)^2, \quad \beta = \frac{\partial^2 x^t_0}{\partial x^2},
\]

and \( \tilde{q} \) is the particular solution to equation (5.1.9) (see section 5.1)

\[
\partial_{t_0}q + a(x, t_0)u_0(x, t_0, t_1)\partial_x q = -a(x, t_0)\partial_x u_0(x, t_0, t_1).
\] (6.2.1)

First of all, this means we need to compute a new characteristic map for each moment in time. Also, it means that our coefficients \( \alpha \) and \( \beta \) will now be time dependent; even though we average out the \( t_0 \) dependence, they now depend on \( t_1 \) (being derived from \( x^t_0 \)). So our effective equation is no longer time-independent, but at least it only depends on the slower scales. However, it gets worse. We want to compute these characteristic maps in order to simulate the function \( F \). But, as just described, the characteristic maps themselves depend on the function \( F \). The only way to break out of this circular trap is to simply compute the ‘true’ solution numerically, using any one of our favorite schemes. Once we have this, assuming it is close to \( u_0 \), we can go back and compute the coefficients \( \alpha, \beta \). This may seem like a defeat - one could ask, what is the use of this method if we
have to solve the PDE in full at the very start? There are two (potential) virtues in doing this. The first is that, we can at least confirm that the multi-scale approximation is accurate, if for no other reason than mathematical curiosity. The second, is that there may yet be qualitative information to extract from averaged equation in the future, and it would therefore be worthwhile for us to know that it is accurate.

Once we compute the characteristic maps, we can compute $\alpha$ and $\beta$ from the formulas above. But how do we compute $\tilde{q}$? Recall from section 1.5 that the particular solution of the inhomogeneous equation governing $\tilde{q}$ will be given by the integral

$$\tilde{q}(x,t_0,t_1) = \int_0^{t_0} g\left(x(x^{t_0}_{t_1}(x,t),s),s\right)ds,$$

Rather than find a clever way to approximate this integral, we can use the fact that, evidently, $\tilde{q}(x,0,0) = 0$. Thus, after we compute $u_0$, we can then compute the solution to the equation governing $\tilde{q}$ (6.2.1) in which $u_0$ appears as a coefficient and simply use $q \equiv 0$ as the initial condition.

However, there is yet another complication. We need to compute $x^{t_1}_{t_0}(x,t_0)$ in order to compute $\alpha, \beta, \tilde{q}$, but recall the definition of $x^{t_1}_{t_0}$; it is the characteristic map of the equation $u_t + a(x,t)uu_x = 0$ with the initial condition $u(x,0) = F(x,t_1)$. Thus, we need access to $F(x,t_1)$, which is the envelope at time $t_1$. However, what we actually will have access to is $u_0$, which is $F(x^{t_1}_{t_0}(x,t_0),t_1)$, the envelope at the right time, but interpolated over the characteristic map. This obstacle can be overcome by realizing that the quantity $F(x^{t_1}_{t_0}(x,t_0),t_1)$ is simply the result of taking the initial condition $F(x,t_1)$ and advancing it through the PDE $u_t + a(x,t)uu_x = 0$ by an amount $t = t_0$. Thus, we have the following strategy:

1. Numerically compute the ‘true’ solution $u$, thereby giving us an approximation for $u_0$. 

98
2. Use $u_0$ to numerically compute $\tilde{q}$, by computing the solution to (5.1.9) with identically zero initial conditions.

3. For each time $t_1$, take $u_0|_{t=t_1}$ as the ‘current condition’ of the PDE $u_t + a(x,t) uu_x = 0$ and numerically solve backwards to $t = 0$, thus giving us $F(x,t_1)$, as well as forward to the final time, giving us as the characteristic map $x_{01}^t(x,t_0)$ defined over the entire interval.

4. Due to the nature of how we compute the characteristics in this case, the characteristic map will actually be stored numerically as $x_{01}^t(x,t_0)$. This is convenient, as we can then compute the quantities $\alpha$ and $\beta$ using the original formulas $\alpha = \left( \frac{\partial x_{01}^t}{\partial x} \right)^2$ and $\beta = \frac{\partial^2 (x_{01}^t)^2}{\partial x^2}$, as opposed to the inverse formulas we used in section 5.1. However, we will now need to invert the characteristic map numerically in order to write $\alpha$ and $\beta$ in terms of $x_0$, not $x$.

