Title
NON-LINEAR DIFFUSION IN HAMILTONIAN SYSTEMS EXHIBITING CHAOTIC MOTION

Author
Abarbanel, Henry D.I.

Publication Date
1980-04-01

Lawrence Berkeley Laboratory
UNIVERSITY OF CALIFORNIA

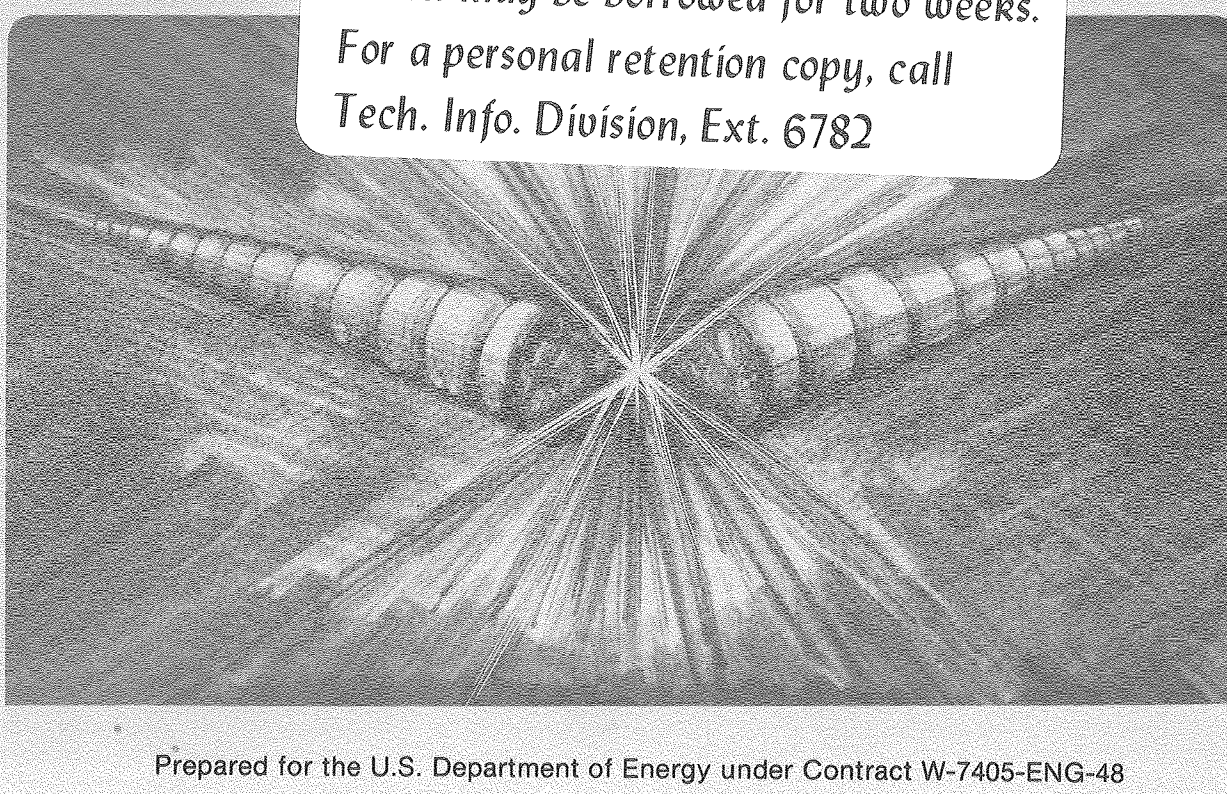
Accelerator & Fusion Research Division

Submitted to Physica D

NON-LINEAR DIFFUSION IN HAMILTONIAN SYSTEMS EXHIBITING CHAOTIC MOTION

Henry D. I. Abarbanel

April 1980



Prepared for the U.S. Department of Energy under Contract W-7405-ENG-48

DISCLAIMER

Non-Linear Diffusion in Hamiltonian Systems Exhibiting Chaotic Motion*

Henry D. I. Abarbanel

Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

Abstract

The exact evolution equation for the angle averaged phase space density in action-angle space is derived from the Liouville equation using projection operator techniques. This equation involves a correlation function of the initial value of the phase space density with the angle dependent part of the Hamiltonian and a correlation function of the angle dependent part of the Hamiltonian and a correlation function of the angle dependent part of the Hamiltonian with itself. Each of these correlation functions develops in time with angle projected dynamics. We show their relation to the correlation functions which develop in time with usual Hamiltonian dynamics. These correlation functions are then studied in the standard model of Chirikov, and we conclude that they behave as $e^{-\sigma t}\cos(\Omega t+\phi))$ in regions of irregular motion. We conjecture that angle averaged correlation functions behave this way in general, and we give an argument based on the mixing property of the Hamiltonian system. Our argument goes beyond the usual mixing, so we regard it as a quasi-mixing hypothesis. Under this hypothesis the equation for the angle averaged phase space density becomes a diffusion equation which incorporates much of the non-linear dynamics of Hamiltonian systems exhibiting chaotic motion.

*Work supported by the U. S. Department of Energy, Office of Fusion Energy, under contract W-7405-ENG-48.

I. Introduction

Mechanical systems of even a few particles are known to show regions of complicated or irregular motion.[1,2] Depending in detail on the strength of the non-integrable perturbation, these regions may be quite small or may occupy the whole of the allowed phase space. A very attractive physical picture of the onset of choatic motion has been developed over the years by Chirikov.[1] He argues that the regularity of motion is most clearly seen in the existence of adiabatic action invariants, and that the overlap of resonances in the non-linear interaction of these invariants causes the system to wander about in phase space from one resonant region to others.[3] Once this resonance overlap sets in, the values of the adiabatic invariants diffuse away from any given region of phase space and we expect the distribution function $F(I,t)$ for the invariants $I = (I_1, I_2, \ldots I_N)$ to satisfy some sort of diffusion equation of the form

$$\frac{\partial}{\partial t} F(\vec{I},t) = \frac{\partial}{\partial I_j} D_{jk}(\vec{I}) \frac{\partial}{\partial I_k} F(\vec{I},t). \tag{1}$$

In the quasi-linear approximation diffusion equation such as this has been considered by Kaufman[4] for particle motion in an axisymmetric plasma, by Rosenbluth, et al.[5] for the diffusion of magnetic field lines and surely by many others.

In this paper we will employ the projection operator method of Zwanzig[6] to investigate the validity and properties of diffusion equations such as (1). Beginning with the Liouville equation for the phase space density function, we project out those variables not of interest--typically the ones varying rapidly. We arrive at an equation much of the form of (1) with an additional term representing a correlation of information about the system at $t = 0$ with information at the later time, t, of interest. On the basis of some model numerical calculations and on the idea of mixing we argue that such correlation functions behave as $e^{-\sigma t}\cos \Omega t$ when the system motion is chaotic. So that to exponential accuracy Equation (1) will hold. Our argument at this stage is a refinement of some observations by Sagdeev and Zaslavskii.[7]

The diffusion tensor $D_{jk}(\vec{I})$ which appears is given by a projected form of dynamics, not by the full dynamics of the problem. We derive a relation between the full diffusion tensor and the projected diffusion tensor for Hamiltonian systems characterized by action-angle variables.

II. Non-linear Diffusion

Since we wish to describe the departure of adiabatic invariants from constancy, we will work with a time independent Hamiltonian system with Hamiltonian given in terms of N actions I_j, $j = 1,\ldots,N$, and N angles θ_j, $j = 1,\ldots,N$

$$H(\vec{I},\vec{\theta}) = H_0(\vec{I}) + H_1(\vec{I},\vec{\theta}). \tag{2}$$

The H_0 term alone would result in I_j = constant, and it is the H_1 term which is responsible for chaotic motion and diffusion of the actions. H_1 is not assumed to be small.

The Liouville equation for the phase space density $f(\vec{I},\vec{\theta},\vec{t})$ takes the two forms

$$\frac{\partial}{\partial t}f + \sum_{j=1}^{N}\left(\frac{\partial}{\partial\theta_j}(\dot{\theta}_j f) + \frac{\partial}{\partial I_j}(\dot{I}_j f)\right) = 0, \tag{3}$$

and

$$\frac{\partial}{\partial t}f = - Lf, \tag{4}$$

where the Liouville operator L is, as usual,

$$L = \sum_{j=1}^{N}\left(\frac{\partial H}{\partial I_j}\frac{\partial}{\partial\theta_j} - \frac{\partial H}{\partial\theta_j}\frac{\partial}{\partial I_j}\right). \tag{5}$$

We want to discuss the evolution of the angle averaged phase space density

$$F(\vec{I},t) \equiv \int \frac{d^N\theta}{(2\pi)^N} f(\vec{I},\vec{\theta},t) \tag{6}$$

$$\equiv Pf \tag{7}$$

which defines the projection operator P. Since P is time independent, from the Liouville equation we have

$$\frac{\partial}{\partial t}F(\vec{I},t) = - PLF - PLG \tag{8}$$

where

$$G(\vec{I},\vec{\theta},t) = (1-P)f(\vec{I},\vec{\theta},t) \tag{9}$$

The function G is made of all the terms of $f(\vec{I},\vec{\theta},t)$ with non-trivial angle dependence. Similarly by multiplication of (4) by (1-P) we arrive at

$$\frac{\partial}{\partial t}G = - (1-P)LF - (1-P)LG \tag{10}$$

Since F depends on the $\vec{I}$'s alone, (repeated indices are summed over now)

$$LF = \frac{\partial H_1}{\partial \theta_a} \frac{\partial}{\partial I_a} F(\vec{I},t), \tag{11}$$

and

$$PLF = -\left(\int \frac{d^N\theta}{(2\pi)^N} \frac{\partial H_1}{\partial \theta_a}\right) \frac{\partial}{\partial I_a} F = 0, \tag{12}$$

since $H_1(\vec{I},\vec{\theta})$ is periodic in $\vec{\theta}$.

