Propagation in Smooth Random Potentials

A thesis presented
by

Scot Elmer James Shaw

to
The Department of Physics
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of

Physics

Harvard University
Cambridge, Massachusetts
May 2002

Thesis advisor | Author

Eric J. Heller | **Scot Elmer James Shaw**

Propagation in Smooth Random Potentials

Abstract

The theoretical study of micron-scale quantum-mechanical systems generally begins with two assumptions about the potential: that there is no background potential, and that any confining potential is hard-walled. In this thesis, we will look at a phenomenon that is seen when these assumptions are not made, in the context of electron conductance through two-dimensional electron gasses (2DEGs).

We begin by setting out two different mathematical frameworks for studying systems with smooth potentials. The discrete variable representation method treats closed systems, where one is solving for eigenstates and eigenvalues. The inverse Green's function method treats open systems, where one is solving for the scattering matrix and steady-state electron flux. It is the latter method that we will apply to the case of 2DEG conductance.

Our study is motivated by recent experiments which probed the spatial pattern of electron flux. In agreement with these experiments, we find that electrons follow narrow branches rather than a diffusive spreading pattern. We conclude that the branches are the result of small-angle scattering off of a weak, smooth disordered background potential, generated by the layer of donor atoms in the 2DEG crystal. We then consider the experimentally observed interference fringes, which persist to ranges of several microns. We present a novel model that explains the persistence and character of these fringes, relying only on first-order scattering off of the sharp potentials generated by impurities in the crystal.

We then turn to the methods of classical mechanics to study the branching pattern. Using classical trajectory stability analysis, we show that the locations of branches can be predicted by a projection of the stability matrix onto the initial manifold. We also consider the scaling laws for various statistical aspects of the classical flow (for example, the momentum-relaxation time). We find that these properties of the branched flow adhere to our theoretical predictions. Finally, we consider what one-dimensional maps can tell us about the dynamics in these systems. The map gives us an understanding of the observed correlation between branch stability properties and turning points in the evolving manifold.

Contents

List of Figures

Acknowledgments

Because this may be my only opportunity to thank these individuals in writing, I may be a bit more verbose in my thanks than necessary.

As my advisor, Rick Heller deserves thanks for many things. Notably, for creating the research environment in which I have performed my graduate studies. I came to Harvard intending to join his group, for two primary reasons: the areas of research in which he was involved, and what I had heard about his personality. I have not been disappointed in either regard. He has provided guidance at key moments in my work while also allowing me to work independently the majority of the time.

Because of the research environment sustained by Rick, I have crossed paths with many graduate students and postdocs who have influenced and enhanced my research. The direction of my graduate work has been strongly influenced by the members of the Westervelt group, notably Mark Topinka and Brian LeRoy. Within the Heller group, speaking with Ragnar Fleischmann helped me as I got my feet under me for 2DEG conductance calculations. Michael Haggerty's example has led me to the current state of my entire programming philosophy. Alex Barnett has provided me the benefits of his own numerical experience. Lev Kaplan and Steve Tomsovic have, through their writings and direct communication, provided me with insights into the more analytic ways to understand these systems.

The influence of those with whom I worked before coming to Harvard also continues to be important. Professors David Cook and John Brandenberger at Lawrence University provided me the foundations that allowed me to hit the ground running at Harvard and complete my thesis in four years. David helped my numerical skills to flourish, and John subjected me to tough examination of my physical intuition. I thank Keith Burnett at Oxford for taking a chance on an undergraduate from the US and giving me my first real-world research experience.

My parents, Michael and Susan Shaw, need to figure in here somewhere: they did provide me with more than just my genetic material, after all. My interest in science and mathematics started early in my life, and they encouraged and educated me in ways both straightforward and devious. Trips to museums, books about science, chemistry experiments with stuff found in the kitchen — most of us who end up going into science probably had these experiences. But how many of us, as young children, were fooled into thinking that a Little Professor[1] was, in fact, the hand-held electronic game for which we had been asking?

As long as I'm writing names down, I cannot neglect that of Arik Martin, my best, most constant friend since high school. He's provided me a lot of support and encouragement over the years; he's a great guy and I don't want to miss an opportunity to get that in the permanent record.

Finally, Moriah Tumbleson, my girlfriend since we were Freshmen at Lawrence University. She came with me to Massachusetts when I came to graduate school, away from her family, friends, and job. Her support can be summed up in the following

---

[1] A device put out by Texas Instruments that looks like a calculator, but which presents the user with simple math problems to answer.

fact: at this very moment, she is proofreading my thesis for me. Though she is not a scientist, she will, perhaps, be the only person other than me who will read the entire thing.

Chapter 1

Discrete Variable Representations

This chapter stands out from much of the rest of this thesis in that it deals with a numerical method in isolation from any results. Though the method of solving differential equations by discrete variable representation (DVR) proved to be one that didn't figure prominently in this research, it is nonetheless important to bring it to the attention of the physics community. It is much better known within the chemistry community.

DVR is a method for solving differential equations. It seems to straddle the boundary between grid discretization and basis-function expansion, two methods well known to the physics community. We will treat it largely with the language of a basis-function expansion, though it carries more baggage than that might imply. Note that some authors refer to it as a variational basis.

The DVR methods are useful in quantum mechanics for solving Schrödinger's equation when we have some useful information about the potential. DVR finds a home in this thesis because it is implicitly a smooth-potential technique. See also [1] for a good introduction to DVR bases.

1.1 Properties of the DVR Basis Functions

For any basis-function expansion, we need first to identify the basis functions that we are going to use. There are many different DVR bases, and the one that you choose will depend on the problem that you desire to solve. This is where some information about the smooth potentials to be encountered is useful. This will become clearer later on; for now, we can state some general properties of the DVR basis functions and how they become useful. We will do so in one dimension, taking higher-dimensional basis sets to be products of one-dimensional sets.

In the most general sense, we will require the following of our DVR basis functions. We are after a set of N basis functions $\{f_i(x)\}$ with the properties

$$f_i(x_j) = \lambda_i^{-1/2}\delta_{ij} \tag{1.1}$$

for a set of N points $\{x_j\}$ and weights $\{\lambda_i\}$, and

$$\int_a^b dx\ f_i(x) f_j(x) = \delta_{ij} \tag{1.2}$$

for an interval $[a, b]$ that contains all of the x_j. We then use these basis functions to approximate wave functions by taking

$$\psi(x) \approx \sum_{i=0}^{N-1} \psi(x_i) f_i(x) \lambda_i^{1/2} \tag{1.3}$$

which, given our requirements for the basis functions, is exact at the grid points.

For two wave functions $\phi(x)$ and $\psi(x)$ approximated as above, we can express the overlap integral by

$$\begin{aligned}\int_a^b dx\ \psi(x)\phi^*(x) &\approx \int_a^b dx\ \left(\sum_i \psi(x_i)f_i(x)\lambda_i^{1/2}\right)\left(\sum_j \phi^*(x_j)f_j(x)\lambda_j^{1/2}\right) && (1.4)\\ &= \sum_{i,j}\lambda_i^{1/2}\lambda_j^{1/2}\psi(x_i)\phi^*(x_j)\int_a^b dx\ f_i(x)f_j(x) && (1.5)\\ &= \sum_i \lambda_i\psi(x_i)\phi^*(x_i). && (1.6)\end{aligned}$$

Motivated by this result, we then take

$$\int_a^b dx\ F(x) \approx \sum_{i=0}^{N-1}\lambda_i F(x_i) \tag{1.7}$$

as an approximate quadrature rule. We will use this quadrature rule frequently in determining DVR functions and weights. At this point, we have made statements about neither the quality of this integral approximation nor any restrictions on the function $F(x)$. As is reasonable, however, the quality of the solutions that we get from our DVR basis is related to the quality of this integral.

1.2 DVR Starting with General Orthonormal Functions

It is, in general, possible to build a basis set $\{f_i(x)\}$ for a DVR beginning with an arbitrary set of orthonormal functions $\{p_i(x)\}$, where

$$\int_a^b dx\ p_i(x)p_j(x) = \delta_{ij}, \tag{1.8}$$

and arbitrary points $\{x_i\}$. Though such a general construction is possible, the quality of the resulting approximation will generally be poor. We look at two cases, one where we have arbitrary points $\{x_i\}$, and one where these points are chosen such that the quadrature rule is trustworthy (though no prescription will be given until the next section for how to achieve that). In either case, we can expand any function on the interval $[a, b]$ in terms of our orthonormal functions; in particular, we can expand the DVR basis functions $f_i(x)$:

$$f_i(x) = \sum_j a_{ij} p_j(x). \tag{1.9}$$

1.2.1 Arbitrary Points

If we pick random points for the $\{x_i\}$, we will generally not generate a good quadrature rule and the resulting DVR will not be particularly useful. One might imagine instances, however, when one would want to follow this path, so we include it for completeness. Since the quadrature rule is inaccurate, we will not use it to determine the expansion coefficients relating the DVR basis to the orthonormal functions. Though we have to compromise on the quadrature rule, we insist on the basic DVR property

$$\begin{aligned}
f_i(x_j) &= \lambda_i^{-1/2}\delta_{ij} && (1.10)\\
\sum_k a_{ik} p_k(x_j) &= \lambda_i^{-1/2}\delta_{ij} && (1.11)\\
\sum_k \lambda_i^{1/2} a_{ik} p_k(x_j) &= \delta_{ij} && (1.12)\\
\lambda_i^{1/2} a_{ik} &= [p_k(x_j)]^{-1} && (1.13)
\end{aligned}$$

where the inverse in the final equation is a *matrix* inversion.

We have no way at the moment to separate the weights λ_i from the expansion coefficients. Looking to the orthogonality of the $\{f_i(x)\}$, we can do so. We have that

$$\int_a^b dx\ f_i(x) f_j(x) = \delta_{ij} \tag{1.14}$$

$$\begin{aligned}
\int_a^b dx \left[\sum_k a_{ik} p_k(x)\right]\left[\sum_l a_{jl} p_j(x)\right] &= \delta_{ij} && (1.15)\\
\sum_k \sum_l a_{ik} a_{jl} \int_a^b dx\ p_k(x) p_l(x) &= \delta_{ij} && (1.16)\\
\sum_k \sum_l a_{ik} a_{jl} \delta_{kl} &= \delta_{ij} && (1.17)\\
\sum_k a_{ik} a_{jk} &= \delta_{ij} && (1.18)\\
\sum_k (a_{ik})^2 &= 1. && (1.19)
\end{aligned}$$

This expression allows us to separate out the λ_i terms after the matrix inversion.

1.2.2 Special Points

Let's assume that, somehow, we have selected points $\{x_i\}$ such that the resulting quadrature rule is good. In this case, we have a much more direct means to find the expansion coefficients a_{ij}. They are given by

$$\begin{aligned}
a_{ij} &\equiv \int_a^b dx\ f_i(x) p_j(x) && (1.20)\\
&= \sum_{k=0}^{N-1} \lambda_k f_i(x_k) p_j(x_k) && (1.21)\\
&= \lambda_i^{1/2} p_j(x_i), && (1.22)
\end{aligned}$$

where we have applied the quadrature rule in Eq. (1.7) to approximate the integral. Using this definition of the a_{ij}, we take

$$f_i(x) = \lambda_i^{1/2} \sum_j p_j(x_i) p_j(x). \qquad (1.23)$$

From Eq. (1.23) we can now determine values of the coefficients λ_i by

$$f_i(x_k) = \lambda_i^{1/2} \sum_j p_j(x_i) p_j(x_k) \qquad (1.24)$$

$$\begin{aligned}\delta_{ik}\lambda_i^{-1/2} &= \lambda_i^{1/2}\sum_j p_j(x_i)p_j(x_k) && (1.25)\\ \delta_{ik}\lambda_i^{-1} &= \sum_j p_j(x_i)p_j(x_k) && (1.26)\\ \lambda_i^{-1} &= \sum_j p_j^2(x_i). && (1.27)\end{aligned}$$

1.3 DVR Starting with Orthogonal Polynomials

Let us narrow our focus by applying the results of the previous section in the case where our orthonormal functions are related to a family of orthogonal polynomials. This will give us a way to determine the points $\{x_i\}$ needed for a good quadrature rule.

Let the functions $\{q_i(x)\}$ be a family of orthogonal polynomials on the interval $[a, b]$ with weighting function $w(x)$; that is,

$$\int_a^b dx\ w(x)q_i(x)q_j(x) = h_i\delta_{ij} \tag{1.28}$$

where h_i is the norm of $q_i(x)$. We begin by defining orthonormal functions $p_i(x) = w(x)^{1/2}h_i^{-1/2}q_i(x)$. These then look like the functions of the previous section, and we can apply those results.

In general, it is the case that there exists a Gaussian quadrature associated with a set of orthogonal polynomials and their weighting function $w(x)$. That is, there exists a set of N points $\{x_i\}$ and weights $\{\lambda_i\}$ such that any polynomial $P(x)$ of order $2N-1$ or less can be integrated exactly by

$$\int_a^b dx\ w(x)P(x) = \sum_i \lambda_i w(x_i)P(x_i). \tag{1.29}$$

We find that the points are $\{x_i\} \equiv \{x : q_N(x) = 0\}$ [2], and the weights are given by the result from the previous section,

$$\lambda_i^{-1} = \sum_j p_j^2(x_i) \tag{1.30}$$

$$= w(x_i)\sum_j q_j^2(x_i)h_j^{-1}. \tag{1.31}$$

This expression for the weights is considerably simpler to apply than that found in [2], though it is found in other references [3, Eq. 8.4.18].

Using the definition in Eq. (1.23), the $f_i(x)$ have the form of polynomials $P_i(x)$ of order $N-1$ times $w(x)^{1/2}$. Hence we have that

$$\int_a^b dx\ f_i(x)f_j(x) = \int_a^b dx\ w(x)P_i(x)P_j(x) \tag{1.32}$$

$$= \sum_k \lambda_k w(x_k)P_i(x_k)P_j(x_k) \tag{1.33}$$

$$= \sum_k \lambda_k f_i(x_k)f_j(x_k) \tag{1.34}$$

where the sum is equal to the integral because $P_i(x)P_j(x)$ is a polynomial of order $2N-2$, and thus exact in the quadrature. We see that the points and weights that we need in the definition of the DVR are just those of the Gaussian quadrature associated with the family of orthogonal polynomials.

Since the quadrature is exact for the functions $p_i(x)$, the approximations in finding $f_i(x)$ in terms of them are now exact. Hence we have that

$$f_i(x) = \lambda_i^{1/2}\sum_j p_j(x_i)p_j(x), \tag{1.35}$$

or, returning to the original set of orthogonal polynomials,

$$f_i(x) = w(x)^{1/2}w(x_i)^{1/2}\lambda_i^{1/2}\sum_j h_j^{-1}q_j(x_i)q_j(x). \tag{1.36}$$

As an interesting (and ultimately useful) aside, since $w(x)$ will, in general, have no zeroes, we have N conditions on the polynomial $P_i(x) = f_i(x)w(x)^{-1/2}$: that it must have zeroes at all $x_j : j \neq i$ and that $P_i(x_i) = w(x_i)^{-1/2}\lambda_i^{-1/2}$. Hence $P_i(x)$ is uniquely defined, and so must be of the form

$$P_i(x) \propto \prod_{j \neq i} \frac{x - x_j}{x_i - x_j}. \tag{1.37}$$

The normalization of this polynomial is left undetermined, but work with the summation in $p_i(x)$ found above shows that it does yield this form. This product form makes it a bit easier to visualize the DVR basis functions than does the summation.

1.4 Using the Expansion

We now use these basis functions and the quadrature that comes with them to solve the Schrödinger equation for a single particle. We use the associated quadrature to approximate all of the integrals that we encounter.

Expanding our wave function, we take

$$\Psi(x) \approx \sum_i c_i f_i(x), \tag{1.38}$$

where

$$\begin{aligned} c_i &= \int_a^b dx\ \Psi(x) f_i(x) && (1.39)\\ &\approx \sum_j \lambda_j \Psi(x_j) f_i(x_j) && (1.40)\\ &= \lambda_i^{1/2} \Psi(x_i). && (1.41) \end{aligned}$$

For the matrix elements of the potential energy operator $\hat{V}$, we have that

$$V_{ij} = \int_a^b dx\ f_i(x) f_j(x) V(x) \tag{1.42}$$

$$\begin{aligned} &\approx \sum_k \lambda_k f_i(x_k) f_j(x_k) V(x_k) && (1.43) \\ &= \sum_k \lambda_k \lambda_i^{-1/2} \delta_{ik} \lambda_j^{-1/2} \delta_{jk} V(x_k) && (1.44) \\ &= \delta_{ij} V(x_i) && (1.45) \end{aligned}$$

so that, insofar as we trust the quadrature, we have a diagonal potential energy. Some authors will refer to this process as choosing a basis that "diagonalizes coordinate space."

The kinetic energy is not diagonal, but in dimensions higher than one we find that it is sparse. The matrix elements K_{ij} are, in general, difficult to determine since they involve derivatives not only of the $p_i(x)$ but also of $w(x)^{1/2}$. In terms of the DVR basis functions, we have

$$\begin{aligned} K_{ij} &= \int_a^b dx\ f_i(x) \hat{K} f_j(x) && (1.46) \\ &= -\int_a^b dx\ f_i(x) f_j''(x) && (1.47) \\ &\approx -\sum_k \lambda_k f_i(x_k) f_j''(x_k) && (1.48) \\ &= -\sum_k \lambda_k \lambda_i^{-1/2} \delta_{ik} f_j''(x_k) && (1.49) \\ &= -\lambda^{1/2} f_j''(x_i). && (1.50) \end{aligned}$$

Our Hamiltonian is represented by the matrix $H_{ij} = K_{ij} + \delta_{ij} V(x_i)$. The eigenvectors give us the c_i coefficients of the expansion. We then find the values of the eigenfunctions at the grid points by $c_i = \lambda_i^{1/2} \Psi(x_i)$. Because of the higher accuracy that Gaussian quadrature gives us for a fixed number of function evaluations, using this DVR basis set allows us to solve Schrödinger's equation with smaller matrices than might otherwise be required.

1.5 Scaling

If we start from a set of orthogonal polynomials, we are given a set of points x_i with which to begin. Though it is not a complicated matter, it is easy to get confused when attempting to re-scale the points for the DVR. We can move the locations of our grid points by scaling the basis functions. Suppose we want to move the points x_i to hx_i. Then we get the matrix elements of the Hamiltonian as $H_{ij} = h^{-2}K_{ij} + \delta_{ij}V(hx_i)$. The eigenvectors are still the c_i, but those are now related to the values of the wave function at the scaled points $c_i = \lambda_i^{1/2}\Psi(hx_i)$.

1.6 Higher Dimensions

An example in three dimensions should be sufficient to demonstrate the principles. Let a three-dimensional DVR basis function be given by

$$f_{ijk}(x, y, z) = f_i(x)f_j(y)f_k(z) \tag{1.51}$$

and the kinetic energy operator be given by

$$\hat{K} = -\nabla^2 = -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \hat{K}_x + \hat{K}_y + \hat{K}_z \tag{1.52}$$

The elements of the potential energy contribution to the Hamiltonian are

$$\begin{aligned}
V_{ijklmn} &= \int_a^b dx \int_a^b dy \int_a^b dz\ f_{ijk}(x, y, z)V(x, y, z)f_{lmn}(x, y, z) && (1.53)\\
&= \int_a^b dx\ f_i(x)f_l(x) \int_a^b dy\ f_j(y)f_m(y) \int_a^b dz\ f_k(z)f_n(z)V(x, y, z) && (1.54)\\
&\approx \sum_a \lambda_a f_i(x_a)f_l(x_a) \sum_b \lambda_b f_j(x_b)f_m(x_b) \sum_c \lambda_c f_k(x_c)f_n(x_c)V(x_a, x_b, x_c) && (1.55)
\end{aligned}$$

$$= \sum_a \lambda_a \lambda_i^{-1/2} \delta_{ia} \lambda_l^{-1/2} \delta_{la} \sum_b \lambda_b \lambda_j^{-1/2} \delta_{jb} \lambda_m^{-1/2} \delta_{mb} \sum_c \lambda_c \lambda_k^{-1/2} \delta_{kc} \lambda_n^{-1/2} \delta_{nc} \times V(x_a, x_b, x_c) \tag{1.56}$$

$$= \delta_{il} \delta_{jm} \delta_{kn} V(x_i, x_j, x_k). \tag{1.57}$$

So they remain diagonal, just as seen in one dimension.

The kinetic energy is again more complicated than the potential energy.

$$K_{ijklmn} = \int_a^b dx \int_a^b dy \int_a^b dz \; f_{ijk}(x,y,z) \hat{K} f_{lmn}(x,y,z) \tag{1.58}$$

$$= \int_a^b dx \int_a^b dy \int_a^b dz \; [f_i(x) f_j(y) f_k(z)] \left(\hat{K}_x + \hat{K}_y + \hat{K}_z \right) [f_l(x) f_m(y) f_n(z)] \tag{1.59}$$

$$= \int_a^b dx \; f_i(z) \hat{K}_x f_l(z) \int_a^b dy \; f_j(y) f_m(y) \int_a^b dz \; f_k(z) f_n(z) + \cdots \tag{1.60}$$

$$= K_{il} \delta_{jm} \delta_{kn} + K_{jm} \delta_{il} \delta_{kn} + K_{kn} \delta_{il} \delta_{jm}. \tag{1.61}$$

So, though the kinetic energy is not diagonal, it is not dense.

This scaling with dimensionality of problems expressed in the DVR basis is an important aspect of the method. Suppose that we have N basis functions along each of d coordinate axes. In that case, the full Hamiltonian is a $N^d \times N^d$ array, so has a total of N^{2d} matrix elements. There are a total of N^d nonzero elements of V to calculate, and dN^{d+1} nonzero elements of K. Due to overlap, there are only dN^{d+1} nonzero elements of H. DVR, then, gives us a sparse matrix equation to solve, and is well suited to higher dimensions.

1.7 Introducing Symmetry

Many of the problems dealt with in quantum mechanics have built-in symmetries that we might wish to exploit in a solution technique. We can do that in DVR by constructing explicitly symmetric versions of the basis function set. We will assume that the DVR basis has grid points that are symmetric about the origin.[1]

Let us begin with some DVR basis set $\{f_i(x)\}$. Let us define the basis functions

$$f_i^+(x) = 2^{-1/2}\left[f_i(x) + f_{N-1-i}(x)\right] \tag{1.62}$$

$$f_i^-(x) = 2^{-1/2}\left[f_i(x) - f_{N-1-i}(x)\right] \tag{1.63}$$

where we now only let $i \in [0, N/2-1]$ and we require that N be even. Thus we have a total of N basis functions, if we take all f_i^+ and f_i^- to replace the f_i, but these new sets are in explicit symmetry classes. It is easily seen that the two bases are equivalent.

With our new basis set, we get almost all of the benefits of the original DVR basis without modification. We have, for example, that

$$\begin{aligned}
\int_a^b dx\ f_i^\pm(x) f_j^\pm(x) &= \frac{1}{2}\int_a^b dx\ \left[f_i(x) \pm f_{N-1-i}(x)\right]\left[f_j(x) \pm f_{N-1-j}(x)\right] && (1.64)\\
&= \frac{1}{2}\Bigg[\int_a^b dx\ f_i(x) f_j(x) \pm \int_a^b dx\ f_i(x) f_{N-1-j}(x) \\
&\quad \pm \int_a^b dx\ f_{N-1-i}(x) f_j(x) + \int_a^b dx\ f_{N-1-i}(x) f_{N-1-j}(x)\Bigg] && (1.65)\\
&= \frac{1}{2}\left(\delta_{ij} \pm \delta_{i,N-1-j} \pm \delta_{N-1-i,j} + \delta_{N-1-i,N-1-j}\right) && (1.66)\\
&= \delta_{ij} \pm \delta_{i,N-1-j}. && (1.67)
\end{aligned}$$

---

[1] I have yet to encounter a useful DVR basis with grid points that are not symmetric about the origin, so this seems a reasonable assumption to make.

We now note that, since we are restricting the range of i, we can never have $i = N-1-j$. Hence we are left with

$$\int_a^b dx\ f_i^{\pm}(x) f_j^{\pm}(x) = \delta_{ij}. \tag{1.68}$$

Note that this notation is meant to indicate that we make the same choice of plus or minus on both functions.

For the potential energy matrix elements, we have

$$\begin{aligned}
V_{ij}^{\pm} &= \int_a^b dx\ f_i^{\pm}(x) V(x) f_j^{\pm}(x) && (1.69)\\
&= \frac{1}{2}\int_a^b dx\ [f_i(x) \pm f_{N-1-i}(x)] V(x) [f_j(x) \pm f_{N-1-j}(x)] && (1.70)\\
&= \frac{1}{2}\left[\int_a^b dx\ f_i(x) V(x) f_j(x) \pm \int_a^b dx\ f_i(x) V(x) f_{N-1-j}(x)\right.\\
&\quad \left.\pm \int_a^b dx\ f_{N-1-i}(x) V(x) f_j(x) + \int_a^b dx\ f_{N-1-i}(x) V(x) f_{N-1-j}(x)\right] && (1.71)\\
&= \frac{1}{2}\left[\delta_{ij} V(x_i) \pm \delta_{i,N-1-j} V(x_i)\right.\\
&\quad \left.\pm \delta_{N-1-i,j} V(x_{N-1-i}) + \delta_{N-1-i,N-1-j} V(x_{N-1-i})\right]. && (1.72)
\end{aligned}$$

We assume that we are working with an even potential. Since the points of the quadrature are symmetric, so that $x_i = -x_{N-1-i}$, we have $V(x_i) = V(x_{N-1-i})$. Exploiting this result and the limitations on i and j,

$$V_{ij}^{\pm} = \delta_{ij} V(x_i). \tag{1.73}$$

Finally, the kinetic energy matrix elements are

$$\begin{aligned}
K_{ij}^{\pm} &= \int_a^b dx\ f_i^{\pm}(x) \hat{K} f_j^{\pm}(x) && (1.74)\\
&= \frac{1}{2}\int_a^b dx\ [f_i(x) \pm f_{N-1-i}(x)] \hat{K} [f_j(x) \pm f_{N-1-j}(x)] && (1.75)
\end{aligned}$$

$$\begin{aligned} &= \frac{1}{2}\left[\int_a^b dx\ f_i(x)\hat{K}f_j(x) \pm \int_a^b dx\ f_i(x)\hat{K}f_{N-1-j}(x)\right. \\ &\qquad \left. \pm \int_a^b dx\ f_{N-1-i}(x)\hat{K}f_j(x) + \int_a^b dx\ f_{N-1-i}(x)\hat{K}f_{N-1-j}(x)\right] \quad (1.76) \\ &= \frac{1}{2}\left(K_{ij} \pm K_{i,N-1-j} \pm K_{N-1-i,j} + K_{N-1-i,N-1-j}\right). \quad (1.77) \end{aligned}$$

This expression can also be simplified, in the case that each of the orthonormal functions $\{p_i(x)\}$ is either symmetric or anti-symmetric (they need not be all the same). In this case, we have that $p_i(x) = \pm p_i(-x)$ and $p_i''(x) = \pm p_i''(-x)$ for all i, as well as the already-stated condition that $x_i = -x_{N-1-i}$.

Looking first to the coefficients λ_i, we note that

$$\begin{aligned} \lambda_i^{-1} &= \sum_j p_j^2(x_i) && (1.78) \\ &= \sum_j \left[\pm p_j(-x_i)\right]^2 && (1.79) \\ &= \sum_j p_j^2(-x_i) && (1.80) \\ &= \sum_j p_j^2(x_{N-1-i}) && (1.81) \\ &= \lambda_{N-1-i}^{-1}. && (1.82) \end{aligned}$$

Now, for the kinetic energy operator, we have that

$$\begin{aligned} K_{ij} &= -\lambda_i^{1/2} f_j''(x_i) && (1.83) \\ &= -\lambda_i^{1/2}\lambda_j^{1/2}\sum_k p_k(x_j)p_k''(x_i) && (1.84) \\ &= -\lambda_i^{1/2}\lambda_j^{1/2}\sum_k \left[\pm p_k(-x_j)\right]\left[\pm p_k''(-x_i)\right] && (1.85) \\ &= -\lambda_{N-1-i}^{1/2}\lambda_{N-1-j}^{1/2}\sum_k p_k(x_{N-1-j})p_k''(x_{N-1-i}) && (1.86) \\ &= -\lambda_{N-1-i}^{1/2} f_{N-1-j}''(x_{N-1-i}) && (1.87) \\ &= K_{N-1-i,N-1-j}. && (1.88) \end{aligned}$$

Similarly, one can show that $K_{i,N-1-j} = K_{N-1-i,j}$. This gives us the simplified expression

$$K^{\pm}_{ij} = K_{ij} \pm K_{N-1-i,j}. \tag{1.89}$$

1.8 Some Useful Examples

Though the discussion above is sufficient to work out the DVR basis from any of the available sets of orthogonal polynomials, there exist very simplified expressions for some of the possible choices. That is, it is possible to create the matrices for the problem without going through the sums described above.

A good example is the choice of a DVR based on Hermite polynomials. The Hermite polynomials are orthogonal over the interval $(-\infty, \infty)$ with the weighting function $w(x) = e^{-x^2}$. The points of the quadrature are found, as usual, as zeros of a Hermite polynomial, and the weights come from a sum over Hermite polynomials at the quadrature points. I won't go through the steps of the derivation here, but it can be shown (see [1]) that the elements of the kinetic energy matrix are given by

$$\hat{K} = -\frac{\partial^2}{\partial x^2} \Rightarrow K_{ij} = \begin{cases} (1/6)(4N - 1 - 2x_i^2) & i = j \\ (-1)^{i-j}\left[2(x_i - x_j)^{-2} - .5\right] & i \neq j \end{cases}. \tag{1.90}$$

Another useful example comes from a trigonometric basis. We take the set of normalized sine functions $2^{1/2}\sin(i\pi x)$ for the range $(0, 1)$. The resulting DVR has the advantage of yielding equally spaced grid points. We find that [4]

$$x_i = (i+1)/(N+1) \tag{1.91}$$

$$\lambda_i = (N+1)^{-1} \tag{1.92}$$

$$w(x) = 1 \tag{1.93}$$

$$f_i(x) = \frac{\lambda_i^{1/2}}{2}\left\{\frac{\sin[(2N+1)(x-x_i)\pi/2]}{\sin[(x-x_i)\pi/2]} - \frac{\sin[(2N+1)(x+x_i)\pi/2]}{\sin[(x+x_i)\pi/2]}\right\} \tag{1.94}$$

$$K_{ij} = \left\{\begin{array}{ll} \frac{\pi^2}{2}\left[\frac{2}{3}(N+1)^2 + \frac{1}{3} - \sin^{-2}(\pi x_i)\right] & i = j \\ (-1)^{i-j}\frac{\pi^2}{2}\left\{\sin^{-2}[(x_i - x_j)\pi/2] - \sin^{-2}[(x_i - x_j)\pi/2]\right\} & i \neq j \end{array}\right. . \tag{1.95}$$

It is interesting to note that some trigonometric DVR bases are actually disguised polynomial bases [5]. For simplified solutions in other special cases, see [1], [4], and [6].

Chapter 2

Calculating Wave Propagation

The systems that we deal with in quantum mechanics are often closed. In such cases, we perform calculations to find eigenstates of the given potential. Though the difficulties that can be encountered in such calculations should not be minimized, this is only one kind of system that can be treated with quantum mechanics. In this chapter, we concern ourselves with developing a technique to treat open systems by calculating propagating, as opposed to standing, waves.

Furthermore, we are after a technique that makes as few simplifications to or assumptions about the potential as possible. Ultimately, we will be calculating the conductance through smooth, disordered potentials, so in particular we will avoid assumptions of flat potentials or infinitely hard walls. We will also avoid limiting ourselves to the "tunneling regime," where a nearly isolated system is connected weakly to leads. Of course, by avoiding these simplifications, we are cutting ourselves off from some very efficient techniques (see, for example, [7, Ch. 6] and [8]) and we will be limited in the sizes of the systems that we can study and the energy scales

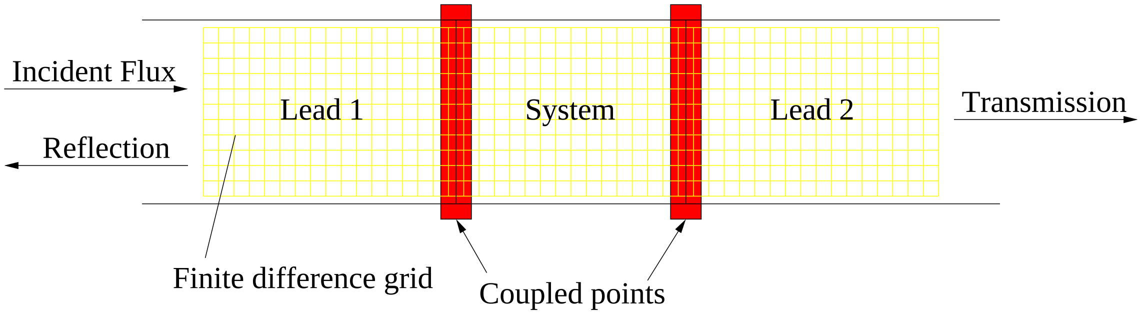


Figure 2.1: Here we see the initial conceptual framework for our numerical technique for calculating wave propagation. We envision the system contained in a trough; to the left and to the right are leads connecting the system to thermal reservoirs, which we do not include explicitly. All of the scattering takes place in a confined region referred to as the "system." No scattering occurs in the leads, which serve to define the modes or channels of the scattering. We discretize space onto a finite difference grid. Points on the grid are coupled only to nearest neighbors, and hence the communication between the scattering system and the leads is accomplished in the two bands indicated. We will eventually extend this setup to include more leads.

that we can reach. There are trade-offs to be made in the choice of any numerical technique.

Techniques to solve problems under these constraints do exist, though they are not as widely applied as the more efficient techniques involving more simplifications. The general outline of our method is as follows. We begin by thinking of the full physical environment as being composed of a "system," where the scattering occurs, and semi-infinite "leads." These leads must support quantized transverse modes. The language of the scattering process is then that of the scattering matrix (S-matrix), coupling incident modes to scattered modes. The S-matrix leads us to a calculation of the conductance, if that quantity is desired, by the familiar Landauer formula [9]. See Figure 2.1 for a schematic representation of this setup.

It should be noted that we will not solve for merely the S-matrix, though in some cases that is all the information desired. Our calculations should also make accessible the scattering states of the system.

For the mathematics required to carry out this method, we begin by finding a matrix representation of the inverse of the Green's function for the system, taking into account the coupling of the scattering system to the leads. Rather than performing a computationally costly matrix inversion, we solve a number of $Ax = b$ matrix multiplication problems to determine the response of the system to the various possible incident modes. If it is desired, we can extract the scattering matrix elements from these system responses.

Our inverse Green's function treats the entire domain of the potential at once. There are other methods in the literature that involve building up the Green's function through the system one "slice" at a time, called recursive Green's-function (RGF) methods [10, 11, 12, 13]. These methods exchange one large problem for many smaller problems. There are two weaknesses to these techniques. First, by discarding information about the interior points of the system, they render inaccessible the propagating states themselves. Second, it does not appear that they can treat systems with more than two leads. There is one major benefit to these methods, however. Though the number of operations needed is the same, by breaking the problem up into many small pieces, one reduces the computer memory requirements. This change would allow one to calculate the conductance of longer systems.

2.1 The Green's Function

Before we delve into our method for finding the Green's function for a system, we should describe this object and its usefulness. The Green's function tells us the response of a system to a delta function source [14, Ch. 8]. For a general differential operator $\hat{F}$, used in the differential equation $\hat{F}f(x) = 0$ in a region without sources, the Green's function satisfies

$$\hat{F}G_F(x, x') = \delta(x - x'). \tag{2.1}$$

If we have a system governed by the differential operator $\hat{F}$ and we have a known source function $g(x)$, we can determine the solution $f(x)$ with that source by

$$f(x) = \int dx'\, G(x, x')g(x') \tag{2.2}$$

as can easily be verified. Our goal is to determine the Green's function for the scattering system; we can then use a given incident wave as the source, and determine the resulting scattered wave.

The differential operator with which we are interested in working is the time-independent Schrödinger equation in two dimensions,

$$\hat{H}(\vec{x})\psi(\vec{x}) = E\psi(\vec{x}), \tag{2.3}$$

or

$$[E - \hat{H}(\vec{x})]\psi(\vec{x}) = 0. \tag{2.4}$$

The Green's function is then a solution to

$$[E - \hat{H}(\vec{x})]G(\vec{x}, \vec{x}') = \delta(\vec{x} - \vec{x}'). \tag{2.5}$$

There are, in general, two possible solution to this equation: the so-called "advanced" and "retarded" Green's functions. We can select one or the other by the inclusion of an infinitesimal imaginary number $\pm i\eta$ into the differential equation,

$$[E - \hat{H}(\vec{x}) \pm i\eta]G^{\pm}(\vec{x}, \vec{x}') = \delta(\vec{x} - \vec{x}'). \tag{2.6}$$

Taking the upper signs gives us the physically acceptable retarded solution. From this point forward, whenever we refer to the Green's function G we will mean the retarded Green's function G^+.

Looking ahead a bit, we want a technique amenable to solution on a computer. This means that we will work with matrix representations of our operators. Taking the Hamiltonian to be a square matrix and I to be the corresponding identity matrix, the matrix equation representation of our differential equation is $[(E+i\eta)I-H]G = I$, or

$$G = [(E + i\eta)I - H]^{-1}. \tag{2.7}$$

2.2 The S-Matrix, Transmission Coefficients, and the Landauer Formula

The quantity frequently of interest for propagating wave problems is the conductance, primarily because this is the most easily accessible experimental quantity. We will now see the connection between the conductance and the scattering matrix, via the Landauer formula. We will then relate the S-matrix to the Green's function.

After the microwave research that was done circa World War II, much of the language for studying scattering became that of wave guides. That is, we speak of incident modes, or channels, and how a scattering event couples those incident channels into scattered channels. It is within this conceptual framework that the Landauer formula is useful, as it connects the transmission coefficients of these channels to the conductance measured through a system. Simply put, it is

$$G = \frac{2e^2}{h} \sum_i T_i, \tag{2.8}$$

where G here is the conductance (not the Green's function), the T_i are transmission coefficients for the incident channels, and the sum is over all channels. e^2/h is the conductance quantum, and the 2 is due to spin degeneracy. We will not derive this expression here (see [9] if you are interested).

Using Eq. (2.8), we get the conductance by calculating transmission coefficients. We can, in turn, determine transmission coefficients by looking for the S-matrix. The S-matrix tells us how a scattering potential connects incident channels to scattered channels, including both amplitude and phase shifts. If a system has n incident channels, we can specify a given incident wave by n amplitudes in a vector $\vec{a}$. Similarly, with m outgoing channels, we can specify the scattered wave with m amplitudes in a vector $\vec{b}$. The S-matrix would then be $m \times n$ and would connect the incident wave to the scattered wave by $\vec{b} = s \cdot \vec{a}$. The properties of the S-matrix reflect the symmetries of the physical system [9]. Indexing all incident channels and all scattered channels, the S-matrix element s_{ij} tells us the scattering into channel j of incident channel i. The transmission coefficient T_i is then

$$T_i = \sum_{j \in \substack{\text{transmission} \\ \text{leads}}} |s_{ij}|^2 . \tag{2.9}$$

We restrict the sum to look only at channels in the "transmission leads" so that we don't include elements s_{ij} corresponding to reflection.

2.3 Reducing the Infinite System to a Finite Matrix

To write a matrix equation for the Green's function, we first need a matrix representation for the Hamiltonian H. We find one by laying down a spatial finite difference grid. This is a regular rectangular spatial grid with spacings a_x and a_y. A notable benefit of the finite differencing basis is that the potential energy component of H is diagonal, requiring only the values of the potential at the grid points. We approximate the second derivatives in the Hamiltonian by

$$\partial_x^2 \psi(x,y) \approx \frac{\psi(x-a_x,y) - 2\psi(x,y) + \psi(x+a_x,y)}{a_x^2}, \tag{2.10}$$

and similarly for ∂_y^2. The approximation used for second derivatives in Eq. 2.10 results in only nearest-neighbor coupling in the kinetic energy component of H.

Our goal is a finite matrix representation of the Green's function in the system. We have, using Eq. 2.7 and the matrix for H generated by finite differencing, an infinite matrix representation for the Green's function (because the finite difference grid covering the leads and the system has an infinite number of grid points). However, we can write Eq. 2.7 symbolically in the form

$$\left(\begin{array}{c|c} G_l & G_{ls} \\ \hline G_{sl} & G_s \end{array}\right) = \left(\begin{array}{c|c} (E+i\eta)I - H_l & \tau \\ \hline \tau^T & (E+i\eta)I - H_s \end{array}\right), \tag{2.11}$$

where we have labeled elements for points in the system with a subscript s and those for points in the leads with a subscript l. We have arranged the rows and columns of the matrix such that we can partition it into sub-matrices dealing separately with the leads, the system, and the coupling between the two. The finite sub-matrix G_s, dealing with the system, is the piece of interest.

