Best Approximation by Free Knot Splines

Günter Meinardus and **Guido Walz**

Nr. 234/1998

Professor Dr. Günter Meinardus
PD Dr. Guido Walz
Department of Mathematics and Computer Science
University of Mannheim
D-68131 Mannheim, Germany
walz@math.uni-mannheim.de
http://www.math.uni-mannheim.de/~math4/walz.htm

Proposed Running Head. Approximation by Free Knot Splines

Abstract. We consider the problem of finding the best (uniform) approximation of a given continuous function by spline functions with free knots.
Our approach can be sketched as follows: By using the Gauß transform with arbitrary positive real parameter t, we map the set of splines under consideration onto a function space, which is arbitrarily close to the spline set, but satisfies the local Haar condition and also possesses other nice structural properties. This enables us to give necessary and sufficient conditions for best approximations (in terms of alternants) and, under some assumptions, even full characterizations and a uniqueness result. By letting $t \to 0$, we recover a best approximation in the original spline space. Our results are illustrated by some numerical examples, which show in particular the nice alternation behavior of the error function.

Keywords. Free Knot Splines, Gauß transform, local Haar condition, nonlinear approximation, alternant theorems

AMS (MOS) Classification. 41A15, 41A30

Correspondence to:
Guido Walz
Department of Mathematics and Computer Science
University of Mannheim
D-68131 Mannheim, Germany
Fax: (Germany) 621-292-1064
walz@math.uni-mannheim.de

1. Introduction

In approximating continuous functions by polynomial spline functions of order m one may expect to gain better accuracies and a faster convergence to zero of the error function by using free knots which are permitted to possess variable multiplicities, fixing the total number of mulitiplicities by n. Of course, one has to deal then with non–linear approximation problems.

In the literature, there exist several direct approaches to this problem, see M. Adam [1] or G. Nürnberger [18] for a survey. In this paper, we propose a new approach, which is based on the fact that each spline set can be mapped bijectively on a set of functions with a Haar–like structure, by using the so–called Gauß transform. In some previous publications (Meinardus [13,14], Walz [22]) this approach was applied to best approximation with fixed knot splines resp. to Schoenberg's problem for a single B–spline. In the present paper, we use it to obtain necessary and sufficient conditions and – under some mild assumptions – characterizations of best uniform approximations. We also give a criterion for detecting those functions which possess a best approximation with single knots only, a situation which is easier to handle. Moreover, since the Gauß transform can be easily evaluated, we also present the results of some numerical tests.

The *main idea* of our approach is as follows: We use the Gauß transform with parameter $t > 0$ cf. [11], to map a given set of free knot splines onto a non–linear function set, which can be shown to satisfy the local Haar condition and also possesses some other nice structural properties. Therefore, we can use general results of nonlinear Approximation Theory to obtain necessary and sufficient conditions for best approximations, which in many cases also coincide and therefore yield a full characterization of the best approximation. Since, for $t \to 0$, the Gauß transform of a spline function converges uniformly to this spline, we obtain, by extrapolating the best approximations in the Gauß transformed space back to $t = 0$, spline approximations with free knots, which are very close to the best ones.

The organization of the paper is as follows: In the rest of this introduction we give an exact definition of spline sets with free knots, and determine the tangential space of this set. In section 2 we investigate the Gauß transform, when applied to a set of free knot splines. In particular we prove that it satisfies the local Haar condition, and give a bound for the number of zeros that the difference of two Gauß transforms can have. These results are used in section 3 to obtain the desired necessary and sufficient criteria for best approximation, and also a de la Vallée Poussin type result. A very attractive situation occurs in the case of splines with single knots, considered in section 4. Here it is possible to give a full characterization of the best approximation (in terms of alternants), and to prove their uniqueness. Moreover, we give a criterion which shows that a large class of functions possesses a best approximation with single knots. Finally, in section 5 we discuss the results of some numerical tests.

Let now $m \in I\!N$, $m \geq 2$, and $n \in I\!N$ be given natural numbers. For given $r \in I\!N, r \leq n$, we denote by

$$S_{m,n,r}(\xi,\lambda) = S_{m,n,r}\begin{pmatrix} \xi_1 & \xi_2 & \dots & \xi_r \\ \lambda_1 & \lambda_2 & \dots & \lambda_r \end{pmatrix}$$

the vector space of all polynomial spline functions $s : I\!R \to I\!R$ of order m with knots ξ_ν, $\nu = 1, 2, \dots, r$, satisfying

$$\xi_1 < \xi_2 < \dots < \xi_r \tag{1}$$

and having the respective multiplicities $\lambda_\nu \in I\!N$, $\nu = 1, 2, \dots, r$, with

$$\sum_{\nu=1}^{r} \lambda_\nu = n. \tag{2}$$

It is well known and easy to prove that

$$\dim S_{m,n,r}(\xi,\lambda) = m + n. \tag{3}$$

Furthermore the truncated power basis of $S_{m,n,r}(\xi,\lambda)$ is given by the functions

$$1, x, \dots, x^{m-1}, (x-\xi_\nu)_+^{m-1}, (x-\xi_\nu)_+^{m-2}, \dots, (x-\xi_\nu)_+^{m-\lambda_\nu}, \quad \nu = 1, 2, \dots, r$$

i.e. this vector space consists of all real functions s on $I\!R$, which have a real representation

$$s(x) = \sum_{\mu=0}^{m-1} a_\mu x^\mu + \sum_{\nu=1}^{r} \sum_{\kappa=1}^{\lambda_\nu} b_{\nu,\kappa} (x-\xi_\nu)_+^{m-\kappa} \tag{4}$$

Here, the truncated power function $(x)_+^k$, $k \in I\!N_0$, is defined by

$$(x)_+^k = \begin{cases} 0 & \text{for } x \leq 0, \\ x^k & \text{for } x > 0, \end{cases}$$

if $k \in I\!N$, and by

$$(x)_+^0 = \begin{cases} 0 & \text{for } x < 0, \\ \frac{1}{2} & \text{for } x = 0, \\ 1 & \text{for } x > 0. \end{cases}$$

In some instances we will later restrict the vector spaces $S_{m,n,r}(\xi,\lambda)$ to a real interval $I = [a,b]$. In those cases we always assume that the knots belong to the interior of I.

The union over all $\xi_1, \xi_2, \dots, \xi_r$, satisfying (1), with fixed multiplicities $\lambda_1, \lambda_2, \dots, \lambda_r$ according to (2) shall be denoted by

$$M_{m,n,r}(\lambda) = \bigcup_{\xi} S_{m,n,r}(\xi,\lambda). \tag{5}$$

The highest flexibility occurs if we consider the set

$$S_{m,n} = \bigcup_{r=1}^{n} \bigcup_{\lambda_1+\ldots+\lambda_r=n}^{\lambda} M_{m,n,r}(\lambda), \tag{6}$$

which usually is called the polynomial spline set with free knots of order m and of total multiplicity n.

