**AXIAL CHANNELING IN PERFECT CRYSTALS,
THE CONTINUUM MODEL AND THE METHOD
OF AVERAGING**

By

H.S. Dumas

J.A. Ellison

and

A.W. Sáenz

IMA Preprint Series # 706

September 1990

AXIAL CHANNELING IN PERFECT CRYSTALS, THE CONTINUUM MODEL AND THE METHOD OF AVERAGING

H.S. DUMAS*† J.A. ELLISON*‡ AND A.W. SÁENZ§

Abstract. We present a mathematically rigorous treatment of axial channeling motions of energetic, positively charged particles based on the classical relativistic perfect crystal model. More specifically, we reduce the study of motions in a six-dimensional phase space to the study of associated motions in a four-dimensional space. Our main mathematical tool is a recently improved version of the classical method of averaging for ordinary differential equations, which we discuss separately. Applying the method at orders one, two, and three, we extract successively better approximations to perfect crystal model motions (the time of validity is the same for each approximation and scales inversely with the square root of incident particle energy). We call these approximations first-, second-, and third-order continuum model solutions, respectively, and our first-order continuum model solutions are precisely those arising in Lindhard's classical continuum model for axial channeling. The second-order solutions introduce the effects of lattice periodicity in what we believe to be the simplest possible way, e.g., in a computationally simpler way than the standard constant longitudinal momentum approximation. The third-order continuum model solutions introduce the first nontrivial relativistic corrections, and are also, we believe, the most precise approximations which may be useful in applications. After introducing these approximations and discussing their use, we conclude by briefly discussing the present state of the mathematics of channeling.

Key Words. channeling, continuum model, averaging methods, perturbation theory

*This research was supported by the National Science Foundation under grants DMR-8214301 and 8704348

†Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455

‡Department of Mathematics, University of New Mexico, Albuquerque, NM 87131

§Naval Research Laboratory, Washington, DC 20375 and Department of Physics, Catholic University, Washington, DC 20064

1 Introduction

Particle channeling in crystals occurs when an energetic beam of charged particles is incident upon a crystal in a direction nearly aligned with a major crystal axis. When particles enter a crystal in a random (non-aligned) direction, their basic motion may be viewed as multiple scattering off lattice atoms; but in the channeling case, the crystal can be viewed as a collection of atomic strings parallel to the crystal axis and the primary particle motion consists of multiple scattering off these atomic strings. The interaction with each string is described by a gentle steering process involving many correlated small-angle scatterings between particles and individual lattice atoms in the string. Positive particles are pushed away from strings, and close encounter processes may be reduced by two orders of magnitude. Lindhard [1] quantified the need for near-alignment by arguing that close collision processes will be sharply reduced if the incident angle ψ between the beam and the axis satisfies

$$\psi \lesssim \psi_1 = (4Z_1Z_2e^2/dmv_0^2)^{1/2}, \quad (1.1)$$

where Z_1 and Z_2 are the projectile and lattice atom atomic numbers respectively, m is the projectile mass and v_0 its initial velocity, d is the periodic lattice spacing along atomic strings and $e^2 = 14.4\text{eV} - \text{\AA}$ is the square of the electronic charge. For m replaced by $m\gamma_0$ (where γ_0 is the incident Lorentz factor), this relation has been verified experimentally for numerous particles and energies ranging from keV to nearly TeV .

In the axial channeling case, the term “channeling” is somewhat of a misnomer because most particles do not stay in a “channel” but instead move rather randomly from string to string as mentioned above. However, a small fraction of particles may be trapped in a channel formed by nearest neighbor strings and these particles are said to be “proper channeled.” Furthermore, if a beam is aligned with a major crystal plane away from a major axis, particles may be trapped in an oscillatory motion between adjacent planes and this is called “planar channeling.”

Historically, channeling has its roots in the study of particle motions in solids, a central theme in physics and a subject to which Niels Bohr returned again and again. It was J. Lindhard, Bohr's collaborator and copioneer in the theory of ion-solid interactions, who in 1965 laid the theoretical foundations for a statistical treatment of channeling in his paper "Influence of Crystal Lattice on the Motion of Energetic Charged Particles" [1]. Lindhard's work introduced the so-called axial and planar continuum models on which essentially all later theoretical investigations of channeling have relied. These continuum models are based on continuum potentials obtained by averaging the lattice potential over atomic strings or planes. Continuum potentials were also introduced by Lehmann and Leibfried [2], Nelson and Thompson [3], and Erginsoy [4]; however, it was Lindhard who first gave convincing and detailed arguments concerning the validity of the axial continuum model. The goal of this article is to present a rigorous mathematical discussion of the axial continuum model for positive particles and its perturbative corrections in the classical case, using a recent version of higher-order averaging theory for periodic systems of ordinary differential equations [5]. Planar channeling and negative particle channeling are also important, but will not be discussed here.

We begin with the relativistic classical perfect crystal model and show that, when an energetic positively charged particle enters a model crystal away from a string in near-alignment with a major axis, it follows a trajectory which is closely approximated by a continuum model trajectory; i.e., a trajectory governed by the crystal potential averaged along the axis. We also give corrections to the continuum model trajectories, along with rigorous estimates of the time intervals on which the corrected trajectories are close to the exact trajectories of the perfect crystal model.

The results of this paper greatly generalize and complete our earlier work using averaging methods in axial channeling theory [6,7], and they serve purposes different from those of the more comprehensive and qualitative theory of particle motion in crystals developed in [8]. Namely, we give here approximate particle trajectories which are explicitly calculable in

terms of continuum-type model trajectories, and yet at the same time are very close to a well-defined class of perfect crystal model trajectories, including the effects of lattice periodicity in the direction of the axis. These effects are difficult to detect experimentally because real crystals are complicated by lattice vibrations and electron multiple scattering; although we do not address these complications here, our results give periodicity corrections which may be checked in specific situations. We stress that these corrections are easy to compute, because they arise as simple transformations of solutions to an *autonomous* differential equation. Furthermore, and most important, the present results provide a mathematical understanding of perfect crystal motions and so constitute a first step in understanding perturbations of these motions.

The reader unfamiliar with the physics of channeling or its numerous applications is urged to consult the literature. Lindhard's seminal paper [1] has been expanded in a set of Aarhus lecture notes which are available [9]. A review article by Gemmel [10] gives general discussions of particle channeling and applications and an extensive bibliography of the literature up to 1974. Five books have appeared more recently. A book representative of Kumakov's contributions to channeling, dechanneling and channeling radiation, which was published in Russian in 1980, has recently appeared in English [11]. Feldman, Mayer and Picraux [12] review channeling as a materials science tool, and their book contains an extensive bibliography of low-energy channeling through 1982. Ohtsuki [13] discusses several theoretical aspects of channeling and includes a detailed analysis of dechanneling. Sáenz and Überall [14] have edited a collection of articles and Kumakhov and Komarov [15] have written a book which together give a detailed review of the theoretical and experimental status of channeling radiation and coherent bremsstrahlung. Results of a recent workshop on relativistic channeling are presented in Ref. [16]. Proceedings from International Conferences on Atomic Collisions in Solids (1979, 1981, 1983, 1985, 1987, 1989) contain many articles related to particle channeling [17], and more recently, articles for a general audience have appeared [18].

The organization of this paper is as follows. Axial channeling in perfect crystals, i.e., crystals modeled as periodic external fields (no thermal vibrations or electron multiple scattering), is considered in Section 2. In that section, using isoenergetic reduction and scaling, the relativistic system of 3 degrees of freedom described by the perfect-crystal Hamiltonian is converted to a system of four first-order ordinary differential equations (ODE's), which is in a standard form for the method of averaging. Here the independent variable is depth into the crystal and the dependent variables are the transverse coordinates and momenta. In an unexpected way, the corresponding nonrelativistic system can be viewed as a special case of the relativistic system. In Section 3, we present an improved third-order version of the Bogoliubov-Mitropolski theory of averaging for ODE's; this is the order needed for the channeling application discussed in Section 4. Section 3 is largely complete in itself, so the reader need not have a detailed knowledge of averaging theory to understand the discussions there. In fact, it is independent of the rest of the paper and is designed as a separate state-of-the-art tutorial on the high-order averaging method. The justification of the continuum model and the calculation of its perturbative corrections is carried out in Section 4, using the relativistic scaled system of ODE's derived in Section 2 and the averaging results obtained in Section 3. We then briefly discuss how these corrections are used, show how to incorporate the time evolution, and show how our results simplify and clarify a commonly used approximation which includes lattice periodicity by assuming a constant longitudinal velocity. Finally, in Section 5 we survey the present state of the mathematical theory of channeling as described by averaging methods and by the classical normal form theory developed by Dumas [8]; we also point to results on the analogous quantum problem in the spirit of this paper and to the need for mathematical results encompassing more complex aspects of channeling (e.g., the effects of lattice vibrations and electron multiple scattering).

2 Axial Channeling in Perfect Crystals

We begin our study of classical axial channeling for positive particles by way of the single-particle relativistic Hamiltonian

$$\begin{aligned} H(p, q) &= (c^2[p_x^2 + p_y^2 + p_z^2] + m^2 c^4)^{1/2} + U(x, y, z) \\ &= m\gamma c^2 + U(x, y, z), \end{aligned} \quad (2.1)$$

where U is a sum of screened Coulomb or thermally averaged screened Coulomb potentials centered on the perfect crystal lattice sites and the second equality defines the Lorentz factor γ . The nonrelativistic case where

$$H(p, q) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + U(x, y, z) \quad (2.2)$$

will be a special case of our analysis. Here the minimum of U is taken to be zero (U is bounded below because the charged particles are positive), the Cartesian coordinate frame (x, y, z) is chosen so that the z -axis is aligned with the axial channeling direction under study, and U is a periodic function of z of period d (see Eq. (1.1)). This direction is a low-order crystallographic axis, where the precise meaning of the term “low-order” is explained in Section 4 (the $\langle 100 \rangle$, $\langle 110 \rangle$, and $\langle 111 \rangle$ axes in a diamond crystal structure are examples of low-order axes, see e.g. [10]). The time for these Hamiltonians will be denoted by τ .

The axial continuum model mentioned in the introduction is obtained by replacing U in (2.1) and (2.2) by its z -average. Another approximation calculates the x and y motion by assuming the z -velocity, v , is constant and replacing the z -coordinate by $v\tau$. These approximations are discussed in Section 4.

Using the *repulsive* nature of the potential it is easy to show that solutions of (2.1) exist for all time. If we assume provisionally that axial channeling motions satisfy $p_z \gg |p_\perp|$ ($p_\perp = (p_x, p_y)$), then the z -coordinate of a channeling motion increases monotonically with time

and may be used in place of time as the independent variable for the equations of motion defined by (2.1). In fact, the existence of motions defined by (2.1) with $p_z \gg |p_\perp|$ for $O(\varepsilon^{-1})$ time intervals will be a result of our analysis. Here $\varepsilon^2 \approx 2K/m\gamma_0 v_0^2$, where K is a typical maximum value of the potential for channeling, v_0 is the magnitude of the incident particle velocity and γ_0 the associated Lorentz factor. Restricting attention to solutions of (2.1) with fixed energy $E = H(p, q)$ and using the equation of motion for z ($\dot{z} = \frac{1}{m\gamma} p_z$), we find that the new equations of motion with independent variable z are the canonical equations generated by the Hamiltonian

$$H_1(p_x, p_y, x, y; z) = - \left(\left[\frac{E - U(x, y, z)}{c} \right]^2 - m^2 c^2 - p_x^2 - p_y^2 \right)^{1/2} = -p_z. \quad (2.3)$$

This is an example of “isoenergetic reduction,” in which the 3 degree-of-freedom (autonomous) Hamiltonian (2.1) has been reduced, for fixed values of H , to the 2 degree-of-freedom nonautonomous Hamiltonian (2.3). The solutions of (2.3) exist only on z -intervals for which the corresponding solutions of (2.1) have positive p_z .