We now derive a matrix identity that we will need. Let's work in terms of a general matrix, then apply the result to our specific case. Assume that we have a square matrix A divided into sub-matrices

$$A = \begin{pmatrix} A_0 & A_1 \\ A_2 & A_3 \end{pmatrix}, \tag{2.12}$$

where A_0 and A_3 are themselves square. Let the matrix $B = A^{-1}$ be similarly subdivided, with sub-matrices B_0 and B_3 of the same dimensions as A_0 and A_3 respectively. Taking the product of A and B, we have

$$A \cdot B = \begin{pmatrix} A_0B_0 + A_1B_2 & A_0B_1 + A_1B_3 \\ A_2B_0 + A_3B_2 & A_2B_1 + A_3B_3 \end{pmatrix} = \mathbf{I}. \tag{2.13}$$

A little algebra with this result allows us to express elements of B purely in terms of elements of A. From the upper right-hand corner,

$$A_0B_1 + A_1B_3 = 0 \tag{2.14}$$

$$B_1 = -A_0^{-1}A_1B_3. \tag{2.15}$$

Combining this result with the lower right-hand corner, we have that

$$A_2B_1 + A_3B_3 = \mathbf{I} \tag{2.16}$$

$$-A_2A_0^{-1}A_1B_3 + A_3B_3 = \mathbf{I} \tag{2.17}$$

$$B_3^{-1} = A_3 - A_2A_0^{-1}A_1. \tag{2.18}$$

Similarly,

$$B_0^{-1} = A_0 - A_1A_3^{-1}A_2. \tag{2.19}$$

Though this result isn't profound, it does allow us to deal with parts of a matrix in isolation from one another, and in particular to perform the inversions separately. For our system, the sub-matrix in Eq. (2.11) dealing with the leads, $(E+i\eta)I-H_l$, has an infinite number of elements in it. The sub-matrix dealing with the system, $(E+i\eta)I-H_s$, has only a finite number of elements. Because we have defined the division between leads and system such that all scattering occurs in the system, the solutions in the leads are traveling waves in constant transverse modes; with this simplification, we can determine the inverse

$$g_l \equiv [(E+i\eta)I-H_l]^{-1} \tag{2.20}$$

analytically; g_l is the combined Green's function for the semi-infinite leads. We then define the matrix

$$\sigma \equiv \tau^T g_l \tau, \tag{2.21}$$

in terms of which

$$G_s^{-1} = EI - H_s - \sigma. \tag{2.22}$$

Note that we have dropped the term $i\eta$ in this expression; it is no longer relevant, as the matrix σ is itself complex and swamps the infinitesimal $i\eta$. Note that σ is only non-zero for points in the system adjacent to the leads, as the finite differencing basis only couples nearest neighbors.

The matrix σ can be divided into individual matrices for each lead, $\sigma=\sum_p \sigma^p$. Let the transverse modes in lead p be $\chi_m^p(y)$. We have that

$$\sigma^p(y_i,y_j) = -\frac{\hbar^2}{2ma_x^2}\sum_m \chi_m^p(y_i)e^{ik_m^p a_x}\chi_m^p(y_j), \tag{2.23}$$

where the y_i and y_j are the y-coordinates of grid points i and j, and k_m^p is the wave number for a wave propagating in mode m of lead p [15, §3.5]. Eq. 2.23 is derived by first using the basis function expansion of the Green's function for a semi-infinite lead, then discretizing the resulting equation. We have written Eq. 2.23 for y as the transverse direction. It can easily be altered to have x as the transverse direction by exchanging x's and y's throughout the expression.

To find the wave vectors k_m^p, we note that the dispersion relation in the approximation of finite differencing gives us

$$E = \epsilon_m^p + \frac{\hbar^2}{ma_x^2}[1 - \cos(a_x k_m^p)], \tag{2.24}$$

with ϵ_m^p the energy of the transverse mode χ_m^p. In the limit $a_x \to 0$, Eq. 2.24 becomes the familiar free-space dispersion relation. Also important is the expression for the velocity of propagation for a wave in one of these modes, $\hbar v = \partial_k E$, which gives us

$$v_m^p = \frac{\hbar}{ma_x}\sin(a_x k_m^p) = \frac{\hbar}{ma_x}\left[1 - \left(1 - \frac{E - \epsilon_m^p}{\hbar^2/ma_x^2}\right)^2\right]^{1/2}. \tag{2.25}$$

2.4 Using G_s^{-1}

Given a finite matrix representation of the Green's function, we can find the elements of the S-matrix needed in the Landauer formula. We do so by looking once again at the transverse lead modes, which define our scattering channels.

Consider an incident wave in channel i impinging on the scattering system. This wave gives us a source term along the edge of the system, allowing us to apply the Green's function to determine the response of the system. We can then project that response onto the scattered modes to determine S-matrix elements. In symbols,

$$s_{ij} = \hbar(-v_i v_j)^{1/2} \int dy \int dy' \, \psi_j^{\text{out}}(y) G_s(x_r, y; x_l, y') \psi_i^{\text{in}}(y'), \tag{2.26}$$

where x_r gives the coordinate of the edge of the transmitting lead and x_l gives the coordinate of the edge of the incident lead. The terms appearing outside of the integrals in this equation are necessary because the S-matrix deals in flux, whereas the integrals deal with amplitude [15]. In terms of matrices, representing the incident and scattered states as vectors, we have

$$\phi_i \equiv G_s \cdot \psi_i^{\text{in}} \Rightarrow s_{ij} = \hbar(-v_i v_j)^{1/2} \phi_i \cdot \psi_j^{\text{out}}. \tag{2.27}$$

Computationally, inverting a large matrix is costly. We choose, therefore, to solve the problem in a manner slightly different from that implied above. We can determine G_s^{-1} using the methods described above; rather than inverting it, we solve the matrix equation $G_s^{-1}\phi_i = \psi_i^{\text{in}}$ with ϕ_i as the unknown. There are computational packages available to efficiently solve this "$Ax = b$" type problem, exploiting the sparseness of G_s^{-1}.[1]

2.5 Extension to Multiple Transmission Leads

Conceptually, the way that we have described this method above is the easiest way to understand it: a single incident lead and a single transmission lead, all in a linear system. However, this method allows for a further expansion to include leads

---

[1] In this work, I used the package "SuperLU." It performs an LU decomposition of the matrix, using row interchanges to preserve sparsity in the decomposition. This allows me to solve for multiple right hand sides without significant additional effort, and thus extract the complete S-matrix in one step.

on the sides of the system. The mathematical extension is easy; the additional leads and their transverse states just couple to other points in the matrix.

Though one can attach additional leads, one should not confuse this with the idea of adding “terminals.” Many experiments on 2DEGs fall generally into the categories: “two terminal” and “four terminal.” In a two-terminal experiment, one applies a bias voltage and measures the current between two leads. In a four-terminal experiment, one passes a current between two leads and measures a voltage difference across two different leads. We are clearly performing a two-terminal numerical experiment with our formalism.

One has much more freedom with the leads than we have explicitly stated so far. For example, none of the leads needs to cover an entire side of the system. More importantly, it is not actually a requirement that the leads be of the form described here, semi-infinite rectangular devices with constant transverse modes. As long as one can determine the Green’s function for the leads and the coupling terms between the leads and the system, one can find the matrix σ in Eq. 2.21 and hence G_s^{-1}.

2.6 An Example

The formulae given above are complete, but without seeing their application they are somewhat abstract. In fact, the implementation of the formulae can be somewhat difficult to carry out, as there are many indices to keep straight. We present here an example with side leads, and to be general we let those side leads have a width different than that of the system. To keep the formulae somewhat tractable, we’ll take the system to be discretized into a three by three grid. This setup is shown in

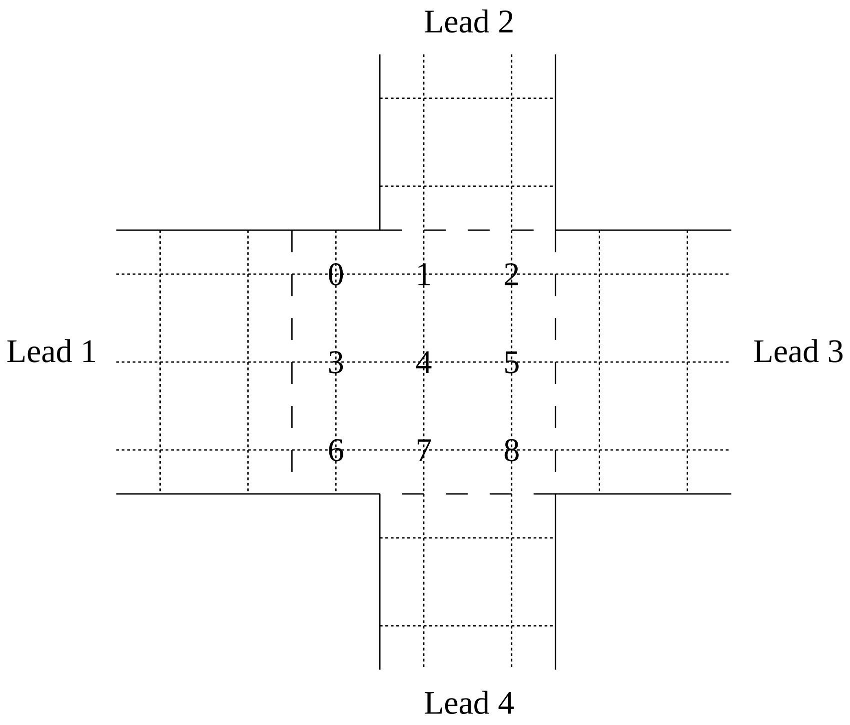


Figure 2.2: Setup for the example in §2.6. The system is discretized into a three-by-three grid, with the nine grid points as labeled. There are four leads on the system, with 1 and 3 three points across and 2 and 4 only two points across. The dashed lines indicate the borders between the system and the leads.

Figure 2.2.

The easiest thing to look at is the potential. Let there be a general potential V in the system. As mentioned above, the only information that we require about this potential is its value at the grid points. We represent the potential as the nine-by-nine diagonal matrix

$$V_s = \begin{bmatrix} V[0] & 0 & 0 & \cdots & 0 \\ 0 & V[1] & 0 & \cdots & 0 \\ 0 & 0 & V[2] & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & V[8] \end{bmatrix}, \tag{2.28}$$

where $V[i]$ indicates the value of the potential at the grid point i.

The next component to consider is the kinetic energy. Let us begin by defining the values

$$t_{x,y} \equiv \frac{\hbar^2}{2ma_{x,y}^2}. \tag{2.29}$$

Using the result in Eq. 2.10, we can express the kinetic energy as

$$\begin{aligned} K &= \frac{p^2}{2m} && (2.30)\\ &= -\frac{\hbar^2}{2m}\left(\partial_x^2 + \partial_y^2\right) && (2.31)\\ &= -t_x\left[\psi(x-a_x,y) - 2\psi(x,y) + \psi(x+a_x,y)\right] \\ &\quad -t_y\left[\psi(x,y-a_y) - 2\psi(x,y) + \psi(x,y+a_y)\right]. && (2.32) \end{aligned}$$

Translated into a matrix formulation, this gives us that

$$K_s = (2t_x + 2t_y)I - \begin{bmatrix} 0 & t_x & 0 & t_y & 0 & 0 & 0 & 0 & 0 \\ t_x & 0 & t_x & 0 & t_y & 0 & 0 & 0 & 0 \\ 0 & t_x & 0 & 0 & 0 & t_y & 0 & 0 & 0 \\ t_y & 0 & 0 & 0 & t_x & 0 & t_y & 0 & 0 \\ 0 & t_y & 0 & t_x & 0 & t_x & 0 & t_y & 0 \\ 0 & 0 & t_y & 0 & t_x & 0 & 0 & 0 & t_y \\ 0 & 0 & 0 & t_y & 0 & 0 & 0 & t_x & 0 \\ 0 & 0 & 0 & 0 & t_y & 0 & t_x & 0 & t_x \\ 0 & 0 & 0 & 0 & 0 & t_y & 0 & t_x & 0 \end{bmatrix}. \tag{2.33}$$

The matrix for the Hamiltonian in the system is then just $H_s = K_s + V_s$.

We now turn our attention to the leads. We need to determine the transverse modes in the leads, using the same discretization used in the system. Here, however, we have only to worry about one dimension. Looking at lead 1, which will later be used as the lead for our incident waves, we again write the Hamiltonian matrix as the sum of a potential energy matrix and a kinetic energy matrix:

$$H_1 = \begin{bmatrix} V[0] & 0 & 0 \\ 0 & V[3] & 0 \\ 0 & 0 & V[6] \end{bmatrix} + \begin{bmatrix} 2t_y & -t_y & 0 \\ -t_y & 2t_y & -t_y \\ 0 & -t_y & 2t_y \end{bmatrix}. \tag{2.34}$$

Note the choice that we have made for the potential in the lead: we have matched that transverse potential to the potential at the edge of the system. Though we are by no means required to make this choice, if we do not then there will be (artificial) reflections off of the discontinuity in the potential.

This three-by-three matrix then represents the one-dimensional Hamiltonian for the transverse leads in the system. Solving for the eigenvectors and eigenvalues of the matrix, we get the set $\{\chi_i^1, \epsilon_i^1 : i \in [0, 2]\}$ of eigenvectors and eigenvalues. Given the energy of our incident wave, Eq. (2.25) gives us additionally a set of velocities $\{v_i^1 : i \in [0, 2]\}$.

Lead 3 looks very similar to 1 as treated above, so let us look now at one of the side leads. We have that

$$H_2 = \begin{bmatrix} V[1] & 0 \\ 0 & V[2] \end{bmatrix} + \begin{bmatrix} 2t_x & -t_x \\ -t_x & 2t_x \end{bmatrix}. \tag{2.35}$$

The solution to this one-dimensional eigenvalue problem gives us the set $\{\chi_i^2, \epsilon_i^2 : i \in [0, 1]\}$, and from these we get $\{v_i^2 : i \in [0, 1]\}$. Note, however, that since this lead has x as the transverse direction we switch x's and y's in our equations.

We won't write out all of the terms of the self-energy matrix, as that would get quite lengthy. Let's look at one of the leads, however, to see where it will contribute. Recall Eq. (2.23) for the self-energy matrix contribution of a single lead,

$$\sigma^p(y_i, y_j) = -\frac{\hbar^2}{2ma_x^2} \sum_m \chi_m^p(y_i) e^{ik_m^p a_x} \chi_m^p(y_j). \tag{2.36}$$

Looking at lead 1, the y_i and y_j will be replaced by the y-coordinates of points 0, 3, and 6 in the system. Or, rather, we take $\chi_m^1(y_i) = \chi_m^1[i]$. The self-energy matrix will couple the three points to one another in all possible combinations. Looking at the full nine-by-nine self energy matrix, we see that the leads contribute to the elements indicated below:

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2,3 & 0 & 0 & 3 & 0 & 0 & 3 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & 3 & 0 & 0 & 3 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 4 \\ 0 & 0 & 3 & 0 & 0 & 3 & 0 & 4 & 3,4 \end{bmatrix} . \tag{2.37}$$

We now have all of the information needed to write our inverse Green's function, $G_s^{-1} = EI - H - \sigma$. To apply it, we look to the $Ax = b$ problem $G_s^{-1}\phi_i = \psi_i^{\text{in}}$. What, however, are we to take as the incoming waves? Since G_s^{-1} is a nine-by-nine matrix, we need both ϕ_i and ψ_i^{in} to be nine-element vectors. Let us begin by ordering all of the transverse lead modes: $\{\chi_0^1, \chi_1^1, \chi_2^1, \chi_0^2, \chi_1^2, \chi_0^3, \chi_1^3, \chi_2^3, \chi_0^4, \chi_1^4\}$. We get to the ψ_i vectors by "padding" these eigenstates with zeros. For example, for the first incident mode we have that

$$\psi_0^{\text{in}} = \left[\chi_0^1[0], 0, 0, \chi_0^1[1], 0, 0, \chi_0^1[2], 0, 0\right] . \tag{2.38}$$

In the elements of ψ_0^{in} that correspond to points adjacent to lead 1, we take the value of χ_0^1 at those adjacent points; for all other points in the system, we take zeros. If we are interested in an outgoing wave in the second state of lead 2, we would write

$$\psi_4^{\text{out}} = \left[0, \chi_1^2[0], \chi_1^2[1], 0, 0, 0, 0, 0, 0\right] , \tag{2.39}$$

using similar logic.

Using these vectors, we can solve first $G_s^{-1}\phi_0 = \psi_0^{\text{in}}$. The S-matrix element is then found by $s_{04} = \hbar(-v_0^1 v_4^2)^{1/2}\phi_0 \cdot \psi_4^{\text{out}}$. To get the full conductance of the system, we will need to determine $s_{ij} : i \in [0,2],\ j \in [3,9]$. Note that this only means solving three $Ax = b$ problems.

2.7 Magnetic Field

Up to this point, we have been assuming that there is no magnetic field present. Of course, there are many experiments where a magnetic field is present, perhaps as the only available independent parameter (e.g., [16]). One can certainly handle a magnetic field in Schrödinger's equation, though there are complications in including one in our calculations. Primarily, one has to deal with the presence of the leads. The presence of a magnetic field in the leads changes a number of the properties of the leads upon which we have depended. For example, unless one is careful, the lead modes may no longer be described by plane waves propagating in fixed transverse modes. Even if they are, then the transverse modes will in general no longer be orthogonal [15], and they will be different for incoming and outgoing waves.

If we want to take the magnetic field to zero in the leads, we need to have a vector potential that points along each lead at that lead. For a rectangular system, that means that $\vec{A}$ must be perpendicular to each wall of the system at that wall. We can then reduce $\vec{A}$ to zero in the leads without introducing any curl, and hence without any spurious magnetic fields [17]. In fact, we can reduce $\vec{A}$ to zero with a step function at the edges of the system, and hence have a constant magnetic field in the system and zero field in the leads.

We have the additional constraint that we should use a continuously defined vector potential. If we use a piecewise-defined $\vec{A}$, there will be false reflections off of the discontinuities, even if the various pieces correspond to the same magnetic field. Following [18], we find a continuous vector potential that satisfies our boundary orientation requirement and gives us a constant magnetic field.

Let us begin with a vector potential $\vec{A}_0(\vec{x}) = By\hat{x}$. This gives us our constant magnetic field, and it is perpendicular to the walls at the left and right sides of our system. We note that the gauge transformation generated by

$$f(\vec{x};\theta) = -Bxy\sin^2(\theta) + \frac{1}{4}B(y^2 - x^2)\sin(2\theta) \tag{2.40}$$

will rotate the vector potential through the angle θ without changing the resulting magnetic field when we take

$$\vec{A}(\vec{x}) = \vec{A}_0(\vec{x}) + \nabla f(\vec{x};\theta). \tag{2.41}$$

We do not want to change the orientation of the vector potential over all of the system, however, as it is already correct along two walls. We therefore need to create a mask for this gauge transformation, one that goes to one along the top and bottom walls and to zero along the left and right walls. Taking the dimensions of the system to be w_x and w_y, we can take

$$m(\vec{x}) = \left(\frac{w_y}{y} + \frac{w_y}{w_y - y}\right)\left(\frac{w_y}{y} + \frac{w_y}{w_y - y} + \frac{w_x}{x} + \frac{w_x}{w_x - x}\right)^{-1}. \tag{2.42}$$

This isn't the prettiest of functions, but it can be re-cast to a form that doesn't rely on canceling infinities to remain finite; this is merely the simplest representation. We then take

$$\vec{A}(\vec{x}) = \vec{A}_0(\vec{x}) + \nabla\left[f(\vec{x};\pi/2)m(\vec{x})\right]. \tag{2.43}$$

Sparing the algebra, we find that vector potential has the components

$$\begin{aligned} A_x &= \frac{Bw_x^2(w_y - y)y^2(w_y^2x^2 + w_x^2w_yy - w_x^2y^2)}{[w_y^2(w_x - x)x + w_x^2(w_y - y)y]^2} && (2.44)\\ A_y &= -\frac{Bw_y^2(w_x - x)x^2(w_x^2y^2 + w_y^2w_xx - w_y^2x^2)}{[w_y^2(w_x - x)x + w_x^2(w_y - y)y]^2}. && (2.45)\end{aligned}$$

It can be verified that this vector potential gives a constant magnetic field in the system and, despite being continuously defined, is perpendicular to each of the walls. One might note that it is undefined at the corners of the system. While this is true, we will see shortly that we only need to evaluate the function between grid points, not at them, so the corners are avoided.

Now that we have a vector potential to apply, let us look at what we will do with it. The vector potential enters the equations as a modification of the coupling terms between lattice sites. The coupling is modified by the exponential

$$\exp\left\{\frac{ie}{\hbar}\int_a^b d\vec{x}\cdot\vec{A}(\vec{x})\right\}, \tag{2.46}$$

where we integrate from point a to point b [19, Vol. 3, p. 21-2]. We will estimate this integral by taking

$$\int_a^b d\vec{x}\cdot\vec{A} \approx (\vec{x}_b - \vec{x}_a)\cdot\vec{A}\left[(\vec{x}_a + \vec{x}_b)/2\right]. \tag{2.47}$$

To introduce this coupling to the Hamiltonian in the most general way, we have

$$H_{ij}(A) = H_{ij}(0)\exp\left\{\frac{ie}{\hbar}(\vec{x}_i - \vec{x}_j)\cdot\vec{A}\left[(\vec{x}_i + \vec{x}_j)/2\right]\right\}. \tag{2.48}$$

We note first that, because of our choice to work in the finite differencing basis, $H_{ij}(0)$ is only non-zero on the diagonal and for nearest-neighbor grid points. This fact is not changed by the vector potential term, so the sparseness of our matrix problem isn't affected. Second, we note that terms on the diagonal aren't altered by the magnetic field.

The simplest exploitation of this gauge transformation, turning the magnetic field off with a step function at the edges of the system, is not ideal. The gauge allows us to do this without introducing additional fields, but we do have a problem with a transverse-mode mismatch at the edges of the system. As a result, there are some reflections at the edges that have no physical meaning. The effect appears to be negligible as long as there is any physically significant scattering going on at the same time (i.e., it is only observable if $V(\vec{x}) = 0$). A more rigorous solution, if one is intent on avoiding magnetic field in the leads, would be to apply this gauge and then include a length of each lead in the domain, using these lengths to decrease the vector potential to zero [18].

Chapter 3

Electron Flow in Simple 2DEG Devices

A good deal of wave propagation studies in the past decade, both theoretical and experimental, have centered on the behavior of mesoscopic devices and two-dimensional electron gasses [20]. A two-dimensional electron gas, or 2DEG, is formed when a crystal heterostructure (typically GaAs/AlGaAs) is made such that free electrons feel a strongly confining potential in one direction [21]. The electrons so confined are effectively reduced to two degrees of freedom. Devices are formed either by etching away parts of the crystal or by capacitively coupling "gates" to the 2DEG and reducing the electron density in areas. Current is then passed through the device, and conductance can be measured as a function of various external parameters (e.g., magnetic field).

In this chapter, we begin to discuss the research that has been done in collaboration with experimentalists in the Westervelt group [22, 23]. These experiments, and the corresponding theory, have looked at the spatial structure of electron flow in a 2DEG. The independent parameter in the experiments is the position of an atomic-force microscope (AFM) tip above the sample, giving spatial resolution to the experiments. We will also touch on other kinds of experiments that can be simulated with these numerical methods.

3.1 The QPC

The prototypical 2DEG device is the quantum point contact (QPC) [24, 25, 26]. Not only are they the simplest devices (other than a completely open system), but they are also components of most more complex devices [27, 28]. This simple device is where we begin.

3.1.1 Smooth Point Contact Model

Before we can do any modeling of the conductance experiments, we need a model of the QPC. As described in Chapter 2, the technique that we will use to find the conductance through the 2DEG requires only that we be able to specify the value of the potential at every grid point. Hence we can define a potential function $V(\vec{r})$ in any way that we wish. To mirror the physical system as closely as possible, however, we construct a smooth, analytic function to model the point contact potential.

Let us choose V such that $V(x, y) = f(x)y^2$. That is, at any fixed x value the transverse potential is parabolic. This is reasonable physically and aids our intuition by making the solutions analytic in one dimension. Since we also wish to make the potential symmetric about the center of the point contact, we can write it in the form

$$V(x, y) = \frac{my^2\omega(x)^2}{2} \tag{3.1}$$

with $\omega(x)^2$ an even function, making the relationship to analytic solutions in the y direction as explicit as possible. We now need to choose an $\omega(x)$; in this work, let us take

$$\omega(x) = \begin{cases} \omega_0\,(1 - 3|2x/l|^2 + 2|2x/l|^3) & |2x| < l \\ 0 & |2x| > l \end{cases}, \tag{3.2}$$

where l is a parameter that determines the "length" of the point contact, and ω_0 determines the "width." We take

$$\omega_0 = \frac{E_F}{\hbar(N - .5)}, \tag{3.3}$$

where N gives the number of open modes for the QPC. This potential has the advantages of being as adiabatic as we wish (via the parameter l), going to zero explicitly beyond a known region, and having an explicitly controllable number of open transverse modes. Figure 3.1 shows contours of this potential function.

The quintessential feature of point contact conductance is quantization [24]. This quantization manifests itself as steps in the conductance as a function of the voltage applied to the gates creating the QPC. Relating this to my computational model, it would be quantized steps in conductance as a function of the parameter ω_0, made very explicit by the formulation with N. Though this aspect of QPC behavior is not our primary focus, it does serve as a useful verification that our method and model are working.

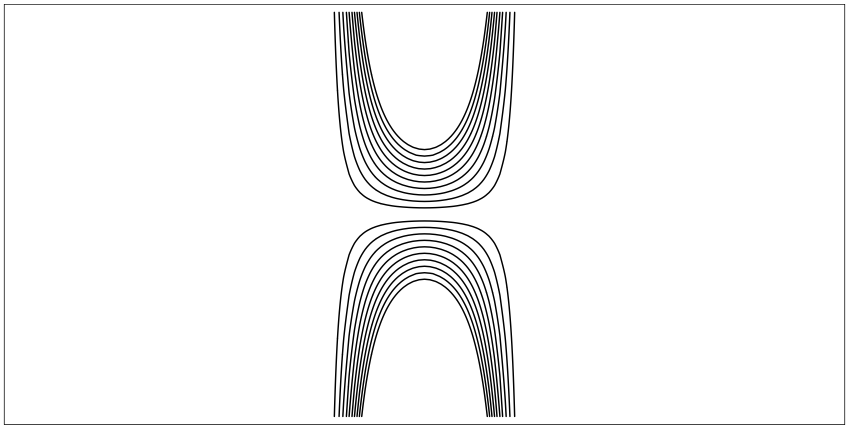


Figure 3.1: Sample contours of the potential function used in this work to model the QPC.

The origin of the conductance steps is quite simple to understand. Recall the Landauer formula, Eq. 2.8, which gives conductance as a sum of transmission coefficients. In the adiabatic limit, there is no scattering between modes as the leads narrow into the point contact; that is, as the transverse modes evolve smoothly as a function of x, a wave in the ith state remains in the ith state. The wave is either reflected in the same mode, or propagates through the QPC. Ignoring tunneling and above-the-barrier reflection (which are also suppressed in an adiabatic point contact), this means that every transmission coefficient T_i is either zero or one. As the point contact opens up, the energies of the transverse modes decrease; as each mode's energy passes the Fermi energy, the corresponding transmission coefficient changes from zero to one and the conductance increases by $2e^2/h$.

In the adiabatic limit, all steps would be sharp and of the same height. For a real potential, we need to deal with both tunneling and mixing of modes. Both effects smooth the conductance steps. In Figure 3.2, we see the steps for our model for one choice of parameters. Note that, whereas the initial steps are quite distinct, the latter

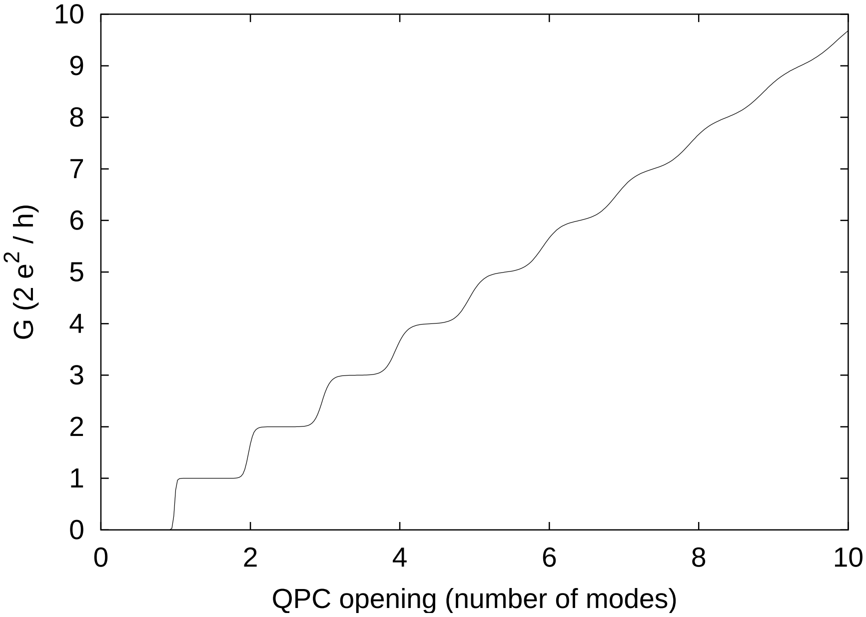


Figure 3.2: A demonstration of the quantized conductance shown by my computational model. For the first several modes, the steps are fairly sharp, though tunneling does generate some rounding. For higher modes, the effects of mode mixing and tunneling overwhelm the regular step pattern. The number of modes that show sharp steps can be changed by changing the adiabaticity (the "length" parameter) of the QPC.

steps are blurred by the non-idealities.

3.1.2 Experimental Method

Now that we have a numerical QPC that can be compared to experiments, we turn to the experimental setup. For simple conductance quantization, the independent parameter is the voltage on the QPC gates. The potential landscape induced in the 2DEG by the gates is static in a sense: though the voltage on the gates can be changed, the gates themselves cannot be moved. Recent advances in experimental technique and equipment have made a spatially resolved measurement possible.

A negatively charged atomic force microscope (AFM) tip can be held above the sample, creating a movable bump in the potential seen by electrons in the 2DEG. The potential induced by the AFM tip is a circularly symmetric Lorentzian bump, centered below the tip [29]. The height of the Lorentzian is determined by the voltage of the tip, and its width by the distance from the tip to the 2DEG. The AFM tip can be moved, and conductance can be measured as a function of tip position. Though the exact connection between this measurement and the wave function of the transmitted electrons needs to be established, it does at least hold the promise of imaging the electron flow; we would expect some dependence on $|\psi|^2$ for the amplitude of the AFM's effect. See Figure 3.3 for a schematic of the setup.

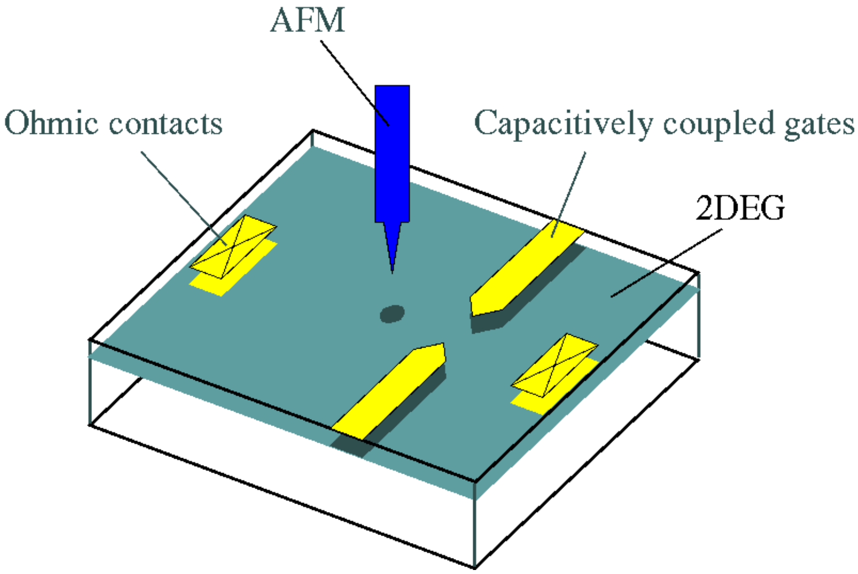


Figure 3.3: In this schematic of a 2DEG device, we see the basic setup for our 2DEG conductance measurements. The ohmic contacts are directly coupled to the 2DEG, allowing a bias voltage to be applied to the system and current measured. The gates, here in a quantum point contact configuration, are patterned on top of the crystal and coupled capacitively to the 2DEG. Their effect can be changed by varying the bias voltage applied between them and the 2DEG. Finally, an atomic force microscope (AFM) tip, also capacitively coupled to the 2DEG, can be moved above the sample.

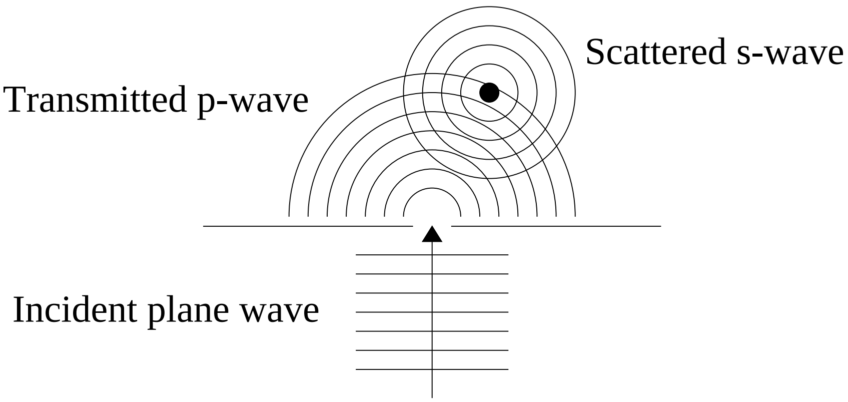


Figure 3.4: A schematic for the simple model under consideration in §3.1.3. A plane wave incident on a small slit in a wall gives rise to a transmitted p-wave. A single point scatterer gives rise to an outgoing s-wave. We consider the interaction of these two waves at the slit to determine flux.

3.1.3 A Simple Model

Before going into the computational predictions and experimental results, it is helpful to look at a simple model to build intuition about what we might expect to see as we move the AFM tip about. The model presented here is not meant to give any quantitative predictions of the results, but rather merely to alert us to some phenomena for which to look.

A schematic of the model is given in Figure 3.4. We consider a plane wave incident on a small slit in a screen. Satisfying the Dirichlet boundary conditions at the screen and taking only the lowest order term (the long wavelength or small slit limit) for the transmitted wave, the transmitted wave is a p-wave. This slit plays the role of the point contact in the model. We then introduce a single point scatterer to play the role of the Lorentzian potential under the AFM tip.

Scattering theory gives us the full transmitted wave by

$$\psi(\vec{r}) = \phi(\vec{r}) + s(k)G(\vec{r},\vec{r}_s)\phi(\vec{r}_s), \tag{3.4}$$

where G is the Green's function, ϕ is the incident wave, $\vec{r}_s$ is the location of the point scatterer, and s is an energy-dependent amplitude (in two-dimensions, $s(k) = -(8\pi k)^{1/2} \exp[i\pi/4] f(k)$ for scattering amplitude $f(k)$ [30, Eq. (2.11)]). We are interested in determining the flux through the slit, with the flux found from the full ψ by

$$\vec{j} = \frac{\hbar}{2im}(\psi^* \nabla \psi - \psi \nabla \psi^*). \tag{3.5}$$

The solution for the full ψ requires the Green's function for the system with a screen and a slit, which is not known analytically. Given that we are interested only in the flux at the slit, however, we can use the Green's function for a simple plane,

$$G(\vec{r}, \vec{r}') = -\frac{i}{4} H_0(k|\vec{r} - \vec{r}'|), \tag{3.6}$$

where H_0 is an outgoing Hankel function [30, Eq. (2.10)]. We take the incident ϕ as the p-wave. If we use this Green's function, the wave coming off of the point scatterer won't satisfy the boundary conditions on the screen. Since we only look at the wave at the slit and are not interested in higher order scattering, this approximation is sufficient.

In Figure 3.5, we show the results with this model. The important result is that, when plotting the change in conductance as a function of scatterer position, we see regular fringes spaced at half the incident wavelength. The incident wave alone has no intensity fringes, and the full wave for any given scatterer location does not have simple radial fringes.

What do we intend to take away from this simple model as we look at real systems? First, that the features seen in the conductance need not be present in the wave being

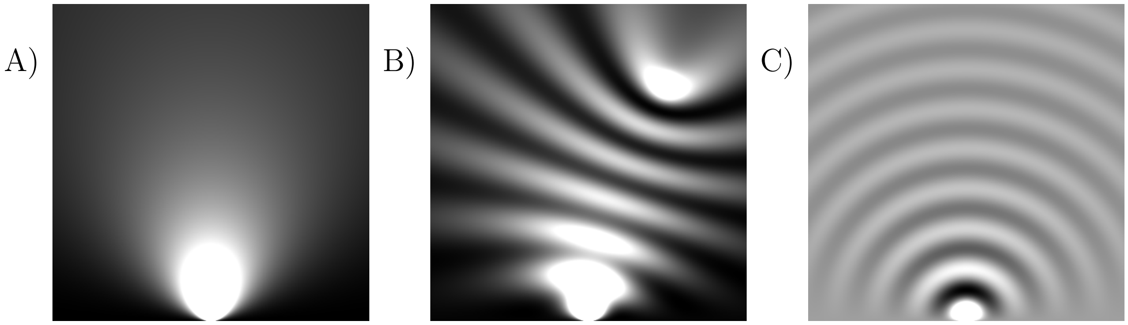


Figure 3.5: These are results calculated for our simple scattering model of QPC conductance. Each image is four wavelengths square. In (A), we show $|\phi|^2$, the incident wave. Note that there are no features in this wave, other than the angular dependence. In (B), we show $|\psi|^2$ for a single location of the scatterer. The resulting wave does have features to it, as the incident wave interferes with the scattered wave. In (C), we show the change in flux as a function of scatterer position (i.e., the intensity at a point is determined by the change in flux through the slit when the scatterer is at that point). In this image, we see the influence of $|\phi|^2$ in the angular dependence, and of interference in the half-wavelength spaced fringes.

probed. As seen in this demonstration, the fringes seen in the scan plot need not be representative of fringes in $|\psi|^2$. Second, we expect to see features on the scale of a half wavelength. Finally, we expect any observed features to have an envelope that is related to $|\psi|^2$.

3.1.4 Tip Scans

The experiments that we are interested in modeling scan the AFM tip over a two-dimensional grid of points and look at conductance as a function of tip position. Given the simple model presented above, we would expect to see a number of things in these results. There may be fringes in the results, with a separation of one half wavelength, as well as an overall envelope due to the spatial profile of the electron flow being imaged. For these effects to show up in experimental data, the experiments must satisfy a number of conditions; among them, that the coherence length in the material is sufficiently long and that the electrons are sufficiently monochromatic. Computationally, we have a single, well-defined energy and infinite coherence length, so these issues do not appear explicitly.[1]

We have already seen the functional form of the QPC potential, but have yet to specify the form of the potential induced by the AFM tip. It is a Lorentzian with variable height and width [29]. The height of the Lorentzian is controlled by the (negative) voltage of the AFM tip. The half-width of the Lorentzian is equal to the distance of the tip from the 2DEG. As a result, the experimental setup imposes a minimum value on this parameter. In the experiments under consideration, the distance from tip to 2DEG has a minimum near 60 nm due to the crystal above the 2DEG. The Fermi wavelength in the system is, in comparison, near 40 nm. This means that the point scatterer used in §3.1.3 is a poor representation of the scattering center created by the tip. With a half-width of $1.5\lambda_f$, the Lorentzian is considerably grosser than the features that we hope to image with it.

In Figure 3.6, we see the computational results of a two-dimensional Lorentzian scan for a point contact operating on the first conductance plateau. These data do show the fringes at half-wavelength intervals, as expected. Note that these data don't begin until 60 nm from the QPC. If the tip is any closer to the point contact, the tail of the Lorentzian extends into the contact, raises the potential at its center, and closes it off. One can see the beginning of this "cross-talk" effect at the left of the scan.

The fringes have two noticeable characteristics. The first is the concentric fringes,

---

[1] To simulate the finite temperature of the experiments, we can perform calculations at a number of energies near the Fermi surface and integrate over the thermal distribution.

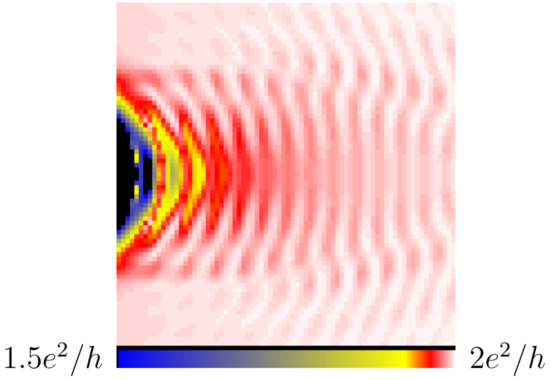


Figure 3.6: These are computational results for the conductance of a QPC as a function of Lorentzian position. The scan region begins on the left 60 nm away from the QPC, and it is 300 nm on a side. The scale bar is linear, but clearly the colors are not applied linearly; to bring out the smaller features over most of the image, we concentrate the color map near $2e^2/h$. There are three things to point out in the image: the region to the left where we can't see much because the AFM is closing off the QPC, the predicted concentric rings in the conductance measurement, and the variation in those rings caused by waves reflecting off of the sides of the QPC.

reminiscent of the simple model with a point scatterer and a slit. The second is the transverse variations of these fringes. This can be explained as the effect of second-order closed orbits. In Chapter 5, we will discuss fringe formation in detail.

