

ROSTOCKER MATHEMATISCHES KOLLOQUIUM

Heft 74

JEAN-LUC VOLÉRY; JEAN-CLAUDE YAKOUBSOHN	<i>α-Theory of Hald's method</i>	3
COLIN BUTTCHEREIT	<i>Monotones Iterationsschema in Risikomodellen</i>	27

UNIVERSITÄT ROSTOCK

INSTITUT FÜR MATHEMATIK

2024

Herausgeber: Universität Rostock
Mathematisch – Naturwissenschaftliche
Fakultät
Institut für Mathematik

Wissenschaftlicher Beirat: Friedemann Brock
Pavel Drábek
Michael Dreher
Gohar Kyureghyan
Jan-Christoph Schlage-Puchta
Dieter Schott
Peter Takáč

Schriftleitung: Raimond Strauß

Herstellung der Druckvorlage: Susann Dittmer

Zitat–Kurztitel: Rostock. Math. Kolloq. **74** (2024)

ISSN 0138-3248

© Universität Rostock, Institut für Mathematik, D-18051 Rostock

BEZUGSMÖGLICHKEITEN: Universität Rostock
Universitätsbibliothek, Schriftentausch
18051 Rostock
Tel.: +49-381-498 86 37
Fax: +49-381-498 86 32
e-mail: maria.schumacher@uni-rostock.de

Universität Rostock
Institut für Mathematik
18051 Rostock
Tel.: +49-381-498 6551
Fax: +49-381-498 6553
e-mail: romako@uni-rostock.de

DRUCK: ProExakt GmbH
Schöppenstedt

JEAN-LUC VOLÉRY, JEAN-CLAUDE YAKOUBSOHN

α -Theory of Hald's method*

ABSTRACT. In this paper, we present new results concerning α -theory for the Hald method to approximate a zero of an analytic function defined on Banach spaces. Hald's method is convenient since there is a quadratic convergence towards the zero and the calculation of the inverse of the operator is not required. We also present a numerical experiment to illustrate our results to solve a classical Fredholm integral equation.

KEY WORDS. α -theory, Hald's method, Newton's method, Ulm's method.

1 Introduction and main Results

We consider F an analytic function defined from an open ball $B(x_0, r) \subset \mathcal{E}$ into \mathcal{F} such that $DF(x_0)^{-1}$ exists where \mathcal{E} and \mathcal{F} are two Banach spaces. We note $L(\mathcal{F}, \mathcal{E})$ the set of linear maps from \mathcal{F} to \mathcal{E} . We denote indifferently by $\|\cdot\|$ a norm in \mathcal{E} or \mathcal{F} . Throughout the text ζ is a zero of F . Newton's method is certainly the most well known to solve non linear equations $F(x) = 0$. It consists in defining the sequence from x_0 by

$$x_{k+1} = x_k - DF(x_k)^{-1}F(x_k), \quad k \geq 0. \quad (1.1)$$

A quick search on Mathscinet with keywords **Newton's method** gives almost 9000 answers. Less known is Hald's method [7] defined by the sequence

$$B_0 \in L(\mathcal{F}, \mathcal{E}) \quad (1.2)$$

$$x_{k+1} = x_k - B_k F(x_k), \quad k \geq 0 \quad (1.3)$$

$$B_{k+1} = 2B_k - B_k DF(x_{k+1}) B_k, \quad k \geq 0. \quad (1.4)$$

Here we will treat the case where B_0 is an approximation of $DF(x_0)^{-1}$. Note that (1.4) is one step of Newton's method applied to the equation $G(B) = B^{-1} - DF(x_{k+1}) = 0$. We see

*Submitted to the editors 2024-06-04.

that Hald's method is "inversion free" and can be interesting when the derivative operator of F is difficult to compute. Note that under the condition $\|BDF(x_{k+1}) - I\| < 1$ we can prove easily the quadratic convergence of the sequence (1.4) from $B_0 = B$ which is based on the following equality :

$$I - B_{k+1}DF(x_{k+1}) = (I - B_kDF(x_{k+1}))^2. \quad (1.5)$$

This method to approximate the inverse is known as Schulz method [13].

The goal of this paper is to do an α -theory of the Hald's method. The α -theory consists in giving conditions obtained from a point which imply the convergence of Hald sequence towards a solution of $F(x) = 0$. In the classical Kantorovich's theory the conditions of convergence are obtained from a behaviour of the second derivative of F in a ball as it is the case for a function of class C^2 . This reduction of punctual conditions is possible only if the function F is analytic. These punctual conditions are established thanks to following quantities :

$$\begin{aligned} \beta(F, B, x) &= \|BF(x)\| \\ \gamma(F, B, x) &= \max_{j \geq 2} \left(\frac{1}{j!} \|BDF^{(j)}(x)\| \right)^{1/(j-1)} \\ \alpha(F, B, x) &= \beta(F, B, x)\gamma(F, B, x) \\ \delta(F, B, x) &= \|I - BDF(x)\|. \end{aligned}$$

We will denote $\beta_k, \gamma_k, \alpha_k$ and δ_k respectively for $\beta(F, B_k, x_k)$ etc... We also denote $\beta = \beta_0, \gamma = \gamma_0, \alpha = \alpha_0$ and $\delta = \delta_0$. Finally we introduce the dominating function

$$h(\tau) = \alpha - (1 - \delta)\tau + \frac{\tau^2}{1 - \tau}.$$

and the associated Hald sequence defined by

$$\begin{aligned} b_0 &= -1, \tau_0 = 0 \\ \tau_{k+1} &= \tau_k - b_k h(\tau_k), \quad k \geq 0. \\ b_{k+1} &= (2 - b_k h'(\tau_{k+1}))b_k. \end{aligned} \quad (1.6)$$

We will also use in the sequel the following quantities :

$$\delta_k = 1 - b_k h'(\tau_k), \quad k \geq 0 \quad (1.7)$$

We also denote by

$$\psi(\tau) = (-\delta + 2)\tau^2 + (2\delta - 4)\tau + 1 - \delta \quad (1.8)$$

the function such that $h'(\tau) = -\frac{\psi(\tau)}{(1-\tau)^2}$ and by

$$d = \alpha^2 + 2(\delta - 3)\alpha + (1 - \delta)^2 \quad (1.9)$$

the discriminant of the polynomial $(1 - \tau)h(\tau)$.

The first result gives a punctual condition of existence of a zero of F . It is a quantitative version of the classical Rouché's theorem using the fact that B_0 approximates $DF(x_0)^{-1}$ and α is less than a certain constant. This result does not depend on any numerical method. Moreover it brings out the dominating function $h(t)$ which plays a central role in the analysis of a Hald sequence.

Theorem 1.1 *Let $0 \leq \delta < 1$ and $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$. Then the strictly convex function $h(\tau)$ has two real roots $0 < \tau_- < \tau_+ < 1$. Suppose that F is analytic in the ball $B\left(x_0, \frac{1 - \delta}{\gamma(2 - \delta)}\right)$ and consider r such that $\tau_- < \gamma r < \tau_+$. Then F has an unique zero $\zeta \in B(x_0, r)$.*

This result generalizes the result given in Proposition 1.11 of [3] that treats the case $B_0 = DF(x_0)^{-1}$: the condition is $\alpha < 3 - 2\sqrt{2}$. We also remark that the use of Rouché's theorem permits to get the same result as in [15] without the study of the Newton sequence.

