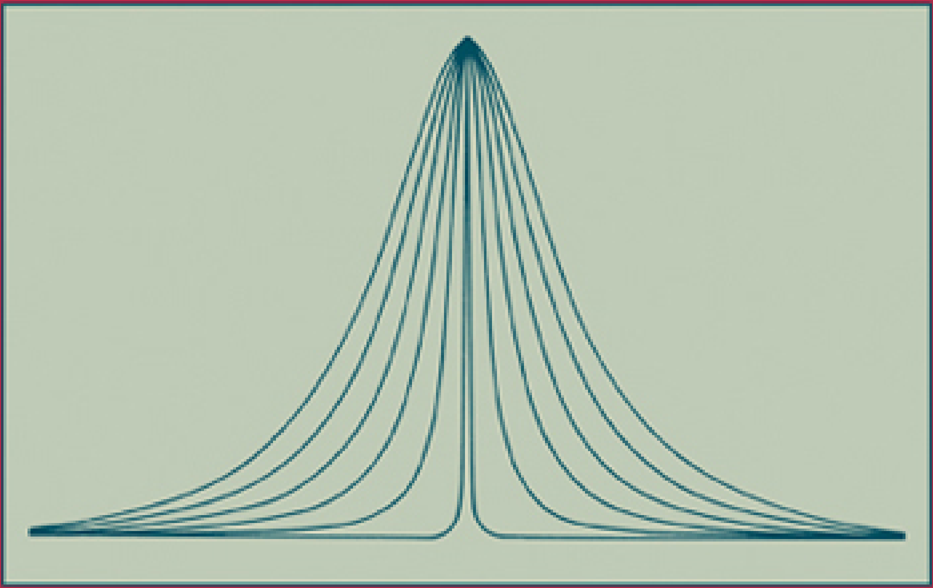
ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

RATIONAL APPROXIMATION OF REAL FUNCTIONS

P. P. PETRUSHEV & V. A. POPOV



CAMBRIDGE

more information - www.cambridge.org/9780521331074

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

EDITED BY G.-C. ROTA

Volume 28

Rational approximation of real functions

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Rational approximation of real functions

P.P. PETRUSHEV, V.A. POPOV

Bulgarian Academy of Sciences

CAMBRIDGE UNIVERSITY PRESS

Cambridge

New York New Rochelle Melbourne Sydney

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
Sao Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521331074

First published 1987

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Petrushev, P.P. (Penco Petrov), 1949-
Rational approximation of real functions.

(Encyclopedia of mathematics and its applications; v. 28)
Bibliography
Includes index.
1. Functions of real variables. 2. Approximation theory.
I. Popov, Vasil A. (Vasil Atanasov) II. Title.
III. Series
QA331.5.P48 1987 515.8 86-34331

ISBN 978-0-521-33107-4 Hardback

CONTENTS

PREFACE

Rational functions are a classical tool for approximation. They turn out to be a more convenient tool for approximation in many cases than polynomials which explains the constant increase of interest in them. On the other hand rational functions are a nonlinear approximation tool and they possess some intrinsic peculiarities creating a lot of difficulties in their investigation. After the classcial results of Zolotarjov from the end of the last century substantial progress was achieved in 1964 when D. Newman showed that $|x|$ is uniformly approximated by rational functions much better than by algebraic polynomials. Newman's result stimulated the appearance of many substantial results in the field of rational real approximations.

Our aim in this book is to present the basic achievements in rational real approximations. Nevertheless, for the sake of completeness we have included some results referring to the field of complex rational approximations in Chapters 6 and 12. Also, in order to stress some peculiarities of rational approximations we have included for comparison some classical and more recent results from the linear theory of approximation. On the other hand, since rational approximations are closely connected with spline approximations, we have included as well some results concerning spline approximations.

As usual the specific topics selected reflect the authors' interests and preferences.

We now sketch briefly the contents of the book. Chapters 1 and 3 contain some basic facts concerning linear approximation theory. A basic problem in approximation theory is to find complete direct and converse theorems. In our opinion the most natural way to obtain such theorems in linear and nonlinear approximations is to prove pairs of adjusted inequalities of Jackson and Bernstein type and then to characterize the corresponding approximations by the K-functional of Peetre. This main viewpoint is given and illustrated at the end of Chapter 3 and next applied to the spline approximation in Chapter 7.

Chapter 2 is devoted to the study of the qualitative theory of rational approximation such as the existence, the uniqueness and the characterization problems, the problem of continuity of metric projection and numerical methods.

The heart of the book is contained in Chapters 4 to 11. Chapter 4 presents the uniform rational approximation of some important functions such as $|x|$, $\sqrt{x}$, e^x. In Chapter 5 the uniform rational approximation of a number of classes is considered. The exact orders of approximation are established. The basic methods for rational approximation are given. In Chapter 6 some converse theorems for rational uniform approximation are proved. In Chapter 7 complete direct and converse theorems for the spline approximation in L_p, C, BMO are proved using Besov spaces. Chapter 8 investigates the relations between the rational and spline approximations. Chapter 9 deals with rational approximation in Hausdorff metric. A characteristic particularity of rational approximation is the appearance of the so-called 'o small' effect in the order of rational approximation of each individual function of some function classes. This phenomenon is investigated and characterized for some function classes in Chapter 10. The exactness of the proved estimates is established and discussed in Chapter 11.

Chapter 12 considers some special problems, connected with Padé approximants – some of the so-called direct and converse problems for convergence of the rows and diagonal of the Padé-table. Finally some numerical results and graphs are presented in the Appendix.

ACKNOWLEDGEMENTS

First of all we owe thanks to Professor G.G. Lorentz for the suggestion to write this book. We are grateful to Academician Bl. Sendov for constant attention and support. In the course of preparing the manuscript we enjoyed the helpful attention of many colleagues. We would especially like to mention A. Andreev and P. Marinov for their aid in numerical methods and calculations, K. Ivanov for the algebraical polynomial approximation in Chapter 3, R. Kovacheva for Padé approximations and E. Moskona for help in writing Chapter 6. Our colleagues A. Andreev, D. Drjanov, K. Ivanov, R. Kovacheva, R. Maleev, Sv.Markov, E. Sendova, S. Tashev read parts of the manuscript and we appreciate their remarks.

We owe special thanks to Professor D. Braess who submitted to us in manuscript the chapter on the rational approximation of his unpublished book on nonlinear approximation. We would like to express our gratitude to all our colleagues who were so kind as to give us additional information for the manuscript: J.-E. Anderson, B. Brosovski, Yu.A. Brudnyi, A.P. Bulanov, R. DeVore, E.P. Dolženko, T. Ganelius, A.A. Gonchar, J. Karlsson, G. López, A.A. Pekarskii, J. Peetre, V.N. Russak, E.A. Sevastijanov, K. Scherer, J. Szabados.

1

Qualitative theory of linear approximation

We shall begin with a short survey of the basic results related to linear approximations (i.e. approximation by means of linear subspaces) so that one can feel better the peculiarities, the advantages as well as some shortcomings of the rational approximation. In this chapter we shall consider the problems of existence, uniqueness and characterization of the best approximation (best polynomial approximation). At the end of the chapter we shall consider also numerical algorithms for finding the best uniform polynomial approximation.

1.1 Approximation in normed linear spaces

Let X be a normed linear space. Recall that X is said to be a normed linear space if:

(i) *X is a linear space, i.e. for its elements sum, and product with real numbers, are defined so that the standard axioms of commutativity and associativity are satisfied;*

(ii) *X is a normed space, i.e. to each $x \in X$ there corresponds a nonnegative real number $\|x\|$ satisfying the axioms*

(a) $\|x\| \geqslant 0$, $\|x\| = 0$ *iff* $x = 0$,

(b) $\|\lambda x\| = |\lambda| \, \|x\|$, λ *a real number,*

(c) $\|x + y\| \leqslant \|x\| + \|y\|$ (the triangle inequality).

Let $\{\varphi_i\}_{i=1}^n$ be a system of n linearly independent elements of X. Let us consider the linear subspace of $X: G = \{\varphi: \varphi = \sum_{i=1}^n a_i\varphi_i,\ a_i$ real numbers$\}$, generated by the system $\{\varphi_i\}_{i=1}^n$. For each element $f \in X$ we denote by $E_G(f)$ the best approximation to f by means of elements of G:

$$E_G(f) = \inf\{\|f - \varphi\| : \varphi \in G\}. \tag{1}$$

The following basic problems (basic not only for linear approximation theory, but for the theory of approximation in general) arise.

(i) *Existence problem: does an element* $\varphi \in G$ *of best approximation for* $f \in X$ *exist, i.e. is there* $\varphi = \varphi(f) \in G$ *such that*

$$\| f - \varphi(f) \| = E_G(f)?$$

(ii) *Uniqueness problem: if there exists an element of best approximation for* $f \in X$*, is it unique?*

(iii) *Characterization problem: in the case where the element of best approximation for* $f \in X$ *exists and is unique, can we characterize it in some way?*

(iv) *Can we estimate how big* $E_G(f)$ *is?*

(v) *Numerical methods: assuming that we know that the answer to the first two (or three) problems is positive, how can we find* $\varphi(f)$ *in practice?*

The whole theory of approximation represents full or partial (for the present, unfortunately) answers to the above problems when we approximate different classes of functions in different normed linear spaces (or, more generally, in metric spaces) with respect to different approximation tools (e.g. algebraic polynomials, trigonometric polynomials, rational functions, spline functions, linear combinations of exponential functions).

In the case of approximation in a normed linear space by a finite dimensional subspace we can give a positive answer to the first question. More precisely the following theorem holds.

Theorem 1.1 (Existence theorem). *Let G be a finite dimensional subspace of the normed linear space X. For every* $f \in X$ *there is an element of best approximation in G.*

Proof. The proof of this theorem is based on the following well-known fundamental property of finite dimensional normed spaces: every bounded closed subset in a finite dimensional normed linear space is compact. The idea of the proof is to show that the inf in (1) may be taken over a compact subset of G.

