1
Introduction

The aim of this chapter is to introduce the concept of Fourier series in an accessible way. We present the analytic setting in which Fourier series arise as the natural generalisation of trigonometric polynomials. We also describe how the problem of the vibrating string and the investigation of heat flow mark the beginning of the theory of Fourier series as a useful approach for solving differential equations of physical relevance. A link between trigonometric polynomials and number theory is also explored.

1.1 Trigonometric polynomials and series

A *trigonometric polynomial* of degree n is an expression of the form

$$p(t) = \sum_{k=-n}^{n} c_k \, e^{2\pi ikt} \tag{1.1}$$

where the c_ks are complex numbers with $|c_{-n}| + |c_n| \neq 0$. Thus p_n is a continuous periodic function of the real variable t, of period 1, determined by its values on $[0, 1)$, or any other interval of length 1. Since

$$\int_0^1 e^{2\pi ikt}\, dt = \begin{cases} 0 & \text{if} \quad k \neq 0, \\ 1 & \text{if} \quad k = 0, \end{cases} \tag{1.2}$$

the constants c_k in the representation (1.1) of the trigonometric polynomial p can be computed by means of

$$c_k = \int_0^1 p(t)\, e^{-2\pi ikt}\, dt, \qquad |k| \leq n. \tag{1.3}$$

The function $e_k(t) = e^{2\pi ikt}$ is sometimes referred to as the *character with frequency* k or as the kth *pure frequency*.

The trigonometric polynomials (1.1) can also be looked at geometrically. Namely, we can interpret the complex number $p(t)$ in (1.1) as the vector sum of its components, each complex number c_k being modified by a supplementary phase $2\pi kt$. In the case of real positive coefficients the visual approach is particularly simple: $p(t)$ is the extremity of a polygonal contour formed by successive straight segments with respective lengths c_k, each one making the same angle $2\pi t$ with the preceding (and following) one. A simple example is depicted in Figure 1.1; for more elaborate examples we refer to the discussion in Lévy-Leblond (1997).

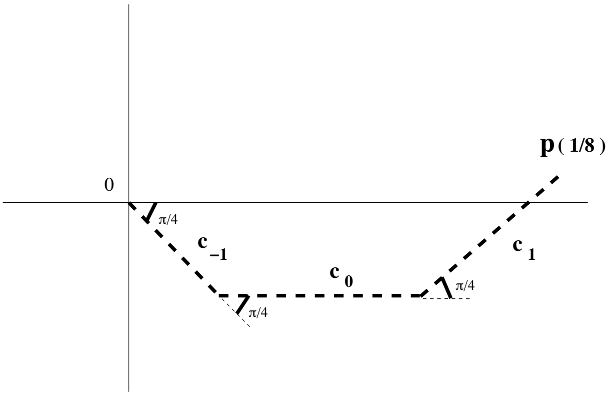

Figure 1.1 The geometric representation of the value at $t = 1/8$ of a trigonometric polynomial $p(t)$ of degree 1 and with positive coefficients.

A fundamental approximation result (to be proved in Chapter 4) is that for any continuous periodic function $f : \mathbb{R} \to \mathbb{C}$ of period 1, given $\varepsilon > 0$, there is a trigonometric polynomial p with

$$|p(t) - f(t)| < \varepsilon, \qquad t \in \mathbb{R}. \tag{1.4}$$

Due to periodicity, it suffices to verify the above inequality for $t \in [0, 1)$.

