Hans J. Stetter

Analysis of Discretization Methods for Ordinary Differential Equations

With 12 Figures

Hans J. Stetter
Institut für Numerische Mathematik, Technische Hochschule Wien
A-1040 Wien/Austria

AMS Subject Classifications (1970): 34 A 50, 39 A 10, 65 J 05, 65 L 05, 65 Q 05

ISBN 3-540-06008-1 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-06008-1 Springer-Verlag New York Heidelberg Berlin

Preface and Introduction

Due to the fundamental role of differential equations in science and engineering it has long been a basic task of numerical analysts to generate numerical values of solutions to differential equations. Nearly all approaches to this task involve a "finitization" of the original differential equation problem, usually by a projection into a finite-dimensional space. By far the most popular of these finitization processes consists of a reduction to a difference equation problem for functions which take values only on a grid of argument points. Although some of these *finite-difference methods* have been known for a long time, their wide applicability and great efficiency came to light only with the spread of electronic computers. This in turn strongly stimulated research on the properties and practical use of finite-difference methods.

While the theory of partial differential equations and their discrete analogues is a very hard subject, and progress is consequently slow, the initial value problem for a system of first order ordinary differential equations lends itself so naturally to discretization that hundreds of numerical analysts have felt inspired to invent an ever-increasing number of finite-difference methods for its solution. For about 15 years, there has hardly been an issue of a numerical journal without new results of this kind; but clearly the vast majority of these methods have just been variations of a few basic themes. In this situation, the classical textbook by P. Henrici has served as a lighthouse: it has established a clear framework of concepts and many fundamental results. However, it appears that now—10 years later—a further analysis of those basic themes is due, considering the immense productivity during this period.

It is the aim of this monograph to give such an analysis. This text is *not* an introduction to the use of finite-difference methods; rather it assumes that the reader has a knowledge of the field, preferably including practical experience in the computational solution of differential equations. It has been my intention to make such a reader aware of the structure of the methods which he has used so often and to help him understand their properties. I really wanted—as the title of the book indicates—to *analyze* discretization methods in general, and particular classes of such methods, *as mathematical objects*. This point of view forced me to neglect to a considerable extent the practical aspects of solving differential equations numerically. (Fortunately, a few recent textbooks—noteably Gear [6] and Lambert [3]—help to fill that gap.)

Another restriction—which is not clear from the title—is the strong emphasis on *initial value problems* for ordinary differential equations. Only in the first chapter have I sketched a *general* theory of the mathematical objects which I have called "discretization methods". This part is applicable to a wide variety of procedures which replace an infinitesimal problem by a sequence of finite-dimensional problems. The remainder of the text refers to systems of first order differential equations only. Originally, I had intended to consider both initial and boundary value problems and to develop parallel theories as far as possible (using the concept of stability also in the boundary value problem context); but the manuscript grew too long and I had to restrict myself to initial value problems.

Within these severe restrictions I hope to have covered a good deal of material. The presentation attempts to make the book largely self-contained. More important, I have used a consistent notation and terminology throughout the entire volume, as far as feasible. At a few places, this has perhaps led to a "twist" in presentation; but I felt it worthwhile to adhere to a common frame of reference, since the exposition of such a frame has been one of the main motivations for writing this book. The numerous examples are never meant as suggestions for practical computation but as illustrations for the theoretical development.

As mentioned above, Chapter 1 deals with the *general structure of discretization methods*. Particular emphasis has been placed on the theory of asymptotic expansions and their application. It appears that they may also be used with various other approximation procedures for infinitesimal problems, procedures which are normally not considered as discretization methods but which may be fitted into our theory. (To facilitate such interpretations I have used a sequence of integers as discretization parameters in place of the stepsizes.) A section on "error analysis" attempts to outline several important aspects of error evaluations in a general manner.

Chapter 2 is devoted to special features of discretization methods for *initial value problems* for ordinary differential equations. The first two sections present mainly background material for an analysis which involves the well-known limit process $h \to 0$: a fixed finite interval of integration is subdivided by grids with finer and finer steps. The last section of Chapter 2, however, deals with a different limit process: the interval of integration is extended farther and farther while the stepsize remains fixed. It is to be expected that a theory of this limit process "$T \to \infty$"

may serve as a basis for an understanding of the behavior of discretization methods on long intervals, with relatively large steps, in the same way that the $h \to 0$ theory has proved to be a good model for the case of small steps on relatively short intervals.

