DOE/ET-53088-363 IFSR #363

Diffusion in Symplectic Maps

Hyung-tae Kook and *James D. Meiss*
Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712

February 1989

Diffusion in Symplectic Maps

Hyung-tae Kook and James D. Meiss
Institute for Fusion Studies
University of Texas
Austin, TX 78712

February 16, 1989

Abstract

The characteristic function method is used to obtain the diffusion tensor for symplectic maps. At lowest order the quasilinear result is obtained, and a series in higher order correlations is developed. Comparison of the theory to numerical experiments is given using a four-dimensional example of Froeschlé. The experiments agree well with the theory for moderately large parameters. Arnol'd diffusion for the 'thick layer' case is discussed. It is shown that the short time correlations in one canonical plane affect the diffusion in the other plane even in the limit of zero coupling. Accelerator modes exist for the Froeschlé example, and cause divergences in the diffusion, but these only appear when the accelerating region is included in the ensemble.

§1. Introduction

Symplectic mappings provide the simplest, non-trivial models of Hamiltonian dynamics. A map can be obtained from a continuous time Hamiltonian system by Poincaré surface of section[1]: for N degrees of freedom the surface of section gives a mapping of a 2N-2 dimensional surface to itself. The map is symplectic as a consequence of the preservation of the Poincaré invariants.

Generically such maps are not integrable, but also not ergodic. The rigorous treatment of the subset of apparently chaotic orbits by statistical methods is a long sought, but elusive goal. One statistical property of interest is the diffusion coefficient. This has been extensively studied for area preserving maps. Notably, for the standard map, Chirikov[1] showed numerically that the momentum diffusion of an ensemble of particles approaches the "quasilinear " value as the nonlinearity parameter increases. Quasilinear diffusion results if the force is a delta correlated random process. For the standard map, the approach to quasilinearity is not monotonic: the oscillations of the diffusion coefficient were explained by Rechester et al[2] using an expansion called the Fourier path method. This was extended and modified to apply to many maps using the characteristic function formalism.[3,4] Here it was shown that the oscillations are due to short time correlations. The characteristic function method allows the exact calculation of the diffusion coefficient for Arnol'd's "cat maps",[5,6] and can be used for maps with nonlinear twist terms.[7,8]

In general these expressions for the diffusion coefficient agree well with numerical experiments, providing the measurement time is moderate. As the time is increased, long time tails in the correlation functions[9,10] become increasingly important, and the diffusion coefficient develops sharp peaks.[11,12] The height of the peaks grows with the computation time; thus the series expressions for D are not convergent. These anomalies are due to the presence of stable "accelerator modes" in the standard map. We will discuss these further in §6 for the four dimensional case.

In this paper we generalize the characteristic function formalism for mappings of the form

$$\mathbf{x}_{t+1} - 2\mathbf{x}_t + \mathbf{x}_{t-1} = -\nabla V(\mathbf{x}_t) \ , \qquad (1)$$

Here $\mathbf{x}$ is an N-dimensional configuration vector, and $V(\mathbf{x})$ is an arbitrary periodic function of $\mathbf{x}$: $V(\mathbf{x} + 2\pi\mathbf{m}) = V(\mathbf{x})$, for any integer vector $\mathbf{m}$. Defining the momentum coordinate by

$$\mathbf{y}_t = \mathbf{x}_t - \mathbf{x}_{t-1} \tag{2}$$

this map is easily shown to be symplectic (it preserves the area $\oint \mathbf{y}\cdot d\mathbf{x}$). Since V is periodic, the configuration space can be taken to be the N-torus. We allow arbitrary values for **y**, thus the phase space is $\mathcal{R}^N \times \mathcal{T}^N$.

For an example we will use a four-dimensional (two degree of freedom) map first studied by Froeschlé[13–14]

$$V(\mathbf{x}) = -a\cos(x^1) - b\cos(x^2) - c\cos(x^1+x^2) \tag{3}$$

where 'a' and 'b' represent the strengths of standard maps for each degree of freedom, and 'c' couples these together. The map appears not to have any global invariants unless all but one of the three parameters are zero. Many of the expressions we will give below will be for the simple case b=0, but this case does not appear to be special in any way.

§2. **Diffusion tensor**

A statistical analysis describes the behavior of an ensemble of initial conditions. We choose a uniform distribution over some invariant region in phase space, denoted as $\mathfrak{R}$. It is convenient to use the Lagrangian form of the map, eq. (1), and to average over an ensemble of initial conditions ($\mathbf{x}_0$, $\mathbf{x}_1$). Averages over this set of initial conditions are denoted $\langle\ \rangle_{\mathfrak{R}}$. Typically we will assume the configuration is arbitrary, so that $\mathfrak{R}$ is $\mathcal{T}^N \times \mathcal{T}^N$. Ideally, one would like to use a uniform distribution on the set of chaotic initial conditions for the ensemble, since the regular initial conditions do not diffuse. However, we know of no analytical techniques to determine this ensemble (which in the two dimensional case is a fat fractal[15]), nor even to show that it occupies a set of positive measure for a typical map.

The momentum diffusion tensor is defined to be the asymptotic rate of spread of the second moment of the momentum distribution

$$\mathbf{D} = \lim_{t\to\infty} \frac{1}{2t} \langle \Delta\mathbf{y}_t\, \Delta\mathbf{y}_t \rangle_{\mathfrak{R}} \tag{4}$$

where $\Delta \mathbf{y}_t(\mathbf{x}_0,\mathbf{x}_1)$ is the momentum change in t iterations for an initial condition $(\mathbf{x}_0,\mathbf{x}_1)$. The factor of two is due to the definition of diffusion in the Fokker-Planck equation[2,7]. Using (1) and (2), this can be easily rewritten as

$$\mathbf{D} = \frac{1}{2}\mathbf{C}(0) + \sum_{t=1}^{\infty}\mathbf{C}(t)$$

$$\mathbf{C}(t) \equiv \left\langle \nabla V(\mathbf{x}_0)\nabla V(\mathbf{x}_t)\right\rangle_{\Re} \tag{5}$$

where $\mathbf{C}$, the force correlation tensor, has been assumed to decay more rapidly than t^{-1}. Finally, the correlation function can be expressed in terms of the Fourier transform of ∇V

$$\nabla V = \sum_{m=-\infty}^{\infty} \mathbf{f}_m e^{i\mathbf{m}\cdot\mathbf{x}} \tag{6}$$

where $\mathbf{f}_m = \mathbf{f}^*_{-m}$, so that

$$\mathbf{C}(t) = \sum_{m,n=-\infty}^{\infty} \mathbf{f}_m\mathbf{f}_n\left\langle \exp\left(i\,\mathbf{m}\cdot\mathbf{x}_t + i\,\mathbf{n}\cdot\mathbf{x}_0\right)\right\rangle_{\Re} \tag{7}$$

Since the region $\Re$ is invariant under the mapping and the mapping is volume preserving, eq. (7) implies the time reversal symmetry $\mathbf{C}(t) = \tilde{\mathbf{C}}(-t)$, where ~ designates transpose.

For the example of eq. (3) the nonzero $\mathbf{f}_m$ are

$$\mathbf{f}_{(1,0)} = (a/2i,0), \quad \mathbf{f}_{(0,1)} = (0,b/2i), \quad \mathbf{f}_{(1,1)} = (c/2i,c/2i)\,, \tag{8}$$

plus their complex conjugates.

