Number Theory and Dynamical Systems
(Preliminary Notes)

J. C. Lagarias

ATT Bell Laboratories
Murray Hill, NJ 07974

(August 5, 1991)

1. Introduction

In the 19th century dynamical systems meant solutions differential equations having analytic coefficients – the word *analytical dynamics* was often used. Much work was motivated by problems of celestial mechanics. In studying limiting motions of solutions of differential equations, Poincaré (1885), (1912) introduced the study of maps of the circle, and the concept of rotation number. In this century the notion of what constitutes a dynamical system has steadily broadened. G. D. Birkhoff (1913), (1927) strongly influenced research on dynamical systems and initiated the development of ergodic theory. Dynamical systems were extended to allow differentiable maps, which can exhibit chaotic behavior, cf. Smale (1967). In addition there was a development of *symbolic dynamics*, a term due to Morse and Hedlund (1983). This area traces back to work of Hadamard (1898) who used symbols to describe the paths of geodesics on a manifold of constant negative curvature. Other symbolic codings were used by Artin (1924) to prove the existence of dense geodesics. Morse and Hedlund, gave spaces of symbols a topology and defined dynamics by shift transformations on them. This led to the development of *topological dynamics* by Hedlund and others and to the viewpoint that iterations of transformations $T : X \to X$ on a compact metric space can be viewed as (discrete) dynamical systems.

Number theoretic problems have appeared repeatedly in dynamical systems. For continuous systems this initially seems surprising since number theory deals with discrete objects. Many connections arise through the occurrence of periodic and quasiperiodic orbits in dynamical systems. The interaction of motions with two different frequencies depends on Diophantine approximation properties of the frequency ratio ω_1/ω_2. Conversely, one-dimensional Diophantine approximation itself can be analyzed as a discrete dynamical system, given by the continued fraction algorithm. This paper describes four connections between number theory and dynamical systems.

First, the dynamical behavior of continuous systems sometimes leads to the appearance of rational numbers. The Solar System supplies numerous examples.

(1) The axis rotation period and orbital rotation period for the moon around the earth have ratio 1:1. The axis rotation period and orbital rotation period for Mercury around the sun have ratio 3:2.

(2) The asteroid Pallas has mean orbital period in the ratio 7:18 with that of Jupiter, as observed by Gauss. Gauss published a note in an encrypted form in the bulletin of the Göttingan academy, to establish priority. In 1812 he mentioned it in a letter to his friend Bessel, asking him to keep it secret ‘‘for the duration.’’ (see Schroeder (1990), p. 171.) In addition, Several asteroids have periods commensurable with that of Jupiter in the ratio 1:1 (Trojans) and 3:2 (Hildas).

(3) In 1866 Kirkwood observed the relative absence of asteroids having periods 3:1, 5:2 and 7:3 with Jupiter, see Figure 1.1. These are called *Kirkwood gaps.*

(4) The moons I_O, Europa and Ganymede of Jupiter have mean periods λ_I, λ_E and λ_G satisfying

$$\lambda_I - 3\lambda_E + 2\lambda_G \doteq 0$$

to within an error of 10^{-9}. (See Peale (1976)).

(5) There is a complicated set of gaps in Saturn’s rings, see Figure 1.2. A set of these gaps are

related to the moon Mimas by the relation

$$\lambda_{Gap} - 2\lambda_M + \omega_{Gap} \doteq 0$$

where ω_{Gap} is the longitude at perihelion of the gap. (See Peale (1976)). Other ring gaps are related to other moons, including a recently found 18th moon, named 1981 S13 or Pan, which orbits inside Encke's gap in the outermost ring.

Most of these commensurabilities are examples of *mode-locking phenomena.* Such phenomena are discussed in §6.

Second, a major dynamical problem where number theory arises concerns the question of stability of the Solar system: Will the planets continue to orbit in well-defined zones without collision? A mathematical abstraction of this question is: Will point masses representing the planets, evolving according to Newtonian dynamics, remain for all time inside disjoint compact sets (invariant tori), in terms of a coordinate system taking the center of gravity of the Solar System as the origin? Recent numerical simulations suggest that Pluto and its satellite Charon do not have this property, and their orbit may undergo major changes, leaving the plane of the ecliptic within 8×10^8 years.

This problem has been approached mathematically by applying perturbation theory to analyze small perturbations of completely integrable Hamiltonian dynamical systems. Completely integrable systems have regular motion confined within invariant tori. The problem is to assess the effect of perturbations on breaking up such tori. This is the subject matter of KAM (Kolmogorov-Arnol'd-Moser) theory, which says that under small enough perturbations many invariant tori persist. Invariant tori are destroyed because orbits with certain rational ratio of periods (''resonances'') can inflate small oscillations. Analyzing the effect of ''resonances'' is the problem of ''small divisors.'' It turns out to be related to Diophantine approximation properties of the ratios of two mean orbital periods in the unperturbed system.

Integrable Hamiltonian systems are discussed in §3 and KAM theory and ‘‘small divisors’’ problems in §4 and §5. The rigorous bounds that have been established, so far, which guarantee the existence of some invariant tori, do not come close to establishing the stability of the major planets in the Solar System. There is a gap between the bounds that are proved and experimental evidence on how far they are valid; this is illustrated for the ‘‘standard map’’ in §4.

Third, it is well-known that the ordinary continued fraction algorithm itself can be treated as a discrete dynamical system on [0,1], which yields a metric theory of continued fractions. Here dynamical systems appear in number theory itself. The ordinary continued fraction finds the set of best Diophantine approximations to a real number θ. There is another, less well-known, variant of the continued fraction, called the *additive continued fraction,* which finds all the best one-sided Diophantine approximations. The additive continued fraction has an analogous dynamical theory, attached to the *Farey shift map* on [0,1]. Here there is more than one natural invariant measure. In §2 we describe two such measures, an absolutely continuous measure and a singular measure associated to the function $?(x)$ of Minkowski (1904). It turns out that the additive continued fraction seems more closely connected to many dynamical problems then the ordinary continued fraction; in particular, the mode-locking problems of §6 are related to singular measures.

Fourth, number theoretic structures occur is in the study of quasicrystals. Quasicrystals are solids whose X-ray diffraction patterns appear to have sharp peaks at points (‘‘Bragg peaks’’) and exhibit symmetrics forbidden in any periodic lattice in R^3, e.g. 5-fold symmetries are observed. Number-theoretic constructions including cut-end-point methods and *substitution sequences,* show that quasi-periodic patterns exist which could exhibit Fourier transforms with the proper patterns. Recent scanning tunneling microscope pictures of pure quasicrystals do indicate that physical quasicrystals sometimes have a quasiperiodic structure. These are discussed in §7. Substitution rules can also be directly used to construct interesting symbolic dynamical systems, some of which have been extensively studied as mathematical objects.

One of the major open directions relating number theory and dynamical systems concerns finding the precise connections between multidimensional dynamical systems, and multidimensional Diophantine approximation. It is apparent that the situation in higher dimensions is complicated and fragmented.

There are many other occurrences of problems arising in dynamical systems that have connections with number theory. In addition number theory contains problems having a dynamical flavor, for example the notorious $3x + 1$ problem, see Lagarias (1985). Some general references are given in §0 in the reference list.

2. Continued Fractions as Dynamical Systems

The ordinary continued fraction expansion $\theta_0 = [a_0, a_1, \ldots, a_n, \ldots]$ of a real number has *partial quotients* a_n and *partial remainders* θ_n computed by the recursion

$$\theta_n = a_n + \frac{1}{\theta_{n+1}} \tag{2.1}$$

where $a_n = \lfloor \theta_n \rfloor$ is the largest integer not exceeding θ_n The process halts if some $\theta_n = a_n$; which occurs exactly when θ is rational. The *convertents* $\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n]$ are given by $\frac{p-1}{q-1} = \frac{1}{0}, \frac{p_0}{q_0} = \frac{a_0}{1}$ and

$$\frac{p_{n+1}}{q_{n+1}} = \frac{a_n p_n + p_{n-1}}{a_n q_n + q_{n-1}} \ . \tag{2.2}$$

They are exactly the set of *best two-sided Diophantine approximations* to θ, i.e. $\frac{p}{q}$ is a convergent if and only if

$$0 \le |q\theta - p| < |q'\theta - p'|$$

holds whenever $0 < q' < q$. The partial remainders θ_n satisfy

$$\theta_{n+1} = T(\theta_n)$$

where $T : [0,1] \to [0,1]$ is the *Gauss map*

$$T(x) = \{\frac{1}{x}\} \quad (\text{mod } 1) \tag{2.3}$$

where $\{y\} = y - \lfloor y \rfloor$ is the fractional part of y. By convention $T(0) = 0$. (See Figure 2.1) Gauss already was aware of the fact that $\frac{dt}{1+t}$ is an invariant measure for T, i.e.,

$$\int_{[a,b]} \frac{dt}{1+t} = \int_{T^{-1}[a,b])} \frac{dt}{1+t} ,$$

for all intervals $[a,b]$. Using current ergodic theory terminology, the continued fraction algorithm is a *skew-product* over the base space $(T, \frac{1}{\log 2} \frac{dt}{1+t})$ together with a fiber map $A : [0, 1] \to \mathbb{Z}^+$ describing the partial quotients, with

$$A(\theta = k \quad \text{if } \frac{1}{k+1} < \theta \le \frac{1}{k} ,$$

and $A(0) = 0$.

