Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Martin Franců

Elementary problems of approximation theory

katedra matematické analýzy

Vedoucí bakalářské práce: doc. RNDr. Luboš Pick, CSc., DSc.,
katedra matematické analýzy

Studijní program: matematika, obecná matematika

2009

Rád bych poděkoval svému vedoucímu doc. RNDr. Lubošovi Pickovi, CSc., DSc. za obětavou péči a pomoc. Zároveň směřují mé díky i všem vynikajícím pedagogům, kteří mě vedli studiem. Samozřejmě děkuji i své rodině za duševní i hmotnou podporu.

Prohlašuji, že jsem svou bakalářskou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce a jejím zveřejňováním.

V Praze dne 6.8. 2009 Martin Franců

Contents

Název práce: Elementární problémy teorie aproximace
Autor: Martin Franců
Katedra (stav): katedra matematické analýzy
Vedoucí bakalářské práce: doc. RNDr. Luboš Pick, CSc., DSc., katedra matematické analýzy
e-mail vedoucího: pick@karlin.mff.cuni.cz

Abstrakt: V této práci studujeme vlastnosti aproximací spojitých funkcí při supremové normě. Ve středu pozornosti stojí dvě věty - alternační věta a de La Vallée Poussinova věta. V prvním případě jsou tvrzení a aplikace této věty uvedeny na specialních případě. V druhém případě je nejprve vysloveno a dokázáno její zobecnění. Důkazy obecně známých verzí jsou pak jednodušší k sledování. Obě věty jsou úzce spjaty s Haarovými systémy. Pojem Haarovy podmínky je samozřejmě také uveden a rozebrán. V dané kapitole jsou také popsány některé základní vlastnosti Haarových systémů.

Klíčová slova: aproximace, Haarovy systémy, alternační věta, de La Vallée Poussinova věta

Title: Elementary problems of approximation theory
Author: Martin Franců
Department: Department of Mathematical Analysis
Supervisor: doc. RNDr. Luboš Pick, CSc., DSc.
Supervisor's e-mail address: pick@karlin.mff.cuni.cz

Abstract: In this work we study properties of approximations of continuous functions with respect to the supremum norm. We focus on two theorems - the alternation theorem and de La Vallée Poussin's theorem. In the first case the statement and applications are illustrated on a special case. In the second case we prove a generalization of the theorem to illustrate the statement and its applications. Proofs of general versions are then easier to follow. Both theorems are connected to Haar systems. Notion of the Haar condition is brought in and examined, too. In the corresponding chapter we also examine some basic properties of the Haar systems.

Keywords: approximation, Haar systems, alternation theorem, theorem of de La Vallée Poussin

Introduction

Approximation is in general described as an inexact representation of something that is still close enough to be useful. The approximation theory then usually solves problems of finding and characterizing the best approximation of a given element, often a function, by elements of a given space. We will assume that the object we strive to approximate is a point in some normed space. That point will be approximated by an object from a given subspace of the normed space mentioned above. For the best approximation we take the nearest point, according to the metric induced by the norm, from our subspace to the given point.

In this thesis we shall deal with problem of approximation of a continuous function on some compact interval, by functions from certain linear space. However most of our conclusions could be easily generalized to the space of continuous functions on any compact metric space. Our interest is directed to approximations which minimize the supremum norm.

Our main goal is to survey theorems on characterization of the best approximation. We shall concentrate on Haar systems and linear spaces spanned by them.

A special attention will be given to a problem of approximation of continuous functions defined on compact interval by polynomials.

Basic notions are introduced in the following section. In the first chapter a special case of the alternation theorem and generalization of de La Vallée Poussin's theorem are proved. The main subject - Haar condition - is studied in the second chapter. Examples of the Haar systems and also some famous theorems will be mentioned.

Basic notions

We shall start with definitions of several terms and symbols and specify facts which will be used in all other chapters.

Notation 0.0.1 (Numbers)**.** *The field of real numbers will be denoted by* $\mathbb{R}$*, the set of positive integers* $\mathbb{N}$*.*

For the sake of completeness let us recall the classical definition of a continuous function.

Definition 0.0.2 (Continuous function)**.** *Function $f : [a, b] \mapsto \mathbb{R}$ is said to be continuous if*

$$\forall x, y \in [a, b], \forall \epsilon > 0, \exists \delta > 0 : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

The space of all continuous functions defined on $[a, b]$ will be denoted as $C[a, b]$.

We shall need the following notions:

Definition 0.0.3 (Error function, Best approximation)**.** *Let f be a function and P its approximation. Then the error function r is defined by*

$$r(x) \equiv f(x) - P(x).$$

The best approximation of f is such approximation P that minimize the norm of the error function r.

Throughout this thesis the supremum norm will be used:

Definition 0.0.4 (Supremum norm)**.** *The space of continuous functions on $[a, b]$ is equipped with the supremum norm, defined as*

$$\|f\| = \max_{x \in [a,b]} |f(x)|.$$

The advantage of using this norm consists in the fact that it gives a very good control over the error of approximation at every point from $[a, b]$.

In this paper we shall use polynomials and generalized polynomials, let us introduce their definition.

Definition 0.0.5 (Polynomial)**.** *A finite linear combination of functions $1, x, x^2, \ldots$ is called a polynomial. Equivalently, let $n \in \mathbb{N}$, let $c_1, c_2, \ldots, c_n \in \mathbb{R}$, then*

$$\sum_{i=0}^{n} c_i x^i$$

is a polynomial.

Chapter 1

General approximation

1.1 Alternation theorem for regular polynomials

This chapter will deal with a modification of the alternation theorem and examples of its usage. The common alternation theorem will be studied in a separate chapter. The version of alternation theorem treated here is weaker than the general one, because it considers only approximations by polynomials. On the other hand, the proof of the simpler theorem is straight and requires only elementary considerations, while a major part of applications are still valid.
We shall prove the common alternation theorem by the same technique used for this simplified version. Having proved the theorem in its simple version, it is easier to follow the proof of the general statement.

Theorem 1.1.1 (alternation theorem)**.** *Let $a, b \in \mathbb{R}$ and let f be continuous on $[a, b]$. In order that a polynomial P is the best approximation on $[a, b]$ to f from polynomials with $\deg P < n$ it is necessary and sufficient that there exist at least $n + 1$ points $x_1, x_2, \ldots, x_{n+1}$ ($x_1 < x_2 < \cdots < x_{n+1}$) such that for the error function $r = f - P$ we have $r(x_i) = -r(x_{i-1})$, $i = 2, \ldots, n + 1$, and $|r(x_j)| = \|r\|$, $j = 1, \ldots, n + 1$.*

Proof. Let P be the best approximation of f. Let $x_1 \in [a, b]$ be the nearest point to a such that $|r(x_1)| = \|r\|$. Then we denote by x_i the point nearest to x_{i-1} from $[x_{i-1}, b]$ such that $r(x_i) = -r(x_{i-1}) = \pm \|r\|$ as far as such point exists. In this way we construct a finite sequence $x_1, \ldots, x_k$. Assume for a contradiction that $k < n + 1$. We denote by $z_i \in [a, x_{i+1}]$ the nearest point to x_{i+1} satisfying $r(z_i) = 0$ for $i = 1, \dots, k-1$, and we add $z_0 = a$ and $z_k = b$. We also write

$$h(x) = \prod_{i=1}^{k-1} (x - z_i).$$

We note that, for all $x \in [z_{i-1}, z_i]$ and for all $i \in \{1, \dots, k\}$, we have $h(x)(-1)^{i+1} > 0$, that is, the function $h(x)$ alternates it's sign. If there is an $x \in [z_{i-1}, z_i]$ such that $r(x) = \pm \|r\|$, then there cannot be $y \in [z_{i-1}, z_i]$ such that $r(y) = \mp \|r\|$. The intervals $[z_{i-1}, z_i]$ are defined in such a way that r on every such interval attains either the value $\|r\|$ or $-\|r\|$, but never both. Therefore, a positive multiple of the polynomial $h(x)$ will change it's sign in the same way. Using this fact, we can minimize the function $r(x)$. The exact multiple shall be found in the following.