We will develop a theory of best uniform approximation with respect to the set $M_{m,n,r}(\lambda)$ and thus to $S_{m,n}$. Let us first consider the tangential space of $M_{m,n,r}(\lambda)$ at a point $s \in M_{m,n,r}(\lambda)$, for fixed r and fixed multiplicities $\lambda_1, \lambda_2, \ldots, \lambda_r$. This tangential space $T_{m,n,r}(\lambda; s)$ is defined as spanned by the partial derivatives of s with respect to the $m+n+r$ parameters

$$\begin{aligned} a_\mu &;\ \mu = 0, 1, \ldots, m-1, \\ b_{\nu,\kappa} &;\ \kappa = 1, 2, \ldots, \lambda_\nu;\ \ \nu = 1, 2, \ldots, r. \\ \xi_\nu &;\ \nu = 1, 2, \ldots, r. \end{aligned}$$

Hence we have to assume differentiability w.r.t. these parameters, which means that all the multiplicities satisfy

$$\lambda_\nu \leq m-2; \quad \nu = 1, 2, \ldots, r. \tag{7}$$

Under this assumption we get

$$\begin{aligned} \frac{\partial s}{\partial a_\mu} &= x^\mu \ ;\ \ \mu = 0, 1, \ldots, m-1, \\ \frac{\partial s}{\partial b_{\nu,\kappa}} &= (x-\xi_\nu)_+^{m-\kappa} \ ;\ \ \kappa = 1, 2, \ldots, \lambda_\nu;\ \ \nu = 1, 2, \ldots, r, \\ \frac{\partial s}{\partial \xi_\nu} &= \sum_{\kappa=1}^{\lambda_\nu} (m-\kappa) b_{\nu,\kappa} (x-\xi_\nu)_+^{m-\kappa-1};\ \ \nu = 1, 2, \ldots, r. \end{aligned}$$

It follows that the tangential space $T_{m,n,r}(\lambda; s)$ is spanned by the functions

$$\begin{aligned} x^\mu \ &;\ \ \mu = 0, 1, \ldots, m-1, \\ (x-\xi_\nu)_+^{m-\kappa} \ &;\ \ \kappa = 1, 2, \ldots, \lambda_\nu;\ \ \nu = 1, 2, \ldots, r, \\ b_{\nu,\lambda_\nu}(x-\xi_\nu)_+^{m-\lambda_\nu-1}&;\ \ \nu = 1, 2, \ldots, r. \end{aligned}$$

Let the defect $\delta_{m,n,r}(\lambda; s)$ of s be defined as the number of coefficients b_{ν,λ_ν} with value zero.

Theorem 1. *The tangential space $T_{m,n,r}(\lambda; s)$ of the spline set $M_{m,n,r}(\lambda)$ at the point s is given by the spline space*

$$T_{m,n,r}(\lambda; s) = S_{m,n,r} \begin{pmatrix} \xi_1 & \xi_2 & \cdots & \xi_r \\ \lambda_1^* & \lambda_2^* & \cdots & \lambda_r^* \end{pmatrix},$$

where, for $\nu = 1, 2, \ldots, r$

$$\lambda_\nu^* = \begin{cases} \lambda_\nu + 1 & \text{if } b_{\nu,\lambda_\nu} \neq 0, \\ \lambda_\nu & \text{if } b_{\nu,\lambda_\nu} = 0, \end{cases}$$

where b_{ν,λ_ν} *is the coefficient of* $(x-\xi_\nu)_+^{m-\lambda_\nu}$ *in the truncated power representation (4) of* s *. We have*

$$\dim T_{m,n,r}(\lambda; s) = m + n + r - \delta_{m,n,r}(\lambda; s). \tag{8}$$

Proof. We only have to compute the dimension of the spline space

$$S_{m,n,r}\begin{pmatrix} \xi_1 & \xi_2 & \ldots & \xi_r \\ \lambda_1^* & \lambda_2^* & \ldots & \lambda_r^* \end{pmatrix}.$$

□

Remark. If we skip the assumption $\lambda_\nu \leq m-2$, i.e. if we permit the knots to have multiplicity $\lambda_\nu = m-1$, we get a space of possibly discontinuous functions. We will still call this spline space a tangential space. The reason will be clear in the next section.

Remark. It seems to be interesting to ask for the number $p_{m,n}$ of subsets $M_{m,n,r}(\lambda)$ in the set $S_{m,n}$. This number is equal to the number of the so–called compositions of n, i.e. the number of representations

$$n = \lambda_1 + \lambda_2 + \ldots + \lambda_k$$

with $\lambda_\nu \in I\!N$, where the arrangement of the parts λ_ν is essential, and where each part satisfies $\lambda_\nu \leq m-1$ (cp.P.A. Mac Mahon [10,11], M.–H. Ostmann [19], G.E. Andrews [2]). By elementary considerations it turns out that $p_{m,n}$ is the coefficient of z^n in the power series of

$$\frac{1}{1 - z - z^2 - \ldots - z^{m-1}} = 1 + \sum_{n=1}^{\infty} p_{m,n} z^n.$$

Its radius of convergence will be denoted by $\frac{1}{\alpha_m}$. Here $\alpha_2 = 1$ and $1 < \alpha_m < 2$ with $\lim_{m\to\infty} \alpha_m = 2$. It is (cp. N. Basu [3]) for $m > 2$:

$$p_{m,n} = \frac{(\alpha_m - 1)\alpha_m^{n+m-1}}{2\alpha_m^{m-1} - m}(1 + o(1)) \quad \text{for } n \to \infty.$$

2. The Gauß Transform

Let H be the vector space of all real and Riemann integrable functions $h : \mathbb{R} \to \mathbb{R}$, which obey for some real number γ, a growth restriction

$$h(x) = O(|x|^\gamma) \text{ for } |x| \to \infty. \tag{9}$$

For $h \in H$ and a parameter $t > 0$ the function

$$u(x;t) = \frac{1}{\sqrt{\pi t}} \int_{-\infty}^{+\infty} h(\tau) e^{-\frac{(\tau-x)^2}{t}} d\tau \tag{10}$$

is called the Gauß transform of h. Because of the restriction (6) it is clear that, for $t > 0$, the function u belongs to the class $C^\infty(\mathbb{R})$ w.r.t. the variable x.

We first apply the Gauß transform to some special functions.

Lemma 2. *For $\mu \in \mathbb{N}_0$ let u_μ be the Gauß transform of the power function*

$$h_\mu(x) = x^\mu.$$

Then

$$u_\mu(x;t) = x^\mu + \sum_{\sigma=1}^{[\mu/2]} \frac{\mu!}{(\mu-2\sigma)!\sigma!2^{2\sigma}} t^\sigma x^{\mu-2\sigma} \tag{11}$$

Proof. We have by definition

$$\begin{aligned} u_\mu(x;t) &= \frac{1}{\sqrt{\pi t}} \int_{-\infty}^{+\infty} \tau^\mu e^{-\frac{(\tau-x)^2}{t}} d\tau \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (x+\sqrt{t}v)^\mu e^{-v^2} dv \\ &= \sum_{\lambda=0}^{\mu} \binom{\mu}{\lambda} x^{\mu-\lambda} t^{\lambda/2} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} v^\lambda e^{-v^2} dv. \end{aligned}$$

Because of the formula

$$\frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} v^\lambda e^{-v^2} dv = \begin{cases} 0 & \text{for odd } \lambda \\ \dfrac{(2\sigma)!}{2^{2\sigma}\sigma!} & \text{for } \lambda = 2\sigma, \quad \sigma \in \mathbb{N}_0, \end{cases}$$

the assertion follows easily. □

The second class of functions consists of the truncated power functions for $k \in \mathbb{N}$.