The second result gives a punctual sufficient condition for the Hald sequence to converge from x_0 . It is an α -theorem.

Theorem 1.2 *For each $0 \leq \delta < 1$ there exists $\bar{\alpha}_\delta$ and $q \in [0, 1[$ such that for all x_0 so that $\alpha \leq \bar{\alpha}_\delta$ we have $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$ and the Hald sequence $(x_k)_{k \geq 0}$ converges quadratically towards a zero ζ of F . More precisely we have for $k \geq 1$,*

$$\gamma \|x_k - \zeta\| \leq \tau_- - \tau_k \leq (\tau_+ - \alpha)q^{2^{k-1}},$$

and

$$\|I - B_k DF(x_k)\| \leq q^{2^{k-1}}.$$

Theorem 4.3 specifies the values of $\bar{\alpha}_\delta$ and q .

We now state a γ -theorem relatively to the Hald sequence. This result gives a ball $B(\zeta, r)$ in which the Hald sequence converges quadratically to ζ for any initial point $x_0 \in B(\zeta, r)$ and from an approximation B_0 of $DF(\zeta)^{-1}$.

Theorem 1.3 Let $z(s) = \frac{4 - 5s - 5s^2 + 4s^3 + 3s^4}{(1 - s - s^2)^2(1 - s)} s$ and $\bar{s} = 0.1741\dots$ such that $z(\bar{s}) = 1$.

Let ζ a zero of F such that $DF(\zeta)^{-1}$ exists. Denote $\gamma_\zeta = \gamma(f, DF(\zeta), \zeta)$ and $\bar{r} = 1 - \frac{\sqrt{2}}{2}$ and suppose that F is analytic in $B(\zeta, \bar{r})$. For all $x_0 \in B\left(\zeta, \frac{\bar{r}}{\gamma_\zeta}\right)$ such that

$$s = a + e := \|B_0 DF(\zeta) - I\| + \gamma_\zeta \|x_0 - \zeta\| < \bar{s} \quad (1.10)$$

then the map $DF(x_0)^{-1}$ exists and the Hald sequence converges quadratically to ζ . More precisely

$$e_k := \gamma_\zeta \|x_k - \zeta\| \leq z(s)^{2^k - 1} s, \quad \text{and,} \quad a_k := \|B_k DF(\zeta) - I\| \leq z(s)^{2^k - 1} s.$$

Moreover, we have

$$\delta_k := \|B_k DF(x_k) - I\| \leq 3z(s)^{2^k - 1} s.$$

Remark 1 It is well known that the quantity $\gamma(F, DF(z), z)^{-1}$ is a lower bound for the radius of convergence of the Taylor series of F in z , see for instance Proposition 6 page 167 of [1]. In the general case it is not easy to estimate the quantity γ . Fortunately this can be done for the polynomial systems of equations or for instance for Fredholm integral operator as in section 9.

Finally we state how Hald's method is related to Newton's method. More precisely we give a condition so that the iterates of Hald sequence and Newton's sequence are close.

Theorem 1.4 Let x_k (respectively τ_k) the k -th term of Hald sequence associated to F (respectively associated to h). We also denote by \bar{x}_k the Newton iterates starting from x_k . The condition

$$\frac{(1 - \tau_k)h(\tau_k)}{\psi(\tau_k)^2} < 3 - 2\sqrt{2}$$

implies that the Newton's sequence defined from x_k converges quadratically towards a root ζ of F . We also have :

$$\gamma \|x_{k+1} - \bar{x}_{k+1}\| \leq -\frac{h(\tau_k)}{h'(\tau_k)}.$$

In the sequel of paper we give in section 2 a short history and some fundamental works in our context. The proof of Theorem 1.1 is done in section 3. Next, in section 4 we study the behavior of Hald sequences associated to the dominating function $h(\tau)$. Theorems 1.2, 1.3 and 1.4 are proved in sections 5, 6 and 7 respectively. In section 8 we study the case of analytic functions such that $D^j F(x) \equiv 0$ for $j > 2$. Finally, section 9 is devoted to a numerical experiment with a classical Fredholm integral equation. We end with a concluding remark by proposing a study that will extend our result.

2 Short overview and discussion

The Hald sequence is a modified version of Moser's method [11] where $2B_k - B_k DF(x_{k+1})B_k$ replaces $2B_k - B_k DF(x_k)B_k$. This difference implies an order of convergence equal to $(1 + \sqrt{5})/2$ for Moser's method. It seems that Moser rediscovered this method first proposed by Ulm [14] who published his article in Russian. The equations $F(x) = 0$ considered comes from the study of stability of solutions of differential equations and/or partial differential equations. More recently in 2008, Ezquerro and Hernández [2] study Hald's method under mild differentiability conditions where there are more precisions and references to the non analytic case. The Kantorovich's classical theory [8] is the beginning of modern studies. Smale [12], drawing inspiration from this theory, had the idea that the complexity of algorithms for solving analytic equations could be expressed only from information given by the function and its derivatives at a point. Then the natural setting of α -theory is that of analytic functions in the ball of convergence of their Taylor series at a point [1]. In this way a criterion of convergence of Newton's method is given by $\alpha(F, DF(x_0)^{-1}, x_0) \leq 3 - 2\sqrt{2}$: this result of semi-local behaviour of Newton's method is named α -theorem. The local behaviour of Newton's method is described thanks the quantity $\gamma(F, DF(\zeta)^{-1}, \zeta)$. More precisely, the convergence of Newton sequence is quadratic for all initial point $x_0 \in B\left(\zeta, \frac{5 - \sqrt{17}}{2\gamma(F, DF(\zeta)^{-1}, \zeta)}\right)$: this result is named γ -theorem. It is the key point in the measure of complexity to find an initial point for Newton's method using an homotopy method [10].

To our knowledge, third papers have studied the analytic case with the goal to apply it to Fredholm integral equation using Hald's method. The first study [4] constructs a system of recurrence relations in order to prove the convergence of the method. The second study [6] uses the technique of dominating functions becoming from [16] but no result of quadratic convergence are presented. Note that the dominating function is of the type $h(\tau) = \alpha - (1 - \delta)\tau + \sum_{j \geq 2} \gamma_j \tau^j$ which well adapted to treat particular case but no universal constant to decide convergence is given. The third study [5] gives a γ -Theorem with the condition $d + 9e < 1/2$ which differs from (1.10). This condition links B_0 with $DF(x_0)^{-1}$ which is not the case in (1.10).

3 Proof of theorem 1.1

To prove the existence of a root of F we use Rouché's theorem. Let us consider the Taylor series of $B_0 F(x)$ in x_0 . We have $B_0 F(x) = B_0 F(x_0) + B_0 G(x)$ with

$$G(x) = \sum_{k \geq 1} \frac{1}{k!} D^k F(x_0) (x - x_0)^k.$$

Rouché's Theorem states that the analytic functions B_0F and B_0G have the same number of zeros counting the respective multiplicities in the ball $B(x_0, r)$ if for all $x \in \partial B(x_0, r)$ we have $\|F(x) - G(x)\| < \|G(x)\|$. Since B_0 is invertible the previous assertion holds with the inequality $\|B_0(F(x) - G(x))\| < \|B_0G(x)\|$.