Let $\varphi_0 \in G$ be arbitrary. Then the set $A \subset G$:

$$A = \{\varphi : \varphi \in G, \| f - \varphi \| \leqslant \| f - \varphi_0 \| \}$$

is nonempty ($\varphi_0 \in A$), closed and bounded (since if $\varphi \in A$ then $\| \varphi \| \leqslant \| \varphi - f \| + \| f \| \leqslant \| \varphi_0 - f \| + \| f \|$). Therefore A is compact and obviously

$$E_G(f) = \inf\{ \| f - \varphi \| : \varphi \in G\} = \inf\{ \| f - \varphi \| : \varphi \in A\}.$$

The norm $\| f - \varphi \|$ is a continuous function of φ (by the triangle inequality $| \| f - \varphi \| - \| f - \psi \| | \leqslant \| \varphi - \psi \|$), therefore $\| f - \varphi \|$ attains its inf on the compact set A at some point $\varphi(f) \in A \subset G$. □

If the set $G \subset X$ has the property that every $f \in X$ has an element of best approximation in G, we shall call G an existence set. Obviously every existence set must be closed (every boundary point of G must belong to G). Theorem 1.1 gives us that every finite dimensional subspace of a linear normed space is an existence set.

Unfortunately the element of best approximation in an existence set G is not always unique. Let us denote by $P_G(f)$ the set

$$P_G(f) = \{\varphi : \varphi \in G, \| f - \varphi \| = E_G(f)\}$$

of all elements of best approximation of f.

Theorem 1.2. *Let X be a normed linear space and G a subspace of X, G an existence set. Then for every $f \in X$ the set $P_G(f)$ is convex and closed.*

Proof. Indeed, if $\varphi \in P_G(f)$ and $\psi \in P_G(f)$ then for every $\alpha \in [0, 1]$ we have

$$E_G(f) \leqslant \| f - (\alpha\varphi + (1 - \alpha)\psi) \| \leqslant \alpha \| f - \varphi \| + (1 - \alpha) \| f - \psi \| = E_G(f).$$

From this it follows that

$$E_G(f) = \| f - (\alpha\varphi + (1 - \alpha)\psi) \|,$$

i.e. $\alpha\varphi + (1 - \alpha)\psi \in P_G(f)$, therefore $P_G(f)$ is convex.

If $\| \varphi_m - \varphi \| \underset{m \to \infty}{\to} 0$, $\varphi_m \in G$, then φ also $\in G$, since G is closed. If $\varphi_m \in P_G(f)$ then

$$E_G(f) \leqslant \| f - \varphi \| \leqslant \| f - \varphi_m \| + \| \varphi_m - \varphi \| \underset{m \to \infty}{\to} E_G(f),$$

i.e.

$$E_G(f) = \| f - \varphi \|,$$

therefore $\varphi \in P_G(f)$. □

We shall see now that, when the normed linear space is strictly normed, there exists a unique element of best approximation in every subspace of X, which is an existence set (in particular in every finite dimensional subspace). Let us recall that a normed linear space X is said to be strictly normed if the equality $\| x + y \| = \| x \| + \| y \|$ implies that $x = \alpha y$, α a real number.

Theorem 1.3 (Uniqueness theorem). *Let X be a strictly normed linear space and G a subspace of X, G an existence set. Then for every $f \in X$ there exists a unique element of best approximation in G, i.e. $P_G(f)$ consists of exactly one element.*

Proof. Let $\varphi \in P_G(f)$ and $\psi \in P_G(f)$. In virtue of theorem 1.2 $(\varphi + \psi)/2 \in P_G(f)$ and therefore

$$E_G(f) = \| f - (\varphi + \psi)/2 \| \leqslant \tfrac{1}{2} \| f - \varphi \| + \tfrac{1}{2} \| f - \psi \| = E_G(f).$$

From this it follows that

$$\| f-(\varphi+\psi)/2 \| = \left\| \frac{f-\varphi}{2} \right\| + \left\| \frac{f-\psi}{2} \right\|.$$

Since X is strictly normed, the last equality implies $f-\varphi=\alpha(f-\psi)$. If $\alpha \neq 1$ it follows that $f \in G$ and in this case $P_G(f)=\{f\}$, i.e. $\varphi=\psi$. If $\alpha=1$ we obtain $\varphi=\psi$. □

Corollary 1.1. *Let G be a finite dimensional subspace of a linear strictly normed space X. Then for every $f \in X$ there exists a unique element of best approximation in G.*

In this book we shall use mostly the following function spaces.

(i) The space $C[a,b]$ of all functions which are continuous in the closed finite interval $[a,b]$. This space becomes a normed one (even a Banach space, i.e. a complete one) if we introduce the so-called uniform or Chebyshev norm,

$$\| f \|_{C[a,b]} = \| f \|_C = \max\{ |f(x)| : x \in [a,b] \}.$$

The approximations in $C[a,b]$ are usually called uniform or Chebyshev approximations.

(ii) The space $L_p(a,b)$, $1 \leqslant p < \infty$, (a,b) a finite or infinite interval,† consisting of all functions f such that $|f|^p$ is Lebesgue-integrable in the interval (a,b). If we consider all equivalent (in the sense of Lebesgue) functions as one, $L_p(a,b)$ becomes a normed (even Banach) space with respect to the so-called L_p-norm

$$\| f \|_{L_p(a,b)} = \| f \|_{L_p} = \| f \|_p = \left\{ \int_a^b |f(x)|^p \mathrm{d}x \right\}^{1/p}. \tag{2}$$

The approximations in $L_p(a,b)$ will be called L_p-approximations.

(iii) We shall use the notation (2) also in the case $0<p<1$ when $\| f \|_p$ is not a norm (since the triangle inequality does not hold), but only a quasinorm

$$\| f+g \|_p \leqslant c(p)(\| f \|_p + \| g \|_p).$$

(iv) The space $L_\infty[a,b]$ consisting of all essentially bounded functions in the interval $[a,b]$ supplied with the norm

$$\| f \|_{L_\infty[a,b]} = \| f \|_{L_\infty} = \| f \|_\infty = \operatorname{ess\,sup.} |f(x)| = \inf\{\lambda : \operatorname{mes}\{x : |f(x)| > \lambda\} = 0\}$$

where mes$\{A\}$ denotes the Lebesgue measure of the set $\{A\}$.

If $f \in L_\infty$ then $\| f \|_p \to \| f \|_\infty$ when $p \to \infty$. Furthermore it is clear that if $f \in C[a,b]$ then $\| f \|_C = \| f \|_\infty$. Sometimes we shall use the notation $\| f \|_C$ also for bounded functions and we shall interpret it as $\sup\{ |f(x)| : x \in [a,b] \}$.

† We shall use also the notation $L_p[a,b]$.

Beside these spaces we shall use in some paragraphs Orlich spaces, Besov spaces, Hardy spaces and BMO spaces.

The spaces $C[a,b]$, $L_p(a,b)$, $1 \leqslant p < \infty$, $L_\infty[a,b]$ are normed linear ones. Therefore, in virtue of theorem 1.1 for each of their elements there exists an element of best approximation with respect to an arbitrary finite dimensional subspace of theirs. The main subspace used is that of algebraic polynomials of nth degree, denoted by P_n. It is the $(n+1)$-dimensional subspace generated by the functions $1, x, \ldots, x^n$. Applying theorem 1.1 in this case we obtain the following.

Theorem 1.4 (E. Borel). *Let $f \in C[a,b]$ (or $L_p[a,b]$, $1 \leqslant p < \infty$). Then for every natural number n there exists an algebraic polynomial $p \in P_n$ of best uniform (or L_p) approximation in P_n.*

It is often necessary to approximate 2π-periodic functions. Without pointing it out explicitly every time, we shall use the notations we introduced in the case of an interval also for linear spaces of 2π-periodic functions, namely $C[0,2\pi]$, $L_p[0,2\pi]$, $1 \leqslant p \leqslant \infty$. The tools used most often in this case are the trigonometric polynomials. We shall denote by T_n the set of all trigonometric polynomials of nth order, i.e. T_n is the $(2n+1)$-dimensional subspace generated by the functions $1, \cos x, \sin x, \ldots, \cos nx, \sin nx$. In the periodic case theorem 1.1 implies the following.

Theorem 1.4′. *Let f be a 2π-periodic function and $f \in C[0,2\pi]$ ($f \in L_p[0,2\pi]$). For every natural number n there exists a trigonometric polynomial $q \in T_n$ of best uniform (L_p) approximation in T_n.*

Let us consider now the question of uniqueness. One can show that the spaces L_p, $1 < p < \infty$, are strictly normed (see for example S.M. Nikol'skij (1969)). Then theorem 1.3 implies the following.

Theorem 1.5. *Let $f \in L_p(a,b)$ (let f be 2π-periodic and $f \in L_p[0,2\pi]$), $1 < p < \infty$. Then for every natural number n there exists a unique algebraic (trigonometric) polynomial of nth degree of best L_p-approximation in P_n (in T_n).*

However, the spaces C, L_∞, $L = L_1$ are not strictly normed. Let us show this for instance for $C[0,1]$. If we consider the functions 1 and x, we have

$$\|1 + x\|_{C[0,1]} = \|1\|_{C[0,1]} + \|x\|_{C[0,1]} = 2$$

but the functions 1 and x are not linearly dependent.

It is easy to see by examples that in the general case in L_1 we do not have uniqueness. Let us consider the function

$$\sigma(x) = \begin{cases} -1, & -1 \leqslant x \leqslant 0, \\ 1, & 0 < x \leqslant 1. \end{cases}$$

In $L_1(-1,1)$ every constant c, $-1 \leqslant c \leqslant 1$, is a polynomial of degree zero of best approximation to σ.

Fortunately enough it turns out that the algebraic polynomial of nth degree of best uniform approximation is unique. This follows from the Chebyshev theorem, which gives a characterization of the algebraic polynomial of best uniform approximation by alternation. This theorem as well as its proof can be modified for the best rational uniform approximation. That is why it will be of special interest of us.

1.2 Characterization of the algebraic polynomial of best uniform approximation

Now we are going to solve the third basic problem of the theory of approximation in the case of uniform approximation by means of algebraic polynomials – characterization of the algebraic polynomial of best uniform approximation. This problem was solved by P.L. Chebyshev in the last century with his famous alternation theorem.

