JOURNAL OF COMPUTATIONAL PHYSICS 120, 15–32 (1995)

A Hyperasymptotic Perturbative Method for Computing the Radiation Coefficient for Weakly Nonlocal Solitary Waves

JOHN P. BOYD

Department of Atmospheric, Oceanic & Space Sciences, University of Michigan, 2455 Hayward Avenue, Ann Arbor, Michigan 48109

Received September 9, 1994

We offer a new computational method to calculate the radiation coefficient α of a weakly nonlocal solitary wave in the limit that the amplitude ε of the core of the structure tends to zero. (A weakly nonlocal solitary wave is a finite amplitude, nondissipative wave which decays to a sinusoidal oscillation of amplitude α at large distance from the core.) The multiple scales series for the solitary wave in powers of ε is asymptotic but divergent. When truncated at optimal order ("superasymptotic" approximation), the series gives no information about the radiation coefficient α (except order-of-magnitude) because the absolute error is $O(\alpha)$. We describe two "hyperasymptotic" approximations which add a second, *different* asymptotic series to the multiple scales expansion so as to compute α itself. The zeroth order hyperasymptotic approximation has a relative error of $O(\varepsilon)$ in α; the sequence of approximations may be extrapolated to give the proportionality constant in the limiting asymptotic expression for α as $\varepsilon \Rightarrow 0$. The first-order hyperasymptotic approximation has a relative error of only $O(\varepsilon^2)$; its extrapolation gives the $O(\varepsilon)$ correction to the limit as $\varepsilon \Rightarrow 0$.

1. INTRODUCTION

A weakly nonlocal solitary wave, like a soliton of the Korteweg–deVries equation, is a coherent structure which steadily translates without change of shape or form with a core which is similar in shape to the hyperbolic secant function. The difference is that classic solitary waves decay to zero as $|x| \Rightarrow 0$. The weakly nonlocal soliton decays instead to a quasi-sinusoidal oscillation of amplitude α where α is the "radiation coefficient." Such weakly nonlocal solitary waves arise in a great diversity of forms including bell solitons, breathers and envelope solitons and a wide range of applications including water waves, particle physics, fiber optics telecommunications, condensed matter physics, and meteorology [1–5].

The radiation coefficient α is important because it measures the lifetime of the solitary wave with respect to radiative decay. Unfortunately, α is very hard to calculate. For the three examples described here and many others, it is easy to derive power series in ε, where ε is the amplitude of the core, with coefficients which are polynomials in hyperbolic functions. Unfortunately, these multiple scales series cannot *directly* yield the radiation coefficient, or indeed even offer a hint of the "far field oscillations" or "wings" whose amplitude is α (Fig. 1).

One reason is that the ε-power series is asymptotic but *divergent* for all $\varepsilon \neq 0$. When truncated at optimal order, as explained below, the minimal error of the power series is $O(\alpha)$, the same order of magnitude as the very quantity α we wish to calculate.

Furthermore, the radiation coefficient is proportional to $\exp(-q/\varepsilon)$ for some constant q. The coefficients of the power series of a function with respect to the origin are proportional to the derivatives of the function at $\varepsilon = 0$. However, all ε derivatives of $\exp(-q/\varepsilon)$ are *zero* at $\varepsilon = 0$, so exponential functions of $1/\varepsilon$ have only the trivial power series whose coefficients are all zero. Such functions (and therefore the radiation coefficient of a nonlocal solitary wave) are *invisible* to the machinery of Taylor expansions about $\varepsilon = 0$.

Another way of arriving at the same pessimistic conclusion is to observe that an exponential function of $1/\varepsilon$ decreases faster than ε^N for *any* finite exponent m as $\varepsilon \Rightarrow \infty$. Mathematically speaking,

$$\exp(-q/\varepsilon) \ll \varepsilon^N, \text{ if } \varepsilon \log(\varepsilon) \ll -q/N, \tag{1.1}$$

which is a condition that can be satisfied for arbitrary finite N for sufficiently small ε. As the perturbation parameter ε decreases, the radiation coefficient "drops off the radar screen" of the ε-power series. By taking ε sufficiently small, the leading error in the N-term power series, which is proportional to ε^{N+1}, can be made as large as we please relative to α. It is unreasonable to expect that the truncated series can directly approximate α when there are other corrections which are so much larger.

In the language of Segur and Kruskal [13], the radiation coefficient lies "beyond all orders" in the perturbation theory in the sense of (1.1): it is smaller than any finite power of ε for sufficiently small ε.

Nevertheless, it is possible to calculate α in a variety of ways. A small cottage industry of perturbation theory "beyond all orders" has developed as catalogued in the NATO Workshop Proceedings edited by Segur, Tanveer, and Levine [9].

The various strategies have a common theme: to recast the problem in such a way that α is *not* exponentially small com-

0021-9991/95 $12.00

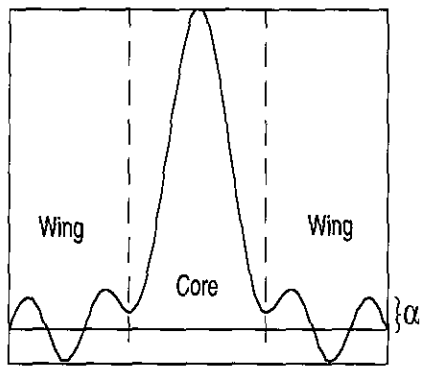

FIG. 1. Schematic of a weakly nonlocal solitary wave. The radiation coefficient α is the amplitude of the oscillatory "wings."

pared to the quantities calculated. The most widely used strategy is to apply matched asymptotic expansions in the complex plane [13–15]. As one moves away from the real x-axis, the "wings" of the solitary wave grow exponentially. In a neighborhood of the off-the-real-axis singularities of the lowest order of the ε-power series, one can match expansions where the "wings" are $O(1)$. One can then extrapolate back to the real axis. Akylas and Yang [16] described an alternative which converts the problem to an integral equation after taking the Fourier transform.

In this work, we develop a novel method which is based on "beyond all orders" methods for correcting asymptotic series for integrals developed by Berry [6–8], Paris and Wood [10, 11], Daalhuis [12], Olver [26], and others too numerous to mention. The guiding principle is that the coefficients of a power series, even if divergent, nevertheless encode information from which α can be computed. In this article, we shall describe a simple procedure for decoding the radiation coefficient.

We will postpone a comparison with competing techniques until the final section, but the algorithm described below offers some interesting insights into asymptotic series. It is also simpler in the sense that it does not use complex arithmetic, Borel summation, or an integral equation.

To explain our algorithm, we need to review some basic concepts in asymptotic series.

Optimal Truncation Rule. For an asymptotic series in the small parameter ε, accuracy is maximized (for a given fixed ε) by truncating the series with the term of smallest magnitude [17], discarding all higher degree (and larger) terms.

This rule has been rigorously proved for many series and numerically seems to be a good heuristic for many others. For our three examples, the optimal truncation N is inversely proportional to ε.

One crucial point is that the error in an optimally truncated series is typically $O(\exp[-q/\varepsilon])$ for some constant q even though the terms of the series—and the usual Poincaré definition of asymptoticity—involve only powers of ε. To call attention to this exponential dependence of error on ε, Berry [7] coined the following term.

DEFINITION 1. A **superasymptotic** approximation is an optimally truncated asymptotic series. The error is $O(\exp(-q/\varepsilon)$ for some constant q where $\varepsilon \ll 1$ is the perturbation parameter.

For the three examples described here, the error in the "superasymptotic" approximation is $O(\alpha)$. That is, to say, given that the ε-power series cannot possibly capture the exponential dependence of the radiation coefficient on ε, it does the next best thing, which is to approximate everything else.

To compute the radiation coefficient from the asymptotic series, we need something better than optimal truncation, something described by the following.

DEFINITION 2. A **hyperasymptotic** approximation is one of higher order accuracy than a superasymptotic approximation, usually obtained by appending one or more terms of a *second* asymptotic series to the optimal truncation of the original asymptotic expansion [7].

What makes a "hyperasymptotic" approximation possible is that a superasymptotic expansion is not merely an approximation. It also can be used to recast the original problem into a different form in which all terms are the same order of magnitude as the radiation coefficient α. That is, if we write

$$u(x;\varepsilon) = \Delta(x;\varepsilon) + \sum_{j=1}^{N} \varepsilon^j u^{(j)}(x) \tag{1.2}$$

and choose $N = N_{\text{opt}}(\varepsilon)$, the superasymptotic approximation is to set $\Delta(x;\varepsilon) = 0$. (Here, x is the spatial coordinate in a frame of reference translating with the wave with the origin chosen so $x = 0$ at the maximum of the core.)

However, we can also substitute the decomposition (1.2) into the differential equation to obtain an exact equation for the new unknown $\Delta(x;\varepsilon)$. If we attempt to calculate $\Delta(x;\varepsilon)$ by using the same scaling and order-of-magnitude assumptions as were used to derive the $u^{(j)}(x)$, then we will fail because these no longer apply to $\Delta(x;\varepsilon)$ when $N = N_{\text{opt}}(\varepsilon)$. A hyperasymptotic approximation is possible if we reexamine the magnitudes of different terms to approximate $\Delta(x;\varepsilon)$ by a second, *different* perturbation series. A hyperasymptotic approximation is the sum of *two* perturbation series for the *same* unknown in the *same* small parameter with *different* assumptions about the scaling and ordering of terms.

To explain these ideas, we use the "forced-KdV" equation [16]

$$u_{xx} + u - u^2 = \varepsilon^2 \operatorname{sech}^2(\varepsilon x). \tag{1.3}$$

When $\varepsilon \ll 1$, the lowest order approximation is

$$u \sim \varepsilon^2 \operatorname{sech}^2(\varepsilon x). \tag{1.4}$$

It is easy to compute higher corrections because both the second derivative and the nonlinear terms are $O(\varepsilon^2)$ relative to u. It is proper to refer to the result as a "multiple scales" perturbation theory because we are implicitly assuming that the solution to all orders depends only upon the "slow" variable

$$X \equiv \varepsilon\, x. \tag{1.5}$$

At each order, we must solve a perturbative equation of the form

$$u_j(X) = r_j(X), \tag{1.6}$$

where u_j and r_j are the coefficients of ε^{2j} in the expansions of $u(x)$ and of the residual of the differential equation, respectively. (For our other examples, (1.6) is linear *differential* equation rather than a mere algebraic identity, but the principle is the same.)

The ε-power series diverges because for a fixed ε, the assumption of "slow" dependence on x breaks down for sufficiently large j. At the optimal truncation $N_{\text{opt}}(\varepsilon)$, we shall show the Fourier transform of $r_N(\varepsilon\, x)$ is peaked at $k = \pm 1$. These are precisely the wavenumbers at which the second derivative of u is the *same* magnitude as u itself, instead of being smaller as assumed by the ε-power series. A hyperasymptotic approximation *must* retain the u_{xx} term on the left-hand side of the perturbation equations to compute the *second* perturbation expansion which is added to the superasymptotic approximation.

Fortunately, the nonlinear term is still small compared to the linear terms. This makes it possible to calculate the radiation coefficient α to lowest *hyperasymptotic* order from

$$u_{xx}^{\text{hyper}} + u^{\text{hyper}} = r_{N_{\text{opt}}}. \tag{1.7}$$

As we shall show in Section 5, it is possible to compute α for any constant coefficient differential equation in closed, analytic form without obtaining the complete explicit solution to (1.7).

We shall further show that $r_{N_{\text{opt}}}$, which is a polynomial in $\text{sech}(\varepsilon\, x)$, can be approximated by a single term to compute α to lowest order, another valuable simplification.

The final step is an extrapolation to $N \Rightarrow \infty$, a simple polynomial fit in $1/N$. This gives an approximation in the form

$$\alpha \sim (\nu_0 + \nu_1\, \varepsilon) \exp\left(-\frac{\pi}{2\, \varepsilon}\right), \qquad \varepsilon \ll 1, \tag{1.8}$$

where ν_0 and ν_1 are constants. Justifying the factor of $\pi/2\ \varepsilon$ inside the exponent is fairly easy, both for our algorithm and for all the alternatives. To compute ν_0 and ν_1 is, as the poet Virgil said of returning from death: "That the labor, that the difficulty."