5. Average $\alpha, \beta, \tilde{q}$ over the desired time interval and use these coefficients (which will depend on $x_0$ and $t_1$) to numerically evolve the envelope function. Compare to $u_0$.

Let us start with something simple,

$$u_t = \cos(t) uu_x = \epsilon u_{xx}. \quad (6.2.2)$$

Then following the procedure outlined above, we obtain the following figures.
Figure 6.11. Simulation of (6.2.2) at time $t = 0$, $\pi/2$ and $t = \pi$. 
Figure 6.12. Simulation of (6.2.2) at time $t = \frac{3\pi}{2}$, and $t = 2\pi$.

It appears to be quite accurate. Figure 12 gives us the decay of the $L^2$ norms.

While it doesn’t look quite as accurate as is suggested by the previous figures, it is interesting to note the oscillatory behavior of the averaged solution. Recall that the coefficients of the averaged equation are no longer time-independent in this case, but rather they only depend on the slower time parameter. Interestingly, they seem to be trying their hardest to mimic the oscillatory behavior of the true solution. It is unclear if the ‘over-enhanced’ decay rate of the averaged equation is actually a short coming of the method, or if it is simply a numerical artifact (given all the problems outlined previously).
This gives us confidence in the effectiveness of the formulas. Unfortunately, if we consider a more exotic advection, the averaged equation doesn’t seem to be as accurate as the previous example. Consider the equation

$$u_t + \cos(t) \sin(2x^2) uu_x = \epsilon u_{xx},$$

where the advection was chosen for both aesthetic novelty as well as numerical stability. Applying the same code to this system gives us the following figures.
Figure 6.14. Simulation of (6.2.3) at times $t = 0, \pi/2, \pi, 3\pi/2, 2\pi$.

A close up view at $t = 2\pi$ (figure 25) reveals that the accuracy in this case does not seem to be as good.

Figure 6.15. Close up view of solution to (6.2.3) at $t = 2\pi$. 
Now we consider using numerical techniques to approach the problem in higher dimensions;

\[ u_t(\vec{x}, t) + \vec{a}(\vec{x}, t) \cdot \nabla u(\vec{x}, t) = \epsilon \Delta u(\vec{x}, t). \]

Specifically, we will compute the characteristic map \( \vec{x}_o(\vec{x}, t) \) as before, and compute the coefficients

\[ \langle \alpha_{k,j} \rangle := \frac{1}{T} \int_{t_1}^{t_2} \nabla x_o^k \cdot \nabla x_o^j \, dt \]
\[ \langle \beta_j \rangle := \frac{1}{T} \int_{t_1}^{t_2} \Delta x_o^j \, dt, \quad (6.3.1) \]

as described in Chapter 4. We then approximate our solution \( u \) as

\[ u_o(x, t_0, t_1) = F\left(x_o(x, t_0), t_1\right), \]

by simulating the evolution of \( F \) according to the equation

\[ \partial_{t_1} F(x_o, t_1) = \sum_{k,j} \langle \alpha_{k,j} \rangle \partial_{x_o^k x_o^j} F(x_o, t_1) + \sum_j \langle \beta_j \rangle \partial_{x_o^j} F(x_o, t_1). \]

However, as in the one dimensional case, there is a slight problem with the direction in which we compute the characteristic map. From the formulas above, it is clear that we want to compute derivatives of \( \vec{x}_o \) in terms of \( \vec{x} \). However, given that we compute the characteristic map from \( t = 0 \) forward in time, what we actually will possess is the map \( \vec{x}(\vec{x}_o, t) \). Thus, we need a way of expressing the derivatives of the inverse maps, analogous to \((6.1.4)\), which will give us the \( \alpha \)'s and \( \beta \)'s in terms of derivatives of \( \vec{x} \) as functions of \( \vec{x}_o \). This higher dimensional case is significantly more complicated than the one dimensional case, and we will need to take a moment to derive some expressions for inverse partial derivatives.
Once again, denoting the components of $\vec{x}$ and $\vec{x}_o$ by $x^i$ and $x^i_o$ respectively, let us write down the statement that $\vec{x} = \vec{x}(\vec{x}_o, t)$ and $\vec{x}_o = \vec{x}_o(\vec{x}, t)$ (suppressing the $t$ dependence) as

