Cambridge University Press & Assessment
978-1-107-04272-8 — Discrete Systems and Integrability
J. Hietarinta , N. Joshi , F. W. Nijhoff
Excerpt
More Information

1

Introduction to difference equations

This chapter serves to introduce the various concepts of discretization and types of equations that we will encounter in the course of this book. We start by recalling basic properties of differential equations and show how some of them translate to difference equations. An important equation that we will meet several times is the discrete Riccati equation, which we discuss here in some detail. Partial difference equations form a major theme of this book, and we introduce linear partial difference equations, which are discrete versions of canonical types that are important in mathematical physics.

For basic methods and tools of the elementary theory of difference calculus and difference equations, we refer the reader to Appendix A.

1.1 A first look at discrete equations

At the most fundamental level, differential equations can be divided into *ordinary differential equations* (ODE) and *partial differential equations* (PDE), depending on the number of independent variables. Their respective theories are quite different in nature. A similar division can be made for difference equations, distinguishing between *ordinary difference equations* (OΔEs) and *partial difference equations* (PΔEs), depending on whether there is one or more than one independent (discrete) variable(s).

1.1.1 Ordinary differential equations (ODEs)

Let us first recall some basic notions about ODEs. The general form of an ODE is

$$\mathcal{F}\left(y(x), y'(x), y''(x), \dots y^{(n)}(x); x \right) = 0, \tag{1.1}$$

in which $\mathcal{F}$ is some given function of its arguments. Here x is the **independent variable**, while $y(x)$ is the **dependent variable** since it depends on x. The primes in (1.1) denote differentiation with respect to x; if there are higher derivatives we use a superscript in parentheses:

$$y'(x) = \frac{dy}{dx}, \quad y''(x) = \frac{d^2y}{dx^2}, \quad \dots \quad y^{(n)} = \frac{d^ny}{dx^n},$$

where the number of primes gives the **order** of the derivative and the order of the equation is given by the order of the highest derivative appearing in the equation, while the **degree** of the equation is the power of the highest derivative. The domain and range of the function y needs to be specified; usually $y : \mathbb{R} \to \mathbb{R}$ or $y : \mathbb{C} \to \mathbb{C}$. The dependent variable can be multi-component, e.g. $y : \mathbb{R} \to \mathbb{R}^M$, in which case there can also be several equations. The equation may depend also on various **parameters**, i.e. unspecified constants (quantities independent of x). If $\mathcal{F}$ does not depend explicitly on x, i.e. x enters only through the function y and its derivatives, then the ODE is called **autonomous**; otherwise it is called **nonautonomous**.

Sometimes we can separate the highest derivative:

$$y^{(n)}(x) = F(y, y', \ldots, y^{(n-1)}; x). \tag{1.2}$$

Such a higher-order equation can be written as a first-order multi-component equation

$$\boldsymbol{y}' = \boldsymbol{F}(\boldsymbol{y}; x), \tag{1.3}$$

by defining (here the superscript T stands for transpose)

$$\boldsymbol{y} := (y, y', \ldots, y^{(n-1)})^T, \quad \boldsymbol{F} := \left(y', y'', \ldots, y^{(n-1)}, F(y, y', \ldots, y^{(n-1)}; x)\right)^T.$$

Usually the aim is to solve for y as a function of x, wherever possible, from (1.1) or (1.2) or (1.3). The solution, if we can find it at all, is not unique, and needs further data in order to render it unique. This can be done by imposing, in addition to the ODE itself, a number of *initial data*; by fixing at a given value of x, say at $x = x_0$, the values of y, y', y'', $\ldots$, $y^{(n-1)}$; or by fixing the data at several boundary points. This leads to solutions $y(x)$ that are functions defined over a domain that is determined by the initial or boundary data. The idea that this process may lead to functions $y(x)$ that have never before been described was very influential in the development of the theory that led to Painlevé transcendents around the turn of the twentieth century. This idea will arise again when we consider functions defined as solutions of difference equations.

Example 1.1.1 Here are some examples of famous differential equations that have been studied widely in the literature and will play a role in the later chapters. The Verhulst equation (or logistic growth model) is given by

$$\frac{dy}{dx} = a\, y(1 - y), \tag{1.4}$$

where a is a constant parameter, and is first order and first degree. Weierstrass' elliptic function $\wp$ satisfies an equation of first order and second degree

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3, \tag{1.5}$$

where g_2 and g_3 are constant parameters. The first Painlevé equation P_I is given by

$$y''(x) = 6y^2(x) + x, \tag{1.6}$$

and is second order and nonautonomous, while the Chazy equation

$$y''' - 2yy'' + 3y'^2 = 0 \tag{1.7}$$

is third order and autonomous. We will later consider discrete versions of some of these equations.