The interval I is said to be *rude* if $I \subset [z_{i-1}, z_i]$ for some $i = 1, \dots, k+1$, and

$$r(x)r(q) < 0 \qquad \text{for every } z \in I,$$

where $q \in [z_{i-1}, z_i]$ is such a point that $r(q) = \pm \|r\|$. Let R be the set of all rude intervals.

We denote

$$d = \sup_{I \in R} \|r\|_{C[I]},$$

$$N = \|h\|_{C[a,b]},$$

and

$$s = \|r\|_{C[a,b]} - d.$$

We claim that if $0 < \lambda < \frac{s}{N}$, then $P + \lambda h$ is a better approximation of f than P. (We are presuming that, for the first point where $|r(x)| = \|r\|$ is valid, we have $r(x) = \|r\|$. Otherwise we take $P - \lambda h$ instead of $P + \lambda h$.)

Denote

$$r_h(x) = f(x) - (P(x) + \lambda h(x)).$$

Let y be any point of $[a, b]$. Then there exists an index i such that $y \in [z_{i-1}, z_i]$. According to our observation, $r(x)$ can attain either the value $\|r\|$ or $-\|r\|$. Let us assume that it is $+\|r\|$. If $r(y) > 0$, then

$$|r_h(y)| = |f(y) - P(y) - \lambda(h(y))| = |r(y) - \lambda h(y)|.$$

If $r(y) > \lambda h(y)$, we get

$$|r(y) - \lambda h(y)| < |r(y)| \leq \|r\|,$$

otherwise

$$|r(y) - \lambda h(y)| \le |\lambda h(y)| \le |\lambda N| \le \left|\frac{s}{N}N\right| < \|r\|.$$

If $r(y) < 0$, then

$$|r_h(y)| = |r(y) - \lambda h(y)| < d + \frac{s}{N}N = \|r\|.$$

The above inequalities are valid for each $y \in [a, b]$, whence $\|r_h\| < \|r\|$ and $P + \lambda h$ is a better approximation to f than P. Since $k < n+1$, $P + \lambda h$ is a polynomial of degree $< n$, and P cannot be the best approximation of f.

Let P be a polynomial with such alternating points. Assume for a contradiction that there exists a polynomial R which is a better approximation. Then the polynomial $P - R$ must alternate its sign $n+1$ times and consequently ought to have n roots. But that contradicts $\deg < n$. □

Problem 1.1.2. *Find the best approximation of the function* e^x *from the space of polynomials with* $\deg < 2$ *on* $[0, 1]$.

Solution. Our desired polynomial of the best approximation will be denoted by $P(x) = a + bx$. Then, by a consequence of the alternation theorem, there must exist three points $x_1, x_2, x_3 \in [0, 1]$ for which the error function $r(x) = e^x - P(x)$ attains, in the absolute value, it's maximum, and it satisfies:

$$r(x_1) = -r(x_2) = r(x_3).$$

At first we will find the points x_1, x_2, x_3, then, using the identities above, we can compute exactly the coefficients a and b.

The function $r(x)$ is a subtract of two continuous functions with continuous derivatives, hence it is also continuous with continuous derivatives. Consequently, $r(x)$ can attain it's maximum, or minimum, only at points where its derivative equals zero or at points 0 or 1. For the points which obey first condition we have

$$r^{(1)}(x) = e^x - b = 0 \Rightarrow x = \ln(b),$$

$$r^{(2)}(\ln(b)) = e^{\ln(b)} = b.$$

If $\ln(b) \in [0, 1]$, then the point $\ln(b)$ is a local minimum of the function $r(x)$. It is clear that inside $[a, b]$ there is no more than one extreme, this fact

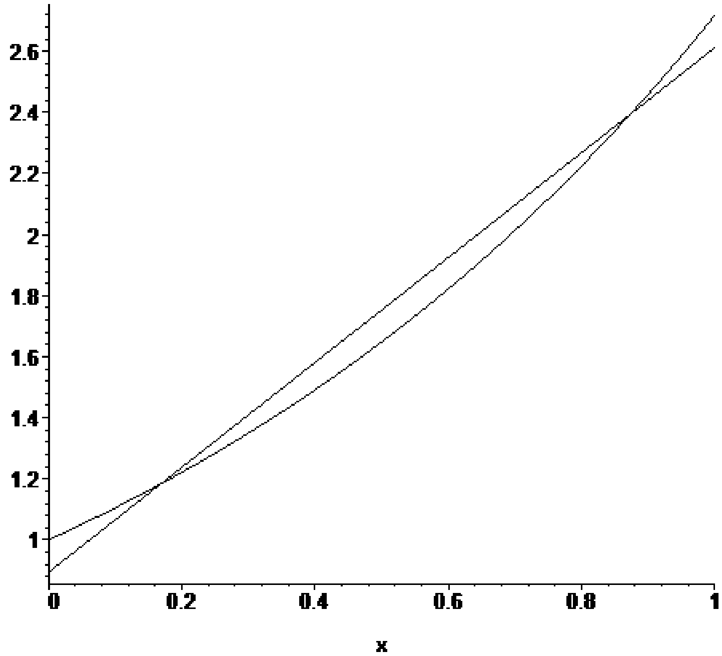

Figure 1.1: The best approximation of e^x by linear function on interval $[0, 1]$

implies that 0 and 1 must be also extremal points. Because $r(x)$ need to have three extremes on $[0, 1]$, therefore $\ln(b) \in [0, 1]$, however we still don't know exact value of b. Consequently:

$$x_1 = 0;\ x_2 = \ln(b);\ x_3 = 1;$$

also:

$$r(0) = -r(\ln(b)) = r(1).$$

Using $r(0) = r(1)$, we can determine value of b:

$$e^0 - a - b0 = e^1 - a - b \ \Rightarrow b = e - 1$$

a can be found from the equation $r(0) = -r(\ln(b))$:

$$e^0 - a = -e^{\ln(e-1)} + a + (e-1)\ln(e-1) \ \Rightarrow a = \frac{1 + (e-1)(1 - \ln(e-1))}{2}.$$

The best approximation of the function e^x from the space of linear functions on $[0, 1]$ is:

$$P(x) = \frac{1 + (e-1)(1 - \ln(e-1))}{2} + (e-1)x.$$

□

Problem 1.1.3. *Let P be the best approximation of $\sqrt{x}$ on $[0,1]$ from the space of polynomials with* $\deg \leq n$. *We can determine the signs of the coefficients of P.*

Solution of this problem needs one more theorem - the Descartes rule of signs - which describes the relationship between signs of coefficients and roots of a polynomial. The following statement is taken from [2].