The remainder of the treatise analyzes particular classes of methods. The two most distinctive features of such methods, the "multistage" and the "multistep" features, are treated separately at first. Thus, in Chapter 3, we consider *one-step methods* which "remember" only the value of the approximate solution at the previous gridpoint but use this value in a computational process which runs through a number of stages, with re-substitutions in each stage (this process may also be strongly implicit). Butcher's abstract algebraic theory of such processes permits a rather elegant approach to various structural investigations, such as questions of equivalence, symmetry, etc. The theory of asymptotic expansions provides a firm basis for the analysis of both the local and the global discretization error. The last section of Chapter 3 is devoted to the $T \to \infty$ limit process; here a number of results are derived and directions for further research are outlined. A number of intuitive conjectures are shown to be false without further assumptions.

In Chapter 4, we analyze discretization methods which use values of the approximate solution at several previous gridpoints but do not permit re-substitutions into the differential equations, i.e., *linear multistep methods*. After an exposition of the wellknown accuracy and stability theory for such methods, we outline a theory of "cyclic" multistep methods which employ different k-step procedures in a cyclic fashion as the computation moves along (as suggested by Donelson and Hansen). The investigations of Gragg on asymptotic expansions for linear multistep methods have been presented and supplemented by a general analysis of the symmetric case. Again the last section of the chapter has been reserved for the $T \to \infty$ theory.

After these preparations, we discuss *general multistage multistep methods* in Chapter 5. This class of methods is so wide that we have concentrated on the analysis of important subclasses, such as predictor-corrector methods, hybrid methods, and others. In agreement with the intentions of the book it appeared relevant to point out typical restrictions which permit farther-reaching assertions. New results concern the principal error function of general predictor-corrector methods and an extension of the concept of effective order to multistep methods. Cyclic methods are taken up again in a more general context.

Methods which explicitly use *derivatives of the right hand side* of the differential equation have been analyzed in the first section of Chapter 6; this class includes power series methods, Lie series methods, Runge-

Kutta-Fehlberg methods and the like. The *Nordsieck-Gear multivalue approach* did not fit into the pattern of Chapter 5; I have tried to give a consistent account of its theory. Last but not least I have included an analysis of *extrapolation methods* which have proved to be most powerful in practical computation; a particular effort was devoted to a clarification of the stability properties of such methods.

Naturally it has not been possible to achieve full coverage of the field. In particular, it seemed premature to force a rigid terminology upon developments which may not yet have found their final form. For this reason I have not included an account of the "variable coefficient" approach of Brunner [1], Lambert [1], and others, although it appears to be very promising. Also, the use of spline functions was omitted together with that of various non-polynomial local approximants (see, e.g., Gautschi [2], Lambert and Shaw [1], Nickel and Rieder [1]). Finally, I did not dare to propose a refined theory of "stiff methods" at this time, although I hope that my theory of strong exponential stability will provide one of the bases for such a theory.

A few remarks concerning the *bibliography* are necessary. I had originally planned to include a comprehensive bibliography of more than a thousand entries. However, the use of such a bibliography in the text would necessarily have been restricted to endless enumerations and thus not have been very helpful. Therefore, I decided to restrict my references essentially to those publications whose results I either quoted without proof or whose arguments I followed exceedingly closely. Thus the bibliography in this book is quite meager and contains—besides a number of "classics"—only a strangely biased selection from the relevant literature. I sincerely hope that the many colleagues whose important contributions I have not quoted will appreciate this reasoning and be reassured by the discovery of how many other important papers have not been quoted. Clearly, the ideas of innumerable papers have influenced the contents of this book though they are not explicitly mentioned.

Similarly, my thoughts about the subject could only mature through personal contact and discussion with many of the prominent numerical analysts all over the world. It would be unfair to enumerate a few of the many names: all colleagues whom I had a chance to ask about their views on discretization methods are silent co-authors of this text, and I wish to thank them accordingly.