The role of the multiplicative factor (2π) in the argument of the fundamental trigonometric monomials $e^{2\pi ikt}$ used in (1.1) is to normalise the period to 1. However, given that (1.1) can be expressed as $p(t) = \sum_{k=0}^{n} a_k \cos(2\pi kt) + \sum_{k=0}^{n} b_k \sin(2\pi kt)$, for some $a_k, b_k \in \mathbb{C}$, it is reasonable to wonder why we de not associate the terminology “trigonometric polynomial" with functions of the form

$$q(t) = \sum_{k=0}^{n} \alpha_k \cos^k(2\pi t) + \sum_{k=0}^{n} \beta_k \sin^k(2\pi t) \tag{1.5}$$

for some $a_k, b_k \in \mathbb{C}$. An exercise in trigonometric identities[1] shows that any function of type (1.5)

[1] In this context, it is comforting to know that, see Borzellino and Sherman (2012), polynomial relations between $\cos(2\pi t)$ and $\sin(2\pi t)$ are always consequences of the Pythagorean identity $\cos^2(2\pi t) + \sin^2(2\pi t) = 1$; there are no hidden tricks.

can be written in the form (1.1), with the same value of the degree. However, not all trigonometric polynomials are expressible in the form (1.5): for example, $t \mapsto \sin(4N\pi t)$ with $N \neq 0$ integer are not expressible, see Borzellino and Sherman (2012). For this reason,[2] expressions of the form (1.5) are not enough to approximate well continuous periodic functions of period 1.

The approximation result expressed by means of (1.4) leads us naturally to the concept of a *trigonometric series* or *Fourier series*, defined in analogy to (1.1) as an expression of the form

$$\sum_{k=-\infty}^{\infty} c_k\, e^{2\pi i k t}\,, \tag{1.6}$$

representing formally a function f of period 1. In light of (1.3), we expect that the constants c_k in (1.6) and the function f are connected by the formula

$$c_k = \int_0^1 f(t)\, e^{-2\pi i k t}\, dt\,, \qquad k \in \mathbb{Z}. \tag{1.7}$$

More generally, the Fourier series associated to a function $f : \mathbb{R} \to \mathbb{C}$ of period $T > 0$ is

$$\sum_{k=-\infty}^{\infty} c_k\, e^{2\pi i k t/T}\,, \tag{1.8}$$

where

$$c_k = \frac{1}{T}\int_0^T f(t)\, e^{-2\pi i k t/T}\, dt\,, \qquad k \in \mathbb{Z}\,. \tag{1.9}$$

The theory of Fourier series studies the classes of periodic functions (of period $T > 0$) and the notions of convergence appropriate for the correspondence $f(t) \approx \sum_{k=-\infty}^{\infty} c_k\, e^{2\pi i k t/T}$, with the constants c_k given by (1.9), expressing the function f in terms of a superposition of oscillations with frequencies $\nu_k = k/T$ that are integer multiples of the fundamental frequency $\nu = 1/T$. As a glimpse at the intricacy of the subject, notice that above we pointed out that for any continuous periodic function $f : \mathbb{R} \to \mathbb{C}$ of period 1 we can find trigonometric polynomials that approximate it uniformly, that is, in the sense of (1.4). Nevertheless, the specific trigonometric polynomials obtained by means of the symmetric partial sums

$$s_n(f,t) = \sum_{k=-n}^{n} c_k\, e^{2\pi i k t} \tag{1.10}$$

with c_k given by (1.9) are not necessarily good approximations: the sequence

[2] The orthogonality considerations made in Chapter 3 show that if we rely only on functions of the form (1.5), then the approximations miss out an infinite-dimensional subspace of the space of square integrable functions.

$\{s_n(f,t)\}_{n\geq 1}$ might diverge for infinitely many values of $t \in [0,1]$; see the discussion in the introduction to Chapter 4. This shows that continuity coupled with the concept of pointwise or uniform convergence is not adequate. The proper setting turns out to be the class of Lebesgue integrable or square integrable functions, with an associated concept of convergence. The need to go beyond the class of continuous functions and the classical theory of Riemann integrable functions is fully justified by the mathematical power and flexibility of the theory within the new setting, and is further emphasised by its wide range of applicability.

