Nuclear Instruments and Methods in Physics Research A 640 (2011) 44–48

Contents lists available at ScienceDirect

Nuclear Instruments and Methods in Physics Research A

journal homepage: www.elsevier.com/locate/nima

Robinson's radiation damping sum rule: Reaffirmation and extension

S.R. Mane *

Convergent Computing Inc., P.O. Box 561, Shoreham, NY 11786, USA

ARTICLE INFO

Article history:
Received 17 January 2011
Received in revised form
20 March 2011
Accepted 20 March 2011
Available online 31 March 2011

Keywords:
Radiation damping
Synchrotron radiation
Storage rings
Circular accelerators

ABSTRACT

Robinson's radiation damping sum rule is one of the classic theorems of accelerator physics. Recently Orlov has claimed to find serious flaws in Robinson's proof of his sum rule. In view of the importance of the subject, I have independently examined the derivation of the Robinson radiation damping sum rule. Orlov's criticisms are without merit: I work through Robinson's derivation and demonstrate that Orlov's criticisms violate well-established mathematical theorems and are hence not valid. I also show that Robinson's derivation, and his damping sum rule, is valid in a larger domain than that treated by Robinson himself: Robinson derived his sum rule under the approximation of a small damping rate, but I show that Robinson's sum rule applies to arbitrary damping rates. I also display more concise derivations of the sum rule using matrix differential equations. I also show that Robinson's sum rule is valid in the vicinity of a parametric resonance.

1. Introduction

Robinson's celebrated radiation damping sum rule [1] from 1958 is one of the classic theorems of accelerator physics. It is by now standard material in textbooks and summer school lectures, e.g. see Ref. [2]. However, Orlov has published a review [3] of Robinson's sum rule [1] as well as Orlov's own contemporary work (joint with Tarasov) [4], and Orlov claims to find "serious flaws" in Robinson's proof and calculations. In view of the importance of the subject, I have independently examined the derivation of the Robinson radiation damping sum rule, and Orlov's principal criticisms. I show that Orlov's criticisms are without merit.

I work through Robinson's derivation below and I demonstrate that Orlov's criticisms (for example, Orlov's statements on matrix determinants) violate well-established mathematical theorems and are hence not valid. Robinson's proof is elegant and the fundamental ideas underlying his derivation are correct.

In passing, I also show that Robinson's damping sum rule is valid under more general circumstances than those treated by Robinson himself. For example, Robinson derived his sum rule under the approximation of a small damping rate. I show that such an approximation is unnecessary, and that Robinson's damping sum rule, and his methodology of proof, are valid for arbitrary damping rates. Furthermore, Orlov criticizes Robinson's derivation as being inapplicable in the vicinity of a parametric resonance (see Example 2 in Ref. [3]). Robinson did not explicitly discuss parametric resonances in Ref. [1], but I show that a parametric resonance, no matter how strong, is *non-dissipative* in the long term, and makes *no* contribution to the sum of the damping decrements, and hence does not affect the Robinson sum rule. I show that Orlov's statements on parametric resonances violate well known mathematical theorems such as Abel's identity [5].

In the rest of this paper, I reproduce the basic derivation of the radiation damping sum rule, following Robinson's ideas [1]. I also offer more concise derivations of the damping sum rule using matrix differential equations and the Liouville–Ostrogradsky formula [6,7]. (Robinson himself did not employ a matrix differential equation.) I then analyze Orlov's criticisms of Robinson's proof, and point out the flaws in those criticisms. For example, I work through Orlov's Examples 1 and 2 (see Ref. [3]), which are offered as disproofs of Robinson's derivation, and demonstrate the flaws in both examples.

2. Recapitulation of Robinson's derivation

I recapitulate Robinson's derivation [1] of his sum rule. Robinson treated a linear dynamical system and employed a transfer matrix formalism. The sum of the damping decrements was obtained from the eigenvalues of the one-turn matrix by calculating the determinant of that matrix. I summarize the principal steps of Robinson's calculation below. (All references to page numbers and paragraphs are from Ref. [1]). First, as stated by Robinson (p. 374 para. 3) "The sixth order transfer matrix for the infinitesimal element will have infinitesimal non-diagonal terms which are first order in the length of the element, and the

* Tel.: +1 631 821 8389.
E-mail addresses: srmane001@gmail.com, srmane@optonline.net

