18 July 1997 | Laboratory of Nuclear Studies, Cornell University | CBN-97-16

THEORY OF HEAD-TAIL CHROMATICITY SHARING

R. Talman

Laboratory of Nuclear Studies
Cornell University
Ithaca, NY 14853

ABSTRACT

Chromaticity compensation ($Q'_x \approx Q'_y \approx 0$) is required to avoid the "head-tail" effect in all high-energy accelerators. A potential advantage of the "Möbius" scheme, or any other strongly coupled lattice, is that, because of cooperation between damping in the two planes, a less-restrictive condition, $Q'_x \approx -Q'_y$, should be satisfactory. This "chromaticity sharing" concept has been tested at CESR and, with qualification, been found valid. During these studies numerous fascinating features have emerged. This report contains theoretical calculations bearing on these phenomena, the most important of which are the influence on the head-tail effect of coupling, decoherence, and sympathetic damping. Most results are obtained independently by the Krylov and Bogoliubov method and by near-symplectic perturbation theory. In spite of the paper's considerable length and complexity, significant observational features are left qualitatively unexplained.

1. Introduction

A simple description, due originally to Pellegrini and to Sands[1], of the "head-tail" effect to which single bunches are subject, is illustrated in Fig. 1.1. The actual charge distribution is approximated by two super-particles, each undergoing betatron and synchrotron oscillations. A "wake field" due to the beam environment trails each of the particles and, though the particle (temporarily) in front feels little force, the tail particle feels a transverse force oscillating at the head particle's (and its own) natural frequency. Being exactly "on-resonance" this force, if continued indefinitely, would cause unlimited betatron growth and eventual loss of the tail particle. But, because of synchrotron oscillation at tune $Q_s = \mu_s/(2\pi) \stackrel{\text{e.g.}}{=} 0.05$, the head and tail particles alternate roles;[†]

$$z_2 = -z_1 = \frac{A}{2}\cos\mu_s t, \quad \frac{\delta p_2}{p} = -\frac{\delta p_1}{p} = \left(\frac{dp}{p}\right)_{\text{typ}} \sin\mu_s t \tag{1.1}$$

where A is shown in Fig. 1.1 and $(\frac{dp}{p})_{\text{typ}} = 0.0006$ is the r.m.s. fractional synchrotron oscillation momentum spread assigned to the head and tail super-particles. This periodic (and in our treatment inexorable) interchange has the possiblility of stabilizing the transverse motion.

This paper describes the theory of a somewhat more complicated process in which, because of intentionally-large cross-plane coupling, betatron oscillations slosh back and forth between horizontal, x and vertical, y. For introductory purposes, even though the uncoupled description can be obtained by specializing formulas appearing later in this paper, the main results in the uncoupled case (all well-known) will now be summarized, say for horizontal motion.

For describing the one-transverse-dimensional motion of the super-particles it is helpful to introduce "σ-mode" $e_1 = (x_1 + x_2)/2$ and "π-mode" $e_2 = (x_1 - x_2)/2$ coordinates. This terminology is borrowed from the eigenmode description of two pendulums coupled by a weak spring, where the in-phase σ-mode has the same frequency as either pendulum by itself, while the frequency of the out-of-phase σ-mode is (doubly) shifted by the connecting spring. In the accelerator, for pure oscillation in the σ/π-mode, the particles' betatron

† The bunch length oscillation in this model is unphysical.

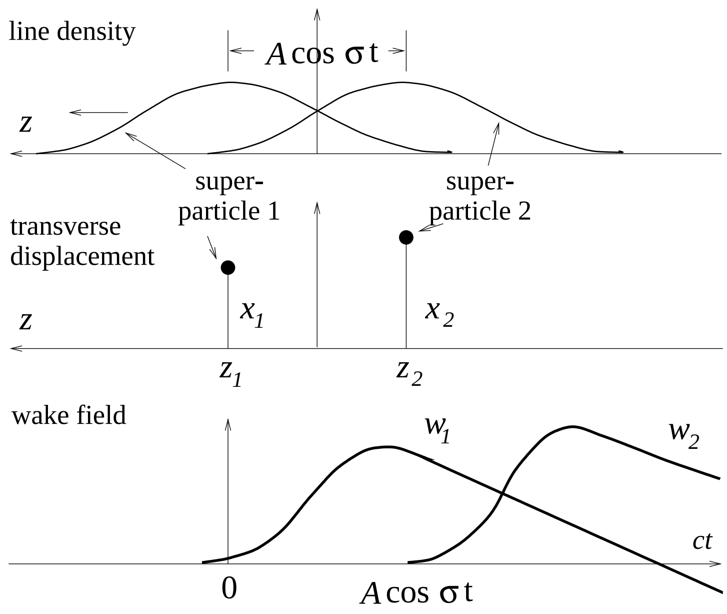

Figure 1.1: Approximation of a bunch by two super-particles.

oscillation amplitudes are predominantly equal/opposite in sign. (Because electrical beam pickups sense only centroid displacement, which vanishes in π-mode, the π-mode will also be said to be "invisible".) Whether the wake force causes betatron growth or decay depends delicately on the phase deviation from these equal or equal-but-opposite amplitude situations. Since "chromatic" (momentum) dependence is the leading cause of such phase shifts, the lattice chromaticity $Q'_x = dQ_x/(dp/p)$ acquires a special importance. Assuming $dp/p \overset{\text{e.g.}}{=} 0.0006$ and $Q'_x \overset{\text{e.g.}}{=} 1$ the maximum betatron phase deviation accumulating between head and tail particle in one betatron cycle is $2\pi\Delta Q_x/Q_x = 2\pi(Q'_x 2dp/p)/Q_x = 0.8\times10^{-3}$. Since this phase deviation continues to accumulate for one quarter cycle of synchrotron oscillation (about $0.25Q_x/Q_s \overset{\text{typ}}{=} 50$ betatron cycles) there is an extreme phase deviation, symbolized by χ, of about 0.038 radians.†

For ordinary electron rings (operated "above transition") stability of the σ-mode requires positive Q'_x. In that case, according to the simple model being described, the

† These estimates are appropriate for an electron machine, in particular CESR. For a proton machine the accumulated phase advance tends to be more than an order of magnitude greater which tends to invalidate the present two particle model. This issue will studied in more detail in a later section where experimental results are interpreted.

π-mode should be unstable. Experimentally it is found that the former prediction is correct, but the latter is false—at least for a range of somewhat positive values of Q'_x there is no evidence of instability or profile distortion that would accompany π-mode instability. This unexpected stability has been ascribed to "Landau damping" or "decoherence" of the π-mode, a form of damping that is ineffective in the σ-mode. Since decoherence is a multi-particle effect, it can obviously not be represented by our two particle model. We will concentrate on the σ modes; unlike π-modes they are externally visible because they have centroid motion.

For sufficiently strong wake fields a qualitatively different phenomenon known as the "fast head-tail effect" sets in; the damage inflicted on the tail particle during the half-period while it *is* in the tail is too great to be stabilized by the longitudinal oscillation.[2] This phenomenon is well described by the same two particle model we are using. To calculate its onset it is necessary to account also for the tune shifts caused by the wake fields. At beam currents for which these tune shifts cause two eigenfrequencies to become equal, certain denominators appearing in the response function vanish, reflecting resonance and particle loss. With the formulation complicated by the inclusion of coupling, the fast head-tail description is somewhat more complicated but not particularly difficult and not essentially different from the uncoupled case. We do not discuss it in this paper.

In this essentially theoretical paper, extending the two particle model to include coupled motion, a few qualitatively striking observed phenomena, some already mentioned, will be considered and referred to by the following labels:

(i) The growth rates of the "visible modes", to be labeled 1 and 2, are observed to be different even though their betatron motions seem to be symmetrically related, plane polarized at ± 45 degrees, when viewed at a symmetry point of the lattice.

(ii) With chromaticities adjusted to make modes 1 and 2 stable, the "invisible modes", 3 and 4, though predicted to be unstable, seem to be harmless.

(iii) "Chromaticity sharing" is observed to occur. That is, in sufficiently coupled lattices, when one chromaticity is decreased and the other increased more or less equally, the betatron damping of visible modes tends to be unchanged.

(iv) The damping of the visible modes is much enhanced by the presence of strongly unequal horizontal and vertical chromaticities.

(v) With horizontal and vertical chromaticities strongly unequal but adjusted to keep one mode barely stable the other visible mode acquires damping as much as fifty times greater than the natural synchrotron radiation induced damping.

(vi) In spite of the fact that (proportional to current) head-tail anti-damping has to be comparable with natural synchrotron damping at quite low currents to account for observed instability there when the chromaticity is (even slightly) negative, the damping rate depends only weakly on beam current when the chromaticity is positive.

This paper emphasizes item (iii), showing it to be understandable as a natural generalization of the two particle head-tail model, and item (iv) which is due to decoherence accompanying the spread of particle momenta. We continue to accept the explanation of item (ii) as being accounted for by Landau damping, even arguing that Hamiltonian requirements cause some of this damping to be inherited by the otherwise-more-weakly-damped visible modes. This is used to account for item (vi). This leaves item (i) which, though it doesn't violate known laws, is not persuasively explained here, even heuristically. Also item (v) remains mysterious and capable only of phenomenological description.

The main theoretical result of this paper, which can be said to contain most of the other results as worked-out special cases, is the "near-symplectic" perturbation series of Eq. (9.36). The main results needed for the description of experimental results are Eqs. (5.15), (6.6), (12.11), and (14.24). All formulas intended for comparing with observations are enclosed in boxes for emphasis.

In proton accelerators, the bunches are typically long enough to have large chromatic betatron phase shifts along the bunch, which tends to ameliorate the wake field induced instability. This "wake field washout" due to destructive interference is enhanced in the large, unbalanced chromaticity region studied here, but the effect is shown to be too weak to account for observations with electron beams at CESR.

If an undamped oscillation mode is coupled to a strongly damped mode it seems that it should acquire some damping as a result. This effect, to be called "sympathetic damping", is analysed here, taking advantage of the near-Hamiltonian nature of the oscillations.† The damping rates acquired by the visible modes 1 and 2 from the Landau damping rate α_π (known only empirically) of modes 3 and 4 is calculated. This effect, though weak at low intensity, is strongly current dependent. It is used here to account for effect (vi).

Damping will be quantified by a quantity α with the sign chosen so that it is a "growth rate"; $\alpha > 0$ implies growth. This will remain true even if α is refered to as a "damping rate".

2. Two particle model, ("unperturbed")

The equation governing x motion of particle 1 is

$$\left(\frac{d^2}{dt^2} + \mu^2\right) x_1 + \frac{\mu\Sigma}{2} y_1 + \mu w_x \iota\, x_2 = R_{x1}, \tag{2.1}$$

where $Q_x = \mu/(2\pi) \overset{\text{e.g.}}{=} 10.77$ is the horizontal tune of the ring, $S \equiv \Sigma/(2\pi) \overset{\text{e.g.}}{=} 0.04$ is the strength of vertical to horizontal cross-plane coupling, expressed as a minimum tune separation, $\iota \overset{\text{e.g.}}{=} 0.02\,\text{A}$ is the single bunch current *magnitude*, and $\mu w_x \iota$ is the "effective", "common mode" horizontal focusing strength acting on particle 1 due to the wake field from particle 2. (The factor μ has been artificially included only for later convenience.) By implicitly defining ι to stand for the magnitude of the current the same form applies to either electrons or positrons. The polarity of the most naïve, pure electric image, no magnetic image, wake force would be attractive toward the closer wall and hence defocusing; in that case, as defined, w_x and w_y would be negative. The wake force will be discussed further below. The first term of (2.1) yields the dominant betatron motion, the remaining terms on the left hand side (though essentially perturbative themselves) will be regarded as being included in the "unperturbed" system. (The purpose of this "device" is to remove degeneracy in lowest order.) It is assumed the ring is run "close to the coupling resonance", meaning that the nominal vertical tune Q_y is also $\mu/(2\pi)$ but that the actual tunes will be split by an amount roughly equal to S. Other perturbative terms, to be explained later, are included in R_{x1}.

† For an RF resonator sympathetic damping often causes large reduction of the "loaded Q".

There are similar equations for y_1, x_2, and y_2. Defining matrices

$$\mathbf{P} = \mu \begin{pmatrix} \mu & \Sigma/2 & w_x\,\iota & 0 \\ \Sigma/2 & \mu & 0 & w_y\,\iota \\ w_x\,\iota & 0 & \mu & \Sigma/2 \\ 0 & w_y\,\iota & \Sigma/2 & \mu \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} R_{x_1} \\ R_{y_1} \\ R_{x_2} \\ R_{y_2} \end{pmatrix}, \tag{2.2}$$

the equations of motion are

$$\left(\mathbf{I}\frac{d^2}{dt^2} + \mathbf{P}\right)\mathbf{X} = \mathbf{R} \tag{2.3}$$

Using

$$\mathbf{E} = \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix}, \quad \mathbf{X} = \mathbf{G}\mathbf{E}, \quad \mathbf{E} = \mathbf{G}^{-1}\mathbf{X}, \quad \mathbf{R}^{\mathbf{g}} = \mathbf{G}^{-1}\mathbf{R}, \tag{2.4}$$

these equations can be transformed to

$$\left(\mathbf{I}\frac{d^2}{dt^2} + \mathbf{P}^{\mathbf{g}}\right)\mathbf{E} = \mathbf{R}^{\mathbf{g}}, \tag{2.5}$$

where $\mathbf{G}$ is to be chosen so that

$$\mathbf{P}^{\mathbf{g}} = \mathbf{G}^{-1}\mathbf{P}\mathbf{G} = \mathrm{diag}\left(\mu_1^2, \mu_2^2, \mu_3^2, \mu_4^2,\right). \tag{2.6}$$

The diagonal elements are "eigenmode" frequencies-squared.

A new set of unperturbed basis vectors $\mathbf{E}_{01} \equiv \mathbf{E}_{\sigma,\sigma}$, $\mathbf{E}_{02} \equiv \mathbf{E}_{\sigma,-\sigma}$, $\mathbf{E}_{03} \equiv \mathbf{E}_{\pi,\pi}$, and $\mathbf{E}_{04} \equiv \mathbf{E}_{\pi,-\pi}$, with indices corresponding to horizontal and vertical respectivel, can be defined by

$$\mathbf{E}_{01} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{E}_{02} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}, \quad \mathbf{E}_{03} = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \quad \mathbf{E}_{04} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}' \tag{2.7}$$

where the elements are x_1, y_1, x_2, y_2 components. In terms of these vectors, $\mathbf{X}$ is given by

$$\mathbf{X} = e_{01}\mathbf{E}_{01} + e_{02}\mathbf{E}_{02} + e_{03}\mathbf{E}_{03} + e_{04}\mathbf{E}_{04}. \tag{2.8}$$

For these to correspond to Eqs. (2.4), the transformation matrix $\mathbf{G}_0$ and its inverse are

$$\begin{aligned} \mathbf{G}_0 &= \left(\mathbf{E}_{01} \quad \mathbf{E}_{02} \quad \mathbf{E}_{03} \quad \mathbf{E}_{04}\right) \\ \mathbf{G}_0^{-1} = \mathbf{G}_0 &= \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}. \end{aligned} \tag{2.9}$$

As a consequence of the choice of basis vectors

$$\mathbf{G}_0^{-1} = \mathbf{G}_0 = \mathbf{G}_0^T. \tag{2.10}$$

As the unperturbed system has been defined, its eigenvectors $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$, and $\mathbf{E}_4$, are close but not identical to the basis vectors just introduced, and $\mathbf{P}^g$ is given by

$$\mathbf{P}^{\mathbf{g}} = \mu\,\mathrm{diag}\left(\mu + \Sigma/2 + \iota\overline{w},\quad \mu - \Sigma/2 + \iota\overline{w},\quad \mu + \Sigma/2 - \iota\overline{w},\quad \mu - \Sigma/2 - \iota\overline{w}\right) \tag{2.11}$$

where

$$\overline{w} \equiv \frac{w_x + w_y}{2}, \qquad \Delta w \equiv \frac{w_x - w_y}{2}. \tag{2.12}$$

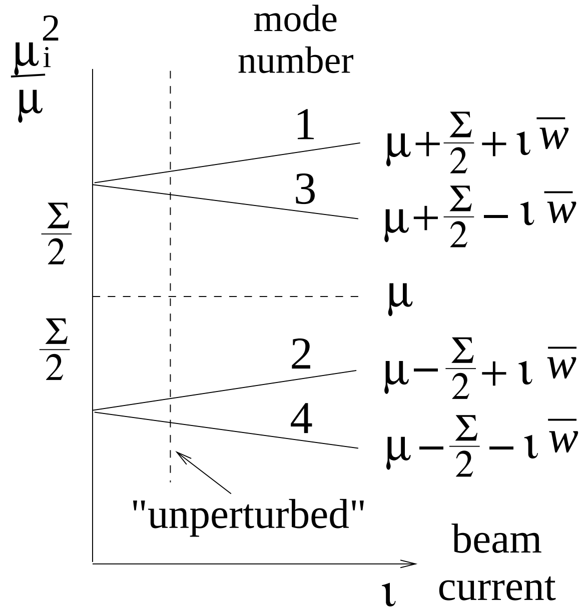

Figure 2.1: Labeling of the eigenmodes with schematic representation of dependence of eigenvalues on "beam current" ι. To break the degeneracy the unperturbed system is taken at slightly positive ι.

Even fixing the order of the eigenvectors and requiring $\Sigma > 0$, Eq. (2.6) does not determine $\mathbf{G}$ uniquely; two satisfactory choices are

$$\begin{aligned}
\mathbf{G} =& \mathbf{G}_0 + \Delta\mathbf{G} = \mathbf{G}_0 + \frac{\iota\Delta w}{\Sigma}\left\{ \begin{pmatrix} 0 & 0 & -1 & 1 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & -1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\}, \\
\mathbf{G}^{-1} =& \mathbf{G}_0 + \Delta\mathbf{G}_I = \mathbf{G}_0 + \frac{\iota\Delta w}{\Sigma}\left\{ \begin{pmatrix} 1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 & -1 \\ 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -1 \end{pmatrix} \right\}
\end{aligned} \tag{2.13}$$

and there are two others. Here, and throughout, terms of order ι^2 are neglected. The alternatives distinguish only among terms linear in ι. The ambiguity will not affect subsequent calculations since the current-dependent terms of (2.13) will turn out to be unimportant. In any case,

$$\mathbf{G}_0\,\Delta\mathbf{G}_I = -\Delta\mathbf{G}\,\mathbf{G}_0, \tag{2.14}$$

From now on all results will assume the choices listed second in Eq. (2.13) have been chosen.

There are numerous perturbing effects. Some are “linear” (meaning linear in $\mathbf{X}$) so they can be treated by matrices like $\mathbf{P}$. For example, a perturbing term $\mathbf{R}^{CH}$ that is due to chromaticity, can be written as

$$\mathbf{R}^{CH} = -\mathbf{P}^{CH}\mathbf{X}, \tag{2.15}$$

and then a corresponding transformed matrix by

$$\mathbf{P}^{CH,g} = \mathbf{G}^{-1}\mathbf{P}^{CH}\mathbf{G}. \tag{2.16}$$

One wake field force component will be treated similarly, but other terms are “constant” (independent of $\mathbf{X}$) or nonlinear. Expanding $\mathbf{R}$ into these terms, Eq. (2.5) becomes

$$\left(\mathbf{I}\frac{d^2}{dt^2} + \mathbf{P}^{\mathbf{g}}\right)\mathbf{E} = -\left(\mathbf{P}^{\mathrm{CH,g}} + \mathbf{P}^{\mathrm{W_1,g}}\right)\mathbf{E} + \mathbf{R}^{\mathrm{W_2,g}}, \tag{2.17}$$

where

$$\mathbf{P}^{\mathrm{W_1,g}} = \mathbf{G}^{-1}\mathbf{P}^{\mathrm{W_1}}\mathbf{G}, \quad \mathbf{R}^{\mathrm{W_2,g}} = \mathbf{G}^{-1}\mathbf{R}^{\mathrm{W_2}}. \tag{2.18}$$

Recall that terms on the right hand side of Eq. (2.17), though time-dependent, are “small”. However, one (inhomogeneous) term, $\mathbf{R}^{\mathrm{W_2,g}}$, has the special property of not vanishing in the small amplitude limit, though it vanishes in the small ι limit. This term will therefore be regarded as also belonging to the unperturbed model; this will be explained shortly.

Spelling out Eqs. (2.4) explicitly,

$$\begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} x_1 + y_1 + x_2 + y_2 \\ x_1 - y_1 + x_2 - y_2 \\ x_1 + y_1 - x_2 - y_2 \\ x_1 - y_1 - x_2 + y_2 \end{pmatrix} + \frac{\iota\Delta w}{\Sigma}\left\{\begin{pmatrix} x_1 + x_2 \\ -x_1 - x_2 \\ y_1 - y_2 \\ y_1 - y_2 \end{pmatrix}, \begin{pmatrix} -y_1 - y_2 \\ -y_1 - y_2 \\ y_1 - y_2 \\ y_1 - y_2 \end{pmatrix}\right\} \tag{2.19}$$

where the second terms need to be retained only when evaluating terms otherwise independent of ι. Similarly

$$\begin{pmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e_1+e_2+e_3+e_4 \\ e_1-e_2+e_3-e_4 \\ e_1+e_2-e_3-e_4 \\ e_1-e_2-e_3+e_4 \end{pmatrix} + \frac{\iota\Delta w}{\Sigma}\left\{ \begin{pmatrix} -e_3+e_4 \\ -e_1-e_2 \\ e_3-e_4 \\ -e_1-e_2 \end{pmatrix}, \begin{pmatrix} e_1-e_2-e_3+e_4 \\ 0 \\ e_1-e_2+e_3-e_4 \\ 0 \end{pmatrix} \right\}. \tag{2.20}$$

The leading approximations to certain sums of variables will be useful;

$$\begin{aligned} x_1+x_2 &= e_1+e_2, & x_1-x_2 &= e_3+e_4, \\ y_1+y_2 &= e_1-e_2, & y_1-y_2 &= e_3-e_4, \\ x_1+y_1 &= e_1+e_3, & x_1-y_1 &= e_2+e_4, \\ x_2+y_2 &= e_1-e_3, & x_2-y_2 &= e_2-e_4 \end{aligned} \tag{2.21}$$

3. Chromatic terms

To describe the effect of lattice chromaticities† the contributions to $\mathbf{R}$ are

$$\begin{aligned} R^{\rm CH}_{x1} &= \quad (D+d)\sin\mu_s t\; x_1 \\ R^{\rm CH}_{y1} &= \quad (D-d)\sin\mu_s t\; y_1 \\ R^{\rm CH}_{x2} &= -(D+d)\sin\mu_s t\; x_2 \\ R^{\rm CH}_{y2} &= -(D-d)\sin\mu_s t\; y_2 \end{aligned} \tag{3.1}$$

where chromaticities are described by

$$\begin{aligned} D+d = 4\pi\mu Q'_x\,(dp/p)_{\rm typ}\,, &\quad D-d = 4\pi\mu Q'_y\,(dp/p)_{\rm typ}\,, \\ D = 2\pi\mu\left(Q'_x+Q'_y\right)(dp/p)_{\rm typ}\,, &\quad d = 2\pi\mu\left(Q'_x-Q'_y\right)(dp/p)_{\rm typ}\,; \end{aligned} \tag{3.2}$$

except for a scale factor, D and d are just sums and differences of the two chromaticities. Combining them this way simplifies certain intermediate formulas. With $\delta p_1/p$ given by Eq. (1.1), Eqs. (3.1) come from the equations of free betatron motion as parametrically altered by energy oscillation; *e.g.*

$$\begin{aligned} &\frac{d^2x_1}{dt^2} + \left(2\pi Q_x - 2\pi Q'_x\,(dp/p)_{\rm typ}\sin\mu_s t\right)^2\; x_1 \\ &\approx \left(\frac{d^2}{dt^2}+\mu^2\right)x_1 - 4\pi\mu Q'_x\,(dp/p)_{\rm typ}\sin\mu_s t\; x_1 = 0. \end{aligned} \tag{3.3}$$

† The chromatic terms derived in this section are the only terms requiring the second terms in Eqs. (2.19) and (2.20).