1.2 The dawn of the theory

Fourier analysis dates back to late eighteenth/early nineteenth century studies of the vibrating string and of heat propagation. Two basic partial differential equations of one-dimensional mathematical physics are the wave equation

$$\frac{\partial^2 U}{\partial T^2} = c^2 \frac{\partial^2 U}{\partial X^2} \tag{1.11}$$

and the heat equation

$$\frac{\partial U}{\partial T} = \kappa \frac{\partial^2 U}{\partial X^2}, \tag{1.12}$$

where $c > 0$ and $\kappa > 0$ are physical constants. In (1.11), $U = U(X,T)$ represents, at the location X and at time T, the displacement of a homogeneous string placed in the (X,Y)-plane and stretched along the X-axis between $X = 0$ and $X = L$, where it is tied. The value of the constant c is $\sqrt{\tau/\rho}$, where τ is the tension coefficient of the string and ρ is its mass density. Equation (1.11) is to be solved for $T > 0$ and X between 0 and L, subject to the boundary conditions

$$U(0,T) = U(L,T) = 0, \qquad T \geq 0, \tag{1.13}$$

which express the fact that the endpoints of the string are fixed. The solution U describes the vibrations of a violin string. On the other hand, in (1.12), $U = U(X,T)$ is the temperature in a homogeneous, straight wire of length L, whose endpoints are held at constant temperature zero. The value of the constant κ in (1.12) is specific to the conducting material. The problem is to describe the temperature at time T from its knowledge at time $T = 0$. Consequently, we seek solutions to (1.12) for $T > 0$ and X between 0 and L, subject to the boundary conditions

$$U(0,T) = U(L,T) = 0, \qquad T \geq 0, \tag{1.14}$$

and with the initial temperature specified by

$$U(X,0) = U_0(X), \qquad 0 \le X \le L. \tag{1.15}$$

For the physical derivation of (1.11) and (1.12) we refer to Dym and McKean (1972), Krantz (1999), Stein and Shakarchi (2003) and Strauss (2008). We now discuss some mathematical aspects of historical interest that provided the motivation for the development of the rigorous theory of Fourier series.

The first natural step in the mathematical investigation of (1.11) consists of scaling the equation: a change of units permits us to write the equation in non-dimensional form, thus reducing the number of physical parameters involved. This can be accomplished by means of the change of variables

$$X = Lx, \quad T = \frac{L}{c}\,t, \quad U(X,T) = c_0\,u(x,t), \tag{1.16}$$

where $c_0 = 1$ m is the reference length. The fact that X takes values between 0 and L translates into $x \in [0,1]$, the constant c is absorbed into (1.16), and all variables (the independent variables x and t, as well as the dependent variable u) are now numbers, whereas X and U were expressed initially in m (metres) and T in s (seconds). Clearly $\frac{\partial U}{\partial X} = \frac{c_0}{L}\frac{\partial u}{\partial x}$, $\frac{\partial U}{\partial T} = \frac{c_0 c}{L}\frac{\partial u}{\partial t}$, $\frac{\partial^2 U}{\partial X^2} = \frac{c_0^2}{L^2}\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 U}{\partial T^2} = \frac{c_0^2 c^2}{L^2}\frac{\partial^2 u}{\partial t^2}$, so that (1.11) and (1.13) are transformed into

$$\begin{cases} u_{tt} = u_{xx}, & t > 0, \quad 0 \le x \le 1, \\ u(0,t) = u(1,t) = 0, & t \ge 0, \end{cases} \tag{1.17}$$

where $u_{tt} = \frac{\partial^2 u}{\partial t^2}$ and $u_{xx} = \frac{\partial^2 u}{\partial x^2}$. Once we solve (1.17), we can return to the original physical variables by making the change of variables inverse to (1.16). Similarly, the nondimensional scaled version of (1.12) coupled with (1.14)–(1.15) is

$$\begin{cases} u_t = u_{xx}, & t > 0, \quad 0 \le x \le 1, \\ u(0,t) = u(1,t) = 0, & t \ge 0, \\ u(x,0) = u_0(x), & 0 \le x \le 1, \end{cases} \tag{1.18}$$

with $u_0 : [0,1] \to \mathbb{R}$ a given continuous function satisfying $u_0(0) = u_0(1) = 0$; here $u_t = \frac{\partial u}{\partial t}$. The issue of finding proper initial data (at time $t = 0$) for (1.17), playing the role that u_0 plays for (1.18), is discussed in Section 1.2.1.