These fringes also appear in the experimental data, a sample of which is shown in Figure 3.7. Here, as well, the data begins some distance from the QPC, for the same reason given for the computational results. The data isn't as "clean" as that from the calculations, due to other, uncontrolled features of the potential landscape in the 2DEG. Nevertheless, the qualitative agreement is evident. Quantitatively, the computations agree with both the spacing of the features and their relative magnitude.

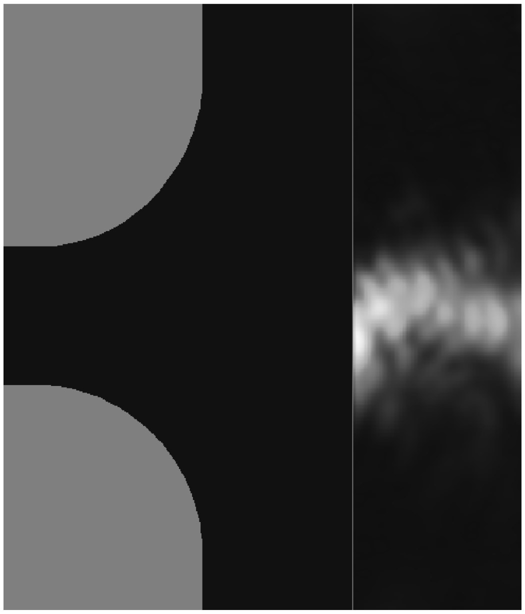


Figure 3.7: These data are from an experiment and are courtesy of the Westervelt group. We show conductance as a function of AFM position, with the QPC on the first conductance plateau. A schematic of the point contact gate is included. No data are shown in the region where the AFM tip closes the point contact. Fringes are separated by $\lambda/2$ ($\approx 20\,\mathrm{nm}$)

3.1.5 Theory of the AFM Tip Scattering

The demonstrated possibility of measuring half-wavelength scale features with our large Lorentzian begs the question of how this is possible. Our first hint comes from something first observed experimentally: the measured features disappear as the height of the Lorentzian is decreased. The experimental control of, and knowledge about, the height of the Lorentzian isn't great, but it was observed that halving the voltage on the tip (and thus halving the height of the Lorentzian) caused the features to go away. Computationally, of course, we have much finer control over the parameters of the potential. In Figure 3.8, we look at conductance as a function of tip position for a series of peak heights, starting at $1.2\,E_f$ and ending at $.8\,E_f$. We see that the features, which begin on the order $.05\,G$, go away as we decrease the Lorentzian height. So clearly, it is something about the presence of a classical turning point on the Lorentzian potential that is critical in changing the conductance of the system.

Another way of testing the dependence on the strength of the tip potential allows for a direct comparison with the experiment. Fixing the voltage on the AFM and changing the distance between it and the 2DEG has the effect of decreasing the strength of the potential that it generates. The experimentalists set up their system and selected an area far enough from the point contact that they could take measurements without the AFM affecting the QPC itself. They then performed two-dimensional scans of the AFM tip and averaged the resulting signals, obtaining a different average value for different AFM heights. We can run a simulation where we reproduce the experimental parameters and perform the same measurements and

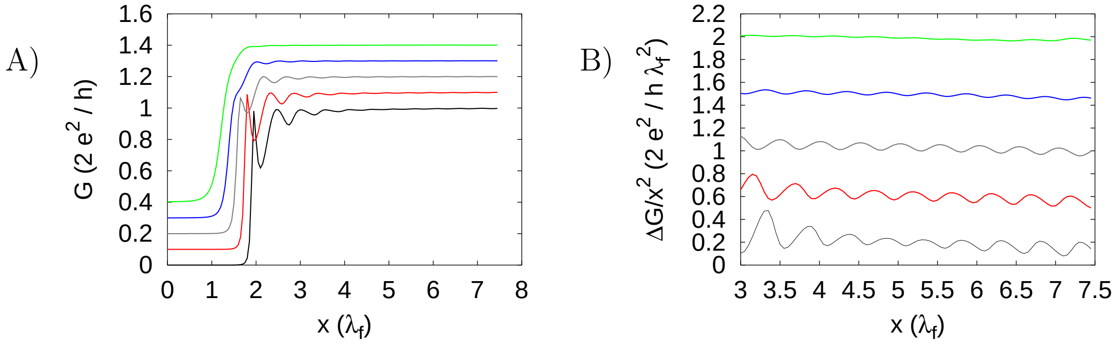


Figure 3.8: These plots show conductance as a function of tip position for a series of tip peak heights. The center of the Lorentzian is located on the $y = 0$ axis and scanned in the x direction, beginning at the center of the QPC. Successive curves are given a vertical offset so that they can all be viewed (the apparent horizontal shift is an artifact of changing tip influence on the QPC). In all cases, the half-width of the Lorentzian is 1.5λ. The highest curve has the AFM tip potential at approximately $.8\,E_f$ and the lowest curve has it at approximately $1.2\,E_f$. In (A), we show the conductance, unmodified. In (B), we show a modified version of the same data, where we show $((2e^2/h) - G)x^2$. The reason for this modification is that the x^{-2} geometric falloff of the signal makes it hard to see what is going on for the latter part of the curves in (A). In both (A) and (B), we note the disappearance of features as the peak of the Lorentzian dips below the Fermi energy.

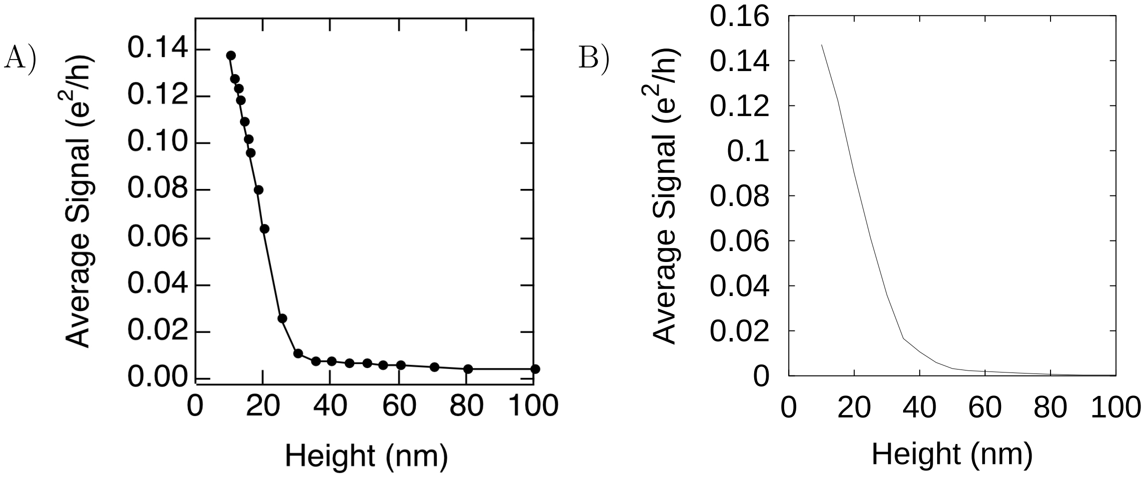


Figure 3.9: Experimental and computational data ((A) and (B), respectively) for the average signal strength as a function of AFM height. The agreement is good, though we note that noise comes to dominate the experimental signal at around .01 e^2/h. In the computational work, the peak of the AFM tip potential equals the Fermi energy at a tip height of 25 nm. This is clearly in the range where the signal is disappearing, and the rounding can be understood as follows. In both theory and experiment, there is a smooth disordered background potential present (discussed in Chapter 4). When the AFM is over a bump in this potential, the tip potential will rise above the Fermi energy for heights above 25 nm. Taking this effect into account, we expect the signal to disappear at a height of 30 nm. Though there is some residual signal for weaker AFM scattering potentials, the disappearance of a classical turning point is clearly a significant event.

averaging. The results, shown in Figure 3.9, show a surprisingly strong agreement. In the case of the simulations, we know that the the observed disappearance of the signal takes place as the potential bump dips below the Fermi energy.

If we are going to see an effect on the conductance, we would reasonably expect to find a scattered wave returning to the QPC. In Figure 3.10, we look at the actual wave functions for electrons coming through the QPC and scattering off of the Lorentzian. We look at three different point contact modes, and at three different Lorentzian heights (one of those heights is zero, to provide a point of reference). We note that the Lorentzian below the Fermi energy clearly scatters the incoming wave, but there is little interference going back into the QPC. When the Lorentzian peak is above the Fermi energy, however, there is strong interference going back to the point contact, which is the signature that we were looking for to indicate when conductance should be strongly affected.

3.1.6 Symmetry

In the physical system, a symmetry is observed under a change in the direction of current flow. That is, if we leave the voltage on both the gates and tip the same, and leave the tip in the same location, then changing the direction of the bias voltage across the 2DEG does not change the measured conductance. This is important to note, as measurements have been made using AC current and lock-in techniques. Some might argue that this symmetry is a result of the scattering of holes flowing opposite the electrons, since in a crystal such as this we can think of current being carried by either electrons or holes (or both). However, the same symmetry appears trivially in a model that is strictly single-electron wave propagation.

Our system exhibits both flux conservation and time reversibility (no magnetic fields have yet been introduced). The former consideration results in a unitary scattering matrix ($(s^T)^* = s^{-1}$), and the latter results in a symmetric scattering matrix ($s^T = s$) [9]. Suppose that, for a given configuration of the system, we have m conducting modes on the left of the QPC and n conducting modes on the right. We calculate the conductance from left to right by the sum

$$G_{l \to r} = \frac{2e^2}{h} \sum_{i=0}^{m-1} \sum_{j=m}^{m+n-1} |s_{ji}|^2 \,, \tag{3.7}$$

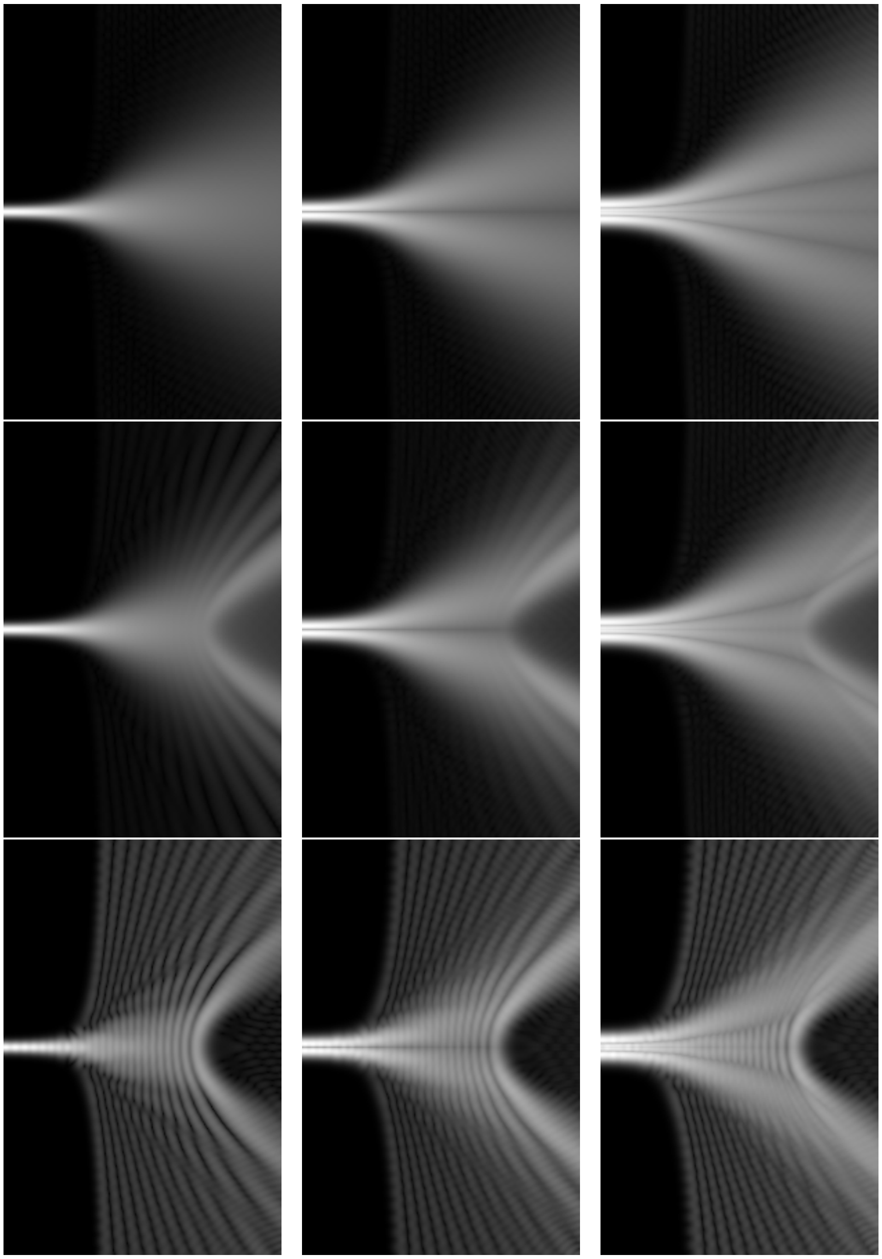


Figure 3.10: These figures show $|\psi|^2$ for electrons coming through the point contact open to one, two, and three transverse modes, with three different Lorentzians. The center of the Lorentzian is centered vertically and right of center horizontally. From the top, we have the peak height at 0, $.75\,E_f$, and $1.25\,E_f$. Lorentzian half-widths are 1.5λ in each case. Note that it is only when the peak height exceeds E_f that we see strong back-scattering to the QPC. Note also that, with the introduction of the second and third modes, we see "lobes" that miss the AFM tip and are bent but not back-scattered.

recalling that the S-matrix connects incoming to outgoing modes. Now, with the same scattering matrix, we get the conductance from right to left by the sum

$$G_{r\to l} = \frac{2e^2}{h} \sum_{i=m}^{m+n-1} \sum_{j=0}^{m-1} |s_{ji}|^2 \tag{3.8}$$

and, since S is symmetric,

$$G_{r\to l} = G_{l\to r}. \tag{3.9}$$

This result isn't profound, but it does have some un-intuitive consequences. Suppose that we have a system with a symmetric point contact and a DC bias voltage. If we place the AFM on the transmission side of the system, it should by now be intuitive that we will have some pattern of regular interference fringes and an envelope related to the transmitted wave function. However, if we place the AFM tip in the symmetric position, we must have the same conductance. Hence the same pattern is seen if we scan the AFM on the *incident* side of the QPC. Our intuition for the physical origin of the pattern is much murkier in this instance.

3.2 A Resonant Cavity

Before the experimental techniques were developed for spatial resolution of electron flow, conductance measurements were made with other independent variables: the voltages on gates, magnetic field, or Fermi energy, for example. One such experiment is to measure the conductance of some cavity as a function of these parameters, to study the spectrum of that cavity, and to extract information about the dynamics. One such cavity has been of particular interest to the Westervelt and Heller groups [30, Chapter 8].

Figure 3.11: This grey-scale image shows the potential used for the resonant cavity in §3.2. The QPC potential is of the same form that we have been considering so far. The mirror part is defined in the following way. First, find the arc desired (here, 90 degrees of arc with the center of curvature at the QPC). Any region to the right of this arc will be held at some given potential higher than the energy of the electrons coming into the resonator. We then sweep a quartic curve around the edge of this region to make the soft wall; the quartic is chosen to match the values and derivatives of the potentials on either side of the soft wall.

The cavity in question has two components (see Figure 3.11). Waves enter through a QPC, just as we have considered previously. On the transmission side of the QPC (here, the right side) we place a mirror, which is a segment of a circle. We can vary the arc length of the mirror, as well as the location of the center of curvature. Classically, the system is closed if the center of curvature is to the left of the QPC, open if it is to the right, and marginal if they coincide.

This system has been considered by the Heller group before. In [31], one finds wavelet calculations for the 2DEG system, showing agreement with some general features of the experimental results. In [30], Hersch explores the same geometry, both theoretically and in a microwave experiment of his own design and execution. He explores the importance of diffractive classical orbits in the open geometry. In such a system, as opposed to 2DEG experiments, one will indeed have the sharp corners that he concentrates on, and the agreement between experiment and theory is very good.

With that background, one might wonder why this geometry should be revisited. The resolution of features in [31] is sufficiently coarse that we couldn't hope to show better agreement than the wavelet calculations, and since we are dealing with smooth potentials the phenomena will be different than those seen by Hersch. There are several reasons to proceed despite these limitations. One is merely to demonstrate the general power of the techniques developed here to treat different experimental situations. In the future, we could also address standing experimental questions, such as how one might use the AFM probe to measure the standing wave in a cavity. Our primary reason, however, is to lay the groundwork for work in a cavity with smooth disorder (explored in Chapter 4).

Computationally, we have more control and freedom in setting up the system than one does experimentally. The drawback is that, with only imperfect information about the experimental system, no amount of control will let us exactly replicate it. We will therefore merely choose a parameter range that produces some interesting results. Let us use the resonator system with the QPC held at just one open mode — that is, at the very beginning of the first conductance plateau. At this point, the wave will come out of the QPC with no longitudinal momentum and we are in a regime where returning waves can interfere with the outgoing wave (as in the tunneling regime).

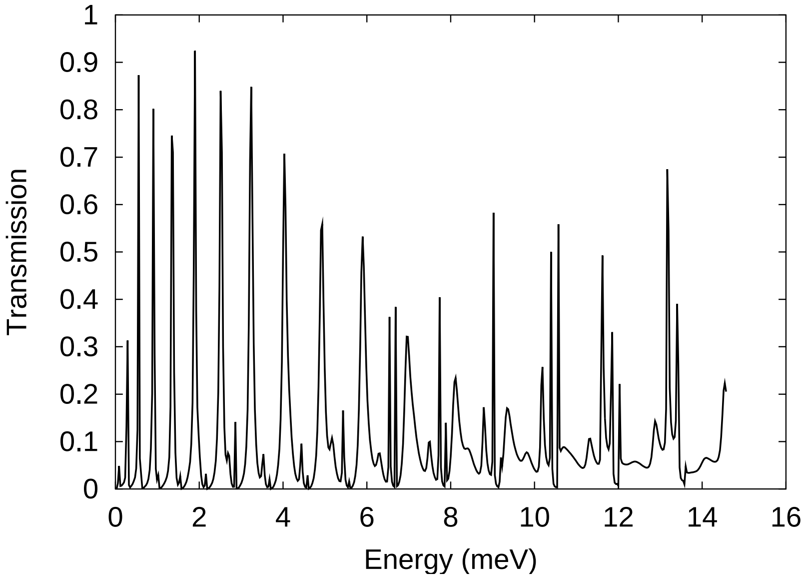


Figure 3.12: Here we see the transmission of the resonant cavity over a wide range of energies, starting at zero. We see the general structure that we expect. At low energies, we see individual peaks corresponding to modes with no angular nodes and an increasing number of radial nodes. For higher energies, we see more structure as the geometry allows for angular nodes as well.

We also make the QPC long enough that all tunneling of other modes is suppressed. Finally, let us take E_F as our independent parameter and look at the spectrum of a fixed cavity. We set the mirror so that the system is marginal; that is, the center of curvature is centered at the opening of the QPC (though we note that, for a smooth potential, this is an ill-defined concept).

In Figure 3.12, we see the conductance curve for a wide sweep of the energy. We see that the conductance has the kind of structure that we would expect, with conductance peaks corresponding to resonances of the structure. For lower energies, we see a regular pattern of individual strong peaks, though there are some smaller peaks present as well. Investigating these strong peaks (see Figure 3.13) we see that they correspond to a sequence of modes with no angular nodes.

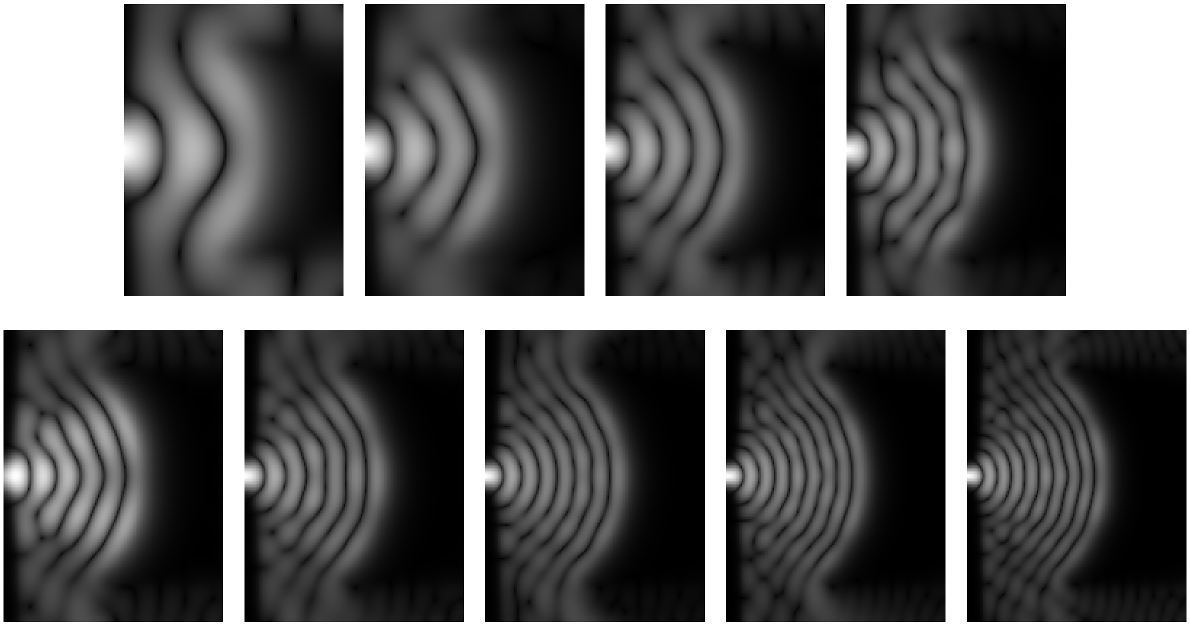


Figure 3.13: In this figure, we look at the wave functions in the cavity at the strongest conductance peaks in Figure 3.12. These strong peaks correspond to the simple cases of no angular nodes. Starting at low energies, we see an incremental progression in the number of radial modes occurring at energies that, properly scaled, give us the progression of squared integers.

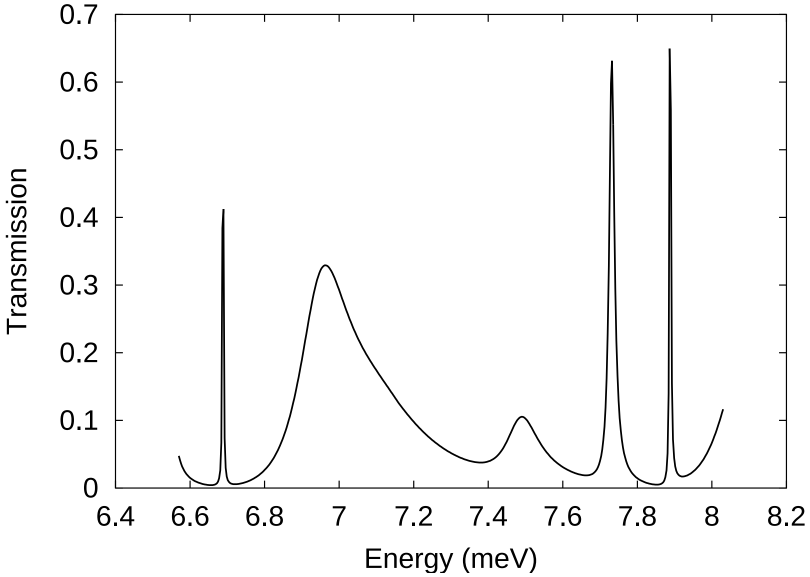


Figure 3.14: Here we look at a finer scan of energies in a narrow band of Figure 3.12, exposing the structure resulting from modes with angular nodes.

At higher energies, the structure becomes more interesting. This results from the shortening Fermi wavelength, and the increasing sustainability of angular nodes in the cavity. We choose one range of energies from Figure 3.12 and do a finer scan of energies, shown in Figure 3.14. We see the additional structure here, and in Figure 3.15 we see the standing waves in the cavity at the peaks in this transmission curve.

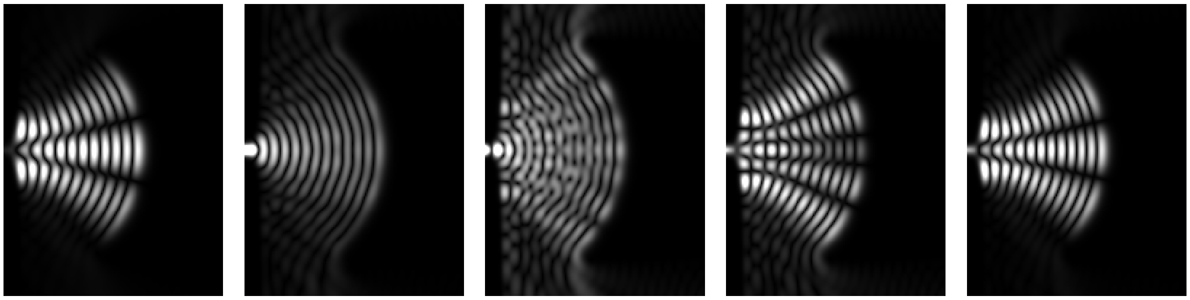


Figure 3.15: These are the wave functions at the peaks in conductance in Figure 3.14. At these energies, we note that the narrowest transmission peaks correspond to modes with angular nodes. There are no modes with a central node, as dictated by the symmetry of the cavity. The weakest transmission peak, the third, is notable in that is has significant amplitude near the edges of the mirror.

Chapter 4

Disorder and Branched Electron Flow

In the previous chapter, the devices that we considered had particularly simple potentials. They took into account only the shape of the potential generated by the electrostatic gates and the AFM tip. What has been left out are components of the potential that are beyond the control of the experimentalist. There is disorder in the 2DEG crystal, leading to a lumpy background potential that will inevitably be present in the experimental system. Since we have been careful to develop techniques that can handle arbitrary smooth potentials, we can develop a model for this disorder and include it in our simulations.

It should be noted that, though the presence and nature of this disorder is generally left out of simulations, experimental results from the Westervelt group have highlighted the importance of considering this disorder. It is not only present, but it has a significant impact on the results seen in the new spatially resolved conductance measurements.

In this chapter, we will begin by developing a model for how this disordered potential should appear. We then see the effects that including the background potential in our calculations has on electron flow.

4.1 Experimental Hints

Though we had not yet discussed it, we have already seen some evidence of disorder. Referring back to the data shown in Figure 3.7, we see that the flow is bending, well away from the QPC, in a way that suggests the presence of additional features in the potential. It wasn't until the experiments probed the region of electron flow well away from the QPC, however, that the full impact of this disorder became evident. Whereas the flow close to the QPC could be understood by a fairly simple theory looking at the transverse modes of the point contact, this picture breaks down on the micron scale.

The flow of electrons over longer distances is seen to have a branched structure, as shown in Figure 4.1. The first suspect for a cause is clearly the disordered background potential. If the 2DEG potential were flat after the QPC, the simple patterns of QPC transverse modes would remain at all length scales. Furthermore, we can see that the effect of the disorder is not merely diffusive. If it were, we would see the "smearing out" of features without the introduction of new, interesting structures.

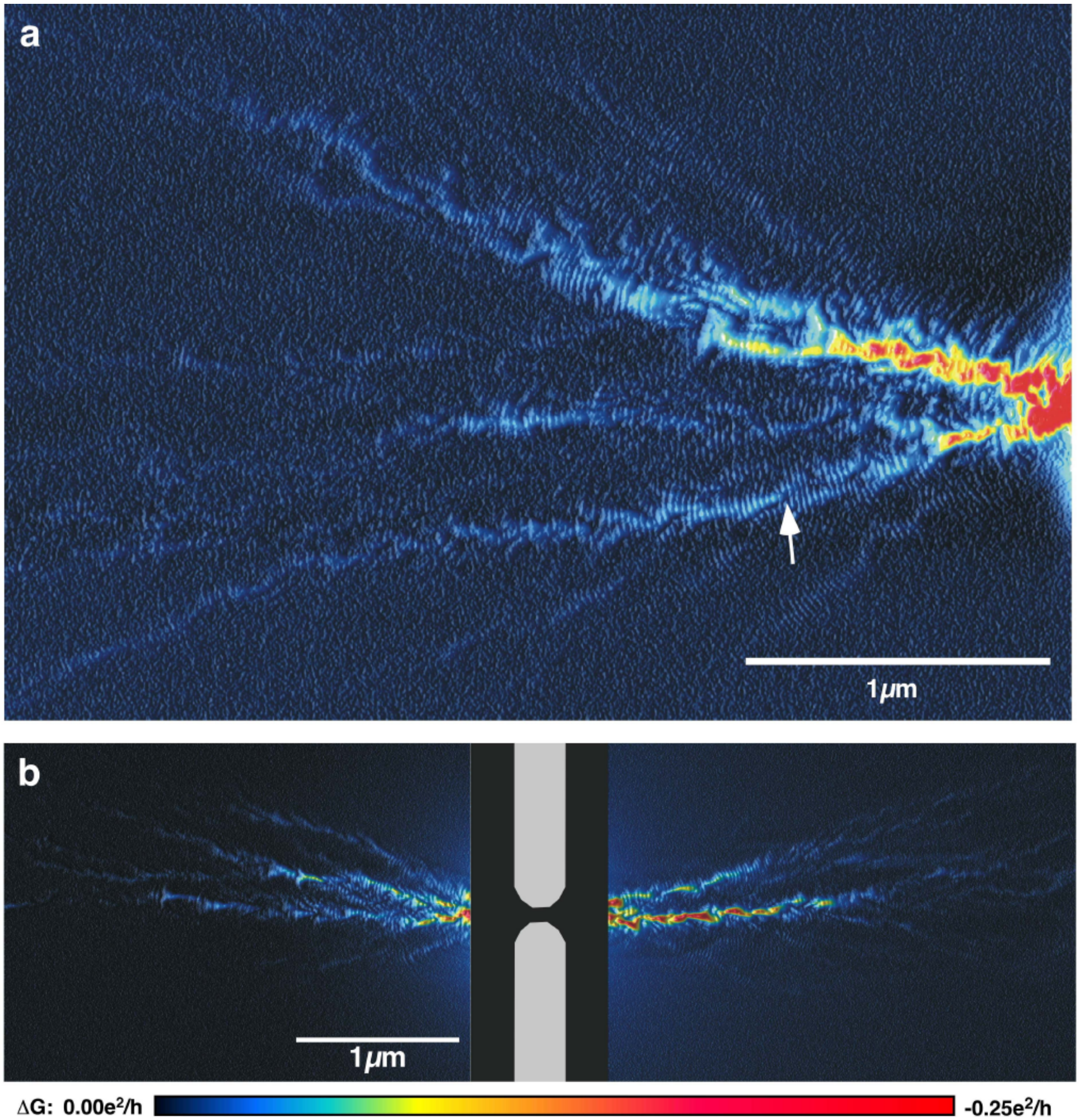


Figure 4.1: In this figure, we see experimental data taken over a relatively large range beyond the point contact. We see immediately that the spatial pattern of the flow is not simply an extension of the patterns of QPC transverse modes. Parts (A) and (B) show the results in different experimental systems, demonstrating that this branched flow is more the rule than the exception. The arrow in (A) indicates the location of an apparent caustic, indicating that the flow has experienced some focusing event.

4.2 Mathematical Model for Disorder

The techniques that allow us to calculate electron propagation with general smooth potentials also allow us to introduce any background potential that we want. We must first determine what form this background should take. There are two physical origins for this potential: impurities throughout the crystal structure, and variations in the density of donor atoms in the doped plane. We will consider both of these contributions. When modeling the disordered potential, we will seek to mimic the physical system as closely as possible. This means designing a model that uses the physical parameters of the experimental system, and it should have some measurable quantities that can be compared with that system.

We will treat both the impurities and the donor atoms as point charges above a conducting plane. The effective potential induced in the 2DEG will be of the form

$$V(r) = -q\frac{\hbar^2}{2m}\frac{d}{(r^2+d^2)^{3/2}}, \tag{4.1}$$

where q is the charge on the impurity or donor atom and d is the distance between the charge and the 2DEG; the distance r is measured in-plane. We arrive at this expression by first approximating the 2DEG as an ideal conducting plane and finding the 2DEG density variations [32, §3.2.2]. From these density variations, we look at the two-dimensional density of states to find the local variation in wave number [21, Ch. 1]. Finally, we determine the effective potential implied by these wave number variations. Due to our assumption of an ideal conducting plane, this is not the full form of the potential that would arise in a 2DEG with limited mobility and carrier density. It does, however, capture the key feature of the screening in the long-range behavior [21, Ch. 9]. Both of our sources of disorder carry a charge $|q| = e$, the electron charge. Donor atoms are positively charged, impurities are negatively charged.[1]

Let us look first at the impurities. These charges are distributed randomly in three-dimensional space with a density ρ_i. The density of impurities is largely a function of the fabrication process. For the samples used in experiments of interest to us, ρ_i is sufficiently small ($\rho_i \sim 10^{14}\,\mathrm{cm}^{-3}$) that we ignore any correlation between charges.

The impurities appear throughout the crystal, both above and below the 2DEG, so there is effectively no limit on how far they can be from the 2DEG. Nor, more importantly, on how close they can be to the 2DEG. Those that are sufficiently far from the 2DEG will have very broad, low potentials. Those that are close to the 2DEG, however, will have high and narrow contributions.

Next we consider the donors. These charges are distributed in a two-dimensional region parallel to the 2DEG (called δ-doping). The sheet density in this layer is σ_d, and for the systems under consideration $\sigma_d \approx 8 \times 10^{12}\,\mathrm{cm}^{-2}$. The dopant layer is 22 nm away from the 2DEG. It should be noted that this is closer than in many experiments because of the particular experimental needs of the Westervelt group. One wants to bring the AFM tip as close as possible to the 2DEG to have a narrow perturbation. As a result, the crystal structure above the 2DEG is made thinner than otherwise necessary.

Comparing the sheet density of donor atoms to the spatial density of impurities, we see that there are a greater number of donor atoms significantly affecting the 2DEG.

---

[1] The origins of the charges on the impurities are localized electron states. Some of the electrons released by the donors find these impurities in the crystal structure rather than making it to the 2DEG.

The distribution of donor atoms in their plane is, however, of great importance. If the donors were distributed with complete uniformity, we could reasonably treat them as a sheet charge with a simple mean-field effect on the 2DEG. It is the inevitable variations in the density of donor atoms that give rise to lumps in the potential felt in the 2DEG.

Though the donors will not be distributed uniformly, the other extreme is also not realized. The density of donor atoms is sufficiently high that we cannot ignore correlations and distribute them completely randomly. These charge correlations limit the variations in density, which, in turn, limits the magnitudes of variations in the effective potential. We find in [33] a self-consistent study of the distribution of donor atoms. The authors find a correlation function for charged donor atoms, and a relationship to apply this correlation function to any donor atom density. We use this correlation function in a simple but sufficient fashion: it provides a maximum local density for the charged donors, which we then use to limit the random distribution of atoms.

4.3 Measuring Disorder

This model for a disordered potential has the advantage of taking into account physical parameters of the system. We can enter the (at least approximately) known densities of donors and impurities, and generate an ensemble of random potentials. An example of a potential from this ensemble is shown in Figure 4.2. Note that, though the potentials are random, we are left with no arbitrarily adjustable param-

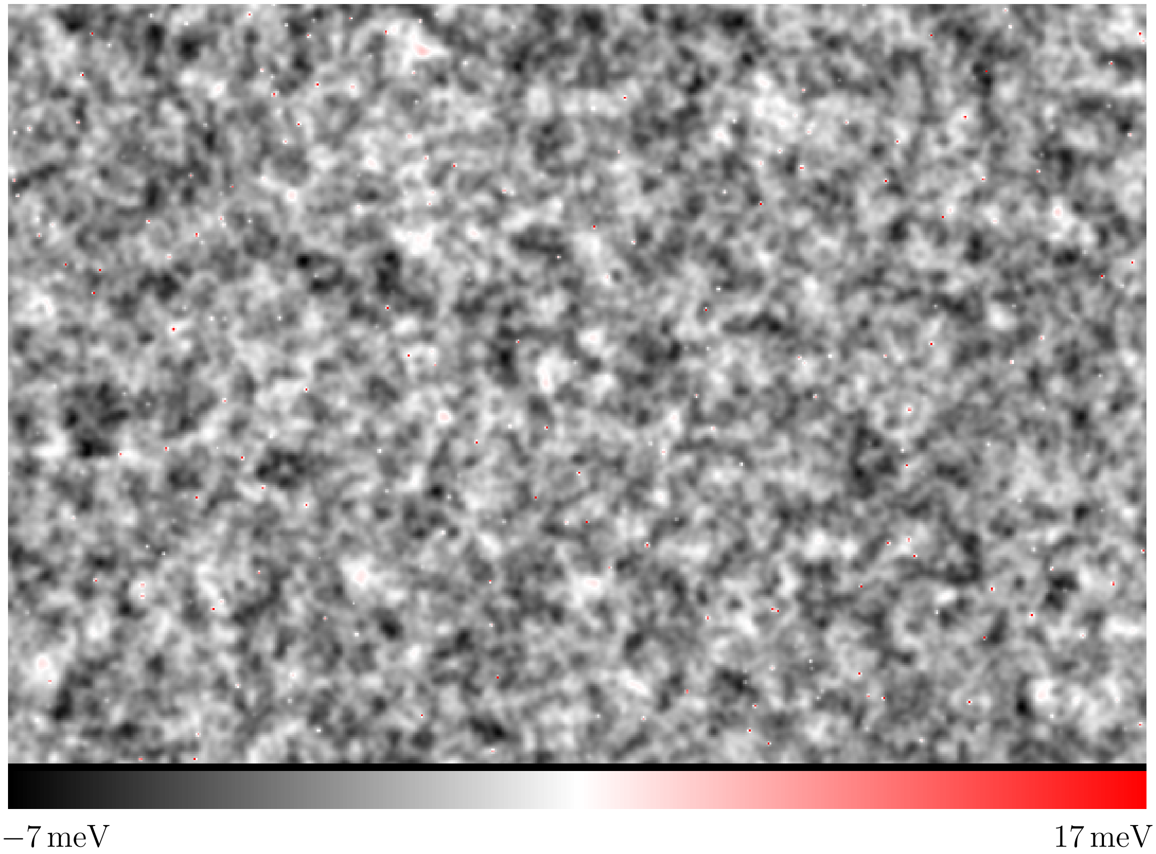


Figure 4.2: Here we see an example of the disordered potential generated by our model. The effects of donor density fluctuations and of impurities are included. The notable effects of impurities can be seen as the occasional sharp spike in the potential, which we have attempted to bring out with the color map. In the actual potential, these spikes go higher than 17 meV, but the data has been chopped to allow us to see the smoother background.

eters.[2] It would be useful, therefore, to test the model by comparing to some other experimentally accessible parameter.

We find such a parameter in the mobility μ. This quantity can be measured experimentally, and we have a means to measure it numerically as well. We do so using classical trajectories and looking at the momentum-momentum correlation function. We define the quantity

$$c(t) = \hat{p}(t) \cdot \hat{p}(0), \tag{4.2}$$

which tells us how well the particle "remembers" its initial direction of propagation. We let the momentum-relaxation time τ for a potential be defined as that time τ for which the ensemble averaged[3] $\langle c(\tau) \rangle$ has dropped to a value of e^{-1}. Note that $\langle c(\tau) \rangle$ decays exponentially with time, as discussed later in this thesis (see, for example, Figure 7.1). The mobility is given in terms of the momentum-relaxation time by $\mu = e\tau/m$ [21, Eq. 9.69]. It is also useful to consider a momentum-relaxation length $l = v_F \tau$.

Our physical system has a mobility in the range 10^6 to 4.5×10^5 $\mathrm{cm}^2/\mathrm{V\,s}$. Using the physically acceptable ranges of system parameters, this potential model consistently produces mobilities in that range.

Next, we would like to characterize the random potentials generated by this method in order to compare them with those from other methods. Two measures that we use are correlation length and potential height distribution. For the former, we begin by assuming that the auto-correlation function of the potential will be

---

[2] Because of the error bars on all experimental parameters, we are given ranges within which to work.

[3] Here, "ensemble" means both our ensemble of random potentials and ensembles of random trajectories in those potentials.

circularly symmetric (a reasonable assumption, as we have not built in any preferred direction). The number that we quote as the correlation length is the distance that one needs to travel radially for the auto-correlation function to drop off to e^{-1} (the value for zero displacement is 1). The correlation length is the answer to the question "in general, how broad are the features of the potential?"

The potential height distribution answers the question "in general, how large are the features in the potential?" To answer this, we collect the value of the potential at points on a regular grid. We then make a histogram of these values. If the potential is sufficiently random (which it is, when taken over many correlation lengths), this distribution will be a Gaussian. We take one standard deviation of this Gaussian distribution as the height scale of the potential.

Putting our experimental values into the model, the resulting potentials have correlation length scales on the order of 40 nm and height scales between ten and fifteen percent of the Fermi energy.

4.4 Electron Flux Results

The planning that has been put into the design of the potential and the matching of experimental parameters was not a waste; when we ran quantum flux through these random potentials, we observed the same branching phenomenon seen experimentally. An example is shown in Figure 4.3.

