Effective bounds on strong unicity in L_1-approximation

Ulrich Kohlenbach Paulo Oliva

May, 2001
BRICS*
Department of Computer Science
University of Aarhus
Ny Munkegade
DK-8000 Aarhus C, Denmark

Abstract

In this paper we present another case study in the general project of Proof Mining which means the logical analysis of prima facie non-effective proofs with the aim of extracting new computationally relevant data. We use techniques based on monotone functional interpretation (developed in [17]) to analyze Cheney's simplification [6] of Jackson's original proof [9] from 1921 of the uniqueness of the best L_1-approximation of continuous functions $f \in C[0,1]$ by polynomials $p \in P_n$ of degree $\leq n$. Cheney's proof is non-effective in the sense that it is based on classical logic and on the non-computational principle WKL (binary König Lemma). The result of our analysis provides the first effective (in all parameters f, n and ε) uniform modulus of uniqueness (a concept which generalizes 'strong uniqueness' studied extensively in approximation theory). Moreover, the extracted modulus has the optimal ε-dependency as follows from Kroo [20]. The paper also describes how the uniform modulus of uniqueness can be used to compute the best L_1-approximations of a fixed $f \in C[0,1]$ with arbitrary precision, and includes some remarks on the case of best Chebycheff approximation.

1 Introduction

This paper is another case study in the general project of proof mining which means the logical analysis of prima facie non-effective proofs with the aim of extracting new computationally relevant data.[1] At the same time we obtain new results in approximation theory. More

*Basic Research in Computer Science, funded by the Danish National Research Foundation.
[1]See [15], [16], [19] and [12] for other case studies as well as more information on Proof Mining in general.

specifically, we analyze a non-effective proof of the uniqueness of best approximations of continuous functions $f \in C[0,1]$ by polynomials $p \in P_n$ of degree $\leq n$ with respect to the L_1-norm[2]

$$\|f\|_1 := \int_0^1 |f(x)|dx.$$

In [15], the first author showed how a quite general class of (non-effective) proofs of uniqueness theorems in analysis can be analyzed such that an effective so-called modulus of uniqueness can be extracted which generalises the concept of 'strong unicity'.[3] In [15] and [16] this technique has been applied to the case of best Chebycheff approximation yielding new uniform bounds on constants of strong unicity and a new quantitative version of the alternation theorem. In this paper we apply this logical approach to investigate the quantitative rate of strong unicity for the quite different case of best L_1-approximation. Like Chebycheff approximation, L_1-approximation, also called 'approximation in the mean', is a classical topic in numerical mathematics and was considered already by Chebycheff in 1859 and has been investigated ever since (see [24] for a comprehensive survey). The uniqueness of the best L_1-approximation of $f \in C[0,1]$ by polynomials of degree $\leq n$ was first proved in [9]. This proof uses measure theoretic arguments. A new uniqueness proof which avoids this and only uses the Riemann integral instead was given in 1965 by Cheney (see [6],[7]). Because of this feature, Cheney called his proof 'elementary'. From a logical point of view, however, it is highly non-constructive relying both on classical logic and non-computational analytical principles which correspond – in logical terminology – to the so-called binary ('weak') König's lemma, a principle which has received considerable attention in various parts of logic in recent years (see [25]). In this paper we carry out a complete logical analysis of Cheney's proof and show how the explicit modulus mentioned above can be extracted from this seemingly hopelessly non-constructive proof. Consequently, our result, like Cheney's proof, does not require any measure theory.

The main result of the present paper is the following effective strong uniqueness theorem:

Main result (Theorem 4.3) *Let*

$$\begin{aligned}
&\Phi(\omega, n, \varepsilon) := \min\{\tfrac{c_n \varepsilon}{3^{n+2}(n+1)^{n+1}}, \tfrac{c_n \varepsilon}{2}\omega_n(\tfrac{c_n \varepsilon}{2})\}, \textit{ where}\\
&c_n := \tfrac{\lfloor n/2 \rfloor ! \lceil n/2 \rceil !}{2^{n+3} 3^{n^2+2n}(n+1)^{n^2+2n+1}} \textit{ and}\\
&\omega_n(\varepsilon) := \min\{\omega(\tfrac{\varepsilon}{4}), \tfrac{\varepsilon}{40(n+1)^4 \lceil \frac{1}{\omega(1)} \rceil}\}.
\end{aligned}$$

[2]For $f \in L_1$ uniqueness in general fails.

[3]The term 'strong unicity' was introduced by Newman and Shapiro [23] in 1963 and has been studied extensively in approximation theory. See e.g. the introduction in [2] and the references given there for a discussion of the crucial importance of estimates of strong unicity for the convergence analysis of iterative algorithms and for stability analysis.

The functional Φ is a uniform modulus of uniqueness for the best L_1-approximation of any function f in $C[0,1]$ having modulus of uniform continuity ω from P_n, i.e.

$$\forall n \in \mathbb{N}; p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}_+^* \Big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) < \Phi(\omega, n, \varepsilon)) \to \|p_1 - p_2\|_1 \le \varepsilon \Big),$$

where $dist_1(f, P_n) := \inf_{p \in P_n} \|f - p\|_1$ and $\omega : \mathbb{Q}_+^* \to \mathbb{Q}_+^*$ is a modulus of uniform continuity for $f \in C[0,1]$ if [4]

$$\forall x, y \in [0,1]; \varepsilon \in \mathbb{Q}_+^* (|x - y| < \omega(\varepsilon) \to |f(x) - f(y)| < \varepsilon).$$

Moreover, this theorem can be proved in Heyting Arithmetic HA^ω in all finite types (and consequently holds in constructive mathematics in the sense of Bishop).

The technical details of this analysis are mainly due to the second author who is using the results in a subsequent paper to determine a complexity upper bound for the sequence $(p_{b,n})_{n \in \mathbb{N}}$ of best approximating polynomials for poly-time computable functions $f \in C[0,1]$ (in the sense of [10],[11]).
Before going into the details of the analysis we need to recall some general logical background from [15].[5] First we introduce a little amount of logical terminology:
Let $\mathcal{A}^\omega$ be a (sub-)system of arithmetic in all finite types (like E-PA^ω from [26] or Feferman's fragment E-PRA^ω with quantifier-free induction and primitive recursion on the type 0 only [8]). Let $\mathcal{A}^\omega_*$ denote the extension of $\mathcal{A}^\omega$ by the schema

$$\text{QF-AC}: \ \forall f^1 \exists x^0 A_{qf}(f, x) \to \exists F^2 \forall f^1 A_{qf}(f, F(f))$$

of quantifier-free choice from functions to numbers (where A_{qf} is quantifier-free) plus certain analytical principles Γ which – described in analytical terms – correspond to applications of Heine-Borel compactness of e.g. $[0,1]^d$. In logical terms, these principles correspond to the so-called binary ('weak') König's lemma WKL which suffices to derive a substantial amount of mathematics relative to weak fragments of arithmetic (see [25]).[6] In this paper the only genuine analytical tool Γ (which goes beyond E-PA^ω+QF-AC) is the attainment of the minimum of $f \in C[0,1]$

$$(*) \ \forall f \in C[0,1] \exists x \in [0,1] \big(f(x) = \inf_{y \in [0,1]} f(y) \big).$$

[4] Note that this notion – used also in constructive mathematics and computable and feasible analysis – differs from the concept of modulus of continuity used in numerical analysis which we will discuss further below.

[5] Readers only interested in the numerical results but not in the general process of proof mining might skip this passage.

[6] E-PRA^ω+QF-AC+WKL is a finite type extension of the system WKL_0 used in reverse mathematics and is (like the latter) Π^0_2-conservative over primitive recursive arithmetic PRA (see [14],[1]).

$(*)$ is known to fail in computable analysis and even for poly-time computable f there will be in general no computable $x \in [0,1]$ satisfying $(*)$. [7]

Now, let X be a Polish space, K a compact Polish space and $F : X \times K \to \mathbb{R}$ a continuous function (moreover all these objects have to be explicitly representable in $\mathcal{A}^\omega$) and assume that we can prove in $\mathcal{A}^\omega_*$ that for every $f \in X$, $F(f,\cdot)$ has at most one root in K, i.e.[8]

$$(1)\ \forall f \in X \forall x_1, x_2 \in K \big(\bigwedge_{i=1}^{2} F(f,x_i) = 0 \to x_1 = x_2 \big).$$

Then by a general logical meta-theorem proved in [15] (theorem 4.3) one can extract from such a proof an explicit bound $\Phi(f,k)$ (given by a closed term of the underlying arithmetical system $\mathcal{A}^\omega$) such that

$$(2)\ \forall f \in X \forall k \in \mathbb{N} \forall x_1, x_2 \in K \big(\bigwedge_{i=1}^{2} (|F(f,x_i)| < 2^{-\Phi(f,k)}) \to d_K(x_1,x_2) < 2^{-k} \big),$$

where d_K denotes the metric on K. Moreover, (2) can be proved without using WKL and even in the intuitionistic variant $\mathcal{A}^\omega_i$ of $\mathcal{A}^\omega$ (and hence in constructive analysis in the sense of Bishop).

The proof of this meta-theorem provides an algorithm for actually extracting Φ. This algorithm is based on the proof-theoretic technique of monotone functional interpretation [17]. It is important to note that $\Phi(f,k)$ does **not** depend on $x_1, x_2 \in K$. Because of this fact, $\Phi(f,k)$ – which we call a **modulus of uniqueness** – can be used to compute the unique root (if existent) from any algorithm $\Psi(f,k)$ computing approximate so-called $\varepsilon(= 2^{-k})$-roots of $F(f,\cdot)$:

$$(3)\ \forall f \in X \forall k \in \mathbb{N} \big(\Psi(f,n) \in K \wedge |F(f,\Psi(f,k))| < 2^{-k} \big).$$

One easily verifies that (2) and (3) imply that $\Psi(f,\Phi(f,k))$ is a Cauchy sequence in K which converges with rate of convergence 2^{-k} to the unique root $x \in K$ of $F(f,\cdot)$. So $x = \lim_{k\to\infty} \Psi(f,\Phi(f,k))$ can be computed with arbitrarily prescribed precision (which can also be proved in $\mathcal{A}^\omega_i$, see [15], theorem 4.4) and the computational complexity of x can be estimated in terms of the complexities of Φ and Ψ.

Remark 1.1 (Important!) *As usual in computable analysis (see [27]), $\Phi(f,k)$ and $\Psi(f,k)$ will depend not only on $f \in X$ in the set theoretic sense but on a (computationally meaningful)*

[7] $(*)$ is known to be equivalent to WKL over systems like E-PRA$^\omega$ even when f is given together with a modulus of uniform continuity, see [25].

[8] We may even have functions $F : X \times Y \to \mathbb{R}$, where X, Y are general Polish spaces and can allow constructively definable families $(K_f)_{f \in X}$ of compact subspaces of Y which are parametrised by $f \in X$ instead of a fixed K. See [15] for details.

representation of f. In the case of $f \in C[0,1]$, the representation of $C[0,1]$ as a Polish space $(C[0,1], \|\cdot\|_\infty)$ in $\mathcal{A}^\omega$ requires that f is endowed with a modulus of uniform continuity ω_f. So when we write $\Phi(f,k)$ we tacitly understand that f is given as a pair (f, ω_f). Actually, it now suffices to use the restriction f_r of f to the rational numbers in $[0,1]$ (which can be enumerated so that f_r can be represented as a number theoretic function), since f can be reconstructed from f_r with the help of ω_f. In this way, the representation (f_r, ω_f) of f can be viewed as an object of type 1 so that computability on f reduces to the well-known type-2 notion of computability (see again [27] for more information on this).

Let us now move to the case of best L_1-approximation treated in the present paper. The uniqueness of the best approximation can be written as follows

$$(4)\ \forall n \in \mathbb{N} \forall f \in C[0,1] \forall p_1, p_2 \in P_n \big(\bigwedge_{i=1}^{2} (\|f - p_i\|_1 = dist_1(f, P_n)) \to p_1 = p_2 \big).$$

Note that in (4) we can without loss of generality replace the non-compact subspace P_n of $C[0,1]$ with the compact one $\tilde{K}_{f,n} := \{p \in P_n : \|p\|_1 \le 2\|f\|_1\}$ since any best approximation p_b has to satisfy $\|f - p_b\|_1 \le \|f\|_1$ because otherwise the zero polynomial would be a better approximation. As a consequence of this, $dist_1(f, P_n) = dist_1(f, \tilde{K}_{f,n})$ can easily be seen to be computable (uniformly in f as represented above and n). We use the slightly larger space $K_{f,n} := \big\{p \in P_n : \|p\|_1 \le \frac{5}{2}\|f\|_1\big\}$ in (4) for technical reasons.

In this paper we analyze the above mentioned proof of Cheney for (4) as given in [6],[7][9] which uses the non-computational principle $(*)$ (together with classical logic) but which can be formalized in $\mathcal{A}^\omega_*$ (as was shown in [13]). So the above mentioned result on the extractability of a modulus of uniqueness is applicable, i.e. the extractability of a (primitive recursive in the sense of Gödel's T) functional Φ satisfying

$$(5) \begin{cases} \forall n, k \in \mathbb{N} \forall f \in C[0,1] \forall p_1, p_2 \in K_{f,n} \\ \quad \big(\bigwedge\limits_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) < 2^{-\Phi(f,n,k)}) \to \|p_1 - p_2\|_1 < 2^{-k} \big) \end{cases}$$

is guaranteed. Moreover, a simple trick (used also in [15] in the Chebycheff case) allows to replace $K_{f,n}$ with P_n so that

$$(6) \begin{cases} \forall n, k \in \mathbb{N} \forall f \in C[0,1] \forall p_1, p_2 \in P_n \\ \quad \big(\bigwedge\limits_{i=1}^{2} (\|f - p_i\|_1 - dist_1(f, P_n) < 2^{-\Phi(f,n,k)}) \to \|p_1 - p_2\|_1 < 2^{-k} \big). \end{cases}$$

[9] This result was first proved in [9] and is also called Jackson's theorem. Cheney's proof (which applies to arbitrary Chebycheff systems) is a simplification of Jackson's proof.

Remark 1.2 *Since* $\|p_1 - p_2\|_\infty \leq 2(n+1)^2\|p_1 - p_2\|_1$ *any upper bound on* $\|p_1 - p_2\|_1$ *gives an bound on* $\|p_1 - p_2\|_\infty$ *and we can use this to get a bound on the coefficients of* $p_1 - p_2$*. Namely, if* $p_1(x) - p_2(x) := a_n x^n + \ldots + a_1 x + a_0$ *and* $\|p_1 - p_2\|_1 < M$ *then* $|a_i| \leq \frac{(2(n+1)^2)^{i+1}}{i!} M$*. The proof of this fact is given in section 3.5.*

The importance of the modulus of uniqueness $\Phi(f, k)$ can also be illustrated by the fact that $\Phi + 1$ is automatically a modulus of pointwise continuity for the operator which maps $f \in X$ to its unique best approximation $f_b \in E \subset X$ (see [15]). For the special cases of Chebycheff resp. L_1-approximation this was shown first in [7] resp. [3]. Therefore,

$$(7) \forall n, k \in \mathbb{N} \forall f, \tilde{f} \in C[0,1] \big(\|f - \tilde{f}\|_1 < 2^{-\Phi(f,n,k)-1} \to \|\mathcal{P}(f,n) - \mathcal{P}(\tilde{f}, n)\|_1 < 2^{-k} \big),$$

where $\mathcal{P}(f, n)$ is the unique best L_1-approximation of $f \in C[0,1]$ from P_n.