Theorem 1.1.4 (Descartes rule of signs)**.** *Let $h(x) = \sum_{i=0}^{n} c_i x^i$ be a polynomial and let k be a number of alternations of signs of nonzero coefficients of the polynomial h ordered by descending variable exponent. Then the number of positive roots of h (where multiple roots of the same value are counted separately) equals either k or k from which we subtract an integral multiple of 2.*

Solution of Problem 1.1.3. If P is the best approximation of $\sqrt{x}$ on $[0,1]$, then, according to the alternation theorem, the function $P(x) - \sqrt{x}$ alternates its signs at least $n+2$ times, hence it has $n+1$ roots. The polynomial $Q(x) = P(x^2) - x = (-1)x + \sum_{k=0}^{n} c_i x^{2k}$ has the same number of roots in $[0,1]$. The polynomial $Q(x)$ has $n+2$ nonzero coefficients and needs to have $n+1$ alternations of their signs, because of the number of its positive roots and by the Descartes rule of signs. In result each of the subsequent coefficients of $Q(x)$ needs to alternate its signs. Because the sign of the coefficient at x^1 is set to -1, the other signs are also determined. Consequently:

$$\operatorname{sign} c_0 = 1; \quad \operatorname{sign} c_i = (-1)^{i-1}; \quad 0 < i < n+1$$

□

1.2 Generalized de La Vallée Poussin's theorem

Theorem of de La Vallée Poussin uses the concept of a Haar condition which we didn't define yet. (This subject will be studied in the section on Haar systems.) However, there exists a generalization of the de La Vallée Poussin theorem which does not need the definition of a Haar system. This theorem will be formulated and proved right now, taking advantage of it's general assumptions. Before we proceed we should introduce two definitions.

Definition 1.2.1 (Generalized polynomial). *Let* $F \equiv \{f_1, f_2, \ldots, f_n\}$, $f_i \in C[a,b]; a, b \in \mathbb{R}\}$ *be a system of functions. Then,*

$$\sum_{i=1}^{n} c_i f_i, \quad c_i \in \mathbb{R},$$

will be called a generalized polynomial over the system of functions F*. The polynomial such that* $c_i = 0$ *for all* i *will be called a trivial polynomial.*

Remark 1.2.2. *It is clear that a finite sum and a finite linear combination of generalized polynomials is still a generalized polynomial.*

Definition 1.2.3. *Let* $x_1, x_2, \ldots, x_n$ *be points of any linear space* M*. Then the convex hull of points* $x_1, x_2, \ldots, x_n$ *is the set* $\{x \in M : \exists \lambda_i$ *for all* $i \in \{1, 2, \ldots, n\};\ x = \sum_i \lambda_i x_i;\ \sum_i \lambda_i = 1;\ \lambda_i \geq 0\}$.

Notation 1.2.4. *If* $G \equiv \{g_1, g_2, \ldots, g_n\}$ *is a system of functions, we introduce the vector*

$$\hat{x}^G = [g_1(x), g_2(x), \ldots, g_n(x)]\,.$$

In situations where it will be clear which system of functions is used to create the vector $\hat{x}^G$ *we shall write just* $\hat{x}$*.*

Theorem 1.2.5 (Generalized de La Vallée Poussin's theorem). *Let* $f \in \mathbf{C}\,[a,b]$ *function,* $\{g_1, g_2, \ldots, g_n\}$ $(g_i \in \mathbf{C}\,[a,b])$ *be a system of functions and let* P *be any generalized polynomial over the given system. Let* $x_1, x_2, \ldots, x_k$ *be any points from* $[a,b]$*. If the origin of* $\mathbb{R}^n$ *lays in the convex hull of vectors*

$$\{r(x_i) \cdot \hat{x}_i = (f(x_i) - P(x_i)) \cdot \hat{x}_i | i \in \{1, 2, \ldots, k\}\},$$

where $\hat{x}_i = [g_1(x_i), g_2(x_i), \ldots, g_n(x_i)] \in \mathbb{R}^n$*, then*

$$\inf_R \|f - R\| \geq \min_{i=1,2,\ldots,k} |r(x_i)|\,,$$

where R *ranges over all generalized polynomials from* $\{g_1, g_2, \ldots, g_n\}$*.*

In order to prove the previous theorem we shall need a criterion for testing whether the origin lies in the convex hull of a given set. The following lemma can be found in [1].

Lemma 1.2.6 (Lemma on linear inequalities)**.** *Let U be a compact subset of $\mathbb{R}^n$. A necessary and sufficient condition for the set*

$$\{z \in \mathbb{R}^n |\, \langle z, u\rangle > 0 \, \forall u \in U\},$$

being empty is that the origin of $\mathbb{R}^n$ lies in the convex hull of U.

Proof of the generalized De la Vallée Poussin theorem 1.2.5. Our proof will be indirect. Starting from the negation of the assertion we arrive at the negation of the assumption. Suppose that

$$\inf_R \|f - R\| < \min_{i \in \{1,2,\ldots,k\}} |r(x_i)| \,,$$

then there exists $d = [d_1, d_2, \ldots, d_n]$ such that, for the generalized polynomial $S = \sum_{i=1}^n (c_i + d_i) g_i(x)$, the following inequality holds:

$$\|f - S\| < \min_{i \in \{1,2,\ldots,k\}} |r(x_i)| \,.$$

Thus, for each $i \in \{1, 2, \ldots, k\}$, we get

$$|f(x_i) - S(x_i)| = \left| f(x_i) - \sum_{i=1}^{n} (c_i + d_i) g_i(x) \right| < |f(x_i) - P(x_i)| \,.$$

Equivalently,

$$\left| r(x_i) - \sum_{j=1}^{n} d_j g_j(x_i) \right| < |r(x_i)| \,.$$

If the previous inequality is valid, then $r(x_i)$ must have the same sign as $\sum_{j=1}^n d_j g_j(x_i)$ for each $i \in \{1, 2, \ldots, k\}$, and we get the following system of inequalities:

$$r(x_i) \sum_{j=1}^{n} d_j g_j(x_i) = r(x_i) \cdot \langle d, \hat{x}_i \rangle = \langle d, r(x_i) \cdot \hat{x}_i \rangle > 0.$$

Now we use the Lemma on linear inequalities (Lemma 1.2.6). In this case the set U from the lemma will be the set $H \equiv \{r(x_i) \cdot \hat{x}_i | i \in \{1, 2, \ldots, k\}\}$. H is a compact subset of $\mathbb{R}^n$, because it is finite. The vector d is a solution of the system of previous inequalities, therefore the origin of $\mathbb{R}^n$ does not belong to the convex hull of H. $\square$

In following example we shall use the previous theorem to prove that a given approximation is the best approximation from the given space. Because any function without root is a Haar system we do not need the generalized theorem, the regular theorem of de La Vallée Poussin would be enough.

Problem 1.2.7. *Find the best approximation of the function x on $[0,2]$ by a constant function.*

Solution. It seems to be clear that the best approximation will be constant 1, but we should prove that no better approximation is available in the given space. We simply take points 0 and 2. The error function $r(x) = x - 1$ takes values -1 and 1, therefore zero is in convex hull of points $-1 \cdot 1$ and $1 \cdot 1$. The assumptions of the previous theorem are fulfilled and we obtain that the distance of the function x from the space of constants is 1. 1 is also the distance of the function x from our approximation, consequently 1 is the best approximation of function x on interval $[0,2]$.

□

Problem 1.2.8. *We have two functions, g_1, g_2, specified below, and we need to obtain some estimate on the best approximation of the function $\cos(x)$ from the space of linear combinations of our functions on $\left[0, \frac{\pi}{2}\right]$. We have*

$$g_1(x) = 1 - \frac{2x}{\pi},\ g_2(x) = \begin{cases} 1 & x \in \left[0, \frac{\pi}{4}\right] \\ 2 - \frac{4x}{\pi} & x \in \left(\frac{\pi}{4}, \frac{\pi}{2}\right] \end{cases} .$$

Solution. We shall use the generalized de La Vallée Poussin's theorem (Theorem 1.2.5) to obtain the estimate of the best approximation.