These Hamiltonians describe individual particle motions in phase space and it is our goal to approximate these motions. It must be stressed that, in adopting such an approach, we probably relinquish the possibility of analyzing the exact trajectories of (2.1) or (2.3) for time intervals longer than $O(\frac{1}{\varepsilon})$. This is because, for a general anharmonic crystal potential U in (2.1), we suspect that there are positive Lyapunov exponents and so neighboring solutions separate exponentially locally. This leads to an exponentially fast separation after a characteristic time of $O(\frac{1}{\varepsilon})$, which means that approximate solutions cannot hope to be valid on longer intervals. Fortunately however, these $O(\frac{1}{\varepsilon})$ intervals are just long enough to provide useful information about the nature of solutions of (2.1), particularly about how they are approximated by continuum model solutions and their higher order corrections. (For example, $O(\frac{1}{\varepsilon})$ intervals are long enough to describe general axial motions which scatter off several strings or the spiral-type motions associated with proper channeling.)

To find approximating solutions which are simple to calculate and close in phase space to the exact solutions of (2.1) on the longest possible time interval, we use the method of averaging for periodic vector fields, a very general procedure for transforming systems of ODE's to a normal form based on averaging with respect to the z -coordinate. In essence, this reduces the problem of studying the 3 degree-of-freedom Hamiltonian system (2.1) to the problem of studying a 2 degree-of-freedom Hamiltonian system in the transverse coordinates (x, y) and transverse momenta (p_x, p_y) . If what is desired is instead an estimate of the precision with which longitudinal momentum, p_z , and an associated transverse energy are conserved, then Hamiltonian perturbation theory (specifically Nekhoroshev-like results) may be used to give conservation results on much longer time intervals [8]. Using these same methods, approximate solutions to (2.1) may also be derived, but their time of validity (i.e., of closeness in phase space to the exact solutions) is similarly limited to intervals of $O(\frac{1}{\varepsilon})$.

The method of averaging is the simplest and most general rigorous long-time perturbation procedure for ODEs, and in the present application, it gives the best possible results. In order to apply the method to the axial channeling problem, we proceed to scale the equations of motion arising from the isoenergetically reduced Hamiltonian (2.3). The scaling proceeds heuristically, on the assumption that the continuum model is a reasonable approximation. This allows the channeling problem to be cast in a form with a small parameter ε to which the averaging theorem can be applied; the resulting channeling theorem justifies, a posteriori, the scaling (and the continuum model).

Since axial channeling motions stay away from strings, channeled particles feel a maximum potential energy K which is on the order of 100 eV for particles with unit positive charge. Furthermore, continuum model motions have the property that the longitudinal momentum, p_z , is constant. If we assume p_z is constant and that channeled particles can move between the extremes $p_\perp = 0$, $U = K$ and $p_\perp = p_{\perp c}$, $U = 0$ then

$$[c^2(p_{\perp c}^2 + p_z^2) + m^2 c^4]^{1/2} = [c^2 p_z^2 + m^2 c^4]^{1/2} + K, \quad (2.4)$$

from conservation of energy. If we use the fact that $E^2 \simeq p_z^2 c^2 + m^2 c^4$, which follows from $|p_\perp| \ll p_z$ and the fact that $K \ll E$, then to a good approximation

$$p_{\perp c}^2 c^2 = 2EK, \quad (2.5)$$

and we take this as our definition of $p_{\perp c}$. The natural length scale is the minimal period, d , of the potential in the z -direction. We thus introduce the scaling transformation

$$P_\perp = p_\perp / p_{\perp c}, \quad Q_\perp = q_\perp / d, \quad t = z / d \quad (2.6)$$

into the Hamiltonian (2.3) where $q_\perp = (x, y)$. It is easy to check that this transformation preserves the Hamiltonian structure and the new Hamiltonian with independent variable t is

$$H_2(P_\perp, Q_\perp, t) = - \left\{ \left[\frac{E - KW(Q_\perp, t)}{p_{\perp c} c} \right]^2 - \frac{m^2 c^2}{p_{\perp c}^2} - P_\perp^2 \right\}^{1/2}, \quad (2.7)$$

which is just $H_1/p_{\perp c}$. Here $U = KW$ defines W . Using Eq. (2.5), the term in curly brackets can be written $\frac{E^2 - m^2 c^4}{2KE} - [P_\perp^2 + W(Q_\perp, t)] + \frac{1}{2} \frac{K}{E} W^2(Q_\perp, t)$ and if we introduce the small parameter

$$\varepsilon = \left(\frac{2KE}{E^2 - m^2 c^4} \right)^{1/2} = O \left(\frac{p_{\perp c}}{p_{zo}} \right) \ll 1 \quad (2.8)$$

and notice that $\frac{K}{E} \leq \frac{1}{2} \varepsilon^2$, then (2.7) can be written

$$\begin{aligned} H_2(P_\perp, Q_\perp, t) &= -\frac{1}{\varepsilon} \left\{ 1 - \varepsilon^2 [P_\perp^2 + W(Q_\perp, t)] + \frac{1}{4} a \varepsilon^4 W^2(Q_\perp, t) \right\}^{1/2} \\ &= -\frac{1}{\varepsilon} + \frac{1}{2} \varepsilon (P_\perp^2 + W(Q_\perp, t)) + \frac{1}{8} \varepsilon^3 [(P_\perp^2 + W(Q_\perp, t))^2 - a W^2(Q_\perp, t)] \\ &\quad + \varepsilon^5 S(P_\perp, Q_\perp, t, \varepsilon, a), \end{aligned} \quad (2.9)$$

where $a = 2 \frac{K}{E} \frac{1}{\varepsilon^2} = (E^2 - m^2 c^4)/E^2$ defines $a \in [0, 1]$. Notice that our scaling creates a distinguished limit in the sense that $P_\perp^2$ and W appear at the same order.

If we let γ_0 denote the incident Lorentz factor, $v_0 = \beta_0 c$ the associated incident velocity, and $U_0 \leq K$ the value of the potential energy on entrance, then $E = m\gamma_0 c^2 + U_0$ and it is easy to show that $0 < \varepsilon^2 \lesssim 2K/m\gamma_0 v_0^2$ and $\beta_0^2 \lesssim a \leq 1$. Here $\lesssim$ means less than

or equal and approximately equal, where the approximate equality results because $E \simeq m\gamma_0 c^2$, since U_0 is much smaller than the rest mass of even the lightest channeled particle (namely the positron). The parameters ε and a both depend on E , and $a = 0$ implies $E = mc^2$ and $\varepsilon = \infty$. Mathematically however, (2.9) is a perfectly well-defined Hamiltonian for ε and a independent and our asymptotic analysis for small ε will be valid for any a fixed in $[0,1]$. Furthermore, if we follow the procedure going from (2.1) to (2.9) for the nonrelativistic Hamiltonian (2.2) taking $p_{\perp c}^2 = 2mK$, then we arrive at (2.9) with $a = 0$ and $\varepsilon^2 = K/E \approx 2K/mv_0^2$. Our analysis will thus include the nonrelativistic case. Conversely, the ultrarelativistic case ($\gamma_0 \gg 1$) is obtained by taking $a = 1$ and $\varepsilon^2 = 2K/m\gamma_0 c^2$. In summary, the relativistic effect appears in (2.9) in two ways; one in terms of ε , where basically m is replaced by $m\gamma_0$, and the other in terms of a , which gives distinctly different Hamiltonians for $a = 0$ and a positive. To give some sense for orders of magnitude, we note that for 1 MeV, 1 GeV/c and 1 TeV/c protons, ε is approximately 1.0×10^{-2} , 5.0×10^{-4} and 2.0×10^{-5} , respectively. Furthermore, since the critical transverse velocity is defined by $p_{\perp c} = m\gamma_0 v_{\perp c}$ we have from (2.5) that $v_{\perp c} \simeq (2K/m\gamma_0)^{1/2}$, which decreases with increasing γ_0 . Also, we can define a channeling critical angle ψ_c by $v_{\perp c} = v_0 \sin \psi_c \simeq v_0 \psi_c$ which gives $\psi_c^2 \simeq 2K/m\gamma_0 v_0^2 \simeq \varepsilon^2$, so the critical angle is essentially ε . If K is chosen to be $2Z_1 Z_2 e^2/d$, which is reasonable for a Thomas-Fermi type screened Coulomb potential, then $\psi_c \simeq \psi_1$ (the critical angle of Eq. (1.1)). These points give further credence to our a priori choice of ε as the natural perturbation parameter.

To obtain the equations of motion, we define R by

$$\begin{aligned} \frac{1}{H_2} &= -\varepsilon \left(1 - \varepsilon^2 (x_1^2 + x_2^2 + W(x_3, x_4, t)) + \frac{1}{4} a \varepsilon^4 W(x_3, x_4, t)^2 \right)^{-1/2} \\ &= -\varepsilon - \frac{1}{2} \varepsilon^3 (x_1^2 + x_2^2 + W(x_3, x_4, t)) + \varepsilon^5 R(x, t, \varepsilon, a), \end{aligned} \quad (2.10)$$

where the conjugate variables collectively form the vector $x \in \mathbf{R}^4$, by defining $(x_1, x_2) = P_\perp$ and $(x_3, x_4) = Q_\perp$. The equations of motion arising from (2.2) are then written as the system

$$\begin{aligned} \dot{x} &= \varepsilon \begin{pmatrix} -\frac{1}{2}W_{x_3} \\ -\frac{1}{2}W_{x_4} \\ x_1 \\ x_2 \end{pmatrix} + \varepsilon^3 \begin{pmatrix} -\frac{1}{4}W_{x_3}[x_1^2 + x_2^2 + (1-a)W] \\ -\frac{1}{4}W_{x_4}[x_1^2 + x_2^2 + (1-a)W] \\ \frac{1}{2}x_1(x_1^2 + x_2^2 + W) \\ \frac{1}{2}x_2(x_1^2 + x_2^2 + W) \end{pmatrix} \\ &\quad + \varepsilon^5 \begin{pmatrix} \frac{1}{2}RW_{x_3} + \frac{a}{8}WW_{x_3}(x_1^2 + x_2^2 + W) - \frac{a}{4}\varepsilon^2 RWW_{x_3} \\ \frac{1}{2}RW_{x_4} + \frac{a}{8}WW_{x_4}(x_1^2 + x_2^2 + W) - \frac{a}{4}\varepsilon^2 RWW_{x_4} \\ -x_1R \\ -x_2R \end{pmatrix} \\ &= \varepsilon f_1(x, t) + \varepsilon^3 f_3(x, t, a) + \varepsilon^4 r(x, t, \varepsilon, a) , \end{aligned} \tag{2.11}$$

where $\dot{x} = dx/dt$ and the last equality defines f_1, f_3 , and r . This system is in a form amenable to the averaging method, and we apply third-order averaging to it in Section 4. In Section 3 we first state and prove the appropriate averaging theorem at third order.

Before proceeding, we note that by dropping terms of order higher than ε and by replacing $W(x_3, x_4, t)$ by its t -average $\overline{W}(x_3, x_4)$ we obtain a scaled form of the so-called axial continuum model

$$\dot{x} = \varepsilon \begin{pmatrix} -\frac{1}{2}\overline{W}_{x_3}(x_3, x_4) \\ -\frac{1}{2}\overline{W}_{x_4}(x_3, x_4) \\ x_1 \\ x_2 \end{pmatrix} \tag{2.12}$$

with its associated conservation of transverse energy

$$e_\perp = x_1^2 + x_2^2 + \overline{W}(x_3, x_4) . \tag{2.13}$$

Also notice that through first order in ε the relativistic effect in (2.11) enters only through ε , whereas it enters also through the linear term in a at third order. If we introduce the original variables into Eq. (2.12), we obtain what is essentially a Hamiltonian flow defined

by (2.2) with U replaced by its z -average and m replaced by $m\gamma_0$. The latter is the usual axial continuum model with relativistic effects as found in the channeling literature.