Since $(C[0,1], \|\cdot\|_1)$ is not a Polish space we have to represent $C[0,1]$ as the space $(C[0,1], \|\cdot\|_\infty)$ to apply the logical meta-theorem mentioned above. As we discussed already, this amounts to enriching the input f by a modulus of uniform continuity ω_f so that Φ will also depend on ω_f.

Note that if $C[0,1]$ is replaced by the (pre-)compact (w.r.t. $\|\cdot\|_\infty$) set $\mathcal{K}_{\omega,M}$ of all functions $f \in C[0,1]$ which have the common modulus of uniform continuity ω and the common bound $\|f\|_\infty \leq M$, then the same logical meta-theorem guarantees the extractability of a modulus of uniqueness Φ which only depends on $\mathcal{K}_{\omega,M}$ i.e. on ω, M (in addition to n, k). Moreover, even the M-dependency can be eliminated as the approximation problem for f can be reduced to that for $\tilde{f}(x) := f(x) - f(0)$ so that only a bound $N \geq \sup_{x \in [0,1]} |f(x) - f(0)|$ is required, which can easily be computed from ω (e.g take $N := \lceil \frac{1}{\omega(1)} \rceil$). Therefore, from the logical meta-theorem and the fact that Cheney's proof can be formalized in E-PA$^\omega$+WKL we obtain already the extractability of a primitive recursive (in the sense of Gödel's T) modulus of uniqueness Φ which only depends on ω_f, n and k: **a-priori information.** Of course, only the actual extraction of Φ by applying the algorithm provided by the logical meta-theorem gives the detailed mathematical form of Φ as presented above: **a-posteriori information.** It is interesting to note that although the proof we analyze here was published in 1965 (by Cheney) only in 1975 Björnestal proved the *existence* of a modulus of uniform uniqueness for the best L_1-approximation having the form $c_{f,n}\varepsilon\, \omega_f(c_{f,n}\varepsilon)$ where the constant $c_{f,n}$ depends only on f (and its modulus of continuity) and n, but no explicit constant was presented. In 1978 Kroo improved such a result (using some amount of measure theory) proving the existence of a constant $C_{\omega_f,n}$ which was independent of any particular value of f (i.e. the modulus of uniqueness would depend on f only through its modulus of continuity) doing the same job but as Björnestal he did not present any constant. In the same paper Kroo (using the method of Björnestal) also proved that the dependency on ε, i.e. $\varepsilon\ \omega_f(\varepsilon)$, is optimal (even for the modulus of pointwise continuity for the projection, see theorem 4.5). Therefore it is quite amazing that the **a-priori information** - the dependencies of the modulus of uniqueness - we obtain immediately by showing that Cheney's proof can be formalized in the system $\mathcal{A}^\omega_*$ (which in some sense means that this information was already implicit in Cheney's proof) was obtained without the use of logic only long after the proof was given. Moreover, the **a-posteriori information** – the actual modulus of uniform uniqueness – presents explicitly the dependencies on ω, n and ε, and the dependency on ε is optimal (as shown by Kroo).

2 Analysing proofs in analysis

The algorithm to be used for proof mining applied in cases like Cheney's proof of Jackson's theorem (as treated in this paper) is based on the proof theoretic technique of monotone functional interpretation combined with negative translation as developed in [17]. Whereas the technical details of this process are of importance to establish general meta-theorems on proof mining, this is not necessary for applications to specific proofs since here all numerical data will explicitly be exhibited and verified. This is because monotone functional interpretation explicitly transforms a given proof into another numerically enriched proof (in the normal mathematical sense). It is the strategy to find that proof (and to guarantee its existence) which is provided by the logical technique.

To approach the problem of proof mining applied to a logically involved proof as Cheney's, one starts off by splitting the proof into small pieces which are analyzed separately. As a consequence of the modularity of monotone functional interpretation one can easily combine the results obtained from the analysis of the pieces into a global result (this only requires functional application and λ-abstraction). Applications of monotone functional interpretation to the lemmas in the given proof at hand consist mostly of two steps,

1) transforming a given lemma L into a variant L^* which has the form

$$(**)\ \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k A_1(n, x, y, k),$$

where X is a Polish space, K a compact Polish space and $A_1 \in \Sigma^0_1$, and

2) extracting a bound $\Phi(n, x)$ for k which is independent of y.

Because of this it is worthwhile to formulate the case of lemmas having the form $(**)$ as a special meta-theorem (2.1 below) which allows us to avoid having to go into the details of the underlying mechanism of functional interpretation each time. Although in the following we perform the transformation $L \mapsto L^*$ "by hand" one should note that this transformation is also usually automatically provided by functional interpretation. Only in the case of "Lemma 1" below, we first simplify the lemma to achieve this.

Theorem 2.1 ([15], theorem 4.1) *Let X, K be $\mathcal{A}^\omega$-definable Polish spaces, K compact and consider a sentence which can be written (when formalized in the language of $\mathcal{A}^\omega$) in the form*

$$A := \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \in \mathbb{N}\, A_1(n, x, y, k),$$

where A_1 is a purely existential. Then the following rule holds:[10]

$$\begin{cases} \mathcal{A}^\omega_* \vdash \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \in \mathbb{N}\, A_1(n, x, y, k) \\ \text{then one can extract an } \mathcal{A}^\omega\text{-definable functional } \Phi \text{ s.t.} \\ \mathcal{A}^\omega_i \vdash \forall n \in \mathbb{N} \forall x \in X \forall y \in K \exists k \leq \Phi(n, x)\, A_1(n, x, y, k). \end{cases}$$

In particular, if

$$\mathcal{A}^\omega_i \vdash (k \leq \tilde{k} \wedge A_1(n, x, y, k)) \rightarrow A_1(n, x, y, \tilde{k})$$

then

$$\mathcal{A}^\omega_i \vdash \forall n \in \mathbb{N} \forall x \in X \forall y \in K\, A_1(n, x, y, \Phi(n, x)).$$

Again it is important to note that Φ does not depend on $y \in K$.[11]

In the following we try to avoid too much reference to logic in the main text and only insert various 'logical remarks' to explain to those readers interested in the process of proof mining in general how the various steps in our concrete 'mining' correspond to steps in the monotone functional interpretation (as used in the general meta-theorems). Readers only interested in the numerical results can skip these remarks.

3 Analysis of Cheney's proof of Jackson's Theorem

3.1 Logical preliminaries on Cheney's proof

In this section we sketch how a slight modification of Cheney's proof can be seen to be formalizable in basic arithmetic like $\mathcal{A}^\omega$:=E-PA$^\omega$ plus the already mentioned analytical principle $(*)$, i.e. WKL. The only part of the proof which cannot be directly formalized in $\mathcal{A}^\omega$ is the so-called 'lemma 1' (see [7], p. 219) which reads as follows

Lemma 3.1 (Lemma 1) *Let $f, h \in C[0, 1]$. If f has at most finitely many roots and if*

[10] As the theorem shows the conclusion can be proved already in $\mathcal{A}^\omega_i$ instead of $\mathcal{A}^\omega_*$. This, however, is not important for the applied aspect of the present paper where only the construction of Φ matters.

[11] As discussed in the previous section, $\Phi(n, x)$ will depend on the representation of $x \in X$.

$\int_0^1 h\, sgn(f) \neq 0$, *then for some* $\lambda \in \mathbb{R}$, $\int_0^1 |f - \lambda h| < \int_0^1 |f|$, *where*

$$sgn(f)(x) =_0 \begin{cases} 1, \text{ if } f(x) >_{\mathbb{R}} 0 \\ 0, \text{ if } f(x) =_{\mathbb{R}} 0 \\ -1, \text{ if } f(x) <_{\mathbb{R}} 0. \end{cases}$$

In the context of the Cheney's proof of Jackson's theorem, h will be a polynomial in P_n. Moreover, it will be shown that if f (for the particular f at hand) has only less than $n+1$ roots one can construct an h such that $\int_0^1 h\, sgn(f) \neq 0$. So we only need the lemma with the stronger assumption that f has fewer than $n+1$ roots. The existence of $sgn(f)$ relies on the existence of the characteristic function $\chi_{=_{\mathbb{R}}}$ for equality between reals which in turn is equivalent to the existence of Feferman's ([8]) non-constructive μ-operator (see [18]) and hence to a strong form of arithmetical comprehension which is not available in $\mathcal{A}^\omega_* := \mathcal{A}^\omega$+WKL. However, the use of sgn can be eliminated as follows: if f has less than $n+1$ roots then there exist points $x_0 < \ldots < x_{n+1}$ in $[0,1]$ (where $x_0 = 0$ and $x_{n+1} = 1$) which contain all the roots of f. By classical logic and induction one shows in $\mathcal{A}^\omega$ the existence of a vector $(\sigma_1, \ldots, \sigma_{n+1}) \in \{-1, 1\}^{n+1}$ such that

$$\sigma_i =_0 \begin{cases} 1, \text{ if } f \text{ is positive on } (x_{i-1}, x_i), \\ -1, \text{ if } f \text{ is negative on } (x_{i-1}, x_i) \end{cases}$$

for $i = 1, \ldots, n+1$. Using this vector, $\int_0^1 h\, sgn(f)$ can be written as $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h$. It will turn out below that it is the precise logical form of this reformulation of lemma 1 which will play a crucial role in the analysis of Cheney's proof. Monotone functional interpretation of (the negative translation of) our version of lemma 1 will automatically introduce the main notion needed for the quantitative analysis of the proof namely the concept of so-called 'r-clusters of δ-roots'. This concept, furthermore, is the key for the elimination of the use of $(*)$ (i.e. WKL) on which Cheney's proof of lemma 1 relies.[12]

3.2 Analysing the structure of the proof

The main goal of the paper is to extract from Cheney's proof [7] of Jackson's theorem [9] an effective modulus of uniqueness which can be used, as it will be shown in section 4.2, to compute the best L_1-approximation, p_b, from P_n of a given function $f \in C[0,1]$ with arbitrary precision.[13] In order to carry out the analysis we need to formalize Cheney's proof. The first

[12] It is the argument that 'δ', in the middle of page 219 in [7], is strictly positive which uses $(*)$. See section 3.10 and Remark 3.10.3 for more information.

[13] P_n is a Haar subspace of $C[0,1]$ of dimension $n+1$.

step we take in this direction is to list the main formulas used in the proof and to show how they are combined into lemmas. As mentioned before, each lemma will be analyzed separately. The functional interpretation of the lemma shows which functionals *can be* extracted from the proof of the lemma. But not all the functionals need to be presented, since some of them will disappear in the analysis of the proof (see the treatment of modus pones in the soundness of functional interpretation, e.g. in [17]). By analyzing the structure of the whole proof we can see which functionals are relevant and *need to be* extracted in order to obtain the final result. Then we construct such functionals and prove that they realize the lemma. In section 4 we show how the final modulus Φ is obtained by combining these functionals.

In the propositions $A - K$ below we omitted the parameters f, n, p_1 and p_2, therefore, instead of A one should read $A(f, n, p_1, p_2)$, where n ranges over $\mathbb{N}$, $f \in C[0,1]$ and $p_1, p_2 \in P_n$, and the same holds for all the others propositions. We also use here and for the rest of this paper the defined functions $p(x) := \frac{p_1(x)+p_2(x)}{2}$ and $f_0(x) := f(x) - p(x)$ as shorthand notation. In the formulas and in the sketch of the proof presented below we use $\overline{x} := x_1, \ldots, x_n$ and $\overline{\sigma} := \sigma_1, \ldots, \sigma_{n+1}$. The following formulas are used in Cheney's proof:

$A := \bigwedge_{i=1}^{2}(\|f - p_i\|_1 - dist_1(f, P_n) = 0)$, i.e.
p_1 and p_2 are best L_1-*approximations* of f from P_n.

$B := \|f - p\|_1 - dist_1(f, P_n) = 0$, i.e. p is a best L_1-*approximation* of f.

$C := \|f_0\| = \frac{1}{2}\|f - p_1\| + \frac{1}{2}\|f - p_2\|$.

$C_1 := \forall \varepsilon \in \mathbb{Q}_+^* \exists \delta \in \mathbb{Q}_+^* \forall x, y \in [0,1](|x - y| < \delta \to |g(x) - g(y)| < \varepsilon)$,
where $g(x) := |f_0(x)| - \frac{1}{2}|f(x) - p_1(x)| - \frac{1}{2}|f(x) - p_2(x)|$.
The formula C_1 states that g is uniformly continuous.

$D := \forall x \in [0,1](|f_0(x)| = \frac{1}{2}(|f(x) - p_1(x)| + |f(x) - p_2(x)|))$.

$E := \exists x_0, \ldots, x_n \in [0,1]\big(\bigwedge_{i=0}^{n} f_0(x_i) = 0 \wedge \bigwedge_{i=1}^{n} x_{i-1} < x_i\big)$, i.e.
f_0 has at least $n+1$ distinct roots.

$F := \exists x_0, \ldots, x_n \in [0,1]\big(\bigwedge_{i=0}^{n} p_1(x_i) = p_2(x_i) \wedge \bigwedge_{i=1}^{n} x_{i-1} < x_i\big)$, i.e.
$p_1 - p_2$ has at least $n+1$ distinct roots.

$G := \forall x \in [0,1](p_1(x) = p_2(x))$, alternatively, $\|p_1 - p_2\|_1 = 0$ or $p_1 = p_2$.

$H(h) := \|f_0 - h\|_1 \geq \|f_0\|_1$.

$I(\overline{x},\overline{\sigma},h) := \sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx > 0$, where $x_0 := 0$ and $x_{n+1} := 1$.

$J(\overline{x}) := \exists y \in [0,1](f_0(y) = 0 \wedge \bigwedge_{i=0}^{n+1} x_i \neq y)$, where $x_0 := 0$ and $x_{n+1} := 1$.

$K := \forall x \in [0,1](f_0(x) = 0 \to p_1(x) = p_2(x))$.