Using the methods set out in Chapter 2, the quantity most easily accessible to us is the quantum-mechanical flux density in the system. The spatially resolved measurements performed in the experiments are, however, of conductance as a function

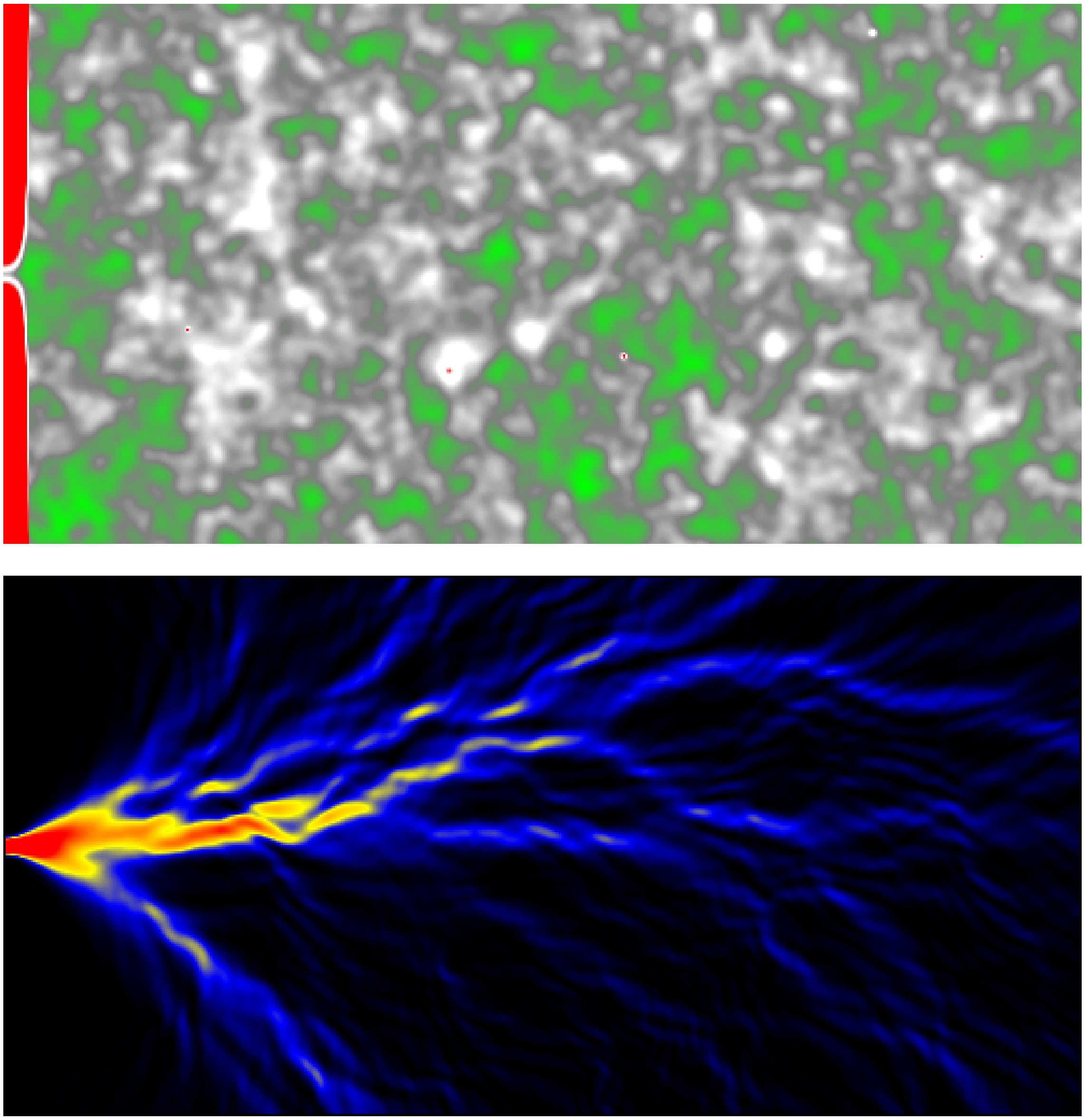


Figure 4.3: Here we see an example of quantum flux through a disordered potential in our model. The top image shows the potential, with a color map that starts at green for low values, through grey and to white. The QPC appears in red at the left. The bottom image shows the quantum flux through the system. These data are over an area two microns long by one micron high.

of AFM tip position. It is reasonable to presume that these two quantities will show similar behavior, since the amount of back-scattering from the AFM tip will necessarily depend on the electron flux at its position. This assumption should, however, be tested. Just as we simulated a tip scan to look at fringes close to the QPC in Figure 3.6, we can also introduce a movable AFM tip potential in a system with branched electron flow and measure conductance. As shown in Figure 4.4, we find strong agreement between the two quantities. This result bolsters both the experimental claim that electron flux has been mapped, and our intention to draw conclusions about the experimental system based on flux calculations.

Since the potentials are generated randomly, the goal was never to reproduce precisely the results of any given experiment. However, we were extremely pleased at how well the qualitative characteristics were reproduced. Not only are the branches present, but we also see the same general length scales for the formation and perseverance of the branches. In the simulated tip scans, we see features in the data similar to those in the experimental data.

A crucial result of these simulations is the following: comparing the potential in Figure 4.3 to the branched flow through it, one cannot make a direct correlation between obvious features of the potential and the locations of the branches. In particular, it is not the case that the branches are merely following low areas in the potential, which we could characterize as “channels” directing the flow. Though there are occasional events where, for example, the flow is split around a sharp impurity scattering center, the branches are the result of accumulated small-angle scattering.

Of the two contributions to the potential, the far more important for branching

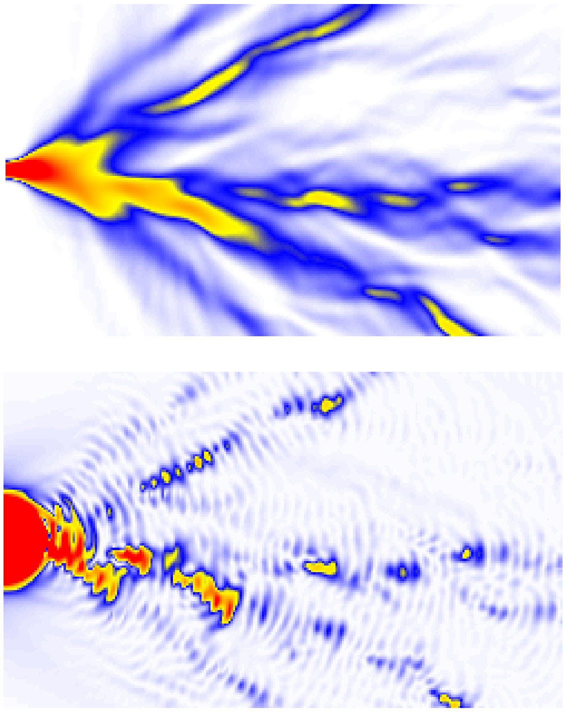


Figure 4.4: In this figure, we see a comparison of quantum-mechanical flux density (top) and conductance as a function of tip position (bottom) for the same disordered potential. As anticipated, the strength of the signal when scanning the AFM tip is dependent on the flux density at the tip position. The fringes seen in the tip scan are not present in the flux, also as anticipated. These data cover an area one micron long by .6 microns high.

is that from the dopant density variations. If one removes the impurities from the system, the branches remain. They may change somewhat in the details, but they are still present and qualitatively the same. Removing the dopants and leaving only the impurities significantly changes the flow, however, and branches are no longer present.

4.5 Branched Flow in a Magnetic Field

Up to this point, we have only considered the branched flow phenomenon in systems without an external magnetic field. The effects of a magnetic field are somewhat intuitive, but are nonetheless interesting to note.

First, let us consider the effect of a magnetic field on the electron flux itself. In a system with a flat potential, the effect of the field would be to bend all electron trajectories to curves with some known cyclotron radius. With disorder present, however, the dynamics will favor some paths over others (hence the branched nature of the flow). The competition between these dynamics for electrons in the magnetic field results in a "ratcheting" of the branches as the field strength is increased, rather than a continuous sweeping. Though this is difficult to convey without a continuous set of flux images, the idea is illustrated in Figure 4.5.

The shifting of branches has greater implications for the experimental situation, since the accessible conductance measurement depends on scattering. We can understand the correlation between the measured conductance and electron flux by considering classical trajectories. To decrease the conductance of the system, a trajectory needs to be scattered by the AFM tip and return to the QPC. In the absence

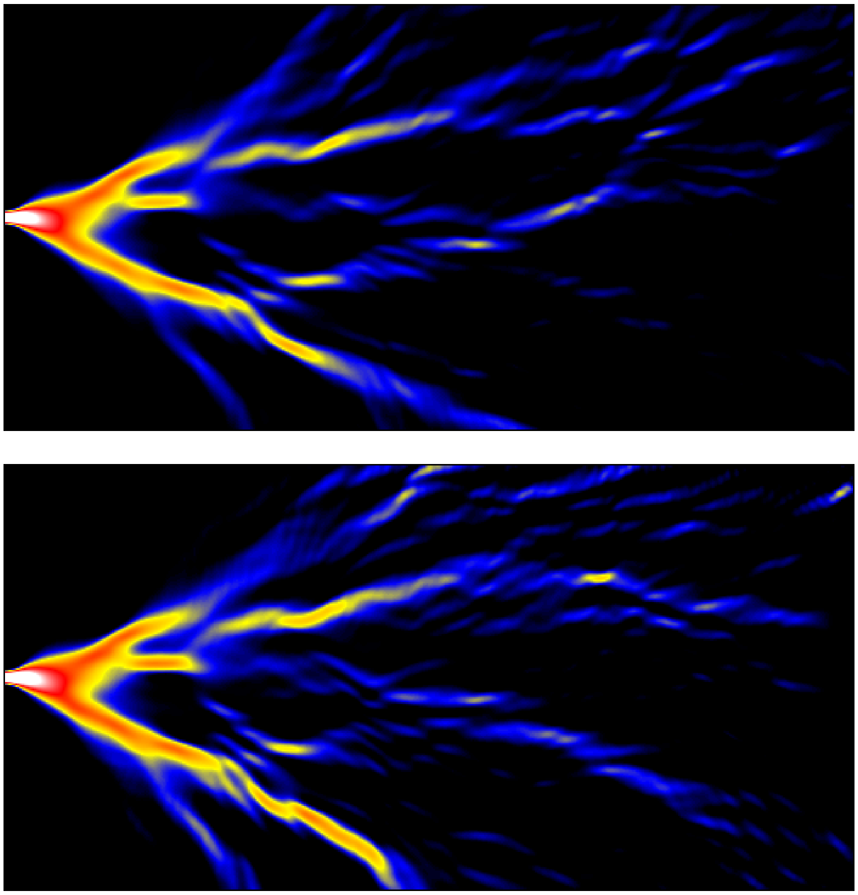


Figure 4.5: Here we see two sets of electron flux density data, taken with the same disordered potential but at two different magnetic field values. For the top image $B = 0$, and for the bottom image $B = 25\,\mathrm{mT}$. Both data sets cover an area two microns long by one micron high. We see that, when the field is increased, the branches do not bend continuously. Rather, the branches are the same out to some distance, at which point the electrons take clearly distinct paths. We understand this as the effect of the magnetic field accumulating until it is sufficient to cause trajectories to jump from one dynamically favored branch to another. This ratcheting effect continues as the flux increases, resulting in a cumulative net bending of trajectories. We also note that flux density has shifted within the branches present, another effect that accomplishes the net bending caused by the magnetic field.

of a magnetic field, this means that the trajectory should impinge on the AFM tip potential at a right angle to the classical turning point and be scattered back along itself. Time reversal invariance guarantees that the trajectory will follow the same branch back to the QPC that it used to get to the AFM. Placing the AFM over a branch ensures both a high density of impinging electrons and a high density of possible return paths into which to scatter.

In the presence of a magnetic field, we lose the time-reversal invariance that we used to describe our scattering process. The existence of a classical trajectory from the QPC to the AFM no longer guarantees the existence of a return path.

Thinking in terms of time reversal, however, points us in the direction of what we should instead expect to see. The set of trajectories that can be used to return to the QPC from the AFM are just those that are seen going to the AFM from the QPC for the opposite magnetic field. The conditions for a strong signal, the presence of both outgoing and return paths, are satisfied by those regions of space where these two branching patterns overlap. We would then predict that the signal measured with a tip scan would correlate to the product of the flux densities calculated at both signs of the magnetic field. This prediction is born out by simulations, as shown in Figure 4.6.

This intuition implies a relationship that we already know to be true: the signal that we measure must be symmetric in magnetic field [16]. Changing the sign of the field merely exchanges the roles of outgoing and return paths.

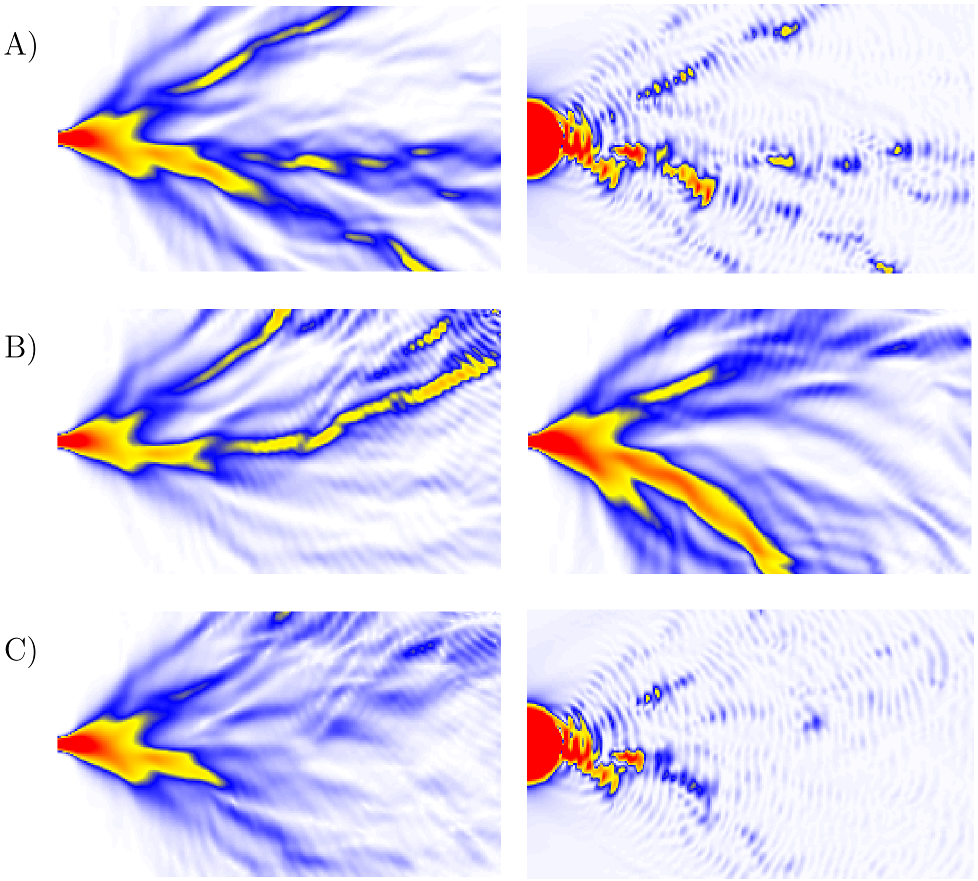


Figure 4.6: In this figure, we show the results of a simulated AFM tip scan in the presence of magnetic field. The data are over a limited range, measuring one micron long by .6 microns high. In row (A), we see the flux and the results of a tip scan at zero magnetic field. The correspondence between the two is as seen before. In (B), we have the flux at a magnetic field of $\pm .1\,\mathrm{T}$. (The interference in the upper-right corner of the first image in (B) is simply the result of a branch hitting the corner of the grid.) In (C), we show the square root of the product of the two sets of data in (B), and the result of a simulated AFM tip scan at $B = .1\,\mathrm{T}$. (We have chosen the square root of the product so that, as we take $B \to 0$, we reduce to the image used in (A).) This tip scan correlates to the product, rather than to the actual flux at either field value. We find the same tip scan data if we take $B = -.1\,\mathrm{T}$. The most interesting feature of the tip scan in (C) is isolated island of signal near the center of the scan where two branches cross one another.

4.6 Disorder in a Resonant Cavity

We can now, as earlier suggested, look at the effect that a disordered potential has on the resonant cavity considered in §3.2. In a cavity, we expect noticeable effects for levels of disorder that might not be seen in an open system at our length scales. The reason is that electrons confined to a cavity will "see" the same disorder many times as they bounce around. There are three things that we consider when we introduce disorder. First is that the resonant energies of the cavity will shift, and so the peaks in the transmission curve will move. Second, we note that the closed status of the cavity, especially as we approach marginality, is a delicate balance. Small levels of disorder in the cavity can easily destroy it for the reasons noted above. As a result, we would expect that the transmission of the cavity will be increased throughout the spectrum. Finally, the symmetry of the cavity is broken by any disorder. With the symmetry broken, waves can couple to a new class of cavity modes and we should see new peaks appear in the transmission spectrum.

We take the same cavity setup used before, and introduce disorder gradually. An example of the resulting potential is shown in Figure 4.7. In Figure 4.8, we see what happens to the spectrum in a narrow band as the disorder is introduced. There are five steps as the disorder strength is increased, ending at $.05\,E_F$. Even at this level of disorder, we see that the extant conductance peaks are shifted and new ones appear. In addition, the whole conductance curve is raised, as the geometric closure of the cavity is lost. In Figure 4.9, we can see what is happening to the wave functions at the same time; for comparison, we have the wave functions at the conductance peaks for zero disorder and for disorder at $.05\,E_F$. Most interestingly, we see that the disorder

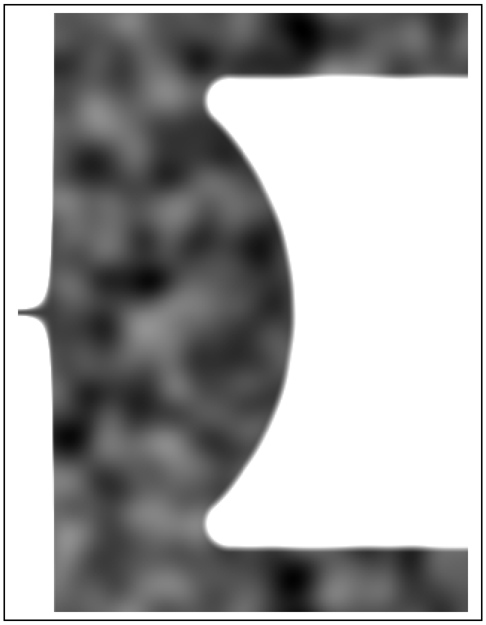


Figure 4.7: This picture shows the potential in our resonant cavity with a disordered background introduced. Because the heights of the peaks in the disorder are small (around 5%) compared to the energy of the electrons that we will send through the system, the grey scale color map is cut off and we don't see the softness of the mirror and QPC. The radius of curvature of the mirror is five correlation lengths of the disorder.

has allowed our previously symmetric system to couple to modes with a central node.

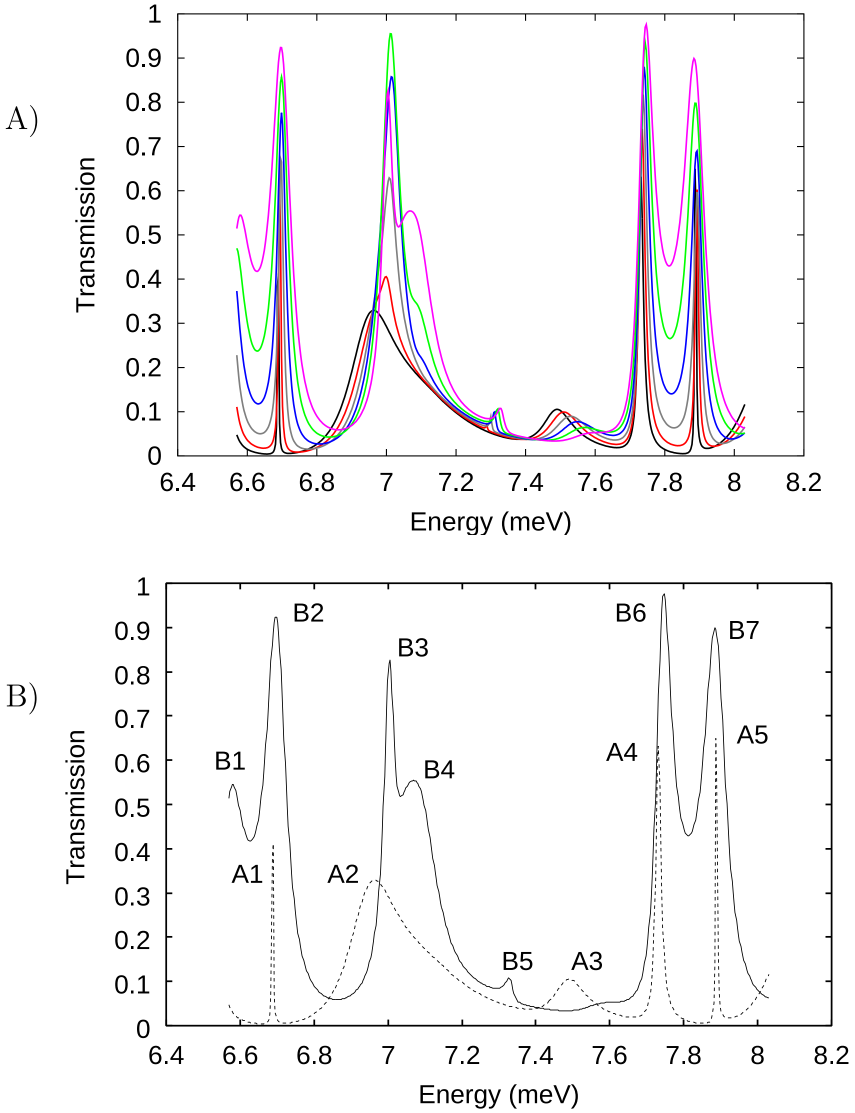


Figure 4.8: Here we see the evolution of the cavity transmission spectrum as the strength of the disorder is increased. In (A), we show six curves, the lowest (black) is for no disordered background and the highest (purple) is for a background at .37 meV (.05 E_F for the center of the spectrum), with even steps between. In (B), we show only the two extremes, and we have labeled the peaks. We see the shifting of peaks (e.g., A4 to B6), the disappearance of peaks (e.g., A3), and the appearance of new ones (e.g., B5) as the disorder is introduced. We also note the overall effect of the disorder to increase the transmission of the system.

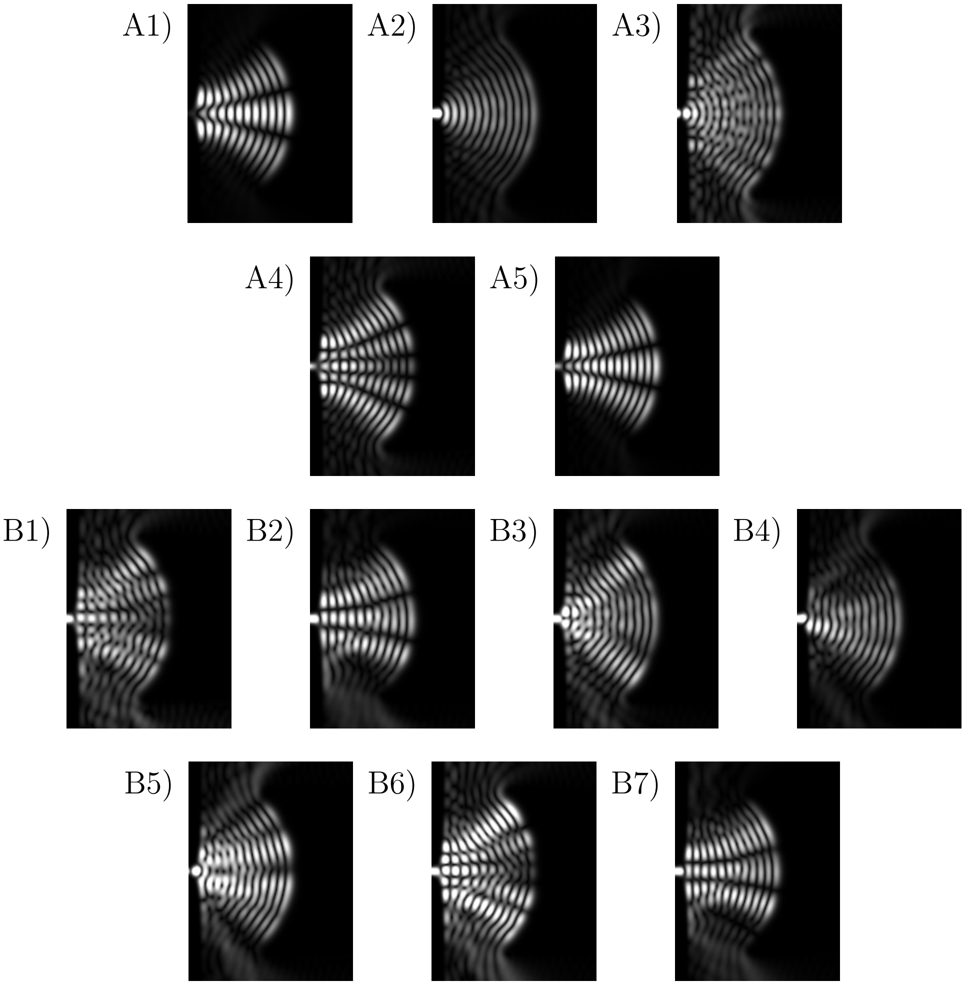


Figure 4.9: Here we see the wave functions at the peaks in the transmission curves of Figure 4.8 (B). Comparing the two data sets, we note two things. First of all, looking at peaks that are present for both curves, the amount that a wave function changes is related to the height and width of the original transmission peak (compare A1 changing to B2 with A2 splitting to B3 and B4). This makes sense, as narrow, high peaks correspond to stronger resonances of the original cavity, and we might expect that these will be harder to destroy. Second, for the wave functions that change significantly and for the one that isn't present at all without disorder (B5 is the only completely new peak since B1 has simply been shifted into this energy range), the difference is primarily the introduction of a component with a central node. In the original, symmetric system, such states weren't allowed.

Chapter 5

Fringes

The results of the experiments mapping electron flow in a 2DEG showed two surprising results. The first was the branched nature of the flow, explained as the effect of a disordered background potential. The second was the perseverance of regular interference fringes over the entire, several-micron range of the scans.

Interference fringes spaced by half of the Fermi wavelength were expected close to the point contact. At distances of several hundred nanometers, the various energies present in the electron flow would remain in phase with one another. As one moves farther from the QPC, however, simple considerations would suggest that the various frequencies would disperse and any interference effects would be washed out.

The survival of the interference fringes is seen in full quantum-mechanical simulations with thermal averaging, as shown in Figure 5.1, so we know that a complete theory will reproduce them. This full solution does not, however, tell us much about the mechanism that allows the fringes to survive.

In this chapter, we begin by explaining why one might expect the fringes to die

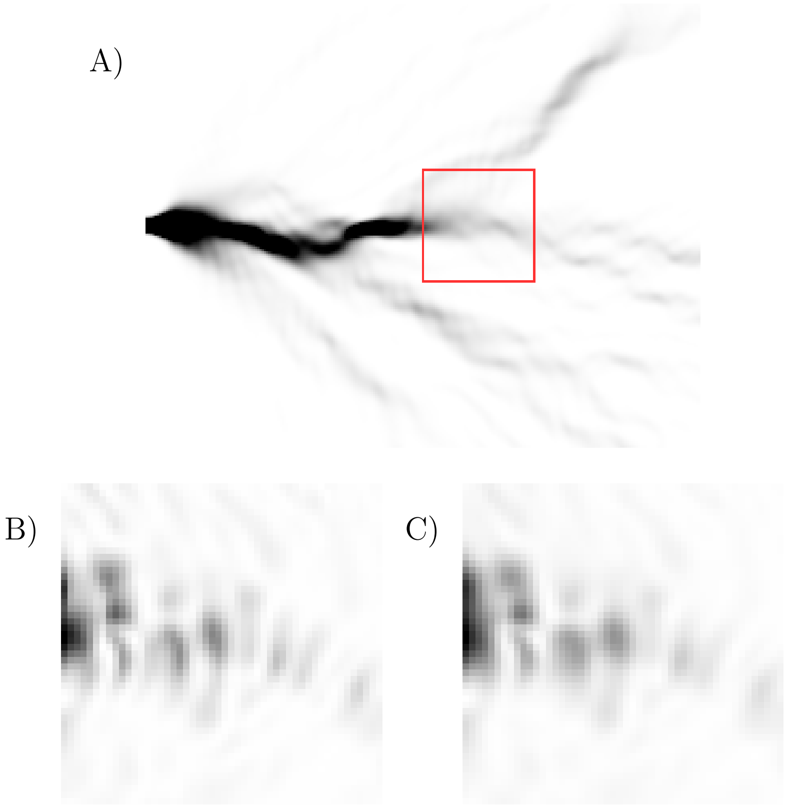


Figure 5.1: Here we see the survival of interference fringes beyond the thermal length in a full quantum-mechanical simulation. (A) shows the quantum-mechanical flux through the system; for a discussion of the "branched" nature of the flux, see Chapter 4. In (B), we have introduced a movable tip potential and plot conductance as a function of the tip position. This scan is taken at the Fermi energy. In (C), we have a thermally averaged scan of the same region. In this simulation, $\ell_T \approx 600\,\mathrm{nm}$ (see §5.1) and the left edge of the scan is approximately $1\,\mu$m from the point contact.

out as we move away from the QPC. Then we present a simple model for fringe formation, appealing only to first-order scattering, that has a very different dependence on thermal averaging and shows that fringes should survive to the phase-coherence length. We will also take a look at an idea, appealing to higher-order scattering, that suggests that fringes may occur beyond the phase-coherence length.

At the root of these discussions is the model that we have for the potential seen by electrons in a 2DEG. Though we found that the donor-atom contribution to the potential was the more important component for branching, it is the impurity contribution that is more important for fringing. The reason is that only the impurities give us strong scattering centers, which are necessary to get waves traveling back towards the QPC.

5.1 The Thermal Length

There are many different quantities referred to as "thermal lengths," one of which plays a role in our understanding of the interference fringes. Before explaining our understanding for the perseverance of the interference fringes, it is worth spending some time to explain why they were unexpected.

Recall that the measurements are of conductance as a function of tip position. We have shown that the signal is dependent on electrons being scattered back to the QPC, so we will consider paths with this result. There are three sources of scattering to consider: the gates, the AFM tip, and the impurities in the crystal structure.[1] The gates are clearly strong scatterers, and we would expect the same

---

[1] For this effect, we ignore the small-angle scattering caused by donor atom density variations. It

of the AFM tip as long as it creates a depletion region. The impurity scatterers, however, have comparatively small cross sections. Though we may consider multiple scattering events from the AFM and the gates, we expect any signal involving multiple scattering from impurities to be negligible compared to a single-scattering signal.

There are two ways to understand the fringes near the QPC. The first is the one that we used in constructing our simple model. If the QPC has no open channels, so that all conductance comes from tunneling, then a wave scattered back to the QPC from the AFM can interfere with the outgoing wave and change conductance. If there are open channels, which is the more common experimental situation, then the outgoing wave is distinguishable from the return wave. Here, we look at the interference of multiple ways of returning to the QPC. A wave scattered from the AFM tip will, in general, be partly transmitted back through the QPC and partially reflected. This reflected wave can scatter from the AFM again and interfere with the first return wave. We thus have, essentially, an open Fabret-Perot cavity and we can understand the interference fringes within that paradigm.

Simple kinematic considerations suggest that these fringes should die out at a thermal length ℓ_T given by

$$\ell_T = \hbar^2 k_0 / 2mkT. \tag{5.1}$$

This length comes from a consideration of the spread of energies present in the experiments. Though they are performed at low temperatures (less than 4.7 K), the implied spread of energies is still noticeable. We found ℓ_T as the distance at which waves differing in energy by kT will drift out of phase by one radian over the round

---

is unlikely that it would cause a path shorter than the phase-coherence length to return to the QPC.

trip (QPC to AFM and back). When this happens, the interference patterns from the various energies present will be sufficiently out of phase with one another that the aggregate signal would have no discernible fringe pattern. The fringes seen experimentally, however, survive well beyond this radius. We look for other paradigms of fringe formation to understand this observation.

5.2 Within the Phase-Coherence Length

Here we use a simple, single-scattering model to predict the fringes seen beyond the thermal length. The result depends on phase-coherent transport at each individual energy present, and therefore does not apply beyond the phase-coherence length. We will consider single-scattering events involving the AFM and the impurities that result in waves returning to the QPC, and the interference between these various paths.

5.2.1 Thermal Averaging

It was the thermal average, a sum over the various energies present in our propagating electrons, that gave us the thermal length and the expectation that fringes would die out. In this model, we will need to perform a thermal averaging integral explicitly. The fully correct thermal average is accomplished by an integral over energy with the derivative of the Fermi function as a weighting function. In order to simplify the mathematics of this model, we seek an approximation that is an integral over wave vector with Gaussian weighting. For the ranges of parameters in this system, such an approximation can be made to an acceptable degree of accuracy.

The thermal distribution of energies begins with the derivative of the Fermi function at the known temperature T and Fermi energy E_F:

$$\begin{aligned} f(E) &= \left[1+e^{(E-E_F)/kT}\right]^{-1} & (5.2)\\ -f'(E) &= \left[1+e^{(E-E_F)/kT}\right]^{-2}\frac{1}{kT}e^{(E-E_F)/kT}. & (5.3)\end{aligned}$$

We wish to approximate this function by a Gaussian while preserving the normalization. We can do so by taking

$$-f'(E) \approx \frac{1}{2\pi^{1/2}kT}e^{-[(E-E_F)/(2kT)]^2}. \tag{5.4}$$

The width of this Gaussian was set by matching the second-order Taylor series about $E = E_F$ for the two functions. If we were to match the value at $E = E_F$ rather than matching the normalization, we would see that the shapes of the two curves are very similar. The normalization doesn't change this fact; it merely changes the overall constant. The majority of the error in this approximation comes from this first step.

To transform this exponential into a Gaussian in k, we perform further simplifications. We take

$$\begin{aligned} \left(\frac{E-E_F}{2kT}\right)^2 &= \left(\frac{k^2-k_0^2}{4mkT/\hbar^2}\right)^2 & (5.5)\\ &= \frac{\hbar^4}{16m^2(kT)^2}(k^2-k_0^2)^2 & (5.6)\\ &\approx \frac{\hbar^4}{16m^2(kT)^2}\left[4k_0^2(k-k_0)^2\right] & (5.7)\\ &= \frac{\hbar^4k_0^2}{4m^2(kT)^2}(k-k_0)^2. & (5.8)\end{aligned}$$

We note that, using the definition of the thermal length in Eq. 5.1,

$$\frac{\hbar^4k_0^2}{4m^2(kT)^2} = \ell_T^2, \tag{5.9}$$

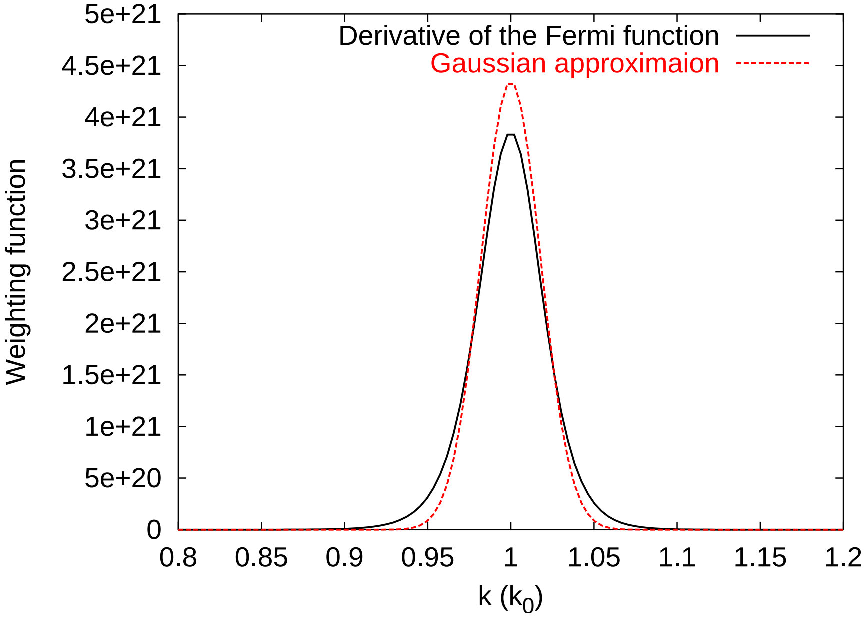


Figure 5.2: Here we show two possible weighting functions for comparison. First is the true thermal distribution function, the derivative of the Fermi function. Second is our approximation, a Gaussian in wave number. The curves are calculated for $E_F = 16\,\mathrm{meV}$ and $T = 4.7\,\mathrm{K}$.

in terms of which we take

$$-f'(E) \approx \frac{1}{2\pi^{1/2}kT} e^{-(k-k_0)^2\ell_T^2}. \tag{5.10}$$

This is the weighting function that we will use in performing the thermal average. We compare it to the original derivative of the Fermi function in Figure 5.2. Though the two curves are not identical, it is reasonable to assert that the key results from the model will not be affected.

We wish to integrate over k rather than E in taking the thermal average. Given the dispersion relation $E = \hbar^2 k^2/2m$, we have $dE = (\hbar^2 k/m)\,dk$. Again appealing to the values that will appear for k_0 and ℓ_T, we can approximate this dispersion relation as linear over the range of the weighting function and take $dE \approx (\hbar^2 k_0/m)\,dk$. Hence, for a signal $s(k,r)$ at fixed wave vector, we have the thermally averaged signal $s(r)$ given by

$$\begin{aligned} s(r) &= \frac{\hbar^2 k_0}{2\pi^{1/2} mkT} \int dk\ e^{-(k-k_0)^2 \ell_T^2} s(k,r) && (5.11) \\ &= \pi^{-1/2} \ell_T \int dk\ e^{-(k-k_0)^2 \ell_T^2} s(k,r). && (5.12) \end{aligned}$$

5.2.2 The Single-Scattering Model

This model is designed to have simple mathematics so that we can express an analytic result. The approximations made have no effect on the qualitative results, and little effect on the quantitative results. We simplify the mathematics by using e^{ikr} rather than Bessel functions for the two-dimensional s-waves. We assume scattering amplitudes proportional to the scattering length for each scatterer, and a phase shift equal to the scattering length times the wave number. The quantity of interest is the flux back through the point contact as a result of the scattering. We will look for constructive or destructive interference of the returning waves at the point contact, and take that as our signal.

Take a random distribution of s-wave scatterers (the impurities) at the points $\{r_i\}$ with scattering lengths $\{a_i\}$, and assume phase-coherent transport over the round-trip distances. Let the wave from the QPC be just $r^{-1/2}e^{ikr}$, and the scattered wave from a point scatterer, measured at the QPC, be $(ca_i/r_i)e^{ik(2r_i+a_i)}$. We have called the constant of proportionality between the scattering length and amplitude c. The actual value of this constant will depend on details of the scattering potentials irrelevant to this model. Note also that there are two factors of $r_i^{-1/2}$, one for the falloff of the wave illuminating the scatterer and one for the falloff of the scattered wave, and that the phase advances by the round-trip distance plus the phase shift. Let the tip be at a radius r_t and have the scattering length a_t, giving a similar return wave. Finally, to simplify the notation, let us define $r_i' \equiv r_i + a_i/2$.

The full return wave at a single energy is

$$\sum_i \frac{ca_i}{r_i} e^{2ikr_i'} + \frac{ca_t}{r_t} e^{2ikr_t'}. \tag{5.13}$$

We are interested in the returning signal, so we take the absolute square of this wave. We concentrate on the cross terms, which will give rise to the oscillations with r_t. The cross terms are

$$s(r,k) = 2 \ \mathrm{Re} \left[\sum_i \frac{c^2 a_i a_t}{r_i r_t} e^{2ik(r_i' - r_t')} \right]. \tag{5.14}$$

We should take note at this point of the terms that we are neglecting. First, there are the terms independent of r_t. Since in any physical system we would be detecting a change in the conductance, this constant background would simply figure into our baseline. Second, there is the term $c^2 a_t^2 / r_t^2$, a generally expected monotonic signal independent of energy.