Lemma 3. *(cp. [13],[14]) For $k \in \mathbb{N}_0$ let $w_k(x;t)$ be the Gauß transform of the truncated power function $(x)_+^k$. Then the following assertions are valid:*

1. For $x \geq 0$ we have

$$0 \leq w_k(x;t) = u_k(x;t) + R_1(x;t); \tag{12}$$

2. For $x \leq 0$ we have

$$0 \leq w_k(x;t) = R_2(x;t); \tag{13}$$

where

$$|R_{1,2}(x;t)| \leq \frac{t^{h/2}}{2\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right) e^{\frac{-x^2}{t}}. \tag{14}$$

3. For all $x \in \mathbb{R}$ and $t > 0$ the recursion formula

$$w_{k+2}(x;t) = x w_{k+1}(x;t) + \frac{t(k+1)}{2} w_k(x;t) \tag{15}$$

is valid for $k \in \mathbb{N}_0$. In particular

$$w_0(x;t) = \frac{1}{\sqrt{\pi}} \int_{-\frac{x}{\sqrt{t}}}^{+\infty} e^{-v^2}\, dv \tag{16}$$

and

$$w_1(x;t) = x w_0(x;t) + \frac{\sqrt{t}}{2\sqrt{\pi}} e^{-\frac{x^2}{t}}. \tag{17}$$

4. It is

$$\frac{\partial}{\partial x} w_{k+1}(x;t) = (k+1) \cdot w_k(x;t) \tag{18}$$

for $k \in \mathbb{N}_0$.

Proof. By definition we have

$$\begin{aligned} w_k(x;t) &= \frac{1}{\sqrt{\pi t}} \int_0^\infty \tau^k e^{-\frac{(\tau-x)^2}{t}}\, d\tau \\ &= e^{\frac{-x^2}{t}} \frac{1}{\sqrt{\pi t}} \int_0^\infty \tau^k e^{-\frac{\tau^2}{t} + \frac{2\tau x}{t}}\, d\tau. \end{aligned}$$

For $x \leq 0$ we get

$$0 \leq w_k(x;t) \leq e^{\frac{-x^2}{t}} \frac{1}{\sqrt{\pi t}} \int_0^\infty \tau^k e^{-\frac{\tau^2}{t}}\, d\tau = \frac{t^{k/2}}{2\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right) e^{\frac{-x^2}{t}},$$

which proves (13) and (14). For $x \geq 0$ we write

$$0 \leq w_k(x;t) = \frac{1}{\sqrt{\pi t}} \int_{-\infty}^{+\infty} \tau^k e^{-\frac{(\tau-x)^2}{t}} d\tau - \frac{1}{\sqrt{\pi t}} \int_{-\infty}^{0} \tau^k e^{-\frac{(\tau-x)^2}{t}} d\tau$$

$$\leq u_k(x;t) - e^{-\frac{x^2}{t}} \int_0^\infty \tau^k e^{-\frac{\tau^2}{t}} d\tau,$$

which yields (12). To prove (15), we use integration by parts to evaluate

$$w_k(x;t) = \frac{1}{\sqrt{\pi t}} \int_0^\infty \tau^k e^{-\frac{(\tau-x)^2}{t}} d\tau.$$

This gives

$$\begin{aligned} w_k(x;t) &= \frac{1}{k+1} \frac{1}{\sqrt{\pi t}} \tau^{k+1} e^{-\frac{(\tau-x)^2}{t}} \Big|_0^\infty \\ &+ \frac{2}{t(k+1)\sqrt{\pi t}} \int_0^\infty \tau^{k+1} (\tau - x) e^{-\frac{(\tau-x)^2}{t}} d\tau \\ &= \frac{2}{t(k+1)} \left(w_{k+2}(x;t) - x w_{k+1}(x;t) \right), \end{aligned}$$

and so (15) is proved. Finally, relation (18) is either proved by explicitly differentiating $w_{k+1}(x;t)$ and using (15), or it can be deduced from Walz [22], where it was proved that differentiation and Gauß transformation can be interchanged. □

For arbitrary but fixed $t > 0$ and $r \in \{1, \ldots, n\}$, we denote by $G_{m,n,r} = G_{m,n,r}(\lambda)$ the set of Gauß transforms of the spline set $M_{m,n,r}(\lambda)$, i.e., $G_{m,n,r}$ consists of all functions which can be written in the form

$$g(x;t) = \sum_{\mu=0}^{m-1} a_\mu u_\mu(x;t) + \sum_{\nu=1}^{r} \sum_{k=1}^{\lambda_\nu} b_{\nu,k} w_{m-k}(x - \xi_\nu; t) . \tag{19}$$

The defect of g is defined in analogy to that of a spline function.

Remark. It is also tempting to use the B-spline basis instead of the truncated power basis of the spline space, for the theoretical investigations as well as for the practical computations: The partial derivative of a B-spline w.r.t. the knots is well-known (see e.g. [20]), and also the recursion formula for the Gauß transform of a B-spline has already be found [13, 14]. However, it turned out that the non-dependency of a specific spline function on a knot, in other words, the defect of this function, comes out much more clearly when using the truncated power basis.

The following theorem gives some insight into the structure of $G_{m,n,r}$. In particular it says that the set $G_{m,n,r}$ satisfies the local Haar condition (Statement 1).

Theorem 4. *Let g denote some arbitrary function from $G_{m,n,r}$ with defect $\delta_{m,n,r}(\lambda, g)$. Then the following statements hold true:*
1. The tangential space $T(g)$ of g is a Haar space of dimension

$$d(g) := m + n + r - \delta_{m,n,r}(\lambda, g)$$

on $\mathbb{R}$.
2. For each function $\tilde{g} \in G_{m,n,r}$, $\tilde{g} \neq g$, the difference function $g - \tilde{g}$ can have at most $m + 2n - \delta_{m,n,r}(\lambda, g) - 1$ zeros in $\mathbb{R}$.

Proof. The proof is based on the following auxiliary

Claim. Consider, for arbitrary r, the spline space $S_{m,n,r}(\xi, \lambda)$ with fixed knots. Then the Gauß transform of each $s \in S_{m,n,r}(\xi, \lambda)$ has at most

$$m + n - 1$$

zeros in $\mathbb{R}$.

Proof of Claim. Since $S_{m,n,r}(\xi, \lambda)$ is a weak Haar space, the statement could be deduced from a general result due to Jones & Karlovitz [7]. However, in this concrete situation we want to give a short self-contained proof. So, consider an arbitrary $g \in G_{m,n,r}$

$$g(x;t) = \sum_{\mu=0}^{m-1} a_\mu u_\mu(x;t) + \sum_{\nu=1}^{r} \sum_{k=1}^{\lambda_\nu} b_{\nu,k} w_{n-k}(x - \xi_\nu; t),$$

and assume that g has at least $m+n$ zeros in $\mathbb{R}$. Then $g^{(m)}$ has at least n zeros, due to Rolle's theorem (here and below, we consider derivatives w.r.t to x).
Obviously, $u_\mu^{(m)}(x;t) \equiv 0$ for $\mu = 0, \ldots, m-1$, due to (11). Consider for arbitrary ν and k the function $w_{m-k}^{(m)}(x - \xi_\nu; t)$. From (18), it is easy to see that

$$w_{m-k}^{(m-k)}(x - \xi_\nu; t) = (m-k)! \cdot w_0(x - \xi_\nu; t),$$

and so, using (16),

$$w_{m-k}^{(m)}(x - \xi_\nu; t) = \tilde{p}_{k-1,\nu}(x) e^{-\frac{(x-\xi_\nu)^2}{t}} = e^{-\frac{x^2}{t}} p_{k-1,\nu}(x) e^{\frac{2\xi_\nu}{t} \cdot x}$$

with polynomials $\tilde{p}_{k-1,\nu}$ and $p_{k-1,\nu}$ of degree $k-1$. It follows that

$$\begin{aligned} g^{(m)}(x;t) &= e^{-\frac{x^2}{t}} \sum_{\nu=1}^{r} \sum_{k=1}^{\lambda_\nu} b_{\nu,k}\, p_{k-1,\nu}(x) e^{\frac{2\xi_\nu}{t} \cdot x} \\ &= e^{-\frac{x^2}{t}} \sum_{\nu=1}^{r} q_\nu(x) e^{\frac{2\xi_\nu}{t} \cdot x} \end{aligned}$$

with polynomials q_ν of degree $\lambda_\nu - 1$. Thus $g^{(m)}$ is a nonzero multiple of a generalized exponential sum, which can have at most

$$\sum_{\nu=1}^{r}(\lambda_\nu - 1 + 1) - 1 = n - 1$$

zeros. This contradicts the assumption and proves the auxiliary claim.