1.2.1 The vibrating string controversy

For the sake of simplicity, let us first drop the restrictions $0 \le x \le 1$ and $t \ge 0$, and suppose that u is twice differentiable and solves the partial differential equation in (1.17) for all real x and t. If we change variables $\xi = x-t$, $\eta = x+t$, and set $\gamma(\xi,\eta) = u(x,t)$, in terms of the new variables the partial differential equation in (1.17) becomes $\dfrac{\partial^2\gamma}{\partial\xi\partial\eta} = 0$. Integrating this relation twice gives $\gamma(\xi,\eta) = f(\xi) + g(\eta)$ for some functions f and g, so that

$$u(x,t) = f(x-t) + g(x+t)\,. \tag{1.19}$$

Note that the graph of the function $x \mapsto f(x-t)$ at time $t = 0$ is simply the graph of the function f, while at time t it becomes the graph of f translated by t: $f(x-t)$ represents a *travelling wave* (a pattern that travels without change of form) which propagates to the right with unit speed; see Figure 1.2. Similarly, $g(x+t)$ represents a travelling wave that propagates to the left with unit speed.[3] The partial differential equation in (1.17) being linear, the *superposition principle* holds: if $u_1(x,t)$ and $u_2(x,t)$ are solutions, then so is $a\,u_1(x,t) + b\,u_2(x,t)$ for any constants a and b. In particular, (1.19) shows that the general solution is a superposition of two waves travelling in opposite directions.

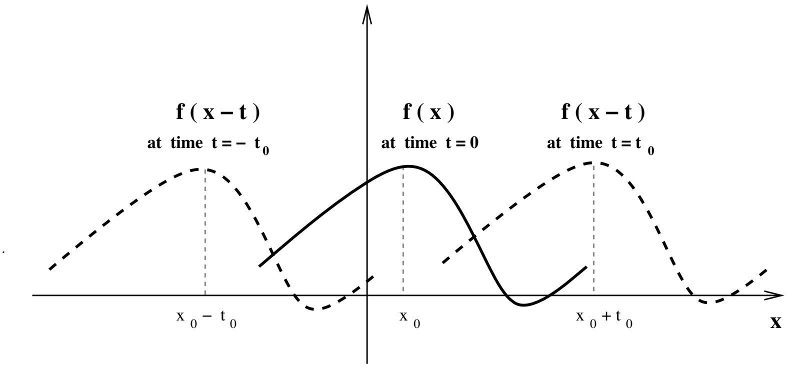

Figure 1.2 A travelling wave propagating to the right: the profile is depicted at three successive times $t = -t_0$, $t = 0$ and $t = t_0$ with $t_0 > 0$.

To connect the obtained result with the original problem (1.17), provided that $u(x,0) = u_0(x)$ for $x \in [0,1]$ specifies the initial shape of the string, extend u_0 to $\mathbb{R}$ as an odd function[4] of period 2. Also, extend the presumed solution

[3] The non-dimensional unit speed corresponds to the speed c in the original physical variables, if we recall the scaling (1.16).

[4] Meaning that $u_0(-x) = -u_0(x)$ and $u_0(x+2) = u_0(x)$ for $x > 0$. Note that if we extend u_0 to the whole real line, relation (1.22) emerges, and this forces oddness. With regard to periodicity, the boundary conditions in (1.17) might seem to indicate the period 1. However, period 2 and oddness combined do not impose any constraint upon $u_0 : [0,1] \to \mathbb{R}$ with $u_0(0) = u_0(1) = 0$, whereas period 1 and oddness require $u_0(1/2) = 0$, for example, due to $u_0(1/2) = u_0(1/2-1) = u_0(-1/2) = -u_0(1/2)$.