We first prove that $B_0G(x)$ has only one zero in the ball $B\left(x_0, \frac{1-\delta}{\gamma(2-\delta)}\right)$. Let us consider another zero $y \neq x_0$ of B_0G . Since

$$\begin{aligned} \|B_0DF(x_0)(y - x_0)\| &\geq \|y - x_0\| - \|I - B_0DF(x_0)\| \|y - x_0\| \\ &\geq (1 - \delta) \|y - x_0\| \end{aligned} \quad (3.1)$$

we can write

$$\begin{aligned} \|B_0G(y)\| = 0 &\geq \left(1 - \delta - \sum_{k \geq 2} \frac{1}{k!} \|B_0D^k F(x_0)\| \|y - x_0\|^{k-1}\right) \|y - x_0\| \\ &\geq \left(1 - \delta - \frac{\gamma \|y - x_0\|}{1 - \gamma \|y - x_0\|}\right) \|y - x_0\| \\ &\geq \frac{1 - \delta - (2 - \delta)\gamma \|y - x_0\|}{1 - \gamma \|y - x_0\|} \|y - x_0\|. \end{aligned}$$

Hence $y \notin B\left(x_0, \frac{1-\delta}{\gamma(2-\delta)}\right)$ since $\gamma \|y - x_0\| \geq \frac{1-\delta}{2-\delta}$.

It is easy to see that this inequality $\|B_0(F(x) - G(x))\| < \|B_0G(x)\|$ is verified for all $x \in \partial B(x_0, r)$ if we have :

$$e := \|B_0F(x_0)\| - \|B_0DF(x_0)(x - x_0)\| + \sum_{k \geq 2} \frac{1}{k!} \|B_0D^k F(x_0)\| \|x - x_0\|^k < 0.$$

Using (3.1) substituting y by x and $\|x - x_0\| = r$, we have

$$\begin{aligned} \frac{1}{\gamma} h(\gamma r) &\geq \beta - (1 - \delta)r + \sum_{k \geq 2} (\gamma r)^{k-1} r \\ &\geq e. \end{aligned}$$

Hence the condition $h(\gamma r) < 0$ implies $e < 0$. From Lemma 4.1 below the function $h(\tau)$ is strictly convex for $\tau \in [0, 1[$ and has two distinct real roots satisfying $0 < \tau_- < \tau_+ < \frac{1-\delta}{2-\delta}$ under the conditions $0 < \delta < 1$ and $\alpha < 3 - \delta - 2\sqrt{2-\delta}$. The inequalities $\tau_- < \gamma r < \tau_+$ imply $h(\gamma r) < 0$. This proves Theorem 1.1.

4 The behavior of a Hald sequence associated to $h(\tau)$

The condition of existence of zeros of $h(\tau)$ is given by the following lemma :

Lemma 4.1 *Let $0 \leq \delta < 1$. Under the condition $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$ the strictly convex function $h(\tau)$ has two real roots*

$$\tau_{\pm} = \frac{\alpha + 1 - \delta \pm \sqrt{d}}{2(2 - \delta)} \quad (4.1)$$

where d is defined by (1.9) and moreover

$$\tau_- < \tau_+ < \frac{1 - \delta}{2 - \delta}.$$

Proof. Under the condition $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$ the computation of roots of $h(t)$ is straightforward and the strict convexity is easy to show. The proof $\tau_+ < \frac{1 - \delta}{2 - \delta}$ reduces to $1 - \alpha - \delta - \sqrt{d} > 0$. It is easy to see that it suffices that $1 - \alpha - \delta > 0$. The condition $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$ implies

$$\begin{aligned} 1 - \alpha - \delta &> 2\sqrt{2 - \delta} - 2 \\ &> 0, \quad \text{since } 2\sqrt{2 - \delta} - 2 > 0 \text{ for } \delta < 1. \end{aligned}$$

We are done. □

Lemma 4.2 *Suppose $0 \leq \delta < 1$ and $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$. Hald sequence (1.6) associated to $h(\tau)$ converges towards τ_- with the following properties for $k \geq 0$:*

1. $b_k < 0$.
2. $0 < b_k h'(\tau_k) < 1$. Consequently $-b_k < -\frac{1}{h'(\tau_k)}$.
3. $0 \leq \tau_k < \tau_{k+1} < \tau_-$.

Then, the sequence $(\delta_k)_{k \geq 0}$ defined by (1.7) satisfies

$$\delta_{k+1} \leq \delta_k < 1.$$

Proof. The proof is inspired by [6]. From Lemma 4.1 $h(\tau)$ has two roots $0 < \tau_- < \tau_+ < 1$. Prove items 1 to 3 inductively. We have $b_0 = -1$, $b_0 h'(0) = 1 - \delta < 1$ and $0 = \tau_0 < \tau_1 = \alpha$. Prove that $\alpha < \tau^-$. We have

$$\tau_1 = \alpha \leq -\frac{h(0)}{h'(0)} = \frac{\alpha}{1 - \delta} = \bar{\tau}_1.$$

The real $\bar{\tau}_1$ is one step of Newton's Method associated to $h(\tau)$ from 0. In other words, the point $(\bar{\tau}_1, 0)$ is the intersection of the x -axis and of the tangent line at point $(0, \alpha)$ to graph of $h(\tau)$. The strict convexity of $h(\tau)$ ensures $\alpha < \tau_-$.

Proof of items 1 to 3 for $k + 1$ if it is true for k . Clearly, since $h(\tau)$ decreases and is strictly

convex on $[0, \tau_-]$ we can see that $\tau_k \leq \tau_{k+1}$ implies $1 - b_k h'(\tau_{k+1}) > 1 - b_k h'(\tau_k) > 0$. We then get

$$b_{k+1} = b_k + (1 - b_k h'(\tau_{k+1}))b_k < b_k < 0.$$

Since $\tau_{k+1} < \tau_-$ one has $h'(\tau_{k+1}) < 0$ and $b_{k+1} h'(\tau_{k+1}) > 0$.

Moreover $b_{k+1} h'(\tau_{k+1}) - 1 = 2b_k h'(\tau_{k+1}) - b_k^2 h'(\tau_{k+1})^2 - 1 = -(1 - b_k h'(\tau_{k+1}))^2 < 0$. Item 3 follows from $\tau_{k+2} - \tau_{k+1} = -b_{k+1} h(\tau_{k+1}) > 0$ and also

$$\tau_- - \tau_{k+2} = (1 - b_{k+1} h'(\lambda_{k+1}))(\tau_- - \tau_{k+1}) > 0.$$

for $\lambda_{k+1} \in]\tau_{k+1}, \tau_{k+2}[$. The inequalities $b_{k+1} h'(\lambda_{k+1}) < b_{k+1} h'(\tau_{k+1})$ permit to conclude $\tau_- - \tau_{k+2} > 0$. Then items 1 to 3 hold.

Now since $b_{k+1} < b_k < 0$ and $h'(\tau_{k+1}) < h'(\tau_k) < 0$ it follows that $\delta_{k+1} = 1 - b_{k+1} h'(\tau_{k+1}) < \delta_k = 1 - b_k h'(\tau_k)$.