Now statement 1 follows immediately from Theorem 1 in connection with our auxiliary claim. Since both g and $\tilde{g}$ are elements of $G_{m,n,r}$ they are images of two spline functions of order m, say s and $\tilde{s}$. Obviously, the difference function $s - \tilde{s}$ is a spline of order m with at most $2n - \delta_{m,n,r}(\lambda, s)$ knots, and so also statement 2 follows from the auxiliary claim. □

Remark. Statement 2 of Theorem 4 can be sharpened in several directions, for example, if not all knots of g and $\tilde{g}$ are different. This can be seen as follows: Let $\{\zeta_1, \dots, \zeta_l\}$ be the set of common knots of g and $\tilde{g}$, and denote their multiplicities as knots of g and $\tilde{g}$ by λ_j resp. $\tilde{\lambda}_j$, $j = 1, \dots, l$. Then the same proof as above shows that $g - \tilde{g}$ can have at most

$$m + 2n - \delta_{m,n,r}(\lambda, g) - 1 - \rho$$

zeros, where

$$\rho := \sum_{j=1}^{l} \min\{\lambda_j, \tilde{\lambda}_j\} \geq l.$$

3. Criteria for Best Approximation by $G_{m,n}$

We now use the results of the previous section, in particular Theorem 4, to obtain necessary and sufficient conditions for best approximations.

Theorem 5. *Let $f \in C[a,b]$ be given. Consider the following statements for $g \in G_{m,n,r}$:*
(a) The error function $(f - g)$ has an alternant of length

$$m + 2n - \delta_{m,n,r}(\lambda, g) + 1 \,.$$

(b) g is a best approximation of f.
(c) The error function $(f - g)$ has an alternant of length

$$m + n + r - \delta_{m,n,r}(\lambda, g) + 1 \,.$$

Then (a) $\Rightarrow$ (b) $\Rightarrow$ (c).

Proof. The basis for this result was built in Theorem 4, such that we now can use standard techniques from nonlinear approximation theory, see e.g. Meinardus [12].

To prove the implication $(a) \Rightarrow (b)$, assume that there is a better approximation of f, say $\hat{g}$. Then $(f-g)-(f-\hat{g}) = \hat{g}-g$ has at least $m+2n-\delta_{m,n,r}(\lambda,g)$ zeros, which is a contradiction to statement 2 of Theorem 4. The implication $(b) \Rightarrow (c)$ follows immediately from the fact that the set $G_{m,n,r}$ satisfies the so–called local Haar condition, as proved in Theorem 4, statement 1. □

As an immediate implication of Theorem 5, we obtain

Corollary 6. *If g has only single knots (i.e. if $r=n$), then g is a best approximation of f if and only if the error function $(f-g)$ has an alternant of length*

$$m+2n-\delta_{m,n,n}(\lambda,g)+1 \ .$$

The following de la Vallée Poussin type result gives us lower and (trivial) upper bounds for the minimal deviation of f from $G_{m,n,r}$, denoted as $\rho_{G_{m,n,r}}(f)$.

Theorem 7. *For $f \in C[a,b]$ and $g \in G_{m,n,r}$, let there exist $m+2n-\delta_{m,n,r}(\lambda,g)+1$ points $\{z_j\} \subset [a,b]$, such that*
1. $(f-g)(z_j) \neq 0$ *for* $j=1,\dots,m+2n-\delta_{m,n,r}(\lambda,g)+1$
and
2. $\operatorname{sgn}\,(f-g)(z_j) = -\operatorname{sgn}\,(f-g)(z_{j+1})$, *for* $j=1,\dots,m+2n-\delta_{m,n,r}(\lambda,g)$.
Then

$$\min_{1\le j\le m+2n-\delta_{m,n,r}(\lambda,g)+1} |(f-g)(z_j)| \le \rho_{G_{m,n,r}}(f) \le \|f-g\|_{[a,b]}.$$

Proof. It follows from our assumptions in connection with Theorem 4 that there is no function $\tilde{g} \in G_{m,n,r}$, such that

$$(f(z_j)-g(z_j))\,(\tilde{g}(z_j)-g(z_j)) > 0 \quad \text{for} \quad j=1,\dots,m+2n-\delta_{m,n,r}(\lambda,g)+1.$$

Therefore the result follows from the nonlinear Kolmogoroff criterion (cf. Meinardus [12], Theorem 85). □

4. Particular Results for Single Knots

We had already seen in Corollary 6 that in the case $r=n$ a full characterization of the best approximation by the length of the alternant is possible. In this section we want to consider the case $r=n$ in some more detail. The basis for our results is the following

Theorem 8. *If $r = n$ then the set $G_{m,n,r}$ has the local and the global Haar property, i.e., for each $g \in G_{m,n,n}$*
1. the tangential space $T(g)$ is a Haar space of dimension

$$d(g) = m + 2n - \delta_{m,n,n}(\lambda, g)$$

on $\mathbb{R}$, and
2. for each function $\tilde{g} \in G_{m,n,n}$, $g \neq \tilde{g}$, the difference function $g - \tilde{g}$ can have at most

$$m + 2n - \delta_{m,n,n}(\lambda, g) - 1$$

zeros in $\mathbb{R}$.

Proof. This result follows immediately from Theorem 4. □

By applying the non–linear theory developed in [12,17], we get as an immediate consequence of Theorem 8 the following characterization and uniqueness result.

Theorem 9. *Let $f \in C[a, b]$. Then the following statements hold.*
1. The function f can have at most one best approximation from $G_{m,n,n}$.
2. A function $g^* \in G_{m,n,n}$ is the best approximation of f, if and only if the difference function $(f - g^*)$ has an alternant of length

$$m + 2n + 1 - \delta_{m,n,n}(\lambda, g) .$$

In the same manner, also the de la Vallée Poussin type result given in Theorem 7 can be reformulated.

Now, the question may arise for which functions f the existence of a best approximation with single knots can be expected. Extending an idea due to Handscomb [5] (see also [6]), we can proof the following criterion. For the fomulation of this result, it is appropriate to consider approximation of the Gauß transform of f, say $F(\cdot\,;t)$, instead of f itself.

Theorem 10. *Consider $f \in C[a, b]$ such that $f^{(m-2)}$ exists and is strictly convex in (a, b). By $F(x;t)$, we denote the Gauß transform of f.*
Then each best approximation $g^* \in G_{m,n,r}$ of F has n distinct knots.
Moreover, the error function $(g^* - F)$ has an alternant of exact length $m + 2n + 1$, and the endpoints a and b of the interval belong to the alternant.

Proof. For convenience, the proof is split up into several claims. We assume troughout that $t > 0$ is arbitrary but fixed.

Claim 1. Let g be an arbitrary Gauß transformed spline of order m with exactly r inner knots, i.e., $g \in G_{m,n,r}$. Then $(g - F)(x;t)$ has at most $m + 2r + 1$ extreme points, this number being attained only if a and b belong to the set of extrema.

Proof of Claim 1. Let g be the Gauß transform of the spline $s \in S_{m,n,r}$ with knots $\xi_1, \dots, \xi_r$. Then $s^{(m-2)}$ is a (possibly discontinuous) piecewise linear function, and $(s^{(m-2)} - f^{(m-2)})$ has at most $(2r+2)$ sign changes in $[a,b]$. Since the Gauß transform is variation–diminishing, $(g^{(m-2)} - F^{(m-2)})$ has at most $(2r+2)$ zeros in $[a,b]$. It follows that $(g' - F')$ has at most $(2r+m-1)$ zeros, and therefore $(g - F)$ has at most $(2r+m-1)$ extrema in (a,b). This proves Claim 1.

