

Strings, Waves, Drums: Spectra and Inverse
Problems

Richard Beals
Yale University

Peter C. Greiner
University of Toronto
City University of Hong Kong

Preface

This volume grew from lectures that the authors gave at City University of Hong Kong. Its aim is to cover a variety of interconnected topics that we have found particularly interesting, with some historical background, some motivation, and enough detail (however sketchy) to convey the main ideas.

We are most grateful to Professors Wong and Ciarlet and to the staff of the Mathematics Department and of the Liu Bie Ju Center for their hospitality and aid.

Contents

1	The vibrating string	1
2	Can one hear the density of a string?	5
3	Can one hear a nuclear potential?	9
4	Can one reconstruct the potential?	15
5	Quantum mechanics in one dimension	19
6	Preserving the spectrum: KdV	25
7	Can one explain the shape of a wave?	29
8	Should one hear white noise – or KdV?	35
9	The soliton revolution	39
10	Strings revisited	43
11	Waves with peaks and troughs	47
12	Can one hear the shape of a drum?	51
13	Inverting an elliptic operator	57
14	Indices and index theorems	63
15	The index and the heat operator	69

Chapter 1

The vibrating string

We begin our discussion of the questions in the title with an example whose generalizations and ramifications will occupy us at length.

Relations among the length, tension, and pitch of a vibrating string of constant density have been known since antiquity. The mathematical formulation of a model of a vibrating string as a partial differential equation is one of the first successes of continuum mechanics. It was accomplished in the 18th century by work of Brooke Taylor, Daniel Bernoulli, Jean d'Alembert, and Leonhard Euler. The problem led to controversy about the possibility of representing an arbitrary function as a trigonometric series, and, indeed about the meaning of the very term “function” controversies not resolved until the middle of the 19th century.

The standard mathematical model of a vibrating string of length L is a function $u(x, t)$, defined for x in the interval $[0, L]$ and for time $t \geq 0$. This function represents the vertical displacement of the string at time t from its resting position $u \equiv 0$. The fixed endpoint condition is

$$u(0, t) = u(L, t) = 0, \quad t \geq 0. \quad (1.1)$$

The equation of motion is

$$m u_{tt} = k u_{xx}; \quad (1.2)$$

(D'Alembert, 1747). This is Newton's equation as interpreted in this case: m is the density, assumed constant, so the left side is mass times acceleration (at a point of the string). The right side represents a force proportional to curvature of the string at that point; the constant k is the tension. To complete the formulation one can impose initial conditions,

$$u(x, 0) = f(x), \quad u_t(x, 0) = g(x)$$

and then fit the general solution

$$u(x, t) = F(x + ct) + G(x - ct), \quad c^2 = \frac{k}{m} \quad (1.3)$$

found by D'Alembert (1747) and Euler (1748).

Instead we start with a more physical point of view: what pure pitches can be associated with this string? Specifically, are there solutions that have the form

$$u(x, t) = a(x) \cos(2\pi\omega t) ? \quad (1.4)$$

Such a solution is periodic with period $1/\omega$ and frequency (pitch) ω . Equation (1.2) then requires that the amplitude a satisfy the second order ODE

$$\frac{d^2 a}{dx^2} = -\frac{4\pi^2\omega^2 m}{k} a. \quad (1.5)$$

Combined with the endpoint condition (1.1), this implies that a nonzero solution with no initial velocity is a multiple of one of the functions

$$\varphi_n(x) = \sin\left(\frac{n\pi x}{L}\right), \quad n = 1, 2, 3, \dots \quad (1.6)$$

Comparing (1.5) and (1.6), we see that the possible frequencies are

$$\omega_n = \frac{n}{2L} \sqrt{\frac{k}{m}}. \quad n = 1, 2, 3, \dots \quad (1.7)$$

Thus the fundamental (lowest) frequency $\sqrt{k/m}/2L$ tells us one of the quantities density, tension, or length if we know the other two. In particular we can *hear* one if we know the other two (and we have perfect pitch). This is the first of several questions that we will discuss about what can be “heard” in various contexts. A less picturesque version of the question is: how much can a spectrum tell us?

Modern mathematicians and physicists would approach the string problem by concentrating first on the *operator* $L = k(d/dx)^2$ that appears on the right in (1.2). Acting on functions that satisfy (1.1) and that are smooth enough, it is symmetric with respect to the L^2 inner product

$$\begin{aligned} (Lu, v) &= \int_0^L k \frac{d^2 u}{dx^2} v \, dx = - \int_0^L k \frac{du}{dx} \frac{dv}{dx} \, dx \\ &= (u, Lv). \end{aligned}$$

In analogy with the finite-dimensional case, one can hope to diagonalize the operator L : find a basis for this space of functions that consists of eigenfunctions of L : $L\varphi_n = \lambda_n\varphi_n$. Up to multiplicative constants, these are exactly the functions (1.6), with $\lambda_n = kn^2\pi^2/L^2$. Still in analogy with the finite-dimensional case, one might hope to write a general function of x as a sum of the eigenfunctions. For a function of x and t we can try

$$u(x, t) = \sum_{n=1}^{\infty} b_n(t) \varphi_n(x), \quad (1.8)$$

and (1.2) leads to the equations

$$m \frac{d^2 b}{dt^2} = -\lambda_n b = -k \frac{n^2 \pi^2}{L^2} b.$$

Note the reversal of point of view here from (1.5). Either way we find the *standing wave* solutions

$$\sin\left(\frac{n\pi x}{L}\right) \cos\left(\sqrt{\frac{k}{m}} \frac{n\pi t}{L}\right), \quad n = 1, 2, 3, \dots \quad (1.9)$$

A few more historical remarks: Brooke Taylor found the $n = 1$ standing wave solution for the vibrating string in 1713, although he did not write the partial differential equation (1.2) found by D'Alembert. Euler noted in 1747 that linear combinations of functions (1.9) are solutions to (1.2), but he did not consider this to be the general form of a solution. Daniel Bernoulli wrote solutions as infinite sums of standing wave solutions in 1753, and *did* consider this to be the general form. Lagrange very nearly found this same form of the general solution in 1759.

The justification for the expansion (1.8) did not come until 1829, when Gustav Lejeune Dirichlet gave the first convergence proof for the Fourier expansion. The form (1.3) was also a source of controversy: do the functions F and G need to be differentiable? D'Alembert said yes, on mathematical grounds. Euler said no, on physical grounds – for example one can pull the string into a starting position that has a corner, and this should still yield an acceptable solution. A full resolution of this question came only in the 20th century, with the notions of weak solution (Leray, Sobolev) and distribution solution (Schwartz). For a full discussion of the early history see [16], [42], [61].

Chapter 2

Can one hear the density of a string?

The title of this chapter is borrowed from the title of a famous lecture by Mark Kac. That lecture and the associated problem will be discussed in Chapter 12. For now we consider a one-dimensional variation.

For a string of *variable* density m with fixed endpoints, vibrating vertically (with units suitably normalized), the position function $u(x, t)$ is again the solution of an equation with Dirichlet boundary conditions:

$$m(x)u_{tt}(x, t) = u_{xx}(x, t), \quad 0 < x < 1; \quad u(0, t) = u(1, t) = 0.$$

The standing wave solutions are

$$u_\nu(x, t) = \cos(\lambda_\nu t) \varphi_\nu(x)$$

where the φ_ν are the solutions of a generalized eigenvalue problem:

$$\varphi_\nu''(x) + \lambda_\nu^2 m(x) \varphi_\nu(x) = 0; \quad \varphi_\nu(0) = \varphi_\nu(1) = 0. \quad (2.1)$$

The density m determines the possible frequencies $\{\lambda_\nu\}$. The question whether the frequencies determine the density was touched on by Lord Rayleigh [78], sections 140, 141. Rayleigh gave a formal argument that suggests that when the density is symmetric about the midpoint, then it is determined uniquely by the frequencies – one can hear the density of a symmetric string (at least if one has perfect pitch and no upper limit to the frequencies one can detect).

If m is strictly positive and sufficiently regular, the Liouville transform, taking as a new independent variable

$$y(x) = \int_0^x m(x')^{1/2} dx',$$

converts the generalized eigenvalue problem (2.1) to a standard eigenvalue problem for an operator with potential term. Changing notation once again by writing x in place of y , the transformed problem can be normalized as an eigenvalue problem

$$\varphi_\nu''(x) + \lambda_\nu^2 \varphi_\nu(x) = q(x) \varphi_\nu(x); \quad \varphi_\nu(0) = \varphi_\nu(1) = 0. \quad (2.2)$$

The associated operator

$$L_q = \frac{d^2}{dx^2} - q(x) \quad (2.3)$$

assumed new importance in the 20th century as the (negative of) the one-dimensional Schrödinger operator of quantum mechanics. This led to renewed interest in the inverse problem.

In 1929 V. Ambarzumian [2] considered the eigenvalue problem for (2.3) with Neumann boundary conditions $\varphi'(0) = 0 = \varphi'(1)$ as a model case for inverse problems of quantum mechanics. When $q \equiv 0$ the eigenvalues are $\lambda_\nu^2 = n^2\pi^2$, $n = 0, 1, 2, \dots$. Ambarzumian showed conversely that if (2.2) has spectrum $\{n^2\pi^2\}$, then $q \equiv 0$. Göran Borg [15] studied this problem in great detail and showed in 1946 that in general the spectra for *two* sets of boundary conditions are necessary and sufficient to determine a unique potential; in particular neither the Dirichlet spectrum ($\varphi = 0$ at the endpoints) nor the Neumann spectrum ($\varphi' = 0$ at the endpoints) alone is sufficient. He also showed that when q is symmetric about the midpoint, then the Dirichlet spectrum suffices, as suggested by Rayleigh. Simpler proofs were found by Norman Levinson [60] and Vladimir Marčenko [62].

Instead of using two spectra, we could ask what additional information besides the $\{\lambda_\nu\}$ for (2.3) or (2.2) would suffice to fix m or q ? For example, one might seek to use the L^2 norms

$$a_\nu = \int_0^1 \varphi(x, \lambda_\nu)^2 dx$$

for solutions to (2.1) or to (2.2) that are normalized by

$$\varphi_\nu'(0) = 1.$$

For problem (2.1) these are the energies of the standing wave solutions. For problem (2.2) they determine the *spectral measure* $d\rho$ associated to the choice of eigenfunctions $\{\varphi(\cdot, \lambda_\nu)\}$, i.e. the measure such that the map U that takes f to

$$\widehat{f}(\lambda) = \int_0^\infty f(x) \varphi(x, \lambda) dx \quad (2.4)$$

is a unitary map

$$U : L^2([0, 1], dx) \rightarrow L^2(\mathbf{R}, d\rho).$$

Equivalently, the inverse map U^* is

$$f(x) = \int_{-\infty}^\infty \widehat{f}(\lambda) \varphi(x, \lambda) d\rho(\lambda) = \sum_\nu \frac{1}{a_\nu} \widehat{f}(\lambda_\nu) \varphi(x, \lambda_\nu). \quad (2.5)$$

In this case the measure ρ is supported on the set $\{\lambda_\nu\}$ and the point λ_ν has measure $1/a_\nu$.

Marčenko [62] noted that the spectral measure uniquely determines the potential, because of an identity due to Jean Delsarte and Aleksander Povsner. This identity is a continuous analog of the Gram-Schmidt orthogonalization process. For any complex λ , let $\varphi(\cdot, \lambda)$ be the solution to

$$\varphi'' + \lambda^2 \varphi = q \varphi; \quad \varphi|_{x=0} = 0, \quad \varphi'|_{x=0} = 1. \quad (2.6)$$

Let $\varphi_0(\lambda) = \sin \lambda x / \lambda$ be the solution to (2.6) for $q_0 \equiv 0$. Then the φ are related to the φ_0 by a Volterra integral equation

$$\varphi(x, \lambda) = \varphi_0(x, \lambda) + \int_0^x V(x, y) \varphi_0(y, \lambda) dy. \quad (2.7)$$

In operator form this is $\varphi = (I + V)\varphi_0$. Operators of the form $I + V$ are a group under composition, so it follows that there is a similar relation between the solutions φ_1 and φ_2 for potentials q_1 and q_2 . If q_1 and q_2 have the same spectral measure, then the corresponding unitary operators U_j have the property that $U_2^* U_1$ is a unitary operator in $L^2([0, 1], dx)$. But $\varphi_2 = (I + V)\varphi_1$ implies $U_2^* = (I + V)U_1^*$, so $I + V = U_2^* U_1$ is unitary. It follows that $(I + V)^{-1} = I + V^*$. But this is impossible unless $V = 0$, because the kernel $W(x, y) = \overline{V(y, x)}$ of V^* is supported where $y \geq x$ while that of $(I + V)^{-1}$ is supported where $y \leq x$.

The proof of (2.7) also involves Volterra integral equations. In fact it is not difficult to see that the conditions on V in order for the expression (2.7) to give solutions to (2.6) are that

$$V_{xx}(x, y) = V_{yy}(x, y) + q(x)V(x, y), \quad 0 < y < x \quad (2.8)$$

with boundary conditions

$$V(x, 0) = 0, \quad V(x, x) = \frac{1}{2} \int_0^x q(x') dx'. \quad (2.9)$$

As we shall see, relations like (2.7) have a number of important applications. Given the usefulness, it is surprising how simple the proof is. To prove the existence of V it is convenient to extend the problem to the wedge $|y| \leq x$ by looking for a solution to (2.8) that satisfies the modified boundary conditions

$$V(x, x) = \frac{1}{2} \int_0^x q(x') dx', \quad V(x, -y) = -V(x, y). \quad (2.10)$$

The function

$$v(x, y) = \frac{1}{2} \int_{(x-y)/2}^{(x+y)/2} q(s) ds$$

satisfies the boundary conditions (2.9) and the pure wave equation $vu_{xx} = v_{yy}$ for $|y| < x$. On the other hand the function

$$w(x, y) = Tf(x, y) \equiv \frac{1}{2} \int_0^x \int_{-x'}^{x'} f(x', y - y') dy' dx'$$

satisfies $w_{xx} = w_{yy} + f$. The Volterra operator T takes functions supported on $|y| \leq x$ to functions that are supported in the same region and vanish for $|y| = x$. Moreover, Tf is an odd function of y if f is. It follows from this that the desired function V is the solution of the Volterra integral equation

$$V = v + T(qV) \equiv v + T_q(V).$$

The solution is given by the convergent series

$$V = v + T_q(v) + T_q^2(v) + T_q^3(v) + \dots$$

This argument is easily adapted to other boundary conditions for the wave functions φ , such as

$$\varphi|_{x=0} = 1, \quad \varphi'|_{x=0} = h,$$

where h is a real constant.

For more historical discussion of the material in this chapter and the next two, see [64].

Chapter 3

Can one hear a nuclear potential?

Although Ambarzumian's study of the problem (2.2) on a finite interval was motivated by questions from quantum mechanics, the more typical quantum-mechanical problem involves an infinite interval. In particular, for a radial potential the Schrödinger operator is (2.3) on the half-line $[0, \infty)$. The following discussion depends on some decay of the (real) potential q , such as

$$\int_{-\infty}^{\infty} (1+x^2) |q(x)| dx < \infty.$$

The variation of constants solution

$$\varphi(x, \lambda) = \frac{\sin \lambda x}{\lambda} + \int_0^x \sin \lambda(x-y) q(y) \varphi(y, \lambda) dy$$

to the problem

$$\varphi'' + \lambda^2 \varphi = q\varphi, \quad x \geq 0; \quad \varphi(0, \lambda) = 0, \quad \varphi'(0, \lambda) = 1, \quad (3.1)$$

leads to the approximation

$$\varphi(x, \lambda) \sim \frac{A(\lambda)}{\lambda} \sin(\lambda x - \Phi(\lambda)), \quad x \rightarrow \infty, \quad (3.2)$$

where it is easy to check that

$$F(\lambda) \equiv A(\lambda) e^{i\Phi(\lambda)} = 1 + \int_0^{\infty} e^{i\lambda x} q(x) \varphi(x, \lambda) dx. \quad (3.3)$$

It is important to note that F extends to a holomorphic function of λ in the upper half-plane $\{\text{Im } \lambda > 0\}$. The *asymptotic phase* Φ can be normalized by choosing it to be continuous, with

$$\lim_{\lambda \rightarrow +\infty} \Phi(\lambda) = 0.$$

Note that (3.2) can be rewritten

$$\varphi(x, \lambda) \sim F(\lambda)[e^{i\lambda x} - S(\lambda)e^{-i\lambda x}]$$

where the *scattering function* S is

$$S(\lambda) = \frac{F(-\lambda)}{F(\lambda)} = e^{-2i\Phi(\lambda)}. \quad (3.4)$$

This is the simplest case of the *scattering matrix* introduced by John Wheeler [89] in 1937 and Werner Heisenberg [44] in 1943. Heisenberg proposed that the scattering matrix was the basic physically observable quantity of elementary particle theory and that it should therefore determine the potential q .

The question of determining the potential q from the phase Φ or, equivalently, the scattering function S was studied by C. E. Fröberg [32] in 1948. Levinson [59] showed in 1949 that the asymptotic phase uniquely determines the potential when there is no discrete spectrum. As we shall see, this is a special case of the extension to the half-line of Marčenko's result that was sketched in the previous chapter.