The first part of the proof (which we call derivation $\mathcal{D}_1$) is very simple and derives K from the assumption A,

$$
\begin{prooftree}
\AxiomC{$[A]$}
\AxiomC{$[A]$}
\AxiomC{$A \to B$}
\BinaryInfC{B}
\BinaryInfC{$A \wedge B$}
\AxiomC{$A \wedge B \to C$}
\BinaryInfC{C}
\AxiomC{C_1}
\BinaryInfC{$C \wedge C_1$}
\AxiomC{$C \wedge C_1 \to D$}
\BinaryInfC{D}
\AxiomC{$D \to K$}
\BinaryInfC{K}
\end{prooftree}
$$

The most involved part of the proof (which includes the application of Lemma 1) is when we want to prove that f_0 has $n+1$ distinct roots. In the derivation below we use $\overline{\sigma}' := \sigma'_1, \ldots, \sigma'_{n+1}$, where $\sigma'_i := sgn\,(f_0)(\frac{x_{i-1}+x_i}{2})$. Moreover, $\forall \overline{x} := \forall x_1 \leq \ldots \leq x_n$, where $\forall x_1 \leq \ldots \leq x_n\ Q(\overline{x})$ is an abreviation for $\forall x_1, \ldots, x_n\ (x_1 \leq \ldots \leq x_n \to Q(\overline{x}))$.

$$
\begin{prooftree}
\AxiomC{$[A]$}
\AxiomC{$A \to B$}
\BinaryInfC{B}
\AxiomC{$B \to \forall h\ H(h)$}
\BinaryInfC{$\forall h\ H(h)$}
\AxiomC{$\forall \overline{x}, \overline{\sigma} \exists h\ I(\overline{x},\overline{\sigma},h)$}
\AxiomC{$\forall \overline{x}, h\ (\forall \lambda H(\lambda h) \wedge I(\overline{x},\overline{\sigma}',h) \to J(\overline{x}))$}
\BinaryInfC{$\forall \lambda H(\lambda \tilde{h}) \to \forall \overline{x} J(\overline{x})$}
\BinaryInfC{$\forall \overline{x}\ J(\overline{x})$}
\end{prooftree}
$$

We call this derivation $\mathcal{D}_2$. An outline of the whole proof in the form of an informal natural deduction derivation is presented below,

$$
\begin{prooftree}
\AxiomC{$\mathcal{D}_1$}
\UnaryInfC{K}
\AxiomC{$\mathcal{D}_2$}
\UnaryInfC{$\forall \overline{x}\ J(\overline{x})$}
\AxiomC{$\forall \overline{x}\ J(\overline{x}) \to E$}
\BinaryInfC{E}
\BinaryInfC{$K \wedge E$}
\AxiomC{$K \wedge E \to F$}
\BinaryInfC{F}
\AxiomC{$F \to G$}
\BinaryInfC{G}
\RightLabel{$[A]$}
\UnaryInfC{$A \to G$}
\end{prooftree}
$$

Remark 3.2 *We assume that real numbers are represented as Cauchy sequences* $(a_n)_{n\in\mathbb{N}}$ *of rational number with fixed rate of convergence (say* 2^{-n}*) i.e.* $\forall k, \tilde{k} \geq n(|(a)_k - (a)_{\tilde{k}}| \leq 2^{-n})$. *In this way, equality* $=_{\mathbb{R}}$ *(similarly* $\leq_{\mathbb{R}}$ *and* $\geq_{\mathbb{R}}$*) between real numbers is a* $\forall$*-statement (for any point* k *in the Cauchy sequence the approximants are close by* 2^{-k}*) and strict inequality* $<_{\mathbb{R}}$ *is a* $\exists$*-statement (there exists a point* $k+1$ *in the sequence such that the approximants are distant by* 2^{-k}*). We call those: 'hidden quantifiers'. For example, let* $a, b \in \mathbb{R}$*, then* $a <_{\mathbb{R}} b$ *is an abreviation for* $\exists k \in \mathbb{N}((a)_{k+1} + 2^{-k} <_{\mathbb{Q}} (b)_{k+1})$*. In the analysis below we avoid going into the representation of the real numbers by observing that* $a <_{\mathbb{R}} b$ *can be written either as* $\exists r \in \mathbb{Q}^*_+(a <_{\mathbb{R}} b + r)$ *or* $\exists r \in \mathbb{Q}^*_+(a \leq_{\mathbb{R}} b + r)$*. The idea is that, if* $a <_{\mathbb{R}} b$ *occurs positively we write it as* $\exists r \in \mathbb{Q}^*_+(a <_{\mathbb{R}} b + r)$ *and if it occurs negatively we write it as* $\exists r \in \mathbb{Q}^*_+(a \leq_{\mathbb{R}} b + r)$*, in this way after prenexing these quantifiers the matrix is purely existential and (given the prenexed quantifiers have a* $\forall\exists$ *form as described in theorem 2.1) we can apply our meta-theorem 2.1. In the beginning of the analysis of each lemma we present the hidden quantifiers that are relevant for the final modulus.*

Remark 3.3 *In general we can only apply our meta-theorem 2.1 if* P_n *is replaced by* $K_{f,n}$*. As it happened, only in section 3.5 this limitation really matters. Nonetheless, as we discussed already, at the end we show that the final result actually holds for* P_n*.*

3.3 Lemma $A \to B$ [Triangle inequality]

The first lemma states,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}\Big(\bigwedge_{i=1}^{2} \|f - p_i\|_1 = dist_1(f, P_n) \to \|f - p\|_1 = dist_1(f, P_n)\Big).$$

As described in the previous section, the first step is to present the hidden quantifiers,

$$\begin{aligned}&\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}\\ &\quad\big(\forall \delta \in \mathbb{Q}^*_+(\textstyle\bigwedge_{i=1}^{2} \|f - p_i\|_1 - dist_1(f, P_n) \leq \delta) \to \forall \varepsilon \in \mathbb{Q}^*_+(\|f - p\|_1 - dist_1(f, P_n) < \varepsilon)\big).\end{aligned}$$

Then we look at the functional interpretation of the lemma,

$$\begin{aligned}&(1)\ \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+ \exists \delta \in \mathbb{Q}^*_+\\ &\qquad\qquad\big(\textstyle\bigwedge_{i=1}^{2} \|f - p_i\|_1 - dist_1(f, P_n) \leq \delta \to \|f - p\|_1 - dist_1(f, P_n) < \varepsilon\big).\end{aligned}$$

We see now that (1) has the same structure as the formula A in theorem 2.1. Therefore, we are sure to find a functional Φ_1, depending at most on n, f and ε, such that,[14]

[14]Since in theorem 2.1 we used 2^{-k} (with $k \in \mathbb{N}$) instead of $\delta \in \mathbb{Q}^*_+$, the upper bound on k guaranteed by the meta-theorem gives a lower bound on δ.

$$(2)\ \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+ \exists \delta \geq \Phi_1(f,n,\varepsilon)$$
$$\big(\textstyle\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \delta) \to \|f-p\|_1 - dist_1(f,P_n) < \varepsilon\big).$$

Since we have monotonicity in δ the functional Φ_1 actually realizes δ. The same phenomenon will happen in all the following lemmas, i.e. the lower bounds will always be realizing functionals for the variables they bound. Here, it is obvious how to construct Φ_1,

Claim 3.4 *The functional* $\boxed{\Phi_1(f,n,\varepsilon) := \Phi_1(\varepsilon) := \varepsilon}$ *does the job.*[15]

Proof: Suppose (1) $\|f-p_1\|_1 - dist_1(f,P_n) < \varepsilon$ and (2) $\|f-p_2\|_1 - dist_1(f,P_n) < \varepsilon$. Multiplying (1) and (2) by 1/2 and adding them together we get $1/2(\|f-p_1\|_1 + \|f-p_2\|_1) - dist_1(f,P_n) < \varepsilon$. By the triangle inequality for the L_1-*norm*, $1/2(\|2f-p_1-p_2\|_1) - dist_1(f,P_n) < \varepsilon$, i.e. $\|f-p\|_1 - dist_1(f,P_n) < \varepsilon$.

□

Remark 3.5 *The reader may have noticed that from* (1) *to* (2) *we changed from* $\leq$ *to* $<$ *in the premise of the implication. The reason we wrote* $\leq$ *first was just to show that the lemma could be written in the form of A (from theorem 2.1) and that a functional realizing* δ *was guaranteed by our meta-theorem. Since* $a \leq b/2$ *implies* $a < b$ *(and the reverse implication holds without the factor* 1/2*) we normally write the relation that yields the optimal bound. When analysing the following lemmas we often claim that some sentence is an instance of our meta-theorem 2.1 without bothering to write it explicitly in the form of A. We hope the reader can see that through the implications mentioned above these lemmas could in fact be written in the form of A.*

3.4 Lemma $A \wedge B \to C$ [Definition of L^1 norm]

The lemma states,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}$$
$$\big(\textstyle\bigwedge_{i=1}^{2}(\|f-p_i\|_1 = dist_1(f,P_n)) \to \|f-p\|_1 - 1/2\|f-p_1\|_1 - 1/2\|f-p_2\|_1 = 0\big).$$

After presenting the hidden quantifiers and performing the functional interpretation we come again to the same logical structure of the formula in theorem 2.1, and again we know that there must exist a functional Φ_2 depending at most on n, f and ε such that,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+$$
$$\big(\textstyle\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \Phi_2(f,n,\varepsilon)) \to$$
$$|\ \|f-p\|_1 - 1/2\|f-p_1\|_1 - 1/2\|f-p_2\|_1\ | < \varepsilon\big).$$

[15]Note that in fact Φ_1 is independent of n and f. We adopt the convention that parameters not used in the definition of the functionals will be dropped.

Again, the choice of Φ_2 is simple,

Claim 3.6 *The functional* $\boxed{\Phi_2(f, n, \varepsilon) := \Phi_2(\varepsilon) := \varepsilon}$ *does the job.*

Proof: Suppose (1) $\|f - p_1\|_1 - dist_1(f, P_n) < \varepsilon$ and (2) $\|f - p_2\|_1 - dist_1(f, P_n) < \varepsilon$. By previous lemma we have (3) $\|f - p\|_1 - dist_1(f, P_n) < \varepsilon$. And $\frac{(1)+(2)}{2}$ gives, (4) $1/2(\|f - p_1\|_1 + \|f - p_2\|_1) - dist_1(f, P_n) < \varepsilon$. From (3) and (4), we have, $|\ \|f - p\|_1 - 1/2\|f - p_1\|_1 - 1/2\|f - p_2\|_1\ | < \varepsilon$, i.e., we used the fact that $a \in [0, m)$ and $b \in [0, m)$ then $|a - b| \in [0, m)$.

$\square$

3.5 Lemma C_1 [Continuity of g(x)]

Let $g(x) := |f_0(x)| - \frac{1}{2}|f(x) - p_1(x)| - \frac{1}{2}|f(x) - p_2(x)|$. Based on the continuity of f, p_1 and p_2 we derive that g is continuous,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+; x, y \in [0,1] \exists \delta \in \mathbb{Q}^*_+ \big(|x - y| \leq \delta \rightarrow |g(x) - g(y)| < \varepsilon\big).$$

Note that here we can again apply the meta-theorem 2.1 and we are sure to find a function Δ depending only f, n and ε such that,[16]

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+; x, y \in [0,1] \big(|x - y| < \Delta(f, n, \varepsilon) \rightarrow |g(x) - g(y)| < \varepsilon\big).$$

We write $\Delta(f, n, \varepsilon)$ as $\omega_{f,n}(\varepsilon)$. In this section we show how the modulus of continuity $\omega_{f,n}(\varepsilon)$ can be computed using only n, the modulus of continuity of f, ω_f, and an upper bound $M_f \geq \|f\|_\infty$ (in section 4 we show that we just need $M_f \geq \sup_{x \in [0,1]} |f(x) - f(0)|$, for instance $\lceil \frac{1}{\omega_f(1)} \rceil$, so that the final result only depends on ω_f).

3.5.1 Modulus of the sum

Given the moduli of continuity ω_f and ω_g for the functions f and g respectively, we find the modulus of continuity for $f + g$, ω_{f+g}, in the following way. We have,

$$|x - y| < \omega_f(\varepsilon/2) \rightarrow |f(x) - f(y)| < \varepsilon/2.$$

$$|x - y| < \omega_g(\varepsilon/2) \rightarrow |g(x) - g(y)| < \varepsilon/2.$$

[16] Here it is fundamental that p_1 and p_2 live in the compact space $K_{f,n}$ otherwise the modulus of continuity for g would depend also on these elements and we would be unable to get a uniform modulus of uniqueness at the end.

Therefore,

$$|x-y| < \min\{\omega_f(\varepsilon/2), \omega_g(\varepsilon/2)\} \to (|f(x)-f(y)| < \varepsilon/2 \wedge |g(x)-g(y)| < \varepsilon/2).$$

$$|x-y| < \min\{\omega_f(\varepsilon/2), \omega_g(\varepsilon/2)\} \to |f(x)+g(x)-f(y)-g(y)| < \varepsilon.$$

Hence, $\omega_{f+g}(\varepsilon) = \min\{\omega_f(\varepsilon/2), \omega_g(\varepsilon/2)\}$.

3.5.2 Modulus of a constant times a function

We show that $\omega_{af}(\varepsilon) = \omega_f(\frac{\varepsilon}{a})$, for all $a \in \mathbb{Q}^*_+$,

$$|x-y| < \omega_f(\tfrac{\varepsilon}{a}) \to |f(x)-f(y)| < \tfrac{\varepsilon}{a},$$

$$|x-y| < \omega_f(\tfrac{\varepsilon}{a}) \to |af(x)-af(y)| < \varepsilon,$$

$$|x-y| < \omega_{af}(\varepsilon) \to |af(x)-af(y)| < \varepsilon.$$

3.5.3 Modulus of p_1 and p_2

Let $p_i \in K_{f,n}$. Then $\|p_i\|_1 \leq \frac{5}{2}\|f\|_1 \leq \frac{5}{2}\|f\|_\infty$. If $p_i(x) = a_n x^n + \ldots + a_1 x + a_0$ and $p_i^*(x) = \frac{a_n x^{n+1}}{n+1} + \ldots + \frac{a_1 x^2}{2} + a_0 x$ then for all $x \in [0,1]$ we have,

$$|p_i^*(x)| = |\textstyle\int_0^x p_i(x)dx| \leq \int_0^x |p_i(x)|dx \leq \|p_i\|_1 \leq \frac{5}{2}\|f\|_\infty.$$

By Markov's inequality (see e.g. [7]),

$$\|p_i\|_\infty = \|(p_i^*)'\|_\infty \leq 2(n+1)^2(\tfrac{5}{2}\|f\|_\infty) = 5(n+1)^2\|f\|_\infty.$$

If we apply Markov's inequality once more we get,

$$\|p_i'\|_\infty \leq 2n^2 5(n+1)^2\|f\|_\infty < 10(n+1)^4\|f\|_\infty.$$

By the mean value theorem this implies that p_i has Lipschitz constant $10(n+1)^4\|f\|_\infty$ on $[0,1]$, i.e. $\frac{\varepsilon}{10(n+1)^4\|f\|_\infty}$ is a modulus of uniform continuity for p_i on $[0,1]$. Given an upper bound M_f on $\|f\|_\infty$ we have,[17]

$$\omega_{p_i}(\varepsilon) := \frac{\varepsilon}{10(n+1)^4 M_f}.$$

[17] It should be clear that given f together with its modulus of continuity, ω_f, there is a simple algorithm to compute M_f, just take for instance $M_f := \max\{|f(i.\omega_f(1))| : 0 \leq i \leq \lfloor\frac{1}{\omega_f(1)}\rfloor\} + 1$.