0168-9002/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.nima.2011.03.037

diagonal terms will differ from unity by a quantity which is proportional to the infinitesimal length of the element."[1] Returning to Robinson's derivation, later in the same paragraph he states: "Consider an element of the accelerator of infinitesimal length. …The determinant of the transfer matrix is given by $1+\sum \delta_{nn}$, where δ_{nn} are the differences of the diagonal terms from unity." Then after Eq. (5) in Ref. [1], "The determinant of the transfer matrix for one complete period is the product of the transfer matrices of the infinitesimal elements of that period." (This is a slight typographical error, it should say "…the product *of the determinants* of the transfer matrices of the infinitesimal elements …" since the overall determinant is a product of individual determinants, not a product of matrices.) Then after Eq. (6) in Ref. [1], "The characteristics of the principal modes of oscillation are determined by solving for the principal values of the transfer matrix for one complete period."

We can express the above statements mathematically as follows. As we see above, Robinson used the term "infinitesimal" and his presentation was rather terse. I proceed in more detail and consider a transfer matrix over a small element of arc-length δs. I say the magnitude of δs is very small but finite. I take the limit $\delta s \to 0$ at the end. For later use, to rebut Orlov's criticisms, I shall retain terms through order $O((\delta s)^2)$ below. Hence I write that the transfer matrix $m(s)$ over a small interval δs has the form:

$$m(s) = I + A(s)\delta s + B(s)(\delta s)^2 + \dots. \tag{2.1}$$

Here I is the unit 6×6 matrix. The determinant of $m(s)$ is

$$\det[m(s)] = 1 + \mathrm{tr}(A(s))\delta s + \mathcal{B}(s)(\delta s)^2 + \dots \equiv 1 + \mathcal{T}(s)\,\delta s. \tag{2.2}$$

The above expression defines $\mathcal{T}(s) = \mathrm{tr}(A(s)) + \mathcal{B}(s)\delta s + \dots$. It is well known that only the trace of $A(s)$ contributes to the first order term of $O(\delta s)$. All contributions from B and the off-diagonal terms in A (as well as higher-order products involving diagonal terms in A) appear only in the term $\mathcal{B}$ (and higher-order terms). Set $\delta s = L/n$, where the ring circumference is L and $n \gg 1$ is a large positive integer. The central mathematical formula in this context is (here $s_i = i\,\delta s$):

$$\lim_{n\to\infty} \prod_{i=1}^{n} \det[m(s_i)] = \exp\left\{\int_0^L \mathrm{tr}(A(s))\,ds\right\}. \tag{2.3}$$

I derive this in more detail as follows, with bookkeeping of higher-order terms. Define the matrix product $M_n = \prod_{i=1}^{n} m(s_i)$. The product of the determinants is

$$\det[M_n] = \prod_{i=1}^{n}[1+\mathcal{T}(s_i)\delta s] = 1 + \sum_{i=1}^{n}\mathcal{T}(s_i)\delta s + \sum_{i=1}^{n}\sum_{j=1}^{i-1}\mathcal{T}(s_i)\mathcal{T}(s_j)(\delta s)^2 + \sum_{i=1}^{n}\sum_{j=1}^{i-1}\sum_{k=1}^{j-1}\mathcal{T}(s_i)\mathcal{T}(s_j)\mathcal{T}(s_k)(\delta s)^3 + \dots \tag{2.4}$$

We obtain a systematic bookkeeping by ordering the terms $i > j > k > \dots$. Now take the limit $n\to\infty$ (hence $\delta s \to 0$) and use the definition of the Riemann integral to express the limits. Denote the one-turn matrix by $M(L) = \lim_{n\to\infty} M_n$ and its determinant by $D(L) = \det[M(L)]$. Then, for example,

$$\lim_{n\to\infty}\sum_{i=1}^{n}\mathcal{T}(s_i)\delta s = \lim_{n\to\infty}\left\{\left[\sum_{i=1}^{n}\mathrm{tr}\,A(s_i)\delta s\right] + O(L\delta s)\right\} = \int_0^L \mathrm{tr}(A(s))\,ds. \tag{2.5}$$