Let $f \in C[a,b]$. We shall denote by $E_n(f)_C$ the best uniform approximation of the function f by means of algebraic polynomials of nth degree:

$$E_n(f)_{C[a,b]} = E_n(f)_C = \inf\{\, \| f-p \|_{C[a,b]} : p \in P_n \}.$$

In what follows in this section we shall write $E_n(f)$ instead of $E_n(f)_C$ and $\| f \|$, $\| f-p \|$ instead of $\| f \|_C$, $\| f-p \|_C$.

Definition 1.1. *Let $f \in C[a,b]$. The polynomial $p \in P_n$ is said to realize Chebyshev alternation (or simply alternation) for f in $[a,b]$ if there exist $n+2$ points x_i, $i=1,\dots,n+2$, $a \leqslant x_1 < \cdots < x_{n+2} \leqslant b$, such that*

$$f(x_i) - p(x_i) = \varepsilon(-1)^i \| f-p \|, \quad i=1,\dots,n+2,$$

where the number ε is $+1$ or -1.

The Chebyshev alternation has the following geometric interpretation: let $p \in P_n$ realize Chebyshev alternation for $f \in C[a,b]$ in $[a,b]$. Let us consider the functions $\varphi(x) = f(x) + \| f-p \|$ and $\psi(x) = f(x) - \| f-p \|$. Then the graph of the polynomial lies in the strip between φ and ψ, touching alternately the upper function φ and the lower function ψ at least $n+2$ times.

Theorem 1.6 (Chebyshev alternation theorem). *Let $f \in C[a,b]$. The necessary and sufficient condition for the algebraic polynomial $p \in P_n$ to be a polynomial of best uniform approximation for f in P_n is that p realizes Chebyshev alternation for f in $[a,b]$.*

Proof. Let $p \in P_n$ realize Chebyshev alternation for f in $[a,b]$. Assume that p is not a polynomial of best uniform approximation, but $q \in P_n$ is. Then

$$E_n(f) = \| f-q \| < \| f-p \|.$$

The above inequality implies that the polynomial $s = p - q \in P_n$ has the sign of $p-f$ in the points x_i, $i=1,\dots,n+2$, since $|p(x_i)-f(x_i)|=\|f-p\|>\|f-q\|$, $p(x_i)-q(x_i)=p(x_i)-f(x_i)-(q(x_i)-f(x_i))$. Therefore $s\in P_n$ will change its sign at least $n+1$ times, i.e. s must have at least $n+1$ zeros in $[a,b]$. Since $s\in P_n$, it follows that $s\equiv 0$, i.e. $p=q$, which is a contradiction with the assumption.

Let now $p\in P_n$ be an algebraic polynomial of best uniform approximation for f in P_n. We shall show that p realizes Chebyshev alternation for f. Let us assume, contrary to this, that $m+2$ is the highest number of points $x_1<x_2<\cdots<x_{m+2}$ in $[a,b]$ such that

$$f(x_i)-p(x_i)=\varepsilon(-1)^i\|f-p\|=\varepsilon(-1)^iE_n(f),\quad i=1,\dots,m+2, \tag{1}$$

where $\varepsilon=1$ or -1 and $m<n$. Then there exist $m+3$ points $\xi_0, \xi_1,\dots,\xi_{m+2}$ which satisfy the inequalities

$$a=\xi_0\leqslant x_1<\xi_1<x_2<\xi_2<\cdots<\xi_{m+1}<x_{m+2}\leqslant\xi_{m+2}=b$$

and are such that for every $x\in[\xi_{i-1},\xi_i]$ we have

$$\varepsilon(-1)^i(f(x)-p(x))>-E_n(f),\quad i=1,\dots,m+2. \tag{2}$$

From (1) it follows that the continuous function $f-p$ changes its sign in the interval $[x_i,x_{i+1}]$, therefore the points $\xi_1, \xi_2,\dots,\xi_{m+1}$ can be chosen so that

$$f(\xi_i)=p(\xi_i),\quad i=1,\dots,m+1. \tag{3}$$

Since $[\xi_{i-1},\xi_i]$, $i=1,\dots,m+2$, are a finite number of closed intervals and $f-p$ is a continuous function in each of them, from (2) it follows that there exists $\delta>0$ such that for every $x\in[\xi_{i-1},\xi_i]$, $i=1,\dots,m+2$, we have the inequality

$$\varepsilon(-1)^i(f(x)-p(x))>\delta-E_n(f). \tag{4}$$

Let us set

$$Q(x)=(-1)^{m+1}\lambda(x-\xi_1)\cdots(x-\xi_{m+1}),$$

where

$$\lambda=\frac{\delta}{2\|(x-\xi_1)\cdots(x-\xi_{m+1})\|_{C[a,b]}}.$$

Since $m<n$, we have $Q\in P_n$.

From this definition of Q it follows also that

$$|Q(x)|\leqslant\delta/2\quad \text{for } x\in[a,b], \tag{5}$$

$$(-1)^iQ(x)>0\quad \text{for } x\in(\xi_i,\xi_{i+1}),\quad i=0,\dots,m+1, \tag{6}$$

$$Q(\xi_0)>0,\quad (-1)^{m+1}Q(\xi_{m+2})>0, \tag{7}$$

$$Q(\xi_i)=0, \quad i=1,\ldots,m+1. \tag{8}$$

Since p is an algebraic polynomial of best approximation to f in P_n, we have

$$-E_n(f)\leqslant f(x)-p(x)\leqslant E_n(f) \quad \text{for } x\in[a,b]. \tag{9}$$

Let us consider the difference

$$f(x)-p(x)-\varepsilon Q(x).$$

In view of (4), (5) and (6), for every $x\in[\xi_i,\xi_{i+1}]$, we have, for $i=0,\ldots,m+1$,

$$\begin{aligned}\varepsilon(-1)^i(f(x)-p(x)-\varepsilon Q(x))&=\varepsilon(-1)^i(f(x)-p(x))-(-1)^iQ(x)\\ &>\delta-E_n(f)-\delta/2=\delta/2-E_n(f).\end{aligned} \tag{10}$$

From (5)–(9) we also obtain that, for every $x\in(\xi_i,\xi_{i+1})$ and $x=\xi_0,\ \xi_{m+2}$, we have

$$\begin{aligned}\varepsilon(-1)^i(f(x)-p(x)-\varepsilon Q(x))=(-1)^i\varepsilon(f(x)-p(x))-(-1)^iQ(x)\leqslant E_n(f)\\ -(-1)^iQ(x)<E_n(f).\end{aligned} \tag{11}$$

For $x=\xi_i$, $i=1,\ldots,m+1$, we have, from (8),

$$f(\xi_i)-p(\xi_i)-\varepsilon Q(\xi_i)=0. \tag{12}$$

Consequently the inequalities (10)–(12) give us that, for every $x\in[a,b]$, we have

$$|f(x)-p(x)-\varepsilon Q(x)|<E_n(f). \tag{13}$$

Since $f-p-\varepsilon Q$ is a continuous function in $[a,b]$, from (13) it follows that

$$\|f-p-\varepsilon Q\|<E_n(f),$$

i.e. a contradiction, since $p+\varepsilon Q\in P_n$. □

From theorem 1.6 there follows easily the uniqueness of the algebraical polynomial of best uniform approximation as follows.

Theorem 1.7. *Let $f\in C[a,b]$. For every natural number n there exists a unique algebraic polynomial $p\in P_n$ of best uniform approximation to f in P_n.*

Proof. Let $p\in P_n$ and $q\in P_n$ be two algebraic polynomials of best uniform approximation to f:

$$\|f-p\|=\|f-q\|=E_n(f). \tag{14}$$

From theorem 1.2 the polynomial $g=(p+q)/2\in P_n$ is also a polynomial of best uniform approximation to f. By theorem 1.6 g realizes Chebyshev alternation for f, i.e. there exist $n+2$ points x_i, $i=1,\ldots,n+2$, $a\leqslant x_1<$

$x_2 < \cdots < x_{n+2} \leqslant b$, such that

$$f(x_i) - \frac{p(x_i) + q(x_i)}{2} = \varepsilon(-1)^i E_n(f), \quad i = 1, \ldots, n+2, \tag{15}$$

where $\varepsilon = 1$ or -1.

From (14) it follows that

$$\left.\begin{aligned} |f(x_i) - p(x_i)| \leqslant E_n(f), \\ |f(x_i) - q(x_i)| \leqslant E_n(f). \end{aligned}\right\} \tag{16}$$

Therefore the equality (15) can be fulfilled only if we have

$$f(x_i) - p(x_i) = f(x_i) - q(x_i),$$

i.e. if $p(x_i) = q(x_i)$ for $i = 1, \ldots, n+2$.

We thus have that the algebraical polynomials $p \in P_n$ and $q \in P_n$ coincide in $n+2$ different points. Consequently $p = q$. □

The following theorem of de la Vallée-Poussin is very useful in the numerical methods for obtaining the polynomial of best uniform approximation.

Theorem 1.8. *Let $f \in C[a, b]$, $p \in P_n$ and x_i, $i = 1, \ldots, n+2$, $a \leqslant x_1 < x_2 < \cdots < x_{n+2} \leqslant b$, be $n+2$ different points in $[a, b]$. If the difference $f - p$ has alternate signs at the points x_i, $i = 1, \ldots, n+2$, then*

$$E_n(f) \geqslant \mu = \min\{|f(x_i) - p(x_i)| : i = 1, \ldots, n+2\}.$$

Proof. Let us assume that $E_n(f) < \mu$. Let $q \in P_n$ be the algebraic polynomial of best uniform approximation to f, i.e. $\| f - q \| = E_n(f) < \mu$.