Having formulated the problem mathematically, we can now restate the goal of this paper. We wish to characterize approximate solutions of (2.11) on $O(\frac{1}{\varepsilon})$ time intervals. We will show that for a certain class of initial conditions the continuum model (2.12) is a good approximation to leading order and we will derive the $O(\varepsilon)$ and $O(\varepsilon^2)$ corrections to the continuum model.

3 Third Order Averaging For Periodic Systems

This section is meant to be self-contained, however the reader wishing a more general background in averaging methods is referred to the introduction in [5] and to two recent books [19,20], which also have interesting examples and references.

We discuss third-order averaging for the initial-value problem (IVP)

$$\begin{aligned}\dot{x} &= \varepsilon f_1(x, t) + \varepsilon^2 f_2(x, t) + \varepsilon^3 f_3(x, t) + \varepsilon^4 r(x, t, \varepsilon) , \\ x(0, \varepsilon) &= \xi ,\end{aligned}\tag{3.1}$$

where the f_i s are one-periodic in t . Here, the f_i s are defined on $\mathcal{U} \times \mathbf{R}$ where $\mathcal{U}$ is an open subset of $\mathbf{R}^n$ and r is defined on $\mathcal{U} \times \mathbf{R} \times [0, \varepsilon_0]$. Smoothness conditions will be discussed later. Our goal is to understand a special class of solutions $x(t, \varepsilon)$ of (3.1) on $O(\frac{1}{\varepsilon})$ time intervals in terms of a simpler autonomous system constructed from the f_i s. Averaging is essentially a transformation procedure which leads to a systematic perturbation expansion of certain solutions of (3.1), and leads to better and better approximations on time intervals of $O(\frac{1}{\varepsilon})$ as discussed in [5]. Here we focus on the third-order case because of its interest in the channeling problem.

In our approach, we look for an autonomous IVP

$$\dot{v} = \varepsilon V_1(v) + \varepsilon^2 V_2(v) + \varepsilon^3 V_3(v) ,\tag{3.2}$$

$$v(0, \varepsilon) = \xi + \varepsilon \eta_1(\xi) + \varepsilon^2 \eta_2(\xi) ,$$

whose solution we denote by $v(t, \varepsilon)$, and a transformation of period 1 in t ,

$$w = v + \varepsilon P_1(v, t) + \varepsilon^2 P_2(v, t) , \quad (3.3)$$

which transforms (3.2) into the IVP

$$\begin{aligned} \dot{w} &= \varepsilon f_1(w, t) + \varepsilon^2 f_2(w, t) + \varepsilon^3 f_3(w, t) + \varepsilon^3 g(v, t) + \varepsilon^4 h(v, t, \varepsilon), \\ w(0, \varepsilon) &= \xi + \varepsilon^3 \alpha(\xi, \varepsilon), \end{aligned} \quad (3.4)$$

whose solution we denote by $w(t, \varepsilon)$, where $g(v, t+1) = g(v, t)$ and has zero t -mean. The term $\varepsilon^4 h$ will contribute $O(\varepsilon^3)$ on $O(\frac{1}{\varepsilon})$ time intervals and a lemma of Besjes [21; cf. also Lemma 5 in 5] shows the same holds for $\varepsilon^3 g$; thus $\varepsilon^3 g + \varepsilon^4 h$ and $\varepsilon^4 r$ can be viewed as perturbations of the same problems and so the IVPs (3.4) and (3.1) are close. We will show, under suitable smoothness conditions, that $x(t, \varepsilon) = w(t, \varepsilon) + O(\varepsilon^3)$ uniformly on $O(\frac{1}{\varepsilon})$ time intervals; thus we have a third-order perturbation procedure based on a solution of the autonomous problem (3.2). To determine the V_i s, P_i s, g , and h , we differentiate (3.3) along the solutions of (3.2) and insert it into (3.4) to obtain

$$\begin{aligned} [1 + \varepsilon D_1 P_1(v, t) + \varepsilon^2 D_1 P_2(v, t)] [\varepsilon V_1(v) + \varepsilon^2 V_2(v) + \varepsilon^3 V_3(v)] + \varepsilon D_2 P_1(v, t) + \varepsilon^2 D_2 P_2(v, t) \\ = \varepsilon f_1(w, t) + \varepsilon^2 f_2(w, t) + \varepsilon^3 f_3(w, t) + \varepsilon^3 g(v, t) + \varepsilon^4 h(v, t, \varepsilon) . \end{aligned} \quad (3.5)$$

Here D_i denotes the derivative of the i th argument be it vector or scalar. Expanding the f_i s in a Taylor series, we obtain

$$\begin{aligned} f_1(w, t) &= f_1(v, t) + \varepsilon D_1 f_1(v, t) \cdot P_1(v, t) + \varepsilon^2 \left[\frac{1}{2} D_1^2 f_1(v, t) \cdot P_1^2(v, t) \right. \\ &\quad \left. + D_1 f_1(v, t) \cdot P_2(v, t) \right] + \varepsilon^3 \rho_1(v, t, \varepsilon) , \end{aligned} \quad (3.6a)$$

$$f_2(w, t) = f_2(v, t) + \varepsilon D_1 f_2(v, t) \cdot P_1(v, t) + \varepsilon^2 \rho_2(v, t, \varepsilon) , \quad (3.6b)$$

$$f_3(w, t) = f_3(v, t) + \varepsilon \rho_3(v, t, \varepsilon) , \quad (3.6c)$$

which can be viewed as defining the ρ_i s. Using (3.6) in (3.5) and equating powers of ε gives

$$V_1(v) + D_2 P_1(v, t) = F_1(v, t) , \quad (3.7a)$$

$$V_2(v) + D_2 P_2(v, t) = F_2(v, t) , \quad (3.7b)$$

$$V_3(v) - g(v, t) = F_3(v, t) , \quad (3.7c)$$

$$h(v, t, \varepsilon) = D_1 P_1(v, t) \cdot V_3(v) + D_1 P_2(v, t) \cdot V_2(v) + \varepsilon D_1 P_2(v, t) \cdot V_3(v) - \sum_{i=1}^3 \rho_i(v, t, \varepsilon) , \quad (3.7d)$$

where

$$F_1(v, t) = f_1(v, t) , \quad (3.8a)$$

$$F_2(v, t) = D_1 f_1(v, t) \cdot P_1(v, t) + f_2(v, t) - D_1 P_1(v, t) \cdot V_1(v) , \quad (3.8b)$$

$$\begin{aligned} F_3(v, t) = & \frac{1}{2} D_1^2 f_1(v, t) \cdot P_1^2(v, t) + D_1 f_1(v, t) \cdot P_2(v, t) , \\ & + D_1 f_2(v, t) \cdot P_1(v, t) + f_3(v, t) - D_1 P_1(v, t) \cdot V_2(v) \\ & - D_1 P_2(v, t) \cdot V_1(v) . \end{aligned} \quad (3.8c)$$

Equations (3.7) are solved sequentially, treating v and t as independent variables. Since P_1 is to be periodic in t , $D_2 P_1$ has zero t -mean; therefore $V_1(v)$ must be the average of f_1 (whence the term “method of averaging”). Equation (3.7a) then determines P_1 up to an additive function of v . In the same manner, (3.7b) and (3.7c) determine V_2 , P_2 (up to an additive function of v), V_3 and g . Finally, h is defined by (3.7d) in combination with (3.6). Thus, we have

$$V_i(v) = \int_0^1 F_i(v, t) dt , \quad i = 1, 2, 3 , \quad (3.9)$$

$$P_i(v, t) = \int_0^t [F_i(v, s) - V_i(v)] ds + p_i(v) , \quad i = 1, 2 , \quad (3.10)$$

where the p_i s are additive functions, and g and h are given by (3.7c) and (3.7d). The initial condition functions are determined by a regular perturbation analysis which yields

$$\eta_1(\xi) = -P_1(\xi, 0) , \quad (3.11a)$$

$$\eta_2(\xi) = -P_2(\xi, 0) + D_1 P_1(\xi, 0) \cdot P_1(\xi, 0) , \quad (3.11b)$$

$$\begin{aligned}\alpha(\xi, \varepsilon) &= \frac{1}{\varepsilon^2}[P_1(\xi + \varepsilon\eta_1(\xi) + \varepsilon^2\eta_2(\xi), 0) - P_1(\xi, 0) - \varepsilon D_1 P_1(\xi, 0) \cdot \eta_1(\xi)] \\ &\quad + \frac{1}{\varepsilon}[P_2(\xi + \varepsilon\eta_1(\xi) + \varepsilon^2\eta_2(\xi), 0) - P_2(\xi, 0)] .\end{aligned}\tag{3.11c}$$

We now have completely determined a candidate w for an approximate solution of the IVP (3.1), namely,

$$x(t, \varepsilon) \simeq w(t, \varepsilon) = v(t, \varepsilon) + \varepsilon P_1(v(t, \varepsilon), t) + \varepsilon^2 P_2(v(t, \varepsilon), t)\tag{3.12}$$

where P_1 and P_2 are defined by (3.10) and $v(t, \varepsilon)$ is the solution of the IVP (3.2) with the V_i s defined by (3.9) and the η s by (3.11a,b).

In order to state the theorem which specifies the relation between $x(t, \varepsilon)$ and $w(t, \varepsilon)$, we need some smoothness conditions on the f_i s and r . For the purpose of this paper, we make the following assumptions:

A) For $i = 1, 2, 3$, $D_1^{4-i} f_i$ is in $C(\mathcal{U} \times \mathbf{R}_+)$ and $f_i(x, t)$ is 1-periodic in t at each $x \in \mathcal{U}$. Here $\mathbf{R}_+ = [0, \infty)$. Moreover, $r(x, t, \varepsilon)$ is jointly continuous in x, t on $\mathcal{U} \times \mathbf{R}_+$ at each $\varepsilon \in (0, \varepsilon_0)$, and bounded and x -Lipschitz on each $K \times \mathbf{R}_+ \times (0, \varepsilon_0)$, where $K \subset \mathcal{U}$ is compact.

B) Each additive function p_i ($i = 1, 2, 3$) in (3.10) is in $C^{4-i}(\mathcal{U})$.

Note that A) implies that the rhs of the differential equation (3.1) has the same continuity and Lipschitz properties assumed for r , and hence that the IVP (3.1) has a unique maximal solution $x(t, \varepsilon)$ in $\mathcal{U}$.

An important concept in our theorem and its proof is the guiding solution of Perko [22; cf. also 5], which is defined by the ε -independent IVP

$$\begin{aligned}\frac{du}{d\tau} &= V_1(u) , \\ u(0) &= \xi .\end{aligned}\tag{3.13}$$

By A), $V_1(u)$ is locally Lipschitz on $\mathcal{U}$. Therefore, there exists a function $T_1 : \mathcal{U} \rightarrow (0, \infty]$ where $[0, T_1(\xi))$ is the maximum forward interval of existence of the solution of the IVP (3.13).

Standard continuation arguments show that for $T_1(\xi) < \infty$, $u(\tau)$ approaches the boundary of $\mathcal{U}$ as $\tau \nearrow T_1(\xi)$. We can now state the basic existence, uniqueness and approximation

Theorem (Third-Order Averaging). Fix ξ in $\mathcal{U}$ and let $T < T_1(\xi)$. Then there exist positive numbers $\varepsilon^* = \varepsilon^*(\xi, T) < \varepsilon_0$ and $C = C(\xi, T)$ such that $0 < \varepsilon \leq \varepsilon^*$ implies that for $0 \leq t \leq T/\varepsilon$, the IVPs (3.1) and (3.2) have unique solutions in $\mathcal{U}$ and

$$\|x(t, \varepsilon) - w(t, \varepsilon)\| \leq C\varepsilon^3. \quad (3.14)$$

Remark: We assume $\varepsilon_0 \leq 1$ in the proof for ease in making estimates.