The task is now to obtain expressions for the ensemble average of the exponential appearing in eq. (7).

§3. **Characteristic function method**

Any statistical property of the ensemble can be obtained from the joint probability distribution $P(\mathbf{x}_0,\mathbf{x}_1,\mathbf{x}_2,.....\mathbf{x}_t)$ which is the probability that the trajectory lands

at the points $\mathbf{x}_k$ for k=0,1,...,t. The t^{th}-characteristic function is defined to be the Fourier transform of this distribution:[3-5]

$$\chi_t(\mathbf{n}_t, \mathbf{n}_{t-1}, ..., \mathbf{n}_1, \mathbf{n}_0) = \left\langle \exp(i \sum_{j=0}^{t} \mathbf{n}_j \cdot \mathbf{x}_j) \right\rangle_{\mathfrak{R}} \tag{9}$$

where the $\mathbf{n}_t$ are arbitrary integer vectors. In particular, the correlation function depends only on the set of characteristic functions with the first and last arguments nonzero. Thus it is convenient to define the reduced function

$$\chi_t(\mathbf{n}_t, \mathbf{n}_0) \equiv \chi_t(\mathbf{n}_t, \mathbf{0}, ..., \mathbf{0}, \mathbf{n}_0) \tag{10}$$

Using eq. (7) with this, the correlation function is

$$\mathbf{C}(t) = \sum_{\mathbf{m},\mathbf{n}=-\infty}^{\infty} \mathbf{f}_m \mathbf{f}_n \chi_t(\mathbf{m}, \mathbf{n}) \tag{11}$$

For the potential (3), one can form linear combinations of the components of the correlation tensor to separate the contributions explicitly due to each term in (3) independently, and those due to coupling between them. For example, when b=0 we define

$$\begin{aligned} C_{aa}(t) \equiv C^{11} - 2C^{12} + C^{22} &= -1/2\, a^2 \{\chi_t[(1,0), (1,0)] - \chi_t[(-1,0), (1,0)]\} \\ C_{ac}(t) \equiv C^{12} - C^{22} &= -1/2\, ac \{\chi_t[(1,0), (1,1)] - \chi_t[(-1,0), (1,1)]\} \\ C_{cc}(t) \equiv C^{22} &= -1/2\, c^2 \{\chi_t[(1,1), (1,1)] - \chi_t[(-1,-1), (1,1)]\} \end{aligned} \tag{12}$$

Note, however, that each of the χ_t in (12) depends implicitly on both parameters 'a' and 'c'.

Evaluation of χ_t requires knowledge of the exact orbit beginning at an arbitrary point ($\mathbf{x}_0$, $\mathbf{x}_1$), and is impossible in general. However, a remarkably simple recursion relation can be obtained, using eq. (1), which relates χ_t to χ_{t-1}:

$$\chi_t(\mathbf{n}_t, \mathbf{n}_{t-1}, ..., \mathbf{n}_1, \mathbf{n}_0) = \sum_m g_m(\mathbf{n}_t)\, \chi_{t-1}(\mathbf{n}_{t-1}+2\mathbf{n}_t-\mathbf{m}, \mathbf{n}_{t-2}-\mathbf{n}_t, \mathbf{n}_{t-3}..., \mathbf{n}_1, \mathbf{n}_0) \tag{13}$$

Here we have defined the Fourier transform of the exponential of ∇V:

$$g_{\mathbf{m}}(\mathbf{n}) \equiv \frac{1}{(2\pi)^N} \int_0^{2\pi} d^N x \exp\left[i\left(\mathbf{m}\cdot\mathbf{x} - \mathbf{n}\cdot\nabla V(x)\right)\right] \tag{14}$$

for an integer vector $\mathbf{m}$. For the Froeschlé example, $g_{\mathbf{m}}$ can be evaluated in terms of Bessel functions:

$$g_{\mathbf{m}}(\mathbf{n}) = \sum_{j=-\infty}^{\infty} J_{m^1-j}(n^1 a)\; J_{m^2-j}(n^2 b)\; J_j\left[(n^1+n^2)c\right] \tag{15a}$$

When b=0 the infinite sum in eq. (15a) collapses to a single term

$$g_{\mathbf{m}}(\mathbf{n}) = J_{m^1-m^2}(n^1 a)\;\; J_{m^2}\left[(n^1+n^2)c\right] \qquad (b=0) \tag{15b}$$

The recursion relation (13) is easily formally iterated to obtain χ_t in terms of χ_1. In particular for the class of eq. (10) we obtain the exact, if somewhat formal, expression:

$$\chi_t(\mathbf{n}_t, \mathbf{n}_0) = \sum_{\mathbf{n}_1,\ldots,\mathbf{n}_{t-1}} \prod_{k=2}^{t} g_{\mathbf{m}_k}(\mathbf{n}_k)\; \chi_1(\mathbf{n}_1, \mathbf{n}_0 - \mathbf{n}_2) \tag{16a}$$

where the $\mathbf{m}_k$ for k = 2,3,...t are defined by

$$\mathbf{m}_k = -\delta^2 \mathbf{n}_k \equiv -\mathbf{n}_{k+1} + 2\mathbf{n}_k - \mathbf{n}_{k-1} \tag{16b}$$

with $\mathbf{n}_{t+1} \equiv 0$. Equation (16) is exactly the same as for the area preserving case[5] except that the integers m_t and n_t have been replaced by integer vectors, and we have reversed the sign of $\mathbf{m}$ and the ordering of the index "k" for later convenience.

If we assume that $\mathfrak{R}$ is $\mathcal{T}^N \times \mathcal{T}^N$, then the first characteristic function is simply a Kronnecker delta

$$\chi_1(\mathbf{m}, \mathbf{n}) = \delta_{m,0}\,\delta_{n,0} \tag{17}$$

Using this in eq. (16) gives the symmetry property $\chi_t(\mathbf{m}, \mathbf{n}) = \chi_t(\mathbf{n}, \mathbf{m})$. The definition (11) then implies that the correlation matrix is symmetric, $\mathbf{C}(t) = \tilde{\mathbf{C}}(t)$, and therefore so is the diffusion matrix.

When we set b=0 in the Froeschlé map and use $\mathfrak{R} = \mathcal{T}^N \times \mathcal{T}^N$, the first three χ_t can be written explicitly:

$$\chi_2(\mathbf{m}, \mathbf{n}) = \delta_{\mathbf{m},\mathbf{n}} J_{2(n^1-n^2)}(an^1)\, J_{2n^2}\left[c(n^1+n^2)\right]$$

$$\chi_3(\mathbf{m}, \mathbf{n}) = J_{m^2-m^1+2(n^1-n^2)}(an^1)\, J_{n^2-n^1+2(m^1-m^2)}(am^1) \times$$
$$J_{2n^2-m^2}\left[c(n^1+n^2)\right] J_{2m^2-n^2}\left[c(m^1+m^2)\right] \qquad (18)$$

The next characteristic function, χ_4, is an infinite sum over an intermediate index of a product of six Bessel functions. In general when all three parameters a,b,c are non-zero, each χ_t contains sums over products of 3t Bessel functions, though it is possible that some of these could turn out to be $J_0(0) = 1$. In fact, as we will see below, terms with a maximal number of such factors dominate when the parameters are large.