In the 1920's P. Levy and A. Khinchin (1930) developed a metric theory of continued fractions, which treats θ as a random variable on $[0,1]$ drawn with the Gauss measure $d\mu_G = \frac{1}{\log 2} \frac{dt}{1+t}$. The Gauss map with the Gauss measure is ergodic, and one has;

Theorem 2.1 *(Khinchin) If* θ *is drawn from [0,1] with the Gauss measure* $\frac{1}{\log 2} \frac{dt}{1+t}$ *then the partial quotients of* $\theta = [0, a_1, a_2, \ldots]$ *are independent and identically distributed with*

$$\text{Prob}[a_k = m] = \log_2 (1 + \frac{1}{m(m+2)}) . \tag{2.3}$$

From this one can deduce such facts as:

Theorem 2.2 *(Khinchin-Levy Theorem) There is a set of Lebesgue measure one in [0,1] such that all* θ *in this set have*

$$\lim_{n \to \infty} \left[\prod_{k=1}^{n} a_k \right]^{1/n} = \prod_{m=1}^{\infty} \left(1 + \frac{1}{m(m+2)}\right)^{\frac{\log m}{\log 2}} = 2.6 \qquad (2.4)$$

and

$$\lim_{n \to \infty} \frac{1}{n} \log(q_n) = \frac{\pi^2}{12 \log 2} . \qquad (2.5)$$

The Gauss measure is the unique absolutely continuous invariant measure for the Gauss map. In particular, if one iterates any other absolutely continuous invariant measure repeatedly by the Gauss map, it will converge to the Gauss measure, a fact which Gauss asserted in a letter to Laplace. It was proven by Kuzmin in 1928. In fact this convergence is exponentially fast.

Theorem 2.3 *(Doeblin) Let* $d\lambda_n$ *be the* n^{th} *iterate of Lebesgue measure* $d\lambda$ *under the Gauss map T. There exists* $\delta < 1$ *and* c_0 *such that*

$$1 - c_0 \delta^n \le \frac{d\lambda_n(\mathbf{J})}{d\mu_G(\mathbf{J})} \le 1 + c_0 \delta^n$$

for any open interval **J** *in [0,1].*

P. Levy showed that $\delta < .68$ and Wirsing (1974) determined the best constant $\delta = .303663\ldots$. It is interesting that this result can be proved by a procedure analogous to ‘‘renormalization’’ in statistical mechanics, see Mayer (1990).

The *additive continued fraction algorithm* is a variant of the continued fraction algorithm that finds all the intermediate convergents to θ as well as the ordinary continued fraction convergents. (The *intermediate convergents* of the ordinary continued fraction algorithm are the fractions $\frac{lp_n + p_{n-1}}{lq_n + q_{n-1}}$ having $1 \le l < a_n$.) These are exactly the *best one-sided Diophantine* *approximations* to θ, i.e. exactly those fractions $\frac{p}{q}$ such that if $\frac{p}{q} < \theta$ then

$$0 < q\theta - p < q'\theta - p'$$

whenever $q' < q$ and $\frac{p'}{q'} < \theta$, while if $\frac{p}{q} > \theta$ then

$$q'\theta - p' < q\theta - p < 0$$

whenever $q' < q$ and $\frac{p'}{q'} > 0$.

The additive continued fraction has a simple description in terms of the Farey tree pictured in Figure 2.2. The Farey tree ^ is an infinite binary tree whose vertices are labelled by all rationals $\frac{p}{q}$ in the open interval $(0, 1)$, as follows. Two fractions $\frac{p_1}{q_1}$ and $\frac{p_2}{q_2}$ are *Farey neighbors* if $p_1 q_2 - p_2 q_1 = \pm 1$, i.e. the interval $[\frac{p_1}{q_1}, \frac{p_2}{q_2}]$ has length $\frac{1}{q_1 q_2}$. For two Farey neighbors, their *mediant*

$$\frac{p_1}{q_1} \oplus \frac{p_2}{q_2} = \frac{p_1 + p_2}{q_1 + q_2} \tag{2.6}$$

is always in lowest terms, and is called their *Farey sum.* The Farey tree starts with the unit interval $[\frac{0}{1}, \frac{1}{1}]$ having mediant $\frac{1}{2} = \frac{0}{1} \oplus \frac{1}{1}$, and the mediant $\frac{1}{2}$ is the label of the root node of the binary tree. Now split the unit interval at its mediant to obtain two subintervals $[\frac{0}{1}, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$, which have mediants $\frac{1}{3}$ and $\frac{2}{3}$. These label the leaves at level one of the Farey tree. Continuing to split at mediants, at level n we obtain a partition of $[0,1]$ into 2^n subintervals whose rediants label the Farey tree at level n. We have:

Lemma 2.4. *Every rational* $\frac{p}{q} \in (0,1)$ *can be expressed uniquely as a Farey sum* $\frac{p}{q} = \frac{p_1}{q_1} \oplus \frac{p_2}{q_2}$. *Every such rational occurs uniquely as a vertex of the Farey tree.*

To each vertex $\frac{p}{q}$ of the Farey tree is associated this interval $[\frac{p_1}{q_1}, \frac{p_2}{q_2}]$. The *additive continued fraction convergents* of θ are exactly the sequence of vertices of the Farey tree which is encountered in taking the sequence of nested intervals that contain θ at each level of the Farey tree. This sequence is unique if θ is irrational; there are exactly two such sequences if θ is rational.

There is a nice symbolic dynamics attached to the Farey tree, in which each edge of the tree is labelled by L (for Left) or R (for Right) according to whether it gives the left or right subinterval of the interval being split. The sequence of edges that one follows from the root in following a set of nested intervals for θ is called a *Farey symbol sequence* for θ. If one sets

$$L = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \ R = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \tag{2.7}$$

then the additive continued fraction convergents can be encoded using matrix multiplication.

There is a metric theory for the additive continued fraction, associated to the *Farey shift map* $U : [0, 1] \to [0, 1]$ is

$$U(x) = \begin{cases} \dfrac{x}{1-x} & 0 \leq x \leq \dfrac{1}{2} \\ \dfrac{1-x}{x} & \dfrac{1}{2} \leq x \leq 1 \ . \end{cases} \tag{2.8}$$

See Figure 2.3. There is a symbolic dynamics associated to iterates of U which keeps track of whether a point falls in $[0, \frac{1}{2})$ (labelled with symbol A) or in the right half (labelled B).

Theorem 2.5. *The symbolic dynamics for the Farey shift map for* $0 \in (0, 1]$ *essentially encodes the additive continued fraction expansion of* θ*, in this sense: There is a finite automaton that converts on* $\{A, B\}$ *symbol sequence for* θ *to an* $\{L, R\}$ *symbol sequence and vice-versa.*

The Farey shift map actually encodes θ using the matrices

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} , \; B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} . \tag{2.9}$$

The Farey shift map has the absolutely continuous invariant density $\frac{dt}{t}$, and it is ergodic with respect to this density, see Parry (1962). This density has infinite mass, which implies that the metric entropy of the mapping $(U, \frac{dt}{t})$ is *zero.* (Metric entropy has a complicated definition, for which see Petersen (1983).)