In the examples given above, the equations contain parameters, which are usually fixed at a given value. However, there is often merit in considering these parameters as *free* (i.e. unspecified) so that the solutions of the differential equations can be studied as functions of both the parameters and the independent variable.

1.1.2 The difference operator

Given a continuous system, an important question is how to find a corresponding discrete equation, which has qualitatively the same properties (such as conserved quantities). This question arises as a core problem in numerical analysis, where there is a vast literature on appropriate ways of discretizing a given differential equation.

The starting point of discretizing an ODE is to recall that the derivative can be considered as a limit of a difference

$$\frac{dy}{dx} = \lim_{\delta\to 0} \frac{y(x+\delta) - y(x)}{\delta}. \tag{1.8}$$

The discretization procedure effectively amounts to going one step back, namely to the object we had before the limit is taken, which is given by the **difference operator** Δ_δ defined by

$$\Delta_\delta y(x) := \frac{y(x+\delta) - y(x)}{\delta}. \tag{1.9}$$

Replacing a derivative by Δ_δ (or by similar discrete objects) often leads to a good approximation to the differential equation, the solution of which can be iterated on a computer. In numerical analysis, more subtle approaches are often needed to capture the correct qualitative behavior of the solution (which may also involve replacing the nonlinear terms by suitably chosen expressions).

Higher derivatives also need to be considered. Since these are also defined by limits, for example

$$\frac{d^2y}{dx^2} = \lim_{\delta\to 0} \frac{y(x+\delta) - 2y(x) + y(x-\delta)}{\delta^2}, \tag{1.10}$$

their discretization requires more points where the function $y(x)$ is evaluated. These operations can be expressed in terms of the **(lattice) shift operator**:

$$T_\delta y(x) = y(x+\delta). \tag{1.11}$$

It is easy to see that the difference operator Δ_δ and its powers can be simply expressed in terms of the shifts T_δ through the following formulae

$$\begin{aligned}
\Delta_\delta y(x) &= \frac{1}{\delta}\,(T_\delta - \mathrm{id})\, y(x)\,, \\
\Delta_\delta^2 y(x) &= \frac{1}{\delta^2}\left(T_\delta^2 - 2T_\delta + \mathrm{id}\right) y(x), \\
&\vdots \\
\Delta_\delta^n y(x) &= \frac{(-1)^n}{\delta^n}\sum_{j=0}^{n}\begin{pmatrix} n \\ j \end{pmatrix}(-T_\delta)^{n-j}\, y(x)
\end{aligned} \tag{1.12}$$

(here and later "id" denotes the identity operator) and thus the nth-order difference operator acting on $y(x)$ can be expressed in terms of the shifted variables $y(x), y(x+\delta), \ldots, y(x+n\delta)$.

One way to obtain discrete dynamics from continuous ones is to take stroboscopic pictures of the flow. We assume that the underlying dynamics is given by an equation such as (1.3), and consider a sequence of values of the solution $y : \mathbb{Z} \to \mathbb{R}^N$ at regularly spaced points, $x + n\delta$, by defining

$$y_n := y(x_0 + n\delta) = y(x). \tag{1.13}$$

Another circumstance in which discrete maps arise is by taking a **Poincaré section** of the orbit of a continuous system; see Figure 1.1. If the orbit revisits the same region of the space of the dependent variable repeatedly, the flow will intersect again and again with some given surface. The intersections give rise to a discrete sequence y_n, where n labels the crossings of the given surface. (Sometimes it is best to keep only those crossings that come from the same side of the surface.)

1.1.3 Ordinary difference equations

In analogy to (1.1) we can, by replacing derivatives with differences, define an equation of the form:

$$\mathcal{F}(y(x), \Delta_\delta y(x), \Delta_\delta^2 y(x), \ldots; x) = 0.$$

Using the replacements (1.12) this can also be rewritten in the form

$$\overline{\mathcal{F}}(y(x), y(x+\delta), y(x+2\delta), \ldots; x) = 0\,, \tag{1.14}$$

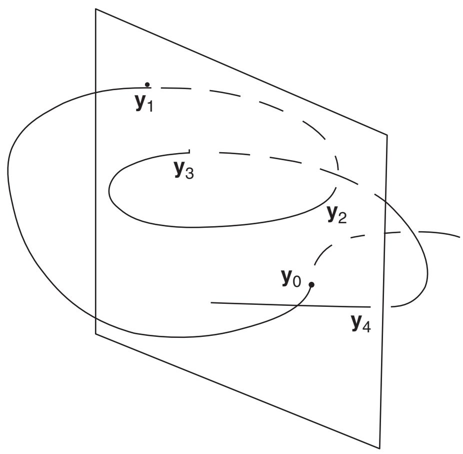

Figure 1.1 Dynamical mapping arising from the Poincaré section of an orbit.