The manuscript was prepared between the summers of 1969 and 1971. It could not have been written had I not been able to spend the year 1969/70—without much teaching obligation—at the University of California at San Diego under a National Science Foundation Senior Fellowship. More than two thirds of the manuscript originated during

that pleasant year in Southern California for which I have to thank above all the chairman of the UCSD Mathematics Department, my friend Helmut Roehrl. The remainder of the manuscript I somehow managed to complete at weekends, evenings, etc. during the following year, despite my many duties in Vienna, and during a "working vacation" with my friends at the University of Dundee.

The typing of Chapters 1 through 4 was done by Lillian Johnson at UCSD, that of Chapters 5 and 6 by Christine Grill at the Technical University of Vienna. Miss Grill also re-typed and rere-typed many pages and did a lot of editorial work on the manuscript. Further editorial work and proofreading was done by a number of my young colleagues at Vienna, principally by W. Baron, R. Frank, and R. Mischak. To all these I wish to express my gratitude.

Another word of thanks is due to my friend and colleague Jack Lambert at Dundee. Although he was hard-pressed by the preparation of a manuscript himself, he kindly read the entire text and corrected many of my blunders with regard to the use of the English language. Further improvements in style were suggested by Kenneth Wickwire. (I ventured to write this book in English because it will be more easily read in poor English than in good German by 90% of my intended readers.)

The co-operation I have received from Springer-Verlag has been most pleasing. The type-setting and the production of the book are of the usual high Springer quality.

Finally, I wish to give praise to my dear wife Christine who suffered, without much complaint, severe restrictions of our home life during more than 2 years. At the same time, her loving care and reassurance helped me to get on with the work. If there is just one person in the world who will rejoice in the completion of this book it will be she.

Vienna, June 72

Hans J. Stetter

Table of Contents

Chapter 1 General Discretization Methods 1

1.1. Basic Definitions 1

1.1.1 Discretization Methods 1
1.1.2 Consistency 5
1.1.3 Convergence 7
1.1.4 Stability 9

1.2 Results Concerning Stability 11

1.2.1 Existence of the Solution of the Discretization 11
1.2.2 The Basic Convergence Theorem 13
1.2.3 Linearization 15
1.2.4 Stability of Neighboring Discretizations 18

1.3 Asymptotic Expansions of the Discretization Errors 21

1.3.1 Asymptotic Expansion of the Local Discretization Error 21
1.3.2 Asymptotic Expansion of the Global Discretization Error 25
1.3.3 Asymptotic Expansions in Even Powers of n 29
1.3.4 The Principal Error Terms 30

1.4 Applications of Asymptotic Expansions 33

1.4.1 Richardson Extrapolation 33
1.4.2 Linear Extrapolation 37
1.4.3 Rational Extrapolation 42
1.4.4 Difference Correction 44

1.5 Error Analysis 49

1.5.1 Computing Error 49
1.5.2 Error Estimates 51
1.5.3 Strong Stability 54
1.5.4 Richardson-extrapolation and Error Estimation 56
1.5.5 Statistical Analysis of Round-off Errors 58

1.6 Practical Aspects 61

Chapter 2 Forward Step Methods 63

2.1 Preliminaries 63

2.1.1 Initial Value Problems for Ordinary Differential Equations . . . 63
2.1.2 Grids 65
2.1.3 Characterization of Forward Step Methods 67
2.1.4 Restricting the Interval 69
2.1.5 Notation 72

2.2 The Meaning of Consistency, Convergence, and Stability with Forward Step Methods 74

2.2.1 Our Choice of Norms in E_n and E_n^0 74
2.2.2 Other Definitions of Consistency and Convergence 76
2.2.3 Other Definitions of Stability 79
2.2.4 Spijker's Norm for E_n^0 81
2.2.5 Stability of Neighboring Discretizations 84

2.3 Strong Stability of f.s.m. 87

2.3.1 Perturbation of IVP 1 87
2.3.2 Discretizations of $\{\text{IVP } 1\}_T$ 92
2.3.3 Exponential Stability for Difference Equations on $[0,\infty)$ 94
2.3.4 Exponential Stability of Neighboring Discretizations 98
2.3.5 Strong Exponential Stability 101
2.3.6 Stability Regions 103
2.3.7 Stiff Systems of Differential Equations 104