§4 **Principal Terms**

Each term in the sum of eq (16) can be viewed as a particular path in the space $(\mathbf{n},k)$. Once the path is chosen, then the $\mathbf{m}_k$ are determined by the second differences through eq (13); thus $\mathbf{m}_k$ is effectively the curvature of $\mathbf{n}_k$ - path. The sum is over every path in this space which begins as $(\mathbf{0},1)$, $(\mathbf{n}_0,2)$, and ends as $(\mathbf{n}_t, t)$, $(\mathbf{0}, t+1)$.

In the area preserving case, the most important terms in this sum are those which have the minimum number of Bessel functions of nonzero argument and index. These are the so-called "principal terms". These appear to dominate[5] for large values of the parameter 'a' since $J_m(na) \propto (a)^{-1/2}$.

The same argument can be given in this case, namely $g_{\mathbf{m}}(\mathbf{n}) \leq 1$ with equality only when $\mathbf{m} = \mathbf{n} = 0$. For the later case the path must have a straight segment which passes through the origin of the $\mathbf{n}$ plane. The principal terms are those with the maximum number of such segments.

Consider the case of $\chi_t[(1,1),s(1,1)]$ where $s = \pm 1$, which is needed to compute C_{cc}. When t is even the principal path is given by the sequence, $\cdots \rightarrow (1,1) \rightarrow (0,0) \rightarrow$

$(-1,-1) \rightarrow (0,0) \rightarrow \cdots$, as shown in Fig. 1(a) by the solid line. The resulting contribution to χ_t for $t \geq 2$ is

$$\chi_t^P [(1,1),s(1,1)] = [J_0(a) J_2(2c)]^{\frac{t}{2}} \delta_{s,(-1)^{\frac{t}{2}+1}} \qquad t \text{ even} \qquad (19)$$

For odd t, the paths necessarily include at least one defect (i.e. an extra point with **m**, **n** $\neq 0$) in order to satisfy the boundary conditions at k=0 and t. Paths with a single defect, such as those shown by dashed lines in Fig. 1(a) will have the minimum number of Bessel functions. The defect can occur at any of (t-1)/2 positions. Each term gives the same result, so we obtain (for $t \geq 3$)

$$\chi_t^P [(1,1),s(1,1)] = \frac{t-1}{2} \{ [J_3(2c)]^2 \delta_{s,(-1)^{\frac{t+1}{2}}} + [J_1(2c)]^2 \delta_{s,(-1)^{\frac{t-3}{2}}} \} \times [J_0(a)]^2 [J_0(a) J_2(2c)]^{\frac{t-3}{2}} \qquad t \text{ odd} \qquad (20)$$

The paths contributing to the principal terms for $\chi_t[(1,0), s(1,0)]$ are shown in Fig. 1(b). When t is even the solid path is dominant, and when t is odd, both the dashed paths must be included. The results are equivalent to the above upon making the substitutions $2c \rightarrow a$ and $a \rightarrow c$.

Finally one would like to compute the principal terms for $\chi_t[(1,1),(\pm 1,0)]$. Such paths must also have at least one defect. However, it is not difficult to see that for each path with one defect contributing to $\chi_t[(1,1),(1,0)]$, there is one which gives the identical result for $\chi_t[(1,1),(-1,0)]$. Equation (15) implies that these terms cancel from C_{ac}. In order to compute the effect of C_{ac} on the diffusion it is necessary to include paths with two or more defects. Thus $C_{ac} = 0$ in the principal term approximation.

§5. **Quasilinear Diffusion and Corrections**

Combining eq. (5) and (11) with eq. (16) gives an explicit, formal expression for the diffusion tensor. As was the case for area preserving maps, the convergence of these series is doubtful due to the long-time correlations induced by regular regions in the phase space (in particular, the accelerator modes, see below). Such effects have not been studied when the number of degrees of freedom (N) is larger than two. However, moderate time and finite precision numerical experiments can be expected to give results close to those given by eq. (5) if a finite number of terms in t are retained.

The first term in the series for the diffusion tensor is the so-called quasilinear term:

$$\mathbf{D}_{ql} = \frac{1}{2}\mathbf{C}(0) = \frac{1}{2}\sum_{m} \mathbf{f}_m \mathbf{f}_{-m} \tag{21}$$

where we have set $\Re$ to the entire torus. For the Froeschlé example this is

$$\mathbf{D}_{ql} = \frac{1}{4}\begin{pmatrix} a^2+c^2 & c^2 \\ c^2 & b^2+c^2 \end{pmatrix} \tag{22}$$

The series in force correlations gives corrections to this which, because of the Bessel functions, tend to oscillate as the parameter values change. The first few terms are easy to write out (with b=0):

$$C_{aa}(1) = C_{cc}(1) = 0$$

$$C_{aa}(2) = -1/2 a^2 J_2(a)J_0(c)\,, \quad C_{cc}(2) = -1/2 c^2 J_0(a)J_2(2c)$$

$$C_{aa}(3) = -1/2 a^2 [\, J_1^2(a) - J_3^2(a)\,] J_0^2(c),$$

$$C_{cc}(3) = -1/2 c^2 [\, J_1^2(2c) - J_3^2(2c)\,] J_0^2(a) \tag{23}$$

The coupling correlation, $C_{ac}(t)$, vanishes identically for $t<4$.

Using the principal terms (19) and (20) for the longer time correlations we can do the sum in eq. (5) to obtain:

$$D^P_{aa} = \frac{a^2}{4}\,\frac{1 - \left[2J_1^2(a) + J_2^2(a) - 2J_3^2(a)\right] J_0^2(c)}{\left[1 + J_2(a)J_0(c)\right]^2}$$

$$D^P_{cc} = \frac{c^2}{4}\,\frac{1 - \left[2J_1^2(2c) + J_2^2(2c) - 2J_3^2(2c)\right] J_0^2(a)}{\left[1 + J_0(a)J_2(2c)\right]^2} \tag{24}$$

Note that these are not separable into terms due to the first degree of freedom, and terms due to the second. In the principal terms approximation, $D_{ac} = 0$, or $D_{12} = D_{21} = D_{22}$. We expect these results to be a good approximation when both a and c are not too small.

An interesting limit is that of $a \gg 1$, implying that the first degree of freedom is highly stochastic and $c \ll 1$, implying that the coupling is weak. In this case the self-diffusion of the first degree-of-freedom is approximately quasilinear. This degree of freedom drives the diffusion in the second, this is called "thick-layer diffusion".[16] Tennyson et al compute such diffusion by treating the second degree of freedom as though it was stochastically driven by the first. This technique yields only the lowest order, quasilinear term, in D_{cc}. The principal terms expression implies that $D_{cc} \to D_{ql}$ in the limit $c \to 0$; however, this is not exact because there are an infinite number of terms which contribute to D at $O(c^4)$ which have been neglected.

Another limit of interest, is $a \lesssim 1$, $c \ll 1$, where the initial conditions are chosen with the first degree of freedom in a narrow stochastic layer. This leads to "thin layer diffusion", or what we think of as Arnol'd diffusion.[13] Our method does not give an effective technique for computing this.

§6. **Accelerator Modes**

Maps of the form (1) are also periodic in the momentum direction: if $(\mathbf{y}_t, \mathbf{x}_t)$ is an orbit on $\mathcal{R}^N \times \mathcal{T}^N$ then $(\mathbf{y}_t + 2\pi\mathbf{j}, \mathbf{x}_t)$ is also an orbit for any integer vector $\mathbf{j}$. This suggests looking for generalized "periodic" orbits satisfying the conditions

$$\mathbf{y}_t = \mathbf{y}_0 + 2\pi\mathbf{j}, \quad \mathbf{x}_t = \mathbf{x}_0 \tag{25}$$

If $\mathbf{j} \neq \mathbf{0}$ then these orbits are called t^{th} order accelerator modes.

