

INTRODUCTORY STATISTICAL MECHANICS FOR ELECTRON STORAGE RINGS*

JOHN M. JOWETT[†]

*Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305*

TABLE OF CONTENTS

1	Introduction	2
2	Statistical Mechanical Principles	5
2.1	Deterministic systems	6
2.2	Fluctuating systems	8
2.3	Distribution functions and the Liouville equation	10
2.4	Fokker-Planck equation	11
2.5	The recipe	13
2.6	Solving Fokker-Planck equations	15
2.7	Expectation values of dynamical variables	19
2.8	Stochastic calculus	21
3	Statistics of Incoherent Synchrotron Radiation	24
3.1	Classical picture	26
3.2	Quantum-mechanical picture	29
3.3	Synchrotron radiation as a stochastic process	30
3.4	Distribution of photon emission time	33
3.5	Mean of radiation power	35
3.6	Campbell's theorem	36
3.7	Representation of fluctuating power	39
3.8	Simulating photon emission	40
4	Electron Dynamics with Radiation Reaction	44
4.1	Coordinate system and Hamiltonian	44
4.2	Radiation reaction forces	47
4.3	Robinson's theorem and generalizations	50

*Work supported by the Department of Energy, contract DE-AC03-76SF00515.

[†]Permanent address and address when lectures were given: CERN, CH-1211 Geneva 23.

5	Closed Orbit and Normal Modes	54
5.1	Dispersion functions	54
5.2	Local synchrotron motion	58
5.3	Off-momentum closed orbit	63
5.4	Betatron motion	66
5.5	Coupled motion with radiation	68
6	Radiation Damping	68
6.1	Damping in longitudinal phase space	69
6.2	Damping of vertical betatron oscillations	72
6.3	Damping of radial betatron oscillations	74
6.4	Damping partition numbers and damping aperture	77
7	Quantum Fluctuations and Fokker-Planck Equations	79
7.1	Quantum fluctuations in longitudinal phase space	79
7.2	Stationary energy distribution	81
7.3	Equilibrium in the nonlinear RF bucket	82
7.4	Fokker-Planck equation in action-angle variables	84
7.5	Longitudinal quantum lifetime	87
7.6	Moment equation for action	89
7.7	A model of injection damping	90
7.8	Correlation functions of dynamical variables	93
7.9	Quantum fluctuations in radial betatron phase space	93
8	The Vlasov Equation	96
8.1	Many particle systems	96
8.2	Phase space density	97
8.3	Two-particle correlations	98
8.4	The Vlasov equation	99
8.5	Correlation effects and BBGKY hierarchy	100
9	References	102

ABSTRACT

These lectures introduce the beam dynamics of electron-positron storage rings with particular emphasis on the effects due to synchrotron radiation. They differ from most other introductions in their systematic use of the physical principles and mathematical techniques of the non-equilibrium statistical mechanics of fluctuating dynamical systems. A self-contained exposition of the necessary topics from this field is included.

Throughout the development, a Hamiltonian description of the effects of the externally applied fields is maintained in order to preserve the links with other lectures on beam dynamics and to show clearly the extent to which electron dynamics is non-Hamiltonian.

The statistical mechanical framework is extended to a discussion of the conceptual foundations of the treatment of collective effects through the Vlasov equation.

1. INTRODUCTION

The physics of particle accelerators is a microcosm of the whole of physics. An adequate understanding of the construction and control of particle accelerators and the behaviour of their beams requires the application of the deepest knowledge of many of the established branches of physics and technology alongside the more topical.

Even if we look only at the area of theoretical beam dynamics, we can find considerable scope for the application of many of the traditional subdivisions of theoretical physics. For example, the somewhat dusty subject of classical analytical dynamics has found a new lease of life in the challenging problems of single particle dynamics in synchrotrons and storage rings. So also with classical electromagnetism and optics. The first step in the synthesis of a description of macroscopic matter from the behaviour of its microscopic constituents is the realm of *statistical mechanics* and we can expect this, too, to play an important rôle in accelerator theory.

However the number of particles in a typical accelerated beam (10^{11} or so), although large, is still many orders of magnitude less than in a macroscopic sample of bulk matter (10^{22} or so); this leads us to expect some differences in the statistical description. At first sight it is not at all clear whether the concept of thermodynamic equilibrium has any relevance; if it does, then it surely only applies in the rest frame of the beam; if it does not, then we are on the less familiar ground of *non-equilibrium* statistical mechanics where we have fewer general principles and mathematical tools to guide us.

In fact an ultra-relativistic electron beam is a system very far from true thermodynamic equilibrium. Electrons continually receive energy from accelerating cavities and promptly lose it again by the phenomenon of *synchrotron radiation*. Since this energy loss has a random component due to the quantised nature of the radiation, the net effect is that of a random external force preventing the beam from attaining equilibrium. Each particle in the beam acts rather like a channel, absorbing ordered energy from the power source and releasing it again in small random bursts. It turns out that, when one averages over many (independent) particles, they are found to be distributed in a stationary non-equilibrium state which nevertheless resembles a genuine equilibrium in some of its particulars.

In proton beams synchrotron radiation effects are not noticeable up to energies of several TeV. However several other kinds of statistical phenomena are important. Fluctuations from thermodynamic equilibrium arise because of the finite number of particles and these can be detected macroscopically (more easily in the case of coasting beams) as the famous *Schottky noise*. This discovery led to the technique of stochastic cooling which has allowed all three of the

world's high energy hadron colliders to collide beams of protons and antiprotons with adequate intensity for high-energy physics experiments. Noise in the radio-frequency (RF) system can also introduce a degree of randomness in the equations of motion which may counteract the beneficial effects of stochastic cooling. A similar emittance blow-up arises due to scattering among the particles in the bunch. This is known as *intra-beam scattering* and may be important in lower energy electron machines.

The mathematical description of statistical phenomena in hadron beams does have some elements in common with electron ring theory. Nevertheless these lectures will concentrate on statistical phenomena in electron storage rings because the phenomena arising in proton machines are sufficiently different in nature to warrant a separate theoretical description. The reader will find thorough reviews of these topics and guidance through the literature in another School of this Series¹ and elsewhere^{2,4,5} but we point out some analogies with and differences from the statistical mechanics of hadron beams where appropriate.

I have tried to introduce the characteristic physics of e^+e^- rings in parallel with a tutorial introduction to a selection of useful techniques from the theory of stochastic processes and the statistical mechanics of non-equilibrium systems. These topics are normally covered in quite different contexts with a view to different applications and I hope it will be found useful to introduce them in this way. Thus, while many of the physical results given may be obtained more quickly by simpler means, they will be seen in a broader context as preparation for more complicated problems. Along the way, we take the opportunity to cover much of the basic physics of electron rings.

The plan of these lectures is as follows.

A stored electron beam is a *dissipative, fluctuating system far from equilibrium*. To describe such systems mathematically we first introduce a selection of concepts and techniques from non-equilibrium statistical mechanics. Among these are the appropriate generalizations of *Liouville's Theorem, stochastic differential equations, distribution functions, Fokker-Planck equations* and some methods for solving them, *stationary states* and the evolution of the *expectation values* of dynamical variables. Our derivation of the Fokker-Planck equation is somewhat unconventional but is designed to point up its physical content and permit later comparison with the derivation of the Vlasov equation. In an aside we briefly discuss the relationship between the Fokker-Planck equations we describe and the one used in plasma physics.

The principal source of randomness in electron dynamics is the *quantum fluctuations* of synchrotron radiation. The statistical properties of incoherent synchrotron radiation are reviewed and adapted to the formalism of stochastic differential equations. From the study of the *photon energy distribution function*, the *Poisson distribution of the photon emission time* and a generalized

version of *Campbell's Theorem*, we construct a stochastic representation of the synchrotron radiation power. As an application, we discuss the problem of how photon emission can be simulated numerically.

To describe the fluctuating radiation reaction forces, we add stochastic terms to Hamilton's equations of motion. The essential ingredient for this is the representation of the instantaneous fluctuating power. By examining the extent to which these equations violate Liouville's theorem, we are led to deduce a generalised form of the famous *Robinson theorem* on the damping rates in an electron storage ring. This version of the theorem is local and includes the quantum fluctuations; in cases where the total energy loss in one turn is small compared to the beam energy, the original Robinson theorem is recovered.

We go on to follow how the radiation reaction forces develop in the transformation to the *normal modes* of particle motion about an *equilibrium orbit*. The idea of an equilibrium orbit in an electron ring is fundamentally statistical since no particle ever follows this orbit. Moreover the equilibrium momentum is not necessarily equal to the reference momentum (used to scale the field in the bending magnets) but we show how can be controlled by making small shifts in the RF frequency.

Leaving out (temporarily) the fluctuations of the radiation power, we derive the *radiation damping* effects around the equilibrium orbit in each of the three degrees of freedom. This includes a detailed treatment of the physics of the *damping aperture* and *damping partition number variation* by means of small changes in the equilibrium momentum. The famous sum rule for the damping partition numbers is derived and recognized as a concrete manifestation of Robinson's theorem.

Putting back the quantum fluctuations of the radiation power, we are led naturally to a unified description of radiation damping and *quantum diffusion* in single-particle phase-space by means of stochastic differential equations from which we can derive either Fokker-Planck equations or moment equations. The regime of linear damping and diffusion is discussed in some detail with examples drawn mostly from longitudinal phase-space dynamics. The Fokker-Planck equation is used to find the equilibrium distribution, including the nonlinearity of the RF bucket, and perform the calculation of quantum lifetime. The moment equation technique is used to completely solve a time-dependent problem describing the damping and diffusion which occur in the process of injection into longitudinal phase space.

A subsidiary aim of our treatment has been to clarify the technical details of exactly how these statistical concepts are applied to accelerator problems. For example, the Fokker-Planck equation is often used in an *ad hoc* way, with coefficients fitted to reproduce familiar results in certain limits. Here it is derived from "first principles" starting from a consideration of the radiation reaction

forces on the electron and ending up with the familiar formulae for the energy spread, emittance, *etc.*

To some extent, we follow the treatment first outlined in Refs. 6 and 8 although much more detail is given and the applications are somewhat different. Parts of these lectures were later used in a more condensed form in another summer school.¹⁰ Some alternative approaches to the mathematical side of the problem of electron motion may be found in Refs. 11, 12 and 13 while the implications of the phenomena discussed here for the parameter choice and design of electron storage rings are discussed more extensively in, *e.g.*, Refs. 14, 15, 17, 19, 20 and 21.

In the final chapter, we derive the Vlasov equation from the same general standpoint as we already derived the Fokker-Planck equation. The formalism of Klimontovich allows us to clarify the physical relation between the Vlasov equation and Liouville's theorem (or the equivalent BBGKY hierarchy of equations). This prepares the ground for the theory of collective phenomena.

Prerequisites for reading these lectures should be about the same as for the other beam dynamics lectures in this school. Beyond the basic knowledge of classical Hamiltonian mechanics and electrodynamics which is assumed, the reader should be familiar with single-particle dynamics in storage rings, as covered in earlier lectures.^{22,23,24} Strictly speaking, no specific knowledge of radiation effects on electron motion is required but, to reinforce understanding of these effects, it is well worth reading Ref. 19.

Certain subsections, marked with a ♠, contain material which, although important, it is probably better to skim through or skip entirely on a first reading. They can be returned to later and their temporary omission should not prejudice the understanding of later sections.

2. STATISTICAL MECHANICAL PRINCIPLES

This chapter supplies a general framework in which we can set the specifically accelerator-related problems of the later chapters. It draws on ideas from *non-equilibrium* statistical mechanics rather than the equilibrium theory which is more familiar to most physicists. However those with an engineering background will recognise elements of noise theory.

A paradigm for many dynamical systems occurring in physics and elsewhere is provided by equations of motion of the general form

$$\dot{\mathbf{X}}(t) = \mathbf{K}(\mathbf{X}, t) + \mathbf{Q}(\mathbf{X}, t)\xi(t), \quad (2.1)$$

where $\mathbf{X}(t) = (X_1, \dots, X_N)$ is a vector of coordinates completely characterising the microscopic state of the system, *e.g.* in a near-Hamiltonian system,

the canonical coordinates and momenta $(q_1, \dots, q_n, p_1, \dots, p_n)$, where $N = 2n$. For the moment we shall think of the "system" as one particle with a few degrees of freedom, perhaps a single particle in a stored beam; there may well be many other particles like it but they are presumed not to interact with the particle under consideration. This assumption will be relaxed later in our discussion of the Vlasov equation.

The first term $\mathbf{K}(\mathbf{X}, t)$ in the equations of motion (2.1) represents the deterministic forces acting on the particle; these may include both conservative and dissipative forces. The second term contains a unit noise source $\xi(t)$ and represents the *fluctuating* forces due to some external source of randomness or noise. The function $\mathbf{Q}(\mathbf{X}, t)$ allows for the possibility that the coupling to the noise source varies with the phase space coordinates \mathbf{X} . We shall discuss all these ingredients in more detail below.

2.1 DETERMINISTIC SYSTEMS

In the absence of the fluctuating term, the time-evolution of the system is completely determined (assuming certain mathematical conditions are satisfied) by the set of coupled ordinary differential equations

$$\dot{\mathbf{X}}(t) = \mathbf{K}(\mathbf{X}, t). \quad (2.2)$$

For the purposes of this discussion, the solution need not be stable with respect to small variations of the initial conditions or external parameters; all that matters is that the equations are supposed to be determined. Thus we naïvely include the case of deterministic chaos. Let us denote the formal solution of these equations by $\mathbf{X}(t)$. Then an entirely equivalent description can be given in terms of the *exact phase-space density* for this solution, also known as the *Klimontovich distribution*,^{25,26}

$$\mathcal{F}(\mathbf{X}, t) = \delta(\mathbf{X} - \mathbf{X}(t)) = \prod_{i=1}^N \delta(X_i - X_i(t)). \quad (2.3)$$

Here \mathbf{X} is a free variable and $\mathbf{X}(t)$ is the realisation of the solution of (2.1) starting from some initial conditions $\mathbf{X}(0)$. Clearly, the Klimontovich distribution is a function which is zero everywhere in phase space except at the point $\mathbf{X}(t)$ where it has a δ -function singularity. This singularity moves according to (2.2). The continuity equation for F is a simple mathematical identity, following from (2.2) and (2.3),

$$\begin{aligned} \partial_t \mathcal{F}(\mathbf{X}, t) + \nabla_{\mathbf{X}} \cdot [\dot{\mathbf{X}}(t) \mathcal{F}(\mathbf{X}, t)] &= \partial_t \mathcal{F}(\mathbf{X}, t) + \nabla_{\mathbf{X}} \cdot [\mathbf{K}(\mathbf{X}, t) \mathcal{F}(\mathbf{X}, t)] \\ &= 0. \end{aligned} \quad (2.4)$$

and is known as the *Liouville equation*; the gradient operator $\nabla_{\mathbf{X}}$ denotes the

vector of derivatives with respect to the components of \mathbf{X} and the notation ∂_t stands for the partial derivative with respect to t .

2.1.1 Hamiltonian systems and Liouville's Theorem.

Now is an appropriate moment to discuss the relationship of the Liouville equation to Liouville's Theorem—they are not the same thing.

In the special case of a *Hamiltonian system*, the coordinate vector is the set of canonical variables

$$\mathbf{X} = (q_1, \dots, q_n, p_1, \dots, p_n) \quad (2.5)$$

and, as a consequence of Hamilton's equations,

$$\mathbf{K} = \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}, -\frac{\partial H}{\partial q_1}, \dots, -\frac{\partial H}{\partial q_n} \right), \quad \mathbf{Q} = 0, \quad (2.6)$$

the phase space flow is divergence-free:

$$\nabla_{\mathbf{X}} \cdot \mathbf{K} = 0. \quad (2.7)$$

The gradient operator $\nabla_{\mathbf{X}} = (\partial/\partial \mathbf{q}, \partial/\partial \mathbf{p})$ in (2.4) may then be commuted to the right, past the $\dot{\mathbf{X}}(t)$, giving *Liouville's theorem*,

$$\partial_t \mathcal{F}(\mathbf{q}, \mathbf{p}, t) + \mathbf{K}(\mathbf{q}, \mathbf{p}, t) \cdot \nabla_{\mathbf{X}} \mathcal{F}(\mathbf{q}, \mathbf{p}, t) = 0, \quad (2.8)$$

that the phase space density is conserved *along the flow lines determined by the equation of motion*. In hydrodynamic terms, the flow is incompressible. This may be rewritten in terms of the Poisson bracket operation²⁷

$$\partial_t \mathcal{F} = [H, \mathcal{F}]. \quad (2.9)$$

Note that this equation has the opposite sign from the equation of motion of a dynamical variable $A(\mathbf{q}, \mathbf{p}, t)$:

$$\dot{A} = \partial_t A + [A, H]. \quad (2.10)$$

Whereas the evolution of A is generated by the Hamiltonian, the evolution of the function \mathcal{F} must be such as to “undo” the effects of the time-evolution of the phase-space coordinates, \mathbf{X} , in order to satisfy Liouville's theorem. We must keep in mind the distinction between (2.8) and the more general (2.4). The latter equation can describe dissipative systems which may not conserve phase space density along the flow. However the total phase space density

$$\int \mathcal{F}(\mathbf{X}, t) d^N \mathbf{X} = 1 \quad (2.11)$$

is conserved in any system (by construction of (2.3)).

2.2 FLUCTUATING SYSTEMS

In our attempts to describe real physical systems we find ourselves forced in many cases to include fluctuating terms in the equations of motion. This approach was pioneered by Einstein in his theory of Brownian motion and is now a well developed branch of the theory of stochastic processes. In many problems, the random forces arise in the following manner.

We begin with a description of a microscopic system with many degrees of freedom in terms of the fundamental laws of physics, in the form of equations of motion. To get something more manageable, we try to achieve a reduction of this description to just a few important macroscopic variables. The random forces are an attempt to describe the effects of the neglected microscopic degrees of freedom. We shall see that the random forces acting on an electron in a storage ring are due to the neglect of the many degrees of freedom of the electromagnetic field which couple to the electron.

In the problem of Brownian motion, for example, we consider only the position and velocity of the massive Brownian particle, subject to a deterministic viscous drag force and a random force whose average is zero. It is interesting to note that, fundamentally, these two forces have exactly the same origin in the thermal motion of the molecules of the background medium. Later we shall see that, in very close analogy to this, the deterministic and fluctuating parts of the dissipative forces acting on an electron in a storage ring have the same physical origin. However the randomness arises, not only from the neglect of a large number of uninteresting degrees of freedom, but also from the intrinsic indeterminism of quantum mechanics.

Brownian motion can be described by *Langevin equations*²⁸ which are a special case (constant $Q(\mathbf{X}, t)$) of the stochastic differential equations (2.1). We shall show in Chapter 4 that the electron equations of motion can be cast in a similar form. Let us therefore explain further how we are to interpret such equations.

The state vector $\mathbf{X}(t)$ evolves in time according to the first-order stochastic differential equations (2.1). Here we consider only the case of a single noise source $\xi(t)$. The generalisation to several noise sources is easy: for M sources, simply replace $Q(\mathbf{X}, t)$ with a $N \times M$ matrix and proceed to make the obvious generalisations of the following. In fact we shall find that one noise source is sufficient for our applications since the changes in the three canonical momenta of the electron due to a particular photon emission are correlated.

When we speak of a *realisation* of $\xi(t)$ we mean just one of the ensemble of possible histories of the stochastic process. Of course this ensemble would contain much more information than we could possibly cope with. To overcome this difficulty, we allow ourselves to consider mathematically what can happen

in a given realisation but bear in mind that we must afterwards average over all realisations with an appropriate weighting. The final results of the theory will therefore always be in the form of expectation values.

Many problems become tractable when the stochastic quantity $\xi(t)$ is taken to be gaussian white noise,^{29,30,31,32} defined to have the formal properties

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t)\xi(t') \rangle = \delta(t - t'). \quad (2.12)$$

with respect to this *ensemble-averaging* process $\langle \dots \rangle$. The two-time correlation function $\langle \xi(t)\xi(t') \rangle$ gives the correlation between values of ξ at two different times. The first of these conditions is no great restriction since we could always redefine the coherent force $\mathbf{K}(\mathbf{X}, t)$ to include any non-zero average of $\xi(t)$. The second condition asserts that the fluctuations at two different times are statistically independent and is a much stronger requirement. Since the correlation function of ξ depends only on the *difference* of the times in question, the process is statistically “constant” or homogeneous in time.

These approximations turn out to work very well in many problems. In particular, it turns out that they are satisfied to an excellent approximation in the electron equations of motion, thanks to the properties of synchrotron radiation. We shall demonstrate this in detail later.

To each realisation of $\xi(t)$, then, there corresponds a realisation of the solution $\mathbf{X}(t)$ which we can consider, at least in a formal sense.

2.2.1 An example: the Wiener process.

To show how the formalism of stochastic differential equations works, we consider the simplest possible example, the *Wiener process* which is, in fact, essentially the spatial representation of the Brownian motion problem.

Consider the case where a single particle coordinate is subject only to random kicks and we can set $K = 0$ and $Q = \text{constant}$:

$$\dot{X}(t) = Q\xi(t). \quad (2.13)$$

Then, imposing the initial condition $X(0) = 0$, we can formally integrate a given realisation to find

$$X(t) = Q \int_0^t \xi(t') dt'. \quad (2.14)$$

Averaging this under the integral with the help of (2.12), we find

$$\langle X(t) \rangle = 0. \quad (2.15)$$

However if we first square (2.14) and then average with the help of the second member of (2.12) we obtain

$$\begin{aligned}
\langle X(t)^2 \rangle &= Q^2 \left\langle \int_0^t dt' \int_0^t dt'' \xi(t') \xi(t'') \right\rangle \\
&= Q^2 \int_0^t dt' \int_0^t dt'' \delta(t' - t'') \\
&= Q^2 t.
\end{aligned} \tag{2.16}$$

This is the well-known result that the mean-square deviation of a Brownian particle from its initial position increases linearly with time. The quantity $Q^2/2$ is the *diffusion coefficient*.

2.3 DISTRIBUTION FUNCTIONS AND THE LIOUVILLE EQUATION

The Liouville, or continuity, equation following from a given realisation of (2.1) is

$$\partial_t \mathcal{F}(\mathbf{X}, t) + \nabla_{\mathbf{X}} \cdot [\mathbf{K}(\mathbf{X}, t) \mathcal{F}(\mathbf{X}, t)] + \nabla_{\mathbf{X}} \cdot [\mathbf{Q}(\mathbf{X}, t) \xi(t) \mathcal{F}(\mathbf{X}, t)] = 0. \tag{2.17}$$

where the $\mathbf{X}(t)$ which enters into the definition of $\mathcal{F}(\mathbf{X}, t)$ in (2.3) is the appropriate realisation of the solution of (2.1) starting from some initial conditions $\mathbf{X}(0)$. The Liouville equation is completely equivalent to (2.1). However since $\mathbf{X}(t)$ now depends on the values of the fluctuations $\xi(t)$, $\mathcal{F}(\mathbf{X}, t)$ is also a fluctuating quantity; by averaging over its realisations, we can separate it into an expectation value and a fluctuating component whose average is zero:

$$\mathcal{F}(\mathbf{X}, t) = F(\mathbf{X}, t) + \hat{F}(\mathbf{X}, t), \quad F(\mathbf{X}, t) \stackrel{\text{def}}{=} \langle \mathcal{F}(\mathbf{X}, t) \rangle. \tag{2.18}$$

It would be quite as difficult to deal with—let alone find—the ensemble of these solutions as it would be to deal with that of (2.1). However we are interested only in $F(\mathbf{X}, t)$ and intuition suggests that it ought to be possible to derive an autonomous equation describing its evolution. This is the subject of the next section.

2.4 FOKKER-PLANCK EQUATION

To derive the *Fokker-Planck equation* describing the time evolution of the average phase-space density of a system, subject to deterministic drift \mathbf{K} and diffusion \mathbf{Q} , one must average (2.17) but take account of the way in which the fluctuations make the sharply-peaked phase space density (2.3) spread out to become, upon averaging, a smooth function $F(\mathbf{X}, t)$. In this process, information about the fine details of the distribution function is lost and we make the transition to an *irreversible description* of the time-evolution.

2.4.1 ♠ Derivation

We present a formal derivation of the Fokker-Planck equation intended to emphasise its physical origin and be easily comparable with the derivation of the Vlasov equation which we shall give later. Numerous alternative derivations may be found in the literature.^{30,29,33}

Averaging the Liouville equation (2.17) over all realisations of $\xi(t)$, we find

$$\partial_t \langle \mathcal{F} \rangle + \nabla_{\mathbf{X}} \cdot [\mathbf{K} \langle \mathcal{F} \rangle] + \nabla_{\mathbf{X}} \cdot [\mathbf{Q} \langle \xi \hat{F} \rangle] = 0, \quad (2.19)$$

where we suppressed the arguments (\mathbf{X}, t) of all the vector functions, remembered that \mathbf{K} and \mathbf{Q} are deterministic, and used the first member of (2.12). Our goal of a closed equation for $F = \langle \mathcal{F} \rangle$ is not yet in sight since (2.19) contains the unknown correlation function $\langle \xi(t) \hat{F} \rangle$. To eliminate this, we have to use the information contained in the fluctuating part of (2.17):

$$\partial_t \hat{F} + \nabla_{\mathbf{X}} \cdot [\mathbf{K} \hat{F}] + \nabla_{\mathbf{X}} \cdot \left[\mathbf{Q} \left(\xi \langle \mathcal{F} \rangle + \xi \hat{F} - \langle \xi \hat{F} \rangle \right) \right] = 0. \quad (2.20)$$

Here we used the identity

$$\widehat{\xi \mathcal{F}} = \xi \mathcal{F} - \langle \xi \mathcal{F} \rangle = \xi \langle \mathcal{F} \rangle + \xi \hat{F} - \underbrace{\langle \xi \rangle \langle \mathcal{F} \rangle}_{=0} - \langle \xi \hat{F} \rangle \quad (2.21)$$

and the first member of (2.12).

Now we must think a little about how we model real processes through the strictly Markov idealization $\xi(t)$. A real process $\xi_{\tau}(t)$ has some finite correlation time τ , requiring the replacement of the δ -function in (2.12) with some rapidly-decaying function of the time difference, e.g.,

$$\langle \xi_{\tau}(t) \xi_{\tau}(t') \rangle \sim \frac{1}{2\tau} e^{-|t-t'|/\tau}. \quad (2.22)$$

In the model process, τ is allowed to tend to zero in an appropriate way. For this approach to be valid, the correlation time must be shorter than all other characteristic time scales of the system.

In (2.21) we have to deal with the terms

$$\xi \hat{F} - \langle \xi \hat{F} \rangle = \widehat{\xi \hat{F}} \quad (2.23)$$

which can be identified as the fluctuating part of the product of two fluctuating quantities $\xi \hat{F}$. We therefore neglect all terms in the fluctuation equation (2.20) except

$$\partial_t \hat{F}(\mathbf{X}, t) \simeq -\nabla_{\mathbf{X}}[\mathbf{Q}(\mathbf{X}, t)\xi(t)\langle \mathcal{F}(\mathbf{X}, t) \rangle]. \quad (2.24)$$

In other words, we are assuming that the evolution of \hat{F} on short time scales is governed by the average quantity $\langle \mathcal{F} \rangle$ which evolves only on the slower scales (determined by \mathbf{K}). One can compare this to the method of multiple scales³⁴ or, perhaps more closely, Bogoliubov's "synchronisation hypothesis" in non-equilibrium statistical mechanics.³⁵ This latter idea will be encountered again in our discussion of the Vlasov equation.

Integrating (2.24) over all time preceding the present t , will guarantee that the influence of the initial conditions is negligible, so that

$$\hat{F}(\mathbf{X}, t) = -\int_{-\infty}^t \nabla_{\mathbf{X}} \cdot [\mathbf{Q}(\mathbf{X}, t')\xi(t')\langle \mathcal{F}(\mathbf{X}, t') \rangle] dt'. \quad (2.25)$$

Introducing this into the average equation, (2.19), we obtain

$$\begin{aligned} \partial_t \langle \mathcal{F}(\mathbf{X}, t) \rangle + \nabla_{\mathbf{X}} \cdot [\mathbf{K}(\mathbf{X}, t)\langle \mathcal{F}(\mathbf{X}, t) \rangle] \\ = \nabla_{\mathbf{X}} \cdot \left\{ \mathbf{Q}(\mathbf{X}, t) \int_{-\infty}^t \nabla_{\mathbf{X}} \cdot [\langle \xi(t)\xi(t') \rangle \langle \mathcal{F}(\mathbf{X}, t') \rangle] dt' \right\} \\ = \frac{1}{2} \nabla_{\mathbf{X}} \cdot \{ \mathbf{Q}(\mathbf{X}, t) \nabla_{\mathbf{X}} \cdot [\mathbf{Q}(\mathbf{X}, t)\langle \mathcal{F}(\mathbf{X}, t) \rangle] \}. \end{aligned} \quad (2.26)$$

where the integral was interpreted with the help of (2.12) and (2.22). This is the *Fokker-Planck equation* describing the time evolution of the average phase-space density of a system, subject to deterministic drift \mathbf{K} and diffusion \mathbf{Q} . At this point we are finished with \mathcal{F} and shall simply write F for $\langle \mathcal{F} \rangle$ from now on.

2.5 THE RECIPE

The vector notation used in the previous section may have hidden some of the complexity of the Fokker-Planck equation for systems with several degrees of freedom. So that we can refer to it later, we rewrite (2.26), the Fokker-Planck equation providing an alternative but equivalent description of a system governed by the stochastic differential equations (2.1), as

$$\partial_t F = - \sum_i \frac{\partial}{\partial X_i} \left\{ \left[K_i + \frac{1}{2} \sum_j \frac{\partial Q_j}{\partial X_j} Q_j \right] F \right\} + \frac{1}{2} \sum_i \sum_j \frac{\partial}{\partial X_i} \frac{\partial}{\partial X_j} [Q_i Q_j F]. \quad (2.27)$$

The second term inside the curly brackets is known as the “spurious drift” and is evidently absent when the diffusion $\mathbf{Q}(\mathbf{X}, t) = \mathbf{Q}(t)$ independently of amplitude. Because it often vanishes, this term is easily overlooked but it is essential for a complete description.

So, whenever we have a set of stochastic differential equations of the form (2.1), the corresponding Fokker-Planck equation is uniquely determined. But the converse is not true! *Different stochastic equations may lead to the same Fokker-Planck equation and are therefore statistically equivalent descriptions of the same physical system.*

When we allow stochastic forces into physics we have to get used to the fact that, even when they are formulated in terms of the same dynamical variables, different equations of motion may yet be physically equivalent. The most obvious test for equivalence is, of course, to write down the Fokker-Planck equation.

2.5.1 ♠ Interpretations of stochastic differential equations

It is possible to avoid the apparent complication of spurious drift terms by writing stochastic differential equations which incorporate them into the definition of \mathbf{K} . One must then work with the so-called *Itô calculus*,^{36,37} in which the ordinary rules of differentiation are replaced by rules which bring in extra terms designed precisely to maintain the simple relationship between the differential equations and the Fokker-Planck equation. A change of variables is then rather more work than it is when one works in the *Stratonovich interpretation* as we have done so far.

Since variable transformations will form a larger part of our work than writing down Fokker-Planck equations, the Stratonovich interpretation will be the better choice, and we shall hold to it in the following chapters. In this scheme, Markovian stochastic differential equations are written down as limits of equations governing real processes. This will be implicit in the formulation of the electron equations of motion. Itô equations of motion can be written down to

describe the same physical system but, since the Fokker-Planck equation has to be the same, they will look different from the equivalent Stratonovich equations. The moral is that one must always be careful to know which interpretation is in use.

2.5.2 Near-Hamiltonian systems.

For a system, like the electron in a storage ring, which is close to being Hamiltonian, (2.1) would take the special form

$$\begin{aligned}\dot{\mathbf{q}} &= \frac{\partial H(\mathbf{q}, \mathbf{p}, t)}{\partial \mathbf{p}} + \epsilon \mathcal{K}_{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) + \sqrt{\epsilon} \mathbf{Q}_{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) \xi(t), \\ \dot{\mathbf{p}} &= -\frac{\partial H(\mathbf{q}, \mathbf{p}, t)}{\partial \mathbf{q}} + \epsilon \mathcal{K}_{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) + \sqrt{\epsilon} \mathbf{Q}_{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) \xi(t).\end{aligned}\tag{2.28}$$

Here the ordering of terms with respect to the small parameter ϵ has been chosen so that the dissipative terms have an average effect proportional to $\epsilon \Delta t$ in a short time interval Δt . That this is so is clearer when, following the scheme of (2.27), we write down the Fokker-Planck equation corresponding to (2.28) for the average distribution $F(\mathbf{q}, \mathbf{p}, t)$

$$\begin{aligned}\partial_t F &= -\frac{\partial H}{\partial p_i} \frac{\partial F}{\partial q_i} + \frac{\partial H}{\partial q_i} \frac{\partial F}{\partial p_i} \\ &\quad - \epsilon \frac{\partial}{\partial q_i} \left\{ \left[\mathcal{K}_{q_i} + \frac{1}{2} \frac{\partial Q_{q_i}}{\partial q_j} Q_{q_j} + \frac{1}{2} \frac{\partial Q_{q_i}}{\partial p_j} Q_{p_j} \right] F \right\} \\ &\quad - \epsilon \frac{\partial}{\partial p_i} \left\{ \left[\mathcal{K}_{p_i} + \frac{1}{2} \frac{\partial Q_{p_i}}{\partial q_j} Q_{q_j} + \frac{1}{2} \frac{\partial Q_{p_i}}{\partial p_j} Q_{p_j} \right] F \right\} \\ &\quad + \frac{\epsilon}{2} \frac{\partial}{\partial q_i} \frac{\partial}{\partial q_j} [Q_{q_i} Q_{q_j} F] + \epsilon \frac{\partial}{\partial q_i} \frac{\partial}{\partial p_j} [Q_{q_i} Q_{p_j} F] \\ &\quad + \frac{\epsilon}{2} \frac{\partial}{\partial q_i} \frac{\partial}{\partial q_j} [Q_{q_i} Q_{q_j} F],\end{aligned}\tag{2.29}$$

where summation over the repeated indices i and j is to be understood. The terms on the first line of the right-hand side represent the incompressible flow in phase space described by Liouville's Theorem; the terms on the second and third lines represent a possibly irreversible drift (*e.g.* damping); finally, the last three terms describe diffusion due to the fluctuations.