Chapter 3 Runge-Kutta Methods 107

3.1 RK-procedures 107

3.1.1 Characterization 107
3.1.2 Local Solution and Increment Function 110
3.1.3 Elementary Differentials 111
3.1.4 The Expansion of the Local Solution 115
3.1.5 The Exact Increment Function 117

3.2 The Group of RK-schemes 120

3.2.1 RK-schemes 120
3.2.2 Inverses of RK-schemes 125
3.2.3 Equivalent Generating Matrices 127
3.2.4 Explicit and Implicit RK-schemes 131
3.2.5 Symmetric RK-procedures 134

3.3 RK-methods and Their Orders 135

3.3.1 RK-methods 135
3.3.2 The Order of Consistency 137

3.3.3 Construction of High-order RK-procedures 141
3.3.4 Attainable Order of m-stage RK-procedures 144
3.3.5 Effective Order of RK-schemes 148

3.4 Analysis of the Discretization Error 150

3.4.1 The Principal Error Function 150
3.4.2 Asymptotic Expansion of the Discretization Error 153
3.4.3 The Principal Term of the Global Discretization Error 157
3.4.4 Estimation of the Local Discretization Error 161

3.5 Strong Stability of RK-methods 165

3.5.1 Strong Stability for Sufficiently Large n 165
3.5.2 Strong Stability for Arbitrary n 170
3.5.3 Stability Regions of RK-methods 174
3.5.4 Use of Stability Regions for General $\{IVP\ 1\}_s$ 178
3.5.5 Suggestion for a General Approach 182

Chapter 4 Linear Multistep Methods 185

4.1 Linear k-step Schemes 185

4.1.1 Characterization 185
4.1.2 The Order of Linear k-step Schemes 191
4.1.3 Construction of Linear k-step Schemes of High Order 195

4.2 Uniform Linear k-step Methods 199

4.2.1 Characterization, Consistency 199
4.2.2 Auxiliary Results 203
4.2.3 Stability of Uniform Linear k-step Methods 206
4.2.4 Convergence 210
4.2.5 Highest Obtainable Orders of Convergence 214

4.3 Cyclic Linear k-step Methods 216

4.3.1 Stability of Cyclic Linear k-step Methods 216
4.3.2 The Auxiliary Method 221
4.3.3 Attainable Order of Cyclic Linear Multistep Methods 225

4.4 Asymptotic Expansions 228

4.4.1 The Local Discretization Error. 228
4.4.2 Asymptotic Expansion of the Global Discretization Error. Preparations 232
4.4.3 The Case of No Extraneous Essential Zeros 234
4.4.4 The Case of Extraneous Essential Zeros 240

4.5 Further Analysis of the Discretization Error 245

4.5.1 Weak Stability 245
4.5.2 Smoothing 248

4.5.3 Symmetric Linear k-step Schemes . . . 250
4.5.4 Asymptotic Expansions in Powers of h^2 . . . 256
4.5.5 Estimation of the Discretization Error . . . 260

4.6 Strong Stability of Linear Multistep Methods . . . 263

4.6.1 Strong Stability for Sufficiently Large n . . . 263
4.6.2 Stability Regions of Linear Multistep Methods . . . 266
4.6.3 Strong Stability for Arbitrary n . . . 270

Chapter 5 Multistage Multistep Methods . . . 272

5.1 General Analysis . . . 272

5.1.1 A General Class of Multistage Multistep Procedures . . . 272
5.1.2 Simple m-stage k-step Methods . . . 275
5.1.3 Stability and Convergence of Simple m-stage k-step Methods . . . 278

5.2 Predictor-corrector Methods . . . 282

5.2.1 Characterization, Subclasses . . . 282
5.2.2 Stability and Order of Predictor-corrector Methods . . . 284
5.2.3 Analysis of the Discretization Error . . . 290
5.2.4 Estimation of the Local Discretization Error . . . 294
5.2.5 Estimation of the Global Discretization Error . . . 297