$$x^i = x^i(x^i_o) \quad x^i_o = x^i_o(x^i).$$

Inserting one into the other gives us

$$x^i = x^i(x^i_o(x^i)).$$

Now, taking the derivative of both sides with respect to $x^l$ gives us

$$\frac{\partial x^i}{\partial x^l} = \sum_j \frac{\partial x^i}{\partial x^j_o} \frac{\partial x^j_o}{\partial x^l},$$

or rather

$$\delta^i_l = \sum_j \frac{\partial x^i}{\partial x^j_o} \frac{\partial x^j_o}{\partial x^l}, \quad (6.3.2)$$

where $\delta$ is the Kroenecker delta. If we define the two matrices

$$(J)_j^i := \frac{\partial x^i}{\partial x^j_o} \quad (J_o)_j^i = \frac{\partial x^i_o}{\partial x^j}, \quad (6.3.3)$$

then we see that

$$\delta^i_l = \sum_j (J)_j^i (J_o)_j^l,$$

which is just an expression for matrix multiplication. This means that

$$\left(J^{-1}\right)_j^l = (J_o)_j^l. \quad (6.3.4)$$
In light of (6.3.3), the relationship (6.3.4) tells us how to write the derivatives of \( x_o \) in terms of \( x \) when all we have is the derivatives of \( x \) in terms of \( x_o \). For example, in two dimensions we would have

\[
\begin{align*}
\frac{\partial x_o}{\partial x} &= \frac{1}{D} \frac{\partial y}{\partial y_o} \\
\frac{\partial x_o}{\partial y} &= -\frac{1}{D} \frac{\partial y}{\partial x_o} \\
\frac{\partial y_o}{\partial x} &= -\frac{1}{D} \frac{\partial x}{\partial y_o} \\
\frac{\partial y_o}{\partial y} &= \frac{1}{D} \frac{\partial x}{\partial x_o}
\end{align*}
\]

(6.3.5)

where \( D = \frac{\partial x}{\partial x_o} \frac{\partial y}{\partial y_o} - \frac{\partial x}{\partial y_o} \frac{\partial y}{\partial x_o} \) is the Jacobian. Now, to get the second derivatives we can take the derivative of equation (6.3.2) with respect to \( x^k \) again:

\[
\frac{\partial}{\partial x^k} \delta^i_j = \frac{\partial}{\partial x^k} \left( \sum \frac{\partial x^i}{\partial x^l} \frac{\partial x^j}{\partial x^l} \right) = \sum \left[ \frac{\partial}{\partial x^k} \left( \frac{\partial x^i}{\partial x^l} \right) \frac{\partial x^0}{\partial x^0} \frac{\partial x^j}{\partial x^l} \right] + \sum \frac{\partial x^i}{\partial x^0} \frac{\partial x^0}{\partial x^k} \frac{\partial x^j}{\partial x^l} + \sum \frac{\partial x^i}{\partial x^0} \frac{\partial x^0}{\partial x^k} \frac{\partial x^j}{\partial x^l}.
\]

But the derivative of the delta function is zero, so we have

\[
\sum \frac{\partial x^i}{\partial x^0} \frac{\partial x^0}{\partial x^k} \frac{\partial x^j}{\partial x^l} = \sum \frac{\partial^2 x^i}{\partial x^0 \partial x^k} \frac{\partial x^0}{\partial x^l}.
\]

Looking back at the definition of the coefficients (6.3.1), we see that we will only be interested in the ‘pure’ second derivatives making up the Laplacian (no mixed second
derivatives). Thus, we can set \( k = l \) in the above expression. Doing so (and relabeling the \( j \) on the right side as as \( n \)) will give us

\[
\sum_n \frac{\partial x^n}{\partial x_0^k} \frac{\partial^2 x_o^n}{(\partial x^k)^2} = - \sum_{j,m} \frac{\partial^2 x^i}{\partial x_0^m \partial x_0^l} \frac{\partial x^n}{\partial x^k} \frac{\partial x^j}{\partial x^k} \frac{\partial x^m}{\partial x^l} \frac{\partial x^j}{\partial x^k} \frac{\partial x^l}{\partial x^k}.
\]