Note that in general there may be discrete spectrum: finitely many values $\lambda_\nu^2 < 0$ for which $\varphi_\nu = \varphi(\cdot, \lambda_\nu)$ has L^2 norm $\|\varphi_\nu\| < \infty$; we take $\text{Im } \lambda_\nu > 0$. These λ_ν are precisely the zeros of F in the upper half plane. V. Bargmann [8] showed that when there is discrete spectrum, the scattering function does not suffice to determine the potential.

In the absence of discrete spectrum, $\log A$ and Φ are the real and imaginary parts of the value on the real line of the function $\log F$, holomorphic in the upper half plane. This implies that either of $\log A$ or Φ is essentially the Hilbert transform of the other. To see this, consider for simplicity a smooth real-valued function p such that

$$\int_{-\infty}^{\infty} |p(s)| ds < \infty.$$

The function

$$f(\lambda) = \frac{1}{\pi i} \int_{-\infty}^{\infty} \frac{p(s) ds}{s - \lambda}$$

is holomorphic in the upper half plane and

$$f(t + \varepsilon i) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\varepsilon p(s) ds}{(t-s)^2 + \varepsilon^2} + \frac{i}{\pi} \int_{-\infty}^{\infty} \frac{(s-t)p(s) ds}{(t-s)^2 + \varepsilon^2}.$$

The real part can be rewritten as

$$\operatorname{Re} f(t + \varepsilon i) = \int_{-\infty}^{\infty} g(s) p(t - \varepsilon s) ds, \quad g(t) = \frac{1}{\pi(1+t^2)}.$$

Since g has integral 1 it follows that on the line, taking the limit $\varepsilon \rightarrow 0$ gives $\operatorname{Re} f = p$. It follows that

$$\operatorname{Im} f(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\operatorname{Re} f(s) ds}{t-s}.$$

Applying this to $\log F$ and to $i \log F$ we get

$$\Phi(\lambda) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{A(s) ds}{\lambda - s}$$

and

$$A(\lambda) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\Phi(s) ds}{\lambda - s}.$$

Thus in the absence of discrete spectrum, the asymptotic amplitude A can be recovered from the asymptotic phase Φ and conversely, so Levinson's result shows that in this case the asymptotic amplitude A also determines q uniquely.

When there *is* discrete spectrum, $\{\lambda_\nu\}$, one can consider the modification

$$\tilde{F}(\lambda) = F(\lambda) \prod_{\nu} \left(\frac{\lambda - \bar{\lambda}_\nu}{\lambda - \lambda_\nu} \right).$$

This function has no zeros in the upper half plane and has the same modulus as F on the real line. In fact

$$\tilde{F}(\lambda) = A e^{i\tilde{\Phi}(\lambda)}, \quad \tilde{\Phi}(\lambda) = \Phi(\lambda) - 2 \sum_{\nu} \arg(\lambda - \lambda_\nu),$$

and either of A or Φ , together with $\{\lambda_\nu\}$, determines the other.

Much earlier (1910), Weyl [84] had computed the spectral measure associated to (3.1). Generalizing the case of a finite interval, this is the measure $d\rho$ on the line with the property that the map from $L^2((0, \infty))$ to $L^2(\mathbf{R}, d\rho)$

$$f \rightarrow Uf(\lambda) = \int_0^{\infty} \varphi(x, \lambda) f(x) dx,$$

where $\varphi(\cdot, \lambda)$ is the normalized wave function of (3.1), has inverse

$$U^*g(x) = \int_{-\infty}^{\infty} \varphi(x, \lambda) g(\lambda) d\rho(\lambda).$$

In $(-\infty, 0)$ we take the measure $d\rho$ to be supported on the points λ_ν^2 of the discrete spectrum, with mass $c_\nu = \|\varphi(\cdot, \lambda_\nu)\|^{-2}$. Then

$$U^*g(x) = \sum_{\nu} c_{\nu} g(\lambda_{\nu}^2) \varphi_{\nu}(x) + \int_0^{\infty} \varphi(x, \lambda) g(\lambda) d\rho(\lambda).$$

We expect $d\rho$ to be absolutely continuous on $[0, \infty)$. It can be computed from the information above. In fact it should satisfy

$$\lim_{N \rightarrow \infty} \int_0^N \varphi(x, \lambda) \varphi(y, \lambda) d\rho(\lambda) = \delta(x - y)$$

for $x, y > 0$. We take x, y very large and use the asymptotics (3.2) to find that the integrand is

$$\sim \frac{A(\lambda^2)}{2\lambda^2} \{ \cos \lambda(x - y) - \cos[\lambda(x + y) + 2\Phi(\lambda)] \}.$$

The second part of the expression on the right oscillates rapidly as a function of λ when $x + y \gg 0$. Therefore the principal contribution to the integrand comes from the part involving $\cos \lambda(x - y)$. Now

$$\int_0^N \cos \lambda x d\lambda = \frac{\sin Nx}{x}$$

and as distribution this converges to $c\delta(x - u)$ where

$$c = \int_{-\infty}^{\infty} \frac{\sin Nx}{x} dx = \pi.$$

Therefore we have

$$d\rho(\lambda) = \frac{2\lambda^2}{\pi A(\lambda)^2} d\lambda, \quad \lambda \geq 0. \quad (3.5)$$

In other words the asymptotic amplitude $A = |F|$ determines the spectral measure ρ modulo the discrete spectrum $\{\lambda_\nu\}$ and the norms $\|\varphi_\nu\|$, and conversely.

As in the case of the problem on a finite interval, the function (2.6) is related to the solution $\varphi_0(x, \lambda) = \sin \lambda x / \lambda$ for $q_0 \equiv 0$ by the Volterra integral equation

$$\varphi(x, \lambda) = \varphi_0(x, \lambda) + \int_0^x V(x, y) \varphi_0(y, \lambda) dy. \quad (3.6)$$

Once again, as observed by Marčenko, this implies that the spectral measure determines the potential uniquely. See [64] for further discussion.

Chapter 4

Can one reconstruct the potential?

There is a large gap between knowing that some data, such as a pair of spectra or a spectral measure, can belong at most to one density or one potential, and being able to reconstruct the density or the potential from that data. Elegant solutions of reconstruction problems for equations like (3.1) were found in the 1950s. In 1951 I. M. Gel'fand and B. M. Levitan [37] showed how to reconstruct q from the spectral measure ρ , while in 1955 Marčenko [63] reconstructed q directly from the asymptotic phase and the discrete part of the spectral measure; see the discussion in [64]. Much work has been done since then on systems, higher dimensional operators, and numerical implementations. We refer to the book by Chadan and Sabatier [21] and its foreword by Roger Newton for an extensive treatment of these developments and their history.

The basic idea of Gel'fand and Levitan for the boundary value problem

$$\begin{aligned}\varphi_{xx}(x, \lambda) + \lambda^2 \varphi(x, \lambda) &= q(x) \varphi(x, \lambda), \quad x > 0; \\ \varphi(0, \lambda) &= a, \quad \varphi_x(0, \lambda) = b\end{aligned}$$

was that the kernel K that converts the solutions φ_0 of the corresponding equation with potential $q_0 \equiv 0$ to φ by the formula

$$\varphi(x, \lambda) = \varphi_0(x, \lambda) + \int_0^x K(x, y) \varphi_0(y, \lambda) dy \quad (4.1)$$

can be computed from the spectral measure ρ by making use of orthogonality relations

$$\int_{-\infty}^{\infty} \varphi(x, \lambda) \varphi(y, \lambda) d\rho(\lambda) = \delta(x - y) \quad (4.2)$$

and the corresponding relations for φ_0 .

The spectral measure $d\rho_0$ that corresponds to $q_0 \equiv 0$ is readily computed. There is an integral equation similar to (4.1) that expresses $\varphi_0(y, \lambda)$ in terms of $\varphi(z, \lambda)$, $z \leq y$. Putting this into (4.1) gives

$$\int_{-\infty}^{\infty} \varphi(x, \lambda) \varphi_0(y, \lambda) d\rho(\lambda) = 0, \quad y < x. \quad (4.3)$$

For $x \neq y$ set

$$f(x, y) = \int_{-\infty}^{\infty} \varphi_0(x, \lambda) \varphi_0(y, \lambda) d\sigma(\lambda), \quad \sigma = \rho - \rho_0.$$

Combining (4.1) and (4.3) for $y < x$ gives

$$\begin{aligned} 0 &= \int_{-\infty}^{\infty} \varphi(x, \lambda) \varphi_0(y, \lambda) d\rho(\lambda) \\ &= f(x, y) + \int_0^x K(x, z) \left\{ \int_{-\infty}^{\infty} \varphi_0(z, \lambda) \varphi_0(y, \lambda) d\rho(\lambda) \right\} dz \\ &= f(x, y) + \int_0^x K(x, z) \{ \delta(z - y) + f(z, y) \} dz \\ &= f(x, y) + K(x, y) + \int_0^x K(x, z) f(z, y) dz. \end{aligned}$$

Thus for each $x > 0$ we have an integral equation for $K(x, \cdot)$ on the interval $[0, x]$,

$$K(x, y) + \int_0^x K(x, z) f(z, y) dz = -f(x, y), \quad (4.4)$$

the famous *Gel'fand–Levitan equation*. The potential is then recovered as

$$q(x) = 2 \frac{d}{dx} \{ W(x, x) \}.$$

It follows from earlier remarks that the Gel'fand–Levitan method also – in principle – recovers the potential from the asymptotic phase Φ (or the scattering function $S = e^{-2i\Phi}$), together with the discrete part of the spectral measure. However the Hilbert transform of Φ is highly sensitive to the detailed behavior of Φ . Marčenko developed a more stable process of recovering q from Φ and the discrete part of ρ by scaling away the amplitude. Let

$$\psi(x, \lambda) = \frac{\lambda}{A(\lambda)} \varphi(x, \lambda) \sim \psi_0(x, \lambda) \equiv \sin(\lambda x - \Phi(\lambda)) \quad \text{as } x \rightarrow \infty.$$

These functions satisfy the orthogonality relation

$$\int_{-\infty}^{\infty} \psi(x, \lambda) \psi(y, \lambda) dr(\lambda) = \delta(x - y),$$

where the continuous part of $dr(\lambda)$ is $(2/\pi)d\lambda$, $\lambda \geq 0$. To simplify the exposition slightly we assume now that there is no discrete spectrum. Note that $(2/\pi)d\lambda$ is also the spectral measure for the functions $\sin \lambda x$ that correspond to $\psi(x, \lambda)$ when $q = 0$.

There is an analogue here, due to Boris Levin, of the formula (3.6):

$$\psi(x, \lambda) = \psi_0(x, \lambda) + \int_x^{\infty} V(x, y) \psi_0(y, \lambda) dx \quad (4.5)$$

and a similar formula expressing ψ_0 in terms of ψ . This second formula leads to the orthogonality relation for $x < y$:

$$\frac{2}{\pi} \int_0^{\infty} \psi(x, \lambda) \psi_0(y, \lambda) d\lambda = 0. \quad (4.6)$$

Combining (4.5) and (4.6),

$$\begin{aligned} 0 &= \frac{2}{\pi} \int_0^{\infty} \psi_0(y, \lambda) \left\{ \psi_0(x, \lambda) + \int_x^{\infty} V(x, z) \psi_0(z, \lambda) dy \right\} d\lambda \quad (4.7) \\ &= \frac{2}{\pi} \int_0^{\infty} \psi_0(y, \lambda) \psi_0(x, \lambda) d\lambda \\ &\quad + \int_x^{\infty} V(x, y) \left\{ \frac{2}{\pi} \int_0^{\infty} \psi_0(y, \lambda) \psi_0(z, \lambda) d\lambda \right\} dz. \end{aligned}$$

Now

$$\begin{aligned} &\frac{2}{\pi} \int_0^{\infty} \psi_0(x, \lambda) \psi_0(y, \lambda) d\lambda \\ &= \delta(x - y) + \frac{2}{\pi} \int_0^{\infty} \{ \psi_0(x, \lambda) \psi_0(y, \lambda) - \sin \lambda x \sin \lambda y \} d\lambda \\ &= \delta(x - y) + \frac{1}{\pi} \int_0^{\infty} \{ \cos(\lambda x + \lambda y) - \cos(\lambda x + \lambda y - 2\Phi) \} d\lambda. \end{aligned}$$

Since Φ is odd, this can be rewritten as

$$\frac{2}{\pi} \int_0^{\infty} \psi_0(x, \lambda) \psi_0(y, \lambda) d\lambda = \delta(x - y) + f(x + y), \quad (4.8)$$

where

$$f(x) = \frac{1}{2\pi} \int_0^\infty \cos(\lambda x) [1 - \cos(2\Phi)] d\lambda$$

is determined from Φ . Substituting (4.8) into (4.7) gives

$$0 = f(x+y) + V(x,y) + \int_x^\infty V(x,z) f(y+z) dz, \quad x < y. \quad (4.9)$$

This is the *Marčenko equation* (in the absence of discrete spectrum), that determines the Levin kernel V from the asymptotic phase Φ . The potential q is

$$q(x) = -2 \frac{d}{dx} [V(x,x)].$$

Including the discrete spectrum is left as an exercise.

Chapter 5

Quantum mechanics in one dimension

Although the operators considered in the previous two chapters were one-dimensional, the fact that the relevant interval was $[0, \infty)$ was due to an assumption that a potential $V(x)$ in higher dimensions was a function of the norm $|x|$. For a truly one-dimensional problem one would have both the potential q and the wave functions $\varphi(\cdot, \lambda)$ defined on $\mathbf{R} = (-\infty, +\infty)$. Although a problem of limited physical interest (for quantum mechanics), this turns out, unexpectedly, to have considerable interest mathematically and has led to many consequences in other physical theories.

The Gel'fand–Levitan and Marčenko methods can be adapted to this case, but it will be of interest to describe a modification based on the *Jost solutions* [49] of the basic equation

$$(L_q + \lambda^2)\varphi(x, \lambda) \equiv \varphi_{xx}(x, \lambda) + [\lambda^2 - q(x)]\varphi(x, \lambda) = 0, \quad x \in \mathbf{R}. \quad (5.1)$$

We continue to assume that q is real and $\int(1+x^2)|q(x)|dx < \infty$. For real λ there are solutions with prescribed asymptotics as $-x \rightarrow \pm\infty$:

$$\begin{aligned} \varphi_+(x, \lambda) &= e^{-i\lambda x} + \int_{-\infty}^x \sin(\lambda(x-y)) q(y) \varphi_+(y, \lambda) dy; \\ \varphi_-(x, \lambda) &= e^{-i\lambda x} - \int_x^{\infty} \sin(\lambda(x-y)) q(y) \varphi_-(y, \lambda) dy. \end{aligned}$$

The functions $\tilde{\varphi}_{\pm}(x, \lambda) = \varphi_{\pm}(x, -\lambda)$ are also solutions. For real λ , these are just the complex conjugates $\tilde{\varphi}_{\pm} = \bar{\varphi}_{\pm}$, but as shall see there are various extensions to complex λ .

The functions $\chi_{\pm}(x, \lambda) = \varphi_{\pm}(x, \lambda)e^{i\lambda x}$ are solutions of the integral equations

$$\begin{aligned}\chi_+(x, \lambda) &= 1 + \int_{-\infty}^x \left[\frac{e^{i\lambda(2x-2y)} - 1}{2i} \right] q(y) \chi_+(y, \lambda) dy; \\ \chi_-(x, \lambda) &= 1 - \int_x^{\infty} \left[\frac{e^{i\lambda(2x-2y)} - 1}{2i} \right] q(y) \chi_-(y, \lambda) dy.\end{aligned}$$

These formulas extend naturally to λ in the upper and lower half-planes respectively, so χ_+ and $\tilde{\chi}_- = \tilde{\varphi}_- e^{-i\lambda x}$ extend to $\{\text{Im } \lambda > 0\}$ while χ_- and $\tilde{\chi}_+$ extend to $\{\text{Im } \lambda < 0\}$, holomorphic in λ . As a consequence, φ_+ and $\tilde{\varphi}_-$ extend to the upper half-plane while φ_- and $\tilde{\varphi}_+$ extend to the lower half-plane. For $\text{Im } \lambda > 0$, $\varphi_+(x, \lambda)$ decays exponentially as $x \rightarrow -\infty$ while $\tilde{\varphi}_-(x, \lambda)$ decays exponentially as $x \rightarrow +\infty$. Unless they are dependent, each grows exponentially in the other direction. It follows that negative eigenvalues are exactly the λ^2 with $\text{Im } \lambda > 0$ for which the Wronskian $W(\varphi_+, \tilde{\varphi}_-)$ vanishes.

The function φ_- and $\tilde{\varphi}_-$ are independent solutions of (5.1) for real λ , so

$$\varphi_+(x, \lambda) = a(\lambda) \varphi_-(x, \lambda) + b(\lambda) \tilde{\varphi}_-(x, \lambda), \quad \lambda \text{ real.} \quad (5.2)$$

Then

$$\begin{aligned}\varphi_+(x, \lambda) &\sim e^{-i\lambda x}, \quad x \rightarrow -\infty; \\ &\sim a(\lambda) e^{-i\lambda x} + b(\lambda) e^{i\lambda x}, \quad x \rightarrow +\infty.\end{aligned}$$

Thus a and b are determined entirely by asymptotics of solutions.