Remark 3.7 *Here we present how one gets a bound on the coefficients of p given* $\|p\|_1$ *(or some bound on* $\|p\|_1$*). Let* p^i *denote the i-th derivative of p. Above we have shown that* $\|p\|_\infty \leq 2(n+1)^2\|p\|_1$ *which by Markov's inequality yields (1)* $\|p^i\|_\infty \leq (2(n+1)^2)^{i+1}\|p\|_1$. *Since* $p^i(x) = \frac{n!}{(n-i)!}a_n x^{n-i} + \ldots + i!a_i$, *from (1) we get* $|i!\ a_i| \leq (2(n+1)^2)^{i+1}\|p\|_1$ *which implies* $|a_i| \leq \frac{(2(n+1)^2)^{i+1}}{i!}\|p\|_1$.

3.5.4 The Modulus of Continuity $\omega_{f,n}$

Now we can present $\omega_{f,n}$ as a function of ω_f and n (note that we can take $\omega_{|f|} := \omega_f$),

$$\begin{aligned}
\omega_{f,n}(\varepsilon) &= \min\{\omega_{|f-p|}(\varepsilon/2), \omega_{1/2|f-p_1|}(\varepsilon/4), \omega_{1/2|f-p_2|}(\varepsilon/4)\} \\
&= \min\{\omega_{f-p}(\varepsilon/2), \omega_{f-p_1}(\varepsilon/2), \omega_{f-p_2}(\varepsilon/2)\} \\
&= \min\{\omega_f(\varepsilon/4), \omega_{p_1}(\varepsilon/4), \omega_{p_2}(\varepsilon/4)\} \\
&= \min\{\omega_f(\frac{\varepsilon}{4}), \frac{\varepsilon}{40(n+1)^4 M_f}\}.
\end{aligned}$$

3.6 Lemma $C \wedge C_1 \to D$ [Integrand is ≤ 0 and continuous]

Let $g(x) := |f(x) - p(x)| - 1/2|f(x) - p_1(x)| - 1/2|f(x) - p_2(x)|$. The lemma says,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}\big(\textstyle\int_0^1 g(x)dx = 0 \to \forall x \in [0,1](g(x) = 0)\big).$$

After presenting the hidden quantifiers and applying functional interpretation we observe that again we can apply theorem 2.1, and we are guaranteed to find a functional $\Phi_3(f, n, \varepsilon)$ such that,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+\big(|\textstyle\int_0^1 g(x)dx| \leq \Phi_3(f, n, \varepsilon) \to \|g\|_\infty \leq \varepsilon\big).$$

Let $\omega_{f,n}(\varepsilon) : \mathbb{Q}^*_+ \to \mathbb{Q}^*_+$ denote the uniform modulus of continuity of the function $g \in C[0,1]$, proved to exist in the analysis of Lemma C_1 (section 3.5).

Claim 3.8 *The functional* $\boxed{\Phi_3(f, n, \varepsilon) := \Phi_3(\omega_{f,n}, \varepsilon) := \frac{\varepsilon}{2}.\min\{\frac{1}{2},\ \omega_{f,n}(\frac{\varepsilon}{2})\}}$ *does the job.*

Proof: Assume $\|g\|_\infty > \varepsilon$, since $\forall x \in [0,1] g(x) \leq 0$ we conclude $\exists x_0 \in [0,1] g(x_0) < -\varepsilon$. By the continuity of g we also have,

$$\forall x \in [0,1]\big(|x - x_0| < \omega_{f,n}(\varepsilon/2) \to g(x) < -\varepsilon/2\big).$$

If $x_0 < 1/2$ then,

$$|\int_0^1 g(x)dx| > |\int_{x_0}^{\min\{1, x_0+\omega_{f,n}(\varepsilon/2)\}} -\varepsilon/2\ dx| = \frac{\varepsilon}{2}\min\{\frac{1}{2}, \omega_{f,n}(\frac{\varepsilon}{2})\},$$

otherwise ($x_0 \geq 1/2$),

$$|\int_0^1 g(x)dx| > |\int_{\max\{0,x_0-\omega_{f,n}(\varepsilon/2)\}}^{x_0} -\varepsilon/2\ dx| = \frac{\varepsilon}{2}\min\{\frac{1}{2},\omega_{f,n}(\frac{\varepsilon}{2})\}.$$

From this we conclude,

$$|\int_0^1 g(x)dx| > \frac{\varepsilon}{2}\min\{\frac{1}{2},\omega_{f,n}(\frac{\varepsilon}{2})\}.$$

□

3.7 Lemma $D \to K$ [If $f_0(x) = 0$ then $p_1(x) = p_2(x)$]

Let $f_1(x) := 1/2(|f(x) - p_1(x)| + |f(x) - p_2(x)|)$, the lemma says,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x \in [0,1]$$
$$\big(\| |f_0| - f_1\|_\infty = 0 \to (|f_0(x)| = 0 \to p_1(x) = p_2(x))\big).$$

Again we are sure to find functionals $\Phi_4(f,n,\varepsilon)$ and $\Phi_5(f,n,\varepsilon)$ such that,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x \in [0,1]; \varepsilon \in \mathbb{Q}^*_+$$
$$\big(\| |f_0| - f_1\|_\infty \leq \Phi_4(f,n,\varepsilon) \to (|f_0(x)| \leq \Phi_5(f,n,\varepsilon) \to |p_1(x) - p_2(x)| \leq \varepsilon)\big).$$

Claim 3.9 *The functionals* $\boxed{\Phi_4(f,n,\varepsilon) := \Phi_4(\varepsilon) := \varepsilon/8}$ *and* $\boxed{\Phi_5(f,n,\varepsilon) := \Phi_5(\varepsilon) := \varepsilon/8}$ *do the job.*

Proof: Trivial.

□

3.8 Lemma $F \to G$ [If p has $n+1$ roots then $p = 0$]

The lemma states that if $p_1(x) - p_2(x)$ has $n+1$ distinct roots then $\forall x \in [0,1] p_1(x) = p_2(x)$,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n} \forall x_0, \ldots, x_n \in [0,1]$$
$$\big(\bigwedge_{i=1}^n (x_i < x_{i+1}) \wedge \bigwedge_{i=0}^n (p_1(x_i) = p_2(x_i)) \to \|p_1 - p_2\|_\infty = 0\big)$$

Then we present the hidden quantifiers and apply functional interpretation,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; r, \varepsilon \in \mathbb{Q}^*_+; x_0, \ldots, x_n \in [0,1] \exists \delta \in \mathbb{Q}^*_+$$
$$\big(\bigwedge_{i=1}^n (x_{i-1} + r \leq x_i) \wedge \bigwedge_{i=0}^n (|p_1(x_i) - p_2(x_i)| \leq \delta) \to \|p_1 - p_2\|_\infty \leq \varepsilon\big)$$

By theorem 2.1 we are sure to find a functional Φ_6 realizing δ.

Claim 3.10 *The functional* $\boxed{\Phi_6(f,n,r,\varepsilon) := \Phi_6(n,r,\varepsilon) := \frac{\lfloor n/2\rfloor!\lceil n/2\rceil! r^n}{(n+1)}\varepsilon}$ *does the job.*

Proof: See [15], pages 82–83.

□

Remark 3.11 *In fact, the functional* Φ_6 *does the job for* $p_1, p_2 \in P_n$ *(not only for* $p_1, p_2 \in K_{f,n}$*).*

3.9 Lemma $B \to \forall h\ H(h)$ [Definition of best L_1-approximation]

This lemma is a trivial consequence of the definition of $dist_1$,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}\big(\|f_0\|_1 = dist_1(f, P_n) \to \forall h \in P_n(\|f_0 - h\|_1 \geq \|f_0\|_1)\big).$$

We can easily find a functional $\Phi_7(f,n,\varepsilon)$ s.t.,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \varepsilon \in \mathbb{Q}^*_+ \\ \big(\|f_0\|_1 - dist_1(f, P_n) \leq \Phi_7(f,n,\varepsilon) \to \forall h \in P_n(\|f_0 - h\|_1 + \varepsilon \geq \|f_0\|_1)\big).$$

Claim 3.12 *The functional* $\boxed{\Phi_7(f,n,\varepsilon) := \Phi_7(\varepsilon) := \varepsilon}$ *does the job.*

Proof: Assume (1) $\|f_0\|_1 - dist_1(f, P_n) \leq \varepsilon$. By the definition of $dist_1$ we have for any $h \in P_n$ (2) $\|f_0 - h\|_1 \geq dist_1(f, P_n)$. From (1) and (2) we have $\|f_0 - h\|_1 + \varepsilon \geq \|f_0\|_1$.

□

3.10 Lemma $\forall \overline{x}, h\ (\forall \lambda H(\lambda h) \wedge I(\overline{x}, \overline{\sigma}', h) \to J(\overline{x}))$ [Lemma 1]

This is the most intricate lemma used in the proof and therefore we analyze it in greater detail. We first rewrite the lemma in order to show how it is used in the proof of Jackson's theorem in [7].

Lemma 1 Let $f \in C[0,1]$ and $h, p_1, p_2 \in P_n$. If f_0 has at most n roots then either $\int_0^1 (h(x)\, sgn(f_0)(x))dx = 0$ or there exists a $\lambda \in \mathbb{R}$ such that $\int_0^1 |f_0(x) - \lambda h(x)|dx < \int_0^1 |f_0(x)|dx$.

Proof: Assume that all the roots of f_0 are among $0 = x_0 \leq x_1 \leq \ldots \leq x_{n+1} = 1$ and w.l.g. assume that $\int_0^1 (h(x)\, sgn(f_0)(x))dx > 0$. Let $B' := \bigcup_{i=0}^{n+1}(x_i - r, x_i + r)$ and $B := B' \bigcap [0,1]$. Let $A := [0,1]\backslash B$. Make r small enough so that $\int_A (h(x)\, sgn(f_0)(x))dx > \int_B |h(x)|dx$. Note that A is a finite union of closed intervals which contain no roots of f_0, therefore $\delta := \min\{|f_0(x)| : x \in A\}$ is positive. Hence we can find a λ such that $0 < \lambda\|h\|_\infty < \delta$, and for points $x \in A$, $sgn(f_0 - \lambda h)(x) = sgn(f_0)(x)$, which implies (see [7] or the proof of claim 3.13 for details) that $\int_0^1 |f_0(x) - \lambda h(x)| dx < \int_0^1 |f_0(x)| dx$.

$\square$

The contraposition of the above stated lemma is then used in the proof. Now we analyze the logical structure of the lemma in order to discover which functionals can be extracted and what their dependencies are:

3.10.1 Logical Analysis

First we replace P_n with $K_{f,n}$ in the statement of lemma 1. Then we rewrite the integral of $h\ sgn(f_0)$ over the interval $[0,1]$ as a sum of integrals over smaller intervals (which are guaranteed by the premise to contain no root of f_0) as described in section 3.1. In this way lemma 1 can be formally written as,

$$\begin{aligned} &\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n \\ &\qquad \big(\forall y \in [0,1](f_0(y) = 0 \to \textstyle\bigvee_{i=0}^{n+1} x_i = y) \ \to \\ &\qquad\qquad\qquad \textstyle\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x) dx = 0 \vee \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 < \|f_0\|_1)\big) \end{aligned}$$

where $\sigma_i := sgn(f_0)(\frac{x_{i-1}+x_i}{2})$, $x_0 := 0$ and $x_{n+1} := 1$. The above formula implies[18] (relative to PA^ω),

$$\begin{aligned} &\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \eta, \delta \in \mathbb{Q}_+^* \\ &\quad \exists r, l \in \mathbb{Q}_+^* \big(\forall y \in [0,1](|f_0(y)| < \delta \to \textstyle\bigvee_{i=0}^{n+1} |x_i - y| \leq r) \ \to \\ &\qquad\qquad\qquad |\textstyle\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x) dx| < \eta \vee \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + l < \|f_0\|_1)\big) \end{aligned}$$

Here we notice that η can be eliminated since $h/\eta \in P_n$,

$$(*) \begin{cases} \forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta \in \mathbb{Q}_+^* \\ \exists r, l \in \mathbb{Q}_+^* \big(\forall y \in [0,1](|f_0(y)| < \delta \to \bigvee_{i=0}^{n+1} |x_i - y| \leq r) \ \to \\ \qquad\qquad |\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x) dx| < 1 \vee \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + l < \|f_0\|_1)\big) \end{cases}$$

3.10.2 Functionals Realizing Lemma 1

By observing that $(*)$ has the same logical form as the formula A in the meta-theorem 2.1 we are sure to find functionals $\Phi_8(f, n, \delta, h)$ and $\Phi_9(f, n, \delta, h)$ such that,

[18] In fact the formula we obtain below is a weakening of Lemma 1. It will turn out in the course of the extraction that it is sufficient to consider this weakened form to get the final result.

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta \in \mathbb{Q}^*_+$$
$$\big(\forall y \in [0,1](|f_0(y)| < \delta \to \bigvee_{i=0}^{n+1} |x_i - y| \leq \Phi_9(f,n,\delta,h)) \;\to$$
$$|\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx| < 1 \vee \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + \Phi_8(f,n,\delta,h) < \|f_0\|_1)\big).$$

Claim 3.13 *The functionals* $\boxed{\Phi_8(f,n,\delta,h) := \Phi_8(n,\delta,h) := \frac{\delta}{3\|h\|_\infty}}$ *and* $\boxed{\Phi_9(f,n,\delta,h) := \Phi_9(n,h) := \frac{1}{6(n+1)\|h\|_\infty}}$ *do the job.*[19]

Proof: We have to prove that,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta \in \mathbb{Q}^*_+$$
$$\big(\forall y \in [0,1](|f_0(y)| < \delta \to \bigvee_{i=0}^{n+1} |x_i - y| \leq \tfrac{1}{6(n+1)\|h\|_\infty}) \;\to$$
$$|\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx| \leq 1 \vee \exists \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + \tfrac{\delta}{3\|h\|_\infty} < \|f_0\|_1)\big).$$

For simplicity we define $r := \frac{1}{6(n+1)\|h\|_\infty}$. Note that now we do not require y (in the above formula) to be a root of f_0 but a δ-root (i.e. $|f_0(y)| < \delta$). As a consequence y will not necessarily be equal to one of the x_i's but only 'r-close' to some x_i, we say that y belongs to the (x_i, r)-cluster.[20] Now we follow the original proof. Take n points, $x_1, \ldots, x_n$, such that (1) $0 = x_0 \leq x_1 \leq \ldots \leq x_{n+1} = 1$ and suppose that (2) all δ-roots of f_0 belong to at least one of the (x_i, r)- clusters. Moreover, suppose w.l.g. that (3) $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx > 1$. Let $B' := \bigcup_{i=0}^{n+1}(x_i - r, x_i + r)$, $B := [0,1] \bigcap B'$ and $A := [0,1] \backslash B$. Note that A can be written as the union of smaller (possibly containing a single point) intervals $A_i := [x_{i-1} + \min\{r, \frac{x_{i-1}+x_i}{2}\}, x_i - \min\{r, \frac{x_{i-1}+x_i}{2}\}]$, for $1 \leq i \leq n+1$. Therefore by assumption (2) we have $sgn(f_0)(x) = \sigma_i$, for $x \in A_i$, where $\sigma_i = sgn(f_0)(\frac{x_{i-1}+x_i}{2})$, and then $\sum_{i=1}^{n+1} \sigma_i \int_{A_i} h(x)\, dx = \int_A (h(x)\, sgn(f_0)(x))\, dx$. Following the original proof we would have to shrink B until $\int_B |h(x)|dx \leq 1/3$, but $\int_B |h(x)|dx \leq |B| \cdot \|h\|_\infty \leq (n+1)\frac{1}{3(n+1)\|h\|_\infty}\|h\|_\infty = 1/3$. Then by (3) we get (4) $\int_A (h(x)\, sgn(f_0)(x))dx > 1/3 + \int_B |h(x)|dx$. And by the definition of A and (2) we have $\min\{|f_0(x)| : x \in A\} \geq \delta$, therefore taking $\lambda := \frac{\delta}{\|h\|_\infty}$ we have (5)

[19] In order to be precise we should have written $\max\{1, \|h\|_\infty\}$ instead of $\|h\|_\infty$ in the definition of Φ_8 and Φ_9, so that they are always defined. This can be seen to be not necessary because we only apply these functionals to an h with uniform norm different from zero (see section 3.12). For another comment on the norm of the function h see remark 3.14.