However the Farey shift has another invariant measure *d*? which is the ‘‘derivative’’ of Minkowski’s ?-function. The ?-function of Minkowski (1904) is pictured in Figure 2.4. Minkowski introduced this function as a monotonic function on [0,1] which gives a one-to-one order-preserving map from all real quadratic irrationals in [0,1] into the set of rationals with denominator not a power of 2, e.g. $?\left[\frac{\sqrt{5} - 1}{2}\right] = \frac{2}{3}$. The simplest definition of $?(\theta)$ is that if $\theta = [0, a_1, a_2, a_3, \ldots]$ then

$$?(\theta) = \sum_{k=1}^{\infty} (-1)^K 2^{-(a_1 + \cdots + a_k - 1)} \tag{2.10}$$

Salem (1943) showed that $?(x)$ is a singular function, so that *d*? is a singular measure. Kinney (1960) shows that *d*? is concentrated on a set of Hausdorff dimension α given by the formula

$$\alpha = \frac{1}{2} \left(\int_0^1 \log_2 (1 + x) \, d?(x) \right)^{-1}$$

in the sense that, for any set *S* of Hausdorff dimension less than α one has $\int_S d?(x) = 0$, while there is a set *S* of Hausdorff dimension α with $\int_S d?(x) = 1$. Computer calculations establish that

$$.8746 < \alpha < .8749 . \tag{2.11}$$

The measure $d?$ is characterized by the property that the dynamical system $(U, d?)$ has maximal entropy. It is known that the metric entropy of a continuous map of the interval never exceeds its *topological entropy*

$$h_{top}(U) = \lim_{n \to \infty} \frac{1}{n} \log(\mathrm{Fix}(U^{(n)})) \tag{2.12}$$

where $\mathrm{Fix}(F)$ counts the number of fixed points of a map F. Clearly $\mathrm{Fix}(U^{(n)}) = 2^n$ so that

$$h_{top}(U) = \log 2 \ .$$

The Minkowski measure $d?$ turns out to assign equal weight to every edge of the Farey tree, and Theorem 2.5 then implies that $(U, d?)$ is just the Bernoulli shift on the two letters $\{R, L\}$, hence it has metric entropy $\log 2$. This has the following consequence.

Theorem 2.6. *If* θ *is drawn from [0,1] according to the singular measure* $d?(x)$*, then the partial quotients of* $\theta = [0, a_1, a_2, \cdots]$ *are independent and identically distributed with*

$$\mathrm{Prob}[a_k = m] = m^{-m} \ .$$

The Farey tree construction arises in renormalization schemes to analyze mode-locking phenomona, described in §6, and also in analyzing rotation orbits of circle maps $f : S^1 \to S^1$ *of degree d* > 1*, see Goldberg and Tresser (1991).*

There are several other connections between continued fractions and dynamical systems. First, a real number θ *can be described by a sequence of symbols describing elements of the lattice* Z^2 *adjacent to the line* $(t, \theta t)$ *in* R^2*. This viewpoint traces back to H. J. S. Smith in the 1870's. It is described in Series (1985a) and in McIlroy's paper, as it is useful in computer graphics. This description also arises as a special case of the cut-and-project method of constructing one-dimensional quasicrystals described in 9sc7. Second, there is a well-known use of continued fractions to code geodesics on the modular surface* $\mathsf{H}/SL(2, \mathsf{Z})$ *and other surfaces ofconstantnegativecurvature. For this see Series (1982), (1985b). Third, Arnoux and Nogueira*

(1991) indicate how to reformulate some continued fraction transformations as one-to-one area-preserving maps, which are analogous to a completely integrable Hamiltonian systems.

A nice description of the additive continued fraction appears in Richards (1981). The facts concerning the Farey shift stated here are mostly known ones; proofs are given in Lagarias (1991).

3. Integrable Hamiltonian Dynamical Systems

The dynamical systems of classical mechanics can be put in Hamiltonian form. The Hamiltonian formation has two important features: it directly identifies conserved quantities of the flow, and it is the classical starting point to obtain quantum mechanical systems, see Dirac (1950). The second order differential equations of Newtonian dynamics can be converted to a Hamiltonian system of first order differential equations. (Hamilton actually developed this formalism in connection with his work in optics, however, rather than mechanics.)

The *Hamiltonian function* $H(\mathbf{p}, \mathbf{q}, t)$ describes the dynamics, where $t \in \mathbb{R}$ is a time variable, and $(\mathbf{p}, \mathbf{q})$ lie on a $2n$-dimensional manifold $\}$, locally resembling $\mathbb{R}^{2n}$, called ‘‘phase space.’’ Here $\mathbf{p} \in \mathbb{R}^n$ are traditionally called *position variables* and $\mathbf{q}$ *momentum variables.* The equations of motion are

$$\begin{aligned} \frac{d\mathbf{p}}{dt} &= \frac{\partial H}{\partial \mathbf{q}} \ , \\ \frac{d\mathbf{q}}{dt} &= -\ \frac{\partial H}{\partial \mathbf{p}} \ . \end{aligned} \tag{3.1}$$

One consequence is

$$\begin{aligned} \frac{\mathrm{H}}{dt} &= \frac{\partial H}{\partial t} + \left\langle \frac{\partial H}{\partial \mathbf{p}} \ , \ \frac{d\mathbf{p}}{dt} \right\rangle + \left\langle \frac{\partial H}{\partial \mathbf{q}} \ , \ \frac{d\mathbf{q}}{dt} \right\rangle \\ &= \frac{\partial H}{\partial t} \ . \end{aligned}$$

In particular if $H = H(\mathbf{p}, \mathbf{q})$ is *autonomous* (independent of t-variable) i.e. the time variable does not appear explicitly then $\frac{H}{dt} = 0$, so that H is a *conserved quantity* under the flow (3.1).

In what follows we generally suppose that $H(\mathbf{p}, \mathbf{q})$ is autonomous. The simplest example is

$$H(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{n} \frac{1}{2} p_n^2 + V(\mathbf{q})$$

where the first term represents kinetic energy and the second term potential energy; this Hamiltonian prescribes conservation of energy.

The phase space } is endowed with a distinguished two-form *(symplectic form),* which locally has the form

$$\omega = \Sigma \, dp_i \wedge dq_i \ , \tag{3.2}$$

and which is *closed* ($d\omega = 0$). Any transformation of coordinates $\Phi(\mathbf{p}, \mathbf{q}) = (\mathbf{p}', \mathbf{q}')$ that preserves this form is called a *canonical transformation;* it preserves the form (3.1) of the differential equations in $(\mathbf{p}', \mathbf{q}')$ coordinates. A key feature of Hamiltonian flows Φ_t defined by $\Phi(\mathbf{p}(0), \mathbf{q}(0))) = (\mathbf{p}(t) \ , \mathbf{q}(t))$ is that they *preserve area,* as given by the symplectic form ω. This implies *Liouville's theorem:* Hamiltonian flows preserve volume in phase space. In consequence one may consider *area-preserving maps* as discrete analogues of Hamiltonian systems.

The symplectic form induces a Lie algebra structure on smooth functions $K(\mathbf{p}, \mathbf{q})$ on the manifold, with the operation of *Poisson bracket,* given by

$$\{K_1, K_2\} = \sum_{i=1}^{n} \left[\frac{\partial K_1}{\partial \mathbf{p}_i} \frac{\partial K_2}{\partial \mathbf{q}_i} - \frac{\partial K_2}{\partial \mathbf{p}_i} \frac{\partial K_1}{\partial \mathbf{q}_i} \right] . \tag{3.3}$$

A function K is said to *commute with H* or be *in involution with H* if

$$\{K, H\} \equiv 0 \ . \tag{3.4}$$

In that case $K(\mathbf{p}\ \mathbf{q})$ is also a conserved quantity of the system (3.1); such K are also sometimes called *integrals of the motion.*

The number of *degrees of freedom* for a time-independent Hamiltonian system $H(\mathbf{p}, \mathbf{q})$ is *n*, and that of a time-dependent Hamiltonian system $H(\mathbf{p}\ \mathbf{q}, t)$ which is periodic in *t*,

$$H(\mathbf{p}, \mathbf{q}, t+1) \equiv H(\mathbf{p}, \mathbf{q}, t)$$

is by convention $n + \frac{1}{2}$.