This lemma is proved. \square

Next, we will use the following quantities :

$$\begin{aligned} q_0 &= \frac{\tau_-}{\tau_+} \text{ and } q_1 = \frac{1 - \tau_+}{1 - \tau_- - \delta q_0} \left(q_0^2 + \frac{1}{1 - \tau_+} \delta q_0 \right) \\ \delta_1 &= 1 - b_1 h'(\alpha) = \left(\delta + \frac{(2 - \alpha)\alpha}{(1 - \alpha)^2} \right)^2 \\ q &= \delta_1 + q_1 \\ \eta &= \frac{1 - \tau_+}{1 - \tau_- - (1 - \alpha)\delta_1 q_1} + (\tau_+ - \alpha)^2 \end{aligned}$$

Theorem 4.3 *For each $\delta < 1$ there exists $\bar{\alpha}_\delta$ root of $\eta - 1$ such that for all $\alpha \leq \bar{\alpha}_\delta$ we have $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$, $q < 1$ and $\eta \leq 1$. Moreover for $\alpha \leq \bar{\alpha}_\delta$ we have :*

$$\delta_k := 1 - b_k h'(\tau_k) \leq q^{2^{k-1}} \text{ and } \tau_- - \tau_k \leq (\tau_+ - \alpha)q^{2^{k-1}}, \quad k \geq 1. \quad (4.2)$$

We give some values of $\bar{\alpha}_\delta$:

δ	0	0.2	0.4	0.6	0.8	1
$\bar{\alpha}_\delta$	0.165	0.106	0.06	0.022	0.004	0

Proof. Writing $h(\tau) = \frac{(2 - \delta)(\tau - \tau_-)(\tau - \tau_+)}{1 - \tau}$ a straightforward calculation shows that

$$\begin{aligned} \tau_- - \tau_{k+1} &= \tau_- - \tau_k + b_k h(\tau_k) \\ &= \tau_- - \tau_k + \frac{h(\tau_k)}{h'(\tau_k)} - (1 - b_k h'(\tau_k)) \frac{h(\tau_k)}{h'(\tau_k)} \\ &= \frac{(2 - \delta)(1 - \tau_+)}{\psi(\tau_k)} (\tau_k - \tau_-)^2 - \delta_k \frac{h(\tau_k)}{h'(\tau_k)}. \end{aligned} \quad (4.3)$$

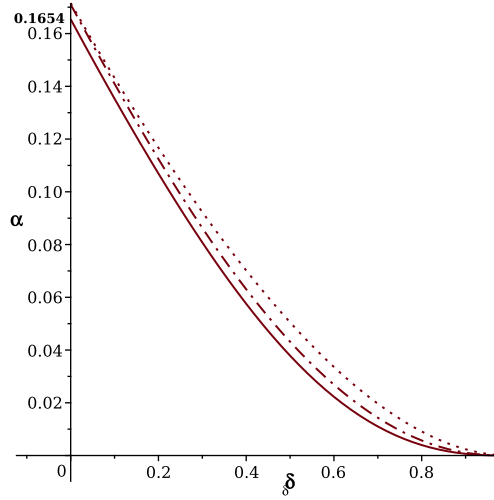


Figure 1: Theorem 4.3.

Curve $-$: $\eta = 1$.

Curve $-.-$: $q = 1$.

Curve \dots : $\alpha = 3 - \delta - 2\sqrt{2 - \delta}$.

In the same way we get :

$$\tau_+ - \tau_{k+1} = \frac{(2 - \delta)(1 - \tau_-)}{\psi(\tau_k)} (\tau_k - \tau_+)^2 - \delta_k \frac{h(\tau_k)}{h'(\tau_k)}. \quad (4.4)$$

Remember that $h'(\tau) = \frac{\psi(\tau)}{(1 - \tau)^2} < 0$ for $\tau \in [0, \tau_-]$. Let $q_k = \frac{\tau_k - \tau_-}{\tau_k - \tau_+}$, $k \geq 0$. It is easy to see that the function $\tau \in [0, \tau_-] \rightarrow \frac{\tau - \tau_-}{\tau - \tau_+}$ strictly decreases. From Lemma 4.2 we know $\tau_k < \tau_{k+1}$. Hence $q_{k+1} < q_k$ from which we get the bound

$$\begin{aligned} - \left(\frac{(2 - \delta)(1 - \tau_-)}{\psi(\tau_k)} (\tau_k - \tau_+)^2 \right)^{-1} \delta_k \frac{h(\tau_k)}{h'(\tau_k)} &\leq \frac{\delta_k (\tau_- - \tau_k)(1 - \tau_k)}{(\tau_+ - \tau_k)(1 - \tau_-)} \\ &\leq \frac{1 - \tau_k}{1 - \tau_-} \delta_k q_k \end{aligned} \quad (4.5)$$

reduces to the lower bound

$$\tau_+ - \tau_{k+1} \geq \frac{(2 - \delta)(1 - \tau_-)}{\psi(\tau_k)} (\tau_+ - \tau_k)^2 \left(1 - \frac{1 - \tau_k}{1 - \tau_-} \delta_k q_k \right). \quad (4.6)$$

Using equality (4.3) and bounds (4.5), (4.6) we get with $q_0 = \frac{\tau_-}{\tau_+}$:

$$\begin{aligned} q_1 &\leq \frac{1}{1 - \frac{1}{1 - \tau_-} \delta q_0} \left(\frac{1 - \tau_+}{1 - \tau_-} q_0^2 + \frac{1}{1 - \tau_-} \delta q_0 \right) \\ &\leq \frac{1 - \tau_+}{1 - \tau_- - \delta q_0} \left(q_0^2 + \frac{1}{1 - \tau_+} \delta q_0 \right). \end{aligned} \quad (4.7)$$

and for $k \geq 1$:

$$\begin{aligned} q_{k+1} &\leq \frac{1}{1 - \frac{1-\alpha}{1-\tau_-} \delta_k q_k} \left(\frac{1-\tau_+}{1-\tau_-} q_k^2 + \frac{1-\alpha}{1-\tau_-} \delta_k q_k \right), \quad \text{since } \tau_1 = \alpha, \\ &\leq \frac{1-\tau_+}{1-\tau_- - (1-\alpha)\delta_1 q_1} \left(q_k^2 + \frac{1-\alpha}{1-\tau_+} \delta_k q_k \right), \end{aligned} \quad (4.8)$$

since $q_k < q_1$ and from Lemma 4.2 $\delta_k < \delta_1$. On another side

$$\delta_{k+1} := 1 - b_{k+1} h'(\tau_{k+1}) = (1 - b_k h'(\tau_{k+1}))^2.$$

One has $0 < 1 - b_k h'(\tau_k) = \delta_k < 1 - b_k h'(\tau_{k+1}) = \delta_k - b_k h'(\lambda_k)(\tau_{k+1} - \tau_k)$ where $\lambda_k \in]\tau_k, \tau_{k+1}[$.