From this it follows that the difference $p - q$ must have the sign of $p(x_i) - f(x_i)$ at the points x_i, $i = 1, \ldots, n+2$. By the conditions of the theorem therefore $p - q$ must have alternate signs at $n+2$ points x_i, $i = 1, 2, \ldots, n+2$, i.e. the algebraic polynomial $p - q \in P_n$ must have at least $n+1$ different zeros in $[a, b]$; consequently $p - q \equiv 0$ which contradicts

$$\| f - q \| = E_n(f) < \mu \leqslant \| f - p \|.$$ □

1.3 Numerical methods

We shall describe in this section the so-called Remez algorithms for numerical solution of basic problem (v) from section 1.1 – finding the polynomial of best uniform approximation. The algorithms are more general and can be used for best uniform approximation by means of arbitrary Haar subspaces of $C[a, b]$.

Definition 1.2. *The system $\{\varphi_i\}_{i=1}^n$ of functions $\varphi_i \in C[a, b]$, $i = 1, \ldots, n$, is said to be a Chebyshev system on the interval $[a, b]$ if every generalized polynomial*

$\varphi = \sum_{i=1}^{n} a_i\varphi_i$ can have at most $n-1$ zeros in $[a,b]$ (every zero calculated with its multiplicity).

Let $C^{(k)}[a,b]$ denote the space of all functions in the interval $[a,b]$ which have kth derivative $f^{(k)}$ in $[a,b]$, which belongs to $C[a,b]$.

We shall say that $x_0 \in [a,b]$ is a zero of $f \in C^{(k)}[a,b]$ of order k (or multiplicity k) if

$$f(x_0) = f'(x_0) = \cdots = f^{(k-1)}(x_0) = 0, \quad f^{(k)}(x_0) \neq 0.$$

Definition 1.3. *A subspace $G \subset C[a,b]$, $G = \{\varphi : \varphi = \sum_{i=1}^{n} a_i\varphi_i\}$, generated by the Chebyshev system $\{\varphi_i\}_{i=1}^{n}$, is said to be a Haar subspace.*

Let $\{\varphi_i\}_{i=1}^{n}$ be a Chebyshev system. In this section we shall use the following notations. Let $f \in C[a,b]$. Then

$$E_n(f) = \inf\left\{ \| f - \varphi \| : \varphi = \sum_{i=1}^{n} a_i\varphi_i \right\},$$

$$\Delta(a) = \left\| \sum_{i=1}^{n} a_i\varphi_i - f \right\|_{C[a,b]},$$

$$\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{R}^n, \quad r(\mathbf{a}, x) = \sum_{i=1}^{n} a_i\varphi_i(x) - f(x),$$

$\mathbb{R}^n$ the n-dimensional Euclidean space.

Our aim is to find real numbers $\{c_i^*\}_{i=1}^{n}$ such that

$$\left\| f - \sum_{i=1}^{n} c_i^*\varphi_i \right\|_{C[a,b]} = E_n(f).$$

First Remez algorithm

The algorithm consists of the following recursive procedure.

(i) *Select $n+1$ points $X^{(0)} = \{x_i\}_0^n$, where $a \leqslant x_0 < x_1 < \cdots < x_n \leqslant b$;*

(ii) *Set $k = 0$;*

(iii) *Given the set $X^{(k)}$ find a vector $\mathbf{c}^{(k)} \in \mathbb{R}^n$ such that if we denote $\Delta^{(k)}(\mathbf{c}) = \max\{|r(\mathbf{c}, x)| : x \in X^{(k)}\}$ then*

$$\Delta^{(k)}(\mathbf{c}^{(k)}) = \inf\{\Delta^{(k)}(\mathbf{c}) : \mathbf{c} \in \mathbb{R}^n\};$$

(iv) *Find a point $x_{n+k+1} \in [a,b]$ such that $\Delta(\mathbf{c}^{(k)}) = |r(\mathbf{c}^{(k)}, x_{n+k+1})|$;*

(v) *Form the set $X^{(k+1)} = X^{(k)} \cup \{x_{n+k+1}\}$;*

(vi) *Set $k = k+1$;*

(vii) *Go to* (iii).

The choice of the initial set X^0 can be done in different ways (equidistant point in the trigonometrical case, the roots of the $(n+1)$-th polynomial of Chebyshev in the algebraic case and so on) and there exist no strong rules for this.

At step (iii) we have to find in fact the polynomial of the best uniform approximation on a set which consists of a finite number of points.

Step (iv) is usually the most laborious point in the algorithm.

The execution of the algorithm stops when the polynomial obtained at the kth iteration satisfies some demands.

The method (i)–(vii) generates a sequence of vectors $\{\mathbf{c}^{(k)}\}_{k=0}^{\infty}$ for which we have the following.

Theorem 1.9. *Let* $\mathbf{c}^*$ *be a cluster point of the sequence* $\{\mathbf{c}^{(k)}\}_{k=0}^{\infty}$. *Then* $E_n(f) = \Delta(\mathbf{c}^*)$.

Proof. Let us set $|\mathbf{c}| = \sum_{i=1}^{n} |c_i|$ and

$$\theta = \min_{|\mathbf{c}|=1} \max_{x \in X^{(0)}} \left| \sum_{i=1}^{n} c_i \varphi_i(x) \right|.$$

Since $X^{(0)}$ contains $n+1$ different points and $\{\varphi_i\}_{i=1}^{n}$ is a Chebyshev system on the interval $[a, b]$ we have $\theta > 0$. From $X^{(k)} \subset X^{(k+1)} \subset [a, b]$ we get for every $\mathbf{c} \in \mathbb{R}^n$ that

$$\Delta^{(k)}(\mathbf{c}) \leqslant \Delta^{(k+1)}(\mathbf{c}) \leqslant \Delta(\mathbf{c})$$

and consequently ($\bar{\mathbf{c}}$: $\Delta(\bar{\mathbf{c}}) = E_n(f)$)

$$\Delta^{(k)}(\mathbf{c}^{(k)}) \leqslant \Delta^{(k)}(\mathbf{c}^{(k+1)}) \leqslant \Delta^{(k+1)}(\mathbf{c}^{(k+1)}) \leqslant \Delta^{(k+1)}(\bar{\mathbf{c}}) \leqslant \Delta(\bar{\mathbf{c}}) = E_n(f).$$

The last inequalities show that the sequence $\{\Delta^{(k)}(\mathbf{c}^{(k)})\}_0^{\infty}$ is monotone nondecreasing and bounded from above. This means that there exists $\varepsilon \geqslant 0$ such that $\lim_{k \to \infty} \Delta^{(k)}(\mathbf{c}^{(k)}) = E_n(f) - \varepsilon$. We shall show that $\varepsilon = 0$.

First we prove that the sequence $\{\mathbf{c}^{(k)}\}_{k=0}^{\infty}$ is bounded. Indeed,

$$\Delta^0(\mathbf{c}) = \max_{x \in X^0} \left| \sum_{i=1}^{n} c_i \varphi_i(x) - f(x) \right| \geqslant \max_{x \in X^{(0)}} \left| \sum_{i=1}^{n} c_i \varphi_i(x) \right| - \| f \|_C \geqslant \theta |\mathbf{c}| - \| f \|$$

and if $|\mathbf{c}| > 2\| f \|/\theta$ then $\Delta^{(k)}(\mathbf{c}) \geqslant \Delta^0(\mathbf{c}) > \| f \| = \Delta^{(k)}(0)$, i.e. $\mathbf{c}$ can not minimize any of the functions $\Delta^{(k)}$. So the sequence $\{\mathbf{c}^{(k)}\}_{k=1}^{\infty}$ generated by the algorithm is bounded.

Further let us set $M = \max_{1 \leqslant i \leqslant n} \| \varphi_i \|_{C[a,b]}$. Then for an arbitrary vector $\mathbf{b}$

$$|r(\mathbf{b}, x) - r(\mathbf{c}, x)| = \left| \sum_{i=1}^{n} (b_i - c_i) \varphi_i(x) \right| \leqslant M |\mathbf{b} - \mathbf{c}|$$

and therefore $|r(\mathbf{b}, x)| \leqslant |r(\mathbf{c}, x)| + M |\mathbf{b} - \mathbf{c}|$, i.e.

$$\Delta(\mathbf{b}) = \| r(\mathbf{b}, \cdot) \|_{C[a,b]} = |r(\mathbf{b}, \bar{x})| \leqslant |r(\mathbf{c}, \bar{x})| + M |\mathbf{b} - \mathbf{c}| \leqslant \Delta(\mathbf{c}) + M |\mathbf{b} - \mathbf{c}|. \tag{1}$$

Let us suppose now that $\varepsilon > 0$ and $\mathbf{c}^* \in \mathbb{R}^n$ is a cluster point of the sequence $\{\mathbf{c}^{(k)}\}_{k=0}^{\infty}$. For every $\delta > 0$ there exists an index k such that $|\mathbf{c}^* - \mathbf{c}^{(k)}| < \delta$ and an index $i > k$ such that $|\mathbf{c}^* - \mathbf{c}^{(i)}| < \delta$. Then $|\mathbf{c}^{(k)} - \mathbf{c}^{(i)}| < 2\delta$ and, using (1). setting $\mathbf{c}^*$ in place of $\mathbf{b}$, we obtain

$$\begin{aligned} E_n(f) &\leqslant \Delta(\mathbf{c}^*) \leqslant \Delta(\mathbf{c}^{(k)}) + M\delta = |r(\mathbf{c}^{(k)}, x^{(k+1)})| + M\delta \\ &\leqslant |r(\mathbf{c}^{(i)}, x^{(k+1)})| + 3M\delta \leqslant \Delta^{(k+1)}(\mathbf{c}^{(i)}) + 3M\delta \\ &\leqslant \Delta^{(i)}(\mathbf{c}^{(i)}) + 3M\delta \leqslant E_n(f) - \varepsilon + 3M\delta. \end{aligned}$$

The number $\delta > 0$ was arbitrary, so for $\varepsilon > 3M\delta$ this leads to a contradiction. Therefore $\varepsilon = 0$ and $\Delta(\mathbf{c}^*) = E_n(f)$. □

Corollary 1.2. *Let $\{\varphi_i(x)\}_1^n = \{x^i\}_0^{n-1}$. Then there exists $\lim_{k\to\infty} \mathbf{c}^{(k)} = \mathbf{c}^*$.*

This follows from the uniqueness of the best uniform algebraic approximation (theorem 1.7).