The condition for the first order accelerator modes, t=1, for eq. (1) is that $-\nabla V(\mathbf{x}) = 2\pi\mathbf{j} = (r,s)$ where r and s are integers, and $\mathbf{y} = \mathbf{0}$. Using the potential (3), with b=0 implies that

$$\sin(x^1) = \frac{2\pi(s-r)}{a} \qquad \sin(x^1 + x^2) = -\frac{2\pi s}{c} \tag{26}$$

Thus these can exist only if $|c| \geq 2\pi|s|$ and $|a| \geq 2\pi|r-s|$. In general there are four such fixed points for each (r,s), of which only one is stable. The requirement for stability (all four eigenvalues of modulus unity) is determined by linearizing the map about the accelerator orbit:

$$\alpha\gamma \geq 0\,, \quad (\alpha\text{-}8)(\gamma\text{-}4) \geq 16 \quad \text{and} \quad 0 \leq \alpha+2\gamma \leq 8$$

$$\text{where} \quad \alpha \equiv a\cos(x^1) \quad = \pm\sqrt{a^2 - 4\pi^2(s-r)^2}$$

$$\gamma \equiv c\cos(x^1+x^2) = \pm\sqrt{c^2 - 4\pi^2 s^2} \tag{27}$$

The regions of parameters where the stable first order accelerator modes exist are shown in Fig. 2.

In the case of area preserving maps, the existence of a stable accelerator mode causes the series for the diffusion coefficient to diverge. This is due to two effects: first, if the region $\Re$ is chosen to be entire phase space, then initial conditions trapped in an accelerator mode (which by the KAM theorem have finite measure) are included in the average.[4-11] Since Δy increases in proportion to t for such initial conditions, D necessarily diverges. The characteristic functions must include this effect; however, extracting it from the formulae appears difficult.

Second, even if initial conditions trapped in an accelerator mode are not included in the average, numerical evidence indicates that the correlation function decays as t^{-z} for large time where z is in the range 0.3 - 0.5.[9-10] This is due to long time trapping of stochastic orbits near "sticky" islands. Consequently, the series (5) must be divergent even if $\Re$ is chosen to be the connected stochastic region.

This divergence might be expected to be less severe in higher dimensional maps because initial conditions in the neighborhood of an elliptic orbit can still escape by the mechanism of Arnol'd diffusion. However, this escape can be slow,[17] and may still give an algebraic decay of the correlation function.

§7. **Numerical Experiments**

The diffusion tensor can be computed directly from eq. (4), interpreting the average as a sum over a grid of initial conditions in phase space. We iterate each initial condition for a time T, and estimate the statistical errors in D by the RMS deviation of the results for the different initial conditions. The RMS deviation is seen to decrease as inverse of the square root of the ensemble size.

Comparison of the numerical results with the principal term theory is shown in Figs. 3 for a=3.0, as c varies from 0.0 to 5.0. Here we chose a grid of 10^4 initial conditions, and iterated each T=50 times. The plots (a) and (b) show the values of D_{aa} and D_{cc} normalized to their quasilinear values, respectively. The contribution of D_{ac} to D_{12} is comparable to D_{cc} when c is small (Fig. 3c). For larger c, D_{ac} is consistent with

zero (statistical errors due to the finite ensemble prevent an accurate measurement). Results for various ranges of parameters show that providing a and c are larger than 2 the principal term results agree extremely well with the numerical experiments.

In the region of parameter space where a first order accelerator mode is stable, the diffusion coefficient is enhanced when the ensemble is $\mathcal{T}^N \times \mathcal{T}^N$. As in the case of the standard map, the peak in the diffusion coefficient grows as T is increased, indicating that convergence has not been obtained. To exhibit the role of the accelerator mode, we choose initial conditions at $\mathbf{y} = \mathbf{0}$, with $\mathbf{x}$ chosen on a 10^2 by 10^2 grid: the location as well as the size of stable islands varies as parameters vary; however, first order accelerator modes are always at $\mathbf{y} = \mathbf{0}$. Figs. 4(a)-(c) demonstrate divergent behavior of D_{aa} in T due to the first order accelerator mode with $\mathbf{j} = (1,0)$.

To attempt to eliminate orbits trapped on invariant tori encircling the stable accelerator mode, we choose an initial condition near the unstable fixed point $\mathbf{y} = \mathbf{0}$, and $\mathbf{x} =$ (0.5,0.5), iterate it a large number of iterations, and break the orbit into n segments of length T to obtain statistics:

$$\mathbf{D} = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{2T} (\mathbf{y}_{iT} - \mathbf{y}_{(i-1)T})\,(\mathbf{y}_{iT} - \mathbf{y}_{(i-1)T}) \tag{28}$$

The result is shown in Fig. 4(d) with $n=10^3$, $T=10^4$ for the same parameters as Fig. 4(c). In contrast to the area preserving case, these results show little dependence of the diffusion coefficients on T. This can be interpreted that the trapping time distribution of a stochastic orbit near the accelerator mode decays rapidly enough that the diffusion series converges. We will study the decay of correlations for this situation in a future paper.

When either of parameters is small, the effect of trapping is conceivably important. Thus, to study small c regime we use the single, long orbit statistics introduced above. Our numerical results imply that as c approaches zero, D_{aa} limits to the diffusion of the standard map for the given value of a, as would be expected. On the other hand, D_{cc} appears to scale like c^2 as c goes to zero, but does not seem to converge to its quasilinear value consistently for all values of a, as would be expected on the basis of Eq. (24). Rather, it is observed that $\lim_{c\to 0} D_{cc}/D_{ql}$ depends on a in an oscillatory fashion, only approaching one as $a\to\infty$. This is shown in Fig. 5, setting $c = 10^{-5}$ (which is small enough to be considered as the $c\to 0$ limit), $n=10^3$ and $T=10^3$. Note that the oscillations in D_{cc}/D_{ql} are in phase with those of D_{aa}/D_{ql} which is also shown in Fig. 5. The relative behavior of D_{aa} and D_{cc} give an insight to correlations

between the motion in the two canonical planes in the parameter regime of small c and large a: the short time correlations in one plane affect even the zero coupling limit of the diffusion in the other plane.

§8. **Conclusions**

We have derived a formal series for the momentum diffusion tensor of a symplectic map with the second difference (or Lagrangian) form (1), in terms of the characteristic functions. The lowest order terms in this series give the quasilinear results for the diffusion tensor. Higher order terms provide oscillations about $\mathbf{D}_{ql}$ as a function of the map parameters, showing that the oscillatory behavior of **D** is due to short time correlations. The force correlation tensor, **C**(t) can be obtained analytically for $t \leq 3$ using the characteristic function method. For larger t, the expressions for the correlations involve infinite sums. The dominant terms in these sums for large parameters, the principal terms, have an easily calculable form, and an approximate summation of the diffusion series retaining only these terms gives a good approximation to **D** as long as the correlation function decays rapidly.