2.5.3 Fokker-Planck equation for the Wiener process

For a simple example, we turn again to the one-dimensional Wiener process. According to the above recipe, its Fokker-Planck equation is simply

$$\frac{\partial F(X, t)}{\partial t} = \frac{Q^2}{2} \frac{\partial^2 F(X, t)}{\partial X^2}, \quad (2.30)$$

which we recognise as the ordinary diffusion equation. Fokker-Planck equations reduce to diffusion equations in cases like this where the coherent forces $\mathbf{K}(\mathbf{X}, t)$ vanish.

2.6 SOLVING FOKKER-PLANCK EQUATIONS

We speak of solving Fokker-Planck equations, rather than *the* Fokker-Planck equation, to stress that each problem has a different equation associated with it and remarks or progress made in connexion with one Fokker-Planck equation cannot always be generalised to others.

In writing down a Fokker-Planck equation, we replace the problem of solving a set of N stochastic, possibly coupled and nonlinear, ordinary differential equations with the problem of solving a single *partial* differential equation in $N + 1$ variables. Despite any nonlinearity of the stochastic differential equations, the Fokker-Planck equation has the advantage of being *always linear in the unknown distribution function* $F(\mathbf{X}, t)$.

This may—or may not—be a simpler problem to solve. For nonlinear systems particularly, it is usually more tractable. And when the functions $\mathbf{Q}(\mathbf{X}, t)$ have any non-trivial dependence on \mathbf{X} , *i.e.* they are not constants, it is certainly easier to work with. In such cases one cannot average the equations of motion, (2.1), directly since there may be some correlation between the values of the noise ξ and the \mathbf{X} values appearing in its coefficient. The lack of such dependence for the Wiener process allowed us to use the approach of moment equations in the solution and will help us in some of our later examples. But it is not something that can be counted upon in general. When we try to write a moment equation for the longitudinal emittance, for example, we shall find that this strategy fails after serving us well in the same problem formulated in cartesian phase space coordinates.

Not surprisingly, it is hard to find cases in which a full time-dependent solution of a Fokker-Planck equation can be obtained analytically—yet there are a few. For the remainder of this chapter we shall restrict ourselves to cases where the functions \mathbf{K} and \mathbf{Q} have no explicit time-dependence.

2.6.1 Time-dependent solutions

In 1943, Chandrasekhar³⁸ published a fundamental solution (and, with a trivial extension, the Green function) for the Fokker-Planck equation associated with linear problems in one degree of freedom (essentially the harmonic oscillator subject to damping and random noise). Recently Kheifets³⁹ took this as the basis of a perturbation theory and applied it successfully to a number of accelerator physics problems. Another approach, applicable in cases where the time and phase space variables are separable, is to expand the solution as a series in the eigenfunctions; in stable systems these will generally be decaying functions of time, except for the constant equilibrium mode towards which most initial conditions evolve. In other cases numerical solutions can often be attempted.

2.6.2 Stationary states

In a broader category of cases it may be possible to find the *equilibrium solution* analytically, if it exists. The equilibrium solution is a function $F_0(\mathbf{X})$ on phase space found by setting $\partial_t F = 0$ in the Fokker-Planck equation. Despite the terminology, it need not represent a true thermodynamic equilibrium; it would be more precise to speak of *stationary solutions* of the Fokker-Planck equation. Not surprisingly, the terminology arose in cases, like the Brownian motion, where the equation did describe the approach to a true equilibrium.

Notice that, in general, there are no such stationary solutions of the Liouville equation. If the system is Hamiltonian and *completely integrable* the transformation to action-angle variables

$$(\mathbf{q}, \mathbf{p}) \mapsto (\phi, \mathbf{J}) \quad (2.31)$$

can be found, in which the actions \mathbf{J} are all invariant. Then any function $F(\mathbf{J})$ of the invariants alone will be stationary. However there is no irreversible evolution towards such a state.

In one phase space dimension, with time-independent coefficients K and Q , the equilibrium solution of the Fokker-Planck equation can always be found in the form

$$F_0(X) = \frac{|Q(X_0)|}{Z(\lambda)|Q(X)|} \exp \left\{ 2 \int_{X_0}^X \frac{K(X')}{Q(X')^2} dX' \right\}. \quad (2.32)$$

where X_0 is some suitable point in the phase space (perhaps the origin) and it is assumed that $Q(X)$ does not vanish. We have supposed here that the system has some dependences on a set of parameters λ ; these would originate as parameters of K and Q although this is not indicated in our notation. Then, through the

analytic behaviour of the “normalisation constant” or *partition function* $Z(\lambda)$ regarded as a function of λ , we may be able to deduce many of the properties of the distribution and even study whole families of solutions generated by changing the parameters. This technique will be exemplified later.

The function

$$\Phi(X; \lambda) = -\log \left(\frac{F_0(X)}{Z(\lambda)} \right) \quad (2.33)$$

is a generalised thermodynamic potential, in terms of which the equilibrium distribution has the simple form

$$F_0(X) = \frac{1}{Z(\lambda)} \exp(-\Phi(X; \lambda)). \quad (2.34)$$

This is reminiscent of the canonical distribution of equilibrium statistical mechanics although the potential Φ need not bear any simple relationship to a Hamiltonian.

Another point worth noting about (2.32) is that the equilibrium solution no longer contains any information about the dominant time scale of the problem. Analytically this means that, in most cases, only the ratio of the functions K and Q^2 enters. Typically we shall find that some relaxation (or damping) time scale cancels from the equations at this point.

In our example of the Wiener process, it is clear that the equilibrium distribution is a constant, $F_0(X) = Z^{-1}$, provided that the phase space is bounded (in the common case of a phase variable, it may be a circle):

$$X_{\min} < X < X_{\max}, \quad Z = X_{\max} - X_{\min}, \quad (2.35)$$

and we impose the boundary condition $F(X_{\min}) = F(X_{\max})$. Otherwise there is no limit on the diffusion of the process X and the distribution cannot be normalised.

More generally, we can see that the existence of an equilibrium solution is intimately connected with the stability of the dynamical system in the absence of fluctuations. The equilibrium solution is a well-behaved non-negative function on the phase space whose features correspond to topological properties of the underlying phase portrait. Its maxima will correspond to attracting fixed points, its minima to unstable fixed points, and it will be close to constant on an invariant surface. Asymptotically, it will tend to zero along the time-reversed trajectories starting from any attracting set (*e.g.*, if there is damping at large amplitude, the tails of the distribution will decay).

However the visibility of these features deteriorates as the intensity of the noise (magnitude of Q) increases. Thus we can view the featureless Wiener process as a limit of all other one-dimensional processes when the amplitude of the

noise overwhelms any deterministic terms in the equation of motion. Sometimes a deterministic system can become sufficiently ergodic that the Wiener process is a good model for it.

2.6.3 ♠ *Diffusion in Hamiltonian systems*

When a Hamiltonian system is non-integrable there can be no true approach to an equilibrium because of the conservation of phase-space density. However it may be that the motion of the system, at least in some regions of phase space, is sufficiently chaotic that it can be regarded as a diffusion process in some sense. In such cases there may be a rôle for a Fokker-Planck equation whose diffusion coefficient may be calculated from theories of modulational diffusion,⁴⁰ resonance trapping,⁴¹ resonance streaming,⁴² or the combination of nonlinear resonances and external fluctuation effects.⁴³ Thus we should keep in mind that, despite the separation between dissipative and Hamiltonian effects which is maintained here, the Hamiltonian of the system may be intimately involved with effects which at the macroscopic level may be indistinguishable from diffusion. Further information on this topic, especially as it relates to the beam-beam effect in hadron or e^+e^- colliders may be found in Refs. 42.

2.6.4 ♠ *The Fokker-Planck equation of plasma physics.*

Since some readers may be familiar with plasma kinetic theory,^{26,44,45,46} it may be helpful to briefly discuss the relationship of the Fokker-Planck equation used there to that given here. The overlapping terminologies often require some clarification.

Mathematically speaking, these equations are not of the same form, but they are nevertheless related. Physically, of course, they describe quite different phenomena—this is one resemblance between high energy electron bunches and plasmas which turns out to be rather superficial.

Here, the most notable difference is that (2.27) is *always linear*, while in plasma physics, the name is commonly applied to an equation (sometimes also called the *Landau equation*) in which the distribution function appears *quadratically* in order to describe collisions between particles of the same, or different, species. Thus, while we have been exclusively concerned with *single particles* subject, in effect, to *random external forces*, the collision terms of the “plasma” Fokker-Planck equation describe forces acting between *pairs of particles*. This additional physical content makes it considerably more difficult to solve. Yet, when the distribution is linearised about the equilibrium Maxwell-Boltzmann form, an equation of the form (2.27) is recovered. This is not surprising since, in that case, we may think of the equation as describing the distribution function of

a single test particle subjected to random forces arising from the thermal background plasma. In this limit, the Rosenbluth potentials⁴⁶ become independent of the perturbation of the distribution function and are related to the functions \mathbf{K} and \mathbf{Q} used here.

2.7 EXPECTATION VALUES OF DYNAMICAL VARIABLES

When we come to compare the results of theory with experiment, we are not so much interested in the distribution function as in the expectation values of dynamical variables. Dynamical variables are just functions $A(\mathbf{X}, t)$ of the phase space variables. Such functions may or may not be macroscopically measurable; whether they are or not does not prevent us from calculating their values. In the case of a Hamiltonian system, we have already given the equation of motion for $A(\mathbf{X}, t)$ in (2.10). For the general case, (2.1), we can calculate the rate of change of $A(\mathbf{X}, t)$ by the chain rule:

$$\begin{aligned}\dot{A}(\mathbf{X}, t) &= \partial_t A(\mathbf{X}, t) + \dot{\mathbf{X}} \cdot \nabla_{\mathbf{X}} A(\mathbf{X}, t) \\ &= \partial_t A(\mathbf{X}, t) + \mathbf{K}(\mathbf{X}, t) \cdot \nabla_{\mathbf{X}} A(\mathbf{X}, t) + \xi(t) \mathbf{Q}(\mathbf{X}, t) \cdot \nabla_{\mathbf{X}} A(\mathbf{X}, t).\end{aligned}\tag{2.36}$$

When we try to average this over the ensemble of realisations of $\xi(t)$, we run into essentially the same problem that we encountered in the derivation of the Fokker-Planck equation: we do not know the correlation function

$$\langle \xi(t) \mathbf{Q}(\mathbf{X}, t) \cdot \nabla_{\mathbf{X}} A(\mathbf{X}, t) \rangle.\tag{2.37}$$

This time the problem is more severe since the expression of $A(\mathbf{X}, t)$ in terms of \mathbf{X} may be quite complicated. For example, if it is an n th degree polynomial in the X_i then we should need to know a whole sequence of correlation functions:

$$\langle X_i \rangle, \quad \langle X_i X_j \rangle, \quad \langle X_i X_j X_k \rangle, \dots, \langle X_{i_1} X_{i_2} X_{i_3} \dots X_{i_n} \rangle.\tag{2.38}$$

It is clear that the problem is intractable in general. Of course, the Fokker-Planck equation was invented precisely in order to avoid this problem. If, having solved it, we know the distribution function, then we can calculate the expectation value in question from

$$\begin{aligned}\langle A(\mathbf{X}, t) \rangle &= \left\langle \int A(\mathbf{X}, t) \mathcal{F}(\mathbf{X}, t) d^N \mathbf{X} \right\rangle \\ &= \int A(\mathbf{X}, t) F(\mathbf{X}, t) d^N \mathbf{X}.\end{aligned}\tag{2.39}$$

It is interesting to see what happens if we try to use the Fokker-Planck equation, in the form (2.26), to differentiate (2.39) under the integral sign. For simplicity,

let us suppose that $\partial_t A(\mathbf{X}, t) = 0$; then

$$\begin{aligned} \frac{d}{dt} \langle A(\mathbf{X}, t) \rangle &= \int d^N \mathbf{X} A(\mathbf{X}) \partial_t F(\mathbf{X}, t) \\ &= \int d^N \mathbf{X} A(\mathbf{X}) \\ &\quad \times \nabla_{\mathbf{X}} \cdot \left[-\mathbf{K}(\mathbf{X}, t) F(\mathbf{X}, t) + \frac{1}{2} \mathbf{Q} \nabla_{\mathbf{X}} \cdot (\mathbf{Q}(\mathbf{X}, t) F(\mathbf{X}, t)) \right]. \end{aligned} \quad (2.40)$$

Assuming that $F(\mathbf{X}, t) \rightarrow 0$ sufficiently fast towards the boundaries of phase space (usually at infinity), we may integrate by parts to find

$$\begin{aligned} \frac{d}{dt} \langle A(\mathbf{X}, t) \rangle &= \underbrace{\int d^N \mathbf{X} F(\mathbf{X}, t) \mathbf{K} \cdot \nabla_{\mathbf{X}} A(\mathbf{X})}_{\langle \mathbf{K} \cdot \nabla_{\mathbf{X}} A \rangle} \\ &\quad - \underbrace{\frac{1}{2} \int d^N \mathbf{X} [\mathbf{Q}(\mathbf{X}, t) \cdot \nabla_{\mathbf{X}} A(\mathbf{X})] \nabla_{\mathbf{X}} \cdot [\mathbf{Q}(\mathbf{X}, t) F(\mathbf{X}, t)]}_{\frac{1}{2} \int d^N \mathbf{X} F(\mathbf{X}, t) \mathbf{Q}(\mathbf{X}, t) \cdot \nabla_{\mathbf{X}} [\mathbf{Q}(\mathbf{X}, t) \cdot \nabla_{\mathbf{X}} A(\mathbf{X})]} \\ &= -\frac{1}{2} \langle \mathbf{Q} \cdot \nabla_{\mathbf{X}} [\mathbf{Q} \cdot \nabla_{\mathbf{X}} A] \rangle \end{aligned} \quad (2.41)$$

Finally, this yields the result that we attempted to obtain by means of (2.36),

$$\frac{d}{dt} \langle A \rangle = \langle \mathbf{K} \cdot \nabla_{\mathbf{X}} A \rangle + \frac{1}{2} \langle \mathbf{Q} \cdot \nabla_{\mathbf{X}} [\mathbf{Q} \cdot \nabla_{\mathbf{X}} A] \rangle, \quad (2.42)$$

showing that, in general, autonomous equations of evolution for the expectation values of the dynamical variables of a fluctuating system do not exist. The evaluation of the right hand side of (2.42) will involve correlation functions of ever higher order. Sometimes it is possible to truncate the hierarchy by means of some approximation scheme, whereby correlation functions above a certain order are expressed in terms of those of lower order. In certain important special cases, there are great simplifications and we shall consider these next.

2.7.1 Amplitude-independent noise

In practice the noise intensity function $\mathbf{Q}(\mathbf{X}, t)$ is often independent of \mathbf{X} . As we noted before, this greatly simplifies the writing down of the Fokker-Planck equation. The moment equation (2.42) reduces to

$$\frac{d}{dt} \langle A \rangle = \langle \mathbf{K} \cdot \nabla_{\mathbf{X}} A \rangle + \frac{1}{2} Q_i Q_j \left\langle \frac{\partial^2 A}{\partial X_i \partial X_j} \right\rangle. \quad (2.43)$$

Things are particularly simple if $A(\mathbf{X})$ is linear or quadratic in \mathbf{X} ,

$$\begin{aligned}\frac{d}{dt}\langle X_1 \rangle &= \langle K_1(\mathbf{X}) \rangle, \\ \frac{d}{dt}\langle X_1 X_2 \rangle &= \langle K_1(\mathbf{X}) X_2 \rangle + \langle K_2(\mathbf{X}) X_1 \rangle + Q_1 Q_2, \\ \frac{d}{dt}\langle X_1^2 \rangle &= 2\langle K_1(\mathbf{X}) X_1 \rangle + Q_1^2, \\ &\dots\dots\end{aligned}\tag{2.44}$$

If, moreover, each of the K s is linear in the X s, it is evident that the system of equations for the first and second moments is linear and can be solved completely (by simple diagonalization) without any truncation.

2.8 STOCHASTIC CALCULUS

We shall have occasion to formally integrate stochastic equations of the form (2.1) and we must therefore establish the rules for doing so; again this is something which turns out differently in the Itô and Stratonovich interpretations. We can establish the Stratonovich integration rule for the average part of (2.1) by substituting $A(\mathbf{X}, t) = \mathbf{X}$ into (2.42), and then integrating formally; this introduces the Q -function and its gradient into the average evolution of \mathbf{X} in the correct manner. Although we shall not demonstrate the result here, it turns out that the fluctuating part of the integral is given in the most obvious way and the final result can be written

$$\begin{aligned}\mathbf{X}(t) &= \mathbf{X}(0) + \int_0^t \mathbf{K}(\mathbf{X}, t') dt' \\ &\quad + \frac{1}{2} \int_0^t \mathbf{Q}(\mathbf{X}, t') \cdot \nabla_{\mathbf{X}} \mathbf{Q}(\mathbf{X}, t') dt' \\ &\quad + \int_0^t \mathbf{Q}(\mathbf{X}, t') \xi(t') dt'\end{aligned}\tag{2.45}$$

For a fuller justification of this formula, the reader is referred to the literature on stochastic differential equations.^{29,30,37}

On the other hand, as we have already mentioned, the Stratonovich interpretation allows us to apply the formal rules of ordinary calculus whenever we have to form the derivative of a function. Thus, taking (2.17) and (2.39), we might have done an integration by parts to justify our use of the chain rule in writing down (2.36). The other rules of stochastic differential calculus can be derived in a similar fashion. The rules of Itô differential calculus are more complicated.

When we come to change variables in stochastic differential equations, then, we can simply apply the chain rule and this is often rather easy.

Sometimes it can be quite complicated but, when dissipative, fluctuating systems are close to being canonical, as in (2.28), their equations of motion contain a Hamiltonian part, upon which it is convenient to make canonical transformations.²⁷ It is therefore useful to extend the formalism of generating functions to transform the dissipative parts too. The more complicated coordinate transformations of electron storage ring theory are facilitated by such a formalism. The radiation terms in the equations of motion are small on average compared with those describing the applied electromagnetic fields.

In the following subsection we shall derive the appropriate extension of the canonical formalism for a particular type of generating function.

2.8.1 ♠ Transformation theory for stochastic equations

Consider the case of a *free* (i.e. such that old and new coordinates are independent) transformation from $(\mathbf{q}, \mathbf{p}) = (q_1, \dots, q_n, p_1, \dots, p_n)$ to new variables $(\mathbf{Q}, \mathbf{P}) = (Q_1, \dots, Q_n, P_1, \dots, P_n)$.

Call the old Hamiltonian $H(\mathbf{q}, \mathbf{p}, t)$ and the new $K(\mathbf{Q}, \mathbf{P}, t)$. Both in the old and new variables, Hamilton's equations have to be supplemented by terms (\mathbf{a}, \mathbf{b}) and (\mathbf{A}, \mathbf{B}) which describe the dissipation:

$$\begin{array}{ccc} \overbrace{\begin{array}{l} \dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}} + \mathbf{a}(\mathbf{q}, \mathbf{p}, t), \\ \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}} + \mathbf{b}(\mathbf{q}, \mathbf{p}, t), \end{array}}^{\text{OLD}} & \iff & \overbrace{\begin{array}{l} \dot{\mathbf{Q}} = \frac{\partial K}{\partial \mathbf{P}} + \mathbf{A}(\mathbf{Q}, \mathbf{P}, t), \\ \dot{\mathbf{P}} = -\frac{\partial K}{\partial \mathbf{Q}} + \mathbf{B}(\mathbf{Q}, \mathbf{P}, t). \end{array}}^{\text{NEW}} \end{array} \quad (2.46)$$

In the case of (2.28) we should make the correspondence

$$\begin{aligned} \mathbf{a}(\mathbf{q}, \mathbf{p}, t) &\equiv \epsilon \mathcal{K}_{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) + \sqrt{\epsilon} \mathbf{Q}_{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t) \xi(t), \\ \mathbf{b}(\mathbf{q}, \mathbf{p}, t) &\equiv \epsilon \mathcal{K}_{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) + \sqrt{\epsilon} \mathbf{Q}_{\mathbf{p}}(\mathbf{q}, \mathbf{p}, t) \xi(t). \end{aligned} \quad (2.47)$$

Since the transformation is canonical, there must exist a generating function $S(\mathbf{Q}, \mathbf{q}, t)$, depending on the old and new coordinates and no momentum,[†] such that the relationship between old and new coordinates and momenta is obtained

[†] This type of generating function, which is useful in the transformation to action-angle variables, is sometimes denoted $F_1(\mathbf{Q}, \mathbf{q}, t)$.

by solving the system of equations

$$\mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}, \quad \mathbf{P} = -\frac{\partial S}{\partial \mathbf{Q}}, \quad K = H + \frac{\partial S}{\partial t}. \quad (2.48)$$

In a system with 3 degrees of freedom, 6 different types²⁷ of generating function are necessary to generate all possible canonical transformations, and in practice each of them comes up sooner or later. The other cases can be worked by analogy to this one; for a different example, see Ref. 8.

The equivalence of the two descriptions of the Hamiltonian system is guaranteed by these relationships. Given that the two sets of equations in (2.46) are supposed to describe the same dissipative system, we need to know how to calculate the new dissipative terms (\mathbf{A}, \mathbf{B}).

[As previously, we must now acknowledge that any stochastic terms are to be interpreted in the Stratonovich, not the Itô, sense. The following results would be a great deal more complicated in the Itô calculus.]

To avoid cumbersome notations, let us denote partial derivatives by subscripts where convenient; thus, for example, we may write a vector or a matrix

$$S_{\mathbf{q}} = \begin{pmatrix} \frac{\partial S}{\partial q_1} \\ \frac{\partial S}{\partial q_2} \\ \vdots \\ \frac{\partial S}{\partial q_n} \end{pmatrix}, \quad S_{\mathbf{Q}\mathbf{q}} = \begin{pmatrix} \frac{\partial^2 S}{\partial Q_1 \partial q_1} & \frac{\partial^2 S}{\partial Q_1 \partial q_2} & \cdots & \frac{\partial^2 S}{\partial Q_1 \partial q_n} \\ \frac{\partial^2 S}{\partial Q_2 \partial q_1} & \frac{\partial^2 S}{\partial Q_2 \partial q_2} & \cdots & \frac{\partial^2 S}{\partial Q_2 \partial q_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 S}{\partial Q_n \partial q_1} & \frac{\partial^2 S}{\partial Q_n \partial q_2} & \cdots & \frac{\partial^2 S}{\partial Q_n \partial q_n} \end{pmatrix}. \quad (2.49)$$

By hypothesis, the arguments of the generating function provide a unique labelling of points in extended phase space.²⁷ Hence, their differentials form a basis for the vector space of differentials of dynamical variables and any such differential can be expressed as a linear combination of them. In particular, from (2.48), we may use the chain rule to calculate

$$d\mathbf{p} = S_{\mathbf{q}\mathbf{q}}d\mathbf{q} + S_{\mathbf{q}\mathbf{Q}}d\mathbf{Q} + S_{\mathbf{q}s}dt. \quad (2.50)$$

When the postulated equations of motion (2.46), are used to project this identity along the local direction of time-evolution of the system, we find

$$\left. \begin{aligned} d\mathbf{p} &= S_{\mathbf{q}\mathbf{q}}(H_{\mathbf{p}} + \mathbf{a})dt \\ &+ S_{\mathbf{q}\mathbf{Q}}(K_{\mathbf{p}} + \mathbf{A})dt + S_{\mathbf{q}s}dt \\ &= -H_{\mathbf{q}}dt + \mathbf{b}dt \end{aligned} \right\} \implies \begin{cases} S_{\mathbf{q}\mathbf{q}}H_{\mathbf{p}} + S_{\mathbf{q}\mathbf{Q}}K_{\mathbf{p}} + S_{\mathbf{q}s} = -H_{\mathbf{q}} \\ S_{\mathbf{q}\mathbf{q}}\mathbf{a} + S_{\mathbf{q}\mathbf{Q}}\mathbf{A} = \mathbf{b}. \end{cases} \quad (2.51)$$

where we separately identified the Hamiltonian and dissipative parts of the equality.

In an analogous way, we can calculate dP and conclude

$$\begin{aligned} S_{Qq}H_p + S_{QQ}K_P + S_{Qs} &= K_P, \\ -S_{Qqa} - S_{QQ}A &= B. \end{aligned} \quad (2.52)$$

Since the canonical transformation is free, the Jacobian determinant of either set of canonical variables with respect to the independent coordinates (Q, q) cannot vanish, *i.e.*

$$\frac{\partial(\mathbf{p}, \mathbf{q})}{\partial(\mathbf{Q}, \mathbf{q})} = \begin{vmatrix} \frac{\partial \mathbf{p}}{\partial \mathbf{Q}} & \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \\ \frac{\partial \mathbf{q}}{\partial \mathbf{Q}} & \frac{\partial \mathbf{q}}{\partial \mathbf{q}} \end{vmatrix} = \begin{vmatrix} \frac{\partial \mathbf{p}}{\partial \mathbf{Q}} & \frac{\partial \mathbf{p}}{\partial \mathbf{q}} \\ 0 & 1 \end{vmatrix} = \det \frac{\partial \mathbf{p}}{\partial \mathbf{Q}} = \det S_{qQ} \neq 0. \quad (2.53)$$

Hence the matrix S_{qQ} is invertible and it follows that the last members of (2.51) and (2.52) may be solved simultaneously to yield a *linear** relationship between the dissipative terms in the old and new representations:

$$\begin{aligned} \mathbf{A} &= [S_{qQ}]^{-1}(\mathbf{b} - S_{qq}\mathbf{a}), \\ \mathbf{B} &= -S_{Qqa} - S_{QQ}[S_{qQ}]^{-1}(\mathbf{b} - S_{qq}\mathbf{a}). \end{aligned} \quad (2.54)$$

By means of these formulae, the dissipative terms can be transformed in parallel with the canonical transformation; (2.54) may be regarded as a convenient recipe for carrying out complicated transformations. It should be remembered that, in the course of this work, all derivatives of the generating function must be expressed in terms of the variables (Q, q) and the expressions for $\mathbf{q}(Q, P)$ and $\mathbf{p}(Q, P)$ should only be substituted afterwards. Of course, exactly the same constraint applies while transforming the Hamiltonian.

3. STATISTICS OF INCOHERENT SYNCHROTRON RADIATION

As an electron is accelerated in a magnetic field, it emits photons. Because this is an intrinsically quantum-mechanical phenomenon, the emission times and the magnitudes of the quanta of energy carried away by the photons are random quantities. However, provided the energies and magnetic fields are not too high, certain average quantities, such as the mean emission rate and the mean radiated power, may be calculated with good accuracy within classical electrodynamics. Only when the magnetic fields and energies become so high that the mean of the classical synchrotron radiation spectrum becomes comparable to the electron energy does it become necessary to include quantum corrections to the total radiated power and the frequency spectrum.

* When these terms have fluctuating parts this property is not at all trivial. It applies only as a consequence of our use of the Stratonovich interpretation.

It is somewhat surprising, but nonetheless true, that the randomness cannot be ignored even in the “classical” synchrotron radiation regime.

Since electron dynamics is profoundly influenced by the phenomenon of synchrotron radiation it *cannot* be described completely in terms of a single-particle Hamiltonian. We must somehow add *dissipative terms* to describe the energy loss through radiation. A purely Hamiltonian description can be recovered only in the context of quantum electrodynamics by including the degrees of freedom of the electromagnetic field together with those of the particle.

It is much easier, and quite adequate for our purposes, to add fluctuating radiation reaction forces to the classical equations of motion.

The properties of incoherent synchrotron radiation,^{19,47,48,49,26,50,20} play a governing rôle in the statistical mechanics of electron motion; we first review them, cast in a notation suited for our purposes. Paying special attention to the statistical properties of the radiation, we shall derive the principal result of this chapter, namely that the instantaneous radiation power may be represented as a stochastic function of the type postulated in (2.1). This provides the ingredients we need to construct the electron equations of motion in the next chapter.

Since the orbital quantum numbers of electrons in typical storage rings are very large, we may use classical arguments to construct the equations of motion provided we do not pretend to describe the emission process itself correctly. We shall use a mathematically convenient model which approximates it as an instantaneous jump in energy. This is acceptable since, for an electron of energy $E = \gamma mc^2$ in a magnetic field $B = E/e\rho$, photon emission occurs within a time

$$\tau_\gamma \simeq \frac{\rho}{\gamma c} \ll \frac{1}{\Omega}, \quad (3.1)$$

where Ω is any frequency characteristic of betatron or synchrotron oscillations. In addition, the fact that

$$\frac{1}{\tau_\gamma} \ll \omega_c, \quad (3.2)$$

where ω_c is the *critical frequency*

$$\omega_c \stackrel{\text{def}}{=} \frac{3}{2} \frac{c\gamma^3}{\rho} \quad (3.3)$$

means that the frequency (or energy) spectrum of the photons is locally well-defined.¹⁹ The critical frequency serves as a useful scale parameter for the frequency distribution; the definition used here is one-half that used in Ref. 47.

We shall restrict ourselves to the orbital degrees of freedom of the stored electron and neglect its spin. The evolution of the spin state is somewhat analogous

to that of the classical canonical variables in that it is primarily determined by the applied electromagnetic fields, in accordance with the Thomas-BMT equation, but important secondary effects arise from the changes in spin state induced by photon emission (actually the spin may, or may not, flip when a photon is emitted). Just as quantum excitation causes a diffusion of orbital amplitudes around the closed orbit, so there is a corresponding spin diffusion around the precession axis. However no analogue of radiation damping exists for the spin.

The quantum effects on spin motion can be taken into account by mathematical methods similar to those employed here. Further details on spin motion and polarization of electron beams in general will be found in Refs. 51 and 52.

3.1 CLASSICAL PICTURE

According to classical electrodynamics, a particle accelerated transversely by a magnetic field emits a continuous beam of radiation in a narrow cone around its momentum vector. Quantum mechanically, the momentum vector of each photon is almost collinear with the particle's momentum.

We start from the relativistic classical theory^{47,48} of radiation from a point charge moving in externally applied electric and magnetic fields \mathbf{E} and \mathbf{B} , according to the Lorentz force equation. We adopt the usual (timelike) convention that, if $A = (A^0, \mathbf{A})$ and $B = (B^0, \mathbf{B})$ are 4-vectors, then their scalar product $A \cdot B$ is given by

$$A \cdot B = A^\mu B_\mu = A^0 B^0 - \mathbf{A} \cdot \mathbf{B}. \quad (3.4)$$

Proper time along the particle's world line will be denoted by τ .

3.1.1 Radiated power

The power radiated by a particle with 4-momentum $p^\mu = (E/c, \mathbf{p})$ is given by the Lorentz-invariant expression⁴⁷

$$\begin{aligned} P &= \left\{ -\frac{dE}{dt} \text{ in any frame} \right\} \\ &= -\frac{2}{3} \frac{e^2}{m^2 c^3} \frac{dp^\mu}{d\tau} \frac{dp_\mu}{d\tau} \end{aligned} \quad (3.5)$$

If there exists a frame in which the electromagnetic field is purely magnetic, then, in that frame, the Lorentz force equation has components

$$\frac{dp^\mu}{d\tau} = \frac{e}{mc} (0, \mathbf{p} \times \mathbf{B}). \quad (3.6)$$

If, further, the magnetic field is transverse to the direction of motion of the particle ($\mathbf{p} \cdot \mathbf{B} = 0$), then (3.5) simplifies to

$$\begin{aligned} P &= \frac{2}{3} \frac{e^4}{m^4 c^3} (\mathbf{p} \times \mathbf{B})^2 \\ &= \frac{2 e^2 r_e p^2}{3 (mc)^3} |\mathbf{B}|^2 = \frac{2 r_e p_0^2}{3 m^3 c} p^2 b^2 \end{aligned} \quad (3.7)$$

where $r_e = e^2/mc^2$ is the classical electron radius. This has been expressed in terms of the magnitude of the 3-momentum, $p = |\mathbf{p}|$ (which we shall later use as a canonical variable) and the normalised magnetic field strength b defined with respect to some reference momentum p_0 by

$$\mathbf{B} = \nabla \times \mathbf{A} = \frac{p_0 c}{e} \mathbf{b} \quad (3.8)$$

We have also written b for \mathbf{b} . Let us pause to observe that the power is a function only of the momentum of the particle, through the factor p^2 , and its position, through the factor b^2 . A similar functional dependence will hold for all other parameters related to the synchrotron radiation. Since it will be very important to evaluate these dependences correctly in terms of the dynamical variables defining the state of the particle, we shall indicate the dependence on the position and momentum variables by a subscript X . In addition, the physical position of the particle includes its azimuthal position s around the ring. This has special status as the independent variable of particle dynamics in a synchrotron (like time in ordinary dynamics) and the corresponding dependence will be indicated separately. Thus for the radiation power, (3.7), we shall write $P_X(s)$ when we need to be explicit; when there is no danger of ambiguity we may still write simply P .