In the next few sections, we shall carefully explain each of these steps using (1.3) for illustration. We shall then compute ν_0 for three examples: (i) forced-KdV equation, (ii) The fifth-order KdV equation [5, 14, 15], and (iii) the third-order nonlinear Schroedinger equation [1, 18, 19].

2. MULTIPLE SCALES PERTURBATION THEORY: OPTIMAL TRUNCATION AND SUPERASYMPTOTICS

In the limit that $\varepsilon \ll 1$, the forced-KdV equation (1.3) can be approximately solved by assuming a power series expansion in ε:

$$u^{(N)}(x;\, \varepsilon) \sim \sum_{j=1}^{N} \varepsilon^{2j}\, u_j(X) \tag{2.1}$$

$$u_j(x) = \sum_{m=1}^{j} a_{jm}\, \text{sech}^{2m}(\varepsilon\, x) \tag{2.2}$$

using hyperbolic identities such as

$$\frac{d\, \text{sech}(z)}{dz} = -\text{sech}(z)\tanh(z); \quad \frac{d\, \tanh(z)}{dz} = \text{sech}^2(z); \quad \tanh^2(z) = 1 - \text{sech}^2(z) \tag{2.3}$$

and matching powers of ε. The algorithm is so simple that the entire computer program is given as Table I; cost grows as $O(N^4)$, where N is the highest order to which the expansion is carried.

At each order, we solve

$$u_j(x) = r_j(x), \tag{1.5 bis}$$

where $r_j(x)$ is the $O(\varepsilon^{2j})$ term in the residual obtained by substituting $u^{(j-1)}(x)$ into the full nonlinear equation (1.3). In words, both the second derivative and the nonlinear term are $O(\varepsilon^2)$ smaller than undifferentiated linear term, collapsing the perturbation theory into a trivial recursion.

The perturbation theory is thus simultaneously an expansion for small amplitude and for large length scale. It is legitimate to refer to the expansion (2.1) as a "multiple scales" perturbation theory [17, 20, 21] because it is the assumption of a large length scale—not the nonlinearity—which is responsible for the divergence of the series [22]. Even when quadratic term u^2 is dropped, the ε-power series for the resulting *linear* equation is still *divergent* [22].

Figure 2 illustrates the error in the series (2.1) as a function of degree j for three different ε. One observes: (i) There is a minimum error for each fixed ε. (ii) The degree j of the minimum increases as ε decreases (iii) The minimum error is equal to the radiation coefficient $\alpha(\varepsilon)$.

The reason that the perturbative series diverges is that the neglect of the second derivative on the left-hand side of (1.5) implicitly assumes that the solution is only a function of the "slow" spatial variable

TABLE I

MATLAB Code to Compute the ε-Power Series for the Forced-KdV Equation

$$u^{(N)}(x) = \sum_{j=1}^{N} \varepsilon^{2j} \sum_{m=1}^{j} a_j \operatorname{sech}^{2m}(\varepsilon x)$$

```
a(1, 1) = 1;  a(2, 1) = −4;  a(2, 2) = 7;  % Initialization
for j = 3: N  % N = user-chosen maximum order
        % Second derivative contributions
        a(j, 1) = − 4 * a(j − 1, 1);
        a(j, j) = (2 * j − 2) * (2 * j − 1) * a(j − 1, j − 1);
        for m = 2: (j − 1)
        a(j, m) = − 4 * m * m * a(j − 1, m) + (2 * m − 2) * (2 * m − 1) * a(j − 1, m − 1);
        end
        % Nonlinear contributions
        for k = 1: (j − 1)
           for m = 1: k
              for p = 1: (j − k)
              a(j, m + p) = a(j, m + p) + a(k, m) * a(j − k, p);
              end
           end
        end
end
```

$$X \equiv \varepsilon x \tag{2.5}$$

and not of the "fast" scale of the homogeneous solutions of the linear part of the different equation, $\cos(x)$ and $\sin(x)$, which vary on a unit length scale. Unfortunately, this assumption of "only-slow" dependence is not entirely true.

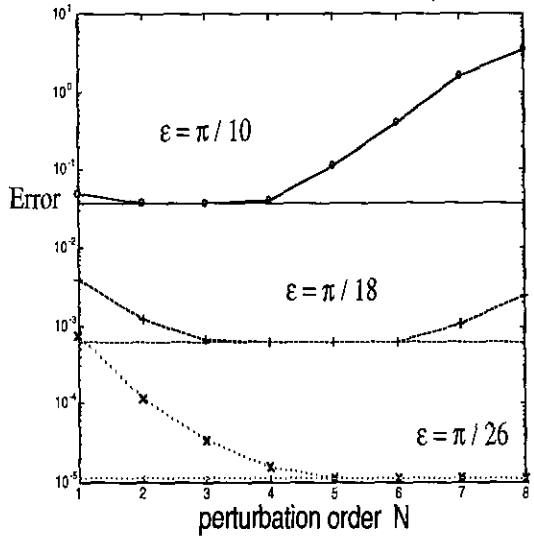

FIG. 2. Maximum pointwise error in the multiple scales perturbation theory as a function of perturbation order N for three different values of ε: solid-with-circles, $\varepsilon = \pi/10$ ($N_{opt} = 3$); dashed-with-pluses, $\varepsilon = \pi/18$ ($N_{opt} = 5$); dotted-with-x's, $\varepsilon = \pi/26$ ($N_{opt} = 7$). The horizontal lines denote the radiation coefficient α for each of three values of ε.

The Fourier transform of $u(x)$ is helpful to see this, where the transform is defined by

$$U(k) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(x) \exp(-ikx)\, dx. \tag{2.6}$$

Transform theory shows that if $u(x)$ decays exponentially for large $|x|$ (or asymptotes to a sinusoidal oscillation) and is analytic in the strip $|\mathrm{Im}(x)| < q$, then its transform $U(k)$ will decay as $\exp(-q|k|)$. Inserting ε shows that the transform of $u(\varepsilon x)$ must decay as $\exp(-q|k|/\varepsilon)$. At $k = 1$, the wavenumber of the far field oscillations, $U(k)$ is proportional to $\exp(-q/\varepsilon)$. Therefore, the dependence of $u(\varepsilon x)$ on the "fast" length scales, equivalent in the Fourier transform to $k \sim O(1)$, or larger, is *exponentially small* in $1/\varepsilon$, but *not zero.*

Retaining more terms in the ε-power series reduces the error in the *slow* components of the solution, i.e., $|k| < 1$. Unfortunately, the error is *magnified* in the higher wavenumbers [1, Chap. 2]. Consequently, the Fourier transform of $r_{N_{opt}}(x)$ $[= u_{N_{opt}}(x)]$ is peaked not at $k = 0$, but rather at $k = \pm 1$ as shown in Fig. 3. For $N > N_{opt}$, the peaks of the transform of $r_N(x)$ move to larger and larger $|k|$.

It is for this reason that adding more terms (beyond $j = N_{opt}$) is counterproductive: The assumption that $u_{j,xx} \sim O(\varepsilon^2)\, u_j$ fails utterly at $j = N_{opt}$. The hyperasymptotic approximation *must* include the second derivative as part of the operator on the left-hand side of the perturbative equation, as in Eq. (1.7).

Figure 3 suggests a second definition of the optimum perturbation order: The order is optimum when the Fourier transform

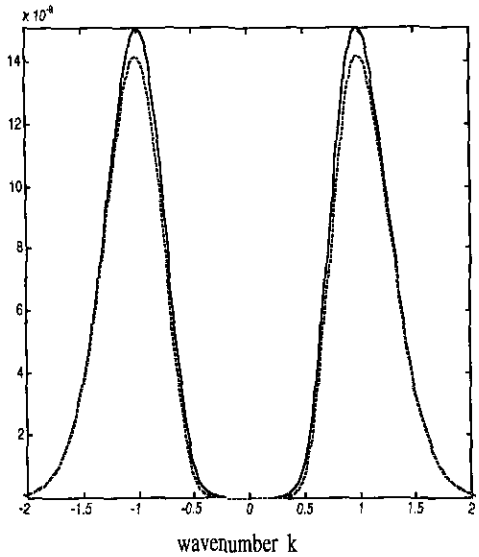

FIG. 3. (Forced-KdV equation). Absolute value of the Fourier transforms for (solid curve) the eighth-order residual, which is also $u^{(8)}$, and for its "leading" approximation (dashed), which is proportional to the 14th derivative of $\mathrm{sech}^2(\varepsilon\, x)$.

of the Nth order residual $r_N(x)$ is peaked at the wavenumber of the far field oscillations [i.e., at $k = 1$ for the forced-KdV equation].

Unfortunately, there is no simple formula for the Fourier transform of $r_N(x)$ *per se*. However, as will be shown later, the *dominant* term in the Nth order residual is equal to a constant times the $(2N\text{-}2)$th derivative of $\mathrm{sech}^2(\varepsilon\, x)$. (The Fourier transform of this "leading" term is the dashed curve in Fig. 3; note how well it approximates the Fourier transform of the whole.) Since differentiation with respect to x multiplies the Fourier transform by $(i\ k)$, it follows that Fourier transform of the leading term in $r_N(x)$ is proportional to

$$k^{2N-1}\exp\left(-\frac{\pi}{2\varepsilon}k\right), \tag{2.7}$$

where we have approximated $1/\sinh(\pi\, k/(2\,\varepsilon))$ by the exponential for $\varepsilon \ll 1$. The transform of the leading term is peaked at $k = 1$ when

$$N_{\mathrm{opt}} = \frac{1}{2} + \frac{\pi}{4\varepsilon} \leftrightarrow \varepsilon_{\mathrm{opt}} = \frac{\pi}{4(N - 1/2)}. \tag{2.8}$$

The right equation of (2.8) was used to pick the ε values for which $N = 3$, 5, and 7 should be optimum; the minima in Fig. 2 are indeed at the predicted N.

It turns out that knowledge of the optimum perturbation order as a function of ε is *not* necessary to compute ν_0, only to compute its $O(\varepsilon)$ correction ν_1. In any event, the minima in error as a function of order are shallow and broad. Of much greater importance is the fact that the transform of the residual is not peaked about $k = 0$, but rather at the wavenumber of the far field oscillations. This defines the new scaling, the alternative method of approximation, which is essential in passing beyond the limitations of the superasymptotic approximation.

3. NEWTON'S ITERATION: $O(\alpha^2)$ IN ONE ITERATION

Newton's iteration is the standard numerical method for nonlinear problems. The differential (or algebraic) equation is linearized with respect to the current iterate to obtain an equation for the correction. For one-dimensional problems, each Newton's iteration requires the solution of a linear, variable coefficient ordinary differential equation. (This linearized ODE is often called the "Newton–Kantorovich" equation where the second name honors the Nobel-winning economist who proved that Newton's method could be applied to differential equations.)

For the forced-KdV equation, the iteration is

$$\Delta^l_{xx} + \Delta^l - 2u^l(x)\,\Delta^l = r^l(x) \equiv$$
$$-\{u^l_{xx} + u^l - (u^l)^2 - \varepsilon^2\mathrm{sech}^2(\varepsilon\, x)\} \tag{3.1a}$$
$$u^{l+1}(x) = u^l + \Delta^l. \tag{3.1b}$$

We repeat this until the correction Δ^l is satisfactorily small.

Sufficiently close to a solution, Newton's iteration has "digit-doubling" convergence. That is to say, the error after each iteration is the *square* of the error at the previous iteration (in order-of-magnitude).

This implies that if the initial iterate is a superasymptotic approximation, which has an error $O(\alpha)$, then a *single* Newton's iteration will reduce the error to $O(\alpha^2)$. Since α is exponentially small, it follows that solving one *linear* equation of the form of (3.1) will calculate the radiation coefficient α to within a *relative* error of $O(\alpha)$, i.e., a relative error which is exponentially small in $1/\varepsilon$.

Figure 4 confirms this prediction. (The order $N_{\mathrm{opt}}(\varepsilon)$ of the superasymptotic perturbation approximation is defined by Eq. (2.8).) Indeed, the error after a single Newton iteration is actually a little smaller than α^2.