We show that stable accelerator modes enhance the diffusion. For the area preserving case these cause the divergence of D; however, for higher dimensions this appears to be no longer true when the ensemble of orbits does not include the accelerating invariant tori. The long time behavior of the correlation function for a stochastic orbit is still under investigation.

The thick-layer model of diffusion is appropriate when a is large and c is small. In the limit $c \to 0$, the motion in the two canonical planes is separable, and when b=0 the (x^2,y^2) motion is integrable. The transport in this limit is described by the quantity $\lim_{c \to 0} D_{cc}/D_{ql}$. While this function approaches one as a increases, it oscillates with a due to short time correlations in the (x^1,y^1) plane.

Acknowledgements

We would like to thank the dynamical systems laboratory at the Mathematics Institute of the University of Warwick for support and the congenial surroundings in which this work was initiated. The work was partially supported by NATO grant RG 85/0461 for international collaboration, and US DOE grant DE-FG05-80ET-53088.

References

1 B.V. Chirikov, "A Universal Instability of Many-Dimensional Oscillator Systems", Phys. Reports, **52** (1979) 265.

2 A.B. Rechester, and R.B. White, "Calculation of Turbulent Diffusion for the Chirikov-Taylor Model", Phys. Rev. Lett. **44** (1980)1586; A.B. Rechester, M.N. Rosenbluth and R.B. White, "Fourier-space Paths Applied to the Calculation of Diffusion for the Chirikov-Taylor Model", Phys. Rev. **23A** (1981) 2664.

3 J. R. Cary, J. D. Meiss, and A. Bhattacharjee, "Statistical Characterization of Periodic Measure-Preserving Mappings," Physical Review **A23** (1981)2744-2746.

4 J. D. Meiss, J. R. Cary, C. Grebogi, J. D. Crawford, A. N. Kaufman and H.D.I. Abarbanel, "Correlations of Periodic Area-Preserving Maps," Physica **6D** (1983) 375-384.

5 J. R. Cary and J.D. Meiss, "Rigorously Diffusive Deterministic Map," Phys. Rev **24A** (1981)2664.

6 T.M. Antonsen and E. Ott , Univ. of Maryland Report, No. PL81-018 (1981).

7 N.W. Murray, M.A. Lieberman, and A.J. Lichtenberg, "Corrections to Quasilinear Diffusion in Area Preserving Maps", Phys. Rev. **32A** (1985) 2413; A.J. Lichtenberg, M. A. Lieberman and N.W. Murray, "The Effect of Quasi-Accelerator Modes on Diffusion", Physica **28D**,371-381 (1987).

8 T. Hatori, T. Kamimura, Y.H. Ichikawa, "Turbulent Diffusion for the Radial Twist Map", Physica **14D** (1985) 193.

9 C.F.F. Karney, "Long Time Correlations in the Stochastic Regime", Physica **8D** (1983) 360.

10 J.D. Meiss and E. Ott, "Markov Tree Model of Transport in Area-Preserving Maps", Physica **20D** (1986) 387.

11 C.F.F. Karney, A. B. Rechester, and R.B. White, "Effect of Noise on the Standard Mapping", Physica **4D** (1982) 425.

12 Y. H. Ichikawa, T. Kamimura, and T. Hatori, "Stochastic Diffusion in the Standard Map", Phys. Rev. **A** (1987) .

13 C. Froeschlé, "Numerical Study of a Four-Dimensional Mapping", Astron. & Astrophys. **16** (1972) 172.

[14] C. Froeschlé and J.-P.Scheideker, "Numerical Study of a Four-Dimensional Mapping II", Astron. & Astrophys. **22** (1973) 431.
[15] D. K. Umberger and J. D. Farmer, "Fat Fractals on the Energy Surface", Phys. Rev. Lett. **55** (1985) 661.
[16] J. L. Tennyson, M.A. Lieberman and A.J. Lichtenberg, "Diffusion in Near-Integrable Hamiltonian Systems with Three Degrees of Freedom," in Nonlinear Dynamics and the Beam-Beam Interaction, M.Month and J.C. Herrera (eds.) AIP Conf. Proc. **57** (1979) 272.
[17] N.N. Nekhoroshev, "An Exponential Estimate for the Time of Stability of Nearly-Integrable Hamiltonian Systems," Russ. Math. Surveys **32**:6, 1 (1965).

Figure captions

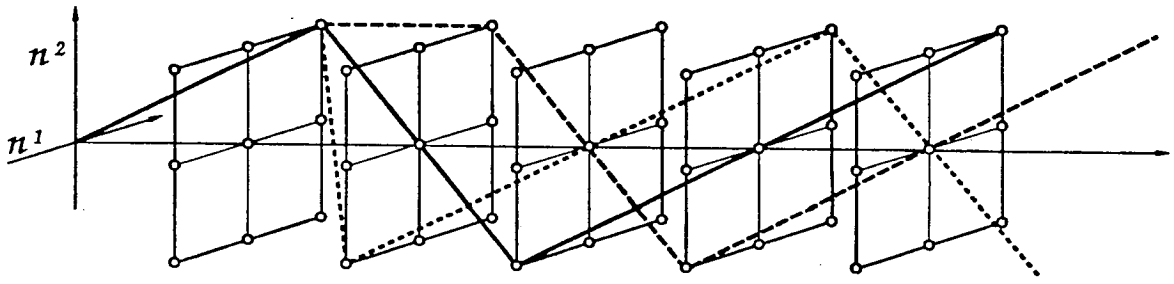
1) Principal terms represented by paths in (**n**,k) space (a) for $\chi_t[(1,1),\pm(1,1)]$ and (b) $\chi_t[(1,0),\pm(1,0)]$. The solid paths are for even t ; the dashed paths for odd t cases.

2) Accelerator mode stability regions. Parameter regions for the existence of the first order accelerator modes with various **j** are shown. An orbit trapped in the **j**=(r,s) mode propagates by $2\pi r$ in y^1 direction and by $2\pi s$ in y^2 direction upon each iteration of the map. The mode **j**=(0,0), the fixed point, does not contribute to acceleration.

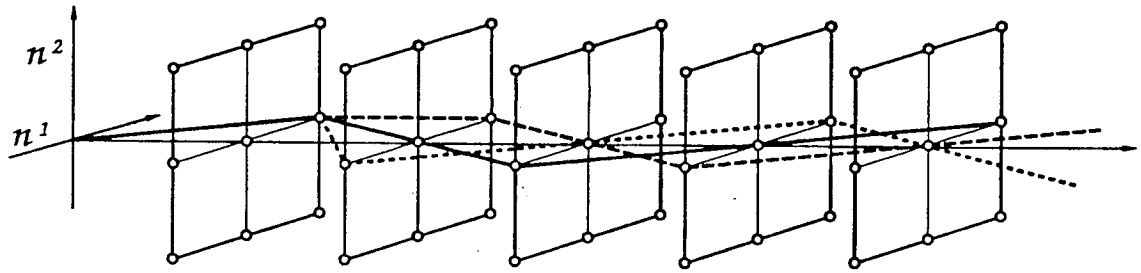
3) Comparison of numerical diffusion coefficients with theoretical results. The dotted curves represent the theory which includes up to the third order correction to the quasilinear result. The principal term results are drawn with the solid curves. The statistical ensemble consists of 10^4 orbits of length T=50 with their initial conditions chosen from an uniform grid over the whole phase space. The error bars give the RMS deviation from the average values. (a) D_{aa} and (b) D_{cc} are normalized to their quasilinear values, respectively. (c) D_{ac} is normalized to D_{cc}. Its fluctuation for $c \geq 2.0$ is of order of D_{ac}; however, considering its contribution to D_{12}, it is still consistent to zero.