The motivation for the above assumptions is rather obvious: along most of a storage ring orbit, an electron moves in a magnetic bending field almost transverse to the direction of motion; synchrotron radiation in the electric fields of the accelerating system is negligible; finally, p_0 corresponds to the value of the momentum on the reference orbit.

The critical frequency (3.3) can now be expressed as

$$\omega_c = \frac{3}{2} \frac{cp_0}{(mc)^3} p^2 b_X(s). \quad (3.9)$$

Because of (3.8), p_0 and b will appear to the same power in each expression.

3.1.2 Frequency spectrum

Following Ref. 19 closely, we express (3.7) in terms of its spectral density

$$P = \int_0^{\infty} \mathcal{P}(\omega) d\omega \quad (3.10)$$

where, again according to the (approximate) classical calculation

$$\mathcal{P}(\omega) = \frac{P}{\omega_c} S(\omega/\omega_c). \quad (3.11)$$

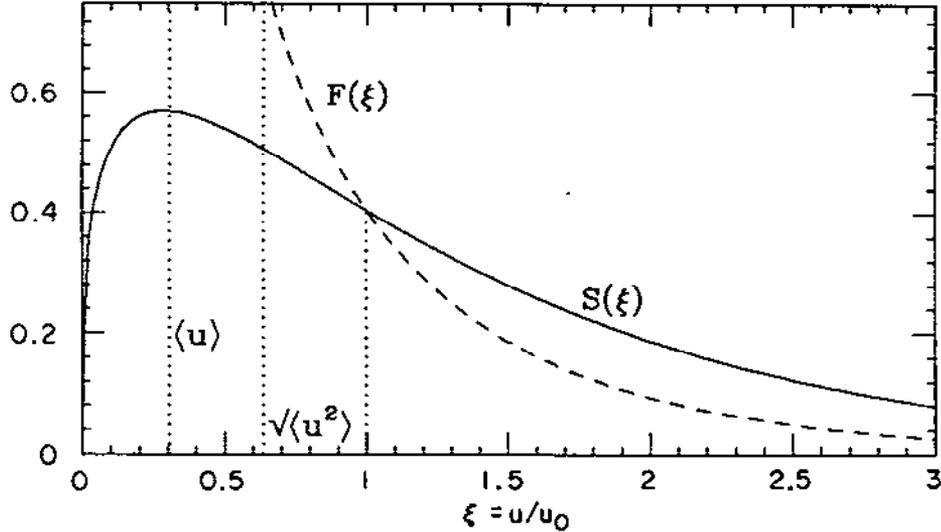


Fig. 1 Frequency or energy distribution of synchrotron radiation

Here we use the standard definitions¹⁹

$$S(\xi) \stackrel{\text{def}}{=} \frac{9\sqrt{3}}{8\pi} \xi \int_{\xi}^{\infty} K_{5/3}(\zeta) d\zeta, \quad F(\xi) \stackrel{\text{def}}{=} S(\xi)/\xi, \quad (3.12)$$

and $K_{5/3}$ is a modified Bessel function (see *e.g.* Refs. 19, 47 and 48 for derivations of (3.11)). The universal functions $F(\xi)$ and $S(\xi)$ are plotted in Fig. 1. Note that a non-algebraic dependence of $\mathcal{P}(\omega)$ on the momentum and magnetic field arises through the factor ω_c in the argument of S .

3.2 QUANTUM-MECHANICAL PICTURE

Knowing that, in vacuum, a classical electromagnetic wave of frequency ω corresponds to a photon of energy $u = \hbar\omega$ we can reinterpret the frequency distribution of synchrotron radiation as the energy distribution of photons.

3.2.1 Photon energy distribution

Defining the energy spectrum of photons emitted per unit time by

$$P = \int_0^{\infty} n_X(u; s) u \, du \quad (3.13)$$

we can compare with (3.10) and make the correspondence

$$n_X(u; s) = \frac{\mathcal{P}(u/\hbar)}{\hbar u}. \quad (3.14)$$

From this we may directly deduce the number of photons emitted per unit time at all energies

$$\begin{aligned} N_X(s) &= \int_0^{\infty} n_X(u; s) \, du \\ &= \frac{5\sqrt{3}}{6} \frac{e^3}{mc^2 \hbar} |\mathbf{B}(x, y, s)| = \frac{5\sqrt{3}}{6} \frac{cr_e p_0}{\hbar} b_X(s) \end{aligned} \quad (3.15)$$

which is independent of p .

Given a photon of unknown energy we may ask: what is the relative likelihood of different energies? This question is answered by the energy spectrum of the photons

$$\begin{aligned} f_X(u; s) &= \frac{n_X(u; s)}{N_X(s)} \\ &= \frac{5\sqrt{3}}{8} \frac{(mc)^3}{\hbar c p_0} \frac{F(u/u_c)}{p^2 b_X(s)}, \end{aligned} \quad (3.16)$$

The total probability of any energy is, of course,

$$\int_0^{\infty} f_X(u; s) \, du = 1. \quad (3.17)$$

3.2.2 Moments of the photon energy

Now we can evaluate the average or expectation value of the photon energy

$$\langle u \rangle_{Xs} = \int_0^{\infty} u f_X(u; s) du = \frac{4}{5\sqrt{3}} \frac{\hbar c p_0}{(mc)^3} p^2 b_X(s) = \frac{8}{15\sqrt{3}} u_c, \quad (3.18)$$

by means of which we can re-express the classical power as the average power

$$\langle P_X(s) \rangle = N_X(s) \langle u \rangle_{Xs}. \quad (3.19)$$

From now on we shall always have to indicate explicitly when we are speaking of the average power and when we mean the complete fluctuating power.

Similarly, the mean-square quantum energy is the second moment of $f_X(u; s)$:

$$\langle u^2 \rangle_{Xs} = \int_0^{\infty} u^2 f_X(u; s) du = \frac{11}{12} \frac{(\hbar c p_0)^2}{(mc)^6} p^4 b_X(s)^2. \quad (3.20)$$

We shall see that, together with the mean emission rate, the mean square quantum energy determines the magnitude of the quantum fluctuations in electron motion.

3.3 SYNCHROTRON RADIATION AS A STOCHASTIC PROCESS

Now that we have succeeded in evaluating some of the statistical properties of the photon energy distribution, we can use this knowledge to construct expressions for a fluctuating radiation power which can be used in stochastic differential equations of the sort discussed in Chapter 2.

An individual photon emission event, in which a photon of energy u_j is emitted at $s = s_j$, is specified by the ordered pair of random variables (u_j, s_j) . As we pointed out after (3.7), the distribution function of (u_j, s_j) depends *only on the local magnetic field and the particle's momentum*. Since these quantities vary as the particle moves, there is no very meaningful way of relating time averages (along the trajectory of a given particle) and ensemble averages (over many hypothetical particles experiencing the same conditions).

For definiteness, let us define the *expectation value* operation on a dynamical variable $\mathcal{A}(x, y, z, p_x, p_y, p; s)$, associated with the instantaneous state of the particle, to be the average of \mathcal{A} over all possible realisations of (u_j, s_j) , that is

to say, all the ways in which the particle's photon emission history might occur, weighted appropriately. As in Chapter 2, we denote such ensemble averages by

$$\langle \mathcal{A}(x, y, z_t, p_x, p_y, p; s) \rangle \quad \text{or, more briefly and generally,} \quad \langle \mathcal{A} \rangle_{Xs} \quad (3.21)$$

as it suits us; here X is a shorthand notation for the set of (usually canonical) variables describing the instantaneous state of the particle in whatever representation we happen to be using and s is the azimuthal position around the machine.

The averaging is understood to be taken while s and the phase-space coordinates X —and thereby the magnetic field felt by the particle—are supposed fixed. Parameters characterising the synchrotron radiation may be regarded as dynamical variables of the particle since they too are determined by X and s . For example, the critical energy may be thought of as a parameter determining the overall scale of the energy distribution of the photons which the particle has a propensity to emit.

The exact density, analogous to the Klimontovich distribution (2.3), of a given realisation (holding X fixed) in (u, s) space is

$$\Omega_X(u, s) = \sum_j \delta(s - s_j) \delta(u - u_j), \quad (3.22)$$

The sum is taken over all events which actually take place in that realisation. The expectation value of (3.22) is the distribution function of (u_j, s_j) (more correctly termed the probability density function)

$$\langle \Omega_X(u, s) \rangle = N_X(s) f_X(u; s) / c, \quad (3.23)$$

The factorization of this expression reflects the fact that u_j and s_j are statistically independent variables. Here, $N_X(s)$ is the distribution function of s_j , or the average photon emission rate, given by (3.15), and $f_X(u; s)$ is the energy spectrum of the photons, given by (3.16).

The instantaneous radiated power is

$$\begin{aligned} P_X(s) &= c \int_0^\infty u \, du \int_{-\infty}^\infty ds' \Omega_X(u, s') \Delta(s - s') \\ &= c \sum_j u_j \Delta(s - s_j). \end{aligned} \quad (3.24)$$

where $\Delta(s)$ is a function which we may think of as describing the shape of the photon emission pulse. It has the properties

$$\int_{-\infty}^{\infty} \Delta(s) ds = 1, \quad \int_{-\infty}^{\infty} \Delta(s)^2 ds \simeq \frac{1}{c\tau_\gamma}, \quad (3.25)$$

i.e. it is normalised to unity and has a pulse width equivalent to the photon emission time (3.1). As $\tau_\gamma \rightarrow 0$, Δ becomes a δ -function.

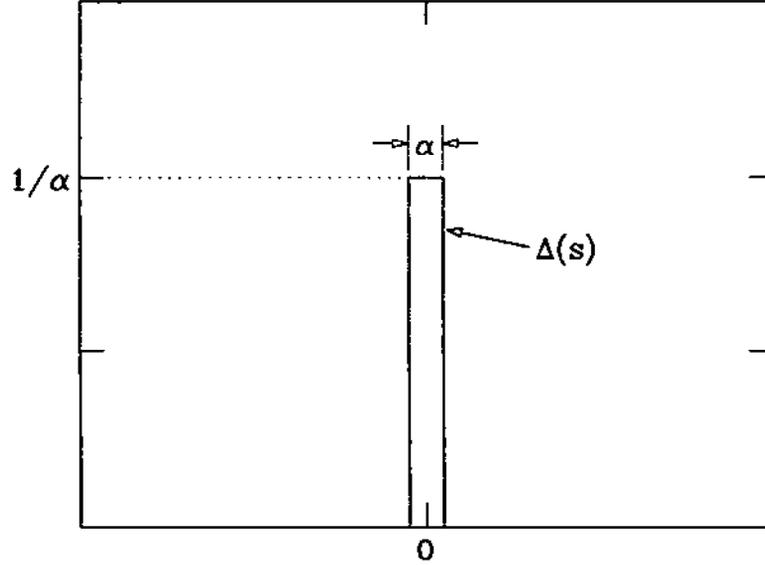


Fig. 2 Representation of the photon emission pulse

Strictly speaking, the value of τ_γ is another fluctuating quantity but the arguments already given show that we need not concern ourselves with this complication. Moreover we may choose any convenient form for the pulse function Δ provided it satisfies (3.25). One such choice is the simple bump:

$$\begin{aligned} \Delta(s) &= \frac{\theta(s + \alpha/2)\theta(\alpha/2 - s)}{\alpha} \\ &= \begin{cases} 0, & \text{for } s < -\alpha/2; \\ 1/\alpha, & \text{for } -\alpha/2 < s < \alpha/2; \\ 0, & \text{for } s > \alpha/2. \end{cases} \end{aligned} \quad (3.26)$$

where $\alpha = c\tau_\gamma$ and $\theta(s)$ is the usual Heaviside step function; this function is shown in Fig. 2.

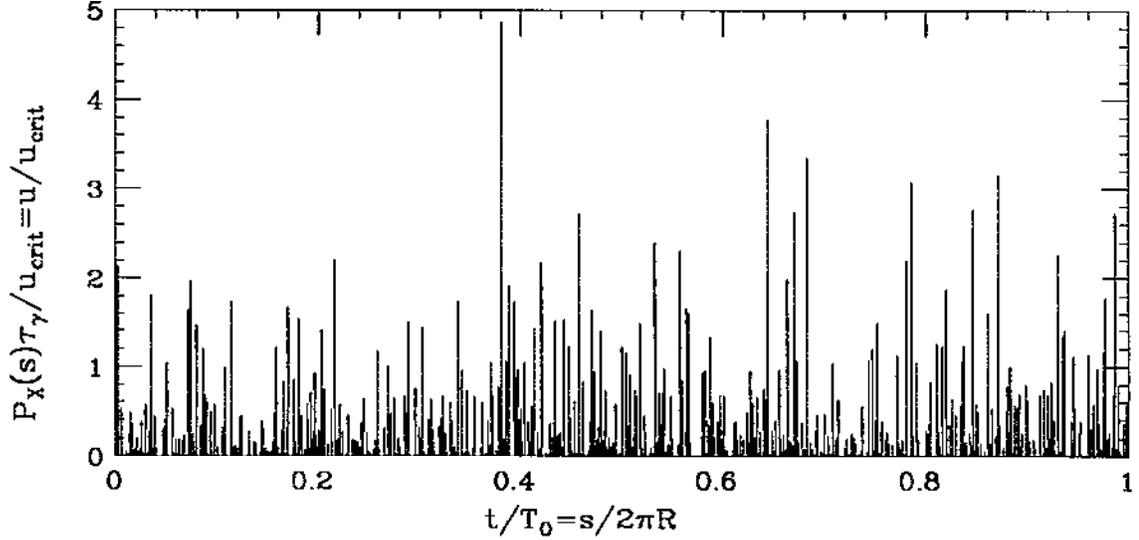


Fig. 3 A realisation of $P_X(s)$

A typical realisation of $P_X(s)$, obtained by simulation, is shown in Fig. 3; the parameters are such that the expectation value of the number of photons emitted in one revolution period is

$$N_X(s)T_0 = 1000 \quad (3.27)$$

Counting the peaks, we find 1049 in this realisation, a 1.55σ deviation. In preparing the figure, τ_γ was been taken to be *fixed* and equal to the width of a line on the printer so that the distribution of the heights of the peaks, displayed in units of the critical energy u_c , is given by the function $f_X(u; s)$ defined in (3.16). It is worth remarking that, although half the energy is carried away by photons with $u > u_c$, there are few such photons. In fact 91% of the photons have $u < u_c$ and 50% have $u < 0.1 u_c$.

3.4 DISTRIBUTION OF PHOTON EMISSION TIME

Let $\text{Pr}(n, s)$ denote the probability that n photons are emitted in an interval of azimuth $(0, s)$. Suppose that the emission rate per metre, $N_X(s)/c = N_X/c$ is constant, *i.e.* there is no significant variation of momentum p or magnetic field \mathbf{B} . Then, for a sufficiently small step Δs (the probability of 2 photons being emitted in Δs is supposed to be negligible), we can write the probability that n photons are emitted in $(0, s + \Delta s)$ as the sum of the probabilities that

- a) n are emitted in $(0, s)$ and none in $(s, s + \Delta s)$, and
- b) $n - 1$ are emitted in $(0, s)$ and 1 in $(s, s + \Delta s)$.

This gives us the equation

$$\Pr(n, s + \Delta s) = \Pr(n, s) \left(1 - \frac{N_X}{c} \Delta s \right) + \Pr(n-1, s, s) \frac{N_X}{c} \Delta s. \quad (3.28)$$

Taking the limit as $\Delta s \rightarrow 0$, we can form the derivative

$$\begin{aligned} \frac{\partial \Pr(n, s)}{\partial s} &= \lim_{\Delta s \rightarrow 0} \frac{\Pr(n, s + \Delta s) - \Pr(n, s)}{\Delta s} \\ &= \frac{N_X}{c} [\Pr(n-1, s) - \Pr(n, s)] \end{aligned} \quad (3.29)$$

A well-known technique for solving this type of differential equation, for all values of n , is to introduce a generating function

$$G(k, s) = \sum_{n=0}^{\infty} k^n \Pr(n, s). \quad (3.30)$$

From this we obtain a differential equation for G ,

$$\begin{aligned} \frac{\partial G(k, s)}{\partial s} &= \frac{N_X}{c} \sum_{n=0}^{\infty} k^n [\Pr(n-1, s) - \Pr(n, s)] \\ &= \frac{N_X}{c} (k-1) \sum_{n=0}^{\infty} k^n \Pr(n, s) \\ &= \frac{N_X}{c} (k-1) G(k, s). \end{aligned} \quad (3.31)$$

Since $\Pr(0, 0) = 1$ and $\Pr(n, 0) = 0$ for $n > 1$, the initial condition is $G(s, 0) = 1$ and we may integrate (3.31) to find the solution

$$\begin{aligned} G(k, s) &= \exp \left(\frac{N_X}{c} (k-1) s \right) \\ &= \exp \left(-\frac{N_X s}{c} \right) \sum_{n=0}^{\infty} \left(\frac{N_X k s}{c} \right)^n \frac{1}{n!}. \end{aligned} \quad (3.32)$$

Equating coefficients between (3.30) and (3.32), we have the required probability

$$\Pr(n, s) = \exp \left(-\frac{N_X s}{c} \right) \left(\frac{N_X s}{c} \right)^n \frac{1}{n!} \quad (3.33)$$

which is the familiar *Poisson distribution*.

From this expression it is straightforward to calculate the expectation value of any function of the number of photons emitted in $(0, s)$ using the formula

$$\langle f(n) \rangle_s = \sum_{n=0}^{\infty} f(n) \text{Pr}(n, s). \quad (3.34)$$

We already know the expectation value of n

$$\langle n \rangle_s = \frac{N_X s}{c}, \quad (3.35)$$

but (3.33) enables us to find its mean-square

$$\langle n^2 \rangle_s = \langle n \rangle_s + \langle n \rangle_s^2. \quad (3.36)$$

This result will be used in the proof of Campbell's Theorem.

3.5 MEAN OF RADIATION POWER

As a prelude to Campbell's Theorem, let us illustrate how we can recover expression (3.19) for the average power from the representation of the instantaneous fluctuating power in (3.24).

Consider all possible realisations of an electron's photon emission events in some large interval $(0, S)$. Again, assume that there is no significant variation of p or \mathbf{B} . Restrict consideration to those realisations in which *exactly* n photon emission events occur, *i.e.* those that can be written in the form

$$P_X(s)_n = c \sum_{j=1}^n u_j \Delta(s - s_j). \quad (3.37)$$

Then averaging over all the realisations in this subset, we have

$$\begin{aligned} \langle P_X(s) \rangle_n &= c \int_0^S \frac{ds_1}{S} \int_0^S \frac{ds_2}{S} \cdots \int_0^S \frac{ds_n}{S} \int_0^{\infty} du_1 f_X(u_1; s_1) \int_0^{\infty} du_2 f_X(u_2; s_2) \\ &\quad \cdots \int_0^{\infty} du_n f_X(u_n; s_n) \sum_{j=1}^n u_j \Delta(s - s_j) \\ &= c \sum_{j=1}^n \int_0^S \frac{ds_j}{S} \Delta(s - s_j) \int_0^{\infty} du_j f_X(u_j; s_j) u_j \\ &= cn \langle u \rangle_{X_s} \int_0^S \frac{ds_1}{S} \Delta(s - s_1), \end{aligned} \quad (3.38)$$

where we recognized that the sum was composed of n identical terms and used (3.17). If we now assume that $S \gg \tau_\gamma$ so that, with high probability, we can

neglect the interference of the pulse $\Delta(s - s_1)$ with the ends of the interval of integration, we may finally average over n , using

$$\lim_{S \rightarrow \infty} \frac{\langle n \rangle}{S} = \frac{N_X}{c}, \quad (3.39)$$

to recover (3.19).

3.6 CAMPBELL'S THEOREM

The exact power (3.24) may be split into its mean and fluctuating parts:

$$P_X(s) = \langle P_X(s) \rangle + \hat{P}_X(s), \quad (3.40)$$

where, in a given realisation, the fluctuating part $\hat{P}_X(s)$ is just the difference between the classical power (3.19) and the instantaneous power. Obviously, it has zero average.

The manipulations of the previous section were a preparation for the proof of a generalized version of Campbell's Theorem,³³ which in the present context gives us an expression for the *two-time correlation function of the fluctuating part of the radiation power*:

$$\langle \hat{P}_X(s) \hat{P}_X(s') \rangle = cN_X(s) \langle u^2 \rangle_{X_s} \Delta(s - s'). \quad (3.41)$$

The importance of this theorem for electron dynamics was first pointed out by Sands⁵³ and a proof may be found in the article of S. Rice in Ref. 33. We give a less rigorous but simpler proof of a more general result in the following subsection.

One important special case of Campbell's Theorem is the famous Schottky formula for shot noise in thermionic tubes; this is important in the Schottky scan diagnostics of coasting proton beams and in stochastic cooling theory. In this case the rôle of the synchrotron radiation photons is played by the protons themselves. The rôle of the photon energy is taken by the proton charge so there is only one random variable in this case.

For purposes of calculation, it is convenient to take the limit $\tau_\gamma \rightarrow 0$ and write (3.41) as

$$\begin{aligned} \langle \hat{P}_X(s) \hat{P}_X(s') \rangle &= cN_X(s) \langle u^2 \rangle_{X_s} \delta(s - s') \\ &= \frac{55}{24\sqrt{3}} \frac{r_e \hbar c^4 p_0^3}{(mc)^6} p^4 b_X(s)^3 \delta(s - s'). \end{aligned} \quad (3.42)$$

The δ -function expresses the fact that $\hat{P}_X(s)$ and $\hat{P}_X(s')$ are uncorrelated at different times ($s \neq s'$).

3.6.1 ♠ Proof

To establish the generalised version of Campbell's Theorem, we again consider the interval $(0, S)$ and restrict our attention to those realisations in which exactly n photons are emitted. Using (3.24), we take the average of

$$P_X(s)_n P_X(s')_n = c^2 \sum_{k=1}^n \sum_{j=1}^n u_j u_k \Delta(s - s_j) \Delta(s - s_k), \quad (3.43)$$

which is

$$\begin{aligned} \langle P_X(s) P_X(s') \rangle_n &= c^2 \sum_{k=1}^n \sum_{j=1}^n \int_0^S \frac{ds_1}{S} \dots \int_0^S \frac{ds_n}{S} \int_0^\infty du_1 f_X(u_1; s_1) \dots \\ &\dots \int_0^\infty du_n f_X(u_n; s_n) u_j u_k \Delta(s - s_j) \Delta(s' - s_k). \end{aligned} \quad (3.44)$$

In the double sum we separate the n terms which have $j = k$ and the $n(n-1)$ which have $j \neq k$ to obtain

$$\begin{aligned} \langle P_X(s) P_X(s') \rangle_n &= c^2 \sum_{j=1}^n \int_0^S \frac{ds_j}{S} \int_0^\infty du_j f_X(u_j; s_j) u_j^2 \Delta(s - s_j) \Delta(s' - s_k) \\ &+ c^2 \sum_{j \neq k=1}^n \int_0^S \frac{ds_1}{S} \dots \int_0^S \frac{ds_n}{S} \int_0^\infty du_1 f_X(u_1; s_1) \dots \\ &\dots \int_0^\infty du_n f_X(u_n; s_n) u_j u_k \Delta(s - s_j) \Delta(s - s_k) \\ &= c^2 n \langle u^2 \rangle_{Xs} \int_0^S \frac{ds_1}{S} \Delta(s - s_1) \Delta(s' - s_1) \\ &+ c^2 n(n-1) \langle u \rangle_{Xs}^2 \int_0^S \frac{ds_1}{S} \Delta(s - s_1) \int_0^S \frac{ds_2}{S} \Delta(s' - s_2). \end{aligned} \quad (3.45)$$

Using the result (3.36) for the Poisson distribution, we can average over n as before and push the limits of the integrations over s and s' to $\pm\infty$,

$$\begin{aligned}
\langle P_X(s)P_X(s') \rangle_n &= cN_X\langle u^2 \rangle_{Xs} \int_{-\infty}^{\infty} ds_1 \Delta(s-s_1)\Delta(s'-s_1) \\
&\quad + c^2 \left(\frac{N_X\langle u \rangle_{Xs}}{c} \right)^2 \underbrace{\int_{-\infty}^{\infty} ds_1 \Delta(s-s_1)}_1 \underbrace{\int_{-\infty}^{\infty} ds_2 \Delta(s'-s_2)}_1 \\
&= cN_X\langle u^2 \rangle_{Xs} \int_{-\infty}^{\infty} ds_1 \Delta(s-s_1)\Delta(s'-s_1) \\
&\quad + \langle P_X(s) \rangle \langle P_X(s') \rangle.
\end{aligned} \tag{3.46}$$

Finally, we deal with the remaining integral, using the representation (3.26):

$$\begin{aligned}
I &= \int_{-\infty}^{\infty} ds_1 \Delta(s-s_1)\Delta(s'-s_1) = \int_{-\infty}^{\infty} d\sigma \Delta(\sigma)\Delta(s'-s+\sigma) \\
&= \frac{1}{\alpha} \int_{-\alpha/2}^{\alpha/2} d\sigma \Delta(s-s'+\sigma) \\
&= \begin{cases} 0, & \text{if } |s-s'| > \alpha/2; \\ \frac{1}{\alpha^2} \int_{-\alpha/2}^{\alpha/2} d\sigma \theta(s'-s-\sigma-\alpha/2)\theta(s'-s-\sigma+\alpha/2) & \text{otherwise.} \end{cases}
\end{aligned} \tag{3.47}$$

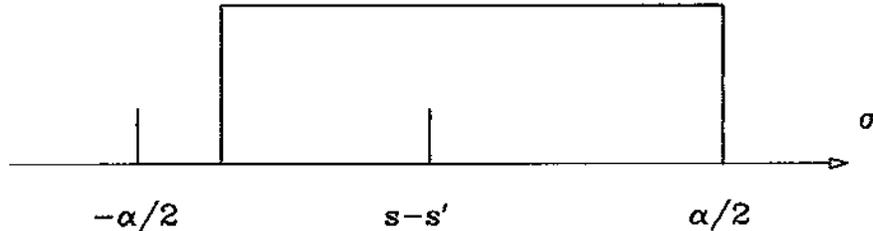


Fig. 4 Evaluation of the integral (3.47)

The integral arising in the second case is evaluated with the help of Fig. 4. The integrand is zero everywhere except in the region shown in the figure where it

is equal to unity. Taking account of both cases in (3.47) by means of more Heaviside functions, we therefore have

$$\begin{aligned}
I &= \frac{1}{\alpha^2} [\alpha - |s - s'|] \theta(s' - s - \alpha/2) \theta(s' - s + \alpha/2) \\
&= \underbrace{\Delta(s' - s)}_{\rightarrow \delta(s' - s)} - \underbrace{\frac{|s' - s|}{\alpha} \Delta(s' - s)}_{\rightarrow 0},
\end{aligned} \tag{3.48}$$

where, finally, we took the limit $\alpha \rightarrow 0$ with $|s' - s| \neq 0$ fixed; the last term vanishes in the case $s = s'$.

This completes the proof of the theorem.

3.7 REPRESENTATION OF FLUCTUATING POWER

In terms of the unit noise source, $\xi(s)$, introduced in (2.12), we can construct the principal result of this chapter, a formal representation of the stochastic power

$$P_X(s) = p^2 c^2 [c_1 b_X(s)^2 + \sqrt{c_2} |b_X(s)|^{3/2} \xi(s)]. \tag{3.49}$$

The constants c_1 and c_2 are defined by

$$c_1 \stackrel{\text{def}}{=} \frac{2r_e p_0^2}{3(mc)^3}, \quad c_2 \stackrel{\text{def}}{=} \frac{55r_e \hbar p_0^3}{24\sqrt{3}(mc)^6}. \tag{3.50}$$

These are not “fundamental” constants because they still depend on the absolute value of the bending field in the ring through the reference momentum p_0 .

With the help of (2.12), it is easily shown that this representation correctly reproduces the essential properties (3.19) and (3.42). Statistically, there is *no way of distinguishing* the conceptual model of discrete random photon emission and this formal representation.

Noting that $c_2 \propto \hbar$, we see that, in this formalism, the classical radiation power has been corrected by a stochastic term of order $\sqrt{\hbar}$. We also observe that, in general, the average radiation power and its quantum fluctuations depend nonlinearly on the particle’s coordinates through the spatial dependences of the magnetic field. This can generate interesting nonlinear effects but we shall not discuss them here; the interested reader is referred to Refs. 8, 6 or 10 for further details.

3.8 SIMULATING PHOTON EMISSION

Some interesting statistical questions arise when we consider the problem of simulating the photon emission process on a computer. The principles involved in solving this illustrate some important conceptual points concerning distribution functions and are a good model for the simulation of other statistical phenomena.

In simulating synchrotron radiation, there are two things which a program must decide, namely the time when a photon emission is to occur and, subsequently, how much energy the photon in question has to carry away. Thus, the two random quantities u_j and s_j have to be generated with the correct distribution.

3.8.1 Generation of an event

Let us consider the emission time first, supposing that the program advances the time, or azimuth variable, in discrete steps. The correct Poisson distribution is easily generated in the following manner:

Divide each unit of azimuth up into $N_X(s)/c$ cells and then further subdivide each cell into I_d steps, where $I_d \approx 10$, say; so the basic time step size will be $\Delta t = \Delta s/c = 1/N_X(s)I_d$. The expectation number of photons emitted in one of these small steps is then

$$N_X(s)\Delta t = \frac{N_X(s)}{c} \times \left(\frac{c}{N_X(s)I_d} \right) \ll 1 \quad (3.51)$$

At each step, generate a random number R_s , uniformly distributed in $[0, 1]$; if $R_s \leq 1/I_d$, an event which has the probability given by (3.51), then emit a photon in this step; otherwise do not.

To generate the random photon energies correctly, we must first calculate the cumulative probability distribution in energy

$$\Phi_X(u; s) = \Pr(u_j \leq u) = \int_0^u f_X(u'; s) du', \quad (3.52)$$

i.e., the probability that the energy of a given photon falls in the interval $[0, u]$. Clearly, this function is monotonic-increasing with the properties

$$\Phi_X(0; s) = 0, \quad \lim_{u \rightarrow \infty} \Phi_X(u; s) = 1. \quad (3.53)$$

Generate a second random number R_u , with the same statistical properties as for R_t and solve the equation

$$\Phi_X(u; s) = R_u \quad (3.54)$$

for the unknown u ; this provides an energy value with the correct statistical properties.

The algorithm just described was used in the program which generated Fig. 3. In practice, of course, the numerical integrations involved in calculating the function Φ_X by means of (3.52), (3.16) and (3.12) need only be performed once to generate a set of numerical values for subsequent interpolation in solving (3.54).

The technique described here for generating random numbers according to a specific distribution is quite standard and can be applied to numerous problems; as an interesting recent example, it has been used⁵⁴ to simulate photon emission in a regime where the quantum corrections to the synchrotron radiation spectrum become significant.

3.8.2 ♠ *Approximate simulation techniques*

In many simulation programs used in accelerator problems, particularly those concerned with single-beam collective effects or the beam-beam effect, the forces under study are localised in a small number of points around the machine. It is common practice in such cases to replace the many random photon emission events occurring in a sector, $[0, S]$ of the machine by the emission of a single "fat" photon whose energy is, on average, equal to

$$\langle U_S \rangle = \int_S P_X(s) ds/c \quad (3.55)$$

and has a fluctuation whose mean square value is chosen so that the correct value for the energy spread is obtained. Generally, (3.55) is expressed as a constant (energy loss for a particle on the reference trajectory) plus terms linear in the deviations of the canonical coordinates from those of the reference particle. Typically, it is also assumed that the orbital dynamics through the same sector of the machine is linear. With these assumptions, the evolution of the dynamical variables of the particle can be broken down into two separate mappings, one for the orbital dynamics (usually a matrix representation) and one for the synchrotron radiation effects.

These simplifications allow substantial economies in running time which, for programs concerned with many particles and many turns of a large storage

ring lattice, can be well worthwhile. Indeed it would very often be practically impossible to run the simulations otherwise.