A Hamiltonian system is said to be *globally completely integrable* if there exist *n* independent commuting integrals of the motion:

$$\{K_i, H\} = 0 , \qquad 1 \le i, j \le n \tag{3.5}$$

with nonvanishing Jacobian everywhere in position space. (Usually one of the K_i is *H*). What is the effect of this? A solution curve of (3.1) satisfies *n* independent conservation laws

$$K_i(\mathbf{p}, \mathbf{q}) = 0 \qquad 1 \le i \le n ,$$

which are all transverse, hence is constrained to an *n*-dimensional submanifold. In fact under a suitable canonical transformation it can be taken to be the **q**-variables, with $H(\mathbf{p}, \mathbf{q}) = H(\mathbf{p})$ and Hamilton's equations become

$$\begin{aligned} \frac{\partial \mathbf{p}}{\partial t} &= 0 \\ \frac{\partial \mathbf{q}}{\partial t} &= \frac{\partial H}{\partial \mathbf{p}} . \end{aligned}$$

Now a further canonical coordinate change gives, locally,

$$\begin{aligned} \frac{\partial p_i}{\partial t} &= 0 \\ \frac{\partial q_i}{\partial t} &= \theta_i p_i \end{aligned} \tag{3.6}$$

for constants $\{\theta_i : 1 \le i \le n\}$.

It is often a difficult problem to find action-angle coordinates for a completely integrable system, that put it into the form (3.6), however.

Consequence is that the flow lines of a completely integrable system fit together smoothly. This contrasts with non-integrable systems, which can exhibit ‘‘chaotic’’ motion, see §4.

Theorem 3.1. *(Arnol'd) A globally completely integrable Hamiltonian system has phase space manifold diffeomorphic to* $\mathsf{R}^n \times \Sigma$ *with* $\Sigma \cong \mathsf{R}^k \times \mathsf{T}^{n-k}$*, for some* $0 \le k \le n$.

The most studied case of this is $\Sigma \cong \mathsf{T}^n$; the variables $\mathbf{q}$ are then all periodic, and are called *action angle variables.* See Arnol'd, Kozlov and Neishtadt (1988), p. 110 for references and a proof sketch.

There is also a form of Theorem 3.1 which applies to *locally completely integrable systems*, which are ones in for which there is a set of n independent commuting integrals $\{K_i : 1 \le i \le n\}$ but these have vanishing Jacobian at some points of phase space. A point $\mathbf{x}$ is *nondegenerate* if the Jacobian of the set $\{K_i : 1 \le i \le n\}$ does not vanish at $\mathbf{x}$. The local version is: *The closure of the orbit of an integrable system passing through a nondegenerate point in phase space is diffeomorphic to* $\mathsf{R}^k \times \mathsf{T}^{n-k}$ *for some* $0 \le k \le n$. Here k depends on the point. For example, the frictionless pendulum is locally completely integrable, and its orbits in phase space are T or R according as the pendulum swings back and forth or goes ‘‘over the top.’’

All Hamiltonian systems with one degree of freedom are locally completely integrable. For two or more degrees of freedom integrable Hamiltonian systems are very rare; generic Hamiltonian systems are not completely integrable (see Arnol'd et al (1988).) Completely integrable systems appear to play some special role in quantization, in the theory of ‘‘geometric quantization.’’

We describe four examples:

Example 1. *(Simple Harmonic Oscillator)* The Hamiltonian is:

$$H = \sum_{i=1}^{n} \frac{1}{2} p_i^2 + \sum_{i=1}^{n} \frac{1}{2} w_i q_i^2 ,$$

where the first term represents kinetic energy and the second potential energy. A set of Hamiltonians in involution is

$$F_i = \frac{1}{2} p_i^2 + \frac{1}{2} w_i q_i^2 , \quad 1 \le i \le n .$$

This Hamiltonian describes a set of n non-interacting particles. For each particle separately one has

$$\begin{aligned} \frac{dp_i}{dt} &= - w_i q_i , \\ \frac{dq_i}{dt} &= p_i . \end{aligned}$$

which is simple harmonic motion, so the phase space is T^n.

More generally, since all 1-dimensional Hamiltonian systems are locally completely integrable, except at isolated points, any system

$$H = \sum_{i=1}^{n} H_i(p_i, q_i)$$

is locally completely integrable.

Example 2. *(Two-body problem)* This system in R^3 has $n = 6$ degrees of freedom; these are accounted for by conservation of energy, of momentum, and of angular momentum $(1 + 3 + 2)$.

This relevance of this integrable system in celestial mechanics is that some N-body problems can be viewed as perturbations of a 2-body problem if two masses are very large relative to the others, e.g. the sun and Jupiter. In the Solar system example (4) in the introduction, the variable ω_{Gap} appears in terms of action angle coordinates (3.6) for the two-body problem (Saturn, Mimas).

Example 3. *(Geodesic Flow on a Riemannian Manifold)* The simplest case is geodesic flow on the torus $\mathsf{T}^n = \mathsf{R}^n/\mathsf{Z}^n$. Here T^n is a flat manifold with the induce metric from R^n. In describing the behavior of geodesics one encounters the problem of multidimensional Diophantine approximations where, the *initial conditions* encode the numbers to be approximated. Geometrically the number of Z-independence conditions for $(\alpha_1, \ldots, \alpha_n)$ determines the dimension of the subtorus which is the closure of the geodesic.

Example 4. (Toda Lattice). The Toda lattice is an infinite chain of particles having exponential nearest neighbor interactions, given by

$$\ddot{q}_n = e^{q_{n+1} - q_n} - e^{q_n - q_{n-1}} .$$

The Hamiltonian form is

$$H = \sum_n \frac{1}{2} p_n^2 + \sum_n (e^{q_n - q_{n-1}} - 1) .$$

It is infinite-dimensional, but has a finite-dimensional analogues obtained by either imposing periodic boundary conditions or by fixing endpoints, see Moser (1975). Flaschka (1974) observed that one of the finite-dimensional analogues can be put into the Lax-pair form

$$\dot{\mathbf{L}}(t) = [\mathbf{B}(t) , \mathbf{L}(t)]$$

where $\mathbf{L}(t)$ is a symmetric tridiagonal matrix, and

$$\mathbf{L}(t) = \begin{bmatrix} a_1(t) & b_1(t) & & 0 \\ b_1(t) & a_2(t) & & 0 \\ & & & \\ & & & b_{n-1}(t) \\ 0 & 0 & b_{n-1}(t) & a_n(t) \end{bmatrix} , \quad B(t) = \begin{bmatrix} 0 & -b_1(t) & & 0 \\ b_1(t) & 0 & & 0 \\ & & & \\ & & & -b_{n-1}(t) \\ 0 & 0 & b_{n-1}(t) & 0 \end{bmatrix} .$$

where

$$a_k(t) = \frac{1}{2} e^{\frac{1}{2}(q_k - q_{k-1})}$$

$$b_k(t) = -\frac{1}{2} p_k = -\frac{1}{2} \frac{d}{dt}(q_k) \ .$$

The matrix $\mathbf{A}(t)$ is isospectral, and the coefficients of its characteristic polynomial form a full set of commuting Hamiltonians for the flow, making it completely integrable. Synes (1982) observed a close connection between this form of the Toda flow and the QR-algorithm for numerically finding the eigenvalues of a symmetric matrix.

Further examples of completely integrable systems are given by certain flows on the Jacobian varieties of algebraic curves, which naturally have the structure of (complex) tori. The solutions of such differential equations can often be written in terms of elliptic functions.

As a final topic, one may relax the usual notion of completely integrable system by not requiring that all conserved quantities commute with each other. Define a *G-completely integrable system* to be a set $\{K_1, \ldots, K_m\}$ of functions on a symplectic manifold $\}$ with $K_1 = K$ which form an *m*-dimensional Lie algebra under Poisson bracket, and whose Jacobian is of full-rank at each point of $\}$. It appears that an analogue of Theorem 3.1 holds, where the closure of each orbit has the structure like a simply connected Lie group modulo the action of a discrete subgroup, see Arnol'd, Koslov and Neishtadt (1988) p. 110.

I believe that the connection of such systems to number theory arises from the *discrete subgroup.* In the case of the torus $\mathsf{T}^n = \mathsf{R}^n / \mathsf{Z}^n$, the flow curves can sometimes be analyzed in terms of Diophantine approximation.

4. KAM Theory and Small Divisors

KAM theory concerns the behavior of the orbits of a ''nearly integrable'' dynamical system. A ''nearly integrable'' system is one with Hamiltonian

$$H_\varepsilon = H_0(\mathbf{p}, \mathbf{q}) + \varepsilon H_1(\mathbf{p}, \mathbf{q}, \varepsilon) , \tag{4.1}$$

where H_0 is completely integrable and ε is a small parameter. We will suppose that all orbits of H_0 have the form T^n. The problem is approached by perturbation theory in the parameter ε. Under what conditions does an orbit for $\varepsilon = 0$ smoothly deform into an orbit for $\varepsilon = \varepsilon_0$? Such problems are directly motivated by celestial mechanics, in studying the question of the stability of the solar system.