We have now

$$\frac{\partial}{\partial t} F(\vec{I},t) = -\ PLG(\vec{I},\vec{\theta},t) \tag{13}$$

and

$$\frac{\partial G}{\partial t} = -\ LF - (1-P)LG \tag{14}$$

This last equation has the solution

$$G(t) = e^{-(1-P)Lt} G(0) - \int_0^t dz\ e^{-(1-P)Lz}\ LF(t-z) \tag{15}$$

which leads to

$$\frac{\partial}{\partial t} F(\vec{I},t) = -\ PLe^{-(1-P)Lt} G(\vec{I},\vec{\theta}\ ,0) + \int_0^t d\ \ PL\ e^{-(1-P)L\tau}\ LF(\vec{I},t-\tau). \tag{16}$$

Consider the first term of this equation. Using the properties of L expressed in (3) we find

$$- PL\; e^{-(1-P)Lt} G(\vec{I},\vec{\theta},t=0) = \frac{\partial}{\partial I_j} P \frac{\partial H_1}{\partial \theta_j} e^{-(1-P)Lt} G(\vec{I},\vec{\theta},t=0) \quad . \tag{17}$$

Next we look at the second term. Using (3) and (5) we cast this into the form

$$\frac{\partial}{\partial I_j} \int_0^t D_{jk}(\vec{I},t) \frac{\partial}{\partial I_k} F(\vec{I},t-\tau)\, d\tau \tag{18}$$

with

$$D_{jk}(\vec{I},\tau) = P\left(\frac{\partial H_1}{\partial \theta_j} e^{-(1-P)L\tau} \frac{\partial H_1}{\partial \theta_k}\right) . \tag{19}$$

Note that $D_{jk}(\vec{I},\tau)$ is an operator of multiplication by I and derivatives with respect to I. We have separated out D_{jk} in this fashion since the projection operator has no effect on the function of $\vec{I}$ alone upon which $D_{jk}(\vec{I},\tau)$ acts.

So altogether we arrive at

$$\begin{aligned} \frac{\partial}{\partial t} F(\vec{I},t) &= \frac{\partial}{\partial I_j} P\left(\frac{\partial H_1}{\partial \theta_j} e^{-(1-P)Lt}\, G(\vec{I},\vec{\theta},t=0)\right) \\ &\quad + \frac{\partial}{\partial I_j} \int_0^t D_{jk}(\vec{I},\tau) \frac{\partial}{\partial I_k} F(\vec{I},t-\tau)\, d\tau . \end{aligned} \tag{20}$$

The diffusion tensor D_{jk} arises from the projected Liouville operator (1-P)L and so its dynamics are a bit unfamiliar. We shall call D the projected diffusion tensor. The dynamics of (1-P)L can be seen by its role in the evolution of initial action-angle variables I_j, θ_j into $\tilde{I}_j(\vec{I},\vec{\theta},t)$ and $\tilde{\theta}_j(\vec{I},\vec{\theta},t)$ via

$$\frac{d\tilde{I}_j}{dt} = (1-P)LI\Big|_{\substack{I=\tilde{I}\\ \theta=\tilde{\theta}}} = -\frac{\partial H_1}{\partial \theta_j}(\vec{I},\vec{\theta},t)\Big|_{\substack{I=\tilde{I}\\ \theta=\tilde{\theta}}} \tag{21}$$

and

$$\frac{d}{dt}\tilde{\theta}_j = (1-P)L\theta\Big|_{\substack{I=\tilde{I}\\ \theta=\tilde{\theta}}} = \frac{\partial H_1}{\partial I_j}(\vec{I},\vec{\theta},t)\Big|_{\substack{I=\tilde{I}\\ \theta=\tilde{\theta}}} \tag{22}$$

The motion of $\tilde{I}$ and $\tilde{\theta}$ differs from the usual motion due to L in that the term due to H_0 - the regular motion - is absent from the $\tilde{\theta}$ time development. So the projected motion is missing the regular, integrable piece most familiar to us and may be expected to take an even more irregular form than the usual motion. the projected motion is a kind of "generalized interaction representation" since the influence of H_0 is absent.

Under the influence of the usual dynamics we would encounter the direct diffusion tensor

$$\Delta_{jk}(\vec{I},t) = P\left(\frac{\partial H_1}{\partial \theta_j}\, e^{-Lt}\, \frac{\partial H_1}{\partial \theta_k}\right) \tag{23}$$

This is also a correlation function between $\partial H_1/\partial\theta_k$ and $\partial H_1/\partial\theta_j$ separated by a time t. If the system is chaotic or irregular, we would expect information at $t = 0$ to be rapidly forgotten by the system on a time scale similar to the time it takes for nearby orbits in phase space to separate exponentially. The direct diffusion tensor may fall very rapidly in t in chaotic motion; its behavior is easily amenable to numerical investigation.

Now we will find the relation between D_{jk} and Δ_{jk}. Introduce the Laplace transforms

$$D_{jk}(\vec{I},s) = \int_0^\infty dt\ e^{-st} D_{jk}(\vec{I},t), \tag{24}$$

and

$$\Delta_{jk}(\vec{I},s) = \int_0^\infty dt\ e^{-st} \Delta_{jk}(\vec{I},t). \tag{25}$$

We have

$$D_{jk}(\vec{I},s) = P\left(\frac{\partial H_1}{\partial\theta_j}\ \frac{1}{s+(1-P)L}\ \frac{\partial H_1}{\partial\theta_k}\right), \tag{26}$$

and

$$\Delta_{jk}(\vec{I},s) = P\left(\frac{\partial H_1}{\partial\theta_j}\ \frac{1}{s+L}\ \frac{\partial H_1}{\partial\theta_k}\right). \tag{27}$$

Using the operator identity

$$\frac{1}{s+(1-P)L} = \frac{1}{s+L} + \frac{1}{s+L}PL\frac{1}{s+(1-P)L} \tag{28}$$

we find

$$D_{jk}(\vec{I},s) = \Delta_{jk}(\vec{I},s) + P\left(\frac{\partial H_1}{\partial \theta_j}\frac{1}{s+L}\,PL\,\frac{1}{s+(1-P)L}\frac{\partial H_1}{\partial \theta_k}\right). \tag{29}$$

In the second term we use

$$PLW(I,\theta) = P\left(\frac{\partial}{\partial \theta_j}(\dot{\theta}_j W) + \frac{\partial}{\partial I_j}(\dot{I}_j W)\right) \tag{30}$$

$$= -P\frac{\partial}{\partial I_j}\left(\frac{\partial H_1}{\partial \theta_j}W\right) \tag{31}$$

for any W periodic in θ_j's, and find

$$D_{jk}(\vec{I},s) = \Delta_{jk}(\vec{I},s) - P\left(\frac{\partial H_1}{\partial \theta_j}\frac{1}{s+L}\frac{\partial}{\partial I_n}P\frac{\partial H_1}{\partial \theta_n}\frac{1}{s+(1-P)L}\frac{\partial H}{\partial \theta_k}\right) \tag{32}$$

or

$$D_{jk}(\vec{I},s) = \Delta_{jk}(\vec{I},s) - P\left(\frac{\partial H_1}{\partial \theta_j}\frac{1}{s+L}\right)D_{nk}(I,s) \tag{33}$$

Note that

$$\frac{1}{s+L} = \frac{1}{s} - \frac{1}{s(s+L)}L \tag{34}$$

and $P\left(\frac{\partial H_1}{\partial \theta_j}\frac{1}{s}\right) = 0$. Since

$$LD(\vec{I},s) = -\frac{\partial H_1}{\partial \theta_a}\frac{\partial D}{\partial I_a}(\vec{I},s), \tag{35}$$

we arrive at the <u>operator</u> equation

$$D_{jk}(\vec{I},s) = \Delta_{jk}(\vec{I},s) - \frac{1}{s}\Delta_{jn}(\vec{I},s)\frac{\partial}{\partial I_n \partial I_\ell} D_{\ell k}(\vec{I},s) \tag{36}$$

or

$$\left(\frac{\partial^2}{\partial I_a \partial I_c} + s(\Delta^{-1})_{ac}\right) D_{cb}(\vec{I},s) = s\delta_{ab} . \tag{37}$$

This result relates the direct diffusion tensor, about which we might chance some guesses or make some approximations; e.g. the quasi-linear approximation;[4] to the projected diffusion tensor which we need for the evolution of the projected phase space density $F(\vec{I},t)$ in Equation (20).

From (37) we infer that for $s \to \infty$, $D_{ab} \to \Delta_{ab}$. More directly we find from (26) and (27) that

$$D_{ab}(\vec{I},s) - \Delta_{ab}(\vec{I},s) \underset{s\to\infty}{\sim} \frac{1}{s^3} P\left(\frac{\partial H_1}{\partial\theta_a}\frac{\partial H_1}{\partial\theta_k}\right)\frac{\partial^2}{\partial I_k \partial I_n} P\left(\frac{\partial H_1}{\partial\theta_n}\frac{\partial H_1}{\partial\theta_b}\right) + O(1/s^4) \quad (38)$$

This result is understood physically by remembering that the Laplace transform variable $s \to \infty$ corresponds to $t \to 0$ in real time. For very short times projected and direct dynamics should correspond. Indeed, (38) shows they differ by $O(t^2)$.