This is a linear system for the unknowns \( \frac{\partial^2 x_o^n}{(\partial x^k)^2} \) and we can solve the left side by contracting both sides with the inverse of \( \frac{\partial x^i}{\partial x_0^k} \):

\[
\frac{\partial^2 x_o^n}{(\partial x^k)^2} = - \sum_{i,j,m} \frac{\partial^2 x^i}{\partial x_0^m \partial x_0^l} \frac{\partial x^n}{\partial x^k} \frac{\partial x^j}{\partial x^k} \frac{\partial x^m}{\partial x^l} \frac{\partial x^j}{\partial x^k} \frac{\partial x^l}{\partial x^k} \frac{J^{-1}}{k} \frac{J^{-1}}{j} \frac{J^{-1}}{n}.
\]

But, of course, we want to write the right hand side in terms of \( x_o \) as the dependent variable, so we use the definition \((6.3.3)\) to write

\[
\frac{\partial^2 x_o^n}{(\partial x^k)^2} = - \sum_{i,j,m} \frac{\partial^2 x^i}{\partial x_0^m \partial x_0^l} \left( J^{-1} \right)_k^m \left( J^{-1} \right)_k^j \left( J^{-1} \right)_m^n.
\]

To recap then, between \((6.3.3)\), \((6.3.4)\), and \((6.3.6)\), we now have all the expressions we need to compute the \( \alpha \)'s and \( \beta \)'s in terms of \( x_o \) as the independent variable:

\[
\langle \alpha_{m,n} \rangle := \frac{1}{T} \int_{T_1}^{T_2} \nabla x_0^m \cdot \nabla x_0^n \, dt_o = \frac{1}{T} \int_{T_1}^{T_2} \left[ \sum_i \left( J^{-1} \right)_i^k \left( J^{-1} \right)_i^n \right] \, dt_o, \quad \text{\((6.3.7)\)}
\]

\[
\langle \beta_n \rangle := \frac{1}{T} \int_{T_1}^{T_2} \Delta x_0^n \, dt_o = \frac{1}{T} \int_{T_1}^{T_2} \left[ - \sum_{i,j,k,m} \frac{\partial^2 x^i}{\partial x_0^m \partial x_0^l} \left( J^{-1} \right)_k^m \left( J^{-1} \right)_k^n \right] \, dt_o, \quad \text{\((6.3.8)\)}
\]
where, again, $(J)_i^j = \frac{\partial y^j}{\partial x_i^o}$. To get a better sense of what these formulas look like, we present them explicitly for the 2-D case:

\[
\Delta = \frac{1}{D^2} \left[ \left( \frac{\partial x}{\partial y_o} \right)^2 + \left( \frac{\partial y}{\partial y_o} \right)^2 \right] \partial_{x_o}^2 + \frac{1}{D^2} \left[ \left( \frac{\partial x}{\partial x_o} \right)^2 + \left( \frac{\partial y}{\partial x_o} \right)^2 \right] \partial_{y_o}^2 - \frac{2}{D^2} \left[ \frac{\partial x}{\partial x_o} \frac{\partial x}{\partial y_o} + \frac{\partial y}{\partial x_o} \frac{\partial y}{\partial y_o} \right] \partial_{x_o y_o}
\]

\[
+ \frac{1}{D^3} \left[ -\frac{\partial y}{\partial y_o} \left( \frac{\partial x}{\partial y_o} \right)^2 + \frac{\partial y}{\partial x_o} \frac{\partial x}{\partial y_o} \right] \partial_{x_o}^2 - \frac{\partial y}{\partial y_o} \left( \frac{\partial y}{\partial x_o} \right)^2 \partial_{y_o}^2 + \frac{\partial x}{\partial x_o} \frac{\partial x}{\partial y_o} \partial_{x_o}^2 + \frac{\partial x}{\partial x_o} \frac{\partial y}{\partial y_o} \partial_{y_o}^2
\]