First, we must choose a testing polynomial, then we need to pick points such that the origin of $\mathbb{R}^2$ lies in the convex hull of $\{r(x_i) \cdot \hat{x}_i\}$. Let the testing polynomial be $\frac{1}{2}(g_1 + g_2)$ and let the points be $\left\{\frac{\pi}{6}, \frac{\pi}{4}, \frac{5\pi}{12}\right\}$. The corresponding vectors are

$$\frac{\hat{\pi}}{6} = \left[\frac{2}{3}, 1\right],\ \frac{\hat{\pi}}{4} = \left[\frac{1}{2}, 1\right],\ \frac{\widehat{5\pi}}{12} = \left[\frac{1}{6}, \frac{2}{6}\right],$$

and the corresponding values of the error function are

$$r\left(\frac{\pi}{6}\right) = \cos\left(\frac{\pi}{6}\right) - \frac{1}{2}\left(\frac{2}{3} + 1\right) \doteq 0.8660 - 0.8334 = 0.0326\,,$$

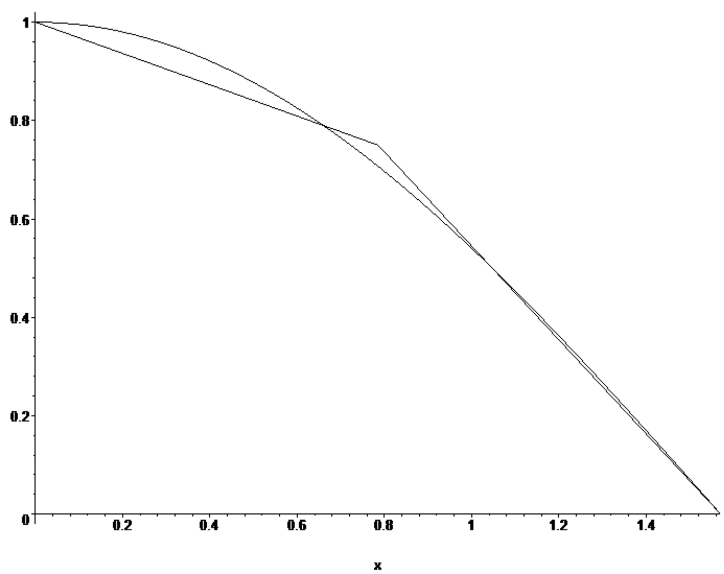

Figure 1.2: The Approximation of $\cos(x)$ by the function $\frac{1}{2}(g_1(x) + g_2(x))$.

$$r\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) - \frac{1}{2}\left(\frac{1}{2} + 1\right) \doteq 0.7071 - 0.7500 = -0.0429\,,$$

$$r\left(\frac{5\pi}{12}\right) = \cos\left(\frac{5\pi}{12}\right) - \frac{1}{2}\left(\frac{1}{6} + \frac{2}{6}\right) \doteq 0.2588 - 0.25 = 0.0088.$$

Because of the linear dependence of the vectors $\hat{\frac{\pi}{4}}$ and $\hat{\frac{5\pi}{12}}$ and the different signs of the error function in mentioned points we do not need the first point $\frac{\pi}{6}$. Using Theorem 1.2.5 we obtain

$$\|\cos(x) - (ag_1(x) + bg_2(x))\| \geq 0.008,$$

where $a, b \in \mathbb{R}$. □

Our estimate could be improved by modifying our testing polynomial and by choosing other points. The example illustrates an application of the theorem.

Chapter 2

Approximation with the Haar condition

2.1 Notion of the Haar system

So far we have either strictly specified the subspace of approximating functions or imposed no restrictions on it's properties. In this chapter we introduce the notion of a Haar system. It is, in principle, a generalization of a concept of linear independence to continuous functions. The system satisfying the Haar condition will be called the Haar system. The Haar systems are in some sense non degenerated.

Definition 2.1.1 (Haar condition)**.** *A system of functions $H \equiv \{f_1, \ldots, f_n\}$, $f_i \in \mathbf{C}[a,b]; a, b \in \mathbb{R}$ is said to satisfy the Haar condition if no nontrivial generalized polynomial has more than $n-1$ roots. The system H could also be called a Haar system.*

Although the previous definition is quite elegant, it is often reasonable to use another one, which is equivalent to the original one. This fact is rather well illustrated by the following lemma.

Lemma 2.1.2. *Let H be a system of functions $\{f_1, f_2, \ldots, f_n\}$, $f_i \in \mathbf{C}[a,b]$; $a, b \in \mathbb{R}$. Then the following conditions are equivalent:*
1) H satisfies the Haar condition.
2) For each $x_1, x_2, \ldots x_n$; $x_j \in [a,b]$;, $\det f_i(x_j) \neq 0$.

Proof. That H satisfies the Haar condition means that no nontrivial generalized polynomial has more than $n-1$ roots. The previous condition is equivalent to saying that for all $c_1, c_2, \ldots, c_n; c_i \in \mathbf{R}$ there do not exist points $x_1, x_2, \ldots, x_n; x_j \in [a, b]$ that $\sum_i c_i f_i(x_j) = 0$ for $j = 1, 2, \ldots n$. This is true if and only if the vectors $[f_1(x_j), f_2(x_j), \ldots, f_n(x_j)]$ are linearly independent for $j = 1, 2, \ldots n$. The previous statement is equivalent to $\det f_i(x_j) \neq 0$.

□

In order to build up a better insight into the meaning of the definition of the Haar condition we will determine in some special cases whether the chosen system satisfies the Haar condition on some specified interval.

Remark 2.1.3 ($S = \{1, x, \ldots, x^n\}$ is a Haar system on any interval)**.** *According to a basic knowledge from algebra, no nontrivial polynomial of* $\deg < n + 1$ *has more than* n *roots, hence* $\{1, x, \ldots, x^n\}$ *is a Haar system on any interval (together with a constant function, this amounts to* $n + 1$ *functions in* S*).*

At this point we should also emphasize, in particular, that satisfaction of the Haar condition is closely associated with the interval on which the functions are defined. This fact will be illustrated by the following two remarks.

Remark 2.1.4 ($\{1, x^2, \ldots, x^{2n}\}$ is a Haar system on $[0, c]\,; c \in \mathbb{R}^+$)**.** *The polynomial satisfying* $\deg = 2n$ *can have at most* $2n$ *roots in* $\mathbb{R}$*. Each polynomial* $P(x)$ *over our system has the following property: if* a *is a root of* $P(x)$*, then also* $-a$ *is a root of* $P(x)$*. Consequently,* $P(x)$ *cannot have* $n+1$ *roots in any of* $[0, c]\,; c \in \mathbb{R}^+$*.*

Remark 2.1.5 ($\{1, x^2, \ldots, x^{2n}\}$ is not a Haar system on $[-c, c]\,; c \in \mathbb{R}^+$)**.** *Consider the polynomial*

$$P(x) = \prod_{i=1}^{n} \left(x^2 - \left(\frac{c}{i}\right)^2 \right).$$

Then $P(x)$ *has evidently* $2n$ *roots:* $-c, -\frac{c}{2}, \ldots, -\frac{c}{n}, \frac{c}{n}, \ldots, \frac{c}{2}, c$*. Next,* $P(x)$ *is a generalized polynomial over the system* $\{1, x^2, \ldots, x^{2n}\}$ *because the power of* x *is always even.*