4) Enhancement of diffusion coefficient D_{aa} due to the presence of the stable accelerator mode with **j**=(1,0). Its divergent behavior in time is shown; (a) $T=10^2$, (b) $T=10^3$ and (c) $T=10^4$. The statistical ensemble consists of 10^4 orbits whose initial conditions are uniformly distributed on the **y=0** plane. (d) shows that the effect of accelerator modes can be avoided for moderately long times by choosing the ensemble in a connected stochastic region. The ensemble consists of 10^3 suborbits of length $T=10^4$, obtained by breaking a single stochastic orbit of length $T=10^7$ into pieces. The initial condition is chosen near the unstable fixed point; $x^1=0.5$, $y^1=0.01$, $x^2=0.5$ and $y^2=0.02$.

5) Oscillatory dependence of $\lim_{c\to 0} D_{cc}/D_{ql}$ on a. Actually $c= 10^{-5}$, which is small enough for D_{cc} to approach its limiting behavior. Both D_{cc} (squares) and D_{aa} (circles) normalized to their quasilinear values are shown. Only D_{cc} is plotted with its RMS deviations. The same statistics as for Figs. 4 is used except a different length for suborbits ($T=10^3$).



(a)



(b)

Figure 1.

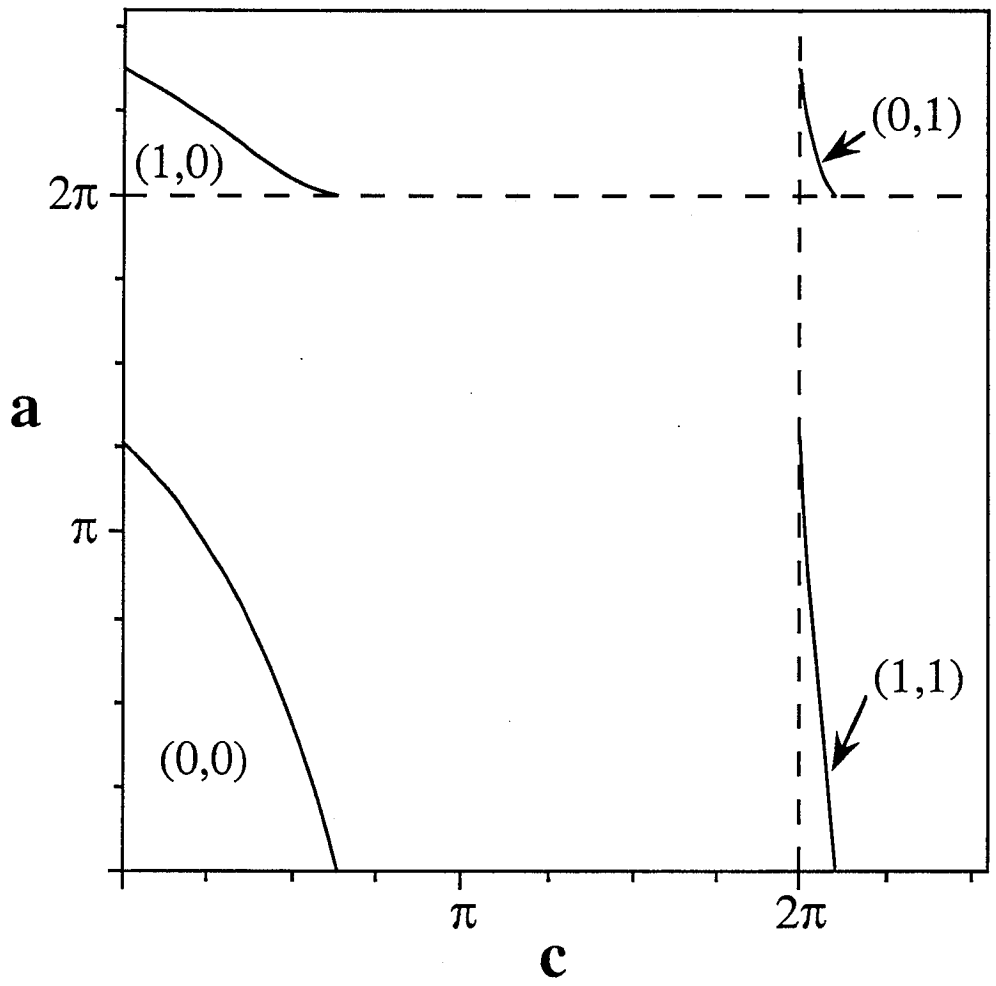


Figure 2.

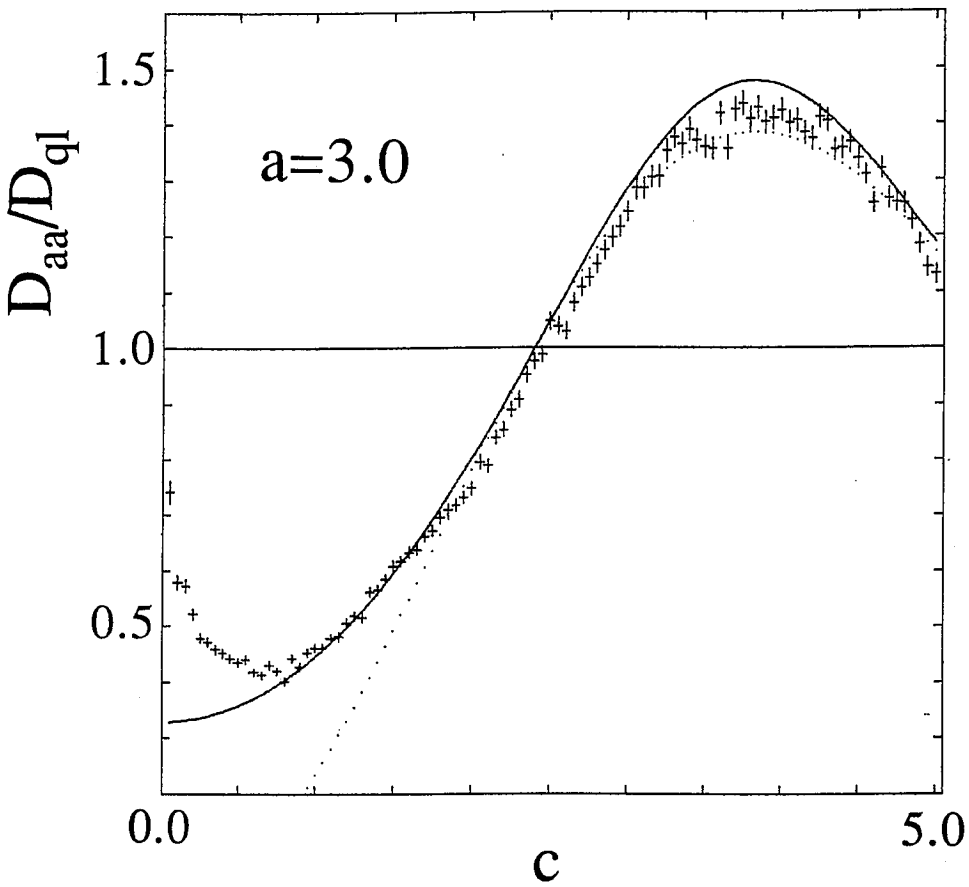


Figure 3(a)