[20] This weakening is fundamental to the elimination of the WKL, as mentioned in section 3.1. We discuss this point in more details in section 3.10.3.

$sgn(f_0 - \lambda h)(x) = sgn(f_0)(x)$, for $x \in A$. Hence,

$$\begin{aligned}
\|f_0 - \lambda h\|_1 &= \int_A |f_0 - \lambda h| - \int_B |f_0 - \lambda h| \\
&\overset{(5)}{=} \int_A (f_0 - \lambda h)\, sgn(f_0) + \int_B |f_0 - \lambda h| \\
&= \int_A f_0\, sgn(f_0) - \lambda \int_A h\, sgn(f_0) + \int_B |f_0 - \lambda h| \\
&\leq \int_A f_0\, sgn(f_0) - \lambda \int_A h\, sgn(f_0) + \int_B |f_0| + \lambda \int_B |h| \\
&= \int_A |f_0| + \int_B |f_0| + \lambda \int_B |h| - \lambda \int_A h\, sgn(f_0) \\
&= \int_0^1 |f_0| + \lambda \int_B |h| - \lambda \int_A h\, sgn(f_0).
\end{aligned}$$

Now we can add $\frac{\delta}{3\|h\|_\infty}$ on both sides of the inequality and put $\lambda = \frac{\delta}{\|h\|_\infty}$ in evidence to get,

$$\begin{aligned}
\|f_0 - \lambda h\|_1 + \frac{\delta}{3\|h\|_\infty} &\leq \|f_0\|_1 + \frac{\delta}{\|h\|_\infty}(1/3 + \int_B |h| - \int_A h\, sgn(f_0)) \\
&\overset{(4)}{<} \|f_0\|_1.
\end{aligned}$$

$\square$

As we mentioned, the lemma is going to be used in the contrapositive form,

$$\begin{aligned}
&\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \leq \ldots \leq x_n \in [0,1]; h \in P_n; \delta \in \mathbb{Q}^*_+ \\
&\quad \big(\forall \lambda \in \mathbb{R}(\|f_0 - \lambda h\|_1 + \tfrac{\delta}{3\|h\|_\infty} \geq \|f_0\|_1) \wedge \textstyle\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx > 1 \rightarrow \\
&\qquad\qquad \exists y \in [0,1](|f_0(y)| < \delta \wedge \textstyle\bigwedge_{i=0}^{n+1} |x_i - y| > \frac{1}{6(n+1)\|h\|_\infty})\big).
\end{aligned}$$

Remark 3.14 *The functionals Φ_8 and Φ_9 should range over $\mathbb{Q}^*_+$, but $\|h\|_\infty \in \mathbb{R}_+$. Therefore, we should have written $\|h\|_{\infty,\mathbb{Q}}$ instead of $\|h\|_\infty$ in the definition of Φ_8 and Φ_9, where $\|h\|_{\infty,\mathbb{Q}}$ is a rational upper bound on $\|h\|_\infty$.*

3.10.3 Elimination of WKL

As we discussed already in the introduction, the logical method of monotone functional interpretation on which the proof of the general logical meta-theorem is based upon not only provides an algorithm for the extraction of the modulus of uniqueness Φ but also a constructive verification of Φ which can be formalized in intuitionistic arithmetic in all finite types HA^ω. In particular, we get from this that Jackson's theorem is provable in HA^ω despite the fact that Cheney's proof heavily relies on classical logic and the non-computational binary König's lemma WKL. We will not carry out the details of this intuitionistic **verification** since we focus in this paper on the **applied** aspect of **constructing** Φ. However, in 3.10.2 above we can see already how the constructivisation of Cheney's proof comes out of our analysis: as said before, WKL is used in the equivalent (see [25])[21] form of

$$(+)\ \forall f \in C[0,1] \forall a, b \in [0,1](a < b \to \exists x_0 \in [a,b](f(x_0) = \inf_{x_0 \in [a,b]} f(x)))$$

to conclude

$$\forall x \in [x_i + r, x_{i+1} - r](f(x) > 0) \to \inf_{x \in [x_i + r, x_{i+1} - r]} f(x) > 0.$$

After our replacement of 'roots x_i' by 'r-clusters of δ-roots' this transforms into

$$\forall x \in [x_i + r, x_{i+1} - r](f(x) > \delta) \to \inf_{x \in [x_i + r, x_{i+1} - r]} f(x) \geq \delta$$

which follows from the constructively valid 'ε-weakening'

$$(+)_\varepsilon \begin{cases} \forall f \in C[0,1] \forall a, b \in [0,1] \\ \quad (a < b \to \forall \varepsilon > 0 \exists x_0 \in [a,b](f(x_0) - \inf_{x_0 \in [a,b]} f(x) < \varepsilon)) \end{cases}$$

version of $(+)$ which eliminates the use of WKL.
Also the use of classical logic to find σ_i such that

$$\sigma_i =_0 0 \leftrightarrow f(\frac{x_i + x_{i+1}}{2}) \geq_{\mathbb{R}} 0$$

is no longer necessary since we now have that

$$f(\frac{x_i + x_{i+1}}{2}) \geq_{\mathbb{R}} \delta \vee f(\frac{x_i + x_{i+1}}{2}) \leq_{\mathbb{R}} -\delta$$

which can easily be decided since $\delta \in \mathbb{Q}_+^*$.

3.11 Lemma $\forall \overline{x}, \overline{\sigma} \exists h\, I(\overline{x}, \overline{\sigma}, h)$

In the second part of Cheney's proof he considers the case where f_0 has less than $n+1$ roots, from this assumption he arrives at a contradiction (using Lemma 1) when assuming that for any $h \in P_n$, $\int h\, sgn(f_0) = 0$. Here we show that for any given n points $x_1 \leq \ldots \leq x_n$ in the interval $[0,1]$ and for any $\sigma_1, \ldots, \sigma_{n+1} \in \{-1, 1\}$ (where σ_i will denote the sign of the function f_0 in the interval $[x_{i-1}, x_i]$) it is possible to find a function $h \in P_n$ such that $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h > 0$, where $x_0 = 0$ and $x_{n+1} = 1$. Formally,

[21] Note that $f \in C[0,1]$ is given together with a modulus of uniform continuity ω_f.

$$\forall n \in \mathbb{N}; x_1 \leq \ldots \leq x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n \big(\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h > 0 \big).$$

In the same way as we did in section 3.10.1 we present the hidden quantifier η in the inequality and since $h/\eta \in P_n$ we have,

$$\forall n \in \mathbb{N}; x_1 \leq \ldots \leq x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n \big(\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h > 1 \big).$$

The sentence above states the existence of a function $h \in P_n$, therefore there exists also a $k \in \mathbb{Q}_+^*$ such that, in the interval $[0,1]$, $k \geq \|h\|_\infty$. Here we can again apply our meta-theorem 2.1 and we are sure to find a function Φ_{10} depending only on n such that,[22]

$$\forall n \in \mathbb{N}; x_1 \leq \ldots \leq x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n$$
$$\big(\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h > 1 \wedge \Phi_{10}(n) \geq \|h\|_\infty \big).$$

Claim 3.15 *The function* $\boxed{\Phi_{10}(n) := 3^{n+1}(n+1)^{n+1}}$ *does the job.*

Before we give the proof for Claim 3.15 we need,

Claim 3.16 *Let $y_1, \ldots, y_m$ be points in $[0,1]$ such that $y_1 < y_2 < \ldots < y_m$. Define $h(x) := a(x - y_1)\ldots(x - y_m)$. For any $a \in \mathbb{R}^*$ and any two points $x_0 < x_1$ in $[0,1]$ such that h has no roots inside $[x_0, x_1]$,*

$$\int_{x_0}^{x_1} |h(x)|dx > \frac{x_1 - x_0}{3} \min\{|h(x_0 + \frac{x_1 - x_0}{3})|, |h(x_1 - \frac{x_1 - x_0}{3})|\}.$$

Proof: It is sufficient to show that for any $x \in (x_0 + \frac{x_1 - x_0}{3}, x_1 - \frac{x_1 - x_0}{3})$,

$$|h(x)| \geq \min\{|h(x_0 + \frac{x_1 - x_0}{3})|, |h(x_1 - \frac{x_1 - x_0}{3})|\}.$$

Let $y_0 = 0$ and $y_{m+1} = 1$. First note that $y_i \leq x_0 < x_1 \leq y_{i+1}$ for some $0 \leq i < m$, where at least one of y_i and y_{i+1} is a root of h (w.l.g. say that y_i is a root of h). Now suppose that there is an $x^* \in (x_0 + \frac{x_1-x_0}{3}, x_1 - \frac{x_1-x_0}{3})$ such that $|h(x^*)| < \min\{|h(x_0 + \frac{x_1-x_0}{3})|, |h(x_1 - \frac{x_1-x_0}{3})|\}$. Then clearly h has a local minimum inside the interval $(x_0 + \frac{x_1-x_0}{3}, x_1 - \frac{x_1-x_0}{3})$. On the other hand since $|h(y_i)| \leq |h(x_0 + \frac{x_1-x_0}{3})|$ and $|h(x^*)| \leq |h(x_0 + \frac{x_1-x_0}{3})|$ (and $y_i < x_0 + \frac{x_1-x_0}{3} < x^*$) the polynomial h has a local maximum in the interval (y_i, x^*). But we know in advance that h has at least one local extremum in each interval (y_i, y_{i+1}), $1 \leq i < m$. Since we have $m-1$ intervals we have found already m local extrema, a contradiction since the derivative of h has degree $m - 1$.

[22]Note that Φ_{10} does not depend on the points $x_1, \ldots, x_n$ neither on the σ's since they are elements from compact spaces.

$\square$

Proof of Claim 3.15: The claim states that for any given points $0 = x_0 \leq x_1 \leq \ldots \leq x_{n+1} = 1$ and for any $\sigma_1, \ldots, \sigma_{n+1} \in \{-1, 1\}$ there exists an $h \in P_n$ such that $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx > 1$, and $\|h\|_\infty \leq 3^{n+1}(n+1)^{n+1}$. Let $0 = x_0 \leq x_1 \leq \ldots \leq x_{n+1} = 1$ and $\sigma_1, \ldots, \sigma_{n+1} \in \{-1, 1\}$ be given. Ignore all the points x_j such that $x_i = x_j$ and $i < j$. We are left with $\tilde{n}+1$ points $0 = x_{a_0} < x_{a_1} < \ldots < x_{a_{\tilde{n}+1}} = 1$ where $a_{i-1} < a_i$, $a_i \in \{0, \ldots, n+1\}$ and $\tilde{n} \leq n$. Let $\tilde{x}_i := x_{a_i}$ and $\tilde{\sigma}_i := \sigma_{a_i}$. Since we have eliminated just empty intervals we have for any function $h \in P_n$, $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)\ dx = \sum_{i=1}^{\tilde{n}+1} \tilde{\sigma}_i \int_{\tilde{x}_{i-1}}^{\tilde{x}_i} h(x)\ dx$. Among the points $\tilde{x}_1, \ldots, \tilde{x}_{\tilde{n}}$ pick only the points $\tilde{x}_i$ for which $\tilde{\sigma}_i \neq \tilde{\sigma}_{i+1}$. Finally, we are left with $m+1$ points $0 = \tilde{x}_{b_0} < \tilde{x}_{b_1} < \ldots < \tilde{x}_{b_{m+1}} = 1$ where $b_{i-1} < b_i$, $b_i \in \{0, \ldots, \tilde{n}+1\}$ and $m \leq \tilde{n}$. Let $y_i := \tilde{x}_{b_i}$ and $\sigma_i^* := \tilde{\sigma}_{b_i}$. Again we have $\sum_{i=1}^{\tilde{n}+1} \tilde{\sigma}_i \int_{\tilde{x}_{i-1}}^{\tilde{x}_i} h(x)dx = \sum_{i=1}^{m+1} \sigma_i^* \int_{y_{i-1}}^{y_i} h(x)dx$, for any $h \in P_n$. Then we define $h(x) := a(x - y_1) \ldots (x - y_m)$ and choose the $sgn(a) = (-1)^j$ ($j \in \{0, 1\}$) so that $\sum_{i=1}^{m+1} \sigma_i^* \int_{y_{i-1}}^{y_i} h(x)dx = \sum_{i=1}^{m+1} \int_{y_{i-1}}^{y_i} |h(x)|dx$. Therefore, for this specific h we have (1) $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx = \sum_{i=1}^{m+1} \int_{y_{i-1}}^{y_i} |h(x)|dx = \int_0^1 |h(x)|dx$.

By the pigeon hole principle for some $i \in \{1, \ldots, n+1\}$ we have $x_i - x_{i-1} \geq \frac{1}{n+1}$. Let $i_0 \in \{1, \ldots, n+1\}$ be such that $x_{i_0} - x_{i_0-1} \geq \frac{1}{n+1}$. Then,

$$
\begin{aligned}
\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx \quad &\overset{(1)}{=} \quad \int_0^1 |h(x)|dx \\
&\geq \quad \int_{x_{i_0-1}}^{x_{i_0}} |h(x)|dx \\
&\overset{\text{Claim 3.16}}{>} \quad \frac{x_{i_0} - x_{i_0-1}}{3} \min\{|h(x_{i_0-1} + \frac{(x_{i_0} - x_{i_0-1})}{3})|, \\
&\qquad\qquad |h(x_{i_0} - \frac{(x_{i_0} - x_{i_0-1})}{3})|\} \\
&\geq \quad \frac{1}{3(n+1)} \frac{|a|}{3^n(n+1)^n} = \frac{|a|}{3^{n+1}(n+1)^{n+1}}.
\end{aligned}
$$

Therefore, by choosing $j \in \{0, 1\}$ appropriately and taking $a := (-1)^j 3^{n+1}(n+1)^{n+1}$ we have $\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h(x)dx > 1$ and $\|h\|_\infty \leq 3^{n+1}(n+1)^{n+1}$.