Hence

$$\begin{aligned} \delta_{k+1} &\leq (\delta_k - b_k h'(\lambda_k)(\tau_{k+1} - \tau_k))^2 \\ &\leq \delta_k^2 + (\tau_{k+1} - \tau_k)^2, \quad \text{since } b_k h'(\lambda_k) > 0. \end{aligned}$$

We then get

$$\begin{aligned} \delta_1 &= 1 - b_1 h'(\alpha) \\ &= \left(\delta + \frac{(2-\alpha)\alpha}{(1-\alpha)^2} \right)^2, \end{aligned} \quad (4.9)$$

and for $k \geq 1$

$$\begin{aligned} \delta_{k+1} &\leq \delta_k^2 + (\tau_- - \tau_k)^2 \\ &\leq \delta_k^2 + q_k^2 (\tau_+ - \tau_k)^2 \\ &\leq \delta_k^2 + (\tau_+ - \alpha)^2 q_k^2. \end{aligned} \quad (4.10)$$

Let $\eta = \frac{1-\tau_+}{1-\tau_- - (1-\alpha)\delta_1 q_1} + (\tau_+ - \alpha)^2$, $\mu = \frac{1-\alpha}{2(1-\tau_+)}(\eta - (\tau_+ - \alpha)^2)$ and $q = q_1 + \delta_1$. We have for $\alpha \leq \bar{\alpha}_\delta$ root of $\eta - 1$:

$$\begin{aligned} \delta_{k+1} + q_{k+1} &\leq \eta q_k^2 + 2\mu \delta_k q_k + \delta_k^2 \\ &\leq (\delta_k + q_k)^2, \quad \text{since from Lemma 4.4 we have } \mu < \eta \leq 1 \text{ for } \alpha \leq \bar{\alpha}_\delta. \end{aligned}$$

By induction we then derive that for $k \geq 1$, $\delta_k + q_k = q^{2^{k-1}}$ using $q < 1$ from Lemma 4.4.

We then deduce

$$\delta_k \leq q^{2^{k-1}} \text{ and } \tau_- - \tau_k \leq (\tau_+ - \alpha) q^{2^{k-1}}, \quad k \geq 1.$$

□

Lemma 4.4 *Let q, η be defined as in Theorem 4.3 and*

$$\mu = \frac{1 - \alpha}{2(1 - \tau_+)}(\eta - (\tau_+ - \alpha)^2).$$

Then we have

1. $\eta > \mu$.
2. *For each $\delta < 1$ there exists $\bar{\alpha}_\delta$ such that $\eta - 1 = 0$. Moreover for all $\alpha \leq \bar{\alpha}_\delta$ we have $\alpha < 3 - \delta - 2\sqrt{2 - \delta}$, $q < 1$ and $\eta \leq 1$.*

Proof. Let d defined by (1.9). We know $\tau_+ = \frac{\alpha + 1 - \delta + \sqrt{d}}{2(2 - \delta)}$ from Lemma 4.1. Then $\eta - \mu > \frac{\mu}{1 - \alpha}(1 + \alpha - 2\tau_+)$. Since $(1 + \alpha - \alpha\delta)^2 - d = (2 - \delta)(4\alpha + \delta - \alpha^2\delta) > 0$ we have $1 + \alpha - 2\tau_+ > 0$. This proves $\eta > \mu$.

To prove item 2 we remark for $\delta \in [0, 1[$ be given the function $\alpha \rightarrow \tau_-$ increases and the function $\alpha \rightarrow \tau_+$ decreases. From this we can show that the functions $\alpha \rightarrow q$ and $\alpha \rightarrow \eta$ increase. When $\alpha = 0$ a direct calculation shows that $q - 1 = \delta - 1 < 0$. When $\alpha = 0.999 \times (3 - \delta - 2\sqrt{2 - \delta})$ a numerical calculation shows that $q - 1 > 0$: see Figure 2. Hence there exists an implicit function $\alpha_1(\delta)$ such that $q - 1 = 0$.

When $\alpha = 0$ and $\delta \in [0, 1[$ be given we have $\eta - 1 = -\frac{1 - \delta}{(2 - \delta)^2} < 0$. On another side when $\alpha = \alpha_1(\delta)$ a numerical calculation shows that $\eta - 1 > 0$: see Figure 3. Hence there exists an implicit function $\alpha_2(\delta)$ such that $\eta - 1 = 0$. Finally we have $\alpha_2(\delta) \leq \alpha_1(\delta) \leq 3 - \delta - 2\sqrt{2 - \delta}$ as it is shown in Figures 1 and 4. Then it is sufficient to consider $\bar{\alpha}_\delta := \alpha_2(\delta)$ and the result follows. \square

5 Proof of α -Theorem 1.2

Theorem 1.2 follows from Theorem 5.1 below and Theorem 4.3.

Theorem 5.1 *If $\alpha \leq 3 - \delta - 2\sqrt{2 - \delta}$ then the scalar Hald sequence defined by (1.6) dominates the Hald sequence $(x_k)_{k \geq 0}$ defined by (1.2), that is,*

$$\gamma \|x_{k+1} - x_k\| \leq \tau_{k+1} - \tau_k, \quad k \geq 0.$$

Then the sequence $(x_k)_{k \geq 0}$ converges towards a zero ζ of F with

$$\gamma \|x_k - \zeta\| \leq \tau_- - \tau_k, \quad k \geq 0.$$

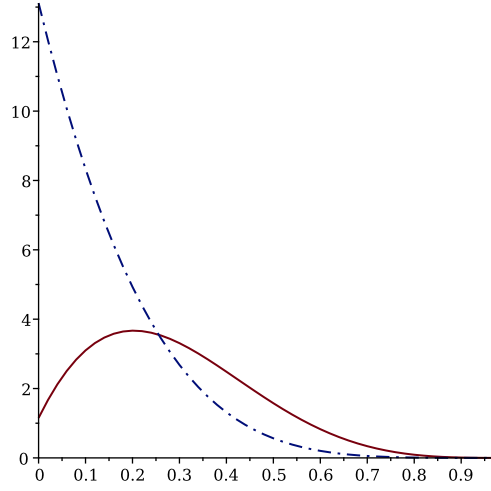


Figure 2: Lemma 4.4.

- : numerator of $q - 1$ when $\alpha = 0.999 \times (3 - \delta - 2\sqrt{2 - \delta})$.
 -.- : denominator of $q - 1$ when $\alpha = 0.999 \times (3 - \delta - 2\sqrt{2 - \delta})$.

Proof. We adapt the proof of Theorem 2 of [6] to our context. We proceed by mathematical induction proving that the following inequalities hold for all $k \geq 0$:

$$\delta_k := \|I - B_k F(x_k)\| \leq 1 - b_k h'(\tau_k) \quad (5.1)$$

$$\gamma \|B_k F(x_k)\| \leq -b_k h(\tau_k) \quad (5.2)$$

$$\|B_k D^{(j)} F(x_k)\| \leq -b_k \gamma^{j-1} h^{(j)}(\tau_k). \quad (5.3)$$

For $k = 0$ we have simultaneously $\delta = 1 + h'(0)$, $\gamma \|x_1 - x_0\| = \alpha$ and $\gamma(F, B_0, x_0) \leq -\gamma b_0 \gamma(h, b_0, 0) = \gamma$. In fact for $j \geq 2$ we have $h^{(j)}(\tau) = \frac{j!}{(1 - \tau)^{j+1}}$ and $|h^{(j)}(0)| = j!$.