Lemma 2.1.6. *Let* f *be a continuous function on* $[a, b]$*,* $f^{(n)}(x) > 0$ *on* $[a, b]$*, then system of functions* $\{1, x, x^1, \ldots, x^{(n-1)}, f\}$ *is a Haar system.*

Proof. Suppose on contrary that $\{1, x, x^1, \dots, x^{(n-1)}, f\}$ is not a Haar system, so there exists a generalized polynomial $h(x) = \sum_{i=0}^{n-1} c_i x^i + cf(x)$ with $n+1$ roots (there are $n+1$ functions in our system). If the coefficient $c = 0$, then $\{1, x, x^1, \dots, x^{(n-1)}\}$ would not be a Haar system, consequently c is nonzero according to Remark 2.1.3. Then, using Rolle's theorem, we determine that $h'(x)$ has at least n roots according to the following argument. Let $x_1, \dots, x_{n+1}$ be the roots of $h(x)$, sorted in such a way that the condition $i < j$ implies $x_i < x_j$. Then Rolle's theorem's presumptions are fulfilled for each interval $[x_i, x_{i+1}]$, therefore there must exist a point $y_i \in [x_i, x_{i+1}]$ such that $h'(y_i) = 0$. Applying a similar argument on $h^{(2)}(x), h^{(3)}(x), \dots, h^{(n)}(x) = (\sum_{i=0}^{n-1} c_i x^i)^{(n)} + cf^{(n)}(x)$ gives that $h^{(n)}$ has at least one root on $[a, b]$. But $h^{(n)}(x) = (\sum_{i=0}^{n-1} c_i x^i)^{(n)} + cf^{(n)}(x) = cf^{(n)}(x) > 0$ for all $x \in [a, b]$, which leads to a contradiction.

□

From the previous examples it is also obvious that not every subset of a Haar system satisfies the Haar condition (consider the example of $\{1, x, x^2, \dots, x^{2n}\}$ and $\{1, x^2, x^4, \dots, x^{2n}\}$).

Lemma 2.1.7 (A transformation of a Haar system)**.** *Let H be a system of functions $\{f_1, f_2, \dots, f_n\}$, $f_i \in \mathbf{C}[a, b]$ which satisfies the Haar condition.*
1) If $A = \{a_{i,j}\}_{i,j=1}^n$ is a regular matrix $n \times n$ of real numbers, then the system of functions

$$\left\{ \sum_{i=1}^n a_{1,i}\, f_i, \sum_{i=1}^n a_{2,i}\, f_i, \dots, \sum_{i=1}^n a_{n,i}\, f_i \right\}$$

is a Haar system too.
2) If Φ is monotone and continuous on $[a, b]$ then the set of functions $\{f_1(\Phi^{-1}(x)), f_2(\Phi^{-1}(x)), \dots, f_n(\Phi^{-1}(x))\}$ satisfies the Haar condition on the interval $[\Phi(a), \Phi(b)]$.

Proof. Ad 1)
We claim that, for any points $x_1, x_2, \dots, x_n \in [a, b]$,

$$\det \left\{ \sum_{k=1}^n a_{i,k}\, f_k(x_j) \right\}_{i,j=0}^n \neq 0.$$

Then, using Lemma 2.1.2, we get that

$$\left\{ \sum_{i=1}^n a_{1,i}\, f_i, \sum_{i=1}^n a_{2,i}\, f_i, \dots, \sum_{i=1}^n a_{n,i}\, f_i \right\}$$

is a Haar system. Let $x_1, x_2, \ldots, x_n$ any points from $[a, b]$, arbitrary but fixed. We consider the matrix $n \times n$ $B = f_i(x_j)_{i,j=0}^n$. Then it is valid that

$$A \cdot B = \{a_{i,k}\}_{i,k=1}^n \cdot \{f_k(x_j)\}_{k,j=0}^n = \left\{ \sum_{k=0}^n a_{i,k}\, f_k(x_j) \right\}_{i,j=0}^n .$$

Because $\det(A \cdot B) = \det(A) \cdot \det(B)$, $\det(A) \neq 0$ and $\det(B) \neq 0$, we have $\det \{\sum_{k=0}^n a_{i,k}\, f_k(x_j)\}_{i,j=0}^n \neq 0$.

Ad 2)

Let us consider the contradictory proposition that there exists a nontrivial generalized polynomial $P(x) = \sum_{k=1}^n c_k f_k(\Phi^{-1}(x))$, where $c_i \in \mathbb{R}$ and $P(x)$ has n (or more) roots in $[\Phi(a), \Phi(b)]$. Then there exist $x_1, x_2, \ldots, x_n$; $x_i \in [\Phi(a), \Phi(b)]$; $x_i \neq x_j$ if $i \neq j$ and $y_1, y_2, \ldots, y_n$; $y_i \in [a, b]$ such that: $\Phi(y_i) = x_i$ and $P(x_i) = 0$ for $i = 1, 2, \ldots, n$. $\Phi(x)$ is monotone, hence $y_i \neq y_j$ if $i \neq j$, therefore the polynomial $\sum_{k=1}^n c_k f_k(y)$ has n roots, which is in a contradiction with the assumption that $\{f_1, f_2, \ldots, f_n | f_i \in \mathbf{C}[a, b]\}$ satisfies the Haar condition. □

Remark 2.1.8 ($\{1, e^x, e^{2x}, e^{3x}, \ldots e^{nx},\}$ is a Haar system on $[a, b]$ for all $n \in \mathbb{N}; a, b \in \mathbb{R}$)**.** *Pick any $n \in \mathbb{N}$ and $a, b \in \mathbb{R}$. According to Remark 2.1.3, $S = \{1, x, \ldots, x^n\}$ is a Haar system on any interval, especially on $\left[e^a, e^b\right]$. Our system can be written also in the following way:*

$$\left\{1, (e^x)^1, (e^x)^2, (e^x)^3, \ldots, (e^x)^n\right\}$$

This notation reminds one of the statement 2) of Lemma 2.1.7, which we are about to use with $\Phi \equiv \ln x$. The function $\ln x$ is monotone and continuous on $\left[e^a, e^b\right]$, therefore the assumptions are valid. Consequently, we get that $\{1, e^x, e^{2x}, e^{3x}, \ldots, e^{nx}\}$ is a Haar system on $[a, b]$ for all $n \in \mathbb{N}; a, b \in \mathbb{R}$.

Lemma 2.1.9 (Interpolation in Haar system)**.** *For each points $a_1, a_2, \ldots, a_n$ and values $y_1, y_2, \ldots, y_n$ there exists a generalized polynomial P over a given Haar system taking prescribed values $P(a_i) = y_i$*

Proof. Let $\{g_1, g_2, \ldots, g_n\}$ satisfy the Haar condition on $[a, b]$. We note that

$$\mathtt{D}[x_1, x_2, \ldots, x_n] \equiv \det \{g_i(x_j)\}_{i,j=1}^n .$$

Then the polynomial given by:

$$P(x) = \sum_{i=1}^n \frac{y_i}{\mathtt{D}[a_1, a_2, \ldots, a_n]} R_i(x) = \sum_{i=1}^n y_i \frac{\mathtt{D}[a_1, a_2, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n]}{\mathtt{D}[a_1, a_2, \ldots, a_n]}$$

is a generalized polynomial obeying all the conditions. In the previous definition, x is only a variable, and $a_1, a_2, \ldots, a_n$ are fixed constant points from $[a, b]$.
The following paragraph will prove that the above definition is correct. The denominator of the fraction in the definition is always nonzero because of the Haar condition. We should also clear up any possible doubts that a given function is a generalized polynomial, and that it is taking prescribed values. Now, $P(x)$ is a finite linear combination of an unusually defined functions $R_i(x)$. Each of these functions is a generalized polynomial, hence $P(x)$ as a linear combination of generalized polynomials is also a generalized polynomial. $R(x)$ is defined as a determinant of matrix, which has one column dependent on the variable x, and all elements in other columns are constant. The easiest way to show that R_i is a generalized polynomial is to expand the determinant $\mathtt{D}[a_1, a_2, \ldots, a_{i-1}, x, a_{i+1}, a_n]$ by the elements in i-th column.