This suggests we introduce yet another diffusion tensor

$$\eta_{jk}(\vec{I},s) \equiv D_{jk}(\vec{I},s) - \Delta_{jk}(\vec{I},s) \quad (39)$$

which we call the anomalous diffusion tensor. It satisfies

$$\left[\frac{\partial^2}{\partial I_a \partial I_c} + s\Delta^{-1}(\vec{I},s)_{ac}\right]\eta_{cb}(\vec{I},s) = -\frac{\partial^2}{\partial I_a \partial I_c}\Delta_{cb}(\vec{I},s). \quad (40)$$

To make any progress in learning about D or η we need some ideas about Δ; we turn now to that.

III. The Direct Diffusion Tensor

We begin our study of the direct diffusion tensor by considering the one dimension system

$$H(I,\theta) = H_0(I) + V \sin\theta \qquad (41)$$

Such systems are always integrable and, thus, regular so they will show no diffusion of actions. However, it does provide an example of what not to expect for chaotic systems. It is straightforward to integrate the equations of motion for $\theta(t)$ from (41)

$$\theta(t) = \theta + \omega(I)t, \ \omega(I) = \frac{\partial H_0}{\partial I} \qquad (42)$$

and then to evaluate

$$\Delta(I,t) = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{\partial H}{\partial\theta} e^{-Lt} \frac{\partial H}{\partial\theta} \qquad (43)$$

$$= \int_0^{2\pi} V^2 \cos\theta \cos(\theta-\omega(I)t), \qquad (44)$$

$$= \frac{V^2}{2} \cos(\omega(I)t). \qquad (45)$$

So, in the case of regular motion we can anticipate that the correlation function (or direct diffusion tensor)

$$P\left(\frac{\partial H_1}{\partial \theta_a} e^{-Lt} \frac{\partial H_1}{\partial \theta_b}\right)$$

will exhibit pure oscillatory behavior. This is a reflection of the regularity of the systems motion - namely the motion in quasi-periodic and "smooth." Orbits which are nearby at some time separate slowly from each other and do not lose information rapidly about having been close.

Next we turn to the other extreme; namely, we consider a system which is mixing.[8] The definition of such systems is in terms of correlation functions of two bounded phase functions $f(\Gamma)$ and $g(\Gamma)$ defined on phase space Γ. The boundedness of such functions is a mathematical requirement of no distinct physical importance, but since we will soon consider unbounded functions it is useful to be clear that proofs of mixing entail bounded, or more precisely square integrable functions on phase space. The correlation function of f and g is given by

$$C_{fg}(t) \equiv \langle f^*(\Gamma)e^{-Lt}g(\Gamma)\rangle - \langle f^*(\Gamma)\rangle\langle g(\Gamma)\rangle \tag{46}$$

where Γ is 2N dimensional phase space and $\langle\rangle$ is defined as an integral over the energy shell:

$$\langle f(\Gamma)\rangle = \int d\Gamma\delta(E-H(\Gamma))f(\Gamma)/\int d\Gamma\delta(E-H(\Gamma)) \quad . \tag{47}$$

The system is said to be mixing, if for all bounded f and g,

$$C_{fg}(t) \xrightarrow[|t| \to \infty]{} 0 \tag{48}$$

{The need for choosing square integrable phase space functions is seen in (46), since we must look at $C_{ff}(t)$ in determining whether a system is mixing. $C_{ff}(0)$ involves $\langle |f(\Gamma)|^2 \rangle$ and this must be finite.} Mixing systems are certainly irregular in the usual sense of the idea.

To get an idea how fast such correlation functions decay to zero for $|t| \to \infty$ we argue in the following heuristic manner. Introduce, in a formal fashion, the eigenfunctions $\psi_\lambda(\Gamma)$ and eigenvalues λ of the Liouville operator

$$L\psi_\lambda(\Gamma) = \lambda\psi_\lambda(\Gamma) \tag{49}$$

the λ are pure imaginary since L is anti-hermitean. Choose f and g so $\langle f \rangle = \langle g \rangle = 0$, and expand $f(\Gamma)$ and $g(\Gamma)$ as

$$f(\Gamma) = \sum_\lambda f_\lambda \ \psi_\lambda(\Gamma), \tag{50}$$

and

$$g(\Gamma) = \sum_\lambda g_\lambda \psi_\lambda(\Gamma). \tag{51}$$

$C_{fg}(t)$ is now given by

$$C_{fg}(t) = \sum_{\lambda} e^{-\lambda t} f^*_{\lambda} g_{\lambda} \tag{52}$$

This sum, or integral along the imaginary axis, is governed for large t by the singularities in the complex λ-plane of the quantity $f^*_{\lambda} g_{\lambda}$. The singularity whith the smallest real part will give the dominant behavior for large t. For systems showing regular motion $C_{fg}(t)$ will not decay for large t but show multiple periodicity. So for such regular systems $f^*_{\lambda} g_{\lambda}$ must have all of its singularity structure on the imaginary axis. When the motion becomes irregular, these singularities will move off into the λ-plane and lie at some position $\lambda = \sigma \pm i\Omega$. Suppose the dominant singularity is a pole at this point, then

$$C_{fg}(t) \underset{t \to +\infty}{\sim} e^{-\sigma t} \cos(\Omega t + \phi); \; \sigma > 0, \; t > 0 \tag{53}$$

where ϕ is a phase reflecting the residue at the pole. For $t \to -\infty$ the λ-plane must contain a "conjugate" pole at $\lambda = -\sigma' \pm i\Omega'$ so $C_{fg} \to 0$ in that limit too. Mixing systems thus may be expected to exhibit exponential decay for correlation functions. This seems rather natural actually, since such systems also have the property that nearby points in phase space diverge exponentilly from each other at later times.[9] It is tempting to guess that the rate of divergence is $e^{\sigma t}$ with the same σ as in (53). In some very simple examples that is indeed the case.[1,10]

We assumed a pole in the λ-plane in order to arrive at (53). A branch point at $\lambda = \sigma + i\Omega$ would yield additional powers of time multiplied by (53).

Our actual problem in the case of the correlation function $\Delta_{ab}(\vec{I},t)$ involves an averaging process over part of phase space only, namely the angle variables. Let us look at $C_{fg}(t)$ above with $\langle f\rangle = \langle g\rangle = 0$ and choose for f the <u>unbounded</u> function

$$F(\vec{I},\vec{\theta}) = \rho(\vec{J},\vec{\theta})\delta^N(\vec{I}-\vec{J}) \tag{54}$$

while g is left unspecified except for $\langle g\rangle = 0$. Now C_{fg} is

$$C_{fg}(t) = \int \frac{d^N\theta}{(2\pi)^N}\,\rho(\vec{J},\vec{\theta})e^{-Lt}g(\vec{J},\vec{\theta})$$

where

$$L = \sum_{k=1}^{N}\left(\frac{\partial H(\vec{J},\vec{\theta})}{\partial J_k}\frac{\partial}{\partial \theta_k} - \frac{\partial H(\vec{J},\vec{\theta})}{\partial \theta_k}\frac{\partial}{\partial J_k}\right) .$$

Clearly $C_{fg}(t)$ now depends on $\vec{J}$ as well as time. The use of the unbounded function f takes us out of the usual realm of mixing and leads us to the <u>quasi-mixing hypothesis</u>: angle averaged functions like

$$\Delta_{ab}(\vec{I},t) = \int \frac{d^N\theta}{(2\pi)^N} \frac{\partial H(\vec{I},\vec{\theta})}{\partial\theta_a} e^{-Lt} \frac{\partial H(\vec{I},\vec{\theta})}{\partial\theta_b}$$

behave as

$$\Delta_{ab}(\vec{I},t) \underset{t\to\infty}{\sim} e^{-\sigma(\vec{I})t}\cos(\Omega(\vec{I})t + \phi(\vec{I})),\ \sigma(\vec{I}) > 0, \qquad (55)$$

when the system motion is irregular. This is a stronger requirement than mixing alone and requires mixing to be local in $\vec{I}$ space and to result from the averaging over only some of the chaotically varying phase space co-ordinates.

On the basis of this hypothesis I propose to approximate $\Delta_{ab}(\vec{I},t)$ for regions of irregular motion by

$$\Delta_{ab}(\vec{I},t) = \Delta_{ab}(\vec{I},t=0)\ e^{-\sigma(\vec{I})t}\cos(\nu(\vec{I},t)) \qquad (56)$$

for all times $t \geq 0$. Here $\nu(\vec{I},t)$ must satisfy

$$\nu(\vec{I},t) \xrightarrow[t\to\infty]{} \Omega(\vec{I})t + \phi(\vec{I}), \qquad (57)$$

and

$$\nu(\vec{I},0) = 0 \qquad (58)$$

Many choices for $\nu(\vec{I},t)$ clearly satisfy these requirements, but an especially simple one is $\nu(\vec{I},t) = \Omega(\vec{I})t + \phi(\vec{I})(1-e^{-\eta(\vec{I})t})$. Another choice for parametrizing $\Delta_{ab}(\vec{I},t)$, indeed one we will use below is

$$\Delta_{ab}(\vec{I},t) = \Delta_{ab}(\vec{I},0)e^{-\sigma(\vec{I})t}\cos(\Omega(\vec{I})t + \phi(\vec{I}))/\cos\phi(\vec{I}), \qquad (59)$$

which clearly supposes $\phi \neq (2n+1)\pi/2$ $n = 0, \pm 1, \ldots$.