\[
+ \frac{1}{D^3} \left[ -\frac{\partial y}{\partial x_o} \left( \frac{\partial x}{\partial x_o} \right)^2 + \frac{\partial y}{\partial y_o} \frac{\partial y}{\partial x_o} \right] \partial_{x_o}^2 - \frac{\partial x}{\partial x_o} \frac{\partial x}{\partial y_o} \partial_{x_o}^2 + \frac{\partial x}{\partial x_o} \frac{\partial y}{\partial y_o} \partial_{y_o}^2
\]

\[
+ \frac{2}{D^3} \left[ \frac{\partial x}{\partial x_o} \frac{\partial x}{\partial y_o} + \frac{\partial x}{\partial x_o} \frac{\partial y}{\partial y_o} \right] \partial_{x_o y_o} - \frac{2}{D^3} \left[ \frac{\partial x}{\partial x_o} \frac{\partial x}{\partial y_o} + \frac{\partial x}{\partial x_o} \frac{\partial y}{\partial y_o} \right] \partial_{x_o y_o}
\]

where, again, $D = \frac{\partial x}{\partial x_o} \frac{\partial y}{\partial y_o} - \frac{\partial x}{\partial y_o} \frac{\partial y}{\partial x_o}$. This example was computed by hand. In general, it is best to let a computer do these types of things. Now that we have an explicit way to compute the $\alpha$'s and $\beta$'s, we can run simulations just as in section 6.1.

6.4 NUMERICS IN HIGHER DIMENSIONS (RESULTS)

Let us begin in two spatial dimensions by replicating existing results. In their paper [23], Schaefer, Vukadinovic and Poje develop an averaging method that requires first a change of coordinates (action-angle), then another change of coordinates (to the characteristic
map) and finally they implement a ‘Lie-averaging’ technique. As an explicit example, they take the advection field

\[ \vec{d}(x, y, t) = \sin(t) \begin{pmatrix} \frac{y}{1+x^2+y^2} \\ -\frac{x}{1+x^2+y^2} \end{pmatrix}, \]

giving us a vortex which diminishes in strength as we move away from the origin. They begin with the initial condition \( u(t = 0) = xe^{-4(x^2+y^2)}; \)

![Figure 6.16. Initial Condition](image)

For 5 periods they simulate the true solution \( u_{true} \), the averaged solution \( u_{ave} \) and the solution to the purely diffusive equation \( u_{diff} \). If we plot the decay of the \( L^2 \) norm for each three solutions, they are nearly indistinguishable, until we zoom in:

Finally, to get a good sense of how well the averaged solution truly captures the net effect of the fast scale advection, they subtract \( u_{diff} \) from the true and averaged solutions to isolate the resulting advection.

The images in Figure 6.18 were produced using the formulas derived in chapter 4, but, to the eye, they precisely match those given in [23], which were obtained through an entirely different averaging technique.
Figure 6.17. Above: $L^2$ decay, Below: $L^2$ decay zoomed in.

Figure 6.18. Left: $u_{\text{true}} - u_{\text{diff}}$, Right: $u_{\text{ave}} - u_{\text{diff}}$ at $t = 10\pi$.

Moving on to three dimensional flows, we consider the ABC flow, given as the solution to the system of ODE

$$\dot{x} = \alpha \sin(z) + \gamma \cos(y)$$
$$\dot{y} = \beta \sin(x) + \alpha \cos(z)$$
$$\dot{z} = \gamma \sin(y) + \beta \cos(x).$$
A corresponding vector field

\[ \vec{a}(x, y, z) = \begin{pmatrix} \alpha \sin(z) + \gamma \cos(y) \\ \beta \sin(x) + \alpha \cos(z) \\ \gamma \sin(y) + \beta \cos(x) \end{pmatrix} \]

is an incompressible solution to the Euler equation, and is known to be a relatively simple example of a fluid flow displaying turbulent, chaotic trajectories \[7\]. To make our flow periodic in time, we multiply by \(3 \cos(t)\) (the \(3\) is there to make things more visually interesting):

\[ \vec{a}(x, y, z, t) = 3 \cos(t) \begin{pmatrix} \alpha \sin(z) + \gamma \cos(y) \\ \beta \sin(x) + \alpha \cos(z) \\ \gamma \sin(y) + \beta \cos(x) \end{pmatrix} \]

We can run a similar experiment as in the two dimensional case, only this time, the resolution will be coarser (using a desktop computer, this is actually pushing the limits of allotted memory) and for fewer periods (for the code to finish in any reasonable time, ie., before a thesis defense). Furthermore, in the 1 and 2 dimensional case the characteristic map was computed over the entire interval, and then averaged down via the integral in \(t_0\). Computing the characteristic map over the entire interval in 3 dimensions, however, was beyond the memory capacity of the desktop computer used in this simulation. Instead, the characteristics were computed, for each point, one time step at a time and added to a cumulative sum (which represents a trapezoidal approximation to the \(t_0\) integral). This way, the characteristic map did not need to be saved all at once.