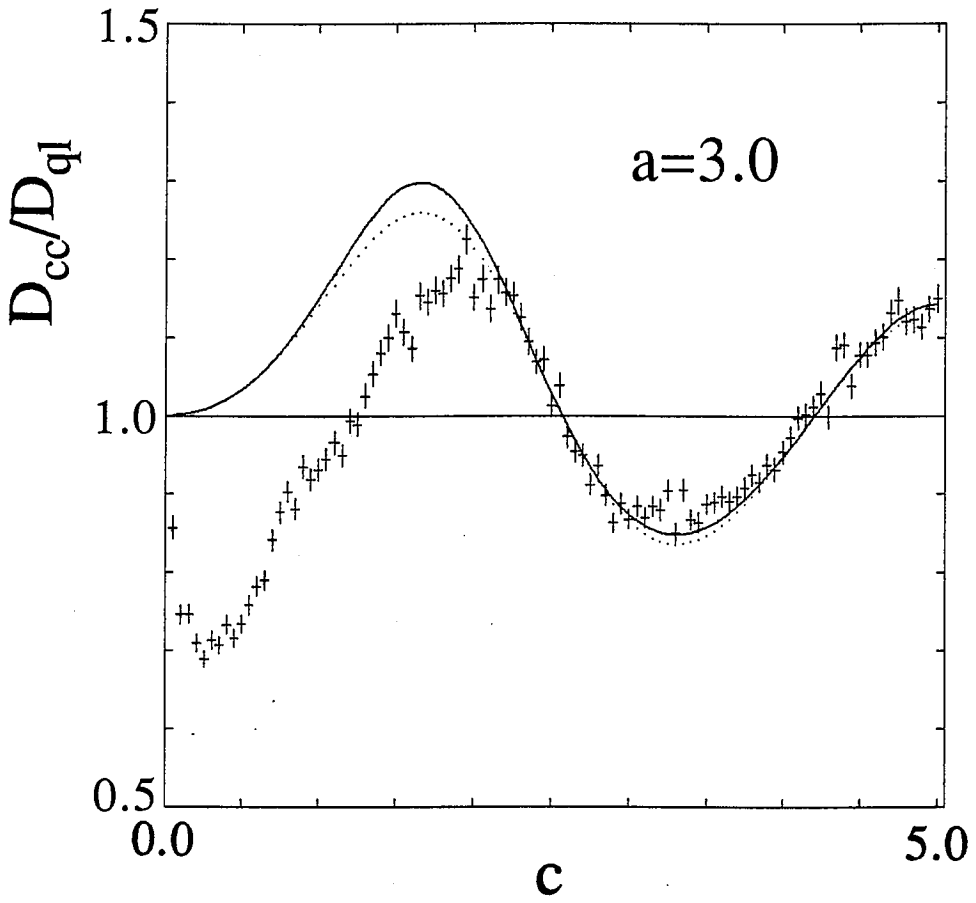


Figure 3(b)

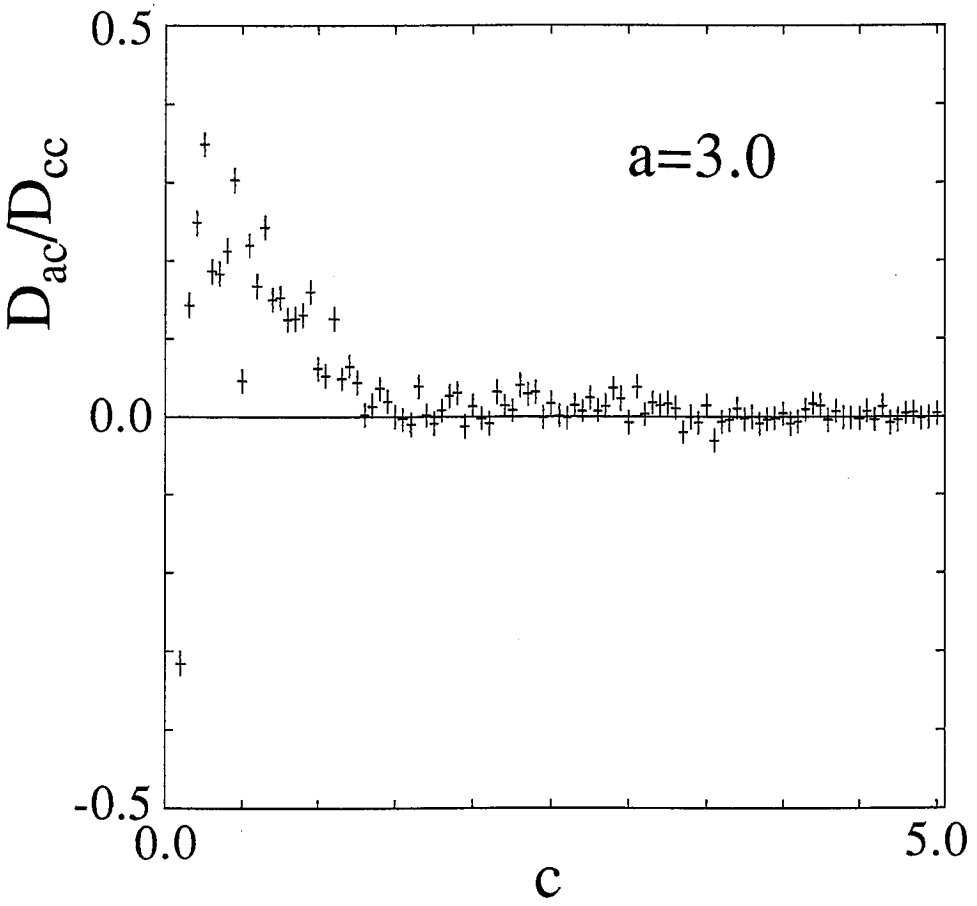


Figure 3(c)

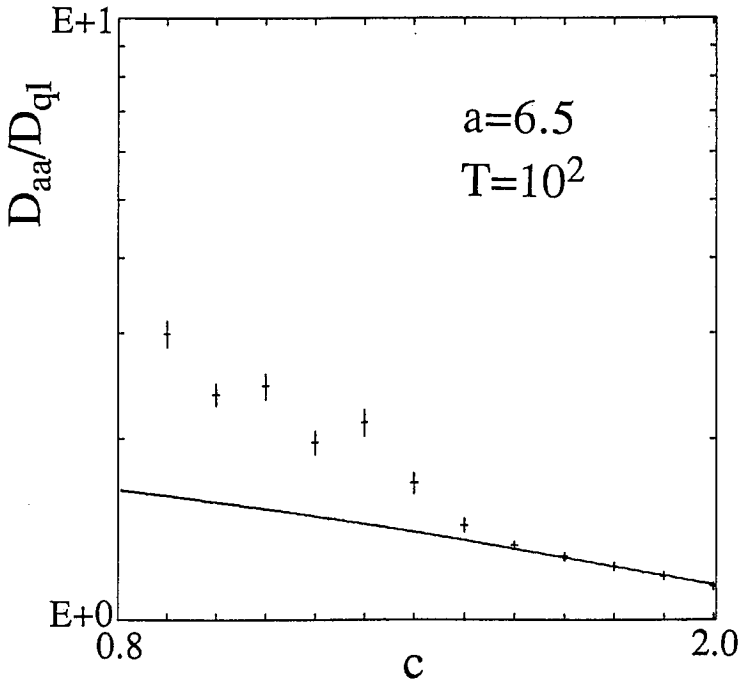


Figure 4(a)