Claim 2. If $g^* \in G_{m,n,r}$ is a best approximation of F, then $n = r + \delta_{m,n,r}(\lambda; g)$.

Proof of Claim 2. According to Corollary 6 , the error function $(g^* - F)$ has at least $m + n + r - \delta_{m,n,r}(\lambda; g) + 1$ alternating extreme points, and so Claim 1 implies

$$m + n + r - \delta_{m,n,r}(\lambda, g) + 1 \leq 2r + m + 1, \tag{20}$$

hence

$$n - \delta_{m,n,r}(\lambda, g) \leq r. \tag{21}$$

On the other hand, the inequality

$$\delta_{m,n,r}(\lambda, g) \leq n - r \tag{22}$$

is obvious.
A combination of (21) and (22) now implies

$$n \leq r + \delta_{m,n,r}(\lambda, g) \leq n\ , \tag{23}$$

which completes the proof of Claim 2.

Claim 3. The error function $(g^* - F)$ (with g^* as in Claim 2) has an alternant of exact length $m + 2r + 1$, and a and b belong to the alternant.

Proof of Claim 3. If g^* is a best approximation Corollary 6 in combination with Claim 2 implies that $(g^* - F)$ has an alternant of length

$$m + n + r - \delta_{m,n,r}(\lambda, g) + 1 = m + 2r + 1 \tag{24}$$

at least. On the other hand, the number of extreme points is also bounded from above by this number, due to Claim 1. This proves Claim 3.

Claim 4. $r = n$.

Proof of Claim 4. Assume $r < n$. Then we may insert one additional knot $\xi_{r+1} = b - \epsilon$ with $\epsilon > 0$ sufficiently small. Since b is an element of the alternant, it is possible to choose a polynomial p_ϵ, such that $\tilde{g}$, defined as

$$\tilde{g}(x;t) := \begin{cases} g^*(x;t) & a \leq x \leq b - \epsilon \\ p_\epsilon(x;t) & b - \epsilon \leq x \leq b \end{cases}$$

has the same error norm as g^*. This contradicts the unicity proved in Theorem 9. □

Remark. The assumptions of Theorem 10 are in particular satisfied for the function $f(x) = x^m$, i.e. for the monospline problem, see e.g. Braess [4]. (Note that

$$\rho_{G_{m,n,r}}(x^m) = \rho_{G_{m,n,r}}(u_m(x;t))$$

for all t, due to Lemma 2).

Moreover, it is known (see e.g. Schumaker [20]) that the error in free knot spline approximation behaves asymptotically like that in segmented approximation by piecewise polynomials of degree $m-1$, which itself is known to be of order n^{-m}, see Meinardus [15]. Thus, altogether we can prove that

$$\rho_{G_{m,n,n}}(x^m) = \rho_{S_{m,n,n}}(x^m) = O\left(n^{-m}\right) \quad \text{for } n \to \infty .$$

At the moment we are looking at the monospline problem in some more detail; the corresponding results will be published in a forthcoming paper.

In rational best approximation, it is well–known that anomalies occur in square blocks (cf. Meinardus [12]). In the present case, a triangular structure appears:

Lemma 11. *Assume that the best approximation $g \in G_{m,n,r}$ of $f \in C[a,b]$ is of exact order m and possesses n distinct knots, i.e., $r = n$, and set $\tilde{G}_{m,n} := G_{m,n,n}$.*
If the error function $(f-g)$ has an alternant of exact length

$$m + 2n + 1 + k$$

with $k \in \mathbb{N}_0$, then g is the best approximation of f w.r.t. the spaces

$$\tilde{G}_{m+\mu,n+\nu}$$

for all $\mu, \nu \in \mathbb{N}_0$ with $\mu + \nu \leq k$ and no others.

Proof. Due to our assumptions, the defect of g w.r.t. the set $\tilde{G}_{m+\mu,n+\nu}$ equals ν. It follows Corollary 6 that g is the best approximation of f from this set if and only if the error function $(f-g)$ has an alternant of length

$$m + \mu + 2(n+\nu) - \nu + 1 = m + 2n + \mu + \nu + 1 .$$

This obviously is the case if and only if $\mu + \nu \leq k$.
Moreover, if either $\mu < 0$ or $\nu < 0$, then g is no element of the set $\tilde{G}_{m+\mu,n+\nu}$. This proves Lemma 11. □

We briefly discuss an example for this effect:
Consider the function

$$f(x) := 2|x| - 1 .$$

On $[-1, 1]$, this function has an alternant of length 3, and so the function $g_0 \equiv 0$, considered as an element of the set $\tilde{G}_{1,0}$, is the best approximation of f; it meets the assumptions of Lemma 11 with $m = 1$, $n = 0$, and $k = 1$. Consequently, g_0 is also the best approximation of f w.r.t. the sets $\tilde{G}_{2,0}$ and $\tilde{G}_{1,1}$, the minimal deviation being 1 in all cases.

However, if the Walsh type tableau would have a rectangular structure, then g_0 would also be the best approximation from $\tilde{G}_{2,1}$. But this is not true, since f itself is a linear spline with one knot, and thus the spaces $\tilde{G}_{2,1}$ can get arbitrarily close to f (for $t \to 0$).

In view of these phenomena, we introduce the following definition, which is well known e.g. in rational best approximation: A function f is called (m, n)-normal, if the defect of its best approximation from $G_{m,n,n}$ is zero. It is easy to see that a function is (m, n)-normal if and only if its minimal deviation satisfies

$$\rho_{G_{m,n,n}}(f) < \rho_{G_{m,n-1,n-1}}(f) .$$

Moreover, we have seen a sufficient normality criterion in Theorem 10.

5. Numerical Examples

In this section we show the results of some numerical examples. Clearly this is only a small selection of a large number of numerical tests that was examined.

We observed that the calculations always were very stable, even if the parameter t was chosen close to zero. In all examples shown below we have $t = 10^{-7}$. This means that we deal with functions that are very close to the original spline space, but still have a Haar-like structure.

As a first example, we look at the function $f(x) = e^x$ on the interval $[0, 1]$. Since the exponential function satisfies the assumptions of Theorem 10, we can expect that the error function of the best free knot approximation always has an alternant of length $(m + 2n + 1)$.

We compare the error functions of the best approximations for fixed (equidistant) knots (Figures 1a and 2a), and for free knots (Figures 1b and 2b). In all cases, the alternant is of length $(m + n + 1)$ for fixed knots, and $(m + 2n + 1)$ for free knots.