None of these specific parametrizations has a fundamental significance; each expresses the asymptotic behavior $e^{-st}\cos(\Omega t = \phi)$ suggested before.

All of these statements refer to $t \geq 0$. For $t \leq 0$, we expect the correlation functions to behave as $e^{\sigma' t}\cos(\Omega' t + \phi')$, $\sigma' \geq 0$. If the correlation functions involved integrals over all of phase space rather than just over angular variables, we could demonstrate a connection between the $t \geq 0$ and $t \leq 0$ behavior of the correlation function using the anti-hermitiean property of L. Here we do not integrate over action variables and have found no general relation between positive t and negative t behavior of $\Delta_{ab}(\vec{I},t)$.

Evidence for this behavior of the direct correlation function comes from several sources. First there is the work by Mo[11] on the onset of stochasticity in Hamiltonian systems. Mo examined the pole positions of the Laplace transform of correlation functions for three Hamiltonians including the well-known Henon-Heiles example. At the parameter values associated with the onset of chaotic behavior in the surface of section plot, the pole positions moved into the complex plane in such a way as to produce behavior like (56).

Secondly, we present some calculations done on correlation functions in the standard mapping of Chirikov.[1] This is an area preserving mapping which takes variables I_n, θ_n each lying between 0 and 1 into I_{n+1}, θ_{n+1} in the same interval:

$$\left.\begin{aligned} I_{n+1} &= I_n + \frac{k}{2\pi}\sin 2\pi\theta_n \qquad (60) \\ \theta_{n+1} &= \theta_n + I_{n+1} \qquad (61) \end{aligned}\right\} \pmod 1$$

The mapping can be derived from a physical system which is a pendulum subject to delta function kicks at unit intervals. I_n and θ_n are the values of momentum and angle after the n^{th} kick.

We have evaluated the correlation function

$$C(N,I_0) = 2\int_0^1 d\theta_0 \sin 2\pi\theta_N(I_0,\theta_0)\sin 2\pi\theta_0;\ C(0,I_0)=1, \qquad (62)$$

for various values of I_0 for $0 \leq N \leq 10$ for values of the mapping parameter k which are known to give significant regions of stochastic behavior. Irregular motion in the standard mapping sets in for $k \approx 1$[1,3,12] and exhibits itself by the filling of large portions of the I,θ plane by a single trajectory.

As an example we show in Figure 1 the trajectory I_n,θ_n for $n = 0,1,2,\ldots,10^4$ for the standard mapping with $k = 3.5$ and $I_0 = 0.4357$ and $\theta_0 = 0.078695$. These initial values were essentially arbitrary except that previous experience with the mapping indicated that they lay outside the large islands seen in Figure 1. Excluding the large islands we would expect 28.25 points/square in Figure 1 if the 10^4 points are uniformly distributed. A survey of the squares reveals a mean number of 26.1 points with $\langle (N-\langle N\rangle)^2\rangle/\langle N\rangle^2 \approx 0.038$ which is very close to $\langle N\rangle^{-1} = 0.027$. This seems to indicate that a uniform distribution with normal fluctuations is what we are seeing in Figure 2.

We calculated $C(N,I_0)$ for $I_0 = 0.4357$ and $k = 3.5$ and this is shown in Figure 2 for $0 \le N \le 10$. All θ_0 integrated over here lie in the apparently chaotic region. The points can be fit rather well by

$$C(N,I_0) = e^{-N\sigma}\cos(\Omega N + \phi)/\cos\phi \tag{63}$$

with $\sigma = 0.25$, $\Omega \approx 1.578$, and $\phi = -0.44$. The comparison between (63) and $C(N,I_0)$ is given in Figure 3. In Figure 4 the same correlation function at $k = 3.5$ is shown for $I_0 = 0.4$, 0.5, and 0.55 to give an indication of the sensitivity of I_0. The variation of $C(N,I_0)$ with respect to k for fixed I_0 is shown in Figure 5 for $I_0 = 0.5$ and $k =$ 3.5, 4.0, 4.5, and 5.0. In Figure 6 we show $C(N,I_0)$ for two closely spaced values of k, $k = 5.00$ and $k = 5.03$, for $I_0 = 0.5$.

There is a sublty in making these calculations. Since the function $\theta_N(I_0,\theta_0)$ becomes very rapidly varying, many integration points are needed in the calculation of $C(N,I_0)$. For the range of N shown and the value of k chosen, we found that for $\ge$ 2000 integration points we reproduced the same $C(N,I_0)$ while for much fewer, the computed values differed as they "settled" into the answer shown. Clearly one must get to the stage where there are at least several integration points between each wiggle of the integrand. As k increases the integrand becomes more and more wiggly, we expect it to be increasingly difficult to numerically evaluate $C(N,I_0)$. Indeed at $k = 5$, some 7000 integration points were needed to get a stable answer over the same range of N. The behavior observed by Smith[13] that the r.m.s. value of $C(N,I_0)$ decreases as $(\text{integration points})^{-1/2}$ is consistent with this view of the way errors will disappear as we settle in on the reproducible answer since we are slowly getting between the rapid variations.

A further complication is connected with the growth of errors in any calculation with the standard mapping for sizeable k and a large number of steps. Errors grow according to the tangent mapping

$$\Delta I_{n+1} = \Delta I_n + (k \cos 2\pi\theta_n)\Delta\theta_n \tag{64}$$

$$\Delta\theta_{n+1} = \Delta\theta_n + \Delta I_{n+1} \tag{65}$$

and for sizeable k should behave as $\Delta\theta_n \approx k^n \Delta\theta_0$, $\Delta I_n \approx k^n \Delta I_0$. This behavior has been verified by Chirikov[1] and discussed by Greene.[12] For large enough k, the errors grow so large, even for a small number of steps, that any numerical calculation of $\theta_N(I_0,\theta_0)$ $C(N,I_0)$ must be examined very carefully.

IV. Projected Diffusion Tensor

Now we turn to the equation which determines the projected diffusion tensor $D_{ab}(\vec{I},s)$

$$\left(\frac{\partial^2}{\partial I_a \partial I_b} + s(\Delta^{-1})_{ab}\right) D_{ab}(\vec{I},s) = s\delta_{ac}. \tag{37}$$

Our replacement of the operator form of $\Delta_{ab}(\vec{I},s)$, Equation (23), by the ansatz of Equation (56) which is an ordinary function, not an operator, means that this equation for $D_{ab}(\vec{I},s)$ is no longer an operator equation.

We know that for $s \to \infty$,

$$D_{ab}(\vec{I},s) \to \Delta_{ab}(\vec{I},s) \tag{66}$$

and have discussed that above. (See before Equation (38)). For $s \to 0$, assuming $\Delta^{-1}_{ab}(\vec{I},s)$ is not singular,[14]

$$\frac{\partial^2}{\partial I_a \partial I_b} D_{bc}(\vec{I},s = 0) = 0. \tag{67}$$

So $D_{ab}(\vec{I},s = o)$ is a polynomial in $\vec{I}$ of first order. The general solution to (67) is

$$D_{ab}(\vec{I},s = 0) = A_{ab} + I_a B_b + I_b C_a, \tag{68}$$

with A_{ab}, B_b, and C_a fixed tensors of the indicated rank which are independent of $\vec{I}$.

We can acquire some handle on the behavior of $D_{ab}(\vec{I},s)$ from the properties of $\Delta_{ab}(\vec{I},s)$. Suppose first that, in some sense, $\Delta_{ab}(\vec{I},s)$ is large and $\Delta^{-1}_{ab}D_{ac}$ is small compared to 1. Then the second term on the left hand side of (37) can be dropped with **respect** to $s\delta_{ac}$, and

$$\frac{\partial^2}{\partial I_a \partial I_b} D_{bc}(\vec{I},s) = s\delta_{ac}. \tag{69}$$

The solution to this is

$$D_{ab}(\vec{I},s) = A_{ab} + I_a B_b + I_b C_a + \frac{s}{N+1} I_a I_b \tag{70}$$

where A_{ab}, B_b, and C_a have the same meaning as before, and N is the dimension of $\vec{I}$ space.

Secondly, if $\Delta_{ab}(\vec{I},s)$ is independent of $\vec{I}$, then an acceptable and consistent solution for D is

$$D_{ab}(\vec{I},s) = \Delta_{ab}(\vec{I},s) = \Delta_{ab}(s). \tag{71}$$

Thirdly, if $\Delta_{ab}(\vec{I},s)$ is only weakly dependent on I, then we write (as in Equation (39))

$$D_{ab} = \Delta_{ab} - \eta_{ab}$$

and use (40) to solve for η_{ab}, the anomalous diffusion tensor, in a perturbative fashion.

Finally, suppose $\Delta_{ab}(\vec{I},s)$ becomes small. Since $(\Delta^{-1})_{ab}(\vec{I},s)$ will become large, we have again $D_{ab} = \Delta_{ab}$ except near $s = 0$.

It seems clear that whenever neither D nor Δ have rapid variation in $\vec{I}$ that $D_{ab} \approx \Delta_{ab}$ and the singularities of D in the s plane will be slightly displaced from those of Δ.

