Nonlinearity **11** (1998) 845–857. Printed in the UK PII: S0951-7715(98)85958-6

Removing the time dependence in a rapidly oscillating Hamiltonian

Ildar Gabitov† and Ian Marshall‡§

† L D Landau Institute for Theoretical Physics, 2 Kosygin str., 117940 Moscow, Russia

‡ Centre des mathématiques et leurs applications, Ecole normale supérieure de Cachan, 61 avenue du président Wilson, 94235 Cachan Cedex, France

Received 10 July 1997
Recommended by S Fauve

Abstract. Hamiltonian systems with rapidly oscillating explicit dependence on time are considered. The wavelength of this oscillation is treated as a small parameter and it is shown how to remove the time dependence up to some order in the small parameter by means of a canonical transformation presented in the form of an asymptotic series. The result has applications for the study of pulse propagation for high bit-rate transmission in optical fibres.

AMS classification scheme numbers: 78-08, 69-00
PACS number 4279Sz

1. Introduction

In this paper we will describe an approach for treating a Hamiltonian system with fast-oscillating time-dependent terms in such a way that it is approximated by a suitable autonomous system to some order in ω^{-1}, where ω is the frequency of the oscillation.

Our motivation for analysing this mathematical problem was to study the model of a particular physical system arising in the design of an optical fibre transmission system which makes use of dispersion–compensation to improve the bit-rate capacity.

The objective will be to find for any given system, a suitable time-dependent canonical transformation, to achieve the removal to the appropriate order, of the explicit time dependence in the equation of motion. The standard approach, known as von Zeipel's method and described for example in [1–3], would be to look for a generating function for the transformation in the form of an asymptotic series. Whilst this will certainly work just as well as the method which we describe, it is far more cumbersome to apply to the examples which we consider.

The method described in the present paper is derived almost directly from results described in [4, 5]. In this sense our work is a kind of comment on those papers. However, the Hamiltonian nature of the entire procedure, which seems to be more evident in our description, provides considerable computational advantage. Note that for the generating function approach the phase space is required to be a symplectic space, whereas in the present paper the phase space could be any Poisson space: in practice this avoids the need to identify explicitly the symplectic leaves.

§ Present address: Department of Mathematics, University of Glasgow, Glasgow G12 8QW, UK. E-mail: amt6im@amsta.leeds.ac.uk

The problem of averaging was treated via canonical transformations in a similar way to that used here, using the integration of a Hamiltonian vector field to define the transformation, in [6]. It is possible to see the result of this paper as a consequence of Hori's work. However, the form that the result takes here is particularly adapted to problems arising in the model of an optical fibre transmission system.

Consider the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x,t). \tag{1.1}$$

Here x may be a point in $\mathbb{R}^n$ or a point in a function space. In the first case (1.1) will be an ODE whilst in the second case it will in general be a PDE. The 'vector field' f is some t-dependent mapping from the space inhabited by x to the tangent space. In general the space we are considering could be any differentiable manifold, but in our examples it will always be a linear space.

Let us suppose that the dependence of f on t is via some fast-oscillating functions such as $\exp \pm \mathrm{i}\omega t$ with $\omega \gg 1$. We then pose the following problem.

Find a transformation $x \mapsto X$ in the form of an asymptotic series

$$X = x + \phi_1(x,t) + \phi_2(x,t) + \phi_3(x,t) + \cdots \tag{1.2}$$

with $\phi_k = \mathrm{O}(\omega^{-k})$ and $\partial\phi_k/\partial t = \mathrm{O}(\omega^{-k+1})$, such that X satisfies a differential equation of the form

$$\frac{\mathrm{d}X}{\mathrm{d}t} = F(X) + \mathrm{O}(\omega^{-n}) \tag{1.3}$$

in which the right-hand side is independent of t up to some order n in ω^{-1}.

There may be supplementary requirements to be met: for example if (1.1) is a t-dependent deformation of an integrable system, we might require the transformed system to be integrable up to the appropriate order in ω^{-1}. Such supplementary requirements will tend to limit the size of n.

1.1. Examples

(a) First, let us discuss a simple linear example. Consider the ODE in $\mathbb{R}^n$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = [A + B(t)]x \qquad x \in \mathbb{R}^n \tag{1.4}$$

where A is a real t-independent $n \times n$ matrix and B is a real t-dependent matrix of the form

$$B(t) = \beta \exp \mathrm{i}\omega t + \bar{\beta} \exp -\mathrm{i}\omega t \tag{1.5}$$

where β is a complex t-independent $n \times n$ matrix and $\bar{\beta}$ is its complex conjugate. Let us suppose that $\omega \gg 1$ and seek a transformation $x \mapsto X$ of the form

$$X = x + \phi_1(x,t) + \phi_2(x,t) + \cdots \tag{1.6}$$

with $\phi_k = \mathrm{O}(\omega^{-k})$ and $\mathrm{d}\phi_k/\mathrm{d}t = \mathrm{O}(\omega^{-k+1})$ in order to remove the t-dependence up to some leading order in ω^{-1}.

It is easy to check that if for example we take

$$\begin{aligned} \phi_1(x,t) &= \omega^{-2}\dot{B}(t)x \\ \phi_2(x,t) &= \omega^{-2}[A, B(t)]x - \tfrac{1}{2}\omega^{-2}B(t)^2 x, \end{aligned} \tag{1.7}$$

where '$\dot{}$' denotes '$\mathrm{d}/\mathrm{d}t$', we will obtain

$$\frac{\mathrm{d}X}{\mathrm{d}t} = (A + \tfrac{1}{2}\omega^{-1}[B, \dot{B}])X + \mathrm{O}(\omega^{-2}). \tag{1.8}$$

(Note that $\mathrm{d}/\mathrm{d}t([B, \dot{B}]) = 0$.) Here $[B, \dot{B}]$ denotes the commutator $B\dot{B} - \dot{B}B$. The exercise can be continued to any order required and we obtain an equation for X of the form

$$\frac{\mathrm{d}X}{\mathrm{d}t} = A_n X + \mathrm{O}(\omega^{-n}) \tag{1.9}$$

with A_n independent of t for any n, without violating the condition $\phi_k = \mathrm{O}(\omega^{-k})$.

The next three examples are much more complicated than the last one. To treat them it is convenient to use the machinery of Hamiltonian systems. We shall return to them again in section 3.