Proof. The assumptions on the f_i 's, r , and p_i 's are sufficient to guarantee that all functions which enter the proof are bounded and/or x -Lipschitz on compact subsets of $\mathcal{U}$. Let $S = \{x \in \mathbf{R}^n : x = u(\tau), 0 \leq \tau \leq T\}$ be the compact set defined by the guiding solution and d be a positive distance less than the distance, $d_1(\xi, T)$, between S and $\partial\mathcal{U}$. Let $\mathcal{U}_i = \{x \in \mathbf{R}^n : x = y + z, y \in S, \|z\| < id/3\}$, $i = 1, 2, 3$, be open sets with compact closure. We let M and L be universal bound and Lipschitz constants on $\overline{\mathcal{U}_3}$. The proof has three steps, each involving restrictions on ε : (i) $v(t, \varepsilon)$ exists uniquely in $\mathcal{U}_1$ for $0 \leq t \leq T/\varepsilon$, $0 < \varepsilon \leq \varepsilon_1$; (ii) $w(t, \varepsilon)$ stays in $\mathcal{U}_2 \supset \mathcal{U}_1$ for $0 \leq t \leq T/\varepsilon$, $0 < \varepsilon < \varepsilon_2$; (iii) $x(t, \varepsilon)$ exists uniquely in $\mathcal{U}_3 \supset \mathcal{U}_2$ for $0 \leq t \leq T/\varepsilon$, $0 < \varepsilon \leq \varepsilon_3$ and satisfies $\|x(t, \varepsilon) - w(t, \varepsilon)\| \leq C\varepsilon^3$ at each such t, ε . Here the ε_i 's are sufficiently small positive constants.

To prove (i), we define $\tilde{u}(\tau, \varepsilon)$ as the solution of the IVP

$$\begin{aligned} \frac{d\tilde{u}}{d\tau} &= V_1(\tilde{u}) + \varepsilon V_2(\tilde{u}) + \varepsilon^2 V_3(\tilde{u}), \\ \tilde{u}(0, \varepsilon) &= \xi + \varepsilon \eta_1(\xi) + \varepsilon^2 \eta_2(\xi), \end{aligned}$$

which is related to (3.2) in the obvious way. Restrict ε so that $\tilde{u}(0, \varepsilon) \in \mathcal{U}_1$; then by standard continuation theory of ODEs, $\tilde{u}(\tau, \varepsilon)$ exists uniquely in $\mathcal{U}_1$ on some maximal forward interval

$[0, \tilde{T}_1(\varepsilon))$. Furthermore,

$$\begin{aligned}\tilde{u}(\tau, \varepsilon) - u(\tau) &= \varepsilon \eta_1(\xi) + \varepsilon^2 \eta_2(\xi) + \int_0^\tau [V_1(\tilde{u}(s, \varepsilon)) - V_1(u(s))] ds + \varepsilon \int_0^\tau V_2(\tilde{u}(s, \varepsilon)) ds \\ &\quad + \varepsilon^2 \int_0^\tau V_3(\tilde{u}(s, \varepsilon)) ds\end{aligned}$$

on $I = [0, \tilde{T}_1(\varepsilon)) \cap [0, T]$ at each such ε . Therefore, $\|\tilde{u}(\tau, \varepsilon) - u(\tau)\| \leq L \int_0^\tau \|\tilde{u}(s, \varepsilon) - u(s)\| ds + M\varepsilon(1 + \varepsilon)(1 + \tau)$ at the τ and ε in question, and the Gronwall inequality (see e.g. [23], p. 36) gives

$$\|\tilde{u}(\tau, \varepsilon) - u(\tau)\| \leq M\varepsilon(1 + \varepsilon)(1 + \tau) \exp L\tau$$

at each such τ, ε . Thus for $\varepsilon < \min(\varepsilon_0, \varepsilon_1)$, where $\varepsilon_1 = d/(6M(1+T) \exp LT)$, the right-hand side of the inequality is less than $d/3$, which ensures that $\tilde{T}_1 > T$ (if $\tilde{T}_1 \leq T$ then $\tilde{u}(\tau)$ will not have reached $\partial\mathcal{U}_1$, contradicting the fact that $[0, \tilde{T}_1)$ is a maximal interval of existence). This part of the proof is completed by noting that $v(t, \varepsilon) = \tilde{u}(\varepsilon t, \varepsilon)$. Part ii) follows easily by noting that the boundedness of P_1 and P_2 on $\mathcal{U}_2$ gives $\|w(t, \varepsilon) - v(t, \varepsilon)\| < d/3$ for $\varepsilon < \min(\varepsilon_0, \varepsilon_1, \varepsilon_2)$, where $\varepsilon_2 = d/6M$. In each case we have used the fact that $\varepsilon \leq 1$.

To prove (iii), we let $[0, \beta(\varepsilon))$ denote the maximal forward interval of existence of the unique solution of (3.1) in $\mathcal{U}_3$ and $J = [0, \beta(\varepsilon)) \cap [0, T/\varepsilon]$ for $0 < \varepsilon < \min(\varepsilon_0, \varepsilon_1, \varepsilon_2)$. A straightforward calculation shows that $w(t, \varepsilon)$ defined by (3.2) and (3.3) satisfies (3.4). Therefore, subtracting (3.4) from (3.1), integrating and using the triangle inequality gives

$$\begin{aligned}\|x(t, \varepsilon) - w(t, \varepsilon)\| &\leq \varepsilon^3 \|\alpha(\xi, \varepsilon)\| + \sum_{i=1}^3 \varepsilon^i \int_0^t \|f_i(x(s, \varepsilon), s) - f_i(w(s, \varepsilon), s)\| ds \\ &\quad + \varepsilon^3 \left\| \int_0^t g(v(s, \varepsilon), s) ds \right\| + \varepsilon^4 \int_0^t \|h(v(s, \varepsilon), s, \varepsilon)\| ds \\ &\quad + \varepsilon^4 \int_0^t \|r(x(s, \varepsilon), s, \varepsilon)\| ds.\end{aligned}$$

By a generalization of a result of Besjes (see [5], Lemma 5) (which uses the fact that g is of zero mean and v is slowly varying), $\|\int_0^t g(v(s, \varepsilon), s) ds\| \leq M(\varepsilon t + 1)$ for $0 \leq t \leq T/\varepsilon$, $0 < \varepsilon \leq \varepsilon_2$, and therefore

$$\|x(t, \varepsilon) - w(t, \varepsilon)\| \leq M\varepsilon^3(2 + 3\varepsilon t) + L(\varepsilon + \varepsilon^2 + \varepsilon^3) \int_0^t \|x(s, \varepsilon) - w(s, \varepsilon)\| ds$$

at these t, ε values. An application of the Gronwall inequality yields

$$\|x(t, \varepsilon) - w(t, \varepsilon)\| \leq M\varepsilon^3(2 + 3\varepsilon t) \exp L(1 + \varepsilon + \varepsilon^2)\varepsilon t, \quad (3.15)$$

for $t \in J$, $0 < \varepsilon \leq \varepsilon_2$. For $\varepsilon < \varepsilon^* = \min(\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3)$, where $\varepsilon_3^3 = d/3M(2 + 3T) \exp 3LT$, the right-hand side of the inequality is less than $d/3$, which assures that $\beta > T/\varepsilon$ (If $\beta \leq T/\varepsilon$ then $x(t, \varepsilon)$ will not have reached $\partial\mathcal{U}_3$, contradicting the fact that $[0, \beta(\varepsilon))$ is the maximal interval of existence of x in $\mathcal{U}_3$). The proof of the theorem is completed by taking $C = (2 + 3T)M \exp 3LT$.

Remarks

1. The first-order averaged system is (3.13) with τ replaced by εt . The first-order averaging theorem gives $\|x(t, \varepsilon) - u(\varepsilon t)\| = O(\varepsilon)$ on $[0, T/\varepsilon]$ with fewer restrictions on ε^* . A simple proof of this can be found in Remark 3 after Theorem 1 in Ref. [5].
2. The proof clearly illustrates the fact that this type of averaging theorem belongs to the class of Gronwall results for ODE's. Note that we have constructed a very conservative estimate of ε^* , since by their nature Gronwall estimates are quite crude, and we have used generic constants M and L .
3. This proof of the Third-Order Averaging Theorem is probably the simplest in the literature and is a complement to Remark 3 after the statement of Theorem 1 in Ref. [5]. The simplicity results from slightly relaxing the assumptions in Ref. [5] (cf. assumptions (A) and (B) here with (A), (B) and (C) in Ref. [5]). The key ideas in our proof are the guiding solution, the transformation (3.3), the ODE (3.4) for w , and the generalization of the Lemma of Besjes. The latter eliminates the need for a $P_3(v, t)$ in the transformation (3.3), which is a considerable simplification in the approximation. The elimination of P_3 results from the inclusion of the g function in Eq. (3.4).

4. If $T_1(\xi) < \infty$, then it is impossible to extend the averaging approximation to an interval longer (asymptotically) than T/ε . If $T_1(\xi) = \infty$, then the approximation may be valid on longer (e.g. $O(\frac{1}{\varepsilon^2})$) intervals, but more refined approaches are necessary except in some linear cases. If $V_1 = 0$, one can easily derive a version of the above approximation theorem which is valid on $O(\frac{1}{\varepsilon^2})$. A similar comment can be made when $V_1 = V_2 = 0$. Generally however, it is very difficult to obtain extensions of this theorem to intervals longer than $O(\frac{1}{\varepsilon})$.
5. We can pick any property of the guiding solution (3.13) which occurs on an arbitrary finite t interval $[0, T]$ (subject, of course, to $T < T_1(\xi) \leq \infty$), and for ε small enough this property is reflected in the exact solution on $0 \leq t \leq T/\varepsilon$. More precisely, the orbit of the guiding solution can be made as close as we like to the exact solution for ε small enough for t as above. This practical point seems to be overlooked in previous averaging theorems, where the interval of validity of the approximation, obtained from a local existence theorem, is shown to be $O(\frac{1}{\varepsilon})$ without optimizing the order constant.
6. For more details and a generalization of the theorem to N th-order averaging under less restrictive conditions, see [5].

4 Application of Averaging to the Axial Channeling Problem

It is now relatively simple to apply the averaging theorem to the axial channeling problem formulated in Section 2. We need only find a suitable domain $\mathcal{U}$ for the application of the averaging theorem to (2.11); that is, a domain of suitable (scaled) initial conditions which contains each orbit and on which the functions f_1 , f_3 and r defined by (2.11) have the behavior required in assumption (A) of the previous section. We first define $\mathcal{U}$ and then construct the approximation which leads to the theorem

4.1 The Domain $\mathcal{U}$

We assume that the (scaled) potential $W = U/K$ has at least four continuous derivatives “outside” the potential energy surface defined by $U(x_3, x_4, t) = K$; i.e., we assume $W \in C^4(D)$, where

$$D(K) = \{(x_3, x_4, t) \in \mathbf{R}^3 \mid W(x_3, x_4, t) < 1\}. \quad (4.1)$$

For example, if W is a sum of screened Coulomb potentials centered at the lattice sites, then it is of class $C^\infty(\mathbf{R}^3 \setminus \mathcal{S})$, where $\mathcal{S}$ is the set of all lattice sites on which W is singular, so it is certainly of class $C^4(D)$ for any finite $K > 0$, since $D \subset \mathbf{R}^3 \setminus \mathcal{S}$. Here K can be chosen to be any positive number; however, from a physical point of view it should be chosen to keep motions sufficiently far from lattice sites. If W is a sum of thermally averaged screened Coulomb potentials, then W is analytic on $\mathbf{R}^3$ and K is chosen so that $K \leq \max U$. In what follows we assume that W is one of these two types of potentials, thus D is typically a subset of $\mathbf{R}^3$ which excludes neighborhoods about the lattice sites, and ∂D is the union of slightly deformed spheres about each lattice site. Once the value of K has been chosen, channeling motions are readily identified as those solutions of (2.1) which remain in D for large times; i.e., as those solutions which do not attain potential energies equal to or greater than K . Since $E \gg K$, it is not clear at the outset that such solutions exist in reasonable abundance; solutions might very quickly convert some of their initial energy into potential energy, penetrating deeply through ∂D into the lattice sites.