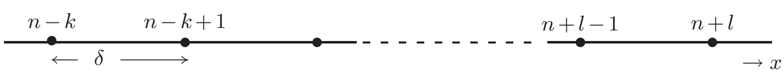

Figure 1.2 An illustration of the points involved in the difference equation (1.15) of order $k+l$.

where the expression $\overline{\mathcal{F}}$ can straightforwardly be obtained from $\mathcal{F}$ by using the above-mentioned substitutions.

At this point, we should ask ourselves what is actually meant by an equation of the form (1.14). In other words, what range of values can the independent variable x take and what would represent a solution of this equation? These are deep questions. For the time being, we could distinguish between the cases where x takes values in a discrete set or where it takes values in a continuous domain. We call the former type of equation a **finite-difference equation** and the latter an **analytic difference equation**.

In the latter case, we can rewrite the equation in different ways; by shifting x, for example

$$\overline{\mathcal{F}}\,(y(x-k\delta), y(x-k\delta+\delta), \ldots, y(x+l\delta-\delta), y(x+l\delta); x) = 0, \tag{1.15}$$

where k, l are fixed integers with $k+l > 0$ (see Figure 1.2). Thus the equation only involves integer multiples of δ in the arguments of the dependent variable $y(x)$, and if $n\delta = x$ and then the point named $n+s$ corresponds to the value $x+s\delta$ of the independent variable.

Analytic difference equations

If the independent variable x is a *continuous* variable (typically x in subdomains of $\mathbb{R}$ or $\mathbb{C}$), solving the difference equation (1.15) would amount to finding a *function* $y(x)$, for x in some appropriate domain. We may want to specify the function class or space to which

y belongs, for example, it may be a meromorphic function of x. However, often we will leave this open, depending on the type of problem we wish to solve.

Clearly, there is also indeterminacy in this problem because a solution is needed for the whole domain, whereas only values at discrete intervals are used in the equation itself. This implies that the initial data has to be given over an entire interval on the real line, and consequently the solution of the difference equations is determined only up to periodic functions, i.e. functions $\pi(x)$ obeying

$$\pi(x+\delta) = \pi(x),$$

which here play the role of "integration constants".

Finite-difference equations

In contrast to the analytic difference case, we may assume that x starts from a fixed base point value (such as 0), in which case it is natural to consider y as a function of the integers $y : \mathbb{Z} \to \mathbb{R}$ or $\mathbb{C}$. We then use the shorthand notation $y_n := y(x + n\delta)$. In this case, we have equations containing the discrete variable y_n and its various shifts $y_{n-k}, y_{n-k+1}, \dots,$ $y_{n+l-1}, \dots\ y_{n+l}$. Thus the canonical form of a difference equation is

$$\mathcal{F}(y_{n-k}, y_{n-k+1}, \dots, y_{n+l-1}, y_{n+l}; n) = 0, \quad \forall n \in \mathbb{Z}. \tag{1.16}$$

This is variously called a *recurrence relation* or *iterative scheme* from which we wish to solve y at discrete points n only. Thus, giving initial values at a sufficient number of points, generically $y_0, y_1, \dots, y_{k+l-1}$, we can hope to iterate the equation and calculate y_{k+l} and subsequent values step by step.

If we can solve (1.16) for y_{n+l} then the equation provides a **map**

$$(y_{n-k}, y_{n-k+1}, \dots, y_{n+l-1}) \mapsto y_{n+l},$$

and if we can also solve for y_{n-k} then we have a **reversible map**. If we cannot uniquely solve for the first or last term from equation (1.16) then it is called a *correspondence* and the possibility of multiple ways to proceed must be studied carefully.

The **order** of equation (1.16) is $k+l$, which is the minimal number of initial data required to uniquely define the evolution of y_n as a function of integer n. We will sometimes refer to an equation of the form (1.16) also as a $(k+l+1)$-**point map**, where it is understood that the map is iterated by composing the map with itself. (The iteration scheme could break down if the composition of the map becomes ill-defined. That may occur if the original difference equation becomes noninvertible at certain points in the domain of the independent variable.)