We assume that action-angle variables $(\mathbf{I}, \boldsymbol{\theta})$ are given, where $\mathbf{I} = (I_1, \ldots, I_n) \subset B \subseteq \mathsf{R}^n$ with B compact and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ (mod 2π) $\in \mathsf{T}^n$, so that the Hamiltonian $H_0 = H_0(\mathbf{I})$ depends only on the action variables. Hamilton's equations are

$$\frac{d\mathbf{I}}{dt} = \mathbf{0} , \quad \frac{d\boldsymbol{\theta}}{dt} = \frac{\partial H_0}{\partial \mathbf{I}} . \tag{4.2a}$$

The perturbation $\varepsilon H_1(\mathbf{I}, \boldsymbol{\theta}, \varepsilon 0$ yields

$$\frac{d\mathbf{I}}{dt} = -\varepsilon \frac{\partial H_1}{\partial \boldsymbol{\theta}} , \quad \frac{d\boldsymbol{\theta}}{dt} = \frac{\partial H_0}{\partial \mathbf{I}} + \varepsilon \frac{\partial H_1}{\partial \mathbf{I}} . \tag{4.2b}$$

In the perturbed system the quantities $(I_1, \ldots, I_n)$ are no longer exactly conserved, and one wishes to know how they change over time, i.e. do they remain bounded? Here the rotational quantities $\boldsymbol{\theta}$ are viewed as *fast variables* and the quantities $\mathbf{I}$ as *slow variables,* which may gradually evolve over time. The *frequencies* of the unperturbed system are

$$\omega_i = \frac{\partial H_0}{\partial I_i}\Big|_{\mathbf{I}} . \tag{4.3}$$

A system is *nonresonant* if $(\omega_1, \ldots, \omega_n)$ are linearly independent over Z.

A much-used approach is the method of *averaging,* which approximates the system (4.2b) averaged over the fast variables.

The detailed approach to the averaging method is to construct suitable changes of coordinates that formally eliminate, to some degree of accuracy, the fast variables. Such coordinate changes started with Lindstedt, with contributions by Delaunay, Poincaré, and many others. One seeks $(\mathbf{I}, \boldsymbol{\theta}) \to (\mathbf{J}, \psi)$ expressed as a formal series

$$\begin{aligned} \mathbf{I} &= \mathbf{J} + \varepsilon \mathbf{u}_1(\mathbf{J}, \psi) + \varepsilon^2 \mathbf{u}_2(\mathbf{J}, \psi) + \ldots \\ \boldsymbol{\theta} &= \boldsymbol{\psi} + \varepsilon \mathbf{v}_1(\mathbf{J}, \psi) + \varepsilon^2 \mathbf{v}_2(\mathbf{J}, \psi) + \ldots \end{aligned} \tag{4.4}$$

where the functions $\mathbf{u}_i$ and $\mathbf{v}_i$ are periodic (mod 2π) in ψ. In the new coordinate system one wants independence of $\boldsymbol{\psi}$ variables:

$$\begin{aligned} \frac{d\mathbf{J}}{dt} &= \varepsilon \mathbf{F}_0(\mathbf{J}) + \varepsilon^2 \mathbf{F}_1(\mathbf{J}) + \ldots \\ \frac{d\boldsymbol{\psi}}{dt} &= \boldsymbol{\omega}(\mathbf{J}) + \varepsilon \mathbf{G}_0(\mathbf{J}) + \ldots \end{aligned} \tag{4.5}$$

given on 2π-periodic function $u(\mathbf{J}, \boldsymbol{\psi})$, expanding it in Fourier series gives

$$u(\mathbf{J}, \boldsymbol{\psi}) = f_0(\mathbf{J}) + \sum_{\substack{\mathbf{k} \in Z^n \\ \mathbf{k} \neq \mathbf{0}}} h_k(\mathbf{J}) \exp(i \langle \mathbf{k}, \boldsymbol{\psi} \rangle) \tag{4.6}$$

Assuming that H_1 is sufficiently smooth function, one can expand it in a power series in ε, then expand each term in Fourier series (4.6), and recursively solve (4.5) to determine explicit expressions for $\mathbf{u}_i$, $\mathbf{v}_i$ in (4.4) up to order r, say. In the resulting expressions the quantities $\langle \mathbf{k}, \boldsymbol{\omega} \rangle^{-1}$ appear, where

$$\langle \mathbf{k}, \boldsymbol{\omega} \rangle = k_1 \omega_1 + \ldots + k_n \omega_n ,$$

where $\mathbf{k} \in Z^n$, and each $|k_i| \leq r$. The occurrence of values $\langle \mathbf{k}, \boldsymbol{\omega} \rangle$ close to 0, is called ‘‘small denominators’’, and may make the formal series (4.4) diverge. Lindstedt’s series usually civerges.

Kolmogorov (1956) and Arnol’d (1963) obtain a modified recursion to get transformations (4.4) which exhibits a ‘quadratic convergence’ in that after steps all terms ε^k with $k < 2^r$ are eliminated. The influence of the small denominators can be controlled for if ω satisfies a bound like

$$|\langle \mathbf{k}, \boldsymbol{\omega} \rangle| > c \|\mathbf{k}\|^{-\nu}$$

holds for all nonzero $\mathbf{k} \in \mathbb{Z}^n$, and $\nu > n - 1$. We say that a Hamiltonian system H_0 is *nondegenerate* at energies I if $\det \left[\frac{\partial^2 H_0}{\partial \mathbf{I}^2} \right] \neq 0$.

Theorem 5.1. *(Kolmogorov's theorem) If in (4.2) the unperturbed system $H_0(\mathbf{I})$ is nondegenerate, then for a sufficiently small Hamiltonian perturbation $\varepsilon H_1(\mathbf{I}, \boldsymbol{\theta}, \varepsilon)$ most nonresonant invariant tori do not vanish but are slightly deformed, so that the phase space of the perturbed system contains invariant tori, which contain dense trajectories which have n independent frequencies of motion. The measure of the complement of the set of invariant tori approaches zero at least as fast as $\sqrt{\varepsilon}$, as $\varepsilon \to 0$.*

Kolnogorov stated this result when H_0 and H_1 are analytic functions, but it has been extended to apply when H_1 is only C^r with $r > 2n$ by Pöschel. The $\sqrt{\varepsilon}$ result on the complementary set is due to later authors, see Arnol'd et al. (1988), Chapter 5. In particular on these invariant tori the variables **I** remain bounded for all time.

KAM-type phenomona can also be studied in the simpler discrete setting of *area-preserving maps*. A much-studied example is the *standard map* or *Chirikov-Taylor map* $S(r, \theta) = (r', \theta')$ on $\mathbb{T}^2$ given by

$$\begin{aligned} r^1 &= r - \frac{K}{2\pi} \sin(2\pi\theta) \quad (\mathrm{mod}\ 1) \\ \theta^1 &= \theta + r' \ (\mathrm{mod}\ 1)\ . \end{aligned} \tag{4.7}$$

Here K is a coupling parameter. For $0 \leq K < 1$ this is an *twist map* since $\frac{\partial \theta'}{\partial r} > 0$, see Mather (1982). For $K = 0$ the closure of the orbits of S consists of *invariant circles* of rotation number θ

for all irrational θ in [0,1], see Figure 4.1. *(Rotation number* is defined in (6.3);) the irrationality of θ is the nonresonance condition here.) As K increases invariant circles are destroyed: At the instant of destruction the closure of orbits becomes a perfect set, like a Cantor sets. Such sets are sometimes called *cantori.* For larger values of K there exist orbits whose closure has positive measure.

The problem of how long an invariant circle of a given rotation number θ exists has been extensively treated. An analogue of KAM theory has been applied to establish that at least one invariant circle exists for $0 \le K < \frac{1}{30}$. The most difficult rotation number to destroy appears to be $\theta = \frac{\sqrt{5} - 1}{2}$, apparently since it is the hardest number in [0,1] to approximate by rationals.