Choosing the second alternative in Eq. (2.13), the matrix $\mathbf{P}^{\mathrm{CH,g}}$ is given by

$$\mathbf{P}^{\mathrm{CH,g}} = -\begin{pmatrix} 0 & 0 & D-\frac{D\iota\Delta w}{\Sigma/2} & d+\frac{D\iota\Delta w}{\Sigma/2} \\ 0 & 0 & d-\frac{D\iota\Delta w}{\Sigma/2} & D+\frac{D\iota\Delta w}{\Sigma/2} \\ D+\frac{D\iota\Delta w}{\Sigma/2} & d-\frac{D\iota\Delta w}{\Sigma/2} & 0 & 0 \\ d+\frac{D\iota\Delta w}{\Sigma/2} & D-\frac{D\iota\Delta w}{\Sigma/2} & 0 & 0 \end{pmatrix} \sin\mu_s t. \tag{3.4}$$

The perturbation vector, expressed in terms of the eigencoordinates, is

$$\mathbf{R}^{\mathrm{CH,g}} = \begin{pmatrix} D\,e_3 + d\,e_4 \\ d\,e_3 + D\,e_4 \\ D\,e_1 + d\,e_2 \\ d\,e_1 + D\,e_2 \end{pmatrix} \sin\mu_s t + \frac{\iota\Delta w D}{\Sigma/2}\begin{pmatrix} -e_3+e_4 \\ -e_3+e_4 \\ e_1-e_2 \\ e_1-e_2 \end{pmatrix} \sin\mu_s t. \tag{3.5}$$

Only the first term here is expected to be important; this can be confirmed once values of D and Δw are known.

4. Wake forces

We next include wake terms;†

$$\begin{aligned} R^{\mathrm{W}}_{x1} &= -\frac{\mu\,\iota\,w_x}{2\pi}\cos\mu_s t\ \left(x_2 + \eta\,\sin\mu_s t\right), \\ R^{\mathrm{W}}_{y1} &= -\frac{\mu\,\iota\,w_y}{2\pi}\cos\mu_s t\ y_2, \\ R^{\mathrm{W}}_{x2} &= \frac{\mu\,\iota\,w_x}{2\pi}\cos\mu_s t\ \left(x_1 - \eta\,\sin\mu_s t\right), \\ R^{\mathrm{W}}_{y2} &= \frac{\mu\,\iota\,w_y}{2\pi}\cos\mu_s t\ y_1, \end{aligned} \tag{4.1}$$

where for brevity a factor $(dp/p)_{\mathrm{typ}}$ has been subsumed into the definition of "dispersion" η, which is assumed to be purely horizontal. The coefficients of the terms proportional to dispersion η are very small compared to 1, but they are comparible in magnitude with the other transverse displacements and the importance of these terms may be amplified synchro-betatron resonance.

Eqs. (4.1) describe the "differential" wake-field effect. As shown in Fig. 4.1 the wake modulation would have a square-toothed sawtooth shape for a step function wake, but it has been smoothed by retaining only the first term in a Fourier series expansion. Though our head and tail particles are treated as points in the model, the fact that the true charge

† Part of the "wake field" force cancels a term artificially included in the "unperturbed equation" (2.3).

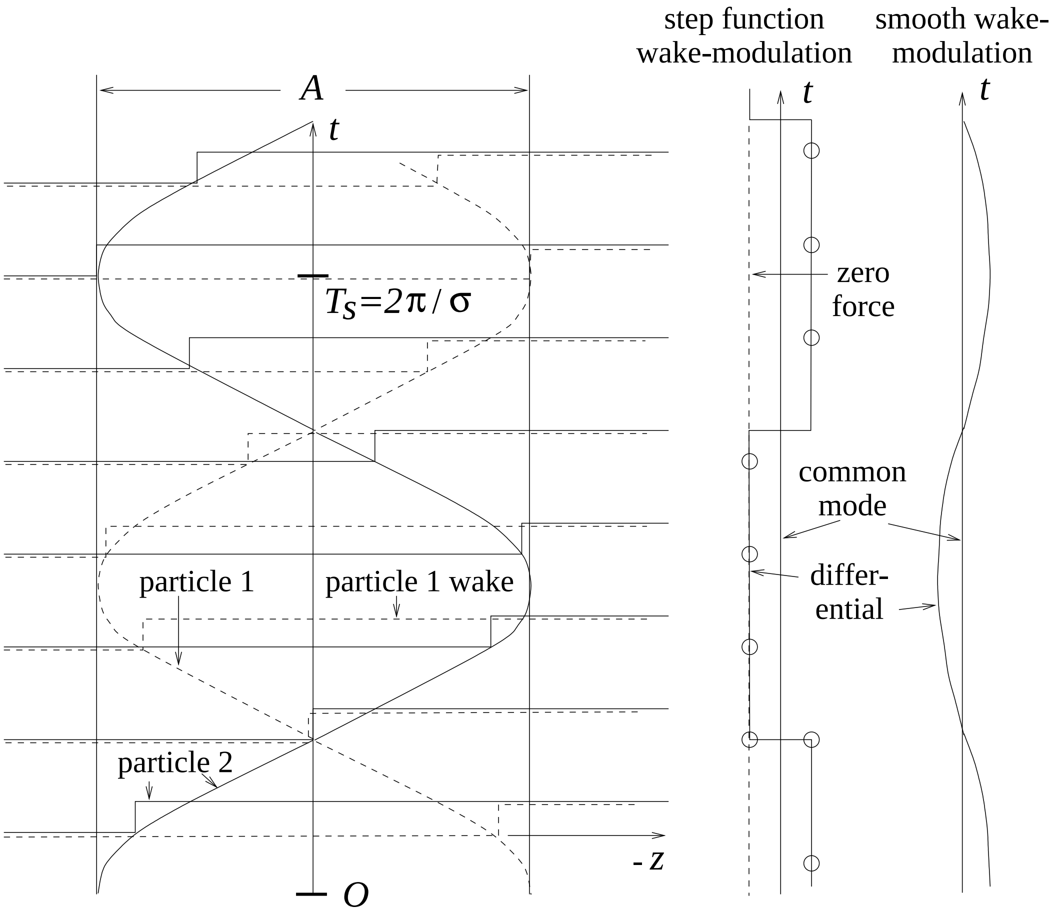

Figure 4.1: The factor modulating the wake force can be separated into a common-mode and a differential part. Less-rapid-than step-function wake variation tends to smooth the modulation factor.

distributions are spread out longitudinally as well as the detailed longitudinal dependence of the wake field will cause unknown multiplicative factors in this smoothing; these are assumed to be subsumed into the definitions of w_x and w_y, which are expected to be comparable in magnitude, but not necessarily equal. The strengths of the wake fields also have to be altered to account for momentum spread. This reduction factor can be estimated with semi-quantitative accuracy but, since our head and tail particles have definite momenta at any time, this reduction factor has to be incorporated in *ad hoc* fashion. Because of it there will be a spreads of head-tail betatron phase deviations of half-widths $\delta\chi_{x,y}$, that can be estimated as being equal to the previously introduced maximum head-tail phase deviations, $\delta\chi_{x,y} \approx \chi_{x,y}$.

For $\delta\chi << 1$ this spread has negligible effect and for $\delta\chi >> 1$, if our calculation were averaged over this phase, the wake would be effectively "washed out" by destructive interference. The function interpolating between these extremes should resemble a "single slit diffraction pattern" which we represent empirically by expressing the wake fields as

$$w_x = w_{x0} e^{-\frac{1}{2}\left(\frac{Q'_x}{Q'_{x0}}\right)^2}, \quad w_y = w_{y0} e^{-\frac{1}{2}\left(\frac{Q'_y}{Q'_{y0}}\right)^2}. \tag{4.2}$$

To be consistent with previous estimates $Q'_{0x} \approx Q'_{0y} \approx 60$. Even after these semi-empirical factors have been incorporated the differential wake-field term averages to zero *by definition* since the average effect has already been included in the unperturbed system as the final term on the right hand side of Eq. (2.1).

Dropping terms quadratic in ι, the matrix $\mathbf{P}^{\mathrm{W}_1,\mathrm{g}}$ defined in Eq. (2.17) is given by

$$\mathbf{P}^{\mathrm{W}_1,\mathrm{g}} = \frac{\mu\iota}{2\pi} \begin{pmatrix} 0 & 0 & -\overline{w} & -\Delta w \\ 0 & 0 & -\Delta w & -\overline{w} \\ \overline{w} & \Delta w & 0 & 0 \\ \Delta w & \overline{w} & 0 & 0 \end{pmatrix} \cos\mu_s t, \tag{4.3}$$

and the perturbation vector $\mathbf{R}^{\mathrm{W},\mathrm{g}}$ is

$$\mathbf{R}^{\mathrm{W}_1,\mathrm{g}} + \mathbf{R}^{\mathrm{W}_2,\mathrm{g}} = \frac{\mu\iota}{2\pi} \begin{pmatrix} \overline{w}\, e_3 + \Delta w\, e_4 \\ \Delta w\, e_3 + \overline{w}\, e_4 \\ -\overline{w}\, e_1 - \Delta w\, e_2 \\ -\Delta w\, e_1 - \overline{w}\, e_2 \end{pmatrix} \cos\mu_s t - \frac{\mu\iota w_x \eta}{2\pi} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \sin 2\mu_s t \tag{4.4}$$

As mentioned above, the term $\mathbf{R}^{\mathrm{W}_2,\mathrm{g}}$ will be regarded as belonging to the unperturbed model. When substituted into Eq. (2.5), this term causes response in the first two modes;

$$e_1 = -\frac{\mu\iota w_x \eta}{4\pi} \frac{\sin 2\mu_s t}{-(2\mu_s)^2 + \mu_1^2} \equiv e_{\mathrm{W}} \sin 2\mu_s t, \quad e_2 = -\frac{\mu\iota w_x \eta}{4\pi} \frac{\sin 2\mu_s t}{-(2\mu_s)^2 + \mu_2^2} \equiv e_{\mathrm{W}} \sin 2\mu_s t. \tag{4.5}$$

In principle, because of the different denominators, the amplitudes are different but, barring chance synchro-betatron resonance, the difference will be minor and we ignore it—especially since this term will eventually be judged unimportant. This motion will be superimposed on each of the "pure" eigenmotions so the force $\mathbf{R}^{\mathrm{W}_2,\mathrm{g}}$ will not otherwise be counted as contributing to the perturbation.

5. Iterative solution of the equations

We consider the eigenmodes one by one, starting with the first. Its unperturbed, or "zero'th order", time dependence, of amplitude a, is

$$e_1 = a\,\cos\mu_1 t + e_{\rm W}\sin 2\mu_s t, \quad e_2 = e_{\rm W}\sin 2\mu_s t, \quad e_3 = e_4 = 0. \tag{5.1}$$

This "steady-state" solution exhibits no damping or growth only because no terms causing them have been incorporated in the model. By substituting this solution into Eq. (3.5) an approximate time variation of the perturbation $\mathbf{R}^{\mathbf{g}}$ is obtained.

Using the trigonometric identities:

$$\begin{aligned} 2\cos A\cos B =&\ \cos(A+B)+\cos(A-B)\,, \\ 2\sin A\sin B =&\ -\cos(A+B)+\cos(A-B)\,, \\ 2\cos A\sin B =&\ \sin(A+B)-\sin(A-B)\,, \\ 2\sin A\cos B =&\ \sin(A+B)+\sin(A-B)\,, \end{aligned} \tag{5.2}$$

the mode 1 perturbations are

$$\begin{aligned} R_1^{(1),g} =&\ R_2^{(1),g} =\ 0 \\ R_3^{(1),g} =&\ \left(\frac{aD}{2}+\frac{aD\iota\Delta w}{\Sigma}\right)(\sin(\mu_1 t+\mu_s t)-\sin(\mu_1 t-\mu_s t)) \\ &+\frac{(D+d)\,e_{\rm W}}{2}(-\cos 3\mu_s t+\cos\mu_s t) \\ &-\frac{a\mu\iota\overline{w}}{4\pi}(\cos(\mu_1 t+\mu_s t)+\cos(\mu_1 t-\mu_s t))\,, \\ R_4^{(1),g} =&\ \left(\frac{ad}{2}+\frac{aD\iota\Delta w}{\Sigma}\right)(\sin(\mu_1 t+\mu_s t)-\sin(\mu_1 t-\mu_s t)) \\ &+\frac{(D+d)\,e_{\rm W}}{2}(-\cos 3\mu_s t+\cos\mu_s t) \\ &-\frac{a\mu\iota\Delta w}{4\pi}(\cos(\mu_1 t+\mu_s t)+\cos(\mu_1 t-\mu_s t))\,. \end{aligned} \tag{5.3}$$

These terms act like sinuisoidal "external drive" terms, to be substituted into Eq. (2.5) which, when solved, yields a "first" or "next" approximation to the motion. We are looking for the "steady-state" response, which is

$$
\begin{aligned}
e_1^{(1)} &= a\cos\mu_1 t + e_{\rm W}\sin 2\mu_s t, \\
e_2^{(1)} &= e_{\rm W}\sin 2\mu_s t, \\
e_3^{(1)} &= \left(\frac{aD}{2} + \frac{aD\iota\overline{w}}{\Sigma}\right)\left(\frac{\sin(\mu_1 t + \mu_s t)}{-(\mu_1+\mu_s)^2 + \mu_3^2} - \frac{\sin(\mu_1 t - \mu_s t)}{-(\mu_1-\mu_s)^2 + \mu_3^2}\right) \\
&\quad + \frac{(D+d)\,e_{\rm W}}{2}\left(\frac{-\cos 3\mu_s t}{-(3\mu_s)^2 + \mu_3^2} + \frac{\cos\mu_s t}{-(\mu_s)^2 + \mu_3^2}\right) \\
&\quad - \frac{a\mu\iota\overline{w}}{4\pi}\left(\frac{\cos(\mu_1 t + \mu_s t)}{-(\mu_1+\mu_s)^2 + \mu_3^2} + \frac{\cos(\mu_1 t - \mu_s t)}{-(\mu_1-\mu_s)^2 + \mu_3^2}\right), \qquad (5.4) \\
e_4^{(1)} &= \left(\frac{ad}{2} + \frac{aD\iota\overline{w}}{\Sigma}\right)\left(\frac{\sin(\mu_1 t + \mu_s t)}{-(\mu_1+\mu_s)^2 + \mu_4^2} - \frac{\sin(\mu_1 t - \mu_s t)}{-(\mu_1-\mu_s)^2 + \mu_4^2}\right) \\
&\quad + \frac{(D+d)\,e_{\rm W}}{2}\left(\frac{-\cos 3\mu_s t}{-(3\mu_s)^2 + \mu_4^2} + \frac{\cos\mu_s t}{-(\mu_s)^2 + \mu_4^2}\right) \\
&\quad - \frac{a\mu\iota\Delta w}{4\pi}\left(\frac{\cos(\mu_1 t + \mu_s t)}{-(\mu_1+\mu_s)^2 + \mu_4^2} + \frac{\cos(\mu_1 t - \mu_s t)}{-(\mu_1-\mu_s)^2 + \mu_4^2}\right).
\end{aligned}
$$

Obtaining this response is slightly subtle since, with no damping present, transient solutions (that depend on initial conditions) remain comparable indefinitely with the driven, steady state, terms. In practice the transients would decay due to inevitable damping. One analytic procedure to accomplish this is to solve the equations by Laplace transform and, before inverse transformation, to discard poles not due to (5.3). Alternatively, since there is no damping and the equations are linear, drive and response are in phase and to get the response it is only necessary to divide the terms of Eq. (5.3) by appropriate denominator factors. Those are the only frequencies present in the response. Furthermore, since R_{g1} has no fundamental component, the fundamental is unperturbed in this order of approximation; that is, its frequency is unshifted and its growth rate continues to vanish. In this approximation the only mode 2 response is at the doubled synchrotron frequency and there is mode 3 and mode 4 response at frequencies $\mu_1 \pm \mu_s$.

To better approximate the fundamental it is necessary to perform another iteration, first substituting Eqs. (5.4) into Eq. (3.5). This time, not only is R_1^g non-zero, it has a part oscillating at frequency μ_1. (One of the sidebands of a sideband-of-the-fundamental is the fundamental itself.) Before continuing with mode 1 analysis we digress into the nature of the next iteration.

The equation of motion is this approximation is

$$\frac{d^2 e_1}{dt^2} + \mu_1^2 e_1 = \gamma_1 \sin \mu_1 t \equiv F(t)\,, \tag{5.5}$$

where, later, $F(t)$ could stand for any time-dependent force. If we think of $e_1 = a \cos \mu_1 t$ as the displacement of a unity-mass point particle executing simple harmonic motion (unperturbed) because of the (restoring) force $-\mu_1^2 e_1$, its total energy is $\frac{1}{2}\dot{e}_e^2 + \frac{1}{2}\mu_1^2 e_1^2 = \frac{1}{2}\mu_1^2 a^2$. According to Eq. (5.5) the mass is also subject to an external force $\gamma_1 \sin \mu_1 t$ which does work and hence changes the total energy by an amount

$$\Delta E = \int_0^{2\pi/\mu_1} (\gamma_1 \sin \mu_1 t)\,(-a\mu_1 \sin \mu_1 t)\,dt = -\pi\gamma_1 a \tag{5.6}$$

during one betatron cycle. On the other hand, if a multiplicative amplitude decay factor $e^{\alpha_1 t}$ is to represent it, this same change in energy during one betatron cycle is given by

$$\Delta E = \left(\frac{1}{2}\mu_1^2 a^2\right)(\alpha_1)\left(\frac{2\pi}{\mu_1}\right) = 2\pi\alpha_1\mu_1 a^2. \tag{5.7}$$

It follows that

$$\alpha_1 = -\frac{1}{2a\mu_1}\gamma_1 = -\frac{1}{2a\mu_1}\frac{1}{\pi}\int_0^{2\pi} F(t')\sin(\mu_1 t')\,d(\mu_1 t')\,. \tag{5.8}$$

This result will be used in a form such that the integral averages away all terms except those proportional to $\sin \mu_1 t$.

Returning to mode 1 analysis, the only term contributing to the growth rate of mode 1 is the first term of Eq. (3.5), and only selected terms of e_3 and e_4, copied here from Eq. (5.4), contribute;

$$\begin{aligned}
R_1^{(1),g} &= (D\,e_3 + d\,e_4)\sin \mu_s t,\\
D\,e_3^{(1)} \sin \mu_s t &= -\frac{\mu\iota}{2\pi}\frac{D\overline{w}}{2}\left(\frac{\cos(\mu_1 t + \mu_s t)}{-(\mu_1+\mu_s)^2 + \mu_3^2} + \frac{\cos(\mu_1 t - \mu_s t)}{-(\mu_1-\mu_s)^2 + \mu_3^2}\right) a \sin \mu_s t,\\
d\,e_4^{(1)} \sin \mu_s t &= -\frac{\mu\iota}{2\pi}\frac{d\Delta w}{2}\left(\frac{\cos(\mu_1 t + \mu_s t)}{-(\mu_1+\mu_s)^2 + \mu_4^2} + \frac{\cos(\mu_1 t - \mu_s t)}{-(\mu_1-\mu_s)^2 + \mu_4^2}\right) a \sin \mu_s t.
\end{aligned} \tag{5.9}$$

With $A = (\mu_1+\mu_s)t$ and $\sin B = \sin \mu_s t$, the only one of Eqs. (5.2) yielding time dependence $\sin \mu_1 t$ is the third; that was the basis for retaining only the terms shown in Eq. (5.9). Note, furthermore, that even these terms average to zero except for their different denominator factors. (*e.g.* $\langle(\cos(\mu_1 t+\mu_s t)+\cos(\mu_1 t-\mu_s t))\sin\mu_s t\rangle = 2\langle\cos\mu_1 t\cos\mu_s t\sin\mu_s t\rangle = 0.$) For manipulating expressions like (5.9), the following identity is useful:

$$
\begin{aligned}
f_{ij} \equiv f\left(\mu_i, \Delta\xi_{ij}\right) &\equiv \frac{-1}{-\left(\mu_i+\mu_s\right)^2+\mu_j^2} + \frac{1}{-\left(\mu_i-\mu_s\right)^2+\mu_j^2} \\
&= \frac{-4\mu_i\mu_s}{\left(\mu_i^2-\mu_j^2+{\mu_s}^2\right)^2-\left(2\mu_i\mu_s\right)^2} = \frac{-\mu_i\mu_s}{\left(\Delta\xi_{ij}+{\mu_s}^2/2\right)^2-\left(\mu_i\mu_s\right)^2};
\end{aligned}
\tag{5.10}
$$

the quantity $\xi_{ij} = (\mu_i^2-\mu_j^3)/2$ can be related to "line half-separations" that can be read off from Fig. 2.1, as illustrated by the following examples (that assume $\iota << 1$):

$$
\begin{aligned}
f_{13} &= \frac{-\mu_1\mu_s}{\left(\mu\iota\overline{w}+{\mu_s}^2/2\right)^2-\left(\mu_1\mu_s\right)^2} \overset{\Sigma<<\mu_s}{\approx} \frac{1}{\mu}\,\frac{1}{\mu_s}, \\
f_{14} &= \frac{-\mu_1\mu_s}{\left(\mu\Sigma/2+\mu\iota\overline{w}+{\mu_s}^2/2\right)^2-\left(\mu_1\mu_s\right)^2} \overset{\Sigma<<\mu_s}{\approx} \frac{1}{\mu}\left(\frac{1}{\mu_s}+\frac{\left(\Sigma/2\right)^2}{{\mu_s}^2-\left(\Sigma/2\right)^2}\right), \\
f_{23} &= \frac{-\mu_2\mu_s}{\left(-\mu\Sigma/2+\mu\iota\overline{w}+{\mu_s}^2/2\right)^2-\left(\mu_2\mu_s\right)^2} \overset{\Sigma<<\mu_s}{\approx} \frac{1}{\mu}\left(\frac{1}{\mu_s}+\frac{\left(\Sigma/2\right)^2}{{\mu_s}^2-\left(\Sigma/2\right)^2}\right), \\
f_{24} &= \frac{-\mu_2\mu_s}{\left(\mu\iota\overline{w}+{\mu_s}^2/2\right)^2-\left(\mu_2\mu_s\right)^2} \overset{\Sigma<<\mu_s}{\approx} \frac{1}{\mu}\,\frac{1}{\mu_s}.
\end{aligned}
\tag{5.11}
$$

and similarly for other possibilities. Note that the final (crude) approximations are not indicative of resonance; the approximation fails to be valid well before the denominator can vanish and certainly therefore before the true resonance at $\mu_s = \Sigma$. The small additive corrections to f_{13} and f_{14} are only retained since below, in Eq. (5.8), they cause minor deviation from exact "chromaticity sharing". Note that

$$f_{ij} = -f_{ji}. \tag{5.12}$$

Continuing to simplify Eq. (5.9), we obtain for small ι,

$$R_1^{(1),g} \overset{\text{eff}}{=} -\frac{\mu\iota}{2\pi}\,\frac{1}{4}\left(D\overline{w}f_{13}+d\Delta w f_{14}\right)a\sin\mu_1 t, \tag{5.13}$$

where "eff" means terms that will not eventually contribute have been dropped. Using Eq. (5.8), α_1, which we now call the "head-tail growth rate" α_{HT}, is given by

$$\alpha_{HT} = \frac{1}{8\mu}\,\frac{\mu\iota}{2\pi}\,\left(D\overline{w}f_{13}+d\Delta w f_{14}\right) \tag{5.14}$$

Substitution from Eq. (4.1) yields *the main result*,

$$\boxed{\alpha_{HT} = \frac{\iota}{8\mu_s}\Big(\frac{dp}{p}\Big)_{\rm typ}\Big(Q'_x w_x + Q'_y w_y + \frac{(Q'_x - Q'_y)(w_x - w_y)}{2}\,\frac{(\Sigma/2)^2}{{\mu_s}^2 - (\Sigma/2)^2}\Big).} \tag{5.15}$$

The terms dropped from the right hand side of Eq. (5.9) would not have contributed to α_{HT}. The term proportional to $\cos\mu_1 t$ would have canceled in the averaging occuring in integral (5.9). This term "shifts the tune", an effect we will look at in section (7). The other dropped terms oscillate at "synchrotron side-band" frequencies. Extending the averaging time in Eq. (5.9) to the synchrotron period these terms also average to zero. Note also that the terms proportional to $e_{\rm W}$ do not contribute, which is the same result as is obtained if they are (incorrectly, it seems to me) treated purely as perturbative terms.