The justification for this procedure can be given in terms of our stochastic representation of synchrotron radiation as follows:

In a given realization of the electron dynamics in $[0, S]$, the energy radiated can be decomposed into its mean and fluctuating parts:

$$\begin{aligned}
 U_S &= \int_0^S P_X(s) ds \\
 &= \underbrace{\int_0^S \langle P_X(s) \rangle ds}_{\langle U_S \rangle} + \underbrace{\int_0^S \hat{P}_X(s) ds}_{\hat{U}_S}.
 \end{aligned} \tag{3.56}$$

Evidently, (3.55) gave the correct value for the mean energy loss and, with the help of (3.40), (3.42) and (2.12), it is easy to see that the r.m.s. value of the fluctuation should be

$$\langle \hat{U}_S^2 \rangle = \int_0^S N_X(s) \langle u^2 \rangle_{X_s} ds / c \tag{3.57}$$

Usually, the quantity U_S is given a gaussian (normal) distribution with mean and variance given by (3.55) and (3.57) respectively. If the number of photons emitted is very large,

$$\mathcal{N}_S = \int_0^S N_X(s) ds / c \gg 1 \tag{3.58}$$

then the Central Limit Theorem⁵⁵ can be invoked to guarantee that the special form, (3.16), of the photon energy spectrum will be “washed out” when the energies of all the individual photons are added and the distribution of U_S will indeed be gaussian to a good approximation.

There are two circumstances in which this common simulation technique will be inappropriate:

- a) *The number of photons emitted in the lumped sector of the machine is small.* This situation can be dealt with by using a more appropriate

distribution according to the number of photons, *e.g.* in the case of 2 photons being emitted, the probability that the sum of their energies lies in a small interval $[u, u + \Delta u]$ is given by the convolution

$$\begin{aligned}
 Pr(u \leq u_1 + u_2 \leq u + \Delta u) &= \Delta u \int_0^\infty du_1 f_X(u_1; s_1) \int_0^\infty du_2 f_X(u_2; s_2) \\
 &\quad \times \delta(u - u_1 - u_2) \\
 &= \int_0^\infty du_1 f_X(u_1; s_1) f_X(u - u_1; s).
 \end{aligned} \tag{3.59}$$

(naturally, we define $f_X(u; s)$ to be zero for negative u) which is quite different from a gaussian. When n photons are emitted, we get an n -fold convolution; multiple convolution integrals of this form are precisely what one deals with in the proof of the Central Limit Theorem (see, *e.g.*, the Appendix of Ref. 55).

- b) If the particle dynamics through the lumped sector of the machine is not linear, either through nonlinearities of the lattice or of the dependence of the synchrotron radiation properties on the dynamical variables of the particle, then the changes caused by the radiation and the other dynamical effects cannot be simply superposed.

The second of these two cases is the main obstacle to the development of a ****tracking**** program for nonlinear orbital dynamics which correctly simulates the effect of synchrotron radiation on electrons. Thus, in most dynamic aperture studies for electron rings, the radiation effects are simply left out. It is quite difficult to achieve even the limited goal of a correct simulation of the deterministic radiation damping effect in this context. For, if the fluctuations are simply left out for lack of an appropriate simulation technique, the longitudinal phase space volume of the beam will simply damp away to zero.

The only truly satisfactory solution to the problem of simulating radiation effects on nonlinear motion would be to use the algorithm described at the beginning of this section to track particles through every single magnet in the ring. Clearly this would greatly increase the running time of any tracking program; however it could be (and is) argued that, since electrons need only be tracked for a few damping times, which may only be a few thousand turns in a high energy machine, the situation is more hopeful than in hadron machines where it is difficult to place any *a priori* upper bound on the number of turns necessary in tracking studies.

4. ELECTRON DYNAMICS WITH RADIATION REACTION

It may turn out that the ultimate solution of the dynamic aperture problem for electrons may come from an analytical approach, perhaps through improved techniques for solving the Fokker-Planck equations, rather than the tracking approach which, despite the attendant practical and theoretical problems, has become so widely used.

In order to establish our notation and conventions we briefly recapitulate the Hamiltonian formulation of particle dynamics, which will be familiar to the reader from other lectures in this School.^{22,24,23}

But Hamilton's equations are not the whole story: the main point of this chapter is to add the *dissipative and fluctuating terms* which describe the energy loss through radiation. The expression for the fluctuating power derived in the previous chapter is the essential ingredient for a formulation of electron dynamics in terms of stochastic differential equations of the type introduced and illustrated in Chapter 2.

4.1 COORDINATE SYSTEM AND HAMILTONIAN

As usual, we employ the curvilinear coordinate system of Courant and Snyder⁵⁶ and follow the conventions of most standard optics programs (MAD,⁵⁷ TRANSPORT,⁵⁸ etc.).

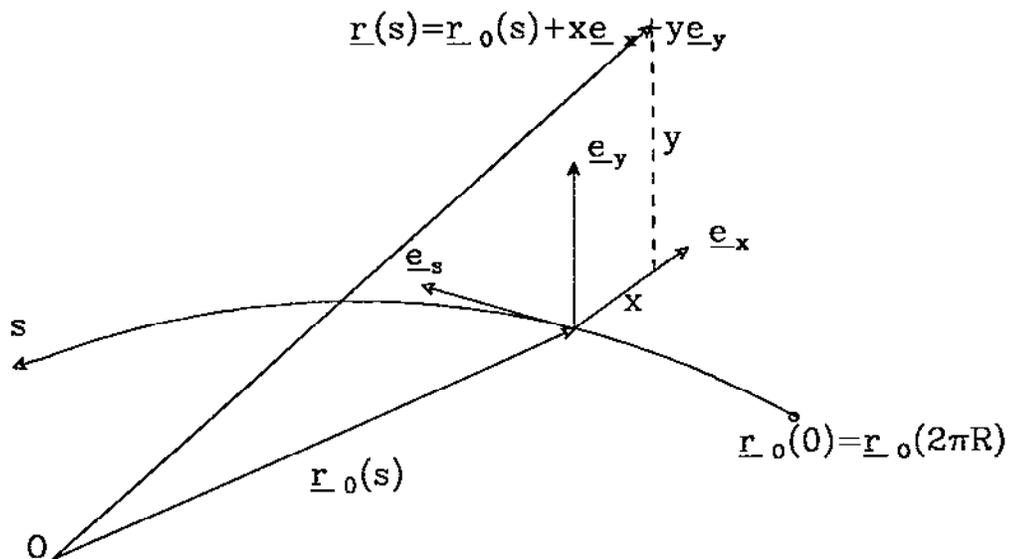


Fig. 5 The reference orbit and Courant-Snyder coordinate system

For algebraic simplicity, we assume that the magnets are perfectly aligned in the sense that there exists a closed planar reference curve $r_0(s)$, passing

through all their centres which is also the *closed orbit* for a hypothetical reference particle of (suitably orientated) constant momentum p_0 which neither radiates nor couples to the RF accelerating fields. There is no difficulty in incorporating a closed orbit deviation from the magnet centres into what follows; even if there is some vertical bending so that the ring is not planar, the expressions simply become somewhat more cumbersome but no less tractable.

The position of a real particle of kinematic momentum \mathbf{p} is then described by giving the azimuthal position s of the closest point on this curve and its radial and vertical deviations x and y from that point, as shown in Fig. 5.

If we neglect edge effects in the magnets and use the Coulomb gauge⁴⁷ then, in many important cases, the fields can be derived from a single scalar function,⁵⁶ the *canonical vector potential*⁵⁹

$$\begin{aligned} A_s(x, y, t, s) &= \mathbf{A} \cdot \mathbf{e}_s (1 + G(s)x) \\ &= -\frac{p_0 c}{e} \left\{ xG(s) \left(1 + G(s)\frac{x}{2} \right) \right. \\ &\quad \left. + \frac{1}{2}K_1(s)(x^2 - y^2) + \frac{1}{6}K_2(s)(x^3 - 3xy^2) + \dots \right\} \quad (4.1) \\ &\quad + \sum_k \frac{c\hat{V}_k}{\omega_{\text{rf}}} \delta_C(s - s_k) \cos(\omega_{\text{rf}}t + \phi_k). \end{aligned}$$

The final term describes the fundamental accelerating mode of a set of “thin lens” RF cavities with peak voltages \hat{V}_k located at positions s_k ; δ_C is a δ -function, periodic on the circumference $C = 2\pi R$. Our assumption about the closed orbit of the reference particle implies that the normalised dipole field strength is equal to the curvature of the reference orbit:

$$G(s) = \frac{\mathbf{r}_0''(s) \cdot \mathbf{r}_0(s)}{|\mathbf{r}_0(s)|} \quad (4.2)$$

(where primes denote derivatives with respect to s). Following Courant and Snyder,^{56,24} we take s as independent variable so that (x, y, t) may be taken as canonical coordinates and $(p_x, p_y, -E)$ as canonical momenta. The Hamiltonian of a particle with kinematic momentum \mathbf{p} is

$$\begin{aligned} H_s(x, y, t, p_x, p_y, -E; s) &= (\mathbf{p} + (e/c)\mathbf{A}) \cdot \mathbf{e}_s (1 + G(s)x) \\ &= -(e/c)A_s(x, y, t, s) \\ &\quad - (1 + G(s)x) \sqrt{E^2/c^2 - m^2c^2 - p_x^2 - p_y^2}. \end{aligned} \quad (4.3)$$

The term m^2c^2 in the square root causes some formal complications later; in hadron machines these give rise to the phenomenon of transition energy and are

related to the velocity of the particle. Since we are working with ultra-relativistic electrons ($v \simeq c$), we do not wish to be concerned with these. Before going any further, it is therefore convenient to replace the energy, E , by the magnitude of the total momentum p ; this requires a canonical transformation of one pair of variables,

$$(t, -E) \mapsto (z_t, p), \quad (4.4)$$

effected by means of the generating function,

$$F_2(p, t) = -ct\sqrt{p^2 + m^2c^2}, \quad (4.5)$$

and resulting in the new Hamiltonian

$$H(x, y, z_t, p_x, p_y, p; s) = -\frac{e}{c}A_s(x, y, t(z_t, p), s) - (1 + G(s)x)\sqrt{p^2 - p_x^2 - p_y^2}. \quad (4.6)$$

The new canonical variables are related to the old by

$$\begin{aligned} p &= \sqrt{E^2/c^2 - m^2c^2}, \\ z_t &= -ct\sqrt{1 - m^2c^4/E^2} \\ &= -(\text{instantaneous velocity}) \times (\text{time particle passes } s). \end{aligned} \quad (4.7)$$

i.e. z_t is *not* immediately related to the path length, except while the energy is constant.

4.1.1 Hamilton's equations

Writing out Hamilton's equations explicitly, we find

$$\begin{aligned} x' &= (1 + Gx)\frac{p_x}{\sqrt{p^2 - p_x^2 - p_y^2}} \simeq (1 + Gx)\frac{p_x}{p}, \\ y' &= (1 + Gx)\frac{p_y}{\sqrt{p^2 - p_x^2 - p_y^2}} \simeq (1 + Gx)\frac{p_y}{p}, \\ z_t' &= -(1 + Gx)\frac{p}{\sqrt{p^2 - p_x^2 - p_y^2}} \simeq -(1 + Gx), \end{aligned} \quad (4.8)$$

for the coordinates and

$$\begin{aligned} p_x' &= -G(p - p_0) - p_0(G^2 + K_1)x - \frac{1}{2}p_0K_2(x^2 - y^2) + \dots, \\ p_y' &= p_0K_1y + \frac{1}{2}p_0K_2xy + \dots, \\ p' &= -\sum_k \frac{e\hat{V}_k}{c} \delta_C(s - s_k) \sin(\omega_{\text{rf}}z_t/c + \phi_k). \end{aligned} \quad (4.9)$$

for the momenta. Our purpose in writing these out is not to revise the fundamentals of single-particle dynamics but to exhibit the dependence on the canonical momenta. In particular the factors $p_{x,y}/p$ in the expressions for x' and y' are very important for the correct evaluation of the radiation damping effects.

In these variables, the reference momentum p_0 factors out of all terms in H except those describing the cavities. Thus, particle motion in magnets will be “geometric”, depending only on $\delta = (p - p_0)/p_0$ and not on the mass or absolute value of p . No approximation was made here so this Hamiltonian can just as well be used for hadron machines with the same benefits, *e.g.* in constructing a particle-tracking program, the algorithms for tracking through magnets will be simplified.

The price paid for these advantages comes through the more complicated expression of the time-dependence of A_s in (4.6). There, $t(z_t, p)$ denotes the solution of (4.7) for t in terms of z_t and p . This does not matter much since the motion of high energy electrons is extreme-relativistic and the third argument of A_s in (4.6) may be set equal to $-z_t/c$. Similarly, we can use the excellent approximation

$$t' \approx (1 + Gx)/c \quad (4.10)$$

Let us recall also the definition of the normalised magnetic field strength \mathbf{b} in (3.8).

4.2 RADIATION REACTION FORCES

Knowing all the essential properties of the synchrotron radiation, we can include its effect on the motion of the electrons by adding radiation reaction forces to (4.9). As a prelude to formulating the correct expressions for these forces, we first consider a single photon emission event.

Since τ_γ is short, the emission of a photon of energy u_j at azimuth $s = s_j$ (when $t = t_j$, say) will produce an abrupt change in the momentum but leave the particle at the same position in space—the particle cannot move instantaneously. At high energy the opening angle of the beam of radiation is⁴⁷

$$\theta_{\max} \simeq \frac{1}{2\gamma} \simeq 0.26 \text{ mrad} \quad \text{at } E_0 = 1 \text{ GeV}. \quad (4.11)$$

The 3-momentum vector $\mathbf{u}_j/c = u_j \mathbf{p}/pc$ of the photon is therefore almost collinear with the particle’s momentum \mathbf{p} . By applying momentum conservation, we can evaluate the changes in the canonical momenta. We are neglecting a minute increase in z_t due to the small reduction in the velocity of the electron (see (4.7)) and a small momentum exchange with virtual photons of the

guide field. The latter momentum exchange can be found by applying energy conservation and is approximately⁶⁰

$$\frac{1}{2} \frac{u}{c} \left(\frac{mc}{p} \right)^2 \simeq \frac{u}{2c\gamma^2}. \quad (4.12)$$

Calculating the change in each canonical momentum variable in its turn, we find

$$\begin{aligned} p &\mapsto p - u_j/c, \\ p_x &\mapsto p_x - \frac{u_j p_x}{c p} = p_x - \frac{u_j}{c} \frac{x'}{\sqrt{1 + (x')^2 + (y')^2}} = p_x - \frac{u_j x'}{c^2 t'}, \\ p_y &\mapsto p_y - \frac{u_j y'}{c^2 t'}. \end{aligned} \quad (4.13)$$

Consider a time-interval surrounding the moment of photon emission; let it be so short that the probability of more than one photon being emitted can be neglected. Then we know that the energy of the emitted photon has to be equal to the time-integral of the fluctuating radiation power (3.24),

$$u_j = \int_{s_j - \varepsilon}^{s_j + \varepsilon} P_X(s) ds/c \quad (\text{with probability} \rightarrow 1 \text{ as } \varepsilon \rightarrow 0^+), \quad (4.14)$$

and we may reinterpret (4.13) as stochastic differential equations (or, if we like, finite difference equations)

$$\begin{aligned} dp &= -P_X(s) dt/c = -P_X(s) t' ds/c = -P_X(s) z'_t ds/c^2 + O(\gamma^{-2}) ds, \\ dp_x &= -P_X(s) (x'/t') dt/c^2 = -P_X(s) x' ds/c^2, \\ dp_y &= -P_X(s) (y'/t') dt/c^2 = -P_X(s) y' ds/c^2. \end{aligned} \quad (4.15)$$

To complete the equations of motion, we must restore the forces due to the direct action of the external fields given by Hamilton's equations, $x' = \partial H/\partial p_x$, *etc.* Assembling everything, we have

$$\begin{aligned} x' &= \frac{\partial H}{\partial p_x}, & p'_x &= -\frac{\partial H}{\partial x} - \frac{P_X(s)}{c^2} \frac{\partial H}{\partial p_x}, \\ y' &= \frac{\partial H}{\partial p_y}, & p'_y &= -\frac{\partial H}{\partial y} - \frac{P_X(s)}{c^2} \frac{\partial H}{\partial p_y}, \\ z'_t &= \frac{\partial H}{\partial p}, & p' &= -\frac{\partial H}{\partial z_t} + \frac{P_X(s)}{c^2} \frac{\partial H}{\partial p}. \end{aligned} \quad (4.16)$$

The Hamiltonian part of these equations has already been written out explicitly in (4.8) and (4.9), but it is instructive to look at the radiation terms in detail:

$$\begin{aligned}
p'_x &= -\frac{\partial H}{\partial x} - (1 + Gx)pp_x[c_1b(x, y, s)^2 + \sqrt{c_2}b(x, y, s)^{3/2}\xi(s)] \\
&\stackrel{\text{def}}{=} -\frac{\partial H}{\partial x} + \frac{p_0}{c}\Pi_x, \\
p'_y &= -\frac{\partial H}{\partial y} - (1 + Gx)pp_y[c_1b(x, y, s)^2 + \sqrt{c_2}b(x, y, s)^{3/2}\xi(s)] \\
&\stackrel{\text{def}}{=} -\frac{\partial H}{\partial y} + \frac{p_0}{c}\Pi_y, \\
p'_z &= -\frac{\partial H}{\partial z_t} - (1 + Gx)p^2[c_1b(x, y, s)^2 + \sqrt{c_2}b(x, y, s)^{3/2}\xi(s)] \\
&\stackrel{\text{def}}{=} -\frac{\partial H}{\partial z_t} - \frac{p_0}{c}\Pi_t,
\end{aligned} \tag{4.17}$$

This also serves to define the *radiation coupling functions* Π_x , Π_y and Π_t .

Notice the interesting dependence of the radiation terms on the canonical momenta: in the radial and vertical phase planes, the factors of $1/p$ to which we drew attention in (4.8) have cancelled a factor of p in the expression (3.49) for $P_X(s)$. But this has not occurred in the longitudinal direction. This has some important consequences which are embodied in *Robinson's Theorem* on the damping partition numbers.^{60,19}

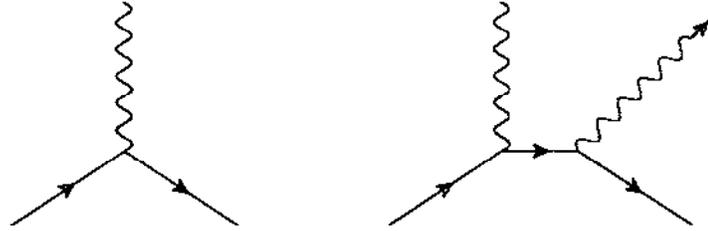


Fig. 6 “Feynman diagrams” corresponding to (4.17).

In a heuristic semi-classical picture, *not* proper quantum electrodynamics,⁶¹ the interpretation of (4.17) would associate the Hamiltonian terms with the (first-order) scattering of the electron off the virtual photons of the guide field. The radiation terms are associated with the (second-order) absorption of a virtual photon and its subsequent re-emission as a real photon (see Fig. 6). This association is manifested when the canonical variables are divided through by $\sqrt{\hbar}$ and the electric charge e is eliminated in favour of the fine structure constant $\alpha = e^2/\hbar c$.

Averaged over time, the radiation terms are small, although at the moment when a photon is radiated, they may dominate the equations of motion. Classical treatments^{47,48} of radiation reaction and radiation damping effects assume that, *in the rest frame of the particle*, the radiation reaction can be represented as a continuous force, small in comparison to the Lorentz force acting on the particle.

4.3 ROBINSON'S THEOREM AND GENERALIZATIONS

In an important early paper,⁶⁰ Robinson proved an elegant theorem concerning the damping rates of the normal modes of oscillation around the equilibrium orbit in circular electron accelerators.

The principal application of this theorem is nowadays found in a technique known as *damping partition number variation*. When we discuss this in a later section we shall see in detail how Robinson's theorem is satisfied in the representation of electron motion that we shall be using. However the theorem is rather more general than will be evident from this derivation alone so we shall pause here to give a more general and very much simpler proof, based on our formulation of the electron equations of motion (4.16).

Since these equations are local and contain the quantum fluctuations of the instantaneous radiation reaction forces, we shall actually deduce a result which is more general than that given by Robinson. Another difference, which simplifies the proof somewhat, is that we avail ourselves of the power of Hamiltonian mechanics.

The equations of motion (4.16) can be regarded as defining a flow in the 6-dimensional phase space of an electron in a storage ring. In the absence of the radiation terms this flow would be incompressible and have other invariants associated with the symplectic structure of a Hamiltonian system.²⁷ These invariants are of course completely independent of the particular canonical coordinates which we use for our calculations; they may be evaluated in terms of $\mathbf{X} = (x, y, z, p_x, p_y, p)$, or any other set. Similarly, if the radiation terms cause any of these invariants to change, then the *rate of change* of any invariant of the underlying Hamiltonian system can be calculated in terms of any set of canonical variables. Let us therefore calculate the total dissipation rate for our system; this is defined to be the (fractional) rate of change of phase-space volumes *i.e.*, the *divergence of the flow*. This gives an invariant measure of the extent to which Liouville's theorem is violated.

If we write (4.16) in the suggestive hydrodynamic form

$$\mathbf{X}' = \mathbf{V}(\mathbf{X}), \tag{4.18}$$

then the components of the local flow velocity are given by the right-hand sides of (4.16) and we may directly compute the divergence

$$\begin{aligned}
\nabla_{\mathbf{X}} \cdot \mathbf{V} &= -\frac{\partial}{\partial p_x} \left(\frac{P_X(s)}{c^2} \frac{\partial H}{\partial p_x} \right) - \frac{\partial}{\partial p_y} \left(\frac{P_X(s)}{c^2} \frac{\partial H}{\partial p_y} \right) \\
&\quad + \frac{\partial}{\partial p} \left(\frac{P_X(s)}{c^2} \frac{\partial H}{\partial p} \right) \\
&= \frac{P_X(s)}{c^2} \left(-\frac{\partial^2 H}{\partial p_x^2} - \frac{\partial^2 H}{\partial p_y^2} + \frac{\partial^2 H}{\partial p^2} \right) \\
&\quad + \frac{2P_X(s)}{c^2 p} \frac{\partial H}{\partial p},
\end{aligned} \tag{4.19}$$

where we recognized that, by virtue of Hamilton's equations, all the terms which do not contain $P_X(s)$ must cancel each other (Liouville's theorem). We also took advantage of the fact that $P_X(s)$, written out in (3.49), is directly proportional to p^2 and independent of p_x and p_y .

Since we want this proof to be as general as possible, we shall compute the derivatives in (4.19) in terms of the exact Hamiltonian (4.6). The only part of this which might give us some trouble is the vector potential term, and this only insofar as it depends on the total momentum p through its functional dependence on $t(z_t, p)$. However these terms only matter if there exist values of s for which the Hamiltonian is time-dependent (*e.g.* an RF cavity) and for which the magnetic field (hence $P_X(s)$) does not vanish. Since the synchrotron radiation induced by the oscillating magnetic fields in RF cavities is normally negligible and RF cavities are not placed in regions of static magnetic field, these terms may be neglected for all practical purposes.

We are left with the evaluation of the terms arising from the kinematic part of (4.6), *i.e.* the square root of a quadratic form in the momenta. This is quite straightforward: we have, *e.g.*

$$\frac{\partial^2 H}{\partial p_x^2} = (1 + G(s)x) \frac{p^2 - p_y^2}{(p^2 - p_x^2 - p_y^2)^{3/2}}, \tag{4.20}$$

and when we assemble all the terms (taking due care over the sign change in the longitudinal component of \mathbf{V}), we get

$$\nabla_{\mathbf{X}} \cdot \mathbf{V} = -(1 + G(s)x) \frac{4P_X(s)}{c^2 \sqrt{p^2 - p_x^2 - p_y^2}}. \tag{4.21}$$

This gives the phase space compression with respect to our independent variable s . To express it in terms of time t as independent variable, as Robinson did, we simply have to derive an exact expression for t' from the Courant-Snyder Hamiltonian (4.3),

$$t' = (1 + G(s)x) \frac{E(p)/c^2}{\sqrt{p^2 - p_x^2 - p_y^2}}. \quad (4.22)$$

Effecting the change of independent variable, we have, for the rate of change of phase space volumes with respect to time, the simple expression

$$\left\{ \begin{array}{l} \text{Phase space compression} \\ \text{per unit time} \end{array} \right\} = -4P_X(s)/E \quad (4.23)$$

Conventionally, the total damping rate is defined to be one-half of the negative of the rate of change of phase-space volumes (this is readily seen by looking at the definition of the damping rate for the standard textbook example of a damped harmonic oscillator) and is therefore given by

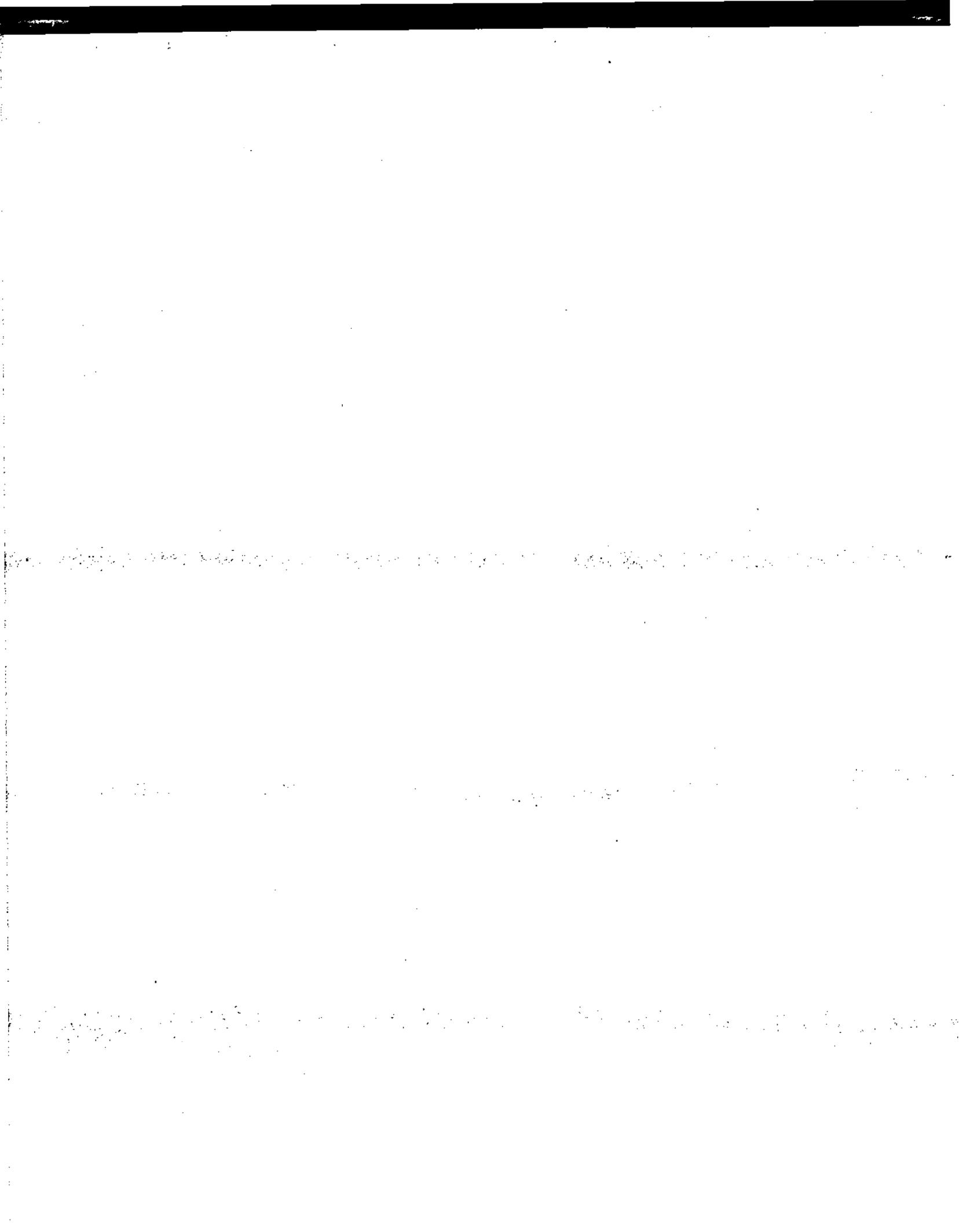
$$\alpha_{\text{tot}} = -2P_X(s)/E. \quad (4.24)$$

This is the promised generalization of Robinson's theorem. It holds *locally* at every point on the circumference of the machine and the only restriction on its validity is the one given above, that synchrotron radiation is not emitted in places where the externally applied fields are time-dependent (or that such effects are small). No kinematic approximations were made in deriving it.

By virtue of the underlying symplectic structure of the phase space, it follows that this quantity will be invariant under all canonical transformations to new sets of coordinates and momenta—provided the independent variable is not transformed. In addition, since it refers to *fractional* changes in phase space volume, it continues to hold after any rescaling of some of the variables which preserves the Hamiltonian form of the equations of motion (we mention this since we shall do some such rescalings later).

In particular we can deduce that, when we later find a canonical transformation to the normal modes of particle motion in the neighbourhood of an equilibrium orbit, then the damping rates of these modes must satisfy this sum rule. This applies even if one or two of the modes are anti-damped (*i.e.* unstable by virtue of radiation effects).

The power $P_X(s)$ appearing in (4.21) is the full fluctuating power so we really now have a version of Robinson's theorem which takes account of quantum fluctuations. We can use it, for example, to evaluate the change in emittances when a single photon is emitted.



Robinson's version of the theorem may be recovered by averaging this result around the ring *with the further assumption that the total energy lost by the particle is small compared to its total energy E* . Then, if we define the total energy loss per turn as

$$U = \int_0^{2\pi R} P_X(s) t' ds, \quad (4.25)$$

we find the approximate result

$$\alpha_{\text{tot}} \simeq 2U f_0 / E, \quad \text{for } U/E \ll 1. \quad (4.26)$$

This is the original version of Robinson's theorem.

4.3.1 ♠ Lorentz-invariant Robinson theorem

From (4.21) it is possible to derive a Lorentz-invariant version of Robinson's theorem. This is a simple matter of the transformation from the time in the laboratory frame of reference into the proper time for the particle. The rate of change of laboratory time with respect to proper time is

$$\frac{dt}{d\tau} = \frac{1}{\gamma} = \frac{E}{mc^2}, \quad (4.27)$$

and therefore the damping rate of phase space coordinates with respect to proper time is

$$\alpha_\tau = \frac{dt}{d\tau} \alpha_{\text{tot}} = -\frac{2P_X(s)}{mc^2}, \quad (4.28)$$

which is manifestly a Lorentz scalar since both the radiation power and the mass of the particle are invariant under Lorentz transformations.

Put into words, this says that the rate of compression of phase-space volume with respect to the particle's proper time is equal to 4 times the radiation power measured in units of the particle's rest energy. Conversely, the factor of E which appears in (4.24) or (4.26) can be seen as arising from a Lorentz transformation.

5. CLOSED ORBIT AND NORMAL MODES

In equations (4.16) we have a complete description of electron motion under the combined influences of the applied electromagnetic fields and synchrotron radiation. Although much of their physical content is evident, the dynamical variables are coupled in complicated ways so that these equations are not the most suitable for practical calculations.

Indeed we know, from the theory of beam optics, that a better way to describe particle trajectories is in terms of the three normal modes of linearised motion around the closed orbit in phase space. In the process one introduces a number of optical functions (dispersion, β -functions and the like), defined at each point of the circumference. These functions are determined by the magnetic structure of the storage ring and also determine the frequencies of the linear modes. In a planar ring, with no skew magnetic elements, such as we have assumed, the normal modes of linearised motion are the betatron and synchrotron oscillations.

In the coordinates (x, y, z_t) , the Hamiltonian contains linear coupling terms between x and p due to the spectrometer effect of the (horizontal) bending magnets, but these may be eliminated by introducing the *dispersion functions*. The main point in describing this here is to show how the energy loss by synchrotron radiation is coupled into the transverse oscillations through these functions.

We shall make the customary canonical transformation whereby the total momentum is replaced, in the list of variables, by its deviation from the reference value p_0 . In a less customary refinement we shall then further split this momentum deviation into two components δ_s and ε ; δ_s will define an off-energy closed orbit and ε will be the canonical momentum for synchrotron oscillations. It is important to keep this distinction clear.