A heuristic method to estimate when circles break up was developed by Greene (1979). The idea is that invariant circles are broken by a sequence of stable resonances nearby. The linear stability of a periodic orbit of a map S of period q is measured by its *residue*

$$R = \frac{1}{4}(1 - Tr(D(S^{(q)})) \ . \tag{4.8}$$

The Poincare-Birkhoff twist theorem says that for an area-preserving twist map, each rational $\frac{p}{q}$ has least two periodic orbits of rotation number $\frac{p}{q}$, one with residue greater than zero, one with residue less than zero. (They are critical points of an associated actions) Generically there are two, having residues $R^{-}_{p/q} \le 0 \le R^{+}_{p/q}$, say. Breene observed empirically that for orbits of period q_n, the n-th ordinary continued fraction convergent to θ, if, as $n \to \infty$:

(i) $R^{\pm}_{p_n/q_n} \to 0$ *then a smooth invariant circle exists.*

(ii) $R^{\pm}_{p_n/q_n} \to \pm\infty$ *then no invariant circle exists.*

For $\theta = \frac{\sqrt{5} - 1}{2}$, computer experiments using this heuristic indicate that the critical value K_∞

at which the invariant circle of rotation number disappears is K_∞ *dot* = .971635406, see Figure 4.2. Above this value of K there are apparently no invariant circles.

5. Iteration of Analytic Functions and Small Divisors

In addition to KAM theory, small divisors problems arise in studying normal forms for the iteration of analytic function around a fixed point. This situation is analogous to a Hamiltonian system with one degree of freedom. Consider an analytic function $F(z)$ having a fixed point at 0, so that

$$F(z) \; = \; a_1 z \, + \, a_2 z^2 \, + \, a_3 z^3 \, + \ldots \; . \tag{5.1}$$

What is the behavior of $F(z)$ under iteration in a neighborhood of $z \; = \; 0$? If $|a_1| \; < \; 1$ then, for small z, $|F(z)| \; < \; |z|$ and $F^{(k)}(Z_0) \; \to \; 0$ as $k \to \infty$, and 0 is an *attracting fixed point.* If $|a_1| \; > \; 1$ then $|F(z)| \; > \; (z)$ for small z and 0 is a *repelling fixed point.* The interesting case is $|a_1| \; = \; 1$, sometimes called an *indifferent fixed point.* Here the higher order terms in (2.1) become relevant. Let $a_1 \; = \; e^{2\pi i \theta}$. The simplest case is the *pure rotation*

$$F_\theta(z) \; = \; e^{2\pi i \theta} z \; . \tag{5.2}$$

The behavior of F_θ under iteration is different according as θ is rational or irrational. In both cases circles $C_r \; = \; \{z : |z| \; = \; r\}$ are invariant sets. However if $\theta \; = \; \frac{p}{q}$ is rational then $F^{(q)}(z) \; \equiv \; z$ and every point lies in a finite orbit, while if θ is irrational, then the orbit of every point $z \in \; C_r$ is dense in C_r, in fact uniformly distributed (Weyl). What happens for general

$$F(z) \; = \; e^{2\pi i \theta} z \, + \, a_2 z^2 \, + \, a_3 z^3 \, + \ldots \quad ? \tag{5.3}$$

For some such F on analytic change of variable

$$\phi(z) \; = \; z \, + \, b_2 z^2 \, + \, b_3 z^3 \, + \ldots \tag{5.4}$$

conjugates $F(z)$ to a rotation, i.e.

$$\phi^{-1}(F(\phi(z))) = e^{i\theta} z \ . \qquad (5.5)$$

When this occurs, the qualitative behavior of iterates of $F(z)$ is the same as the rotation $F_\theta(z)$; there are ‘‘invariant circles,’’ which are analytic curves.

The equation (5.5) is called the *Schröder functional equation* after Schröder (1871). Schröder came to this problem from studying general iterative schemes to find solutions of equations, cf. Schröder (1870). Newton’s method is a special case, of an iterative scheme. This is interesting since Newton’s method proves important in solving the Schröder functional equation itself.

In the case of rational $\theta = \frac{p}{q}$ the functional equation (5.5) cannot necessarily be solved, even formally, while for any irrational θ it always has a formal power series solution (5.4) where the coefficients b_i are found recursively:

$$\begin{aligned} b_2 &= \\ b_3 &= \\ b_4 &= \end{aligned}$$

However the resulting power series sometimes has radius of convergence zero. The problem in determining the radius of convergence arises from the ‘‘small denominators’’ $p\theta - q$ appearing the formal coefficients b_i. When $\phi(z)$ has a positive radius of convergence one says that $F(z)$ is *analytically conjugate to a rotation.* For certain values of θ every convergent series (5.3) is analytically conjugate to a rotation, while for other values of θ this is not so. Siegel (1942) showed that the condition: There is a constant $\beta > 0$ such that

$$|\theta - \frac{p}{q}| \geq \frac{1}{q^\beta} \quad \text{for all } (p, q) \in \mathbb{Z}^2 \ , \ q \neq 0 \ , \qquad (5.6)$$

guarantees that all convergent series (5.3) are analytically conjugate to $e^{2\pi i\theta} z$. (Such a condition is often called a *Diophantine condition* in the literature.) Siegel also gave examples of an irrational θ where there exists some $F(z)$ not analytically conjugate to a rotation.

Brjuno (1979) gave a stronger sufficient condition on θ.

Theorem 5.1. *(Brjuno) Let* θ *have the ordinary continued fraction expansion* $[0, a_1, a_2, a_3, \ldots]$ *with convergents* $\frac{p_n}{q_n}$. *All functions* $F(z) = e^{2\pi i\theta} z + a_2 z^2 + \ldots$ *analytic in a neighborhood of 0, are analytically conjugate to the rotation* $e^{i\theta} z$ *, provided that* θ *satisfies*

$$\sum_{i=1}^{\infty} q_n^{-1} \log q_{n+1} < \infty \ . \tag{5.7}$$

For such θ Brjuno showed that a sequence of successive polynomial approximations $\phi_n(z)$ converge uniformly to a solution $\phi(z)$ to (5.5) near 0, thus guaranteeing analyticity of $\phi(z)$. The $\phi_n(z)$ are constructed by a Newton-method like process.

Recently J.-C. Yoccoz (1988) showed that Brjuno's condition is also necessary, thus completely solving this problem.

How do the iterates of a point near 0 behave for $F(z)$ when it is not analytically conjugate to a rotation? I. N. Baker and P. J. Rippon (1984) show that there exist an irrational θ such that the entire function

$$F(z) = e^{i\theta}(e^z - 1)$$

where the iterates of some point z_0 are *dense* in $\mathbb{C}$; such z_0 can clearly be found in any neighborhood of 0.

Further information can be found in Martinet (1981) and Herman (1987).

6. Mode-Locking

The *mode-locking phenomenon* is that weakly coupled oscillators tend to synchronize their motion, e.g. their apparent modes of oscillation acquire $\mathbb{Z}$-linear dependencies. This arises when there are two or more periodic motions with competing frequencies with some sort of a coupling force between them. The motion of the three moons of Jupiter (given in §1) is an example of mode-locking. Mode-locking occurs also in some simple Josephson junction circuits, see Bak, Bohr and Høgh-Jensen (1985).

Typical motions of such dissipative systems may include a mixture of quasi-periodic motion, mode-locked motion, and chaotic motion, depending on the initial frequencies and the degree of nonlinearity.

An extensively studied simple model is the *sine circle map*

$$x_{n+1} = f_{K,\,\Omega}(x_n) \tag{6.1}$$

where

$$f_{K,\,\Omega}(x) = x + \Omega - \frac{K}{2\pi}\sin(2\pi x) \ , \tag{6.2}$$

where $K \geq 0$ is a parameter; this map was introduced in Arnol'd (1965). Here x_n (mod 1) may be thought of as an angle, measuring motion around the circle $\mathsf{T} = \mathsf{R}/\mathsf{Z}$, (with units chosen so that a full angle is 1 rather than 2π.) The map (6.1) is a lift of it to the universal cover R. For $K = 0$, the map is

$$x_{n+1} \equiv x_n + \Omega \pmod 1$$

which is a pure rotation. The two competing frequencies of the motion are the *bare rotation rate* Ω and the rotation rate 1 associated to the driving term $\sin 2\pi x$, and K is the coupling constant. For $0 \leq K < 1$ the function $f_{K,\,\Omega} : \mathsf{R} \to \mathsf{R}$ is monotone increasing, with positive derivative everywhere. The *critical value* is $K = 1$, where $f_{1,\Omega}$ is still monotone increasing but has inflection points at all $x \in \mathsf{Z}$.