$$R_i(x) = \mathtt{D}[a_1, a_2, \ldots, a_{i-1}, x, a_{i+1}, a_n] = \sum_{k=1}^{n} (-1)^{i+k} g_k(x) \det A_{i,k}.$$

Here $A_{i,k}$ is the adjoint matrix to the element in the i-th column and the k-th row. In conclusion, $P(x)$ is a generalized polynomial over the system $\{g_1, g_2, \ldots, g_n\}$. Now, it is simple to prove that $P(a_i) = y_i$. At the point a_j, all terms of the sum from the definition of $P(x)$ vanish except of the i-th one, because the determinants $\mathtt{D}[a_1, a_2, \ldots, a_{i-1}, a_j, a_{i+1}, a_n]$, where $j = 1, 2, \ldots, i-1, i+1, \ldots, n$, are all zero.

$$P(a_i) = y_i \frac{\mathtt{D}[a_1, a_2, \ldots, a_{i-1}, a_i, a_{i+1}, a_n]}{\mathtt{D}[a_1, a_2, \ldots, a_n]} + \sum_{j \neq i; j > 0; j \leq n} \frac{y_i}{\mathtt{D}[a_1, a_2, \ldots, a_n]} \cdot 0 = y_i$$

□

Lemma 2.1.10 (Polynomial with given roots)**.** *Let $a, b \in \mathbb{R}$ and assume that $c_1, c_2, \ldots, c_{n-1} \in [a, b]$. If $G \equiv \{g_1, g_2, \ldots, g_n | g_i\, [a, b]\}$ satisfies the Haar condition on $[a, b]$, then there exists a generalized polynomial P over the system G having zeroes at these and only these points. We can write $P(x) \equiv D[c_1, c_2, \ldots, c_{n-1}, x]$. Moreover, any other polynomial with the given property is a multiple of polynomial P.*

Proof. First we should verify that P is a generalized polynomial, but it is clear from the proof of the previous lemma where we were facing the same problem. The polynomial P meets our condition because for any point c_i

$P(c_i) = D[c_1, c_2, \dots, c_{n-1}, c_i] = 0$. (Determinant of a singular matrix.) Let R be a polynomial obeying the conditions of the lemma. If $R \neq P$, then there must exist a point $y \in [a, b]$ and $\lambda \in \mathbb{R}, \lambda \neq 1$ such that $R(y) = \lambda P(y)$. Now we consider polynomial $R - \lambda P$. This polynomial has roots at the points $c_1, c_2, \dots, c_{n-1}$ and y, therefore it must be a trivial polynomial. Consequently, $R(x) = \lambda P(x)$ for all $x \in [a, b]$. □

Definition 2.1.11. *An ordered system of functions* $\{f_1, f_2, \dots\}$; $f_i \in C[a, b]$ *is called a Markoff system if for each* $n \in \mathbb{N}$ *system* $\{f_1, f_2, \dots, f_n\}$ *satisfies the Haar condition on* $[a, b]$.

From some previous remarks it is clear that $\{1, x^1, x^2, \dots\}$ is a Markoff system on any $[a, b]$, $\{1, e^x, e^{2x}, \dots\}$ is also a Markoff system according to the previous remark on the arbitrariness of $[a, b]$.

2.2 Theorem of de La Vallée Poussin

After bringing in the concept of the Haar condition we are able to examine a standard version of the de La Vallée Poussin theorem.

Theorem 2.2.1 (Theorem of de La Vallée Poussin). *Let* $G \equiv \{g_1, g_2, \dots, g_n\}$ *be a system of continuous functions satisfying the Haar condition on some interval* $[a, b]$*, let* f *be continuous on* $[a, b]$. *If* P *is a generalized polynomial over the given system such that* $f - P$ *assumes alternatively positive and negative values at* $n + 1$ *consecutive points* x_i *of* $[a, b]$*, then*

$$\inf_R \|f - R\| \geq \min_i |f(x_i) - P(x_i)|,$$

where R *ranges over all generalized polynomials over system* G.

Proof. We claim that if the error function $r \equiv f - P$ assumes alternatively positive and negative values at $n + 1$ consecutive points x_i of $[a, b]$, then the origin of $\mathbb{R}^n$ lies in the convex hull of points $\{r(x_i) \cdot \hat{x_i}\}$ (using notation 1.2.4). If we prove this, then we can simply use the Generalized de La Vallée Poussin's theorem (Theorem 1.2.5). Achieving that, we find very helpful lemma on linear inequalities (Lemma 1.2.6).
Suppose for a contradiction that $\{g_1, g_2, \dots, g_n\}$ is a Haar system, the error function r assumes alternating values at $n + 1$ points and the origin of $\mathbb{R}^n$ is not in a convex hull of the corresponding points. According to lemma on

linear inequalities (Lemma 1.2.6), the last fact is equivalent to saying that there exists a vector $d \equiv [d_1, d_3, \ldots, d_n]$, solution of the following inequalities

$$\{\langle d, r(x_i) \cdot \hat{x}_i \rangle > 0\},$$

where $i \in \{1, 2, \ldots, n\}$. Equivalently we get for all i

$$r(x_i) \sum_{j=1}^{n} d_j g_j x_i > 0.$$

Hence $r(x_i)$ and $\sum_j d_j g_j(x_i)$ have the same signs for all i and the generalized polynomial $\sum_j d_j g_j(x)$ must alternate $(n+1)$-times sign on $[a, b]$ and consequently ought to have n roots on $[a, b]$. Which is in a contradiction with the Haar condition. (No nontrivial polynomial over G have n roots and d cannot be $O = [0, 0, \ldots, 0]$.)
We have proved that in our case the origin lies in the convex hull of the points $\{r(x_i) \cdot \hat{x}_i | i = 1, 2, \ldots, n\}$ and now using the Generalized theorem of de La Vallée Poussin (Theorem 1.2.5) we obtain the whole statement.

□

Problem 2.2.2. *Is there any approximation of the function* $\sin(x)$ *from the space of linear functions on* $\left[0, \frac{\pi}{2}\right]$ *such that* $|r(x)| < 0.05$*?*

Solution. The system $\{1, x\}$ satisfies the Haar condition. Consider the polynomial

$$P(x) = \frac{1}{10} + \frac{2x}{\pi}.$$

At points $0, \frac{\pi}{6}, \frac{\pi}{2}$, the error function assumes alternative values;

$$r(0) = 0 - \frac{1}{10} = -0.1\,,$$

$$r\left(\frac{\pi}{6}\right) = \frac{1}{2} - \frac{1}{10} - \frac{1}{3} = 0.5 - 0.43\bar{3} = 0.06\bar{6},$$

$$r\left(\frac{\pi}{2}\right) = 1 - \frac{1}{10} - 1 = -0.1\,.$$

According to theorem of de La Vallée Poussin, there does not exist any polynomial such that $\|r\| < 0.05$. □

2.3 The alternation theorem

We have already introduced the statement of the alternation theorem, but only for the Haar system $\{1, x, x^2, \ldots, x^n\}$. In this section we are about to prove the general version of the alternation theorem for any system obeying the Haar condition.