Now we need to thermally average this signal using Eq. (5.12). The thermally averaged signal, averaged after the absolute square so that it is an incoherent sum, is

$$s(r) = 2\pi^{-1/2} \ell_T \ \mathrm{Re} \left[\int dk \ e^{-(k-k_0)^2 \ell_T^2} \sum_i \frac{c^2 a_i a_t}{r_i r_t} e^{2ik(r_i' - r_t')} \right]. \tag{5.15}$$

Note that we can bring the selection of the real part outside of the integral, since all other terms in the expression are real. We carry the imaginary part through the thermal average, since it makes the integral easier. Performing the resulting Gaussian integral, we have

$$s(r) \quad = \quad 2\pi^{-1/2} \ell_T \ \mathrm{Re} \left[\sum_i \frac{c^2 a_i a_t}{r_i r_t} \int dk \ e^{-(k-k_0)^2 \ell_T^2} e^{2i(k-k_0)(r_i' - r_t')} e^{2ik_0(r_i' - r_t')} \right]$$

$$(5.16)$$

$$= \quad 2\pi^{-1/2}\ell_T \ \mathrm{Re}\ \left[\sum_i \frac{c^2 a_i a_t}{r_i r_t} e^{2ik_0(r_i'-r_t')} \int dk\ e^{-(k-k_0)^2\ell_T^2} e^{2i(k-k_0)(r_i'-r_t')}\right] \tag{5.17}$$

$$= \quad 2\ \mathrm{Re}\ \left[\sum_i \frac{c^2 a_i a_t}{r_i r_t} e^{2ik_0(r_i'-r_t')} e^{-(r_i'-r_t')^2/\ell_T^2}\right]. \tag{5.18}$$

Taking the real part, we are left with

$$s(r) = 2\sum_i \frac{c^2 a_i a_t}{r_i r_t} \cos[2k_0(r_i' - r_t')] e^{-(r_i'-r_t')^2/\ell_T^2}. \tag{5.19}$$

We see the following in this result. The wave scattered from the tip interferes with the background of waves scattered from the impurities. After the thermal average, most of the resulting signal is lost. The pieces that survive the average are contributions from those scatterers that are close to the same radius[2] from the QPC as is the tip. Though we can have a signal when $r_t' > \ell_T$, the thermal length still plays a role in that it determines the width of the band around r_t' that contributes to the thermally averaged signal.

Note that the fringes predicted by this model are at half the Fermi wavelength, as observed. Furthermore, the fringes will be oriented perpendicular to the direction of electron flow, also as observed.

In Figure 5.3, we show examples of $s(r)$ at two temperatures for the same distribution of scatterers. To make the signal easier to observe, we divide out the overall radial dependence of the signal strength. From Eq. 5.19 we determine that there is a clear r^{-2} dependence. However, there is a less obvious factor of $r^{1/2}$ that appears as well, giving the signal strength an overall $r^{-3/2}$ dependence. The $r^{1/2}$ comes from

---

[2] To be completely correct, we should substitute "path length" for "radius."

a density dependence of the signal. If we simply take the sum σ of N cosines of random phase, we find that $\langle|\sigma|\rangle \propto N^{1/2}$. Noting that the number of scatterers in our ℓ_T-wide band increases approximately linearly with radius, we have a resulting $r^{1/2}$ modification of the fringe strength.

The radial variation in fringe strength suggested by this model is, unfortunately, made difficult to observe experimentally by other variations of signal strength. Experiments are currently being planned and performed that should provide more direct evidence of this model [34].

5.2.3 Single-Scattering Continuum Limit

If we change the sum in Eq. (5.19) to an integral over the plane and let $a_i = a\ \forall i$, we go from a system of discrete scattering centers to a continuum approximation with constant density. We have

$$\begin{aligned}
& \int_0^\infty dr\, r \int_0^\pi d\theta\, 2\frac{c^2 a a_t}{r r_t} \cos[2k_0(r'-r'_t)] e^{-(r'-r'_t)^2/\ell_T^2} \\
= \;& \mathrm{Re}\left\{\int_0^\infty dr \int_0^\pi d\theta\, 2\frac{c^2 a a_t}{r_t} e^{2ik_0(r'-r'_t)} e^{-(r'-r'_t)^2/\ell_T^2}\right\} && (5.20)\\
= \;& 2\pi\frac{c^2 a a_t}{r_t}\, \mathrm{Re}\left\{\int_0^\infty dr\, e^{2ik_0(r'-r'_t)} e^{-(r'-r'_t)^2/\ell_T^2}\right\} && (5.21)\\
= \;& 2\pi\frac{c^2 a a_t}{r_t}\, \mathrm{Re}\left\{\int_0^\infty dr\, e^{2ik_0[r-(r'_t-a/2)]} e^{-[r-(r'_t-a/2)]^2/\ell_T^2}\right\} && (5.22)\\
= \;& \pi^{3/2}\ell_T\frac{c^2 a a_t}{r_t} e^{-k_0^2\ell_T^2}\left\{1 + \mathrm{Re}\,\left[\mathrm{erf}\left(ik_0\ell_T + (r'_t - a/2)/\ell_T\right)\right]\right\}. && (5.23)
\end{aligned}$$

The expression with the error function is oscillatory, but bounded in magnitude. As a result, the dominant term for the magnitude of this signal is the exponential $e^{-k_0^2\ell_T^2}$. For the values applicable in the physical systems that we have considered,

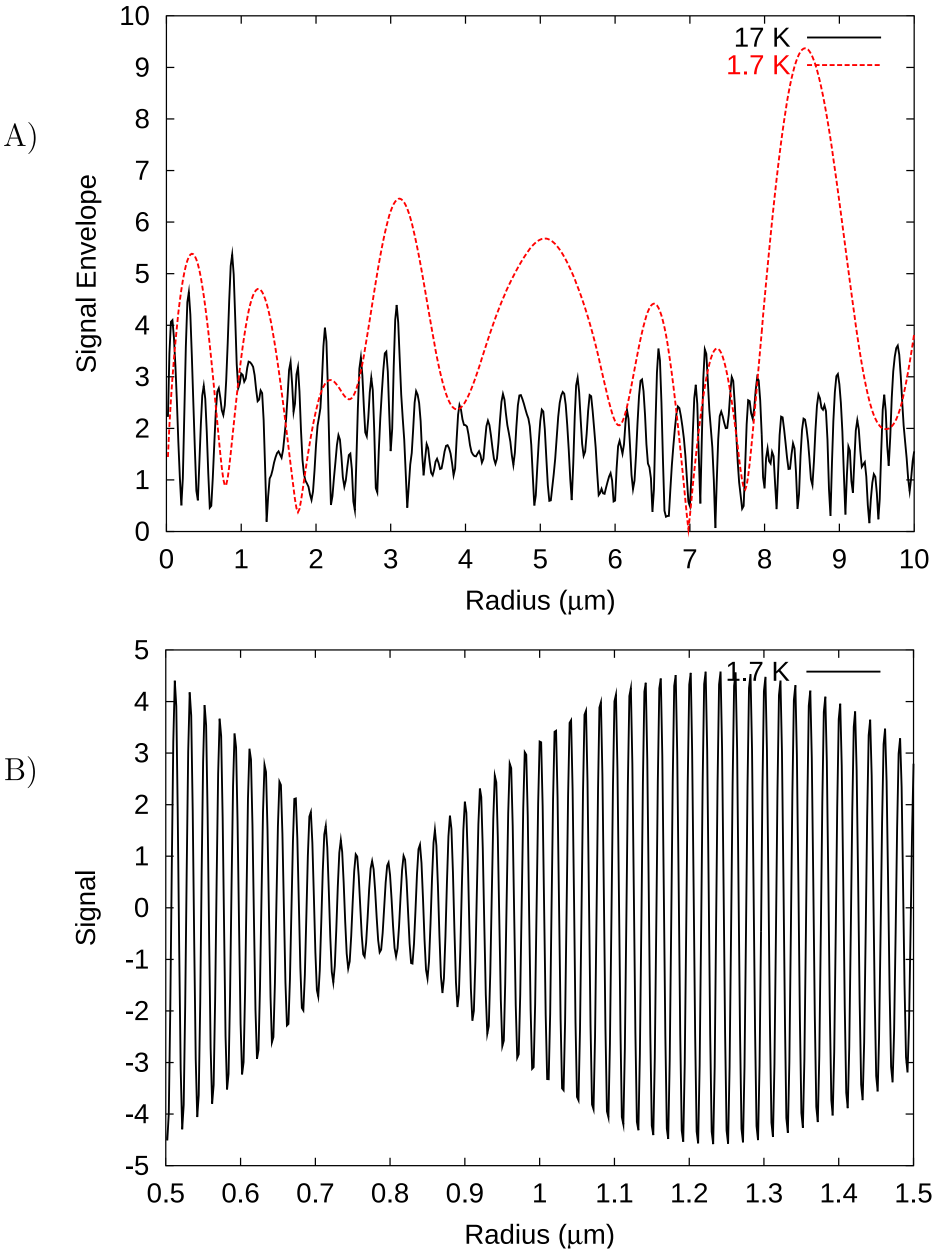


Figure 5.3: Here we see examples of the signal generated by Eq. 5.19. In all cases, we show $r^{3/2}s(r)$. In (A), we have the envelopes of the signal at two different temperatures for the same distribution of scatterers. The scatterers have a density of 40 μm^{-2}, and $\lambda_0 = 40$ nm. These demonstrate how the decreased thermal length at higher temperatures results in a smaller signal with more rapid variations with radius. In (B), we show the signal at 1.7 K over a range more like that scanned experimentally.

this is vanishingly small; in this continuum limit, the signal disappears.[3]

There are reasons to be skeptical about this continuum limit. Notably, it assumes a completely uniform distribution of scatterers over the plane. This assumption is in opposition to the randomness of the systems of interest. We already know from multiple-scattering theory that such a regular distribution can simply appear as a mean field, effectively "raising the floor" of the potential over which the electrons propagate.

Furthermore, this continuum-limit calculation tells us nothing about how that limit is approached. As noted previously, simply taking the single-scattering model in Eq. (5.19) and increasing the density d of scatterers without any correlation in their placement results in a $d^{1/2}$ dependence of the signal strength. Clearly, however, at some density the single-scattering model ceases to be a useful one. Multiple-scattering calculations have shown that, for the densities of scatterers considered here, the single-scattering approximation is appropriate. As the density of scatterers increases and multiple scattering becomes more important, the signal ceases to be as simple to understand [35].

5.3 Multiple Scattering and the Phase-Coherence Length

Another model that could account for fringing beyond the thermal length is one that appeals to higher-order scattering and the formation of local resonances. If

---

[3] Note that the math is simpler but the answer is incorrect if you take the integral over r over the range $(-\infty, \infty)$.

the tip is within the thermal length of impurity (and, implicitly, within the phase-coherence length), then a resonance can form between the two. The nature of this resonance could then either enhance or suppress back-scattering to the QPC, thus changing the measured conductance.

This model was ultimately rejected as an explanation for the fringes that we were seeing because it could not adequately explain the orientation of those fringes perpendicular to the local electron flux; fringes from this model would be centered on the local impurity participating in the resonance. A consideration of this model led us to the conclusion, however, that it may be possible to observe interference fringes not only beyond the thermal length, but also beyond the phase-coherence length. The loss of phase information between the QPC and the tip has no effect on the local resonance under consideration. The intuition behind this should be familiar, as we are accustomed to the idea of measuring the properties of coherent systems with equipment in the lab, which involves much more incoherent transport than coherent.

Note that, in the limit considered, a signal dependent upon multiple scattering off of impurities should be considerably weaker than one from single scattering. In the region where the single-scattering model discussed above is applicable, we would expect that mechanism to dominate. Only beyond the phase-coherence length, where the single-scattering model fails, would we expect this multiple-scattering contribution to be seen.

In one dimension, we can work out explicitly what is happening with scattering beyond the phase-coherence length. The results found this way are simple but sufficient to make our point. We will work in the transfer matrix (T-matrix) formalism (see

[21, Ch. 5] for a more complete discussion). This is another matrix matrix method, in addition to our S-matrix, for describing scattering in terms of channels. Whereas the S-matrix connects incoming modes to outgoing modes, the T-matrix connects incoming and outgoing modes on one side of the system to incoming and outgoing modes on the other side. For example, considering the scatterer Q in Figure 5.4, we would have

$$\begin{pmatrix} B \\ C' \end{pmatrix} = S_Q \begin{pmatrix} A \\ D' \end{pmatrix}, \tag{5.24}$$

but

$$\begin{pmatrix} C' \\ D' \end{pmatrix} = \mathrm{T}_Q \begin{pmatrix} A \\ B \end{pmatrix}. \tag{5.25}$$

The same information is contained in the S-matrix and the T-matrix, and one can transform between the two. The advantage of the transfer-matrix formalism is that the transfer matrix for a composite system can be constructed by matrix multiplication of the transfer matrices for the components.

If one solves for the transmission and reflection coefficients of a system, requiring flux conservation, one finds that the elements of the transfer matrix can be written in the form

$$\mathrm{T} = \begin{pmatrix} 1/t^* & r^*/t^* \\ r/t & 1/t \end{pmatrix}, \tag{5.26}$$

where t and r are complex and include phase shift information, and the transmission and reflection coefficients are given by $T = |t|^2$ and $R = |r|^2$. (Note that we are writing the transfer matrix as “T” to avoid confusion with the transmission coefficient “T.”)

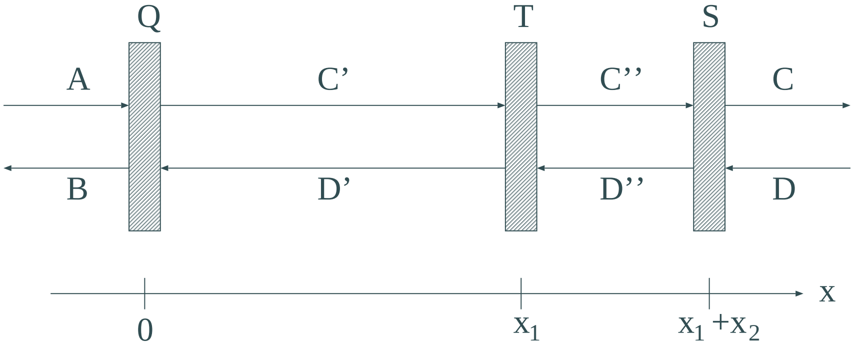


Figure 5.4: Here we see the one-dimensional scattering setup described in §5.3. The scatterers labeled Q, A, and S are playing the roles of the QPC, AFM tip, and impurity scatterer respectively.

Let us define the matrix

$$A(\phi) = \begin{pmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{pmatrix}. \tag{5.27}$$

This matrix is the transfer matrix for propagation over free space, where the wave accumulates a phase ϕ. This matrix is useful because it allows us to treat any scatterer as if it were located at the origin; we merely define the matrix $\mathrm{T}_i(x_i)$ for a scatterer located at position x_i and take

$$\mathrm{T}_i(x_i) = A^{-1}(kx_i)\mathrm{T}_i(0)A(kx_i). \tag{5.28}$$

We note that the simple form $A(\phi)$ allows us to write $A^{-1}(\phi) = A(-\phi)$ and $A(\phi_0 + \phi_1) = A(\phi_0)A(\phi_1)$. Assuming phase-coherent transport, we can now write the T-matrix for the full system in Figure 5.4 as

$$\mathrm{T}_{STQ} = A^{-1}(kx_1 + kx_2)\mathrm{T}_S(0)A(kx_1 + kx_2)A^{-1}(kx_1)\mathrm{T}_T(0)A(kx_1)\mathrm{T}_Q(0) \tag{5.29}$$

$$= A(-kx_1)\left[A(-kx_2)\mathrm{T}_S(0)A(kx_2)\mathrm{T}_T(0)\right]A(kx_1)\mathrm{T}_Q(0) \tag{5.30}$$

$$\equiv A^{-1}(kx_1)\mathrm{T}_{ST}(0)A(kx_1)\mathrm{T}_Q(0), \tag{5.31}$$

working from right to left as dictated by the definition of the T-matrices. We see that the formalism allows us to treat the scatterer and tip system separately from their interaction with the QPC. We can now introduce de-phasing in the following simple manner: we add an additional, random phase to that accumulated between the QPC and the tip. That is, we replace the matrix $A(kx_1)$ with $A(kx_1+\phi)$ and $A^{-1}(kx_1)$ with $A^{-1}(kx_1+\phi)$, where ϕ is some random variable with distribution $P(\phi, x_1)$. We include a dependence on x_1 because we expect the distribution to widen as we propagate over greater distances. We assume that we have phase-coherent transport over the distance x_2 between the tip and the scatterer.

Let's look first at the generic case,

$$\mathrm{T}_{BA} = A^{-1}(\alpha)\mathrm{T}_B A(\alpha)\mathrm{T}_A. \tag{5.32}$$

With the general phase accumulation α, we'll be able to apply this to our sub-system T_{ST} and, in turn, the full system. We find that

$$t_{BA}^{-1} = (t_B t_A) + \left(e^{2i\alpha} r_B r_A^*\right)(t_B t_A^*)^{-1} \tag{5.33}$$

by taking the matrix multiple and simplifying. To simplify notation, let us define $t_j = T_j^{1/2} e^{i\delta_j}$ and $r_j = R_j^{1/2} e^{i\rho_j}$. We then find the transmission coefficient as

$$T_{BA} = |t_{BA}|^2 = \frac{T_A T_B}{1 + R_A R_B + 2(R_A R_B)^{1/2}\cos\left(2\alpha + 2\delta_A + \rho_B - \rho_A\right)}. \tag{5.34}$$

Let us define $\Phi_{BA} \equiv 2\delta_A + \rho_B - \rho_A$ to shorten the expression somewhat to

$$T_{BA} = \frac{T_A T_B}{1 + R_A R_B + 2(R_A R_B)^{1/2}\cos\left(2\alpha + \Phi_{BA}\right)}. \tag{5.35}$$

Using Eq. 5.35, we can find the expression for the conductance of either a coherent or an incoherent system. Let us look first at the coherent part,

$$\mathrm{T}_{ST}(0) = A(-kx_2)\mathrm{T}_S(0)A(kx_2)\mathrm{T}_T(0). \tag{5.36}$$

The transmission of this system is Eq. 5.35 with $\alpha = kx_2$ and the appropriate substitutions for transmission and reflection coefficients:

$$T_{ST} = \frac{T_T T_S}{1 + R_T R_S + 2(R_T R_S)^{1/2} \cos\left(2kx_2 + \Phi_{ST}\right)}. \tag{5.37}$$

This is the usual resonant-tunneling expression, as we should expect. Looking now at the incoherent part, we need to take $\alpha = kx_1 + \phi$ and perform our integral over the random phase distribution function $P(\phi, x_1)$. Working from the definition of T_{STQ} in Eq. 5.31,

$$T_{STQ} = \int_{-\infty}^{\infty} d\phi \, \frac{T_Q T_{ST} P(\phi, x_1)}{1 + R_Q R_{ST} + 2(R_Q R_{ST})^{1/2} \cos\left(2\phi + 2kx_1 + \Phi_{STQ}\right)}. \tag{5.38}$$

Let us assume that we are well beyond the phase-coherence length, and we can reasonably approximate $P(\phi, x_1)$ as a constant; that is, any phase is equally likely. Taking advantage of the angular nature of the phase, we can write this as

$$T_{STQ} = \int_{0}^{2\pi} d\phi \, \frac{T_Q T_{ST}/2\pi}{1 + R_Q R_{ST} + 2(R_Q R_{ST})^{1/2} \cos\left(2\phi + 2kx_1 + \Phi_{STQ}\right)} \tag{5.39}$$

$$= \frac{T_Q T_{ST}}{1 - R_Q R_{ST}}. \tag{5.40}$$

This expression should also look familiar; we could write it down from first principles, considering the transmission of two barriers ignoring interference effects (i.e., an infinite sum of terms for multiple bounces between the barriers).

So, the transfer-matrix formalism shows us that the system as a whole behaves as we would expect: we have resonance between scatterer and tip that doesn't care about incoherent transport back to the QPC. The combination of resonant system and QPC behaves like a simple transmission system without interference effects, and the fringes from the resonance will survive.

Chapter 6

Stability and the Stability Matrix

There are several tools that become available when one deals with classical trajectories rather than quantum-mechanical waves. One of those tools is the analysis of the stability matrix, and the information that it carries about local behavior [36, Course 9].

The stability matrix is generally seen as applied in systems with periodic orbits. Part of the goal of this chapter, apart from spelling out a derivation of the stability matrix equations that may be clearer than those usually seen, is to highlight the errors that can be made if one takes periodic-orbit stability and naively applies it in a system without periodic orbits.

We begin by looking at what the stability matrix $\mathbf{M}$ tells us about a classical trajectory. With this understanding, we find an expression for the stability matrix in a simple context with discretized time. Taking the discretization to zero, we arrive at the differential equation to be solved for the elements of the stability matrix. We then take a second look, arriving at the same differential equation by a more direct but perhaps less clear route. Next, we look at some properties of the stability matrix that will be useful in understanding it. Finally, we look at what this matrix tells us about periodic trajectories and non-periodic trajectories.

6.1 Local Dynamics

The stability matrix tells us about the dynamical behavior local to a given trajectory. The idea of “local dynamics” is a fairly simple one. Given a reference trajectory propagating through some complicated potential, we ask what is happening to other trajectories “near” the reference. The notion of “nearness” here is in the sense of the neighborhood of the reference trajectory in phase space. However complicated the potential, there will, in general, be some neighborhood small enough that we can make linear approximations of the dynamics. The stability matrix will describe those local, linearized dynamics. One must be careful to note that the stability matrix will, therefore, be a property associated with an individual trajectory. Furthermore, it is a time-dependent quantity.

Let us make the question more concrete. We define a trajectory in N spatial dimensions as

$$\vec{x}(t) = (q_0(t), \ldots, q_{N-1}(t), p_0(t), \ldots, p_{N-1}(t)) \,. \tag{6.1}$$

That is, we track the location of the trajectory in a full $2N$-dimensional phase space. The specification of all $2N$ coordinates at any time is sufficient to specify a trajectory. Let a trajectory $\vec{x}_0(t)$ be our reference trajectory, and let a second trajectory $\vec{x}(t)$

begin displaced from $x_0(t)$ by some small amount:

$$\vec{x}(0) = \vec{x}_0(0) + \delta\vec{x}(0). \tag{6.2}$$

We will assume that the initial displacement $\delta\vec{x}(0)$ is sufficiently small that $\vec{x}(t)$ remains sufficiently close to $\vec{x}_0(t)$ for all times of interest.

Given this definition, we can take the natural definition of a time-dependent displacement

$$\delta\vec{x}(t) = \vec{x}(t) - \vec{x}_0(t). \tag{6.3}$$

We use the evolution of this displacement vector to define our stability matrix for the trajectory $\vec{x}_0(t)$. Take

$$\delta\vec{x}(t) = \mathbf{M}(t) \cdot \delta\vec{x}(0) \tag{6.4}$$

to define the stability matrix $\mathbf{M}(t)$. Note that the matrix is dependent upon our choice of reference trajectory, but not dependent upon our choice of initial displacement. That is, the stability matrix describes the behavior of all trajectories in an infinitesimal neighborhood of the reference. Using the expression above, we can define the elements of the stability matrix as

$$\mathbf{M}_{ij}(t) \equiv \frac{\partial\, \delta x_i(t)}{\partial\, \delta x_j(0)}. \tag{6.5}$$

6.2 The Stability Matrix via Discrete Time

An intuitive way to approach the derivation of the stability matrix differential equations is by first looking at the evolution of a trajectory under the usual classical dynamics with discrete time step τ. Let us proceed in this fashion, eventually taking $\tau \rightarrow 0$ to arrive at the differential equations.

Let us define a second matrix now, $\mathbf{DM}(t_2, t_1)$ such that

$$\delta\vec{x}(t_2) = \mathbf{DM}(t_2, t_1) \cdot \delta\vec{x}(t_1). \tag{6.6}$$

Our original stability matrix can be written as $\mathbf{M}(t) = \mathbf{DM}(t, 0)$. We can also use $\mathbf{DM}$ matrices to break up the time interval to find $\mathbf{M}$. For example, taking two intervals, we can write

$$\mathbf{M}(t) = \mathbf{DM}(t, t')\mathbf{DM}(t', 0). \tag{6.7}$$

The elements of $\mathbf{DM}$ are found in a manner very similar to those of $\mathbf{M}$,

$$\mathbf{DM}_{ij}(t_2, t_1) \equiv \frac{\partial\, \delta x_i(t_2)}{\partial\, \delta x_j(t_1)}. \tag{6.8}$$

The matrix $\mathbf{DM}$ is useful, at the moment, for tracking the evolution of $\mathbf{M}$ in a single time interval without reference to any other time intervals.

The simplest way to discretize time for a trajectory, going only to first order in τ and working with a one-dimensional system, is to write

$$\begin{aligned} q(t+\tau) &= q(t) + \tau m^{-1} p(t) && (6.9) \\ p(t+\tau) &= p(t) - \tau V_q[q(t)] && (6.10) \end{aligned}$$

where by V_q we mean the q derivative of the potential V.[1] Using these and Eq. (6.8), we find the elements of $\mathbf{DM}(t+\tau, t)$ to be

$$\mathbf{DM}(t+\tau, t) = \begin{bmatrix} 1 & \tau m^{-1} \\ -\tau V_{qq} & 1 \end{bmatrix}. \tag{6.11}$$

---

[1] We use these expressions so that we can, later, take τ to zero and get a differential equation. If one wanted to turn the continuous system into a discrete time map, this isn't the way that one should do it – the resulting map would not be symplectic (area preserving).

To get to a differential equation for $\mathbf{M}$ we note that

$$\begin{aligned}
\mathbf{M}(t+\tau) &= \mathbf{DM}(t+\tau,t)\mathbf{DM}(t,0) && (6.12)\\
&= \mathbf{DM}(t+\tau,t)\mathbf{M}(t) && (6.13)\\
\dot{\mathbf{M}}(t) &\approx \left[\mathbf{M}(t+\tau)-\mathbf{M}(t)\right]\tau^{-1} && (6.14)\\
&\approx \left[\mathbf{DM}(t+\tau,t)\mathbf{M}(t)-\mathbf{M}(t)\right]\tau^{-1} && (6.15)\\
&\approx \left[\mathbf{DM}(t+\tau,t)-\mathbf{I}\right]\mathbf{M}(t)\tau^{-1} && (6.16)\\
&\approx \mathbf{K}(t)\mathbf{M}(t), && (6.17)
\end{aligned}$$

where we have defined the matrix $\mathbf{K}(t)$ in the final step to be

$$\mathbf{K}(t) = \left[\mathbf{DM}(t+\tau,t)-\mathbf{I}\right]\tau^{-1} \tag{6.18}$$

which, writing out the matrix elements, we find to be

$$\mathbf{K}(t) = \begin{bmatrix} 0 & m^{-1} \\ -V_{qq} & 0 \end{bmatrix}. \tag{6.19}$$

We note that τ has dropped out of this expression, so we can take the limit $\tau \to 0$ exactly and write $\dot{\mathbf{M}}(t) = \mathbf{K}(t)\mathbf{M}(t)$.

6.3 A Fuller, More Complicated Derivation

For two reasons, we will now work through a second derivation of the differential equation for $\mathbf{M}(t)$. One reason to do this is that the extension to higher dimensions is not immediately obvious from the one-dimensional results above. The second reason is that, though more opaque, this derivation is perhaps more satisfying.

As before, we are after the elements of $\dot{\mathbf{M}}(t)$. These can be written most intuitively as

$$\dot{\mathbf{M}}_{ij}(t) = \frac{\partial}{\partial t}\left(\mathbf{M}_{ij}\right) = \frac{\partial}{\partial t}\left(\frac{\partial x_i(t)}{\partial x_j(0)}\right). \tag{6.20}$$

We now change the order of the differentiation to have

$$\dot{\mathbf{M}}_{ij}(t) = \frac{\partial}{\partial x_j(0)}\left(\frac{\partial x_i(t)}{\partial t}\right). \tag{6.21}$$

The terms inside the parentheses are entirely "local" in time, and we would rather have a local derivative outside as well. To get that, we expand the derivative using

$$\frac{\partial}{\partial x_j(0)} = \sum_k \frac{\partial x_k(t)}{\partial x_j(0)}\frac{\partial}{\partial x_k(t)}. \tag{6.22}$$

We recognize the coefficients in the summation as elements of $\mathbf{M}$.

We need also to look at the term inside the parentheses and seek to simplify it. We can write it in terms of the Hamiltonian function, though we must be careful about whether $x_i(t)$ is a spatial coordinate or a momentum coordinate – there is a sign difference. We have that

$$\frac{\partial x_i}{\partial t} = \begin{cases} \partial H/\partial p_i & x_i = q_i \\ -\partial H/\partial q_i & x_i = p_i \end{cases}. \tag{6.23}$$

Putting it all together, we have that

$$\dot{\mathbf{M}}(t) = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{bmatrix} \cdot \left[\frac{\partial^2 H}{\partial x_i \partial x_j}\right] \cdot \mathbf{M}(t) = \mathbf{K}(t)\mathbf{M}(t), \tag{6.24}$$

where $\mathbf{I}$ is the identity matrix of rank equal to the dimensionality of the system. The matrix given by $\partial^2 H/\partial x_i \partial x_j$ is the Jacobian of the gradient of H. In most cases,[2]

---

[2] If there are terms in the Hamiltonian involving products of position and momentum coordinates, there will be terms in the diagonal blocks of $\mathbf{K}$.

this gives us

$$\mathbf{K}(t) = \begin{bmatrix} 0 & \mathbf{I}m^{-1} \\ -\partial^2 V/\partial q_i \partial q_j & 0 \end{bmatrix} \tag{6.25}$$

In one dimension, this definition reproduces the matrix **K** that we saw in our previous derivation.

6.4 The Unit Determinant of the Stability Matrix

There is one restriction on the stability matrix that needs to be noted: it must maintain a determinant of 1. This comes from the area-preserving requirements of classical mechanics as set out in Liouville's theorem, as we will now show. Since it is not always obvious, let's look first at a proof of Liouville's theorem. Let $f(\vec{x}, t)$ be a density function in phase space. That is, it gives the density of trajectories in phase space as a function of the $2N$ dimensional phase-space position $\vec{x}$ and time. The total time derivative df/dt gives the density following the flow, and

$$\frac{df}{dt} = \partial_t f + \dot{\vec{x}} \cdot \nabla f, \tag{6.26}$$

where

$$\nabla \equiv \sum_{i=1}^{N} \hat{q}_i \frac{\partial}{\partial q_i} + \hat{p}_i \frac{\partial}{\partial p_i} \tag{6.27}$$

is the del operator in phase space. By the continuity theorem,

$$\partial_t f = -\nabla \cdot (f \dot{\vec{x}}). \tag{6.28}$$

Using this as a substitution, we have

$$\frac{df}{dt} = -\nabla \cdot (f \dot{\vec{x}}) + \dot{\vec{x}} \cdot \nabla f \tag{6.29}$$

$$= -\dot{\vec{x}} \cdot \nabla f - f \nabla \cdot \dot{\vec{x}} + \dot{\vec{x}} \cdot \nabla f \tag{6.30}$$

$$= -f \nabla \cdot \dot{\vec{x}} \tag{6.31}$$

$$= -f \sum_{i=1}^{N} \frac{\partial}{\partial q_i} \dot{q}_i + \frac{\partial}{\partial p_i} \dot{p}_i \tag{6.32}$$

$$= -f \sum_{i=1}^{N} \frac{\partial}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial}{\partial p_i} \frac{\partial H}{\partial q_i} \tag{6.33}$$

$$= 0. \tag{6.34}$$

where we have used Hamilton's equations. This shows that, for a Hamiltonian system, the density is preserved in phase space.

Put another way, phase-space volumes are preserved. We now relate this statement to our stability matrix. Recall that the N-dimensional volume defined by N vectors is given by the determinant of a matrix with the vectors as columns.[3] If we define a matrix $\mathbf{D}(0)$ with N displacement vectors as the columns, we can act on all of them simultaneously by taking $\mathbf{D}(t) = \mathbf{M}(t)\mathbf{D}(0)$. By the properties of the determinant, we have $|\mathbf{D}(t)| = |\mathbf{M}(t)||\mathbf{D}(0)|$, so the preservation of the N-dimensional volume defined by $\mathbf{D}(t)$ requires that $\mathbf{M}(t)$ have unit determinant.

6.5 A First Look at Stability

Before we define stability in a mathematical sense, it is useful to discuss what we mean by it in a general sense. Stability is a property of a given trajectory, and comes from the local, linearized dynamics as represented by the stability matrix. We consider a trajectory to be stable if, in some time scale of interest, trajectories that start close to it tend to remain close to it. One may recall that one definition of chaos

---

[3] In two dimensions, this reduces to the magnitude of the cross product, and in three dimensions to the scalar triple product.

is that trajectories that begin close to one another diverge exponentially with time; this is an example of instability.

However, as we shall see, this consideration leads to quite different mathematical descriptions depending on the type of system being considered. Note that, in higher dimensions, there are multiple, orthogonal choices for the displacement vector. As a result, there might be some ambiguity in this definition of stability. In different cases, we will resolve this ambiguity in various ways.

6.6 Reducing Dimensionality

Before we go further, we must make a seeming digression for a moment and mention the method used to reduce the dimensionality of a system. Note that we will always be speaking in terms of potentials in two dimensions, though these notions are by no means confined to low-dimensional systems.

If we have a potential in two dimensions, there are four phase-space coordinates to track for a given trajectory, all of them time dependent. As noted above, the specification of all phase-space coordinates at any time completely determines a trajectory. In most physical systems, however, we have additional information. If we assume a static potential (i.e., time independent), then the energy of our particle will be conserved. This conservation of energy reduces the dimensionality of accessible phase space by one dimension. Hence, for example, a trajectory can be completely specified by giving the two spatial coordinates and a direction of propagation; the magnitude of the momentum can be determined by energy conservation.

A three-dimensional phase space still presents problems for visualization. The usual practice is to reduce the dimensionality further by taking two-dimensional slices of phase space. Suppose, for example, we take a plane of fixed x-coordinate and look at where all trajectories pass through this plane. Since the x-coordinate is given by the location of the plane, all of our trajectories are completely determined by specifying the y-coordinate where they pass though the plane and the angle of propagation. It is generally the practice to take only trajectories that pass through the plane from one given side to the other, so that the y momentum can be used unambiguously as the second coordinate rather than angle.

This technique is called a “surface of section,” and it is useful in both closed and open systems. It is discussed in many books that deal with chaotic classical dynamics (see, for example, [37]).

6.7 Stability and Closed Systems

The usual system where one hears about stability analysis is a closed system with period orbits. Here, one finds the stability matrix for a single period of a periodic orbit, calls the result the “monodromy matrix,” and extracts stability information from it.

The first thing to do is to select a surface of section in the domain of the system. The surface of section defines a discrete dynamics on top of our continuous Newtonian dynamics. Recalling that a point on the surface of section is the intersection of a trajectory with that surface, we have a map that takes a trajectory from one intersection point to its next.

The stability is now observed on the surface of section, and the following question is asked: given a point on the surface of section that corresponds to a periodic reference trajectory, what happens to neighboring points under n iterations of the map? Note that, by definition, no matter how many times we iterate the map, the reference point is always mapped onto itself.

By looking at the underlying dynamics, we can determine a monodromy matrix $\mathbf{M}$ for our reference trajectory for a single iteration of the map; this is essentially the stability matrix for the trajectory for a single period, projected into the subspace of the surface of section. The standard measure of stability, called the Lyapunov exponent for the trajectory, is the logarithm of the largest eigenvalue of this matrix. Why this measure? Take any general displacement (in the surface of section) $\delta\vec{x}_0$ and take n iterations by

$$\delta\vec{x}_n = \mathbf{M}^n \cdot \delta\vec{x}_0. \tag{6.35}$$

We are interested in how the initial displacement is growing, so we want the length of $\delta\vec{x}_n$. Because the effect of the stability matrix at each iteration is multiplicative, we can determine the net effect of a single iteration by taking

$$h \;=\; \lim_{n\to\infty}\left[\frac{1}{n}\ln|\delta\vec{x}_n|\right]. \tag{6.36}$$

The eigenvalues λ_i and eigenvectors $\vec{v}_i$ of the matrix $\mathbf{M}$ allow us to expand the initial displacement

$$\delta\vec{x}_0 = \sum_i c_i\vec{v}_i. \tag{6.37}$$

Under n iterations of the map, we then have that

$$\mathbf{M}^n \cdot \delta\vec{x}_0 = \sum_i c_i\lambda_i^n\vec{v}_i. \tag{6.38}$$

If we order our eigenvalues such that $\lambda_0 \geq \lambda_i$, we can say that

$$\begin{aligned} \mathbf{M}^n \cdot \delta\vec{x}_0 &= \lambda_0^n \sum_i c_i \left(\frac{\lambda_i}{\lambda_0}\right)^n \vec{v}_i && (6.39) \\ \lim_{n\to\infty} \mathbf{M}^n \cdot \delta\vec{x}_0 &\approx \lambda_0^n c_0 \vec{v}_0 && (6.40) \end{aligned}$$

for a generic $\delta\vec{x}_0$. (In particular, it will not be true for those particular $\delta\vec{x}_0$ that happen to be orthogonal to $\vec{v}_0$.) Using this result, we have that

$$\begin{aligned} h &= \lim_{n\to\infty}\left[\frac{1}{n}\ln|\delta\vec{x}_n|\right] && (6.41) \\ &\approx \frac{1}{n}\ln|\lambda_0^n \vec{v}_0 c_0| && (6.42) \\ &= \frac{1}{n}\ln(\lambda_0^n c_0) && (6.43) \\ &= \ln\lambda_0 + \frac{1}{n}\ln c_0 && (6.44) \\ &\approx \ln\lambda_0. && (6.45) \end{aligned}$$

We see that, if we are considering the limit of iterations of a periodic trajectory, it is useful to consider the eigenvalues of the monodromy matrix. Trajectories can be put into three general categories by this method by looking at the traces, or sums of the eigenvalues, of their monodromy matrices. Note that our original stability matrix had a unit determinant, and this property carries over to the monodromy matrix. This is true, even though the surface of section formulation does not guarantee that trajectories will take the same time for an iteration of the map, as a result of the Poincaré-Cartan theorem [38]. For the two-dimensional phase space defined by the surface of section, there are only two eigenvalues to consider and, as a result, they must be reciprocals.

If the trace of the stability matrix is greater than 2, then we know that both eigenvalues are real. One is greater than 1, the other less than 1. In this case, there will be a positive Lyapunov exponent, and the trajectory is considered to be unstable. Neighboring trajectories will tend to run away from the reference trajectory exponentially with n.

If the trace of the stability matrix is less than 2, then the eigenvalues must be complex; they will be complex conjugates of one another. In this case, the Lyapunov exponent is zero.

If the trace of the monodromy matrix is exactly 2, then both eigenvalues are equal to 1 and the trajectory is considered marginally stable. In this case, trajectories can move away from the reference trajectory, but they do so with a linear dependence on n rather than exponential. As an example, consider the case of "shearing":

$$\mathbf{M} = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix} \tag{6.46}$$

$$\mathbf{M}^n \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a + ndb \\ b \end{pmatrix}. \tag{6.47}$$

6.8 Stability and Open Systems

The measure of stability found above for periodic trajectories in closed systems becomes fairly meaningless when we consider open systems where the trajectories of interest are not periodic. Here, trajectories will not traverse the same history of the potential multiple times, so there is no true sense of iterations or taking limits as $n \to \infty$.

There may still be a temptation, however, to use the Lyapunov exponent as a measure of stability. It is, after all, a comfortable concept. Let us consider an extreme example to see why this notion fails. Let

$$\mathbf{M}(t) = \begin{pmatrix} 0 & a \\ -1/a & 0 \end{pmatrix} \tag{6.48}$$

for some time t. In this case, the determinant is 1 (as required), the eigenvalues are $\pm i$, the trace is 0, and the trajectory would be declared stable by the Lyapunov exponent measure. However, we see that if we take $\delta\vec{x}(0) = (0, 1)$, we have $\delta\vec{x}(t) = (a, 0)$. We can make this final displacement as large as we wish by our choice of a, but the Lyapunov exponent will remain 0.

What do we use to replace the notion of Lyapunov exponent stability? The concept that we are after is the same: how much do neighboring trajectories move away from our reference trajectory? Let us look at this first in a general sense, where we have no special displacement directions to consider (we will soon consider manifolds of trajectories where we will have special directions). We state the question mathematically as follows: given a stability matrix $\mathbf{M}$, we look for the direction of initial displacement that results in the greatest stretching. We can take our initial displacement as $\delta\vec{x}(0) = (\cos\theta, \sin\theta)$ and take[4]

$$\mathbf{M}(t) \cdot \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} = \delta\vec{x}(t; \theta). \tag{6.49}$$

We then want to maximize $|\delta\vec{x}(t)|$ as a function of θ. If we label the elements of the

---

[4] We still want to be working within the infinitesimal neighborhood around the reference trajectory where the linear dynamics in $\mathbf{M}$ are a reasonable approximation. Since the action of $\mathbf{M}$ is linear, however, we can work with vectors of norm 1 to make life simpler; sticklers can multiply all equations by some infinitesimal δ without changing anything else.

stability matrix

$$\mathbf{M}(t) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{6.50}$$

we find that

$$\max|\delta\vec{x}(t;\theta)|^2 = \frac{1}{2}\left\{a^2+b^2+c^2+d^2+\left[\left(a^2-b^2+c^2-d^2\right)^2+4\left(ab+cd\right)^2\right]^{1/2}\right\}. \tag{6.51}$$

If we have no reason to be interested in any initial displacement over any other, we should then take

$$\frac{1}{2}\ln\left(\max|\delta\vec{x}(t;\theta)|^2\right) \tag{6.52}$$

as a replacement for the Lyapunov exponent.

Suppose that we are not interested in the behavior of an isolated trajectory, but rather in the behavior of a collection of trajectories, or manifold. For example, we may take trajectories emanating from a single point at a distribution of angles. In the case of a manifold, it is not likely that we are interested in the general statement above. Eq. (6.51) determines the maximum possible stretching, taking into account all possible initial displacement directions. In our example, all possible initial displacements about a given reference trajectory are not represented by the manifold. In fact, we can specify a single vector $\hat{d}$ for a given trajectory that points along the initial manifold, and only initial displacements along $\hat{d}$ appear in our infinitesimal neighborhood. For a picture to go along with this situation, see Figure 6.1.