$u(x,t)$ from $[0,1]\times[0,\infty)$ to $\mathbb{R}^2$ by requiring that for every fixed $t \geq 0$, the map $x \mapsto u(x,t)$ is odd and periodic of period 2, while for $t < 0$ we simply solve (1.17) backwards in time: we seek a twice differentiable solution $u : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ to (1.17). Consequently, the solution u must be of the form (1.19) for some functions f and g. We get

$$f(x) + g(x) = u_0(x), \qquad x \in \mathbb{R}, \tag{1.20}$$

by evaluating (1.19) at $t = 0$. To accommodate the boundary conditions in (1.17), (1.19) yields $f(-t) + g(t) = 0$ for all $t \in \mathbb{R}$, so that (1.19) and (1.20) take the form

$$u(x,t) = g(x+t) - g(t-x), \qquad x, t \in \mathbb{R}, \tag{1.21}$$

$$g(x) - g(-x) = u_0(x), \qquad x \in \mathbb{R}, \tag{1.22}$$

respectively. The formula (1.19) was first obtained in 1747 by d'Alembert, who was concerned with finding the general solution of the partial differential equation and ignored the physical context,[5] in particular, the significance of (1.22). A closer look at (1.22) reveals that the form of its left side encodes the fact that the function u_0 is odd, but this relation by itself does not determine the function g. For example, if g is a solution to (1.22), so will be $g + g_0$ for any even function[6] g_0 of period 1. The underlying physics indicates that perhaps the initial velocity[7] $v_0(x) = \frac{\partial u}{\partial t}(x,0)$ for $x \in [0,1]$, might be relevant. Indeed, if v_0 is given on $[0,1]$, we extend it to $\mathbb{R}$ by requiring it to be odd and periodic of period 2. Differentiating (1.21) with respect to the time variable[8] and evaluating the outcome at $t = 0$, we get

$$g'(x) - g'(-x) = v_0(x), \qquad x \in \mathbb{R}. \tag{1.23}$$

Now (1.22) and (1.23) yield $g(x) = \frac{1}{2}\left[u_0(x) + \int_0^x v_0(s)\,ds\right] + \alpha$ for some constant α. Since u_0 and v_0 are odd,[9] using (1.21), we obtain

$$u(x,t) = \frac{u_0(x-t) + u_0(x+t)}{2} + \frac{1}{2}\int_{x-t}^{x+t} v_0(s)\,ds, \tag{1.24}$$

[5] It was not unusual for d'Alembert to sacrifice physical reality for a purely philosophical viewpoint, see Wheeler and Crummett (1987).

[6] Meaning that $g_0(-x) = g_0(x)$ for all $x \in \mathbb{R}$.

[7] $u_t = \frac{\partial u}{\partial t}$ is the rate of change of the displacement of a particular point on the string and (generally) differs from the speed c of propagation of the wave along the string. This situation is also encountered for waves in media other than strings.

[8] Since u is twice differentiable, $t \mapsto u(t,t)$ is differentiable. From (1.21) with $x = t$ we then infer that g is differentiable.

[9] So that $-u_0(t-x) = u_0(x-t)$ and $\int_{t-x}^0 v_0(s)\,ds = \int_{x-t}^0 v_0(y)\,dy$, the latter as a consequence of the change of variables $y = -s$.

as shown in 1748 by Euler. One can check directly that for a general twice differentiable function $u_0 : \mathbb{R} \to \mathbb{R}$ and for a general differentiable function $v_0 : \mathbb{R} \to \mathbb{R}$, (1.24) provides us with a classical solution to the wave equation $u_{tt} - u_{xx} = 0$, with $u(x,0) = u_0(x)$ and $u_t(x,0) = v_0(x)$. In our particular setting, observe that the extensions performed for u_0 and v_0 ensure the validity of the boundary conditions in (1.17). The above discussion illustrates the fruitful interplay between abstract mathematics and its relation to nature: physical intuition can provide a feeling for mathematical facts and the other way around.