Theorem 2.3.1 (alternation theorem)**.** *Let $a, b \in \mathbb{R}$, let f be continuous on $[a, b]$ and let $G \equiv \{g_1, g_2, \ldots, g_n\}$ be a Haar system. In order that a polynomial P is the best approximation on $[a, b]$ to f from space of generalized polynomials over G, it is necessary and sufficient that there exist at least $n+1$ points $x_0, x_1, x_2, \ldots, x_n$ ($x_0 < x_1 < x_2 < \cdots < x_n$) such that for the error function $r = f - P$ we have $r(x_i) = -r(x_{i-1})$, $i = 1, \ldots, n$, and $|r(x_j)| = \|r\|$, $j = 0, \ldots, n$.*

In our proof of the alternation theorem we shall need one special property of Haar systems.

Lemma 2.3.2. *Let $G \equiv \{g_1, g_2, \ldots, g_n\}$ be a Haar system on $[a, b]$, let $x_1, x_2, \ldots, x_{n-1} \in [a, b]$ be points such that $a = x_0 \leq x_1 < x_2 < \cdots < x_{n-1} \leq x_n = b$. If P is a nontrivial generalized polynomial with roots at points $x_1, x_2, \ldots, x_{n-1}$, then for all $i \in \{1, 2, \ldots, n-1\}$ we have*

$$\operatorname{sign} P(z) \neq \operatorname{sign} P(y),$$

where $z \in (x_{i-1}, x_i)$ and $y \in (x_i, x_{i+1})$ are arbitrary points.

Proof of Lemma 2.3.2. We note that P cannot have any other roots because G is a Haar system. Let us assume the contradictory proposition that there exists $i \in \{1, 2, \ldots, n-1\}$ such that there exists $z \in (x_{i-1}, x_i)$ and $y \in (x_i, x_{i+1})$ with property

$$\operatorname{sign} P(z) = \operatorname{sign} P(y).$$

Using the lemma on interpolation in Haar system (Lemma 2.1.9), we get that there exists a generalized polynomial R such that $R(x_j) = 0$ for all $j \in \{1, 2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n-1}\}$, $R(z) = \operatorname{sign} P(z)$ and $R(y) = -\operatorname{sign} P(y)$. R is nontrivial and x_i is not a root of R, because, according to lemma on polynomial with given roots (Lemma 2.1.10), any polynomial with such roots must be a multiple of P. The polynomial R is continuous, consequently it has to have a root $m \in [z, y]$. Let us assume without any loss on generality that $m \in [x_i, y]$, the situation is illustrated on Figure 2.1. (If our presumption is not valid, we can instead use the polynomial $-R$.) Then there must exist points $h_1 \in [z, x_i]$ and $h_2 \in [x_i, m]$ such that $R(h_k) = P(h_k), k = 1, 2$. Consequently, the polynomial $P - R$ has roots at least at points $x_1, x_2, \ldots, x_{i-1}, h_1, h_2, x_{i+1}, \ldots, x_{n-1}$. The polynomial $P-R$ is nontrivial because at the point y we have $|P(y)-R(y)| = |P(y)-(-\operatorname{sign} P(y))| > 1 > 0$. This is in a contradiction with the Haar condition. $\square$

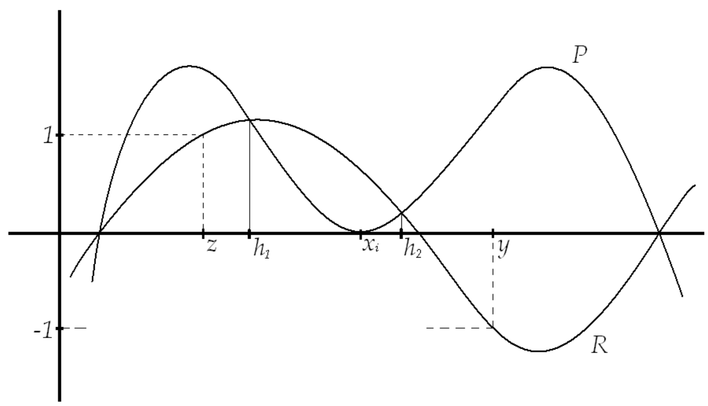

Figure 2.1: Situation from Lemma 2.3.2

Proof of the alternation theorem. The regular alternation theorem can be proved analogously to that of the version Theorem 1.1.1. Like there, we shall suppose for a contradiction that there exists a polynomial which doesn't meet our conditions and it is still the best approximation to f from the given space. Then we shall find an improved polynomial which gives a better approximation to f.

Let P be the best approximation and let $r \equiv f - P$ be the error function. Let $x_1, x_2, \ldots, x_k \in [a, b]$ be a maximal system of points, such that x_1 is the nearest point to a obeying

$$|r(x_1)| = \|r\|$$

and $x_{i+1} \in [x_i, b]$ is always the nearest point to x_i with property

$$r(x_i) = -r(x_{i+1}),$$

for all $i \in \{1, 2, \ldots, k-1\}$. Assume that $n+1 > k = n$. We denote points z_j for $j \in \{1, 2, \ldots, k-1\}$ as points nearest to x_{j+1} from $[x_j, x_{i+j}]$ which meet the condition $r(z_j) = 0$. According to lemma on polynomial with given roots (Lemma 2.3.1), there exists a polynomial $S(x) = \sum_{j=1}^{n} s_j g_j(x)$ such that

$$S(z_i) = 0.$$

We can also manage to fulfill the condition that $\operatorname{sign} S(x_i) = \operatorname{sign} r(x_i)$ for all $i \in \{1, 2, \ldots, n-1\}$ on applying Lemma 2.3.2 and taking S or $-S$.

Now we define *rude* interval analogously as in the proof of the weaker theorem. The interval $[v, w]\,, v, w \in [a, b]$ is said to be rude if it is fulfilling: $S(v) = S(w) = 0$, if $x \in [v, w]$ is any arbitrary point, then $r(x)S(x) \leq 0$ (they have different signs) and there exists $\epsilon > 0$ such that for any point y outside the rude interval with property $|y - v| < \epsilon$ or $|y - w| < \epsilon$ previous condition is not valid.

Let R be a set of all rude intervals. It is clear that any two different intervals from R are disjoint. Because intervals are subsets of a compact interval $[a, b]$, R is finite.

We denote

$$M = \sup_{I \in R} \|r\|_{C[I]},$$

$$N = \sup_{x \in [a,b]} |S(x)| = \|S\|\,,$$

and

$$C = \|r\| - M.$$

It is important that $C > 0$, because inside a rude interval $[v, w]$ there does not exist any point q such that $|r(q)| = \|r\|$. For sure we should comment on this fact. Let q be such point. Then there must exist $i \in \{1, 2, \ldots, n\}$ with property $q \in [x_i, x_{i+1}]$. If $r(q) = r(x_i)$, then q cannot lie inside a rude interval, because $S(q)r(q) > 0$. If $r(q) = -r(x_i)$ then $q = x_{i+1}$ and again, q is not inside any rude interval.

Consequently there must exist $\lambda > 0$ such that

$$\lambda N < \frac{1}{2} C.$$

Now we consider the polynomial $T(x) = P(x) + \lambda S(x)$, T is a generalized polynomial over G. Our goal is to prove that T is a better approximation to f than P. We assign

$$r_T(x) \equiv f(x) - T(x) = f(x) - P(x) - \lambda S(x) = r(x) - \lambda S(x)$$

First let $z \in [a,b]$ not from any rude interval, then we get

$$|r_T(z)| = |r(z) - \lambda S(z)| < |r(z)| < \|r\|,$$

because z is not in rude interval and therefore S(z) and r(z) have same sign and $S(z) \neq 0$. If $z \in [v,w]$ where $[v,w]$ is a rude interval, we obtain

$$|r_T(z)| = |r(z) - \lambda S(z)| < M + \frac{1}{2}C < \|r\|.$$

T is a better approximation than P, but P was assumed to be the best approximation.