5.3 Predictor-corrector Methods with Off-step Points . . . 300

5.3.1 Characterization . . . 300
5.3.2 Determination of the Coefficients and Attainable Order . . . 302
5.3.3 Stability of High Order PC-methods with Off-step Points . . . 306

5.4 Cyclic Forward Step Methods . . . 308

5.4.1 Characterization . . . 308
5.4.2 Stability and Error Propagation . . . 311
5.4.3 Primitive m-cyclic k-step Methods . . . 316
5.4.4 General Straight m-cyclic k-step Methods . . . 322

5.5 Strong Stability . . . 324

5.5.1 Characteristic Polynomial, Stability Regions . . . 324
5.5.2 Stability Regions of PC-methods . . . 326
5.5.3 Stability Regions of Cyclic Methods . . . 329

Chapter 6 Other Discretization Methods for IVP 1 . . . 332

6.1 Discretization Methods with Derivatives of f . . . 332

6.1.1 Recursive Computation of Higher Derivatives of the Local Solution . . . 332
6.1.2 Power Series Methods . . . 335

6.1.3 The Perturbation Theory of Groebner-Knapp-Wanner 336
6.1.4 Groebner-Knapp-Wanner Methods 339
6.1.5 Runge-Kutta-Fehlberg Methods 343
6.1.6 Multistep Methods with Higher Derivatives 347

6.2 General Multi-value Methods 349

6.2.1 Nordsieck's Approach 349
6.2.2 Nordsieck Predictor-corrector Methods 354
6.2.3 Equivalence of Generalized Nordsieck Methods 357
6.2.4 Appraisal of Nordsieck Methods 361

6.3 Extrapolation Methods 362

6.3.1 The Structure of an Extrapolation Method 362
6.3.2 Gragg's Method 364
6.3.3 Strong Stability of $\mathfrak{M}_G$ 369
6.3.4 The Gragg-Bulirsch-Stoer Extrapolation Method 372
6.3.5 Extrapolation Methods for Stiff Systems 375

Bibliography 380

Subject Index 385

Chapter 1

General Discretization Methods

In this introductory chapter we consider general aspects of discretization methods. Much of the theory is applicable not only to standard discretization methods for ordinary differential equations (both initial and boundary value problems) but also to a great variety of other numerical methods as indicated in the Preface (see also the end of Section 1.1.1). It should be emphasized that this is also true for the material on asymptotic expansions and their applications, although we have not elaborated on this. The chapter is concluded by a few remarks on the practical aspects of "solving" ordinary differential equations by discretization methods.

1.1 Basic Definitions

1.1.1 Discretization Methods

Def. 1.1.1. Throughout, the problem whose solution is to be approximated by a discretization method will be called the *original problem*. It is specified by a triple $\{E, E^0, F\}$ where E and E^0 are Banach-spaces and $F: E \to E^0$, with 0 in the range of F. A *true solution* of the original problem is an element $z \in E$ such that

(1.1.1) $$Fz = 0.$$

We will always assume that a true solution of the original problem *exists* and is *unique* (the uniqueness may often be obtained by suitably restricting the domain of F); we will not consider existence and uniqueness questions for the original problem.

Since we are dealing with the numerical solution of (real) ordinary differential equations, E und E^0 will normally be spaces of continuous functions from an interval in $\mathbb{R}$ to $\mathbb{R}^s$ and the mapping F will be defined by a differential operator. Initial and/or boundary conditions are included in the definitions of E, F or the domain of F as seems appropriate.

Example.

$$E = C^{(1)}[0,1], \quad \text{with} \quad \|y\|_E := \max_{t\in[0,1]} |y(t)|;$$

$$E^0 = \mathbb{R} \times C[0,1], \quad \text{with} \quad \left\|\begin{pmatrix} d_0 \\ d \end{pmatrix}\right\|_{E^0} := |d_0| + \max_{t\in[0,1]} |d(t)|;$$

(1.1.2) $$Fy := \begin{pmatrix} y(0) - z_0 \\ y'(t) - f(y(t)) \end{pmatrix} \in E^0 \quad \text{for } y \in E,$$

where $z_0 \in \mathbb{R}$, $f \in C(\mathbb{R} \to \mathbb{R})$ (f Lipschitz-continuous) are the data of the problem.