(b) Consider the time-dependent perturbation of the Garnier system

$$\begin{gathered}\frac{\mathrm{d}^2 q}{\mathrm{d}t^2} = -Aq - 2\left(\sum_{k=1}^{n} q_k^2\right)q - E(t)q \\ q \in \mathbb{R}^n, \qquad A \in gl_n(\mathbb{R}), \qquad A = A^T \\ E = \mathcal{E} \exp \mathrm{i}\omega t + \bar{\mathcal{E}} \exp -\mathrm{i}\omega t \qquad \omega \gg 1, \qquad \mathcal{E} \in gl_n(\mathbb{C}),\end{gathered} \tag{1.10}$$

where $\bar{\mathcal{E}}$ is the complex conjugate of $\mathcal{E}$. The Garnier system, which is integrable, is recovered from equation (1.10) when $\mathcal{E} = 0$. Let us try to find a transformation to new variables Q such that Q satisfies an ODE with no explicit t-dependence, say to order ω^{-2}.

The following pair of examples are related to the design of optical-fibre transmission systems. See the appendix for more details. Here as well as in Sections 2 and 3, '$\dot{}$' denotes '$\mathrm{d}/\mathrm{d}t$' and '$'$' denotes '$\mathrm{d}/\mathrm{d}x$'.

(c) Consider the time-dependent perturbation of the nonlinear Schrödinger equation (NLS) given by

$$\mathrm{i}\frac{\partial u}{\partial t} + \frac{\partial^2 u}{\partial x^2} + (1 + c(t))|u|^2 u = 0 \tag{1.11}$$

with c a periodic function satisfying $\max|c| \sim 1$ and $\max|\dot{c}| \gg 1$. We shall call this the NLS with order zero rapidly oscillating nonlinear perturbation and we shall use the abbreviation NRO(0) (for '*nonlinear rapidly oscillating of order zero*') to refer to it. It is obtained from the simplest model of an optical fibre transmission system in which the amplifiers are spaced periodically and the dispersion is constant, see [4].

(d) Consider the time-dependent perturbation of the NLS given by

$$\mathrm{i}\frac{\partial u}{\partial t} + (1 + d(t))\frac{\partial^2 u}{\partial x^2} + |u|^2 u = 0 \tag{1.12}$$

with d periodic, satisfying $1 \ll \max|d| \ll \max|\dot{d}|$. We shall call this the NLS with large rapidly oscillating dispersive perturbation and we shall refer to it by means of the abbreviation DRO(1/r) (for '*dispersion rapidly oscillating of order* $1/r$'), where r characterises the order of magnitude by which $|\dot{d}|$ dominates $|d|$, i.e. $\max|d| \sim (\max|\dot{d}|)^{1/r}$. (We shall look in detail only at the case $r = 2$, but in general r could be any rational number greater than 1.) It is obtained from the model of a system in which a short length of fibre having dispersion of opposite sign to and with much larger magnitude than that in the transmission fibre is placed at each amplifier station, whilst the amplifiers are still periodically spaced and the dispersion in the transmission fibre is constant, see [7].

In the next section we will describe a method based on canonical transformations which can be used for solving problems such as those described in examples (b), (c) and (d). In the following section we will return to them as an illustration of the method.

2. Time-dependent canonical transformations

For details concerning definitions the reader is recommended to consult the textbooks [1, 2, 8, 9].

We will use the following result.

Lemma 1. *Let A be a linear operator, depending on a real variable t. For $G = \mathrm{e}^A$, we have*

$$\dot{G}G^{-1} = \dot{A} + \tfrac{1}{2}[A, \dot{A}] + \tfrac{1}{6}[A[A\dot{A}]] + \cdots = \int_0^1 (\dot{A} + \tau[A\dot{A}] + \tfrac{1}{2}\tau^2[A[A\dot{A}]] + \cdots)\,\mathrm{d}\tau$$

$$= \int_0^1 (\exp(\tau \mathrm{ad}_A)\dot{A})\,\mathrm{d}\tau = \int_0^1 \mathrm{e}^{\tau A}\dot{A}\mathrm{e}^{-\tau A}\,\mathrm{d}\tau \qquad (2.1)$$

where $[,]$ *denotes the commutator given by* $[A, B] = AB - BA$ *and* ad_A *denotes the operator given by* $\mathrm{ad}_A B = [A, B]$.

Proof. $X(y, t) = \mathrm{e}^{yA(t)}$ is the unique solution to the matrix ODE $\partial X/\partial y = AX$ with $X(y = 0) = \mathrm{Id}\ \forall t$. The derivative $\dot{X} = \partial X/\partial t$ satisfies $\dot{X}(y = 0) = 0\ \forall t$ and $\partial/\partial y(\mathrm{e}^{-yA(t)}\dot{X}(y, t)) = \mathrm{e}^{-yA(t)}\dot{A}(t)X(y, t) = \mathrm{e}^{-yA(t)}\dot{A}(t)\mathrm{e}^{yA(t)}$. It follows that

$$(\dot{X}X^{-1})(y, t) = \int_0^y \mathrm{e}^{\tau A(t)}\dot{A}(t)\mathrm{e}^{-\tau A(t)}\,\mathrm{d}\tau \qquad (2.2)$$

and hence that

$$\dot{G}G^{-1} = (\dot{X}X^{-1})(y = 1) = \int_0^1 \mathrm{e}^{\tau A}\dot{A}\mathrm{e}^{-\tau A}\,\mathrm{d}\tau. \qquad (2.3)$$

□

We will also make crucial use of the following standard results.

Lemma 2. *Let $(P, \{,\})$ be a Poisson space and let $\Phi \in C^\infty(P)$. Let $\mathbb{X}_\Phi$ denote the Hamiltonian vector field corresponding to Φ. Then for any fixed $T \geqslant 0$, the mapping $m_T : P \to P$ (formally) defined by*

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mathbb{X}_\Phi(x) \qquad x(0) = \xi, m_T(\xi) = x(T), \qquad (2.4)$$

known as the 'time-T map' corresponding to Φ is always a Poisson map. That is

$$\mathrm{d}m_T(\xi) \cdot (\mathbb{X}_{H \circ m_T}(\xi)) = \mathbb{X}_H(m_T(\xi)) \qquad \forall H \in C^\infty(P), \forall \xi \in P. \qquad (2.5)$$

Here $\mathrm{d}m_T$ denotes the Jacobian of the map m_T, i.e. $\mathrm{d}m_T$ is the linear transformation defined on the tangent space which is compatible with m_T. Note that in general m_T is not well defined for all T and on the whole of P and this is the reason for the parenthetic qualifier 'formally', but see Remark (i) below.