The last term of Eq. (5.14) diverges as $\Sigma/2 \to \mu_s$ which is presumably unphysical. This is discussed further in section (8). For almost all the data that has been taken this distinction is somewhat academic since $\Sigma/2 << \mu_s$ and $w_x \approx w_y$.

Analysis of mode 2 is very similar to that for mode 1 and the resulting growth rate is the same. We proceed then to mode 3.

6. Head-tail damping of π-modes

Later, in fitting to measured damping rates, a phenomenological contribution due to Landau damped will be added fo them, but the π-modes are also subject to head-tail damping or anti-damping. The formula will turn out to be dominated by the same terms as are α_1 and α_2 but, since that may not be obvious, small terms will be retained temporarily. In zero'th approximation,

$$e_1 = e_{\rm W}\sin 2\mu_s t, \quad e_2 = e_{\rm W}\sin 2\mu_s t, \quad e_3 = a\cos\mu_3 t, \quad e_4 = 0, \tag{6.1}$$

and the perturbation terms are:

$$\mathbf{R}^{(3),\mathbf{g}} = \begin{pmatrix} Da\cos\mu_3 t \\ da\cos\mu_3 t \\ (D+d)\,e_{\rm W}\sin 2\mu_s t \\ (D+d)\,e_{\rm W}\sin 2\mu_s t \end{pmatrix}\sin\mu_s t + \frac{\iota\Delta w D}{\Sigma/2}\begin{pmatrix} -a\cos\mu_3 t \\ -a\cos\mu_3 t \\ 0 \\ 0 \end{pmatrix}\sin\mu_s t, \tag{6.2}$$

$$\mathbf{R}^{\mathrm{W,g}} = \frac{\mu\iota}{2\pi}\begin{pmatrix} \overline{w} a\cos\mu_3 t \\ \Delta w a\cos\mu_3 t \\ -(\overline{w}+\Delta w)\sin 2\mu_s t \\ -(\Delta w+\overline{w})\sin 2\mu_s t \end{pmatrix}\cos\mu_s t. \tag{6.3}$$

Fourier expanding the third components, no term proportional to $\sin\mu_3 t$ is found so there is no contribution to the damping in this approximation. As in (5.9), the response to these terms will be substituted into

$$R_3^{(3),g} = (D\,e_1 + d\,e_2)\sin\mu_s t, \tag{6.4}$$

where some terms linear in e_1 and e_2, presumably small compared to the first term, have been dropped.

$$\begin{aligned} D\,e_1^{(3)}\sin\mu_s t &\overset{\mathrm{eff}}{=} \frac{D\,a\mu\iota\overline{w}}{4\pi}\left(\frac{\cos(\mu_3 t+\mu_s t)}{-(\mu_3+\mu_s)^2+\mu_1^2} + \frac{\cos(\mu_3 t-\mu_s t)}{-(\mu_3-\mu_s)^2+\mu_1^2}\right)\sin\mu_s t, \\ d\,e_2^{(3)}\sin\mu_s t &\overset{\mathrm{eff}}{=} \frac{d\,a\mu\iota\Delta w}{4\pi}\left(\frac{\cos(\mu_3 t+\mu_s t)}{-(\mu_3+\mu_s)^2+\mu_2^2} + \frac{\cos(\mu_3 t-\mu_s t)}{-(\mu_3-\mu_s)^2+\mu_2^2}\right)\sin\mu_s t. \end{aligned} \tag{6.5}$$

At this point it can be seen that no contributions to damping depending on either η survive and, except for sign, the growth rate is the same as for the previous modes;

$$\boxed{\alpha_{\pi,\mathrm{HT}} = -\alpha_{\sigma,\mathrm{HT}}.} \tag{6.6}$$

7. Tune variation at zero current

We distill from Eq. (5.4) only the leading "drive terms" needed to obtain mode 1 and mode 2 response in the next approximation;

$$\begin{aligned} De_3^{(1)}\sin\mu_s t &= \frac{aD^2}{2}\left(\frac{\sin(\mu_1 t+\mu_s t)}{-(\mu_1+\mu_s)^2+\mu_3^2} - \frac{\sin(\mu_1 t-\mu_s t)}{-(\mu_1-\mu_s)^2+\mu_3^2}\right)\sin\mu_s t \overset{\mathrm{eff}}{=} \frac{D^2}{4}g_{13}e_1^{(0)}, \\ de_4^{(1)}\sin\mu_s t &= \frac{ad^2}{2}\left(\frac{\sin(\mu_1 t+\mu_s t)}{-(\mu_1+\mu_s)^2+\mu_4^2} - \frac{\sin(\mu_1 t-\mu_s t)}{-(\mu_1-\mu_s)^2+\mu_4^2}\right)\sin\mu_s t \overset{\mathrm{eff}}{=} \frac{d^2}{4}g_{14}e_1^{(0)}, \\ de_3^{(2)}\sin\mu_s t &= \frac{ad^2}{2}\left(\frac{\sin(\mu_2 t+\mu_s t)}{-(\mu_2+\mu_s)^2+\mu_3^2} - \frac{\sin(\mu_2 t-\mu_s t)}{-(\mu_2-\mu_s)^2+\mu_3^2}\right)\sin\mu_s t \overset{\mathrm{eff}}{=} \frac{d^2}{4}g_{23}e_2^{(0)}, \\ De_4^{(2)}\sin\mu_s t &= \frac{aD^2}{2}\left(\frac{\sin(\mu_2 t+\mu_s t)}{-(\mu_2+\mu_s)^2+\mu_4^2} - \frac{\sin(\mu_2 t-\mu_s t)}{-(\mu_2-\mu_s)^2+\mu_4^2}\right)\sin\mu_s t \overset{\mathrm{eff}}{=} \frac{D^2}{4}g_{24}e_2^{(0)}, \end{aligned} \tag{7.1}$$

where, like the quantities f_{ij} defined in Eq. (5.10),

$$g_{ij} \equiv \frac{1}{-\left(\mu_i+\mu_s\right)^2+\mu_j^2}+\frac{1}{-\left(\mu_i-\mu_s\right)^2+\mu_j^2}=2\frac{\mu_j^2-\mu_i^2-{\mu_s}^2}{\left(\mu_j^2-\mu_i^2-{\mu_s}^2\right)^2-4\mu_i^2{\mu_s}^2}\overset{\Sigma<<\mu_s}{\approx}\frac{1}{2\mu^2}, \tag{7.2}$$

where the approximation assumes $|\iota| << 0$ and $\Sigma << \mu_s$. Having selected terms oscillating at the natural frequencies, the perturbing terms have been written in terms of amplitudes expressed in lower approximation. Similar expressions could be written giving the mode 1 and mode 2 components that accompany nominally pure mode 3 and mode 4 eigenmotions. The effect of these terms is to shift the tunes; the diagonal elements of matrix $\mathbf{P}^g$ become

$$\begin{aligned}
P^g_{11} &= \mu\left(\mu+\Sigma/2+\iota\overline{w}\right)-\frac{D^2}{8\mu}\,g_{13}-\frac{d^2}{8\mu}\,g_{14},\\
P^g_{22} &= \mu\left(\mu-\Sigma/2+\iota\overline{w}\right)-\frac{d^2}{8\mu}\,g_{23}-\frac{D^2}{8\mu}\,g_{24},\\
P^g_{33} &= \mu\left(\mu+\Sigma/2-\iota\overline{w}\right)-\frac{D^2}{8\mu}\,g_{31}-\frac{d^2}{8\mu}\,g_{32},\\
P^g_{44} &= \mu\left(\mu-\Sigma/2-\iota\overline{w}\right)-\frac{d^2}{8\mu}\,g_{41}-\frac{D^2}{8\mu}\,g_{42}.
\end{aligned} \tag{7.3}$$

These frequency shifts depend quadratically both on the chromaticities and on $(dp/p)_{\rm typ}$. According to Eq. (7.2) all the frequency shifts are approximately the same and hence are unimportant. In any case it would be inconsistent to use these to infer dependence of tune on amplitude of synchrotron oscillation, since the parameter Σ was implicitly introduced as being independent of $\hat{\delta}$. This is discussed further in section (13).

8. Reformulation in Hamiltonian terms

Of the qualitative observations listed in section (1), so far we have only succeeded in accounting for effect (iii) chromaticity sharing. Items (i,ii,iv,v) have not been accounted for. We have however alluded to a mysterious "Landau damping" of π-modes, to be symbolized by (a negative, current dependent quantity) $\alpha_\pi(\iota)$. We will show later that under certain conditions decoherence can also cause an appreciable rate (negative quantity) $\alpha_{\sigma,\mathrm{dec}}$. These are inherently multi-particle effects, but we wish to include their description into the otherwise-quantitative 2-particle picture. The purpose is to include all effects in a semi-empirical model whose parameters can be determined by matching to observations.

It is not entirely obvious how to input empirical damping terms consistently into a theory in which some damping is being calculated as output. This is especially true because of "sympathetic damping" in which an otherwise weakly-damped mode acquires damping by virtue of being coupled to a strongly-damped mode. Another complication is the $\mu_s \approx \Sigma$ resonance. To address these problems, in order to take advantage of sophisticated mathematical procedures analysing near-linear, near-Hamiltonian systems with time-varying but periodic coefficients, we re-formulate the problem in Hamiltonian form, in preparation for the perturbative treatment of the following section.

The equation of free motion Eq. (2.5), now expressed in eigencoordinates, is

$$\left(\mathbf{I}\frac{d^2}{dt^2} + \mathbf{Q}^{\mathbf{s},\mathbf{1}}\frac{d}{dt} + \mathbf{P}^0 + \mathbf{P}^{\mathbf{s},\mathbf{1}} + \mathbf{P}^{\mathbf{a},\mathbf{1}}\right)\mathbf{E} = 0, \tag{8.1}$$

where

$$\mathbf{P}^0 = \begin{pmatrix} \mu_1^2 & 0 & 0 & 0 \\ 0 & \mu_2^2 & 0 & 0 \\ 0 & 0 & \mu_3^2 & 0 \\ 0 & 0 & 0 & \mu_4^2 \end{pmatrix}, \tag{8.2}$$

$$\mathbf{P}^{s,1} = \begin{pmatrix} 0 & 0 & D & d \\ 0 & 0 & d & D \\ D & d & 0 & 0 \\ d & D & 0 & 0 \end{pmatrix} \frac{e^{i\mu_s t} - e^{-i\mu_s t}}{2i} \equiv \begin{pmatrix} \mathbf{0} & \mathcal{D}\mathrm{s} \\ \mathcal{D}\mathrm{s} & \mathbf{0} \end{pmatrix}, \tag{8.3}$$

$$\mathbf{P}^{a,1} = \frac{\mu\iota}{2\pi} \begin{pmatrix} 0 & 0 & -\overline{w} & -\Delta w \\ 0 & 0 & -\Delta w & -\overline{w} \\ \overline{w} & \Delta w & 0 & 0 \\ \Delta w & \overline{w} & 0 & 0 \end{pmatrix} \frac{e^{i\mu_s t} + e^{-i\mu_s t}}{2} \equiv \begin{pmatrix} \mathbf{0} & -\mathcal{W}\mathrm{c} \\ \mathcal{W}\mathrm{c} & \mathbf{0} \end{pmatrix}, \tag{8.4}$$

where s and c are abbreviations for $\sin\mu_s t$ and $\cos\mu_s t$, where the superscripts s and a differentiate symmetric and anti-symmetric matrices, where 1 indicates a term one "order of smallness" less than a term with superscript 0, and where the superscripts g have now been dropped even though the equation is expressed in eigencoordinates. For algebraic convenience the time-varying factors $\sin\mu_s t$ and $\cos\mu_s t$ will eventually be expanded in complex exponential form and the perturbation matrices have been abreviated to take advantage of their zero elements. Only the leading chromatic terms from Eq. (3.4) and wake terms from Eq. (4.3) have been retained. The remaining term in Eq. (8.1) is

$$\mathbf{Q}^1 = 2\,\mathrm{diag}\left(\alpha_{\mu_s} = 0, \alpha_{\mu_s} = 0, \alpha_\pi, \alpha_\pi\right) \tag{8.5}$$

which contains the empirically-added damping.

All the terms in Eq. (8.1) except $\mathbf{P^{a,1}}$ and $\mathbf{Q^1}$ are derivable from a Hamiltonian, the sum of "kinetic" plus "potential" plus "coupling" energies, namely

$$H(\mathbf{z}) = \frac{1}{2}\sum_i p_i^2 + \frac{1}{2}\sum_i \mu_i^2 e_i^2 + D\left(e_1e_3 + e_2e_4\right) + d\left(e_1e_4 + e_2e_3\right), \tag{8.6}$$

where all particle "masses" are 1, and $\mathbf{e} \equiv (e_1, e_2, e_3, e_4)^T$, and

$$\mathbf{p} = \dot{\mathbf{e}}, \quad \text{and} \quad \mathbf{z} = \begin{pmatrix} \mathbf{e} \\ \mathbf{p} \end{pmatrix}. \tag{8.7}$$

In these terms the 8 unperturbed equations of motion, now written in first-order or Hamiltonian, matrix form, are

$$\frac{d\mathbf{z}}{dt} = \mathbf{C}\mathbf{z}, \quad \text{where} \quad \mathbf{C} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{P}^0 & \mathbf{0} \end{pmatrix}. \tag{8.8}$$

They can be also be written

$$\mathbf{C} = \widetilde{\mathbf{J}}^{-1}\mathbf{H}_0, \quad \text{where} \quad \widetilde{\mathbf{J}} \equiv \begin{pmatrix} \mathbf{0} & -\mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix}, \quad \text{and} \quad \mathbf{H}_0 \equiv \begin{pmatrix} \mathbf{P}^0 & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}. \tag{8.9}$$

$\mathbf{C}$ will be referred to as the "unperturbed Hamiltonian matrix" even though the term might seem to refer more naturally to $\mathbf{H}_0$.

It can be seen that symmetric perturbations added to $\mathbf{P}^0$, such as $\mathbf{P^{s,1}}$, leave the system Hamiltonian, while anti-symmetric contributions such as $\mathbf{P^{a,1}}$ do not and, because of that, may lead to damping or anti-damping. It is possible for "velocity-dependent potential" terms to be Hamiltonian as well, but only if their matrix $\mathbf{Q}$ is anti-symmetric. Since $\mathbf{Q}^1$ defined in Eq. (8.5) is symmetric it causes "resistive" damping.

So far all quantities appearing in the equations are real. But there are well-known formalisms, such as the impedance description of AC circuits in which the coefficients of linear equations can be complex. In that description damping is reflected by the "frequencies" μ_i acquiring imaginary parts. The present discussion can be generalized to encompass this possibility by altering what has so far been taken as the requirement that the Hamiltonian matrix be symmetric, to the requirement that it be Hermitean, which correctly treats terms of both $\mathbf{P}$ and $\mathbf{Q}$ type. Even though this possibility of introducing complex coefficients will not be exploited here, complex quantities will inevitably intrude when eigenvalues and eigenvectors are sought. In anticipation of this we define the "scalar product" of vectors $\mathbf{x}$ and $\mathbf{y}$ by†

$$(\mathbf{x}, \mathbf{y}) \equiv \sum_i y_i^* x_i \equiv \mathbf{y}^{*T}\mathbf{x} \equiv \mathbf{y}^\dagger \mathbf{x}, \tag{8.10}$$

where $*$ indicates complex conjugation, T indicates matrix transposition, and $\dagger$ indicates "Hermitean conjugation. With the Hermitean conjugate of a matrix $\mathbf{A}$ being defined similarly, we have

$$\mathbf{A}^\dagger \equiv \mathbf{A}^{*T}, \quad \text{and} \quad (\mathbf{A}\mathbf{x}, \mathbf{y}) = \left(\mathbf{x}, \mathbf{A}^\dagger \mathbf{y}\right). \tag{8.11}$$

After this digression into notation we return to the perturbed equations of motion (replacing $\mathbf{z}$ by $\mathbf{x}$);

$$\frac{d\mathbf{x}}{dt} = (\mathbf{C} + \mathbf{B})\,\mathbf{x}, \quad \text{where} \quad \mathbf{B} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^{\mathbf{s},\mathbf{1}} - \mathbf{P}^{\mathbf{a},\mathbf{1}} & -\mathbf{Q}^1 \end{pmatrix}; \tag{8.12}$$

$\mathbf{B}$ will be referred to as the "perturbation matrix". These equations can be said to be "near-Hamiltonian" as the non-Hamiltonian terms have superscript 1, indicating smallness.

For solving Hamiltonian linear equations with periodic coefficients such as Eq. (8.1) there is a formalism due to Floquet and Lyapunov—the traditional β-function formalism is the best-known example of this in accelerator physics. For near-Hamiltonian systems there is a perturbation theory described by Yakubovich and Starzhinski[3], which we now describe.

---

† The unconventional ordering of the symbols in the definition of scalar product is the result of following notation used by Yakubovich and Starzhinski[3] so that formulas can be copied unchanged from that source.

They begin with eigenvectors of the matrix $\mathbf{C}$ which therefore satisfy

$$\mathbf{Cc} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{P}^0 & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} = \lambda \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} = \lambda \mathbf{c}. \tag{8.13}$$

These require $\mathbf{b} = \lambda \mathbf{a}$ and

$$\begin{pmatrix} \mu_1^2 & 0 & 0 & 0 \\ 0 & \mu_2^2 & 0 & 0 \\ 0 & 0 & \mu_3^2 & 0 \\ 0 & 0 & 0 & \mu_4^2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix} = -\lambda^2 \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{pmatrix}. \tag{8.14}$$

As a result, the eigenvalues are

$$\begin{aligned} &\lambda_{-1} = -i\mu_1, \ \lambda_{-2} = -i\mu_2, \ \lambda_{-3} = -i\mu_3, \ \lambda_{-4} = -i\mu_4, \\ &\lambda_1 = i\mu_1, \qquad \lambda_2 = i\mu_2, \qquad \lambda_3 = i\mu_3, \qquad \lambda_4 = i\mu_4, \end{aligned} \tag{8.15}$$

and the corresponding eigenvectors are

$$\mathbf{c}_{\pm 1} = \begin{pmatrix} \mathbf{a}_1 \\ \pm i\mu_1 \mathbf{a}_1 \end{pmatrix}, \ \mathbf{c}_{\pm 2} = \begin{pmatrix} \mathbf{a}_2 \\ \pm i\mu_2 \mathbf{a}_2 \end{pmatrix}, \ \mathbf{c}_{\pm 3} = \begin{pmatrix} \mathbf{a}_3 \\ \pm i\mu_3 \mathbf{a}_3 \end{pmatrix}, \ \mathbf{c}_{\pm 4} = \begin{pmatrix} \mathbf{a}_4 \\ \pm i\mu_4 \mathbf{a}_4 \end{pmatrix}, \tag{8.16}$$

where

$$\mathbf{a}_1 = \frac{1}{\sqrt{2\mu_1}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{a}_2 = \frac{1}{\sqrt{2\mu_2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ \mathbf{a}_3 = \frac{1}{\sqrt{2\mu_3}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ \mathbf{a}_4 = \frac{1}{\sqrt{2\mu_4}} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{8.17}$$

The normalizing coefficients will be justified shortly. Later it will be useful to express these also as

$$\mathbf{a}_j = \begin{pmatrix} \mathbf{a}_j^{(u)} \\ \mathbf{a}_j^{(l)} \end{pmatrix}. \tag{8.18}$$

It is also necessary to introduce the "adjoint" equation

$$\mathbf{C}^\dagger \mathbf{d}_{\pm i} = \lambda_{\pm i}^* \mathbf{d}_{\pm i}; \tag{8.19}$$

its eigenvectors are

$$\mathbf{d}_{\pm h} = \begin{pmatrix} \mu_h \mathbf{a}_h \\ \pm i \mathbf{a}_h \end{pmatrix}, \tag{8.20}$$

where the labeling is such that $\mathbf{c}_j$ and $\mathbf{d}_h$ vectors satisfy mutual orthogonality relations

$$(\mathbf{c}_j, \mathbf{d}_h) = \mathbf{d}_h^\dagger \mathbf{c}_j = \delta_{jh}. \tag{8.21}$$

As the coefficients have been chosen, the normalizing symplectic forms[†]

$$\gamma_{jh} \equiv i\left(\widetilde{\mathbf{J}}\mathbf{c}_j, \mathbf{c}_h\right) = i\left(\begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix}\begin{pmatrix} \mathbf{a_j} \\ \pm i\mu_j \mathbf{a_j} \end{pmatrix}, \begin{pmatrix} \mathbf{a_h} \\ \pm i\mu_h \mathbf{a_h} \end{pmatrix}\right) \tag{8.22}$$

satisfy

$$\gamma_{jh} = \pm\delta_{jh} = \begin{cases} -1 & \text{for } j = h < 0 \\ 1 & \text{for } j = h > 0 \\ 0 & \text{for } j \neq h \end{cases} . \tag{8.23}$$

Here and in formulas to follow, the upper of the $\pm$ or $\mp$ signs is to go with positive values of indices such as j or h.

This perturbation theory will be reminiscent of the perturbation theory of quantum mechanics. A mnemonic device is to think of an index such as j as standing for "principle quantum number" $|j|$ of a state and $m = \text{sign}(j)$ as its "magnetic quantum number".

To take advantage of the periodicity of $\mathbf{B}$, the standard "Floquet procedure" is to seek solutions of Eq. (8.12) in the form

$$\mathbf{z}(t) = e^{\alpha t}\mathbf{u}(t), \quad \text{where} \quad \mathbf{u}(t+T) = \mathbf{u}(t), \tag{8.24}$$

and the period T is given by $T = 2\pi/\mu_s$.