Let us assume $\gamma \|x_{j+1} - x_j\| \leq \tau_{j+1} - \tau_j$ and that these inequalities hold for $j \leq k$. We prove they hold for $k + 1$.

We know that

$$\begin{aligned} I - B_{k+1} DF(x_{k+1}) &= (I - B_k DF(x_{k+1}))^2 \\ I - B_k DF(x_{k+1}) &= I - B_k DF(x_k) - \sum_{j \geq 2} \frac{j}{j!} B_k D^{(j)} F(x_k) (x_{k+1} - x_k)^{j-1} \end{aligned}$$

It follows

$$\begin{aligned} \|I - B_k DF(x_{k+1})\| &\leq 1 - b_k h'(\tau_k) - \sum_{j \geq 2} \frac{j}{j!} b_k h^{(j)}(\tau_k) (\gamma \|x_{k+1} - x_k\|)^{j-1} \\ &\leq 1 - b_k h'(\tau_{k+1}). \end{aligned} \quad (5.4)$$

Hence the inequality (5.1) holds since

$$\|I - B_{k+1} DF(x_{k+1})\| \leq (1 - b_k h'(\tau_{k+1}))^2 = 1 - b_{k+1} h'(\tau_{k+1}).$$

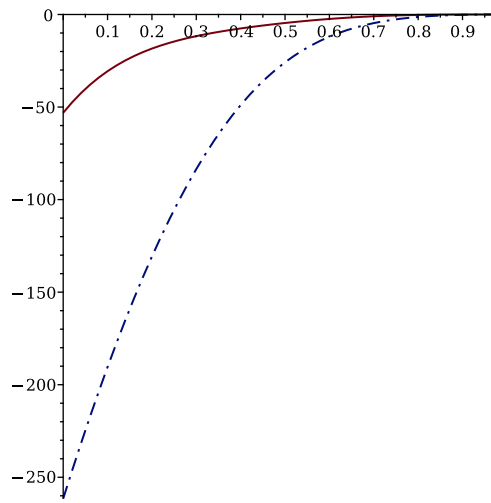


Figure 3: Lemma 4.4.

— : numerator of $\eta - 1$ when $\alpha = \alpha_1(\delta)$
 -.- : denominator of $\eta - 1$ when $\alpha = \alpha_1(\delta)$.

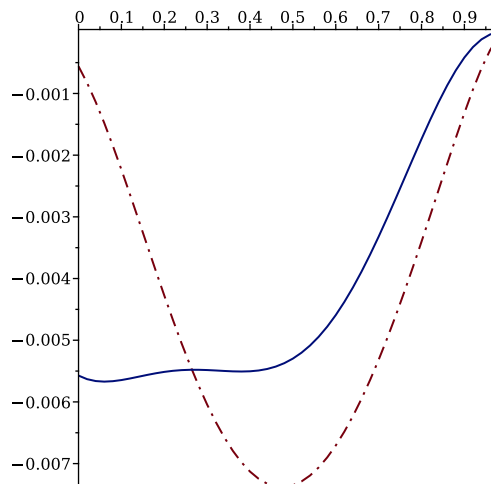


Figure 4: Lemma 4.4.

— : $\bar{\alpha}_\delta - \alpha_1(\delta)$.
 -.- : $\alpha_1(\delta) - (3 - \delta - 2\sqrt{2 - \delta})$.

Next, since $B_k F(x_k) + B_k DF(x_k)(x_{k+1} - x_k) = -(I - B_k DF(x_k))(x_{k+1} - x_k)$ we get

$$\gamma B_k F(x_{k+1}) = (I - B_k DF(x_k))\gamma B_k F(x_k) + \sum_{j \geq 2} \frac{1}{j!} B_k D^{(j)} F(x_k) \gamma (x_{k+1} - x_k)^j.$$

We deduce

$$\begin{aligned} \gamma \|B_k F(x_{k+1})\| &\leq -(1 - b_k h'(\tau_k)) b_k h(\tau_k) - \sum_{j \geq 2} \frac{1}{j!} b_k h^{(j)}(\tau_k) (\tau_{k+1} - \tau_k)^j \\ &\leq -b_k h(\tau_{k+1}). \end{aligned}$$

Then using $B_{k+1} F(x_{k+1}) = (2I - B_k F(x_{k+1})) B_k F(x_{k+1})$ the inequality (5.2) follows from

$$\begin{aligned} \gamma \|B_{k+1} F(x_{k+1})\| &\leq -(2 - b_k h'(\tau_{k+1})) b_k h(\tau_{k+1}) \\ &\leq -b_{k+1} h(\tau_{k+1}). \end{aligned}$$

From this, it follows $\gamma \|x_{k+2} - x_{k+1}\| \leq \tau_{k+2} - \tau_{k+1}$.

Thirdly, inequality (5.3) is satisfied since for $j \geq 2$ we have

$$\begin{aligned} \|B_{k+1} D^{(j)} F(x_{k+1})\| &\leq (2 - b_k h'(\tau_k)) \sum_{l \geq 0} \frac{1}{l!} (-b_k h^{(l+j)}(\tau_k)) \gamma^{j-1} (\tau_{k+1} - \tau_k)^l \\ &\leq -\gamma^{j-1} (2 - b_k h'(\tau_k)) b_k h^{(j)}(\tau_{k+1}) = -\gamma^{j-1} b_{k+1} h^{(j)}(\tau_{k+1}). \end{aligned}$$

From Lemma 4.2 the scalar Hald sequence strictly increases and converges to τ_- . This implies the convergence of the sequence $(x_k)_{k \geq 0}$. Theorem 5.1 follows easily. \square

6 Proof of γ -theorem 1.3

Here we have $F(\zeta) = 0$ and $\gamma_\zeta = \gamma(F, DF(\zeta)^{-1}, \zeta)$. Let $s_k = e_k + a_k$ where $e_k = \gamma_\zeta \|x_k - \zeta\|$ and $a_k = \|B_k DF(\zeta) - I\|$. We denote $e = e_0$, $a = a_0$ and $s = s_0$.

To prove the convergence of Hald sequence from x_0 , we proceed by mathematical induction assuming $s_k \leq z(s)^{2^k - 1} s$ for a given index k . For $k = 0$, evidently $s_0 \leq s$, but we need to prove that $DF(x_0)$ is an invertible map. Lemma 2 (b) page 156 of [1] states the map $DF(x_0)$ is invertible for all $x_0 \in B\left(\zeta, \frac{1}{\gamma_\zeta} \bar{r}\right)$ with $\bar{r} = (1 - \sqrt{2})/2$. Consequently there exists linear map B_0 which approximates $DF(x_0)^{-1}$. For instance $B_0 = DF(x_0)^{-1} + \lambda I$ with $\|\lambda DF(x_0)\| < 1$ satisfies $\|B_0 DF(x_0) - I\| < 1$.