We now proceed to define $\mathcal{U}$. Eventually, we show that each initial condition starting in $\mathcal{U}$ leads to a solution which (1) remains in D on time intervals of order $1/\varepsilon$ and (2) is approximated in phase space by a solution of an autonomous problem (if $D = \mathbf{R}^3$ (1) is trivial).

It is easy to see that Eq. (3.13), which defines the guiding solution, corresponds to the continuum model of (2.12). Thus it is natural to define $\mathcal{U}$, in terms of (2.12) and (2.13), to

be the subset of phase space defined by

$$\mathcal{U}(K) = \{x \in \mathbf{R}^4 \mid (x_1^2 + x_2^2) + \overline{W}(x_3, x_4) < e_{\perp c}\} , \quad (4.2)$$

where $e_{\perp c} = \min_{(x_3, x_4, t) \notin D} \overline{W}(x_3, x_4) \leq 1$ ensures that $x \in \mathcal{U}$ implies $(x_3, x_4, t) \in D$. To see this, simply note that $(x_3, x_4, t) \notin D$ implies $\overline{W}(x_3, x_4) \geq e_{\perp c}$ by definition, and this implies $x \notin \mathcal{U}$. The channeling critical angle is built into $\mathcal{U}$ and this will be discussed after the Axial Channeling Theorem.

If we define the “projection” $\mathcal{U}_p$ of $\mathcal{U}$ on configuration space by

$$\mathcal{U}_p(K) = \{(x_3, x_4, t) \in \mathbf{R}^3 \mid \overline{W}(x_3, x_4) < e_{\perp c}\} ,$$

then we have $\mathcal{U}_p \subset D$, and the bounding surfaces of $\mathcal{U}_p$ are cylindrical tubes parallel to the z -axis just containing the slightly deformed spheres comprising ∂D (recall t corresponds to z via (2.6)). In other words, the fact that $\mathcal{U}_p \subset D$ amounts to saying that there are “axial corridors,” with boundaries given as equipotential surfaces of the axial continuum model, which avoid the high-potential regions excluded from D . The significance of the requirement, mentioned in Section 2, that t be directed along a “low-order” axis is now clear: For fixed K , the axis must be of sufficiently low order to ensure that $e_{\perp c} > \min \overline{W}$ so that $\mathcal{U}$ is nonempty. (For fixed K , the order of the axis may be so high that the domain D is “opaque”; i.e., has no axial corridors.) The situation can easily be understood in Fig. 1 where we illustrate the axial continuum potential for a $\langle 100 \rangle$ cubic crystal structure. The atomic strings are at the (x_3, x_4) positions where the potential is unbounded and if K is chosen to correspond to the plane labeled a then a channeled particle stays away from the corresponding cross-hatched region at the right. The critical order above which no $\mathcal{U}$ exists depends on K , W , and the geometry of the crystal, and will not be discussed further here (although see the figure caption). Suffice it to say that a number of crystal potential contour calculations attest to reasonable domains $\mathcal{U}$ with $K \simeq 100$ eV [see, e.g. 10, p. 157].

Having chosen $\mathcal{U}$ in this way, $\varepsilon < \varepsilon_0 = \frac{1}{\sqrt{1+e_{\perp c}}}$ and $\mathcal{U}_p \subset D$ guarantee that f_1 , f_3 and r defined in (2.11) meet the requirements of assumption (A). The ε_0 is chosen so that the

term inside the square root in (2.9) is positive in $\mathcal{U}$; this in turn restricts consideration of the solutions of (2.1) to those with positive p_z . The averaging theorem therefore applies to the solution $x(t, \varepsilon, a)$ of the initial value problem (2.11), recalled here as

$$\dot{x} = \varepsilon f_1(x, t) + \varepsilon^3 f_3(x, t, a) + \varepsilon^4 r(x, t, \varepsilon, a) , \quad (4.3a)$$

$$x(0, \varepsilon, a) = \xi , \quad \xi \in \mathcal{U} , \quad (4.3b)$$

and its corresponding third-order averaged system, which we call the “third-order continuum model”:

$$\dot{v} = \varepsilon V_1(v) + \varepsilon^2 V_2(v) + \varepsilon^3 V_3(v, a) , \quad (4.4a)$$

$$v(0, \varepsilon, a) = \xi + \varepsilon \eta_1(\xi) + \varepsilon^2 \eta_2(\xi) , \quad (4.4b)$$

where V_1, V_2, V_3 and η_1, η_2 are given by (3.9) and (3.11). Since the continuum model (2.12) is precisely the first-order averaged system (as mentioned in Remark 1 of Section 3) derived from the perfect crystal model (4.3), we use “ n th-order continuum model” to denote the n th-order averaged system derived from (4.3) by expanding r in powers of ε and applying the general theory in [5]. We will show shortly that V_2 can be chosen to be 0, and that the existence time $T_1(\xi)$ for the guiding solution is ∞ . It will then follow from the averaging theorem that given any $\xi \in \mathcal{U}$ and $T > 0$, there exist $\varepsilon^* > 0, C > 0$ such that for all $0 < \varepsilon \leq \varepsilon^*$ and for all $0 \leq t \leq T/\varepsilon$, the solutions of (4.3) and (4.4) exist uniquely in $\mathcal{U}$ and the approximate solution $w(t, \varepsilon, a) = v(t, \varepsilon, a) + \varepsilon P_1(v(t, \varepsilon, a), t) + \varepsilon^2 P_2(v(t, \varepsilon, a), t)$ satisfies $\|x(t, \varepsilon, a) - w(t, \varepsilon, a)\| \leq C\varepsilon^3$ for $0 \leq t \leq T/\varepsilon$, where P_1, P_2 are given by (3.10). It should be noted that for fixed T , $\inf \varepsilon^*$ over $\xi \in \mathcal{U}$ is zero, but an ε^* could be chosen to govern any closed subset of initial conditions contained in $\mathcal{U}$.

4.2 Construction of the Approximation

We now construct the approximation, including the third-order continuum model (4.4) and the transformation which nearly converts it to (4.3), by following the algorithm given in Eqs. (3.7)-(3.11). In what follows we use the notation

$$\overline{F}(x) = \int_0^1 F(x, t) dt, \quad (4.5a)$$

$$\tilde{F}(x, t) = F(x, t) - \overline{F}(x), \quad (4.5b)$$

$$(I\tilde{F})(x, t) = \int_0^1 s \tilde{F}(x, t + s) ds, \quad (4.5c)$$

where $F(x, t+1) = F(x, t)$. These define the average, oscillating part and zero-mean integral, respectively, of F . Since $F_1 = f_1$, we have from (3.9)

$$V_1 = \overline{f}_1 = \begin{pmatrix} -\frac{1}{2}\overline{W}_{x_3} \\ -\frac{1}{2}\overline{W}_{x_4} \\ x_1 \\ x_2 \end{pmatrix}, \quad (4.6)$$

which is the vector field for the continuum model. A simple calculation using (3.10) gives

$$P_1 = -\frac{1}{2} \begin{pmatrix} I\tilde{W}_{x_3} \\ I\tilde{W}_{x_4} \\ 0 \\ 0 \end{pmatrix}, \quad (4.7)$$

where for simplicity we have chosen p_1 so that P_1 has zero mean, leading to $V_2 = 0$ below.

Next,

$$F_2 = D_1 f_1 \cdot P_1 - D_1 P_1 \cdot V_1,$$

since $f_2 = 0$. Now $V_2 = \overline{F}_2 = 0$ since $\overline{P}_1 = 0$ by the choice of p_1 . Thus

$$P_2 = I\tilde{F}_2 = -\frac{1}{2} \begin{pmatrix} I^2(x_1 \tilde{W}_{x_3 x_3} + x_2 \tilde{W}_{x_3 x_4}) \\ I^2(x_1 \tilde{W}_{x_3 x_4} + x_2 \tilde{W}_{x_4 x_4}) \\ I^2 \tilde{W}_{x_3} \\ I^2 \tilde{W}_{x_4} \end{pmatrix}, \quad (4.8)$$

where again we have taken P_2 to have zero mean for simplicity. We next calculate

$$F_3 = f_3 + \frac{1}{2}D_1^2 f_1 P_1^2 + D_1 f_1 \cdot P_2 - D_1 P_1 \cdot V_2 + D_1 f_2 \cdot P_1 - D_1 P_2 \cdot V_1 .$$

The fourth and fifth terms clearly vanish since $V_2 = f_2 = 0$. It is easy to show that the second term also vanishes, so that the preceding expression reduces to

$$F_3 = f_3 + D_1 f_1 \cdot P_2 - D_1 P_2 \cdot V_1 =$$

$$\begin{aligned} & \begin{pmatrix} -\frac{1}{4}W_{x_3}[x_1^2 + x_2^2 + (1-a)W] + \frac{1}{4}W_{x_3x_3}I^2(\tilde{W}_{x_3}) + \frac{1}{4}W_{x_3x_4}I^2(\tilde{W}_{x_4}) \\ -\frac{1}{4}W_{x_4}[x_1^2 + x_2^2 + (1-a)W] + \frac{1}{4}W_{x_3x_4}I^2(\tilde{W}_{x_3}) + \frac{1}{4}W_{x_4x_4}I^2(\tilde{W}_{x_4}) \\ \frac{1}{2}x_1(x_1^2 + x_2^2 + W) \\ \frac{1}{2}x_2(x_1^2 + x_2^2 + W) \end{pmatrix} \\ & + \begin{pmatrix} \frac{1}{4}[\overline{W}_{x_3}I^2(\tilde{W}_{x_3x_3}) + \overline{W}_{x_4}I^2(\tilde{W}_{x_3x_4})] - \frac{1}{2}x_1I^2(x_1\tilde{W}_{x_3x_3x_3} + x_2\tilde{W}_{x_3x_3x_4}) - \frac{1}{2}x_2I^2(x_1\tilde{W}_{x_3x_3x_4} + x_2\tilde{W}_{x_3x_4x_4}) \\ \frac{1}{4}[\overline{W}_{x_3}I^2(\tilde{W}_{x_3x_4}) + \overline{W}_{x_4}I^2(\tilde{W}_{x_4x_4})] - \frac{1}{2}x_1I^2(x_1\tilde{W}_{x_3x_3x_4} + x_2\tilde{W}_{x_3x_4x_4}) - \frac{1}{2}x_2I^2(x_1\tilde{W}_{x_3x_4x_4} + x_2\tilde{W}_{x_4x_4x_4}) \\ 0 \\ 0 \end{pmatrix} \end{aligned}$$

hence

$$V_3 = \overline{F}_3 =$$

$$\begin{pmatrix} -\frac{1}{4}(x_1^2 + x_2^2)\overline{W}_{x_3} - \frac{1}{4}(1-a)\overline{W}\overline{W}_{x_3} + \frac{1}{4}\overline{I^2(\tilde{W}_{x_3})\tilde{W}_{x_3x_3}} + \frac{1}{4}\overline{I^2(\tilde{W}_{x_4})\tilde{W}_{x_3x_4}} \\ -\frac{1}{4}(x_1^2 + x_2^2)\overline{W}_{x_4} - \frac{1}{4}(1-a)\overline{W}\overline{W}_{x_4} + \frac{1}{4}\overline{I^2(\tilde{W}_{x_3})\tilde{W}_{x_3x_4}} + \frac{1}{4}\overline{I^2(\tilde{W}_{x_4})\tilde{W}_{x_4x_4}} \\ \frac{1}{2}x_1(x_1^2 + x_2^2 + \overline{W}) \\ \frac{1}{2}x_2(x_1^2 + x_2^2 + \overline{W}) \end{pmatrix}, \quad (4.9)$$

$g = -\tilde{F}_3$, and h_3 is defined by (3.7d). It follows that the third-order continuum model is a Hamiltonian system with its “averaged” Hamiltonian given by

$$\begin{aligned} \overline{H}(x_1, x_2, x_3, x_4) &= \frac{\varepsilon}{2}[x_1^2 + x_2^2 + \overline{W}] + \frac{\varepsilon^3}{8}[(x_1^2 + x_2^2)^2 + 2(x_1^2 + x_2^2)\overline{W} \\ &\quad + (1-a)\overline{W}^2 + \overline{(I\tilde{W}_{x_3})^2} + \overline{(I\tilde{W}_{x_4})^2}]. \end{aligned} \quad (4.10)$$

The fact that we obtain a Hamiltonian vector field is probably a consequence of our choice of P_1 and P_2 to have zero mean. Note that the terms involving W are functions of (x_3, x_4) obtained in the indicated nested way. Having constructed the V_i s and P_i s, the η_i s are easy to compute from (3.11), and we now have all the ingredients for the approximation w of Eq. (3.12).