$\square$

3.12 Eliminating the function h in Lemma 1

We have just shown that,

$$(1)\ \forall x_1 \le \ldots \le x_n \in [0,1]; \sigma_1, \ldots, \sigma_{n+1} \in \{-1,1\} \exists h \in P_n \quad \Big(\sum_{i=1}^{n+1} \sigma_i \int_{x_{i-1}}^{x_i} h > 1 \wedge \Phi_{10}(n) \ge \|h\|_\infty \Big).$$

Let $f \in C[0,1], n \in \mathbb{N}, p_1, p_2 \in K_{f,n}$ and $x_1 \le \ldots \le x_n \in [0,1]$ be fixed, and let $\tilde{h}$ be the function from (1) when $\sigma_i := f_0(\frac{x_{i-1}+x_i}{2})$, where $x_0 := 0$ and $x_{n+1} := 1$. Applying Lemma 1 to $\tilde{h}$ we get,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \le \ldots \le x_n \in [0,1]; \delta \in \mathbb{Q}^*_+$$
$$\big(\forall \lambda \in \mathbb{R}(\|f_0 - \lambda \tilde{h}\|_1 + \Phi_8(n, \delta, \tilde{h}) \ge \|f_0\|_1) \to$$
$$\exists y \in [0,1](|f_0(y)| < \delta \wedge \textstyle\bigwedge_{i=0}^{n+1} |x_i - y| > \Phi_9(n, \tilde{h}))\big).$$

We have $\|\tilde{h}\|_\infty \le 3^{n+1}(n+1)^{n+1}$. Let $\boxed{\tilde{\Phi_8}(n,\delta) := \frac{\delta}{3^{n+2}(n+1)^{n+1}}}$ and $\boxed{\tilde{\Phi_9}(n) := \frac{1}{2 \cdot 3^{n+2}(n+1)^{n+2}}}$. By the monotonicity of the functionals Φ_8 and Φ_9 in $\|h\|_\infty$ we have, $\tilde{\Phi_8}(n,\delta) \le \Phi_8(n, \delta, \tilde{h})$ and $\tilde{\Phi_9}(n) \le \Phi_9(n, \tilde{h})$. Then,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; x_1 \le \ldots \le x_n \in [0,1]; \delta \in \mathbb{Q}^*_+$$
$$\big(\forall \lambda \in \mathbb{R}(\|f_0 - \lambda \tilde{h}\|_1 + \tilde{\Phi}_8(n, \delta) \ge \|f_0\|_1) \to$$
$$\exists y \in [0,1](|f_0(y)| < \delta \wedge \textstyle\bigwedge_{i=0}^{n+1} |x_i - y| > \tilde{\Phi_9}(n))\big).$$

And we can conclude,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \delta \in \mathbb{Q}^*_+$$
$$\big(\forall h \in P_n(\|f_0 - h\|_1 + \tilde{\Phi}_8(n, \delta) \ge \|f_0\|_1) \to$$
$$\forall x_1 \le \ldots \le x_n \in [0,1] \exists y \in [0,1](|f_0(y)| < \delta \wedge \textstyle\bigwedge_{i=0}^{n+1} |x_i - y| > \tilde{\Phi_9}(n))\big).$$

We can actually replace the conclusion of the implication above with the actual existence of $n+1$ roots in the following way (Lemma $\forall \overline{x} J(\overline{x}) \to E$): Assume

$$(1)\ \forall x_1 \le \ldots \le x_n \in [0,1] \exists y \in [0,1](|f_0(y)| < \delta \wedge \textstyle\bigwedge_{i=0}^{n+1} |x_i - y| > \tilde{\Phi_9}(n))\big).$$

If $m < n+1$ is the biggest number of δ-roots of f_0 which are pairwise apart from each other by at least $\tilde{\Phi_9}(n)$ then by (1) we have a contradiction. Hence,

$$\exists x_0, \ldots, x_n \in [0,1](\textstyle\bigwedge_{i=0}^{n} |f_0(x_i)| < \delta \wedge \bigwedge_{i=1}^{n}(x_{i-1} + \tilde{\Phi_9}(n) < x_i)).$$

Therefore, we have,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in K_{f,n}; \delta \in \mathbb{Q}^*_+$$
$$\big(\forall h \in P_n(\|f_0 - h\|_1 + \tilde{\Phi}_8(n,\delta) \geq \|f_0\|_1) \rightarrow$$
$$\exists x_0, \ldots, x_n \in [0,1](\textstyle\bigwedge_{i=0}^n |f_0(x_i)| < \delta \wedge \bigwedge_{i=1}^n x_{i-1} + \tilde{\Phi}_9(n) < x_i)\big).$$

4 The uniform modulus of uniqueness for L_1-approximation

In this section we show how the computed functionals are combined in order to obtain the uniform modulus of uniqueness. Let $f \in C[0,1]$, $n \in \mathbb{N}, p_1, p_2 \in K_{f,n}$ and $\varepsilon \in \mathbb{Q}^*_+$ be fixed. Assume ($i \in \{1,2\}$),

$$(1)\ \|f - p_i\|_1 - dist_1(f, P_n) < \min\{\Phi_1(\Phi_7(\tilde{\Phi}_8(n, \Phi_5(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon))))),$$
$$\Phi_1(\Phi_2(\Phi_3(\omega_{f,n}, \Phi_4(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon)))))\}.$$

By section 3.3 we have, (where $f_0(x) = f(x) - \frac{p_1(x)+p_2(x)}{2}$)

$$\|f_0\|_1 - dist_1(f, P_n) < \Phi_2(\Phi_3(\omega_{f,n}, \Phi_4(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon)))).$$

By section 3.4 (and since Φ_1 is the identity),

$$|\ \|f_0\|_1 - 1/2\|f - p_1\|_1 - 1/2\|f - p_2\|_1\ | < \Phi_3(\omega_{f,n}, \Phi_4(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon))).$$

By section 3.6 (since $\int |f_0| - \frac{1}{2}|f - p_1| - \frac{1}{2}|f - p_2| = \|f_0\|_1 - \frac{1}{2}\|f - p_1\|_1 - \frac{1}{2}\|f - p_2\|_1$),

$$\|\ |f_0| - 1/2|f - p_1| - 1/2|f - p_2|\ \|_\infty \leq \Phi_4(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon)).$$

Hence, by section 3.7,

$$(2)\ \forall x \in [0,1](|f_0(x)| \leq \Phi_5(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon)) \rightarrow |p_1(x) - p_2(x)| \leq \Phi_6(n, \tilde{\Phi}_9(n), \varepsilon)).$$

By the same assumption (1) and section 3.3 we also have,

$$\|f_0\|_1 - dist_1(f, P_n) < \Phi_7(\tilde{\Phi}_8(n, \Phi_5(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon)))).$$

And by section 3.9,

$$\forall h \in P_n\big(\|f_0 - h\|_1 + \tilde{\Phi}_8(n, \Phi_5(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon))) \geq \|f_0\|_1\big).$$

Hence, by section 3.12 (taking $\delta = \Phi_5(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon))$),

$$\exists x_0, \ldots, x_n \in [0,1](\textstyle\bigwedge_{i=0}^n |f_0(x_i)| < \Phi_5(\Phi_6(n, \tilde{\Phi}_9(n), \varepsilon)) \wedge \bigwedge_{i=1}^n x_{i-1} + \tilde{\Phi}_9(n) \leq x_i).$$

And by (2),

$$\exists x_0,\dots,x_n\in[0,1](\bigwedge_{i=0}^{n}|p_1(x_i)-p_2(x_i)|\le\Phi_6(n,\tilde{\Phi}_9(n),\varepsilon)\wedge\bigwedge_{i=1}^{n}x_{i-1}+\tilde{\Phi}_9(n)\le x_i).$$

Therefore, by section 3.8 (taking $r=\tilde{\Phi}_9(n)$) we conclude,

$$(3)\ \|p_1-p_2\|_\infty\le\varepsilon.$$

If we substitute the linear functionals, $\Phi_1,\Phi_2,\Phi_4,\Phi_5$ and Φ_7, to make the conclusion more legible, we have (1) $\to$ (3),

$$\|f-p_i\|_1-dist_1(f,P_n)<\min\{\tilde{\Phi}_8(n,\tfrac{\Phi_6(n,\tilde{\Phi}_9(n),\varepsilon)}{8}),\Phi_3(\omega_{f,n},\tfrac{\Phi_6(n,\tilde{\Phi}_9(n),\varepsilon)}{8})\}\to\|p_1-p_2\|_\infty\le\varepsilon.$$

After applying $\tilde{\Phi}_8$ and $\tilde{\Phi}_9$ we get,

$$\|f-p_i\|_1-dist_1(f,P_n)<\min\{\frac{\Phi_6(n,\frac{1}{2.3^{n+2}(n+1)^{n+2}},\varepsilon)}{8\cdot3^{n+2}(n+1)^{n+1}},\Phi_3(\omega_{f,n},\frac{\Phi_6(n,\frac{1}{2\cdot3^{n+2}(n+1)^{n+2}},\varepsilon)}{8})\}\to$$
$$\|p_1-p_2\|_\infty\le\varepsilon.$$

Then we apply Φ_6,

$$\|f-p_i\|_1-dist_1(f,P_n)<\min\{\frac{\frac{\lfloor n/2\rfloor!\lceil n/2\rceil!\varepsilon}{2^{n+3}3^{n^2+2n}(n+1)^{n^2+2n+1}}}{3^{n+2}(n+1)^{n+1}},\Phi_3(\omega_{f,n},\frac{\lfloor n/2\rfloor!\lceil n/2\rceil!\varepsilon}{2^{n+3}3^{n^2+2n}(n+1)^{n^2+2n+1}})\}\to$$
$$\|p_1-p_2\|_\infty\le\varepsilon.$$

Let $c_n:=\frac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2^{n+3}3^{n^2+2n}(n+1)^{n^2+2n+1}}$ then we can rewrite the above formula as,

$$\|f-p_i\|_1-dist_1(f,P_n)<\min\{\tfrac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}},\Phi_3(\omega_{f,n},c_n\varepsilon)\}\to\|p_1-p_2\|_\infty\le\varepsilon.$$

And finally we apply the definition of Φ_3,

$$\|f-p_i\|_1-dist_1(f,P_n)<\min\{\tfrac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}},\tfrac{c_n\varepsilon}{2}\omega_{f,n}(\tfrac{c_n\varepsilon}{2})\}\to\|p_1-p_2\|_\infty\le\varepsilon.$$

Let $\tilde{\Phi}(f,n,\varepsilon)=\min\{\frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}},\frac{c_n\varepsilon}{2}\omega_{f,n}(\frac{c_n\varepsilon}{2})\}$, where $\omega_{f,n}:=\min\{\omega_f(\frac{\varepsilon}{4}),\frac{\varepsilon}{40(n+1)^4M_f}\}$ and $M_f\ge\|f\|_\infty$. We have shown that,

$$\forall f\in C[0,1];n\in\mathbb{N};p_1,p_2\in K_{f,n}$$
$$\big(\bigwedge_{i=1}^{2}(\|f-p_i\|_1-dist_1(f,P_n)<\tilde{\Phi}(f,n,\varepsilon))\to\|p_1-p_2\|_\infty\le\varepsilon\big).$$

Proposition 4.1 *The functional $\tilde{\Phi}(f,n,\varepsilon)$ is a uniform modulus of uniqueness for the best L_1-approximation of $C[0,1]$ from $K_{f,n}$, i.e.*

$$\forall f\in C[0,1];n\in\mathbb{N};p_1,p_2\in K_{f,n};\varepsilon\in\mathbb{Q}_+^*$$

$$\left(\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \tilde{\Phi}(f,n,\varepsilon)) \to \|p_1-p_2\|_1 \le \varepsilon\right).$$

Proof: Above, plus the fact that $\|p_1-p_2\|_1 \le \|p_1-p_2\|_\infty$.

$\square$

Claim 4.2 $\tilde{\Phi}(f,n,\varepsilon) \le \frac{\varepsilon}{8}$

Proof: Trivial.

$\square$

Now we show that Proposition 4.1 can be generalised to the whole space P_n (i.e. we can replace $K_{f,n}$ with P_n). Moreover, we notice that the dependency on particular values of the function f can be eliminated so that the modulus of uniqueness depends on f only through its modulus of continuity.

Theorem 4.3 *Let* $\Phi(\omega,n,\varepsilon) = \min\{\frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}}, \frac{c_n\varepsilon}{2}\omega_n(\frac{c_n\varepsilon}{2})\}$*, where the constant* $c_n := \frac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2^{n+3}3^{n^2+2n}(n+1)^{n^2+2n+1}}$, $\omega_n(\varepsilon) := \min\{\omega(\frac{\varepsilon}{4}), \frac{\varepsilon}{40(n+1)^4\lceil\frac{1}{\omega(1)}\rceil}\}$*. For all* $f \in C[0,1]$ *with modulus of continuity* ω

$$\forall n \in \mathbb{N}; p_1,p_2 \in P_n; \varepsilon \in \mathbb{Q}_+^*\left(\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \Phi(\omega,n,\varepsilon)) \to \|p_1-p_2\|_1 \le \varepsilon\right).$$

Proof: Actually, we prove the stronger version of the theorem where instead of $\lceil\frac{1}{\omega(1)}\rceil$ in the definition of ω_n we have any upper bound on $\sup_{x\in[0,1]}|f(x)-f(0)|$. First we show that in Proposition 4.1 we can replace $K_{f,n}$ with P_n. Suppose without loss of generality that $p_1 \in P_n\backslash K_{f,n}$. Then $\|p_1\|_1 > \frac{5}{2}\|f\|_1$ and hence $\|f-p_1\|_1 > \frac{3}{2}\|f\|_1 \ge \frac{3}{2}dist_1(f,P_n)$. Assume that $\|f-p_i\|_1 < dist_1(f,P_n) + \tilde{\Phi}(f,n,\varepsilon)$. By Claim 4.2, $\|f-p_i\|_1 < dist_1(f,P_n) + \frac{\varepsilon}{8}$. Then, $\frac{\varepsilon}{8} > \frac{1}{2}dist_1(f,P_n)$, i.e. $dist_1(f,P_n) < \frac{\varepsilon}{4}$. Therefore $\|f-p_i\|_1 < dist_1(f,P_n) + \frac{\varepsilon}{8} < \frac{\varepsilon}{2}$ and we have $\|p_1-p_2\|_1 \le \varepsilon$.
The second point is that some upper bound $M_f \ge \|f\|_\infty$ is used to define $\omega_{f,n}$ in Proposition 4.1. We claim that an upper bound $N_f \ge \sup_{x\in[0,1]}|f(x)-f(0)|$ is sufficient. For any function $f \in C[0,1]$ and polynomials $p_1,p_2 \in P_n$ let $\tilde{f}$, $\tilde{p_1}$ and $\tilde{p_2}$ be the functions obtained by the transposition of f, p_1 and p_2 respectively by $f(0)$ (i.e. $\tilde{f}(x) := f(x) - f(0)$ and $\tilde{p_i}(x) := p_i(x) - f(0)$). It is clear that (1) $\|f-p_i\|_1 = \|\tilde{f}-\tilde{p_i}\|_1$, (2) $dist_1(f,P_n) = dist_1(\tilde{f},P_n)$ and (3) $\|p_1-p_2\|_1 = \|\tilde{p_1}-\tilde{p_2}\|_1$. Let ω be the modulus of continuity for f and assume,

$$\|f-p_i\|_1 < dist(f,P_n) + \Phi(\omega,n,\varepsilon),$$

by (1) and (2) we have,

$$\|\tilde{f} - \tilde{p_i}\|_1 < dist(\tilde{f}, P_n) + \Phi(\omega, n, \varepsilon).$$

Since ω is also a modulus of continuity for $\tilde{f}$ and $\|\tilde{f}\|_\infty = \sup_{x\in[0,1]} |f(x) - f(0)| \le N_f$ we have $\tilde{\Phi}(\tilde{f}, n, \varepsilon) = \Phi(\omega, n, \varepsilon)$, therefore,

$$\|\tilde{f} - \tilde{p_i}\|_1 < dist(\tilde{f}, P_n) + \tilde{\Phi}(\tilde{f}, n, \varepsilon),$$

which implies, by Proposition 4.1, the first part of this proof and (3), $\|p_1 - p_2\|_1 \le \varepsilon$. Since $\lceil \frac{1}{\omega(1)} \rceil \ge \sup_{x\in[0,1]} |f(x) - f(0)|$ if ω is a modulus of uniform continuity for f the theorem follows.