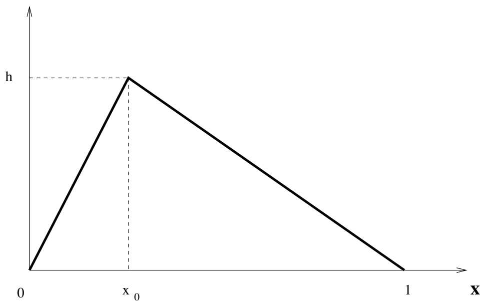

Figure 1.3 Initial position of a plucked string.

Euler's solution (1.24) differed from d'Alembert's (1.21) on the specification[10] of the function g. Furthermore, Euler proclaimed that the function g does not need to be differentiable, but may be any curve drawn by hand.[11] Euler had in mind the plucked string: taking as the initial position of the string the triangular shape given, for some constants $x_0 \in (0,1)$ and $h > 0$, by

$$u_0(x) = \begin{cases} \frac{xh}{x_0} & \text{for} \quad 0 \le x \le x_0, \\ \frac{h(1-x)}{1-x_0} & \text{for} \quad x_0 \le x \le 1, \end{cases} \tag{1.25}$$

(see Figure 1.3), and choosing zero initial velocity $v_0 \equiv 0$, Euler declared that the subsequent positions of the string are given by

$$u(x,t) = \frac{u_0(x-t) + u_0(x+t)}{2}, \qquad t \ge 0, \quad 0 \le x \le 1, \tag{1.26}$$

obtained formally from (1.24). Euler used a physical observation (the fact that the violin string could be released from an initial position with a corner) to impose a mathematical formula. The unsatisfactory aspect of the solution (1.26)

[10] Due to (1.20), the knowledge of g determines f uniquely in terms of the initial position u_0.
[11] The mathematical formalism which we take for granted today was not available at that time: Euler did not perceive a function to be an arbitrary rule that assigns to every point of the domain of definition a single point of the range, see Krantz (1999). This explains the vague formulation.

is that it does not satisfy the partial differential equation we set out to solve: since u_0 is not differentiable at $x = x_0 \in (0, 1)$, the function $u(x, t)$ defined by (1.26) is not differentiable. In light of this, d'Alembert objected to physical arguments for solutions to a partial differential equation and called for the other researchers to engage in mathematics, see Wheeler and Crummett (1987). Euler defended his solutions with corners with mathematically unconvincing arguments. His position was later on vindicated: it turns out that u does solve the equation in an appropriate generalised sense, the understanding of which requires the theory of distributions.

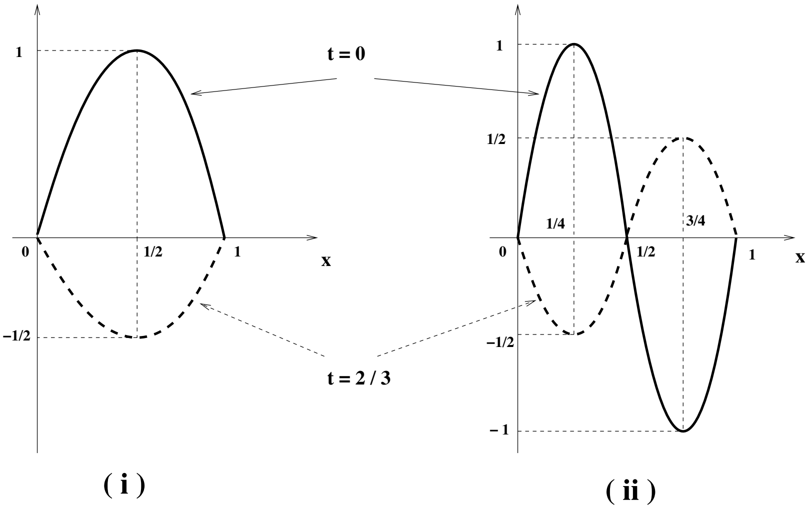

Figure 1.4 Fundamental tone (i) and the first overtone (ii) at two instants in time.