We can introduce the eigenfunctions $\psi_a^{(\lambda)}(\vec{I})$ and eigenvalues $\lambda(s)$ of the operator

$$O_{ab} \equiv \frac{\partial^2}{\partial I_a \partial I_b} + s(\Delta^{-1})_{ab} \, , \tag{72}$$

and then write

$$D_{ab}(\vec{I},s) = s\sum_{\lambda} \frac{\psi_a^{(\lambda)}(I)\int \bar{\psi}_b^{(\lambda)}(\vec{I}')d^N I'}{\lambda(s)} \tag{73}$$

From this it is clear that the zeroes of $\lambda(s)$ are what determine the singularity structure in the s-plane for $D_{ab}(\vec{I},s)$. $\lambda(s)$ may be taken from the variational principle

$$\Lambda[\chi,\phi] = \int d^N I\ \chi_a(\vec{I}) O_{ab} \phi_b(\vec{I}) / \int d^N I\ \chi_a(\vec{I}) \phi_a(\vec{I}) \tag{74}$$

Varying $\Lambda[\chi,\phi]$ with respect to χ_a or ϕ_b yields

$$\frac{\partial \Lambda[\chi,\phi]}{\partial \chi_c(\vec{I})} = \frac{1}{\int d^N \chi_a \phi_a} \{O_{cb}\phi_b - \Lambda\phi_c\} \tag{75}$$

$$\frac{\partial \Lambda[\chi,\phi]}{\partial \phi_c(\vec{I})} = \frac{1}{\int d^N I \chi_a \phi_a} \{\tilde{O}_{cb}\chi_b - \Lambda\chi_c\} \tag{76}$$

where

$$\tilde{O}_{cb} = \frac{\partial^2}{\partial I_c \partial I_b} + s(\Delta^T)^{-1}{}_{cb} \tag{77}$$

and T means transpose. In our ansatz, (56), Δ_{ab} is symmetric so if the variations of Λ with respect to χ and ϕ are set to zero, we have that Λ is the eigenvalue of O_{ab} we seek.

Henceforth in this paper, we will assume Δ_{ab} is almost independent of $\vec{I}$ and $D_{ab} = \Delta_{ab}$. Small corrections lie in η_{ab}.

Now we turn our attention back to the evolution equation for $F(\vec{I},t)$, Equation (20), the angle averaged phase space distribution function. We must first consider the correlation function

$$\gamma_k(\vec{I},t) \equiv P\left(\frac{\partial H_1}{\partial \theta_k} e^{-(1-P)Lt} G(\vec{I},\vec{\theta},t=o)\right). \tag{78}$$

One can derive a connection between this projected correlation function of initial information, $G(\vec{I},\vec{\theta},t{=}0)$, and the perturbation Hamiltonian H_1 and the direct correlation

$$C_k(\vec{I},t) \equiv P\left(\frac{\partial H_1}{\partial \theta_k} e^{-Lt} G(\vec{I},\vec{\theta},t = 0)\right). \tag{79}$$

Using the methods of Section II we find

$$O_{ab}(\vec{I},s)\gamma_b(\vec{I},s) = s(\Delta^{-1})_{ak}(\vec{I},s)C_k(\vec{I},s), \tag{80}$$

so the eigenvalues of O_{ab}, Equation (72), play a key role here as well, since in terms of $\psi_a^{(\lambda)}(I)$ we have

$$\gamma_b(\vec{I},s) = s\sum_\lambda \frac{\psi_b^{(\lambda)}(\vec{I},s)}{\lambda(s)} \int d^N I' \bar{\psi}_\ell^{(\lambda)}(\vec{I}',s)\Delta^{-1}(\vec{I}',s)_{\ell k} C_k(\vec{I}',s). \tag{81}$$

The behavior of both D_{ab} and γ_b is thus governed by the zeroes of $\lambda(s)$. Let us take them to be simple zeroes at $s = s_o = -\sigma_o \pm i\Omega_o$. Then each of D_{ab} and γ_b behave as $(\exp - \sigma_o t)$ $\cos(\Omega_o t + \phi_o)$ which is similar to the behavior of Δ_{ab}. σ_o and Ω_o we can find directly from $\lambda(s)$ or the variational principle for it. For ϕ_o we need some knowledge of the residue at the pole in s at s_o.

Our equation for $F(\vec{I},t)$ reads in the present notation

$$\frac{\partial}{\partial t} F(\vec{I},t) = \frac{\partial}{\partial I_k}\gamma_k\ (\vec{I},t) + \frac{\partial}{\partial I_a}\int_o^t D_{ab}(\vec{I},\tau)\frac{\partial}{\partial I_b}F(\vec{I},t-\tau). \qquad (82)$$

Our suppositions to this point have led us to the <u>same</u> exponential decay in time for each of γ_k and D_{ab}. The decay has a time scale σ_o^{-1}. For $t \gg \sigma_o^{-1}$, the first term in (82) is exponentially negligible. In the second term only $\tau \lesssim \sigma_o^{-1}$ contributes to the integral, and we may extend the integration to $\tau \to \infty$ with exponentially small error. We now arrive at

$$\frac{\partial}{\partial t} F(\vec{I},t) = \frac{\partial}{\partial I_a}\int_o^\infty D_{ab}(\vec{I},\tau)\frac{\partial}{\partial I_b} F(\vec{I},t-\tau)d\tau, \qquad (83)$$

which is our desired diffusion equation for $F(\vec{I},t)$. The non-linear dynamics lies in $D_{ab}(\vec{I},t)$ for which we have been making the ansatz

$$D_{ab}(\vec{I},t) = e^{-\sigma_o(\vec{I})t}\ \frac{\cos(\Omega_o(\vec{I})t + \phi_o(\vec{I}))}{\cos\phi_o(I)}\Delta_{ab}(\vec{I},t = o), \qquad (84)$$

noting $$D_{ab}(\vec{I},\ t = 0) = \Delta_{ab}(\vec{I},t = 0). \tag{85}$$

To this stage we have said nothing about the size of $H_1(\vec{I},\vec{\theta})$–the non-integrable piece of the underlying Hamiltonian. If, on the scale of H_0, H_1 is large, we must use the full Equation (83) to determine $F(\vec{I},t)$.

If H_1 is small, then D_{ab} is $O(H_1^2)$ and from (83) we see $\frac{\partial F}{\partial t}$ is $O(H_1^2)$. Under the integral in (83) we may write

$$F(\vec{I},t-\tau) = F(\vec{I},t) - \tau \frac{\partial F}{\partial t}(\vec{I},t) + \ldots \tag{86}$$

and, since H_1 is small, so is

$$\frac{1}{\sigma_0} \frac{\partial F}{\partial t} \frac{1}{F} \ll 1. \tag{87}$$

This allows us to drop the τ dependence of $F(\vec{I},t-\tau)$ in (83) and arrive

$$\frac{\partial}{\partial t} F(\vec{I},t) = \frac{\partial}{\partial I_a} D_{ab}(\vec{I}) \frac{\partial}{\partial I_b} F(\vec{I},t) \tag{88}$$

with $$D_{ab}(\vec{I}) = \int_0^\infty dt\ D_{ab}(\vec{I},t), \tag{89}$$

which is the form of diffusion equation encountered in the quasi-linear approximation.[4]

We call (83) the evolution equation for <u>non-linear diffusion</u>. The equation is of course linear in $F(\vec{I},t)$, but through D it incorporates the non-linearity of the inderlying dynamical problem governed by $H(\vec{I},\vec{\theta})$.

V. Conclusions

In this paper we have studied the evolution equation for the angle averaged phase space density

$$F(\vec{I},t) = \int \frac{d^N\theta}{(2\pi)^N} f(\vec{I},\vec{\theta},t)$$

where $f(\vec{I},\vec{\theta},t)$ – the density in action-angle space – satisfies the usual Liouville equation. $F(\vec{I},t)$ satisfies an exact equation, Equation (20), which we derived using the projection operator technique of Zwanzig.[6] By studying the connection between the diffusion tensor D_{ab} entering the evolution equation for $F(\vec{I},t)$ and the direct diffusion tensor $\Delta_{ab}(\vec{I},t)$ (Equations (23) and (37)) we argued that for times large compared to the exponential decay of correlation functions of the general form

$$C_{AB}(\vec{I},t) \equiv \int \frac{d^N\theta}{(2\pi)^N} A(\vec{I},\vec{\theta}) e^{-Lt} B(\vec{I},\vec{\theta})$$

$$- \left[\int \frac{d^N\theta}{(2\pi)^N} A(\vec{I},\vec{\theta})\right]\left[\int \frac{d^N\theta}{(2\pi)^N} B(\vec{I},\vec{\theta})\right] ,$$

the equation of evolution for $F(\vec{I},t)$ becomes the usual diffusion equation to exponential accuracy with the diffusion tensor governed by angle projected dynamics, see Equations (21) and (22).

The exponential fall off in time of $C_{AB}(\vec{I},t)$ which we tested numerically in the so-called standard mapping of Chirikov[1,3] we called the <u>quasi-mixing</u> <u>hypothesis</u>. More specifically we suggested writing $C_{AB}(\vec{I},t)$ as

$$C_{AB}(\vec{I},t) = \frac{e^{-\sigma(\vec{I})t}\cos(\Omega(\vec{I})t + \phi(\vec{I}))}{\cos\,\phi(\vec{I})} C_{AB}(\vec{I},t = o).$$

This quasi-mixing can only be true in the regions of $\vec{I}$, $\vec{\theta}$ space where irregular or chaotic motion is occurring. Outside those regions we generally anticipate $\sigma = 0$ and the correlation functions will show oscillations characteristic of regular motion.