Lemma 3. *Consider the dynamical system in $\mathbb{R}^n$*

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = v(x) \qquad x(0) = \xi. \qquad (2.6)$$

A formal solution to (2.6) is given by

$$x(\tau) = \xi + \tau v(\xi) + \tfrac{1}{2}\tau^2(v \cdot v)(\xi) + \tfrac{1}{6}\tau^3(v \cdot v \cdot v)(\xi) + \cdots \qquad (2.7)$$

where $v \cdot v$ *means*

$$(v \cdot v)(\xi) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0} v(\xi + tv(\xi)) = \sum_{i=1}^{n} v_i(\xi)\frac{\partial v}{\partial \xi_i}(\xi), \qquad (2.8)$$

$v \cdot v \cdot v$ *means*

$$(v \cdot v \cdot v)(\xi) = \frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t=0}\frac{\mathrm{d}}{\mathrm{d}s}\bigg|_{s=0} v(\xi + tv(\xi + sv(\xi))) = \sum_{i=1}^{n}\sum_{j=1}^{n} v_i(\xi)\frac{\partial}{\partial \xi_i}v_j(\xi)\frac{\partial v}{\partial \xi_j}(\xi), \qquad (2.9)$$

and so on.

It is convenient for us to give a different formulation of lemma 3, which is in fact also the generalization to the case where the dynamical system moves on a differentiable manifold M. In this case the formula in (2.6) is replaced by

$$\frac{\mathrm{d}}{\mathrm{d}t}(F(x(t)) = v(F)(x(t)) \qquad \forall F \in C^{\infty}(M), = \forall x \in M, x(0) = \xi \qquad (2.10)$$

where now v has to be understood as the vector field which acts as a derivative on functions, i.e. in local coordinates $\{x^i\}$ on M,

$$v(x) = v(x)^i \frac{\partial}{\partial x^i}. \qquad (2.11)$$

Lemma 3 can be deduced from the following.

Lemma 3a. *For any* F *a formal solution to (2.10) is given by*

$$F(x(\tau)) = F(\xi) + \tau(vF)(\xi) + \tfrac{1}{2}\tau^2(v^2F)(\xi) + \tfrac{1}{6}\tau^3(v^3F)(\xi) + \cdots. \qquad (2.12)$$

Let us note that Lemma 3a is Taylor's theorem for functions from $\mathbb{R}$ to $\mathbb{R}$ in disguise. To see this just consider the composition $f = F \circ x$ which defines the function f from $\mathbb{R}$ to $\mathbb{R}$.

Combining Lemmas 2 and 3, we have the following

Proposition 4. *Let* $(P, \{,\})$ *be a Poisson vector space. Let* $\Phi \in C^{\infty}(P)$*. Then the map* $m : P \to P$ *given formally by*

$$m = 1 + \mathbb{X}_{\Phi} + \tfrac{1}{2}\mathbb{X}_{\Phi} \cdot \mathbb{X}_{\Phi} + \cdots = \exp \mathbb{X}_{\Phi} \qquad (2.13)$$

is a Poisson map.

Proof. m is the time-1 map corresponding to Φ. □

We can formulate a coordinate-independent version of Proposition 4, which is more useful for us.

Proposition 4a. *Let* $(P, \{,\})$ *be a Poisson manifold and let* $\Phi \in C^{\infty}(P)$*. Then the map* $m : P \to P$ *given formally by*

$$m^*F = F + \mathbb{X}_{\Phi}(F) + \tfrac{1}{2}\mathbb{X}_{\Phi}^2(F) + \cdots = \exp \mathbb{X}_{\Phi}(F) \qquad \forall F \in C^{\infty}(P) \qquad (2.14)$$

is a Poisson map. Here $m^* : C^{\infty}(P) \to C^{\infty}(P)$ *denotes the pull-back map on functions, given by*

$$m^*F(x) = F(m(x)). \qquad (2.15)$$

Remarks. (i) If we have a small parameter ε and Φ has the property that $\Phi = \mathrm{O}(\varepsilon)$ in the region of P which concerns us, then we can replace the qualifier 'formally' in the above lemmas and propositions by 'in the form of an asymptotic series in ε'.

(ii) There is a subtle difference in the literature between the terms 'Poisson map' and 'canonical transformation', but for present purposes they are considered to be the same.

Suppose that we have a function $\Phi : \mathbb{R} \times P \to \mathbb{R}$, so that we introduce a continuous parameter or *time dependence*. Let us suppose that we define now the t-dependent map $m(t) : P \to P$ by

$$m(t) = \exp[\mathbb{X}_{\Phi(t)}]. \tag{2.16}$$

Then we have the following.

Proposition 5.

$$\dot{m}(t)m(t)^{-1} = \mathbb{X}_{K(t)} \qquad \textit{where } K = \dot{\Phi} + \tfrac{1}{2}\{\Phi, \dot{\Phi}\} + \tfrac{1}{6}\{\Phi, \{\Phi, \dot{\Phi}\}\} + \cdots. \tag{2.17}$$

Proof. Using Lemma 1 we have

$$\dot{m}m^{-1} = \dot{A} + \tfrac{1}{2}[A, \dot{A}] + \tfrac{1}{6}[A[A, \dot{A}]] + \cdots \tag{2.18}$$

with $A = \mathbb{X}_\Phi$. Now, making use of the standard property of Hamiltonian vector fields, that

$$[\mathbb{X}_F, \mathbb{X}_G] = \mathbb{X}_{\{F,G\}} \qquad \text{for any } F, G \in C^\infty(P), \tag{2.19}$$

we get the result. □

Suppose that we have the Hamiltonian system on $(P, \{,\})$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mathbb{X}_h(x) \tag{2.20}$$

and that we make a t-dependent canonical transformation of the form given by (2.16). Then for $X = m(t)(x)$ we obtain

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\mathrm{d}m}{\mathrm{d}t}(x) + \mathrm{d}m(t) \cdot \frac{\mathrm{d}x}{\mathrm{d}t} = \left(\frac{\mathrm{d}m}{\mathrm{d}t} \cdot m^{-1}\right)(X) + \mathbb{X}_{h\circ m^{-1}}(X) = \mathbb{X}_H(X) \tag{2.21}$$

with

$$H = (m^{-1})^* h + K = h \circ m^{-1} + K \tag{2.22}$$

where K is given by (2.17).

The result given by equations (2.20), (2.21) and (2.22) is the most important one for us. Let us therefore state it as a theorem.