Daniel Bernoulli enters the debate in 1753 in the midst of the d'Alembert–Euler disagreement. His starting point, see Benedetto (1997), is Brook Taylor's observation from 1715 that for any integer $m \geq 1$ the tone[12]

$$u_m(x, t) = \sin(\pi m x)\, \cos(\pi m t) \tag{1.27}$$

represents a solution to (1.17) with zero initial velocity. In contrast to the travelling waves discussed before, (1.27) represents a *standing wave*. The terminology comes from looking at the graph of $x \mapsto u_m(x, t)$ as t varies (see Figure 1.4): the endpoints $x = 0$ and $x = 1$, as well as all points $x = \frac{k}{m}$ with

[12] The case $m = 1$ is called the *fundamental tone* or *first harmonic* of the vibrating string, and $m \geq 2$ are the *overtones* or *higher harmonics*, $m = 2$ being the *first overtone* or *second harmonic*, see Stein and Shakarchi (2003).

$k \in \{1, \cdots, m-1\}$, remain motionless in time and are called nodes. The points $x = \frac{2k-1}{2m}$ with $k \in \{1, \cdots, m\}$, where the amplitude is maximal, are named anti-nodes. Bernoulli argued formally in terms of the physics of sound and provided no mathematical support for his arguments, see Wheeler and Crummett (1987), claiming that the solution to (1.17) with initial velocity $v_0 \equiv 0$ must be an infinite sum of the tones:

$$u(x,t) = \sum_{m=1}^{\infty} a_m \sin(\pi m x) \cos(\pi m t)\,. \tag{1.28}$$

Using the trigonometric identity $\sin(\alpha)\cos(\beta) = \frac{\sin(\alpha+\beta)+\sin(\alpha-\beta)}{2}$, we can express (1.28) equivalently in the form

$$u(x,t) = \sum_{m=1}^{\infty} a_m \frac{\sin(\pi m[x-t]) + \sin(\pi m[x+t])}{2}. \tag{1.29}$$

Setting $v_0 \equiv 0$ in (1.24), we see that the problem of reconciling Bernoulli's formal solution to d'Alembert's rigorous solution reduces to the question of whether a twice differentiable odd periodic function u_0 of period 2 may be written in the form $u_0(x) = \sum_{m=1}^{\infty} a_m \sin(\pi m x)$. d'Alembert objected to Bernoulli's solution on physical grounds, believing (erroneously) that there was only one possible frequency associated with a vibration, while Euler felt that Bernoulli's solution was too special.

Lagrange entered the debate in 1759, supporting Euler's admission of functions with corners and dismissing Bernoulli's solution, see Wheeler and Crummett (1987). He proposed a completely new approach that avoided the wave equation by viewing the string as a collection of n equally spaced point masses, connected by a light cord. This model leads to a set of n equations of the form $\frac{d^2 y_k}{dt^2} = y_{k-1} - 2y_k + y_{k+1}$. After solving the system for a finite number of masses, Lagrange generated, for the initial position u_0 and the initial velocity v_0 of the string (both odd and of period 2), the solution

$$\begin{aligned} u(x,t) &= 2\sum_{m=1}^{\infty}\Big(\int_0^1 \sin(\pi m s)\, u_0(s)\, ds\Big) \cos(\pi m t)\, \sin(\pi m x) \\ &\quad + \frac{2}{\pi}\sum_{m=1}^{\infty}\Big(\int_0^1 \frac{\sin(\pi m s)}{m}\, v_0(s)\, ds\Big) \sin(\pi m t)\, \sin(\pi m x)\,. \end{aligned} \tag{1.30}$$

Note that if we set $t = 0$ in (1.30) and if we interchange summation and integration, then (1.30) gives rise to the Fourier series expansion of the function u_0, while differentiating (1.30) with respect to t and subsequently setting $t = 0$ yields the Fourier series expansion of the function v_0.