Although we did not explore any applications of our diffusion formalism we have two in mind which will be pursued in future articles: (1) the diffusion away from adiabatically constant values of the adiabatic invariants of a particle in a magnetically confined plasma when one tries to heat the plasma with electrostatic waves.[13] The formulation of this problem by Smith and Kaufman[15] is precisely adapted to the techniques presented here. (2) The braiding or destruction of magnetic field lines[5] in the presence of small perturbations is also amenable to the analysis given here. With various physical mechanisms in mind for the perturbations of the lines, we can proceed to use our evolution equation for $F(\vec{I},t)$ to estimate the transport of energy from the confining regions.

A last point: we have talked in this paper about the behavior of $F(\vec{I},t)$ for times long compared to the decay time of angle averaged correlation functions and only in the region of chaotic motion. Since the evolution equation for $F(\vec{I},t)$ is exact, our ansatz for $C_{AB}(\vec{I},t)$ may prove useful for all $\vec{I}$ and t recalling $\sigma(\vec{I}) = 0$ in domains of regular motion.

Acknowledgments

I have enjoyed many conversations with A. M. DeSpain, C. Grebogi, andf R. M. Kulsrud on the subjects in this paper. A. N. Kaufman read the manuscript carefully and made many key suggestions for its improvement.

References

1. B. V. Chirikov, "A Universal Instability of Many-Dimensional Oscillator Systems,", Physics Reports 52, 263 (1979).
2. M. V. Berry, "Regular and Irregular Motion," pp. 16-121 in Topics in Non-Linear Dynamics, ed. S. Jorna (American Institute of Physics, N. Y., 1978).
3. The situation has recently been reviewed by A. J. Lichtenberg in "Determination of the Transition between Adiabatic and Stochastic Motion," Proceedings of the International Workshop on "Intrinsic Stochasticity in Plasmas" available from "Les Editions de Physique, BP-112, 91402 Orsay Cedex, France.
4. A. N. Kaufman, Phys. Fluids 15, 1063 (1972).
5. M. N. Rosenbluth, R. Z. Sagdeev, J. B. Taylor, and G. M. Zaslavsky, Nucl. Fusion 6, 297 (1966); N. N. Filonenkov, R. Z. Sagdeev, and G. M. Zaslavsky, Nucl. usion 7, 253 (1967).
6. R. W. Zwanzig, Lectures in Theoretical Physics, Vol III, ed. W. E. Brittin, B. W. Downs, and J. Downs (J. W. Wiley, New York 1961) pp. 106-141; H. Mori, Prog. Theor. Phys. 33, 423 (1965).
7. G. M. Zaslavskii and R. Z. Sagdeev, Soviet Physics-JETP 25, 718 (1967).
8. Ya. G. Sinai, "Ergodic Theory," Acta Physica Austriaca, Suppl. X, 575-608 (1973); V. I. Arnold and A. Avez, Ergodic Problems of Statistical Mechanics, (W. A. Benjamin, New York, 1968).
9. J. Ford, "The Statistical Mechanics of Classical Analytical Dynamics," p. 215-255 in Fundamental Problems in Statistical Mechanics, Vol. 3, ed. E. D. B. Cohen (North-Holland, Amsterdam, (1975).

10. G. M. Zaslavskii and B. V. Chirikov, Soviet Physics-Uspekhi 14, 549 (1972).

11. K. C. Mo, Physica 57, 445 (1972).

12. J. M. Greene, J. Math. Phys. (to be published).

13. G. R. Smith, "Stochastic Acceleration by a Single Wave in a Magnetized Plasma," Lawrence Berkeley Laboratory Report LBL-6824, September, 1977. (Ph.D. thesis, Univ. Microfilm).

14. If $\Delta_{ab}(\vec{I},t)$ falls as $e^{-\sigma t}$ then $(\Delta^{-1})_{ab}$ will be regular at $s = o$, but if $\Delta_{ab}(\vec{I},t)$ only falls as a power of t, Δ^{-1} will be singular at $s = o$.

15. G. R. Smith and A. N. Kaufman, Phys. Fluids 21, 2230 (1978).

Figure Captions

Figure 1. 10^4 points along one trajectory for the standard mapping, Equations (60) and (61) of the text, for $k = 3.5$. The initial point was $I_0 = 0.4357$, $\theta_0 = 0.0798695$.

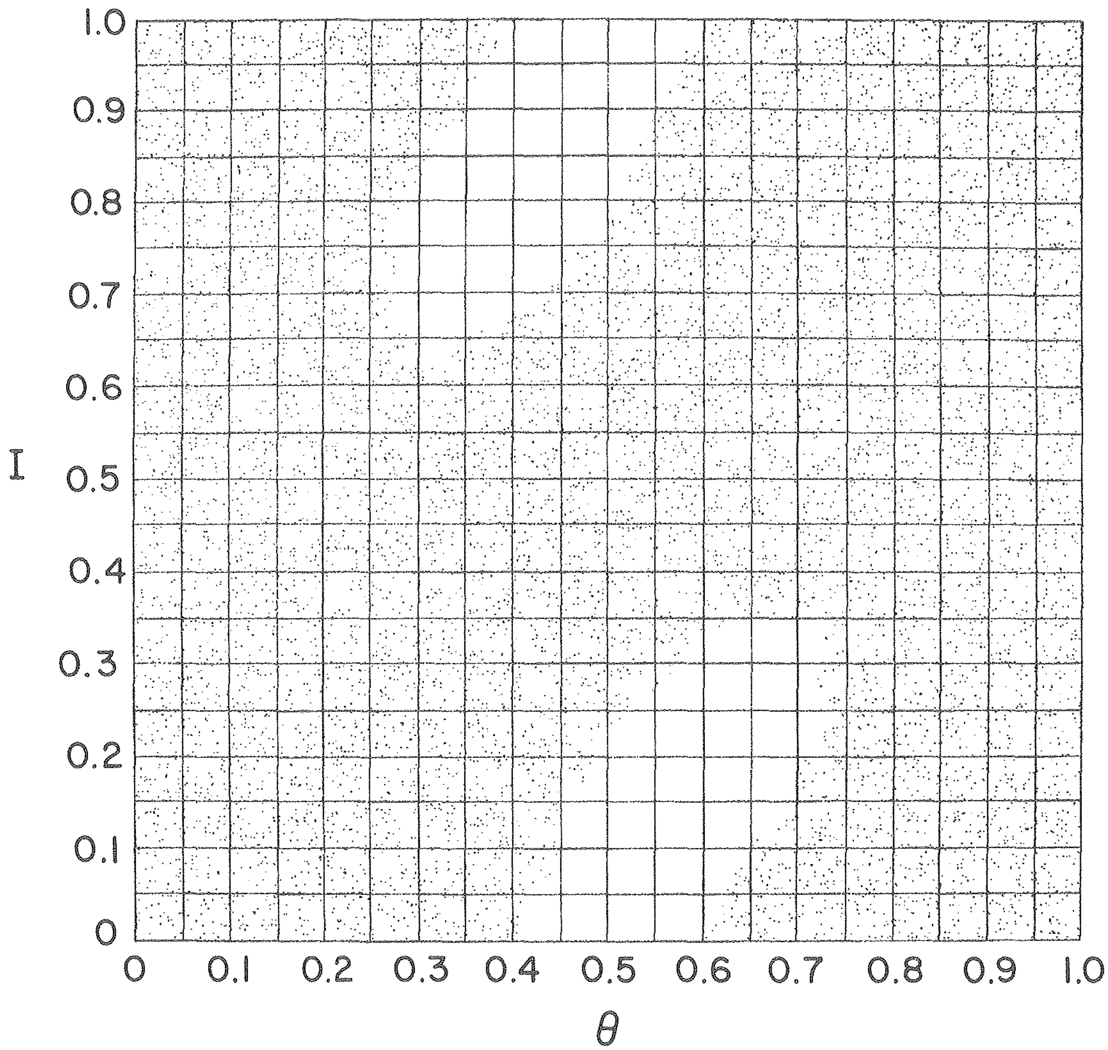
Figure 2. The correlation function $C(N,I_0) = 2 \int_0^1 d\theta_0 \sin 2\pi\theta_N(I_0,\theta_0) \sin 2\pi\theta_0$ for the standard mapping with $k = 3.5$, $I_0 = 0.4357$. 2000 integration points were used. The use of more integration points reproduces the values shown.

Figure 3. Comparison between $C(N,I_0)$ in Figure 2 and $e^{-N\sigma} \cos(N\Omega + \phi)/\cos \phi$.

Figure 4. $C(N,I_0)$ for $k = 3.5$ and $I_0 = 0.4$, 0.5, and 0.55, 2000 integration points were used in each case.

Figure 5. $C(N,I_0)$ for $I_0 = 0.5$ and $k = 3.5$, 4.0, 4.5, and 5.0, 7000 integration points were used in each case.

Figure 6. $C(N,I_0)$ for $I_0 = 0.5$ and $k = 5.00$ and $k = 5.03$. 7000 integration points were used in each case.