5.1 DISPERSION FUNCTIONS

To simplify the Hamiltonian (4.6) we neglect higher order kinematic terms in the transverse momenta and in the momentum deviation δ ; more explicitly, we approximate the square root term by

$$(1 + Gx)\sqrt{p^2 - p_x^2 - p_y^2} \simeq (1 + Gx)p - \frac{p_x^2 + p_y^2}{2p} + \dots \quad (5.1)$$

Then we perform a simple rescaling of the Hamiltonian and canonical momenta with the constant value p_0 :

$$\begin{aligned} H &\mapsto H_1 = H/p_0, & p_x &\mapsto P_x = p_x/p_0, \\ p &\mapsto P = p/p_0 & p_y &\mapsto P_y = p_y/p_0, \end{aligned} \quad (5.2)$$

After these manoeuvres, all the momenta are dimensionless and the canonical coordinates and the independent variable all have dimensions of length; the Hamiltonian is

$$\begin{aligned}
H_1(x, y, z_t, p_x, p_y, p; s) \simeq & -Gx(P - 1) + \frac{P_x^2 + P_y^2}{2P} \\
& + \frac{G^2 x^2}{2} + \frac{1}{2}K_1(x^2 - y^2) + \frac{1}{6}K_2(x^3 - 3xy^2) + \dots \\
& - \sum_k \frac{e\hat{V}_k}{p_0\omega_{rf}} \delta_C(s - s_k) \cos(\omega_{rf}z_t/c + \phi_k).
\end{aligned} \tag{5.3}$$

One way of looking at the dispersion function η is to think of it as a device which eliminates the linear coupling appearing in the first term of (5.3). To this extent, its definition is universal. By allowing it to depend on momentum, however, it is possible to do away with some higher-order couplings at the same time; such an approach is very useful in nonlinear optics studies when synchrotron motion is neglected, *e.g.* for hadron colliders where the synchrotron oscillation tune is so small that p can be taken as a constant. There are at least two ways of setting up this dependence.⁵⁹

In an electron ring, on the other hand, the value of p oscillates relatively rapidly although still much more slowly than the transverse coordinates. Nevertheless p must certainly be included as a dynamical variable and it appears at first sight that the simplest approach would be to define η with respect to the reference momentum p_0 ; indeed this is the "ideal" dispersion function printed out by most optics programs. From the analytical point of view, this has the advantage that the Hamiltonian does not depend on a canonical momentum through functions which can only be calculated (and, eventually, differentiated) numerically. From the experimental point of view there is the convenience of a direct measure of the deviation of the off-momentum equilibrium orbits from the reference closed orbit, here presumed to pass through the reference points of the beam position monitors at the centres of the magnet apertures. This dispersion function will be denoted $\eta(0, s)$ in what follows.

We shall see that, depending on the precise value of the RF frequency, the equilibrium momentum of the beam may not be equal to p_0 ; synchrotron oscillations will then take place around a slightly different value of the momentum which we shall denote as $p_0(1 + \delta_s)$, with $\delta_s \ll 1$. In electron ring theory, then, a compromise definition of η as a function of δ_s seems the most useful. The dispersion $\eta(\delta_s, s)$ and its conjugate function $\zeta(\delta_s, s)$ may be introduced by means of a canonical transformation

$$(x, y, z_t, P_x, P_y, P) \mapsto (x_\beta, y, z, p_\beta, p_y, \varepsilon) \quad (5.4)$$

of the type whose generating function,

$$F_2(p_\beta, p_y, \varepsilon, x, y, z_t) = p_\beta [x - \eta(\delta_s, s)(\delta_s + \varepsilon)] + x\zeta(\delta_s, s)(\delta_s + \varepsilon) + p_y y + (1 + \delta_s + \varepsilon) [z_t + Z_0(s)], \quad (5.5)$$

depends on the new momenta and the old coordinates.

Here $Z_0(s)$, $\eta(\delta_s, s)$ and $\zeta(\delta_s, s)$ are to be interpreted as unspecified functions of s . We are at liberty to choose them as we will but natural choices will emerge in due course. In order to take proper account of chromatic effects in cases where the equilibrium value of p is other than p_0 , they have also been allowed to depend *parametrically* on the constant δ_s . Later we shall show how the value of δ_s is determined naturally by the RF frequency. When the equations of motion are constructed from the new Hamiltonian there is no need to differentiate η or ζ with respect to δ_s .

According to the familiar rules,²⁷ the relations between the old and new coordinates and momenta are given by

$$\begin{aligned} x_\beta &= \frac{\partial F_2}{\partial p_\beta} = x - \eta(\delta_s, s)(\delta_s + \varepsilon), & P_x &= \frac{\partial F_2}{\partial x} = p_\beta + \zeta(\delta_s + \varepsilon), \\ y &= \frac{\partial F_2}{\partial p_y} = y, & P_y &= \frac{\partial F_2}{\partial y} = p_y, \\ z &= \frac{\partial F_2}{\partial \varepsilon} = z_t + Z_0(s) - \eta p_\beta + \zeta x_\beta + \eta(\delta_s + \varepsilon), & P &= \frac{\partial F_2}{\partial z_t} = 1 + \delta_s + \varepsilon, \end{aligned} \quad (5.6)$$

and the new Hamiltonian is

$$H_2 = H_1 + \frac{\partial F_2}{\partial s} = H - p_\beta(\delta_s + \varepsilon)\eta' + [x_\beta + \eta(\delta_s + \varepsilon)](\delta_s + \varepsilon)\zeta'. \quad (5.7)$$

The splitting of the radial displacement x into its betatron and “energy” components should be familiar. It is perhaps less well-known that, in order to preserve the canonical structure (symplecticity), one must also use a new longitudinal coordinate, z , which takes account of local changes in the length of the particle’s orbit due to its betatron and synchrotron oscillations. Since we have assumed no vertical bending and a perfect machine, there is no vertical dispersion and the transformations in the y - p_y phase plane are trivial.

Expressing the new Hamiltonian H_2 in terms of the new coordinates, we can eliminate coupling terms linear in x_β by imposing the conditions

$$\frac{\partial H_2}{\partial x_\beta} = \frac{\partial H_2}{\partial p_\beta} = 0 \quad \text{for } x_\beta = p_\beta = y = p_y = z = \varepsilon = 0. \quad (5.8)$$

Writing these out explicitly, we find that η and ζ must satisfy a pair of first-order differential equations and a periodicity condition,

$$\eta' = \frac{\zeta}{1 + \delta_s}, \quad \zeta' = G - (K_1 + G^2)\eta - \frac{1}{2}K_2\eta^2\delta_s, \quad (5.9)$$

$$\eta(\delta_s, s + 2\pi R) = \eta(\delta_s, s),$$

which are nothing but the familiar equations defining the dispersion function. The separate notation for $\zeta(\delta_s, s)$ has been retained in order to show that (5.9) can themselves be derived from a Hamiltonian. To avoid confusion with other systems of notation, we stress again that δ_s appears as a *parameter* and that primes denote differentiation with respect to s .

A common practical means of determining these functions for a range of values of δ_s is to expand them as

$$\eta(\delta_s, s) = \eta_0(s) + \eta_1(s)\delta_s + \dots, \quad \zeta(\delta_s, s) = \zeta_0(s) + \zeta_1(s)\delta_s + \dots, \quad (5.10)$$

and equate coefficients of δ_s in the equations (5.9). Then each function may be evaluated once and for all, independently of δ_s .

Since the canonical transformation was linear, we can work out the new equations of motion from (4.17) and (5.6) with little trouble:

$$\begin{aligned} x'_\beta &= \frac{\partial H_2}{\partial p_\beta} + \Pi_t \eta / c, & p'_\beta &= -\frac{\partial H_2}{\partial x_\beta} + (\Pi_x + \Pi_t \zeta) / c, \\ y' &= \frac{\partial H_2}{\partial p_y}, & p'_y &= -\frac{\partial H_2}{\partial y} + \Pi_y / c, \\ z' &= \frac{\partial H_2}{\partial \varepsilon} - \Pi_x \eta / c - \Pi_t \zeta \eta / c, & \varepsilon' &= -\frac{\partial H_2}{\partial z} - \Pi_t / c. \end{aligned} \quad (5.11)$$

In more complicated cases we would need the analogue of the transformation formulae given in (2.54).

Defining an effective quadrupole gradient for a particle with the reference momentum at the position of the off-momentum orbit,

$$k_1(\delta_s, s) = K_1(s) + \frac{1}{2}K_2(s)\eta(\delta_s, s)\delta_s \quad (5.12)$$

we find, after a good deal of algebra, exploiting the cancellations implied by (5.8) and dropping several constant terms, that the Hamiltonian is

$$\begin{aligned}
H_2(x_\beta, y, z, p_\beta, p_y, \varepsilon; s) = & \frac{1}{2} \left\{ \frac{\zeta^2}{1 + \delta_s} - (G^2 + k_1)\eta^2 \right\} (\delta_s + \varepsilon)^2 \\
& + [Z_0'(s) - 1] \varepsilon + \frac{1}{6} K_2 (\varepsilon - \delta_s/2) (\delta_s + \varepsilon)^2 \\
& + \frac{p_\beta^2 + p_y^2}{2(1 + \delta_s)} + \frac{G^2 x^2}{2} \\
& + \frac{1}{2} k_1 (x_\beta^2 - y^2) + \frac{1}{6} K_2 (x_\beta^3 - 3x_\beta y^2) \\
& - \sum_k \frac{e\hat{V}_k}{p_0\omega_{\text{rf}}} \delta_C(s - s_k) \\
& \times \cos \left\{ \frac{\omega_{\text{rf}}}{c} [z - Z_0(s) + \eta p_\beta - \zeta x_\beta - \zeta \eta (\delta_s + \varepsilon) + \phi_k] \right\}.
\end{aligned} \tag{5.13}$$

5.2 LOCAL SYNCHROTRON MOTION

Although the terms describing betatron motion are simplified, the local formulation of synchrotron motion appears fairly complicated. However we recall that, in order to avoid dangerous synchro-betatron coupling effects, storage rings are now usually designed to make the dispersion functions vanish at the locations of the RF cavities:

$$\eta(s_k) = \zeta(s_k) = 0, \quad \text{for each } k. \tag{5.14}$$

(In practice, of course, imperfections will usually create some horizontal and vertical dispersion in the cavities.) Then, thanks to the δ -functions, the phase of the cosine describing the RF waveform simplifies to

$$\Phi(z, s) = \frac{\omega_{\text{rf}}}{c} [z - Z_0(s) + \phi_k], \tag{5.15}$$

and we have succeeded in eliminating all coupling effects between the longitudinal and transverse motions. We remain free to choose the function $Z_0(s)$ to our best advantage. Analogy with (5.8) suggests that we should impose the condition

$$\frac{\partial H_2}{\partial \varepsilon} = 0 \quad \text{for } x_\beta = p_\beta = y = p_y = z = \varepsilon = 0. \tag{5.16}$$

To interpret this physically, we note that (in the absence of radiation effects) the change in z around an orbit will be

$$z(2\pi R) - z(0) = \int_0^{2\pi R} z' ds = \int_0^{2\pi R} \frac{\partial H_2}{\partial \varepsilon} ds, \quad (5.17)$$

and the condition (5.16) determines a shift in the origin of phase space to a *fixed point* of the mapping which describes the evolution of the phase space coordinates over one turn. Furthermore this must hold true at every point on the circumference. In other words, we have made the canonical transformation to a reference frame moving with the synchronous particle.

Working out (5.16) explicitly, we find that the effects of the sextupole terms in the longitudinal part of the Hamiltonian cancel, leaving us with

$$\left\{ \frac{\zeta^2}{1 + \delta_s} - (G^2 + k_1)\eta^2 \right\} \delta_s + Z_0'(s) - 1 = 0. \quad (5.18)$$

Integrating this, we are led to

$$Z_0(s) = -z_s + s - \delta_s \int_0^s \Gamma(\delta_s, s) ds, \quad (5.19)$$

where z_s is a constant, related to the *stable phase angle*. The local path length slippage function is defined by

$$\Gamma(\delta_s, s) \stackrel{\text{def}}{=} \frac{\zeta^2}{1 + \delta_s} - (G^2 + k_1)\eta^2. \quad (5.20)$$

In the following subsection we provide details of the Fourier analysis of this function. It turns out that its average value is just the negative of the *momentum compaction factor*, α_c . By suitably averaging the Hamiltonian around the ring we shall construct the *smooth approximation* to synchrotron motion from this local description.

5.2.1 Equilibrium momentum

Neglecting some unimportant sextupole terms, the Hamiltonian for local synchrotron motion is now

$$\begin{aligned}
 H_s(z, \varepsilon, s) = & \frac{1}{2} \Gamma(\delta_s, s) \varepsilon^2 \\
 & - \sum_k \frac{e \hat{V}_k}{p_0 c \omega_{\text{rf}}} \delta_C(s - s_k) \\
 & \times \cos \left\{ \frac{\omega_{\text{rf}}}{c} (z + z_s) - \frac{\omega_{\text{rf}}}{c} s + \frac{\omega_{\text{rf}} \delta_s}{c} \int_0^s \Gamma(\delta_s, \sigma) d\sigma + \phi_k \right\}.
 \end{aligned} \tag{5.21}$$

Again, the Hamiltonian has to be periodic in s :

$$H_s(z, \varepsilon, s + 2\pi R) = H_s(z, \varepsilon, s). \tag{5.22}$$

From this condition it follows that the argument of the cosine must advance by an integer multiple of 2π per revolution. In the limit $\delta_s \rightarrow 0$, the RF frequency is an integer multiple of the revolution frequency on the reference orbit:

$$\omega_{\text{rf}} = 2\pi f_{\text{rf}} = 2\pi h f_0 = \frac{hc}{R}, \quad (\delta_s = 0), \tag{5.23}$$

where h is called the *harmonic number*. Therefore (5.22) is equivalent to

$$\begin{aligned}
 & -\frac{\omega_{\text{rf}}}{c} (2\pi R) + \frac{\omega_{\text{rf}} \delta_s}{c} \int_0^{2\pi R} \Gamma(\delta_s, s) ds = -2\pi h \\
 \Leftrightarrow & \quad \delta_s = -\frac{1}{\alpha_c(\delta_s)} \left(1 - \frac{hc}{\omega_{\text{rf}} R} \right)
 \end{aligned} \tag{5.24}$$

showing that the equilibrium momentum of the beam may be adjusted by making small shifts of the RF frequency.

We shall come back to (5.24) after we analyse the terms in (5.21) individually.

5.2.2 ♠ Local momentum compaction

Saying that the synchrotron tune is small means that a single synchrotron oscillation takes many turns of the machine. The largest synchrotron tunes are realised in large machines like LEP and are of the order of 0.1. To simplify the description of this motion, we shall make a Fourier decomposition of the Hamiltonian on the circumference and retain only the most slowly-varying terms.

From (5.21) the “kinetic energy” of synchrotron oscillations is given by the term

$$\frac{1}{2}\Gamma(\delta_s, s)\varepsilon^2 = \frac{1}{2}\left\{\frac{\zeta^2}{1+\delta_s} - (G^2 + k)\eta^2\right\}\varepsilon^2 = \frac{1+\delta_s}{2}\{(\eta\eta')' - G\eta\}\varepsilon^2, \quad (5.25)$$

where we used (5.9). We make a Fourier analysis of the function

$$\Gamma(\delta_s, s) = \sum_{n=-\infty}^{\infty} \Gamma_n(\delta_s)e^{ins/R}, \quad (5.26)$$

where the coefficients will be given by

$$\Gamma_n(\delta_s) = \frac{1}{2\pi R} \int_0^{2\pi R} ds \Gamma(\delta_s, s)e^{-ins/R}. \quad (5.27)$$

Integrating by parts twice and using periodicity arguments, we see that

$$\Gamma_n(\delta_s) = \Gamma_{-n}^*(\delta_s) = -\frac{1+\delta_s}{2\pi R} \int_0^{2\pi R} \eta(\delta_s, s) \left[\frac{n^2\eta(\delta_s, s)}{2R^2} + G(s) \right] e^{ins/R} ds. \quad (5.28)$$

In particular, the constant term is simply

$$\begin{aligned} \Gamma_0(\delta_s) &= \frac{1+\delta_s}{2\pi R} \int_0^{2\pi R} \left\{ \frac{\zeta^2}{1+\delta_s} - (G^2 + K_1)\eta^2 \right\} ds \\ &= -\frac{1+\delta_s}{2\pi R} \int_0^{2\pi R} G\eta ds \\ &= -\alpha_c(\delta_s), \end{aligned} \quad (5.29)$$

where $\alpha_c(\delta_s)$ is called the *momentum compaction factor*. Like the dispersion functions, α_c may be expanded in powers of δ_s . Generally, $\alpha_c \ll 1$.

Neglect of the higher harmonics of $\Gamma(\delta_s, s)$ (and the RF voltage terms) gives the *smooth approximation*. In addition, simple scaling arguments show that the first term in the integral (5.28) is much smaller than the second.

To simplify the analysis of the RF voltage distribution we define the function

$$\Sigma(\delta_s, s) = - \int_0^s \Gamma(\delta_s, \sigma) d\sigma, \quad (5.30)$$

which gives the increase in path length per unit momentum deviation δ_s in the sector of the ring between azimuths 0 and s .

5.2.3 ♠ Effective RF voltage

The Fourier analysis of the RF voltage term in (5.21), may be made by substituting the identity

$$\delta_C(s) = \frac{1}{2\pi R} \sum_{n=-\infty}^{\infty} e^{ins/R} \quad (5.31)$$

and expanding the cosine into complex exponentials. The result is

$$-\frac{1}{2\pi R} \sum_k \frac{e\hat{V}_k}{p_0\omega_{rf}} \frac{1}{2} \sum_{n=-\infty}^{\infty} \left\{ \begin{aligned} &\exp \left[\frac{in(s-s_k)}{R} + \frac{h(z+z_s)}{\mathcal{R}(\delta_s)} - \frac{h}{\mathcal{R}(\delta_s)} [s + \Sigma(\delta_s, s)\delta_s] + \phi_k \right] \\ &\exp \left[\frac{in(s-s_k)}{R} - \frac{h(z+z_s)}{\mathcal{R}(\delta_s)} + \frac{h}{\mathcal{R}(\delta_s)} [s + \Sigma(\delta_s, s)\delta_s] - \phi_k \right] \end{aligned} \right\}. \quad (5.32)$$

Bearing in mind the rôle of the RF harmonic number, we recognize that nearly all the terms in the expansion are rapidly oscillating functions of the independent variable s and will not produce significant average effects on the beam. The terms which *do* count are those in which the s -dependence can be made to cancel from the arguments of the exponentials. From (5.23), it follows that there are precisely two of these, namely the term with $n = +h$ in the first group and that with $n = -h$ in the second. Happily, these two terms combine to reassemble a slowly-varying cosine function:

$$-\frac{1}{2\pi R} \sum_k \frac{e\hat{V}_k}{p_0\omega_{rf}} \cos \left\{ h \left(\frac{s}{R} - \frac{s}{\mathcal{R}(\delta_s)} - \frac{\Sigma(\delta_s, s)\delta_s}{\mathcal{R}(\delta_s)} \right) - \frac{hs_k}{R} + \frac{h(z+z_s)}{\mathcal{R}(\delta_s)} + \phi_k \right\}. \quad (5.33)$$

We separate the function Σ into a contribution from the $n = 0$ term in (5.26) and a remainder:

$$\Sigma(\delta_s, s) = \alpha_c(\delta_s)s + \tilde{\Sigma}(\delta_s, s) \quad (5.34)$$

where, by virtue of (5.25) and (5.26),

$$\tilde{\Sigma}(\delta_s, s) = \sum_{n=1}^{\infty} \frac{1}{n\pi R} \int_0^{2\pi R} \Gamma(\delta_s, \sigma) \sin \left[\frac{n(s - \sigma)}{R} \right]. \quad (5.35)$$

Then (treating the two terms in (5.34) differently with respect to the original δ -functions) the s -dependence cancels and the argument of the cosine becomes

$$-\frac{hs_k}{R} + \frac{h(z + z_s)}{\mathcal{R}(\delta_s)} - \frac{h\delta_s}{\mathcal{R}(\delta_s)} \tilde{\Sigma}(\delta_s, s_k) + \phi_k. \quad (5.36)$$

It is clear that for the most efficient use of the RF system (we ignore all collective effects here) we should choose the relative phases of the cavities according to

$$\phi_k = \frac{hs_k}{R} + \frac{h\delta_s}{\pi R^2(1 + \alpha_c\delta_s)} \sum_{n=1}^{\infty} \frac{1}{n} \int_0^{2\pi R} \Gamma(\delta_s, s) \sin \left[\frac{n(s - s_k)}{R} \right] ds, \quad (5.37)$$

so that (5.32) finally simplifies to

$$-\frac{e\hat{V}}{2\pi p_0 c h} \cos \left[\frac{h(z + z_s)}{\mathcal{R}(\delta_s)} \right]. \quad (5.38)$$

The total effective peak voltage is

$$\hat{V} = \sum_k \hat{V}_k. \quad (5.39)$$

which is clearly the greatest value it can have.

5.3 OFF-MOMENTUM CLOSED ORBIT

By means of the smooth approximation, we have replaced an arbitrary set of discrete, but properly phased, cavities with h periods of a standing wave, distributed over the whole circumference of the (off-momentum) equilibrium orbit.

Writing $f_{rf} = hf_0 + \Delta f_{rf}$, we may write the familiar linearised version of (5.24),

$$\delta_s \approx -\frac{1}{\alpha_c(0)} \frac{\Delta f_{rf}}{f_{rf}}, \quad (5.40)$$

and exhibit the dependence of the average radius of the equilibrium orbit on the momentum compaction factor:

$$\begin{aligned} \mathcal{R}(\delta_s) &\stackrel{\text{def}}{=} \frac{hc}{\omega_{rf}} \sqrt{1 + \frac{m^2 c^2}{p_0^2 (1 + \delta_s)^2}} \\ &\simeq \frac{hc}{\omega_{rf}} = R (1 + \alpha_c(\delta_s) \delta_s) \\ \Rightarrow \frac{1}{\mathcal{R}} \frac{d\mathcal{R}}{d\delta_s} &= \alpha_c(\delta_s). \end{aligned} \quad (5.41)$$

In the previous two subsections it was shown that, in smooth approximation, (5.13) may be replaced by the simplified Hamiltonian

$$\begin{aligned} H_3(x_\beta, y, z, p_\beta, p_y, \varepsilon; s) &= -\frac{\alpha_c \varepsilon^2}{2} - \frac{e\hat{V}}{2\pi p_0 c h} \cos(h(z + z_s)/\mathcal{R}) \\ &+ \frac{p_\beta^2 + p_y^2}{2(1 + \delta_s)} + \frac{G^2 x^2}{2} + \frac{1}{2} k_1 (x_\beta^2 - y^2) \\ &+ \frac{1}{6} K_2 (x_\beta^3 - 3x_\beta y^2). \end{aligned} \quad (5.42)$$

Including the radiation reaction effects, the equations of motion are given by (5.11) with H_3 substituted for H_2 . To make the radiation terms explicit, we have to work forward through the chain of variable substitutions from the original forms of Π_x and Π_t as functions of (x, p, s) . These operations are deferred to the next section.

First however, it is natural to make a further rescaling of the momenta with respect to $p_s = p_0(1 + \delta_s)$, the momentum of the equilibrium orbit, instead of the reference momentum p_0 . Thus, if we make the substitutions

$$\begin{aligned} H_3 \mapsto H_4 &= H_3/(1 + \delta_s) & p_y \mapsto \tilde{p}_y &= p_y/(1 + \delta_s) \\ p_\beta \mapsto \tilde{p}_\beta &= p_\beta/(1 + \delta_s) & \varepsilon \mapsto \tilde{\varepsilon} &= \varepsilon/(1 + \delta_s) \end{aligned} \quad (5.43)$$

and define renormalized gradient and momentum compaction functions:

$$\begin{aligned} K_x(s) &= \frac{G(s)^2 + k_1(s)}{1 + \delta_s} & \tilde{K}_2(s) &= \frac{K_2(s)}{1 + \delta_s} \\ K_y(s) &= -\frac{k_1(s)}{1 + \delta_s} & \tilde{\alpha}_c &= (1 + \delta_s)\alpha_c \end{aligned} \quad (5.44)$$

we can include longitudinal and transverse momentum-dependent effects in a renormalized Hamiltonian which takes the simple form

$$\begin{aligned}
H_4(x_\beta, y, z, \tilde{p}_\beta, \tilde{p}_y, \tilde{\epsilon}; s) = & -\frac{\tilde{\alpha}_c \tilde{\epsilon}^2}{2} - \frac{e\hat{V}}{2\pi p_s c h} \cos(h(z + z_s)/R) \\
& + \frac{\tilde{p}_\beta^2 + \tilde{p}_y^2}{2} + \frac{1}{2}(K_x x_\beta^2 + K_y y^2) \\
& + \frac{1}{6}\tilde{K}_2(x_\beta^3 - 3x_\beta y^2).
\end{aligned} \tag{5.45}$$

Having achieved this simplification, we can now drop the tildes on the quantities introduced in (5.43) and (5.44); it will only be necessary for us to remember to substitute p_s for p_0 in some of the expressions for the radiation effects.

The physical significance of this last transformation was to shift our point of reference from the original centred reference orbit to the true equilibrium orbit determined by the RF frequency. All momenta are now scaled with respect to p_s , the equilibrium momentum on this orbit instead of p_0 . Of course the scaling could have been made in (5.2) by using p_s instead of p_0 but the physical motivation would not have been obvious at that point and the algebra of the separation of the total momentum deviation into δ_s and $\tilde{\epsilon}$ would have been somewhat more complicated.

5.3.1 ♠ Relation to transition energy

The formulation of synchrotron motion given here may appear to have somehow neglected the possibility of the revolution frequency's increasing with momentum as it does below transition energy. Indeed electron storage rings are always above transition but this appearance is only a consequence of our having used s , and not time t , as independent variable. The phenomena associated with transition energy can indeed be recovered from (5.42) when the increase in *time* taken to cover a greater orbit length due to a momentum deviation is exactly compensated by the greater velocity of the particle on that orbit. When the transformation of independent variable is made, including some higher order terms in (4.10), the velocity (and hence the familiar γ^{-2} factor) enters explicitly. A canonical transformation⁶² then restores variable sign in the coefficient of ϵ^2 .

5.4 BETATRON MOTION

To study the transverse beam dynamics with radiation, we first select the terms constituting the Hamiltonian for *linear* betatron motion from (5.45)

$$H_\beta(x_\beta, y, p_\beta, p_y; s) = \frac{p_\beta^2 + p_y^2}{2} + \frac{1}{2}(K_x x_\beta^2 + K_y y^2) \quad (5.46)$$

where the sextupole terms have been dropped since, for the remainder of our discussion we shall not concern ourselves with chromatic or nonlinear effects.

Substituting H_β for H_2 , the equations of motion are given by (5.11) corrected for our last rescaling:

$$\begin{aligned} x'_\beta &= \frac{\partial H_\beta}{\partial p_\beta} + \Pi_t \eta / c, & p'_\beta &= -\frac{\partial H_\beta}{\partial x_\beta} + (\Pi_x + \Pi_t \zeta) p_0 / (p_s c), \\ y' &= \frac{\partial H_\beta}{\partial p_y}, & p'_y &= -\frac{\partial H_\beta}{\partial y} + \Pi_y p_0 / (p_s c). \end{aligned} \quad (5.47)$$

5.4.1 Action-angle variables

Action-angle variables for the Hamiltonian are obtained through a canonical transformation which introduces the β - and α -functions of Courant and Snyder^{56,59}:

$$(x_\beta, y, p_\beta, p_y) \mapsto (\psi_x, \psi_y, I_x, I_y). \quad (5.48)$$

The generating function for this transformation is

$$\begin{aligned} S(\psi_x, x_\beta) &= -\frac{x_\beta^2}{2\beta_x(s)} \{ \alpha_x(s) + \tan \phi_x(\psi_x, s) \} \\ &\quad - \frac{y^2}{2\beta_y(s)} \{ \alpha_y(s) + \tan \phi_y(\psi_y, s) \}, \end{aligned} \quad (5.49)$$

where we have defined two phase functions depending on the phases $\psi_{x,y}$ and s :

$$\phi_{x,y}(\psi_{x,y}, s) = \psi_{x,y} + \int_0^s \left\{ \frac{1}{\beta_{x,y}(\sigma)} - \frac{\nu_{x,y}}{R} \right\} d\sigma. \quad (5.50)$$

Note that the actual dynamical variables are $\psi_{x,y}$ and that $\phi_{x,y}$ are just convenient auxiliary functions.

Solving the transformation equations, we can express the old coordinates in terms of the new; the results are the familiar expressions

$$\begin{aligned}
x_\beta &= \sqrt{2\beta_x(s)I_x} \cos \phi_x(\psi_x, s), \\
p_\beta &= -\sqrt{\frac{2I_x}{\beta_x(s)}} \{ \alpha_x(s) \cos \phi_x(\psi_x, s) + \sin \phi_x(\psi_x, s) \}, \\
y &= \sqrt{2\beta_y(s)I_y} \cos \phi_y(\psi_y, s), \\
p_y &= -\sqrt{\frac{2I_y}{\beta_y(s)}} \{ \alpha_y(s) \cos \phi_y(\psi_y, s) + \sin \phi_y(\psi_y, s) \},
\end{aligned} \tag{5.51}$$

Provided we choose the arbitrary functions $\alpha_{x,y}(s)$, $\beta_{x,y}(s)$ to be the periodic solutions of the differential equations

$$\alpha_{x,y} = -\beta'_{x,y}/2, \quad \alpha'_{x,y} = K_{x,y}\beta_{x,y} - (\alpha_{x,y}^2 + 1)/\beta_{x,y}, \tag{5.52}$$

the coefficients of the terms in $\cos(2\phi_{x,y})$ and $\sin(2\phi_{x,y})$ in the new Hamiltonian will vanish. Provided also that we choose the constants $\nu_{x,y}$ to be the *betatron tunes*,

$$\nu_{x,y} = \frac{1}{2\pi} \int_0^{2\pi R} \frac{ds}{\beta_{x,y}(s)}, \tag{5.53}$$

the s -dependence of the of the Hamiltonian for linear betatron motion is eliminated and we have, simply,

$$H_5(\psi_x, \psi_y, I_x, I_y) = \nu_x I_x / R + \nu_y I_y / R. \tag{5.54}$$

In (2.54) we gave the rules for transforming dissipative terms alongside a canonical transformation with generating function of type $S(\mathbf{Q}, \mathbf{q})$. Applying them is a straightforward but tedious exercise in matrix algebra, best carried out with the help of a computer algebra program. When all is said and done we have the equations of motion:

$$\begin{aligned}
\psi'_x &= \frac{\nu_x}{R} - \frac{1}{\sqrt{2\beta_x I_x}} \{ (\beta_x \zeta + \alpha_x \eta) \cos \phi_x + \eta \sin \phi_x \} \frac{\Pi_x p_0}{p_s c} - \sqrt{\frac{\beta_x}{2I_x}} \cos \phi_x \frac{\Pi_x p_0}{p_s c}, \\
I'_x &= -\sqrt{\frac{2I_x}{\beta_x}} \{ (\beta_x \zeta + \alpha_x \eta) \sin \phi_x - \eta \cos \phi_x \} \frac{\Pi_x p_0}{p_s c} - \sqrt{2\beta_x I_x} \sin \phi_x \frac{\Pi_x p_0}{p_s c}, \\
\psi'_y &= \frac{\nu_y}{R} - \sqrt{\frac{\beta_y}{2I_y}} \cos \phi_y \frac{\Pi_y p_0}{p_s c}, \\
I'_y &= -\sqrt{2\beta_y I_y} \sin \phi_y \frac{\Pi_y p_0}{p_s c}.
\end{aligned} \tag{5.55}$$

Note that, apart from dissipative terms, the phase variable increases linearly with s . The usual description in terms of a betatron phase advance $\mu_{x,y}(s)$, with $\mu'_{x,y} = 1/\beta_{x,y}$, is absorbed into the definition of the variables; this trick allowed us to eliminate s from the Hamiltonian.

5.5 COUPLED MOTION WITH RADIATION

The absence of vertical dispersion means that the radiation terms have a much smaller effect in the vertical plane. Although unintentional vertical dispersions are generally small from the point of view of beam optics in a machine which has been designed to be planar, this may not be the case when one considers radiation effects. In practice there is always some x - y coupling, *e.g.* through a skew-quadrupole component. This couples quantum excitation effects from the horizontal plane into the vertical.^{63,64}

In a coupled machine, the normal modes of betatron motion are no longer horizontal and vertical oscillations and must be found before the transformation to appropriate action-angle variables via coupled β -functions *etc.* Although this process is straightforward in principle, there are practical difficulties in knowing how much coupling there is in a ring since the main sources of coupling are usually accidental. Statistical studies with programs such as PETROS⁶⁵ are one tool for analysing such effects. Another interesting approach to linear radiation effects in coupled machines uses a matrix formalism without knowledge of the coupled dispersion functions⁶⁶.

Since the coupled theory is beyond the scope of these lectures we shall say little more about the vertical plane and concentrate on the horizontal where, even in the presence of errors, we usually know a good approximation to the dispersion function.

Further complications arise in cases where the longitudinal and transverse oscillations are strongly coupled, *e.g.* when the assumption (5.14) is invalid.

6. RADIATION DAMPING

The deterministic parts of the equations (5.11) show, first of all, that *all three of the canonical momenta are directly damped by photon emission*. Secondly, the appearance of the longitudinal coupling function Π_t in the radial betatron equations shows that further dissipative effects are due to the dispersion function coupling the longitudinal damping into the radial phase space. In this chapter we discuss the radiation damping effects due to the average radiation reaction force in detail; the next chapter will include the fluctuations.