A "characteristic exponent" α of the equation $d\mathbf{z}/dt = \mathbf{C}\mathbf{z}$ (which is the unperturbed equation) is defined to be a complex number for which a T-periodic function $\mathbf{u}(t)$ can be found with the property that $e^{\alpha t}\mathbf{u}(t)$ is a solution. Since the function

$$\mathbf{z}(t) = e^{\alpha t}\mathbf{c}_{\pm i} = e^{\pm i\mu_i t}\mathbf{c}_{\pm i}, \tag{8.25}$$

satisfies $d\mathbf{z}/dt = \mathbf{C}\mathbf{z}$, and because the requirement concerns only periodicity, all constants of the form

$$\lambda = \alpha + \frac{2\pi i}{T}m, \quad m = 0, \pm1, \pm2, \dots. \tag{8.26}$$

are also valid characteristic exponents.

It can happen that two or more of the unperturbed mode tunes $\pm\mu_j$ are related as in Eq. (8.26) for some value of the integer m. Verbally this would be expressed by the

[†] In more modern discussions of symplectic geometry, for example p. 219, Arnold, *Mathematical Methods of Classical Mechanics*, eigenbasis vectors are self-orthogonal. The difference here is due to the complex conjugation in definition (8.10). Arnold defines the symplectic form without the complex conjugation, with the result that eigenbasis vectors are "orthogonal" to themselves and all other basis vectors except their "companion" vector having complex conjugate eigenvalue. *e.g.* $\mathbf{c}_{+1}$ and $\mathbf{c}_{-1}$ are companions in this sense.

statement that there is a "sum resonance" or a "difference resonance", meaning that an integral multiple of frequency μ_s matches a sum or difference of the normal mode tunes. When this happens it leads to algebraic degeneracy. (Had the cross plane coupling Σ and the common mode wakefield not been included in the definition of the unperturbed model, we would now be facing an eight-fold degeneracy of this sort.) The following analysis can be adapted to the degenerate situation, but it become much more complicated—rather than ending up with explicit formulas for the damping rates one obtains a (much simpler) eigenvalue problem.

9. Nonresonant perturbation theory

For now we exclude the possibility of resonance and assume there are no integer values of m for which two different normal mode tunes are related as in Eq. (8.26). Using perturbation theory we seek the normal mode of the perturbed equation $d\mathbf{x}/dt = (\mathbf{C}+\mathbf{B})\mathbf{x}$ that is "close to" a particular, non-degenerate, unperturbed, normal mode which, without loss of generality, can be assigned label 1;

$$\mathbf{x} = e^{\alpha_0 t}\mathbf{c_1}; \tag{9.1}$$

the symbol for its characteristic exponent is $\alpha_0 t$ to indicate "zero'th" approximation.† The substitution

$$\mathbf{x} = e^{\alpha_0 t}\mathbf{y} \tag{9.2}$$

transforms Eq. (8.9) to "homogeneous" equation

$$\frac{d\mathbf{y}}{dt} - (\mathbf{C} - \alpha_0 \mathbf{1})\,\mathbf{y} = 0, \tag{9.3}$$

which is satisfied by the time-independent vector

$$\mathbf{y}_0\,(t) = \mathbf{c}_1. \tag{9.4}$$

† If it really is the first mode that is being analysed then $\alpha_0 = i\mu_1$ or, since it will be convenient to have symbol μ_1 stand for the perturbed tune of mode 1 and $\mu_{1,0}$ its unperturbed value, then $\alpha_0 \equiv i\mu_{1,0}$. The notation in this section would probably be made clearer and more consistent by replacing α_0 by $i\mu_{1,0}$ throughout.

The equation "adjoint" to Eq. (9.3),

$$\frac{d\mathbf{z}}{dt} + \left(\mathbf{C}^{\dagger} - \alpha_0^{*}\mathbf{1}\right)\mathbf{z} = 0, \tag{9.5}$$

is satisfied by

$$\mathbf{z}_0\left(t\right) = \mathbf{d}_1, \tag{9.6}$$

defined in Eq. (8.20). Consider the "inhomogeneous" equation

$$\frac{d\mathbf{y}}{dt} - \left(\mathbf{C} - \alpha_0\mathbf{1}\right)\mathbf{y} = \mathbf{u}\left(t\right). \tag{9.7}$$

Though $\mathbf{u}(t)$ as it appears in (9.7) is time dependent, it is assumed to be T-periodic and in practice it will be a (short) Fourier sum of terms. For reasons to be explained immediately, we introduce the "time-averaged scalar product" of any two functions $\mathbf{u}(t)$ and $\mathbf{v}(t)$,

$$((\mathbf{u}, \mathbf{v})) \equiv \frac{1}{T}\int_0^T \left(\mathbf{u}\left(t\right), \mathbf{v}\left(t\right)\right) dt. \tag{9.8}$$

This form of averaging will figure prominently in the sequel. (This may cause the unjustified impression that the procedure is only heuristic, but in fact it is rigorous for the sinuisoidal perturbations that enter. As with all transform methods, the strategy here is to replace differential equations by algebraic equations, and the purpose of time-averaging is to extract Fourier coefficients. For linear functions this yields exact coefficients, for nonlinear functions it extracts "leading" terms in a Fourier series that may or may not converge.) Two complications inevitably arise in seeking a periodic solution $\mathbf{y}(t)$ of an *inhomogeneous* equation such as Eq. (9.7):

(i) A periodic solution of (9.7) exists if and only if $\mathbf{u}(t)$ is "orthogonal" to solution $\mathbf{d}_1$ of adjoint Eq. (9.3);

$$((\mathbf{u}, \mathbf{d}_1)) = 0. \tag{9.9}$$

We must therefore require that $\mathbf{u}$ satisfy (9.9).

(ii) Adding any solution $\mathbf{y}^{(1)}(t)$ of (9.3) to some solution $\mathbf{y}^{(0)}(t)$ of (9.7) yields another solution of (9.7). To obtain a unique solution requires the specification of as many extra conditions as there are independent solutions of (9.3). To make the solution of Eq. (9.7) unique we choose to take

$$\mathbf{y} = \mathbf{y}^{(0)} - \left(\left(\mathbf{y}^{(0)}, \mathbf{d}_1\right)\right)\mathbf{c}_1, \tag{9.10}$$

which, because of normalization (8.21), is equivalent to requiring

$$((\mathbf{y}, \mathbf{d}_1)) = 0. \tag{9.11}$$

If a function $\mathbf{u}(t)$ is a tentative candidate for appearing on the right hand side of Eq. (9.7), but does not satisfy Eq. (9.9), it can be replaced by $\mathcal{P}\mathbf{u}(t)$ where $\mathbf{u}$ has been "operated on" by an operator $\mathcal{P}$, defined by

$$\mathcal{P}\mathbf{u}\,(t) = \mathbf{u}\,(t) - ((\mathbf{c}_1, \mathbf{d}_1))\ \mathbf{c}_1, \tag{9.12}$$

which can therefore be written

$$\mathcal{P} = \mathbf{1} - \mathbf{c}_1\,((\,\cdot, \mathbf{d}_1))\,. \tag{9.13}$$

This causes $\mathcal{P}$ to have the property that

$$\mathcal{P}\mathbf{c}_1 = 0. \tag{9.14}$$

If $\mathbf{u}$ *is time-independent* then $((\mathbf{u}, \mathbf{d}_1)) = (\mathbf{u}, \mathbf{d}_1)$ and

$$\mathcal{P}\mathbf{u} = (\mathbf{1} - \mathbf{c}_1\,(\,\cdot, \mathbf{d}_1))\,\mathbf{u} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 & \frac{i}{2\mu_{1,0}} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{-i\mu_{1,0}}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{u}. \tag{9.15}$$

When operating on a function $\mathbf{u}$ that time-averages to zero, $\mathcal{P} = \mathbf{1}$. Returning to Eq. (9.7), to assure that the inhomogeneous term appearing on the right hand side satisfies Eq. (9.9), we write that term as $\mathcal{P}\mathbf{u}$ where $\mathbf{u}$ can therefore be arbitrary;

$$\frac{d\mathbf{y}}{dt} - (\mathbf{C} - \alpha_0\mathbf{1})\,\mathbf{y} = \mathcal{P}\mathbf{u}\,(t)\,. \tag{9.16}$$

The solution to this equation, made unique by condition (9.11), can be said to be the result of "operating on" $\mathcal{P}\mathbf{u}(t)$ with some linear operator $\mathcal{S}$; that is

$$\mathbf{y}\,(t) = \mathcal{S}\mathcal{P}\mathbf{u}\,(t)\,. \tag{9.17}$$

By definition then, the operator $\mathcal{S}$ satisfies

$$((\mathcal{SP}\mathbf{u}, \mathbf{d}_1)) = 0, \tag{9.18}$$

for arbitrary $\mathbf{u}$.†¶ A tentative, arbitrary solution $\mathbf{y}_0$ of Eq. (9.16) can be made unique by applying Eq. (9.10);

$$\mathcal{SP}\mathbf{u} = \mathbf{y}_0 - ((\mathbf{y}_0, \mathbf{d}_1))\, \mathbf{c}_1. \tag{9.20}$$

We now seek the coefficient α (close to α_0) in a solution to Eq. (8.12) of the form

$$\mathbf{x} = e^{\alpha t}\mathbf{y}, \quad \text{where} \quad \mathbf{y}\,(t+T) = \mathbf{y}\,(t)\,. \tag{9.21}$$

Substitution into (8.12) yields

$$\frac{d\mathbf{y}}{dt} - (\mathbf{C} - \alpha_0 \mathbf{1})\,\mathbf{y} = [(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}]\,\mathbf{y}. \tag{9.22}$$

As stated above, this can only be valid if the right hand side satisfies

$$(([(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}]\,\mathbf{y}, \mathbf{d}_1)) = 0. \tag{9.23}$$

As a result, it is legitimate to replace Eq. (9.22) by

$$\frac{d\mathbf{y}}{dt} - (\mathbf{C} - \alpha_0 \mathbf{1})\,\mathbf{y} = \mathcal{P}\,[(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}]\,\mathbf{y}. \tag{9.24}$$

The function $\mathbf{y}(t)$ may or may not satisfy condition (9.11), but in any case there is a related function

$$\mathbf{y}^{(0)} = \mathbf{y} - \zeta\, \mathbf{c}_1, \tag{9.25}$$

that satisfies both Eq. (9.24) and condition (9.11). By the definition of $\mathcal{S}$ this function is also given by

$$\mathbf{y}^{(0)} = \mathcal{SP}\,[(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}]\,\mathbf{y}. \tag{9.26}$$

---

† Were it valid to assume $\frac{d\mathbf{y}}{dt} = \mathbf{\Omega y}$ for some constant matrix or linear operator $\mathbf{\Omega}$ then $\mathcal{S}$ could be written as

$$\mathcal{S} = (\mathbf{\Omega} - \mathbf{C} + \alpha_0 \mathbf{1})^{-1} \tag{9.19}$$

plus an appropriate solution of the homogeneous equation. In practice it will typically not be possible to write $\mathcal{S}$ in closed form; rather it will be necessary to evaluate $\mathcal{SP}\mathbf{u}(t)$ for each particular value of $\mathbf{u}(t)$ that arises. To avoid undefined expressions $\mathcal{S}$ must always be "preceeded" by $\mathcal{P}$, as in $\mathcal{SP}$.

¶ In the earlier treatment (section (5)) the need for operator $\mathcal{P}$ was avoided by interpreting "secular" terms as sources of tune shifts, as in section (7). The special condition (9.11) on $\mathcal{S}$ was satisfied, below Eq. (5.4), by keeping only "driven" terms, thereby (artificially) suppressing transient terms. This simpler procedure was possible because velocity-proportional terms and complex exponentials had not yet entered.

In other words, $\mathbf{y}$ can be written as

$$\mathbf{y} = \mathcal{SP}\left[(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}\right]\mathbf{y} + \zeta\,\mathbf{c}_1 \tag{9.27}$$

and hence as

$$\mathbf{y}\,(t) = \left(\mathbf{1} - \mathcal{SP}\left[(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}\right]\right)^{-1}\zeta\,\mathbf{c}_1. \tag{9.28}$$

By construction this solution $\mathbf{y}(t)$ satisfies Eq. (9.23) which, assuming $\zeta \neq 0$, therefore implies

$$\left(\left([(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}]\left(\mathbf{1} - \mathcal{SP}\left[(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}\right]\right)^{-1}\mathbf{c}_1, \mathbf{d}_1\right)\right) \equiv \phi_{11} = 0; \tag{9.29}$$

The vanishing of this "matrix element" ϕ_{11} yields an implicit formula for the eigenvalue α being sought.

To proceed further it is necessary to take advantage of the smallness of $\mathbf{B}$, sorting terms by "orders of smallness" equal to the number of factors of $\mathbf{B}$. Introducing the temporary abbreviations

$$\mathcal{Q} = (\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}, \quad \text{and} \quad \mathcal{N} = \mathcal{SPQ}, \tag{9.30}$$

both being quantities of "first order of smallness" the middle factor in Eq. (9.29) can be transformed using

$$(\mathbf{1} - \mathcal{N})^{-1} = \mathbf{1} + \mathcal{N}\,(\mathbf{1} - \mathcal{N})^{-1}, \tag{9.31}$$

to yield

$$\phi_{11} = ((\mathcal{Q}\mathbf{c}_1, \mathbf{d}_1)) + \left(\left(\mathcal{QN}\,(\mathbf{1} - \mathcal{N})^{-1}\mathbf{c}_1, \mathbf{d}_1\right)\right), \tag{9.32}$$

where the second term is of "second order of smallness". In the second term, because of Eq. (9.17), it is legitimate to keep only the $\mathbf{B}$ part of the factor $\mathcal{Q}$

$$\left(\left(\mathcal{QN}\,(\mathbf{1} - \mathcal{N})^{-1}\mathbf{c}_1, \mathbf{d}_1\right)\right) = \left(\left(\mathbf{B}\mathcal{N}\,(\mathbf{1} - \mathcal{N})^{-1}\mathbf{c}_1, \mathbf{d}_1\right)\right) = \left(\left(\mathbf{B}\,(\mathbf{1} - \mathcal{N})^{-1}\mathcal{N}\mathbf{c}_1, \mathbf{d}_1\right)\right), \tag{9.33}$$

where the latter step is allowed because $\mathcal{N}$ commutes with $(\mathbf{1}-\mathcal{N})^{-1}$. Further simplification results from using

$$\mathcal{N}\mathbf{c}_1 = \mathcal{SP}\mathbf{B}\mathbf{c}_1, \tag{9.34}$$

which follows from Eq. (9.14). Collecting results we have

$$\phi_{11} = (\alpha_0 - \alpha)\,((\mathbf{c}_1, \mathbf{d}_1)) + ((\mathbf{B}\mathbf{c}_1, \mathbf{d}_1)) + \left(\left(\mathbf{B}\left(\mathbf{1} - \mathcal{SP}\left[(\alpha_0 - \alpha)\,\mathbf{1} + \mathbf{B}\right]\right)^{-1}\mathcal{SP}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1\right)\right). \tag{9.35}$$

This is a big improvement over Eq. (9.29) since it yields an explicit expansion for α in ascending powers of $\mathbf{B}$;

$$\boxed{\alpha = \alpha_0 + ((\mathbf{Bc}_1, \mathbf{d}_1)) + ((\mathbf{B}\mathcal{SP}\mathbf{Bc}_1, \mathbf{d}_1)) + ((\mathbf{B}\mathcal{SP}[(\alpha_0 - \alpha)\mathbf{1} + \mathbf{B}]\mathcal{SP}\mathbf{Bc}_1, \mathbf{d}_1)) + \cdots,} \tag{9.36}$$

where the factor $(\alpha_0 - \alpha)$ in the last term has to be approximated by $-((\mathbf{Bc}_1, \mathbf{d}_1))$. This is the master formula on which everything else is based.

Throughout this discussion, though it has not been indicated explicitly, the factor $\mathbf{B}$ has been allowed to be time dependent. But the leading correction term, $((\mathbf{Bc}_1, \mathbf{d}_1))$, because it is first order in $\mathbf{B}$, averages to zero for all but the time-independent part of $\mathbf{B}$. In the higher order terms it is possible for the products of time-varying factors to have non-zero average values. Though Eq. (9.36) is pleasingly compact, there is still a great deal of work to do to evaluate the higher order terms, mainly because the operator $\mathcal{S}$ has been defined only implicitly (by Eqs. (9.7), (9.9), and (9.10)).

Since the perturbing terms are periodic it is possible to Fourier-expand them in complex exponentials;

$$\mathbf{P}^{s,1}_{jh} = \sum_m P^{s,1(m)}_{jh}\, e^{im\mu_s t}, \quad \mathbf{P}^{a,1}_{jh} = \sum_m P^{a,1(m)}_{jh}\, e^{im\mu_s t}, \quad \mathcal{Q}^{1}_{jh} = \sum_m Q^{1(m)}_{jh}\, e^{im\mu_s t}. \tag{9.37}$$

In our case the $\mathbf{P}$-elements are non-vanishing only for $m = \pm 1$ and the $\mathbf{Q}$-elements are non-vanishing only for $m = 0$.

10. Pure damping

Purely damped motion, with no coupling between modes, though elementary, provides practice in evaluating the matrix elements of Eq. (9.36). Consider a weakly-damped, one-dimensional oscillator with equation of motion

$$\frac{d^2 e}{dt^2} + 2\alpha_{\mu_s}\frac{de}{dt} + \mu_0^2 e = 0. \tag{10.1}$$

Free motion of this oscillator is known to be described by the real part of

$$e\,(t) = e^{\left(-\alpha_{\mu_s} + i\sqrt{\mu_0^2 - \alpha_{\mu_s}^2}\right)t} \tag{10.2}$$

The "complex frequency" is

$$\alpha = i\sqrt{\mu_0^2 - \alpha_{\mu_s}^2} - \alpha_{\mu_s} \approx i\mu_0 - \alpha_{\mu_s} - i\frac{\alpha_{\mu_s}^2}{2\mu_0} + \cdots. \tag{10.3}$$

To describe this system by the previous formalism, treating the damping perturbatively, define

$$\mathbf{C} = \begin{pmatrix} 0 & 1 \\ -\mu_0^2 & 0 \end{pmatrix},\ \mathbf{B} = \begin{pmatrix} 0 & 0 \\ 0 & -2\alpha_{\mu_s} \end{pmatrix},\ \mathbf{c}_1 = \begin{pmatrix} a_1 \\ i\mu_0 a_1 \end{pmatrix},\ \mathbf{d}_1 = \begin{pmatrix} \mu_0 a_1 \\ i a_1 \end{pmatrix},\ a_1 = \sqrt{\frac{1}{2\mu_0}}. \tag{10.4}$$

(Temporariy these are 2×2 matrices and 2 component vectors.) The only time-independent perturbing term among Eqs. (9.37), $\mathbf{Q}^1$, corresponds to pure damping. Its first order matrix elements are

$$\mu_{sjh} = (\mathbf{B}\mathbf{c}_j, \mathbf{d}_h) = \begin{pmatrix} \mu_h \mathbf{a}_h^\dagger & -i\mathbf{a}_h^\dagger \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Q}^1 \end{pmatrix} \begin{pmatrix} \mathbf{a}_j \\ i\mu_j \mathbf{a}_j \end{pmatrix} = -\mu_j \mathbf{a}_h^\dagger \mathbf{Q}^1 \mathbf{a}_j. \tag{10.5}$$

With $\mathbf{Q}^1 = 2\,\mathrm{diag}(\alpha_\sigma, 0, 0, 0)$

$$\sigma_{11} = -\alpha_\sigma, \tag{10.6}$$

and Eq. (9.36) reduces

$$\alpha = i\mu_{1,0} - \alpha_\sigma, \tag{10.7}$$

that agrees to a first approximation with Eq. (10.3).

Proceeding to the next approximation, the operator $\mathcal{P}$ is obtained by striking all but rows and columns 1 and 5 from Eq. (9.15). Then

$$\mathcal{P}\mathbf{B}\mathbf{c}_1 = \begin{pmatrix} \frac{1}{2} & \frac{i}{2\mu_0} \\ \frac{-i\mu_0}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -2\alpha_\sigma \end{pmatrix} \begin{pmatrix} a_1 \\ i\mu_0 a_1 \end{pmatrix} = \begin{pmatrix} 1 \\ -i\mu_0 \end{pmatrix} \alpha_\sigma a_1. \tag{10.8}$$

The result of operating on this vector with $\mathcal{S}$ is defined operationally by Eq. (9.17);

$$(\mathbf{C} - i\mu_0\mathbf{1}) \begin{pmatrix} e \\ p \end{pmatrix} = \begin{pmatrix} -i\mu_0 & 1 \\ -\mu_0^2 & -i\mu_0 \end{pmatrix} \begin{pmatrix} e \\ p \end{pmatrix} = \begin{pmatrix} -1 \\ i\mu_0 \end{pmatrix} \alpha_\sigma a_1. \tag{10.9}$$

The attempt to solve this equation by inverting $\mathbf{C} - i\mu_0\mathbf{1}$ fails because its determinant vanishes, but solving the upper equation directly yields

$$\begin{pmatrix} 0 \\ -\alpha_\sigma a_1 \end{pmatrix} + \begin{pmatrix} 1 \\ i\mu_0 \end{pmatrix} e \tag{10.10}$$

as a vector satisfying Eq. (10.9) for any value of e. The second term is proportional to $\mathbf{c}_1$ (as, according to the general theory, it must be.) Finally, condition (9.11),

$$0 = (\mathbf{y}, \mathbf{d}_1) = (\,\mu_0 a_1 \quad -i a_1\,) \left(\begin{pmatrix} 0 \\ -\alpha_\sigma a_1 \end{pmatrix} + \begin{pmatrix} 1 \\ i\mu_0 \end{pmatrix} y_1 \right) \tag{10.11}$$

fixes $e = -\frac{i\alpha_\sigma a_1}{2\mu_0}$ and

$$\mathcal{SP}\mathbf{Bc}_1 = -\frac{\alpha_\sigma}{2\mu_0}\begin{pmatrix} ia_1 \\ \mu_0 a_1 \end{pmatrix}. \tag{10.12}$$

Finally we evaluate

$$(\mathbf{B}\mathcal{SP}\mathbf{Bc}_1, \mathbf{d}_1) = -\begin{pmatrix} \mu_0 a_1 & -ia_1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & -2\alpha_\sigma \end{pmatrix}\frac{\alpha_\sigma}{2\mu_0}\begin{pmatrix} ia_1 \\ \mu_0 a_1 \end{pmatrix} = -\frac{i\alpha_\sigma^2}{2\mu_0}, \tag{10.13}$$

which, when substituted into Eq. (9.36) agrees with Eq. (10.3).