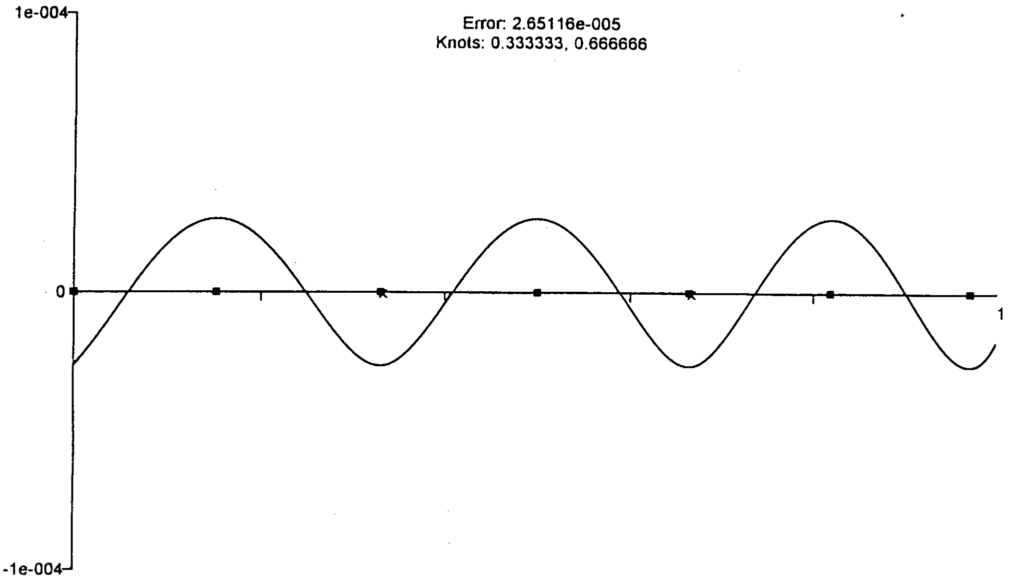

Figure 1a. Approximation of e^x : Fixed knots, $m = 4$, $n = 2$.

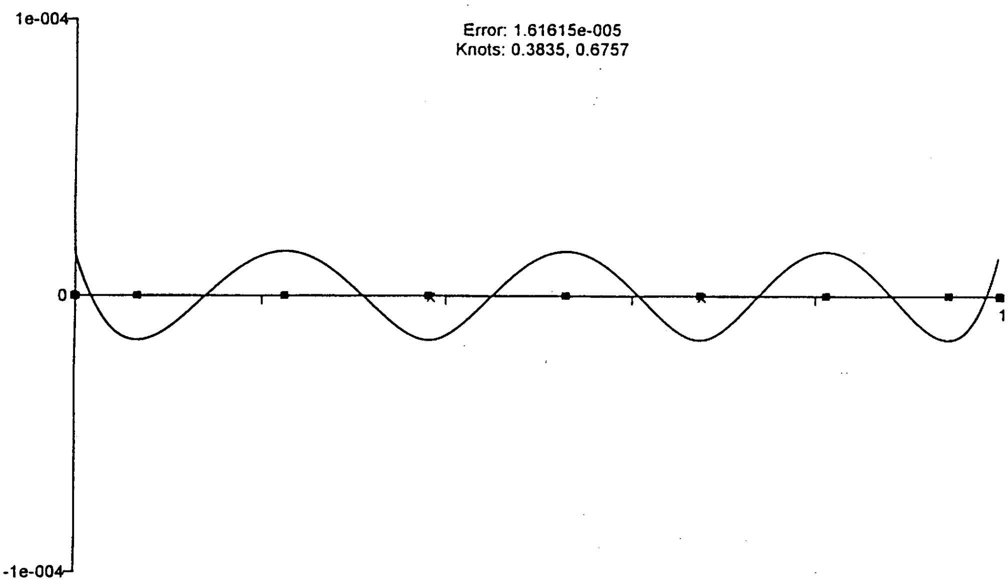

Figure 1b. Approximation of e^x : Free knots, $m = 4$, $n = 2$.

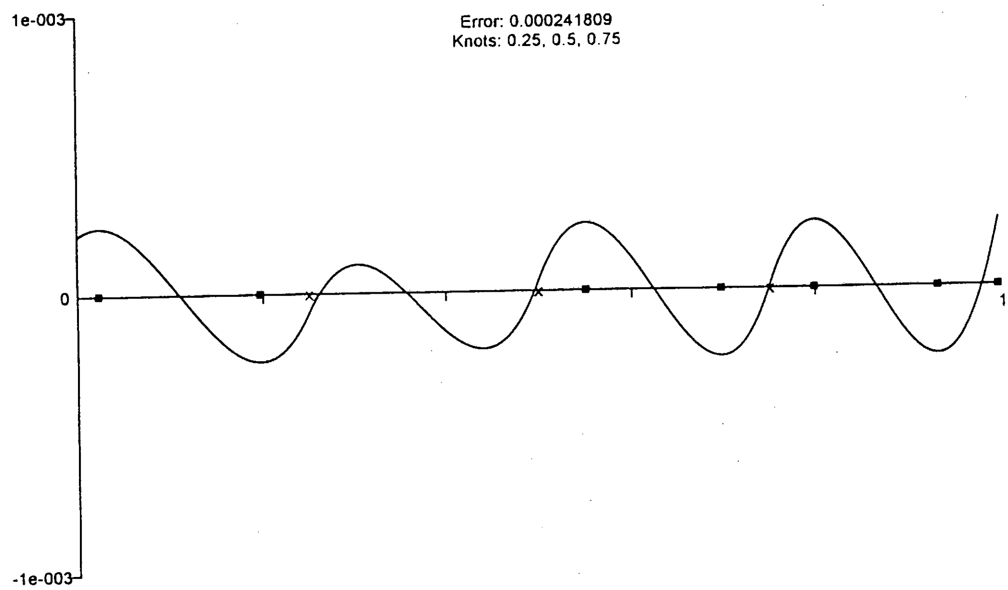

Figure 2a. Approximation of e^x : Fixed knots, $m = 3$, $n = 3$.

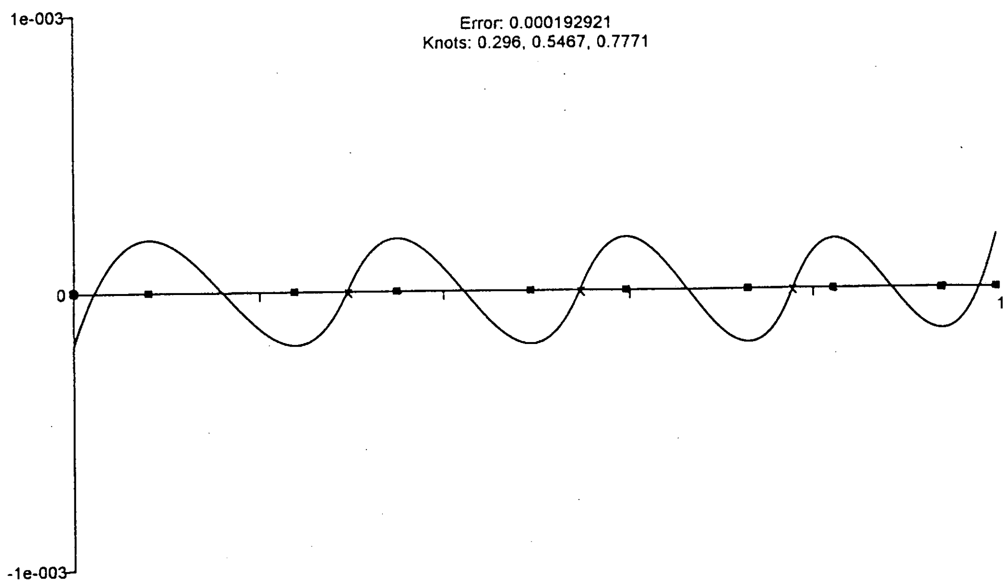

Figure 2b. Approximation of e^x : Free knots, $m = 3$, $n = 3$.

Our second example is the square root function, also on the interval $[0,1]$. Here we have the effect that the optimal knots are very close to the left endpoint of the interval, and so a part of the alternant simply becomes invisible when the values of m or n are larger than 2. We therefore show the output for the cases $m = n = 2$ (Figure 3) and $m = 1$, $n = 2$ (Figure 4). Note that in the latter case the error function is indeed continuous (since t is greater than zero), although with $t = 10^{-7}$ we are very close to the spline space of piecewise constants.