In the previous argument we have assumed that $k = n$. If $k < n$, then we want to find points $z_k, \ldots, z_{n-1}$ to have $n-1$ roots of our improving polynomial in total. We will consider three cases.

If $r(a) \neq \|r\|$ or $r(b) \neq \|r\|$, let τ be an interval $[a,q]$ or $[q,b]$ obeying that $|r|$ does not take the value $\|r\|$ on it. Then we can simply pick any points $z_k, \ldots, z_{n-1} \in \tau$ and use the polynomial with roots at the points $z_1, z_2, \ldots, z_{n-1}$. The argument will be the same like as before.

If $r(a) = \|r\|$ and $r(b) = \|r\|$ but $n-k$ is even, then we can find our missing roots in any interval where $|r|$ does not attain it's maximum (such an interval exists), and our polynomial will have the required properties.

The last case where $r(a) = \|r\|$, $r(b) = \|r\|$ and $n-k$ is odd is more difficult. For the last root we cannot pick any point from (a,b), because the last (odd) root always change the sign of the polynomial to the opposite one and therefore the point b would be in a rude interval. (Consequently, we would have $C = 0$.) We add an even number of roots like in the previous case and let the last root be b. Then the constructed polynomial has to alternate its signs on intervals $[z_i, z_{i+1}]$ for each i according to Lemma 2.3.2. We can find λ similar to proof of the case with n points ignoring the last point. But in a contrary to that case, the polynomial $Q \equiv P - \lambda S$ is not a better approximation than P because at the point b there is $f(b) - P(b) + \lambda 0 = \|r\|$. However, Q is a better approximation everywhere except at b. There exists a polynomial O such that the linear combination of Q and O is a better approximation to f than P. The exact polynomial shall be found in the following.

Let $r_Q = f - Q$ be the corresponding error function, let $z \in [a,b)$ be the nearest point to b with property $r_Q(z) = 0$, let $\iota = \operatorname{sign} r_Q(b)$. It is clear that $\|r_Q\| = \|r\|$ and the only point where r_Q attain it's maximum is the point b.

Let $c_1, c_2, \ldots, c_{n-2} \in [a, z]$ be any points. According to Lemma 2.1.9, there exists a polynomial O with property $O(c_i) = 0$, $O(z) = 0$, $O(b) = \iota$.

We denote

$$\nu = \|O\|,$$

$$\epsilon = \inf_{x\in[a,z]} \|r_Q\| - |r_Q(x)|.$$

We note that $\epsilon > 0$. Let $\kappa \in \mathbb{R}$ be such number that

$$0 < \kappa < \frac{\epsilon}{2\nu}.$$

We claim that the polynomial $Q + \kappa O$ is a better approximation than Q and consequently better than P.

Let x be any point from $[a, z]$, then

$$|r_{Q+\kappa O}(x)| = |f(x) - Q(x) - \kappa O(x)| \leq \|r_Q\| + |\kappa O(x)| < \|r_Q\| - \epsilon + \kappa\nu.$$

Consequently

$$\|r_Q\| - \epsilon + \kappa\nu = \|r_Q\| - \epsilon + \frac{\epsilon\nu}{2\nu} = \|r_Q\| - \frac{\epsilon}{2} < \|r_Q\| = \|r\|.$$

Let x be any point from $(z, b]$, then

$$|r_{Q+\kappa O}(x)| = |f(x) - Q(x) - \kappa O(x)| < |r_Q(x) - \kappa O(x)|.$$

Because both $r_Q(x)$ and $\kappa O(x)$ have the same sign and because $r_Q(x)$ is greater in absolute value, we get

$$|r_Q(x) - \kappa O(x)| \leq |r_Q(x)| - |\kappa O(x)| \leq \|r_Q\| - |\kappa O(x)| < \|r_Q\| = \|r\|,$$

the last strict inequality is valid because $O(x) \neq 0$ on $(z, b]$. Consequently, P is not the best approximation.

Proof of the sufficiency part is build upon the theorem of de La Vallée Poussin (Theorem 2.2.1), however the proof from the special version 1.1.1 mentioned before would work too.

Let P be a polynomial fulfilling the conditions of the theorem, then P also obey presumptions of de La Valleée Poussin's theorem ($x_1, \ldots, x_n$ are points at which function $f - P$ assumes alternative values). Consequently, we get

$$\max_{x\in[a,b]} |r(x)| = \|r\| = \min_i |f(x_i) - P(x_i)| \leq \inf_R \|R - f\|.$$

In conclusion, the minimal norm of the error function of any polynomial is greater than or equal to the norm of the error function of the polynomial P, therefore P is the best approximation. □

In the following problem, which is a modification of Problem 1.1.2, we shall see applications of the previous theorem.

Problem 2.3.3 (modification of Problem 1.1.2)**.** *Find the best approximation of the function* x *by a linear combination of functions* 1 *and* e^x *on* $[0,1]$.

Solution. We first observe that $\{1, e^x\}$ satisfies the Haar condition on $[0,1]$ and consequently the presumptions of the theorem are fulfilled. Let P be our desired best approximation and let r be the error function

$$r \equiv x - P = x - a - be^x.$$

We need to find points $x_1, x_2, x_3 \in [0,1]$ with property

$$r(x_1) = -r(x_2) = r(x_3);\ |r(x_1)| = \|r\|,$$

previous equalities also determine coefficients a and b.

Like in Problem 1.1.2 we find out that $x_1 = 0$ and $x_3 = 1$ because there is just one point of extreme inside $[0,1]$. This knowledge comes from examining of the derivative of r:

$$r^{(1)}(x_2) = 0 \Rightarrow 1 - 0a - be^{x_2} = 0 \Rightarrow b = e^{-x_2};\ r^{(2)}(x_2) = -be^{x_2} < 0.$$

Using equality $r(0) = r(1)$, we get the exact value of b:

$$0 - a - b = 1 - a - be \Rightarrow b = \frac{1}{e-1}.$$

Now using equalities which we obtained from the derivative of r, we can determine the point x_2:

$$be^{x_2} = 1 \Rightarrow x_2 = \ln(e-1).$$

Finally we can get a:

$$0 - a - b = -x_2 + a + be^{x_2} \Rightarrow a = \frac{1}{2}\left(\ln(e-1) - \frac{e}{e-1}\right).$$

Consequently, the best approximation of x from space of linear combinations of 1 and e^x is

$$P(x) \equiv \frac{1}{2}\left(\ln(e-1) - \frac{e}{e-1}\right) + \frac{e^x}{e-1}.$$

□

Conclusion

Our interest was in approximation by generalized polynomials over Haar systems. The properties of Haar systems and polynomials build over them are examined in examples and some conclusions are generalized into the form of a lemma. The properties established are then used in proofs of following theorems. In this way we were able to prove difficult theorems using just elementary knowledge. We have re-proved the alternation theorem and the theorem of de La Vallée Poussin. All the proofs in this thesis are due to the author. The theorems are then aplicated to some problems of the approximation theory.

Bibliography

[1] Cheney E. W.: *Introduction to APPROXIMATION THEORY*, American Mathematical Society, United States of America, 2000.

[2] Wikipedia: Descartes rule of signs
http://en.wikipedia.org/wiki/Descartes%27_rule_of_signs