6.1 DAMPING IN LONGITUDINAL PHASE SPACE

To study the effects of radiation on the longitudinal phase space in smooth approximation, we set $x_\beta = p_\beta = y = p_y = 0$ and work from (4.17), (5.6), (5.11) and (5.42). Carrying through the changes of variables, we find that the deterministic parts of the equations of synchrotron motion are

$$\begin{aligned} \varepsilon' &= \frac{e\hat{V}}{2\pi\mathcal{R}(\delta_s)p_s c} \sin(h(z+z_s)/R) - \frac{1}{2\pi R} \int_0^{2\pi R} ds \langle \Pi_t p_0 / (p_s c) \rangle, \\ z' &= -\alpha_c \varepsilon - \underbrace{\frac{1}{2\pi R} \int_0^{2\pi R} ds \langle \Pi_x \eta / c - \Pi_t \zeta \eta / c \rangle}_{\text{zero on average}}. \end{aligned} \quad (6.1)$$

All the dissipative terms average out in smooth approximation, except the one which appears in the equation for ε' . This term is simply related to the average energy loss. Integrating the definition of Π_t contained in (4.16) and (4.17), we find that

$$p_0 \int_0^{2\pi R} \langle \Pi_t \rangle ds = \int_0^{2\pi R} \langle P_X(s) \rangle t' ds = U(\delta_s + \varepsilon) \quad (6.2)$$

is the average energy of a particle with total momentum $p_0(1 + \delta_s + \varepsilon)$ during one turn of the ring.

The normalised magnetic field strength at a displacement x in the median plane is

$$b(x, 0, s) = G(s) + K_1(s)x + \frac{1}{2}K_2(s)x^2 + \dots \quad (6.3)$$

Taking into account the energy lost in the lattice dipoles and quadrupoles, we can write out the first few terms in the expansion of the expectation value in powers of δ_s and ε :

$$\begin{aligned} U(\delta_s + \varepsilon) &\simeq c_1 p_0^2 c \int_0^{2\pi R} ds \{ 1 + (2 + G(s)\eta)(\delta_s + \varepsilon) \} c_1 b(\eta(\delta_s + \varepsilon), 0, s)^2 \\ &= c_1 p_0^2 c \left\{ I_2 + \delta_s(2I_2 + I_4) + \varepsilon(2I_2 + I_4) + \right. \\ &\quad \left. (\delta_s^2 + 2\varepsilon\delta_s)(I_2 + 2I_4 + I_8) + O(\varepsilon^2) \right\}. \end{aligned} \quad (6.4)$$

To simplify the notation, the arguments of η have been suppressed and we have used the definitions of the synchrotron radiation integrals I_2 , I_4 and I_8

Definition	Enters into:
$I_1 = \int_0^{2\pi R} G\eta ds$	momentum compaction factor
$I_2 = \int_0^{2\pi R} G^2 ds$	energy loss, energy spread, damping times, emittances, damping partition numbers
$I_3 = \int_0^{2\pi R} G^3 ds$	energy spread, polarization time, polarization level
$I_{3a} = \int_0^{2\pi R} G^3 ds$	polarization level
$I_4 = \int_0^{2\pi R} (G^2 + 2K_1)G\eta ds$	energy spread, emittances, damping partition numbers
$I_5 = \int_0^{2\pi R} G ^3 \mathcal{H} ds$	emittance
$I_{6z} = \int_0^{2\pi R} K_1^2 \beta_x ds$	energy loss in quadrupoles, nonlinear radiation damping
$I_8 = \int_0^{2\pi R} K_1^2 \eta^2 ds$	damping partition number variation

Table 1 Definitions of common synchrotron radiation integrals

which are summarised in Table 1. The other entries in the table will be discussed in due course.

Integrals 1–5 were defined in Ref. 67 which also contains useful formulae for evaluating them over a given magnet. Algorithms based on these formulae have been implemented in several computer programs (*e.g.* BEAMPARAM,⁶⁸ COMFORT,⁶⁹ PATRICIA⁷⁰ ...). These integrals are widely used; some samples of their use and further information will be found in Refs. 19, 8, 71, 21 and, especially, 67.

6.1.1 Stable phase angle

Returning to the equations (6.1) we see that they have a fixed point (*i.e.* a point where $\epsilon' = z' = 0$) which lies on the line $\epsilon = 0$ in the phase plane. The still undetermined constant z_s can then be chosen to move the origin to this natural position; this is achieved by imposing the condition

$$\begin{aligned} \frac{e\hat{V} \sin(hz_s/\mathcal{R}(\delta_s))}{2\pi \mathcal{R}(\delta_s) p_s c} &= \frac{U(\delta_s)}{2\pi R p_s c} \\ \Rightarrow \frac{e\hat{V} \sin(hz_s/\mathcal{R}(\delta_s))}{(1 + \alpha_c \delta_s)} &= c_1 p_0^2 c \{ I_2 + (2I_2 + I_4)\delta_s + (I_2 + 2I_4 + I_8)\delta_s^2 \}, \end{aligned} \quad (6.5)$$

where we used (5.41) and (6.4).

The interpretation of this condition is that, for a particle arriving at the cavities at a certain phase of the RF wave, the acceleration (or deceleration) due to the electric field just balances its average energy loss by synchrotron radiation. In the case of a stable fixed point, $\phi_s = h z_s / \mathcal{R}(\delta_s)$ is called the *stable phase angle*. Moving around the circumference of the ring, we find h solutions of (6.5) corresponding to the stable fixed points at the centres of the RF buckets.

6.1.2 Damping partition number

Synchrotron oscillations are, of course, nonlinear²³ but we can understand the nature of radiation damping much more easily by linearising (6.1) and looking at the small oscillations around the equilibrium orbit determined by δ_s ; if we introduce the *longitudinal damping partition number*^{19,60,71}

$$J_\epsilon(\delta_s) \stackrel{\text{def}}{=} \frac{d \log U(\delta_s)}{d \delta_s} = 2 + \frac{I_4}{I_2} + \delta_s \underbrace{\left(\frac{2I_8}{I_2} - \frac{I_4^2}{I_2^2} - 2 \right)}_{\text{negligible}} + O(\delta_s^2), \quad (6.6)$$

then (6.1) become

$$z' = -\alpha_c \epsilon, \quad \epsilon' = \frac{e \hat{V} \cos \phi_s}{2\pi \mathcal{R}(\delta_s)^2 h p_0 c} z - J_\epsilon(\delta_s) \frac{U(\delta_s)}{p_0 c^2} f_0 \epsilon. \quad (6.7)$$

The fluctuation terms are still being neglected. These equations are immediately recognizable as those of a damped linear oscillator with natural frequency Ω_s given by

$$\frac{\Omega_s^2}{c^2} = \frac{h \alpha_c e \hat{V}}{2\pi \mathcal{R}(\delta_s)^2 p_s c (1 + \alpha_c \delta_s)} \cos \phi_s, \quad (6.8)$$

and are equivalent to a single second-order differential equation in time,

$$\ddot{z} + 2\bar{\alpha}_\epsilon \dot{z} + \Omega_s^2 z = 0, \quad (6.9)$$

where the *damping rate* $\bar{\alpha}_\epsilon$ and the damping time $\bar{\tau}_\epsilon$ are defined by

$$\begin{aligned} \bar{\alpha}_\epsilon &\stackrel{\text{def}}{=} \frac{1}{\bar{\tau}_\epsilon} \stackrel{\text{def}}{=} \frac{J_\epsilon}{2\bar{\tau}_\epsilon} \stackrel{\text{def}}{=} \frac{J_\epsilon(\delta_s) \alpha_\epsilon}{2} \\ &\stackrel{\text{def}}{=} \frac{J_\epsilon(\delta_s) U(\delta_s)}{2p_s c} f_0 \\ &= \frac{r_e f_0}{3} \left(\frac{p_0}{mc} \right)^3 (2I_2 + I_4). \end{aligned} \quad (6.10)$$

and, in the very last expression, we used the formulae (3.50) and (6.4) and neglected the higher-order terms in δ_s . This plethora of definitions has been

introduced in order to establish an unambiguous notation which is as close as possible to the common practice of workers in the field. Ambiguity arises sometimes because, as we have seen, the damping partition number $J_\epsilon(\delta_s)$ is a variable quantity. Therefore the damping time, in which oscillation amplitudes reduce by a factor $1/e$, is a variable quantity. However in practice the damping time is often quoted for the reference case

$$I_4 = 0, \quad \delta_s = 0, \quad J_\epsilon = 2, \quad (\text{separated function lattice}), \quad (6.11)$$

which is an excellent approximation for most modern storage rings with beam circulating on their design orbits. The actual damping rate $\bar{\alpha}_\epsilon$ coincides with the quantity α_ϵ when $J_\epsilon = 2$.

6.2 DAMPING OF VERTICAL BETATRON OSCILLATIONS

Clearly, the lack of dispersion in the vertical plane results in a considerable simplification of the description of radiation damping. For this reason, we treat the vertical plane first.

The ingredients needed for writing out the vertical equations in (5.55) in full detail are simply the relations between the current canonical variables and those in which (4.17) were expressed; these are found in (5.2), (5.43) and (5.51). We also note the relation

$$\Pi_y = p_y \Pi_t / p = \tilde{p}_y \Pi_t, \quad (6.12)$$

which now gives us

$$\Pi_y = \sqrt{\frac{2I_y}{\beta_y}} (\alpha_y \cos \phi_y + \sin \phi_y) \Pi_t. \quad (6.13)$$

To restrict attention to the y -plane, we have to set $x = \eta \delta_s$, $I_x = \epsilon = z = 0$. Leaving out the fluctuations again, the equations of motion are

$$\begin{aligned} \psi_y' &= \frac{\nu_y}{R} - \cos \phi_y (\alpha_y \cos \phi_y + \sin \phi_y) c_1 p_0 (1 + \delta_s) (1 + G\eta \delta_s) \\ &\quad \times \left(G + K_1 \eta \delta_s - K_1 \sqrt{2\beta_y I_y} \cos \phi_y \right)^2, \\ I_y' &= -2I_y \sin \phi_y (\alpha_y \cos \phi_y + \sin \phi_y) c_1 p_0 (1 + \delta_s) (1 + G\eta \delta_s) \\ &\quad \times \left(G + K_1 \eta \delta_s - K_1 \sqrt{2\beta_y I_y} \cos \phi_y \right)^2 \end{aligned} \quad (6.14)$$

Scrutiny of the equation for the phase reveals that, after averaging over phases and around the ring, all values of ψ_y appear on an equal footing. There is no net

radiation damping effect towards any special value. Therefore the phase distribution will rapidly become uniform by simple mixing. However if we linearise in I_y and carry out the averaging operations in the equation for the action I_y , we find that the only term which contributes is the one containing the average of $(\sin \phi_y)^2$, namely

$$I'_y = -2I_y \frac{c_1 p_0 (1 + \delta_s)}{2(2\pi R)} \int_0^{2\pi R} (1 + G\eta\delta_s)(G + K_1\eta\delta_s)^2 ds. \quad (6.15)$$

Comparing with (6.4) or (6.5), we recognize the essentials of the expression for the energy loss on the equilibrium orbit, enabling us to rewrite this as

$$\begin{aligned} I'_y &= -\frac{U(\delta_s)}{p_s c^2} f_0 I_y \\ &= -\frac{2\alpha_y}{c} I_y. \end{aligned} \quad (6.16)$$

In this case there is no need to distinguish $\tilde{\alpha}_y$ and α_y since

$$\alpha_y \stackrel{\text{def}}{=} \frac{1}{\tau_y} \stackrel{\text{def}}{=} \frac{J_y(\delta_s) U(\delta_s)}{2p_s c} f_0 \quad (6.17)$$

where the vertical damping partition number is

$$J_y = 1, \quad (6.18)$$

independently of δ_s . We note that since there is no damping of the phase, the damping rate of the action variable carries all of the phase space compression and is twice the conventional damping rate for the cartesian coordinate y , *c.f.* the remarks following (4.23). Since the action variables give the volume of the invariant tori of the unperturbed linear system their rate of decrease due to the radiation terms is equal to the phase-space divergence.

Although tradition requires us to use the same notation α_y for the damping rate as for the optical function $\alpha_y(s)$ confusion is unlikely to arise. Changing the RF frequency does not affect the vertical damping partition number in a machine free of vertical dispersion. It should be remembered that some approximations have been made to derive this result. (Indeed there are some effects on the higher order nonlinear damping terms).

6.3 DAMPING OF RADIAL BETATRON OSCILLATIONS

From (5.6) we see that the radial deviation of a particle from the equilibrium orbit, $\eta(\delta_s, s)\delta_s$, has components due to the betatron oscillation amplitude and the energy deviation. Although the position of the particle in space, x itself, cannot change when the energy of the particle changes through photon emission, the part ascribed to the energy deviation will change and therefore the part ascribed to the betatron amplitude must undergo an equal and opposite change. Through this mechanism, the dissipative terms describing the radiation reaction forces in the longitudinal phase space are coupled into the betatron oscillations. The magnitude of the coupling is characterized locally by the values of the dispersion functions.

A part of the evaluation of the terms in the radial equations of (5.55) proceeds by analogy with the previous section. However the non-zero dispersion function means that we have a few more terms—and their averages—to deal with. The expression for the radiation coupling function, Π_x , is entirely analogous to that of Π_y in (6.13). Again, the equation for the phase ψ_x is uninteresting and we shall not write it out, but that for the action variable is

$$\begin{aligned}
 I'_x = & -\sqrt{\frac{2I_x}{\beta_x}} \{(\beta_x \zeta + \alpha_x \eta) \sin \phi_x - \eta \cos \phi_x\} c_1 p_0 (1 + \delta_s) \\
 & \times \left(1 + G\eta\delta_s + G\sqrt{2I_x\beta_x} \cos \phi_x\right) \\
 & \times \left(G + K_1\eta\delta_s + K_1\sqrt{2I_x\beta_x} \cos \phi_x\right)^2 \\
 & - 2I_x \sin \phi_x (\alpha_x \cos \phi_x + \sin \phi_x) c_1 p_0 (1 + \delta_s) \\
 & \times \left(1 + G\eta\delta_s + G\sqrt{2I_x\beta_x} \cos \phi_x\right) \\
 & \times \left(G + K_1\eta\delta_s + K_1\sqrt{2\beta_x I_x} \cos \phi_x\right)^2.
 \end{aligned} \tag{6.19}$$

In general, the best way to deal with expressions like this is to use a computer algebra program to combine the trigonometric functions and perform the averaging but, if we limit ourselves to the lowest order terms in I_x and perform a phase averaging, it is quite easy to pick out the terms which contribute. In order to make it clearer which terms are being included, we write them out, denoting the phase average with a bar,

$$\begin{aligned}
 I'_x = & -(2I_x) c_1 p_0 (1 + \delta_s) \frac{1}{2} \left\{ -\eta G (G + K_1\eta\delta_s)^2 \overline{\cos^2 \phi_x} \right. \\
 & - 2\eta G K_1 (1 + G\eta\delta_s) (G + K_1\eta\delta_s)^2 \overline{\cos^2 \phi_x} \\
 & \left. + (1 + G\eta\delta_s) (G + K_1\eta\delta_s)^2 \overline{\sin^2 \phi_x} \right\}
 \end{aligned} \tag{6.20}$$

Careful scrutiny of this expression reveals that it can be written in the suggestive form

$$I_x' = -(2I_x) \frac{c_1 p_0 (1 + \delta_s)}{2} \times \left(1 - \frac{d}{d\delta_s}\right) \{(1 + G\eta\delta_s)(G + K_1\eta\delta_s)^2\}. \quad (6.21)$$

Apart from a missing factor of $(1 + \delta_s)^2$, the expression contained in the last pair of curly braces is the integrand appearing in the formula for the total energy loss per turn of a particle on the equilibrium orbit. Following the practice of smoothing around the ring, we therefore find that this, too, can be expressed in terms of the total energy loss

$$\begin{aligned} I_x' &= -\frac{(2I_x)(1 + \delta_s)}{2p_0(2\pi R)} \left(1 - \frac{d}{d\delta_s}\right) \left\{ \frac{U(\delta_s)}{(1 + \delta_s)^2} \right\} \\ &= -\frac{(2I_x)U(\delta_s)f_0}{2cp_s} \left\{ 3 - \frac{d \log U(\delta_s)}{d\delta_s} - \frac{2\delta_s}{1 + \delta_s} \right\} \\ &= \frac{-2\bar{\alpha}_x}{c} I_x, \end{aligned} \quad (6.22)$$

where, in close analogy to (6.10), the *damping rate* $\bar{\alpha}_x$ and the *damping time* $\bar{\tau}_x$ are defined by

$$\begin{aligned} \bar{\alpha}_x &\stackrel{\text{def}}{=} \frac{1}{\bar{\tau}_x} \stackrel{\text{def}}{=} \frac{J_x(\delta_s)}{\tau_x} \stackrel{\text{def}}{=} \frac{J_x(\delta_s)\alpha_x}{1} \\ &\stackrel{\text{def}}{=} \frac{J_x(\delta_s)U(\delta_s)}{2p_s c} f_0 \\ &= \frac{\tau_e f_0}{3} \left(\frac{p_0}{mc}\right)^3 (I_2 - I_4). \end{aligned} \quad (6.23)$$

The radial damping partition number has been defined as

$$\begin{aligned} J_x(\delta_s) &= 3 - \frac{d \log U(\delta_s)}{d\delta_s} - \frac{2\delta_s}{1 + \delta_s} \\ &= 1 - \frac{I_4}{I_2} - \delta_s \left(\frac{2I_8}{I_2} - \frac{I_4^2}{I_2^2} \right) + O(\delta_s^2), \end{aligned} \quad (6.24)$$

Clearly, the three damping partition numbers are intimately related to each other. The explanation of this will be given in the following section.

We have dropped terms in I_x^2 which arise principally from the quadrupole fields and give rise to *nonlinear radiation damping*. Such effects can be important, particularly in large storage rings and are discussed, *e.g.*, in Reference 8.

6.3.1 ♠ *Where does radiation damping occur?*

The answer to the question posed in the title of this subsection is obvious from the equations of motion of electrons, expressed in any of the sets of canonical variables that we have used up to now:

Radiation damping of all three modes of oscillation about the equilibrium orbit occurs when the particle passes through the static magnetic fields of the storage ring and emits synchrotron radiation.

Provided we understand “damping” as a contraction of phase space volumes (which can only occur when the equations of motion contain a non-Hamiltonian part) this is a trivial consequence of the Hamiltonian nature of the equations of motion in the absence of radiation.

However in many discussions of the subject it is stated that the “damping” of radial betatron oscillations occurs in the RF cavities. In fact there is no contradiction with the present discussion for the simple reason that these treatments use a *non-symplectic phase space* which has been, and remains, very popular in accelerator physics. In this phase space the slope variable $\theta_x = x'$ is used in place of the canonical momentum p_x and although these quantities are very close to each other numerically, it is easy to see that when the energy of the particle changes they behave in different ways.

Specifically, the value of x' *does not* change during photon emission because the photon recoil is collinear with the total momentum vector although the transverse momentum *does* change. On the other hand, only the longitudinal momentum changes as the particle traverses an RF cavity while the slope variable x' is reduced. This explains the shift of the apparent damping effect from the magnets to the cavities!

It is perfectly possible to develop the whole theory of electron dynamics in the non-symplectic phase space. Indeed this is what was done in the original treatments of the problem^{53,60,12,19} and these variables do have the advantage of a more immediate connection with the quantities which can be measured experimentally—it is a little easier to deduce the angle of a beam, in mrad, from beam-position monitor data than it is to work out its transverse momentum. The choice between this and the Hamiltonian approach is ultimately a matter of personal taste.

6.4 DAMPING PARTITION NUMBERS AND DAMPING APERTURE

While the damping is linear, it is clear that, to a very good approximation, the three damping partition numbers (6.6), (6.18) and (6.24) satisfy the sum rule

$$J_e(\delta_s) + J_x(\delta_s) + J_y(\delta_s) = 4. \quad (6.25)$$

for all values of δ_s such that an off-momentum closed orbit given by $\eta(\delta_s, s)$ exists. This sum rule holds even if one of the partition numbers is negative and is a concrete manifestation of *Robinson's Theorem*, discussed in Section 4. For we can rewrite (6.25) as the sum of the damping rates of the normal modes of oscillation about the equilibrium orbit

$$\bar{\alpha}_x + \bar{\alpha}_y + \bar{\alpha}_e = \alpha_{\text{tot}} = \frac{2U(\delta_s)}{p_s c}, \quad (6.26)$$

which is essentially the same as (4.26).

By varying the RF frequency, and thereby δ_s through (5.24) or (5.40), it is therefore possible to redistribute the damping between the longitudinal and radial modes. This has become an important operational tool in modern storage rings since it allows the operator to put more damping into the mode which needs it more, *e.g.* to help counter an instability or increase injection efficiency. Even more importantly, it can be used to decrease the horizontal beam size at the expense of an increased energy spread (or vice versa) in order to stretch the limits of performance of the machine. Several different ways in which this will be done in operation of LEP are described in Ref. 21.

From (6.6) we can immediately deduce the formula for the rate of change of J_e or J_x with respect to the momentum deviation on the equilibrium orbit

$$\begin{aligned} \frac{dJ_e(\delta_s)}{d\delta_s} &\simeq -\frac{dJ_x(\delta_s)}{d\delta_s} = \frac{d^2 \log U(\delta_s)}{d\delta_s^2} \\ &= \left(\frac{2I_8}{I_2} - \frac{I_4^2}{I_2^2} \right) + O(\delta_s) \\ &\simeq \frac{2 \int_0^{2\pi R} K_1^2 \eta^2 ds}{\int_0^{2\pi R} G^2 ds}. \end{aligned} \quad (6.27)$$

A variety of higher-order approximations to this quantity have been published in several papers.^{73,74,75} In comparing them, the different regimes of application and definitions of the synchrotron integrals should be carefully noted. Although we have stopped at the first term (which is enough for most fairly large rings), the principles given here are sufficient to derive further approximations.

To have stable beams in a storage ring, one must ensure that each damping partition number remains positive. The range of values of δ_s in which this is true is called the *damping aperture*, and is determined by the values of the synchrotron integrals I_2 , I_4 and I_8 :

$$-\frac{2I_2 + I_4}{2I_8} < \delta_s < \frac{I_2 - I_4}{2I_8}. \quad (6.28)$$

Together with (5.40), this translates directly into an allowable range of variation of RF frequencies or an allowable displacement of the equilibrium orbit. For further details, including the use of *Robinson wigglers* to shift the damping aperture, the reader may consult Ref. 20.

In small storage rings, the integral I_8 is usually so small that the physical aperture is smaller than the damping aperture; this situation is generally reversed in large rings.

The damping aperture is easily measured by varying the RF frequency and watching for horizontal beam blow-up on a synchrotron light monitor. Since the synchrotron light is usually generated in a bending magnet in which the dispersion is non-zero, the beam size has components due to both the betatron and synchrotron amplitudes in the beam so that both ends of the damping aperture will manifest themselves by inflation of the horizontal beam profile.

6.4.1 Separated vs. combined function lattices

In a separated function lattice, the bending and focusing functions for particles on the reference orbit are performed by separate dipole and quadrupole magnets, so that

$$G(s)K_1(s) = 0, \quad \text{for all } s \quad (6.29)$$

Since the dispersion function is generally much smaller than the bending radius, $\eta G \ll 1$, it turns out that the integral $I_4 \ll I_2$. However most of the earlier proton synchrotrons and a few electron machines were built with the two functions combined into the so-called *gradient magnets*. Most of what we have said applies also to this type of structure although the radial damping partition number (to be introduced later) will probably be negative since $I_4 > I_2$.

This radial anti-damping can be tolerated, as in the DESY electron synchrotron, if the beam does not stay in the machine long enough for it to do much damage. Or it can be corrected, as in the early CEA machine or, more recently, the CERN PS in its rôle as LEP injector, by the use of the so-called *Robinson wigglers*.^{60,20} These wigglers contain a combined function dipole-quadrupole field designed to make a large negative contribution to I_4 . For further details on combined function electron machines, see Ref. 72.

7. QUANTUM FLUCTUATIONS AND FOKKER-PLANCK EQUATIONS

We now go on to study the truly statistical side of electron dynamics with radiation. Inclusion of the quantum fluctuations of the synchrotron radiation power will give us stochastic differential equations for the normal modes of the linearised motion around the equilibrium orbit. The fluctuations make the orbit unstable but the radiation damping effects can prevent the deviations of the orbits of most of the particles from becoming too large. The balance between the radiation damping and quantum excitation of orbit amplitudes can be described in terms of Fokker-Planck equations. We should not forget, however, that these two effects have the same physical origin and that one cannot exist without the other.

Fokker-Planck equations have been applied to several problems in accelerator physics, most of them more complicated than those described here; for some examples, see Refs. 7 (several articles), 76, 12, 39, 77 and the reference lists which they contain.

7.1 QUANTUM FLUCTUATIONS IN LONGITUDINAL PHASE SPACE

Let us reintroduce the fluctuating part of the radiation power from (6.1) into the longitudinal equations of motion, (6.7). To prepare the ground for writing down the Fokker-Planck equation according to the recipe embodied in (2.1) and (2.27), we write them as*

$$z' = K_z(z, \varepsilon) + Q_z(z, \varepsilon)\xi(s), \quad \varepsilon' = K_\varepsilon(z, \varepsilon) + Q_\varepsilon(z, \varepsilon)\xi(s) \quad (7.1)$$

where the K - and Q -functions are

$$\begin{aligned} K_z(z, \varepsilon) &= -\alpha_c \varepsilon, & K_\varepsilon(z, \varepsilon) &= (\Omega_s/c)^2 z - \frac{J_\varepsilon \alpha_\varepsilon}{c} \varepsilon \\ Q_z(z, \varepsilon) &= 0, & Q_\varepsilon(z, \varepsilon) &= -p_0 \sqrt{c_2 G(s)^3}. \end{aligned} \quad (7.2)$$

Strictly speaking, these are in something of a hybrid form since the smooth approximation has not yet been applied to the fluctuation terms. These, by their nature, must be approximated in a *root-mean-square* fashion, rather than directly. This is easier to understand in terms of the Fokker-Planck equation which contains each K -function to the first power but products of two Q -functions.

* We are neglecting some small terms of higher order in δ_s .

Since

$$\frac{\partial Q_\varepsilon}{\partial \varepsilon} = \frac{\partial Q_\varepsilon}{\partial z} = 0 \quad (7.3)$$

the recipe for writing it down, (2.27), simplifies by virtue of the lack of “spurious drift” terms and we find

$$\begin{aligned} \frac{\partial F(z, \varepsilon, s)}{\partial s} = & \alpha_c \varepsilon \frac{\partial F(z, \varepsilon, s)}{\partial z} - \left(\frac{\Omega_s}{c} \right)^2 z \frac{\partial F(z, \varepsilon, s)}{\partial \varepsilon} + \frac{J_\varepsilon \alpha_\varepsilon}{c} \frac{\partial}{\partial \varepsilon} [\varepsilon F(z, \varepsilon, s)] \\ & + \frac{c_2 p_0^2 I_3}{2(2\pi R)} \frac{\partial^2 F(z, \varepsilon, s)}{\partial \varepsilon^2}, \end{aligned} \quad (7.4)$$

where $F(z, \varepsilon, s)$ is the distribution function in longitudinal phase space and we have made the smooth approximation of the diffusion term (i.e. the one with second derivatives) in terms of the synchrotron integral I_3 defined in Table 1.

This equation can be solved completely in terms of its Green function^{39,33} or by eigenfunction expansions²⁹ but we can simplify it further by making a phase-mixing assumption

$$\langle z \rangle = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} d\varepsilon z F(z, \varepsilon, s) = 0 \quad (7.5)$$

which will be true in many situations, including that of equilibrium. Then we can integrate (7.4) over z to get an equation for the reduced distribution function

$$F(\varepsilon, s) = \int_{-\infty}^{\infty} F(z, \varepsilon, s) dz, \quad (7.6)$$

namely

$$\frac{\partial F(\varepsilon, s)}{\partial s} = \frac{J_\varepsilon \alpha_\varepsilon}{c} \frac{\partial}{\partial \varepsilon} [\varepsilon F(\varepsilon, s)] + \frac{c_2 p_0^2 I_3}{2(2\pi R)} \frac{\partial^2 F(\varepsilon, s)}{\partial \varepsilon^2}. \quad (7.7)$$

This equation provides a rough description of the evolution of the energy distribution in the beam. It is incomplete to the extent that (7.5) fails to hold.

7.2 STATIONARY ENERGY DISTRIBUTION

To find the stationary (or, loosely speaking, equilibrium) solution $F_0(\varepsilon)$, we simply set $\partial_s F = 0$ and integrate once to find

$$-\varepsilon F_0(\varepsilon) = \frac{c_2 p_0^2 I_3 f_0}{2J_\varepsilon \alpha_\varepsilon} F_0'(\varepsilon). \quad (7.8)$$

When we integrate this again, choosing the constant of integration to normalise the distribution to unity, we find the familiar gaussian distribution of momentum (or energy) deviations

$$F_0(\varepsilon) = \frac{1}{\sqrt{2\pi\bar{\sigma}_\varepsilon^2}} \exp\left(-\frac{\varepsilon^2}{2\bar{\sigma}_\varepsilon^2}\right), \quad (7.9)$$

where $\bar{\sigma}_\varepsilon$ is the r.m.s. energy spread in the beam for a linear damping rate determined by the value of J_ε :

$$\langle \varepsilon^2 \rangle \stackrel{\text{def}}{=} \bar{\sigma}_\varepsilon^2 = \frac{2\sigma_\varepsilon^2}{J_\varepsilon} = \frac{55}{32\sqrt{3}} \frac{\hbar}{mc} \left(\frac{p_0}{mc}\right)^2 \frac{I_3}{J_\varepsilon I_2}. \quad (7.10)$$

The quantity σ_ε , which can be regarded as a measure of the strength of quantum excitation, is defined to be the energy spread for the reference case $J_\varepsilon = 2$.*

Since $I_2/I_3 \approx \rho$, the bending radius, in an isomagnetic ring there is very little which can be done, beyond varying J_ε , to reduce the energy spread in the beams of an electron storage ring. Moreover, the energy spread is directly proportional to E_0 . A very small decrease of σ_ε can in principle be achieved with wiggler magnets but their usual effect is to increase it.

This is an important limitation since it determines the energy resolution of particle physics experiments which may be trying, for example, to detect, or measure the widths of, narrow resonances in the mass spectrum.²¹ In fact, since design considerations for colliding beam rings usually imply $\rho \propto E_0^2$, it almost always turns out that $\sigma_\varepsilon \approx 0.1\%$ at the top energy of a given ring. Nevertheless e^+e^- rings still provide a much finer energy resolution than foreseeable linear colliders or hadron colliders (although the beam energy spread is small the quarks have a very broad energy distribution) and there remains the possibility of enhancing it still further with the so-called "monochromator" insertions.⁷⁸

The gaussian distribution of energy deviation is by no means inevitable—nonlinear terms^{10,8} (usually dissipative) may well change it, especially in the tails. Arguments based on the "Central Limit Theorem" should only be applied in linear approximation and the analogy with the Maxwell-Boltzmann velocity distribution in a gas is not a complete one.

* Often no notational distinction is made between σ_ε and $\bar{\sigma}_\varepsilon$, so one should always be careful to understand which is meant.

7.3 EQUILIBRIUM IN THE NONLINEAR RF BUCKET

It is possible to find the equilibrium solution of the two-dimensional Fokker-Planck equation (7.4) even when the nonlinearity of the RF waveform is included. Although we assumed a sinusoidal waveform in the previous chapters, the solution can be found for an arbitrary RF waveform, provided the smooth approximation is made.