Almost all maps we consider in this book are **rational** maps; that is, the latest iterate is a rational function of previous iterates. We will see that integrable maps turn out to be **birational**; that is, the map is rational in both directions.

Example 1.1.2 A simple example of a finite-difference equation is

$$y_{n+1} = a\, y_n, \tag{1.17}$$

where a is a constant parameter. The explicit general solution is

$$y_n = y_0\, a^n,$$

where y_0 is a free constant. On the other hand, if it is an analytic difference equation

$$y(x+1) = a(x)\, y(x), \tag{1.18}$$

where $a(x+1) = a(x)$, we get the general solution

$$y(x) = \pi_0(x)\, a(x)^x,$$

where $\pi_0(x)$ is a periodic function $\pi_0(x+1) = \pi_0(x)$ but is otherwise free.

The periodic functions that arise in the solution of analytic difference equations may be subject to further conditions that may depend on the context in which the equation arises, such as in physical models.

As before, if $\mathcal{F}$ is explicitly independent of n (i.e. n does not appear in the equation other than through the dependence of y on n), then the equation is called *autonomous*. Otherwise, it is called *nonautonomous*.

Example 1.1.3 The McMillan map is defined by the difference equation

$$x_{n+1} + x_{n-1} = \frac{2a x_n}{1 - x_n^2}, \tag{1.19}$$

which is an autonomous birational equation.

Example 1.1.4 The first Painlevé equation (1.6) can be discretized as

$$y_{n+1} + y_n + y_{n-1} = \frac{\alpha + \beta n}{y_n} + \gamma, \tag{1.20}$$

which is called "discrete Painlevé I" ($\mathrm{dP_I}$). The "alternate" $\mathrm{dP_I}$ is given by

$$\frac{n+1}{y_{n+1} + y_n} + \frac{n}{y_n + y_{n-1}} = n + a + b y_n^2. \tag{1.21}$$

Both maps are clearly birational (i.e. rational in both directions) and nonautonomous.

Example 1.1.5 The recurrence relation for Hermite polynomials is given by

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \tag{1.22}$$

which can be interpreted as a (nonautonomous) difference equation with independent variable n and parameter x. Hermite polynomials also solve a differential equation with x as the independent variable and n as a parameter

$$H_n'' - 2xH_n' + 2nH_n = 0, \tag{1.23}$$

where the prime refers to derivation with respect to x.

Finite-difference equations of higher order can also be viewed as dynamical mappings by converting them to multicomponent first-order maps in the same way as is done with ODEs. Assuming that we can solve y_{n+l} uniquely from (1.16), leading to an expression of the form

$$y_{n+l} = F(y_{n-k}, \ldots, y_{n+l-1}),$$

we can write it as a dynamical map

$$\boldsymbol{y}_n \quad \mapsto \quad \boldsymbol{y}_{n+1} = \boldsymbol{F}(\boldsymbol{y}_n),$$

where we have introduced the $(k+l)$-component vectors

$$\boldsymbol{y}_n := (y_{n-k}, \ldots, y_{n+l-1})^t, \quad \boldsymbol{F}(\boldsymbol{y}) := (y_{n-k+1}, \ldots, y_{n+l-1}, F(y_{n-k}, \ldots, y_{n+l-1})).$$

This is one of many ways in which a scalar difference equation, of order $N \geq 2$, can be transformed to a system of N first-order equations. (Note that the converse statement, i.e. that a system of N first-order equations leads to a scalar equation of order N, is not always true.)

1.1.4 Continuum limits

In many circumstances, difference equations arise as discretizations of differential equations, such as in numerical analysis. It makes sense in this context to ask how to recover the differential equation from the difference equation. This can be done by taking a **continuum limit** of the difference equation. Even for those cases where we study the difference equation in its own right, its continuum limit can be useful for identifying the equation (e.g. as a physical model) and to get further insight into the behavior of solutions by comparing the continuous and discrete equations.

Before we can compute a continuum limit we must connect y_n and $y(x)$. We usually think of x as generic point on the real line and n as a generic point in the lattice; see

Figure 1.2. They are connected by taking x to be n steps of length δ from some fixed point x_0, i.e. $x = x_0 + \delta n$. Thus we have the identifications

$$y(x) = y(x_0 + \delta n) = y_n, \tag{1.24a}$$

and as a consequence,

$$y_{n+1} = y(x_0 + \delta(n+1)) = y(x+\delta), \text{ etc.} \tag{1.24b}$$

Given a difference equation of the form (1.15) or (1.16), asking for its continuum limit amounts to making an approximation in which a parameter related to the spacing between successive points becomes infinitesimally small. Thus for $y_{n+k} = y(x + k\delta)$, the limit $\delta \to 0$ leads to a continuum limit of the difference equation. Performing the limit involves expanding $y(x + k\delta)$ as a Taylor series

$$y(x + k\delta) = y(x) + k\delta y'(x) + \tfrac{1}{2}(k\delta)^2 y''(x) + \dots$$

taking into account the possibility that the parameters of the equation, and the dependent variable y itself for that matter, may explicitly depend on δ or that they may depend on δ in a hidden way. (The latter case corresponds to the point of view where the parameters are yet to be specified.)