Let us then define a "rarefaction exponent" $r(t)$ by

$$r(t) = \ln\left|\mathbf{M}(t)\cdot\hat{d}\right|. \tag{6.53}$$

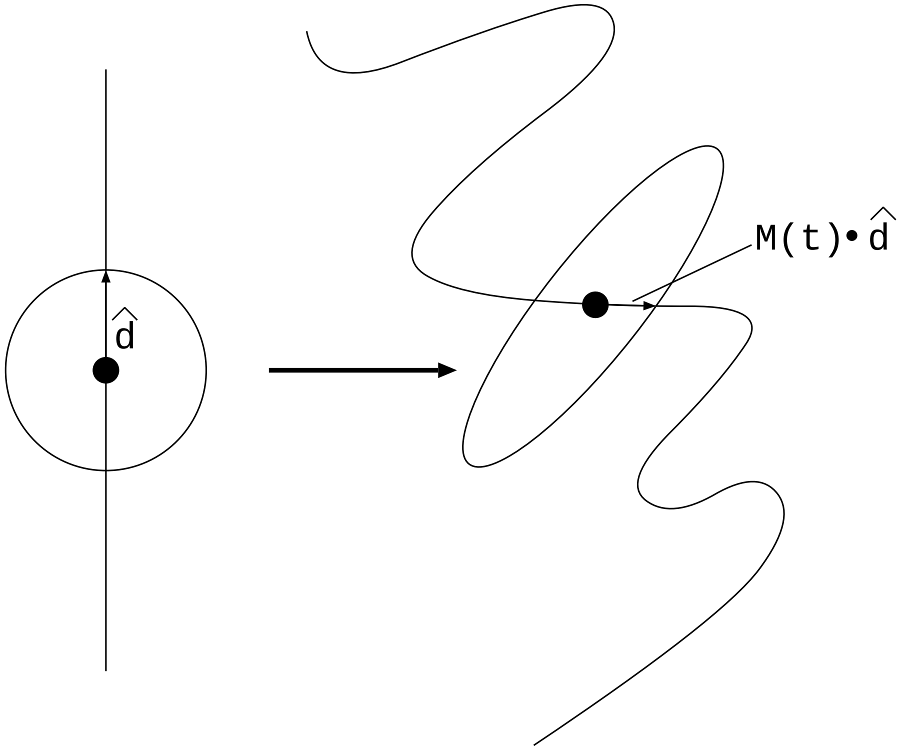


Figure 6.1: This figure is illustrative of the local dynamics around a reference trajectory when it is a member of a manifold of trajectories. The manifold is a one-dimensional object in our phase space, beginning on the left as a simple line (for example, all trajectories starting at the same point with a distribution of angles). Over time, it will evolve to some complex but still one-dimensional object (ignoring a possible fractal dimension as time goes to infinity). The action of the local, linear dynamics will turn some circle of displacements about our reference trajectory into some ellipse. We can see, however, that the maximum stretching of this ellipse isn't as relevant as the stretching of the vector $\hat{d}$ that lies along the manifold.

The rarefaction exponent tells us how much trajectories displaced along the initial manifold have moved away from the reference trajectory. Put another way, it tells us how the dynamics have compressed or expanded the manifold itself around the reference trajectory (hence the name "rarefaction exponent"). It is a property of a given trajectory, and is clearly dependent on the choice of initial manifold. This limited notion is not stability in the usual sense of the word, but in many classical systems it may be significantly more meaningful.

Chapter 7

Ensemble Averages

One of the themes of the research in this thesis is that, when dealing with random systems, there is interesting behavior to be found by looking at individual realizations of the random potentials. It was only by doing so that we came upon the branched flow structure. There are, however, reasons to look at the statistical behavior seen in ensemble averages as well. The behavior of any individual member of the ensemble (i.e., the branched flow in a given potential realization with given initial conditions) can only be determined by actually running the trajectories. The behavior of the ensemble averages, however, are in some cases amenable to analytic treatment. This is part of the strength that classical simulations bring to the table, with the wealth of information about individual trajectories that is available to us.

In this chapter, we will look at several ways that we can characterize the statistics of our random potentials and the classical branched flow running through them. There are two equivalent ways to arrive at these statistical measures, to which we are referring generally as "ensemble averages." The first is to run trajectories in many different potentials generated in the same way (i.e., using the same functions but with different random seeds). The other is to run many independent trajectories in the same potential. Here, "independence" means that the trajectories should see uncorrelated histories as they propagate through the potential. This requirement can be realized by launching trajectories from random locations separated on average by the potential correlation length, and giving them random initial orientations. In practice, we use a combination of these two approaches to ensure that we aren't picking up particular features of any one realization of the random potential.

A word should perhaps be said about how this procedure relates to the pictures of branched flow that appear elsewhere in this thesis. Those pictures arise from launching many trajectories that are about as far from "independent" as one can get; they are meant to be related to physical sources. Can one expect that the ensemble averages will at all apply to these systems? The answer is yes, and we could arrive at the same results for the ensemble averages just by looking at calculations of branched flow. The problem is that, since the trajectories are not independent in any one realization of branched flow, we would need to average over many more trajectories than otherwise necessary. We use randomly distributed trajectories because they are more efficient at achieving convergence of the averages.

7.1 Random Potentials Revisited

The random potential described in Chapter 4 for studies of branched quantum flux was modeled very closely after the physical potential in the system. The end result was a random potential that could be characterized by two quantities: a length scale, determined by the auto-correlation function, and a height scale, determined by the distribution of values in the potential. Qualitatively, the donor-atom potential responsible for the branching is made up of gentle hills well below the energy of the scattered electrons.

At this point, as we step away from the quantum mechanics to look at classical mechanics, it is reasonable to re-examine the potential as well. Is the branching phenomenon peculiar to a random potential generated in this fashion (donor density fluctuations), or is it a general phenomenon more broadly dependent on these statistical properties of the potential? The answer, as one might anticipate, is that the branching phenomenon is not dependent on the origin of the potential, and we are free to examine potentials generated in other ways.

We have considered several distinct means of generating such random potentials. The simplified version most closely related to the model that we initially considered is the convolution of a two-dimensional Gaussian bump with a grid of random values. The resulting potential has Gaussian auto-correlation, and with fast Fourier transform methods it is simple to generate. This is the form that we have used most frequently, for reasons that will be further discussed below. Another potential that is simple to generate and that also exhibits branched flow is the superposition of randomly oriented cosine waves. We have generated such a potential both by taking a single wave number and by taking a normal distribution of wave numbers about a mean.[1]

The problem of classical propagation through random media has certainly been

---

[1] Though one might assume that taking a distribution of wave numbers would lead to a more "random" potential, the opposite is true for a fixed number of component waves. With a fixed wave number, we need only achieve a sufficiently dense distribution on a circle in Fourier space; for a range of wave numbers, we need to achieve a sufficiently dense distribution on an annulus.

studied in the past, so it is important that we make clear at the outset what regime we are studying. Two types of random media dominate the literature to date. First are systems of point scatterers separated by free space. In this regime, the scatterers are generally considered to be strong, so that the rare scattering events are large-angle [39, 13]. Second are systems with a "white noise" potential. In such a system, the potential is taken to have a value uncorrelated from point to point and with some (often Gaussian) distribution [11]. The systems that we consider here fall into the no-man's-land between these two extremes.

7.2 Potential Correlations

As discussed in §4.3, our original model potential was characterized by a Gaussian auto-correlation and a Gaussian height distribution. As we look to other potentials that exhibit branched flow, and especially as we seek to make some analytic statements about trajectories in these potentials, we need to look at these properties. Though we see the branching phenomenon in many different kinds of smooth random potentials, all smooth random potentials are not created equally. Here, we look at the auto-correlation functions for our potentials. We look first at a concrete mathematical foundation for characterizing the differences between various random potentials. We then apply this to two kinds of potentials that yield branched flow.

We will arrive at the following conclusion. Though there are many smooth random potentials that exhibit branched flow, we do not want to stray too far from the original system. The Gaussian correlations, which we will not only see in the value of the potential but also in it's derivatives, are invaluable for our analytic treatment.

7.2.1 Auto-Correlation and the Spectral Function

There are two concepts that we use to characterize the spatial dependence of our potentials: auto-correlation and the spectral function. As developed in this section, these two concepts are simply related to one another. Note that, unless otherwise indicated, all integrals in this section are over the range $(-\infty, \infty)$.

Let us begin by defining the auto-correlation function for a function $f(x, y)$ to be

$$c(\alpha, \beta) = \int dx \int dy \; f(x + \alpha, y + \beta) f^*(x, y), \tag{7.1}$$

where we assume an implicit normalization such that $c(0, 0) = 1$. In looking at our potentials, it will also be useful to look at the auto-correlation of derivatives of the potential. We further define

$$c_{m,n}(\alpha, \beta) \;\; = \;\; \int dx \int dy \; f^{(m,n)}(x + \alpha, y + \beta) f^{(m,n)*}(x, y) \tag{7.2}$$

so that our auto-correlation function is $c(\alpha, \beta) = c_{0,0}(\alpha, \beta)$.

Noting that the Fourier transform of a derivative is given by (see [14, Ch. 15])

$$f^n(t) = \frac{1}{2\pi} \int d\omega \; (-i\omega)^n F(\omega) e^{-i\omega t} \tag{7.3}$$

and that one can displace a function by taking

$$f(t + \tau) = \frac{1}{2\pi} \int d\omega \; F(\omega) e^{-i\omega(t+\tau)}, \tag{7.4}$$

we have

$$\begin{aligned} c_{m,n}(\alpha, \beta) &= \int dx \int dy \; f^{(m,n)}(x + \alpha, y + \beta) f^{(m,n)*}(x, y) \qquad (7.5) \\ &= \int dx \int dy \; \Big[\frac{1}{4\pi^2} \int dk \int dl \; (-ik)^m (-il)^n F(k, l) \times \end{aligned}$$

$$
\begin{aligned}
& e^{-ik(x+\alpha)}e^{-il(y+\beta)}\Big] f^{(m,n)*}(x,y) && (7.6)\\
= \;& \frac{1}{4\pi^2}\int dk\int dl\ (-ik)^m(-il)^n F(k,l)e^{-ik\alpha}e^{-il\beta}\ \times \\
& \left[\int dx\int dy\ f^{(m,n)*}(x,y)e^{-ikx}e^{-ily}\right] && (7.7)\\
= \;& \frac{1}{4\pi^2}\int dk\int dl\ (-ik)^m(-il)^n F(k,l)e^{-ik\alpha}e^{-il\beta}(ik)^m(il)^n F^*(k,l) \\
& && (7.8)\\
= \;& \frac{1}{4\pi^2}\int dk\int dl\ k^{2m}l^{2n}F(k,l)F^*(k,l)e^{-ik\alpha}e^{-il\beta} && (7.9)\\
= \;& \frac{1}{4\pi^2}\int dk\int dl\ k^{2m}l^{2n}S(k,l)e^{-ik\alpha}e^{-il\beta} && (7.10)
\end{aligned}
$$

where, in the final step, we have defined a "spectral function" $S(k,l) \equiv F(k,l)F^*(k,l)$.

Though this may not appear to be real progress, it is. Let us return for a moment to $c_{0,0}(\alpha,\beta)$,

$$c_{0,0}(\alpha,\beta) = \int dx\int dy\ f(x+\alpha,y+\beta)f^*(x,y). \tag{7.11}$$

In most situations, one will either know this function or one will know its Fourier transform. Given either, we can find the derivative correlations. Let's assume for the moment that we are given $c_{0,0}$. Then we have

$$\frac{1}{4\pi^2}\int dk\int dl\ S(k,l)e^{-ik\alpha}e^{-il\beta} = c_{0,0}(\alpha,\beta) \tag{7.12}$$

$$S(k,l) = \int d\alpha\int d\beta\ c_{0,0}(\alpha,\beta)e^{ik\alpha}e^{il\beta}. \tag{7.13}$$

Using this expression,

$$c_{m,n}(\alpha,\beta) \;=\; \frac{1}{4\pi^2}\int dk\int dl\ k^{2m}l^{2n}\left\{\int d\alpha'\int d\beta'\ c_{0,0}(\alpha',\beta')e^{ik\alpha'}e^{il\beta'}\right\}e^{-ik\alpha}e^{-il\beta} \tag{7.14}$$

7.2.2 Circular Symmetry

For most of the potentials that we deal with, there is no preferred direction. That is to say, we expect the correlation function to be circularly symmetric. With that assumption, we will also find that the spectral function is circularly symmetric. Let us take

$$\begin{aligned} \alpha &= r\cos\theta && (7.15) \\ \beta &= r\sin\theta && (7.16) \\ k &= \kappa\cos\phi && (7.17) \\ l &= \kappa\sin\phi. && (7.18) \end{aligned}$$

We take our functions to be in terms of these variables also,

$$\begin{aligned} c_{m,n}(\alpha,\beta) &\rightarrow c_{m,n}(r,\theta) && (7.19) \\ S(k,l) &\rightarrow S(\kappa,\phi). && (7.20) \end{aligned}$$

Transforming our equations, working with the spectral function first, we have

$$\begin{aligned} S(k,l) &= \int d\alpha \int d\beta \; c_{0,0}(\alpha,\beta) e^{ik\alpha} e^{il\beta} && (7.21) \\ S(\kappa,\phi) &= \int_0^\infty dr \int_0^{2\pi} d\theta \; r c_{0,0}(r,\theta) e^{i\kappa r\cos(\theta-\phi)}. && (7.22) \end{aligned}$$

By the assumption that we have a circularly symmetric correlation function, we can take $c_{0,0}(r,\theta) = c_{0,0}(r,0)$ and move it out of the θ integral. Since the cosine has period 2π and there are no other θ dependent terms, the integral is unaffected by the presence of the ϕ term and we have

$$\begin{aligned} S(\kappa,\phi) &= \int_0^\infty dr \; r c_{0,0}(r,0) \int_0^{2\pi} d\theta \; e^{i\kappa r\cos(\theta-\phi)} && (7.23) \\ &= 2\pi \int_0^\infty dr \; r c_{0,0}(r,0) J_0(\kappa r) && (7.24) \end{aligned}$$

which, we note, is also circularly symmetric. Using this fact, we can write

$$c_{m,n}(\alpha,\beta) = \frac{1}{4\pi^2}\int dk \int dl\ k^{2m} l^{2n} S(k,l) e^{-ik\alpha} e^{-il\beta} \tag{7.25}$$

$$c_{m,n}(r,\theta) = \frac{1}{4\pi^2}\int_0^\infty d\kappa \int_0^{2\pi} d\phi\ \kappa(\kappa\cos\phi)^{2m}(\kappa\sin\phi)^{2n} S(\kappa,\phi) e^{-i\kappa r\cos(\phi-\theta)} \tag{7.26}$$

$$= \frac{1}{4\pi^2}\int_0^\infty d\kappa\ \kappa^{2m+2n+1} S(\kappa,0)\int_0^{2\pi} d\phi\ \cos(\phi)^{2m}\sin(\phi)^{2n} e^{-i\kappa r\cos(\phi-\theta)}. \tag{7.27}$$

We cannot, in general, say that $c_{m,n}(r,\theta)$ is circularly symmetric, because the derivatives lie along a Cartesian grid. This manifests itself in the additional terms in the angular integral. However, given our assumptions, we can treat the $m = n = 0$ case and find the inverse of our equation for the spectral function.

$$c_{0,0}(r,\theta) = \frac{1}{4\pi^2}\int_0^\infty d\kappa\ \kappa S(\kappa,0)\int_0^{2\pi} d\phi\ e^{-i\kappa r\cos(\phi-\theta)} \tag{7.28}$$

$$= \frac{1}{4\pi^2}\int_0^\infty d\kappa\ \kappa S(\kappa,0)\int_0^{2\pi} d\phi\ e^{-i\kappa r\cos(\phi)} \tag{7.29}$$

$$= \frac{1}{2\pi}\int_0^\infty d\kappa\ \kappa S(\kappa,0) J_0(\kappa r). \tag{7.30}$$

7.2.3 Gaussian Correlation

The model that we have for the potential in a 2DEG, generated by the donor atoms, is characterized by a circularly symmetric, Gaussian auto-correlation function. As a result, this has been one of the defining characteristics of the potentials used to examine branched flow. This is the first potential that we examine.

The known auto-correlation function is

$$c(r,\theta) = e^{-r^2/L^2} \tag{7.31}$$

for some length scale L. Using the circularly symmetric results above, we then have that

$$\begin{aligned} S(\kappa,\phi) &= 2\pi \int_0^\infty dr\; re^{-r^2/L^2} J_0(\kappa r) && (7.32)\\ &= \pi L^2 e^{-\kappa^2 L^2/4}. && (7.33) \end{aligned}$$

Using this result to find the higher correlation functions,

$$c_{m,n}(r,\theta) = \frac{L^2}{4\pi}\int_0^\infty d\kappa\; \kappa^{2m+2n+1} e^{-\kappa^2 L^2/4} \int_0^{2\pi} d\phi\; \cos(\phi)^{2m} \sin(\phi)^{2n} e^{-i\kappa r\cos(\phi-\theta)}. \qquad (7.34)$$

The general equation resulting from this integral is unpleasant. We can, however, limit our consideration to transverse derivatives. Since we have rotational symmetry, we are free to choose our direction of "propagation" as the x axis without loss of generality. As a result, we can limit our consideration to functions $c_{0,n}(r,0)$, giving

$$\begin{aligned} c_{0,n}(r,0) &= \frac{L^2}{4\pi}\int_0^\infty d\kappa\; \kappa^{2n+1} e^{-\kappa^2 L^2/4} \int_0^{2\pi} d\phi\; \sin(\phi)^{2n} e^{-i\kappa r\cos\phi} && (7.35)\\ &= \frac{L^2}{4\pi}\int_0^\infty d\kappa\; \kappa^{2n+1} e^{-\kappa^2 L^2/4} \left\{ 2\pi(\kappa r)^{-n} \frac{(2n-1)!}{(n-1)!2^{n-1}} J_n(\kappa r)\right\} && (7.36)\\ &= \frac{L^2}{2} r^{-n} \frac{(2n-1)!}{(n-1)!2^{n-1}} \int_0^\infty d\kappa\; \kappa^{n+1} e^{-\kappa^2 L^2/4} J_n(\kappa r) && (7.37)\\ &= \frac{L^2}{2} r^{-n} \frac{(2n-1)!}{(n-1)!2^{n-1}} \left\{ 2^{n+1} e^{-r^2/L^2} L^{-2n-2} r^n \right\} && (7.38)\\ &= \frac{(2n-1)!}{(n-1)!} 2L^{-2n} e^{-r^2/L^2} && (7.39)\\ &= \frac{(2n)!}{n!} L^{-2n} e^{-r^2/L^2}. && (7.40) \end{aligned}$$

This result is important. The analytic work to follow, looking at ensemble average properties of flow through these random potentials, depends on the ability to divide the flow into segments that are independent of one another. This depends on the finite range of the auto-correlation function, not only of the potential but of its derivatives as well.

For the long-range behavior, there are theories that depend on the integrals of these auto-correlation functions [40]. Though the time and length scales that we will consider are generally too short for these theories, we will find the necessary integral. We take

$$\begin{aligned}\int_0^\infty dr\ c_{0,n}(r,0) &= \frac{(2n)!}{n!}L^{-2n}\int_0^\infty dr\ e^{-r^2/L^2} && (7.41)\\ &= \frac{(2n)!}{n!}L^{-2n}\left\{\frac{L\pi^{1/2}}{2}\right\} && (7.42)\\ &= \frac{(2n)!}{n!\,2}L^{1-2n}\pi^{1/2}. && (7.43)\end{aligned}$$

7.2.4 Plane Waves

From work in quantum chaos, the properties of overlapping plane waves are fairly well known [41]. Potentials generated by such a sum also result in branched flow, so we have reason to include them in our discussion here. It is also necessary to get the results for the auto-correlation of the derivatives of the potential, as above.

Let us consider a sum of plane waves of fixed wave number k but random orientation. In this case, we start from the spectral function and work towards the auto-correlation function. From the information just given, we can write

$$S(\kappa,\phi) = 2\pi L\ \delta\left(\kappa - \frac{1}{L}\right). \tag{7.44}$$

The parameter L determines the length scale of the potential by this relationship with the wave number. The factor of $2\pi L$ preserves normalization when we make our Fourier transform, as will soon be shown. To get from this expression to our correlation function, we have

$$\begin{aligned} c_{0,0}(r,\theta) &= \frac{1}{4\pi^2}\int_0^\infty d\kappa\;\kappa\int_0^{2\pi} d\phi\; S(\kappa,\phi)e^{-i\kappa r\cos(\phi)} && (7.45)\\ &= \frac{L}{2\pi}\int_0^\infty d\kappa\;\kappa\int_0^{2\pi} d\phi\;\delta\left(\kappa-\frac{1}{L}\right)e^{-i\kappa r\cos(\phi)} && (7.46)\\ &= \frac{1}{2\pi}\int_0^{2\pi} d\phi\; e^{-ir\cos(\phi)/L} && (7.47)\\ &= J_0(r/L), && (7.48)\end{aligned}$$

as expected. The reverse process, going from Bessel function auto-correlation to the spectral information, is similarly simple and makes use of Bessel function orthogonality.

Looking for the general functions, we have

$$c_{m,n}(r,\theta) = \frac{1}{4\pi^2}\int_0^\infty d\kappa\;\kappa^{2m+2n+1}\int_0^{2\pi} d\phi\;\cos(\phi)^{2m}\sin(\phi)^{2n}S(\kappa,\phi)e^{-i\kappa r\cos(\phi-\theta)} \tag{7.49}$$

$$= \frac{L}{2\pi}\int_0^\infty d\kappa\;\kappa^{2m+2n+1}\int_0^{2\pi} d\phi\;\cos(\phi)^{2m}\sin(\phi)^{2n}\delta\left(\kappa-\frac{1}{L}\right)e^{-i\kappa r\cos(\phi-\theta)} \tag{7.50}$$

$$= \frac{1}{2\pi}L^{-2m-2n}\int_0^{2\pi} d\phi\;\cos(\phi)^{2m}\sin(\phi)^{2n}e^{-ir\cos(\phi-\theta)/L}. \tag{7.51}$$

In general, this is not tractable. However, let us confine our attention as before to taking transverse derivatives. That is, we look at $c_{0,n}(r,0)$, taking y derivatives and looking at correlations in the $\hat{x}$ direction.

$$\begin{aligned} c_{0,n}(r,0) &= \frac{1}{2\pi}L^{-2m-2n}\int_0^{2\pi} d\phi\;\sin(\phi)^{2n}e^{-ir\cos(\phi)/L} && (7.52)\\ &= \frac{(2n)!}{n!\,2^n}(rL)^{-n}J_n(r/L). && (7.53)\end{aligned}$$

Hence we see that the potential generated in this way, as well as all derivatives, has much longer range correlations than the Gaussian case of the previous section.

In fact, we have reason to suspect that the analytic development to follow may break down for such potentials as a result of the long-range correlations that exist.

We take a quick look, once again, at the integral of the auto-correlation functions. Because of the properties of Bessel functions, namely

$$\int_0^\infty dr\ r^{-n} J_n(r) = \frac{n!\,2^n}{(2n)!}, \tag{7.54}$$

we have that

$$\int_0^\infty dr\ c_{0,n}(r,0) = L^{1-2n}. \tag{7.55}$$

7.3 Momentum Relaxation Times

As mentioned earlier in this thesis, one of the ways to characterize motion through a random potential is the momentum relaxation time. Another name for this could be the momentum-momentum correlation time. Essentially, this is a measure of how long it takes a particle to "forget" the direction that it was initially traveling. We make the definition more mathematical by defining the function

$$c(t) = \hat{p}(t) \cdot \hat{p}(0). \tag{7.56}$$

The average of this function over many independent trajectories, $\langle c(t) \rangle$, is the momentum correlation function for the system. We further define the momentum relaxation time τ_r by

$$\langle c(\tau_r) \rangle = e^{-1} \tag{7.57}$$

The usual picture for this momentum relaxation is in a potential with occasional large-angle scattering events, rather than the continuous small-angle scattering that we are seeing. In that picture, the time that comes out is a measure of how long we can expect a particle to propagate before suffering a collision. At each collision, it is assumed that the direction of propagation is randomized. Taking some fixed probability of a collision per unit time, an exponential decay of $\langle c(t) \rangle$ falls out immediately.

In the systems for branched flow, we instead have a gradual loss of the memory of the initial propagation direction. Here, however, we can apply that ubiquitous concept in random processes, the random walk. We will consider the change of the momentum of the particle as a random walk in momentum space. Since energy is conserved and the average value of the potential is zero, we will take our random walk to be in one dimension: the angle of the momentum vector.

For simplicity, we take the initial angle to be $\theta(0) = 0$. We then have

$$\begin{aligned} c(t) &= \hat{p}(t) \cdot \hat{p}(0) && (7.58) \\ &= \cos[\theta(t) - \theta(0)] && (7.59) \\ &= \cos[\theta(t)]. && (7.60) \end{aligned}$$

To apply our random walk, we need to discretize time in some way. It matters very little how we do this; let us assume that we take some time step τ long enough that we can consider $\Delta\theta_i$ and $\Delta\theta_{i+1}$ to be independent random variables. They will come from the same distribution; it doesn't matter what the details of this distribution are, as long as it has zero mean. This assumption is reasonable, given the lack of any preferred direction in our random potentials. Let the standard deviation of the distribution of $\Delta\theta$ be σ_1.

A well-known result from probability theory is the central limit theorem. It states that, for sufficiently large n, the sum of n random numbers from the same distribution will go to a Gaussian. The mean of this Gaussian will be n times the mean of the original distribution, and the standard deviation will be

$$\sigma_n = \sigma_1 n^{1/2}. \tag{7.61}$$

Since we take our distribution of $\Delta\theta$ to have zero mean, the sum of changes will also have a zero mean.

We then ask questions about the distribution of $c_n = c(n\tau)$. We have that

$$\begin{aligned} \langle c_n \rangle &= \int d\theta \, \cos(\theta) \frac{1}{(2\pi)^{1/2}\sigma_n} e^{-\theta^2/2\sigma_n^2} && (7.62) \\ &= e^{-\sigma_n^2/2} && (7.63) \\ &= e^{-\sigma_1^2 n/2}. && (7.64) \end{aligned}$$

From this result, we return to the continuous time case by taking $n = t/\tau$,

$$\langle c(t) \rangle = e^{-(\sigma_1^2/2\tau)t}. \tag{7.65}$$

Just as in the case with discrete large-angle scattering events, we expect an exponential decay of the momentum-momentum correlation function. This is shown in Figure 7.1.

We define the momentum relaxation time τ_r by

$$\langle c(\tau_r) \rangle = e^{-1} \tag{7.66}$$

$$\Rightarrow \tau_r = 2\tau\sigma_1^{-2}. \tag{7.67}$$

We would expect, and indeed observe, that τ_r has a dependence on the strength of the potential. To relate this to the formalism developed above, we would expect a dependence of σ_1 on the strength of the potential. The distribution of $\Delta\theta$ is most

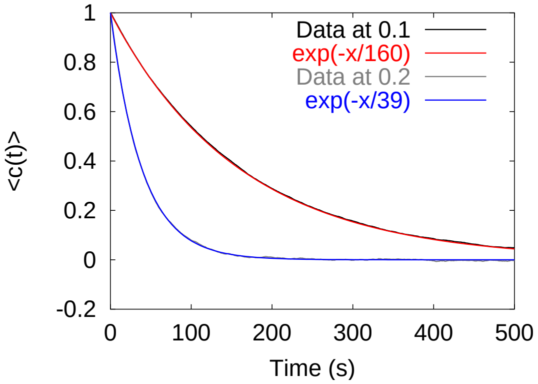


Figure 7.1: Here we see the decay of the momentum correlation function in our random potentials. The potentials used have Gaussian auto-correlation with a correlation length of 1. These data were generated using ten thousand trajectories in each of five realizations of the potential (this number of trajectories was only necessary to get convergence for the stronger potentials tested). We have shown the computational results, along with exponential fits, for two different potential strengths. We see excellent agreement between the calculated momentum correlation function and the fitted exponentials.

directly related to the distribution of gradients in the potential. We do not have direct access to this distribution function, but we do know that, since differentiation is a linear function, the standard deviation of the gradient distribution will be linear in the strength of the potential. We would then expect σ_1 itself to be linear in the strength of the potential, and τ_r to have an inverse square dependence on the strength of the potential. This also is as observed, shown in Figure 7.2. Davies arrives at essentially the same dependence by working explicitly in Fourier space [21].

In a potential with Bessel function auto-correlation (i.e., our sum of plane waves), we still see a nearly exponential decay of the momentum correlation function for any given potential strength. However, the dependence on the potential height is not as nicely inverse-square as it is when we look at potentials with Gaussian auto-correlations. In Figure 7.3, we show the result in this type of potential.

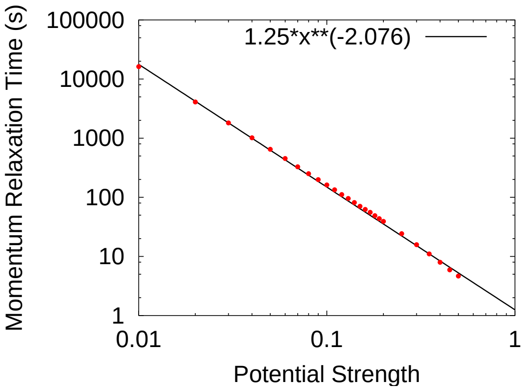
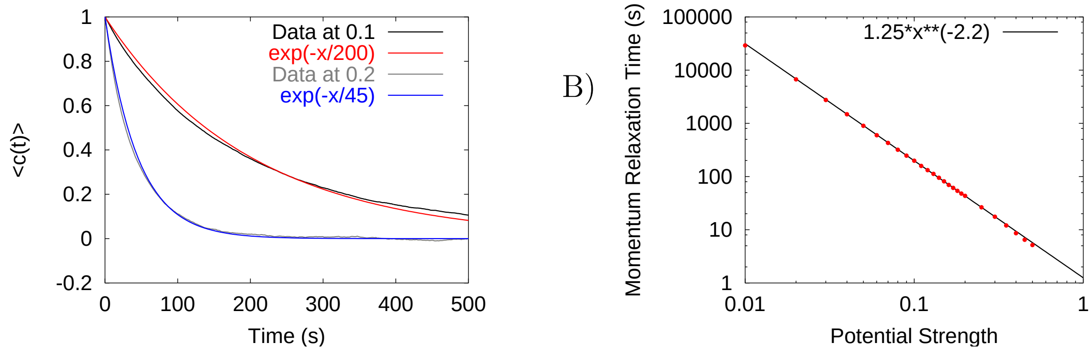


Figure 7.2: Here we see the inverse square dependence of the momentum relaxation time on the strength of the potential. The data points show the e^{-1} times for the momentum correlation functions at various potential strengths, always using a potential with Gaussian auto-correlation. The line is a fitted function to these points, where both the power and the coefficient were allowed to vary. We see that the best fit to the points is close to the predicted inverse square dependence. The data is shown on a log-log scale to make it easier to see the behavior at all scales.

A)



B)

Figure 7.3: Here we see the momentum relaxation in a potential generated as the sum of plane waves. In (A), we see that, at two different values of the potential strength, we have nearly exponential decay of the momentum correlation function. In (B), however, we see that the potential strength dependence of the momentum relaxation time is best fit by $t^{-2.2}$, noticeably different from the inverse square expected. The failure of the momentum correlation function to be exponential in time, and the more rapid falloff of the relaxation time, are presumably the result of longer-range correlations in the potential.

7.4 The Growth of the Exponents

One traditional way to characterize ensemble averages of trajectories is by looking at the ensemble average of Lyapunov exponent. The usual prediction is that, at long times, this average will increase linearly with time (or range). In our systems, however, there are two problems. First, the branched flow structures that we have looked at are not the long-range behavior. Second, we are more interested in the rarefaction exponent than the Lyapunov exponent.

In this section, we get predictions for the growth of the ensemble average rarefaction and Lyapunov exponents at all time scales. As has been argued before, the propagation of the stability matrix can be looked at as the product of random matrices. We will exploit that fact in this section, getting predictions based on the behavior of such matrices and then demonstrating that the full system exhibits the predicted behavior.

The matrices that we work with are not, of course, completely random. Going back to the definition of the stability matrix, we we mimic the dynamics by taking

$$\mathbf{M}_i = \prod_{j=1}^{i} \mathbf{DM}_i \tag{7.68}$$

where

$$\mathbf{DM}_i = \begin{bmatrix} 1 - \tau^2 V_i'' & \tau \\ -\tau V_i'' & 1 \end{bmatrix}. \tag{7.69}$$

This is not the form that we used in Chapter 6 when deriving the differential equation for $\mathbf{M}$. Whereas that definition only preserved the unit determinant of $\mathbf{M}$ in the limit $\tau \to 0$, this form of $\mathbf{DM}$ ensures that we will maintain the unit determinant of $\mathbf{M}$

for any τ. We will take random values of V_i'' in a uniform distribution of zero mean. Without loss of generality, we have taken the mass variable to be 1.

7.4.1 Rarefaction

For part of this section, we will find it easier to speak of a "rarefaction coefficient" r^c rather than a rarefaction exponent r. The relationship between the two is simply $r = \ln r^c$.

Let's look at $\mathbf{M}_2$:

$$\begin{aligned}
\mathbf{M}_2 &= \mathbf{DM}_2\mathbf{DM}_1 && (7.70)\\
&= \begin{bmatrix} 1-\tau^2 V_1'' & \tau \\ -\tau V_1'' & 1 \end{bmatrix}\begin{bmatrix} 1-\tau^2 V_2'' & \tau \\ -\tau V_2'' & 1 \end{bmatrix} && (7.71)\\
&= \begin{bmatrix} -\tau^2 V_1'' + (1-\tau^2 V_1'')(1-\tau^2 V_2'') & \tau + \tau(1-\tau^2 V_2'') \\ -\tau V_1'' - \tau V_2''(1-\tau^2 V_1'') & 1-\tau^2 V_2'' \end{bmatrix}. && (7.72)
\end{aligned}$$

Now, let's look only to first order in τ. The remaining terms are

$$\mathbf{M}_2 = \begin{bmatrix} 1 & 2\tau \\ -\tau(V_1''+V_2'') & 1 \end{bmatrix}. \tag{7.73}$$

We can correctly extrapolate that, to first order in τ,

$$\mathbf{M}_i = \begin{bmatrix} 1 & i\tau \\ -\tau\sum_{j=1}^{i} V_j'' & 1 \end{bmatrix}. \tag{7.74}$$

From this form of $\mathbf{M}_i$, we have a rarefaction coefficient of

$$r_i^c = \left[1 + (i\tau)^2\right]^{1/2}. \tag{7.75}$$

For $i\tau \ll 1$, this becomes $r_i^c \approx 1 + (i\tau)^2/2$, so we have quadratic growth with i. For $i\tau \gg 1$, $r_i^c \approx i\tau$, so for intermediate values we expect a transition from quadratic to linear growth in i.

For long times, it is not appropriate to use our first-order approximation for $\mathbf{M}_i$. Let us assume that the elements of $\mathbf{M}_i$ are eventually randomized with some average magnitude α_i. How does α_i scale with i? Well,

$$\begin{bmatrix} 1-\tau^2 V_i'' & \tau \\ -\tau V_i'' & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a(1-\tau^2 V_i'') + c\tau & b(1-\tau^2 V_i'') + d\tau \\ c - a\tau V_i'' & d - b\tau V_i'' \end{bmatrix}. \tag{7.76}$$

The elements of the new matrix are all, to first order in τ, of the form $a + c\tau$. With $\langle a \rangle = \langle c \rangle = 0$ and $\langle a^2 \rangle = \langle c^2 \rangle = \alpha_i^2$, we have

$$\begin{aligned} \alpha_{i+1} &= \left\langle (a+c\tau)^2 \right\rangle^{1/2} && (7.77) \\ &= \left\langle a^2 + c^2\tau^2 + 2ac\tau \right\rangle^{1/2} && (7.78) \\ &= \left(\left\langle a^2 \right\rangle + \tau^2 \left\langle c^2 \right\rangle + 2\tau \left\langle ac \right\rangle \right)^{1/2} && (7.79) \\ &= \left(\alpha_i^2 + \tau^2 \alpha_i^2 \right)^{1/2} && (7.80) \\ &= \alpha_i \left(1 + \tau^2 \right)^{1/2}. && (7.81) \end{aligned}$$

So we have exponential growth of the elements of the matrix. The expectation for r_i^c will itself be linear in α_i, so we also see exponential growth in r_i^c, or linear growth in r_i. These three regimes of rarefaction exponent growth can be seen in Figure 7.4. For comparison, in Figure 7.5 we show the results for the random matrix multiplication used in this development.

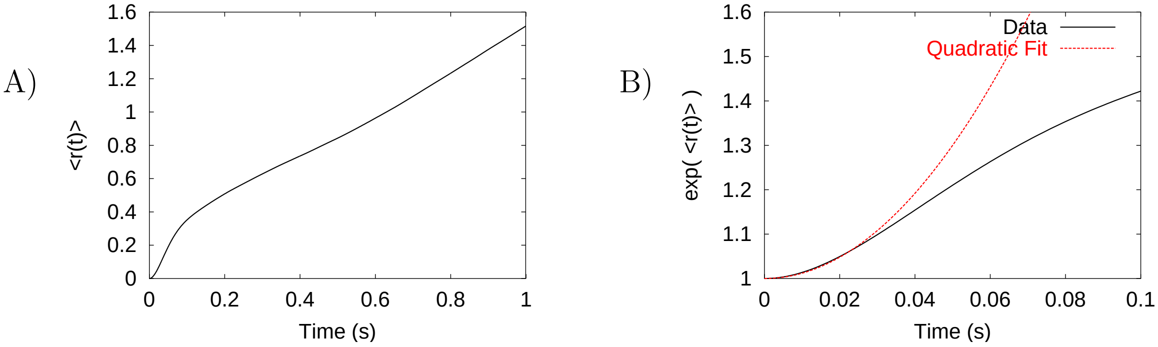


Figure 7.4: Here we see the growth of the rarefaction exponent as a function of time in the real system. Though our predictions for how this should behave were made using random matrix multiplication, those predictions carry through to the full system. In (A), we see the ensemble averaged rarefaction exponent; it is easy here to see that, fairly quickly, we get into the regime of linear growth with time. In (B), we show the rarefaction coefficient (rather than exponent), making it easier to see that quantity's initial quadratic growth and mid-range linear growth.

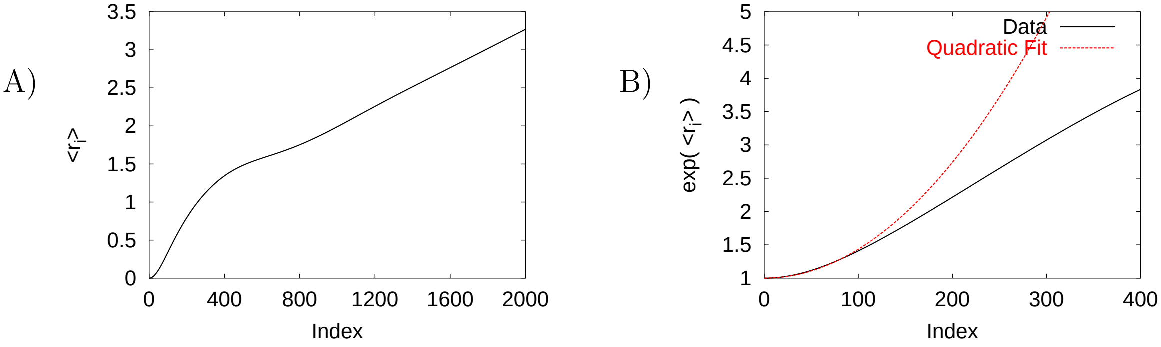


Figure 7.5: These show the growth of the average rarefaction exponent for random matrix multiplication. Here, the random matrices were as described in the text, using $\tau = .01$ and potential second derivatives distributed uniformly in the range $(-5, 5)$. We see the same general shape as that for the full system. In (A), we show the data far enough to see the linear growth with iteration index, and in (B) we show the rarefaction coefficient for the early data. The crossover from the regime shown in (B) to the long-time behavior is dependent on the choice of input parameters.

7.4.2 Lyapunov Exponents

Though we are not primarily interested in the Lyapunov exponent growth, it is interesting to look at because it is another test of the model used above. The general argument for the long-term linear growth of the rarefaction exponents still holds, and we do indeed observe this dependence for Lyapunov exponents. The short-term growth fits our model well, but shows small deviations from that predicted (and observed) for the random matrices.