Following Ref. 19 quite closely, let us suppose that, in smooth approximation, the generalised equations of motion are given by making the replacement

$$\frac{e\hat{V}}{2\pi\mathcal{R}(\delta_s)p_s c} \sin \left[\frac{h(z+z_s)}{\mathcal{R}(\delta_s)} \right] \rightarrow \frac{U(\delta_s)}{2\pi R p_s c} - \frac{d\Phi(z)}{dz} \quad (7.11)$$

so that the definition of the effective potential Φ already absorbs the energy loss by synchrotron radiation and (with appropriate choice of the arbitrary constant)

$$\Phi(0) = 0, \quad \frac{d^2\Phi(0)}{dz^2} = \Omega_s^2/c^2. \quad (7.12)$$

Then the Hamiltonian of synchrotron motion is

$$H_s = -\frac{\alpha_c \varepsilon^2}{2} - \Phi(z) + \frac{U(\delta_s)z}{2\pi\mathcal{R}(\delta_s)p_s c}. \quad (7.13)$$

This Hamiltonian does not really contain the radiation loss—it appears only as a result of the definition of Φ . It cancels out again in the equations of motion:

$$\begin{aligned} z' &= -\alpha_c \varepsilon, \\ \varepsilon' &= -\frac{d\Phi(z)}{dz} - 2\frac{\tilde{\alpha}_\varepsilon}{c}\varepsilon - 2\sqrt{\frac{\tilde{\alpha}_\varepsilon}{c}}\sigma_\varepsilon \xi(s). \end{aligned} \quad (7.14)$$

We are assuming that the linear approximation may still be applied to the damping term even though we have included the nonlinear terms in the RF waveform.¹⁹ This is usually a reasonable approximation since ε generally remains small even when $hz/\mathcal{R}(\delta_s) \simeq 1$. For very large rings like LEP, where the ratio U_0/E_0 can be as large as a few percent, some new nonlinear effects may become noticeable.⁷⁹

This form of the Hamiltonian is more general than the longitudinal part of (5.42) and could include, for example, the case where a second RF system is operated at a harmonic number which may be an integer multiple of h .⁸⁰

The Fokker-Planck equation in the nonlinear RF bucket is then

$$\begin{aligned} \frac{\partial F(z, \varepsilon, s)}{\partial s} = & \underbrace{\alpha_c \varepsilon \frac{\partial F(z, \varepsilon, s)}{\partial z} - \frac{d\Phi(z)}{dz} z \frac{\partial F(z, \varepsilon, s)}{\partial \varepsilon}}_{\text{Liouvillian terms}} \\ & + \underbrace{\frac{\bar{\alpha}_\varepsilon}{c} \frac{\partial}{\partial \varepsilon} [\varepsilon F(z, \varepsilon, s)]}_{\text{damping}} + \underbrace{\frac{2\bar{\alpha}_\varepsilon \bar{\sigma}_\varepsilon^2}{c} \frac{\partial^2 F(z, \varepsilon, s)}{\partial \varepsilon^2}}_{\text{diffusion}}, \end{aligned} \quad (7.15)$$

and, since the terms arising from the Hamiltonian vanish by virtue of Hamilton's equations, it is easy to check that the equilibrium solution is

$$F_0(z, \varepsilon) = \frac{1}{Z} \exp \left(-\frac{\varepsilon^2}{\bar{\sigma}_\varepsilon^2} - \frac{\Phi(z)}{\alpha_c \bar{\sigma}_\varepsilon^2} \right). \quad (7.16)$$

Notice that, despite a superficial resemblance, this is not a Maxwell-Boltzmann or Gibbsian equilibrium distribution since even the sign of the part of the Hamiltonian which appears in the exponential is opposite to what it should be (if the temperature is to be positive). Later we shall see that, since the other sign prevails in the equilibrium distributions in betatron phase space, the distribution in the complete 6-dimensional phase space is not in a form proportional to $\exp(-H/kT)$ and is therefore *not a genuine thermodynamic equilibrium*.

Nevertheless, it is clear that the reduced distributions in each of the three 2-dimensional phase spaces have enough features in common with Maxwell-Boltzmann distributions that we can speak loosely of them as equilibrium distributions.

The normalization constant, which we refer to (loosely again) as the *partition function* is given by the integral

$$\begin{aligned} Z &= \sqrt{\frac{J_\varepsilon(\delta_s)}{2\pi\sigma_\varepsilon^2}} \int_{-\infty}^{\infty} \exp \left(-\frac{J_\varepsilon(\delta_s)\Phi(z)}{\alpha_c 2\sigma_\varepsilon^2} \right) \\ &= F_0(0, 0)^{-1} \end{aligned} \quad (7.17)$$

and is a function of the parameters in the Hamiltonian and of J_ε . If this integral can be evaluated analytically, and if the explicit form of Φ is known, then differentiation with respect to these parameters is a useful technique in the calculation of equilibrium expectation values. For example, from (7.16) and (7.17), it is easy to show that the expectation value of the function Φ itself is

$$\langle \Phi(z) \rangle = -\alpha_c \sigma_\varepsilon^2 \left(\frac{\partial Z}{\partial J_\varepsilon} - \frac{Z}{2J_\varepsilon} \right). \quad (7.18)$$

Further analogies with equilibrium statistical mechanics are readily found.

7.4 FOKKER-PLANCK EQUATION IN ACTION-ANGLE VARIABLES

As usual, it is useful to be able to work with the action-angle variables of the problem. This section is devoted to the technicalities of deriving the Fokker-Planck equation in these variables.

We transform to action-angle variables of linearised synchrotron motion and make a rescaling to variables (χ, I_z) with

$$z = -\frac{1}{\kappa_z} \sqrt{2I_z} \cos(\kappa_z \chi), \quad \varepsilon = \sqrt{2I_z} \sin(\kappa_z \chi), \quad \kappa_z \stackrel{\text{def}}{=} \frac{\Omega_s}{\alpha_c c}. \quad (7.19)$$

The constant κ_z can be thought of as a conversion factor between energy deviation and length units:

$$\sqrt{\langle \varepsilon^2 \rangle} = \kappa_z \sqrt{\langle z^2 \rangle} = \frac{Q_s}{\alpha_c R} \sqrt{\langle z^2 \rangle} \quad (7.20)$$

and $Q_s = \Omega_s/2\pi f_0$ is the synchrotron tune. With these variables, the longitudinal Hamiltonian reduces to

$$H_s = -\alpha_c I_z \quad (7.21)$$

and, by applying (2.54) (or otherwise), we can derive stochastic equations of motion equivalent to (7.4)

$$\begin{aligned} \chi' &= K_\chi(\chi, I_z) + Q_\chi(\chi, I_z) \xi(s), \\ I_z' &= K_I(\chi, I_z) + Q_I(\chi, I_z) \xi(s), \end{aligned} \quad (7.22)$$

where

$$\begin{aligned} K_\chi(\chi, I_z) &= -\alpha_c - \frac{J_\varepsilon(\delta_s) \alpha_\varepsilon}{2c \kappa_z} \sin(2\kappa_z \chi), \\ Q_\chi(\chi, I_z) &= -\frac{2\sigma_\varepsilon}{\kappa_z} \sqrt{\frac{\alpha_\varepsilon}{2c I_z}} \cos(\kappa_z \chi), \\ K_I(\chi, I_z) &= -\frac{J_\varepsilon(\delta_s) \alpha_\varepsilon}{c} [1 - \cos(2\kappa_z \chi)], \\ Q_I(\chi, I_z) &= -2\sigma_\varepsilon \sqrt{\frac{\alpha_\varepsilon}{2c I_z}} \sin(\kappa_z \chi). \end{aligned} \quad (7.23)$$

Now Q_I and Q_χ depend on I_z and χ and their derivatives will contribute spurious drift terms—and a certain amount of algebraic complication—to the Fokker-Planck equation:

$$\begin{aligned} \frac{\partial F(I_z, \chi, s)}{\partial s} = & -\frac{\partial}{\partial I_z} [D_I F(I_z, \chi, s)] - \frac{\partial}{\partial \chi} [D_\chi F(I_z, \chi, s)] \\ & + \frac{1}{2} \frac{\partial^2}{\partial I_z^2} [Q_I^2 F(I_z, \chi, s)] + \frac{\partial^2}{\partial I_z \partial \chi} [Q_I Q_\chi F(I_z, \chi, s)] \\ & + \frac{1}{2} \frac{\partial^2}{\partial \chi^2} [Q_\chi^2 F(I_z, \chi, s)]. \end{aligned} \quad (7.24)$$

Here, the complete drift terms are

$$\begin{aligned} D_I = K_I + \frac{1}{2} \frac{\partial Q_I}{\partial I_z} Q_I + \underbrace{\frac{1}{2} \frac{\partial Q_I}{\partial \chi} Q_\chi}_{\text{important!}} &= -\frac{\alpha_\epsilon}{c} (J_\epsilon I_z - 2\sigma_\epsilon^2) + \frac{J_\epsilon \alpha_\epsilon I_z}{c} \cos(2\kappa_z \chi), \\ D_\chi = K_\chi + \frac{1}{2} \frac{\partial Q_\chi}{\partial I_z} Q_I + \frac{1}{2} \frac{\partial Q_\chi}{\partial \chi} Q_\chi &= -\alpha_\epsilon - \frac{\alpha_\epsilon}{\kappa_z c} \left(J_\epsilon + \frac{\sigma_\epsilon^2}{2I_z} \right) \sin(2\kappa_z \chi), \end{aligned} \quad (7.25)$$

and the diffusion terms are

$$\begin{aligned} Q_I^2 &= \frac{4\alpha_\epsilon \sigma_\epsilon^2 I_z}{c} [1 - \cos(2\kappa_z \chi)], \\ Q_I Q_\chi &= \frac{2\alpha_\epsilon \sigma_\epsilon^2}{\kappa_z c} \sin(2\kappa_z \chi), \\ Q_\chi^2 &= \frac{\alpha_\epsilon \sigma_\epsilon^2}{\kappa_z^2 c I_z} [1 + \cos(2\kappa_z \chi)]. \end{aligned} \quad (7.26)$$

In action-angle variables, it is easy to make a step beyond the smooth approximation and apply the averaging method.^{27,81,34} The result of the phase-averaging of (7.24) is an averaged Fokker-Planck equation for the action variable which will be valid on time-scales longer than that of a synchrotron oscillation, notably the damping time scale:

$$\begin{aligned} \tau_\epsilon \frac{\partial F}{\partial t} = & \underbrace{-\frac{\partial}{\partial I_z} [(-J_\epsilon I_z + 2\sigma_\epsilon^2) F]}_{\text{damping of amplitude}} + \underbrace{c\tau_\epsilon \alpha_\epsilon \frac{\partial F}{\partial \chi}}_{\text{phase advance}} \\ & + \underbrace{2\sigma_\epsilon^2 \frac{\partial^2}{\partial I_z^2} (I_z F)}_{\text{amplitude diffusion}} + \underbrace{\frac{\sigma_\epsilon^2}{2\kappa_z^2 \alpha_\epsilon I_z} \frac{\partial^2 F}{\partial \chi^2}}_{\text{phase diffusion}}. \end{aligned} \quad (7.27)$$

The absence of a damping term for the phase χ guarantees that the phase diffusion term superposed on the rapid oscillatory phase advance will lead to

a uniform distribution in χ on $[0, 2\pi]$. Moreover, we have not included the dependence of the synchrotron frequency on amplitude which produces a filamentation effect, further accelerating the phase-mixing.

7.4.1 Stationary distribution and projections

We can solve (7.27) for the stationary distribution

$$F_0(I_z) = \frac{1}{\bar{\sigma}_\epsilon^2} \exp\left(-\frac{I_z}{\bar{\sigma}_\epsilon^2}\right) \quad (7.28)$$

from which we can evaluate the longitudinal emittance (in these units!)

$$\langle I_z \rangle = \int_0^\infty I_z F_0(I_z) dI_z = \bar{\sigma}_\epsilon^2 = \left\langle \frac{\epsilon^2 + \kappa_z^2 z^2}{2} \right\rangle. \quad (7.29)$$

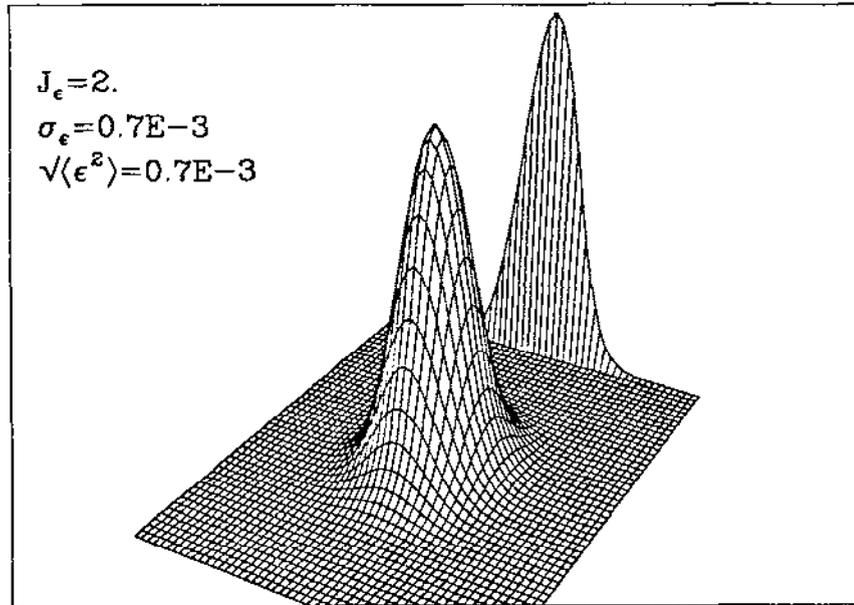


Fig. 7 Gaussian longitudinal phase space distribution and its projection

The distribution (7.28) is equivalent to a joint gaussian in z and ϵ :

$$F_0(z, \epsilon) = F_0(\kappa_z^2 z^2 / 2 + \epsilon^2 / 2) = \frac{1}{2\pi \bar{\sigma}_\epsilon \bar{\sigma}_z} \exp\left(-\frac{\epsilon^2}{2\bar{\sigma}_\epsilon^2} - \frac{z^2}{2\bar{\sigma}_z^2}\right), \quad (7.30)$$

where the natural (or zero-current) bunch length is

$$\bar{\sigma}_z = \bar{\sigma}_\varepsilon / \kappa_z. \quad (7.31)$$

This distribution is shown in Fig. 7 for some typical values of J_ε and σ_ε . The figure also shows the projection of the action distribution along one axis, given by the integral

$$\lambda(\kappa_z z) = \int_{-\infty}^{\infty} F_0(z, \varepsilon) d\varepsilon = \int_0^{\infty} \frac{F_0(I_z)}{\sqrt{2I_z - \kappa_z^2 z^2}} dI_z. \quad (7.32)$$

By virtue of the special properties of the gaussian, this is just the same as (7.9). In general it need not be.¹⁰ Such a projection corresponds to the energy distribution or the longitudinal current density profile of the bunch. It is *not* the shadow of the phase space distribution. The second equality in (7.32) holds only for rotationally symmetric distributions.

7.5 LONGITUDINAL QUANTUM LIFETIME

As we have seen, the quantum fluctuations make the equilibrium closed orbit unstable but* the radiation damping establishes a stationary distribution in the neighbourhood of the closed orbit whose tails decay rapidly to zero at large amplitude. However no account has yet been taken of how far out in amplitude we can go without meeting some boundary of the stable phase space. Such boundaries always exist and may be due to the vacuum chamber, the finite height of the RF bucket or nonlinear terms in the Hamiltonian which make the orbits unstable (dynamic aperture).

Particles exceeding some limiting amplitude are therefore almost always lost rather rapidly; by continuity, therefore, the distribution function must vanish at the boundary. However it is clear from the Fokker-Planck equation that quantum diffusion will continually replenish the region just within this amplitude limit, resulting in a constant flux of particles out of the core of the distribution. The inverse of the loss rate, or *quantum lifetime*, is the time in which, by this effect, the total number of particles in the beam decays by a factor $1/e$.

Separate loss rates are usually calculated in the radial, vertical and longitudinal phase planes. The total loss rate is the sum of all the individual loss rates. In terms of lifetimes, the net beam lifetime is therefore given by

$$\frac{1}{\tau_b} = \frac{1}{\tau_{q\varepsilon}} + \frac{1}{\tau_{qx}} + \frac{1}{\tau_{qy}} + \dots \quad (7.33)$$

where the dots indicate that we should add the loss rates corresponding to

* Provided the damping partition number is positive.

other loss mechanisms. Clearly, the net lifetime is always less than the shortest individual quantum lifetime.

The simplest method for calculating the longitudinal quantum lifetime applies only when the loss rate is very small so that we can regard the equilibrium distribution as *quasi-stationary* in the sense that, over most of the stable phase space, the equilibrium form maintains but the overall normalisation constant of the distribution slowly changes because of the slow loss of particles. The arguments which follow are a simplified version of arguments which seem originally to have been given by Chandrasekhar.³³ A more detailed presentation of quantum and other lifetimes may be found in Ref. 77.

To evaluate the lifetime, we analyse the balance which holds in an equilibrium state with the help of (7.27). Like all others, this Fokker-Planck equation can be written in the form of a *continuity equation*

$$\frac{\partial F(I_z, t)}{\partial t} = -\frac{\partial}{\partial I_z} J(I_z) \quad (7.34)$$

where the particle current in action space can be split into damping and diffusive components:

$$\begin{aligned} J(I_z) &= \underbrace{-2\bar{\alpha}_\epsilon I_z F(I_z, t)}_{\text{damping}} - \underbrace{2\bar{\alpha}_\epsilon \bar{\sigma}_\epsilon^2 I_z \frac{\partial F(I_z, t)}{\partial I_z}}_{\text{diffusion}} \\ &= J_{\text{damp}} + J_{\text{diff}}. \end{aligned} \quad (7.35)$$

Observe that the damping component always constitutes a drift of particles towards the origin $I_z = 0$ while the diffusive component represents a tendency of particles to move away from regions where the density is higher. In an equilibrium distribution extending to $I_z = \infty$, the two components cancel each other exactly at every value of I_z .

Now suppose that the acceptance limit occurs for a value $I_z = I_{\text{max}}$. Although, by the quasi-stationary assumption, the cancellation of the two components continues to maintain for $I_z < I_{\text{max}}$, the balance is upset just below the boundary I_{max} . Here the diffusive current continues to push particles outwards while the compensating damping current of particles returning from $I_z > I_{\text{max}}$ is missing.*

The loss rate is therefore given by evaluating the diffusive current at $I_z =$

* If the potential barrier at I_{max} does not fall away too sharply beyond I_{max} then there is some probability that particles which have passed to the other side may return to mingle with the survivors. This effect is usually neglected—but see Chandrasekhar.^{33,79}

I_{\max} in an equilibrium distribution,

$$\frac{1}{\tau_{q\varepsilon}} = \frac{1}{J_{\text{diff}}(I_{\max})} = -\frac{1}{J_{\text{damp}}(I_{\max})} \quad (7.36)$$

where the second equality holds as a consequence of the Fokker-Planck equation as it applies in the absence of the boundary. The result is therefore

$$\begin{aligned} \tau_{q\varepsilon} &= \frac{\bar{\sigma}_\varepsilon^2}{2\bar{\alpha}_\varepsilon J_{\max}} \exp\left(-\frac{I_{\max}}{\bar{\sigma}_\varepsilon^2}\right) \\ &= \frac{\tau_\varepsilon \sigma_\varepsilon^2}{\varepsilon_{\max}^2} \exp\left(-\frac{J_\varepsilon \varepsilon_{\max}}{2\sigma_\varepsilon^2}\right). \end{aligned} \quad (7.37)$$

The second expression applies when the boundary is constituted by the RF voltage limitation, ε_{\max} being the bucket half-height.

It should always be remembered that this result only applies in the case where the lifetime is very long compared to the damping time. This is not a serious restriction in practice since this is a minimal requirement for good beam storage. Moreover this formula will never mislead us—if it predicts a long quantum lifetime then the answer is a very good approximation and if it predicts a short lifetime the answer may be wrong but we can be sure that the lifetime is unacceptably short.

7.6 MOMENT EQUATION FOR ACTION

As an illustration of the technique of moment equations, we use the averaged Fokker-Planck equation (7.27) to calculate the time-evolution of the longitudinal emittance

$$\langle I_z \rangle = \int_0^\infty dI_z \int_0^{2\pi} d\chi I_z F(I_z, \chi, t). \quad (7.38)$$

Taking the first moment of (7.27) with respect to I_z and imposing periodicity in the phase, $F(I_z, 0, t) = F(I_z, 2\pi, t)$, we find

$$\begin{aligned} \frac{d\langle I_z \rangle}{dt} &= -2\bar{\alpha}_\varepsilon \int_0^\infty dI_z \int_0^{2\pi} d\chi I_z \frac{\partial}{\partial I_z} [(-I_z + \bar{\sigma}_\varepsilon^2) F(I_z, \chi, t)] \\ &\quad + 2\bar{\alpha}_\varepsilon \bar{\sigma}_\varepsilon^2 \int_0^\infty dI_z \int_0^{2\pi} d\chi I_z \frac{\partial^2}{\partial I_z^2} [I_z F(I_z, \chi, t)]. \end{aligned} \quad (7.39)$$

Integration by parts shows that the diffusion term does not contribute except through the spurious drift terms, and we find the moment equation for the longitudinal emittance (valid on time scales longer than a damping time)

$$\frac{d\langle I_z \rangle}{dt} = -2\bar{\alpha}_\epsilon (\langle I_z \rangle - \bar{\sigma}_\epsilon^2) \quad (7.40)$$

The solution is very simple; supposing that $\langle I_z \rangle = I_0$ at $t = 0$, we have

$$\langle I_z \rangle = I_0 e^{-2\bar{\alpha}_\epsilon t} + \bar{\sigma}_\epsilon^2 (1 - e^{-2\bar{\alpha}_\epsilon t}). \quad (7.41)$$

This shows how the energy spread and bunch length will damp down to their equilibrium values determined by the balance between radiation damping and quantum fluctuations. Since we deal with the action which is the sum of the squares of these quantities, the damping rate is twice the damping rate of the phase space coordinates themselves.

We should observe that the moment equation (7.40) cannot be obtained by naïvely averaging the stochastic differential equation for I_z given by (7.22) and (7.23) because, without the spurious drift terms generated from Q_I , we would not obtain the term containing $\bar{\sigma}_\epsilon^2$ in (7.40). This is a case where we have to be aware that our stochastic differential equations are interpreted according to the recipe of Stratonovich and the prescription (2.42) has to be used to calculate the evolution of expectation values.

7.7 A MODEL OF INJECTION DAMPING

The solution obtained in the previous section assumed that the oscillation phases were already completely mixed. As further illustration of the moment equation technique, let us show how we can extract further information from the stochastic differential equations for z and ϵ . These include the oscillatory part of the time evolution. Specifically, we shall consider a simple model of the process of injection of particles into longitudinal phase space.

Suppose that particles are injected into the RF bucket with the correct equilibrium energy for the storage ring but with some offset $hz_0/\mathcal{R}(\delta_s)$ with respect to the stable phase angle on the RF wave. Translated into initial conditions for the equations (7.1), we have

$$\epsilon = 0, \quad z = z_0 \quad \text{at } t = 0 \quad (7.42)$$

Since there are no spurious drift terms to worry about, it is convenient to work with a single second-order stochastic differential equation equivalent to these equations:

$$\ddot{z} + 2\bar{\alpha}_\varepsilon \dot{z} + \Omega_s^2 z = 2\alpha_c \sqrt{\alpha_\varepsilon} \sigma_\varepsilon \xi(t), \quad (7.43)$$

where

$$\langle \xi(t) \rangle = 0, \quad \langle \xi(t) \xi(t') \rangle = \delta(t - t'). \quad (7.44)$$

Solving (7.43) formally as a driven oscillator equation, we have

$$\begin{aligned} z(t) &= z_0 e^{-\bar{\alpha}_\varepsilon t} \cos(\Omega_s t) \\ &\quad + \frac{2\alpha_c \sqrt{\alpha_\varepsilon} \sigma_\varepsilon}{\Omega_s} \int_0^t \xi(\tau) e^{-\bar{\alpha}_\varepsilon(t-\tau)} \sin(\Omega_s(t-\tau)) \\ &= \langle z(t) \rangle + \hat{z}(t), \end{aligned} \quad (7.45)$$

where the splitting into the expectation value of $z(t)$ and its fluctuation is the obvious one.

Taking the square of (7.45), we have

$$\begin{aligned} \langle z(t)^2 \rangle &= z_0^2 e^{-2\bar{\alpha}_\varepsilon t} [\cos(\Omega_s t)]^2 \\ &\quad + \frac{4\alpha_c^2 \alpha_\varepsilon \sigma_\varepsilon^2}{\Omega_s^2} \int_0^t d\tau \int_0^t d\mu \langle \xi(\tau) \xi(\mu) \rangle \\ &\quad \times e^{-\bar{\alpha}_\varepsilon(t-\tau)} e^{-\bar{\alpha}_\varepsilon(t-\mu)} \sin(\Omega_s(t-\tau)) \sin(\Omega_s(t-\mu)) \end{aligned} \quad (7.46)$$

Combining the trigonometric functions, neglecting a small correction to the frequency due to the damping and evaluating the single integral which remains after using the Markov property (7.44), we finally obtain

$$\begin{aligned} \langle z(t)^2 \rangle &= \frac{1}{2} z_0^2 e^{-2\bar{\alpha}_\varepsilon t} \{1 + \cos(2\Omega_s t)\} \\ &\quad + \bar{\sigma}_z \left\{ 1 - e^{-2\bar{\alpha}_\varepsilon t} \left(1 + \frac{\bar{\alpha}_\varepsilon^2}{\Omega_s^2} \right) \right\} \\ &\quad + \bar{\sigma}_z \frac{\bar{\alpha}_\varepsilon^2}{\Omega_s^2} e^{-2\bar{\alpha}_\varepsilon t} \cos(2\Omega_s t) \\ &\quad + \bar{\sigma}_z \frac{\bar{\alpha}_\varepsilon}{\Omega_s} e^{-2\bar{\alpha}_\varepsilon t} \sin(2\Omega_s t) \end{aligned} \quad (7.47)$$

This solution has the properties that we expect:

$$\langle z(0)^2 \rangle = z_0^2, \quad \lim_{t \rightarrow \infty} \langle z(t)^2 \rangle = \bar{\sigma}_z^2 \quad (7.48)$$

and, with the solution for the bunch centroid $\langle z(t) \rangle$ contained in (7.45) we have a complete description of the time-dependence of both the centre of gravity of the bunch and the r.m.s. bunch length. The centre of gravity behaves like an oscillator with frequency Ω_s and damping time $\bar{\tau}_e$. Since the first term in (7.47) is just the square of this solution, the time-dependent r.m.s. bunch length,

$$\langle z(t)^2 \rangle - \langle z(t) \rangle^2, \quad (7.49)$$

is given by the three terms proportional to the equilibrium value $\bar{\sigma}_z^2$. Starting from zero at $t = 0$ this gradually grows until the natural bunch length is attained with the centre of gravity at the value of z corresponding to the stable phase angle.

It is tedious, but not difficult, to extend this solution to cover more general initial conditions than (7.42). This is one way to derive the fundamental solution and Green function of the Fokker-Planck equation.

Since we used the linear approximation to the RF potential well, this solution does not apply, of course, to initial conditions lying outside the RF bucket. The corrections which have to be applied to take account of the nonlinearities inside the bucket are not too important since the bunch rapidly damps down to small amplitudes. However, if the initial condition lies too close to the separatrix there may be some loss of particles and a consequent inefficiency in the injection process.

In the usual mode of beam accumulation in e^+e^- colliding beam rings or synchrotron light sources, small amounts of charge are injected one at a time from the injector linac or synchrotron. Each such packet damps down to join the main bunch circulating around the equilibrium orbit according to the solution we have derived.

In reality the *coherent damping rate* of the phase oscillations of the centre-of-gravity of the *main bunch* will often be different from the radiation damping rate given here because of collective effects^{82,83,84} such as *Robinson damping*, *head-tail damping* or *Landau damping* which are beyond the scope of these lectures. This is not the case for the damping of small injected bunches towards an accumulated bunch containing the bulk of the particles.

Finally we should also mention that injection into betatron phase space is often used as an alternative to the injection into synchrotron phase space described here. The mathematics is similar.

7.8 CORRELATION FUNCTIONS OF DYNAMICAL VARIABLES

We can extract even more important physical information from the solution (7.45) of the stochastic oscillator equation (7.43) by forming the correlation function

$$\langle z(t)z(0) \rangle = z_0 \langle z(t) \rangle = z_0^2 e^{-\bar{\alpha}t} \cos(\Omega_s t). \quad (7.50)$$

This shows how the initial phase and amplitude coherence of particles in the injected bunch are lost with a time constant equal to the damping time. This property is quite general and forms the basis for many arguments concerning how long particles “remember” their initial oscillation phases and amplitudes. For example one may argue that, in nonlinear orbit tracking studies, it is not meaningful to track particles for much longer than a damping time.

It is left as an exercise for the reader to generalise (7.50) to a general initial oscillation phase ($\varepsilon(0) = \varepsilon_0 \neq 0$). As a further step one can then average the initial conditions ε_0, z_0 over the stationary distribution to find all the *two-time correlation functions* of ε and z in the equilibrium state of the bunch.

7.9 QUANTUM FLUCTUATIONS IN RADIAL BETATRON PHASE SPACE

Since it will turn out that the effects of quantum fluctuations in radial betatron phase space are quite analogous to longitudinal phase space, we shall limit ourselves to deriving the Fokker-Planck equation in action-angle variables and deducing the formula for the equilibrium emittance. The other results and techniques discussed in the previous sections can then be applied, with appropriate substitutions, to the betatron oscillations.

We have already studied the deterministic parts of the equations of motion (5.55), which led us to the radiation damping effects. It only remains for us to evaluate the fluctuating parts of the radiation terms. We can make the same approximation as before, neglecting the higher order terms in δ_s , to find

$$\begin{aligned} \hat{\Pi}_t &= \sqrt{c_2} p_0 c |G(s)|^{3/2} \xi(s), \\ \hat{\Pi}_x &= \sqrt{c_2} p_0 c \sqrt{\frac{2I_x}{\beta_x}} (\alpha_x \cos \phi_x + \sin \phi_x) |G(s)|^{3/2} \xi(s). \end{aligned} \quad (7.51)$$

For the sake of brevity, we shall not display the expressions for the Fokker-Planck coefficients before the final application of the averaging method. Since the algebra begins to get tedious, it is best to use a computer algebra program to obtain these terms, particularly when higher order approximations are desired.

After averaging and making the smooth approximation, we find the equations of motion in the standard form

$$\begin{aligned}\psi_x' &= K_\psi(\psi_x, I_x) + Q_\psi(\psi_x, I_x)\xi(s), \\ I_x' &= K_I(\psi_x, I_x) + Q_I(\psi_x, I_x)\xi(s).\end{aligned}\tag{7.52}$$

From these, after smoothing, averaging and neglect of higher order terms, we obtain the coefficients of the Fokker-Planck equation

$$\begin{aligned}\frac{\partial F(I_x, \psi_x, s)}{\partial s} &= -\frac{\partial}{\partial I_x} [D_I F(I_x, \psi_x, s)] - \frac{\partial}{\partial \psi_x} [D_\psi F(I_x, \psi_x, s)] \\ &+ \frac{1}{2} \frac{\partial^2}{\partial I_x^2} [Q_I^2 F(I_x, \psi_x, s)] + \frac{\partial^2}{\partial I_x \partial \psi_x} [Q_I Q_\psi F(I_x, \psi_x, s)] \\ &+ \frac{1}{2} \frac{\partial^2}{\partial \psi_x^2} [Q_\psi^2 F(I_x, \psi_x, s)].\end{aligned}\tag{7.53}$$

Here, the averaged drift terms are

$$D_I = K_I + \frac{1}{2} \frac{\partial Q_I}{\partial I_x} Q_I + \underbrace{\frac{1}{2} \frac{\partial Q_I}{\partial \psi_x} Q_\psi}_{\substack{\text{important} \\ \text{again!}}} = -\frac{2J_x \alpha_x}{c} (I_x - \epsilon_x) + \dots,\tag{7.54}$$

$$D_\psi = K_\psi + \frac{1}{2} \frac{\partial Q_\psi}{\partial I_x} Q_I + \frac{1}{2} \frac{\partial Q_\psi}{\partial \psi_x} Q_\psi = \frac{\nu_x}{R} + \dots,$$

and the diffusion terms are

$$\begin{aligned}Q_I^2 &= \frac{4J_x \alpha_x \epsilon_x I_x}{c} + \dots, \\ Q_I Q_\psi &= \{\text{oscillating terms only}\}, \\ Q_\psi^2 &= \frac{J_x \alpha_x \epsilon_x}{c I_x} + \dots,\end{aligned}\tag{7.55}$$

The analogy with (7.25) and (7.26) is obvious. We pause only to note the definition of the beam emittance in the radial plane

$$\begin{aligned}\epsilon_x = \langle I_x \rangle &= \frac{c_2 p_0^2 f_0 I_5}{4 \alpha_x J_x} \\ &= \frac{55}{32\sqrt{3}} \frac{\hbar}{mc} \left(\frac{p_0}{mc} \right)^2 \frac{I_5}{J_x I_2}.\end{aligned}\tag{7.56}$$

This, of course is very closely comparable to (7.10), the only differences being the replacement of J_e by the radial damping partition number and the replacement

of the synchrotron integral I_3 by I_5 . The definition of I_5 in Table 1 uses the function

$$\mathcal{X}(s) = \{\eta(s)^2 + (\beta_x(s)\zeta(s) + \alpha_x(s)\eta(s))^2\}/\beta_x(s). \quad (7.57)$$

As a function of the dispersions η and ζ this is identical in form to the expression of the action variable I_x in terms of x_β and p_β .