Corollary 1.2 gives that the first Remez algorithm is convergent for the case of approximation by means of algebraic polynomials.

Remark. The uniqueness theorem is also valid for approximation in the uniform metric by means of a Chebyshev system. So we have convergence of the first Remez algorithm also in the general case of a Chebyshev system.

Second Remez algorithm

We shall describe the second Remez algorithm again for an arbitrary Chebyshev system and we shall prove the order of convergence for the case of uniform approximation by means of algebraic polynomials.

(i) *Take $n+1$ different points x_i, $i = 0,\dots,n$, $a \leqslant x_0 < x_1 < \cdots < x_n \leqslant b$;*

(ii) *Solve the linear system*

$$f(x_j) - \sum_{i=1}^{n} c_i\varphi_i(x_j) = (-1)^j\lambda, \quad j = 0, 1, \dots, n,$$

with respect to the unknowns $c_1,\dots,c_n$ and λ;

(iii) *Find the points $\{z_i\}_{i=0}^{n+1}$ such that $z_0 = a$, $z_{n+1} = b$ and $r(z_i) = 0$ for $i = 1,\dots,n$;*†

(iv) *Select the points $y_i \in [z_i, z_{i+1}]$, $i = 0, 1,\dots,n$, such that*

$$(\operatorname{sign} r(x_i))r(y_i) = \max\{r(x)\operatorname{sign} r(x_i) : x \in [z_i, z_{i+1}]\},$$

(v) *If $\|r(\mathbf{c};\cdot)\|_{C[a,b]} > \max\{|r(\mathbf{c}; y_i)| : 0 \leqslant i \leqslant n\}$ then there exists a point $y \in [a,b]$ such that $|r(\mathbf{c}; y)| = \|r(\mathbf{c};\cdot)\|_{C[a,b]}$ – we put the point y in place of some point among $y_0, y_1,\dots,y_n$ so that the function $r(\mathbf{c}; x)$ would preserve the alternating signs on the newly obtained points which we denote again by $y_0, y_1,\dots,y_n$;*

(vi) *Go to* (ii) *and instead of the points $\{x_i\}_{i=0}^n$ consider the points $\{y_i\}_{i=0}^n$.*

† $r(x) \equiv r(\mathbf{c}; x)$

This procedure can be easily carried out using computers and numerical experiments show that it is not very sensitive to the choice of the initial points.

Usually we go out of the iterative process and stop the calculation when on the kth step $\|r(\mathbf{c};\cdot)\|$ differs negligibly from $|\lambda|$. This stop-condition comes from the Chebyshev theorem of alternation.

The second Remez algorithm has quadratic convergence under some restrictions on the smoothness of the function f (see L. Veidinger (1960)). We shall prove here the linear convergence of the algorithm for every $f \in C[a,b]$ in the case of polynomial approximations.

Theorem 1.10. *Let $\{\varphi_i\}_{i=1}^{n} = \{x^i\}_{i=0}^{n-1}$ and let $f \in C[a,b]$. The polynomial $p^{(k)} = \sum_{i=0}^{n-1} c_i x^i$ generated on the kth step by the second Remez algorithm satisfies the condition $\|p^{(k)} - p\|_{C[a,b]} \leqslant c\theta^k$, where p is the algebraic polynomial of best uniform approximation for f of $(n-1)$-th degree, $0 < \theta < 1$ and c is a constant, independent of k.*

Proof. We again use the abbreviation $r(x) \equiv r(\mathbf{c}; x)$. Since we described a single cycle of the second Remez algorithm let us denote $\alpha = |r(x_0)| = \cdots = |r(x_n)| = |\lambda|$, $\beta = \max\{|r(y_i)| : i = 0, \ldots, n\} = \|r(\mathbf{c};\cdot)\|_C$, $\gamma = \min\{|r(y_i)| : i = 0, \ldots, n\}$, $\bar{\beta} = \|f - p\|_C$.

From de la Vallée-Poussin's theorem (theorem 1.8) we get $\alpha \leqslant \gamma \leqslant \bar{\beta} \leqslant \beta$. Let us agree that on the next cycle of the algorithm the constants corresponding to α, β, γ, λ and the coefficient vector $\mathbf{c}$ will be denoted by α', β', γ', λ' and $\mathbf{c}'$. According to this convention it is clear that the vector $\mathbf{c}'$ is selected by the system

$$(-1)^i \lambda' + \sum_{j=0}^{n-1} c'_j y_i^j = f(y_i), \quad i = 0, \ldots, n,$$

and

$$\lambda' = \frac{\begin{vmatrix} f(y_0) & 1 & y_0 \cdots y_0^{n-1} \\ & \cdots & \\ f(y_n) & 1 & y_n \cdots y_n^{n-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & y_0 \cdots y_0^{n-1} \\ -1 & 1 & y_1 \cdots y_1^{n-1} \\ & \cdots & \\ (-1)^n & 1 & y_n \cdots y_n^{n-1} \end{vmatrix}} = \frac{\sum_{i=0}^{n} (-1)^i f(y_i) M_i}{\sum_{i=0}^{n} M_i},$$

where M_i are the minors corresponding to the first column of the matrix in the denominator.

If f has the form $f = \sum_{j=0}^{n-1} a_j x^j$ then the approximation has to be exact and $\lambda' = 0$, i.e.

$$\sum_{i=0}^{n} (-1)^i \sum_{j=0}^{n-1} a_j y_i^j M_i = 0.$$

Thus we may replace $f(y_i)$ by

$$r(y_i)=f(y_i)-\sum_{j=0}^{n-1} c_j y_i^j$$

in the expression for λ'. Taking into account that $\operatorname{sign} r(y_i)=-\operatorname{sign} r(y_{i+1})$ we obtain

$$\alpha'=|\lambda'|=\left(\sum_{i=0}^{n} M_i|r(y_i)|\right)\Big/\sum_{i=0}^{n} M_i.$$

Since $M_i>0$,

$$M_i=\prod(y_k-y_j), \tag{2}$$

where the product is taken over all k,j such that $k>j$, $k,j=0,1,\ldots,i-1$, $i+1,\ldots,n$ and $y_k>y_j$ for $k>j$.

Now let $\theta_i=M_i/\sum_{i=0}^{n} M_i$. Then

$$\alpha'=\sum_{i=0}^{n}\theta_i|r(y_i)|\geqslant\gamma\sum_{i=0}^{n}\theta_i=\gamma\geqslant\alpha. \tag{3}$$

We shall show that there exists θ, $0<\theta<1$, such that for all numbers θ_i generated at the kth iteration of the algorithm we have

$$1-\theta<\theta_i<1,\quad i=0,1,\ldots,n. \tag{4}$$

From (2) it follows that this will be true if there exists $\delta>0$ independent of k such that

$$y_{i+1}^{(k)}-y_i^{(k)}\geqslant\delta>0,\quad i=0,\ldots,n-1,\quad k=1,2,\ldots. \tag{5}$$

Let us assume that this inequality is not true. Then the sequence $\{y_0^{(k)},\ldots,y_n^{(k)}\}_{k=1}^{\infty}$ will have a cluster point $(\bar{y}_0,\bar{y}_1,\ldots,\bar{y}_n)$, where at least two points $\bar{y}_i$ coincide. Consequently there exists an algebraic polynomial $q(x)=\sum_{i=0}^{n-1} a_i x^i$ which interpolates f at the points $\bar{y}_0$, $\bar{y}_1,\ldots,\bar{y}_n$ (the number of the different points is at most n). By definition $\alpha^{(k+1)}$ is the best approximation of f at the points $y_0^{(k)}$, $y_1^{(k)},\ldots,y_n^{(k)}$ at the kth iteration and

$$\begin{aligned}\alpha^{(k+1)}&\leqslant\max\{|f(y_i^{(k)})-q(y_i^{(k)})|: i=0,\ldots,n\}\\&=\max\{|f(y_i^{(k)})-q(y_i^{(k)})-f(\bar{y}_i)+q(\bar{y}_i)|: i=0,\ldots,n\},\end{aligned} \tag{6}$$

since $f(\bar{y}_i)=q(\bar{y}_i)$, $i=0,\ldots,n$. This inequality contradicts the fact that $\alpha'\geqslant\alpha$ (see (3)), i.e. $\alpha^{(1)}\leqslant\alpha^{(2)}\leqslant\cdots\leqslant\alpha^{(k+1)}\leqslant\cdots$. Really, if k is such that $\max\{|y_i^{(k)}-\bar{y}_i|: i=0,\ldots,n\}$ is small enough, then $(\alpha^{(1)}>0)$

$$\max\{|f(y_i^{(k)})-f(\bar{y}_i)-(q(y_i^{(k)})-q(\bar{y}_i))|: i=0,\ldots,n\}<\alpha^{(1)} \tag{7}$$

since q and f are continuous functions. From (6) and (7) we get the contradiction $\alpha^{(k+1)}<\alpha^{(1)}$.