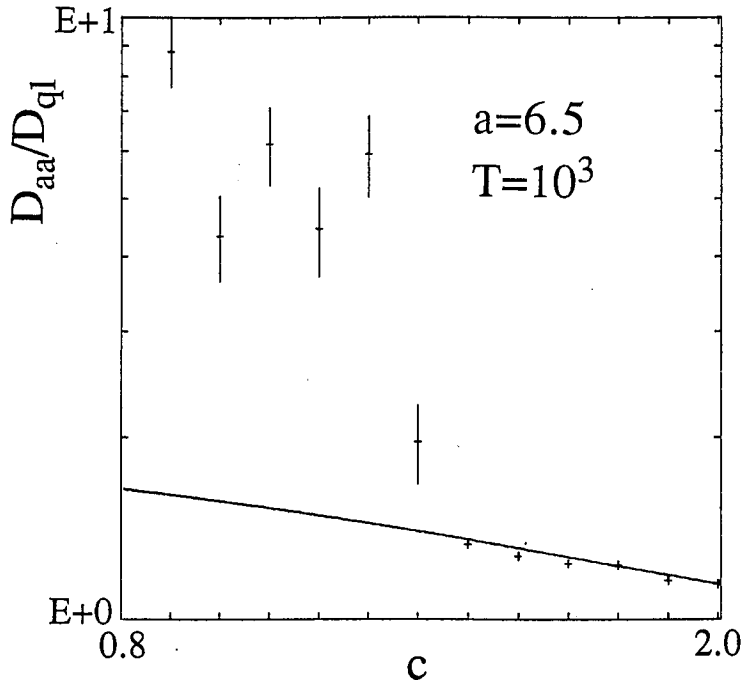


Figure 4(b)

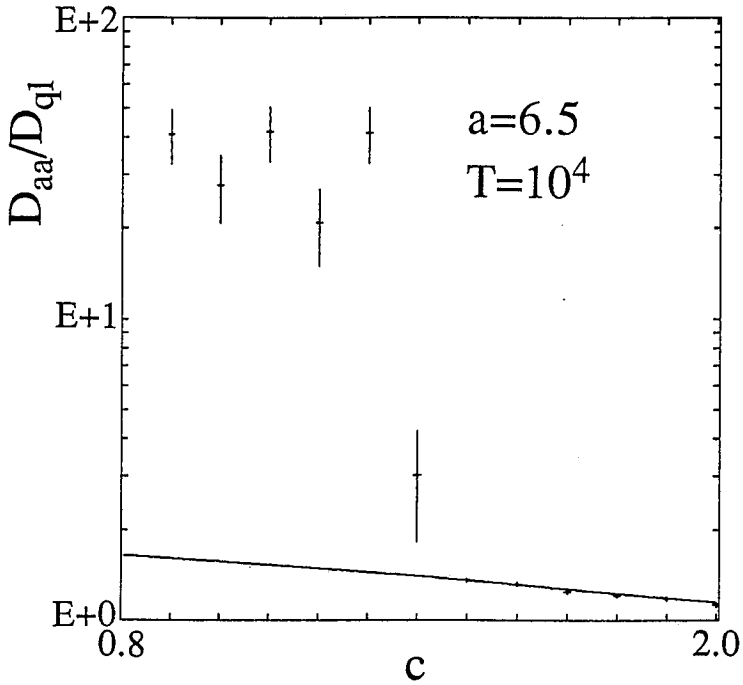


Figure 4(c)

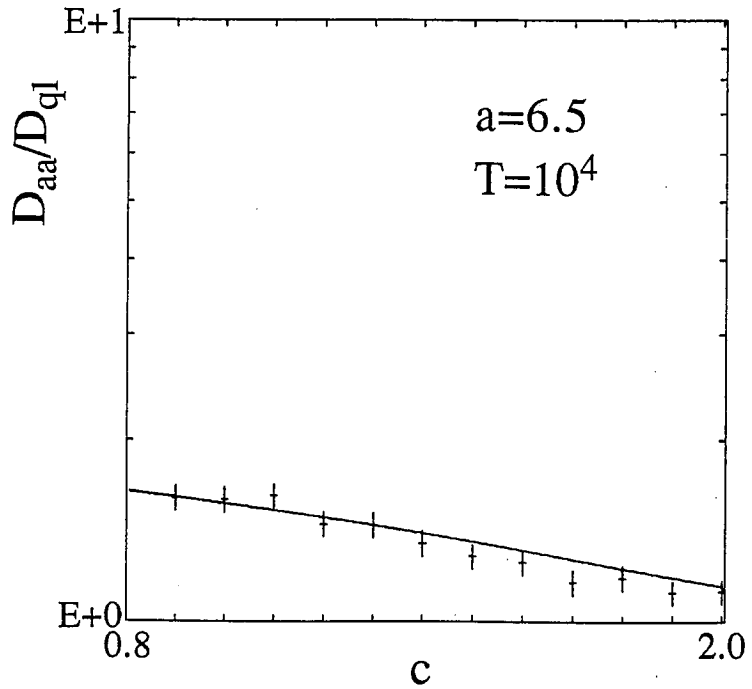


Figure 4(d)

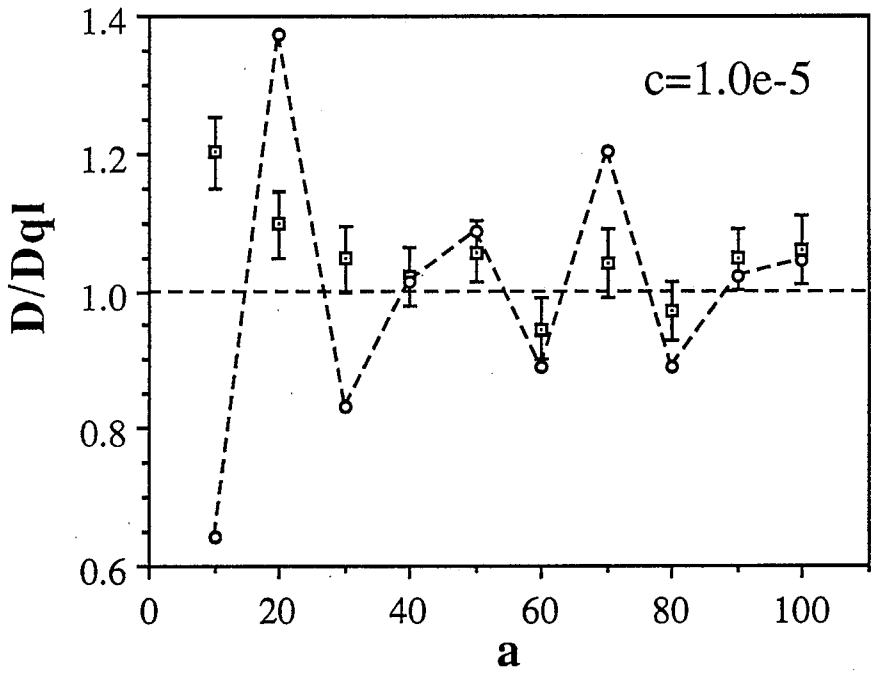


Figure 5