The Wronskians $W(\varphi_-, \tilde{\varphi}_-) \equiv \varphi_- \tilde{\varphi}'_- - \varphi'_- \tilde{\varphi}_-$ and $W(\varphi_+, \tilde{\varphi}_-)$ are constant; the asymptotics imply that each is $2i\lambda$. Using (5.2) and its complex conjugate to compute $W(\varphi_+, \tilde{\varphi}_+)$ in two different ways, we obtain the relation

$$|a(\lambda)|^2 - |b(\lambda)|^2 = 1.$$

Thus $a \neq 0$, Using this fact and the definition of χ_{\pm} , (5.2) can be rewritten as

$$\frac{\chi_+(x, \lambda)}{a(\lambda)} = \chi_-(x, \lambda) + r(\lambda) \tilde{\chi}_+(x, \lambda), \quad r = \frac{b}{a}. \quad (5.3)$$

The function r is called the *reflection coefficient*.

The coefficient a can be written as a Wronskian,

$$a(\lambda) = \frac{W(\varphi_+, \tilde{\varphi}_-)}{2i\lambda},$$

so a is holomorphic in the upper half-plane and the squares of its zeros $\{\lambda_\nu\}$ are the negative eigenvalues.

The Fourier transform is essential in what follows. The Fourier transform of an integrable function g on the line is defined by

$$\widehat{g}(x) = \int_{-\infty}^{\infty} e^{-ixy} g(y) dy. \quad (5.4)$$

If \widehat{g} is also integrable, then g is its inverse Fourier transform:

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ixy} \widehat{g}(y) dy. \quad (5.5)$$

Conversely, if \widehat{g} is integrable and g defined by (5.5) is integrable, then (5.4) recovers \widehat{g} from g . These identities, suitably interpreted, carry over to $g \in L^2$ and \widehat{g} in L^2 , and

$$\int_{-\infty}^{\infty} |g(x)|^2 dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\widehat{g}(x)|^2 dx.$$

For each x , $\chi_-(x, \cdot) - 1$ is holomorphic in the lower half-plane and is $O(1/|\lambda|)$ for large λ . Then for $y < 0$ the integrand in

$$A(x, y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{iyz} [\chi_-(x, z) - 1] dz \quad (5.6)$$

is holomorphic for $\text{Im } z < 0$ and is both $O(1/|z|)$ and $O(\exp(-|y \text{Im } z|))$ in the lower half-plane. By Cauchy's theorem the integral vanishes. Thus $A(x, y) = 0$ for $y < 0$, and the inversion formula gives

$$\begin{aligned} \chi_-(x, \lambda) &= 1 + \int_0^{\infty} e^{-i\lambda y} A(x, y) dy \\ &= 1 + \int_x^{\infty} e^{-i\lambda(x+y)} V(x, y) dy, \quad V(x, y) \equiv A(x, y - x). \end{aligned} \quad (5.7)$$

It follows from the identity $\chi_-(x, -\lambda) = \bar{\chi}_-(x, \lambda)$ for real λ that A is real.

We are now in a position to derive the Marčenko equation for the problem on the line. It is convenient to distinguish three cases. First suppose that there are no bound states: a has no zeros in $\{\text{Im } \lambda > 0\}$. The identity (5.3) can be written as

$$\frac{\chi_+}{a} - 1 = (\chi_- - 1) + r e^{2i\lambda x} + (\tilde{\chi}_- - 1) r e^{2i\lambda x}.$$

Once again the inverse Fourier transform of the function on the left vanishes on $[0, \infty)$. Therefore

$$0 = A(x, y) + f(2x + y) + \int_0^\infty f(2x + y + z)A(x, z) dz, \quad y > 0,$$

where f is the inverse Fourier transform of the reflection coefficient r . This can be rewritten as the Marčenko equation

$$0 = f(x + y) + V(x, y) + \int_x^\infty V(x, z)f(y + z) dz, \quad y < x. \quad (5.8)$$

Thus V is determined by the reflection coefficient. As before, the potential $q(x) = 2dV(x, x)/dx$.

Next, suppose that there are bound states with eigenvalues $-\lambda_\nu^2$. (The number of bound states is necessarily finite.) Suppose also that the reflection coefficient $r \equiv 0$: q is a *reflectionless potential*. At each zero λ_ν of a , χ_+ is a multiple c_ν of $\tilde{\chi}_- e^{2i\lambda_\nu}$. Therefore the rational function

$$\eta(\lambda) = \sum_\nu \frac{c_\nu}{a'(\lambda_\nu)(\lambda - \lambda_\nu)}$$

has the property that the function

$$\frac{\chi_+}{a} - 1 + \eta \tilde{\chi}_-$$

is holomorphic and $O(1/|\lambda|)$ in the upper half plane. Writing this function as

$$\frac{\chi_+}{a} - 1 + \eta \tilde{\chi}_- = (\chi_- - 1) + \eta(\tilde{\chi}_- - 1)e^{2i\lambda x} + \eta e^{2i\lambda x},$$

we obtain (5.8) with f as the inverse Fourier transform of η .

Finally, suppose that there are both bound states and a non-vanishing reflection coefficient. Choose the rational function η as in the second case. Then

$$\frac{\chi_+}{a} - 1 + \tilde{\chi}_- = (\chi_- - 1) + (r + \eta)(\tilde{\chi}_- - 1)e^{2i\lambda x} + (r + \eta)e^{2i\lambda x},$$

and we get (5.8) with f as the inverse Fourier transform of $r + \eta$.

Let us return to the case of a reflectionless potential. In this case f is a linear combination of decaying exponentials on $[0, \infty)$ and the problem of

determining V is purely algebraic; the potential q can be written explicitly as a rational combination of exponentials. In fact, with the notation above,

$$f(x) = \sum_{\nu} \alpha_{\nu} e^{i\lambda_{\nu}x}, \quad \alpha_{\nu} \equiv \frac{ic_{\nu}}{a'(\lambda_{\nu})}.$$

Therefore we may write $f(x+y) = \sum_{\nu} b_{\nu}(x) a_{\nu}(y)$, where $b_{\nu} = \alpha_{\nu} a_{\nu}$, and it follows that V has the form

$$V(x, y) = \sum_{\nu} c_{\nu}(x) a_{\nu}(y).$$

Inserting these expressions into (5.8) we obtain

$$c_{\mu} = -\sum_{\nu} (A^{-1})_{\mu\nu} b_{\nu}(x), \quad A_{\mu\nu}(x) = \delta_{\mu\nu} + \int_x^{\infty} b_{\mu}(z) a_{\mu}(x) dz.$$

It follows that

$$\begin{aligned} V(x, x) &= -\sum_{\nu} (A^{-1})_{\mu\nu} b_{\nu}(x) \alpha_{\mu}(x) \\ &= \sum_{\nu} (A^{-1})_{\mu\nu} \frac{dA_{\nu\mu}}{dx} = \text{tr}(A^{-1}A_x) \\ &= \frac{1}{\det(A)} \frac{d}{dx} \det(A) = \frac{d}{dx} \log \det A. \end{aligned}$$

Therefore the potential is the second logarithmic derivative

$$q(x) = 2 \frac{d^2}{dx^2} (\log \det A).$$

Chapter 6

Preserving the spectrum: KdV

Borg's result (see Chapter 2) and the results on the one-dimensional Schrödinger operator show that knowing the spectrum for one self-adjoint realization of an operator $L_q = (d/dx)^2 - q$ in an interval $J \subset (-\infty, +\infty)$ is not enough to determine the potential uniquely. This raises questions about changing the operator without changing its spectrum, for example finding a one-parameter family of operators unitarily equivalent to L_q . This problem was addressed in the following way by Peter Lax [57], for reasons that will be discussed later.

Suppose that $\{U(t)\}$ is a one-parameter family of unitary operators in $L^2(J)$ that preserve the domain of a particular self-adjoint realization of L_q , say with $U(0) = I$, and consider $L_q(t) = U(t)^* L_q U(t)$. The question of interest here is whether the family can be chosen so that $L_q(t)$ has the same form as L_q , but with a potential $q(\cdot, t)$ depending on t . Such a family of potentials is called an *isospectral flow*. Formally, the family of unitary operators is differentiable in t if it is generated by a family of skew hermitian operators $A(t)$:

$$U'(t) = U(t)A(t), \quad U(0) = I.$$

If so, then the derivative of $L_q(t)$ is a commutator:

$$\frac{d}{dt}L_q(t) = [L_q(t), A(t)] \equiv L_q(t)A(t) - A(t)L_q(t).$$

If $L_q(t)$ is to be $(d/dx)^2 - q(\cdot, t)$, then

$$q_t = [A(t), L_{q(\cdot, t)}], \tag{6.1}$$

This is known as the *Lax equation*. (We have identified the function q_t with the operator of multiplication by q_t .) Equation (6.1) puts a strong constraint on the form of $A(t)$: its commutator with a second order differential operator is an operator of order zero – a multiplication operator. It is natural to take $A(t)$ to be, like $L_{q(\cdot,t)}$, a differential operator. In order for the commutator to have order zero, there are strong constraints on the coefficients, the first of which is that highest order and next highest order coefficient be constant. The choice $A(t) = d/dx$ leads to

$$q_t + q_x = 0$$

so $q(x, t) = q(x - t)$. This already illustrates a point: endpoints and explicit boundary conditions are a problem. Therefore we restrict the discussion here to the problem on the full line: one-dimensional quantum mechanics. Then the solution $q(x, t) = q(x - t)$ works, but rather trivially: just a change in choice of origin.

Continuing, if A is of order 2, it can be checked that A must be a linear combination of d/dx and L_q itself, with constant coefficients, so again the solution is $q(x, t) = q(x - ct)$. At third order we may subtract a constant multiple of L_q and a constant multiple of d/dx and assume that

$$A(t) = \frac{d^2}{dx^3} + b(\cdot, t) \frac{d}{dx} + c(\cdot, t),$$

with coefficients that vanish at $\pm\infty$. (We assume throughout the rest of this chapter that q is real and that q and its derivatives vanish sufficiently rapidly at $\pm\infty$.) Working out the commutator shows that

$$\begin{aligned} 3q_x + 2b_x &= 0, \\ 3q_{xx} + b_{xx} + 2c_x &= 0, \\ q_t + q_{xxx} + bq_x + c_{xx} &= 0, \end{aligned}$$

so

$$A(t) = \frac{d^3}{dx^3} - \frac{3}{2} q(x, t) \frac{d}{dx} - \frac{3}{4} q_x(x, t) \tag{6.2}$$

and the evolution equation for the potential q is

$$q_t + \frac{3}{2} q q_x - \frac{1}{4} 0q_{xxx} = 0. \tag{6.3}$$

There are corresponding nonlinear evolution equations of every higher odd order in x that determine isospectral flows of L_q .

Let us combine these observations with the solution of the corresponding inverse problem from the previous chapter. The Lax equation (6.1) can be viewed as the compatibility condition for the pair of equations

$$\begin{aligned}\varphi_t(x, t, \lambda) + A(t)\varphi(x, t, \lambda) - i\lambda^3\varphi(x, t) &= 0; \\ \varphi_{xx}(x, t, \lambda) + \lambda^2\varphi(x, t, \lambda) &= q(x, t, \lambda).\end{aligned}\tag{6.4}$$

Conversely, suppose (6.1) is true and that φ is a solution of the first of these equations. Then if the second equation holds at $t = 0$, it holds for all t .

With $A(t)$ as in (6.2) and $q \equiv 0$, the function $e^{-i\lambda x}$ is a solution of both these equations. Since this function defines the asymptotics as $x \rightarrow \pm\infty$ of φ_{\pm} , it follows that the functions $\varphi_{\pm}(\cdot, t, \lambda)$ for potentials $q(x, t)$ that satisfy (6.3) are solutions of (6.4). Taking the time derivative in (5.2) as $x \rightarrow +\infty$ we find that the coefficients a and b satisfy

$$a_t(\lambda, t) = 0, \quad b_t(\lambda, t) = -2i\lambda^3 b(\lambda, t).$$

Thus

$$a(\lambda, t) \equiv a(\lambda), \quad b(\lambda, t) = e^{-2i\lambda^3 t} b(\lambda), \quad r(\lambda, t) = e^{-2i\lambda^3 t} r(\lambda),$$

In the absence of discrete spectrum we can in principle recover $q(x, t)$ from $r(\lambda, t)$. Therefore the initial value problem for the *nonlinear* evolution equation (6.3) can be found by solving a linear integral equation, the Marčenko equation (5.8)!

What about the general case, with discrete spectrum? Since a is independent of t the discrete spectrum $\{-\lambda_{\nu}^2\}$ is time-invariant. The residues of the corresponding rational function η are constant multiples of the $c_{\nu}(t) = \tilde{\varphi}_{-}(x, \lambda)/\varphi_{+}(x, \lambda_{\nu})$, and just as for b we find that $c_{\nu}(t) = e^{-2i\lambda_{\nu}^3 t} c_{\nu}$. Therefore the solution of the initial value problem may again be found by linear methods.

When the reflection coefficient r is initially 0, the potentials $q(\cdot, t)$ continue to be reflectionless and, as we noted in the previous chapter, constructing them is an algebraic process that can be carried out explicitly. This gives a class of explicit solutions of (6.3), corresponding to arbitrary choices of parameters $\lambda_{\nu} = il_{\nu}$, $l_{\nu} > 0$ and c_{ν} , $1 \leq \nu \leq n$.

All this would be an only interesting curiosity if the equation (6.3) were of no interest. As it happens, it is the *Korteweg-de Vries equation* (KdV) and has an interesting history that comes in two parts, sketched in the next two chapters. We remark here that by rescaling x and multiplying q by a

constant, one can transform (6.3) to an equation of the same form with any choice of constants; a common choice is

$$q_t + 6qq_x + q_{xxx} = 0.$$

In this form there is a one-parameter family of solutions

$$q_c(x, t) = \frac{c}{2 \cosh^2[c^{1/2}(x - ct)]}, \quad c > 0. \quad (6.5)$$

These are *travelling wave solutions*: a wave with a single maximum amplitude $c/2$ travelling to the right at velocity c . Suitably transformed, they correspond to a reflectionless potential with a single eigenvalue. Rather than resort to the method sketched above, however, one can derive these solutions using the following observations. First, if q is a solution of (6.3) and $a \neq 0$ is real, then

$$q_a(x, t) = a^2 q(ax, a^3 t)$$

is also a solution. Second, $q(x, t) = f(x - t)$ is a travelling wave solution with velocity one if and only if $-f' + 6ff' + f''' = 0$. Assuming that f and derivatives decay rapidly at infinity this can be integrated to $-f + 3f^2 + f'' = 0$. Multiplying by f' and integrating again gives $-f^2 + 2f^3 + (f')^2 = 0$ or

$$\frac{f'}{f} = \pm \sqrt{1 - 2f}. \quad (6.6)$$

It is easy to check that $f(x) = 1/[2 \cosh^2(x/2)]$ is a solution.

To this point in our exposition it is clear that equation (6.3) might have been discovered by pure thinking, starting with the question of finding a one-parameter family of unitarily equivalent one-dimensional Schrödinger operators. Combined with the known inversion method, this would give an interesting, but not compelling, example of finding explicit solutions to a nonlinear PDE. But this is not how (6.3) was discovered, and not how it was found to be connected to the Schrödinger operator.

Chapter 7

Can one explain the shape of a wave?

In 1834 a young scientist named John Scott Russell made the first known observation of a phenomenon that was to be debated for the rest of the century:

I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses when the boat suddenly stopped – not so the mass of water in the channel which it had put into motion: it accumulated around the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth, and well defined heap of water, which continued in its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished and after a chase of one or two miles I lost it in the windings of the channel. [79]

Russell's observation and his subsequent study of the "Great Solitary Wave" were controversial. The equations for long waves in a shallow channel, a specialization of Euler's equations for incompressible irrotational flow, were well known but not easy to integrate, and some argued that a wave with elevation but no depression of the surface was not possible. The possibility of a travelling wave solution was examined in 1871 by Joseph de Boussinesq [17]

and in 1876 by Lord Rayleigh [77] who, after some approximations, derived essentially the equation ((6.6) and noted that its solutions would have the shape described by Russell. Nevertheless, controversy continued until the 1895 paper by Diederik Korteweg and Gustav de Vries that introduced equation (6.3). For historical discussions, see Craik [23] and the introduction to [18].

The KdV equation for long one-dimensional waves in water of uniform depth can be derived as follows, cf [52], [48]. First, suppose that fluid of density $\rho = \rho(x, t)$ is moving with velocity vector $\mathbf{u} = \mathbf{u}(x, t)$ through a region of \mathbf{R}^3 that contains a subregion Ω with smooth boundary $\partial\Omega$. (We take \mathbf{u} and other vectors here to be 3×1 column vectors.) The rate of change of the mass of fluid in Ω is the same as the rate of flow inward through the boundary, so

$$\frac{d}{dt} \int_{\Omega} \rho \, d\mathbf{x} + \int_{\partial\Omega} \rho(\mathbf{u} \cdot \mathbf{n}) \, d\sigma = 0,$$

where \mathbf{n} is the outer unit normal. The divergence theorem converts this to the vanishing of an integral over Ω , and since it must hold for each region Ω we get

$$\rho_t + \nabla \cdot (\rho\mathbf{u}) = 0,$$

where ∇ denotes the \mathbf{x} -gradient. Let \mathbf{F} be the external force on the fluid and p the pressure. By Newton's second law the rate of change of momentum $\rho\mathbf{u}$ in Ω plus the rate of flow of momentum in through the boundary should equal the external force plus the force on the boundary due to pressure p :

$$\frac{d}{dt} \int_{\Omega} \rho\mathbf{u} \, d\mathbf{x} + \int_{\partial\Omega} \rho(\mathbf{u} \cdot \mathbf{n}) \, d\sigma = - \int_{\partial\Omega} p\mathbf{n} \, d\sigma + \int_{\Omega} \mathbf{F} \, d\mathbf{x}.$$

As before this leads to a vector equation

$$(\rho\mathbf{u})_t + (\mathbf{u} \cdot \nabla)(\rho\mathbf{u}) + \rho(\nabla \cdot \mathbf{u})\mathbf{u} + \nabla p = \rho\mathbf{F}.$$

Now consider boundary conditions. On a fixed surface, like the bottom of a channel, the normal component of \mathbf{u} must be zero: $\mathbf{u} \cdot \mathbf{n} = 0$. On a free surface given by $f(\mathbf{x}, t) = 0$ the normal velocity of the surface is the normal component of u . The unit normal to the surface is $|\nabla f|^{-1} \nabla f$, so

$$0 = \frac{d}{dt} f(\mathbf{x}(t), t) = f_t + \mathbf{u} \cdot \nabla f.$$

Finally, ignoring surface tension, the pressure at the free surface should equal the air pressure, which can be taken to be the constant p_0 .