We begin in a way similar to the previous section, with our expression for $\mathbf{M}_i$ to order τ,

$$\mathbf{M}_i = \begin{bmatrix} 1 & i\tau \\ -\tau\sum_{j=1}^{i} V_j'' & 1 \end{bmatrix}. \tag{7.82}$$

From this expression, we have eigenvalues $\lambda_\pm$ satisfying the equation

$$(\lambda_\pm - 1)^2 + i\tau^2 \sum_{j=1}^{i} V_j'' = 0. \tag{7.83}$$

To simplify the notation, let $S_i = \sum_{j=1}^{i} V_j''$. Solving this equation for the eigenvalues $\lambda_\pm$, we have

$$\lambda_\pm = 1 \pm i^{1/2}\tau(-S_i)^{1/2}. \tag{7.84}$$

We take the potential terms to come from a uniform distribution of zero mean. By the central limit theorem, we know that S_i will go to a Gaussian distribution with zero mean and standard deviation proportional to $i^{1/2}$. Now, what happens if $S_i > 0$? Then we have a complex $\lambda_\pm$, which means a Lyapunov exponent of 0. Hence, for any value with $S_i > 0$, we might as well let $\lambda_\pm = 1$. Half of the time, we have $S_i < 0$, which are the only remaining interesting cases. We note that

$$\int_{-\infty}^{0} dS\ (-S)^{1/2} \frac{1}{(2\pi)^{1/2}\sigma} e^{-S^2/2\sigma^2} \approx 0.411\,\sigma^{1/2} \propto i^{1/4}. \qquad (7.85)$$

Putting these into our expression for λ_+, we expect

$$\lambda_+ = 1 + \tau O(i^{3/4}). \qquad (7.86)$$

Finally, for the Lyapunov exponent, we note that

$$\begin{aligned} h_+ &= \ln \lambda_+ && (7.87) \\ &= \ln[1 + \tau O(i^{3/4})] && (7.88) \\ &\approx \tau O(i^{3/4}). && (7.89) \end{aligned}$$

So, at short times, we expect the Lyapunov exponent to grow as $i^{3/4}$.

As shown in Figure 7.6, the full simulations deviate somewhat from this prediction for the earliest times. The 3/4 power law is still, however, a decent first-order description of the short-time behavior. The results for random matrix multiplication, shown in Figure 7.7, follow this power law somewhat more closely. In both systems, we see the linear growth of Lyapunov exponent at long times.

7.5 Long Range Distributions

The analysis of stability matrix evolution gave us access to the functional form of the growth of the means. It can also tell us something about the long-time distributions of variables.

Multiplication of random matrices, such as those used above, is amenable to an application of the same central-limit theorem already employed for momentum

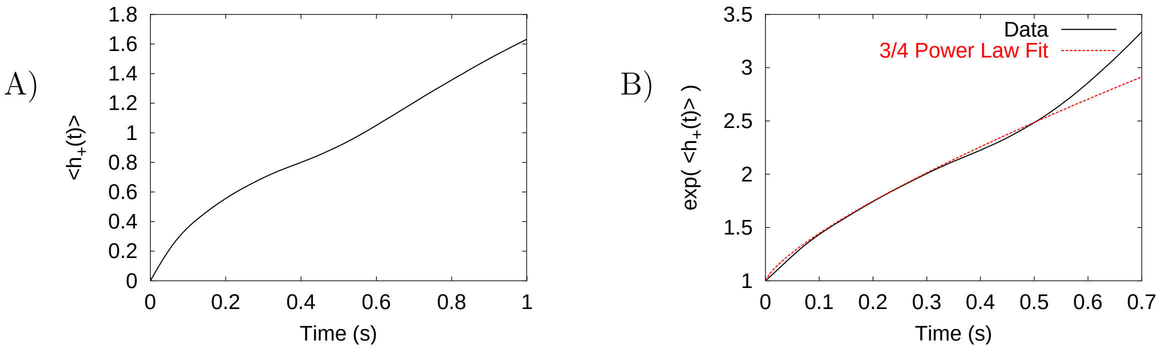


Figure 7.6: Here we see the growth of the Lyapunov exponent as a function of time in the real system. In (A), we see the ensemble averaged Lyapunov exponent, which goes to linear growth with time even sooner than the rarefaction exponent did. In (B), we show the corresponding eigenvalue (i.e., e^{h_+}) making it easier to see the initial behavior. Included is a fit curve showing $t^{3/4}$ growth. The agreement is fairly good, though the very beginning of the data seems to have closer to linear dependence.

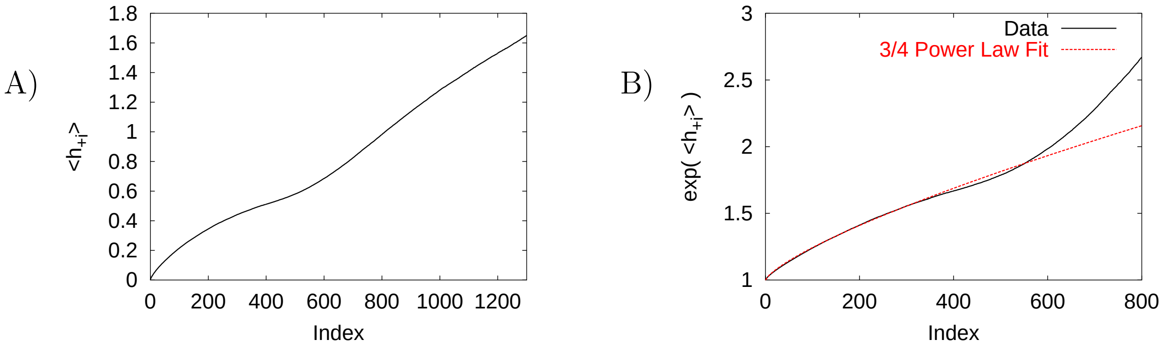


Figure 7.7: These plots show the growth of the average Lyapunov exponent for random matrix multiplication, as described in the text. The features are clearly the same as those for the full system. In (A), we show iterations far enough to establish the long-time linear growth. In (B), we focus on the initial behavior of e^{h_+}. Though the agreement with the 3/4 power law predicted isn't exact, the very early agreement seems better than that for the full system.

relaxation. Theory predicts that the elements of the stability matrices should have a log-normal distribution for sufficiently long times; that is, the logs of the elements should have normal distributions.

As the elements of the stability matrix go to a log-normal distribution, so will simple functions of those elements: notably, the trace and the rarefaction coefficient, in which we are interested. When the mean of the distribution for the trace has grown sufficiently, we can reasonably make the approximation that the Lyapunov exponent is just the log of the trace, so that quantity would have a normal distribution. Let us make the argument for a log-normal distribution in terms of the absolute value of the trace of the stability matrix, and keep in mind that what we're saying could also be applied to the rarefaction coefficient or the elements of the stability matrix themselves.

Why should the elements of the stability matrix, and its trace, have log-normal distributions? Consider again the history of the stability matrix for a trajectory. We again divide this history into independent pieces, so that we can consider the stability matrix to be the product of uncorrelated random matrices. The elements of the matrix resulting from such a product are themselves products of random numbers taken from some distribution. The trace, a simple function of these elements, inherits the property as well. The logarithm of the trace is then a sum of the logarithms of random numbers. The logarithm of a random number is itself a random number; the central limit theorem doesn't care what the actual distribution for the random number is. Thus, the central limit theorem predicts that, as the number of terms in the series increases, the distribution of $\ln|\mathrm{Tr}(\mathbf{M})|$ will tend to a normal distribution;

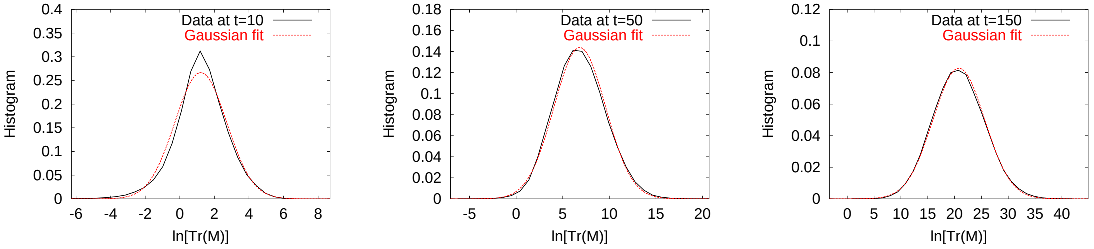


Figure 7.8: Here we see the distribution of $\ln\left[\mathrm{Tr}(\mathbf{M}(t))\right]$ for classical trajectories with $m = 1$ and $E = 1$ in a potential with correlation length 1 and feature height .1. We take the data at three times, $t = 10$, 50, and 150 from left to right. We see that, even at short times, the distribution is approximately log-normal, and that the agreement with this form improves as time progresses.

$|\mathrm{Tr(M)}|$ thus has a log-normal distribution by definition.

A study of the distributions for various quantities for our classical trajectories bear out this prediction of log-normal behavior. Though it is explicitly a long-time prediction, we can see the log-normal distributions forming at times short enough to be relevant to our branching phenomenon. A few examples are shown in Figure 7.8.

Though our disordered potential is stronger than that generally considered by theory, some predictions from the weak-disorder cases carry through to our parameter regime. For our potential with Gaussian correlations, we find the prediction [40, 42]

$$\nu_0 = \left\langle \frac{\ln|\mathrm{Tr}(\mathbf{M})|}{r} \right\rangle \tag{7.90}$$

$$\bar{\nu} = \frac{\ln\left(\left\langle |\mathrm{Tr}(\mathbf{M})|^2 \right\rangle\right)}{2r} \tag{7.91}$$

$$\sigma_\nu^2 = \frac{\bar{\nu} - \nu_0}{r}. \tag{7.92}$$

Here, r is the range coordinate and we can take it as being approximately proportional to time by $r \approx (2E/m)^{1/2}t$. We find ν_L and ν' from our data, and compare the predicted variance with that observed. For ranges of up to approximately 150 correlation lengths with potential features at .1 E_F, we find agreement with this theory to within $\sim 10\%$. Though the distribution of $\mathrm{Tr}(\mathbf{M})$ becomes increasingly log-normal for longer ranges, the agreement with this theory actually worsens; this is presumably the result of the potential features being so strong.

For weak disorder, we also find in [40] the prediction

$$\nu' \approx \left(\frac{1}{2} \int_0^\infty d\xi \left\langle \left. \frac{\partial^2 U(z; r-\xi)}{\partial z^2} \right|_{z=z_0,\, p=p_0} \left. \frac{\partial^2 U(z; r)}{\partial z^2} \right|_{z=z_0,\, p=p_0} \right\rangle \right)^{1/3} \tag{7.93}$$

where ν' provides an upper bound for $\lim_{r\to\infty} \nu_0$. In our terms, this translates to

$$\nu' \approx \left(\frac{1}{2} \sigma^2 \int_0^\infty dr \; c_{0,2}(r, 0) \right)^{1/3}. \tag{7.94}$$

For a potential with Gaussian correlations,

$$\nu' \approx (3\sigma^2 \pi^{1/2})^{1/3} L^{-1}. \tag{7.95}$$

Out to our ranges of about 150 correlation lengths, this ν' upper bound is about a factor of two larger than the numerical ν_0. This factor is in accordance with results quoted in [40].

Chapter 8

Classical Branched Flow

We have already introduced the branched nature of electron flow seen experimentally in 2DEG devices, and have shown that the full quantum-mechanical results exhibit the same phenomenon. There are so many places in physics where we find a strong correlation between classical and quantum behaviors that it should not come as a surprise that the branching appears in classical calculations as well. In Figure 8.1, we see an example of quantum and classical flux through the same potential.

As set out in Chapters 6 and 7, there are tools of analysis available in classical mechanics that are not available in quantum mechanics. We have already explored the expected ensemble-averaged behavior; at this point, we turn to the more interesting behavior in individual realizations of the random potentials.

What will we gain in this discussion? First of all, it opens up for us the language of individual trajectories in describing the behavior of the flow. Needless to say, particle trajectories are closer to our daily intuition than is wave mechanics. Second, we can use the stability matrix analysis already introduced to study the formation and

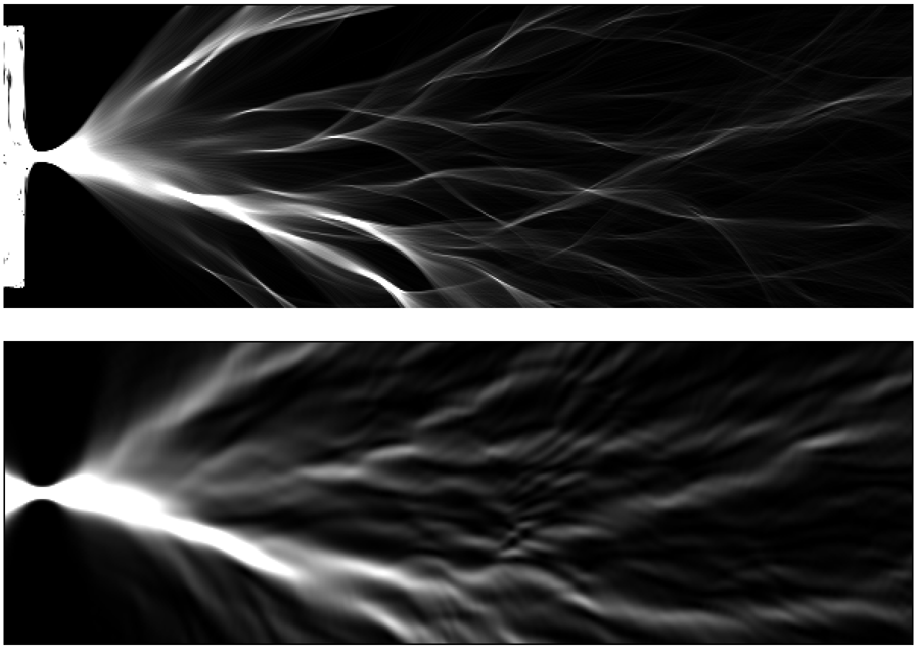


Figure 8.1: Here we see an example of quantum- and classical-mechanical flux calculations, carried out in the same potential. In both cases, a QPC determines the shape of the manifold sent into the random background potential. The agreement between the two calculations is not perfect, but it is sufficiently good to conclude that branching is largely a classical phenomenon. Though this is our motivation, we will generally not make direct comparisons between classical and quantum solutions in the same potentials.

persistence of branches. Third, it is simple to study the branching as a phenomenon in four-dimensional phase space rather than merely two-dimensional coordinate space. The structures seen in phase space further illuminate the underlying behavior.

8.1 Physical Systems

The characteristics that we have described for our random potentials are found in nature, and not only in the 2DEG electron propagation problem. Recent work on the propagation of sound waves through the ocean has used a similar potential, and the beginnings of structures very like our branches have been reported [40]. Similarly, work on gravitational micro-lensing has seen evidence of such structure on a significantly larger length scale [43].

We believe that one reason that these structures haven't been studied in the past is that it requires one to look beyond the statistical properties of trajectories in the potential. Those statistical properties are as expected, as we saw in Chapter 7. The branching phenomenon is seen as a result of choosing not only a single potential from the ensemble, but also choosing a special initial manifold (i.e., one-dimensional) of trajectories. One then must look at the full coordinate space history of the propagated ensemble.

8.2 Manifolds

For the quantum-mechanical flux calculations, we studied the propagation of electrons through the 2DEG potential after they had passed through a point contact. The effect of this point contact was to turn the incident plane wave into something more like a p-wave: the electrons flowing out into the 2DEG originated from approximately the same place, but with a distribution of angles and a node along a wall.

In the classical system, we have more freedom in determining how we send electrons through the system. Whatever we do, we need to choose some initial "manifold," or collection, of trajectories to launch at the system. We should not expect that our results will necessarily be independent of this choice. To more closely simulate the quantum-mechanical experiments, we could launch trajectories that initially fill phase space (all positions, all angles) and let them pass through a point contact (as we have done in Figure 8.1). The QPC would have the effect of forming a new manifold, with trajectories close to one another in space but still with a distribution of angles. Alternatively, we could dispense with the QPC and form this vertical manifold in phase space ourselves.

The latter approach is the one that we will generally take, though whenever direct comparisons are made between classical and quantum flow through the same potential we would use the former. The theory used to deal with the classical manifold is clearer if it is truly one-dimensional in phase space, which we would never achieve using the QPC method.

Note that the classical mechanics opens up other possibilities as well. We can orient our one-dimensional manifold any way that we wish, relative to the phase-space axes. If we were to consider higher-dimensional systems, three spatial dimensions and up, we would have additional, more interesting choices to make. With two spatial dimensions, our only option, if we wish to see interesting evolution of the manifold, is to use a one-dimensional manifold. If we have three spatial dimensions, we can consider not only lines of trajectories, but variously oriented sheets as well. The behaviors that might be seen in such systems have not yet been explored.

8.3 Dimensionless Parameters

Two quantities in our system highlight the fundamentally independent parameters that we can vary for our ensemble of potentials, and are important as we consider the branched flow. First, we normalize all energies by the energy of the particles that we are sending though the system. Hence the value $V(x, y)$ of the potential at a point in space gives us the height relative to the energy of our particle. Only this ratio, and not any independent energy scale, is important for the nature of our branches.

Second, we can measure the length scale of our branches in units of the correlation length of the potential. It is a simple but not well-known result that the spatial re-scaling of the potential, while keeping the height of the potential the same and the energy of the particles the same, results in nothing more than a re-scaling of the classical trajectories through the system. It is worth going through this result below; we will, at the same time, show the energy-scaling result.

Let our potential be given by $V(x, y)$, and take initial conditions $x(0)$, $y(0)$, $p_x(0)$, and $p_y(0)$. Suppose that the solution is the trajectory given by $x(t)$, $y(t)$, $p_x(t)$, and $p_y(t)$. Now consider the same potential stretched spatially by the factor α, and scaled in energy by a factor μ^2. That is, we have $V'(x, y) = \mu^2 V(x/\alpha, y/\alpha)$. We now consider initial conditions scaled by the same amounts: $x'(0) = \alpha x(0)$, $p'_x(0) = \mu p_x(0)$, and similarly for the y components.

The solution to the equations of motion in this scaled potential with these initial conditions is $x'(\tau) = \alpha x(t)$ and $p'_x(\tau) = \mu p_x(t)$ with $\tau \equiv \alpha t/\mu$, and again similarly for the y components. We demonstrate this by beginning with the equations of motion for the scaled system and find that

$$\begin{aligned}
\partial_\tau x'(\tau) &= \frac{1}{m} p'_x(\tau) && (8.1)\\
\partial_\tau \left[\alpha x(t)\right] &= \frac{1}{m} \mu p_x(t) && (8.2)\\
\alpha(\partial_\tau t)\partial_t x(t) &= \frac{1}{m} \mu p_x(t) && (8.3)\\
\alpha(\mu/\alpha)\partial_t x(t) &= \frac{1}{m} \mu p_x(t) && (8.4)\\
\partial_t x(t) &= \frac{1}{m} p_x(t) && (8.5)
\end{aligned}$$

which is the equation of motion for the original system and, by our assumptions, we know it to be satisfied. For the momentum, we have

$$\begin{aligned}
\partial_\tau p'_x(\tau) &= -\partial_{x'(\tau)} V'\left(x'(\tau), y'(\tau)\right) && (8.6)\\
\partial_\tau \left[\mu p_x(t)\right] &= -\partial_{\alpha x(t)} V'\left(\alpha x(t), \alpha y(t)\right) && (8.7)\\
\partial_\tau \left[\mu p_x(t)\right] &= -\partial_{\alpha x(t)} \mu^2 V\left(x(t), y(t)\right) && (8.8)\\
\mu(\mu/\alpha)\partial_t p_x(t) &= -\alpha^{-1}\partial_{x(t)} \mu^2 V\left(x(t), y(t)\right) && (8.9)\\
\partial_t p_x(t) &= -\partial_{x(t)} V\left(x(t), y(t)\right) && (8.10)
\end{aligned}$$

which again is known to be true based on our assumptions.

Clearly, we have performed a time scaling as well as the energy and spatial scalings. To study the branched flow, however, we are interested in the time-independent coordinate-space projection of the trajectories. Hence we aren't sensitive to the changes in time scale introduced in the formal manipulations above.

8.4 Branch Characteristics

Observing and describing the branched flow is only the first step toward understanding it. We also seek a way to describe the formation of these branches, and how to distinguish the regions of phase space where we find branches from those where we do not.

8.4.1 Branch Stability Analysis

Going beyond the statistical measures in Chapter 7, we would like to look at what is happening with trajectories in our systems. Can we characterize those that end up in branches, differentiating them from those that do not?

As argued in Chapter 6, the Lyapunov exponent measure of stability is not particularly meaningful in these open systems without periodic trajectories. We can now make a direct comparison between the Lyapunov exponent measure of stability and the more specialized notion of the rarefaction exponent.

Using the standard Lyapunov exponent measure of stability, shown in Figure 8.2, we see that the trajectories in our branches go through periods of stability and instability. Here we understand "stability" to mean that the Lyapunov exponents go to zero. The branches do not, however, remain stable for even a majority of their histories, nor do the branches represent the only regions with such stability. Hence this measure of stability is a poor descriptive tool for the branching phenomenon.

If we use our rarefaction exponent notion of stability, we find a more useful measure. Here we see a strong correspondence between branches and stability, shown in Figure 8.3. The branches represent not only the least unstable regions of the flow, but

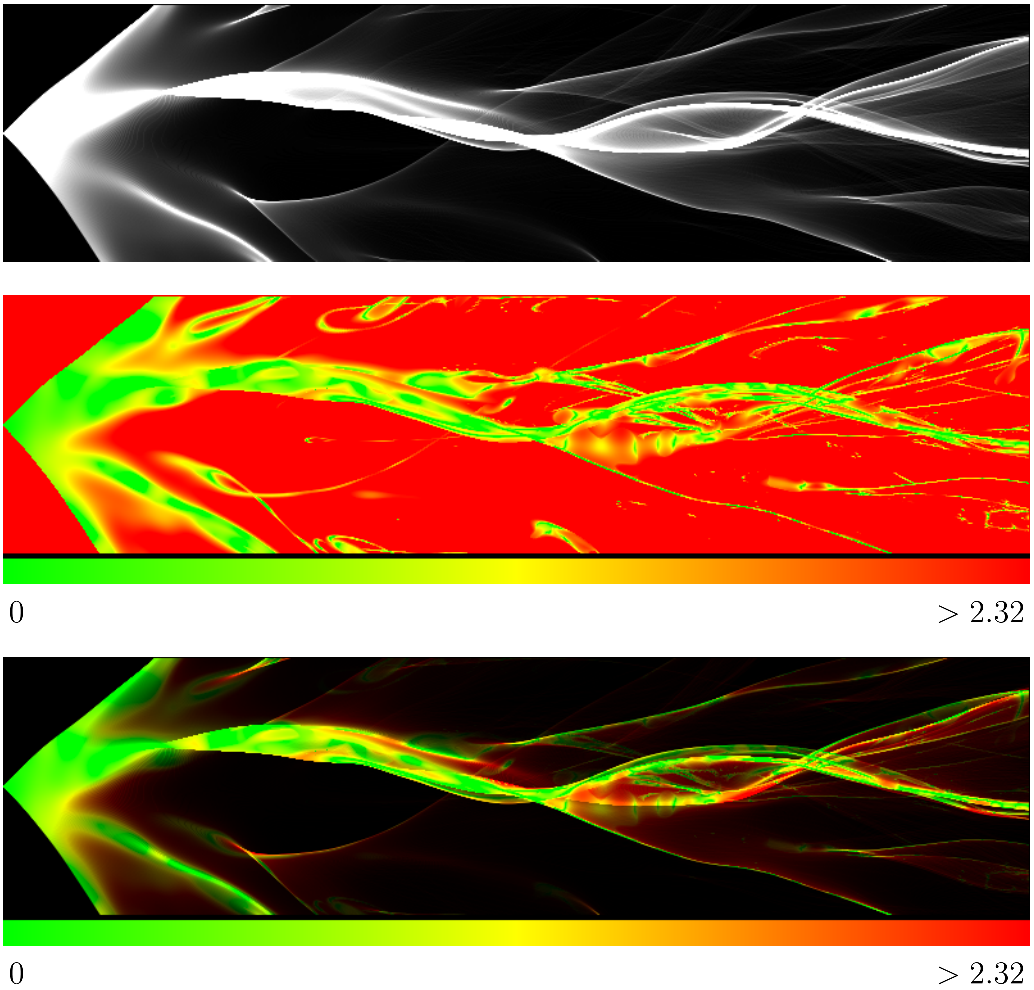


Figure 8.2: Here we see the flux through a potential and the corresponding Lyapunov exponent measure of stability. The top figure shows the flux density in greyscale. We note that the areas of low Lyapunov exponent (green and yellow) have poor correlation to branch locations. We see areas of high flux with high Lyapunov exponent, and areas of low flux with low Lyapunov exponent. At each point of coordinate space, we have shown the stability of the most stable trajectory passing through that point. Higher values in the data have been clipped to show more detail in this range. In the bottom image, we combine both sets of information, with color determined by Lyapunov exponent and intensity determined by flux density. This representation makes it easy to see that the long-range, strong branches pass in and out of areas of low Lyapunov exponent.

in fact the strongest branches remain stable, with negative rarefaction exponents. The weaker branches are not stable in this global sense, but do represent areas of relative local stability. We see no regions of phase space that are local minima of rarefaction exponents that do not form branches.

8.4.2 Intensity Distribution

Though one has a strong qualitative feel for the branched nature of the flux, there is a certain difficulty in characterizing it quantitatively. One way to do so is to look at the distribution of intensities in the flux as a function of the distance from the launching point. We do this by taking a one-dimensional slice of the flux data at a fixed radius and applying the appropriate analysis. Qualitatively, what we see in this slice is that most of space has very little flux, and some regions (where branches pass through our slice) have very high flux. Fixing a radius and aggregating flux data from many different potentials in the ensemble, we see an approximately log-normal distribution of intensities, as shown in Figure 8.4.

This intensity distribution is theoretically predicted, given some reasonable assumptions about the behavior of the system. We return for a moment to our rarefaction measure described above. Let us first take a radial slice and ask about the distribution of rarefaction coefficients (recall that the rarefaction exponent and coefficient are related by $r = \ln r^c$). The arguments laid forth in §7.5 for the log-normal distribution of $|\mathrm{Tr}(\mathbf{M})|$ apply here as well. Our rarefaction coefficient is merely the length of the vector $\mathbf{M} \cdot \hat{d}$, and the log-normal distribution of the elements of $\mathbf{M}$ give us a log-normal distribution of $|\mathbf{M} \cdot \hat{d}|$. This can easily be verified with random

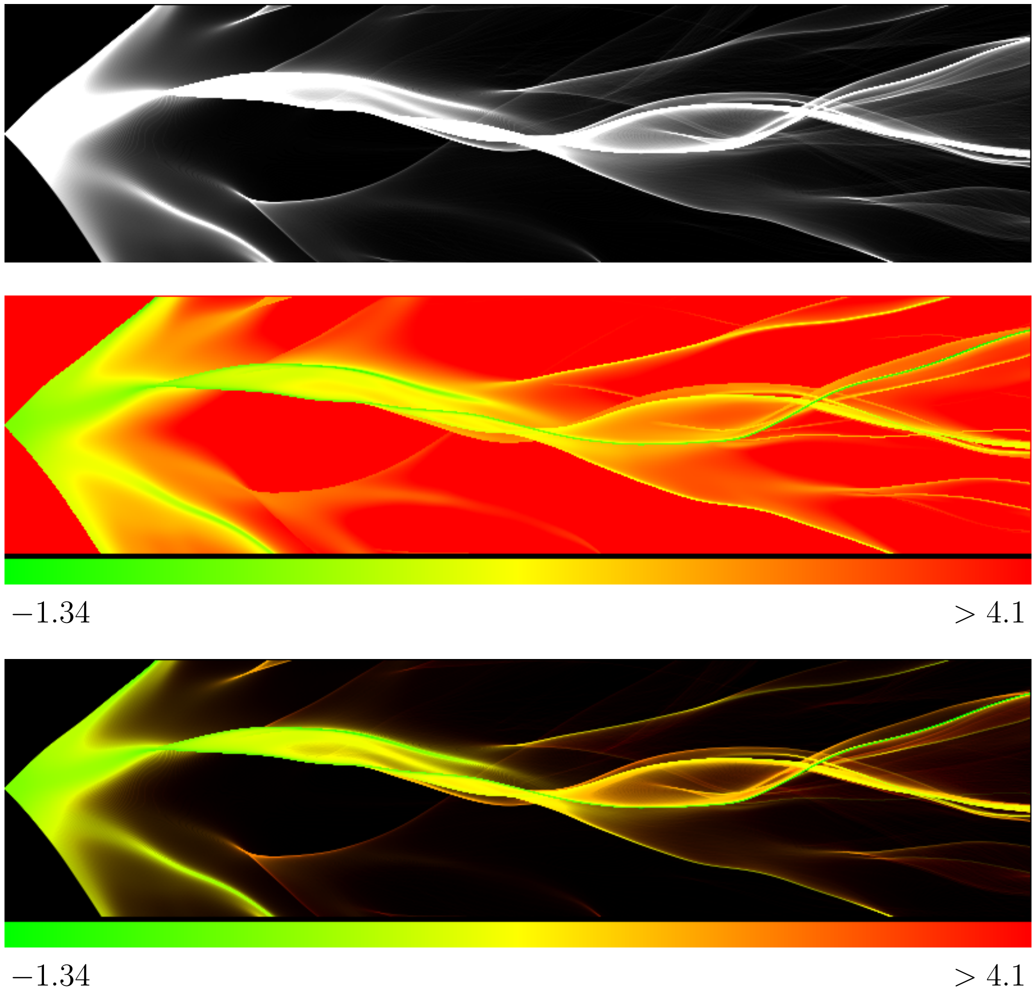


Figure 8.3: Here we see the flux through a potential and the corresponding rarefaction exponent measure of stability for the same system as shown in Figure 8.2. The top figure shows the flux density in greyscale. The middle figure shows the rarefaction exponent for the same flow. We note that the areas of low rarefaction exponent (green and yellow) have a strong correlation to areas of high flux density, as contrasted to the behavior of the Lyapunov exponents. Higher values in the data have been clipped to show more detail in this range. In the bottom image, we combine both sets of information, with color determined by rarefaction and intensity determined by flux density.

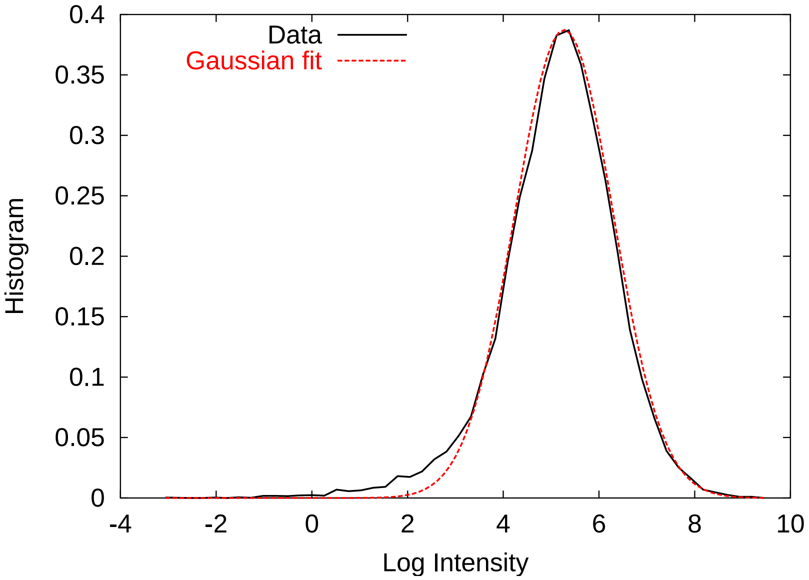


Figure 8.4: This figure shows the distribution of intensities in our classical branched flow. The data are taken from trajectories run in two hundred different realizations of the random potential, looking at the intensities at a fixed radius. The random potential height was $.1E_F$, and the slice was taken at a radius of fifty correlation lengths. What is plotted is the histogram of the logarithms of the intensities, and a Gaussian fit. Were the distribution of intensities truly log-normal, the plotted histogram would itself be normal. Though the bulk of the data clearly fits a Gaussian, we have notable deviation from log-normal starting at about 2σ left of the mean. The Gaussian fit isn't perfect, but at least points in the direction of an emerging log-normal intensity distribution. The quality of this fit improves with radius.

matrices, and is not peculiar to our choice of $\hat{d}$ or the dynamics that generate $\mathbf{M}$.

We now need to connect the distribution of rarefaction coefficients to the observed distribution of intensities. Let r^c be the rarefaction coefficient of a trajectory at our chosen distance, and let $p(r^c)$ be the (log-normal) probability distribution of r^c. If we let the density of trajectories in our initial manifold be unity, then the density in the neighborhood of our reference trajectory is now $(r^c)^{-1}$ by the definition of r^c. We will assume that trajectories sufficiently close to one another will have similar rarefactions, so that whereas $p(r^c)$ gives the probability distribution of r^c in terms of trajectories, $r^c p(r^c)$ gives the distribution in terms of spatial locations (that is, we have divided by the trajectory density). Finally, the intensity i at a location in space is directly proportional to the density of trajectories at that point, and therefore inversely proportional to r^c. Hence the probability distribution of i is given by $i^{-1}p(i^{-1})$. Given the log-normal form of p, this is itself log-normal.[1]

8.4.3 Length Scales

Though a spatial re-scaling of our potential does change the length scale of the features that we see in the flow, we saw that it does so in a trivial fashion. We are left, then, with the task of observing the changes to the flow when we change the ratio of the energy of our particles to the height of the features in the potential.

The qualitative changes are not surprising. The flow is seen to form branches more quickly the stronger we make the potential. Stronger potentials cause larger-angle scattering events, so the spreading trajectories are noticeably perturbed more quickly.

---

[1] A useful, but occasionally annoying, property of the log-normal distribution is it's robustness to taking powers of the random variable.

We have looked for a quantitative measure of this change, and find the most useful to be the distance to a self-intersection, or conjugate, point. This concept comes from semi-classical physics and the tracking of Maslov indices [44].

A self-intersection point is defined as a point where a trajectory and an infinitesimally displaced neighbor will cross one another. We again only concern ourselves with displacements that are along the initial distribution. We find that, with such a definition, the self-intersection points correspond closely to macroscopic foci. We find the self-intersection points by looking for a zero of the quantity

$$(-p_y(t), p_x(t), 0, 0) \cdot \mathbf{M}(t) \cdot \hat{d}. \tag{8.11}$$

The vector on the right is an initial displacement along the distribution, and the vector on the left is one perpendicular to the direction of propagation for the particle at time t. When this quantity goes to zero, a trajectory has intersected a neighboring trajectory infinitesimally displaced along the initial distribution.[2]

Let us immediately specialize to the case where $\hat{d}$ lies entirely in the momentum subspace, corresponding to trajectories starting from the same point but with different momenta. With $\hat{d}$ specified, the shortest distance to a self-intersection point is a property of the random potential rather than of any specific manifold, since all starting positions and directions are equivalent.

Let us take l as the correlation length of our potential, d as the shortest path length to a self-intersection point, and ϵ/E as the feature height in the potential divided by the energy of the particles in our manifold. We are after a relationship between the dimensionless quantities d/l and ϵ/E. We can find one analytically in

---

[2] If you don't immediately see this, it shouldn't be difficult to convince yourself with pen and paper. Consider what it means for this quantity to change sign.

the limit $\epsilon/E \to 0$, though we will see that result still applies in our regime of interest ($\epsilon/E \sim .1$).

We will begin by working in terms of a time τ to a self-intersection, rather than path length d, since time is the fundamentally independent parameter. For weak potentials, however, we can accurately approximate $d \approx \tau(2E/m)^{1/2}$. The theory that we are about to work through is not difficult, but we need to take care to not lose important terms.

Let us take, without loss of generality, y as the direction transverse to the propagation of the trajectory. When the slope dy/dp_y of the manifold is zero, we are at a self-intersection. Considering the inverse, we want dp_y/dy to go to $\pm\infty$. At time $t = 0$, we have $dp_y/dy = \infty$. If we can find the time for dp_y/dy to go from ∞ to zero, it should be roughly half the time scale for it to go from ∞ to $-\infty$. This is the time scale for which we solve.[3]

In the limit $\epsilon/E \to 0$, the dynamics are dominated by shearing. Hence we can take, to first order,

$$\begin{aligned}
\frac{dy}{dt} &= \frac{p_y(t)}{m} &\quad (8.12)\\
&\approx \frac{p_y(0)}{m} &\quad (8.13)\\
y(t) &\approx p_y(t)\frac{t}{m} &\quad (8.14)\\
\frac{dy}{dp_y(0)} &\approx \frac{t}{m} &\quad (8.15)\\
\frac{dp_y(0)}{dy} &\approx \frac{m}{t} &\quad (8.16)
\end{aligned}$$

---

[3] To see that this is the appropriate time scale to find, consider what happens to the manifold to form a caustic.

We then have that

$$\begin{aligned}
\frac{dp_y}{dt} &= -\partial_y V(\vec{x}(t)) && (8.17)\\
p_y(t) &= p_y(0) - \int_0^t dt'\, \partial_y V(\vec{x}(t')) && (8.18)\\
\frac{dp_y(t)}{dy} &= \frac{dp_y(0)}{dy} - \partial_y \left[\int_0^t dt'\, \partial_y V(\vec{x}(t'))\right] && (8.19)\\
\frac{dp_y(t)}{dy} &= \frac{dp_y(0)}{dy} - \int_0^t dt'\, \partial_{yy} V(\vec{x}(t')) && (8.20)\\
\frac{dp_y(t)}{dy} &\approx \frac{m}{t} - \int_0^t dt'\, \partial_{yy} V(\vec{x}(t')) && (8.21)
\end{aligned}$$

Let us consider the integral. In our limit, the trajectory will propagate over many correlation lengths of the potential, which we can consider it to be a series of uncorrelated events. With l as the correlation length of the potential, the time for the particle to cover a correlation length is $t_1 = l/(2E/m)^{1/2}$. We then propagate over $N = t/t_1$ correlation lengths, and approximate the integral as a sum over these independent pieces:

$$\int_0^t dt'\, \partial_{yy} V(\vec{x}(t')) \approx \sum_{i=1}^{N} t_1 \left(\partial_{yy} V\right)_i . \qquad (8.22)$$

We now need to consider these $(\partial_{yy}V)_i$ terms. As we will now argue, they are given by $(\partial_{yy}V)_i = \epsilon r_i / l^2$, where the r_i are random numbers with zero mean and standard deviation σ_r. The ϵ is obvious, since differentiation is a linear operation. The l dependence comes in from the second derivative and the fact that l sets the length scale in the potential. Other constants of proportionality go into σ_r. One can verify that the units are correct. By the central limit theorem, as $N \to \infty$,

$$\begin{aligned}
\int_0^t dt'\, \partial_{yy} V(\vec{x}(t')) &\approx \frac{t_1 \epsilon N^{1/2} \sigma_r}{l^2} && (8.23)\\
&= \frac{t^{1/2} t_1^{1/2} \epsilon \sigma_r}{l^2} && (8.24)
\end{aligned}$$

$$= \frac{t^{1/2}\epsilon\sigma_r}{l^{3/2}}\left(\frac{m}{2E}\right)^{1/4} \tag{8.25}$$

where we have approximated the integral by the standard deviation of the sum, which is the appropriate place in the distribution to look for our first self-intersection. Plugging this result into Eq. 8.21 and setting $dp_y/dy = 0$ at the time τ, we have

$$\frac{m}{\tau} - \frac{\tau^{1/2}\epsilon\sigma_r}{l^{3/2}}\left(\frac{m}{2E}\right)^{1/4} = 0 \tag{8.26}$$

$$\tau^{-3/2} = \frac{\epsilon\sigma_r}{l^{3/2}}\left(\frac{1}{2Em^3}\right)^{1/4} \tag{8.27}$$

$$\tau^{-3/2} = \frac{\epsilon}{E}\frac{\sigma_r}{l^{3/2}}\left(\frac{E^3}{2m^3}\right)^{1/4} \tag{8.28}$$

$$\tau = \left(\frac{\epsilon}{E}\right)^{-2/3} l \left(\frac{2m^3}{E^3\sigma_r^4}\right)^{1/6}. \tag{8.29}$$

The essential piece of this result is the $(\epsilon/E)^{-2/3}$ dependence. The additional energy dependence is merely a result of the time scaling already noted in §8.3. The result is perhaps clearer when we write the distance $d = \tau(2E/m)^{1/2}$,

$$d = \left(\frac{2E}{m}\right)^{1/2}\left(\frac{\epsilon}{E}\right)^{-2/3} l \left(\frac{2m^3}{E^3\sigma_r^4}\right)^{1/6} \tag{8.30}$$

$$\frac{d}{l} = \left(\frac{\epsilon}{E}\right)^{-2/3}\left(\frac{2}{\sigma_r}\right)^{2/3}. \tag{8.31}$$

With this result in hand, we now turn to numerics to test it. Again, we most directly look at $\tau(\epsilon)$, time being the independent parameter. Fixing E, however, we are clearly probing the same relationship as that for d. In Figure 8.5, we see results in the range $\epsilon/E \in (0, .2)$ and a numerical fit to the data. The $-2/3$ power law succeeds to much higher values of ϵ/E than we had a right to expect.

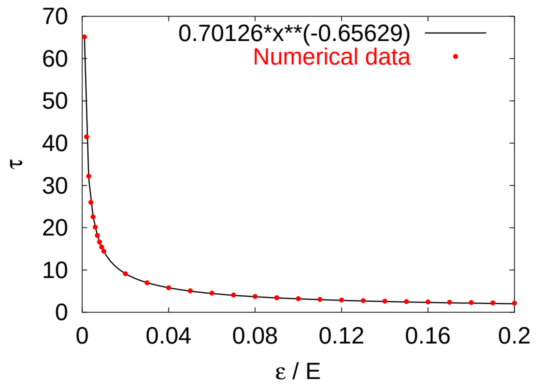


Figure 8.5: In this figure, we see the values of τ, the minimum time to a self-intersection event, as a function of potential strength ϵ/E. The fit is performed numerically, and has an exponent of $-.656$, very close to the predicted $-2/3$. Because the function is changing more rapidly for small ϵ/E, we have taken more points in that region.