Thus, the superasymptotic approximation has made Newton's method *non*-iterative in the sense that one solution of the Newton–Kantorovich equation will suffice for all practical purposes (if $\varepsilon \ll 1$). Unfortunately, (3.1a) is a *variable* coefficient equation. The secret of our new algorithm is to approximate the true Newton–Kantorovich linearized differential equation (3.1a) by something simpler.

4. A QUASI-NEWTON METHOD: APPROXIMATING THE NEWTON–KANTOROVICH EQUATION BY A CONSTANT COEFFICIENT ODE

In practical arithmurgy, Newton's iteration is often prohibitively expensive. The main culprit is the need to compute and

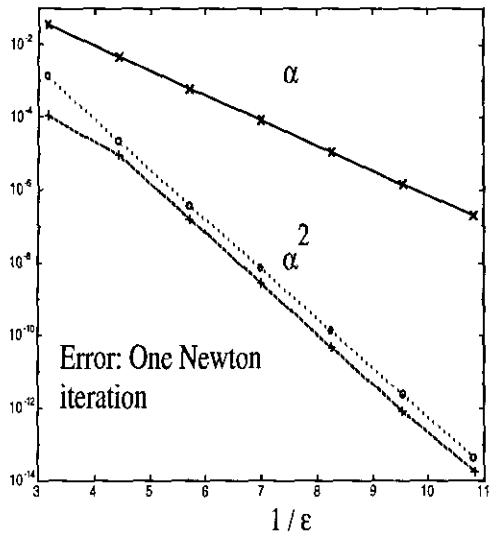

FIG. 4. Solid-with-x's; $\alpha(\varepsilon)$; dotted-with-circles, $\alpha^2(\varepsilon)$; dashed-with-pluses, error in α after one Newton iteration, beginning from a first guess which is the optimal perturbation theory. Forced-KdV equation spatial period = 63 $\pi/100$ collocation points.

factor the Jacobian matrix, which is the discretization of the (always linear) differential operator of the Newton–Kantorovich equation. Most books on solving algebraic systems of equations are really tomes on so-called "quasi-Newton" methods, a catch-all that embraces a great diversity of algorithms that use Newton's linearization but cheat on the Jacobian.

Our quasi-Newton strategy is to *neglect* the *variable coefficient term* in Newton's iteration, that is, to simplify the Newton–Kantorovich equation to the *constant coefficient,* inhomogeneous problem

$$\Delta^l_{xx} + \Delta^l = r^l(x). \tag{4.1}$$

This is a profound simplification because in a Fourier–Galerkin representation, the discretization of the simplified differential operator is a *diagonal* matrix. A single iteration of (4.1) requires only two fast Fourier transforms (FFTs) at a cost of $O(M \log_2 M)$ operations per iteration instead of the $O(M^3)$ cost of factorizing a large, dense spectral matrix [21], where M is the number of Fourier coefficients in the truncated expansion of $u(x)$. Even though the rate of convergence is much slower than an unmodified Newton iteration, the cost per iteration is reduced so drastically that the quasi-Newton approximation is cheaper. (It is for similar reasons that quasi-Newton algorithms have become so popular in numerical analysis.)

The obvious reason to neglect $-2u^l(x)\,\Delta^l(x)$ is that $u(x)$ is $O(\varepsilon^2)$, so this term is apparently $O(\varepsilon^2)$ in comparison to the undifferentiated term on the left in (4.1). (Indeed, this is the same rationale for why the equation which is solved at each order of the multiple scales perturbation theory, (1.6), is also free of this term.) Reality is a little more complicated.

First, our primary goal is the radiation coefficient α. Because α is exponentially small in $1/\varepsilon$, it is possible for an approximation to have small absolute error for $u(x)$ and yet still miss the radiation coefficient. (Indeed, this is precisely the failing of the multiple scales series, which approximates $u(x)$ to $O(\varepsilon^N)$, but approximates α by zero for all orders N!). Even when the second derivative is retained, as in the quasi-Newton approximation, disaster is still possible. For example, Akylas and Yang [16] show that the relative error in α made by solving the "linear" approximation

$$u_{xx} + u = \operatorname{sech}^2(\varepsilon x), \tag{4.2}$$

instead of the full forced-KdV equation, is about 54% for all small ε—an $O(1)$ error.

It *is* true, however, that when the RHS of the Newton–Kantorovich equation is the *residual* of the *optimally truncated* multiple scales series, one can neglect the $O(\varepsilon^2)$ linearized term and obtain an accurate approximation to α as $\varepsilon \Rightarrow 0$. For this special case, the inhomogeneous term in the differential equation is itself of $O(\alpha)$. It follows that an $O(\varepsilon)$ relative error in solving this differential equation yields an absolute error of $O(\alpha\,\varepsilon)$, or in other words a relative error of $O(\varepsilon)$ in α itself.

This claim of $O(\varepsilon)$ error in $\Delta^1(x)$ can be justified by an "envelope" perturbation theory (Appendix). However, *experimentally,* the relative error in the *radiation coefficient* is actually $O(\varepsilon^2)$.

Figure 5a shows the maximum pointwise error in $\Delta^1(x)$ (top curve) and also the error in the radiation coefficient α (bottom). On this log–log plot, a quantity which is proportional to a power of ε should approximate a straight line. Both curves in Fig. 5a asymptote to lines, but the slope of the error in the radiation coefficient in approximately double that for the maximum error in $\Delta^1(x)$. Figure 5b shows the relative error in α divided by ε^2. This ratio varies roughly linearly with ε, showing that for the indicated range of ε,

$$\left|\frac{\alpha - \alpha^{\text{one-quasi-Newton}}}{\alpha}\right| \approx \varepsilon^2 - \tfrac{3}{2}\varepsilon^3, \qquad \varepsilon \ll 1 \text{ (empirical curve-fit).} \tag{4.3}$$

This happy difference between the two errors arises because the maximum absolute error in $\Delta^1(x)$, which is $O(\varepsilon\,\alpha)$, occurs in the *core* of $u(x)$. The error in $\Delta^1(x)$ for *large* $|x|$ is $O(\varepsilon)$ smaller, and this translates to an $O(\varepsilon^2\,\alpha)$ absolute error in α.

One final comment: although all the illustrations in this article use just *one* quasi-Newton iteration, it is quite acceptable (in a numerical, non-hyperasymptotic computation) to iterate many times with a $O(\varepsilon)$ reduction in error at each step. Since $\alpha \sim O(\varepsilon^{N_{\text{opt}}(\varepsilon)})$, N iterations reduce the error from α to (roughly) α^2. N quasi-Newton iterations are equivalent to a single standard Newton's iteration.

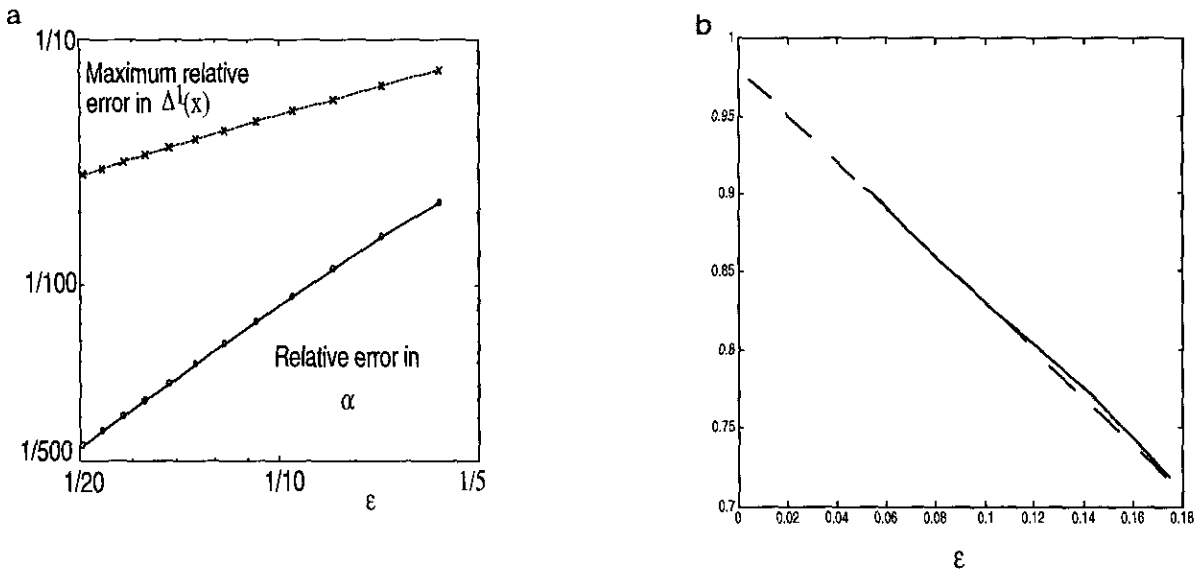

FIG. 5. (a) Relative error, $|\alpha_{exact} - \alpha_{cc}|/\alpha_{exact}$, versus ε is shown as the solid curve with circles, where α_{cc} is the error made in replacing the Newton–Kantorovich equation by a constant coefficient equation. The top curve (dashed with x's) shows the maximum error, again relative to α, in $\Delta^1(x)$. Both errors are after a single quasi-Newton iteration, begun from a superasymptotic approximation. (b) Solid curve: Relative error in α divided by ε^2, after one quasi-Newton iteration begun from a superasymptotic approximation. Dashed line: linear interpolant of the first and last data points; this shows that the error/ε^2 is well-approximated by a straight line, or equivalently, that $|\alpha_{exact} - \alpha_{cc}|/\alpha_{exact} \approx \varepsilon^2 - \frac{8}{2}\varepsilon^3$.

5. COMPUTING THE RADIATION COEFFICIENT FROM THE SOLUTION TO A CONSTANT COEFFICIENT DIFFERENTIAL EQUATION

In the previous section, we showed that it was a good approximation [$O(\varepsilon^2)$ for the radiation coefficient α] to simplify the Newton–Kantorovich equation. Here, we show how to compute α.

For generality, assume that the linear, constant coefficient equation may be written in the generic form

$$P\left(-i\frac{\partial}{\partial x}\right)u(x) = f(\varepsilon x), \tag{5.1}$$

where $P(z)$ is a polynomial and $f(\varepsilon x)$ is the inhomogeneous term, which in applications will be the residual of the optimally truncated perturbation series. Let the Fourier transforms of $u(x)$ and $f(x)$ be denoted by $U(k)$ and $F(k)$, respectively, where

$$F(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-ikx)f(x)\,dx \tag{5.2}$$

and similarly for $U(k)$. Taking the transform of both sides of (5.1) gives the formal solution

$$u(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(ikx)\frac{F(k/\varepsilon)}{\varepsilon P(k)}\,dk. \tag{5.3}$$

We assert (and prove below) that the radiation coefficient comes from the points where the integrand is singular, that is, where $P(k) = 0$.

We make two assumptions. First, $f(x)$ is a function which (i) decays exponentially fast for large $|x|$ and (ii) whose singularities (poles, branch points, etc.) are separated from the real x-axis by a finite distance, and these properties apply to its Fourier transform $F(k)$, too. (It is easy to prove these properties when $f(\varepsilon x)$ is a polynomial in sech(εx), as true of the residuals of the forced-KdV, FKdV, and TNLS equations.) Second, we assume that $P(k)$ has m *simple* roots along the real k-axis.