For water, the density ρ is essentially constant and we take it to be 1. If the external force is due to gravity, the preceding equations simplify to

$$\begin{aligned} \mathbf{u}_t + (\mathbf{u} \cdot \nabla)\mathbf{u} + \nabla p &= (0, 0, -g)^t; & (7.1) \\ \nabla \cdot \mathbf{u} &= 0, \\ f_t + \mathbf{u} \cdot \nabla f &= 0; \\ p &= p_0 \text{ where } f = 0, \\ \mathbf{u} \cdot \mathbf{n} &= 0 \text{ on the fixed surface.} \end{aligned}$$

Let us simplify still further, to water moving in one horizontal direction over a flat bottom, with units scaled so that the equilibrium depth is one and the gravitational constant is one. Take coordinates (x, y) with x the horizontal and y the vertical position, and velocity vector (u, v) . The free surface is $y = 1 + \eta(x, t)$ and $\tilde{p} = p - (p_0 - y)$ is the difference between the pressure and the undisturbed (hydrostatic) pressure. The equations become Euler's equations

$$\begin{aligned} u_t + uu_x + vv_y &= -\tilde{p}_x; & (7.2) \\ v_t + uv_x + vv_y &= -\tilde{p}_y; \\ u_x + v_y &= 0, \quad 0 < y < 1 + \eta(x, t); \\ v_{y=0} &= 0; \quad \tilde{p}|_{y=1+\eta} = \eta, \quad (\eta_t + \eta_x u - v)|_{y=1+\eta} = 0. \end{aligned}$$

In order to get a reasonable simplification of this system for long waves of small amplitude, we consider a depth scale ε and a length scale $1/\delta$, δ and ε small and positive. Since we want to consider travelling waves, we also scale time by $1/\delta$. To do so we introduce functions \bar{u} , \bar{v} , $\bar{\eta}$, \bar{p} of variables \bar{x} , y, \bar{t} . Considering the amplitude scale and the boundary conditions we take

$$\eta(x, t) = \varepsilon \bar{\eta}(\delta x, \delta t), \quad \tilde{p}(x, y, t) = \varepsilon \bar{p}(\delta x, y, \delta t).$$

In order to preserve the linear parts of the equations above we must take

$$u(x, y, t) = \varepsilon \bar{u}(\delta x, y, \delta t), \quad v(x, y, t) = \varepsilon \delta \bar{v}(\delta x, y, \delta t).$$

Putting these into (7.2), dividing out by common factors, and changing notation once more we obtain

$$\begin{aligned} u_t + \varepsilon(uu_x + vv_y) &= -p_x; & (7.3) \\ \delta^2[v_t + \varepsilon(uv_x + vv_y)] &= -p_y; \\ u_x + v_y &= 0, \quad 0 < y < 1 + \varepsilon\eta(x, t); \\ v_{y=0} &= 0; \quad p|_{y=1+\varepsilon\eta} = \eta, \quad (\eta_t + \varepsilon\eta_x u - v)|_{y=1+\eta} = 0. \end{aligned}$$

Taking the long wave, small amplitude limits $\delta \rightarrow 0$, $\varepsilon \rightarrow 0$ we obtain the linear system

$$\begin{aligned} u_t + p_x &= 0, & p_y &= 0; & u_x + v_y &= 0; \\ v|_{y=0} &= 0, & p|_{y=1} &= \eta; & (\eta_t - v)|_{y=1} &= 0. \end{aligned}$$

It follows that u_t is independent of y and it is consistent to assume that u is independent of y ; then the same is true of v_y . Under these assumptions the boundary conditions imply that

$$p \equiv \eta, \quad v = -u_x y, \quad \eta_t + u_x = 0$$

and the first equation becomes

$$u_t + \eta_x = 0.$$

If we assume that $u(x, t) \rightarrow 0$ as $|x| \rightarrow \infty$ then everything is determined by η , which must satisfy the classical wave equation

$$\eta_{tt} = \eta_{xx}.$$

It follows that η is a sum of a wave moving to the right and one moving to the left:

$$\eta(x, t) = \varphi(x - t) + \psi(x + t).$$

When $\psi = 0$ and $\varphi(x) \rightarrow 0$ as $|x| \rightarrow \infty$ we have $u = \eta$.

Considering unidirectional movement for the full system (7.3), we take $\delta^2 = \varepsilon$ and look for a solution that has an additional time scale to allow for the development of nonlinear effects:

$$\eta(x, t) = \varphi(x - t, \varepsilon t).$$

We follow the convention that the arguments of φ and its derivatives are $(x - t, t)$, so, for example, $\eta_t = -\varphi_x + \varepsilon\varphi_t$. We write

$$u = \varphi + \varepsilon u_1, \quad v = -\varphi_x y + \varepsilon v_1, \quad p = \varphi + \varepsilon p_1,$$

and try to choose φ so that equations (7.3) are satisfied to order ε^2 .

The boundary conditions in(7.3) become, to order ε^2 ,

$$v_1|_{y=1} = 0; \quad p_1|_{y=1} = 0; \quad (\varphi_t + 2\varphi\varphi_x - v_1)|_{y=1} = 0.$$

The second equation in (7.3), to order ε^2 , is $p_{1y} + \varphi_{xx}y = 0$. Combining this with the boundary condition on p gives

$$p_1 = \frac{1}{2}(1 - y^2)\varphi_{xx}.$$

Therefore the first equation becomes, to order ε^2

$$u_{1t} + \varphi_t + \varphi\varphi_x = \frac{1}{2}(y^2 - 1)\varphi_{xxx}.$$

It is reasonable to assume that $u_{1t} + u_{1y} = O(\varepsilon)$, so the third equation gives $u_{1t} = v_{1y} + O(\varepsilon)$ and we have

$$v_{1y} = -\left(\varphi_t + \varphi\varphi_x + \frac{1}{2}\varphi_{xxx}\right) + \frac{y^2}{2}\varphi_{xxx}.$$

Thus v_{1y} is independent of y . Integrating to $y = 1$ and using the boundary condition, we obtain

$$\begin{aligned} v_1|_{y=1} &= -\left(\varphi_t + 2\varphi\varphi_x + \frac{1}{2}\varphi_{xxx}\right) + \frac{1}{6}\varphi_{xxx} \\ &= \varphi_t + 2\varphi\varphi_x, \end{aligned}$$

or

$$\varphi_t + 2\varphi\varphi_x + \frac{1}{6}\varphi_{xxx} = 0,$$

the KdV equation. Thus the equation is of genuine physical interest. Nevertheless, it attracted little attention for 60 years after the publication of the Korteweg–de Vries paper. When it did re-emerge, the context was not water waves and an actual observation, but statistical mechanics and puzzling computer experiments.

Chapter 8

Should one hear white noise – or KdV?

The advent of the digital computer made it possible to test numerically various predictions of physical theory that could not be calculated in closed form. This applies especially to those that are modelled by nonlinear equations. We begin with a familiar problem, the vibrating string. Let us consider a periodic discrete version, with n points on a circle:

$$x_k = e^{2\pi ik/n}, \quad k = 0, \dots, n-1; \quad x_{n+k} = x_k.$$

Let $u_k(t)$ be the displacement from equilibrium at time t . The governing equations for the linear string model are

$$\frac{d^2 u}{dt^2} = (u_{k+1} - u_k) - (u_k - u_{k-1}). \quad (8.1)$$

Any function of the x_k is a linear combination of the power functions, or *modes* $\varphi_m(x) = x^m$, $m = 0, 1, \dots, n-1$. In particular, any solution of (8.1) is a linear combination of the standing wave or standing mode solutions

$$u_m(x_k, t) = \varphi_m(x_k) \cos(\lambda_m t) = x_k^m \cos(\lambda_m t),$$

where the frequencies are given by $\lambda_m = 2 \cos(\pi m/n)$. The *energy* associated to the mode φ_m is proportional to λ_m^2 .

Consider now a nonlinear version of (8.1):

$$\frac{d^2 u}{dt^2} = [f(u_{k+1} - u_k) - f(u_k - u_{k-1})], \quad (8.2)$$

where f is not linear. Any solution can be decomposed uniquely into a time-dependent linear combination of modes:

$$u(x_k, t) = \sum_{m=0}^{n-1} a_m(t) \varphi_m(x_k).$$

However, even if only one mode is involved at time $t = 0$, other coefficients will become nonnegative for $t > 0$. Physically this is thought of as interaction among the various modes, and there are reasons to expect that the interactions will result in “equipartition of energy”: the time average of the energy in each mode will be the same:

$$\lim_{T \rightarrow \infty} \left\{ \frac{1}{T} \int_0^T \lambda_m^2 a_m(t) dt \right\} \text{ independent of } m.$$

This is called “thermalization” – indeed it is this process that makes it possible to introduce the concept of temperature and measure it.

In a famous experiment, Enrico Fermi, John Pasta, and Stanislaw Ulam [30] sought to gain information about how long it takes for thermalization to occur in the model (8.2). They took f to contain quadratic or cubic terms, corresponding to cubic or quartic terms in the potential energy. Then an initial condition involving only one or two modes would be expected over time to involve all modes, with approximately equal average energies.

The result of the computer calculations was quite contrary to these expectations. All though energy did pass significantly among a few different modes, only a few modes were essentially involved at all and the system eventually returned very close to its initial state: it behaved as though it were periodic. (Poincaré’s recurrence theorem is not the issue here: if the system were ergodic the expected time to return so close to its initial state would be many orders of magnitude greater than the observed time.)

Faster computers made it possible for Norman Zabusky and Martin Kruskal to take up the problem in more detail later. If now the function $u(x_k, t)$ is extended to be a smooth function of x , then the evolution of the system is given by an equation of the form

$$u_{tt} = u_{xx} + 2\alpha h u_x u_{xx} + h^2 u_{xxxx} + O(h^4),$$

where h is the spacing between particles and α is a measure of the strength of the nonlinear interaction force. Thus with no nonlinear interaction the

obvious continuum limit $h \rightarrow 0$ is the wave equation $u_{tt} = u_{xx}$, while if the nonlinear force term is increased so that $\alpha h \rightarrow \varepsilon$ in the limit, we get

$$u_{tt} = (1 + \varepsilon u_x)u_{xx}.$$

This is a well-known equation that develops shocks on a time scale of order $1/\varepsilon$, which, for the Fermi–Pasta–Ulam model would have been much shorter than the observed time period. Zabusky suggested keeping the next term:

$$u_{tt} = u_{xx} + 2\alpha u_x u_{xx} + \beta u_{xxxx}.$$

This is the *Boussinesq* equation, derived by Boussinesq [17] in his study of water waves. The long wave length, small amplitude limit is just the wave equation, which whose solutions are superpositions of a wave moving to the right and a wave moving to the left.

Specializing to waves moving to the right leads to the KdV equation, and this is the reason that Kruskal and Zabusky did a number of numerical simulations of initial value problems for KdV. They discovered that, typically, an initial profile broke up into a what appeared to be a superposition of the travelling wave solutions (“solitary waves”) described earlier, whose velocities are proportional to their amplitudes. In the periodic case this meant that a larger wave would eventually catch up with a smaller one – and, after a period of interaction, both waves would emerge in their original forms, with the larger one ahead.

This preservation of identities lead Kruskal and Zabusky to coin the term “soliton” for the solitary wave solutions. It also led them to study a modification of KdV with a cubic nonlinearity, since called the “modified KdV” or “mKdV” equation. Here it is convenient to use the normalizations

$$\begin{aligned} u_t - 6uu_x + u_{xxx} &= 0, \\ v_t - 6v^2v_x + v_{xxx} &= 0 \end{aligned}$$

for KdV and mKdV respectively. Very similar behavior was found for solutions of mKdV, which raised the question whether they were in any sense the “same” equation. There is a history of functional transformations from one PDE to another, e.g. [24], [7]. Typically the transformation takes the form $u = f(x, t, v)v_x + g(x, t, v)$, and indeed Robert Miura [69] found that the map

$$v \rightarrow u = v_x + v^2$$

takes solutions of mKdV to solutions of KdV. It is natural to invert this map. The standard procedure to solve the resulting Riccati equation for v is to take $v = \varphi_x/\varphi$, which leads to the equation

$$\varphi_{xx} = u\varphi$$

Now the KdV is invariant under Galilean transformations: if u is a solution, so is

$$\tilde{u}(x, t) = u(x - ct) - c/6.$$

The corresponding solution of *mKdV* is obtained by solving the Schrödinger eigenvalue problem

$$\varphi_{xx} + \frac{c}{6}\varphi = u\varphi.$$

Once this connection to the Schrödinger eigenvalue problem was made, it was soon discovered that the discrete spectrum is invariant under the KdV flow, while the remaining scattering data evolves as derived above. Therefore, as noted above, the nonlinear KdV equation could be integrated by a linear method: the “method of the inverse scattering transform” [35], [70]. In particular, as noted earlier, for a reflectionless potential the solutions can be obtained quite explicitly as second logarithmic derivatives of determinants.

These special solutions are the soliton and multisoliton solutions, and the analytical formulas confirm the computer experiments: for large negative times the solution is very close to a superposition of solitary wave solutions travelling to the right, with the smaller but slower waves ahead, for large positive times one has the same near-superposition of solitary waves, but with the larger, faster waves ahead. In the periodic version these solutions are almost periodic – an analytic confirmation of the behavior found in the original Fermi–Pasta–Ulam experiment.

For a full historical account of the material in this chapter and the next, see the survey article by Palais [75].

We have discussed here the case of KdV on the line, the case that models John Scott Russell’s original observation. Fermi, Pasta, and Ulam, for reasons of computational convenience, dealt with the *periodic* case. The analogous situation for KdV would be to consider only solutions periodic in x , say $u(x+1, t) = u(x, t)$. This leads naturally to questions of an algebraic–geometric nature. The analogue of an n -soliton solution now comes from a hyperelliptic curve; see [58], [26], [73].

Chapter 9

The soliton revolution

Striking as it was, the discovery of KdV solitons and the inverse scattering method seemed at first to be a mysterious coincidence mathematically as well as an isolated and anomalous example among physical phenomena. The mathematical mystery was solved by the elegant explanation of Lax [57] described above. The physical view changed dramatically in 1972, when Vladimir Zakharov and Alexei Shabat studied the cubic nonlinear Schrödinger equation

$$iq_t + q_{xx} + |q|^2 q = 0, \quad (9.1)$$

which models the propagation of spins in various physical systems, including resonant short optical pulses [91]. Zakharov and Shabat showed that (9.1) could be solved in a way very similar to KdV. In fact consider the system of equations for a pair of functions $\psi_1(x, t)$, $\psi_2(x, t)$:

$$\begin{aligned} \psi_{1x} &= \frac{i\lambda}{2}\psi_1 + q\psi_2, \\ \psi_{2x} &= -\frac{i\lambda}{2}\psi_1 - \bar{q}\psi_2 \end{aligned} \quad (9.2)$$

with time derivatives

$$\begin{aligned} \psi_{1t} &= \frac{i\lambda^2}{2}\psi_1 + \lambda q\psi_2 + a_1\psi_1 + a_2\psi_2, \\ \psi_{2t} &= -\frac{i\lambda^2}{2}\psi_1 - \lambda\bar{q}\psi_2 + b_1\psi_1 + b_2\psi_2. \end{aligned} \quad (9.3)$$

Then (9.1) is the compatibility condition between (9.2) and (9.3): the equality of the mixed partial derivatives $(\psi_j)_{xt} = (\psi_j)_{tx}$. The system (9.2) plays the role here that the equation

$$\varphi_{xx} + \lambda^2\varphi - q\varphi = 0$$

played with respect to the KdV evolution of q , and it has the same type of scattering and inverse scattering theory as the (linear) Schrödinger operator on the line. Once again the appropriate asymptotic data evolve linearly under the flow (9.1). Moreover there are special (soliton) solutions that correspond to purely discrete data and can be calculated algebraically.

The discovery of this second physically interesting equation connected to a linear spectral problem like (9.2) led to an explosive development of the subject: very many new “completely integrable” equations were found in one and two space dimensions, some of them modelling interesting physical problems. Connections were established with differential geometry, infinite-dimensional Lie theory, algebraic geometry, and quantum field theory.