Therefore (5), and consequently (4), hold true. Using (4) we obtain

$$\gamma' - \gamma \geqslant \alpha' - \gamma = \sum_{i=0}^{n} \theta_i(|r(y_i)| - \gamma) \geqslant (1-\theta)(\beta - \gamma) \geqslant (1-\theta)(\bar{\beta} - \gamma)$$

and

$$\bar{\beta} - \gamma' = (\bar{\beta} - \gamma) - (\gamma' - \gamma) \leqslant (\bar{\beta} - \gamma) - (1-\theta)(\bar{\beta} - \gamma) = \theta(\bar{\beta} - \gamma),$$

i.e. $\bar{\beta} - \gamma^{(k)} \leqslant \theta^k(\bar{\beta} - \gamma^{(0)})$ and

$$\beta^{(k)} - \beta \leqslant \beta^{(k)} - \gamma^{(k)} \leqslant \frac{\gamma^{(k+1)} - \gamma^{(k)}}{1-\theta} \leqslant \frac{\bar{\beta} - \gamma^{(k)}}{1-\theta} \leqslant \frac{\theta^k(\bar{\beta} - \gamma^{(0)})^\dagger}{1-\theta}. \tag{8}$$

Finally we shall apply the strong uniqueness theorem 2.5 from Chapter 2 (obviously the theorem remains true for P_{n-1}, i.e. when $m = 0$). By this theorem if p is the polynomial of best uniform approximation for f of $(n-1)$-th degree, then there exists a constant $c(f) > 0$, depending only on f, such that for every polynomial $q \in P_{n-1}$ we have:

$$\| f - q \| \geqslant \| f - p \| + c(f) \| q - p \|. \tag{9}$$

Denoting by $p^{(k)}$ the algebraic polynomial generated at the kth step of the algorithm (at the kth iteration), we obtain, from (8) and (9),

$$\begin{aligned} \| p^{(k)} - p \| &\leqslant \frac{1}{c(f)}(\| f - p^{(k)} \| - \| f - p \|) = \frac{1}{c(f)}(\beta^{(k)} - E_n(f)) \\ &= \frac{1}{c(f)}(\beta^{(k)} - \bar{\beta}) \leqslant \frac{\bar{\beta} - \gamma^{(0)}}{c(f)(1-\theta)} \theta^k \end{aligned}$$

which completes the proof. □

Remark. Theorem 1.10 remains valid also for an arbitrary Chebyshev system.

1.4 Notes

The classical theorems for characterization and uniqueness of the best polynomial uniform approximation are given by P.L. Chebyshev (see P.L. Tchebycheff (1899), see also Ch.de la Vallée-Poussin (1910)).

The abstract theory of linear approximations is a very developed domain. We recommend the following books, which contain some more details than given here: I. Singer (1970), E.W. Cheney (1966), J. Rice (1964), (1969), Collatz, Krabs (1973).

Usually uniform approximation by means of a Chebyshev system is considered. We shall give only the formulations of some theorems.

Let K be compact and let $C(K)$ be the set of all continuous functions on

† $\gamma^{(k)}$, $\beta^{(k)}$ are γ, β at the kth iteration.

K (real- or complex-valued). The following characterization theorem is known as the Kolmogorov criterion (A.N. Kolmogorov, 1948).

Let $f \in C(K)$ and let G be a linear subspace of $C(K)$. A function $\varphi_0 \in G$ is a best approximation of f with respect to G if and only if the inequality

$$\min_{x \in A} \operatorname{Re} \overline{(f(x) - \varphi_0(x))}\, \varphi(x) \leqslant 0$$

holds for every $\varphi \in G$, where A is the set of the extremal points of $f - \varphi_0$, i.e.

$$A = \{x : x \in K, |f(x) - \varphi_0(x)| = \| f - \varphi_0 \|_{C(K)}\},$$

and $\bar{\alpha}$ is the conjugate of α.

The uniqueness theorem 1.7 has the following form.

Let G be a Haar subspace of $C(K)$ (see section 1.3). *Then for every $f \in C(K)$ there is exactly one best uniform approximation of f with respect to G* (A. Haar, 1918, A.N. Kolmogorov, 1948).

The theorem (1.6) of Chebyshev also is true for Chebyshev systems (Haar subspaces), as follows.

Let G be a Haar subspace of $C[a,b]$ with dimension n. Let $\varphi \in G$ be the best uniform approximation to $f \in C[a,b]$ with respect to G. Then there exist $n+1$ points x_i, $i = 1, \ldots, n+1$, $a \leqslant x_1 < \cdots < x_{n+1} \leqslant b$, such that

$$f(x_i) - \varphi(x_i) = \varepsilon(-1)^i \| f - \varphi \|_{C[a,b]}, \quad i = 1, \ldots, n+1, \varepsilon = \pm 1.$$

For the first and second Remez algorithms see Remez (1969). There are many modifications of these algorithms, see the books of Cheney (1966), Rice (1964, 1969), Meinardus (1967). We have used in section 1.3 the book of Cheney (1966).

2

Qualitative theory of the best rational approximation

The most essential problems in the qualitative theory of the best approximation are the problems of existence, uniqueness and characterization of the best approximation. Finally the problems connected with the continuity of the operator of the best approximation, or, as is mainly used, the continuity of the metric projection, are considered. In this chapter we shall consider these questions for the best rational approximation. The difficulties arise from the fact that the set R_{nm} of all rational functions of order (n, m) (see the exact definition in section 2.1) is not a finite dimensional linear space and the bounded sets in R_{nm} are not compact in $C[a, b]$ or in $L_p(a, b)$. Nevertheless we shall see that there exists an element of best approximation in $C[a, b]$ and $L_p(a, b)$ (section 2.1). Moreover in $C[a, b]$ we have uniqueness and characterization of the best approximation by means of an alternation, as in the linear case (see section 2.2). Unfortunately in $L_p(a, b)$, $1 \leqslant p < \infty$, we do not have uniqueness (section 2.3). In section 2.4 we consider the problem of continuity of the metric projection in $C[a, b]$ – the metric projection is continuous only in the so-called 'normal points' (see section 2.4). In section 2.5 we consider numerical methods for obtaining the rational function of best uniform approximation. We should like to remark that we examine only the usual rational approximation. Some references for the qualitative theory of generalized rational approximations are given in the notes at the end of the chapter.

2.1 Existence

We shall denote by R_{nm} the set of all real-valued rational functions with numerator an algebraic polynomial of degree at most n and denominator an algebraic polynomial of degree at most m, i.e. $r \in R_{nm}$ if r has the form

$$r(x) = \frac{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0}{b_m x^m + b_{m-1} x^{m-1} + \cdots + b_0} \tag{1}$$

where a_i, $i = 0, \ldots, n$, b_i, $i = 0, \ldots, m$, are real numbers.

If $r \in R_{nm}$ has the form (1) with $a_n \neq 0$, or $b_m \neq 0$, we say that r is nondegenerate.

If $r = p/q$, p and q algebraic polynomials without common zeros, we say that r is a reduced rational function, or r has a reduced form, or r is irreducible.

Since the set R_{nm} is nonlinear when $m \geqslant 1$, we cannot apply the general theory of linear approximation to obtain the existence of the best rational approximation in the spaces $C[a, b]$ and $L_p(a, b)$, $1 \leqslant p < \infty$. So we shall prove its existence directly.

We define the best rational approximation in $C[a, b]$ and $L_p(a, b)$, $1 \leqslant p < \infty$, of order (n, m) as usual:

$$R_{nm}(f)_{C[a,b]} = \inf\{\, \| f - r \|_{C[a,b]} : r \in R_{nm} \},$$

$$R_{nm}(f)_{L_p(a,b)} = \inf\{\, \| f - r \|_{L_p(a,b)} : r \in R_{nm} \}.$$

When it is clear we shall write briefly $R_{nm}(f)_C$ or $R_{nm}(f)$ and $R_{nm}(f)_{L_p}$ or $R_{nm}(f)_p$. When $m = n$ we shall use the notations $R_n(f)_{C[a,b]}$, $R_n(f)_C$ or $R_n(f)$ and $R_n(f)_{L_p(a,b)}$, $R_n(f)_{L_p}$ or $R_n(f)_p$.

Theorem 2.1 (Existence theorem). *Let $f \in C[a, b]$ (or $f \in L_p(a, b)$, $1 \leqslant p < \infty$). Then there exists a rational function $r \in R_{nm}$ (respectively $r_p \in R_{nm}$) such that*

$$\| f - r \|_{C[a,b]} = R_{nm}(f)_{C[a,b]}$$

(respectively

$$\| f - r_p \|_{L_p[a,b]} = R_{nm}(f)_{L_p[a,b]}).$$

Remark. The rational function r, respectively r_p, is called a rational function of best approximation to f in $C[a, b]$, or of best uniform approximation to f, respectively a rational function of the best L_p-approximation to f, of order (n, m).

Proof of theorem 2.1. Let X denote the space $C[a, b]$ or $L_p(a, b)$, $1 \leqslant p < \infty$. Let $f \in X$ and $r_N \in R_{nm}$ be such that

$$\| f - r_N \|_X \leqslant R_{nm}(f)_X + 1/N, \quad N = 1, 2, \ldots. \tag{2}$$

Then it follows that

$$\| r_N \|_X \leqslant R_{nm}(f)_X + \| f \|_X + 1 = A, \quad N = 1, 2, \ldots. \tag{3}$$

Let $r_N = p_N/q_N$, where $p_N \in P_n$, $q_N \in P_m$. We can assume that r_N is normalized so that

$$\|q_N\|_{C[a,b]} = 1, \quad N = 1, 2, \ldots. \tag{4}$$

Now (3) and (4) give us

$$\|p_N\|_X = \|q_N r_N\|_X \leqslant \|r_N\|_X \|q_N\|_{C[a,b]} \leqslant A. \tag{5}$$

From (4) and (5) it follows that the sets $\{p_N: N = 1, 2, \ldots\} \subset P_n$ and $\{q_N: N = 1, 2, \ldots\} \subset P_m$ are sequences in compact sets (P_n, P_m are finite dimensional spaces), so there exists a subsequence N_i, $i = 1, 2, \ldots, \infty$, and $p \in P_n, q \in P_m$ such that

$$\left.\begin{array}{l} \|p - p_{N_i}\|_X \underset{N_i \to \infty}{\longrightarrow} 0; \quad \|p - p_{N_i}\|_{C[a,b]} \underset{N_i \to \infty}{\longrightarrow} 0; \\ \|q - q_{N_i}\|_{C[a,b]} \underset{N_i \to \infty}{\longrightarrow} 0. \end{array}\right\} \tag{6}$$

(all norms in a finite dimensional linear normed space are equivalent).