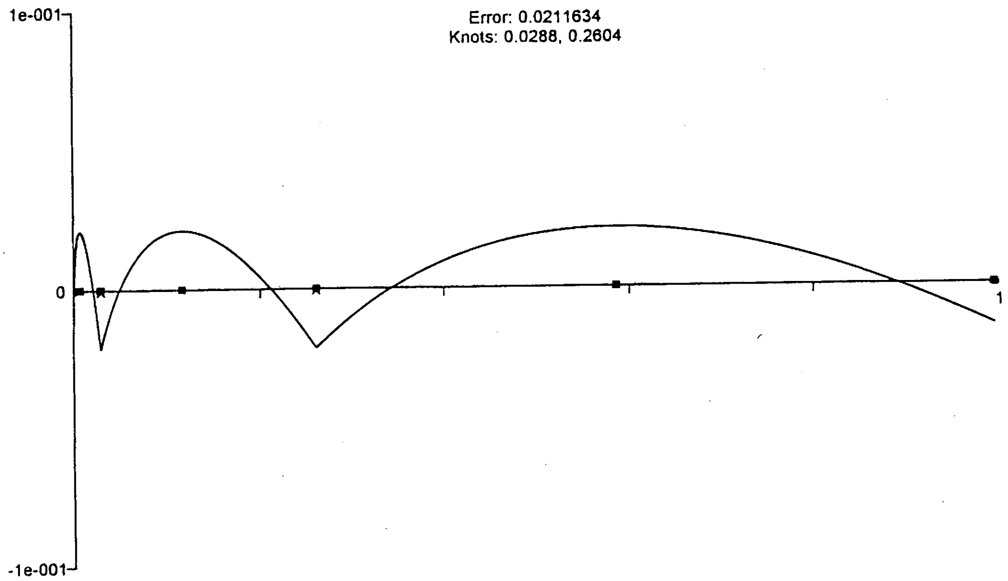

Figure 3. Approximation of $\sqrt{x}$: Free knots, $m = 2$, $n = 2$.

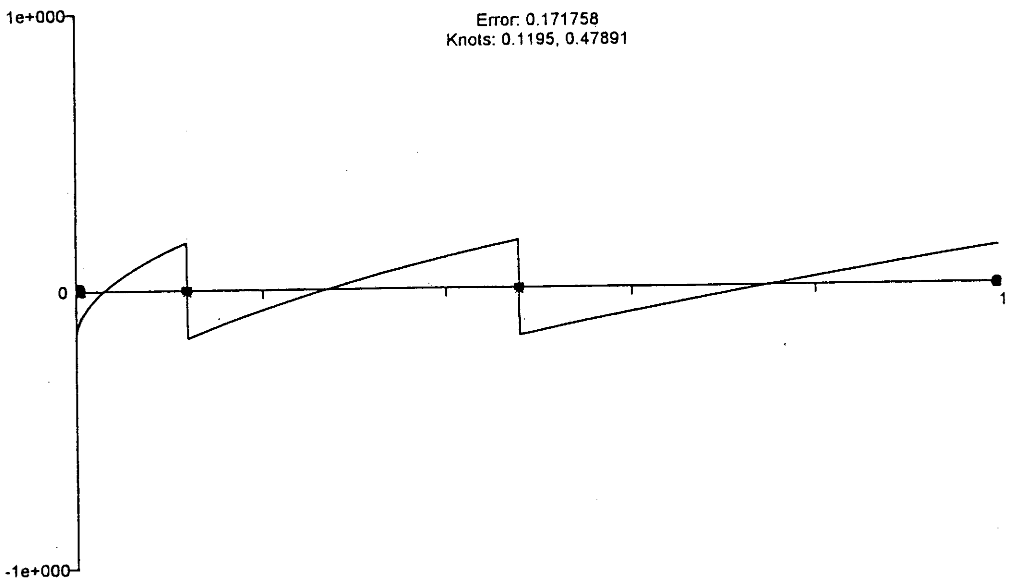

Figure 4. Approximation of $\sqrt{x}$: Free knots, $m = 1$, $n = 2$.

The Runge function

$$f_R(x) = \frac{1}{1+x^2}$$

on the interval $[-5,5]$ provides an example for non-normality, since here situations with $\delta_{m,n,r} = 1$ occur. More precisely, we have the

Proposition. *For an arbitrary odd number m, denote by g_m the best approximation of f_R from $G_{m,1,1}$. Then $\delta_{m,1,1}(g_m) = 1$. Consequently, the error function $f_R - g_m$ has an alternant of length $m+2$.*

Proof. Consider the best polynomial approximation $p_m \in \Pi_m$ of f_R on $[-5,5]$. Then $f_R - p_m$ has an alternant of length $(m+2)$. Since f_R is an even function, the leading coefficient of p_m vanishes, and so $p_m \in \Pi_{m-1} \subset G_{m,1,1}$. Considered as an element of $G_{m,1,1}$, p_m has no "active" knot, and therefore $\delta_{m,1,1}(p_m) = 1$.

Theorem 5 implies that p_m is the best approximation w.r.t. $G_{m,1,1}$ iff the error function has an alternant of length

$$m + 2n + 1 - \delta_{m,1,1}(p_m) = m + 2 .$$

But this is true, as shown above, and so, setting $g_m = p_m$, the proposition is proved. □

As an example, Figure 5 shows the error function of the best approximation from $G_{11,1,1}$ with an alternant of length 13. Note that this was computed by setting $n = 1$, and as expected the coefficient of the truncated power function vanishes.

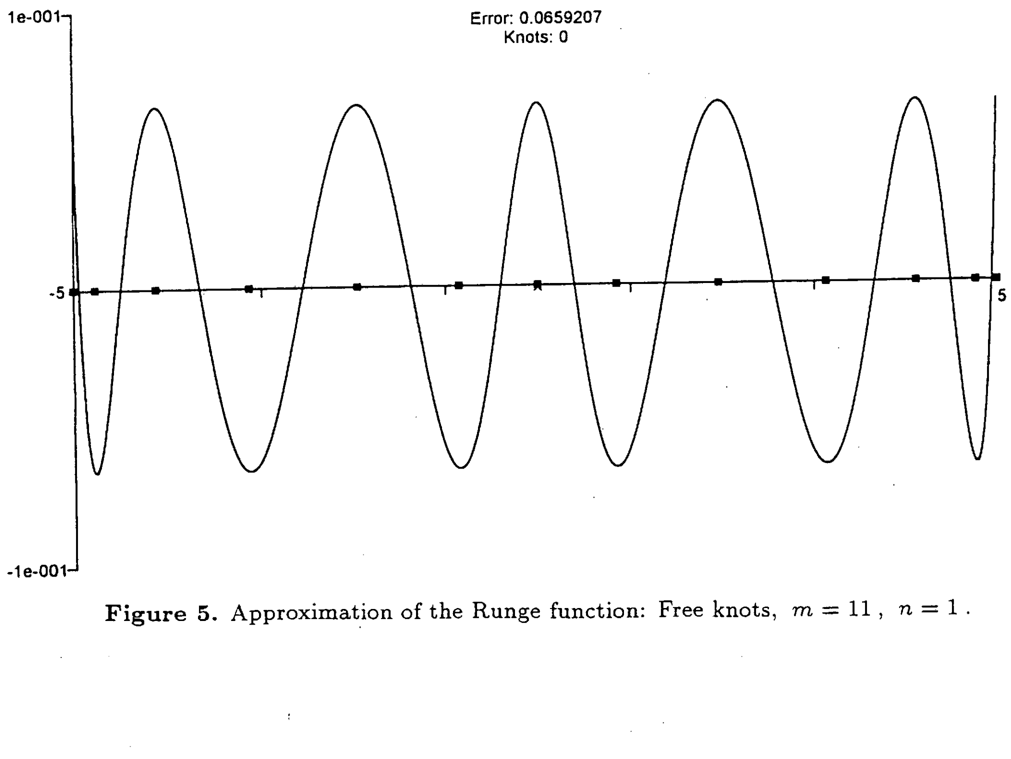

Figure 5. Approximation of the Runge function: Free knots, $m = 11$, $n = 1$.