□

A function $f \in C[0,1]$ is said to *be Lipschitz continuous with Lipschitz constant* $\lambda \in \mathbb{R}^*_+$ if $|f(x)-f(y)| \le \lambda|x-y|$ (i.e. $\frac{\varepsilon}{\lambda}$ is a modulus of continuity for f) and is *Lipschitz-α continuous with constant* λ, $0 < \alpha \le 1$, if $|f(x) - f(y)| \le \lambda|x-y|^\alpha$ (equivalently, $(\frac{\varepsilon}{\lambda})^{1/\alpha}$ is a modulus of continuity in our sense for f).[23] In this way, if a function f is Lipschitz continuous (or Lipschitz-α continuous) with constant λ then $\sup\limits_{x\in[0,1]} |f(x) - f(0)| \le \lambda$ (and we can take λ instead of $\lceil \frac{1}{\omega(1)} \rceil$ in Theorem 4.3). Hence, theorem 4.3 implies,

Corollary 4.4 *For any $f \in C[0,1]$,*

i) let $\Phi_L(\lambda, n, \varepsilon) = \min\{\frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}}, \frac{c_n^2\varepsilon^2}{160(n+1)^4\lambda}\}$. If f is Lipschitz continuous with constant λ then the functional Φ_L is a modulus of uniqueness for f.

ii) let $\Phi_{L_\alpha}(\lambda, \alpha, n, \varepsilon) = \min\{\frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}}, \frac{c_n\varepsilon}{2}(\frac{c_n\varepsilon}{8\lambda})^{1/\alpha}, \frac{c_n^2\varepsilon^2}{160(n+1)^4\lambda}\}$. If f is Lipschitz-α continuous with constant λ then the functional Φ_{L_α} is a modulus of uniqueness for f.

And as a corollary of proposition 5.4 from [15] and theorem 4.3 above we get,

Theorem 4.5 *Let $\mathcal{P}(f,n)$ denote the operator which assigns to any given function $f \in C[0,1]$ and any $n \in \mathbb{N}$ the best L_1-approximation of $f \in C[0,1]$ from P_n. Then $\Phi_P(\omega_f, n, \varepsilon) := \frac{\Phi(\omega_f, n, \varepsilon)}{2}$, Φ as defined in Theorem 4.3, is a modulus of pointwise continuity for the operator $\mathcal{P}(f,n)$, i.e.,*

$$\forall f, \tilde{f} \in C[0,1]; n \in \mathbb{N}; \varepsilon \in \mathbb{Q}^*_+ (\|f - \tilde{f}\|_1 < \Phi_P(\omega_f, n, \varepsilon) \to \|\mathcal{P}(f,n) - \mathcal{P}(\tilde{f}, n)\|_1 \le \varepsilon).$$

[23] In analysis the condition '$|f(x) - f(y)| \le \lambda|x-y|^\alpha$ for some λ' is called *Hölder condition with exponent* α.

Proof: For completeness we reproduce here the proof as given in [15]. One easily verifies that $dist_1(f, P_n)$ is Lipschitz continuous in f (with respect to the L_1-norm) with $\lambda = 1$, i.e.

$$(*)\ \|f - \tilde{f}\|_1 < \varepsilon \to |dist_1(f, P_n) - dist_1(\tilde{f}, P_n)| < \varepsilon.$$

Assume now that $\|f - \tilde{f}\|_1 < \Phi_P(\omega_f, n, \varepsilon) = \frac{1}{2}\Phi(\omega_f, n, \varepsilon)$. Then,

$$\begin{aligned}
\|f - \mathcal{P}(\tilde{f}, n)\|_1 &\leq \|\tilde{f} - \mathcal{P}(\tilde{f}, n)\|_1 + \|f - \tilde{f}\|_1 = dist_1(\tilde{f}, P_n) + \|f - \tilde{f}\|_1 \\
&\overset{(*)}{<} dist_1(f, P_n) + \frac{1}{2}\Phi(\omega_f, n, \varepsilon) + \|f - \tilde{f}\|_1 \\
&< dist_1(f, P_n) + \Phi(\omega_f, n, \varepsilon).
\end{aligned}$$

Since, furthermore, $\|f - \mathcal{P}(f, n)\|_1 = dist_1(f, P_n)$, we obtain from theorem 4.3 that $\|\mathcal{P}(f, n) - \mathcal{P}(\tilde{f}, n)\|_1 \leq \varepsilon$.

□

4.1 Related results

The first proof of the uniqueness of the best L_1-approximation was given in 1921 by Jackson, [9]. The proof we analysed was published by Cheney [6] in 1965 and reprinted in his book [7] from 1966. Only in 1975 Björnestal [3], by analyzing the qualitative (relative to the dependencies) aspect of the continuity of the projection operator for arbitrary normed linear spaces X into a closed linear subspace of X, obtained the following result,

Theorem 4.6 (Björnestal, 75) *Let $f \in C[0,1]$ and Ω_f be a modulus of uniform continuity of $f - p_b$, where p_b is the best L_1-approximation of f from P_n. Then, for $p \in P_n$, ε sufficiently small and for some constant c depending on f and n,*

$$\|p - p_b\|_1 \geq \varepsilon \to \|f - p\|_1 - \|f - p_b\|_1 \geq 2\int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon - \Omega_f(x)\, dx.$$

We want to show that our theorem 4.3 implies an effective version of Björnestal's theorem. First we can rewrite his theorem in the form we have been working with,

$$(1)\ \|f - p\|_1 < dist_1(f, P_n) + 2\int_0^{\Omega_f^{-1}(c\varepsilon)} c\varepsilon - \Omega_f(x)\, dx \to \|p - p_b\|_1 < \varepsilon.$$

The modulus of uniform continuity referred by Björnestal is defined as,

$$\Omega_f(\varepsilon) = \sup_{|x-y|<\varepsilon} |f(x) - p_b(x) - f(y) + p_b(y)|,$$

which means that $\Omega_f^{-1}(\varepsilon)$ (for ε small enough so that $\Omega_f^{-1}(\varepsilon)$ is defined) is the optimal modulus of continuity for $f - p_b$ in our sense.

First we show that $\int_0^{\Omega_f^{-1}(c\,\varepsilon)} c\,\varepsilon - \Omega_f(x)\,dx$ can be written as $c'\,\varepsilon\Omega_f^{-1}(c'\,\varepsilon)$, for some constant $\frac{c}{2} \leq c' \leq c$. For that purpose note that,

$$\int_0^{\Omega_f^{-1}(c\,\varepsilon)} c\,\varepsilon - \Omega_f(x)\,dx \leq \int_0^{\Omega_f^{-1}(c\,\varepsilon)} c\,\varepsilon\,dx = c\,\varepsilon\Omega_f^{-1}(c\,\varepsilon).$$

On the other hand we have,

$$\int_0^{\Omega_f^{-1}(c\,\varepsilon)} c\,\varepsilon - \Omega_f(x)\,dx \geq \int_0^{\Omega_f^{-1}(\frac{c}{2}\,\varepsilon)} c\,\varepsilon - \Omega_f(x)\,dx \geq \int_0^{\Omega_f^{-1}(\frac{c}{2}\,\varepsilon)} \frac{c}{2}\,\varepsilon\,dx = \frac{c}{2}\,\varepsilon\Omega_f^{-1}(\frac{c}{2}\,\varepsilon).$$

Therefore, for some $\frac{c}{2} \leq c' \leq c$, (1) is equivalent to,

$$\|f - p\|_1 < dist_1(f, P_n) + 2c'\,\varepsilon\Omega_f^{-1}(c'\,\varepsilon) \to \|p - p_b\|_1 < \varepsilon.$$

The constant c, however, is not presented by Björnestal and moreover the function Ω_f^{-1} is normally non-computable. We can give an effective modulus of continuity for $f - p_b$ following section 3.5 (and taking $M_f = \lceil\frac{1}{\omega_f(1)}\rceil$ as suggested in the proof of 4.3),

$$\begin{aligned} \omega_{f-p_b}(\varepsilon) &\geq \min\{\omega_f(\frac{\varepsilon}{2}), \omega_{p_b}(\frac{\varepsilon}{2})\} \\ &\geq \min\{\omega_f(\frac{\varepsilon}{2}), \frac{\varepsilon}{20(n+1)^4\lceil\frac{1}{\omega_f(1)}\rceil}\} \end{aligned}$$

Therefore, let $\omega^*_{f-p_b}(\varepsilon) := \min\{\omega_f(\frac{\varepsilon}{2}), \frac{\varepsilon}{20(n+1)^4\lceil\frac{1}{\omega_f(1)}\rceil}\}$, we can restate our theorem 4.3 and see how it relates to Björnestal's result:

Corollary 4.7 *Let $f \in C[0,1]$, ω_f be some modulus of uniform continuity of f, and $p \in P_n$. Then for $\varepsilon \leq 1$,*

$$\|f - p\|_1 < dist_1(f, P_n) + \tilde{c}_n\,\varepsilon\,\omega^*_{f-p_b}(\tilde{c}_n\,\varepsilon) \to \|p - p_b\|_1 \leq \varepsilon,$$

where $\tilde{c}_n := \frac{c_n}{3^{n+2}(n+1)^{n+1}}$ and $c_n := \frac{\lfloor n/2\rfloor!\lceil n/2\rceil!}{2^{n+3}3^{n^2+2n}(n+1)^{n^2+2n+1}}$.

Proof: From theorem 4.3 we have,

$$\|f - p\|_1 < dist_1(f, P_n) + \min\{\frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}}, \frac{c_n\varepsilon}{2}\omega^*_{f-p_b}(\frac{c_n\varepsilon}{4})\} \to \|p - p_b\|_1 \leq \varepsilon,$$

which implies,

$$\|f-p\|_1 < dist_1(f,P_n)+\min\{\frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}}, \frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}}\omega^*_{f-p_b}(\frac{c_n\varepsilon}{4})\} \to \|p-p_b\|_1 \le \varepsilon.$$

For $\varepsilon \le 1$ we have $\omega^*_{f-p_b}(\frac{c_n\varepsilon}{4}) \le 1$. Hence,

$$\|f-p\|_1 < dist_1(f,P_n) + \frac{c_n\varepsilon}{3^{n+2}(n+1)^{n+1}}\omega^*_{f-p_b}(\frac{c_n\varepsilon}{4}) \to \|p-p_b\|_1 \le \varepsilon.$$

Since $3^{n+2}(n+1)^{n+1} > 4$ we get our result.

□

Some years later, in 1978, Kroo [20] showed that the constant c in Björnestal's result needed not to depend on any particular point of the function f but only on its modulus of continuity. [24] We got an effective version of Björnestal's result where our constant c is completely independent of the function f and only depends on the dimension of the space P_n.

Remark 4.8 *In Kroo [20] the problem of L_1-approximation of continuous functions is considered for arbitrary Haar subspaces of $C[0,1]$ containing the constant functions. Kroo [22] treats uniqueness subspaces of $C[0,1]$ but in that case the constant c also depends on values of the function f and not only on its modulus of continuity. Since Cheney's proof which we analyzed works for arbitrary Haar subspaces we are also guaranteed to extract uniform moduli of uniqueness in the general setting. As done by Jackson [9] in his original proof, in the present work we focused on the specific Haar subspace P_n in order to get fully explicit results. One can observe that only section 3.8 (Lagrange interpolation formula used to show that P_n is a Haar space), section 3.5 (Markov inequality used to show that $K_{f,n}$ is compact by constructing a common modulus of uniform continuity) and section 3.11 (the construction of a polynomial which changes sign in each x_i) made reference to the particular Haar space P_n. From results in [4](lemma 4.3), [5](lemma) and [13](after lemma 9.32) is follows that there exist effective and quantitative substitutes for each of these constructions for arbitrary (effectively given) Haar spaces. So it is clear that the analysis carried out in this paper can be extended to general Haar spaces H containing the constant functions[25]. Details will be presented in a subsequent paper.*

4.2 How p_b is computed

An operator $B_{f,n}(\varepsilon) : \mathbb{Q}^*_+ \to P_n$ computes the unique best L_1-approximation, $p_b \in P_n$, of a function $f \in C[0,1]$ from P_n if for any given $\varepsilon \in \mathbb{Q}^*_+$ it generates a polynomial of degree

[24] As in Björnestal [3], Kroo does not present the actual constant.

[25] We only need the constant functions to belong to H if we want to get rid of the f dependency in c, i.e. obtain a constant c in the uniform modulus of uniqueness depending only on n and ω_f.

$\leq n$ with rational coefficients (i.e. a $n+1$-vector of rational coefficients) $B_{f,n}(\varepsilon)$ such that, $\|B_{f,n}(\varepsilon)-p_b\|_1 \leq \varepsilon$. We show how this can be achieved using the uniform modulus of uniqueness, $\Phi(f,n,\varepsilon)$,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p_1, p_2 \in P_n; \varepsilon \in \mathbb{Q}_+^* \\ \big(\bigwedge_{i=1}^{2}(\|f-p_i\|_1 - dist_1(f,P_n) < \Phi(f,n,\varepsilon)) \rightarrow \|p_1-p_2\|_1 \leq \varepsilon\big).$$

First we substitute p for p_1 and p_b for p_2,

$$\forall f \in C[0,1]; n \in \mathbb{N}; p \in P_n; \varepsilon \in \mathbb{Q}_+^* \big(\|f-p\|_1 - dist_1(f,P_n) < \Phi(f,n,\varepsilon) \rightarrow \|p-p_b\|_1 \leq \varepsilon\big).$$

Now we just need to find a $B_{f,n}(\varepsilon)$ such that, $\|f-B_{f,n}(\varepsilon)\|_1 - dist_1(f,P_n) < \Phi(f,n,\varepsilon)$. Note that now there is no explicit reference to p_b, only implicit in $dist_1(f,P_n)$.