For $0 \leq K \leq 1$ the map $f_{K,\,\Omega}$ has a well-defined *rotation number*

$$r(\Omega, K) = \lim_{n \to \infty} \frac{1}{n}(x_n - x_0) \ , \tag{6.3}$$

which measures the average speed of a point around the circle T. This limit exists and is independent of the starting point x_0, but depends on Ω and K, by a result of Herman (1977), because $f_{K,\Omega}$ is a monotone twist map. The mode-locking phenomenon here is that for $0 < K \leq 1$ the rotation number $r(\Omega,K)$ takes rational values on intervals, while a given irrational value is taken at a point. (See Figure 6.1) As the nonlinearity parameter increases, the intervals become larger, and as indicated in Fig. 6.1 they assume the shape of tongues, and are sometimes called ‘‘Arnol’d tongues’’. A rational rotation

$$r(\Omega,K) = p/q$$

is a $\mathbb{Z}$-linear dependency between the rotation number r and the driving frequency 1.

Mode-locking occurs because the monotone increasing nature of the function (6.2) means that whenever f_K has a periodic point (mod 1), i.e.

$$f_{K,\Omega}^{(p)}(x_0) = x_0 + q \tag{6.4}$$

then by (6.3) it must have

$$r(\Omega,K) = \frac{p}{q} .$$

One sees by plotting $f_{K,\Omega}^{(p)}(x)$ (mod 1) versus x (mod 1), that since the graph changes continuously (on the torus) as Ω moves, there is an interval of values of Ω where (6.4) holds.

For $0 < K < 1$ there is thus a mixture consisting of mode-locked motion for Ω in the locked intervals, and quasi-periodic motion at an irrational rotation rate $r(\Omega, K)$ otherwise. What happens at $K = 1$? Høgh-Jensen, Bak and Bohr (1983) and also Ostlund et al (1983) studied a devil’s staircase structure of the rotation rate $r(\Omega, 1)$ as Ω increases, see Figure 6.2. They observed an apparent self-similar structure in this staircase, and conjectured based on simulations, that the mode-locked intervals have Lebesgue measure one, and that the set of Ω such that $r(\Omega,1)$ is irrational had Hausdorff dimension strictly between zero and one. They numerically estimated this Hausdorff dimension to be .8700 ± .0002 by a ‘‘renormalization’’ method.

Swiatek (1988) rigorously proved the first part of their conjectures.

Theorem 6.1. *(Swiatek) For the sine circle map with* $K = 1$,

$$x_{n+1} = x_n + \Omega + \frac{1}{2\pi}\sin(2\pi x_n)$$

the set of Ω *having rational rotation number* $r(\Omega, 1)$ *has Lebesgue measure one.*

Swiatek's proof uses projective invariants (cross ratios) to keep track of scaling.

Several authors have introduced ''renormalization schemes'' to estimate universal parameters associated to the critical value $K = 1$ such as the Hausdorff dimension of the unlocked set, cf. Cvitanovic et al (1985), Ostlund and Kim (1985), Procaccia et al (1987), Feigenbaum (1988), Sinai and Khanin (1988), Kim and Ostlund (1989). These schemes look at the behavior of the locked intervals at the *n*-th level of the Farey tree $\wedge_n$, and examine asymptotics as $n \to \infty$. The Farey shift map U of §2 appears explicitly in Feigenbaum (1988). One can consider these schemes as assigning a measure to the nodes of $\wedge_n$, which leads to a limiting singular measure μ_∞ on [0,1] which encodes ''universal'' information. It appears likely that this measure is an invariant measure for the Farey shift map.

What happens for $K > 1$? In that case the maps $f_{K,\Omega}$ are no longer monotone, and iterates of (5.1) can exhibit chaotic motion, in the following sense. Define

$$r_+(\Omega, K) = \sup_{x_0}\{\limsup_{n \to \infty} \frac{1}{n}(x_n - x_0)\} \tag{6.5a}$$

$$r_-(\Omega, K) = \inf_{x_0}\{\liminf_{n \to \infty} \frac{1}{n}(x_n - x_0)\} \ . \tag{6.5b}$$

There exist values with $r_+(\Omega, K) > r_-(\Omega, K)$, i.e. maps have *rotation intervals* rather than rotation numbers. There are still self-similar structures for $K > 1$, see Jakobson (1988).

The width of Arnol'd tongues of rotation number $\frac{p}{q}$ near $K = 0$ decreases like K^q, see Arnol'd (1983), Jonker (1990). This is related to ''forbidden zones'' in solid state physics.

An important open problem concerns what happens for mode-locking in systems with 3 or more competing frequencies? Is multidimensional Diophantine approximation relevant to analysis of this question? An initial discussion appears in Kim and Ostlund (1985).

A general discussion of mode-locking occurs in Schroeder (1990), p. 171 ff.

7. Quasicrystals

References

§0. General references

[1] T. Bedford, M. Keane, C. Series, Eds. (1991), *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces,* Oxford Univ. Press, Oxford.

[2] M. Dodson, J. M. Vickers, Ed. (1989), *Number Theory and Dynamical Systems,* Cambridge U. Press, Cambridge.

[3] M. Jaric, Ed. (1989), *Introduction to the Mathematics of Quasicrystals,* Academic Press.

[4] J-M. Luck, P. Moussa, M. Waldschmidt, Eds. (1990), *Number Theory and Physics,* Proc. les Houches, France 1989, Springer-Verlag, New York.

[5] M. Schroeder (1990) *Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise,* Freeman, New York.

§1. Introduction

[1] A. D. Brjuno (1970), Instability in a Hamiltonian system and the Distribution of Asteroids, Math USSR Sb. **12,** 271-312.

[2] C. R. Chapman, J. G. Williams, W. K. Hartman (1978), The Asteroids, Annual Rev. Astron. Astrophys **16,** 33-75.

[3] J. Koiller, J. M. Balthazar and T. Yokoyama (1987), Relaxation – Chaos Phenomena in Celestial Mechanics I. On Wisdom’s Model for the 3/1 Kirkwood Gap, Physica **26D,** 85-122.

[4] S. J. Peale (1976), Orbital resonances in the Solar System, Annual Rev. Astron. Astrophys. **14,** 215-246.

[5] J. Wisdom (1982), The origin of the Kirkwood gaps: a mapping for asteroidal motion near

the 3/7 acommensurability, Astron. J. **87,** 577-593.

§2. Continued Fractions

[1] P. Arnoux and A. Nogueira (1991), Mesures de Gauss pour des algorithmes de fractions continues multidimensionnelles, preprint (U. Paris VII)

[2] A. Ya. Khinchin (1930), Metrische Kettenbruchprobleme, Compositio Math. **1,** 361-382.

[3] A. Ya Khinchin (1964), *Continued Fractions,* U. of Chicago Press.

[4] J. C. Lagarias (1991), The Farey Shift, in preparation.

[5] D. Meyer (1990), Continued Fractions and Related Transformations, in: *Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces* (T. Bedford, M. Keane, C. Series, Eds.), Oxford U. Press, 175-222.

[6] M. Pollicott (1986), Distribution of Closed Geodesics on the Modular Surface and Continued Fractions, Bull. Math. Soc. France **114,** 431-446.

[7] I. Richards (1981), Continued Fractions Without Tears, Math. Mag. **54,** 163-171.

[8] C. Series (1982), Non-Euclidean Geometry, Continued Fractions, and Ergodic Theory Math. Intelligencer **4,** 24-31.

[9] C. Series (1985a), The Geometry of Markoff Numbers, Math. Intelligencer **7,** 20-29.

[10] C. Series (1985b), The Modular Surface and Continued Fractions, J. London Math. Soc. **31,** 69-80.

§3. Integrable Hamiltonian Dynamical Systems

[1] V. I. Arnol’d, V. V. Kozlov, A. I. Neishtadt (1988), Mathematical Aspects of Classical and Celestial Mechanics, in: *Dynamical Systems III* (V. Arnol’d, Ed.), Springer-Verlag, London.

[2] P. A. M. Dirac (1950), Generalized Hamiltonian Dynamics, Canad. Math. J. **2,** 129-148.

[3] L. D. Faddeev and L. A. Takhtajan (1987), *Hamiltonian Methods in the Theory of Solitons,* Springer-Verlag, New York.

[4] H. Flaschka (1974), The Toda Lattice I, Phys. Rev. B **9,** 1924-1925.

[5] J. Moser (1975), Finitely many mass points on the line under the influence of an exponential potential – an integrable system, Lecture Notes in Physics No. 38, Springer-Verlag, 467-497.