11. Pure head-tail oscillation

As observed previously, their is no first order contribution to Eq. (9.36) from the $\mathbf{P}^1$ perturbation terms defined in Eqs. (8.3) and (8.4). Suppppressing the $\mathbf{Q}$ perturbation for the time being, the second order matrix elements due to $\mathbf{P}^1$ are

$$((\mathbf{B}\mathcal{SP}\mathbf{Bc}_j, \mathbf{d}_h)) = ((\mathbf{B}\mathcal{S}\mathbf{Bc}_j, \mathbf{d}_h)) = \left(\left(\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0} \end{pmatrix}\mathcal{S}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0} \end{pmatrix}\mathbf{c}_j, \mathbf{d}_h\right)\right); \tag{11.1}$$

here the factor $\mathcal{P}$ has been set to $\mathbf{1}$ because it operates on a quantity that averages to zero. By its definition in Eq. (9.16) the factor $\mathcal{S}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0} \end{pmatrix}\mathbf{c}_1$ is a solution to the equation

$$\frac{d\mathbf{y}}{dt} - (\mathbf{C} - i\mu_{1,0}\mathbf{1})\,\mathbf{y} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0} \end{pmatrix}\mathbf{c}_1 = \mathbf{u}\,(t)\,; \tag{11.2}$$

or

$$\begin{pmatrix} \left(\frac{d}{dt} + i\mu_{1,0}\right)\mathbf{1} & -\mathbf{1} \\ \mathbf{P}^0 & \left(\frac{d}{dt} + i\mu_{1,0}\right)\mathbf{1} \end{pmatrix}\mathbf{y} = \mathbf{u}\,(t)\,. \tag{11.3}$$

Seeking a solution with time variation $e^{i\rho t}$ this becomes

$$\begin{pmatrix} (i\rho + i\mu_{1,0})\,\mathbf{1} & -\mathbf{1} \\ \mathbf{P}^0 & (i\rho + i\mu_{1,0})\,\mathbf{1} \end{pmatrix}\mathbf{y} = \mathbf{u}\,(t)\,. \tag{11.4}$$

Taking advantage of the fact that this matrix is block-by-block diagonal by multiplying by the matrix shown on the right in the next equation yields

$$\begin{pmatrix} -\left(\rho + \mu_{1,0}\right)^2\mathbf{1} + \mathbf{P}^0 & \mathbf{0} \\ \mathbf{0} & -\left(\rho + \mu_{1,0}\right)^2\mathbf{1} + \mathbf{P}^0 \end{pmatrix}\mathbf{y} = \begin{pmatrix} (i\rho + i\mu_{1,0})\,\mathbf{1} & \mathbf{1} \\ -\mathbf{P}^0 & (i\rho + i\mu_{1,0})\,\mathbf{1} \end{pmatrix}\mathbf{u}\,(t)\,. \tag{11.5}$$

The quantity ρ is not simple here, since it depends on the time dependence of the vector being operated on, but it can be otherwise be treated as an ordinary parameter. Define

$$\begin{aligned}\mathcal{T}_{1\rho} = & -\left(\rho+\mu_{1,0}\right)^2\mathbf{1}+\mathbf{P}^0 \\ = & -\operatorname{diag}(\rho^2+2\mu_{1,0}\rho, \quad \rho^2+2\mu_{1,0}\rho+\mu\Sigma, \\ & \qquad \rho^2+2\mu_{1,0}\rho-2\iota\rho\overline{w}, \quad \rho^2+2\mu_{1,0}\rho+\mu\Sigma-2\iota\rho\overline{w})\end{aligned} \tag{11.6}$$

Also we will use the abbreviation and partitioning

$$\mathcal{T}_{1\rho I}\equiv\left(\mathcal{T}_{1\rho}\right)^{-1}\equiv\begin{pmatrix}\mathcal{T}_{1\rho I}^{(u)} & \mathbf{0} \\ \mathbf{0} & \mathcal{T}_{1\rho I}^{(l)}\end{pmatrix}. \tag{11.7}$$

Continuing with (11.5)

$$\mathbf{y}=\mathcal{S}\mathbf{u}\left(t\right)=\begin{pmatrix}\mathcal{T}_{1\rho I}\left(i\rho+i\mu_{1,0}\right) & \mathcal{T}_{1\rho I} \\ -\mathcal{T}_{1\rho I}\mathbf{P}^0 & \mathcal{T}_{1\rho I}\left(i\rho+i\mu_{1,0}\right)\end{pmatrix}\mathbf{u}\left(t\right), \tag{11.8}$$

which serves to define $\mathcal{S}$. (It should perhaps be symbolized by $\mathcal{S}_{1\rho}$ since it depends on the particular eigenvalue being evaluated and on ρ, which in our case will be $\pm\mu_s$ or 0.) We have then

$$\begin{aligned}\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B} = & \begin{pmatrix}\mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathcal{T}_{1\rho I}\left(i\rho+i\mu_{1,0}\right) & \mathcal{T}_{1\rho I} \\ -\mathcal{T}_{1\rho I}\mathbf{P}^0 & \mathcal{T}_{1\rho I}\left(i\rho+i\mu_{1,0}\right)\end{pmatrix}\begin{pmatrix}\mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0}\end{pmatrix} \\ = & \begin{pmatrix}\mathbf{0} & \mathbf{0} \\ \mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1 & \mathbf{0}\end{pmatrix}.\end{aligned} \tag{11.9}$$

Then we have

$$\begin{aligned}\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1 = & \begin{pmatrix}\mathbf{0} & \mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c} \\ \mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathcal{T}_{1\rho I}^{(u)} & \mathbf{0} \\ \mathbf{0} & \mathcal{T}_{1\rho I}^{(l)}\end{pmatrix}\begin{pmatrix}\mathbf{0} & \mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c} \\ \mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c} & \mathbf{0}\end{pmatrix} \\ = & \begin{pmatrix}\left(\mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\right)\mathcal{T}_{1\rho I}^{(l)}\left(\mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c}\right) & \mathbf{0} \\ \mathbf{0} & \left(\mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\right)\mathcal{T}_{1\rho I}^{(u)}\left(\mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c}\right)\end{pmatrix}\end{aligned} \tag{11.10}$$

The elements of this matrix have to be time-averaged; for example

$$\begin{aligned}&\left(\left(\left(\mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\right)\mathcal{T}_{1\rho I}^{(l)}\left(\mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c}\right)\right)\right) \\ &\quad = \frac{1}{2}\left(\left(\left(\mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\right)\left(\mathcal{T}_{+,I}^{(l)}\left(-i\mathcal{D}+\mathcal{W}\right)e^{i\mu_s t}+\mathcal{T}_{-,I}^{(l)}\left(i\mathcal{D}+\mathcal{W}\right)e^{-i\mu_s t}\right)\right)\right) \\ &\quad = \frac{1}{4}\left(i\mathcal{D}-\mathcal{W}\right)\mathcal{T}_{+,I}^{(l)}\left(-i\mathcal{D}+\mathcal{W}\right)+\frac{1}{4}\left(-i\mathcal{D}-\mathcal{W}\right)\mathcal{T}_{-,I}^{(l)}\left(i\mathcal{D}+\mathcal{W}\right) \\ &\quad = \frac{1}{4}\left(\mathcal{D}\left(\mathcal{T}_{+,I}^{(l)}+\mathcal{T}_{-,I}^{(l)}\right)\mathcal{D}+\mathcal{W}\left(\mathcal{T}_{+,I}^{(l)}+\mathcal{T}_{-,I}^{(l)}\right)\mathcal{W}\right) \\ &\qquad +\frac{i}{4}\left(\mathcal{D}\left(\mathcal{T}_{+,I}^{(l)}-\mathcal{T}_{-,I}^{(l)}\right)\mathcal{W}+\mathcal{W}\left(\mathcal{T}_{+,I}^{(l)}-\mathcal{T}_{-,I}^{(l)}\right)\mathcal{D}\right).\end{aligned} \tag{11.11}$$

The *real part* is given by

$$\frac{1}{4}\begin{pmatrix} D & d \\ d & D \end{pmatrix}\begin{pmatrix} \mathcal{T}_{3+,I}+\mathcal{T}_{3-,I} & 0 \\ 0 & \mathcal{T}_{4+,I}+\mathcal{T}_{4-,I} \end{pmatrix}\begin{pmatrix} D & d \\ d & D \end{pmatrix} + \text{similar for } \mathcal{W}. \tag{11.12}$$

The diagonal elements have appeared before in Eq. (7.2). Substituting from there the 1,1 element is

$$\frac{1}{4}\left(g_{13}D^2 + g_{14}d^2\right) + \text{similar for } \mathcal{W}, \tag{11.13}$$

The *imaginary part* of Eq. (11.11), using the constants f_{ij} from Eq. (5.10), is

$$-\frac{\mu\iota}{2\pi}\frac{i}{4}\left(\begin{pmatrix} D & d \\ d & D \end{pmatrix}\begin{pmatrix} f_{13} & 0 \\ 0 & f_{14} \end{pmatrix}\begin{pmatrix} \overline{w} & \Delta w \\ \Delta w & \overline{w} \end{pmatrix} + \begin{pmatrix} \overline{w} & \Delta w \\ \Delta w & \overline{w} \end{pmatrix}\begin{pmatrix} f_{13} & 0 \\ 0 & f_{14} \end{pmatrix}\begin{pmatrix} D & d \\ d & D \end{pmatrix}\right). \tag{11.14}$$

Its 11 element is

$$-\frac{\mu\iota}{2\pi}\frac{i}{2}\left(f_{13}D\overline{w} + f_{14}d\Delta w\right). \tag{11.15}$$

Collecting results we have

$$\begin{aligned}((\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1)) = & \left(\left(\begin{pmatrix} \mu_{1,0}\mathbf{a}_1^\dagger & -i\mathbf{a}_1^\dagger \end{pmatrix}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1 & \mathbf{0} \end{pmatrix}\begin{pmatrix} \mathbf{a}_1 \\ i\mu_{1,0}\mathbf{a}_1 \end{pmatrix}\right)\right) \\ = & -i\left(\left(\mathbf{a}_1^\dagger\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\right)\right) \\ = & \frac{-i}{2\mu}\left(\left((\mathcal{D}\mathrm{s} - \mathcal{W}\mathrm{c})\,\mathcal{T}_{1\rho I}^{(l)}\,(\mathcal{D}\mathrm{s} + \mathcal{W}\mathrm{c})\right)\right)_{11}\end{aligned} \tag{11.16}$$

For example, the $\iota = 0$ imaginary part of this is given by

$$\Im((\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1)) \overset{\iota=0}{=} \frac{-1}{8\mu}\left(g_{13}D^2 + g_{14}d^2\right). \tag{11.17}$$

When substituted into Eq. (9.36) this agrees with Eq. (7.3). The overall real part is

$$\Re((\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1)) = \frac{1}{4\mu}\frac{\mu\iota}{2\pi}\left(f_{13}D\overline{w} + f_{14}d\Delta w\right), \tag{11.18}$$

which (except for a factor of 2) agrees with Eq. (5.14).

12. Sympathetic damping

Suppose that, in lowest approximation, one mode is damped but a second is not. If there is coupling between modes there is the possibility, in a next approximation, that the second mode will acquire some damping due to its interaction with the first. This phenomenon will be called "sympathetic damping". The need to analyse this phenomenon provided the main motivation for developing the Hamiltonian formalism of the previous several sections. Sympathetic damping may influence the head-tail strongly because, in lowest approximation, the π-modes are thought to be strongly Landau damped (growth rate α_π) while the σ-modes are not.

It is usually easiest to calculate sympathetic damping indirectly, using energy conservation. But here, since the longitudinal motion (treated here as inexorable) is a potential source of energy, conservation of energy is not necessarily applicable and the effect will be calculated directly.

Since we have no reliable calculation or measurement of α_π it will have to be regarded as another parameter to be determined empirically by fitting to the data. Here we are primarily concerned with the small chromaticity region where $\alpha_\sigma \approx 0$. According to our current understanding of the effect of decoherence, α_σ becomes appreciable in the region of highly-unbalanced chromaticities, which suggests that sympathetic damping plays a relatively much smaller roll there. For simplicity, continuing to specialize to mode 1, we take $\alpha_\sigma = 0$ as the mode 1 unperturbed growth rate and seek its first non-vanishing contribution proportional to α_π which is the unperturbed growth rate of modes 3 and 4.

Because $\mathbf{Q}^1$ is independent of time and both of the $\mathbf{P}^1$ perturbations are sinuisoidally varying, the first contribution to sympathetic damping is the last term of Eq. (9.36),

$$((\mathbf{B}\mathcal{S}\mathcal{P}\left[(\alpha_0-\alpha)\,\mathbf{1}+\mathbf{B}\right]\mathcal{S}\mathcal{P}\mathbf{B}\mathbf{c}_1,\mathbf{d}_1)) = ((\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathbf{c}_1,\mathbf{d}_1))\,. \tag{12.1}$$

Here, following the recipe mentioned below Eq. (9.36), the factor $\alpha_0-\alpha$ has been obtained from the σ-mode damping rate α_σ using Eq. (10.6). Since we are assuming $\alpha_\sigma = 0$, the first term in Eq. (12.1) has been dropped.

With the remaining term being cubic in the perturbation $\mathbf{B}$, one can take the time-independent $\mathbf{Q}^1$ as one of the factors, and take one of $e^{\pm i\mu_s t}$ for the $\mathbf{P}^{s,1}$ and the other for

$\mathbf{P}^{a,1}$, to end up with a non-zero, time-independent matrix element. Since both $\mathbf{P}^{s,1}$ and $\mathbf{P}^{a,1}$ are block off-diagonal, individually they "toggle" modes 1 and 2 to modes 3 and 4 or *vice versa*, but acting twice they mix modes 1 and 2 between themselves and modes 3 and 4 between themselves. Since the extra factor $\mathbf{Q}^1$ does not mix states at all, the overall matrix element (12.1) can therefore be expected to be non-vanishing (barring accidental cancellation.) Unfortunately the matrix element is complicated and there are numerous possible ordering of the factors. To take advantage of calculations of the previous sections Eq. (12.1) can be factorized as

$$\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B} = (\mathbf{B}\mathcal{S}\mathbf{B})\left(\mathcal{S}\mathcal{P}\mathbf{B}'\right) + \mathbf{B}\left(\mathcal{S}\mathbf{B}'\right)(\mathcal{S}\mathbf{B}) + \left(\mathbf{B}'\mathcal{S}\mathcal{P}\right)(\mathbf{B}\mathcal{S}\mathbf{B}), \tag{12.2}$$

where now $\mathbf{B}$ contains both of the $\mathbf{P}^1$ perturbations and $\mathbf{B}'$ contains the $\mathbf{Q}^1$ perturbation; the factor $\mathcal{P}$ has been retained only when it operates on a quantity with time-independent part. These terms will be evaluated one-by-one, specializing to operand $\mathbf{c}_1$ in cases where $\mathcal{P}$ has had to be retained.

$$\begin{aligned}
(\mathbf{B}\mathcal{S}\mathbf{B})\left(\mathcal{S}\mathcal{P}\mathbf{B}'\mathbf{c}_1\right) &= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1 & \mathbf{0}\end{pmatrix}\left(-\frac{\alpha_\sigma}{2\mu_{1,0}}\right)\begin{pmatrix} i\mathbf{a}_1 \\ \mu_{1,0}\mathbf{a}_1\end{pmatrix} = -\frac{i\alpha_\sigma}{2\mu_{1,0}}\begin{pmatrix} \mathbf{0} \\ \mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\end{pmatrix} = 0 \\
\mathbf{B}\left(\mathcal{S}\mathbf{B}'\right)(\mathcal{S}\mathbf{B}) &= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0}\end{pmatrix}\begin{pmatrix} \mathcal{T}_{1\rho I}\left(i\rho + i\mu_{1,0}\right) & \mathcal{T}_{1\rho I} \\ -\mathcal{T}_{1\rho I}\mathbf{P}^0 & \mathcal{T}_{1\rho I}\left(i\rho + i\mu_{1,0}\right)\end{pmatrix}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Q}^1\end{pmatrix} \\
&\quad \begin{pmatrix} \mathcal{T}_{1\rho I}\left(i\rho + i\mu_{1,0}\right) & \mathcal{T}_{1\rho I} \\ -\mathcal{T}_{1\rho I}\mathbf{P}^0 & \mathcal{T}_{1\rho I}\left(i\rho + i\mu_{1,0}\right)\end{pmatrix}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1 & \mathbf{0}\end{pmatrix} \\
&= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{Q}^1\mathcal{T}_{1\rho I}\left(i\rho + i\mu_{1,0}\right)\mathbf{P}^1 & \mathbf{0}\end{pmatrix} \\
\left(\mathbf{B}'\mathcal{S}\mathcal{P}\right)(\mathbf{B}\mathcal{S}\mathbf{B})\,\mathbf{c}_1 &= \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{Q}^1\end{pmatrix}\begin{pmatrix} \mathcal{T}_{1\rho I}\left(i\rho + i\mu_{1,0}\right) & \mathcal{T}_{1\rho I} \\ -\mathcal{T}_{1\rho I}\mathbf{P}^0 & \mathcal{T}_{1\rho I}\left(i\rho + i\mu_{1,0}\right)\end{pmatrix}\mathcal{P}\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1 & \mathbf{0}\end{pmatrix}\mathbf{c}_1
\end{aligned} \tag{12.3}$$

The first term vanishes because of our assumption $\alpha_\sigma = 0$. Had this approximation not been made the result would be

$$\left(\left(\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}'\mathbf{c}_1, \mathbf{d}_1\right)\right) = \frac{-i\alpha_\sigma}{2\mu}\left(\left(-i\mathbf{a}_1^\dagger\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\right)\right) = \frac{-i\alpha_\sigma}{2\mu}\left(\left(\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1\right)\right). \tag{12.4}$$

Here the extremely small factor $\frac{-i\alpha_\sigma}{2\mu}$ multiplies the eigenvalue shift calculated in the previous order of approximation in Eq. (11.16). This term is therefore entirely negligible whether or not α_σ is negligible.

The third contribution to Eq. (12.2) includes the following factor, which "starts" with a factor $\mathbf{B}\mathcal{SPB}$ already calculated in Eq. (11.9);

$$
\begin{aligned}
&\left(\mathbf{B}'\mathcal{SP}\right)\left(\mathbf{B}\mathcal{SP}\mathbf{B}\mathbf{c}_1\right) \approx \begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & -\mathbf{Q}^1\end{pmatrix}\begin{pmatrix}\mathcal{T}_{1\rho I}\left(i\mu_{1,0}\right) & \mathcal{T}_{1\rho I}\\ -\mathcal{T}_{1\rho I}\mathbf{P}^0 & \mathcal{T}_{1\rho I}\left(i\mu_{1,0}\right)\end{pmatrix}\mathcal{P}\begin{pmatrix}\mathbf{0}\\ \mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\end{pmatrix}\\
&= \frac{1}{2}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & -\mathbf{Q}^1\end{pmatrix}\begin{pmatrix} i\mathcal{T}_{1\rho I} & \mathcal{T}_{1\rho I}/\mu_{1,0}\\ -\mathcal{T}_{1\rho I}\mathbf{P}^0/\mu_{1,0} & i\mathcal{T}_{1\rho I}\end{pmatrix}\begin{pmatrix} i\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\\ \mu_{1,0}\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\end{pmatrix}\\
&= \frac{1}{2}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{Q}^1\mathcal{T}_{1\rho I}\mathbf{P}^0/\mu_{1,0} & -i\mathbf{Q}^1\mathcal{T}_{1\rho I}\end{pmatrix}\begin{pmatrix} i\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\\ \mu_{1,0}\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\end{pmatrix}\\
&= \frac{i\mu}{2}\begin{pmatrix}\mathbf{0}\\ \mathbf{Q}^1\mathcal{T}_{1\rho I}\left(\frac{\mathbf{P}^0}{\mu_{1,0}^2}-\mathbf{1}\right)\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\end{pmatrix} = 0,
\end{aligned}
\tag{12.5}
$$

which can be seen to vanish since all operators preceeding $\mathbf{Q}^1$ (reading from right to left) feed nothing from lower to upper components, and $\mathbf{Q}^1$ annihilates upper elements. (Inclusion of the factor $\mathcal{P}$ does not alter this.)

The second (and only non-vanishing) term of Eq. (12.2) yields

$$
\begin{aligned}
&\left(\left(\mathbf{B}\mathcal{SP}\mathbf{B}'\mathcal{SP}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1\right)\right) \approx -\mu\left(\left(\mathbf{a}_1^\dagger\mathbf{P}^1\mathcal{T}_{1\rho I}\mathbf{Q}^1\mathcal{T}_{1\rho I}\mathbf{P}^1\mathbf{a}_1\right)\right)\\
&= -2\alpha_\pi\mu\left(\left(\mathbf{a}_1^\dagger\begin{pmatrix}\mathbf{0} & \mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\\ \mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathcal{T}_{1\rho I}^{(l)2}\end{pmatrix}\begin{pmatrix}\mathbf{0} & \mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\\ \mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c} & \mathbf{0}\end{pmatrix}\mathbf{a}_1\right)\right)\\
&= -2\alpha_\pi\mu\left(\left(\mathbf{a}_1^\dagger\begin{pmatrix}\left(\mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\right)\mathcal{T}_{1\rho I}^{(l)2}\left(\mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c}\right) & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\mathbf{a}_1\right)\right)\\
&= -\alpha_\pi\left(\left(\left(\mathcal{D}\mathrm{s}-\mathcal{W}\mathrm{c}\right)\mathcal{T}_{1\rho I}^{(l)2}\left(\mathcal{D}\mathrm{s}+\mathcal{W}\mathrm{c}\right)\right)\right)_{11}
\end{aligned}
\tag{12.6}
$$

where $|\rho| << |\mu_{1,0}|$ has been assumed. This expression resembles the matrix element appearing in Eq. (11.16). But now the elements of $\mathcal{T}_{1\rho I}^{(l)2}$ will be approximated by $1/(2\mu\mu_s)^2$.

$$
\begin{aligned}
&\left(\left(\mathbf{B}\mathcal{SP}\mathbf{B}'\mathcal{SP}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1\right)\right) \approx -\frac{\alpha_\pi}{4\mu^2{\mu_s}^2}\left(\left(\mathcal{D}^2\mathrm{s}^2-\mathcal{W}^2\mathrm{c}^2\right)\right)_{11} = -\frac{\alpha_\pi}{8\mu^2{\mu_s}^2}\left(\left(\mathcal{D}^2-\mathcal{W}^2\right)\right)_{11}\\
&= -\frac{\alpha_\pi}{8\mu^2{\mu_s}^2}\left(D^2+d^2-\left(\overline{w}^2+\Delta w^2\right)\left(\frac{\mu\iota}{2\pi}\right)^2\right)\\
&\approx -\alpha_\pi\left(\frac{\pi^2}{{\mu_s}^2}\left(Q'^2_x+Q'^2_x\right)\left(\frac{dp}{p}\right)^2_{\mathrm{typ}} - \frac{w_x^2+w_y^2}{16\mu^2{\mu_s}^2}\left(\frac{\mu\iota}{2\pi}\right)^2\right).\\
&\overset{\mathrm{typ}}{=} -\alpha_\pi\left(\iota\right)\left(0.58\times10^{-3}\left(Q'^2_x+Q'^2_x\right) - 0.32\times10^{-3}\left(\overline{w}\frac{\mu\iota_{SR}}{2\pi}\right)^2\left(\frac{\iota}{\iota_{SR}}\right)^2\right).
\end{aligned}
\tag{12.7}
$$

where the current ι is referred to a current ι_{SR} defined to be the current for which, with $Q'_x + Q'_y = 1$, the wake field anti-damping just cancels $\alpha_{SR} \overset{\text{typ}}{=} 10^{-4}$. In principle ι_{SR} is measurable, but in practice it should perhaps be regarded as an adjustable parameter. In the factor multiplying $\alpha_\pi(\iota)$, the second term vanishes for $\iota = 0$ and, barring unrealistically large chromaticities, the first term is small compared to 1. Though α_π could plausibly be 10 or 100 times greater than α_{SR} the zero current damping from this source still is probably negligible. The factor $Q'^2_x + Q'^2_x$ can perhaps change this but in the high chromaticity regime the sympathetic damping is probably still negligibly small in comparison to the damping that the σ mode itself acquires due to decoherence. To complicate matters further, the π-mode growth rate $\alpha_\pi(\iota)$ is itself strongly current dependent, because of space charge or wake field forces.