We then write for $k \geq 0$:

$$\begin{aligned}
 e_{k+1} &= \gamma_\zeta \|x_k - \zeta - B_k F(x_k)\| \\
 &\leq \|I - B_k DF(\zeta) - B_k DF(\zeta) \sum_{j \geq 2} \frac{1}{j!} DF(\zeta)^{-1} D^j F(\zeta) (x_k - \zeta)^{j-1}\| e_k \\
 &\leq \left(\|I - B_k DF(\zeta)\| + \|B_k DF(\zeta)\| \sum_{j \geq 2} e_k^{j-1} \right) e_k \\
 &\leq \left(a_k + (1 + a_k) \frac{e_k}{1 - e_k} \right) e_k \\
 &\leq \frac{(a_k + e_k) e_k}{1 - e_k} \\
 &\leq \frac{s_k^2}{1 - s_k}.
 \end{aligned} \tag{6.1}$$

On another side a straightforward calculation gives

$$\begin{aligned}
 B_{k+1} DF(\zeta) - I &= B_k DF(\zeta) - I + (I - B_k DF(x_{k+1})) B_k DF(\zeta) \\
 &= -(B_k DF(\zeta) - I)^2 - \sum_{j \geq 2} \frac{j}{j!} B_k D^j F(\zeta) (x_{k+1} - \zeta)^{j-1} B_k DF(\zeta) \\
 &= -(B_k DF(\zeta) - I)^2 - E
 \end{aligned}$$

with $E = -B_k DF(\zeta) \sum_{j \geq 2} \frac{j}{j!} DF(\zeta)^{-1} D^j F(\zeta) (x_{k+1} - \zeta)^{j-1} B_k DF(\zeta)$.

We deduce using (6.1) :

$$\begin{aligned}
 a_{k+1} &\leq a_k^2 + (1 + a_k)^2 \frac{(2 - e_{k+1}) e_{k+1}}{(1 - e_{k+1})^2} \\
 &\leq \frac{3 - 4s_k^2 - 2s_k^3}{(1 - s_k - s_k^2)^2} s_k^2.
 \end{aligned} \tag{6.2}$$

From (6.1) and (6.2), a straightforward calculation shows that

$$s_{k+1} \leq \frac{4 - 5s_k - 5s_k^2 + 4s_k^3 + 3s_k^4}{(1 - s_k - s_k^2)^2 (1 - s_k)} s_k^2 := z(s_k) s_k.$$

Under the condition $s < \bar{s} = 0.1741\dots$ which implies $z(s) < 1$, we easily deduce with $s_k \leq z(s)^{2^k - 1} s$ that $s_{k+1} \leq z(s)^{2^{k+1} - 1} s$.

On another side we have from the fundamental property (1.5) of $B'_k s$

$$B_{k+1} DF(x_{k+1}) - I = -(B_k DF(x_{k+1}) - I)^2. \tag{6.3}$$

Moreover

$$B_k DF(x_{k+1}) - I = B_k DF(\zeta) - I + B_k DF(\zeta) \sum_{j \geq 2} \frac{j}{j!} DF(\zeta)^{-1} D^j F(\zeta) (x_{k+1} - \zeta)^{j-1}. \tag{6.4}$$

Combining (6.3) and (6.4) we get

$$d_{k+1} \leq \left(a_k + (1 + a_k) \frac{(2 - e_{k+1})e_{k+1}}{(1 - e_{k+1})^2} \right)^2. \quad (6.5)$$

Using $a_k, e_k \leq z(s)^{2^k-1}s$, it follows

$$\begin{aligned} d_{k+1} &\leq \left(1 + (1 + s) \frac{(2 - z(s)s)z(s)}{(1 - z(s)s)^2} \right)^2 z(s)^{2^{k+1}-2}s^2 \\ &\leq \left(1 + (1 + \bar{s}) \frac{2 - \bar{s}}{(1 - \bar{s})^2} \right)^2 \bar{s} z(s)^{2^{k+1}-1}s \\ &\leq 3z(s)^{2^{k+1}-1}s, \end{aligned}$$

since the function $s \in [0, \bar{s}] \rightarrow \left(1 + (1 + s) \frac{(2 - z(s)s)z(s)}{(1 - z(s)s)^2} \right)^2 z(s)^{-1}s$ increases. Then the theorem holds.

7 Proof of theorem 1.4

We use the following α -theorem [15] which established the Newton sequence is well defined and converges quadratically from an initial point x_0 provided that $\alpha(F, DF(x_0), x_0) < 3 - 2\sqrt{2}$. We consider the $(k + 1)$ -th step in the iteration. We remark

$$x_{k+1} - ((B_k DF(x_k))^{-1} - I)B_k F(x_k) = x_k - DF(x_k)^{-1}F(x_k). \quad (7.1)$$

We denote by \bar{x}_{k+1} the left hand side of (7.1) that is the result of one step of Newton's method. We now bound $\alpha(F, DF(x_k), x_k)$ with respect to $\alpha(F, B_k, x_k)$. We first have

$$DF(x_k)^{-1}F(x_k) = (B_k DF(x_k))^{-1}B_k F(x_k) \quad (7.2)$$

$$= (I - (I - B_k DF(x_k)))^{-1}B_k F(x_k). \quad (7.3)$$

Using inequality (5.4) of the proof of Theorem 1.2 we find

$$\beta(F, DF(x_k), x_k) \leq (b_k h'(\tau_k))^{-1} \beta(F, B_k, x_k) \quad (7.4)$$

In the same way we get

$$\begin{aligned} \gamma(F, DF(x_k), x_k) &\leq \|(B_k DF(x_k))^{-1}\| \gamma(F, B_k, x_k) \\ &\leq (b_k h'(\tau_k))^{-1} \gamma(F, B_k, x_k). \end{aligned} \quad (7.5)$$

Hence

$$\alpha(F, DF(x_k), x_k) \leq (b_k h'(\tau_k))^{-2} \alpha(F, B_k, x_k) \quad (7.6)$$

Always from (5.2) and (5.3) we get

$$\alpha(F, B_k, x_k) \leq b_k^2 \frac{h(\tau_k)}{(1 - \tau_k)^3}.$$

Hence the condition

$$(b_k h'(\tau_k))^{-2} b_k^2 \frac{h(\tau_k)}{(1 - \tau_k)^3} = \frac{(1 - \tau_k)h(\tau_k)}{\psi(\tau_k)^2} < 3 - 2\sqrt{2} \quad (7.7)$$

implies that the Newton's sequence defined from x_k converges quadratically towards a root ζ of F . From (7.1) we get

$$\gamma \|x_{k+1} - \bar{x}_{k+1}\| \leq -\frac{h(\tau_k)}{h'(\tau_k)}. \quad (7.8)$$

We are done.

8 Study of a particular case

We give without proof the particularities of previous results when $D^{(j)}F(x) \equiv 0$ for $j > 2$. In this way the dominating function becomes

$$h(\tau) = \alpha - (1 - \delta)\tau + \tau^2.$$

with zeros $\tau_{\pm} = \frac{1}{2} (1 - \delta \pm \sqrt{d})$ where $d = \sqrt{(1 - \delta)^2 - 4\alpha}$.