Let us return to the system (9.2), which can be written in matrix form with constant matrix J as

$$\Psi_x = (\lambda J + Q)\Psi, \quad (9.4)$$

while (9.3) is the case $n = 2$, $J_1 = J$ of

$$\Psi_t = B(\lambda)\Psi \equiv (\lambda^n J_1 + \lambda^{n-1} B_{n-1} + \cdots + B_n)\Psi. \quad (9.5)$$

Here J_1 is also constant, and commutes with J . Then the general compatibility condition is the Lax equation

$$Q_t = B(\lambda)_x + [B(\lambda), \lambda J + Q].$$

In the case (9.2),

$$J = \frac{1}{2} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & q \\ -\bar{q} & 0 \end{bmatrix}.$$

Systems like (9.4), (9.5) for 2×2 matrices were examined in detail by Mark Ablowitz, David Kaup, Alan Newell, and Harvey Segur [1], and by Zakharov and Shabat [92]; they are now commonly known as AKNS–ZS systems. These papers develop an inverse theory very similar to the Gel’fand–Levitan–Marčenko methods described above.

The scattering theory for systems like (9.4) can be also be looked at in different way, that is equally applicable to $m \times m$ systems [9]. Consider (9.4) in general form. Suppose that J is a constant diagonal matrix with distinct non-zero entries on the diagonal and that Q is an off-diagonal matrix with entries that, together with their derivatives, converge rapidly to zero as $|x| \rightarrow \infty$. If $Q \equiv 0$ then (9.4) has the fundamental matrix solution

$\exp(x\lambda J)$. It is natural to look for a solution in the general case in the Jost form

$$\Phi(x, \lambda) = M(x, \lambda)e^{x\lambda J}, \quad (9.6)$$

where we normalize by requiring that

$$\lim_{x \rightarrow -\infty} M(x, \lambda) = I, \quad \sup_x \|M(x, \lambda)\| < \infty, \quad (9.7)$$

where we use any convenient matrix norm, such as $\|A\| = \text{trace}(A^*A)^{1/2}$. We assume that $\|Q(\cdot)\|$ has finite L^1 norm.

When $M(\cdot, \lambda)$ exists it satisfies the equation

$$M_x = \lambda[J, M] + QM, \quad (9.8)$$

in which multiplication by J in (9.4) is replaced by the commutator $[J, \cdot]$. One consequence of (9.8) is that $\det M$ is independent of x and therefore is $\equiv 1$. A second consequence is that if $\widetilde{M}(\cdot, \lambda)$ is a second solution of (9.4), (9.5) then

$$(\widetilde{M}^{-1}M)_x = \lambda[J, \widetilde{M}^{-1}M]. \quad (9.9)$$

In particular, the jk element $f = (\widetilde{M}^{-1}M)_{jk}$ satisfies $f_x = \lambda(J_{jj} - J_{kk})f$. Therefore the diagonal entries are constant. If $j \neq k$ and $\lambda(J_{jj} - J_{kk})$ is not purely imaginary, then the boundedness condition (9.7) implies that $f \equiv 0$.

Let us assume J has the form above, with diagonal entries $i/2$ and $-i/2$. It follows that $M(\cdot, \lambda)$ is unique (if it exists) whenever λ is not real. Suppose that $\text{Im } \lambda > 0$. Then, if it exists, $M(\cdot, \lambda)$ is the solution to the matrix integral equation

$$M(x, \lambda) = \begin{bmatrix} 1 + \int_{-\infty}^x [QM]_{11} dy & \int_{-\infty}^x e^{i\lambda(x-y)} [QM]_{12} dy \\ - \int_x^{\infty} e^{-i\lambda(x-y)} [QM]_{21} dy & 1 + \int_{-\infty}^x [QM]_{22} dy \end{bmatrix}, \quad (9.10)$$

where we used the shorthand $[QM]_{ij} = [Q(y)M(y, \lambda)]_{ij}$ for matrix elements. A similar equation holds for $\text{Im } \lambda < 0$, with different limits for the off-diagonal terms.

It is not difficult to show that if

$$\int_{-\infty}^{\infty} \|Q(y)\| dy < 1 \quad (9.11)$$

then (9.10) has a unique solution for each λ in the upper half plane, the corresponding equation has a unique solution when λ is in the lower half

plane. Moreover the solutions are holomorphic as functions of λ and have continuous limits M_+ and M_- as λ approaches the real line from the upper and lower half planes, respectively. Applying the previous remarks to $M_-^{-1}M_+$, we obtain

$$M_+(x, \lambda) = M_-(x, \lambda)e^{-x\lambda J}A(\lambda)e^{x\lambda J}, \quad x \in \mathbf{R}. \quad (9.12)$$

for some matrix-valued function $A(\lambda)$. It follows from (9.12) that $A(\lambda)$ can be determined from the asymptotics of $M_{\pm}(\cdot, \lambda)$. Thus it plays the role here of scattering data.

If we assume also that Q is smooth, it be shown that M has an asymptotic expansion

$$M(x, \lambda) = I + \frac{M_1(x)}{\lambda} + \frac{M_2(x)}{\lambda^2} + \dots,$$

and it follows that $Q = [J, M_1]$. Thus the potential matrix Q can be recovered by solving the *Riemann-Hilbert problem*: given a matrix-valued function A defined on the real line, find for each real x a factorization (9.10) where M_{\pm} are the limiting values of functions holomorphic in the upper and lower half-planes respectively, having limit I as $\lambda \rightarrow \infty$.

If we drop the smallness assumption (9.11) then it remains true that (9.10) has a solution for most non-real values of λ . Indeed the solution is a meromorphic function of λ for λ not real. For “most” choices of Q the solution has finitely many simple poles and has continuous limits M_+ from the upper half-plane and M_- from the lower half-plane. The scattering data must be enlarged to include a matrix associated to each pole, and the Riemann-Hilbert factorization problem must be reformulated: the factors are meromorphic with specified behavior at the poles.

To consider the evolution (9.5), say with $J = J_1$ and $n = 2$ we set

$$M(x, t, \lambda) = \Psi(x, t, \lambda)e^{x\lambda J + t\lambda^2 J}.$$

Arguing as before, we have

$$M_+(x, \lambda) = M_-(x, \lambda)e^{-x\lambda J} [e^{-t\lambda^2 J} A(\lambda) e^{t\lambda^2 J}] e^{x\lambda J},$$

so the scattering data on the line evolves linearly; the same can be shown to be true for the scattering data associated with the poles.

There is an extensive literature; we mention here [9], [74], [31]. As for periodic KdV, there is a rich algebraic-geometric structure connected to these problems; see for example [56], [25].

Chapter 10

Strings revisited

In our earlier discussion of the string density problem we used the Liouville transform to change variables and put the generalized eigenvalue problem (2.1) into the form of an eigenvalue problem (2.2) for an operator with potential function q . This tacitly assumes a certain amount of smoothness of the density function m ; in fact writing q in terms of m requires the second derivative m'' . As we shall see, it is of interest to consider the problem when m is only assumed to be a measure. The inverse problem in this case was studied extensively by Mark Kreĭn [53]–[55] and taken up again in [28],[27]. It will be convenient in discussing it to change the notation slightly from (2.2) so that the generalized eigenvalue problem is

$$D^2\varphi + \lambda\varphi m = 0, \quad D = \frac{d}{dx}. \quad (10.1)$$

We consider solutions $\varphi(\cdot, \lambda)$ on the interval $I = [0, 1]$, normalized by

$$\varphi(0, \lambda) = 0, \quad D\varphi(0, \lambda) = 1. \quad (10.2)$$

We assume that m is a finite positive measure on the interval and that the endpoints have zero mass. Then (10.1), (10.2) can be interpreted as the integral equation

$$\varphi(x, \lambda) = x - \lambda \int_0^x \int_0^{y^-} \varphi(z, \lambda) dm(z) dy. \quad (10.3)$$

For fixed x this is an entire function of λ . The generalized spectrum with respect to Dirichlet boundary conditions $\varphi(0) = 0 = \varphi(1)$ is the finite or countable set $\{\lambda_\nu\}$ of zeros of $\varphi(1, \cdot)$. Set $\varphi_\nu = \varphi(\cdot, \lambda_\nu)$.

If $f = D^2u + \lambda um$ then

$$\begin{aligned} \int_0^1 [Du(x)Dv(x) + f(x)v(x)] dx & \quad (10.4) \\ &= vDu \Big|_0^1 + \lambda \int_0^1 u(x)v(x) dm(x). \end{aligned}$$

This specializes to a collection of identities that prove that the roots of $\varphi(\cdot, \lambda)$ are positive and simple, while the derivatives $D\varphi_\nu$ are orthogonal in $L^2(I, dx)$. In analogy with the situations discussed above, we expect additional information such as $\{a_\nu\}$, $a_\nu = \|D\varphi_\nu\|^{-2}$ to be necessary, in addition to $\{\lambda_\nu\}$, in order to determine the measure m uniquely.

The fact that the data $\{\lambda_\nu\}$, $\{a_\nu\}$ determines the measure m uniquely is a consequence of the work of Kreĭn; see also [12], [13]. The theory lies much deeper than the corresponding theory for the one-dimensional Schrödinger operator sketched above. With a uniqueness result in hand, one way to reconstruct the measure from the data is the following.

If the set $\{\lambda_\nu\}$ is finite, then m is supported on finitely many points and the reconstruction problem is algebraic. Suppose for the moment that for any choice of a pair of n -tuples of positive constants, one can find a corresponding discrete measure. In the countable case, for each n , compute the measure m_n that corresponds to the truncated data

$$\{\lambda_1, \lambda_2, \dots, \lambda_n\}, \quad \{a_1, a_2, \dots, a_n\}.$$

Then the m_n converge to m in the sense that

$$\lim_{n \rightarrow \infty} \int_0^1 u(x) dm_n(x) = \int_0^1 u(x) dm(x)$$

for every continuous function u [14]. Thus the general inverse problem is reduced to the inverse problem for a discrete measure m with support $\{x_j\}_1^n$ and masses $\{m_j\}_1^n$. This can be put into a form that was solved by Thomas Stieltjes in 1894 [82]. We assume that

$$x_0 = 0 < x_1 < x_2 < \dots < x_n < 1 = x_{n+1}$$

and set $l_j = x_{j+1} - x_j$. The problem is to determine $\{l_\nu\}_1^n$ and $\{m_\nu\}_1^n$. Given the wave function $\varphi(\cdot, \lambda)$, define

$$\begin{aligned} q_j(\lambda) &= \varphi(x_j, \lambda), \\ p_j(\lambda) &= D_- \varphi(x_j, \lambda) \equiv \frac{q_j(\lambda) - q_{j-1}(\lambda)}{l_{j-1}}, \quad j = 1, \dots, n+1. \end{aligned}$$

Then the equation (10.1) becomes

$$p_{j+1}(\lambda) - p_j(\lambda) = -\lambda q_j(\lambda) m_j, \quad j = 1, \dots, n.$$

Suppose that the generalized eigenvalues are $0 < \lambda_1 < \dots < \lambda_n$, Set $\lambda_0 = 0$, $a_0 = 1$. We define the Weyl function associated to this problem to be

$$W(\lambda) = \frac{D_- \varphi(1, \lambda)}{\lambda, \varphi(1, \lambda)} = \frac{p_{n+1}(\lambda)}{\lambda q_{n+1}(\lambda)}.$$

It can be shown (using consequences of the identity (10.4)) that W has the partial fractions decomposition

$$W(\lambda) = \sum_{\nu \geq 0} \frac{a_\nu}{\lambda - \lambda_\nu}. \quad (10.5)$$

Thus knowing the data $\{\lambda_\nu\}$, $\{a_\nu\}$ is equivalent to knowing W . Now

$$\begin{aligned} \frac{p_{n+1}}{\lambda q_{n+1}} &= \frac{p_{n+1}}{\lambda(q_n + p_{n+1}l_n)} = \frac{1}{\lambda l_n + \frac{\lambda q_n}{p_{n+1}}} \\ &= \frac{1}{\lambda l_n + \frac{\lambda q_n}{-\lambda q_n m_n + p_n}} = \frac{1}{\lambda l_n + \frac{1}{-m_n + \frac{p_n}{\lambda q_n}}}. \end{aligned} \quad (10.6)$$

Continuing, we can express W as a (finite) continued fraction, with respective coefficients $\lambda l_n, -m_n, \lambda l_{n-1}, -m_{n-1}, \dots$. By relating the asymptotics as $\lambda \rightarrow \infty$ of the partial fractions expansion (10.5) and the continued fractions expansion begun in (10.6), Stieltjes established that

$$l_{n-k} = \frac{(\Delta_k^1)^2}{\Delta_k^0 \Delta_{k+1}^2}, \quad m_{n-k} = \frac{(\Delta_{k+1}^0)^2}{\Delta_k^1 \Delta_{k+1}^1}, \quad (10.7)$$

where $\Delta_0^j = 1$ and Δ_k^i is the determinant of the $k \times k$ submatrix of the Hankel matrix

$$\begin{bmatrix} A_0 & A_1 & A_2 & \dots & A_n \\ A_1 & A_2 & A_3 & \dots & A_{n+1} \\ & & & \ddots & \\ A_n & A_{n+1} & \dots & A_{2n} & \end{bmatrix}$$

starting with 11-entry A_i and the A_j are moments:

$$A_j = \sum_{\nu=0}^n \lambda_\nu^j a_\nu = \int_0^\infty \lambda^j d\rho(\lambda),$$

where ρ is the measure supported on the λ_ν with masses a_ν .

For any positive formal data $\{\lambda_\nu\}_1^n$, $\{a_\nu\}_1^n$, supplemented with $\lambda_0 = 0$, $a_0 = 1$, the various minors Δ_k^0 and Δ_k^1 of the moment matrix (10.8) are positive. Therefore there is a unique measure on the interval with support and masses determined by (10.7).

Chapter 11

Waves with peaks and troughs

The nicely rounded shape of the $1/\cosh^2$ solution of KdV fits the wave profile actually observed by John Scott Russell in 1834, but observed waves often are much less smooth. One might picture instead a wave with elevation

$$u(x, t) = e^{-2|x-t|}, \quad (11.1)$$

travelling to the right at unit speed and with a corner at $x = t$. This function is not a solution of KdV, but it is a weak or generalized solution of another model shallow water equation derived by Robert Camassa and Daryl Holm [20]. This equation,

$$4u_t - 4u_{xxt} + 12uu_x - 2u_xu_{xx} - uu_{xxx} = 0 \quad (11.2)$$

can be written more compactly as the system

$$m = 2u - \frac{1}{2}u_{xx}, \quad m_t + (um)_x + u_xm = 0. \quad (11.3)$$

It is the compatibility condition for the overdetermined system

$$\left[\frac{\partial^2}{\partial x^2} - 1 + \lambda m \right] \varphi = 0, \quad (11.4)$$

$$\left[\frac{\partial}{\partial t} + \left(\frac{1}{\lambda} + u \right) \frac{\partial}{\partial x} - \frac{u_x}{2} \right] \varphi = 0. \quad (11.5)$$

Camassa and Holm found that with the ansatz

$$u(x, t) = \frac{1}{2} \sum_{j=1}^n m_j(t) e^{-2|x-x_j(t)|} \quad (11.6)$$

the CH system (11.3), suitably interpreted, are Hamilton's equations

$$\frac{\partial x_j}{\partial t} = \frac{\partial H}{\partial m_j}, \quad \frac{\partial m_j}{\partial t} = -\frac{\partial H}{\partial x_j},$$

where the Hamiltonian is

$$H(x_1, \dots, x_n, m_1, \dots, m_n) = \frac{1}{2} \sum_{j,k=1}^n m_j m_k e^{-2|x_j - x_k|}.$$

When $n = 1$ the solutions are

$$u_{a,x_0}(x, t) = a e^{-2|x - x_0 - at|}, \quad a \in \mathbf{R}. \quad (11.7)$$

For $a > 0$ this is a peaked wave moving to the right, called a *peakon*. For $a < 0$ it is a depression or trough moving to the left, called an *antipeakon*. Note that in this case $m(\cdot, t)$ is a measure supported on $\{x_j(t)\}$ with masses $\{m_j(t)\}$. In particular, it need not be positive. Note also that, like the KdV case, the speed of the travelling wave (11.7) is proportional to its height or depth.

It can be shown that the general solution (11.8) is asymptotic for $t \rightarrow \pm\infty$ to a superposition

$$u^\pm(x, t) \sim \sum_{j=1}^n u_{a_j, x_j^\pm}(x, t), \quad (11.8)$$

where the phase differences $x_j^+ - x_j^-$ can be calculated. Thus in the long run, individual peakons and antipeakons can be identified.

Again the key to finding explicit solutions of (11.2) on the line is to study the problem (11.4) and the corresponding inverse problem. Consider (11.4) with no t -dependence. If the function m is positive and sufficiently smooth, then again there is a Liouville transformation that converts (11.4) to the eigenvalue problem for a one-dimensional Schrödinger operator $D^2 - q$. As we have seen, any smoothness assumption on m rules out the solutions (11.7), and a positivity assumption rules out antipeakons. Moreover, even with smoothness and positivity, it is difficult to extract desired information on (11.2) from the transformed problem.