We can then derive explicit formulas for the radiation coefficients by adding and subtracting functions whose denominators vanish at the real roots of $P(k)$, thereby splitting $u(x)$ into two parts:

$$u(x) = u_{\text{local}}(x) + u_{\text{far field}}(x), \tag{5.4}$$

where, writing P' for dP/dk,

$$u_{\text{local}}(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{\exp(ikx)}{\varepsilon}\left\{\frac{F(k/\varepsilon)}{P(k)} - \sum_{j=1}^{m}\frac{F(k_j/\varepsilon)\exp(-\lambda(k-k_j)^2)}{P'(k_j)(k-k_j)}\right\}dk \tag{5.5}$$

$$u_{\text{far field}}(x) = \sum_{j=1}^{m}\frac{1}{P'(k_j)}\frac{F(k_j/\varepsilon)}{\varepsilon\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(ikx)\frac{\exp(-\lambda(k-k_j)^2)}{(k-k_j)}\,dk, \tag{5.6}$$

TABLE II

The Radiation Coefficient for Linear, Constant Coefficient Differential Equations of Various Forms

Eq.	
fKdV	$u_{xx} + K^2 u = f(\varepsilon x) \Rightarrow u \sim \alpha\ \mathrm{sgn}(x)\sin(Kx),\ \|x\| \to \infty$ $\alpha = \frac{1}{K}\int_0^\infty \cos(Kx) f(\varepsilon x)\, ds$
FKdV	$u_{4x} + u_{xx} = f(\varepsilon x) \Rightarrow u \sim \alpha\ \mathrm{sgn}(x)\sin(x),\ \|x\| \Rightarrow \infty$ $\alpha = -\int_0^\infty \cos(x) f(\varepsilon x)\, dx$
TNLS	$-\frac{1}{2}\varepsilon^2 Q + i\varepsilon^1 Q_x + \frac{1}{2}Q_{xx} - iQ_{xxx} = f(\varepsilon x) \Rightarrow Q \sim \alpha\, i\ \mathrm{sgn}(x)\exp(-ix/2),\ \|x\| \to \infty$ $\alpha = -2\int_{-\infty}^\infty \exp(ix/2) f(\varepsilon x)\, dx$

where λ is a non-negative constant. By construction, $u_{\text{local}}(x)$ is the Fourier transform of a function of k which is analytic everywhere within a finite strip about the real k-axis and is exponentially decaying for large $|k|$. A standard transform theorem then implies that $u_{\text{local}}(x)$ must decay exponentially fast with $|x|$. Thus, $u_{\text{local}}(x)$ is spatially localized; the far field oscillations are contained entirely within $u_{\text{far field}}(x)$, which by construction contains all the poles of the integrand of $u(x)$ on the real k-axis.

By using the identity

$$\int_{-\infty}^{\infty} \exp(ikx)\frac{\exp(-\lambda(k-k_j)^2)}{k-k_j}\,dk = i\pi \exp(i k_j x)\,\mathrm{erf}\left(\frac{x}{2\sqrt{\lambda}}\right) \tag{5.7}$$

and recalling that

$$\mathrm{erf}(x) \sim \mathrm{sgn}(x), \quad |x| \to \infty, \tag{5.8}$$

we can easily deduce the asymptotic form of the solution to the constant coefficient differential equation (Table II).

For the forced-KdV equation, polynomial $P(k) = K^2 - k^2$, that is, the linearized equation is

$$u_{xx} + K^2 u = f(\varepsilon x), \tag{5.9}$$

where K is the far field wavenumber ($K = 1$ in the rest of the article). There are two roots at $k = \pm K$; the contributions can be added to give the real-valued asymptotic solution which is listed in the table.

For the FKdV equation, there is one modest subtlety. Its linearized, fourth-order form is

$$u_{4x} + u_{xx} = f(\varepsilon x). \tag{5.10}$$

The quartic polynomial $P(k)$ is

$$P(k) = k^4 - k^2. \tag{5.11}$$

The two roots at $k = \pm 1$ give a far field identical to that of (5.9), except for an overall sign. However, there is also a double root at $k = 0$. This is an artifact of the neglect of the $-c\,u$ term. If we included this $O(\varepsilon^2)$ correction, the roots at $k = \pm 1$ would be modified only by $O(\varepsilon^2)$, which is consistent with our other approximations; these roots remain on the real axis. The double root at $k = 0$, however, is shifted *off* the real axis to a pair of roots on the imaginary axis at $k \approx \pm i\,\varepsilon$. The contributions from these roots are not sinusoidal in the far field, but rather decay exponentially on an $O(1/\varepsilon)$ length scale. Therefore, only the two roots of $P(k)$ at $k = \pm 1$ contribute to the far field of the FKdV nanopteron, and only these contributions are included in Table II (with neglect of the $-cu$ correction).

The linear, constant coefficient equation derived from the third-order non-linear Schroedinger (TNLS) equation is

$$-\tfrac{1}{2}\varepsilon^2 Q + i\varepsilon^2 Q_x + \tfrac{1}{2}Q_{xx} - iQ_{xxx} = f(\varepsilon x). \tag{5.12}$$

The associated polynomial is a cubic which can be exactly factored:

$$P(k) = -\tfrac{1}{2}\varepsilon^2 - \varepsilon^2 k - \tfrac{1}{2}k^2 - k^3 = -(k + \tfrac{1}{2})(k^2 + \varepsilon^2). \tag{5.13}$$

The two complex roots are irrelevant to the far field. The single contribution from $k = -1/2$ gives asymptotic oscillations which are complex-valued, but this is consistent with the complex-valued coefficients and solution of the TNLS equation.

It is important to note that the splitting of $u(x)$ into two

pieces and the asymptotic solutions listed in Table II involve no approximations (except taking large $|x|$). Except for (5.8), every equation in this section is *exact* for the *constant coefficient, linear* differential equation. When we apply Table II to perturbation theory, we will obtain only an approximation. The error is not due to the (non-existent) approximations of this section, but rather the approximations of (i) neglecting the variable coefficients in the Newton–Kantorovich equation (Section 4) and (ii) the simplifications of the residual function $f(\varepsilon\, x)$, which shall be justified in the next section.

6. SIMPLIFYING THE RESIDUAL

When the first guess for Newton's iteration is $u^{(N)}(x)$, the total residual,

$$r_{\text{total}}(x;\, u^{(N)}) \equiv -\{u_{xx}^{(N)} + u^{(N)} - (u^{(N)})^2 - \varepsilon^2 \operatorname{sech}^2(\varepsilon\, x)\}, \quad (6.1)$$

is a mess for large N. For example, the residual of $u^{(4)}(\varepsilon\, x)$ is, with the abbreviation $s = \operatorname{sech}(\varepsilon\, x)$,

$$\begin{aligned} r_{\text{total}}(u^{(4)}) &= \varepsilon^{10}\{256s^2 - 39808s^4 + 370880s^6 - 825840s^8 \\ &\quad + 507206s^{10}\} \\ &+ \varepsilon^{12}\{768s^4 - 25088s^6 + 130080s^8 - 221704s^{10} + 119266s^{12}\} \\ &+ \varepsilon^{14}\{96768s^6 - 1005312s^8 + 3497568s^{10} - 4666144s^{12} \\ &\quad + 2102100s^{14}\} \\ &+ \varepsilon^{16}\begin{Bmatrix} 4096s^4 - 313344s^6 + 7115008s^8 - 43801728s^{10} \\ +110293024s^{12} - 119683200s^{14} + 46580625s^{16} \end{Bmatrix}. \end{aligned} \quad (6.2)$$

For general N, the residual is a bivariate polynomial in ε and s with integer coefficients,

$$r(u^{(N)}) = \sum_{j=N+1}^{2N} \varepsilon^{2j} \sum_{k=1}^{2N} r_{N,jk} \operatorname{sech}^{2k}(\varepsilon\, x), \quad (6.3)$$

where the $r_{N,jk}$ are constants. The number of terms grows quadratically with N. It follows that simplifying the residual is highly desirable.

Since the radiation coefficient α is merely the Fourier transform of the residual, evaluated at the resonant wavenumbers ($k = \pm 1$ for the forced KdV equation), the test of a good approximation to the residual must be that the approximation's Fourier transform agrees closely with the transform of the residual at the resonant k.

The Fourier transform of the residual at any order can be expressed in terms of

$$I_j(\varepsilon;\, k) = \int_0^\infty \varepsilon^{2j} \operatorname{sech}^{2j}(\varepsilon\, x) \cos(k\, x)\, dx \quad (6.4a)$$

$$I_j(\varepsilon;\, k) = \begin{cases} \dfrac{\pi k}{2 \sinh\left(\dfrac{\pi k}{2\varepsilon}\right)}, & j = 1 \\[2ex] \dfrac{\pi k(k^2 + 2^2\varepsilon^2)(k^2 + 4^2\varepsilon^2)\cdots(k^2 + [2j-2]^2\varepsilon^2}{2 \sinh(\pi k/2\varepsilon)(2j-1)!}, & j \geq 2 \\[2ex] \sim \dfrac{\pi k^{2j-1}}{(2j-1)!} \exp\left(-\dfrac{\pi k}{2\varepsilon}\right), & \varepsilon \ll 1/j. \end{cases} \quad (6.4b)$$

The asymptotic form suggests that only the highest powers in the residual, of the form $\varepsilon^{2j} \operatorname{sech}^{2j}(\varepsilon\, x)$, will contribute to the radiation coefficient. This is true, but the argument is shaky because ε is not small compared to $1/j$ for $j = N_{\text{opt}}$.

To see what really happens, we used the top line of (6.4b) to evaluate the contribution of each term in the residual as Table III. The second column of the table is the sum of the numbers to the right in each row; it represents the net contribution of all terms in the residual which are proportional to a given power of ε, that is, the contributions of the $r_{4,jk}$ for fixed j. We see from this column that the lowest order terms in the residual, i.e., those terms proportional to ε^{10}, are much larger than all terms proportional to higher powers of ε.

The second striking feature of the table is that the $O(\varepsilon^{10})$ terms (top row of numbers) are large and alternating in sign. Their sum, 1.53, is the small difference of large numbers. As the order N increases (not illustrated), one finds that this tendency to self-cancellation becomes more and more pronounced. It would seem that the simplest approximation with small errors is

$$r_{\text{total}}(u^{(N)}) \approx u_{N+1}(x) \quad (6.5)$$

which is equivalent to truncating the outer sum in (6.3) to $j = N + 1$. For the forced KdV equation, these coefficients of the residual are automatically computed by the ε-power series; recall that the coefficients of the expansion of $u_j(x)$ are given by a_{jk} (Eq. (2.2)),

$$r_{N,N+1,k} \equiv a_{N+1,k}. \quad (6.6)$$

Fortunately, this pessimistic assessment is false. For the linear boundary value problem

$$u_{xx} + u = \varepsilon^2 \operatorname{sech}^2(\varepsilon\, x) \quad (6.7)$$

the residual is given exactly by the $2N$th derivative of the square of the hyperbolic secant function,

$$r_{\text{total}}(u^{(N)}) = (-1)^N \varepsilon^{2N+2} \frac{d^{2N}}{dX^{2N}}\{\operatorname{sech}^2(X)\}. \quad (6.8)$$

The residual for the nonlinear generalization of (6.7) is not so

TABLE III

Contributions of Each Term in the Residual of $u^{(4)}(x)$ to α

Order	Row sum	s^2	s^4	s^6	s^8	s^{10}	s^{12}	s^{14}	s^{16}
ε^{10}	1.53607	0.0002	−0.2104	4.7854	−17.462	14.4230	—	—	—
ε^{12}	0.00675	—	0.0001	−0.0099	0.0838	−0.1920	0.1247	—	—
ε^{14}	0.00968	—	—	0.0012	−0.0197	0.0923	−0.1487	0.0759	—
ε^{16}	0.00015	—	0.0000	−0.0001	0.0043	−0.0352	0.1070	−0.1317	0.0558

Note: (i) These are scaled by multiplication by $(2/\pi)\sinh(\pi/2\varepsilon)$; (ii) $\varepsilon = 0.1745$, which is optimum for $N = 5$; (iii) $s = \mathrm{sech}(\varepsilon x)$.

simple. However, as N increases, the nonlinear term, which is the difference between the forced-KdV equation and (6.7), becomes less and less important. In Section 4, we have already justified the neglect of the quadratic term on the left-hand side of the Newton–Kantorovich equation. This suggests the alternative approximation for the forced-KdV equation

$$r_{\text{total}}(u^{(N)}) \approx \mu_N \varepsilon^{2N+2} \frac{d^{2N}}{dX^{2N}}\{\mathrm{sech}^2(X)\} \tag{6.9}$$

for some constant μ_N. This constant can be chosen in a variety of ways; our choice is to match the highest power of $\mathrm{sech}(\varepsilon x)$ in $u^{(N+1)}(x)$. A single term in the residual then completely determines the approximation (6.7). By using the identity

$$\frac{d^2}{dX^2}\{\mathrm{sech}^{2j}(X)\} = 4j^2\mathrm{sech}^{2j}(X) - 2j(2j+1)\,\mathrm{sech}^{2j+2}(X) \tag{6.10}$$

which implies that the highest coefficient of the $2N$th derivative of $\mathrm{sech}^2(X)$ is $(-1)^N(2N+1)!$, our choice for μ_N is

$$\mu_N = (-1)^N \frac{a_{NN}}{(2N-1)!}. \tag{6.11}$$

The Fourier transform of a derivative is just (ik) times the Fourier transform of the undifferentiated function. This implies that (6.7) plus the replacement of the full Newton–Kantorovich equation by a constant coefficient equation yields the zeroth-order hyperasymptotic approximation,

$$\alpha_N^{(h0)} \equiv \frac{a_{NN}}{(2N-1)!}\frac{\pi}{2\sinh(\pi/2\varepsilon)}. \tag{6.12}$$

Neglecting only higher powers of ε in the residual gives the more complicated but more accurate first-order hyperasymptotic approximation,

$$\alpha_N^{(h1)} \equiv \sum_{j=1}^{N} I_j(\varepsilon; 1)\, a_{Nj}. \tag{6.13}$$

Figure 6 compares the accuracy of these two approximations. We find that the lowest hyperasymptotic approximation has an error of $O(\varepsilon)$ while using the whole of r_N $(=u_N)$ gives an error which appears to scale as $O(\varepsilon^2)$.