For $\iota >> \iota_{SR}$ the second term of Eq. (12.7) dominates the first. Since the (roughly equal) factors w_x and w_y are entirely phenomenological, they must be estimated by fitting to measured damping rates using Eq. (5.14);

$$\alpha_{SR} = D\frac{1}{8\mu}\frac{1}{\mu\mu_s}\left(\overline{w}\frac{\mu\iota_{SR}}{2\pi}\right) = \frac{\pi}{4\mu\mu_s}\left(\frac{dp}{p}\right)_{\text{typ}}\left(\overline{w}\frac{\mu\iota_{SR}}{2\pi}\right) \tag{12.8}$$

$$\overline{w}\frac{\mu\iota_{SR}}{2\pi} = \alpha_{SR}\frac{4\mu\mu_s}{\pi}\left(\frac{dp}{p}\right)^{-1}_{\text{typ}} \overset{\text{typ}}{=} 4.2 \tag{12.9}$$

If we accept this estimate, and simply drop the first term of Eq. (12.7), the resulting "sympathetic damping rate" is

$$\alpha_{\sigma,\text{symp}} = \big((\mathbf{B}\mathcal{S}\mathcal{P}\mathbf{B}'\mathcal{S}\mathcal{P}\mathbf{B}\mathbf{c}_1, \mathbf{d}_1)\big) \approx \alpha_\pi(\iota) \times 0.56 \times 10^{-2}\left(\frac{\iota}{\iota_{SR}}\right)^2 \tag{12.10}$$

Apart from the head-tail damping rate α_{HT} itself, this is likely to be the strongest source of current dependence of visible modes. Of course this formula can have only a restricted range of validity. For sufficiently large ι presumably $\alpha_{\sigma,\text{symp}} \to \alpha_\pi$ with both being approximately half the value of α_π in the absence of coupling. We therefore replace Eq. (12.10) by

$$\boxed{\alpha_{\sigma,\text{symp}} \approx \tfrac{1}{2}\alpha_\pi(\iota)(1 - e^{-1.12\times 10^{-2}\iota^2/\iota^2_{SR}});} \tag{12.11}$$

the only basis for this form is that it "plausibly" interpolates between weak damping and strong damping extremes.

13. Computer simulation

Using the accelerator modeling program *TEAPOT* some of the phenomena that enter into chromaticity sharing have been simulated. Simulation has the potential advantage of exhibiting behavior that is too complicated to be described by idealized models, for example because of conspiracy of more than one complicating effect, and the complementary disadvantages of possibly masking simple causes and of too-faithfully mirroring the obscurity so characteristic of storage rings.

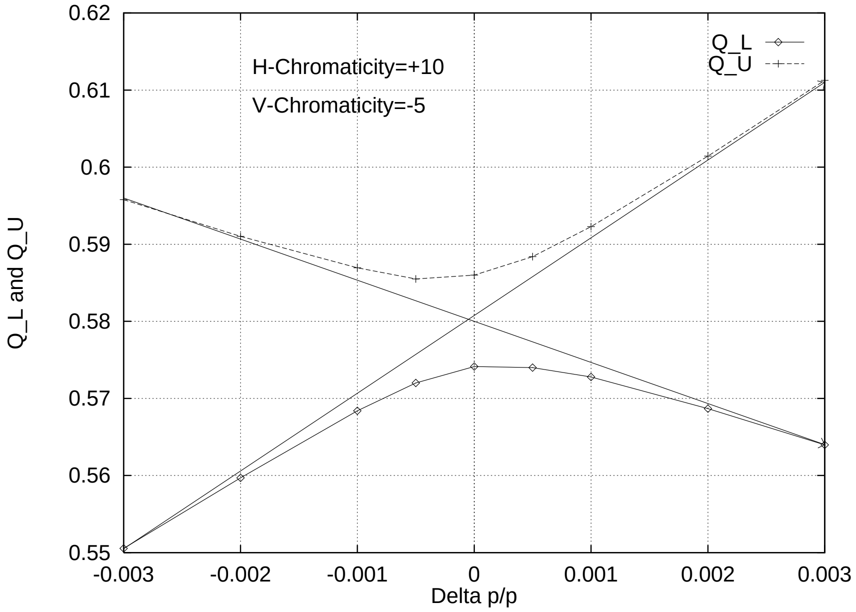

Figure 13.1: Dependence of visible eigentunes (fractional) on $\delta p/p$ as determined by *TEAPOT* analysis of the actual CESR lattice used for machine studies of chromaticity sharing. The straight lined show what the variation would be with no coupling.

Dependence of eigentunes on momentum deviation is shown in Fig. 13.1. The cross-plane coupling causes the lines to "repel" and hence not cross. It must be realized that these are the tunes exhibited by a constant-momentum particle executing betatron oscillations about the closed orbit appropriate for its momentum—they have nothing simple to say about the betatron motion of a particle executing longitudinal oscillations. The tunes at a particular abscissa in Fig. 13.1 can be measured operationally by setting the central momentum (the RF frequency actually) and then spectral analysing a beam position monitor signal. In the computer simulation the same thing is done. In particular, the particle's momentum is not altered by RF cavities present in the lattice because its longitudinal phase is such that it suffers no net longitudinal impulse. The tunes are extracted as eigenvalues of the once-around, on-momentum transfer matrix or, with excellent agreement, by FFT of turn-by-turn data.

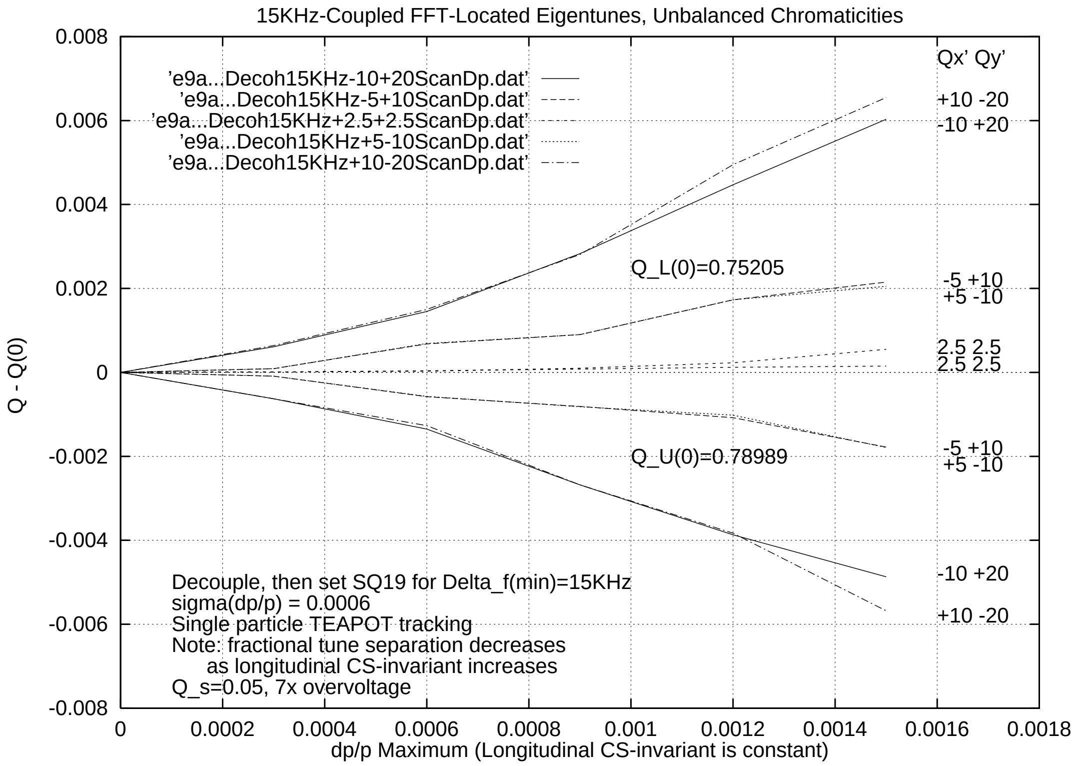

Figure 13.2: Average transverse tunes exhibited by a single particle executing longitudinal oscillations. Data for *unbalanced* chromaticities Q'_x and Q'_y are shown in this plot, balanced in the next.

A more relevant single particle calculation is to follow a single particle which is oscillating longitudinally after being launched off-momentum. To obtain the tunes it is necessary to find the line centers of the relevant lines in the spectrum obtained by applying FFT to the turn-by-turn data. Tunes obtained this way, for various chromaticity combinations, are plotted versus the peak longitudinal momentum offset (which is proportional to the square root of the longitudinal Courant-Snyder invariant) in Fig. 13.2 and Fig. 13.3. The first figure shows unbalanced chromaticities, the second, balanced.

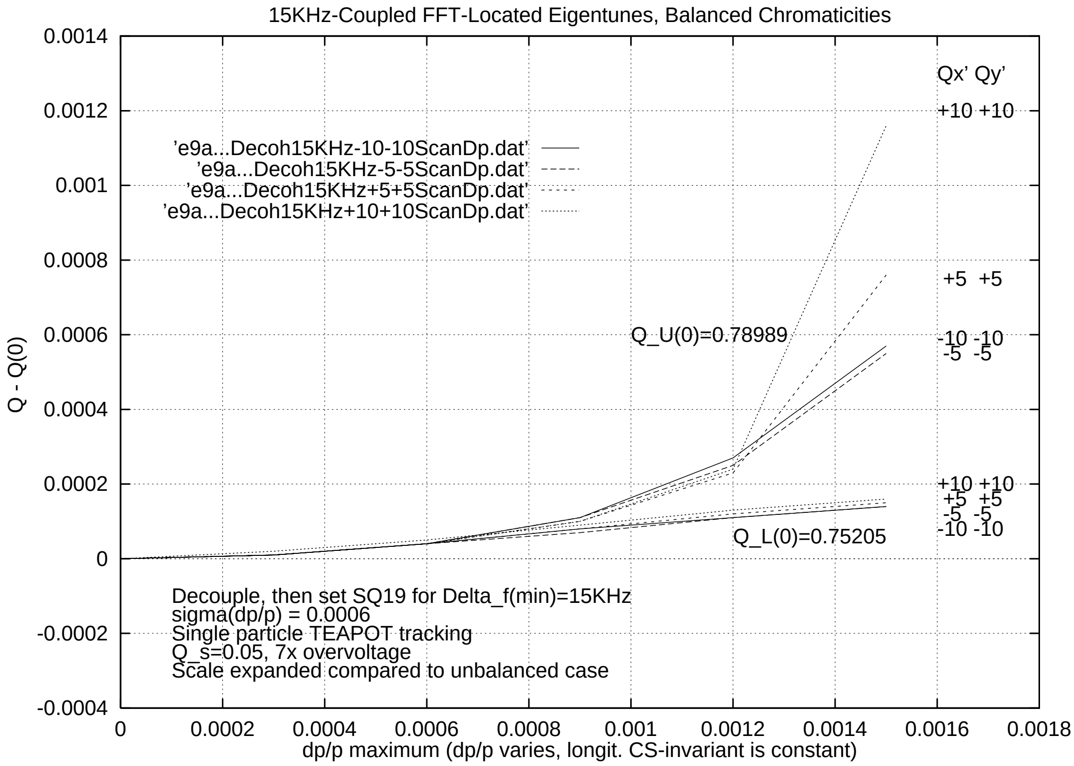

Figure 13.3: Average transverse tunes exhibited by a single particle executing longitudinal oscillations. Data for *balanced* chromaticities Q'_x and Q'_y are shown in this plot, unbalanced in the previous one.

A rough parameterization of the data in these two graphs is

$$\Delta Q_U = 0.4\times 10^{-4}\frac{\hat\delta}{\sigma_\delta} - \left[1.2\times 10^{-3}\frac{\hat\delta}{\sigma_\delta} + 0.4\times 10^{-3}\left(\frac{\hat\delta}{\sigma_\delta}\right)^2\right]\left(\frac{Q'_x-Q'_y}{30}\right)^2 + \\ + \left(1.0\times 10^{-5}\left(\frac{\hat\delta}{\sigma_\delta}\right)^4 ?\right), \tag{13.1}$$

$$\Delta Q_L = 0.4\times 10^{-4}\frac{\hat\delta}{\sigma_\delta} + \left[1.2\times 10^{-3}\frac{\hat\delta}{\sigma_\delta} + 0.4\times 10^{-3}\left(\frac{\hat\delta}{\sigma_\delta}\right)^2\right]\left(\frac{Q'_x-Q'_y}{30}\right)^2 .$$

The term shown in parenthesis with question mark is poorly determined and (for all but the highest momenta) probably negligible but, being asymmetric between upper and lower modes, it could contribute to differences in their behavior. By itself this data does not demand the quadratic dependence, $(Q'_x - Q'_y)^2$. If anything a higher power is suggested. When it comes to fitting our observations in a later section, since it will turn out that a $(Q'_x - Q'_y)^3$ dependence fits our data better, we will shamelessly accept adopt it.

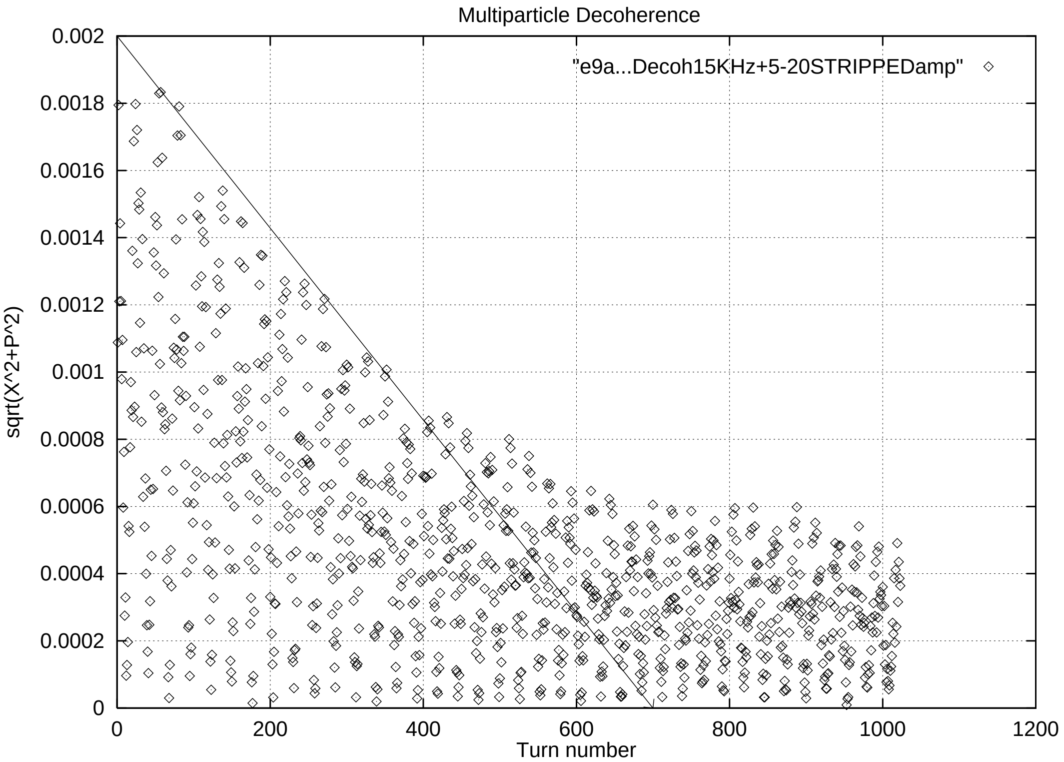

Figure 13.4: Multiparticle decoherence. Using the modeling program *TEAPOT*, an appropriately distributed "beam" consisting of 200 macroparticles is tracked for 1024 turns and the horizontal centroid coordinate is plotted each turn. The chromaticities are highly unbalanced: $Q'_x = +5$, $Q'_y = -20$. The "apparent damping rate" inferred using the straight line is $(700/0.39 \times 10^6)^{-1} = 560\,\mathrm{s}^{-1}$.

The second terms are normally dominant. Note that their signs are such that they cause the tunes to pull toward each other as $\hat{\delta}$ increases. Raphael Littauer has explained why this is to be expected and why the pulling should be proportional to $(Q'_x - Q'_y)^2$. Briefly, copying from his note[4], the argument proceeds as follows. Assume the lattice is perfectly decoupled with $Q_x = Q_y$ before a single skew quadrupole is powered to a strength causing minimum tune split $S \equiv \Sigma/(2\pi)$. To simplify the argument assume the skew quad is located at an equal-β waist and describe the motion at its center. With the skew quad still off, because of the equal tunes, any perpendicular axes can serve as normal mode axes. But the best choice is to define the two normal modes as being "plane-polarized" along the ± 45 degree diagonals. This is because, as the skew quad is turned on, a particle passing through the quad in say the $+45$-degree plane suffers a deflection leaving it in the same plane and, the x and y phase advances due to the rest of the lattice being equal, after a complete turn, the particle returns in the same plane. As a result the normal mode axes, no longer arbitrary, lie along these ± 45 degree diagonals. The reason the tunes are split is that the skew quad, as is always true for an *erect* quadrupole (which the skew quad *is* relative to the normal mode axes) shifts the two tunes in opposite directions. Using the standard tune shift formula, the mode tune split S and the skew quadrupole strength q_s are related by

$$S/2 = \frac{\Sigma/2}{2\pi} = \frac{q_s \beta}{4\pi}, \tag{13.2}$$

which is the maximum tune shift that a quad of strength q_s can cause. If the lattice chromaticities are the same in both planes the presence of longitudinal oscillations has no effect on this picture since after passage around the rest of the ring even an off-momentum particle executing normal-mode oscillation returns in the same plane. But if the chromaticities are unequal the normal mode oscillation "wobbles" about the diagonal axis. This wobbling has to *reduce* the minimum tune split since the value given by Eq. (13.2) is maximal. The importance of this effect is quantified by the difference between accumulated horizontal and vertical phase advances during one quarter synchrotron oscillation period, a quantity much like the parameter χ introduced in section (1), but now proportional the difference $Q'_x - Q'_y$ (and no longer doubled to account for head-tail phase difference). Taking $Q'_x - Q'_y = 30$ and other quantities as in section (1), the result is $\Delta\phi_{\max} = 0.57$ and $\cos(\Delta\phi_{\max}) = 0.84$. At a guess, $< \cos(\Delta\phi) > \approx 0.92$ is the factor by which S is reduced. With $S = 0.04$, its reduction is therefore estimated to be $0.04 \times 0.08 = 0.0032$ at $\hat{\delta} = 0.0006$. This agrees well with Fig. 13.2, which both corroborates the discussion so far and suggests that Littauer's simple picture accounts for most of the dependence on $\hat{\delta}$.

The tune dependence $\Delta Q(\hat{\delta})$ has an important effect on the damping rates being analysed because of decoherence but it is too small to influence significantly any of the other phenomena under consideration. $\Delta Q(\hat{\delta})$ could be calculated by the perturbative method of section 9, but with Σ having been taken as constant, independent of $\hat{\delta}$ this dependence is not included in the equations as they are written. But this is on no consequence since the effect has been reliably estimated and, as mentioned the effect is otherwise negligible.

14. Analytic treatment of decoherence

Betatron (x,p) phase space is shown in Fig. 14.1. The scales have been adjusted so that (linearized) motion in phase space is along circles centered on the origin, with phase advance per turn μ_0.† Assuming the particle distribution is Gaussian and isotropic,‡ it can be expressed either as $P_R(R)$ or as $P_x(x)P_p(p)$, depending on whether polar or cartesian coordinates are employed in phase space. The particles are also distributed with distribution $P_{\hat{\delta}}(\hat{\delta})$ in $\hat{\delta}$ which is the maximum value (as the particle oscillates longitudinally) of its fractional momentum deviation $\delta p/p$. The distributions are given by

$$\begin{aligned}
&P_x\left(x\right) = \frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{x^2}{2\sigma_x^2}}, \quad P_p\left(p\right) = \frac{1}{\sqrt{2\pi}\sigma_x}e^{-\frac{p^2}{2\sigma_x^2}}, \\
&P_{x,y}\left(x,y\right) = P_x\left(x\right)P_y\left(y\right) = \frac{P_R\left(R\right)}{2\pi R}, \quad \text{and hence} \\
&P_R\left(R\right) = \frac{R}{\sigma_x^2}e^{-\frac{R^2}{2\sigma_x^2}}.
\end{aligned} \tag{14.1}$$

This distribution can also be expressed as a joint probability distribution $P_{R,\Phi}(R,\Phi) = P_R(R)/(2\pi)$. A "kick" Δp is administered to every particle in the beam at $i=0$ and hence also to the beam centroid. In this section the subsequent motion of the centroid position is to be studied.

Motion of a particle initially at point **P** is shown in Fig. 14.1. If every particle advances at rate μ_0 the centroid does the same and its radius remains constant. But, in general, since $\mu(R,\Phi,\delta)$ depends on the location of **P** as well as on δ, the particle motions "decoherere" causing the centroid amplitude to "damp". There may or may not be a subsequent recoherence. The predominant decoherence/recoherence occurs through each cycle of synchrotron oscillation due to μ's dependence on δ. In the present discussion, since we will be concerned with times long compared to the synchrotron period, we will average over the longitudinal motion, leaving any surviving tune dependence expressable by the dependency $\mu(\hat{\delta})$.

† For brevity, the phase advance per tune μ will often be referred to as the "tune" even though, technically, that name should be reserved for $\mu/(2\pi)$.

‡ The dimension of p is length.

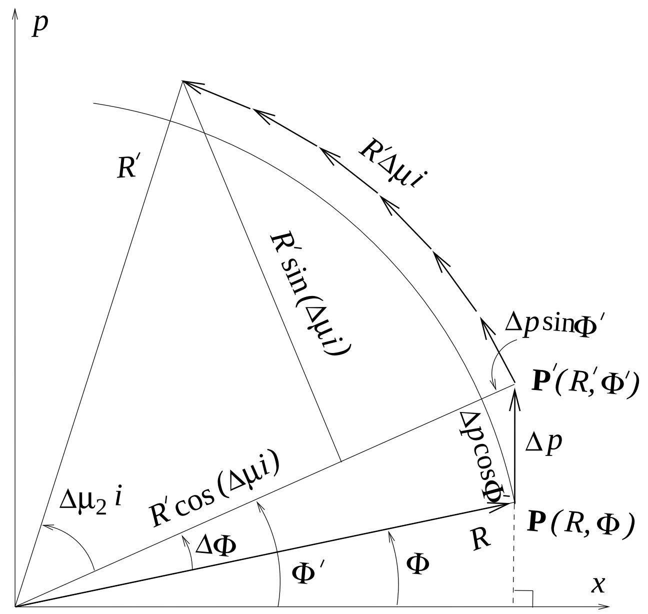

Figure 14.1: Evolution with turn number i of a point $\mathbf{P}$ in betatron phase space, as viewed from a frame rotating at nominal phase advance per turn μ_0. The trigonometry of this figure only makes sense for $\Delta p << R$, which is assumed.

We assume the decoherence is due entirely to the "shearing" motion along circles of different radius in phase space and different values of $\hat{\delta}$, thereby neglecting the fact that, because of nonlinearity at large R, the phase space curves, even while remaining regular become distorted (though not chaotic).