The true solution of the original problem thus specified (viz. the solution of the differential equation $y' = f(y)$, with initial condition $y(0) = z_0$) exists and is unique.

The basic idea of a discretization method is to replace the original problem by an infinite sequence of finite-dimensional problems each of which can be solved "constructively" in the sense of numerical mathematics. The replacement is to be such that the solutions of these finite-dimensional problems approximate, in a sense to be defined, the true solution z of the original problem better and better the further one proceeds in the sequence. Thus, one can obtain an arbitrarily good approximation to z by taking the solution of a suitably chosen problem in the sequence.

Of course, it will generally not be possible to obtain the solution of this finite-dimensional problem with arbitrary accuracy on a given computing tool and with a given computational effort (see Section 1.5). Nevertheless, the construction of discretization methods commonly follows the reasoning in the above paragraph.

Def. 1.1.2. *A discretization method* $\mathfrak{M}$ *(applicable to a given original problem* $\mathfrak{P} = \{E, E^0, F\}$*)* consists of an infinite sequence of quintuples $\{E_n, E_n^0, \Delta_n, \Delta_n^0, \varphi_n\}_{n\in\mathbb{N}'}$ where
E_n and E_n^0 are finite-dimensional Banach spaces;
Δ_n: $E \to E_n$ and Δ_n^0: $E^0 \to E_n^0$ are linear mappings with

$$\lim_{n\to\infty} \|\Delta_n y\|_{E_n} = \|y\|_E \quad \text{for each fixed } y \in E,$$

$$\lim_{n\to\infty} \|\Delta_n^0 d\|_{E_n^0} = \|d\|_{E^0} \quad \text{for each fixed } d \in E^0;$$

and φ_n: $(E \to E^0) \to (E_n \to E_n^0)$, with F in the domain of all φ_n.

$\mathbb{N}'$ is an infinite subset of $\mathbb{N}$.

Def. 1.1.3. A *discretization* $\mathfrak{D}$ is an infinite sequence of triples $\{E_n, E_n^0, F_n\}_{n\in\mathbb{N}''}$ where E_n and E_n^0 are finite-dimensional Banach spaces; F_n: $E_n \to E_n^0$.

A *solution of the discretization* $\mathfrak{D}$ is a sequence $\{\zeta_n\}_{n\in\mathbb{N}''}$, $\zeta_n \in E_n$, such that

(1.1.3) $$F_n \zeta_n = 0, \qquad n \in \mathbb{N}''.$$

$\mathbb{N}''$ is again an infinite subset of $\mathbb{N}$.

Def. 1.1.4. The discretization $\mathfrak{D}=\{\bar{E}_n, \bar{E}_n^0, \bar{F}_n\}_{n\in\mathbb{N}''}$ is called the *discretization of the original problem* $\mathfrak{P}=\{E, E^0, F\}$ *generated by the discretization method* $\mathfrak{M}=\{E_n, E_n^0, \Delta_n, \Delta_n^0, \varphi_n\}_{n\in\mathbb{N}'}$ if $\mathfrak{M}$ is applicable to $\mathfrak{P}$ and

$$\mathbb{N}''\subset\mathbb{N}',$$
$$\left.\begin{array}{ll}\bar{E}_n=E_n, & \bar{E}_n^0=E_n^0\\ \bar{F}_n=\varphi_n(F)\end{array}\right\}\quad\text{for } n\in\mathbb{N}''.$$

In this case $\mathfrak{D}$ is denoted by $\mathfrak{M}(\mathfrak{P})$.

Remark. In the following, we will assume throughout (without loss of generality) that $\mathbb{N}''=\mathbb{N}'$ and that the sequences $\{E_n\}$ and $\{E_n^0\}$ of $\mathfrak{M}$ and $\mathfrak{M}(\mathfrak{P})$ resp. are identical.

Furthermore, we will always assume that the dimensions of E_n and E_n^0 are the same. This is a trivial necessary condition for the existence of a unique ζ_n which satisfies (1.1.3).