4.3 The Axial Channeling Theorem

We are ready to summarize our results on the axial continuum model by recasting the Third-Order Averaging Theorem of Section 3 as a theorem on axial channeling. Mathematically, it is simply a version of the Third-Order Averaging Theorem tailored to the case where f is given by (4.3).

Theorem (Axial Channeling) Choose $K > 0$ appropriately and consider an axial direction of sufficiently low order so that $e_{\perp c} > \min \overline{W}$ as discussed at the beginning of this section. Define $\mathcal{U}$ by (4.2), let $\xi \in \mathcal{U}$ and $T > 0$. Then there is an $\varepsilon^* = \varepsilon^*(\xi, T) > 0$ and there are order constants $C_1, C_2, C_3 > 0$ such that for all $0 < \varepsilon \leq \varepsilon^*$, the four IVP's

$$\dot{x} = \varepsilon f_1(x, t) + \varepsilon^3 f_3(x, t, a) + \varepsilon^4 r(x, t, \varepsilon, a), \quad x(0, \varepsilon, a) = \xi \quad (4.3)$$

(the perfect crystal model),

$$\dot{u} = \varepsilon V_1(u), \quad u(0, \varepsilon) = \xi \quad (4.11)$$

(the continuum model),

$$\dot{y} = \varepsilon V_1(y), \quad y(0, \varepsilon) = \xi + \varepsilon \eta_1(\xi) \quad (4.12)$$

(the second-order continuum model), and from (3.2)

$$\dot{v} = \varepsilon V_1(v) + \varepsilon^3 V_3(v, a), \quad v(0, \varepsilon, a) = \xi + \varepsilon \eta_1(\xi) + \varepsilon^2 \eta_2(\xi) \quad (4.13)$$

(the third-order continuum model),

have unique solutions in $\mathcal{U}$ for $0 \leq t \leq T/\varepsilon$ and, on the same t -interval,

$$\|x(t, \varepsilon, a) - u(t, \varepsilon)\| \leq C_1 \varepsilon, \quad (4.14)$$

$$\|x(t, \varepsilon, a) - y(t, \varepsilon) - \varepsilon P_1(y(t, \varepsilon), t)\| \leq C_2 \varepsilon^2, \quad (4.15)$$

and

$$\begin{aligned} & \|x(t, \varepsilon, a) - w(t, \varepsilon, a)\| = \\ & \|x(t, \varepsilon, a) - v(t, \varepsilon, a) - \varepsilon P_1(v(t, \varepsilon, a), t) - \varepsilon^2 P_2(v(t, \varepsilon, a), t)\| \leq C_3 \varepsilon^3. \end{aligned} \quad (4.16)$$

Here the V_i s are given by (4.6) and (4.9) and the P_i s are given by (4.7) and (4.8).

Remarks:

1. The continuum model (4.11) and the ensuing conclusion (4.14) are a consequence of the first-order averaging theorem mentioned in Remark 1 of Section 3.
2. Eq. (4.12) and the conclusion (4.15) do not appear in the averaging theorem of Section 3.
3. But the proof of this result (which is simply the averaging theorem at second order) goes through in the same way as the proof of the third-order estimate (4.16); and we have included this weaker estimate because of its special significance in channeling (to be described below).

3. T can be chosen so that any property of the continuum model is reflected in the perfect crystal model.
4. $\xi \in \mathcal{U}$, i.e. $\xi_1^2 + \xi_2^2 + \overline{W}(\xi_3, \xi_4) < e_{\perp c}$, represents a particle with an admissible initial position, i.e. $\overline{W}(x_3, x_4) < e_{\perp c}$, and an admissible initial angle, i.e. $\xi_1^2 + \xi_2^2 < e_{\perp c}$. Thus the initial position is in an axial corridor away from lattice sites. Furthermore, if ψ denotes the incident angle with respect to the axis under study then $\psi^2 \simeq \varepsilon^2(\xi_1^2 + \xi_2^2) \simeq \psi_c^2(\xi_1^2 + \xi_2^2) < \psi_c^2 e_{\perp c}$. However, $e_{\perp c} \leq 1$, therefore $\psi < \psi_c$. (Recall the discussion of the critical angle ψ_c and its relation to ψ_1 of Eq. (1.1) just before Eq. (2.10).) Also note that the angle ψ to the axis is given, to a very good approximation, by $\left| \frac{dq_{\perp}}{dz} \right|$ because all angles are small.
5. Since $\varepsilon^2 \lesssim K/\mathcal{E}$ where $\mathcal{E} = \frac{1}{2}m\gamma_0 v_0^2$ (recall discussion after (2.9)), the theorem asserts the existence of an energy $\mathcal{E}^* = K/\varepsilon^{*2}$ such that with the other assumptions the theorem holds for all incident energies $\mathcal{E} \geq \mathcal{E}^*$. Note the difference between $\mathcal{E}$ and $E = m\gamma_0 c^2 + U_0$.
6. The nonrelativistic approximations are obtained by taking $a = 0$ in (4.3) and (4.13) and using $\varepsilon^2 = K/E$ as discussed after (2.9). Furthermore, since a appears only in x in (4.14) and (4.15), these approximations essentially come from (2.2) with m replaced by $m\gamma_0$.

With these remarks, the Axial Channeling Theorem may now be summarized in physical terms by saying that given an axis of sufficiently low order and given a prescribed $T > 0$, there exists a set of initial directions and positions and an energy $\mathcal{E}^*$ sufficiently large to ensure that any particle with a greater energy beginning in $\mathcal{U}$ (nearly aligned with the axis and positioned away from axial strings) will penetrate the crystal to a depth of at least dT/ε (d is the lattice spacing), during which time it will remain in $\mathcal{U}$ (and thus in D , avoiding close encounters with nuclei), and during the same time its trajectory $x(t, \varepsilon, a)$

will be approximated to $O(\varepsilon)$, $O(\varepsilon^2)$, and $O(\varepsilon^3)$, respectively, by the approximate solutions $u(t, \varepsilon)$, $y(t, \varepsilon) + \varepsilon P_1(y(t, \varepsilon), \varepsilon)$, and $v(t, \varepsilon, a) + \varepsilon P_1(v(t, \varepsilon, a), \varepsilon) + \varepsilon^2 P_2(v(t, \varepsilon, a), \varepsilon)$. These approximate solutions are based on the continuum model solution u and the second- and third-order continuum model solutions y and v , all of which are solutions of autonomous IVP's.

The second-order continuum model deserves special comment: because the $O(\varepsilon^2)$ term is missing from (4.3), it was possible to exclude the $O(\varepsilon^2)$ terms from the ODE's of the generalized continuum models in (4.12) and (4.13). This means that the second-order continuum model differs from the continuum model only by an $O(\varepsilon)$ change in the initial condition. We thus have the somewhat surprising estimate (4.15): examining the form of P_1 in (4.7), we see that $y(t, \varepsilon) + \varepsilon P_1(y(t, \varepsilon), \varepsilon)$ includes the effects of lattice periodicity in the transverse momenta [see (4.7)]. In other words, at second order the averaging theory provides a rigorous approximate solution with lattice-period effects approximated to $O(\varepsilon^2)$ precision, and requires only knowledge of the solution y of the continuum model (with adjusted initial condition) and the perfect crystal potential W (appearing in P_1). We believe this to be the simplest way to recover the leading-order effects of lattice periodicity as discussed for example in [24], where it takes a *tour de force* using continued fractions to calculate this effect. This may also be compared with Ref. [7], where we derive essentially the same result for a simple two-dimensional model; however, estimate (4.15) not only extends our earlier lattice-period effects to the three-dimensional problem, but also demonstrates their higher ($O(\varepsilon^2)$) precision by way of the improved averaging method. The higher precision comes essentially through use of Besjes' Lemma as described in Section 3.

Of perhaps equal importance is our introduction of the third-order continuum model (3.2) and its approximate solution $w = v + \varepsilon P_1 + \varepsilon^2 P_2$ satisfying the rigorous estimate (4.13c) on $O(1/\varepsilon)$ time intervals. This gives the first nontrivial correction due to relativistic effects (the trivial correction essentially replaces m by $m\gamma_0$ in the nonrelativistic formalism starting from (2.2)), because V_3 depends on a . We believe this to be the most precise approximation

to axial channeling solutions which is explicitly calculable. We also believe that any higher order effects will not be detected in experiment; in fact, detection of these third-order effects (e.g., in the radiation spectrum of axially channeled positrons) presents a challenge.

4.4 The Time - Based Approximation

In the channeling theorem the independent variable, t , is the scaled depth z/d ; thus our results give the evolution of the transverse coordinates and momenta in terms of the depth. Here we discuss the time evolution in transverse phase space (recall that τ is the time-variable for the Hamiltonians (2.1) and (2.2)). This is required, for example, for studies of the radiation from channeled particles. If we introduce the scaled time

$$\zeta = \tau/\tau_0, \quad \tau_0 = d/c\sqrt{a}, \quad (4.17)$$

note from (2.9) that

$$\frac{p_z}{p_{\perp c}} = \frac{1}{\varepsilon} \{1 - \varepsilon^2 [P_{\perp}^2 + W(Q_{\perp}, t)] + \frac{1}{4} a \varepsilon^4 W^2(Q_{\perp}, t)\}^{1/2}, \quad (4.18)$$

and use the fact that

$$p_z = m\gamma \frac{dz}{d\tau},$$

then the scaled depth as a function of the scaled time, $t(\zeta)$, satisfies the differential equation

$$\frac{dt}{d\zeta} = \{1 - \varepsilon^2 [P_{\perp}^2(t) + W(Q_{\perp}(t), t)] + \frac{1}{4} a \varepsilon^4 W^2(Q_{\perp}(t), t)\}^{1/2} / \{1 - \frac{1}{2} \varepsilon^2 a W(Q_{\perp}(t), t)\}. \quad (4.19)$$

Since $(x_1, x_2) = P_{\perp}$ and $(x_3, x_4) = Q_{\perp}$ and since $x(t, \varepsilon, a)$ exists uniquely in $\mathcal{U}_3$ (see definition of $\mathcal{U}_3$ after (3.14)) for $t \in I = [0, T/\varepsilon]$, Eq. (4.19) and the definition of ε_0 give $0 < (1 - \frac{\varepsilon^2}{\varepsilon_0^2})^{1/2} < \frac{dt}{d\zeta} \leq 1$ for $t \in I$. The standard existence-uniqueness-continuation theorem gives the existence of a unique solution $t(\zeta)$ for $\zeta \in I$ and it follows that

$$\frac{dt}{d\zeta} = 1 + O(\varepsilon^2), \quad (4.20)$$

uniformly for $\zeta \in I$. Therefore,

$$t = \zeta + O(\varepsilon) \quad (4.21)$$

uniformly for $\zeta \in I$. Now from the channeling theorem we have

$$\begin{aligned} x(t, \varepsilon, a) &= u(t, \varepsilon) + O(\varepsilon) \\ &= y(t, \varepsilon) + \varepsilon P_1(y(t, \varepsilon), t) + O(\varepsilon^2) \end{aligned} \quad (4.22)$$

uniformly for $t \in I$. It follows from (4.22) that

$$\begin{aligned} x(t(\zeta), \varepsilon, a) &= u(\zeta, \varepsilon) + O(\varepsilon) \\ &= y(\zeta, \varepsilon) + \varepsilon P_1(y(\zeta, \varepsilon), \zeta) + O(\varepsilon^2), \end{aligned} \quad (4.23)$$

uniformly for $\zeta \in I$. The first equality is obvious and the second follows from $y(\zeta + O(\varepsilon), \varepsilon) = y(\zeta, \varepsilon) + O(\varepsilon^2)$ which is perhaps surprising, but true since y is a slowly varying function of its first argument. A similar argument in the nonrelativistic case also leads to Eq. (4.23), however τ_0 in (4.17) should be defined as $d/(2E/m)^{1/2}$ where E is the energy for (2.2). Thus for the first- and second-order approximations time and depth can be interchanged without changing the order of the approximation. Furthermore, it is a consequence of (4.20) that $\frac{dz}{d\tau} = c\sqrt{a}[1 + O(\varepsilon^2)]$, i.e., the z -velocity is constant in this approximation.