Theorem 6. *Let $(P, \{,\})$ be a Poisson space and let $\Phi \in C^\infty(\mathbb{R} \times P)$. Let $h \in C^\infty(\mathbb{R} \times P)$. Let the map $m(t) : P \to P$ be given by $X = m(t)(x) = \exp[\mathbb{X}_{\Phi(t)}]\cdot x$. Then the Hamiltonian vector field $\dot{x} = \mathbb{X}_h(x)$ is mapped by m to $\dot{X} = \mathbb{X}_H(X)$, where $H = h \circ m^{-1} + K$ and the formula for K is given by (2.17).*

Let us now look at how we can use the above Theorem to transform away a rapidly oscillating term in the Hamiltonian—at least up to some order in ω^{-1}. Suppose that we have on some Poisson space P, a time-dependent Hamiltonian function h of the form given by

$$h(t) = F + c(t)G, \tag{2.23}$$

where $F, G \in C^\infty(P)$ and c is a periodic real function on $\mathbb{R}$ with the properties

$$\max\left|c^{-1}\frac{\mathrm{d}c}{\mathrm{d}t}\right| \gg 1 \qquad \text{and} \qquad \langle c\rangle = \int_\tau^{\tau+T} c(t)\frac{\mathrm{d}t}{T} = 0. \tag{2.24}$$

Here T is the period of c and so the integral in (2.24) is independent of τ. A function obeying the properties in (2.24) has to be a rapidly oscillating function.

Remark. The generalisation of the method to the treatment of a Hamiltonian function of the form $h(t) = F + c_1(t)G_1 + \cdots c_r(t)G_r$ is trivial. We will consider separately two different cases:

(i) $\max|c| \sim 1$,

(ii) $\max|c| \gg 1$, but $\max|\mathrm{d}c/\mathrm{d}t| \sim \max|c|^2$.

These two cases are by no means exhaustive of all possibilities, and indeed we expect the same method to work for cases of the type:

(iii) $\max|c| \gg 1$, but $\max|\mathrm{d}c/\mathrm{d}t| \sim \max|c|^r$

for $r = 3, 4, \ldots$.

We seek a function $\Phi \in C^\infty(\mathbb{R}\times P)$ in the form of an asymptotic series in a small parameter ε_1 such that the transformed Hamiltonian H constructed by means of Theorem 6 is independent of t to some order in ω^{-1} where ω is a large real number which characterises the frequency of the oscillating function c. In fact in both cases (i) and (ii) we will be able to require that $H = F$ to some order in ω^{-1}, and this is important for applications in which F is integrable.

Let

$$\Phi = \phi_1 + \phi_2 + \phi_3 + \cdots \tag{2.25}$$

where $\phi_k = \mathrm{O}(\varepsilon_1^k)$ and $|\phi_k^{-1}\dot{\phi}_k| \sim \omega \gg 1$. We will write $\omega^{-1} = \varepsilon$ and the small parameter ε_1 will be given by $\varepsilon_1 = \varepsilon$ in case (i) and by $\varepsilon_1^2 = \varepsilon$ in case (ii).

The rest of the procedure involves substitution of Φ from (2.25) into the formulae specified in the Theorem. The only difference between the two cases is in how terms are to be collected together.

We have

$$\begin{aligned} H = h &+ \{\Phi h\} + \tfrac{1}{2}\{\Phi, \{\Phi, h\}\} + \tfrac{1}{6}\{\Phi, \{\Phi\{\Phi, h\}\}\} + \cdots \\ &+ \dot{\Phi} + \tfrac{1}{2}\{\Phi, \dot{\Phi}\} + \tfrac{1}{6}\{\Phi, \{\Phi, \dot{\Phi}\}\} + \cdots. \end{aligned} \tag{2.26}$$

Case (i). We collect terms of the same order in $\varepsilon_1 = \varepsilon$. We get

$$\begin{aligned} H = F &+ cG + \dot{\phi}_1 + \{\phi_1, F\} + c\{\phi_1, G\} + \dot{\phi}_2 + \tfrac{1}{2}\{\phi_1, \dot{\phi}_1\} + \{\phi_2, F\} \\ &+ c\{\phi_2, G\} + \tfrac{1}{2}\{\phi_1, \{\phi_1, F\}\} + \tfrac{1}{2}c\{\phi_1, \{\phi_1, G\}\} + \dot{\phi}_3 \\ &+ \tfrac{1}{2}\{\phi_1, \dot{\phi}_2\} + \tfrac{1}{2}\{\phi_2, \dot{\phi}_1\} + \tfrac{1}{6}\{\phi_1, \{\phi_1, \dot{\phi}_1\}\} + \cdots. \end{aligned} \tag{2.27}$$

We have the requirement on the terms $\phi_1, \phi_2, \ldots$ that $\phi_k(t) \sim \mathrm{O}(\varepsilon^k)\ \forall t$. We therefore choose

$$\begin{aligned} \dot{\phi}_1 &= -cG + \langle cG\rangle, \\ \dot{\phi}_2 &= -\tfrac{1}{2}\{\phi_1, \dot{\phi}_1\} - \{\phi_1, F\} - c\{\phi_1, G\} + \langle \tfrac{1}{2}\{\phi_1, \dot{\phi}_1\} + \{\phi_1, F\} + c\{\phi_1, G\}\rangle, \end{aligned} \tag{2.28}$$

and so on. In this way we are able to make a t-dependent transformation on the phase space in such a way as to transform the system to a Hamiltonian system having a Hamiltonian function independent of t up to any order in ε.