Example. Consider the original problem of the Example following Def. 1.1.1. The well-known Euler method (polygon method) may be characterized as a discretization method applicable to this problem as follows:

For $n\in\mathbb{N}'=\mathbb{N}$, let $\mathbb{G}_n:=\{\nu/n, \nu=0(1)n\}$ and

$$E_n=(\mathbb{G}_n\to\mathbb{R}),\quad\text{with } \|\eta\|_{E_n}:=\max_{\nu=0(1)n}\left|\eta\left(\frac{\nu}{n}\right)\right|;$$

$$E_n^0=(\mathbb{G}_n\to\mathbb{R}),\quad\text{with } \|\delta\|_{E_n^0}:=|\delta(0)|+\max_{\nu=1(1)n}\left|\delta\left(\frac{\nu}{n}\right)\right|;$$

$$(\Delta_n y)\left(\frac{\nu}{n}\right)=y\left(\frac{\nu}{n}\right)\quad\text{for } y\in E;$$

$$(\Delta_n^0 d)\left(\frac{\nu}{n}\right)=\begin{cases} d_0, & \nu=0,\\ d\left(\frac{\nu-1}{n}\right), & \nu=1(1)n,\end{cases}\quad\text{for } d=\begin{pmatrix} d_0\\ d(t)\end{pmatrix}\in E^0;$$

$$(1.1.4)\qquad [\varphi_n(F)\eta]\left(\frac{\nu}{n}\right)=\begin{cases}\eta(0)-z_0, & \nu=0,\\ \dfrac{\eta\left(\frac{\nu}{n}\right)-\eta\left(\frac{\nu-1}{n}\right)}{1/n}-f\left(\eta\left(\frac{\nu-1}{n}\right)\right), & \nu=1(1)n.\end{cases}$$

The existence of a unique solution sequence $\{\zeta_n\}$ of the discretization is trivial since the values of the ζ_n are defined recursively.

A discretization method $\mathfrak{M}$ applicable to the original problem $\mathfrak{P}$ not only generates a discretization $\mathfrak{D}=\mathfrak{M}(\mathfrak{P})$ of $\mathfrak{P}$ but establishes at the same time, through its mappings Δ_n and Δ_n^0, relations between the spaces of $\mathfrak{P}$ and its discretization $\mathfrak{M}(\mathfrak{P})$ which permit an interpretation of $\mathfrak{M}(\mathfrak{P})$ and its solution $\{\zeta_n\}$ as an approximation of $\mathfrak{P}$ and its true solution z. This will be elaborated in the following sections.

The relations between the spaces and mappings involved in the original problem and its discretization can be visualized by the following diagram:

$$\begin{array}{ccccl} E & \xrightarrow{\quad F \quad} & E^0 & & \text{original problem} \\ \Big\downarrow{\scriptstyle \Delta_n} & & \Big\downarrow{\scriptstyle \Delta_n^0} & \Big\downarrow{\scriptstyle \varphi_n} & \\ E_n & \xrightarrow{\quad F_n \quad} & E_n^0 \cdot & & \text{discretization,} \quad n \in \mathbb{N}'' \end{array}$$

Fig. 1.1. Relation between the elements of a discretization method

The existence and uniqueness of a true solution z for the original problem does not automatically imply the existence and uniqueness of ζ_n which satisfy (1.1.3). Therefore, we will explicitly establish the unique solvability of $F_n\zeta_n=0$ for many of the discretization methods under discussion. Furthermore, we will prove a result (Theorem 1.2.3) which guarantees the existence of a unique ζ_n for all sufficiently large n for reasonable discretization methods. For many classes of problems and discretization methods, the unique solvability of (1.1.3) is trivial.

As in the example following Def. 1.1.4, in our applications the spaces E_n and E_n^0 will normally be spaces of functions mapping discrete and finite subsets of intervals (so-called "grids") into $\mathbb{R}^s$. The mappings Δ_n and Δ_n^0 will "discretize" the continuous functions of E and E^0 into grid functions. The mappings $F_n=\varphi_n(F)$ will be defined by difference operators relative to the grids.