From (4) and (6) we obtain

$$\|q\|_{C[a,b]} = 1. \tag{7}$$

If x is not a zero of q, in view of (6) $q_{N_i}(x) \to q(x)$ and therefore $q_{N_i}(x) \neq 0$ for sufficiently large N_i. Using (6) we obtain ($r = p/q$):

$$|r(x) - r_{N_i}(x)| \leqslant \frac{1}{|q(x) q_{N_i}(x)|} \{\|p\|_C \|q - q_{N_i}\|_C + \|q\|_C \|p - p_{N_i}\|_C\} \underset{N_i \to \infty}{\longrightarrow} 0. \tag{8}$$

Therefore, for every $x \in [a, b]$, x not a zero of q, we get from (2) and (8)

$$|r(x) - f(x)| \leqslant |r(x) - r_{N_i}(x)| + |r_{N_i}(x) - f(x)| \underset{N_i \to \infty}{\longrightarrow} R_{nm}(f)_C \tag{9}$$

or

$$|r(x) - f(x)| \leqslant R_{nm}(f)_{C[a,b]}. \tag{10}$$

On the other hand we have from (3) for every $x \in [a, b]$

$$\left|\frac{p_{N_i}(x)}{q_{N_i}(x)}\right| \leqslant A \quad \text{or} \quad |p_{N_i}(x)| \leqslant A |q_{N_i}(x)|.$$

The last inequality together with (6) gives us

$$|p(x)| \leqslant A |q(x)|, \quad x \in [a, b]. \tag{11}$$

The inequality (11) shows that every zero of q in $[a, b]$ is also a zero of p with at least the same multiplicity. Therefore $r = p/q$ is a continuous function in $[a, b]$. Then, since (10) is valid for $x \in [a, b]$ which are not zeros of q, (10) is valid for all $x \in [a, b]$, so (10) gives us

$$\|f - r\|_C = R_{nm}(f)_C.$$

Now let $X = L_p[a,b]$. Let K be some collection of intervals $\Delta_i = [\alpha_i, \beta_i] \subset [a,b]$ such that Δ_i does not contain a zero of q. Then, by (10), (6), (2), we have

$$\left\{\int_K |f(x)-r(x)|^p \mathrm{d}x\right\}^{1/p}$$
$$\leqslant \|f - r_{N_i}\|_p + \left\{\int_K |r(x) - r_{N_i}(x)|^p \mathrm{d}x\right\}^{1/p}$$
$$\leqslant \|f - r_{N_i}\|_p + (\mathrm{mes}(K))^{1/p}\{\|p\|_C \|q - q_{N_i}\|_C + \|q\|_C \|p - p_{N_i}\|_C\}$$
$$\xrightarrow[N_i \to \infty]{} R_{nm}(f)_p,$$

i.e.

$$\left\{\int_K |f(x)-r(x)|^p \mathrm{d}x\right\}^{1/p} \leqslant R_{nm}(f)_p$$

for every such compact K. Since the number of the zeros of q is finite, it follows from the definition of the Lebesgue integral that $\|f - r\|_p \leqslant R_{nm}(f)_p$, and since $r \in R_{nm}$ we must have

$$\|f - r\|_p = R_{nm}(f)_p. \qquad \square$$

The proof of this existence theorem shows the difficulties which arise when we work with rational functions. Roughly speaking, we must think in terms of the poles of the rational function – the proof of theorem 2.1 is so long because we have to consider the poles of r. Indeed it follows from the proof that in the uniform case it is not possible that r has poles on $[a,b]$, because, if q has a zero, on $[a,b]$, p should have the same zero at least with the same multiplicity. But from here follows the possibility for the best rational approximation r to be degenerate: this means that $p \in P_{n-1}$, $q \in P_{m-1}$ if $r = p/q \in R_{nm}$.

We shall see that in questions connected with the continuity of the metric projection in $C[a,b]$ on R_{nm} this possibility of degeneracy will be the main problem.

2.2 Uniqueness and characterization of the best uniform approximation

We have seen that if $f \in C[a,b]$ then there exists a rational function $r \in R_{nm}$ of best uniform approximation. The set of rational functions R_{nm} is a nonlinear one; nevertheless it still has uniqueness of the rational function of best uniform approximation and also characterization of this best approximation by means of alternation. In order to formulate this theorem we shall need the notion of the defect of a rational function.

Let $r \in R_{nm}$ and the reduced form of r be $r = p/q$, i.e. p and q have no common zeros. The defect $d(r)$ of r is given by

$$d(r) = \begin{cases} \min\{n - \deg p, m - \deg q\}, & r \not\equiv 0 \\ m, & r \equiv 0, \end{cases}$$

where $\deg p$ denotes the exact degree of the algebraic polynomial p ($\deg p = k$ if $p \in P_k$ and $p \notin P_{k-1}$).

It follows directly from the definition that:

(a) *r is degenerate if and only if $d(r) > 0$;*

(b) *$d(r)$ is the greatest number s for which $r \in R_{(n-s)(m-s)}$.*

Theorem 2.2 *Let $f \in C[a,b]$. For all natural numbers n and m the rational function $r \in R_{nm}$ is a rational function of best uniform approximation to f of order (n,m) if and only if there exist $N = n + m + 2 - d(r)$ points x_i, $i = 1, \ldots, N$, $a \leqslant x_1 < x_2 < \cdots < x_N \leqslant b$, such that*

$$f(x_i) - r(x_i) = \varepsilon(-1)^i \| f - r \|_{C[a,b]}, \quad i = 1, \ldots, N, \ \varepsilon = \pm 1.$$

Moreover the rational function of order (n,m) of best uniform approximation to f is unique.

In other words r is the rational function of order (n,m) of best uniform approximation to f if and only if $f - r$ alternates at least $n + m + 2 - d(r)$ times in the interval $[a,b]$.

Before proving theorem 2.2 we shall give some lemmas.

Lemma 2.1. *Let $\varphi \in C^1[a,b]$ and let x_i, $i = 1, \ldots, k+1$, $a \leqslant x_1 < x_2 < \cdots < x_{k+1} \leqslant b$, be $k+1$ different points in the interval $[a,b]$ such that*

$$\begin{aligned} &\varphi(x_1) \neq 0, \varphi(x_2) = \cdots = \varphi(x_k) = 0, \quad \varphi(x_{k+1}) \neq 0, \\ &\operatorname{sign} \varphi(x_1) = (-1)^k \operatorname{sign} \varphi(x_{k+1}). \end{aligned} \tag{1}$$

Then φ has at least k zeros on (x_1, x_{k+1}), if we compute every zero with its multiplicity.

Proof. The function φ has $k-1$ zeros on (x_1, x_{k+1}) $x_2, x_3, \ldots, x_k$. We must show that there exists in (x_1, x_{k+1}) a zero z of φ, different from $x_2, x_3, \ldots, x_k$, or that one of the zeros $x_2, \ldots, x_k$ has multiplicity at least 2.

If there does not exist a zero of φ in (x_1, x_{k+1}) different from $x_2, \ldots, x_k$, then in each interval (x_i, x_{i+1}), $i = 1, \ldots, k$, the function φ has constant sign. If the sign of φ is the same in two adjacent intervals (x_i, x_{i+1}), (x_{i+1}, x_{i+2}), then x_{i+1} must be at least a double zero of φ, since $\varphi \in C^1[a,b]$. If we assume that in all adjacent intervals (x_i, x_{i+1}), (x_{i+1}, x_{i+2}), $i = 1, \ldots, k-1$, φ has a different sign, we obtain that $\operatorname{sign} \varphi(x_1) = (-1)^{k+1} \operatorname{sign} \varphi(x_{k+1})$ and we come to contradiction with the condition (1) of the lemma. □

Lemma 2.2. *Let $\{\varphi_i\}_{i=1}^n$ be a Chebyshev system on the interval $[a,b]$,*

$\varphi_i \in C^1[a,b]$, $i=1,\ldots,n$, and $G=\{\varphi: \varphi\sum_{i=1}^{n} a_i\varphi_i\}$ be the Haar subspace, generated by $\varphi_1,\ldots,\varphi_n$. Let $x_i, i=0,\ldots,n$, $a \leqslant x_0 < x_1 < \cdots < x_n \leqslant b$, be $n+1$ different points on $[a,b]$. If for $\varphi\in G$ we have

$$(-1)^i\varphi(x_i) \geqslant 0, \quad i=0,\ldots,n,$$

or

$$(-1)^i\varphi(x_i) \leqslant 0, \quad i=0,\ldots,n,$$

then $\varphi \equiv 0$.

Proof. Let us assume that $\varphi \not\equiv 0$. Let us have for example

$$(-1)^i\varphi(x_i) \geqslant 0, \quad i=0,\ldots,n. \tag{2}$$

We shall prove that φ has at least n zeros in the interval $[a,b]$, every zero counted with its multiplicity, which contradicts the assumption of the lemma, that $\{\varphi_i\}_{i=1}^{n}$ is a Chebyshev system on $[a,b]$.

If $\varphi(x_i) \neq 0$, $i=0,\ldots,n$, from (2) and the continuity of the function φ it follows at once that φ has at least n zeros in $[a,b]$. Let now $\varphi(x_i)=0$ for some i. If $\varphi(x_i)\neq 0$ for only one value of i, then the same result follows. There remains the case when $\varphi(x_i)\neq 0$ for at least two values of i. Let the first two be j and $j+k$, i.e.

$$\left.\begin{aligned} \varphi(x_0) = \cdots = \varphi(x_{j-1}) = 0, \quad \varphi(x_j) \neq 0, \\ \varphi(x_{j+1}) = \cdots = \varphi(x_{j+k-1}) = 0, \quad \varphi(x_{j+k}) \neq 0. \end{aligned}\right\} \tag{3}$$

From the hypothesis (2) it follows that

$$\operatorname{sign}\varphi(x_j) = (-1)^k \operatorname{sign}\varphi(x_{j+k}). \tag{4}$$

Since $\varphi\in C^1[a,b]$, from (3), (4) and lemma 2.1 it follows that φ has at least k zeros in the interval (x_j, x_{j+k}) and therefore φ has at least $j+k$ zeros in the interval $[a, x_{j+k}]$. Going on in this way, we obtain that there exist n zeros of φ in $[a,x_n]$, every zero counted with its multiplicity. □

In the proof of theorem 2.2 we shall use also the following modification of the well-known Vallée-Poussin theorem for polynomials.