Thus

$$\phi_1 = f_1 - c_1 G,$$
$$\phi_2 = f_2 - \tfrac{1}{2}c_1\{f_1, G\} - c_2\{F, G\}$$
$$\phi_3 = f_3 - c_3\{\{G, F\}, F\} + \tfrac{1}{2}\rho\{\{F, G\}, G\} - \tfrac{1}{2}c_2\{\{G, F\}f_1\} + \tfrac{1}{6}c_1\{\{G, f_1\}, f_1\}$$
$$-\tfrac{1}{2}c_1\{f_2, G\} + \tfrac{1}{12}c_1^2\{\{f_1, G\}, G\}, \qquad \text{with } \dot{\rho} = cc_2 - < cc_2 >, \tag{2.29}$$

and so on, where $c_1, c_2, c_3, \ldots$ are chosen such that

$$c_0 = c, \dot{c}_i = c_{i-1} \qquad \text{and} \qquad \langle c_i \rangle = 0\ \forall i \tag{2.30}$$

and $f_1, f_2, f_3, \ldots$ are arbitrary functions on P, with $f_i \sim \mathrm{O}(\varepsilon^i)$. We have

$$H = F + \langle cG\rangle + \langle \tfrac{1}{2}\{\phi_1, \dot{\phi}_1\} + \{\phi_1, F\} + c\{\phi_1, G\}\rangle + \langle\{\phi_2, F\} + c\{\phi_2, G\} + \tfrac{1}{2}\{\phi_1, \{\phi_1, F\}\}$$
$$+\tfrac{1}{2}c\{\phi_1, \{\phi_1, G\}\} + \tfrac{1}{2}\{\phi_1, \dot{\phi}_2\} + \tfrac{1}{2}\{\phi_2, \dot{\phi}_1\} + \tfrac{1}{6}\{\phi_1, \{\phi_1, \dot{\phi}_1\}\}\rangle + \cdots \tag{2.31}$$

and hence finally

$$H = F + \{f_1, F\} + \{f_2, F\} + +\tfrac{1}{2}\{\{F, f_1\}, f_1\} + \tfrac{1}{2}\langle c_1^2\rangle\{\{F, G\}, G\} + \cdots. \tag{2.32}$$

Case (ii). We collect terms of the same order in $\varepsilon_1 = \varepsilon^{1/2}$. We get

$$H = cG + \dot{\phi}_1 + F + c\{\phi_1, G\} + \dot{\phi}_2 + \tfrac{1}{2}\{\phi_1, \dot{\phi}_1\} + \{\phi_1, F\} + c\{\phi_2, G\}$$
$$+\tfrac{1}{2}c\{\phi_1, \{\phi_1, G\}\} + \dot{\phi}_3 + \tfrac{1}{2}\{\phi_1, \dot{\phi}_2\} + \tfrac{1}{2}\{\phi_2, \dot{\phi}_1\} + \tfrac{1}{6}\{\phi_1, \{\phi_1, \dot{\phi}_1\}\} + \cdots. \tag{2.33}$$

We have the requirement on the terms $\phi_1, \phi_2, \ldots$ that $\phi_k(t) \sim \mathrm{O}(\varepsilon_1^k)\ \forall t$. We therefore choose

$$\dot{\phi}_1 = -cG + \langle cG\rangle,$$
$$\dot{\phi}_2 = -(c\{\phi_1, G\} + \tfrac{1}{2}\{\phi_1, \dot{\phi}_1\}) + \langle c\{\phi_1, G\} + \tfrac{1}{2}\{\phi_1, \dot{\phi}_1\}\rangle,$$
$$\dot{\phi}_3 = -(\{\phi_1, F\} + c\{\phi_2, G\} + \tfrac{1}{2}c\{\phi_1, \{\phi_1, G\}\} + \tfrac{1}{2}\{\phi_1, \dot{\phi}_2\} + \tfrac{1}{2}\{\phi_2, \dot{\phi}_1\}$$
$$+\tfrac{1}{6}\{\phi_1, \{\phi_1, \dot{\phi}_1\}\}) + \langle\{\phi_1, F\} + c\{\phi_2, G\} + \tfrac{1}{2}c\{\phi_1, \{\phi_1, G\}\}$$
$$+\tfrac{1}{2}\{\phi_1, \dot{\phi}_2\} + \tfrac{1}{2}\{\phi_2, \dot{\phi}_1\} + \tfrac{1}{6}\{\phi_1, \{\phi_1, \dot{\phi}_1\}\}\rangle, \tag{2.34}$$

and so on. Thus

$$\phi_1 = f_1 - c_1 G,$$
$$\phi_2 = f_2 - \tfrac{1}{2}c_1\{f_1, G\} \tag{2.35}$$
$$\phi_3 = f_3 - c_3\{F, G\} - \tfrac{1}{2}c_1\{f_2, G\} + \tfrac{1}{12}c_1^2\{\{f_1, G\}, G\} - \tfrac{1}{12}c_1\{\{G, f_1\}, f_1\},$$

and so on, where $c_1, c_2, c_3, \ldots$ are chosen such that

$$c_{-1} = c, \qquad \dot{c}_i = c_{i-2} \qquad \text{and} \qquad \langle c_i\rangle = 0\ \forall i \tag{2.36}$$

and $f_1, f_2, \ldots$ are arbitrary functions on P, with $f_i \sim \mathrm{O}(\varepsilon_1^i)$. We have

$$H = F + \langle cG\rangle + \langle c\{\phi_1, G\} + \tfrac{1}{2}\{\phi_1, \dot{\phi}_1\}\rangle + \langle\{\phi_1, F\} + c\{\phi_2, G\}$$
$$+\tfrac{1}{2}c\{\phi_1, \{\phi_1, G\}\} + \tfrac{1}{2}\{\phi_1, \dot{\phi}_2\} + \tfrac{1}{2}\{\phi_2, \dot{\phi}_1\} + \tfrac{1}{6}\{\phi_1, \{\phi_1, \dot{\phi}_1\}\}\rangle + \cdots \tag{2.37}$$

and hence finally

$$H = F + \{f_1, F\} + \{f_2, F\} + \tfrac{1}{2}\langle c_1^2\rangle\{\{F, G\}, G\} + \cdots. \tag{2.38}$$

Note that the full details of the computation have not been shown to save space, but H has been computed up to terms involving $\dot{\phi}_4$.

Remark. In passing from (2.28) to (2.29) and from (2.34) to (2.35), we have included the t-independent 'constants of integration' f_1, f_2, etc. Another way of getting the same result would be to insist that all such constants of integration are to be zero, ensuring that $\langle\phi_i\rangle = 0$ holds for all i, and to compose the t-dependent canonical transformation given by (2.16) with Φ given by (2.25), with a t-independent one also given by (2.16) now with $\Phi = \hat{f}_1 + \hat{f}_2 + \cdots$, where $\hat{f}_1$, $\hat{f}_2$ are related to f_1, f_2 etc. but are not equal to them. In both cases (i) and (ii) it is natural to choose f_1 to be zero, as in both cases the highest-order term in the perturbation is $\{f_1, F\}$. However, in any given situation, it may be useful to leave f_2, f_3, etc. open to subsequent choice in order to simplify the problem further.

3. Examples

In this section we shall look at the examples from the Introduction, for which the results of the previous section are useful.

First of all we shall discuss the two perturbations of the NLS.