The accuracy of the first-order hyperasymptotic approximation is hardly surprising because in the multiple scales theory, all the terms in the total residual which do not appear in $u_N(x)$ are *supposed* to be smaller by at least $O(\varepsilon^2)$. However, this argument is inadequate by itself for two reasons. First, when N is the optimal truncation for a given ε, the multiple scales theory is breaking down, which would seem to invalidate arguments based on the ε-power series. However, the perturbation series breaks down at $N = N_{\text{opt}}(\varepsilon)$ because the *second derivative* is no longer negligible, not because of any difficulty with the ordering of terms in the residual.

The second worry is because α is very small, small errors in the residual could have large effect on the radiation coeffi-

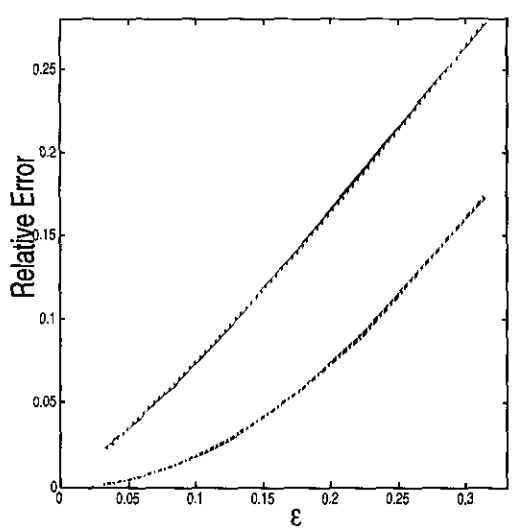

FIG. 6. (Forced-KdV equation). Solid (upper curve): $\alpha^{(h0)}$, relative to use of the total residual for $u^{(N_{\text{opt}}(\varepsilon)-1)}(x)$, where $N_{\text{opt}}(\varepsilon) = \frac{1}{2} + \pi/(4\varepsilon)$. Dotted: quadratic least squares fit to relative error where the fitted polynomial $= -0.003 + 0.72\varepsilon + 0.56\varepsilon^2$. Dashed (lower curve): error in α when the residual is approximated by all $O(\varepsilon^{2N})$ terms. Dash–Dot: quadratic fit of the error by $-0.0011 + 0.031\varepsilon + 1.68\varepsilon^2$. (In the limit $N \Rightarrow \infty$, we conjecture that only the quadratic term would survive.)

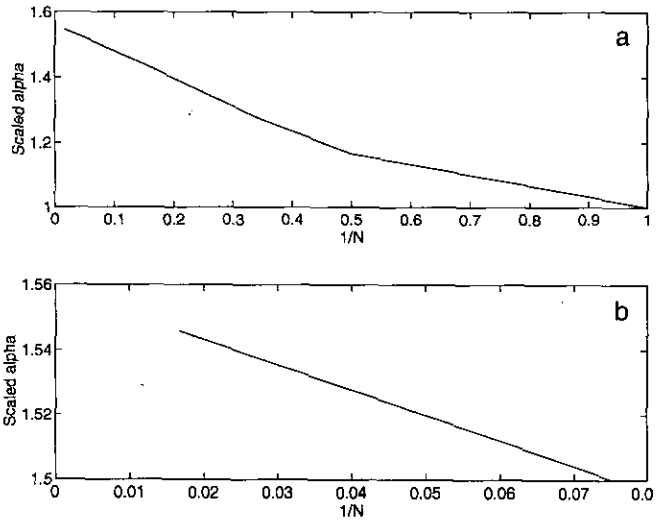

FIG. 7. (a) Scaled radiation coefficient in the zeroth-order hyperasymptotic approximation versus $1/N$, $\alpha_N/\{\pi/[2\ \sinh(\pi/(2\varepsilon))]\}$. This equals $a_{NN}/(2N-1)!$, where a_{NN} is the coefficient of $\mathrm{sech}^{2N}(\varepsilon\, x)$ in $u_N(x)$, the Nth-order term in the multiple scales perturbation theory. (b) Same as (a), but on a shorter range in $1/N$.

cient. This, too, is a false worry because at the optimal truncation, $r_{\text{total}}(u^{(N-1)}) \sim O(\alpha)$. This implies that a relative error in the residual of $O(\varepsilon^2)$ will give an error of only $O(\varepsilon^2)$ in α too.

The zeroth-order hyperasymptotic approximation (6.12) has a more heuristic justification. Nevertheless, it is useful not only because of its simplicity, but also because it contains no explicit dependence on ε except for the $\sinh(\pi/(2\,\varepsilon))$ factor. The accuracy of our formula for the optimum order N as a function of ε is irrelevant to (6.12), so long as it is correct in order of magnitude. As we shall show in the next section, this approximation generates a sequence which converges to ν_0 as $N \Rightarrow \infty$.

7. NUMERICAL RESULTS

1. The forced-Korteweg–deVries equation:

$$u_{xx} + u - u^2 = \varepsilon^2 \mathrm{sech}^2(\varepsilon\, x) \tag{1.1bis}$$

Figure 7a is a graph of the scaled zeroth-order hyperasymptotic approximation as a function of $1/N$; Fig. 7b is the same plot on a narrower interval. We plot $(\pi\ \exp(\pi/(2\ \varepsilon)\ \alpha_N(h\ 0))$ versus $1/N$, rather than N itself, because we have shown above that the error in approximating $\alpha(\varepsilon)$ via the zeroth-order hyperasymptotic approximation is $O(\varepsilon)$. Grimshaw and Joshi [15] have shown, not for (1.1bis) but for the closely related FKdV equation, that the deviation of $\alpha(\varepsilon)$ from $\nu_0 \exp(-\pi/(2\ \varepsilon))$ is also $O(\varepsilon)$. Recalling that $N_{\text{opt}} \sim O(1/\varepsilon)$, this implies that the scaled numbers shown in Fig. 7 will differ from the constant ν_0 by an amount which is $O(\varepsilon) \sim O(1/N)$. For similar reasons, we estimate ν_0 by polynomial fitting in the variable $1/N$.

Figure 7b confirms that for small $1/N$, the radiation coefficient does, indeed, vary roughly linearly with $1/N$. We can make better estimates by fitting polynomials of various orders to the large N terms of the sequence for various N:

$$\alpha \sim \pi\ 1.558823 \exp\left(-\frac{\pi}{2\,\varepsilon}\right), \quad \varepsilon \Rightarrow 0. \tag{7.2}$$

The rightmost digit was obtained by using the recurrence of Akylas and Yang [16]. This, although derived by very different means as discussed below, gives exactly the *same* sequence as our $\alpha_N(h\ 0)$. However, because they have derived an explicit recurrence, one can easily compute the sequence up to $N = 500$ without roundoff difficulties.

Our method is more general but requires more computations. Table IV shows some differences between 1.558823 and the linear coefficients in the fitted polynomials of various degrees. Computing to $N = 60$ via the algorithm in Table I required only a couple of minutes in MATLAB on an Apple Quadra 700. Roundoff does become a problem for larger N, but we still can compute ν_0 to six decimal places in 14 place arithmetic. The cubic and quartic polynomials in $1/N$ give significantly better estimates for ν_0 than the linear and quadratic fits.

The improved approximation,

$$\alpha \sim 1.558823\,(1 + \nu_1\varepsilon)\exp\left(-\frac{\pi}{2\varepsilon}\right), \quad \varepsilon \ll 1, \tag{7.3}$$

can be calculated via

$$\nu_1 = \lim_{N\to\infty} \rho_N, \tag{7.4}$$

where

TABLE IV

Differences between the Constant in a Polynomial Fit to $a_{NN}/(2N-1)!$ (Scaled Zeroth-Order Hyperasymptotic Approximation) and 1.558823, Which Is Our Best Estimate for the Proportionality Constant in the Asymptotic Form of the Radiation Coefficient for the Forced-KdV Equation

Range of N fitted by polynomial in $1/N$	Linear polynomial	Quadratic polynomial	Cubic polynomial	Quartic polynomial
16:20	0.00016	−0.00010	0.00003	0.00000
11:20	0.00037	−0.00026	0.00003	−0.00001
36:40	0.00002	0.00000	0.00000	0.00000
31:40	0.00002	−0.00001	0.00000	0.00000
21:40	0.00004	−0.00002	0.00001	0.00000
56:60	0.000004	−0.000002	0.000000	0.000000

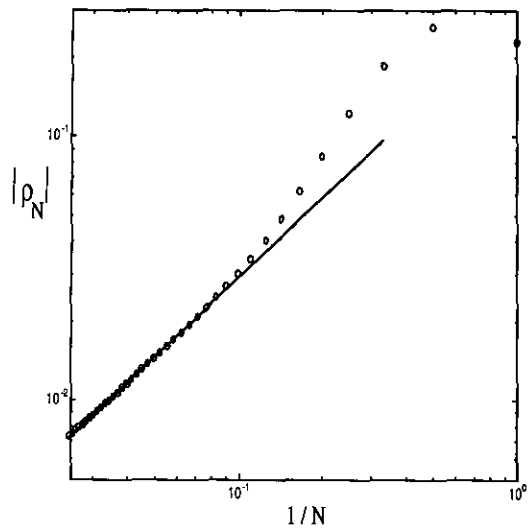

FIG. 8. (Forced-KdV equation). Circles: ρ_N plotted versus $1/N$ for $N = 1, 2, ..., 40$. The solid line is a linear fit for large N, showing that the sequence is asymptoting to this line.

$$\rho_N = \left[\frac{\alpha_N^{(h1)}}{1.558823\exp(-\pi/2\varepsilon(N)} - 1\right] \Big/ \varepsilon(N) \qquad (7.5)$$

with $\varepsilon(N)$ given by the optimal truncation rule:

$$\varepsilon = \frac{\pi}{4(N - 1/2)}. \qquad (2.8\text{bis})$$

Note that we must use the first row residual, i.e., the whole residual at each order as employed in the multiple scales theory. The one-term approximation is inadequate because it has an $O(\varepsilon) \sim O(1/N)$ error which would be added to ν_1 if we calculated (7.5) from (6.12) instead of (6.13).

Grimshaw and Joshi [15] have explicitly calculated the (nonzero) value of ν_1 for the FKdV equation (next subsection), but their improved complex-plane matched asymptotics has not yet been applied to (1.1). Thus, the extrapolation of the sequence ρ_N generates new information, hitherto uncalculated.

One technical complication is that computation in the usual 16-decimal-place floating point arithmetic gives unreliable values for ρ_N for $N > 30$ because these elements are the small difference of large terms. We solved this by evaluating the series coefficients and the integrals I_j in exact rational arithmetic in Maple, retaining π as a symbolic unknown, and then evaluating the result in high precision (25 decimal places for N up to 40) before fitting polynomials in $1/N$. The algorithm of Table I plus computation of the integrals required less than an hour on an Apple Quadra 700.