In order to incorporate the time into the third-order continuum model based on v of (4.13), the $O(\varepsilon)$ term in (4.21) must be used and time and depth cannot be interchanged. The correction term is easy to compute in terms of the first-order continuum model based on the fact that

$$\frac{dt}{d\zeta} = 1 - \frac{1}{2}\varepsilon^2[u_1^2(t, \varepsilon) + u_2^2(t, \varepsilon) + (1 - a)W(u_3(t, \varepsilon), u_4(t, \varepsilon), t)] + O(\varepsilon^3), \quad (4.24)$$

uniformly for $t \in I$.

4.5 Comparison with Constant z -Velocity Approximation

A commonly used approximation in channeling which includes the effects of short-term periodicity is to assume that the z -velocity, v_z , is constant and to solve the equations of motion defined by the time-dependent nonrelativistic Hamiltonian

$$h(p_\perp, q_\perp, \tau) = \frac{1}{2m} p_\perp^2 + U(q_\perp, v_z \tau), \quad (4.25)$$

related to (2.2) in an obvious way. Here we show that this approximation has the same error as our approximation based on the second-order continuum model. However, our approximation is much easier to use because it is based on an autonomous problem.

Equation (4.21) is valid in the case of (2.2) with $\tau_0 = d/(2E/m)^{1/2}$, thus $\frac{dz}{d\tau} = (\frac{2E}{m})^{1/2}(1 + O(\varepsilon^2))$ and it is natural to take $v_z = (\frac{2E}{m})^{1/2}$ (Here E is the energy for (2.2) and $\varepsilon^2 = K/E$). Taking $Q_\perp = q_\perp/d$, $P_\perp = p_\perp/p_{\perp c}$ ($p_{\perp c}^2 = 2mK$), $\zeta = v_z \tau/d$ and $U(q_\perp, v_z \tau) = KW(Q_\perp, \zeta)$, the equations of motion for (4.25) become

$$\frac{dP_\perp}{d\zeta} = -\frac{1}{2}\varepsilon \frac{\partial W(Q_\perp, \zeta)}{\partial Q_\perp}, \quad \frac{dQ_\perp}{d\zeta} = \varepsilon P_\perp, \quad (4.26)$$

which is Eq. (2.4) with t replaced by ζ and $f_3 = r = 0$. If we denote the solution of (4.26) by $\chi(\zeta, \varepsilon)$ then a simple application of the Gronwall inequality gives $x(\zeta, \varepsilon) = \chi(\zeta, \varepsilon) + O(\varepsilon^2)$ uniformly for $\zeta \in I = [O, T/\varepsilon]$. Furthermore, since x is a slowly varying function of its first argument, it follows that $x(t(\zeta), \varepsilon) = x(\zeta, \varepsilon) + O(\varepsilon^2)$; therefore

$$x(t(\zeta), \varepsilon) = \chi(\zeta, \varepsilon) + O(\varepsilon^2). \quad (4.27)$$

Comparing (4.27) and (4.23) we see that $\chi(\zeta, \varepsilon)$ and $y(\zeta, \varepsilon) + \varepsilon P_1(y(\zeta, \varepsilon), \zeta)$ give the same $O(\varepsilon^2)$ -approximation to x . Since $y(\zeta, \varepsilon)$ is defined by an autonomous problem and P_1 is easy to determine, the approximation $y + \varepsilon P_1$ is much easier to analyze than $\chi(\zeta, \varepsilon)$. For example, the step size necessary to numerically calculate y can be $O(\frac{1}{\varepsilon})$ times as large as the step size for χ .

A similar comparison can probably be made in the full relativistic case, although it is not then clear what Hamiltonian is analogous to (4.25), since a constant z -velocity does not

imply a constant z -momentum. However, the exercise is unnecessary since the approximation (4.22) would be as accurate and easier to use.

5 Present State of Mathematics of Channeling

When modeled as a deterministic classical system, the problem of directional effects in perfect crystals may be treated with a wide array of techniques from the modern theory of dynamical systems, and the results form a satisfying – although highly idealized – picture. As the model is made more complex by adding more realistic assumptions, the attending mathematics quickly grows to where the problem escapes fully rigorous treatment, at least with present methods. To place the axial channeling theory presented here in broader context, we contrast it briefly with some of the other mathematical approaches (both actual and potential).

We first want to stress that, in applications to axial channeling, the averaging method is especially appealing because of its relative simplicity and practicality. The equations defining the approximations are derived in explicit form and with only a few restrictions on ε . This simplicity is partly the result of looking only at the *axial* problem, so that we need only consider a single direction (an axis) along which the crystal is periodic. This in turn allows us to use a single-phase averaging technique, without the difficulties associated with resonances in multiphase averaging. This practicality can be better appreciated by contrasting the axial-averaging theory with the general channeling theory developed in [8] (see also [25] for a summary of this theory). As mentioned previously, the latter theory is a much more general one, embracing both axial and planar channeling motions as well as certain nonchanneling motions in perfect crystals. To obtain this generality, the approach of [8] confronts both multiphase resonance problems and the question of “optimal order of approximation.” Using the normal forms from the proof of Nekhoroshev’s theorem (a very general result in Hamiltonian perturbation theory; cf [26]), it is shown that both axial and planar channeling

motions are stable on time intervals which are exponentially long in an inverse power of ε , and that transverse energy and longitudinal momentum are nearly conserved on these intervals. This technique also generates approximate solutions, but for the reasons explained in Section 2, they remain close in phase space to exact solutions only on $O(1/\varepsilon)$ time intervals. Furthermore, the approximate solutions, though calculable in principle, are intractable in practice, and the ε values for which the approximations are proved valid are astronomically small and probably unrealistic in all but the most energetic channeling cases (however, one expects the ε restrictions to be extremely conservative). To summarize, the results in [8] provide a satisfying theoretical overview of particle motions in perfect crystals, and certainly give the longest possible stability times for such motions in the general case. Averaging, on the other hand, gives a practical procedure for calculating approximate solutions on much shorter time intervals, which are however still asymptotically long in $\mathcal{E} = \frac{1}{2}m\gamma_0 v_0^2$. In fact, in terms of actual times spent by particles in crystals, these $O(1/\varepsilon)$ intervals are satisfactory and probably $O(1/\varepsilon^2)$ would be ideal. It is possible that the perturbation results of [8], or other canonical perturbation techniques, could be carried to second or third order in the planar case; this would give practical planar approximations analogous to the axial approximations we present here, although we expect resonance problems would reduce their precision in the general case of uniform approximation for an entire plane (all directions in a single plane sufficiently far from low-order axial directions). A slightly different approach to the planar problem is possible in which single directions (away from axes) in a given plane are considered; this leads to a single-phase averaging method for quasiperiodic systems. Finally, to complete the mathematical picture of the classical perfect crystal model, we would need to understand motions for incident directions “between” axial and planar directions as well as the related possible transition between or escape from axial and planar motions on time intervals longer than the ones considered here.

A quantum mechanical treatment of axial channeling in the spirit of this paper has also been developed by Ben Lemlih and Ellison [27, 28] using the Hamiltonian (2.2); this makes

use of averaging techniques extended to PDE's by the last two authors [28,29]. Work is in progress on the quantum-relativistic case.

Once the perfect crystal model is understood, the next step from a channeling point of view is to incorporate the effect of lattice vibrations. This leads to an analysis of a Hamiltonian system with a random potential—a Hamiltonian which is a perturbation of (2.1). Cogburn, Ellison and Tapia [30,31] are pursuing this using a variety of approaches including Markov chains, stochastic averaging [32], and random evolutions [33].

Other channeling phenomena warranting eventual mathematical treatment include the effects of electron multiple scattering and radiative energy loss, which may be considerable at high incident particle energies. The latter leads to interesting questions involving the classical Lorentz-Dirac equations for the radiation reaction and strong field quantum electrodynamics.

It is our ultimate aim to bring together these various approaches and results into a coherent mathematical structure, with roots in the perfect crystal model and branches extending to embrace the observed features of channeling phenomena.

Acknowledgements

We thank Gloria Lopez for her careful and cheerful preparation of this manuscript. One of us (JAE) acknowledges the support, both financial and intellectual, of the Physics Institute of Aarhus University. I am grateful to J.U. Andersen, J. Lindhard and E. Uggerhøj for many discussions related to the validity of the continuum model and, even though the point of view in this paper differs considerably from theirs, I appreciated their willingness to discuss other viewpoints and valued their constructive criticism.

References

1. J. Lindhard, *Math. Fys. Medd. Dan. Vid. Selsk.* **34**, No. 14 (1965).
2. C. Lehmann and G. Leibfried, *J. Appl. Phys.* **34** (1963), 2821.
3. R. S. Nelson and M. W. Thompson, *Phil. Mag.* **8** (1963), 1677.
4. C. Erginsoy, *Phys. Rev. Lett.* **15** (1965), 360.
5. J. A. Ellison, A. W. Sáenz, and H. S. Dumas, *J. Diff. Equations* **84** (1990), 383.
6. T. J. Burns and J. A. Ellison, *Phys. Rev. B* **29** (1984), 2790.
7. H. S. Dumas and J. A. Ellison, in "Local and Global Methods of Nonlinear Dynamics," pp. 200-230 (R. Cawley, A. W. Sáenz, and W. W. Zachary, Eds.), Springer-Verlag, New York, 1986.
8. H. S. Dumas, "A Mathematical Theory of Particle Channeling in Crystals," Ph.D. thesis, University of New Mexico, 1988. To be published (in revised form) in *Dynamics Reported*.
9. J. U. Andersen, Institute of Physics, Aarhus University, private communication.
10. D. S. Gemell, *Rev. Mod. Phys.* **46** (1974), 129.
11. M. A. Kumakhov and G. Shirmer, "Atomic Collisions in Crystals," Gordon and Breach Science Publishers, New York, 1989.
12. L. C. Feldman, J. W. Mayer, and S. T. Picraux, "Material Analysis by Ion Channeling: Submicron Crystallography," Academic Press, New York, 1982.
13. Y. Ohtsuki, "Charged Beam Interaction with Solids," Taylor and Francis, New York, 1983.
14. "Coherent Radiation Sources" (A. W. Sáenz and H. Überall, Eds.), Topics in Current Physics, Vol. 38, Springer-Verlag, Berlin, 1985.
15. M. A. Kumakhov and F. F. Komarov, "Radiation From Charged Particles in Solids," American Institute of Physics, Translation Series, New York, 1989.
16. "Relativistic Channeling" (R. A. Carrigan, Jr. and J. A. Ellison, Eds.), NATO ASI Series, Series B: Physics, Vol. 165, Plenum Press, New York, 1987.