Finally we want to investigate the dependency of the approximation error on the parameter t. To do this, we computed the best approximations of the exponential function on $[0,1]$ by the spaces $G_{4,1}$ (piecewise cubics with one inner knot) for decreasing values of the parameter t. Table 1 shows the optimal knots and the corresponding errors for $t = 2^{-i}$, $i = 1, \ldots, 12$.

i	Optimal Knot	$\rho_{G_{4,1}}(e^x)$
1	0.7736	$0.4262 \cdot 10^{-6}$
2	0.6491	$0.8058 \cdot 10^{-6}$
3	0.5875	$0.1452 \cdot 10^{-5}$
4	0.5576	$0.2408 \cdot 10^{-5}$
5	0.5437	$0.3545 \cdot 10^{-5}$
6	0.5376	$0.4585 \cdot 10^{-5}$
7	0.5348	$0.5342 \cdot 10^{-5}$
8	0.5335	$0.5814 \cdot 10^{-5}$
9	0.5329	$0.6085 \cdot 10^{-5}$
10	0.5325	$0.6234 \cdot 10^{-5}$
11	0.5324	$0.6313 \cdot 10^{-5}$
12	0.5323	$0.6354 \cdot 10^{-5}$
∞	0.5322	$0.6396 \cdot 10^{-5}$

Table 1. Approximation of e^x by $G_{4,1}$ on $[0,1]$ for various values of $t = 2^{-i}$.

Obviously, in both sequences the deviation from the limit value (for $t = 0$ resp. $i = \infty$) is roughly halved in each step. This already shows that these values behave asymptotically like t does. However, based on the fact that the basis functions u_μ and w_k of the Gauß transformed space possess an asymptotic expansion for $t \to 0$, as shown in Lemma 2 and Lemma 3, we strongly conjecture that this is also true for the other parameters of the best approximation, i.e., the coefficients of the best approximating function, the optimal knots, and the minimal deviations.

This is also supported by the numerical results: We applied linear extrapolation (Richardson extrapolation) to the sequence of errors from above, which improved the results significantly. Table 2 shows the output of the extrapolation process in the usual triangular Romberg scheme (cf. [21]). The convergence acceleration effect is obvious.

.4262 e(-5)							
	.1185 e(-4)						
.8058 e(-5)		.2403 e(-4)					
	.2099 e(-4)		.3983 e(-4)				
.1452 e(-4)		.3785 e(-4)		.5401 e(-4)			
	.3364 e(-4)		.5312 e(-4)		.6130 e(-4)		
.2408 e(-4)		.5121 e(-4)		.6107 e(-4)		.6327 e(-4)	
	.4682 e(-4)		.6058 e(-4)		.6324 e(-4)		.6371 e(-4)
.3545 e(-4)		.5941 e(-4)		.6317 e(-4)		.6370 e(-4)	
	.5626 e(-4)		.6301 e(-4)		.6369 e(-4)		.6385 e(-4)
.4585 e(-4)		.6256 e(-4)		.6368 e(-4)		.6385 e(-4)	
	.6098 e(-4)		.6364 e(-4)		.6385 e(-4)		.6395 e(-4)
.5342 e(-4)		.6350 e(-4)		.6384 e(-4)		.6395 e(-4)	
	.6287 e(-4)		.6383 e(-4)		.6395 e(-4)		.6396 e(-4)
.5814 e(-4)		.6379 e(-4)		.6395 e(-4)		.6396 e(-4)	
	.6356 e(-4)		.6394 e(-4)		.6396 e(-4)		.6396 e(-4)
.6085 e(-4)		.6392 e(-4)		.6396 e(-4)		.6396 e(-4)	
	.6383 e(-4)		.6396 e(-4)		.6396 e(-4)		
.6234 e(-4)		.6395 e(-4)		.6396 e(-4)			
	.6392 e(-4)		.6396 e(-4)				
.6313 e(-4)		.6396 e(-4)					
	.6395 e(-4)						
.6354 e(-4)							

Table 2. Extrapolated Errors $\rho_{G_{4,1}}(e^x)$

References

1. M. Adam, *Beste Approximation durch Splines mit freien Knoten, Diploma Thesis.* Universität Mannheim, 1992.

2. G. E. Andrews, *The Theory of Partitions.* Encyclopedia of Mathematics and its Applications, Vol. 2., Addison–Wesley Publ., Reading, Mass., 1976.

3. N. Basu, *A Note on Partitions.* Bull. Calcutta M. S. **44**, (1952), 27–30.

4. D. Braess, *Nonlinear Approximation Theory.* Springer Verlag, Heidelberg, 1986.

5. D. C. Handscomb, *Characterization of Best Spline Approximation with Free Knots.* In: A. Talbot (ed.), *Approximation Theory.* Proc. Symp. Lancaster, Academic Press, 1969, 63–70.

6. R.S. Johnson, *On Monosplines of Least Deviation,* Trans. AMS **96** (1960), 458 – 477.

7. R.C. Jones & L.A. Karlovitz, *Equioscillation under Nonuniqueness in the Approximation of Continuous Functions,* J. Approx. Theory **3** (1970), 138 – 145.

8. S. Karlin, *Total Positivity.* Stanford University Press, Stanford, 1968.

9. S. Karlin & W.J. Studden, *Tchebycheff Systems: With Applications in Analysis and Statistics.* Interscience, New York, 1966.

10. P. A. Mac Mahon, *Memoir on the theory of the compositions of numbers.* Philos. Trans. Roy. Soc. London **A 184**, (1894), 835–901.

11. P. A. Mac Mahon, *Second Memoir on the compositions of numbers.* Philos. Trans. Roy. Soc. London **A 207**, (1908), 65–134.

12. G. Meinardus, *Approximation of Functions: Theory and Numerical Methods.* Springer Verlag Heidelberg, 1967.

13. G. Meinardus, *Splines und die Wärmeleitungsgleichung,* In: J.W.Schmidt, H.Späth (eds.), *Splines in Numerical Analysis,* Akademie Verlag, Berlin, (1989), 87–96.

14. G. Meinardus, *On the Gauß Transform of Polynomial Spline Spaces,* Results in Mathematics **16** (1989), 290 – 298.

15. G. Meinardus, *Some Results in Segmented Approximation,* Comput. Math. Appl. **33**, (1997), 165–180.

16. G. Meinardus & D. Schwedt, *Nicht-lineare Approximationen.* Arch. Rat. Mech. Anal. **17**, (1964), 297–326.

17. G. Meinardus & G. Walz, *Approximation Theory.* Springer Verlag Heidelberg, in preparation.

18. G. Nürnberger, *Approximation by Spline Functions.* Springer–Verlag, Heidelberg, 1989.

19. M.–H. Ostmann, *Additive Zahlentheorie 1.* Ergebnisse der Mathematik und ihrer Grenzgebiete Heft 7. Springer Verlag, Heidelberg, 1956.

20. L. Schumaker, *Spline Functions: Basic Theory.* John Wiley & Sons, New York, 1981.

21. G. Walz, *Asymptotics and Extrapolation,* Akademie–Verlag, Berlin, 1996.

22. G. Walz, *Splines with Free Knots, the Heat Equation, and the Gauß Transform,* Comm. Appl. Anal. **2** (1998), 169–182.