8.4.4 Phase Space Structures

The branches that we see in coordinate space are, of course, merely projections of structures in phase space. For example, as noted in §8.4.1, when we look at the stability of branches we only see the stability of one trajectory through each point. The evolution of these structures can provide some insight into the formation and persistence of these branches.

To see the structures in phase space, we use the method known as the surface of section, though we do so in a somewhat non-standard fashion. The usual case is to study a system with an energy hyper-surface (that is, the surface in phase space to which energy conservation restricts our trajectories) that is bounded in coordinate space. One then defines a single surface of section and looks at successive intersections of trajectories with this surface [37]. In our system, there is no spatial bound on the trajectories and we are not looking for periodic behavior. We define instead a series of surfaces of section at equally spaced intervals, and look at intersections of trajectories with these surfaces. In some sense, we take surface number as a discretized time coordinate.

We use the rarefaction information that we are already collecting to color the phase space slices, allowing us to compare the structures with the stabilities. At the same time, we track the coordinate space projection of the phase space information, so that we can further correlate structures and stabilities to branches.

Looking at the full information in phase space, the best correlation to branch location still appears to be rarefaction exponent. We see, however, an unanticipated correlation between areas of low rarefaction and turning points in the manifold; this notion is developed further in Chapter 9. For the strong, persistent branches, we see stable "kinks" in the phase space distribution forming on a length scale small compared to the correlation length of the potential. These knots then contain locally high densities of trajectories, and are unlikely to be broken up by the action of the potential until they have spent much time shearing. The regions of phase space that do not project to branches do contain complex structures, including the cusp catastrophes characteristic of focusing. These structures are, however, sufficiently sparsely populated that they do not produce globally significant flux in coordinate space. See Figure 8.6 for an example of the phenomena mentioned here.

8.5 Branch Management

In the rarefaction exponent, we have a measure that seems to correlate to the location of branches. We would like to show that, in addition to being descriptive,

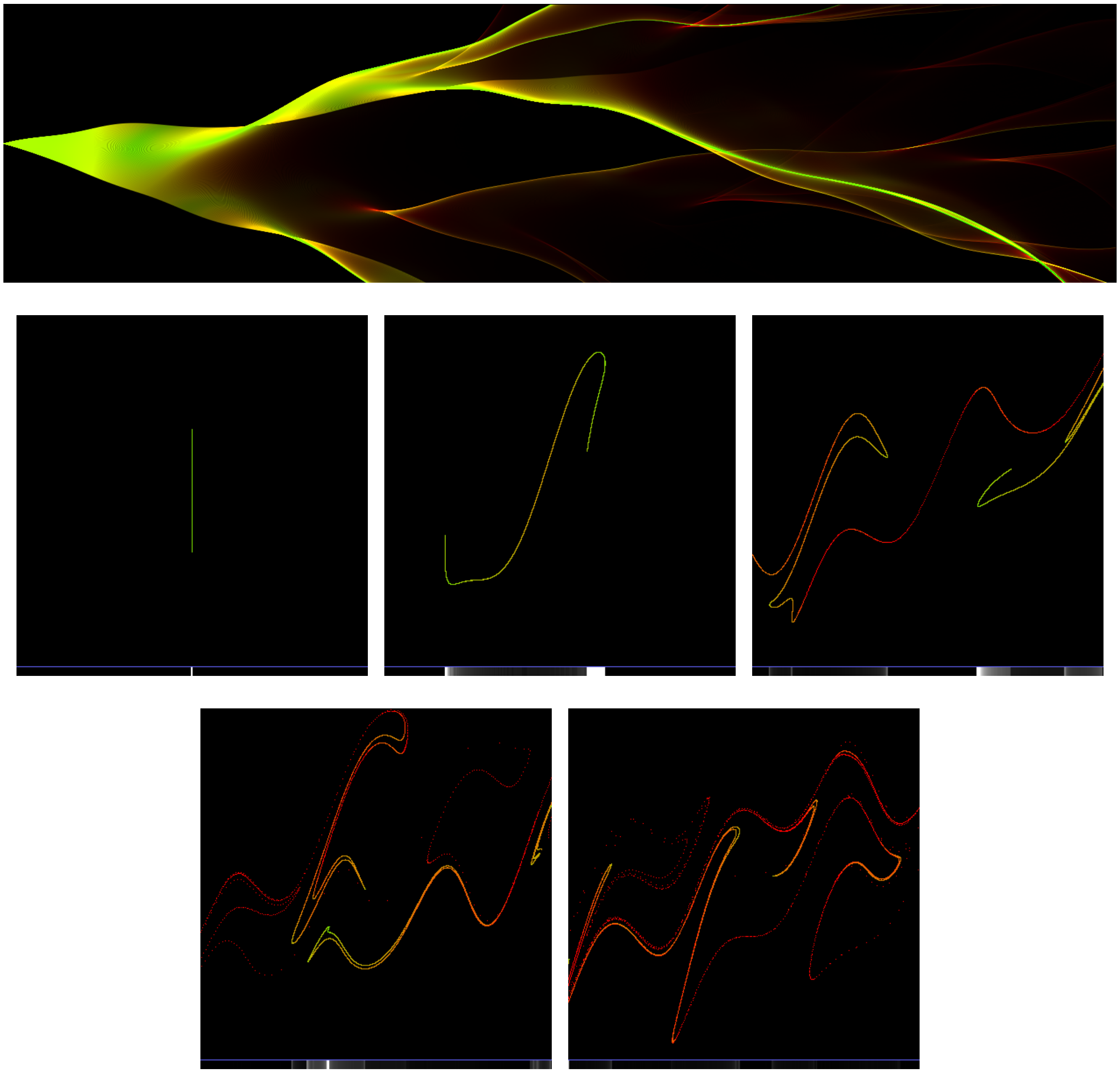


Figure 8.6: In this figure, we see the information that is available to us in the full phase space. At the top, we see the usual picture of branched flow, colored by rarefaction exponent. At the bottom, we have five surface of section images taken from evenly spaced points along the image. In this sequence, color is determined by rarefaction exponent and, in the bar at the bottom, we show the coordinate-space projections. We note first that the regions of high density in coordinate space correlate well to low rarefaction exponent, as noted before. We also see that the regions of low rarefaction exponent tend to correlate to the vertical parts of the manifold. Finally, we see that the strong branch going to the right of the top image is the projection of a small kink in the manifold.

there is a useful predictive power to the rarefaction exponent. The stability matrix for a trajectory is independent of the initial manifold, but the rarefaction exponent isn't; we can thus manufacture a manifold that will minimize the rarefaction exponent for a given trajectory. If we are correct about this measure, the resulting flow should have a branch along that trajectory.

We begin by generating a potential, running a single trajectory, and finding $\vec{x}(t)$ and $\mathbf{M}(t)$. We then select some time τ. If we had complete freedom in the orientation of the initial manifold, we would find the eigenvectors of $\mathbf{M}(\tau)$ and align the manifold along the eigenvector with the smallest eigenvalue. We do not, however, have complete freedom. We clearly restrict ourselves to creating an initial manifold that lies within the energy hyper-surface. Additionally, we choose to restrict ourselves to points that lie in a given surface of section.

To deal with our restriction to a subspace of the full phase space, we choose vectors $\{\hat{v}_i\}$ that span the subspace. We then find the "reduced" stability matrix $\mathbf{M}'(t)$, where

$$\mathbf{M}'(t)_{ij} = \hat{v}_i \cdot \mathbf{M}(t) \cdot \hat{v}_j. \tag{8.32}$$

We can then find the eigenvectors of $\mathbf{M}'(\tau)$ and choose an initial manifold in the full phase space that lies along the eigenvector with the smallest eigenvalue. Because of expansion or compression along the directions orthogonal to our subspace, the eigenvalue of $\mathbf{M}'(\tau)$ will not give the rarefaction exponent for this trajectory with this manifold. This selection of manifold does, however, give us the minimum rarefaction exponent possible with our restrictions.

In practice, there are reasons that this prescription may not work as expected.

Since the rarefaction exponent of a trajectory is a function of time, we are only assured of a small rarefaction exponent at the selected time τ; we know nothing about the rest of the history of the trajectory. What we may see is a trajectory with a focus at time τ but no nearby trajectories for $t \neq \tau$. It may also be the case that the rarefaction exponent for some time $t < \tau$ is sufficiently positive that the nearby trajectories leave the region of linear dynamics. These problems, of course, compound as τ grows.

Despite these problems, numerical experiments have shown good results using this method. In Figure 8.7, we see representative results.

At the moment, this control method is merely an interesting theoretical result, adding credence to our belief that rarefaction is a good tool for describing branches. It has not been turned into an experimental control method, and I don't see any obvious way to achieve this kind of manifold control in the context of 2DEGs. Were it possible to do so, the ability to direct a branch to a desired location, with the consequent increase in signal strength over the mean, might find practical application.

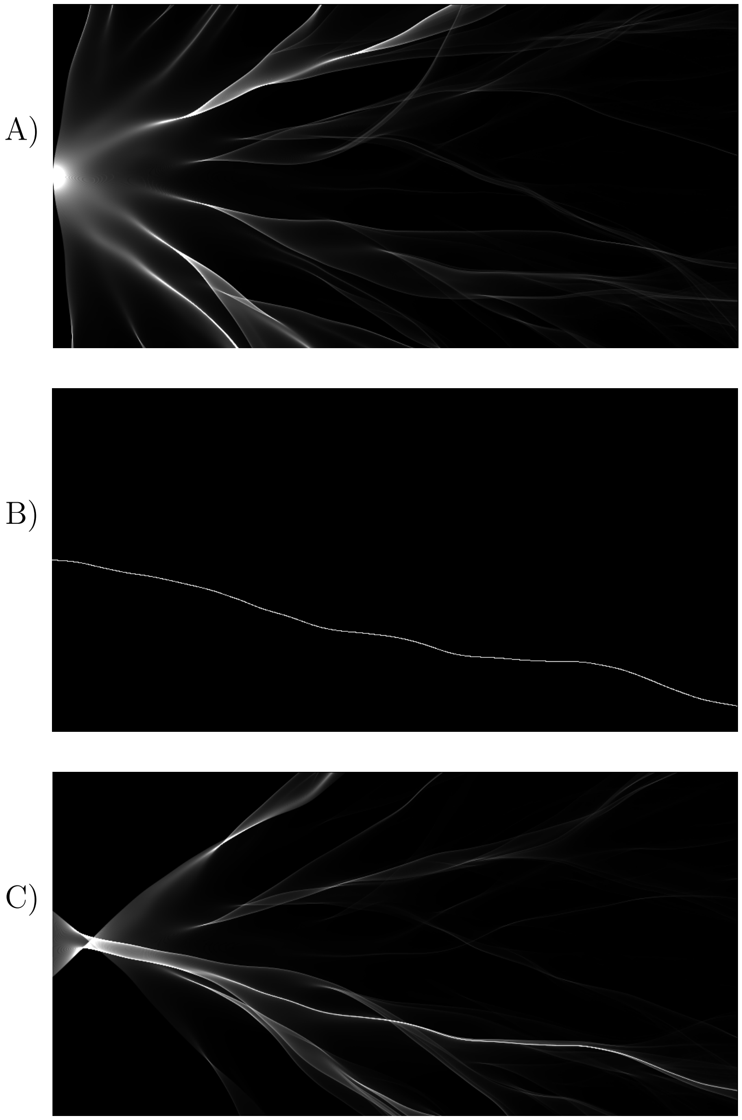


Figure 8.7: Here we see the construction of a manifold to put a branch along a given trajectory. In (A), we see flux over a random potential. The initial manifold is our usual case of trajectories coming from a single point with a distribution of angles. In (B), we have isolated a trajectory from the manifold in (A) that clearly doesn't lie in a branch. In (C), we show the flux with an initial manifold rotated from that in (A), selected so as to minimize the rarefaction exponent of the trajectory in (B). We see that the result is a strong branch along the selected trajectory.

Chapter 9

One-Dimensional Maps

In moving from quantum-mechanical to classical calculations, we stepped away from the original system of interest but gained some analytical tools. We can take another step by moving from the two-dimensional classical dynamics to a one-dimensional mapping system. Though we find that not all of the properties seen in the two-dimensional dynamics carry over to the one-dimensional map, we can take advantage of additional analysis tools in the one-dimensional system. In this chapter, we explore some of what we can learn from the mapping system.

The primary advantage of the one-dimensional mapping systems is speed. One can calculate the evolution of many more trajectories in such a system, and therefore get better statistical information in shorter periods of time. Additionally, the tools that we used to look at the evolution of phase-space structures classically are simpler to apply for mapping systems. Though we note them here, we will not be using these advantages in this thesis. In this chapter, we begin by laying out the method for converting the two-dimensional system to a one-dimensional mapping system.

We then note some differences that crop up between this system and the original. Finally, we extract an analytical result to back up observations made in phase space in Chapter 8.

9.1 Mapping Equations

We must begin by laying out the equations that we will use for our one-dimensional mapping system. There are two fundamental differences from the full two-dimensional dynamics. First, we explicitly move from differential equations with a continuous time variable to a system of discrete time steps. Second, to mimic the two-dimensional propagation, we will have a changing (i.e., time-dependent) potential in our map. By creating a system with a time-dependent potential, we give up energy conservation.

In writing our one-dimensional equations of motion, we cannot use the simplest form of Newton's equations, out of a concern for area preservation in phase space. Though we are giving up many other properties, area preservation is a sufficiently important one that we will make some efforts to retain it. Thinking back to Chapter 6, recall that area preservation is reflected in the determinant of $\mathbf{M}$ remaining 1. In the derivation of the differential equations for $\mathbf{M}$, we at one point utilized discrete time to find a matrix $\mathbf{DM}(t+\tau,t)$ in terms of which $\mathbf{M}(t+\tau) = \mathbf{DM}(t+\tau,t)\mathbf{M}(t)$. Taking $\tau \rightarrow 0$ led us to the differential equations.

Now, rather than taking the limit $\tau \rightarrow 0$, we are stuck with a finite τ between steps of the map. For the simplest set of one-dimensional equations, we had

$$\mathbf{DM}(t+\tau,t) = \begin{bmatrix} 1 & \tau m^{-1} \\ -\tau V_{qq}(t,q(t)) & 1 \end{bmatrix}. \tag{9.1}$$

If we use this form for finite τ, $\det(\mathbf{DM}) \neq 1$. We can solve this with an additional term in the upper left-hand corner of the matrix; there are several ways to arrive at this term, but let's just see what it needs to be and then interpret the results. If we take

$$\mathbf{DM}(t+\tau,t) = \begin{bmatrix} 1-\tau^2 m^{-1}V_{qq} & \tau m^{-1} \\ -\tau V_{qq} & 1 \end{bmatrix} \tag{9.2}$$

we have the desired result. Translating this matrix back into equations of motion, we have that

$$\begin{aligned} p(t+\tau) &= p(t) - \tau V_q(q(t)) && (9.3) \\ q(t+\tau) &= q(t) + \tau m^{-1} p(t) - \tau^2 m^{-1} V_q(q(t)) && (9.4) \end{aligned}$$

which can be re-written as

$$\begin{aligned} p(t+\tau) &= p(t) - \tau V_q(q(t)) && (9.5) \\ q(t+\tau) &= q(t) + \tau m^{-1} p(t+\tau). && (9.6) \end{aligned}$$

We get exact area preservation for any τ if we use these modified equations. The modification only requires that we update the momentum first, then use the new momentum value to update the position. Clearly, there are other ways that we could have achieved area preservation, but this method results in the most sensible alteration to our equations of motion.

As mentioned earlier, our mapping system is explicitly one of discrete time steps, so we can eliminate the "time" coordinate entirely in favor of a step index. At this point, we also need to make our potential explicitly step-dependent. Thus we arrive at

$$\begin{aligned} p_{i+1} &= p_i - \tau \partial_q V_i(q_i) && (9.7) \\ q_{i+1} &= q_i + \tau m^{-1} p_{i+1}. && (9.8) \end{aligned}$$

9.2 Mapping Potentials

The mapping system gives us even more flexibility in determining the potential. Let us first explore the version most closely related to the two-dimensional dynamics already studied. Let $V(x, y)$ be a two-dimensional potential that we might use in the full dynamics. We can now take slices of this potential at discrete steps in x, letting $V_i(q) = V(i\,\Delta x, q)$. The choice of step size Δx can be related to our choice of time step τ in the mapping system. Though we cannot preserve energy in the mapping system, we can at least start with the trajectory energies near some value E. We can then take

$$\frac{\Delta x}{\tau} = v = \left(\frac{2E}{m}\right)^{1/2}. \qquad (9.9)$$

One clear step that we could take away from the full classical system is to abandon this choice of Δx. By varying the value used, we can change the amount of correlation between the potential at one step and the next.

There is also no reason that we need to choose the potential so closely related to that of an actual two-dimensional system. To go to the other extreme, we could generate new random one-dimensional potentials from some ensemble at each time step. If we want to maintain some degree of correlation between potentials at each step, we can "mix in" a new random potential at each step with some mixing factor to determine the degree of correlation.

For the most part, we will use potentials generated by taking slices of an actual two-dimensional potential. One should note that this is a logical extension of work done with the two-dimensional dynamics where we looked at successive surfaces of section.

9.3 Branches in Mapping Systems

The branched flow has proven so robust to date, it should come as no surprise that is persists when we move to a one-dimensional mapping system. It does so in any kind of potential that we use, both those that come from slices of two-dimensional potentials and those that are generated as one-dimensional potentials with forced correlations. In Figure 9.1, we have an example.

The relationship between branch location and low rarefaction exponent, however, doesn't seem as strong in the mapping systems as it is in the full two-dimensional dynamics. For those potentials that have their origins in full two-dimensional potentials, the correspondence is still there. Looking at our more general potentials, however, it is largely broken, as shown in Figure 9.2. What we see in those cases is a greater dependence of stability on index number. The failure of Lyapunov exponents to correlate with branch location, however, is still true in all cases.

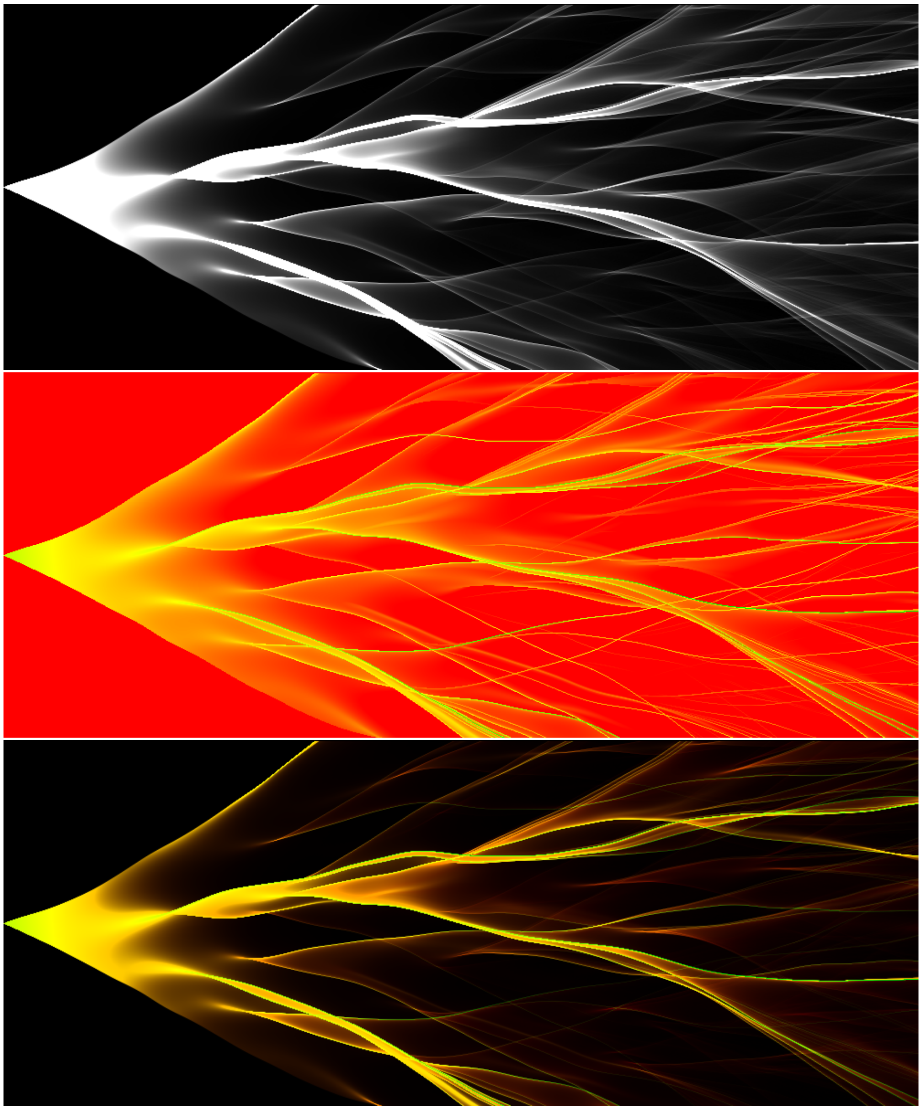


Figure 9.1: Here we have three images of branched flow in a mapping system where the potential is a true two-dimensional one (here, a superposition of plane waves). The vertical axis of each picture is the q coordinate, the horizontal axis is the time, or step index. Trajectories are launched from a single point with a distribution of momenta. In the first image, we see the greyscale image of the flux, which shows the branching phenomenon. In the second image, we have the rarefaction information for the same flow. In the third, we have the flux information again, now colored by the rarefaction information. We note that, though there is generally good agreement between flux and rarefaction, there are areas of small rarefaction exponent that do not seem to show high flux.

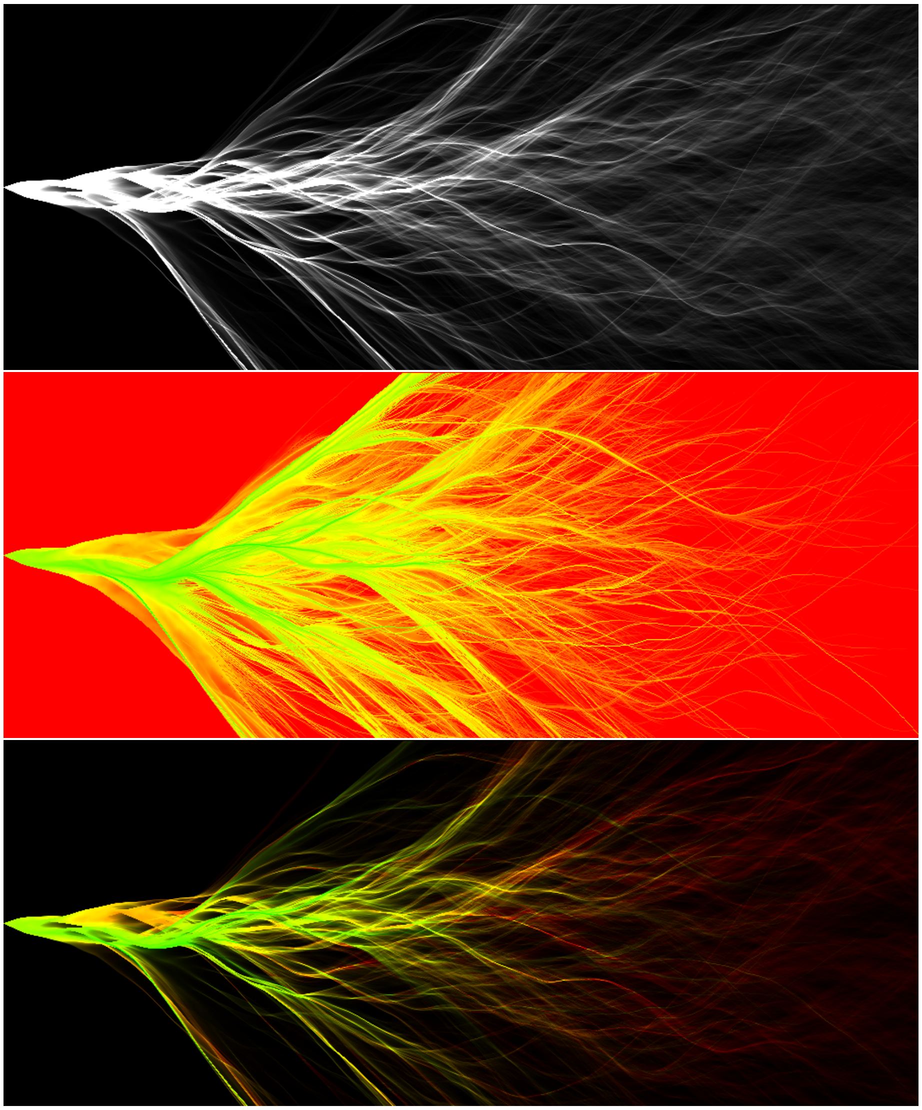


Figure 9.2: Here we have three images of branched flow in a mapping system where the potential is one-dimensional, with a new random potential mixed in at every step. This does not correspond to a sliced two-dimensional potential. The vertical axis of each picture is the q coordinate, the horizontal axis is the time, or step index. Trajectories are launched from a single point with a distribution of momenta. In the first image, we see the greyscale image of the flux, which shows the branching phenomenon. In the second image, we have the rarefaction information for the same flow. In the third, we have the flux information again, now colored by the rarefaction information. In this system, the relationship between rarefaction exponent and branch location is poor at best. Beginning very early in the flow, we see areas of high flux but high rarefaction exponents, and areas of low flux with low rarefaction exponents.

9.4 Stability Analysis

In light of the apparent breakdown of the relationship between flux density and rarefaction exponent, it is dubious to draw conclusions about that relationship based on mapping systems. Nevertheless, the simplifications that come with the mapping system allow analytic work that we can only do in the full two-dimensional dynamics with some hand waving. In particular, we use the mapping system to explore the noted relationship between low rarefaction exponent and "vertical" manifolds. The "verticality" of the manifold in phase space was one early candidate for determining where branches form, as such areas of phase space will have many trajectories built up at the same coordinate space location. Though that relationship didn't pan out, we do see that areas of low rarefaction exponent are frequently areas where the manifold is vertical (the converse fails to be true). We will explore this relationship.

Let us consider the following mapping system. For n_0 steps, the manifold propagates through free space. Then there is one time step with a potential $V(q)$. After passing over this potential, there are n_1 more steps of free space. This gives us

$$\mathbf{DM}^{n_{0,1}} = \begin{bmatrix} 1 & n_{0,1}\tau m^{-1} \\ 0 & 1 \end{bmatrix} \tag{9.10}$$

$$\mathbf{DM}^{V} = \begin{bmatrix} 1-\tau^2 m^{-1} V_{qq} & \tau m^{-1} \\ -\tau V_{qq} & 1 \end{bmatrix}. \tag{9.11}$$

Putting this together, we have $\mathbf{M} = \mathbf{DM}^{n_1}\mathbf{DM}^{V}\mathbf{DM}^{n_0}$. Now, we find a condition on the potential that yields a vertical manifold. We require

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \cdot \mathbf{M} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0 \tag{9.12}$$

This gives us

$$\left(1-\tau^2 m^{-1} V^v_{qq}\right) n_0 \tau m^{-1} + \tau m^{-1} + \left(1 - n_0 \tau^2 m^{-1} V^v_{qq}\right) n_1 \tau m^{-1} = 0 \tag{9.13}$$

$$V^v_{qq} = \frac{1}{\tau^2 m^{-1}} \left(\frac{1}{n_0} + \frac{1}{n_1+1}\right) \tag{9.14}$$

where the superscript v indicates that this is the condition for verticality of the manifold.

Plugging this potential value into the stability matrix, we get the simplified version

$$\mathbf{M}^v = \begin{bmatrix} -(n_1+1)n_0^{-1} & 0 \\ -m\tau^{-1}\left(n_0^{-1} + (n_1+1)^{-1}\right) & -n_0(n_1+1)^{-1} \end{bmatrix}. \tag{9.15}$$

It is now simple to find the rarefaction exponent as well, which gives us

$$\ln \left| \mathbf{M}^v \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right| = \ln \left[\frac{n_0}{n_1+1}\right]. \tag{9.16}$$

Suppose that, rather than looking for a vertical manifold, we seek to minimize the rarefaction exponent. That is equivalent to finding

$$\min \left| \mathbf{M} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right| \tag{9.17}$$

as a function of V_{qq}. Much algebra, of which we will spare the reader, yields

$$V^r_{qq} = \frac{1}{\tau^2 m^{-1}} \left[\frac{1}{n_0} + \frac{n_1+1}{\tau^{-2}m^2 + (n_1^2+1)^2}\right] \tag{9.18}$$

and the minimum rarefaction exponent is

$$\ln \left| \mathbf{M}^r \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right| = \frac{1}{2} \ln \left[\frac{n_0^2}{\tau^{-2}m^2 + (n_1+1)^2}\right] \tag{9.19}$$

where we have let $\mathbf{M}^r \equiv \mathbf{M}|_{V_{qq} \to V^r_{qq}}$.

We note that, in the limit $n_1 + 1 \gg \tau^{-1}m$, $V^r_{qq} = V^v_{qq}$. Additionally, in the same limit we find that the rarefaction exponents in each case go to the same value (which happens to approach $-\infty$ as $n_1 \to \infty$). So, as one propagates farther from the focusing event, the condition to be vertical becomes the same as the condition to have the minimum rarefaction exponent.

What about the Lyapunov exponent for these vertical regions? From Eq. (9.15), we see that the eigenvalues are $[-(n_1{+}1)/n_0]^{\pm 1}$. Thus, in the same large n_1 limit where verticality corresponds to small rarefaction exponent, we find the Lyapunov exponent increasing without bound. These results are shown graphically in Figures 9.3 and 9.4.

We conclude from this result that the process of shearing, in the absence of any potential, will concentrate trajectories in the turning points of the manifold. This results in a correlation between rarefaction exponent and vertical sections of the manifold, a correlation observed in Chapter 8. Even in regions where there are no strong branches, we expect that turning points in the manifold will correspond to locally low rarefaction exponent.

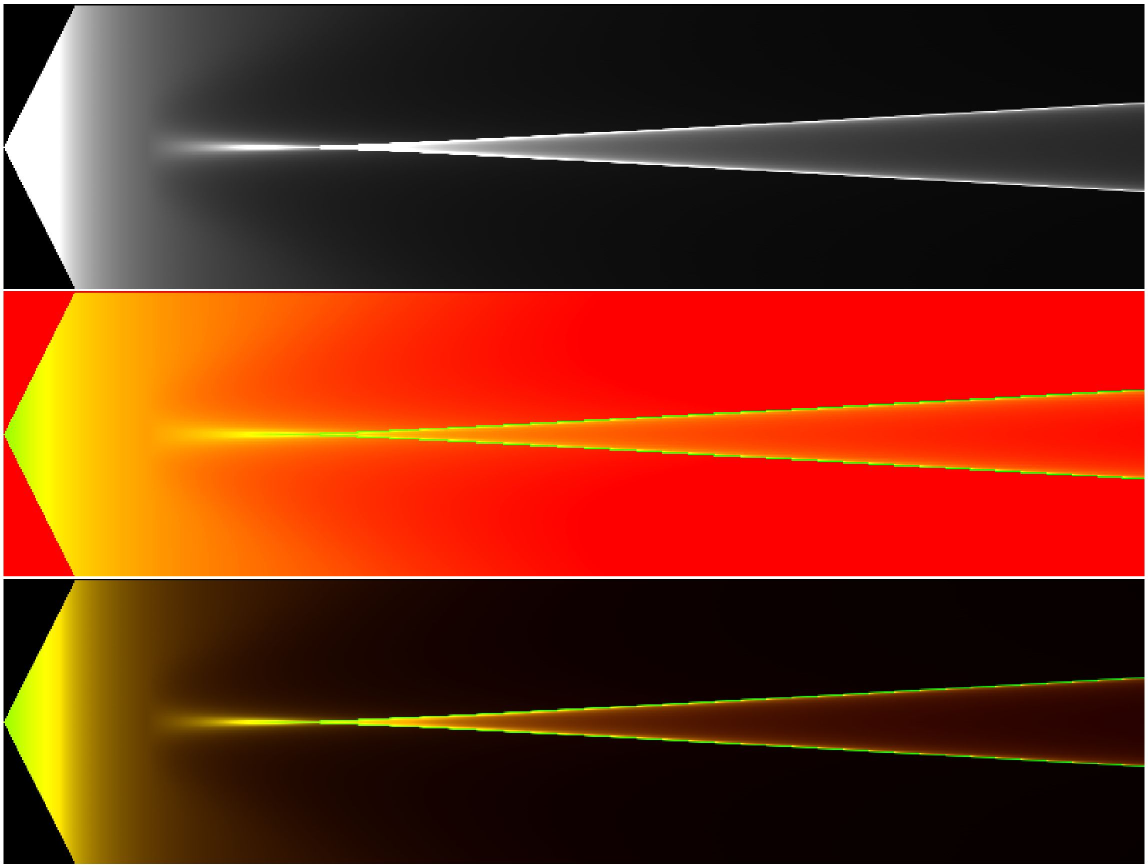


Figure 9.3: Here we see flux in a simple mapping system. The trajectories pass over 100 steps of flat potential before hitting a single focusing dip. They then pass over another 700 steps of flat potential. From top to bottom we have the flux, the rarefaction exponent, and the flux colored by the rarefaction exponent. We see that the two concentrations of flux are also highly localized areas of low rarefaction exponent, a correspondence that gets better as we move from left to right.

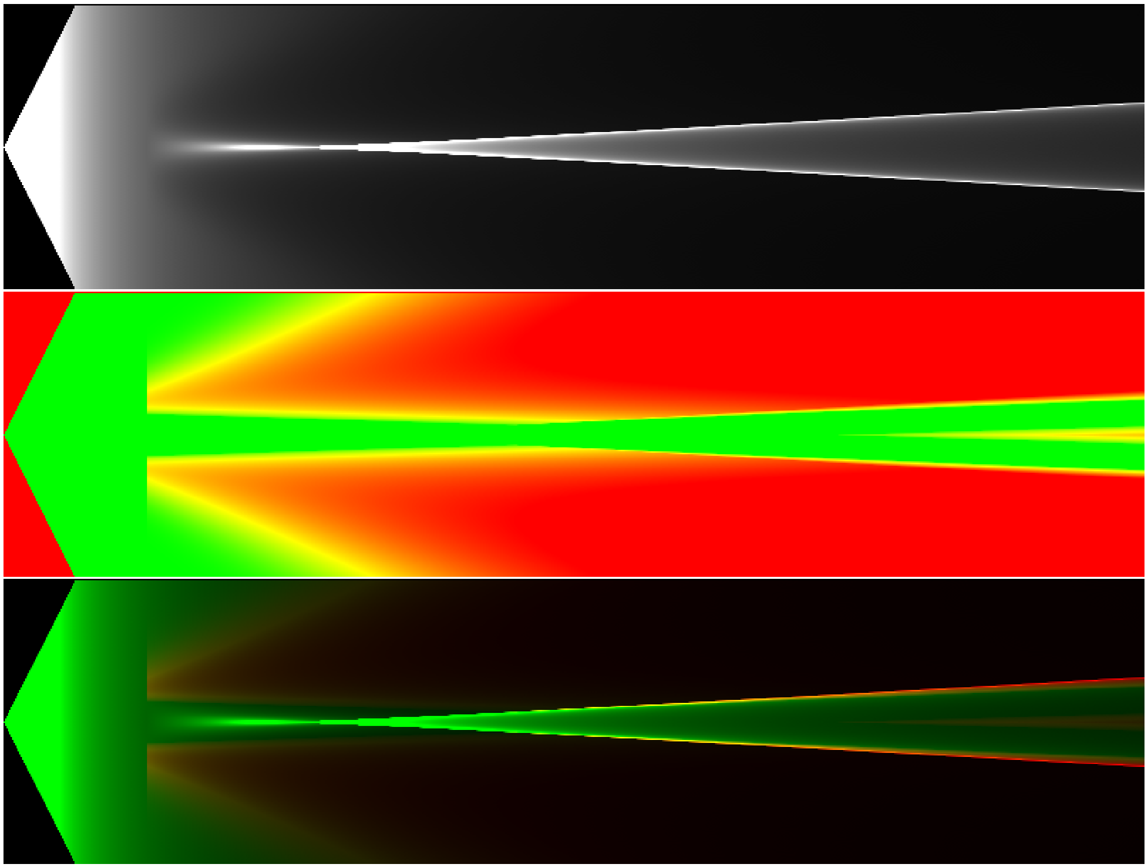


Figure 9.4: Here we see flux in the same simple mapping system as Figure 9.3. The trajectories pass over 100 steps of flat potential before hitting a single focusing dip. They then pass over another 700 steps of flat potential. From top to bottom we have the flux, the Lyapunov exponent, and the flux colored by the Lyapunov exponent. We see that the areas of low Lyapunov exponent do not correspond to the concentrations of flux. Furthermore, these concentrations are growing less stable in this measure the farther we go to the right.

Bibliography

[1] D. Baye and P.-H. Heenen, Journal of Physics A **19**, 2041 (1986).

[2] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, New York, 1997).

[3] F. B. Hildebrand, *Introduction to Numerical Analysis* (McGraw-Hill, New York, 1974).

[4] D. Baye, Journal of Physics B **28**, 4399 (1995).

[5] H. Karabulut and E. L. Sibert, Journal of Physics B **30**, L513 (1997).

[6] V. Szalay, Journal of Chemical Physics **99**, 1978 (1993).

[7] A. H. Barnet, Ph.D. thesis, Harvard University Department of Physics, 2000.

[8] E. J. Heller, Physical Review Letters **53**, 1515 (1984).

[9] M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Physical Review B **31**, 6207 (1985).

[10] D. J. Thouless and S. Kirkpatrick, Journal of Physics C **14**, 235 (1981).

[11] P. A. Lee and D. S. Fisher, Physical Review Letters **47**, 882 (1981).

[12] T. Kawamura and J. P. Leburton, Physical Review B **48**, 8857 (1993).

[13] K. Nikolić and A. MacKinnon, Physical Review B **50**, 11008 (1994).

[14] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, fourth ed. (Academic Press, New York, 1995).

[15] S. Datta, *Electronic Transport in Mesoscopic Systems* (Cambridge University Press, New York, 1995).

[16] C. M. Marcus *et al.*, Physical Review Letters **69**, 506 (1992).

[17] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999).

[18] H. U. Baranger and A. D. Stone, Physical Review B **40**, 8169 (1989).

[19] R. Feynman, *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1964).

[20] *Mesoscopic Electron Transport*, No. 345 in *NATO Asi Series E, Applied Science*, edited by L. L. Sohn, L. P. Kouwenhoven, and G. Schon (Kluwer Academic Publishers, Boston, 1997).

[21] J. H. Davies, *The Physics of Low-Dimensional Semi-Conductors: An Introduction* (Cambridge University Press, New York, 1997).

[22] M. A. Topinka *et al.*, Science **289**, 2323 (2000).

[23] M. A. Topinka *et al.*, Nature **410**, 183 (2001).

[24] B. J. van Wees *et al.*, Physical Review Letters **60**, 848 (1988).

[25] D. A. Wharam *et al.*, Journal of Physics C **21**, L209 (1988).

[26] C. W. J. Beenakker and H. van Houten, Solid State Physics **44**, 1 (1991).

[27] L. W. Molenkamp *et al.*, Physical Review B **41**, 1274 (1990).

[28] K. L. Shepard, M. L. Roukes, and B. P. van der Gaag, Physical Review Letters **68**, 2660 (1992).

[29] M. Eriksson *et al.*, Applied Physics Letters **69**, 671 (1996).

[30] J. S. Hersch, Ph.D. thesis, Harvard University Department of Physics, 1998.

[31] J. A. Katine *et al.*, Physical Review Letters **79**, 4806 (1997).

[32] D. J. Griffiths, *Introduction to Electrodynamics*, 2nd ed. (Prentice Hall, Englewood Cliffs, New Jersey, 1989).

[33] R. Grill and G. H. Dohler, Physical Review B **59**, 10769 (1999).

[34] B. J. LeRoy, private communication.

[35] A. Kalben, private communication.

[36] *Les Houches, Session LII: Chaos and Quantum Physics*, edited by M.-J. Giannoni, A. Voros, and J. Zinn-Justin (North-Holland, New York, 1991).

[37] E. Ott, *Chaos in Dynamical Systems* (Cambridge University Press, New York, 1993).

[38] V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York, 1978).

[39] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (IEEE Press, New York, 1997).

[40] M. A. Wolfson and S. Tomsovic, Journal of the Acoustical Society of America **109**, 2693 (2001).

[41] P. O'Connor, J. Gehlen, and E. J. Heller, Physical Review Letters **58**, 1296 (1987).

[42] M. G. Brown *et al.*, Ray Dynamics in Ocean Acoustics, 2001, nlin.CD/0109027, available at `xxx.lanl.gov`.

[43] K. P. Rauch, S. Mao, J. Wambsganss, and B. Paczynski, Astrophysical Journal **386**, 30 (1992).

[44] B. Eckhardt and D. Wintgen, Journal of Physics A **24**, 4335 (1991).