17. See the corresponding issues of *Nucl. Instr. Meth. B*.
18. A. H. Sørensen and E. Uggerhøj, *Nature* **324** (1987), 311 and *Sci. Am.* **260** (1989), 96.
19. J. A. Sanders and F. Verhulst, "Averaging Methods in Nonlinear Dynamical Systems," Springer-Verlag, New York, 1985.
20. P. Lochak and C. Meunier, "Multiphase Averaging for Classical Systems," Springer-Verlag, New York, 1988.
21. B. J. Besjes, *J. Mécanique* **8** (1969), 357.
22. L. M. Perko, *SIAM J. Appl. Math.* **17** (1968), 698.
23. J. K. Hale, "Ordinary Differential Equations," Wiley Interscience, New York, 1969.
24. G. Kurizki and J. K. McIver, *Phys. Lett.* **89A** (1982), 43.
25. H. S. Dumas and J. A. Ellison, in "Essays on Classical and Quantum Dynamics: A Festschrift in Honor of A. W. Sáenz" (J. A. Ellison and H. Überall, Eds.), Gordon and Breach Science Publishers, to appear.
26. N. N. Nekhoroshev, *Functional Anal. Appl.* **5**, (1971), 82-83; *Russ. Math. Surveys* **32** (1977), 1; G. Benettin, L. Galgani, and A. Giorgilli, *Celestial Mechanics* **37** (1985), 1.
27. A. Ben Lemlih and J. A. Ellison, in preparation.
28. A. Ben Lemlih, "An Extension of the Method of Averaging to Partial Differential Equations," Ph.D. Thesis, University of New Mexico, 1986.
29. A. Ben Lemlih and J. A. Ellison, *Phys. Rev. Lett.* **55** (1985), 1950.
30. R. Cogburn, J. A. Ellison and J. Tapia, private communication.
31. J. Tapia, Ph.D. Thesis, University of New Mexico, 1990.
32. R. Z. Khas'minskii, *Theory Prob. Appl.* **11** (1966), 211 and **11** (1966), 390; M. I. Freidlin and A. D. Wentzell, "Random Perturbations of Dynamical Systems," Springer-Verlag, New York, 1984.
33. R. Cogburn and R. Hersh, *Indiana Univ. Math. J.* **22** (1973), 211; G. C. Papanicolaou and S. R. S. Varadhan, *Comm. Pure Appl. Math.* **26** (1973), 497.

Figure 1. AXIAL CONTINUUM POTENTIAL U AS A FUNCTION OF TRANSVERSE COORDINATES x_3, x_4 (e.g., cubic crystal structure in $\langle 100 \rangle$ direction)

The constant energy planes a , b , and c cut the graph of U in the equipotential contours shown at right. Cross-hatching indicates regions in the transverse plane inaccessible to positive particles with transverse energies less than the corresponding constant energy plane. Particles with transverse energies up to the energy of plane b are restricted to proper channeling motions; particles with transverse energies up to the energy of plane a can in addition execute ordinary axial channeling motions .

In a schematic way, the figure illustrates the meaning of the constant K (cf. §§ 2 and 4.1). With K fixed at a particular transverse energy, the hatched regions in planes a , b , and c correspond to forbidden regions ($U(x_3, x_4) > K$) in the transverse plane for continuum potentials in three axial directions of successively higher order (as the order increases, axial strings become more closely spaced in the transverse plane, but less densely filled with atoms along their length). Plane c corresponds to an “opaque” direction—a direction of such high order that channeled particles’ transverse energy would exceed K over the entire transverse plane.

Recent IMA Preprints

#	Author/s	Title
621	J. David Logan ,	Wave propagation in a qualitative model of combustion under equilibrium conditions
622	M.L. Zeeman ,	Hopf bifurcations in competitive three-dimensional Lotka-Volterra Systems
623	Allan P. Fordy ,	Isospectral flows: their Hamiltonian structures, Miura maps and master symmetries
624	Daniel D. Joseph, John Nelson, Michael Renardy, and Yuriko Renardy ,	Two-Dimensional cusped interfaces
625	Avner Friedman and Bei Hu ,	A free boundary problem arising in electrophotography
626	Hamid Bellout, Avner Friedman and Victor Isakov ,	Stability for an inverse problem in potential theory
627	Barbara Lee Keyfitz ,	Shocks near the sonic line: A comparison between steady and unsteady models for change of type
628	Barbara Lee Keyfitz and Gerald G. Warnecke ,	The existence of viscous profiles and admissibility for transonic shocks
629	P. Szmolyan ,	Transversal heteroclinic and homoclinic orbits in singular perturbation problems
630	Philip Boyland ,	Rotation sets and monotone periodic orbits for annulus homeomorphisms
631	Kenneth R. Meyer ,	Apollonius coordinates, the N-body problem and continuation of periodic solutions
632	Chjan C. Lim ,	On the Poincaré–Whitney circuitspace and other properties of an incidence matrix for binary trees
633	K.L. Cooke and I. Györi ,	Numerical approximation of the solutions of delay differential equations on an infinite interval using piecewise constant arguments
634	Stanley Minkowitz and Matthew Witten ,	Periodicity in cell proliferation using an asynchronous cell population
635	M. Chipot and G. Dal Maso ,	Relaxed shape optimization: The case of nonnegative data for the Dirichlet problem
636	Jeffery M. Franke and Harlan W. Stech ,	Extensions of an algorithm for the analysis of nongeneric Hopf bifurcations, with applications to delay-difference equations
637	Xinfu Chen ,	Generation and propagation of the interface for reaction–diffusion equations
638	Philip Korman ,	Dynamics of the Lotka–Volterra systems with diffusion
639	Harlan W. Stech ,	Generic Hopf bifurcation in a class of integro-differential equations
640	Stephane Laederich ,	Periodic solutions of non linear differential difference equations
641	Peter J. Olver ,	Canonical Forms and Integrability of BiHamiltonian Systems
642	S.A. van Gils, M.P. Krupa and W.F. Langford ,	Hopf bifurcation with nonsemisimple 1:1 Resonance
643	R.D. James and D. Kinderlehrer ,	Frustration in ferromagnetic materials
644	Carlos Rocha ,	Properties of the attractor of a scalar parabolic P.D.E.
645	Debra Lewis ,	Lagrangian block diagonalization
646	Richard C. Churchill and David L. Rod ,	On the determination of Ziglin monodromy groups
647	Xinfu Chen and Avner Friedman ,	A nonlocal diffusion equation arising in terminally attached polymer chains
648	Peter Gritzmann and Victor Klee ,	Inner and outer j- Radii of convex bodies in finite-dimensional normed spaces
649	P. Szmolyan ,	Analysis of a singularly perturbed traveling wave problem
650	Stanley Reiter and Carl P. Simon ,	Decentralized dynamic processes for finding equilibrium
651	Fernando Reitich ,	Singular solutions of a transmission problem in plane linear elasticity for wedge-shaped regions
652	Russell A. Johnson ,	Cantor spectrum for the quasi-periodic Schrödinger equation
653	Wenxiong Liu ,	Singular solutions for a convection diffusion equation with absorption
654	Deborah Brandon and William J. Hrusa ,	Global existence of smooth shearing motions of a nonlinear viscoelastic fluid
655	James F. Reineck ,	The connection matrix in Morse–Smale flows II
656	Claude Baesens, John Guckenheimer, Seunghwan Kim and Robert Mackay ,	Simple resonance regions of torus diffeomorphisms
657	Willard Miller, Jr. ,	Lecture notes in radar/sonar: Topics in Harmonic analysis with applications to radar and sonar
658	Calvin H. Wilcox ,	Lecture notes in radar/sonar: Sonar and Radar Echo Structure
659	Richard E. Blahut ,	Lecture notes in radar/sonar: Theory of remote surveillance algorithms
660	D.V. Anosov ,	Hilbert’s 21st problem (according to Bolibruch)
661	Stephane Laederich ,	Ray–Singer torsion for complex manifolds and the adiabatic limit

- 662 **Geneviève Raugel and George R. Sell**, Navier-Stokes equations in thin 3d domains: Global
regularity of solutions I
- 663 **Emanuel Parzen**, Time series, statistics, and information
- 664 **Andrew Majda and Kevin Lamb**, Simplified equations for low Mach number combustion with
strong heat release
- 665 **Ju. S. Il'yashenko**, Global analysis of the phase portrait for the Kuramoto-Sivashinsky equation
- 666 **James F. Reineck**, Continuation to gradient flows
- 667 **Mohamed Sami Elbially**, Simultaneous binary collisions in the collinear N-body problem
- 668 **John A. Jacquez and Carl P. Simon**, Aids: The epidemiological significance of two different mean
rates of partner-change
- 669 **Carl P. Simon and John A. Jacquez**, Reproduction numbers and the stability of equilibria of SI
models for heterogeneous populations
- 670 **Matthew Stafford**, Markov partitions for expanding maps of the circle
- 671 **Ciprian Foias and Edriss S. Titi**, Determining nodes, finite difference schemes and inertial manifolds
- 672 **M.W. Smiley**, Global attractors and approximate inertial manifolds for abstract dissipative equations
- 673 **M.W. Smiley**, On the existence of smooth breathers for nonlinear wave equations
- 674 **Hitay Özbay and Janos Turi**, Robust stabilization of systems governed by singular integro-differential
equations
- 675 **Mary Silber and Edgar Knobloch**, Hopf bifurcation on a square lattice
- 676 **Christophe Golé**, Ghost circles for twist maps
- 677 **Christophe Golé**, Ghost tori for monotone maps
- 678 **Christophe Golé**, Monotone maps of $T^n \times R^n$ and their periodic orbits
- 679 **E.G. Kalnins and W. Miller, Jr.**, Hypergeometric expansions of Heun polynomials
- 680 **Victor A. Pliss and George R. Sell**, Perturbations of attractors of differential equations
- 681 **Avner Friedman and Peter Knabner**, A transport model with micro- and macro-structure
- 682 **E.G. Kalnins and W. Miller, Jr.**, A note on group contractions and radar ambiguity functions
- 683 **George R. Sell**, References on dynamical systems
- 684 **Shui-Nee Chow, Kening Lu and George R. Sell**, Smoothness of inertial manifolds
- 685 **Shui-Nee Chow, Xiao-Biao Lin and Kening Lu**, Smooth invariant foliations in infinite dimensional spaces
- 686 **Kening Lu**, A Hartman-Grobman theorem for scalar reaction-diffusion equations
- 687 **Christophe Golé and Glen R. Hall**, Poincaré's proof of Poincaré's last geometric theorem
- 688 **Mario Taboada**, Approximate inertial manifolds for parabolic evolutionary equations via Yosida approxi-
mations
- 689 **Peter Rejto and Mario Taboada**, Weighted resolvent estimates for Volterra operators on unbounded inter-
vals
- 690 **Joel D. Avrin**, Some examples of temperature bounds and concentration decay for a model of solid fuel
combustion
- 691 **Susan Friedlander and Misha M. Vishik**, Lax pair formulation for the Euler equation
- 692 **H. Scott Dumas**, Ergodization rates for linear flow on the torus
- 693 **A. Eden, A.J. Milani and B. Nicolaenko**, Finite dimensional exponential attractors for semilinear wave
equations with damping
- 694 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Inertial sets for dissipative evolution equations
- 695 **A. Eden, C. Foias, B. Nicolaenko & R. Temam**, Hölder continuity for the inverse of Mañé's pro-
jection
- 696 **Michel Chipot and Charles Collins**, Numerical approximations in variational problems with potential
wells
- 697 **Huanan Yang**, Nonlinear wave analysis and convergence of MUSCL schemes
- 698 **László Gerencsér and Zsuzsanna Vágó**, A strong approximation theorem for estimator processes in
continuous time
- 699 **László Gerencsér**, Multiple integrals with respect to L -mixing processes
- 700 **David Kinderlehrer and Pablo Pedregal**, Weak convergence of integrands and the Young measure
representation
- 701 **Bo Deng**, Symbolic dynamics for chaotic systems
- 702 **P. Galdi, D.D. Joseph, L. Preziosi, S. Rionero**, Mathematical problems for miscible, incompressible fluids
with Korteweg stresses
- 703 **Charles Collins and Mitchell Luskin**, Optimal order error estimates for the finite element approximation
of the solution of a nonconvex variational problem
- 704 **Peter Gritzmann and Victor Klee**, Computational complexity of inner and outer j -radii of polytopes in
finite-dimensional normed spaces
- 705 **A. Ronald Gallant and George Tauchen**, A nonparametric approach to nonlinear time series analysis:
estimation and simulation
- 706 **H.S. Dumas, J.A. Ellison and A.W. Sáenz**, Axial channeling in perfect crystals, the continuum model and
the method of averaging