A set $N := \{p_1, p_2, \ldots\} \subset P_n$ is said to be an *ε-net of P_n* if $\forall p \in P_n \exists p_i \in N(\|p-p_i\|_1 \leq \varepsilon)$. The algorithm for computing p_b consists in evaluating $\|f-p_i\|_1$ for each p_i in some $\Phi(f,n,\varepsilon)$-net of P_n and taking the p_i which gives the minimum value. This *ε-net* has a priori infinitely many elements, but since we know that the best L_1-approximation p_b of f lives in a compact subspace of P_n, namely $p_b \in K_{f,n}$, we have $\|p_b\|_1 \leq 5/2\|f\|_1$ which gives us a bound on the coefficients of p_b (see remark 3.7). In this way, we can restrict the search to the *ε-net of $K_{f,n}$*, chosen according to these bounds on the coefficients, which is finite. The complexity analysis of the whole algorithm will be carried out in a future work.

5 Concluding remarks on the extraction of Φ

We emphasize again the two important roles played by logic in the extraction of the modulus of uniqueness for best L_1-approximation presented here. First, by showing that Cheney's proof could be formalized in the system $\mathcal{A}_*^\omega$ (and by the logical meta-theorem 2.1) we were guaranteed that such a modulus Φ would exist and that it could be extracted from the mentioned proof. Moreover, the fact that Φ depends only on ω_f, n and ε (which was proved by Kroo years after Cheney's proof) is obtained immediately from the meta-theorem 2.1. The second important role is that logic not only guaranteed the existence of the modulus but it went even further and supplied a procedure (monotone functional interpretation) to extract the modulus, which enabled us to provide for the first time an explicit dependency on n and ω_f. And, as it happened, the extracted modulus of uniqueness has the optimal ε-dependency established by Kroo.

We hope it is transparent that all the mathematical tools used in our analysis were already present in Cheney's proof,[26] which can be noticed for instance in the analysis of Lemma 1

[26]Except Markov's inequality which was used to show that the set $K_{f,n}$ is compact and Lagrange interpolation formula used to prove that P_n is a Haar space. These tools, however, are standard in approximation theory.

(section 3.10) where in order to prove that the functionals presented realized the lemma (see Claim 3.13) we followed line by line the original proof from [7], the only difference being that we considered the ε-version of the propositions. This visibly shows that the uniform modulus of uniqueness here extracted was really implicitly present in Cheney's proof but could only be made explicit with the help of logic. The difficulty to extract ad hoc such information can be understood because Cheney's proof (although very simple from the mathematical point of view and even called 'elementary' by the author) is logically very intricate due to the use of proof by contradiction and principles that fail in computable analysis.

6 Remarks on the case of Chebycheff approximation

In this section we recall some of the results obtained in [16] on effective moduli of uniqueness for the best Chebycheff approximation and show how they can be used to get moduli of uniqueness which – like our modulus for the L_1-approximation in theorem 4.3 – only depend on $f \in C[0,1]$ via a modulus of uniform continuity ω for f. [16] extracts from a classical proof (due to Young [28]) for the uniqueness of the best Chebycheff approximation uniform constants of strong unicity (see [23]) and moduli of uniqueness both for the polynomial case P_n and for general Haar spaces H. Whereas the constants of strong unicity depend on a positive lower bound for $dist(f, P_n)$ (resp. $dist(f, H)$), this dependency can be removed (as shown in [16]) on the expense of losing the linearity of the modulus of uniqueness. In fact, in [16] non-linear moduli of uniqueness (and moduli of pointwise continuity for the Chebycheff projection) which only depend on the modulus of uniform continuity[27] of f, the Haar space, ε and an upper bound for $\|f\|_\infty$ are obtained.

The following constructions are due to [4],[5]:

Definition 6.1 *Let* $\phi := \{\phi_1, \ldots, \phi_n\}$ *be a Chebycheff system over* $[0,1]$, $\underline{\phi}(x) := \big(\phi_1(x), \ldots, \phi_n(x)\big) \in \mathbb{R}^n$, $\|\underline{\phi}\| := \sup\limits_{x\in[0,1]} \|\underline{\phi}(x)\|_2$, *where* $\|\cdot\|_2$ *denotes the Euclidean norm on* $\mathbb{R}^n$.

1) $\beta, \gamma, \kappa : (0, \frac{1}{n}] \to \mathbb{R}^*_+$ *are defined by*

$$\beta(\alpha) := \begin{cases} \inf\limits_{x\in[0,1]} |\phi_1(x)|, & \text{if } n = 1 \\ \inf\left\{ |\det(\phi_j(x_i))| : 0 \le x_1, \ldots, x_n \le 1, \bigwedge\limits_{i=1}^{n-1} (x_{i+1} - x_i \ge \alpha) \right\}, & \text{if } n > 1 \end{cases}$$

[27] Here we refer again to the notion as defined in the introduction and not to the concept of modulus of continuity as used in numerical analysis.

and

$$\gamma(\alpha) := \min\left(\|\underline{\phi}\|, \frac{\beta(\alpha)}{n^{\frac{1}{2}}(n-1)!\prod\limits_{i=1}^{n}(1+\|\phi_i\|_\infty)}\right), \ \kappa(\alpha) := \gamma(\alpha)^{-1}\cdot\|\underline{\phi}\|$$

for $\alpha \in (0, \frac{1}{n}]$. *Since* ϕ *is a Chebycheff system it follows that* $\beta(\alpha) > 0$.
$H :=$ *Lin*$_{\mathbb{R}}(\phi_1, \ldots, \phi_n)$.

2) *Let* $M \in \mathbb{R}^*_+, \omega : \mathbb{R}^*_+ \to \mathbb{R}^*_+$ *and let* $\omega_{\underline{\phi}}$ *be a modulus of uniform continuity for* $\underline{\phi}$. *Then we define for* $\varepsilon \in \mathbb{R}^*_+$

$$\omega_{M,H}(\varepsilon) := \min\left(\omega(\frac{\varepsilon}{2}), \omega_{\underline{\phi}}\left(\frac{\varepsilon\cdot\beta(\frac{1}{n})}{4Mn^{\frac{3}{2}}(n-1)!\prod\limits_{i=1}^{n}(1+\|\phi_i\|_\infty)}\right)\right).$$

Theorem 6.2 ([16]) *Let* $\phi, H, \beta, \gamma, \kappa$ *bas as in definition 6.1 and let* $A \subset C[0,1]$ *be a set of functions which have* $\omega : \mathbb{R}^*_+ \to \mathbb{R}^*_+$ *as a common modulus of uniform continuity and* $M \in \mathbb{R}^*_+$ *as a common upper bound* $M \geq \|f\|_\infty$ *for all* $f \in A$.

1) *Let* $l_{H,A} \in \mathbb{R}^*_+$ *such that* $l_{H,A} < \inf\limits_{f\in A} dist(f,H)$ *and* $0 < \alpha \leq \min\left(\frac{1}{n}, \omega_{M,H}(2\cdot l_{H,A})\right)$.
Then $\frac{\gamma(\alpha)}{n\cdot\|\underline{\phi}\|}$ *as a common constant of strong unicity for all* $f \in A$, *i.e.*

$$\forall f \in A, \varphi \in H\Big(\|f-\varphi\|_\infty \geq \|f-\varphi_b\|_\infty + \left(\frac{\gamma(\alpha)}{n\cdot\|\underline{\phi}\|}\right)\cdot\|\varphi-\varphi_b\|_\infty\Big),$$

where φ_b *is the best approximation of* f *in* H. *Furthermore* $2n\cdot\kappa(\alpha)$ *is a common Lipschitz constant for the Chebycheff projection for all points* $f \in A$.

2)

$$\Phi_{M,H}(\omega,\varepsilon) := \min\left(\frac{\varepsilon}{4}, \frac{\varepsilon}{2n\kappa\big(\min\left(\frac{1}{n}, \omega_{M,H}(\frac{\varepsilon}{2})\right)\big)}\right)$$

is a common modulus of uniqueness (and a common modulus of continuity for the Chebycheff projection in f*) for all* $f \in A$.

Remark 6.3 The bounds in the theorem improve numerically bounds obtained prior by D. Bridges ([4],[5], see [16] for a detailed discussion).

Using the transformation $f \mapsto \tilde{f}$, with $\tilde{f}(x) := f(x) - f(0)$, an argument similar to the one used in the proof of theorem 4.3 allows one to conclude:

Theorem 6.4 *Let* $\{\phi_1, \ldots, \phi_n\}$ *be a Chebycheff system s.t.* $1 \in H := Lin_{\mathbb{R}}(\phi_1, \ldots, \phi_n)$ *and let* $\omega : \mathbb{R}_+^* \to \mathbb{R}_+^*$ *be any function. Then*

$$\Phi_H(\omega, \varepsilon) := \min\left(\frac{\varepsilon}{4}, \frac{\varepsilon}{2n\kappa\big(\min\left(\frac{1}{n}, \omega^H(\frac{\varepsilon}{2})\right)\big)}\right)$$

with

$$\omega^H(\varepsilon) := \min\left(\omega(\frac{\varepsilon}{2}), \omega_{\underline{\phi}}\left(\frac{\varepsilon \cdot \beta(\frac{1}{n})}{4\lceil\frac{1}{\omega(1)}\rceil n^{\frac{3}{2}}(n-1)! \prod\limits_{i=1}^{n}(1 + \|\phi_i\|_\infty)}\right)\right)$$

is a common modulus of uniqueness (and a common modulus of continuity for the Chebycheff projection) for all functions $f \in C[0,1]$ *which have* ω *as a modulus of uniform continuity.*

For the polynomial case we get (see [16](cor.3.8)*)*

$$\Phi(\omega, n, \varepsilon) := \min\Big(\varepsilon/4, \frac{\lfloor\frac{n}{2}\rfloor!\,\lceil\frac{n}{2}\rceil!}{2(n+1)} \cdot (\omega_n(\varepsilon/2))^n \cdot (\varepsilon/4)\Big),$$

with

$$\omega_n(\varepsilon) := \begin{cases} \min\left(\omega\left(\frac{\varepsilon}{2}\right), \frac{\varepsilon}{10n^2\lceil\frac{1}{\omega(1)}\rceil}\right), & \text{if } n \geq 1 \\ 1 & \text{if } n = 0 \end{cases}$$

as such a common modulus.

As a corollary we obtain that for arbitrary Haar spaces having the constant function 1 the continuity behaviour of the Chebycheff projection is uniform for any class of equicontinuous functions which generalizes a result of [21] for the case of (trigonometric) polynomials.

References

[1] J. Avigad and S. Feferman. Gödel's functional ("Dialectica") interpretation. In S.R. Buss, editor, *Handbook of proof theory*, volume 137, pages 337–405. Elsevier, North-Holland, Amsterdam, 1998.

[2] M. Bartelt and W. Li. Error estimates and Lipschitz constants for best approximation in continuous function spaces. *Computers and Mathematics with Application*, 30(3-6):255–268, 1995.

[3] B.O. Björnestal. Continuity of the metric projection operator i-iii. *The preprint series of Department of Mathematics. Royal Institute of Technology. Stockholm, TRITA-MAT*, 17, 1975.

[4] D.S. Bridges. A constructive development of Chebychev approximation theory. *Journal of Approximation Theory*, 30:99–120, 1980.

[5] D.S. Bridges. Lipschitz constants and moduli of continuity for the Chebychev projection. *Proc. Amer. Math. Soc.*, 85:557–561, 1982.

[6] E.W. Cheney. An elementary proof of Jackson's theorem on mean-approximation. *Mathematics Magazine*, 38:189–191, 1965.

[7] E.W. Cheney. *Approximation Theory.* AMS Chelsea Publishing, 1966.

[8] S. Feferman. Theories of finite type related to mathematical practice. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 913–972. North-Holland, Amsterdam, 1977.

[9] D. Jackson. Note on a class of polynomials of approximation. *Transactions of the American Mathematical Society*, 22:320–326, 1921.

[10] K.-I. Ko. On the computational complexity of best Chebycheff approximation. *Journal of Complexity*, 2:95–120, 1986.

[11] K.-I. Ko. *Complexity theory of real functions.* Birkhäuser, Boston-Basel-Berlin, 1991.

[12] U. Kohlenbach. A quantitative version of a theorem due to Borwein-Reich-Shafrir. To appear in: Numerical Functional Analysis and Optimization, 15 pages.

[13] U. Kohlenbach. *Theory of majorizable and continuous functionals and their use for the extraction of bounds from non-constructive proofs: effective moduli of uniqueness for best approximations from ineffective proofs of uniqueness (german).* PhD thesis, Frankfurt, pp. xxii+278, 1990.

[14] U. Kohlenbach. Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. *Journal of Symbolic Logic*, 57:1239–1273, 1992.

[15] U. Kohlenbach. Effective moduli from ineffective uniqueness proofs. an unwinding of de La Vallée Poussin's proof for Chebycheff approximation. *Annals of Pure and Applied Logic*, 64:27–94, 1993.

[16] U. Kohlenbach. New effective moduli of uniqueness and uniform a–priori estimates for constants of strong unicity by logical analysis of known proofs in best approximation theory. *Numerical Functional Analysis and Optimization*, 14:581–606, 1993.

[17] U. Kohlenbach. Analysing proofs in Analysis. In W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, editors, *Logic: from Foundations to Applications*, pages 225–260. European Logic Colloquium (Keele, 1993), Oxford University Press, 1996.

[18] U. Kohlenbach. Higher order reverse mathematics. Preprint, 14 pages, 2000.

[19] U. Kohlenbach. On the computational content of the Krasnoselski and Ishikawa fixed point theorems. In J. Blanck, V. Brattka, P. Hertling, and K. Weihrauch, editors, *Computability and Complexity in Analysis, CCA 2000.* Springer LNCS 2064, 2001.

[20] A. Kroo. On the continuity of best approximations in the space of integrable functions. *Acta Mathematica Academiae Scientiarum Hungaricae*, 32:331–348, 1978.

[21] A. Kroo. On the uniform modulus of continuity of the operator of best approximation in the space of periodic functions. *Acta Math. Academiae Scientiarum Hungaricae*, 34:185–203, 1979.

[22] A. Kroo. On strong unicity of L_1-approximation. *Proceedings of the American Mathematical Society*, 83(4), 1981.

[23] D. J. Newman and H. S Shapiro. Some theorems on Chebycheff approximation. *Duke Mathematical Journal*, 30:673–681, 1963.

[24] A. Pinkus. *On L_1-Approximation*, volume 93 of *Cambridge Trats in Mathematics.* Cambridge University Press, 1989.

[25] S.G. Simpson. *Subsystems of Second Order Arithmetic.* Perspectives in Mathematical Logic. Springer-Verlag, 1999.

[26] A.S. Troelstra. *Metamathematical investigation of intuitionistic Arithmetic and Analysis, Springer Lecture Notes in Mathematics*, volume 344, pages 1–323. Springer Verlag, Berlin, 1973.

[27] K. Weihrauch. *Computable Analysis.* Springer-Verlag, 2000.

[28] J.W. Young. General theory of approximnation by functions involving a given number of arbitrary parameters. *Trans. Amer. Math. Soc.*, 8:331–344, 1907.