Standard mapping k = 3.5 10,000 iterations

$I_0 = 0.4357 \quad \theta_0 = 0.078695$

Fig. 1

XBL 804-578

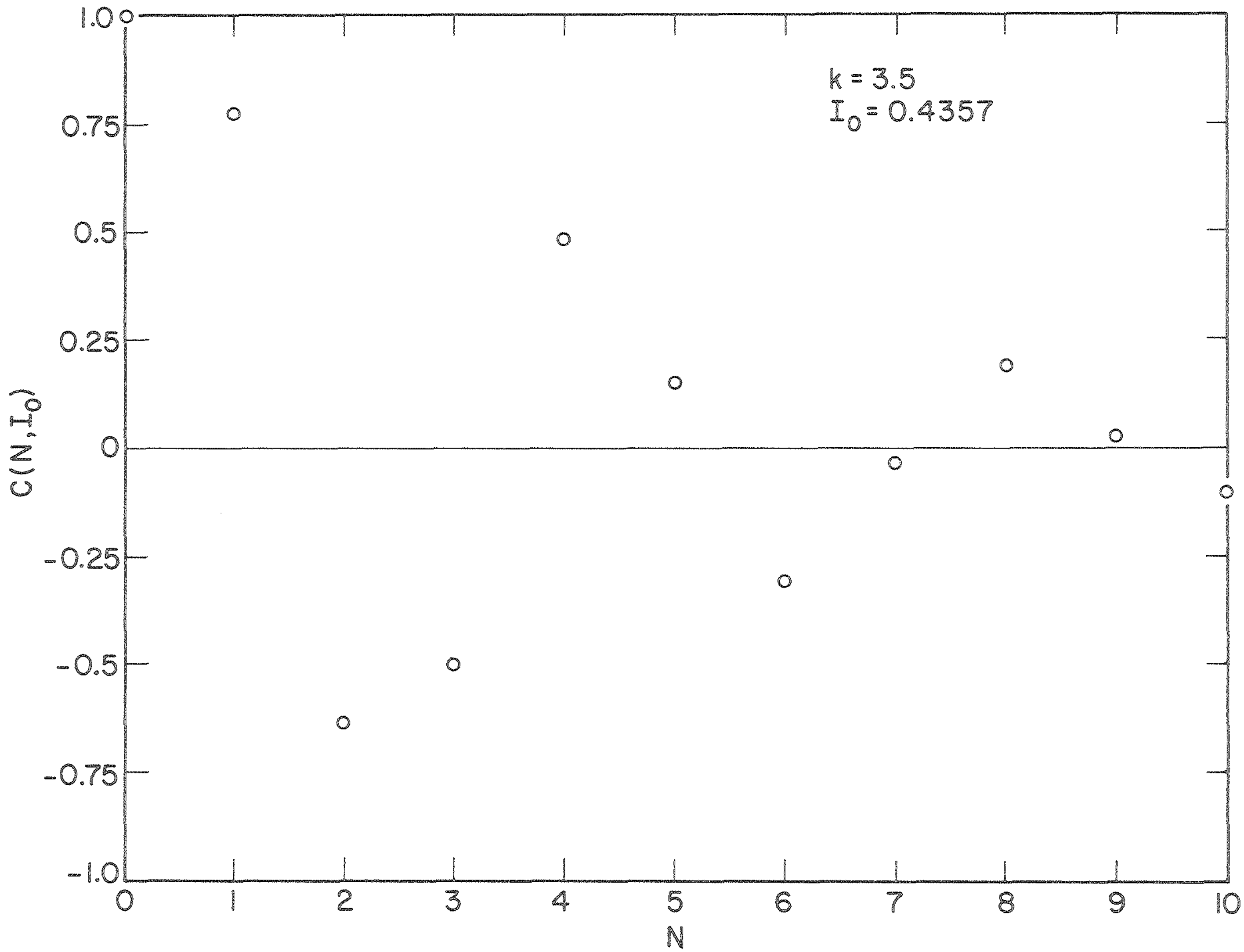


Fig. 2

XBL 804-581

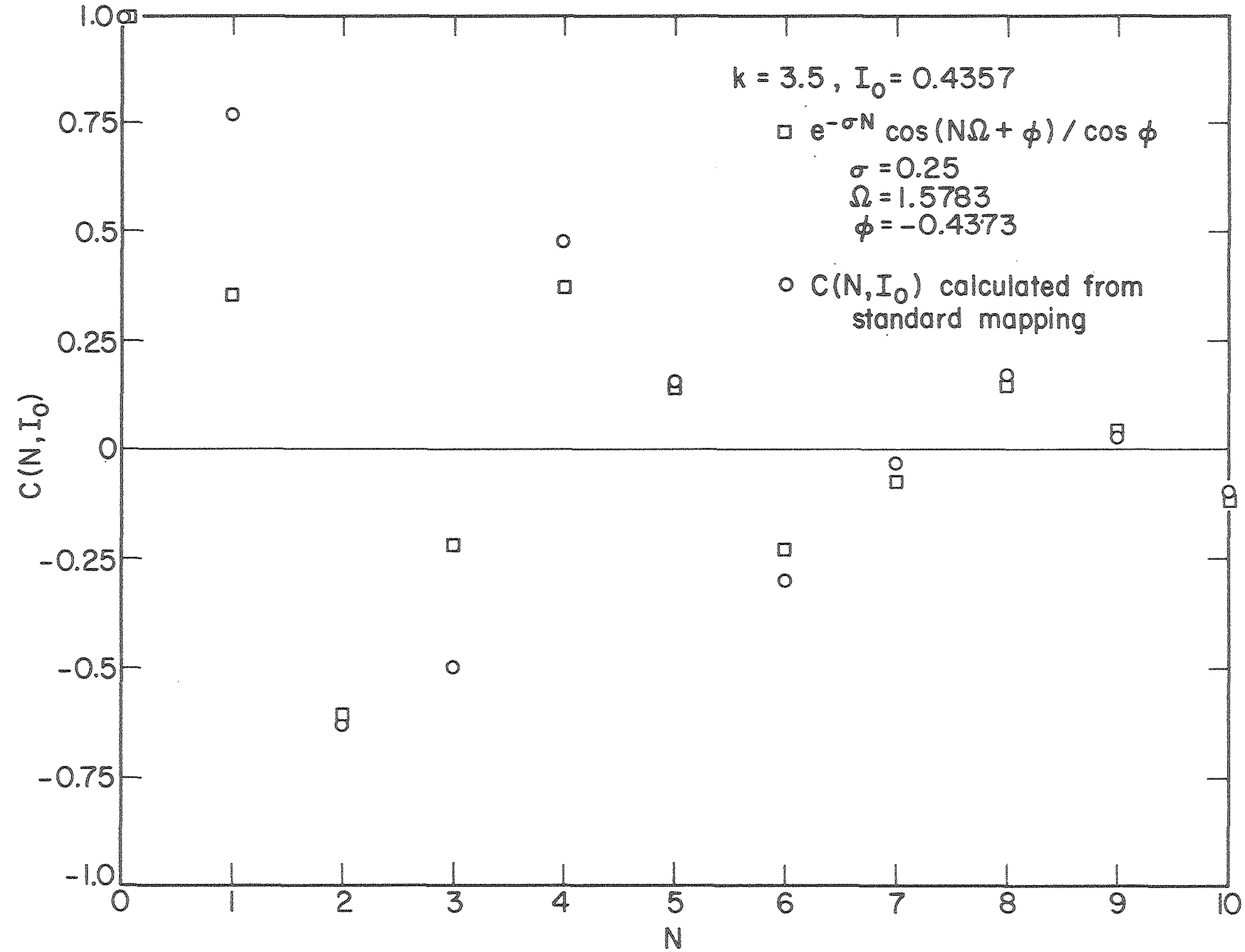