Let us consider (11.4) directly, under a decay assumption: $|m(x)| \leq C e^{-4|x|}$. Then there will be solutions

$$\varphi(x, \lambda) \sim e^x, \quad x \rightarrow -\infty; \quad \psi(x, \lambda) \sim e^{-x}, \quad x \rightarrow +\infty.$$

For a discrete set of values $\{\lambda_\nu\}$ these are multiples of each other:

$$\varphi(x, \lambda_\nu) = c_\nu \psi(x, \lambda_\nu).$$

Under the CH flow, it can be seen from (11.4), (11.5) that the λ_ν are fixed, while

$$\frac{dc_\nu}{dt} = \frac{2c_\nu}{\lambda_\nu}. \quad (11.9)$$

Thus the analogue of the KdV solution procedure is to recover m from the data $\{\lambda_\nu\}$, $\{c_\nu\}$.

The simple Liouville transformation $y = (\tanh x + 1)/2$ converts the line to the interval $I = (0, 1)$ and the operator $D^2 - 1 + \lambda m$ to a multiple of the operator $D^2 + \lambda \tilde{m}$. If m is nonnegative then so is \tilde{m} , and thus (11.4) is converted to the string density problem. The $\{\lambda_\nu\}$ are just the generalized eigenvalues for the string density problem, and the coupling constants $\{c_\nu\}$ are related to the remaining scattering data $\{a_\nu\}$ for the string density problem by

$$a_j = c_j \prod_{k \neq j} \left(1 - \frac{\lambda_j}{\lambda_k}\right)^{-1}.$$

Therefore these also evolve by (11.9) under the CH flow. It follows that the multipeakon case (all $m_j > 0$) of (11.8) can be integrated explicitly by using the method of the previous chapter. The same is true of the multi-antipeakon case (all $m_j < 0$) by working with $-u$ in place of u .

Note that the solution of the inverse problem for a discrete measure in Chapter 10 is purely algebraic: the masses $\{\tilde{m}_\nu\}$ and positions $\{l_\nu\}$ are expressed as rational functions of the (positive) data $\{l_\nu\}$ and $\{a_\nu\}$, and conversely. In particular, for generic choice of formal data, *independent of sign*, there will be a unique corresponding signed discrete measure. Letting it evolve according to (11.9) and pulling back to the line by $x = \tanh^{-1}(2y-1)$, we obtain an explicit solution (11.8) valid until one of the denominators in (10.7) becomes zero. When this happens at time t_0 , although two of the m_j blow up like $\pm(t-t_0)^{-1}$, the corresponding difference $l_j = 0((t-t_0)^2)$ and it follows that the solution (11.8) remains well-behaved. Thus the method of the previous chapter gives global explicit expressions for solutions (11.8) for any choice of signs.

The equation (11.2) was first written down in [33], [34] as one of a number of examples of “completely integrable” equations, but was not given a physical interpretation. The physical interpretation is a matter of some discussion; see Johnson [48].

Chapter 12

Can one hear the shape of a drum?

As we noted at the start, the title of this chapter is borrowed from a famous article by Mark Kac [50]. Consider the surface of a drum as a membrane fixed along the boundary of a bounded region Ω in the plane. If $u(x, t)$ is the vertical displacement at time t after the membrane is set in motion, then the (linearized) equation of motion is the wave equation

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} \right) \equiv c^2 \Delta u. \quad (12.1)$$

The assumption that the membrane is fixed at the boundary is expressed by the Dirichlet boundary condition

$$u(x, t) = 0, \quad x \in \partial\Omega. \quad (12.2)$$

Here c is a constant that incorporates the physical properties of the membrane, as well as its tension. As in the case of a string, either the mathematical goal of diagonalizing the operator $c^2 \Delta$ and separating variables, or the physical goal of finding the pure tones that may be produced, leads us to look for standing wave solutions – the (real and imaginary parts of) solutions of the form

$$u(x, t) = u(x) e^{i\omega t}. \quad (12.3)$$

For convenience, we assume that units have been chosen so that $c^2 = 1/2$. In addition to the boundary condition (12.2), the amplitude function u here must be an eigenfunction with eigenvalue $-\omega^2$:

$$-\omega^2 u = \frac{1}{2} \Delta u. \quad (12.4)$$

On physical grounds, and by analogy with the problem in one dimension – the vibrating string – we expect that there should be a discrete sequence

$$0 < \omega_1 \leq \omega_2 \leq \omega_3 \leq \dots, \quad \omega_n \rightarrow \infty$$

for which there are nontrivial solutions of (12.4) that satisfy

$$u(x) = 0, \quad x \in \partial\Omega.$$

This was one of the great problems of 19th century mathematical physics. The existence of such a discrete spectrum was proved early in the 20th century with the help of the then new theory of integral equations.

We can now make precise the question posed as the title of this chapter: given two such regions Ω_1 and Ω_2 , if the corresponding sequences of eigenvalues (for the same constant c^2) are the same, *are Ω_1 and Ω_2 congruent* (up to reflection)? The answer is not obvious. In 1964, two years before Kac’s article, John Milnor [67] had constructed two noncongruent tori whose Laplace-Beltrami operators had the same eigenvalues – but this was in 16 dimensions. As it turns out, the answer is also “no” for some plane domains. In 1992, Carolyn Gordon, David Webb, and Scott Wolpert [41] exhibited two noncongruent polygonal domains – assemblages of five simple pieces – which have the same Dirichlet spectrum and the same Neumann spectrum.

Given the result just mentioned – that even two different spectra may not be enough to identify the domain – it is a bit surprising how much information is already contained in the *asymptotics* of the spectrum and related quantities. We begin with the asymptotics for plane domains cited by Kac.

The English physicist James Jeans used thermodynamic principles to give a heuristic derivation of the asymptotic behavior of the number $N(\lambda)$ of eigenvalues $\leq \lambda$ for standing electromagnetic waves in an enclosure with perfectly reflecting boundary. The challenge of providing a proof of this result was presented to the mathematicians in Göttingen by the Dutch physicist H. A. Lorentz in a lecture in 1910. One of those present at the lecture was Hermann Weyl, who used Hilbert’s theory of integral equations to provide a proof the next year; [85], [86]. Weyl’s result applied to the vibrating membrane gives

$$N(\lambda) \sim \frac{\lambda|\Omega|}{2\pi}, \tag{12.5}$$

where $|\Omega|$ is the area.

Thus (if we can hear infinitely high frequencies!) we can hear the area of a drum. In 1954 Ake Pleijel [76] refined Weyl's result and obtained

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} \sim \frac{|\Omega|}{2\pi t} - \frac{L}{\sqrt{4\pi t}} \quad \text{as } t \rightarrow 0+. \quad (12.6)$$

Here L is the length of the boundary $\partial\Omega$. Thus one can also hear the length of the circumference of the drum. According to the isoperimetric inequality $L^2 \geq 4\pi A$, with equality if and only if Ω is circular. Therefore circular drum is recognizable by its frequencies. This is a close analogue of Ambarzumian's result about the frequencies of a string of uniform density; see Chapter 2.

In his lecture, Kac showed how to prove that for a region bounded by one or more polygons,

$$\sum_{n=1}^{\infty} e^{-\lambda_n t} = \frac{|\Omega|}{2\pi t} - \frac{L}{\sqrt{4\pi t}} + \frac{1-h}{6} + O(\sqrt{t}) \quad (12.7)$$

as $t \rightarrow 0+$, where h is the number of holes. Thus one can hear the connectivity of a polygonal drum: a topological invariant. (Since $1-h$ is an integer, it remains the same if we deform the region Ω continuously.) A year later McKean and Singer [65] proved (12.7) for a general bounded plane region with smooth boundary.

The exponential sum in (12.6), (12.7) may look mysterious at first, but it is a natural quantity for the problem: the trace of the "heat operator."

The heat problem associated with the region Ω is the following. Suppose that the boundary of the region is kept at a fixed temperature, and now let $u(x, t)$ denote the difference between the temperature at $x \in \Omega$ and the boundary temperature. Then an initial temperature distribution evolves according to

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u, \quad (12.8)$$

with Dirichlet boundary condition

$$u(x, t) = 0, \quad x \in \partial\Omega.$$

The same model describes diffusion in a region with an absorbing boundary, which is the guiding principle used by Kac in his discussion of the membrane problem.

To solve the problem, let us diagonalize the operator as before, with normalized eigenfunctions:

$$\frac{1}{2}\Delta\varphi_n = -\lambda_n\varphi_n, \quad (\varphi_n, \varphi_n) = 1,$$

where (\cdot, \cdot) denotes the $L^2(\Omega)$ inner product

$$(u, v) = \int_{\Omega} u(x) v(x) dx$$

for real-valued functions. The φ_n are orthogonal, so a function $u(x, t)$ has an expansion

$$u(x, t) = \sum_{n=1}^{\infty} a_n(t) \varphi_n(x)$$

with

$$a_n(t) = (u(\cdot, t), \varphi_n).$$

Then (12.8) implies $a_n(t) = a_n(0)e^{-\lambda_n t}$. It follows that the solution to (12.8) with initial value $u(x, 0) = f(x)$ is

$$u(x, t) = \int_{\Omega} P_t(x, y) f(y) dy$$

where the *heat kernel* P_t is

$$P_t(x, y) = \sum_{n=1}^{\infty} e^{-\lambda_n t} \varphi_n(x) \varphi_n(y). \quad (12.9)$$

In particular, its *trace*, in analogy with the trace of a matrix, is

$$\int_{\Omega} P_t(x, x) dx = \int_{\Omega} \sum_{n=1}^{\infty} e^{-\lambda_n t} \varphi_n(x)^2 dx = \sum_{n=1}^{\infty} e^{-\lambda_n t}. \quad (12.10)$$

Thus one way to extract geometric information is to examine the small-time asymptotics of the heat kernel and its trace.

As we indicate later, one can get significant information about the heat kernel in any number of dimensions without any calculation of the actual spectrum. (On the other hand, any information about the spectrum is highly encoded in the heat trace.) By “significant information” we mean a sequence of approximations to the heat kernel that are sufficient for expansions like (12.8). This process will be described in the next three chapters, starting with approximating the inverse of an operator like the Laplacian Δ itself.

We return here to the integer

$$\chi(\Omega) = 1 - h.$$

For the polygonal domains considered by Kac it is the *Euler characteristic*

$$\chi(\Omega) = V - E + F \quad (12.11)$$

where V , F , and E are the number of vertices, faces, and edges in any decomposition of Ω into triangles with disjoint interiors. For a connected, simply connected polyhedral surface in three dimensions, $V - E + F = 2$, a result that may have been known to Descartes and that was proved for convex polyhedra by Euler in 1752. If S is a smooth closed surface, drawing triangles on it gives a determination of $\chi(S)$. Gauss's "Theorema Egregia" gives $\chi(S)$ in terms of a local geometric invariant K , the Gauss curvature:

$$\chi(S) = \frac{1}{2\pi} \iint_S K(x) dV(x) \quad (12.12)$$

where dV is the Riemannian volume element. The identity (12.12) is known as the Gauss–Bonnet formula.

Kac ended his article [50] with the comment

One can, of course, speculate on whether in general one can hear the Euler-Poincaré characteristic and raise all sorts of interesting questions.

The Euler-Poincaré characteristic of an n -dimensional manifold M is Poincaré's generalization of (12.11):

$$\chi(M) = b_0 - b_1 + b_2 - \cdots + (-1)^n b_n, \quad (12.13)$$

where b_k is the number of k -dimensional simplexes in a decomposition of M ; it can also be taken to be the k -th Betti number of the simplicial complex, which is itself a topological invariant. It is a consequence of Poincaré duality that $\chi(M) = 0$ if the dimension n is odd (a result that we return to in Chapter 14). For even n , Shiing-Shen Chern [22] proved the following generalization of the Gauss-Bonnet formula (12.12) to a Riemannian manifold in 1944:

$$\chi(M) = \frac{1}{2\pi} \int_M C(x) dV(x), \quad (12.14)$$

where V is the volume form and C is a homogeneous polynomial of degree $\frac{n}{2}$ in the components of the curvature tensor, one of the Chern classes.

To summarize a bit: at this point we have several analytic formulas for geometric and topological invariants, involving manifolds with boundary, in

the case of (12.8), or without in the case of (12.12) and (12.14). In one case a Laplacian, indeed a heat operator, plays a role, in the other cases it does not seem to. In the remaining chapters we sketch the way in which these matters can be tied together.

The 1948 Gibbs Lecture by Hermann Weyl [88] contains a discussion of various approaches to the eigenvalue distribution question in an historical context.

Chapter 13

Inverting an elliptic operator

The most widely occurring type of elliptic operator in n variables is a second order operator with an expression in local coordinates of the form

$$P = p\left(x, -i\frac{\partial}{\partial x}\right) = \sum_{j,k=1}^n a_{jk}(x)\frac{\partial^2}{\partial x_j\partial x_k} + \sum_{j=1}^n b_j(x)\frac{\partial}{\partial x_j} + c(x), \quad (13.1)$$

where for each x , $p(x, \cdot)$ is the polynomial

$$\begin{aligned} p(x, \xi) &= \sum_{j,k=1}^n a_{jk}(x)(-i\xi_j)(-i\xi_k) + \sum_{j=1}^n b_j(x)(-i\xi_j) + c(x), \\ &= -\sum_{j,k=1}^n a_{jk}(x)\xi_j\xi_k - i\sum_{j=1}^n b_j(x)\xi_j + c(x), \end{aligned}$$

where we may assume that $a_{jk} = a_{kj}$. The mysterious-looking choice of $-i$ here will turn out to be convenient; it is connected with the fact that

$$p\left(x, -i\frac{\partial}{\partial x}\right)\{e^{ix\cdot\xi}\} = p(x, \xi)e^{ix\cdot\xi}.$$

“Ellipticity” here is the condition that for each x , the top-order part has no real zeros:

$$p_2(x, \xi) \equiv -\sum_{j,k=1}^n a_{jk}\xi_j\xi_k \neq 0, \quad \xi \in \mathbf{R}^n \setminus \{0\}.$$

For example, if M is a compact Riemannian manifold, then the associated Laplace–Beltrami operator is elliptic. Locally it can be taken to have the

form

$$P = \frac{1}{2} \sum_{j=1}^n X_j^2 + \text{lower order}$$

where the vector fields

$$X_j = \sum_{k=1}^n a_{jk} \frac{\partial}{\partial x_k}$$

are, locally, an orthonormal frame for the tangent bundle of M .

How would one go about inverting such an operator? Suppose first that we complicate matters by taking a noncompact manifold, $M = \mathbf{R}^n$, but then simplify by assuming the operator P has constant coefficients and no lower order terms. The associated *characteristic polynomial* or *symbol* is then the quadratic

$$p(\xi) = \sum_{j,k=1}^n a_{jk} \xi_j \xi_k$$

If $Pu = f$, then taking the Fourier transform

$$\widehat{u}(\xi) = \int e^{-ix \cdot \xi} u(x) dx,$$

whose inverse is given by

$$u(x) = \frac{1}{(2\pi)^n} \int e^{ix \cdot \xi} \widehat{u}(\xi) d\xi$$

leads to

$$p(\xi) \widehat{u}(\xi) = \widehat{f}(\xi).$$

In turn, taking the inverse Fourier transform gives

$$u(x) = \frac{1}{(2\pi)^n} \int e^{ix \cdot \xi} \frac{1}{p(\xi)} \widehat{f}(\xi) d\xi.$$

In the case of an operator P with variable coefficients, we could take as first approximation to a solution of $Pu = f$ the function

$$\int e^{ix \cdot \xi} q(x, \xi) \widehat{f}(\xi) d\xi, \quad (13.2)$$

where $q(x, \xi) = 1/p(x, \xi)$. This is called a *pseudodifferential operator* with *symbol* $q = q(x, \xi)$.

It takes some time and space to give a proper development of the calculus of pseudodifferential operators; a good starting point is the expository article [71]. Here we simply make a number of remarks.

First, there is the technical problem that $p(\xi)$ vanishes for $\xi = 0$, and $p(x, \xi)$ will also vanish for some values of ξ . On the other hand, the inverse Fourier transform of a function with compact support is a smooth function, so we can modify the symbol near the zeroes if we are willing to make smooth errors (which we are).

Second, the approximation given by (13.2) can be improved, in the sense the error will be smoother, if we systematically modify the symbol. Note that

$$p\left(x, \frac{\partial}{\partial x}\right) = \sum_{j=0}^2 p_j\left(x, \frac{\partial}{\partial x}\right), \quad (13.3)$$

with $p_j(x, \xi)$ homogeneous of degree j in ξ . We look for the approximate inverse to have a symbol with an asymptotic expansion

$$q(x, \xi) \sim \sum_{m=2}^{\infty} q_{-m}(x, \xi), \quad |\xi| \rightarrow \infty, \quad (13.4)$$

where q_{-m} is homogeneous in ξ of degree $-m$.

