The steepest descent method for integrable systems

Extended abstract

Mathematical description of the n-phase waves

Explicit expressions for *n-phase wave* solutions of many integrable systems are given with the aid of the n-fold series

$$\sum_{\mathbf{m}\in\mathbb{Z}^n} \exp\{2\pi i\mathbf{z}\cdot\mathbf{m} + \pi i(B\mathbf{m},\mathbf{m})\},$$

which is summed over the multi-integer $\mathbf{m}$. B is an $n\times n$ matrix with positive definite imaginary part that gives the series exponential quadratic convergence. The series is known as a *Riemann theta function*. It has n complex arguments $\mathbf{z} = (z_1, \cdots z_n)$. In its implementation in the representation of nonlinear waves, each argument is linear in the space and time variables x and t and represents a phase of the wave. The expression of the actual waveform is not the theta function itself; this is given by formulae, easily computable in terms of the theta function and its derivatives or shifts. For many integrable systems, the matrix B arises from periods of the Riemann surface defined by the square root $R(z)$ of a polynomial. In other words, the elements of the matrix B are linear combinations of the hyperelliptic integrals $\oint z^k dz/R(z),\ k = 0, 1\cdots n-1$ along appropriate closed contours on the Riemann surface of R. For the *Korteweg de-Vries Equation*,

$$\text{KdV:}\quad \partial_t u + 6u\partial_x u + \varepsilon^2\partial_x^3 u = 0, \tag{1}$$

the radical is

$$R(z) = \left((z-\beta_0)\prod_{i=0}^{2n+1}(z-\alpha_i)(z-\beta_i)\right)^{\frac{1}{2}}, \quad \text{where all } \alpha_i \text{ and } \beta_i \text{ are real.} \tag{2}$$

For the *focusing cubic nonlinear Schrödinger Equation*,

$$\text{NLS:}\quad i\varepsilon\partial_t u + \varepsilon^2\partial_x^2 u + |u|^2 u = 0, \tag{3}$$

it is

$$R(z) = \left(\prod_{i=0}^{2n+1}(z-\alpha_i)\right)^{\frac{1}{2}}, \quad \text{where all } \alpha_i \text{ are complex and } \alpha_{2j+1} = \overline{\alpha}_{2j}. \tag{4}$$

The n-phase wave is, thus, parametrized by the $2n+1$ real parameters $\beta_0,\alpha_1,\beta_1, \alpha_2,\beta_2,\cdots,\alpha_n,\beta_n$ for KdV and by the $n+1$ complex parameters $\alpha_0, \alpha_2, ..., \alpha_{2n}$ for NLS. They serve as the mathematically most convenient wave parameters. The wavenumbers and frequencies are functions of these parameters, as they are expressed in terms of hyperelliptic integrals involving the radical R. The wave baseline and mean-square amplitude are also computable from the parameters.

The importance of these parameters lies at the root of integrability in the sense of the Lax pair. A linear operator (the first one in the Lax pair of operators) is associated with the nonlinear wave equation and varies isospectrally during the time evolution of the system. The conservation of the spectrum provides sufficiently many conservation laws to the nonlinear equation to make it solvable. The parameters at hand arise as spectra of this operator. Associated to KdV, is a Schrödinger operator for a scalar field $\psi(x,t)$. Its eigenvalue problem is (think of t as fixed)

$$-\varepsilon^2\partial_x^2\psi(x) + u(x,t)\psi = z^2\psi, \tag{5}$$

where $u(x,t)$ is the KdV waveform. The associated linear operator of NLS (Zakharov-Shabat) is a Dirac operator. Its eigenvalue problem is a first order 2×2 ODE system for the vector valued field $\psi(x,t) = (\psi_1, \psi_2)$. The NLS waveform $u(x,t)$ is a (nonconstant) coefficient in this ODE..

The second operators in the Lax pairs for KdV and NLS are differential operators in time that govern the time evolution of the fields $\psi(x,t)$. The nonlinear wave equations tehmselves (KdV and NLS in our context) are identical to the compatibility condition

$$\psi_{xt} = \psi_{tx} \tag{6}$$

The separation of scales: wave modulation and the modulational instability

Wave modulation is a perturbation of an n-phase wave, in which

1. The parameters in the wave formula are no longer constant but vary in space-time, *i.e.* (x,t), in a scale that is asymptotically larger than the scale of the wavelengths and periods of the wave.

2. The wave formula with the varying parameters, rather than being an exact solution of the wave equation, gives the leading asymptotic behavior of an exact solution.

The existence of such solutions is referred to as the *modulational ansatz*. Finding such solutions is posed naturaly as an initial value problem. One prescribes as initial data the expression of a wave, in which the constant parameters are replaced by quantities varying slowly in space. One then seeks the time evolution of these parameters so that the wave expression provides a leading behavior of an exact solution of the nonlinear wave equation. In the case of a 1-phase modulated KdV wave, the relevant parameters are the real numbers β_0, α_1 and β_1. In this case, the evolution equations (or *modulation equations*) consist of a first order nonlinear hyperbolic PDE system in Riemann invariant form first obtained by Whitham. He obtained the result in a perturbation calculation by averaging out the oscillatory part from three conservation laws of KdV and posing the averaged quantities as conservation laws of the modulation system. The (very nontrivial) averaging for higher phase waves was obtained by Forest, Flaschka and McLaughlin. All modulation systems in PDE form are often referred to as Whitham systems.