The phase space $(P, \{,\})$ is a subspace of $C^\infty(\mathbb{R}, \mathbb{C})$. We do not specify further what this subspace is, other than to say that it consists of those functions of x which decay sufficiently fast as $x \to \pm\infty$ that all integrals which arise in our computations are well defined (for details see [10]). We define the scalar product $\langle,\rangle : P \times P \to \mathbb{R}$ by

$$\langle a, b\rangle = 2\,\mathrm{Re}\int_{-\infty}^{\infty} a(x)b(x)\,\mathrm{d}x = \int_{-\infty}^{\infty}[a(x)b(x) + \bar{a}(x)\bar{b}(x)]\,\mathrm{d}x, \tag{3.1}$$

where we use a bar to denote the complex conjugate. For an element F of $C^\infty(P)$ and for a point $u \in P$, we define the variational derivatives $\delta_u F, \delta_{\bar{u}} F \in P$ by

$$\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} F(u + t\Delta) = \int_{-\infty}^{\infty}(\Delta\delta_u F + \bar{\Delta}\delta_{\bar{u}} F)(x)\,\mathrm{d}x \qquad \forall\Delta \in P. \tag{3.2}$$

Note that from $F : P \to \mathbb{R}$ we can deduce that $\delta_{\bar{u}} F$ is the complex conjugate of $\delta_u F$.

The Poisson bracket $\{,\} : C^\infty(P) \times C^\infty(P) \to C^\infty(P)$ is defined by

$$\{F, G\}(u) = \langle\delta_u F, \mathrm{i}\delta_{\bar{u}} G\rangle \qquad F, G \in C^\infty(P), u \in P. \tag{3.3}$$

Let the functions A and B in $C^\infty(P)$ be given by

$$A(u) = \tfrac{1}{2}\int_{-\infty}^{\infty}|u(x)|^4\,\mathrm{d}x \qquad \text{and} \qquad B(u) = -\int_{-\infty}^{\infty}|u'(x)|^2\,\mathrm{d}x. \tag{3.4}$$

The function $A + B$ is the Hamiltonian function for the NLS equation with respect to the above Poisson bracket. The order zero rapidly oscillating nonlinear perturbation of the NLS, or NRO(0) system, is then given by

$$h_{\mathrm{NRO}(0)} = A + B + cA \tag{3.5}$$

where c is a rapidly oscillating function of t and $\max|c| \sim \mathrm{O}(1)$, i.e. the NRO(0) system is an example of case (i).

The large rapidly oscillating disperion perturbation of the NLS, or DRO(1/r) system is given by

$$h_{\mathrm{DRO}(1/r)} = A + B + dB \tag{3.6}$$

where d is a rapidly oscillating function of t and $\max|d| \gg 1$. We shall confine ourselves here to the detailed study of DRO(1/2) where we have $\max|d|^2 \sim \omega$, i.e. the DRO(1/2) system is an example of case (ii).

We have to compute the functions

$$\begin{aligned}\mathcal{H}_{\mathrm{NRO}(0)} &= \{\{A+B, A\}, A\} = \{\{B, A\}, A\} \\ \mathcal{H}_{\mathrm{DRO}(1/2)} &= \{\{A+B, B\}, B\} = \{\{A, B\}, B\}\end{aligned} \tag{3.7}$$

in order to obtain correction terms at order ω^{-2} and at order ω^{-1} for the two respective systems. We get

$$\begin{aligned}\{A, B\}(u) = \langle |u|^2\bar{u}, \mathrm{i}u''\rangle = \mathrm{i}\int_{-\infty}^{\infty}(|u|^2(\bar{u}u'' - u\bar{u}''))(x)\,\mathrm{d}x \Rightarrow \delta_u\{A, B\} \\ = \mathrm{i}[\bar{u}^2u'' + (|u|^2\bar{u})'' - 2|u|^2\bar{u}'']\end{aligned} \tag{3.8}$$

and hence we obtain

$$\begin{aligned}\mathcal{H}_{\mathrm{NRO}(0)}(u) &= \langle -\mathrm{i}[\bar{u}^2u'' + (|u|^2\bar{u})'' - 2|u|^2\bar{u}''], \mathrm{i}|u|^2u\rangle \\ &= -\int_{-\infty}^{\infty}[|u|^4\bar{u}u'' + |u|^4u\bar{u}'' + 2|(|u|^2u)'|^2](x)\,\mathrm{d}x\end{aligned} \tag{3.9}$$

and

$$\begin{aligned}\mathcal{H}_{\mathrm{DRO}(1/2)}(u) &= \langle \mathrm{i}[\bar{u}^2u'' + (|u|^2\bar{u})'' - 2|u|^2\bar{u}'], \mathrm{i}u''\rangle \\ &= \int_{-\infty}^{\infty}[4|u|^2|u''|^2 - \bar{u}^2u''^2 - u^2\bar{u}''^2 - |u|^2\bar{u}u^{(4)} - |u|^2u\bar{u}^{(4)}](x)\,\mathrm{d}x.\end{aligned} \tag{3.10}$$

We still have some freedom in the correction terms in the transformed Hamiltonian function H. Whilst it is natural to choose f_1 in both cases to be zero, it may be convenient to retain the freedom in the choice of f_2. We have therefore in the case of the NRO(0) system

$$\begin{aligned}H_{\mathrm{NRO}(0)} &= A + B + \tfrac{1}{2}\langle c_1^2\rangle\mathcal{H}_{\mathrm{NRO}(0)} + \{f,\ A+B\} + \mathrm{O}(\omega^{-3}) \\ &\quad \text{with } f \in C^{\infty}(P) \text{ arbitrary and order } \mathrm{O}(\omega^{-2})\end{aligned} \tag{3.11}$$

and in the case of the DRO(1/2) system

$$\begin{aligned}H_{\mathrm{DRO}(1/2)} &= A + B + \tfrac{1}{2}\langle d_1^2\rangle\mathcal{H}_{\mathrm{DRO}(1/2)} + \{f,\ A+B\} + \mathrm{O}(\omega^{-3/2}) \\ &\quad \text{with } f \in C^{\infty}(P) \text{ arbitrary and order } \mathrm{O}(\omega^{-1}),\end{aligned} \tag{3.12}$$

where $c_1 \sim \mathrm{O}(\omega^{-1})$ and $d_1 \sim \mathrm{O}(\omega^{-1/2})$ are given by

$$\langle c_1\rangle = 0, \qquad \dot{c}_1 = c \qquad \text{and} \qquad \langle d_1\rangle = 0, \qquad \dot{d}_1 = d. \tag{3.13}$$

Remark. To compare (in the case of NRO(0)) the result given here with the transformation given in [4], one has to compute $\exp[\mathbb{X}_\Phi]\cdot u$ with $\Phi = \phi_1 + \phi_2 + \cdots$ and $\phi_1, \phi_2, \phi_3, \ldots$ are given by (2.29).