Finally let us record the form of the averaged Fokker-Planck equation for betatron oscillations in smooth approximation

$$\begin{aligned} \tau_x \frac{\partial F}{\partial t} = & \underbrace{-\frac{\partial}{\partial I_x} [J_x(-I_x + \epsilon_x)F]}_{\text{damping of emittance}} + \underbrace{2\pi f_0 \nu_x \tau_x \frac{\partial F}{\partial \psi}}_{\text{phase advance}} \\ & + \underbrace{2J_x \epsilon_x \frac{\partial^2}{\partial I_x^2} (I_x F)}_{\text{emittance diffusion}} + \underbrace{\frac{J_x \epsilon_x}{2I_x} \frac{\partial^2 F}{\partial \psi^2}}_{\text{phase diffusion}}, \end{aligned} \quad (7.58)$$

which, just like (7.28), has the stationary solution

$$F_0(I_x) = \frac{1}{\epsilon_x} \exp\left(-\frac{I_x}{\epsilon_x}\right). \quad (7.59)$$

The most important physical consequence of this is the formula for the equilibrium beam size (in real space) which contains components arising both from the distribution in synchrotron oscillation amplitudes and betatron amplitudes. Taking the expression for x from (5.6), we first find the expression for the mean horizontal position of the beam

$$\langle x \rangle = \langle x_\beta + \eta \delta_s + \eta \epsilon \rangle = \eta \delta_s \quad (7.60)$$

and then the mean-square beam size

$$\begin{aligned} \langle x^2 \rangle &= \langle x_\beta^2 \rangle + \langle \eta \epsilon^2 \rangle \\ &= \sigma_{x_\beta}^2 + \sigma_{x_\epsilon}^2 \\ &= \beta_x \epsilon_x + \eta^2 \sigma_\epsilon^2. \end{aligned} \quad (7.61)$$

The r.m.s. beam size in the radial direction at a given point in the storage ring therefore depends on the optical functions η and β_x as well as on the values of both the energy spread and emittance determined by the balance between radiation damping and quantum fluctuations.

8. THE VLASOV EQUATION

In the preceding sections we have been considering statistical phenomena in electron rings which arise due to randomness in the equations of motion of *single particles*. We underline the contrast with Gibbsian statistical mechanics, where a statistical description becomes necessary because of the sheer number of particles—not because of any intrinsic randomness in their dynamics; the equations of motion are known differential equations although their solutions may still exhibit all the properties of random functions.

In accelerator physics we consider beams whose intensity may become so high that it is no longer valid to neglect the influence of the particles upon each other. It is then that the theory of collective phenomena comes into play and tools like the Vlasov equation are required.

The Vlasov equation describes collective effects of many particles on a given particle in terms of the evolution of a one-particle distribution function which is a reduced description of the complete N -particle distribution function. The solution of this equation is a subject of the lectures on collective phenomena in this School⁸² and elsewhere.⁸³ Our purpose in this chapter is rather to lay the statistical mechanical foundation for these discussions and to clarify the relations and differences between the different kinetic equations (Fokker-Planck and Vlasov for example) that are encountered in the literature of theoretical accelerator physics.

8.1 MANY PARTICLE SYSTEMS

In Chapter 2, we described the exact state of a single particle by means of the Klimontovich phase space density. Now we have to make the obvious, if somewhat daunting, generalization from the 6-dimensional phase space of a single particle to the $6N$ -dimensional phase space of a system on N particles.

Let the state of the system at time t be given by the set of N single particle coordinates $\{\mathbf{X}_1(t), \dots, \mathbf{X}_N(t)\}$ which, by analogy with (2.1), evolve according to some equations of motion of the form

$$\dot{\mathbf{X}}_i(t) = \mathbf{K}(\mathbf{X}_i, t) + \sum_{j \neq i} \mathbf{L}(\mathbf{X}_i, \mathbf{X}_j, t) + \mathbf{Q}(\mathbf{X}_i, t) \xi(t), \quad \text{for } i = 1, \dots, N. \quad (8.1)$$

Here, as before, the function $\mathbf{K}(\mathbf{X}, t)$ describes the deterministic external fields acting on a particle at point \mathbf{X} in phase space; these include the magnets, RF cavities and electrostatic devices that exert external electromagnetic forces on particles in an accelerator or storage ring. It also contains the average components of dissipative forces like radiation damping or energy loss by scattering on residual gas in the vacuum chamber, multiple scattering as a beam traverses a foil, and so on.

The effects of random forces which act on one particle independently of any other are taken into account by the function $\mathbf{Q}(\mathbf{X}, t)$. Our primary examples in this category were the quantum fluctuation effects but fluctuations in the other dissipative effects may also be important.

The new ingredient in (8.1) is the function $\mathbf{L}(\mathbf{X}_i, \mathbf{X}_j, t)$ which describes *two-body interactions* of individual particles in the beam with others. Since the average quantities which result from the statistical analysis will be required to provide *self-consistent* calculations of the parameters of the beam as a whole, the equations of motion (8.1) have to include all the necessary coupling effects between every pair of phase space coordinates \mathbf{X}_i and \mathbf{X}_j . Each $\mathbf{X}_i(t)$ feels the influence of every one of the other particles with the appropriate strength. Such forces are mainly direct Coulomb (space charge) and other electromagnetic interactions like plasma effects or two-particle scattering. These interactions may be modified by boundaries (wakefield or impedance effects), or external agencies (feedback systems, stochastic cooling, *etc.*).

8.2 PHASE SPACE DENSITY

Thus, the appropriate exact microscopic phase-space density for a system of N particles is

$$\mathcal{F}(\mathbf{X}; t) = \frac{1}{N} \sum_{i=1}^N N \delta(\mathbf{X} - \mathbf{X}_i(t)), \quad (8.2)$$

where, in analogy with (2.1), $(\mathbf{X}_1(t), \dots, \mathbf{X}_N(t))$ denotes the solution of the equations of motion (8.1) for some given initial conditions.

It should be carefully noted that (8.2) is a function on a 6-dimensional—*not* a $6N$ -dimensional—single particle phase space. The difference from (2.3) is that now there is a swarm of N δ -function points moving around in this space under the influence of their mutual interactions.

Now the continuity equation takes the form

$$\begin{aligned} \partial_t \mathcal{F}(\mathbf{X}, t) + \nabla_{\mathbf{X}} \cdot [\mathbf{K}(\mathbf{X}, t) \mathcal{F}(\mathbf{X}, t)] + \nabla_{\mathbf{X}} \cdot [\mathbf{Q}(\mathbf{X}, t) \xi(t) \mathcal{F}(\mathbf{X}, t)] \\ + \nabla_{\mathbf{X}} \cdot \int d\mathbf{X}' L(\mathbf{X}, \mathbf{X}', t) \mathcal{F}(\mathbf{X}, t) \mathcal{F}(\mathbf{X}', t) \\ = 0. \end{aligned} \quad (8.3)$$

Again, this microscopic Liouville-Klimontovich equation contains too much information for us to cope with. To clarify the following, let us leave out the noise terms \mathbf{Q} and concentrate on how we can deal with the two-body interaction terms.

Instead of averaging over all realisations of a random function, let us now average over the *Liouville distribution*. This is an ensemble of systems which give rise to identical macroscopic averages. With respect to this average, we can define the *one-particle distribution function*

$$f_1(\mathbf{X}, t) = \langle \mathcal{F}(\mathbf{X}, t) \rangle. \quad (8.4)$$

and hope that it may be a smooth function for which we can derive an autonomous equation of evolution in some approximation. Unfortunately we cannot average (8.3) directly since the average of the product of two microscopic distribution functions $\mathcal{F}(\mathbf{X}, t)\mathcal{F}(\mathbf{X}', t)$ would appear in the two-body interaction term.

To deal with this obstacle, we calculate the appropriate average as follows:

$$\begin{aligned} \langle \mathcal{F}(\mathbf{X}, t)\mathcal{F}(\mathbf{X}', t) \rangle &= \left\langle \frac{1}{N^2} \sum_{i=1}^N \delta(\mathbf{X} - \mathbf{X}_i(t)) \sum_{j=1}^N \delta(\mathbf{X}' - \mathbf{X}_j(t)) \right\rangle \\ &= \frac{1}{N^2} \sum_{i=1}^N \langle \delta(\mathbf{X} - \mathbf{X}') \delta(\mathbf{X} - \mathbf{X}_i(t)) \rangle \\ &\quad + \frac{1}{N^2} \sum_{i \neq j=1}^N \langle \delta(\mathbf{X} - \mathbf{X}_i(t)) \delta(\mathbf{X}' - \mathbf{X}_j(t)) \rangle \\ &= \frac{1}{N} \delta(\mathbf{X} - \mathbf{X}') f_1(\mathbf{X}, t) + f_2(\mathbf{X}, \mathbf{X}', t) \end{aligned} \quad (8.5)$$

where, in the double sum, we separated out the terms with $i = j$. The *two-particle distribution function*, $f_2(\mathbf{X}, \mathbf{X}', t)$, is defined by this relationship.

8.3 TWO-PARTICLE CORRELATIONS

The decomposition (8.5) may be substituted into the average of the Liouville equation (8.3) to obtain

$$\begin{aligned} \partial_t f_1(\mathbf{X}, t) + \nabla_{\mathbf{X}} \cdot [\mathbf{K}(\mathbf{X}, t) f_1(\mathbf{X}, t)] \\ = - \nabla_{\mathbf{X}} \cdot \underbrace{\int d\mathbf{X}' \mathbf{L}(\mathbf{X}, \mathbf{X}', t) \delta(\mathbf{X} - \mathbf{X}') f_1(\mathbf{X}, t)}_{=0 \text{ by symmetry}} \\ - \nabla_{\mathbf{X}} \cdot \int d\mathbf{X}' \mathbf{L}(\mathbf{X}, \mathbf{X}', t) f_2(\mathbf{X}, \mathbf{X}', t) \end{aligned} \quad (8.6)$$

which is known as the *first equation of the BBGKY hierarchy*, the acronym standing for the names of Born, Bogoliubov, Green, Kirkwood and Yvon.

It is customary to decompose $f_2(\mathbf{X}, \mathbf{X}', t)$ into a part which is simply the uncorrelated product of two one-particle distribution functions and a part which contains all the information about the correlations:

$$f_2(\mathbf{X}, \mathbf{X}', t) = f_1(\mathbf{X}, t) f_1(\mathbf{X}', t) + g_2(\mathbf{X}, \mathbf{X}', t). \quad (8.7)$$

The new function $g_2(\mathbf{X}, \mathbf{X}', t)$ is called the *two-particle correlation function*.

8.4 THE VLASOV EQUATION

Substitution of (8.7) into (8.6) leads to

$$\begin{aligned} \partial_t f_1(\mathbf{X}, t) + \nabla_{\mathbf{X}} \cdot [\mathbf{K}(\mathbf{X}, t) f_1(\mathbf{X}, t)] = & -\nabla_{\mathbf{X}} \cdot \int d\mathbf{X}' \mathbf{L}(\mathbf{X}, \mathbf{X}', t) f_1(\mathbf{X}, t) f_1(\mathbf{X}', t) \\ & - \nabla_{\mathbf{X}} \cdot \int d\mathbf{X}' \mathbf{L}(\mathbf{X}, \mathbf{X}', t) g_2(\mathbf{X}, \mathbf{X}', t) \end{aligned} \quad (8.8)$$

If g_2 is neglected, this becomes the *Vlasov equation*:

$$\partial_t f_1(\mathbf{X}, t) + \nabla_{\mathbf{X}} \cdot [\mathbf{K}(\mathbf{X}, t) f_1(\mathbf{X}, t)] = -\nabla_{\mathbf{X}} \cdot \int d\mathbf{X}' \mathbf{L}(\mathbf{X}, \mathbf{X}', t) f_1(\mathbf{X}, t) f_1(\mathbf{X}', t) \quad (8.9)$$

The interpretation of the left-hand side of this equation should be familiar by now. It describes the so-called “free-streaming” or evolution of the single particle distribution due to the direct action of external fields. In electron rings it also includes the radiation damping effects. Exactly the same terms appeared earlier in the Fokker-Planck equation.

The term on the right-hand side is new and is known as the *Vlasov or self-consistent field* term. It describes the additional evolution of $f_1(\mathbf{X}, t)$ due to the *average field of all the particles in the system acting back on themselves*. The strength of this field depends on $f_1(\mathbf{X}, t)$ itself with the result that (8.9) is second-order in the single-particle distribution function. The \mathbf{L} appearing in the integral ensures that the particles at position \mathbf{X}' act on the particles at \mathbf{X} with the appropriate strength according to the two-particle coupling term which appeared in the many-body Hamiltonian or the equivalent equations of motion (8.1).

Looking at it this way, it may appear that this is *all* one has to include. However we must not forget that the Vlasov equation only includes the force acting on the average value, $f_1(\mathbf{X}, t)$ of the particle distribution at \mathbf{X} due to the average value $f_1(\mathbf{X}', t)$ of the distribution function at \mathbf{X}' . In representing the dynamics by the Vlasov equation one completely neglects the possibility that the presence of a particle at \mathbf{X} may be correlated with the presence (or absence) of a particle at \mathbf{X}' . For example we know that when two particles

approach sufficiently closely a two-body scattering event will occur. Such effects introduce correlations between the phase space coordinates of pairs of particles and are precisely those which we neglected in dropping the last term of (8.8).

From this, and from the fact that there is no difficulty in incorporating radiation damping into it, it should be clear that the Vlasov equation is both physically and logically distinct from Liouville's theorem. It should also be clear that, had we carried along the noise terms from (8.1), we should by now have an equation containing both the Vlasov self-consistent field and the Fokker-Planck diffusion terms to describe quantum fluctuations. Fortunately, it is rarely necessary to attempt to solve this combined equation since the characteristic rise or damping time scales of the instabilities described by the Vlasov term are generally so much shorter than the radiation damping times that the phenomena can be treated separately. Indeed the starting point of instability theory is often to assume that a stationary distribution has been established by the radiation effects and then to examine how the growth or decay rates of small perturbations of this equilibrium are influenced by the collective forces.

In plasma physics, the Vlasov equation describes the collective Coulomb interaction of a charged plasma with itself and much of that subject is devoted to the analysis of the consequences of this interaction. In accelerator physics, on the other hand, the Vlasov equation appears mainly in the context of single-beam collective effects where the dominant interaction force arises not from the direct Coulomb (or "space-charge") forces but from the electromagnetic interaction of the beam with itself *via its environment*. This makes it much harder to write down an explicit expression for the interaction operator \mathbf{L} . This now depends on the solutions of Maxwell's equations for an arbitrary charge distribution traversing the various structures (RF cavities, separator tanks, vacuum chamber non-uniformities, *etc.*) which compose the beam's environment.

In practice the interaction operator is represented either in terms of *wake fields* (or potentials) or, when one works in a Fourier representation ("frequency domain"), in terms of *impedances*. Further discussion of this topic^{84,83} is beyond the scope of these lectures

8.5 CORRELATION EFFECTS AND BBGKY HIERARCHY

Going beyond the Vlasov approximation of neglecting the correlations among particles, one can derive the second equation of the BBGKY hierarchy which would describe the evolution of $g_2(\mathbf{X}, \mathbf{X}', t)$ in terms of the three-particle correlation function $g_3(\mathbf{X}, \mathbf{X}', \mathbf{X}'')$ and so on. Clearly some means has to be found to truncate the infinite chain of coupled equations. One approach, pioneered by Bogoliubov,³⁵ is to express the higher order correlations as functionals of the lower order ones. Thus, for example, one can try to express the two-particle

correlation function as a functional of the one-particle distribution function:

$$g_2(\mathbf{X}, \mathbf{X}', t) = G[f_1; \mathbf{X}, \mathbf{X}', t]. \quad (8.10)$$

This *ansatz* is known as *Bogoliubov's synchronization hypothesis*. It assumes that, on the short time scale during which the direct pair interaction takes place, one may reasonably approximate the collision integral in (8.8) in terms of f_1 ; in this way it becomes a closed equation for f_1 :

$$\begin{aligned} \partial_t f_1(\mathbf{X}, t) + \nabla_{\mathbf{X}}[\mathbf{K}(\mathbf{X}, t) f_1(\mathbf{X}, t)] = & \\ - \underbrace{\nabla_{\mathbf{X}} \int d\mathbf{X}' \mathbf{L}(\mathbf{X}, \mathbf{X}', t) f_1(\mathbf{X}, t) f_1(\mathbf{X}', t)}_{\text{Vlasov term}} & \\ - \underbrace{\nabla_{\mathbf{X}} \int d\mathbf{X}' \mathbf{L}(\mathbf{X}, \mathbf{X}', t) G(f_1(\mathbf{X}, t), f_1(\mathbf{X}', t))}_{\text{collision term}} & \end{aligned} \quad (8.11)$$

In the kinetic theory of dilute gases^{44,85} it is reasonable to keep only those terms which are of first order in the density. Then the Vlasov term vanishes and (8.11) becomes the famous kinetic equation of *Boltzmann*. For the case of weak-coupling or, equivalently, in the small-angle scattering approximation, it reduces to the *Landau equation*^{44,26} (sometimes known as the *Fokker-Planck equation*—see Section 2). In plasma theory, one can include frequency-dependent dielectric screening of the Coulomb interaction through the plasma dielectric function and obtain the *Balescu-Lenard-Guernsey equation*.^{44,26} This also reduces to the Landau equation (with static Debye screening) in the low frequency limit.

In the context of particle accelerator theory, it should be possible to derive kinetic equations analogous to Boltzmann's which would describe the intra-beam scattering effect. To the writer's knowledge, this theoretical program has yet to be carried through although alternative theories of intra-beam scattering are available.^{77,86} For electron storage rings at low energy (where this effect is important) it may be necessary to include terms of the Fokker-Planck type in the same equation in order to describe the radiation effects. The equilibrium emittances will be determined by a balance between radiation damping, quantum fluctuations and blow-up due to intra-beam scattering. In proton rings above transition energy there is an inexorable blow-up which can only be countered by stochastic cooling.

ACKNOWLEDGEMENTS

It would be very difficult to list all the colleagues at CERN, SLAC, and several other laboratories and universities who have influenced my understanding of what is set forth here; however I must make particular mention of John Dougherty, Albert Hofmann, Eberhard Keil, Bryan Montague, Phil Morton and Ron Ruth. I hope that the list of references will serve as the extension of this list to those from whom I have learned through their writings.

These lectures were written up during ten months spent at SLAC on leave from CERN. With them, I should like to express my gratitude to SLAC for the opportunity to work in its uniquely stimulating and exciting environment. This especially applies to Ewan Paterson and my colleagues in the Accelerator Department for the hospitality, shared wisdom and friendship which made my visit so enjoyable and instructive. My special thanks are due to Melvin Month whose vision and continual encouragement were of great value to me during the writing.

REFERENCES

1. J. Bisognano and S. Chattopadhyay in M. Month, P.F. Dahl and M. Dienes (Eds.), *Physics of High Energy Particle Accelerators* (Brookhaven Summer School 1983), AIP Conference Proceedings No. 127, (1984).
2. H.G. Hereward and W. Schnell in M.H. Blewett (Ed.), *Theoretical aspects of the behaviour of beams in accelerators and storage rings*, Proceedings of the International School of Particle Accelerators, CERN 77-13 (1977).
3. P. Bryant and S. Newman (Eds.), CERN Accelerator School, *Antiprotons for colliding beam facilities*, CERN 84-15 (1984).
4. D. Möhl in Ref. 3.
5. G. Dôme in Ref. 3.
6. J.M. Jowett, *Dynamics of electrons in storage rings including non-linear damping and quantum excitation effects*, Proc. 12th International Conference on High-Energy accelerators, Batavia, 1983 and CERN LEP-TH/83-43, 1983.
7. J.M. Jowett, M. Month and S. Turner (Eds.), *Nonlinear Dynamics Aspects of Particle Accelerators*, Proceedings of the Joint US/CERN School on Particle Accelerators, Sardinia 1985, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986.
8. J.M. Jowett, *Non-linear dissipative phenomena in electron storage rings*, in Ref. 7.

9. CERN Accelerator School, *Advanced Accelerator Physics*, (Oxford, 16–27 September 1985), proceedings to be published as a CERN Yellow Report (1986).
10. J.M. Jowett, *Electron dynamics with radiation and nonlinear wigglers*, in Ref. 9.
11. C. Bernardini and C. Pellegrini, *Linear theory of motion in electron storage rings*, *Ann. Phys.* **46**, 174 (1968).
12. H. Bruck, *Accélérateurs circulaires des particules*, Presses Universitaires de France, Paris, 1966.
13. M. Bell and J.S. Bell, *Radiation damping and Lagrange invariants*, *Particle Accelerators*, **13**, 13 (1983).
14. H. Wiedemann, these proceedings.
15. A. Chao, these proceedings.
16. R.A. Carrigan, F.R. Huson, M. Month (Eds.), *Physics of High Energy Particle Accelerators* (Fermilab Summer School 1981), AIP Conference Proceedings No. 87, (1982)
17. E. Keil, *Storage ring design* in Ref. 16.
18. P. Bryant and S. Turner (Eds.), CERN Accelerator School, *General Accelerator Physics*, (Orsay, 3–14 September 1984), CERN 85-19 (1985).
19. M. Sands, *The Physics of electron storage rings*, SLAC-121 (1970).
20. K. Hübner, *Synchrotron radiation and Radiation damping* in Ref. 18.
21. J.M. Jowett, *Luminosity and energy spread in LEP*, CERN-LEP-TH/85-04 (1985).
22. E.J.N. Wilson, these proceedings.
23. F.T. Cole, these proceedings.
24. R.D. Ruth, these proceedings.
25. Iu. L. Klimontovich, *Statisticheskaya Teoriya Neravnovesnykh Protsessov v Plazma*, Izd. Moska. Univ. (1964), and *Statistical theory of non-equilibrium processes in plasmas*, Pergamon, Oxford, 1967.
26. P.C. Clemmow and J.P. Dougherty, *Electrodynamics of particles and plasmas*, Addison-Wesley, Reading, Mass., 1969.
27. V. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, New York, 1978.
28. P. Langevin, *Sur la théorie du mouvement brownien*, *C.R. Acad. Sci. Paris* **146**, 530 (1908).
29. R.L. Stratonovich, *Topics in the Theory of Random Noise*, Vols. I and II, Gordon and Breach, New York, 1967.

30. C.W. Gardiner, *Handbook of stochastic methods for chemistry and the natural sciences*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1983.
31. L. Garrido, P. Seglar and P.J. Shepherd (Eds.), *Stochastic processes in non-equilibrium systems*, Sitges International School of Statistical Mechanics, June 1978, Springer-Verlag, Berlin, 1978.
32. H. Haken, *Synergetics*, Springer-Verlag, Berlin, 1977.
33. N. Wax (ed.), *Selected papers on noise and stochastic processes*, Dover, New York, 1954.
34. A.H. Nayfeh, *Perturbation methods in nonlinear dynamics*, in Ref. 7.
35. N.N. Bogoliubov, *Problemy Dinamicheskoy Teorii v Statisticheskaya Fizika*, Moscow, 1946 and *Problems of a dynamical theory in statistical physics* in J. de Boer and G.E. Uhlenbeck (Eds.), *Studies in statistical mechanics*, Vol. 1, North Holland, Amsterdam, 1962.
36. K. Itô, *On stochastic differential equations*, Amer. Math. Soc., New York (1951).
37. H. Haken, *Advanced Synergetics*, Springer-Verlag, Berlin, 1983.
38. S. Chandrasekhar, *Stochastic problems in physics and astronomy*, *Rev. Mod. Phys.* **15**, 1 (1943), reprinted in Ref. 33.
39. S. Kheifets, *Blowup of a weak beam due to interaction with a strong beam in an electron storage ring*, Proc. 12th International Conference on High-Energy Accelerators, Batavia, 1983.
40. B.V. Chirikov, *Dokl. Acad. Nauk.*, 125, 5 (1959)
41. A.W. Chao and M. Month, *Particle trapping during passage through a high order nonlinear resonance*, *Nucl. Inst. and Meth.* **121** 129 (1974)
42. J.L. Tennyson, *The dynamics of the beam-beam interaction*, in Ref. 16.
43. H.G. Hereward, *Diffusion in the presence of resonances*, CERN ISR-DI/72-26 (1972).
44. R. Balescu, *Equilibrium and non-equilibrium statistical mechanics*, Wiley, New York, 1975.
45. S. Ichimaru, *Basic Principles of Plasma Physics*, W.A. Benjamin, New York, 1973.
46. C.N. Lashmore-Davies, *Kinetic theory and the Vlasov equation*, in Ref. 9.
47. J.D. Jackson, *Classical electrodynamics*, Wiley, New York, 1975.
48. L.D. Landau and E.M. Lifshitz, *The classical theory of fields*, Pergamon, Oxford (1975).
49. A. Hofmann, these proceedings.

50. A.A. Sokolov and I.M. Ternov, *Synchrotron radiation*, Pergamon Press, Berlin, 1968.
51. B.W. Montague, *Polarized beams in high energy storage rings*, *Physics Reports* **113** No. 1 (1984).
52. A.W. Chao, *Polarization of a stored electron beam*, in Ref. 16.
53. M. Sands, *Phys. Rev.* **97**, 470, (1955), and SLAC-SPEAR Note 9 (1969).
54. R.J. Noble, *Beamstrahlung from colliding electron-positron beams with negligible disruption*, SLAC-PUB 3871 (1986).
55. A.I. Khinchin, *Mathematical foundations of statistical mechanics*, Dover, New York (1949).
56. E.D. Courant and H.S. Snyder, *Theory of the alternating gradient synchrotron*, *Ann. Phys.* **3**, 1 (1958).
57. F.C. Iselin, E. Keil and E. Nordmark, *The MAD program, Version 5, Users' Reference Manual*, CERN/LEP-TH/85-15 (1986).
58. K.L. Brown, D.C. Carey, Ch. Iselin and F. Rothacker, *TRANSPORT: a computer program for designing charged particle beam transport systems*, CERN-80-04 (1984).
59. R.D. Ruth, *Single particle dynamics and nonlinear resonances in circular accelerators*, in Ref. 7.
60. K.W. Robinson, *Radiation effects in circular accelerators*, *Phys. Rev.* **111**, 373 (1958).
61. P. Chen, private communication.
62. P.L. Morton, private communication.
63. G. Ripken, *Untersuchungen zur Strahlführung und Stabilität der Teilchenbewegung in Beschleunigern und Storage-Ringen unter strenger Berücksichtigung einer Kopplung der Betatrons chwingungen*, DESY R1-70/4 (1970).
64. A. Piwinski, *Strahlparameter bei Kopplungen, horizontaler und vertikaler Krümmung und Raumladungseffekt*, unpublished (1971).
65. K. Steffen and J. Kewisch, *Study of integer difference resonances in distorted PETRA optics*, DESY PET-76/09 (1976), see also G. Guignard and Y. Marti, *PETROC users' guide*, CERN/ISR-BOM-TH/81-32 (1981).
66. A.W. Chao, *Study of beam distribution parameters in an electron storage ring*, *J. Appl. Phys.* **50**, 595 (1979).
67. R.H. Helm, M.J. Lee, P.L. Morton and M. Sands, *Evaluation of synchrotron radiation integrals*, *IEEE Trans. Nucl. Sci.* **NS-20**, 900 (1973).

68. M. Hanney and E. Keil, *BEAMPARAM - a program for computing beam dynamics and performance of e^+e^- storage rings*, CERN ISR/TH/79-29 (1979)
69. M.D. Woodley, M.J. Lee, J. Jaeger and A.S. King, *Control of machine functions for transport systems*, *IEEE Trans. Nucl. Sci.* NS-30, (1972).
70. H. Wiedemann, SLAC PEP Note 220 (1976).
71. A. Hofmann, *Attempt to change the longitudinal particle distribution by a dipole-octupole wiggler*, CERN-LEP Note 192 (1979).
72. A. Hofmann and B. Zotter, *Combined function electron storage rings*, *IEEE Trans. Nucl. Sci.* NS-24, 1875 (1977).
73. G. Guignard and J.M. Jowett, *Damping aperture of LEP*, LEP Note 407 (1982).
74. T. Risselada, *Calcul de J_e et $dJ_e/d\delta$* , CERN LEP Note 488 and PS/PSR Note 84-3 (1984).
75. J.-P. Delahaye and J. Jäger, *Variation of the damping partition numbers with particle energy deviation*, SLAC AP-36 (1984).
76. A. Hofmann and J.M. Jowett, *Theory of the Dipole-Octupole Wiggler, Part I: Phase Oscillations*, CERN/ISR-TH/81-23 (1981).
77. A. Piwinski, *Beam losses and lifetime* in Reference 18.
78. A. Renieri, *Possibility of achieving very high energy resolution in electron-positron storage rings*, Frascati Preprint /INF-75/6(R), (1975).
79. F. Ruggiero, private communication.
80. A. Hofmann and S. Myers, *Beam dynamics in a double RF system*, Proc. 11th International Conference on High-Energy Accelerators, Geneva, 1980, Birkhäuser, Basel, 1980
81. N.N. Bogoliubov and Y.A. Mitropolsky, *Asymptotic methods in the theory of non-linear oscillations*, Gordon and Breach, New York, 1961.
82. J.M. Wang, *these proceedings*.
83. A.W. Chao, *Coherent instabilities of a relativistic bunched beam*, in M. Month (Ed.), *Physics of High Energy Particle Accelerators* (SLAC Summer School 1982), AIP Conference Proceedings No. 105, (1983).
84. B. Zotter, *these proceedings*.
85. P. Résibois and M. de Leener, *Classical kinetic theory of fluids*, Wiley, New York, 1977.
86. J.D. Bjorken and S.K. Mtingwa, *Particle Accelerators*, 115 (1983).

SLAC - PUB - 4033
July 1986
(A)

ERRATA

INTRODUCTORY STATISTICAL MECHANICS FOR ELECTRON STORAGE RINGS*

JOHN M. JOWETT†

*Stanford Linear Accelerator Center
Stanford University, Stanford, California 94305*

Please insert the attached page 33 (Figure 3 did not print clearly on the original).

There is no Figure 8. Page 107 was inadvertently included; it is an accidental and blemished copy of Figure 1.

*Work supported by the Department of Energy, contract DE-AC03-76SF00515.

†Permanent address and address when lectures were given: CERN, CH-1211 Geneva 23.

Lectures given at the 5th U.S. Summer School on the Physics of High Energy Particle Accelerators Stanford, California, July 15-26, 1985

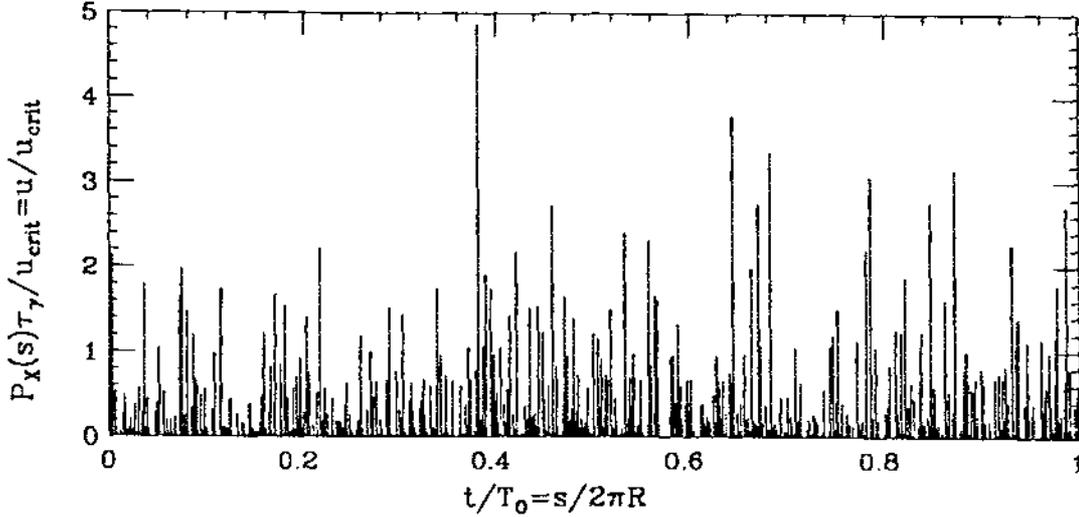


Fig. 3 A realisation of $P_X(s)$

A typical realisation of $P_X(s)$, obtained by simulation, is shown in Fig. 3; the parameters are such that the expectation value of the number of photons emitted in one revolution period is

$$N_X(s)T_0 = 1000 \quad (3.27)$$

Counting the peaks, we find 1049 in this realisation, a 1.55σ deviation. In preparing the figure, τ_γ was been taken to be fixed and equal to the width of a line on the printer so that the distribution of the heights of the peaks, displayed in units of the critical energy u_c , is given by the function $f_X(u; s)$ defined in (3.16). It is worth remarking that, although half the energy is carried away by photons with $u > u_c$, there are few such photons. In fact 91% of the photons have $u < u_c$ and 50% have $u < 0.1 u_c$.

3.4 DISTRIBUTION OF PHOTON EMISSION TIME

Let $\text{Pr}(n, s)$ denote the probability that n photons are emitted in an interval of azimuth $(0, s)$. Suppose that the emission rate per metre, $N_X(s)/c = N_X/c$ is constant, *i.e.* there is no significant variation of momentum p or magnetic field B . Then, for a sufficiently small step Δs (the probability of 2 photons being emitted in Δs is supposed to be negligible), we can write the probability that n photons are emitted in $(0, s + \Delta s)$ as the sum of the probabilities that

- a) n are emitted in $(0, s)$ and none in $(s, s + \Delta s)$, and
- b) $n - 1$ are emitted in $(0, s)$ and 1 in $(s, s + \Delta s)$.