Figure 8 shows that the ρ_N decrease rapidly as $1/N$ goes to zero. The elements asymptote to a straight line on the log–log plot whose extrapolant as $1/N \Rightarrow 0$ is zero. Fitting $\{\rho_{21}, \rho_{22}, ..., \rho_{40}\}$ by polynomials in $1/N$ approximates ν_1 as the constant in the polynomial. Linear, quadratic, cubic, and quartic polynomials give -0.00015, -0.000093, -3.7×10^{-6}, and -1.8×10^{-6}, respectively.

These numbers are so small that it seems likely that

$$\nu_1 \equiv 0, \qquad (7.7)$$

but this conjecture can be proved only by analytical methods [15]. From our numerical estimates, we can say only that ν_1 is very, very small in comparison to $\nu_0 \approx 1.5588$.

2. *The Fifth-Order Korteweg–DeVries Equation*

We solve this equation in its fourth-order form,

$$u_{xxxx} + u_{xx} + \tfrac{1}{2}u^2 - cu = 0, \qquad (7.8)$$

where the phase speed c is given to all orders in ε [15] by

$$c = 4\varepsilon^2 + 16\varepsilon^4. \qquad (7.9)$$

The solution $u(x; \varepsilon)$ is expanded exactly as for the forced-KdV equation,

$$u^{(N)}(x; \varepsilon) \sim \sum_{j=1}^{N} \varepsilon^{2j} u_j(X) \qquad (2.1\text{bis})$$

$$u_j(x) = \sum_{m=1}^{j} a_{jm}\text{sech}^{2m}(\varepsilon x). \qquad (2.2\text{bis})$$

However, at each order we must solve a differential equation rather than an algebraic equation,

$$u_{N,xx} - 4\varepsilon^2 u_N = r_N. \qquad (7.12)$$

Fortunately, this can be accomplished merely by matching powers of $\text{sech}(\varepsilon x)$, just as for our other equations. The Nth-order perturbative residual is

$$r_N(x) = \sum_{k=1}^{N+1} r_{Nk}\text{sech}^{2k}(\varepsilon x). \qquad (7.13)$$

Note that the lowest order residual of $u^{(N-1)}$ is $O(\varepsilon^{2N+2})$ $r_N(x)$, not $O(\varepsilon^{2N})$ $r_N(x)$, and note also that the upper limit on the sum in (7.9) is $(N + 1)$ rather than N. (The pseudocode to compute the series is given as a table in Chapter 9 of [1].)

The zeroth-order and first-order hyperasymptotic approximations to α are

$$\alpha_N^{(h0)} \equiv -\frac{r_{N,N+1}}{(2N + 1)!} \qquad (7.14)$$

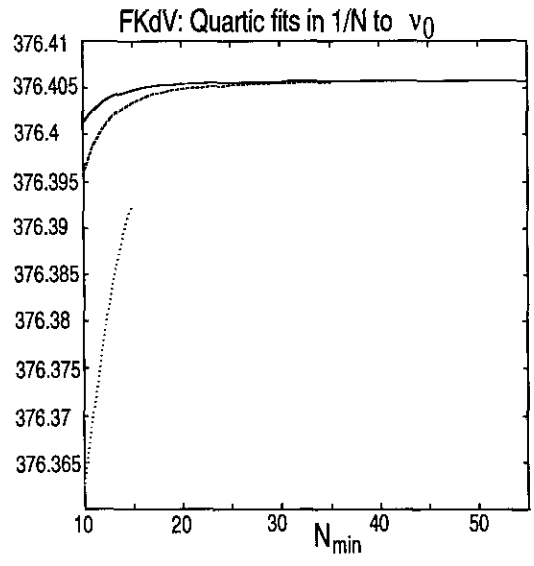

FIG. 9. The constant in the fit of a quartic polynomial in $1/N$ to the array $\alpha_N^{(h0)}(N_{min}, N_{max})$ for the FKdV equation. N_{min} is the abscissa; $N_{max} = 20$ (dotted curve), $N_{max} = 40$ (dashed), and $N_{max} = 60$ (solid curve).

$$\alpha_N^{(h1)} \equiv \sum_{j=1}^{N+1} I_j(\varepsilon; 1) a_{Nj}, \tag{7.15}$$

where we have approximated the exact far field wavenumber, $k_f = (1 + 4\varepsilon^2)^{1/2}$, by 1, consistent with the other $O(\varepsilon^2)$ errors inherent in (7.15) and where it is assumed that N and ε are related by the optimal truncation relations $\varepsilon(N) = \pi/\{4(N - \frac{1}{2})\}$, i.e., Eq. (2.8).

From (7.14), we can easily extrapolate via polynomial fitting in the polynomial argument $1/N$ that

$$\alpha \sim 376.406 \exp\left(-\frac{\pi}{2\varepsilon}\right), \quad \varepsilon \Rightarrow 0. \tag{7.16}$$

Figure 9 is a visualization of the fitting. Since low degree polynomials give less accurate approximations to the constant ν_0, both here (not shown) and for the forced-KdV equation as shown above, the graph illustrates the constant in a fourth-degree polynomial which is fit via least squares to the subsequence $\alpha_j^{(h0)}$ defined by (7.11), where $j = N_{min}, N_{min} + 1, ..., N_{max}$.

Obviously, the approximation to the prefactor in (7.14) is poorest when N_{min} is small; it is best to fit the *tail* of the sequence rather than to approximate many terms of the sequence by the least squares polynomial. As N_{max} increases, the approximation also improves as one would expect. When both N_{min} and N_{max} are large, the constant in the fitting polynomial is approximately 376.406 *independent* of the precise values of N_{min} and N_{max}.

The first-order hyperasymptotic approximation can similarly be extrapolated to compute the constant ν_1 in the improved approximation

$$\alpha \sim 376.406(1 + \nu_1\varepsilon + O(\varepsilon^2))\exp\left(-\frac{\pi}{2\varepsilon}\right), \quad \varepsilon \ll 1, \tag{7.17}$$

$$\nu_1 = \lim_{N\to\infty} \rho_N; \quad \rho_N \equiv \frac{\alpha_N^{(h1)}}{376.406 \exp(-\pi/2\varepsilon(N))} - 1/\varepsilon(N). \tag{7.18}$$

As for the forced-KdV equation, roundoff forced us to compute $\alpha_N(h\ 1)$ in Maple in exact symbolic form for $N > 20$. One must be careful to evaluate the exact Maple expressions for ρ_N, which involve only rational numbers and powers of π, with high precision floating point arithmetic. Using Maple's default precision (10 decimal places) to convert the exact ρ_N to floating point numbers gave rather large errors for large N. We found it sufficient to specify 25 decimal place accuracy for N up to 40.

Grimshaw and Joshi [15], using matched asymptotic expansions in the complex-plane, have shown that the exact value for ν_1 is

$$\nu_1 = -\pi. \tag{7.19}$$

We made a least squares fit of polynomials in $1/N$ to $\{\rho_{21}, \rho_{22}, ..., \rho_{40}\}$. Polynomials of degrees 1 to 4 give $\{-3.1497, 3.1452, -3.1417, -3.1418\}$ with relative errors of 0.25%, 0.11%, 0.0025%, and 0.09%. The linear polynomial is shown in Fig. 10 along with a subsequence of ρ_N. The thin solid line should hopefully intersect the left axis at the thick horizontal line, but it does not. However, the relative error of the extrapolation to $1/N \Rightarrow 0$ is only 1 part in 400.

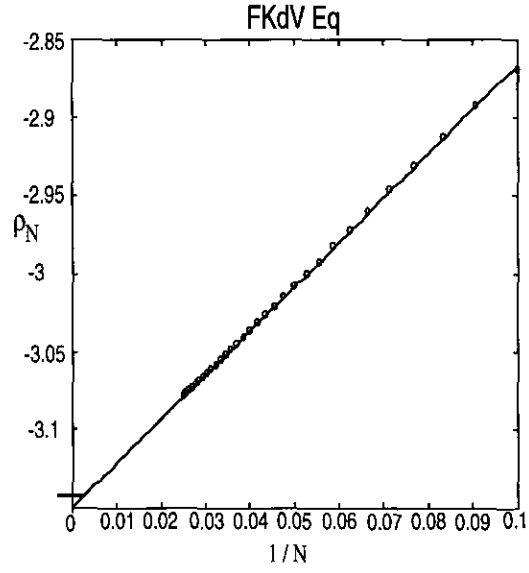

FIG. 10. (FKdV equation). Circles: ρ_N for $N = 11, 12, ..., 40$, plotted versus $1/N$. Thin solid line: The linear polynomial which is the least squares fit to $\{\rho_{21}, ..., \rho_{40}\}$: $p_1(N) = -3.1497 + 2.83/N$. Thick solid horizontal line: The exact $\nu_1 = -\pi$.

3. *Third-Order Nonlinear Schroedinger (TNLS) Equation*

The equation is

$$-\tfrac{1}{2}\varepsilon^2 Q + i\varepsilon^2 Q_x + \tfrac{1}{2}Q_{xx} + |Q|^2 Q - iQ_{xxx} = 0. \quad (7.20)$$

This complex-valued equation has complex-valued envelope solitary waves. The first few terms of the perturbation theory are

$$\begin{aligned} Q(x) \sim\ & \varepsilon\,\mathrm{sech}(\varepsilon x) - 3i\varepsilon^2\,\mathrm{sech}(\varepsilon x)\tanh(\varepsilon x) \\ & + \varepsilon^3\,\mathrm{sech}(\varepsilon x)[21\,\mathrm{sech}^2(\varepsilon x) - \tfrac{39}{2}] \\ & + \varepsilon^4 i\,\mathrm{sech}(\varepsilon x)\tanh(\varepsilon x)[\tfrac{107}{2} - 151\,\mathrm{sech}^2(\varepsilon x)] \\ & + \varepsilon^5\mathrm{sech}(\varepsilon x)[-\tfrac{1013}{8} - \tfrac{2489}{2}\,\mathrm{sech}^2(\varepsilon x) + 1529\,\mathrm{sech}^4(\varepsilon x)]. \end{aligned} \quad (7.21)$$

Strictly speaking, the one-parameter family of solutions approximated by (7.21) is only a subset of a broader parameter family, but [1] shows that the general family can be generated from the solutions we will discuss by trivial rescalings.

Wai, Menyuk, Lee, and Chen [19] have derived general recurrence relations to compute the TNLS multiple scales series to all orders; we follow their spirit with minor changes of notation and implementation.

Our goal is to compute approximations to the radiation coefficient α. The asymptotic form of $Q(x)$ for the nonlocal envelope solitary wave is

$$\begin{aligned} Q(x) \sim \alpha(\varepsilon) i\,\mathrm{sign}(x)\exp(-ix/2 + 8i\,\varepsilon\,\mathrm{sign}(x) \\ + O(\varepsilon^2)), \quad |x| \gg 1/\varepsilon. \end{aligned} \quad (7.22)$$

The zeroth-order hyperasymptotic approximation will yield a sequence of approximations to ν_0, where

$$\alpha(\varepsilon) \sim (\nu_0 + \nu_1\varepsilon + \ldots)\exp\left(-\frac{\pi}{4\varepsilon}\right), \quad \varepsilon \Rightarrow 0. \quad (7.23)$$

To compute the residuals, we expand

$$\begin{aligned} Q(\varepsilon x) = \sum_{j=1}^{\infty} \varepsilon^j \Bigg\{ & \sum_{k=1}^{j} u_{jk}\mathrm{sech}^k(\varepsilon x) \\ & + i\tanh(\varepsilon x)\sum_{k=1}^{j-1} v_{jk}\mathrm{sech}^k(\varepsilon x)\Bigg\}. \end{aligned} \quad (7.24)$$

(This form simplifies notation at the expense of wasted storage: all $u_{2j,k}$ and $v_{2j-1,k}$ are zero.)