Third, calculating the composition of differential operators suggests the following composition rule for a pair of pseudodifferential operators P and Q with symbols p and q : the symbol of PQ should have an expansion

$$\sum_{\alpha} \frac{1}{\alpha!} \left[\left(\frac{\partial}{\partial \xi} \right)^{\alpha} p \right] \left(-i \frac{\partial}{\partial x} \right)^{\alpha} q.$$

Here $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$ is an n -tuple of nonnegative integers and

$$\xi^{\alpha} = \xi_1^{\alpha_1} \xi_2^{\alpha_2} \cdots \xi_n^{\alpha_n}, \quad \alpha! = \alpha_1! \alpha_2! \cdots \alpha_n!.$$

It follows that in our expansion (13.4) of the symbol for the inverse of (13.3) we want the first two terms to be

$$\begin{aligned} q_{-2} &= \frac{1}{p_2}, \\ q_{-3} &= -\frac{1}{p_2} \left[p_1 q_{-2} - i \sum_{j=1}^n \frac{\partial p_2}{\partial \xi_j} \frac{\partial q_2}{\partial x_j} \right]. \end{aligned}$$

In general we want

$$q_{-r} = -\frac{1}{p_2} \left[\sum_{s < r} \sum_{s+|\alpha|-t=-r} \frac{1}{\alpha!} \left[\left(\frac{\partial}{\partial \xi} \right)^\alpha p_t \right] \left(-i \frac{\partial}{\partial x} \right)^\alpha q_{-s} \right],$$

where $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$. It can be shown that there is a symbol q (not unique) with expansion (13.4) and that the corresponding operator Q is an inverse for P modulo smoothing operators:

$$PQ = I + R_1, \quad QP = I + R_2 \quad (13.5)$$

where the operators R_j map distributions to smooth functions. (We are tacitly assuming that the coefficients of P are smooth functions.) The first of these equations follow directly from the construction, since we constructed Q to be essentially a *right* inverse. To obtain the second we modify the construction in the obvious way to obtain an approximate *left* inverse Q_1 , so that

$$Q_1 P = I + R,$$

with R smoothing. Then

$$\begin{aligned} Q &= (I + R)Q - RQ = Q_1 P Q - RQ \\ &= Q_1 (I + R_1) - RQ = Q_1 + [Q_1 R_1 - RQ]. \end{aligned}$$

The term in brackets is a smoothing operator, so Q_1 differs from Q by a smoothing operator and can be replaced by Q .

A final remark: we really wanted to work on a general manifold, not on \mathbf{R}^n . However, the commutator of a pseudodifferential operator with multiplication by a smooth function is an operator of lower order, so by using a smooth partition of unity one can work locally and then glue things together on the manifold.

Some history: the Hilbert transform, mentioned in Chapter 3, is the simplest example of a singular integral operator

$$Tu(x) = \int K(x, x-y) u(y) dy.$$

From our current perspective, such operators are pseudodifferential operator of order zero. In 1936 S. G. Mikhlin [66] worked out the composition of singular integral operators in the plane and defined their multiplicative symbol. Georges Giraud [40] carried this out in higher dimensions. The nature of the symbol – as the partial Fourier transform of the kernel K of the

singular integral operator – was clarified by the work of Alberto Calderón and Antoni Zygmund [19].

The theory of pseudodifferential operators is a refinement of the algebra of singular integral operators introduced by Calderón and Zygmund. It was developed in 1965 independently by Joseph Kohn and Louis Nirenberg [51], Robert Seeley [80] and André Unterberger and Juliane Bokobza [83]; see the survey article by Seeley [81].

With more effort, these procedures can be adapted to a class of subelliptic operators; see [10].

Chapter 14

Indices and index theorems

Continuing the discussion of the previous chapter, suppose that P is an elliptic operator on a compact manifold M without boundary, and let Q be an approximate inverse as constructed above. It will sometimes be convenient to use Hilbert space concepts here by choosing a measure on M (with smooth density) and working with the corresponding L^2 space. We take the domain of P as an operator in $L^2(M)$ to be

$$\{u \in L^2(M) : Pu \in L^2(M)\},$$

where Pu is taken in the sense of distributions. Equivalently, this realization of P is the closure of the restriction of P to smooth functions.

Recall (13.5):

$$PQ = I + R_1, \quad QP = I + R_2.$$

Compactness of the manifold implies that the remainder terms R_j are compact as operators acting in $L^2(M)$. It follows from this that $I + R_j$ has a finite-dimensional kernel (nullspace) and closed range with finite codimension. This immediately implies that P has finite-dimensional nullspace $N(P)$ and closed range $\text{ran}(P)$ with finite codimension. Therefore the *index*

$$\text{index } P = \dim \ker P - \text{codim } \text{ran } P \tag{14.1}$$

is well-defined. It is, of course, an integer and therefore invariant under continuous perturbations (of the coefficients of the operator, so long as ellipticity is preserved, and also of the manifold itself). In some cases it is very closely related to our earlier integer-valued invariants, the Euler–Poincaré characteristics (12.11) and (12.13), though we shall need to extend our notion of ellipticity to systems of operators.

Given a smooth compact manifold without boundary, M , we consider the DeRham complex: it consists of the spaces $\mathcal{E}^{(k)}$ of smooth k -forms on M , together with the exterior derivative d . Its restriction to k -forms is denoted by

$$d_k : \mathcal{E}^{(k)} \rightarrow \mathcal{E}^{(k+1)}.$$

If we equip M with a smooth measure, we can introduce inner products in the spaces $\mathcal{E}^{(k)}$ and then define an adjoint operation d^* ,

$$d_k^* : \mathcal{E}^{(k+1)} \rightarrow \mathcal{E}^{(k)}.$$

The first-order system $d + d^*$ is selfadjoint (in the L^2 completion of the space of smooth forms). Because $d^2 = 0$, the second-order system $\Delta = (d + d^*)^2$ maps k -forms to k -forms:

$$\Delta_k = d_k^* d_k + d_{k-1} d_{k-1}^*.$$

Both $d + d^*$ and Δ are elliptic systems, which means that in local coordinates the matrix $P(\xi)$ of symbols of the highest order terms is invertible – the analogue for systems of the ellipticity condition for real second order operators. The method of the previous chapter can be adapted easily to obtain approximate inverses for these systems.

It follows that each system has an index – which turns out not to be very interesting for these operators. In fact if A is any selfadjoint operator then the nullspace is the orthogonal complement of the range, so if A has an index, the index is zero.

To make life more interesting, we suppose that n is even and take D to be the restriction of $d + d^*$ to forms of even degree. These are mapped to forms of odd degree, and D^* maps forms of odd degree to forms of even degree. The kernel of D is the same as the kernel of

$$D^* D = \Delta_0 + \Delta_2 + \cdots + \Delta_n,$$

while the codimension of the range is the dimension of the kernel of D^* , which is the same as the kernel of

$$D D^* = \Delta_1 + \Delta_3 + \cdots + \Delta_{n-1}.$$

According to the work of W. V. D. Hodge in the 1930s, the dimension of the null space of Δ_k (the space of harmonic k -forms) as an operator in the associated L^2 space is the dimension of the k -th (real) cohomology group and therefore, by Poincaré duality, equal to the Betti number b_{n-k} ; [47],

but see also [87]. This means that the index of D is the Euler-Poincaré characteristic (12.13):

$$\begin{aligned} \text{Index}(D) &= (b_n + b_{n-2} + \cdots + b_0) && (14.2) \\ &\quad - (b_{n-1} + b_{n-3} + \cdots + b_1) \\ &= \chi(M). \end{aligned}$$

Without going into detail, we note that a complex analytic manifold M carries a more refined version of the structure above, in that the exterior derivative d decomposes as the sum $\partial + \bar{\partial}$, where $\bar{\partial}$ annihilates the holomorphic forms. There is a corresponding complex – actually a family of complexes – the Dolbeault complex. The index of this complex that is analogous to that of the DeRham complex, (14.2), was calculated by Friedrich Hirzebruch in 1954 [45], [46]. It is the general form of the Riemann-Roch theorem for Riemann surfaces.

Every elliptic complex has an associated index. Since it is invariant under continuous deformations, it depends only on the principal symbol (highest order terms) of the elliptic system and, one might expect, on topological invariants of the manifold. The problem of determining the index from this data was formulated by Gel'fand in 1960 [36]. The definitive solution was obtained by Michael Atiyah and Isidore Singer [5] in the early 1960s: the previous index, the *analytic index* of an elliptic complex, is equal to the *topological index*, an expression involving product of the Chern character of a bundle associated to the symbol of the complex and a Todd class. For a brief outline, see [46].

The original proofs of the Atiyah Singer theorem used functoriality properties of both mappings and showed that these properties characterized the invariant uniquely. An important feature on the analytic side was the ability to work with elliptic complexes associated to homotopy deformations of the symbol, which is only possible if one goes beyond the class of differential operators to that of pseudodifferential operators.

The topological index is a generalization of the left side of the identities (12.12) or (12.14). The corresponding generalization of the right side of these identities, at least in the case of a Riemannian manifold, would be an integral over the manifold of an expression in local geometric invariants and derivatives of the principal symbol. One way to arrive at such an analytic-geometric expression was proposed by Atiyah and Bott [4]. Suppose D is an elliptic complex, realized as a closed operator in an appropriate L^2 space. The operators D^*D and DD^* are selfadjoint. Restricting them to

the subspaces $(\ker D)^\perp$ and $(\ker D^*)^\perp$ respectively, they are invertible. The operator

$$U = D(D^*D)^{1/2} : (\ker D)^\perp \rightarrow (\ker D^*)^\perp$$

is unitary, and

$$U^*DD^*U = D^*D.$$

Therefore, on the complements of their kernels, the elliptic operators D^*D and DD^* are unitarily equivalent. In particular, they have the same non-zero eigenvalues $\{\lambda_n\}$. It follows that the traces of the heat kernels for $-D^*D$ and $-DD^*$ are

$$\begin{aligned} \text{trace } e^{-tD^*D} &= \dim \ker D^*D + \sum_{n=1}^{\infty} e^{-t\lambda_n}; \\ \text{trace } e^{-tDD^*} &= \dim \ker DD^* + \sum_{n=1}^{\infty} e^{-t\lambda_n}, \end{aligned} \quad (14.3)$$

and

$$\text{index } D = \text{trace } e^{-tD^*D} - \text{trace } e^{-tDD^*}. \quad (14.4)$$

Minakshisundaram and Pleijel [68] showed in 1949 that the trace of the heat operator for the Laplace–Beltrami operator on a compact manifold has an asymptotic expansion

$$\text{trace } e^{-t\Delta} \sim t^{-n/2} \sum_{m=0}^{\infty} c_m t^m, \quad \text{as } t \rightarrow 0+, \quad (14.5)$$

where n is the dimension of the manifold. Expansions like this are true for general elliptic operators. It follows from this that if D is a first order elliptic system, then the c_m coefficients in the expansions of the two traces (14.3) are the same for $m < n/2$. Moreover, the index is zero if n is odd, and is the difference of the $c_{n/2}$ coefficients if n is even. In particular, this gives what has been called the “analyst’s proof” that the Euler–Poincaré characteristic of an odd–dimensional manifold is zero.

In the final chapter we outline a derivation of expansions like (14.5) for general heat kernels and traces.

Some history: The study of the index seems to have originated in work on a singular integral equation on a line by Fritz Noether [72], who had discovered a point overlooked by Poincaré in studying tidal motion. It was further investigated in the context of singular integral operators by F. V. Atkinson [3] in 1951, and in more abstract form by a number of authors.

Gel'fand's 1960 article drew much attention to the index problem in the context of elliptic equations. This led to a number of investigations by analysts and topologists that culminated in the work of Atiyah and Singer. For further references, see the survey article by Seeley [81].

Chapter 15

The index and the heat operator

One way to obtain the expansion of the heat trace of a general elliptic operator is to redo the pseudodifferential calculus of Chapter 13 for heat operators. We sketch the procedure here, assuming for convenience that the elliptic operator P is of second order and is scalar, rather than a system. As in Chapter 13 we start with an operator with constant coefficients, and assume that the principal symbol p_2 is a negative definite quadratic form (as is the case for the Laplacian). To solve

$$\frac{\partial u}{\partial t} - Pu = f$$

we take the Fourier transform in all variables,

$$\widehat{u}(\xi, \tau) = \int_{\mathbf{R}^n \times \mathbf{R}} e^{-ix \cdot \xi - i\tau t} u(x, t) dx dt$$

to obtain

$$[i\tau - p(\xi)] \widehat{u}(\xi, \tau) = \widehat{f}(\xi, \tau).$$

Passing to the variable coefficient version in $\mathbf{R}^n \times \mathbf{R}$, it is natural to begin the process with the symbol

$$q_{-2}(x, \xi, \tau) = [i\tau - p_2(x, \xi)]^{-1},$$

which is homogeneous of degree -2 with respect to the weighted dilations $(\xi, \tau) \rightarrow (\lambda\xi, \lambda^2\tau)$, and to compute further terms by the method of Chapter 13. This will produce an approximate inverse whose symbol q has an

asymptotic expansion

$$q(x, \xi, \tau) \sim \sum_{m=2}^{\infty} q_{-m}(x, \xi, \tau),$$

where each term q_{-m} is a linear combination of terms of the form

$$\frac{r_{m,m-k}(x, \xi)}{[i\tau - p_2(x, \xi)]^k}, \quad \frac{m}{2} \leq k < m, \quad (15.1)$$

where $r_{m,m-k}(x, \cdot)$ is a polynomial of degree $2k - m$.

Formally, the pseudodifferential operator $q(x, -i\partial/\partial x, -i\partial/\partial t)$ is an integral operator with kernel

$$\begin{aligned} K(x, t; y, s) &= k(x, y, t - s) \\ &= \frac{1}{(2\pi)^{n+1}} \int_{\mathbf{R}^n \times \mathbf{R}} e^{i(x-y)\cdot\xi + i(t-s)\tau} q(x, \xi, \tau) d\xi d\tau. \end{aligned}$$

The kernel of e^{-tP} is then $k_t(x, y) = k(x, y, t)$. Corresponding to the expansion of q is an expansion

$$\begin{aligned} k_t(x, y) &\sim \sum_{m=2}^{\infty} k_{t,-m}(x, y), \\ k_{t,-m}(x, y) &= \frac{1}{(2\pi)^{n+1}} \int_{\mathbf{R}^n \times \mathbf{R}} e^{i(x-y)\cdot\xi + it\tau} q_{-m}(x, \xi, \tau) d\xi d\tau. \end{aligned}$$

The form (15.1) shows that each q_{-m} extends to a function holomorphic in τ for $\text{Im } \tau < 0$, and homogeneous of degree $-m$. It follows that the inverse transform satisfies

$$k_{t,-m} = 0, \quad t < 0;$$

$$k_{t,-m}(x, t; x, 0) = t^{1+(m-n)/2} k_{t,-m}(x, 1; x, 0), \quad t > 0.$$

Taking $\lambda = -1$, the homogeneity condition implies that q_{-m} is an odd function of ξ when m is odd, which implies $k_{t,-m} = 0$ when m is odd. Thus we obtain the expansion

$$k_{t,-m}(x, x) \sim t^{-n/2} \sum_{m=0}^{\infty} c_m(x) t^m. \quad (15.2)$$

Again, these considerations extend from the case $\mathbf{R}^n \times \mathbf{R}$ to the case $M \times \mathbf{R}$, where M is a compact manifold. If k_t were the *exact* heat kernel, i.e. the

kernel of the integral operator e^{-tP} for $t > 0$, then we could integrate this expansion and obtain (14.5). But in fact the expansion (15.2) does give (14.5). This follows from the fact that for any given integer N , if N_1 is large enough then

$$P \sum_{m=2}^{N_1} Q_m = I - R$$

where R is an integral operator with continuous kernel $r(x, y, t)$ that satisfies

$$|r(x, y, t)| \leq C t^N, \quad 0 \leq t \leq 1.$$

Let

$$|r|_t = \sup_{0 \leq s \leq t} \sup_{x \in M} |r(x, y, s)|.$$

The kernel of R^2 is

$$r^{(2)}(x, y, t) = \int_0^t \int_M r(x, z, t-s) r(z, y, s) dz ds.$$

and for $0 \leq t \leq 1$,

$$|r^{(2)}|_t \leq C^2 |M|^2 \int_0^t (t-s) s^N ds = \frac{C^2 |M|^2 t^{N+2}}{(N+2)(N+1)},$$

where $|M|$ is the volume of M . Inductively, the kernel of R^k has norm

$$|r^{(k)}|_t \leq \frac{N! C^k |M|^k t^{N+2k}}{(N+2k)!}.$$

Therefore the kernels of the partial sums of the Neumann series $\sum_{k=0}^{\infty} R^k$ converge uniformly to the kernel of $(1 - R)^{-1}$ and give the heat kernel

$$\left[\sum_{m=2}^N Q_{-m} \right] (I - R)^{-1}.$$

Combining the previous results, we find that the difference between the kernel of e^{-tP} and

$$t^{-n/2} \sum_{m=2}^{N_1} c_m(x) t^m$$

is $O(t^N)$ as $t \rightarrow 0_+$.

For the details of the construction for P of any order, see [43]. Again, an analogous procedure can be carried out for certain subelliptic P : [11].

Let us return to questions like those in Chapter 12, which involve a manifold with boundary as well as the heat operator associated with an elliptic operator. Suppose the compact manifold M has a smooth boundary ∂M . Given a second order elliptic operator P on M , we may find a function ρ defined near the boundary such that $d\rho \neq 0$ on ∂M and such that in local coordinates (ρ, x) the principal part of P has the form

$$P = \frac{\partial^2}{\partial \rho^2} + P_{tan}, \quad (15.3)$$

where P_{tan} involves only differentiation in the tangential directions x . In fact we may replace P by the similar operator $e^{-f} P e^f$ obtained by conjugation with multiplication by a nonvanishing smooth function e^f , with f chosen so that (15.3) is true for the full operator (not just the principal part).