Def. 1.1.1—1.1.4 also admit essentially different interpretations. Consider, e.g., the original problem of the previous example, with all functions arbitrarily many times differentiable, and choose

$$E_n = \{\text{polynomials of degree } \le n\}:$$

$$\Delta_n y = \text{polynomial in } E_n \text{ obtained by the truncation of the formal power series of } y \in E\,;$$

$$E_n^0 = \mathbb{R} \times E_{n-1}\,;$$

$$\Delta_n^0 \begin{pmatrix} d_0 \\ d(t) \end{pmatrix} = \begin{pmatrix} d_0 \\ \Delta_{n-1} d(t) \end{pmatrix};$$

$$F_n \eta = \begin{pmatrix} \eta(0) - z_0 \\ \eta'(t) - \Delta_{n-1} f(\eta(t)) \end{pmatrix}, \quad \text{for } \eta \in E_n\,.$$

For the generated discretization, (1.1.3) consists of a system of $n+1$ equations for the $n+1$ coefficients of ζ_n which is considered as an approximation to $\Delta_n z$.

A number of the general concepts and results also make sense for a variety of such unorthodox "discretizations" (e.g. Ritz and Galerkin's methods for boundary value problems). Also it is immediately clear that most results apply to standard discretizations of partial differential equations, at least when the side conditions are not too complicated. However, in the more specialized parts of this treatise we will restrict ourselves exclusively to genuine discretizations of ordinary differential equation problems.

1.1.2 Consistency

To be useful as a tool for obtaining an approximation to the solution z of (1.1.1), a discretization method should generate a discretization which approximates the original problem in the following sense:

Def. 1.1.5. A discretization method $\mathfrak{M}=\{E_n, E_n^0, \Delta_n, \Delta_n^0, \varphi_n\}_{n\in\mathbb{N}'}$ applicable to the original problem $\mathfrak{P}=\{E, E^0, F\}$ is called *consistent with* $\mathfrak{P}$ *at* $y\in E$ if y is in the domain of F and of $\varphi_n(F)\Delta_n$, $n\in\mathbb{N}'$, and

(1.1.5)[1]
$$\lim_{\substack{n\to\infty\\ n\in\mathbb{N}'}} \|\varphi_n(F)\Delta_n y-\Delta_n^0 F y\|_{E_n^0}=0\,.$$

$\mathfrak{M}$ is called *consistent with* $\mathfrak{P}$ if it is consistent with $\mathfrak{P}$ at each $y\in E$.

If $\mathfrak{M}$ is consistent with $\mathfrak{P}$ (at y) the discretization $\mathfrak{M}(\mathfrak{P})$ is also called consistent with $\mathfrak{P}$ (at y).

Def. 1.1.6. In the situation of Def. 1.1.5, $\mathfrak{M}$ and $\mathfrak{M}(\mathfrak{P})$ are called *consistent* with $\mathfrak{P}$ *of order p* at y if

(1.1.6)
$$\|\varphi_n(F)\Delta_n y-\Delta_n^0 F y\|_{E_n^0}=O(n^{-p}) \quad \text{as } n\to\infty\,.$$

In connection with the order of consistency, the reference "at z" may be omitted.

Example. As in Section 1.1.1. For any $y\in C^{(1)}[0,1]$, we have from (1.1.2) and (1.1.4).

$$[\varphi_n(F)\Delta_n y-\Delta_n^0 F y]\left(\frac{\nu}{n}\right)=\begin{cases}(y(0)-z_0)-(y(0)-z_0), & \nu=0,\\[2mm] \dfrac{y\left(\frac{\nu}{n}\right)-y\left(\frac{\nu-1}{n}\right)}{1/n}-f\left(y\left(\frac{\nu-1}{n}\right)\right) & \\ \qquad -\left[y'\left(\frac{\nu-1}{n}\right)-f\left(y\left(\frac{\nu-1}{n}\right)\right)\right], & \nu=1(1)n.\end{cases}$$
$$=\begin{cases}0, & \nu=0,\\ y'(\tilde{t}_\nu)-y'\left(\dfrac{\nu-1}{n}\right), & \nu=1(1)n, \quad \text{with } \tilde{t}_\nu\in\left(\dfrac{\nu-1}{n},\dfrac{\nu}{n}\right).\end{cases}$$

---

[1] *All* limit processes $n\to\infty$ carry the restriction $n\in\mathbb{N}'$ which we will omit from now on.