[6] J. Moser (1980), Various Aspects of Integrable Hamiltonian Systems, in: *Dynamical Systems, CIME Lectures,* Birkhäuser, Boston, 233-290.

[7] I. Percival (1986), Integrable and Nonintegrable Hamiltonian Systems, in: *Nonlinear Dynamics Aspects of Particle Accelerators,* Lecture Notes in Physics No. 247, Springer-Verlag, New York, 12-36.

[8] W. W. Symes (1982), The QR-algorithm and scattering for the finite non-periodic Toda lattice, Physica D **4,** 275-280.

§4. **KAM Theory**

[1] V. I. Arnol'd (1963), Small Denominators and Problems of Stability of Motion in Classical and Celestial Mechanics, Russian Math Surveys **18,** 85-191.

[2] J. Bellissard (1985), Small Divisors in Quantum Mechanics, *Proc. NATO ASI Workshop: Chaotic Behavior in Quantum Systems* (G. Casati, Ed.), Plenum, 11-35.

[3] H. Rüssmann (1989), Nondegeneracy in the perturbation theory of dynamical systems, in: *Number Theory and Dynamical Systems* (M. M. Dodson and J. M. Vickers, Ed.), Cambridge U. Press, 5-19.

§5. **Iteration of Analytic Functions and Small Divisors**

[1] I. N. Baker and P. J. Rippon (1984) Iteration of Exponential Functions, Ann. Acad. Sci. Fennicae ser A1 **9,** 49-77.

[2] A. D. Brjuno (1971), Analytic form of Differential Equations, Trans. Moscow Math. Soc. **25,** 131-288.

[3] M. R. Herman (1987), Recent results and some open questions on Siegel’s linearization theorem of germs of complex analytic diffeomorphisms near a fixed point, in: *VIIth International Congress on Mathematical Physics* (Marseille 1986), World Scientific, 138-184.

[4] J. Martinet (1981), Normalisation des champs le vecteurs, d’apres A. D. Brjuno, Séminaire Bourlaki Exp. 582, Lecture Notes in Math. **901,** Springer-Verlag, 55-70.

[5] E. Schröder (1870), Ueber unendliche viele Algorithmen zur Aüflosung der Gleichungen, Math. Annalen **2,** 317-365.

[6] E. Schröder (1871), Ueber iterirte Functionen, Math. Annalen **3,** 296-322.

[7] C. L. Siegel (1942), Iteration of analytic functions, Ann. of Math. **43,** 807-812.

[8] J.-C. Yoccoz (1988), Linéarisation des germes de diffeomorphisms holomorphes de $(\mathbb{C}, O)$, C. R. Acad. Sci. Paris. Ser I. **306,** 55-58.

§6. Mode-locking

[1] V. I. Arnol’d (1965), Small Denominators I. Mappings of the Circumference onto Itself, AMS Translations, Series 2, Vol. **46,** 213-284.

[2] V. I. Arnol’d (1983), Remarks on perturbation theory for problems of Mathieu type, Russian Math. Surveys **38,** 215-233.

[3] P. Bak, T. Bohr, M. Høgh-Jensen (1985), Mode-locking and the Transition to Chaos in Dissipative Systems, Physica Scripta T **9,** 50-58.

[4] P. Cvitanović, M. H. Jensen, L. P. Kadanoff, I. Procaccia (1985), Renormalization, Unstable Manifolds and the Fractal Structure of Mode Locking, Phys. Rev. Lett. **55,** 343-346.

[5] P. Cvitanović, B. Shraiman, B. Söderberg (1985), Scaling Laws for Mode Locking in Circle Maps, Phys. Scripta **32,** 263-270.

[6] M. J. Feigenbaum (1988), Presentation Functions, Fixed Points, and a Theory of Scaling Function Dynamics, J. Stat. Phys. **52,** 527-569.

[7] M. Høgh-Jensen, P. Bak, T. Bohr (1983), Complete Devil's Staircase, Fractal Dimension and Universality of Mode-Locking Structure in the Circle Map, Phys. Rev. Lett. **50,** 1637-1639.

[8] M. R. Herman (1977), Mesure de Lebesgue et nombre de rotation, in: Geometry and Topology Lecture Notes in Math. **597,** Springer-Verlag, 271-

[9] M. V. Jakobson (1988), Onset of Stochasticity for Some One-Dimensional Systems, Physica D **33,** 157-164.

[10] L. Jonker (1990), The Scaling of Arnol'd Tongues at Degenerate Fold Bifurcations, Physica D **41,** 275-281.

[11] S. Kim and S. Ostlund (1985), Renormalization of Mappings of the Two-Torus, Phys. Rev. Lett. **55,** 1165-1168.

[12] S. Kim and S. Ostlund (1989), Universal Scaling in Circle Maps, Physica D **39,** 365-392.

[13] S. Ostlund, D. Rand, J. Sethna, E. Siggia (1983), Universal Properties of the Transition from Quasi-Periodicity to Chaos in Dissipative Systems, Physica **80,** 303-342.

[14] S. Ostlund and S.-H. Kim (1985), Renormalization of Quasiperiodic Mappings, Physica Scripta **T9,** 193-198.

[15] I. Procaccia, S. Thomae and C. Tresser (1987), First-return maps as a unified renormalization scheme for dynamical systems, Phys. Rev. A **35,** 1884-1900.

[16] B. Simon (1982), The Rotation Number for Almost Periodic Potentials, Comm. Math. Phys. **84,** 403-438.

[17] Ya. G. Sinai, K. M. Khanin (1988), Renormalization Group Method in the Theory of Dynamical Systems, Int. J. Mod. Phys. **B 2,** 147-165.

[18] G. Swiatek (1988), Rational Rotation Numbers for Maps of the Circle, Comm. Math. Phys. **119,** 109-128.

§7. **Quasicrystals**

[1] J. P. Allouche and M. Mendes-France (1986), Finite Automata and Zero Temperature Quasicrystal Ising Chain, J. de Physique, Colloque **C3**, Tome 47, C3-63 – C3-73.

[2] A. E. Burkov (1987), One-dimensional model of the Quasicrystalline Alloy, J. Stat. Phys. **47,** 409-438.

[3] M. Jarić, Ed. (1989), *Introduction to the Mathematics of Quasicrystals,* Academic Press.

[4] M. Mendes-France, A von der Poorten (1981), Arithmetic and Analytic Theory of Paper Folding Sequences, Bull. Australian Math. Soc. **24,** 123-130.

[5] F. Queffélec (1987), *Substitution Dynamical Systems – Spectral Analysis,* Lecture Notes in Math. No. 1294, Springer-Verlag.

Number Theory and Dynamical Systems
(Preliminary Notes)

J. C. Lagarias

AT>T Bell Laboratories
Murray Hill, NJ 07974

(August 5, 1991)

ABSTRACT

This paper describes the occurrence of number-theoretic problems in dynamical systems. These include Hamiltonian dynamical systems, dissipative dynamical systems and discrete dynamical systems. Diophantine approximations and continued fractions play an important role.

Figure 1.1. Kirkwood Gaps [From Arnol'd (1990), p. 78]

Figure 1.2. Saturn's rings [From Arnol'd (1990), p. 79]

Figure 2.1. Gauss map

Figure 2.2. Farey tree

Figure 2.3. Farey shift map

Figure 2.4. Minkowski ?-function

Figure 4.1. Invariant circles for standard map $K = 0$

Figure 4.2. Some orbits of the standard map at $K = 0.971635406$. The golden circle orbit starts at 0.4 at left [From MacKay et al (1984), p. 67]

Figure 6.1. Devil's staircase – rotation numbers for sine circle map at $K = 1$. Enlarged box shows self-similar structure [From Høgh-Jensen, Bak and Bohr (1983), p. 1638]

Figure 6.2. Arnol'd tongues for $f(x) = x + a + \varepsilon \cos \pi x \pmod 1$

Figure 7.1. Diffraction patterns exhibiting icosahedral symmetry [From Schechtman et al. (1984), p. 1952]

Figure 7.2. Scanning tunneling microscope image of $AL_{65}Co_{20}Cu_{15}$ quasicrystal, in 2-D decagonal phase (Enhanced picture) [From Kortan et al. (1990), p. 201]

Figure 7.3. Pentagonal quasilattice model simulating Fig. 7.1 [From Kortan et al. (1990), p. 202]

Figure 7.4. Fibonacci quasicrystal