Fig. 3

XBL-804-583

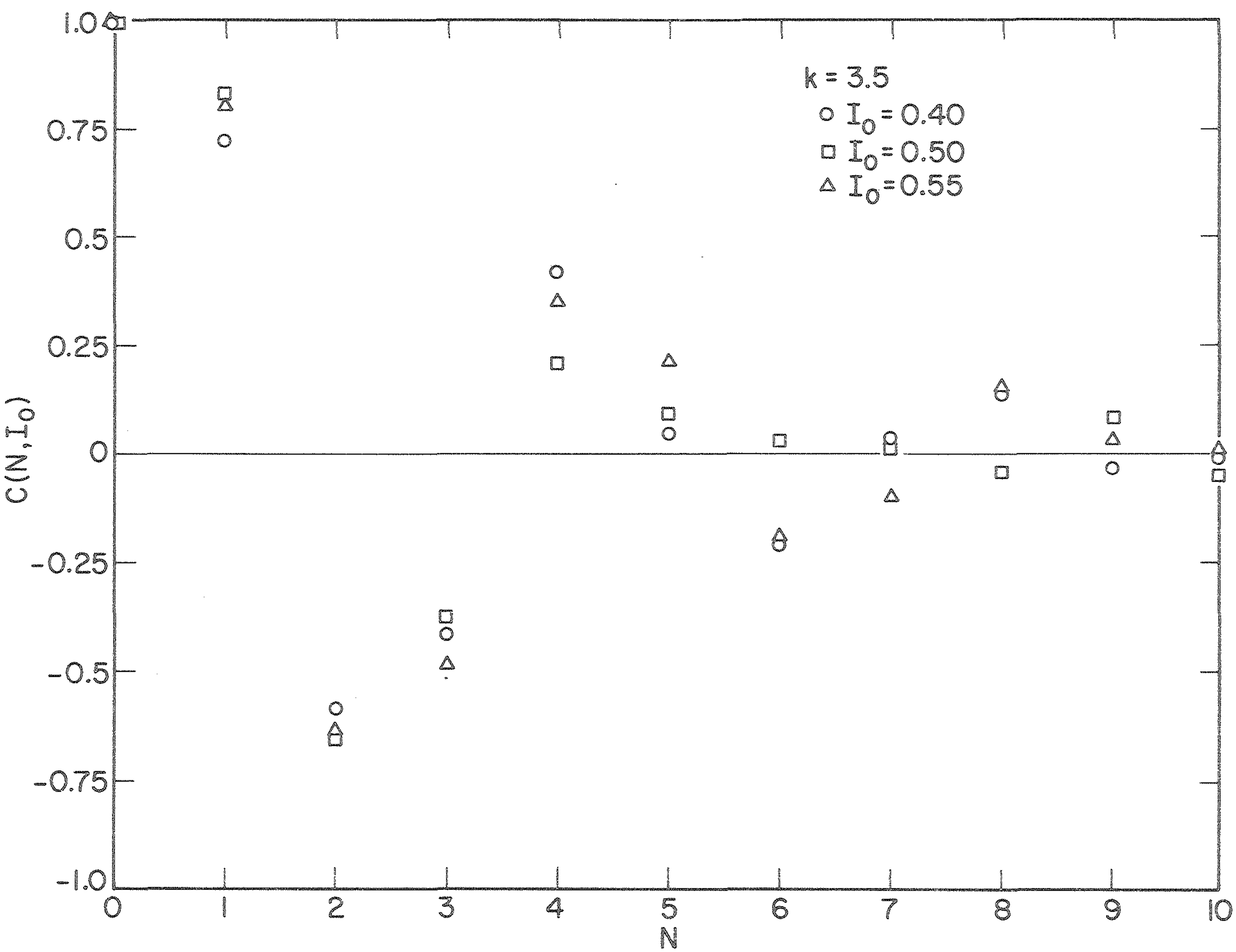


Fig. 4

XBL 804-582

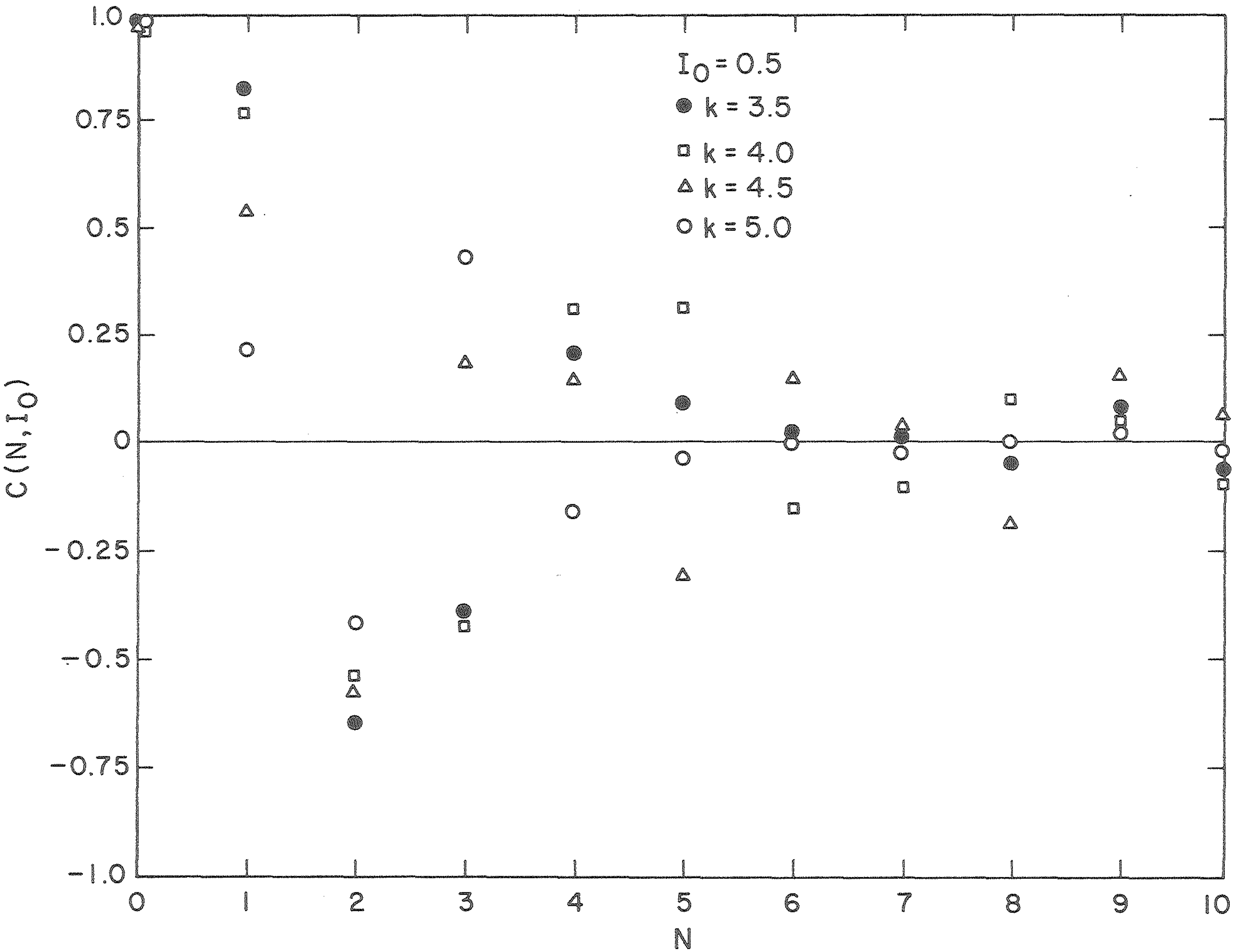
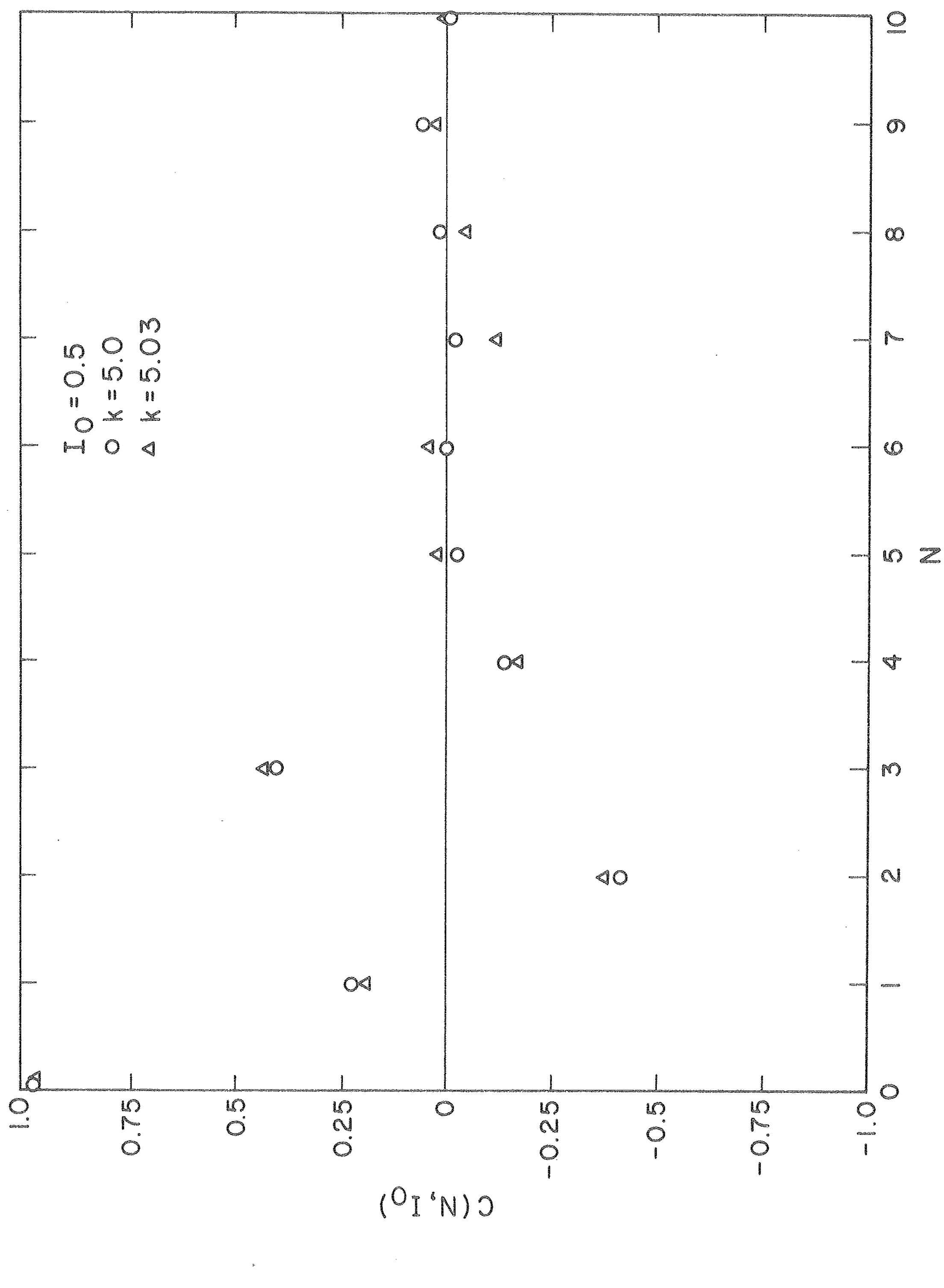


Fig. 5

XBL 804-580



XBL 804-579

Fig. 6