The residual is expanded as

$$\begin{aligned} r_{j+2}(x) = & \sum_{k=1}^{j+2} r_{jk}\mathrm{sech}^k(\varepsilon x) \\ & + i\tanh(\varepsilon x)\sum_{k=1}^{j+1} \rho_{jk}\mathrm{sech}^k(\varepsilon x), \end{aligned} \quad (7.25)$$

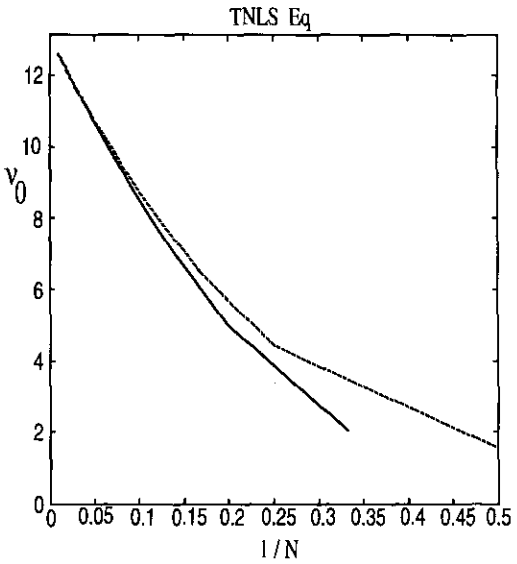

FIG. 11. Elements of $\alpha_N^{(h0)}$, divided by $\exp(-\pi/[4\varepsilon(N)])$, for the TNLS nonlocal solitary wave. Solid: subsequence for odd N, beginning with $N = 3$, 5, ..., 101; the residuals which generate this subsequence are powers of $\mathrm{sech}(\varepsilon x)$. Dashed: even N, from $N = 2$ to $N = 100$; the residuals are polynomials in $\mathrm{sech}(\varepsilon x)$ multiplied to $\tanh(\varepsilon x)$. The top axis is at 13.1273, our best estimate for $\omega(0)$.

where $r_{j+2}(x)$ is $O(\varepsilon^{j+2})$, but serves as the forcing in the perturbation equation that determines the $O(\varepsilon^j)$ term in $Q(x)$. Thus, the $\varepsilon^2\,\mathrm{sech}(\varepsilon x)\tanh(\varepsilon x)$ term in $Q(x)$, whose coefficient is v_{21}, is found from ρ_{21} in $r_4(x)$. The corresponding approximation to the radiation coefficient follows by applying the last line of Table II. In so doing, we replace the multiple scales perturbation equation by a differential equation which *includes* the *third derivative* but is *constant coefficient* so that α may be easily found by evaluating an integral as explained in Section 5.

The zeroth-order approximation is

$$\begin{aligned} \alpha_N^{(h0)} = & \exp\left(-\frac{\pi}{4\varepsilon}\right)(-1)^{N+1}\frac{\pi}{2^{2N}(2N+1)!} \\ & \begin{cases} r_{2N,2N+2}, N\,\text{odd}, \\ \rho_{2N,2N+1}, N\,\text{even}, \end{cases} N = 2, 3, 4, \ldots. \end{aligned} \quad (7.26)$$

When rescaled by dividing out the exponential, (7.26) gives a sequence of numbers that converge to ν_0 as $N \Rightarrow \infty$.

One minor complication is that the TNLS residual is a polynomial in $\mathrm{sech}(\varepsilon x)$ only at odd order; the residual (and the perturbation terms it forces) at even order are of similar form, except for multiplication by $\tanh(\varepsilon x)$. As shown in Fig. 11, the even and odd N subsequences are best plotted separately. However, both converge to a common limit as $N \Rightarrow \infty$.

The convergence of the polynomial extrapolants is illustrated in Fig. 12. The horizontal axis is N, an odd integer, where the polynomial is fitted to $\{\alpha_N^{(h0)}, \alpha_{N+2}^{(h0)}, \ldots, \alpha_{101}^{(h0)}\}$. As the fit is restricted

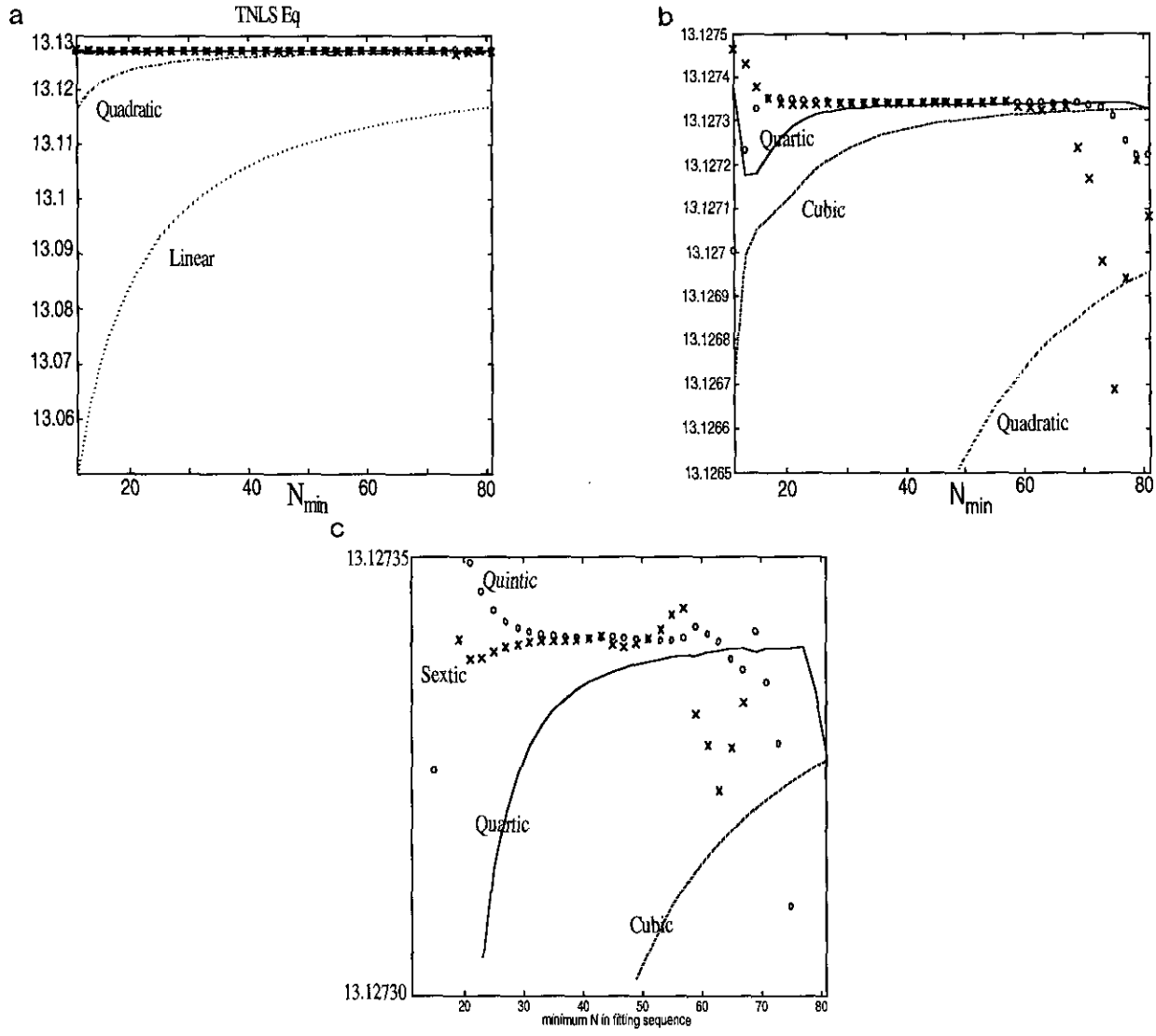

FIG. 12. Hyperasymptotic approximations to ν_0, the TNLS proportionality constant in (7.23), as obtained by least-squares fitting a polynomial to the sequence of (odd-numbered) approximations $\{\alpha_N^{(h0)}, \alpha_{N+2}^{(h0)}, \ldots \alpha_{101}^{(h0)}\}$, where the latter have been scaled by dividing out the exponential $\exp(-\pi/(4\varepsilon))$ and where N, the lower limit of the sequence of fitted approximations, is the horizontal axis of the graph. Dotted: linear polynomial. Dot-dash: quadratic polynomial. Dashed: cubic. Solid: quartic polynomial. Circles: quintic (fifth-degree) polynomial. x's: sextic polynomial. All three graphs (a, b, c) are identical, except for different vertical scales.

to a smaller and smaller subsequence, the linear, quadratic, cubic, and quartic polynomials converge to $\nu_0 \approx 13.12734$. The fifth and sixth degree polynomials (symbols) agree closely with the limit for all lower limits N until N is close to 60. This is not surprising; the high degree polynomials for this N use only one or two more points than the corresponding interpolants, and the failure of high degree, equispaced interpolation is described in most numerical analysis texts as the "Runge phenomenon." Until this instability sets in, however, Fig. 12c shows that all polynomials of degrees three to six seem to be converging on a common limit:

$$\alpha(\varepsilon) \sim 13.1273 \exp\left(-\frac{\pi}{4\varepsilon}\right), \quad \varepsilon \Rightarrow 0. \tag{7.27}$$

This is close to the numerical calculations of Cai *et al.* [19], who find that $\nu_0 \approx 13.24$, a difference of a bit under 1%. The reason for this small difference is unknown.

We shall postpone a computation of the first-order correction ν_1 to a future article devoted to a variety of TNLS results.

8. *Summary and Comparisons with Related Studies*

In this article, we have described a new procedure for computing the prefactors ν_0 and ν_1 in the asymptotic approximation to the radiation coefficient α of a weakly nonlocal solitary wave, which is typically of the form

$$\alpha(\varepsilon) \sim (\nu_0 + \nu_1\varepsilon + \ldots)\exp\left(-\frac{q}{\varepsilon}\right), \quad \varepsilon \ll 1.$$

The constant q is usually easy to determine [1, 4, 13, 15]; q/ε is simply the distance from the real x-axis to the nearest poles or branch points of the lowest order multiple scales approximation. For the forced-KdV equation, $u_1(x) = 12\,\mathrm{sech}^2(\varepsilon x)$, which has double poles at $x = \pm i\pi/(2\varepsilon)$ (plus additional irrelevant singularities farther from the real axis). Therefore, $q = \pi/2$ for this equation.

The task of determining ν_0, ν_1, ... in (8.1) is very much harder. The earliest method, developed by Segur and Kruskal [13], requires matching asymptotic expansions in the complex plane. Even then, the final step is numerical.

Pomeau, Ramani, and Grammaticos [14] simplified the matched asymptotics method by use of Borel summation, which replaced the numerical solution of a differential equation by the extrapolation of a nonlinear recurrence. However, this extrapolation, too, must be performed numerically.

Akylas and Yang [16, 23] have proposed an ingenious alternative: converting the problem to a nonlinear integral equation by taking the Fourier transform and then applying the method of multiple scales to the transform $U(k)$ rather than to $u(x)$. This method does not explicitly require adventures in the complex plane and yields a simple nonlinear recurrence which can be extrapolated to give ν_0—but only numerically.

Our new procedure shares the common defect of requiring a final numerical extrapolation. However, it does not require an integral equation, Borel summation, or matched asymptotics in the complex plane.

Instead, our zeroth-order method for ν_0 has only three steps:

(i) Compute the multiple scales series to high order.

(ii) Extract the term from Nth order residual which is proportional to the highest power of $\mathrm{sech}(\varepsilon x)$. Divide this by an order-dependent constant which is a function only of the linearized but constant coefficient form of the differential equation.

(iii) Extrapolate the sequence of approximations to $N \Rightarrow \infty$ by fitting a polynomial of low degree in $1/N$.

The first step is easy because for the three examples discussed here, and for many other problems catalogued in [1], the multiple scales series is simply a power series in ε with coefficients which are polynomials in $\mathrm{sech}(\varepsilon x)$, perhaps multiplied by a single factor of $\tanh(\varepsilon x)$. By using hyperbolic identities, the series can be calculated to very high order via recurrence relations as illustrated for the forced-KdV equation in Table I.

The second step is easy; Section 5 gives the general method for computing the radiation coefficient for the solution of a forced, linear differential equation. The answer for our three examples is given in Table II.

The final step requires only a library routine for polynomial least-squares approximation. For our three examples, we obtain (we think) the first six digits for ν_0.