As it is a useful exercise which helps to clarify the method, let us look at the finite dimensional example which was example (b) in the Introduction.

The setting for this example is the phase space $P = \mathbb{R}^{2n}$, with canonical coordinates (q, p) and canonical Poisson bracket given by

$$\{F, G\} = \sum_{i=1}^{n}\left(\frac{\partial F}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial F}{\partial p_i}\frac{\partial G}{\partial q_i}\right). \tag{3.14}$$

Let E be a t-dependent, symmetric, real, $n \times n$ matrix, whose entries are all $\mathrm{O}(1)$ in ω^{-1}, given by

$$E(t) = \mathcal{E} \exp \mathrm{i}\omega t + \bar{\mathcal{E}} \exp -\mathrm{i}\omega t \tag{3.15}$$

with $\mathcal{E}$ a symmetric, complex, $n \times n$ matrix of $\mathrm{O}(1)$ in ω^{-1} and $\bar{\mathcal{E}}$ its complex conjugate. We will assume that $\omega \gg 1$ and consider the t-dependent Hamiltonian function h given by

$$\begin{aligned} h_E(t)(q, p) &= \tfrac{1}{2}\left(\sum_{i=1}^{n} p_i^2 + \left(\sum_{i=1}^{n} q_i^2 \right)^2 + q^T A q + q^T E(t) q \right) \\ &= \tfrac{1}{2}(p^2 + q^4 + q^T A q + q^T E q). \end{aligned} \tag{3.16}$$

Hamilton's equation gives the system in (1.10).

Let us apply the approach described in Section 2 to remove the t-dependence up to some order in ω^{-1}. For $E = 0$, (3.16) defines an integrable system. We expect therefore, by analogy with the first two examples, to proceed only up to terms which guarantee integrability of the transformed Hamiltonian to the given order in ω^{-1}. Let us denote by F and G the functions given by

$$F(q, p) = h_0(q, p) = \tfrac{1}{2}(p^2 + q^4 + q^T A q) \tag{3.17}$$

and

$$G(t)(q, p) = (h_E - h_0)(q, p) = \tfrac{1}{2} q^T E(t) q. \tag{3.18}$$

Now let Φ be given by

$$\Phi = \phi_1 + \phi_2 + \phi_3 + \cdots, \tag{3.19}$$

with

$$\begin{aligned} \phi_1 &= \omega^{-2} \dot{G} \\ \phi_2 &= \omega^{-2} \{F, G\} \\ \phi_3 &= \omega^{-4} \{\{F, \dot{G}\}, F\} + \tfrac{1}{4} \omega^{-4} \{\{F, G\}, \dot{G}\} \\ \phi_4 &= -\omega^{-4} \{\{\{F, G\}, F\}, F\} - \tfrac{3}{8} \omega^{-4} \{\{\{F, G\}, G\}, F\}. \end{aligned} \tag{3.20}$$

Then we get

$$H(q, p) = F(q, p) - \omega^{-2} q^T \mathcal{E} \bar{\mathcal{E}} q + \mathrm{i}\omega^{-3} p^T [\bar{\mathcal{E}}, \mathcal{E}] q + \mathrm{O}(\omega^{-4}), \tag{3.21}$$

where $[,]$ denotes the commutator as in (1.8) and in (2.1). To simplify the computation of the term at $\mathrm{O}(\omega^{-4})$ in H, we may observe that any term containing an *odd* number of Gs and $\dot{G}$s can be cancelled by an appropriate choice of ϕ_5 and if we stop at this order, then we shall not need the explicit form of ϕ_5. The only terms which cannot be cancelled are those arising from taking the average over a period $2\pi/\omega$ of terms containing an even number of G and $\dot{G}$. We obtain for the $\mathrm{O}(\omega^{-4})$ term,

$$\langle -\tfrac{1}{2} \omega^{-4} \{\{\{\{F, G\}, F\}, F\}, G\} - \tfrac{3}{8} \omega^{-4} \{\{\{\{F, G\}, G\}, F\}, F\} \rangle. \tag{3.22}$$

We have then

$$\begin{aligned} H(q, p) = {} & \tfrac{1}{2} p^2 + \tfrac{1}{2} q^4 + \tfrac{1}{2} q^T A q - \omega^{-2} q^T \mathcal{E} \bar{\mathcal{E}} q + \mathrm{i}\omega^{-3} p^T [\bar{\mathcal{E}}, \mathcal{E}] q + \omega^{-4} (7 q^2 q^T \mathcal{E} \bar{\mathcal{E}} q \\ & + 4 (q^T \mathcal{E} q)(q^T \bar{\mathcal{E}} q) - \tfrac{3}{2} p^T \mathcal{E} \bar{\mathcal{E}} p + q^T (\mathcal{E} A \bar{\mathcal{E}} + \tfrac{9}{4} \mathcal{E} \bar{\mathcal{E}} A + \tfrac{9}{4} \bar{\mathcal{E}} \mathcal{E} A) q) + \mathrm{O}(\omega^{-5}). \end{aligned} \tag{3.23}$$

If we now make a further canonical transformation $(q, p) \mapsto (Q, P)$ given by

$$\begin{aligned} P &= (1 - \tfrac{3}{4} \omega^{-4} (\mathcal{E} \bar{\mathcal{E}} + \bar{\mathcal{E}} \mathcal{E}))(p + \mathrm{i}\omega^{-3} [\mathcal{E}, \bar{\mathcal{E}}] q) \\ q &= (1 - \tfrac{3}{4} \omega^{-4} (\mathcal{E} \bar{\mathcal{E}} + \bar{\mathcal{E}} \mathcal{E})) Q, \end{aligned} \tag{3.24}$$

which has the generating function S given (to order ω^{-5}) by

$$S(q,P) = q^T(1 + \tfrac{3}{4}\omega^{-4}(\mathcal{E}\bar{\mathcal{E}} + \bar{\mathcal{E}}\mathcal{E}))(P - \tfrac{1}{2}\mathrm{i}\omega^{-3}[\mathcal{E},\bar{\mathcal{E}}]q), \tag{3.25}$$

then we have

$$H(q,p) = K(Q,P) = \tfrac{1}{2}P^2 + \tfrac{1}{2}Q^4 + \tfrac{1}{2}Q^T\hat{A}Q$$
$$+4\omega^{-4}(Q^2Q^T\mathcal{E}\bar{\mathcal{E}}Q + (Q^T\mathcal{E}Q)(Q^T\bar{\mathcal{E}}Q)) + \mathrm{O}(\omega^{-5}). \tag{3.26}$$

Here we denote by $\hat{A}$ the t-independent symmetric matrix which arises in the computation; it is not necessary to write it out explicitly. As the Garnier system is integrable for any A, it follows that the transformed Hamiltonian H above is integrable up to order $\mathrm{O}(\omega^{-3})$, in the sense that the Hamiltonian is an order ω^{-4} perturbation of an integrable one.