The fact that hyperbolic systems are well posed as initial value problems suggests an abundance of KdV waves that satisfy the modulational ansatz. It also suggests that they are, in some sense, modulationally stable. This is not true for the focusing NLS, for which the modulation system turns out to be elliptic (it has complex characteristics), thus, NLS waves are *modulationally unstable.* Nevertheless, as we prove below, modulated solutions of NLS appear naturally in a restricted space of perturbations.

The wavelengths and periods, achieved in KdV and NLS as scaled above, are of order ε. Thus, a separation of scales between the oscillatory part of the wave and the modulating parameters is achieved in the asymptotic limit $\varepsilon \to 0$. The limit is often characterized as *small dispersion* or as *semiclassical*, terms that reflect the dispersive character of the third derivative in the linearized KdV equation $u_t + \varepsilon^2 u_{xxx}$ and the role of ε as a Planck constant in the corresponding Schrö dinger or Dirac assocoated operators acting on the field ψ. The phenomenology in this limit provides a model for the nonlinear analogue to geometrical optics and diffraction theory.

Wave breaking, dispersive shocks and nonlinear caustics

As we saw, the KdV modulation equations are a nonlinear hyperbolic system. It is well known that such systems may break in finite time. When this happens, the modulational ansatz and the above perturbative theory collapses. The shock theory of fluid-dynamic-type systems is not at all applicable in advancing beyond the break time. What actually happens at a break-point in space-time in KdV and NLS is that a new oscillatory phase is generated and spreads in space behind fronts that travel with finite speed. It is also possible that, as time increases, a spatial interval of a certain wave phase n collapses to a point and the particular wave phase disappears. In the large scale, the fronts (or tails) are seen as curves in space-time. These *breaking curves* divide space-time into regions. Inside each region, the solution is described by a modulated n-phase wave, in which $n = 0, 1, 2, 3, ...$ wave-phases interact nonlinearly. The value of n jumps across the breaking curves, reflecting the generation of a new phase across the curve. We sometimes call these curves *(nonlinear) caustics* in analogy to linear theory.

A picture of the modulation and breaks of the wave is provided by the multivalued function that gives the parameter values at each space-time point (x, t). In the simplest case of an exact (not modulated) wave, the parameters are independent of x and t.

The initial value problem for KdV and NLS as $\varepsilon \to 0$

When initial conditions are posed for the KdV and NLS equations, numerics typically display emergent wave structures. The analytical object is to understand the mathematical mechanism for the spontaneous generation of the waves and calculate the wave parameters and the waveform in the asymptotic limit $\varepsilon \to 0$.

The initial value problems considered are

$$\text{KdV:} \quad \partial_t u + 6u\partial_x u + \varepsilon^2 \partial_x^3 u = 0, \qquad u(x,0) = v(x), \quad v(x) \text{ is real}, \tag{7}$$

and

$$\text{NLS:} \quad i\varepsilon\partial_t u + \varepsilon^2 \partial_x^2 u + |u|^2 u = 0, \qquad u(x,0) = A(x)e^{iS(x)/\varepsilon}, \quad S(x) \text{ is real}. \tag{8}$$

The method of scattering and inverse scattering are employed in solving these problems. Scattering is applied to the linear eigenvalue problem (ODE) associated with the nonlinear wave equation. The independent variable is x. The field quantity is $\psi(x, t; z)$. The scatterer is the nonconstant coefficient of the ODE $u(x, t)$, that is assumed to decay as $x \to \pm\infty$. The first challenge is the calculation of the scattering data at $t = 0$. They constitute the values of the variable z that are in the spectrum plus one piece of information needed at each value: the reflection coefficient on continuous spectrum and a norming constant at each proper eigenvalue. The evolution of the scattering data with time is calculated trivially due to integrability. The spectrum is conserved and the additional piece of information is modified by a simple exponential factor. The more significant challenge is the *inverse scattering problem* (IST), or the determing of the scatterer $u(x, t)$ from the scattering data at time t.

The issue in the asymptotic calculation is making the solution as explicit as possible in the limit $\varepsilon \to 0$. This includes proving the generation of wave-phases with $n > 0$, and calculating the wave parameters in the various regions. Observe that the initial data for KdV are nonoscillatory. The same is essentially true for NLS. In spite of the fast phase, the NLS initial data, like the ones for KdV, have the form of modulated 0-phase waves. The necessity for a break in the evolution of such a wave is clearly visible in KdV. Setting $\varepsilon = 0$ in the equation, one obtains

$$u_t + 6uu_x = 0,$$

which typically breaks for finite time. The issue in the asymptotic ($\varepsilon \to 0$) calculation is rendering the solution as explicit as possible. This includes proving the generation of n-phase waves with n ¿ 0, and calculating the wave parameters in the various spatio-temporal regions. The key method for this is the asymptotic solution of the Riemann-Hilbert problem. It is achieved via the steepest descent method and its key ingredient, the g-function mechanism. Outlining the emergence of inverse scattering, of its associated Riemann-Hilbert problems and their asymptotic solutions as well as their relation to modern random matrix theory are the main goals of this presentation.

Overview

1. Introduction to nonlinear waves of integrable systems and numerical results.

2. The zero dispersion limit of the KdV equation

(a) The calculation of the weak limit and equilibrium measures (Lax-Levermore theory)

(b) The calculation of the strong asymptotic limit

3. Inverse scattering for the NLS equation as a Riemann-Hilbert problem

4. The method of steepest descent for the asymptotic solution of Riemann-Hilbert problems.

5. The g-function mechanism for the systematic determination of the modulation parameters.

6. Application to NLS