Example 1.1.6 (i) Consider the difference equation $y_{n+1} = \alpha y_n(1 - y_n)$. Letting $y_n = y(x)$ we expand $y_{n+1} = y(x + \delta) = y + \delta y' + \dots$. Then for w, defined by $y(x) = a\delta w(x)$, and with α depending on the expansion parameter δ by $\alpha = 1 + \delta a$, we get

$$\begin{aligned} a\delta w + a\delta^2 w' + \dots &= (1 + \delta a) a\delta w(1 - a\delta w) \\ &= a\delta w + a\delta^2(aw(1 - w)). \end{aligned}$$

Now the first-order terms cancel and equation (1.4) is recovered for w at order δ^2.

(ii) For $\mathrm{dP_I}$ (1.20), the continuum limit is more involved. We need to take $y_n = \frac{\gamma}{6}(1 - 2\delta^2 w(x_0 + n\,\delta))$, $\alpha + \beta n = -\frac{\gamma^2}{36}(3 + 2\delta^4 x)$, $(x = x_0 + n\,\delta)$. Then at order δ^4 we get $w'' = 6\,w^2 + x$, which is (1.6) for $w(x)$.

1.1.5 Functional equations

Another way to look at equations of the type (1.14) is to consider the step-size δ no longer to be fixed, but to consider the equation as a problem posed for arbitrary values of x and δ. This point of view leads to **functional equations**, which are equations that relate the value of a function (or more than one function) at one point to its values at any other point. These are equations of the form

$$\mathcal{F}(f(x), g(x), f(y), \dots, f(x+y), \dots, g(x+y+z), \dots; x, y, z, \dots) = 0\,, \tag{1.25}$$

where x, y, z, ... take all values in a given domain. This is not just a superficial change in perspective; it changes the problem fundamentally from that of a difference equation. In fact, under general assumptions on these functions (such as continuity, differentiability, etc.), the requirement that the equation holds for all values of the arguments of the functions is often sufficient to almost uniquely fix the functions that solve the functional equation. The notion of initial values is irrelevant in this context.

Example 1.1.7 Consider the functional equation

$$\mathcal{F}(f(x), f(y), f(x+y)) \equiv f(x+y) - f(x)f(y) = 0.$$

Evaluating the equation at $y = 0$ we find $f(x) = f(x)f(0)$ and therefore $f(0) = 1$. Let us denote $f'(0) = \alpha \neq 0$. Computing the y derivative and then evaluating the result at $y = 0$ yields $f'(x) = \alpha f(x)$ so that $f = e^{\alpha x}$.

Example 1.1.8 Consider the functional equation

$$\mathcal{F}(f(x), f(y), g(x), g(y), f(x+y)) \equiv f(x+y) - f(x)f(y) + g(x)g(y) = 0.$$

Under the assumption that f, g are both differentiable for all real values of their arguments, and that f is an *even* function of its argument, one can show that either f, g are trivial or $f(x) = \cos(\alpha x)$ and $g(x) = \sin(\alpha x)$. This is left as an exercise.

Many integrable difference equations can also be interpreted as functional equations. An example is provided by the addition formulas for elliptic functions, which are discussed further in Section 2.2.2. The discrete Painlevé equations have also been interpreted as functional equations; see e.g. Gromak and Tsegel'nik (1994).

1.1.6 Delay-difference equations

Equations that involve iterates in one independent variable and derivatives in another are referred to as differential-difference equations. However, there are dynamical processes that depend on finite time-steps and instantaneous changes in the *same* variable. (This happens, for example, in population models where the population growth in one generation may depend on the numbers in a previous generation as well as the birth rate of the current population.) The resulting equations contain iterates and derivatives of a function in the same variable.

When such a model only contains iterates from a previous time-step, the resulting equation is called a *delay*-difference equation, or delay equation for short. In general, equations may contain both backward and forward iterates as well as derivatives in the same variable. For simplicity, we also describe such equations as "delay" equations, but they can also be called ordinary differential-difference equations. As a first example we mention (Quispel et al., 1992)