The form (15.3) allows one to adapt the procedure of Chapter 13 to construct approximate inverses for boundary value problems associated to P . Specifically we consider the problem

$$Pu = f \quad \text{on } M, \quad (15.4)$$

with either Dirichlet boundary condition

$$u = 0 \quad \text{on } \partial M \quad (15.5)$$

or Neumann boundary condition

$$\frac{\partial u}{\partial \rho} = 0 \quad \text{on } \partial M. \quad (15.6)$$

Let M_1 be a manifold without boundary that contains M , in such a way that ∂M is a smooth submanifold. Extend P to an elliptic operator on M_1 , and let Q be an approximate inverse on M_1 . Given a function f on M , we extend it to M_1 and look for an approximate solution to (15.4) of the form

$$u = Qf + w$$

where w is a solution of

$$Pw = 0 \quad \text{on } M \quad (15.7)$$

with a boundary condition fitted to (15.5) or (15.6). Formally, because of (15.3), the solution of (15.7) should have the form

$$v = \exp\left(-\rho\sqrt{-P_{tan}}\right) g$$

for some function g on ∂M . Thus we want

$$g = -u \quad \text{on } \partial M$$

in the Dirichlet case and

$$g = -(-P_{tan})^{-1/2}u \quad \text{on } \partial M$$

in the Neumann case. Altogether, then, the approximate solution is the restriction to M of

$$[I - \exp(-\rho(-P_{tan})^{1/2})J]Q$$

in the Dirichlet case and

$$[I - \exp(-\rho(-P_{tan})^{1/2})(-P_{tan})^{-1/2}J]Q$$

in the Neumann case, where J is the operator of restriction to the boundary. Approximations to the operators

$$(-P_{tan})^{-1/2}, \quad e^{-\rho(-P_{tan})^{1/2}}$$

can be constructed by the methods in Chapter 13 and this chapter.

For details of this approach and its application to the drum problem, see [43]. The general index theorem has been extended to manifolds with boundary; see [6]. For an extension to certain subelliptic cases, see [11].

In cases where P is defined geometrically, e.g. as the Laplace–Beltrami operator for M , an examination of the construction of the approximate heat kernel shows that the functions $c_m(x)$ above are built from combinations of geometrically defined quantities such as curvature and its covariant derivatives. One can then use invariant theory to narrow down or identify the c_m exactly; see, for example, [38] and [11]. This gives, finally, the generalization of the right side of the identities (12.12) and (12.14).

For the current state of affairs in heat kernel asymptotics in elliptic boundary value problems, see [39].

Bibliography

- [1] M. Ablowitz, D. J. Kaup, A. Newell, and H. Segur. *The inverse scattering transformation – Fourier analysis for nonlinear problems*. Studies Appl. Math. **53** (1974), 249-315.
- [2] V. Ambarzumian. *Über eine Frage der Eigenwerttheorie*. Zeit. Phys. **53** (1929), 690-695.
- [3] F. V. Atkinson. *The normal solubility of linear equations in normed spaces*. Mat. Sb. N.S. **28** (1951), 3-14.
- [4] M. F. Atiyah. *Global aspects of the theory of elliptic differential operators*. Proc. Intern. Congress of Mathematicians, Moscow 1966, pp. 7-14.
- [5] M. F. Atiyah and I. M. Singer. *The index of elliptic operators on compact manifolds*. Bull. Amer. Math. Soc. **69** (1963), 422-433.
- [6] M. F. Atiyah and R. Bott. *The index problem for manifolds with boundary*. Differential Analysis, Bombay Colloq. 1964. Oxford University Press, London.
- [7] A. V. Bäcklund. *Zur Theorie der Flächentransformationen*. Math. Ann. **19** (1882), 387-422.
- [8] V. Bargmann. *On the connection between phase shifts and scattering potential*. Rev. Modern Phys. **21** (1949), 488-493.
- [9] R. Beals and R. R. Coifman. *Scattering and inverse scattering for first order systems*. Comm. Pure Appl. Math. **37** (1984), 39-90.
- [10] R. Beals and P. C. Greiner. *Calculus on Heisenberg Manifolds*. Annals of Math. Studies 119. Princeton University Press, Princeton 1988.

- [11] R. Beals, P. C. Greiner, and N. K. Stanton. *The heat equation on a CR manifold*. J. Diff. Geom. **20** (1984), 343-387.
- [12] R. Beals, D. H. Sattinger, and J. Szmigielski. *Multipeakons and the classical moment problem*. Adv. Math. **154** (2000), 229-257.
- [13] R. Beals, D. H. Sattinger, and J. Szmigielski. *Peakons, strings, and the finite Toda lattice*. Comm. Pure Appl. Math. **54** (2001), 91-106.
- [14] R. Beals, D. H. Sattinger, and J. Szmigielski. *The string problem and the Camassa–Holm equation*. Phil. Trans. Royal Soc. A **365** (2007), 2299-2312.
- [15] G. Borg. *Eine Umkehrung der Sturm–Liouvilleschen Eigenweraufgabe. Bestimmung der Differentialgleichungen durch die Eigenwerte*. Acta Math. **78** (1946), 1-96.
- [16] U. Bottazzini. *The Higher Calculus: a History of Real and Complex Analysis from Euler to Weierstrass*. Springer-Verlag, New York 1986.
- [17] J. de Boussinesq. *Théorie de l'intumescence liquide appelée onde solitaire ou de translation se propageant dans un canal rectangulaire*. C. R. Acad. Sci. Paris **72** (1871), 755-759.
- [18] R. K. Bullough and P. J. Caudrey, eds. *Solitons. Topics in Current Physics 17*, Springer, Berlin 1980.
- [19] A. P. Calderón and A. Zygmund. *Algebras of certain singular integral operators* Amer. J. Math. **78** (1956), 310-320.
- [20] R. Camassa and D. D. Holm. *An integrable shallow water equation with peaked solitons*. Phys. Rev. Letters **71** (1993), 1661-1664.
- [21] K. Chadan and P. C. Sabatier. *Inverse Problems in Quantum Scattering Theory*. Springer-Verlag, New York, 1989.
- [22] S. S. Chern. *A simple intrinsic proof of the Gauss–Bonnet formula for closed Riemannian manifolds*. Ann. Math. **45** (1944), 747-752.
- [23] A. D. D. Craik. *The origins of water wave theory*. Annu. Rev. Fluid Mech. **36** (2004), 1-28.
- [24] G. Darboux. *Leçons sur la Théorie Générale des Surfaces et les Applications Géométriques du Calcul Infinitésimal*, vol. 2, 2nd ed. Gauthier–Villars, Paris 1915.

- [25] B. A. Dubrovin. *Theta-functions and non-linear equations*. Uspekhi Mat. Nauk **36** (1981), 121-184.
- [26] B. A. Dubrovin, V. B. Matveev, and S. P. Novikov. *Non-linear equations of the Korteweg–de Vries type, finite gap operators, and Abelian varieties*. Uspekhi Mat. Nauk **31** (1976), 55-136.
- [27] H. Dym and N. Kravitsky. *On the inverse spectral problem for the string equation*. Int. Eqns Oper. Theory **1** (1978), 270-277.
- [28] H. Dym and H. P. McKean. *Gaussian Processes, Function Theory, and the Inverse Spectral Problem*. Academic Press, New York–London 1976.
- [29] L. D. Faddeev. *The inverse problem in the quantum mechanical theory of scattering*. Uspekhi Mat. Nauk **14** (1959), 57-119; J. Math. Phys. **4** (1963) 72-104.
- [30] E. Fermi, J. Pasta, and S. Ulam. *Studies of nonlinear problems*. Non-linear Wave Motion. Lect. Appl. Math. vol. 15, Amer. Math. Soc. Providence 1974, pp.143-155.
- [31] A. S. Fokas and V. E. Zakharov. *Important Developments in Soliton Theory*. Springer. Berlin 1993.
- [32] C. E. Fröberg. *Calculation of the potential from the asymptotic phase*. Ark. Mat. Astr. Physik **34A** (1948), 16pp.
- [33] B. Fuchssteiner. *The Lie algebra structure of nonlinear evolution equations admitting infinite dimensional abelian symmetry groups*. Progr. Theor. Phys. **65** (1981), 861-876.
- [34] B. Fuchssteiner, B. and A. S. Fokas. *Symplectic structures, their Bäcklund transformations and hereditary symmetries*. Phys. D **4** (1981/82), 47-66.
- [35] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura. *Method for solving the Korteweg–de Vries equation*. Phys. Rev. Lett. **19** (1967), 1095-1097.
- [36] I. M. Gel'fand. *On elliptic equations*. Uspekhi Mat. Nauk **15**, 121-132; Russian Math. Surveys **15** (1960), 113-123.

- [37] I. M. Gel'fand and B. M. Levitan. *On the determination of a differential equation by its spectral function*. Izv. Akad. Nauk SSSR ser. Math. **15**, (1951), 309-360; Amer. Math. Soc. Transl. (2) **1** (1955), 253-304.
- [38] P. B. Gilkey. *Curvature and eigenvalues of the Laplacian for elliptic complexes*. Adv. Math. **10** (1973), 344-382.
- [39] P. B. Gilkey. *Asymptotic Formulas in Spectral Geometry*. Chapman and Hall/CRC, Boca Raton 2003.
- [40] G. Giraud. C. R. Acad. Sci. Paris **200** (1935), 1651-1653.
- [41] C. Gordon, D. Webb, and S. Wolpert. *Isospectral plane domains and surfaces via Riemannian orbifolds*. Invent. Math. **110** (1992), 1-22.
- [42] I. Grattan-Guinness. *The Development of the Foundations of Mathematical Analysis from Euler to Riemann*. MIT Press, Cambridge 1970.
- [43] P. C. Greiner. *An asymptotic expansion for the heat equation*. Arch. Rat. Mech. Anal. **41** (1971), 163-218.
- [44] W. Heisenberg. *Der beobachtaren Grossen in der Theorie der Elementarteilchen*. Z. Phys. **120** (1943), 513, 673.
- [45] F. Hirzebruch. *Arithmetic genera and the theorem of Riemann-Roch for algebraic varieties*. Proc. Nat. Acad. Sci. U.S.A. **40** (1954), 110-114.
- [46] F. Hirzebruch. *Topological Methods in Algebraic Geometry*. Springer Verlag, Berlin 1966.
- [47] W. V. D. Hodge. *The Theory and Application of Harmonic Integrals*. Cambridge University Press, Cambridge 1941.
- [48] R. S. Johnson. *Camassa-Holm, Korteweg-de Vries, and related models for water waves*. J. Fluid Mech. **455** (2002), 63-82.
- [49] R. Jost. *Über die falschen Nullstellen der Eigenwerte der S-Matrix*. Helv. Phys. Acta **20** (1947), 256-266.
- [50] M. Kac. *Can one hear the shape of a drum?*. Amer. Math. Monthly **73** 1966, 1-23.
- [51] J. J. Kohn and L. Nirenberg. *An algebra of pseudo-differential operators*. Comm. Pure Appl. Math. **18** (1965), 269-305.

- [52] D. J. Korteweg and G. de Vries. *On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves*. Phil. Mag. **39** (1895), 422-443.
- [53] M. G. Kreĭn. *On inverse problems for a nonhomogeneous cord*. Doklady Akad. Nauk SSSR **82** (1952), 669-672.
- [54] M. G. Kreĭn. *Solution of the inverse Sturm-Liouville problem*. Doklady Akad. Nauk SSSR **76** (1951), 21-24.
- [55] M. G. Kreĭn. *On a method of effective solution of an inverse boundary problem*. Doklady Akad. Nauk SSSR **94** (1954), 987-990.
- [56] I. M. Krichever. *The integration of non-linear equations with the help of algebraic-geometrical methods*. Funk. Anal. Pril. **11** (1977), 15-31.
- [57] P. D. Lax. *Integrals of nonlinear equations of evolution and solitary waves*. Comm. Pure Appl. Math. **21** (1968), 467-490.
- [58] P. D. Lax. *Periodic solutions of the KdV equation*. Comm. Pure Appl. Math. **28** (1975), 141-188.
- [59] N. Levinson. *On the uniqueness of the potential in a Schrödinger equation for a given asymptotic phase*. Danske Vid. Selsk. Math.-fys. Medd. **25** (1949), 29 pp.
- [60] N. Levinson. *The inverse Sturm-Liouville problem*. Mat. Tidskr. B **1949** (1949), 25-30.
- [61] J. Lützen. *The solution of partial differential equations by separation of variables: a historical survey*. Studies in the History of Mathematics, MAA Studies in Math. vol. 26. Math. Assoc. of America, 1987.
- [62] V. A. Marčenko. *Concerning the theory of a differential operator of the second order*. Dokl. Akad. Nauk SSSR **72** (1950), 457-460.
- [63] V. A. Marčenko. *On reconstruction of the potential energy from phases of the scattered waves*. Dokl. Akad. Nauk SSSR **104** (1955), 695-698.
- [64] V. A. Marčenko. *The generalized shift, transformation operators, and inverse problems*. A. A. Bolibruch et al., eds, Mathematical Events of the Twentieth Century. Springer, Berlin 2006.
- [65] H. P. McKean and I. M. Singer. *Curvature and eigenvalues of the Laplacian*. J. Diff. Geom. **1** (1967), 43-69.

- [66] S. G. Mikhlin. Composition des intégrales singulières doubles. Dokl. Akad. Nauk SSSR **2** (1936), 3-6.
- [67] J. Milnor. *Eigenvalues of the Laplace operator on certain manifolds*. Proc. Nat. Acad. Sci. USA **51** (1964), 542.
- [68] S. Minishisundaram and A. Pleijel. *Some properties of the eigenfunctions of the Laplace-operator on Riemannian manifolds*. Can. J. Math. **1** (1949), 242-256.
- [69] R. M. Miura. *Korteweg-de Vries equation and generalizations. I A remarkable explicit nonlinear transformation*. J. Math. Phys. **9** (1968), 1202-1204.
- [70] R. M. Miura, C. S. Gardner, and M. D. Kruskal. *Korteweg-de Vries equation and generalizations. II. Existence of conservation laws and constants of motion*. J. Math. Phys. **9** (1968), 1204-1209.
- [71] L. Nirenberg. Pseudo-differential operators. Global Analysis, Proc. Symp. Pure Math. vol. XVI, pp. 179-199. Amer. Mat. Soc., Providence 1968.
- [72] F. Noether. Über eine Klasse singulärer Integralgleichungen. Math. Ann. **82** (1921), 42-63.
- [73] S. P. Novikov. *The periodic problem for the Korteweg-De Vries equation, I*. Funk. Anal. Prilozh. **8** (1974), 54-66.
- [74] S. P. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov. *The Theory of Solitons*. Consultants Bureau, New York 1984.
- [75] R. Palais. *The symmetries of solitons*. Bull. Amer. Math. Soc. **34** (1997), 339-403.
- [76] A. Pleijel. *On the problem of improving Weyl's law for the asymptotic eigenvalue distribution*. Convegno Internazionale sulle Equazioni Lineari alle Derivate Parziali, Trieste 1954, pp. 69-75. Edizioni Cremonese, Roma 1955.
- [77] Lord Rayleigh (J. W. Strutt). *On waves*. Philos. Mag. **257** (1876).
- [78] Lord Rayleigh (J. W. Strutt). *The Theory of Sound*, 1877; Dover, New York, 1945.

- [79] J. S. Russell. *Report on Waves*. 14th Mtg. British Assn. Adv. Sci., John Murray, London, 1844, pp. 311-390.
- [80] R. T. Seeley. *Refinement of the functional calculus of Calderón and Zygmund*. Proc. Dutch. Acad. Ser. A **68** (1965), 521-531.
- [81] R. T. Seeley. *Elliptic singular integral operators*. Singular Integrals. Proc. Symp. Pure Math. vol. X. Amer. Math. Soc., Providence 1967.
- [82] T. J. Stieltjes. *Recherches sur les fractions continues*. Ann. Fac. Sci. Toulouse **8** (1894), 1-122.
- [83] A. Unterberger and J. Bokobza. *Les opérateurs de Calderón–Zygmund précisés* C. R. Acad. Sci. Paris **260** (1965), 34-37.
- [84] H. Weyl. *Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen* Math. Ann. **68** (1910), 220-269.
- [85] H. Weyl. *Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen*. Math. Ann. **71** (1911), 441-469.
- [86] H. Weyl. *Ueber die Abhängigkeit der Eigenschwingungen einer Membran von deren Begrenzung*. J. Reine Angew. Math. **141** (1912), 1-11.
- [87] H. Weyl. *On Hodge's theory of harmonic integrals*. Ann. Math. **44** (1943), 1-6.
- [88] H. Weyl. *Ramifications, old and new, of the eigenvalue problem*. Bull. Amer. Math. Soc. **56** (1950), 115-139.
- [89] J. Wheeler. *On the mathematical description of light nuclei by means of resonating group structure*. Phys. Rev. **52** (1937), 1107.
- [90] N. J. Zabusky and M. D. Kruskal. *Interaction of "solitons" in a collisionless plasma and the recurrence of initial states*. Phys. Rev. Letters **15** (1965), 240-243.
- [91] V. E. Zakharov and A. B. Shabat. *Exact theory of two-dimensional self-focussing and one-dimensional self-modulation of waves in nonlinear media*. Sov. Phys. JETP **34** (1972), 62-69.
- [92] V. E. Zakharov and A. B. Shabat. *A plan for integrating the nonlinear equations of mathematical physics by the method of the inverse scattering problem. I.* Funk. Anal. Pril. **8** (1974), 43-53.