For visual representations, we will print 2-dimensional slices at \(z = 0\). Using a similar initial condition, \(u(t = 0) = xe^{-4(x^2+y^2+z^2)}\), we choose the parameter values \(\alpha = -1.2, \beta = 0.1, \gamma = -0.5\) (chosen by trial and error to produce visually interesting results) and \(\epsilon = 0.005\). The significantly smaller value of \(\epsilon\) (as compared to the 1-D examples) is chosen not for accuracy but for numerical stability. Presumably, we don’t
need to take $\epsilon$ quite this small in order for the averaged equation to look respectably accurate.

Figure 6.19. Initial Condition

We plot 4 moments in time over the course of 5 periods:

Figure 6.20. 4 time/space slices for dissipative ABC flow (true solution)
Compare this to the averaged equation:

![Image of the averaged equation]

Figure 6.21. 4 time/space slices for dissipative ABC flow (averaged solution)

Though they look wildly different, we are viewing moments within the interval. The real triumph comes when we compare moments at the end of the interval, for example, at $t = 10\pi$:

![Image of comparison at $t = 10\pi$]

Figure 6.22. Left: $u_{true}$, Right: $u_{avg}$, at $t = 10\pi$.

Or, zooming in, we can see in a little more detail:
In the spirit of Schaefer et al., we plot the true and averaged solution but subtract off the purely diffusive solution, to get an idea of the underlying similarity in the advective dynamics:

Finally, we compare the decay of the $L^2$ norms in Figure 6.24, which confirms that the averaged equation really does capture the long term behavior of the original equation (far better, at least, than the pure diffusion equation).
Figure 6.25.  Decay of $L^2$ norm for ABC flow.

The extent to which the averaged equation captures the long term behavior over the pure diffusion equation is, rather unambiguously, impressive.
Conclusion

What have we done? In this work, we have derived a general framework for approximating the solutions to the advection-diffusion equation on an unbounded domain. In particular, we have developed explicit formulas defining the zeroth order term in the multi-scale asymptotic expansion of the solution. As it was discussed in section 4.2, similar formulas have been developed in the past [16] but it is the intention of this author that the results presented here be more general and explicit, and, in particular, lead directly to numerical simulations. They are more general in the sense that they not only facilitate any type of linear advection field, in any dimension, but they also provide a path for generalizing to non-linear advection. Furthermore, not only do our formulas provide an approximation to the true solution, they provide us with an obvious way of averaging out the faster time scales and arrive at an effective equation governing the slower, dissipative time scales. Access to these effective equations can have both theoretical and practical benefits. Theoretically, it could be useful to have explicit expressions governing the long term behavior of a solution to a PDE, especially considering that we have certain convergence estimates discussed in section 4.2. Practically speaking, the numerics almost speak for themselves.

The biggest drawback of this technique, so far however, is that it is limited to unbounded domains. An obvious next step in research would be to extend these ideas to bounded domains subject to various boundary conditions. As mentioned in section 4.2, due to the effect of boundary layers, one would probably need to employ a multiple scale
approach for the \textit{spatial} variables as well. How the decomposition on multiple time and space scales behaves is sure to be complicated and truly fascinating. Another next step to take would be to consider higher order terms in the asymptotic expansion, although deeper analysis would have to be undertaken in the theory of asymptotic expansions themselves. Specifically, while the zeroth order term is usually unique (the solution to the unperturbed problem), the higher order terms are not. In pursuing the first or second order terms of the multi-scale expansion, we therefore might need to impose additional conditions.

There are yet other directions in which one could depart from here. The one certain thing is that this is only the beginning.