For points close to the origin and having small $\hat{\delta}$ the shear is negligible and the distribution rotates undistorted, as if it were rigid. To take advantage of this, Fig. 14.1 is a snapshot (of the i'th turn) from a frame of reference rotating at rate μ_0. The effect of kick Δp is to change the initial phase space location of point $\mathbf{P}$ to (approximately)

$$R' = R + \Delta R = R + \Delta p \sin\Phi', \quad \Phi' = \Phi + \Delta\Phi = \Phi + \frac{\Delta p \cos\Phi'}{R}. \tag{14.2}$$

After the kick, its tune is $\mu_0 + \Delta\mu(R', \hat{\delta})$, and its positions on subsequent turns are indicated by short arrows in Fig. 14.1. After i turns its coordinates are

$$\begin{pmatrix} x_i\left(\Delta p, R, \Phi, \hat{\delta}\right) \\ p_i\left(\Delta p, R, \Phi, \hat{\delta}\right) \end{pmatrix} = R' \cos\Delta\mu\, i \begin{pmatrix} \cos\Phi' \\ \sin\Phi' \end{pmatrix} + R' \sin\Delta\mu\, i \begin{pmatrix} -\sin\Phi' \\ \cos\Phi' \end{pmatrix}. \tag{14.3}$$

The centroid coordinates are given then by

$$\begin{pmatrix} \overline{x_i}\,(\Delta p) \\ \overline{p_i}\,(\Delta p) \end{pmatrix} = \int_0^{\infty} dR \int_0^{2\pi} d\Phi \int_0^{\infty} d\hat{\delta} \begin{pmatrix} x_i\left(\Delta p, R, \Phi, \hat{\delta}\right) \\ p_i\left(\Delta p, R, \Phi, \hat{\delta}\right) \end{pmatrix} P_{R,\Phi}\,(R,\Phi)\,P_{\hat{\delta}}\left(\hat{\delta}\right). \tag{14.4}$$

These formulas are impractical for calculation because of the complicated dependence of Φ' on position $\mathbf{P}$. Since the trigonometry of Fig. 14.1 breaks down near the origin, we assume

$$\Delta p << R. \tag{14.5}$$

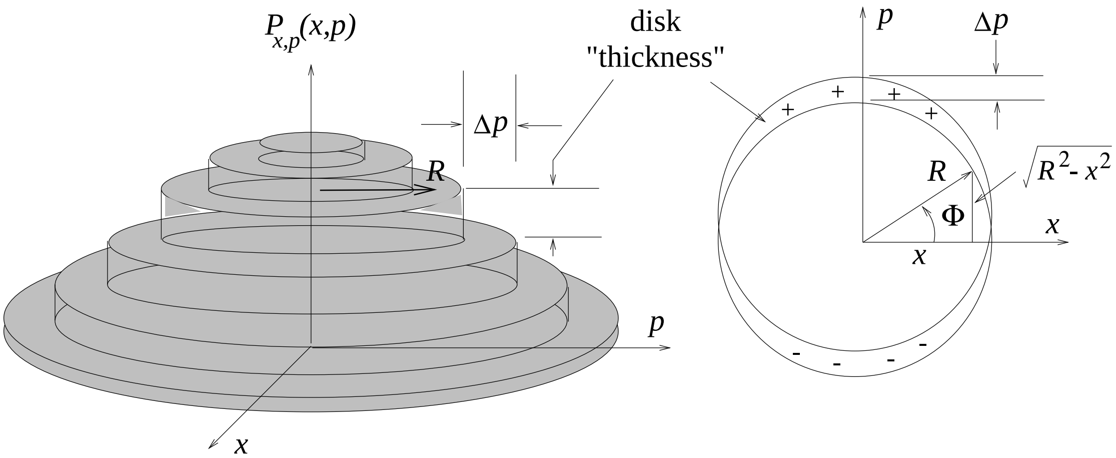

Figure 14.2: On the left the betatron phase space distribution is visualized as a sum of distributions, uniform over disks of radii successively increasing in steps of Δp. This permits deviations from the unkicked distribution to be represented by positive and negative distributions uniform over the "lunes" shown on the right.

In spite of assumption (14.5), it is not legitimate to approximate Φ' by Φ; if this approximation is made, Eqs. (14.3) and (14.4) give a seriously incorrect answer even for $\Delta\mu = 0$ and $i = 0$. This failure is at least partly due to the extravagance of following the evolution of every particle, not taking advantage of the strong tendency for cancellation in pairs of particles symmetric about the origin. To take advantage of this cancellation we follow instead the evolution of *deviations* from the unperturbed distributions as shown in Fig. 14.2. (For the time being we suppress indications of δ dependency from the formulas, since they will be easily restored later.) Since volumes in the plot on the left correspond to probabilities, the units along the vertical axis are length^{-2} and the total "volume" is 1. On the one hand, the volume can be visualized as nested "collars" of inner radius $R - \Delta p/2$, wall thickness Δp, and "height" $P_x(0)P_p(R)$. On the other hand, it can be visualized as a pile of stacked disks of radius $R + \Delta p/2$, with

$$\text{disk "thickness"} = P_x(0)\left(-\frac{dP_p(p)}{dp}\right)_{p=R}\Delta p = \frac{R}{2\pi\sigma_x^4}e^{-\frac{R^2}{2\sigma_x^2}}\Delta p = \frac{P_R(R)\,\Delta p}{2\pi\sigma_x^2}. \tag{14.6}$$

Having units of length^{-2}, this is appropriate as a joint differential probability distribution in the (x,p) plane.

When the beam is displaced by Δp along the p axis most of the probability in any particular one of the stacked disks, for example the one with radius R, can be regarded as unchanged; the entire change can be ascribed to an increase in probability density in the positive-p "lune" shown on the right in Fig. 14.2 and a corresponding reduction in the negative-p lune. (Though the latter probability density is negative the total probability density in the region remains positive.) Since the entire deviation in this region comes from this particular disk and is accounted for by these lunes, and the subsequent shearing motion respects ring boundaries, it is sufficient to work out the subsequent evolution on a ring-by-ring basis. From these distributions the ring centroids will then be found and finally the overall centroid location.

Toward this end the lune (two dimensional) density can be squashed into an angular (one dimensional) distribution. Furthermore the negative lune can be dropped, compensating by doubling the positive-lune probability. With the area of one lune being $2R\Delta p$, the deviation probability it represents is $4R(\Delta p)^2P_R(R)/(2\pi\sigma_x^2)$. Letting $P_R^{\text{dev}}dR$ stand for the deviation probability in range dR we have

$$P_R^{\text{dev}}(R) = \frac{2\Delta p}{\pi\sigma_x^2}RP_R(R), \tag{14.7}$$

which is independent of i. When distributed in x, the just-kicked deviation probability $P_R^{\text{dev}}(R)dR$ is uniform. Therefore, when distributed in Φ, which is related to x by $x = R\cos\Phi$, the distribution is proportional to $dx/d\Phi = R\sin\Phi$. We therefore define a (normalized) angular probability distribution

$$P_{\Phi,0}(\Phi) = \begin{cases} 0 & \text{for } \Phi < 0 \\ (1/2)\sin\Phi & \text{for } 0 < \Phi < \pi \\ 0 & \text{for } \pi < \Phi \end{cases}, \tag{14.8}$$

which is a universal initial angular distribution, independent of R. Then the joint probability distribution $P^{\text{dev}}_{R,\Phi}$ (defined so that $P^{\text{dev}}_{R,\Phi}dRd\Phi$ stands for the deviation probability in range $dRd\Phi$) can be factorized

$$P^{\text{dev}}_{R,\Phi}\left(R,\Phi,i\right) = P^{\text{dev}}_{R}\left(R\right) P_{\Phi}\left(\Phi,R,i\right). \tag{14.9}$$

Initially it is given by

$$P^{\text{dev}}_{R,\Phi}\left(R,\Phi,i=0\right) = P^{\text{dev}}_{R}\left(R\right) P_{\Phi}\left(\Phi,R,i=0\right) = \frac{\Delta p}{\pi\sigma_x^2} R P_R\left(R\right)\sin\Phi. \tag{14.10}$$

Except for the eventual integration over R, all that is required is to evaluate $P_{\Phi}(\Phi,R,i)$ as it evolves away from $P_{\Phi,0}(\Phi)$—a one dimensional calculation. Furthermore the R dependence allowed for notationally by the second argument of $P_{\Phi}(\Phi,R,i)$, will be present only if the betatron motion is nonlinear.

The centroid coordinates are obtained as the averages of $x = R\cos\Phi$ and $p = R\sin\Phi$ weighted by $P^{\text{dev}}_{R,\Phi}(R,\Phi,i)$;

$$\begin{pmatrix} \overline{x_i}\left(\Delta p\right) \\ \overline{p_i}\left(\Delta p\right) \end{pmatrix} = \int_0^{\infty} R\, P^{\text{dev}}_{R}\left(R\right) dR \int d\Phi\ P_{\Phi}\left(\Phi,R,i\right) \begin{pmatrix} \cos\Phi \\ \sin\Phi \end{pmatrix}. \tag{14.11}$$

Here the limits of the Φ integration are not indicated. They can safely be set to $-\infty$ and $+\infty$ since, for finite i, the integrand vanishes exactly outside a finite range. At $i=0$ the non-vanishing range is from 0 to π and for other values of i the range needs to be extended only by $\Delta\mu_{\text{max}} i$ where $\Delta\mu_{\text{max}}$ is the maximum possible tune deviation from nominal.

To check for consistency, let us calculate the $i=0$ centroid location;

$$\overline{p_0}\left(\Delta p\right) = \int_0^{\infty} dR \int_0^{\pi} d\Phi\ P^{\text{dev}}_{R,\Phi}\left(R,\Phi,i=0\right) R\sin\Phi = \Delta p \frac{4}{\pi}\int_0^{\pi}\frac{\sin^2\Phi}{2}d\Phi = \Delta p, \tag{14.12}$$

as expected.

The only dependence on i in Eq. (14.11) is introduced via

$$\Phi_i = \Phi_0 + \Delta\mu\left(R,\hat{\delta}\right)\, i, \tag{14.13}$$

which, for a particle with initial phase Φ_0, gives its phase after i turns. Again we observe that if $\Delta\mu(R,\hat{\delta})$ is constant, independent of R and $\hat{\delta}$, its only effect is to cause the entire distribution to rotate rigidly at a tune shifted by $\Delta\mu$ from the unperturbed tune μ_0. Since, in that case, its effect could have been included in μ_0, we may as well assume that $\Delta\mu(R,\hat{\delta})$ has no part independent of R and $\hat{\delta}$;

$$\Delta\mu\left(R,\hat{\delta}\right)=r_1R+r_2R^2+\cdots+d_1\hat{\delta}+d_2\hat{\delta}^2+\cdots. \tag{14.14}$$

After i turns the distribution originally given by $P_{\Phi,0}$, having precessed through angle $\Delta\mu(R,\hat{\delta})\,i$, becomes $P_{\Phi,i}(\Phi,R,\hat{\delta})=(1/2)\sin(\Phi-\Delta\mu(R,\hat{\delta})\,i)$, (and zero outside the central lobe.) This along with Eq. (14.11) are exact in the small kick limit where approximation (14.5) is valid, and they are simple enough for easy and accurate numerical evaluation, but because of various other uncertainties, great precision is rarely justified. This makes it seem sensible to approximate the angular distribution in a way that will simplify subsequent calculations. Also we take the opportunity to introduce a more convenient angle Θ in terms of which the starting distribution is symmetric about $\Theta=0$;

$$\Theta=\Phi-\frac{\pi}{2}. \tag{14.15}$$

The approximate form to be used is

$$P_{\Theta,i}\left(\Theta,R,\hat{\delta}\right)\approx\frac{1}{\sqrt{2\pi}\sigma_{\text{fit}}}e^{-\frac{\left(\Theta-\Delta\mu(R,\hat{\delta})\,i\right)^2}{2\sigma_{\text{fit}}^2}}. \tag{14.16}$$

This form eliminates the need for the multiple cases of Eq. (14.8) and permits an infinite Φ integration range. The quantity σ_{fit} is simply a dimensionless number (an angle in radians) chosen to make the approximation in Eq. (14.15) as accurate as possible. The value $\sigma_{\text{fit}}=(2\pi)^{-1/6}=0.736$ would match the quadratic variation at $\Phi=\pi/2$, but we choose instead

$$e^{-\frac{\sigma_{fit}}{2}}=\frac{\pi}{4},\quad\text{or}\quad\sigma_{\text{fit}}=0.695, \tag{14.17}$$

which causes Eq. (14.12) to be satisfied, thereby avoiding a (small but inelegant) error in the just-kicked centroid location. Substituting Eq. (14.16) into Eq. (14.11) yields

$$\begin{aligned}
\begin{pmatrix}\overline{x_i}\left(\Delta p,\hat{\delta}\right)\\ \overline{p_i}\left(\Delta p,\hat{\delta}\right)\end{pmatrix} &= \Delta p\int_0^\infty dR\frac{2}{\pi\sigma_x^4}R^3e^{-\frac{R^2}{2\sigma_x^2}}\int_{-\infty}^\infty d\Theta\,\frac{1}{\sqrt{2\pi}\sigma_{\text{fit}}}e^{-\frac{\left(\Theta-\Delta\mu(R,\hat{\delta})\,i\right)^2}{2\sigma_{\text{fit}}^2}}\begin{pmatrix}-\sin\Theta\\ \cos\Theta\end{pmatrix}\\
&= \Delta p\,\frac{2}{\pi\sigma_x^4}\,\frac{1}{\sqrt{2\pi}\sigma_{\text{fit}}}\int_0^\infty dRR^3e^{-\frac{R^2}{2\sigma_x^2}}\begin{pmatrix}-\sin\Delta\mu\,i\\ \cos\Delta\mu\,i\end{pmatrix}\int_{-\infty}^\infty d\Theta\ \cos\Theta\ e^{-\frac{\Theta^2}{2\sigma_{\text{fit}}^2}}\\
&= \Delta p\,\frac{1}{2\sigma_x^4}\int_0^\infty dRR^3e^{-\frac{R^2}{2\sigma_x^2}}\begin{pmatrix}-\sin\Delta\mu\left(R,\hat{\delta}\right)\,i\\ \cos\Delta\mu\left(R,\hat{\delta}\right)\,i\end{pmatrix}
\end{aligned} \tag{14.18}$$

where the dependence on $\hat{\delta}$ has again been acknowledged. This formula, with $\Delta\mu(R,\hat{\delta})$ expressed, for example, as in Eq. (14.14), is the main formula describing the effect of decoherence due to R-dependent tune caused by nonlinear betatron motion. For small i evaluating the integral numerically should be easy. For large i the method of stationary phase may be applicable.[5]

Since there has been no averaging over $\hat{\delta}$ as yet, Eq. (14.18) should also be valid with $\hat{\delta}$ replaced by δ. The major effect of this would be evident in Fig. 14.1 where the phasor amplitudes would vary sinuisoidally because of chromaticity and synchrotron oscillation. Whatever shearing this causes is exactly undone over a complete cycle, causing periodic decoherence/recoherence each period of synchrotron oscillation. By performing these calculations it would be possible to compare to a formula due to Bob Meller[6] but that has not been done. The feature distinguishing the present calculation from his is that he assumed no systematic dependence of tune on $\hat{\delta}$. It is not easy to compare formulas here with that paper since the order of integation is different and he does not make the approximation Eq. (14.16), (which should cause only small numerical differences.)

The calculations have assumed motion in only one transverse plane; it would be possible for the horizontal and vertical decoherence rates to differ. The essential feature though is not the distinction between the two planes but the distinction between two normal mode motions. In this paper, where coupled motion is of the essence, we assme the same calculations are valid when applied mode by mode.

We continue, but now keeping just the term $\Delta\mu(R,\hat{\delta}) = d_1\hat{\delta}$ (which permits the R integration to be performed) and assume that $\hat{\delta}$ is distributed according to

$$P_{\hat{\delta}}\left(\hat{\delta}\right) = \frac{\hat{\delta}}{\sigma_{\delta}^2} e^{-\frac{\hat{\delta}^2}{2\sigma_{\delta}^2}}, \tag{14.19}$$

and average over $\hat{\delta}$ to obtain†

$$\begin{aligned} \frac{\overline{\overline{p_i}}\,(\Delta p)}{\Delta p} &= \frac{1}{\sigma_{\delta}^2}\int_0^{\infty} d\hat{\delta}\;\hat{\delta}\;e^{-\frac{\hat{\delta}^2}{2\sigma_{\delta}^2}}\cos d_1\hat{\delta}\,i \\ &= 1-(d_1\sigma_{\delta}i)^2+\frac{1}{3}(d_1\sigma_{\delta}i)^4-\frac{1}{7.5.3}(d_1\sigma_{\delta}i)^6+\frac{1}{9.7.5.3}(d_1\sigma_{\delta}i)^8+\cdots, \end{aligned} \tag{14.20}$$

† Formula (14.20) is poorly convergent and can only be used for values of the argument less than 2 or so.

and

$$\frac{\overline{\overline{x_i}}\,(\Delta p)}{\Delta p} = -\frac{1}{\sigma_\delta^2}\int_0^\infty d\hat{\delta}\;\hat{\delta}\;e^{-\frac{\hat{\delta}^2}{2\sigma_\delta^2}}\sin d_1\hat{\delta}\,i = -\sqrt{\frac{\pi}{2}}\,(d_1\sigma_\delta i)\,e^{-\frac{(d_1\sigma_\delta i)^2}{2}}. \tag{14.21}$$

Perhaps the most nearly observable quantity is the "decoherence factor"

$$F_i\,(d_1\sigma_\delta) = \sqrt{\left(\frac{\overline{\overline{x_i}}\,(\Delta p)}{\Delta p}\right)^2 + \left(\frac{\overline{\overline{x_i}}\,(\Delta p)}{\Delta p}\right)^2}. \tag{14.22}$$

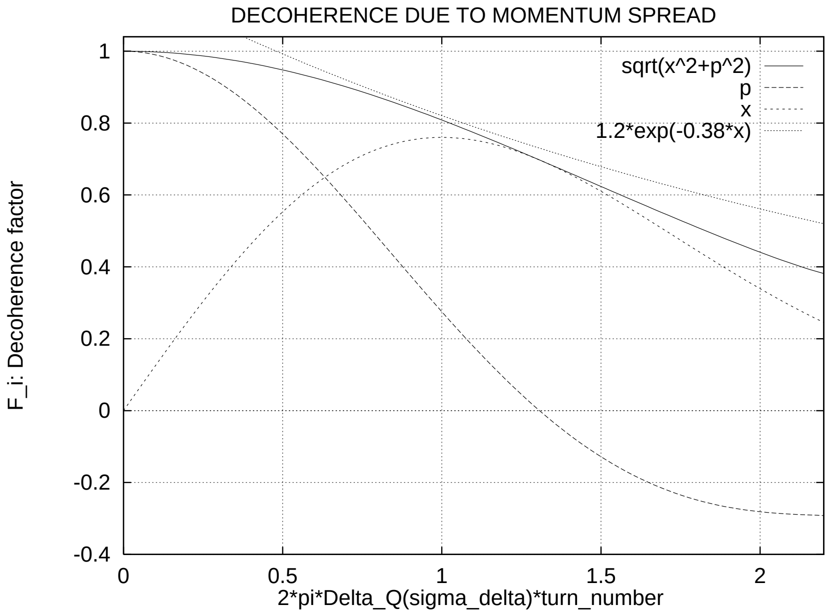

Figure 14.3: Time evolution of (fractional) centroid position $x/\Delta p$, slope $p/\Delta p$, and $\sqrt{x^2+p^2}/\Delta p$ after initial deflection Δp, viewed in a frame of reference rotating at the small amplitude tune, as given by Eqs. (14.20)-(14.22). An exponentially decaying function $1.2e^{-0.38 d_1\sigma_\delta i}$ is also shown for comparison.

These functions are plotted in Fig. 14.3. As has been explained previously, the quantities $\overline{\overline{x_i}}$ and $\overline{\overline{p_i}}$ tend to vary slowly because they refer to a frame of reference rotating at the small amplitude tune μ_0. The corresponding invariant amplitude $\sqrt{\overline{\overline{x_i}}^2 + \overline{\overline{p_i}}^2}$ is presumably even more slowly varying, but when it is viewed in a stationary frame it rotates rapidly and is interpreted as the betatron oscillation of the centroid. Furthermore, its magnitudes in stationary and rotating frames are the same. That was the basis for the statement made above that the "decoherence factor" F_i is the theoretical quantity that can most easily be correlated with experimental observations.[†] From Fig. 14.3 it can be seen that the time evolution of $\overline{\overline{x_i}}$ and $\overline{\overline{p_i}}$ are very different, the latter falls off in more or less Gaussian fashion while the former rises, then falls. Though neither of these behaviors seems deserving of the name "damping", the function F_i falls off more nearly as the decaying exponential that is normally associated with damping. To illustrate this point a pure exponential decay curve that crudely matches F_i is also shown in Fig. 14.3.

When damping rates are measured experimentally the observed response is not a pure exponential decay. Rather, an initial transient (that is hard to interpret and may be instrumental in nature) is followed by a curve well fit by a pure exponential. A recipe for measuring clean "relative" dependence of damping rate on the various parameters has been to select the range over which the log plot is most linear as the signal falls by $1/e$—normally from about 0.8 to about 0.3 of the just-kicked signal. This is not very different from the range over which the exponential described in the previous paragraph gives a tolerable fit to the theoretical response curve. Considering the only-semi-quantitave "absolute" accuracy of the measurements and the lack of accuracy with which the various parameters influencing the phenomenon are known, we therefore judge the exponential fit described in the previous paragraph as a reasonable representation of theory for comparison with data.

Accepting the numerical factor 0.38 from the exponential fit just described, these considerations can be distilled down to a simple prescription for predicting the "damping rate" $\alpha_{\sigma,\text{dec}}$ with which the centroid will be observed to damp after the beam has been pinged. Let us assume that, by particle tracking in the lattice under consideration (CESR in our case) the betatron tune shift from nominal, for a particle with invariant longitudinal invariant equal to the r.m.s. value $\hat{\delta} = \sigma_\delta \overset{\text{e.g.}}{=} 0.6 \times 10^{-3}$ has the value $\Delta Q(\sigma_\delta)$. For one of the CESR lattices on which these effects were investigated $\Delta Q(\sigma_\delta) = \pm 1.5 \times 10^{-3}$, where the sign ambiguity occurs because the tune shifts of the two visible modes are equal and opposite. This sign difference has no effect on the predicted decoherence (unless there are previously neglected tune shifts having the same sign for both modes.) Assuming the tune shift dependence to be linear, this fixes the d_1 coefficient in Eq. (14.14) so that

$$d_1 = \frac{2\pi \Delta Q\,(\sigma_\delta)}{\sigma_\delta}. \tag{14.23}$$

[†] Though more detailed information about the beam is measurable in principle, we are mainly concerned with signals from the beam position monitors; they contain only information about the centroid.