The contribution of the term $\mathcal{B}(s)\delta s$ (and higher powers in δs) in $\mathcal{T}(s)$ *vanishes* in the limit $\delta s \to 0$. This includes the contributions from all off-diagonal elements in $A(s)$ (contrary to claims by Orlov [3] that such off-diagonal terms can make significant contributions). A similar analysis applies to all the other sums in Eq. (2.4). The overall limit is a sum of nested integrals:

$$D(L) = \lim_{n\to\infty}\det[M_n] = 1 + \int_0^L \mathrm{tr}(A(s))\,ds + \int_0^L du\,\mathrm{tr}(A(u))\int_0^u dv\,\mathrm{tr}(A(v)) + \int_0^L du\,\mathrm{tr}(A(u))\int_0^u dv\,\mathrm{tr}(A(v))\int_0^v dw\,\mathrm{tr}(A(w)) + \dots \tag{2.6}$$

Then I use the following identity, for k nested integrals of commuting variables:

$$\underbrace{\int_0^L du\,\mathrm{tr}(A(u))\int_0^u dv\,\mathrm{tr}(A(v))\int_0^v dw\,\mathrm{tr}(A(w))\dots}_{k\ \text{traces}} = \frac{1}{k!}\left(\int_0^L \mathrm{tr}(A(s))\,ds\right)^k. \tag{2.7}$$

Basically, there are $k!$ permutations of the orderings of the traces, and the integral over each permutation is equal, and the sum over all permutations spans the k-dimensional hypercube $[0,L]^k$. For brevity define $\mathcal{F} = \int_0^L \mathrm{tr}(A(s))\,ds$. Hence the overall limit is an exponential series in powers of $\mathcal{F}$. As stated above, the answer is

$$D(L) = \sum_{k=0}^{\infty}\frac{\mathcal{F}^k}{k!} = \exp\left\{\int_0^L \mathrm{tr}(A(s))\,ds\right\}. \tag{2.8}$$

Suppose that we make the *further* (and, in fact, unnecessary) approximation of a small damping rate, so we say that $\mathcal{F}$ is a small quantity. Then we can sum the series in Eq. (2.6) approximately, dropping all terms in powers of $\mathcal{F}$ beyond the first, to obtain the approximate result:

$$D(L) \simeq 1 + \int_0^L \mathrm{tr}(A(s))\,ds. \tag{2.9}$$

This is (equivalent to) the result Robinson actually wrote, as Eq. (6) in Ref. [1], Robinson stated, before Eq. (6) in Ref. [1], "The determinant of the transfer matrix for one complete period is the product of the transfer matrices of the infinitesimal elements of that period. Since the fractional radiation loss in one period is very small, only first order terms need be considered and the determinant of the transfer matrix is given by …" Let us be clear about the usage of the term "first order" in this context: here "first order" means the first power of the integral $\mathcal{F}$ (using my notation), which Robinson showed to be proportional to the radiated energy per turn. This is not the same as "first order" in connection with the elements of a "sixth order" transfer matrix of an "infinitesimal element."

Following Robinson's notation [1], we can write the eigenvalues of the one-turn matrix $M(L)$ as $e^{\gamma_i'}$ where $\gamma_i' = \alpha_i' \pm i\beta_i'$ come in three complex conjugate pairs.[2] Then the determinant of the one-turn matrix is

$$D(L) = \prod_{i=\pm1,\pm2,\pm3} e^{\gamma_i'} = \exp\left\{\sum_{i=1}^{3} 2\alpha_i'\right\}. \tag{2.10}$$

Hence the sum rule for the damping decrements is

$$\sum_{i=1}^{3}\alpha_i' = \frac{1}{2}\int_0^L \mathrm{tr}(A(s))\,ds. \tag{2.11}$$

[1] Note that Robinson was rather casual in his use of the term "order." I clarify his usage in this paper. Hence "sixth order transfer matrix" means a six-dimensional transfer matrix, but later in the same sentence "first order in the length of the element" means "proportional to the first power in the length of the element." I shall clarify another of Robinson's usages of "order" later in this paper.

[2] Robinson's parameterization assumes that, in the absence of damping, the eigenvalues of the one-turn matrix lie on the unit circle. However the ideas expressed above are applicable even if the undamped one-turn matrix is an arbitrary symplectic matrix. For a general classification of the structure of symplectic matrices, see e.g. Dragt [8]. In the general case it is better to work directly with the fractional rate of change of the determinant $(dD/ds)/D$, averaged over one turn.