To compute the first-order correction ν_1 requires two additional steps. First, we need to explicitly estimate the optimal truncation $N_{\mathrm{opt}}(\varepsilon)$. (We used the formula for $N_{\mathrm{opt}}(\varepsilon)$ to motivate some of the approximations for the zeroth-order hyperasymptotic method, but it drops out of the final sequence of approximations to ν_0.) This question is easily answered in inverse form: given N, for what value of ε does the Fourier transform of the leading term in the residual have its peak at the far field wavenumber k_f, that is, at the wavenumber of the "wing" oscillations of the nonlocal solitary wave? This is simpler than one might suppose because the residual's "leading term" is proportional to the jth derivative of $\mathrm{sech}(\varepsilon x)$, where j is chosen so that the highest power in this derivative is of the same degree as the highest power of $\mathrm{sech}(\varepsilon x)$ in the residual; the Fourier transform of the leading term is just $(ik)^j$ times that of $\mathrm{sech}(\varepsilon x)$ itself.

Second, we need to use the full Nth-order perturbative residual to compute α from the formula of Table II. As illustrated for the forced-KdV equation, however, this is relatively easy.

Our "hyperasymptotic" perturbation theory is appealing because of its simplicity. Nothing is manipulated except powers of ε and $\mathrm{sech}(\varepsilon x)$.

There are many similarities between our method and themes previously developed in the literature. The recurrence derived from the Borel summation method, for example, is a recursively computed high order perturbative approximation, just like the multiple scales series that we use. The difference is that our expansion is on the real axis in powers of $\mathrm{sech}(\varepsilon x)$, whereas the Borel-derived series is in inverse powers of y, where y is the distance from a pole of the lowest order multiple scales approximation, situated off the real axis.

Although Akylas and Yang's sequence of approximations is derived through a nonlinear integral equation and a multiple scales approximation in wavenumber, rather than in physical space, nevertheless the elements of their sequence are *identical* with our one-term approximation to ν_0 for the forced-KdV and FKdV equations. Obviously, there are some close but subtle connections between their method and ours, despite the very dissimilar appearances.

Akylas and Yang [16] also solved the forced-KdV equation with a Gaussian forcing. Although this is a rather special and unusual example because the Gaussian is an entire function, it nevertheless should be noted that their integral equation/Fourier transform method is successful for this case (albeit with some messy generalization of their algorithm). It is unclear whether our method can be similarly extended.

Another restriction is that some problems, such as the intermediate long wave (ILW) equation discussed by Yang and Akylas [23], have a multiple scales series which is not a simple power series, but includes terms in $\varepsilon \log(\varepsilon)$ and so on. We have not yet attempted to extend our hyperasymptotic method to power-and-log series.

The notion of extracting α by adding the highest derivative back to the left-hand side of the multiple scales perturbation equation has been around for at least five years as reviewed by Hakim [9]. However, the *first-order* residual gives an $O(1)$ error in ν_0, although the $\exp(-q/\varepsilon)$ factor is computed correctly.

We have shown above that the only way to obtain an accurate answer from this strategy is to use a *higher order* multiple scales residual as the inhomogeneous term in the constant coefficient differential equation which is solved for α. Even then, the Nth-order residual generates an approximation which has the large relative error $O(1/N)$. What makes the hyperasymptotic method really useful is that this sequence of mediocre approximations for different N can be extrapolated to $N \Rightarrow \infty$ to give ν_0 to six decimal places.

The computation of ν_1 has much less precedent in the literature. Grimshaw and Joshi [15] have derived this correction for the fifth-order KdV equation by matching asymptotic expansions in the complex plane. We have been able to confirm Grimshaw and Joshi's analytical result, $\nu_1 = -\pi\ \nu_0$, to within about 0.1%.

Our one completely novel result is that $\nu_1 \approx 0$ for the forced-KdV equation. Our calculations of ν_1 for the third-order nonlinear Schroedinger equation will be published shortly in another place.

APPENDIX: ENVELOPE PERTURBATION THEORY

When the initialization for the Newton–Kantorovich equation (3.1) is a superasymptotic approximation of high order $N_{opt}(\varepsilon)$, the Fourier transform of the inhomogeneous term has the shape of a two-humped camel's back, where the peaks are at $k = \pm 1$ (Fig. 4). To analytically solve the Newton–Kantorovich equation, we apply an envelope perturbation theory which is a cousin of that used to derive the Cubic Schroedinger and TNLS equations [24, 25].

The goal is to solve, in the limit $\varepsilon \ll 1$,

$$\nu_{xx} + \nu - \varepsilon^2 q(\varepsilon x)\, \nu = r(x). \tag{A.1}$$

on either an infinite or spatially periodic interval subject to the restriction that $r(x)$ is a function whose Fourier transform, $R(k)$, is peaked at $k = \pm 1$. (In our application, $\nu(x) = \Delta^1(x)$, the first quasi-Newton correction to the optimally truncated multiple scales series for $u(x)$, $\varepsilon^2 q(\varepsilon x) = 2 u^1(x)$ for the forced-KdV equation and $r(x) = r_N(x)$, the residual at the optimal order. However, the method is general as long as $r(x)$ has a two-humped Fourier transform.)

It is very convenient to solve for each hump separately, so we split the transform $R(k)$ into two pieces via

$$R(k) \equiv H(k)\, R(k) + H(-k)\, R(k), \tag{A.2}$$

where $H(k)$ is a smoothed approximation to the step function. (A.2) is free of approximations if

$$H(k) + H(-k) = 1, \quad \text{all } k. \tag{A.3}$$

Therefore, $H(k)$ should satisfy the following three conditions:

$$\text{(i)} \quad \lim_{k \to \infty} H(k) = 1 \tag{A.4a}$$

$$\text{(ii)} \quad \lim_{k \to -\infty} H(k) = 0 \tag{A.4b}$$

$$\text{(iii)} \quad \{H(k) - \tfrac{1}{2}\} = -\{H(-k) - \tfrac{1}{2}\}. \tag{A.5}$$

The first two conditions ensure that $H(k)\ R(k)$ is a function with only a single large peak at $k = 1$; the other is cut off by the decay of $H(k)$ for negative k. The third condition, antisymmetry about the value of $\frac{1}{2}$ with respect to $k = 0$, is just a rearrangement of (A.3).

There are many possible choices for the quasi-step function including

$$H_1(k) = \frac{1}{2}\operatorname{erfc}(-b\, k), \tag{A.6}$$

where b is an arbitrary positive constant $\gg 1$.

We can then write, without approximation,

$$\nu(x) \equiv \nu_+(x) + \nu_-(x), \tag{A.7}$$

where the two parts solve

$$\nu_{\pm,xx} + \nu_\pm - q(\varepsilon x)\, \nu_\pm = r_\pm(x) \tag{A.8}$$

$$r_\pm(x) = \rho_\pm(\varepsilon x) \exp(\pm ix), \tag{A.9}$$

where $\rho_\pm(\varepsilon x)$, the "envelope of the forcing," is the inverse Fourier transform of the forcing with a unit shift in wavenumber; that is,

$$\rho_\pm(\varepsilon x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(1 \pm k)\, R(k \pm 1) \exp(-ikx)\, dx \tag{A.10}$$

To solve (A.8), define the envelope $\lambda(x)$ through

$$\nu_\pm(x) = \lambda_\pm(\varepsilon x) \exp(\pm i x). \tag{A.11}$$

Introducing the slow variable, $X = \varepsilon x$ and dividing out the common exponential factor gives

$$+2i\, \varepsilon\, \lambda_{\pm,X} + \varepsilon^2\, \lambda_{\pm,XX} - \varepsilon^2\, q(X)\lambda_\pm = \rho_\pm(X). \tag{A.12}$$

So far, we have made no approximations, even though we have inserted ε as a scaling factor in several places. If we assume that our scalings are correct, that is, if $\rho_\pm$ is truly a function only of the slow variable $X = \varepsilon x$, then (A.12) simplifies to

$$\pm 2i\, \varepsilon\, \lambda_{\pm,X} = \rho_\pm(X) \tag{A.13}$$

which implies that to lowest order

$$\lambda_{\pm}(X) \sim \mp \frac{i}{2\varepsilon} \int_0^X \rho_{\pm}(Y)\, dY, \qquad \text{(A.14)}$$

independent of the x-dependent coefficient, $q(\varepsilon\, x)$.

This justifies the quasi-Newton approximation (Section 4) which neglects $q(\varepsilon\, x)$.

There are two limitations on (A.14). The first is that we have not proved that the envelope of the residual varies only on the slow length scale, we have merely assumed it. Second, the argument above does not explain why the relative error in α is $O(\varepsilon^2)$ instead of $O(\varepsilon)$. However, both these conclusions are supported by the experimental evidence of Fig. 5.

ACKNOWLEDGMENTS

This work was supported by the NSF through Grants OCE8812300, DMS8716766, ECS9012263, and OCE9119459 and by the Department of Energy through Grant KC070101. I thank Harvey Segur, Martin Kruskal, Herbert Levine, and Saleh Tanveer for organizing the 1991 NATO Workshop, which was a very good stimulus to my own work. I also thank two anonymous referees for helpful comments.

REFERENCES

1. J. P. Boyd, *Weakly Nonlocal Solitary Waves and Other Exponentially Small Phenomena,* to appear.
2. J. P. Boyd, in *Advances in Applied Mechanics,* Vol. 27, edited by T.-Y. Wu and J. W. Hutchinson (Academic Press, New York, 1989), p. 1.
3. J. P. Boyd, in *Nonlinear Topics in Ocean Physics,* edited by A. R. Osborne and L. Bergamasco (North-Holland, Amsterdam, 1991), p. 527.
4. H. Segur, S. Tanveer, and H. Levine, Eds., *Asymptotics Beyond All Orders,* (Plenum, New York, 1991).
5. J. P. Boyd, *Physica D* **48,** 129–146 (1991).
6. M. V. Berry, *Proc. R. Soc. London A* **422,** 7 (1989).
7. M. V. Berry, in *Asymptotics Beyond All Orders,* edited by H. Segur, S. Tanveer, and H. Levine (Plenum, New York, 1991), p. 15.
8. M. V. Berry and C. J. Howls, *Proc. R. Soc. London A* **434,** 657 (1991).
9. V. Hakim, in *Asymptotics Beyond All Orders,* edited by H. Segur, S. Tanveer, and H. Levine (Plenum, New York, 1991), p. 15.
10. R. B. Paris, *Proc. R. Soc. London A* **436,** 165 (1992).
11. R. B. Paris and A. D. Wood, *J. Comput. Appl. Math.,* **41,** 135 (1992).
12. A. B. O. Daalhuis, 1992, *IMA J. Appl. Math.,* **49,** 203 (1992).
13. H. Segur and M. D. Kruskal, *Phys. Rev. Lett.,* **58,** 747 (1987).
14. Y. Pomeau, A. Ramani, and G. Grammaticos, *Physica D* **21,** 127 (1988).
15. R. H. J. Grimshaw and N. Joshi, *SIAM J. Appl. Math.* **55** (1995), 124–135.
16. T. R. Akylas and T.-S. Yang, *Stud. Appl. Math.* **94** (1995), 1–20.
17. C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw–Hill, New York, 1978).
18. P. K. A. Wai, C. R. Menyuk, Y. C. Lee, H. H. Chen, *Opt. Lett.* **11,** 464 (1986).
19. P. K. A. Wai, H. H. Chen, and Y. C. Lee, *Phys. Rev. A* **41,** 426 (1990).
20. A. H. Nayfeh, *Perturbation Methods* (Wiley, New York, 1973).
21. J. P. Boyd, *Chebyshev and Fourier Spectral Methods,* (Springer-Verlag, Heidelberg, 1989).
22. J. P. Boyd, *Appl. Numer. Math.* **7** (1991), 453–479.
23. T.-S. Yang and T. R. Akylas, *Physica D,* submitted.
24. J. P. Boyd, *J. Phys. Oceangr.* **13,** 428 (1983).
25. R. S. Johnson, *Proc. R. Soc. London A* **357,** 131 (1977).
26. F. W. J. Olver, *Asymptotics and Special Functions,* (Academic Press, New York, 1974).
27. F. W. J. Olver, *SIAM J. Math. Anal.* **24** 756 (1993).