Then the centroid amplitude is proportional to $e^{-\alpha_{\sigma,\mathrm{dec}}t} = e^{-0.38\times 2\pi\times|\Delta Q(\sigma_\delta)|i}$, with time t and turn number i being related by $t = i/f_0$ where $f_0 \stackrel{\text{e.g.}}{=} 0.39\times 10^6\,\mathrm{Hz}$. Then we obtain

$$\boxed{\alpha_{\sigma,\mathrm{dec}} = 2.39|\Delta Q(\sigma_\delta)|f_0 \sim (Q'_x - Q'_y)^2} \qquad (14.24)$$

where the dominant expected dependence on chromaticities has been inferred from section 13. For an extreme situation in which $Q'_x = 10$, $Q'_y = -20$, $\alpha_{\sigma,\mathrm{dec}} \stackrel{\text{e.g.}}{=} 1400\mathrm{s}^{-1}$. For other situations the empirical formulas (13.1) can be used to estimate $|\Delta Q(\sigma_\delta)|$. When the parameters appropriate to Fig. 13.4 are used, the predicted damping rate is $1160\,\mathrm{s}^{-1}$. This is approximately twice that inferred from the multiparticle simulation. Considering all the uncertainties, this is probably as good agreement as can be hoped for.

15. Damping rate formulas collected

Various sources of damping have been identified and estimated. Not even mentioned so far, because it is small and reliably known, is the synchrotron-radiation induced growth rate (a negative quantity) α_{SR}. Unlike decoherence, which only causes the centroid amplitude to decay, without affecting individual particles, α_{SR} is a true dissipative (or incoherent) damping rate to which every particle is subject. For CESR $\alpha_{SR} = -38\,\mathrm{s}^{-1} = 1.0\times 10^{-4}$ per turn.

Since this paper deals only with the "decay" of externally detectable beam signals that depend only on centroid variables, we do not distinguish between coherent and incoherent damping rates. The conjectured "Landau damping" of (unobservable) π-modes is symbolized by $\alpha_{\pi,\mathrm{LD}}$. We have little theoretical guidance considering $\alpha_{\pi,\mathrm{LD}}$ and it is not directly measureable. We can only hope to infer its value indirectly. To account for its ability, independent of current, to stabilize the π-mode in spite of the head-tail destabilization, it is plausible to suppose that $\alpha_{\pi,\mathrm{LD}}$ is proportional to current, and it will be written $\iota\,\alpha'_{\mathrm{LD}}$. This is consistent with ascribing the damping to space charge forces. The damping rate of (observable) σ-modes due to decoherence will be expressed by $\alpha_{\sigma,\mathrm{dec}}$, presumably independent of current. In the only-semi-quantitative description we are attempting, we regard $\alpha_{\sigma,\mathrm{dec}}$ as reasonably well-known, for example from Eq. (14.24) with ΔQ given by

Eq. (13.1).† It will be regarded as appropriate to build all these damping rates into the unperturbed model as coefficients of velocity-proportional terms.

Combining all growth rates discussed so far, the growth rates of the two σ-modes and one of the π-modes can be expressed as

$$\begin{aligned}
\alpha_{\sigma,1} &= \alpha_{\mathrm{SR}} + \alpha_{\mathrm{HT}} + \alpha_{\sigma,\mathrm{dec}} + \alpha_{\sigma,\mathrm{symp}}, \\
\alpha_{\sigma,2} &= \alpha_{\mathrm{SR}} + \alpha_{\mathrm{HT}} + \Delta\alpha + \alpha_{\sigma,\mathrm{dec}} + \alpha_{\sigma,\mathrm{symp}}, \\
\alpha_{\pi} &= \alpha_{\mathrm{SR}} - \alpha_{\mathrm{HT}} + \alpha_{\pi,\mathrm{LD}}.
\end{aligned} \tag{15.1}$$

It has unfortunately been necessary to treat the two π-modes separately because it is found experimentally that their damping rates are different. At present the source of this difference is unknown. In our leading approximation they are equal but, with their tunes split appreciably (by $S = \Sigma/2\pi$) different distances to nearby resonances permits these rates to be different. We add the term $\Delta\alpha$ to allow for this difference, hoping it to be constant, independent of the other parameters. Formulas for these quantities are given in Eqs. (5.15), (6.6), (12.11) and (14.24). Copying from there, introducing fitting parameters A, B, and C, and assuming that $\alpha_{\pi,\mathrm{LD}}$ depends linearly on ι yields

$$\begin{aligned}
\alpha_{HT} &= \iota\, A \left(Q'_x \frac{w_x}{w_y} + Q'_y + \frac{\left(Q'_x - Q'_y\right)\left(\frac{w_x}{w_y} - 1\right)}{2} \, \frac{(\Sigma/2)^2}{{\mu_s}^2 - (\Sigma/2)^2} \right), \\
\alpha_{\sigma,\mathrm{dec}} &= B \left(Q'_x - Q'_y\right)^2, \\
\alpha_{\sigma,\mathrm{symp}} &= \iota\, C \left(1 - e^{-0.0112\,\iota^2/\iota_{SR}^2}\right).
\end{aligned} \tag{15.2}$$

The ratio w_x/w_y and the coefficient B are not entirely free since the former is probably approximately 1 (in one fit it is 0.53) and the latter is calculable to within a factor of perhaps 2, but we treat them as free nevertheless. A, proportional to the wake field, and C, the Landau damping coefficient, are completely empirical, as are $\Delta\alpha$ and ι_{SR}. The measurable mode 1 damping rate, then, is

$$\boxed{\alpha_{\sigma,1} = \ \alpha_{\mathrm{SR}} + B(Q'_x - Q'_y)^2 + \iota[A(Q'_x \tfrac{w_x}{w_y} + Q'_y) + C(1 - e^{-0.0112\,\iota^2/\iota_{SR}^2})].} \tag{15.3}$$

† It should not be considered significant that our notation appears to distinguish betweem Landau damping and decoherence. The physics of these phenomena are certainly similar and, for all we know, identical. In our notation "dec" means "calculable", and "LD" means "purely phenomenological".

The last term of α_{HT} has been dropped as being too small to justify its complicated appearance.

In the small chromaticity region, $Q'_x \approx Q'_y \approx 0$, the chromaticities enter predominantly in the combination $Q'_x\, w_x + Q'_y\, w_y$; this can be said to characterize the phenomenon of "chromaticity sharing". In the large chromaticity region the chromaticities enter primarily through $\alpha_{\sigma,\mathrm{dec}}$, in the combination $(Q'_x - Q'_y)^2$. Absent a better name, this could be called the "chromaticity dominated" region. Another mechanism by which the rates can acquire nonlinear dependence on Q'_x and Q'_y is through "wake-field washout" as described by Eqs. (4.2), but this effect was found to be negligible.

16. Comparison of observations with the model

16.1. Sample data

Results from some of our observations of this phenomenon are plotted in Fig. 16.1 which shows, in the Q'_x, Q'_y plane, measured contours of equal damping rate of the *less stable* (in fact *barely stable*) mode, for two different values of beam current ι. The straight line in Fig. 16.1 is tangent to the contours at their point closest to $Q'_x = Q'_y = 0$. From the slope one infers $w_x/w_y = 0.54$. The damping rates of the *more stable* mode at the same values of Q'_x and Q'_y are plotted in Fig. 16.2.

16.2. Matching the semi-empirical formula to observations

Theoretical contours to be compared with the observations at CESR are shown in Fig. 16.3, Fig. 16.4, and Fig. 16.5. Though the closest fitting contour of Fig. 16.3 qualitatively resembles the measured 4 mA contour of Fig. 16.1, it lacks the near-straight-line segment in the low chromaticity region. When decoherence was discussed in an earlier section it was acknowledged that the simulation data was consistent with a higher than quadratic dependence on $|Q'_X - Q'_y|$. In Fig. 16.4 a $|Q'_X - Q'_y|^3$ dependence is tried, using the parameters shown in the caption. Since the best fit contour is in much better agreement with the data, we accept the cubic dependence.

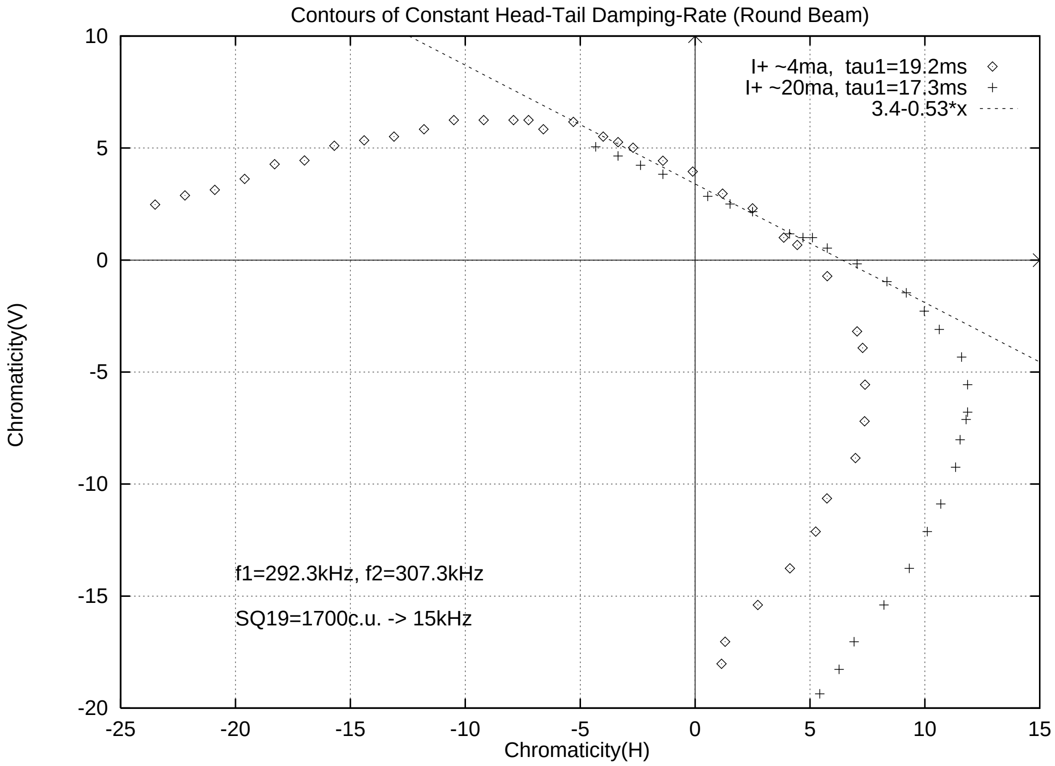

Figure 16.1: Measured contour of equal (barely stable) growth rate on the Q'_x, Q'_y plane for two values of beam current ι. Since the damping times are 19.2 ms and 17.3 ms, the growth rates are $-52\,s^{-1}$ and $-57.8\,s^{-1}$.

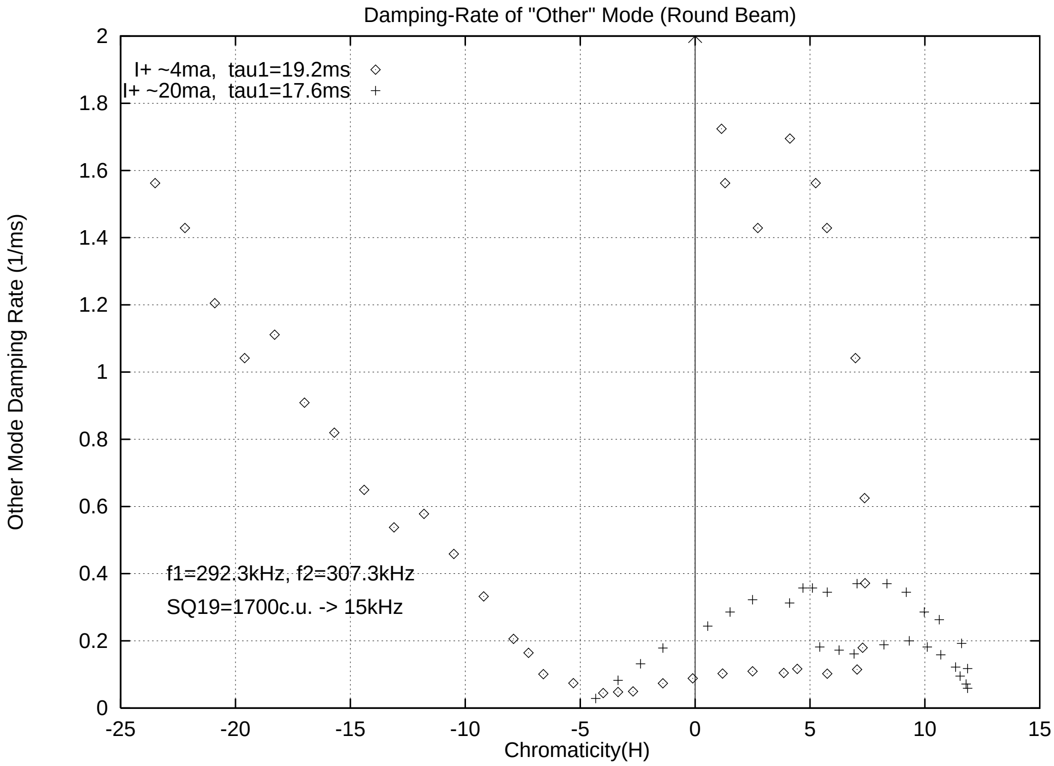

Figure 16.2: Measured growth rates of the "other" mode, whose growth rate was *not* held constant while the data of Fig. 16.1 was taken. The points in this plot match the points in Fig. 16.1, making it possible to infer the (projected-out) coordinate Q'_y of every point. Note that full scale (vertically) on this graph is $2000\,s^{-1}$, or about 40 times the (held constant) damping rates of the previous plot.

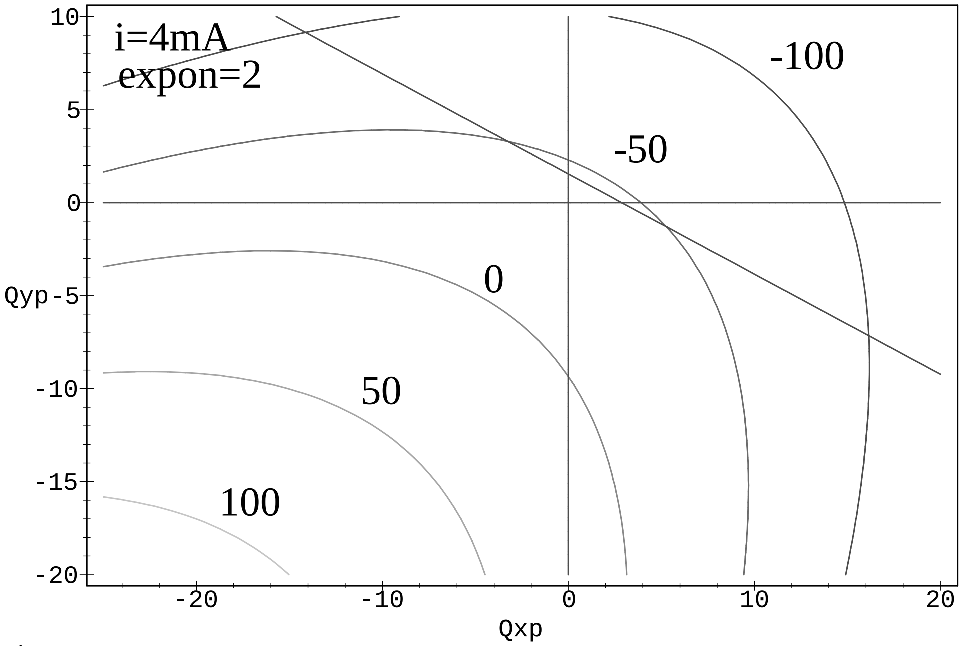

Figure 16.3: Theoretical contours of constant damping rate from Eq. (15.3), for beam current $\iota = 4\,\text{ma}$. Parameters have been adjusted to match the data of Fig. 16.1: $\alpha_{\text{SR}} = -38\text{s}^{-1}$, $B = -0.1\text{s}^{-1}$, $A = -5.0\text{s}^{-1}/\text{mA}$, $w_x/w_y = 0.54$, $C = 0$. The decoherence term is assumed to be proportional to $|Q'_X - Q'_y|^2$. The diagonal line is the same as in that figure (and in subsequent figures as well).

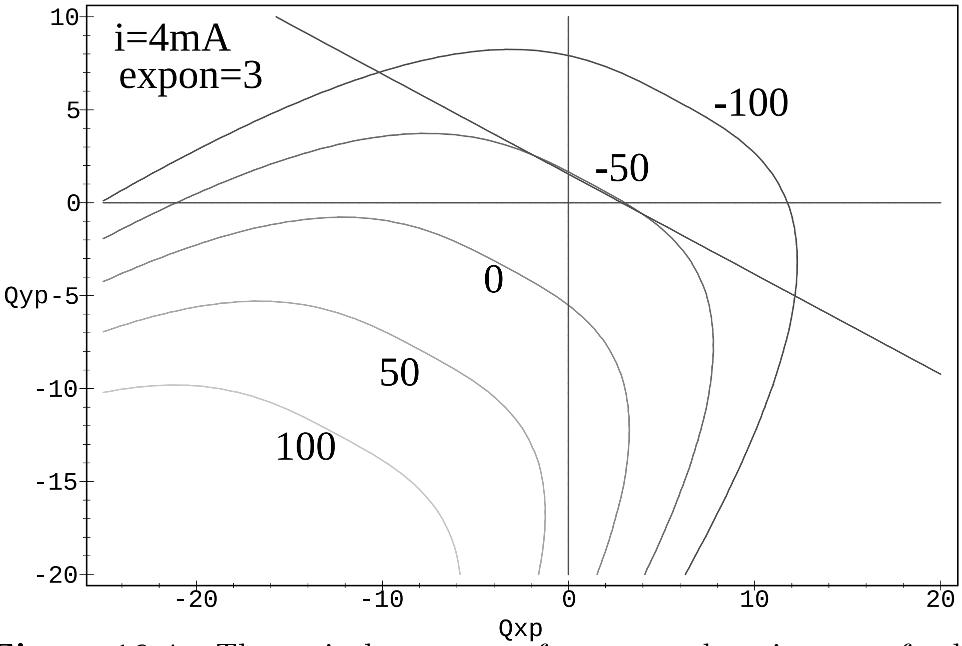

Figure 16.4: Theoretical contours of constant damping rate for beam current $\iota = 4\,\text{ma}$. Parameters have been adjusted to match the data of Fig. 16.1: $\alpha_{\text{SR}} = -38\,\text{s}^{-1}$, $B = -0.01\,\text{s}^{-1}$, $A = -1.8\,\text{s}^{-1}/\text{mA}$, $w_x/w_y = 0.54$, $C = 2.0\,\text{s}^{-1}/\text{mA}$, $\iota_{\text{SR}} = 10\,\text{mA}$. The decoherence term is assumed to be proportional to $|Q'_X - Q'_y|^3$.

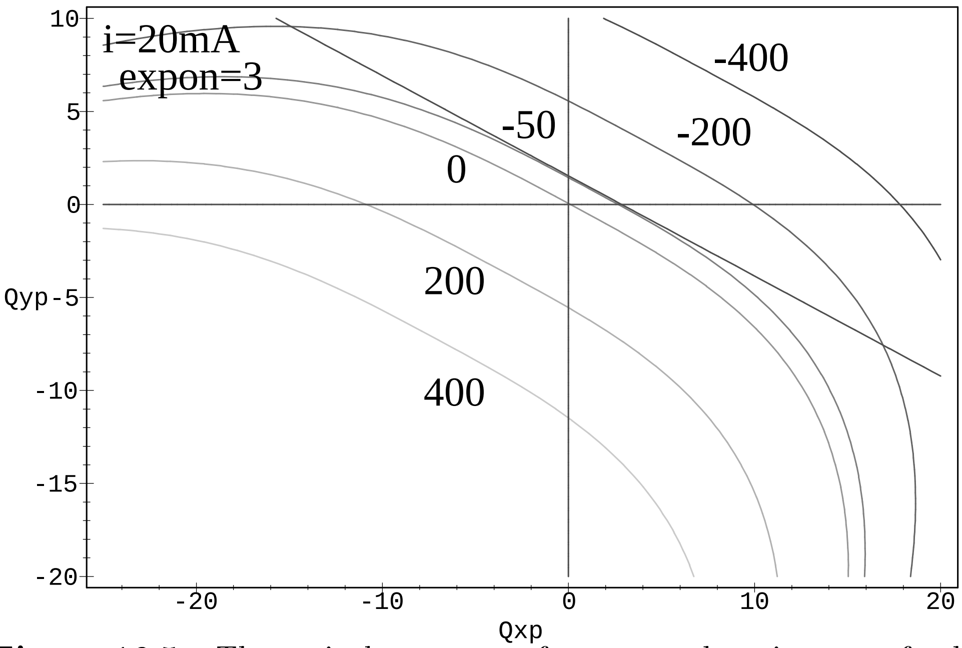

Figure 16.5: Theoretical contours of constant damping rate for beam current $\iota = 20\,\text{ma}$. All other parameters are the same as in Fig. 16.4. The decoherence term is assumed to be proportional to $|Q'_X - Q'_y|^3$.

17. Conclusions

In spite of its considerable complexity, Eq. (15.3) is still not general enough to match the *other mode* data of Fig. 16.2. The natural way to get such extremely strong damping is via the decoherence term $B|Q'_x - Q'_y|^3$ and we have no basis for expecting that to be strongly different between the σ-modes. The extreme current dependence visible in Fig. 16.2 can only be due to the last term of Eq. (15.3), which is however also the same for both σ-modes. The data suggests that either the head-tail damping or the decoherence (or both) are strongly different for the two visible modes. The only-marginal stability at high current (20 mA) suggests that the wake field anti-damping is beginning to overcome the decoherence damping.

Returning to the data of Fig. 16.1, the good agreement between experiment and theory persuasively supports the validity of chromaticity sharing (because of the straight line segments) and decoherence (because of the "rolling over" of the curves.) The support for "sympathetic damping" is far more conjectural; it is based on fitting the observed current dependence in the low chromaticity region, which could be fortuitous, given that the semi-empirical formula has so many free parameters. Nevertheless, it is scarcely revolutionary to suggest that a weakly damped system beomes more strongly damped when coupled to a strongly damped system. In the present context the implication is that for sufficiently large currents ($\iota > \iota_{\rm SR} \approx 10\,$mA) the effective intermixing of modes causes amelioration of the head-tail effect because of the superposition of opposite-sign rates of individual modes.

Perhaps the most striking feature of the data is the extremely strong decoherence damping that can develop with strong chromaticities. This has the potential for improving operations by stabilizing coherent resonances. (Incoherent resonances are of course unaffected.) Concerning the exploitation of this in round beam operations, it remains to be seen whether the decoherence is even nearly as strong in the Möbius lattice as it is in the resonant-coupled lattice analysed in this paper—in the resonant-coupled lattice the decoherence is dominated by the nonlinear tune dependence forced by the near-equal tunes, an effect probably absent in the Möbius lattice. On the other hand, the Möbius lattice may tolerate extremely high chromaticities.

18. References

1. M. Sands, SLAC-TN-69-8, C. Pellegrini, Nuovo Cim. **64A**, 477 (1969).
2. R. Talman, NIM, **193**, 423 (1982), R.D. Kohaupt, DESY 80/22 and DESY M-80/19, (1980).
3. V.A. Yakubovich and V.M. Starzhinskii, *Linear Differential Equations With Periodic Coefficients*, Halsted Press, John Wiley and Sons, 1975.
4. R. Littauer, *Chromaticity in a Coupled Lattice*, CON97-06, June, 1997.
5. Leonard Mandel and Emil Wolf, *Optical Coherence and Quantum Optics*, Cambridge, 1995, p. 128.
6. R. Meller, A. Chao, J. Peterson, S. Peggs, and M. Furman, *Decoherence of Kicked Beams*, SSC-N-360 (1987).

Table of Contents