Theorem 2.3. *Let $f\in C[a,b]$. Let $p\in P_n$, $q\in P_m$ and let q have no zeros on $[a,b]$. Let there exist $N=n+m+2-d(p/q)$ points $\{x_i\}_{i=1}^{N}$, $a\leqslant x_1 < x_2 < \cdots < x_N \leqslant b$, in $[a,b]$ such that*

$$f(x_i) - \frac{p(x_i)}{q(x_i)} = \varepsilon(-1)^i\lambda_i, \quad \varepsilon = \pm 1, \lambda_i > 0, i=1,\ldots,N. \tag{5}$$

Then

$$R_{nm}(f)_{C[a,b]} \geqslant \min\{\lambda_i : i=1,\ldots,N\}.$$

Proof. Let us assume that there exists a rational function $r_1 = p_1/q_1 \in R_{nm}$, (p_1/q_1)-irreducible, such that

$$\| f - r_1 \|_{C[a,b]} < \min \{\lambda_i : i = 1, \ldots, N\}. \tag{6}$$

Let us consider the values of the difference $s = p/q - r_1$ at the points x_i, $i = 1, \ldots, N$. We obtain from (5) and (6) that

$$\operatorname{sign} s(x_i) = \operatorname{sign} \left\{ \left(\frac{p(x_i)}{q(x_i)} - f(x_i) \right) - (r_1(x_i) - f(x_i)) \right\}$$

$$= \operatorname{sign} \left(\frac{p(x_i)}{q(x_i)} - f(x_i) \right) = \varepsilon(-1)^{i+1}, \quad i = 1, \ldots, N.$$

Hence s has at least $N - 1$ different zeros y_i, $i = 1, \ldots, N - 1$, in the interval $[a, b]$, i.e.

$$s(y_i) = 0, \quad i = 1, \ldots, N - 1.$$

Let us note now that $r_1 = p_1/q_1$ has a reduced form and $\| r_1 \|_{C[a,b]} < \infty$, and consequently q_1 has no zeros on $[a, b]$. So from

$$s(y_i) = \frac{p(y_i)}{q(y_i)} - \frac{p_1(y_i)}{q_1(y_i)} = 0, \quad i = 1, \ldots, N - 1,$$

it follows that

$$p(y_i)q_1(y_i) - p_1(y_i)q(y_i) = 0, \quad i = 1, \ldots, N - 1,$$

i.e. the algebraic polynomial $pq_1 - p_1 q \in P_M$, $M \leqslant n + m - d(p/q) = N - 2$, has at least $N - 1 > M$ different zeros in the interval $[a, b]$. This contradiction proves the theorem. □

Let us mention that later on we shall use theorem 2.3 in the numerical method of Remez for finding the rational function of best approximation (see section 2.5).

Proof of theorem 2.2. First we shall prove that if $r \in R_{nm}$ realizes an alternation, then r is a rational function of best uniform approximation to f of order (n, m). If we apply theorem 2.3 to f and r with

$$\lambda_i = \lambda = \| f - r \|_{C[a,b]},$$

we obtain that $\lambda \leqslant R_{nm}(f)_{C[a,b]}$, and since $r \in R_{nm}$ we must really have $\lambda = \| f - r \|_{C[a,b]} = R_{nm}(f)$, i.e. r is a rational function of best uniform approximation to f of order (n, m).

Now let r be a rational function of best uniform approximation to f of order (n, m). We shall prove that $f - r$ must alternate at least $N = n + m +$

$2-d(r)$ times in $[a,b]$. Let us assume the opposite, that $M \leqslant N-1$ is the highest number of points $x_1 < x_2 < \cdots < x_M$ in $[a,b]$ such that

$$f(x_i)-r(x_i)=\varepsilon(-1)^i \| f-r \|_{C[a,b]} = \varepsilon(-1)^i R_{nm}(f)_{C[a,b]}, \quad i=1,\ldots,M,\ \varepsilon=\pm 1. \tag{7}$$

Then there exist $M+1$ points $\xi_i, i=0,\ldots,M$, $a=\xi_0<\xi_1<\cdots<\xi_M=b$ such that for every $x\in[\xi_{i-1},\xi_i]$ we have

$$\varepsilon(-1)^i(f(x)-r(x)) > -R_{nm}(f)_{C[a,b]}, \quad i=1,\ldots,M. \tag{8}$$

In view of (7) the continuous function $f-r$ changes its sign in $[x_i,x_{i+1}]$, therefore we can assume, as in section 1.2, that the points $\xi_i, i=1,\ldots,M-1$, are such that

$$f(\xi_i)-r(\xi_i)=0, \quad i=1,\ldots,M-1. \tag{9}$$

Let us consider the algebraic polynomial

$$s(x)=(-1)^M(x-\xi_1)\cdots(x-\xi_{M-1})\in P_{M-1}. \tag{10}$$

Let $r=p/q$ and p and q have no common zeros. Since $s\in P_{N-2}$, $p\in P_{n-d(r)}$, $q\in P_{m-d(r)}$, there exist two algebraic polynomials $p_1\in P_m, q_1\in P_n$, such that

$$s=pp_1-qq_1.$$

Let us consider the rational function

$$\tilde{r}=\frac{p-\varepsilon\delta q_1}{q-\varepsilon\delta p_1}\in R_{nm}, \tag{11}$$

where ε ($\varepsilon=1$ or -1) is the same as in (7), and δ, $\delta>0$, will be chosen later.

Since $\| f-r \|_{C[a,b]}<\infty$, p and q have no common zeros, and q has no zeros in $[a,b]$, we can find δ_1 so that for $0<\delta<\delta_1$ the polynomial $q-\varepsilon\delta p_1$ has the same sign as q in $[a,b]$.

Let us consider the difference $f-\tilde{r}$. We have

$$f-\tilde{r}=f-r+\frac{p}{q}-\frac{p-\varepsilon\delta q_1}{q-\varepsilon\delta p_1}=f-r-\frac{\varepsilon\delta(pp_1-qq_1)}{q(q-\varepsilon\delta p_1)}=f-r-\frac{\varepsilon\delta s}{q(q-\varepsilon\delta p_1)}.$$

Let $\delta_2\leqslant\delta_1$ be such that for δ, $0<\delta<\delta_2$, we have for $x\in[\xi_{i-1},\xi_i]$

$$\begin{aligned}\varepsilon(-1)^i(f(x)-\tilde{r}(x)) &= \varepsilon(-1)^i(f(x)-r(x)) \\ &\quad +(-1)^{i+1}\frac{\delta s(x)}{q(x)(q(x)-\varepsilon\delta p_1(x))} > -R_{nm}(f)_{C[a,b]}.\end{aligned} \tag{12}$$

This is possible in view of (8), since $f-r$ is a continuous function in $[a,b]$.

On the other hand for $x\in(\xi_{i-1},\xi_i)$, $i=1,\ldots,M$, $x=\xi_0$, $x=\xi_M$, we have for $0 < \delta < \delta_2$

$$\begin{aligned}\varepsilon(-1)^i(f(x)-\tilde{r}(x)) &= \varepsilon(-1)^i(f(x)-r(x)) - \varepsilon(-1)^i \frac{\varepsilon\delta s(x)}{q(x)(q(x)-\varepsilon\delta p_1(x))}\\ &\leqslant R_{mn}(f)_{C[a,b]} - \frac{\delta(-1)^i s(x)}{q(x)(q(x)-\varepsilon\delta p_1(x))}\\ &< R_{nm}(f)_{C[a,b]},\end{aligned} \tag{13}$$

since by (10) we have $(-1)^i s(x) > 0$ for $x \in (\xi_{i-1}, \xi_i)$, $i = 1, \ldots, M$, $x = \xi_0$, $x = \xi_M$.

In view of (9) we have also (13) for $x = \xi_i$, $i = 1, \ldots, M-1$, i.e. for all $x \in [a,b]$. Since $f - \tilde{r}$ is a continuous function on $[a,b]$, the inequalities (12) and (13) give us

$$\| f - \tilde{r} \|_{C[a,b]} < R_{nm}(f)_{C[a,b]}. \tag{14}$$

By (11) $\tilde{r} \in R_{nm}$, and therefore (14) is a contradiction. Consequently $f - r$ must alternate at least N times.

Now let us prove the uniqueness of the best rational approximation of order (n, m).

Let us assume that there exist two different rational functions $r_1 = p_1/q_1 \in R_{nm}$ and $r_2 = p_2/q_2 \in R_{nm}$ such that

$$\| f - r_1 \|_{C[a,b]} = \| f - r_2 \|_{C[a,b]} = R_{nm}(f)_{C[a,b]}.$$

We can assume that $r_1 = p_1/q_1$ and $r_2 = p_2/q_2$ have a reduced form and q_1, q_2 have no zeros in $[a,b]$.

Let $N_1 = n + m + 2 - d(r_1)$, $N_2 = n + m + 2 - d(r_2)$ and let us assume for definiteness that $N_1 \geqslant N_2$, or, which is the same, $d(r_1) \leqslant d(r_2)$. Let x_i, $i = 1, \ldots, N_1$, $a \leqslant x_1 < \cdots < x_{N_1} \leqslant b$, be the points of alternation for r_1, i.e.

$$f(x_i) - r_1(x_i) = \varepsilon(-1)^i R_{nm}(f)_{C[a,b]}, \quad i = 1, \ldots, N_1, \varepsilon = \pm 1. \tag{15}$$

Let us consider the difference $s = r_1 - r_2$ at the points x_i, $i = 1, \ldots, N_1$. There are two possibilities:

(a) $s(x_i) = 0$, $i = 1, \ldots, N_1$,
(b) $s(x_i) \neq 0$ for some i.

In case (b), since $|f(x_i) - r_1(x_i)| = R_{nm}(f)$, we must have

$$\tilde{\varepsilon}(r_2(x_i) - f(x_i)) < \tilde{\varepsilon}(r_1(x_i) - f(x_i)), \quad \tilde{\varepsilon} = \operatorname{sign}(r_1(x_i) - f(x_i)),$$

and therefore

$$\operatorname{sign} s(x_i) = \operatorname{sign}(r_1(x_i) - f(x_i)). \tag{16}$$

From (15) and (16) it follows that

$$\varepsilon(-1)^{i+1} s(x_i) \geqslant 0, \quad i = 1, 2, \ldots, N_1. \tag{17}$$