Appendix. Derivation of the cases NRO and DRO for an optical fibre transmission system

In an optical fibre transmission system, the envelope of the electrical field E is described by the equation

$$\mathrm{i}\frac{\partial E}{\partial Z} + d(Z)\frac{\partial^2 E}{\partial T^2} + \alpha|E|^2E = \mathrm{i}G(Z)E, \tag{A.1}$$

where $G(Z) = -\gamma + A\sum_{k=1}^{N}\delta(Z - kZ_a)$.

Here T is time and Z is the space variable in a frame which moves at the group velocity of the wave. There are several parameters in this equation: γ describes fibre losses; α describes the Kerr nonlinearity (the dependence of the refractive index in the fibre on light intensity); $d(Z)$ describes fibre chromatic dispersion; the function $A\sum_{k=1}^{N}\delta(Z-kZ_a)$ represents the amplifiers, which compensate the fibre losses, and Z_a is the distance between amplifiers (which is a constant in the simplest model).

Let the function c be given by

$$c(Z) = \exp\left(2\int_0^Z G(\zeta)\,\mathrm{d}\zeta\right). \tag{A.2}$$

Let the positive constants k and a be given by

$$k^2 = \langle d\rangle^{-1} \qquad \text{and} \qquad a^2 = \alpha\langle c\rangle \tag{A.3}$$

and introduce the rescaled variables t and x, which are related to Z and T by

$$t = Z, \qquad x = kT. \tag{A.4}$$

Now if we let the field u be defined by

$$u(x,t) = a\exp\left(-\int_0^t G(\zeta)\,\mathrm{d}\zeta\right)E(t,x/k), \tag{A.5}$$

we obtain the following equation for u

$$\mathrm{i}\frac{\partial u}{\partial t} + \left(1 + \frac{\tilde{d}(t)}{\langle d\rangle}\right)\frac{\partial^2 u}{\partial x^2} + \left(1 + \frac{\tilde{c}(t)}{\langle c\rangle}\right)|u|^2u = 0, \tag{A.6}$$

where $\tilde{d} = d - \langle d\rangle$ and $\tilde{c} = c - \langle c\rangle$. On lengthscales significantly larger than Z_a (recall that *length* is t in the transformed equation), $\tilde{c}$ is a rapidly oscillating function and in addition we have $\max|\langle c\rangle^{-1}\tilde{c}| \sim 1$. When $\tilde{c} = 0$ and $\tilde{d} = 0$ (A.6) is the Nonlinear Schrodinger Equation (NSL).

In [4] it is shown how to transform the above equation to the NLS with lower-order correction terms for $\max|\langle c\rangle^{-1}\tilde{c}| \sim 1$ and $\max|\langle d\rangle^{-1}\tilde{d}| \sim 1$. This is the physical system in which the nonlinearity α and the dispersion d are both of order 1. By analogy with a phenomenon in plasma physics, the authors of [4] coined the name *guiding centre soliton* for the solutions of the transformed equation in this case. Note that the method used by the authors of [4], although rather general, is considerably more complicated at a computational level than the one used in this paper.

In the case of dispersion compensation, the effect of the fibre chromatic dispersion is compensated by interspersing at each amplifier, between the lengths of the transmission fibre, short sections of fibre with chromatic dispersion of opposite sign to—and relatively much larger absolute value than—that of the chromatic dispersion of the transmission fibre (see [7]). Hence the modulus of $\tilde{d}(z)$ is not small and $\tilde{d}(z)$ is a periodic function with period z_a, whilst $\langle d\rangle$ is close to zero. This means that on a lengthscale of significantly higher order than Z_a we have, in addition to the condition on $(\langle c\rangle^{-1}\tilde{c})$, that $\tilde{d}$ is a rapidly oscillating function and also that we have $\max|\langle d\rangle^{-1}\tilde{d}| \gg 1$.

In these two cases we are dealing with time-dependent rapidly oscillating perturbations of the NLS. The two cases are equivalent to NRO(0) and DRO(1/r) respectively, due to the fact—following the remark immediately after (2.24) and as is apparent from Section 2—that successive averaging of different t-dependent terms in the Hamiltonian is the same as direct averaging of all terms at one go.

References

[1] Landau L D and Lifschitz E M 1960 *Mechanics* (Oxford: Pergamon)
[2] Goldstein H 1950 *Classical Mechanics* (Reading, MA: Addison-Wesley)
[3] Arnold V I 1988 *Dynamical Systems* vol III (Berlin: Springer)
[4] Hasegawa A and Kodama Y 1990 Guiding-center soliton in optical fibers *Opt. Lett.* **15** 1443–5
Hasegawa A and Kodama Y 1990 Guiding-center soliton *Phys. Rev. Lett.* **66** 161
[5] Kodama Y 1985 *Physica* D **16** 14–26
[6] Hori G 1966 Theory of General perturbations with unspecified canonical variables *Publ. Astron. Soc. Japan* **18** 287–96
[7] Gabitov I R and Turitsyn S K 1995 Averaged pulse dynamics in the cascaded transmission system with passive dispersion compensation *Opt. Lett.* **21** 327
Gabitov I R and Turitsyn S K 1995 Breathing solitons in optical fiber links *Pisma JETP* **63** 814–19
[8] Arnold V I 1989 *Mathematical Methods of Classical Mechanics* (New York: Springer)
[9] Abraham R and Marsden J E 1978 *Foundations of Mechanics* (Reading, MA: Benjamin)
[10] Faddeev L D and Takhtajan L A 1987 *Hamiltonian Methods in the Theory of Solitons* (Berlin: Springer)