We have successively

1. Theorems 1.1 and 1.2 hold with $\delta < 1$ and $\alpha < \frac{1}{4}(1 - \delta)^2$.

2. Theorem 4.3 holds with :

$$(a) \quad q_0 = \tau_- / \tau_+ \text{ and } q_1 = \frac{1 - \tau_+}{1 - \delta q_0} (q_0^2 + \frac{1}{1 - \tau_+} \delta q_0) \text{ and}$$

$$\delta_1 = \left(\delta + \frac{(2 - \alpha)\alpha}{(1 - \alpha)^2} \right)^2.$$

$$(b) \quad q = \delta_1 + q_1$$

$$(c) \quad \eta = \frac{1 - \tau_+}{1 - \tau_- - (1 - \alpha)\delta_1 q_1} + (\tau_+ - \alpha)^2.$$

$$(d) \quad \alpha < \frac{1}{4}(1 - \delta)^2.$$

$$(e) \quad \bar{\alpha}_\delta = \alpha_2(\delta) \text{ root of } \eta - 1.$$

We also give some values of $\bar{\alpha}_\delta$:

δ	0	0.2	0.4	0.6	0.8	1
$\bar{\alpha}_\delta$	0.241	0.148	0.077	0.03	0.005	0

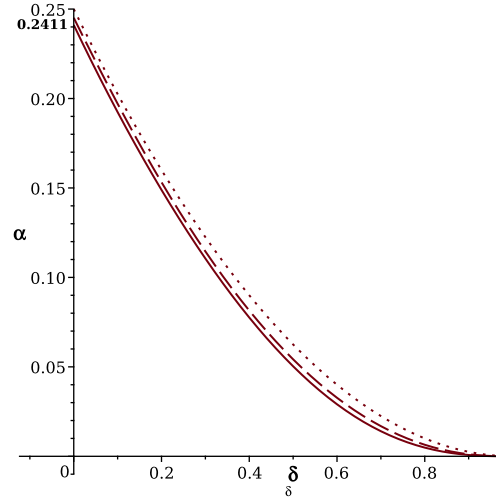


Figure 5: Particular case.

Curve $-$: $\eta = 1$.Curve $--$: $q = 1$.Curve \dots : $\alpha = 3 - \delta - 2\sqrt{2 - \delta}$.

3. Theorem 1.3 holds with $\bar{r} = \frac{1}{2}$, $z(s) = (4s + 3)(2s^2 + 1)s$ and $\bar{s} = 0.23\dots$ solution of $z(s) - 1 = 0$.

4. Theorem 1.4 holds under the condition $\frac{h(\tau_k)}{h'(\tau_k)^2} < \frac{1}{4}$.

Figure 5 shows respectively the curves $\eta - 1 = 0$, $q - 1 = 0$ and $4\alpha - (1 - \delta)^2 = 0$. Figure 6 illustrates numerically that $\bar{\alpha}_\delta = \alpha_2(\delta) < \alpha_1(\delta) < 3 - \delta - 2\sqrt{2 - \delta}$.

9 Numerical experiments

We illustrate the previous results considering the following Fredholm integral operator

$$F(x)(s) = x(s) - 1 - \frac{1}{2} \int_0^1 \sin(st)x(t)^2 dt \quad (9.1)$$

which appears in [9] page 552. We are proving that the following two functions

$$x_{1,0}(s) = 1 \quad (9.2)$$

$$x_{2,0}(s) = 1 + \frac{1}{2} \int_0^1 \sin(st)dt = 1 + \frac{1 - \cos(s)}{2s} \quad (9.3)$$

are close to a solution of (9.1). We use the classical max norm in the space of continuous functions. A straightforward calculation shows that with $B_0 = I$

$$1 - (I - B_0 DF(x))y(s) = \int_0^1 \sin(st)x(t)y(t)dt.$$

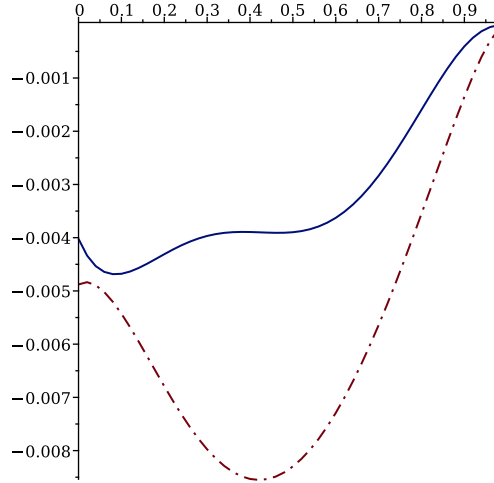


Figure 6: Lemma 4.4.

— : $\alpha_2(\delta) - \alpha_1(\delta)$.
 -.- : $\alpha_1(\delta) - (3 - \delta - 2\sqrt{2 - \delta})$.

$$2- \frac{1}{2}B_0D^2F(x)y^2(s) = -\frac{1}{2} \int_0^1 \sin(st)y(t)^2 dt.$$

Since the maximum of $\int_0^1 \sin(st)dt = \frac{1 - \cos(s)}{s}$ on $[0, 1]$ is obtained for $s = 1$ we can numerically compute the quantities $\beta_i, \delta_i, \gamma_i, \alpha_i, q_i, \eta_i, \eta_i q_i$ corresponding to $x_{1,0}$ and $x_{2,0}$ respectively. The results are given in Table 1. From Theorem 1.1 we deduce the existence of

	β_i	δ_i	γ_i	$\alpha_i < (1 - \delta_i)^2/4?$	η_i	q_i	$\eta_i q_i < 1?$
$x_{1,0}$	0.23	0.46	0.23	$0.05 < 0.07$	1.29	0.49	$0.64 < 1$
$x_{2,0}$	0.08	0.27	0.11	$0.009 < 0.13$	1.51	0.08	$0.11 < 1$

Table 1: Quantities $\beta_i, \delta_i, \gamma_i, \alpha_i, q, \eta_i, \eta_i q_i$ relatively to $x_{1,0}$ and $x_{2,0}$.

only one solution in the closed ball $B(x_{1,0}, \tau_{-,i}/\gamma_i)$, $i = 1, 2$ where the $\tau_{-,i}$'s are respectively the first positive root of $h_i(\tau) = \alpha_i - (1 - \delta_i)\tau + \tau^2$. Here we have $\tau_{-,1}/\gamma_1 = 0.56$ and $\tau_{-,2}/\gamma_2 = 0.11$.

We perform the computation of Hald sequence in the space of functions defined by

$$\left\{1 + a_1 \frac{1 - \cos(s)}{s} + \sum_{i=1}^{n-1} a_{i+1} s^i \quad \text{where } (a_1, \dots, a_n) \in \mathbb{R}^n \right\}.$$

We take here $n = 5$. Then the equation $F(x)(s) = 0$ reduces to a polynomial system whose solution approximates that of (9.1). Finally, Hald sequence from $x_{1,0}$ or $x_{2,0}$ associated at this polynomial system converges to the function

$$\zeta(s) = 1 + 0.86 \frac{1 - \cos(s)}{s} - 0.04s + 4 \times 10^{-7} s^2 + 9 \times 10^{-4} s^3 + 3 \times 10^{-6} s^4. \quad (9.4)$$

Thanks to this solution we determine the ball of local quadratic convergence given by γ -Theorem 1.3. To do that we compute $\gamma(F, DF(\zeta), \zeta) = \frac{1}{2} \|DF(\zeta)^{-1} D^2 F(\zeta)\|$.

Proposition 9.1 *We have*

1. $\|D^2 F(\zeta)\| = \left\| \int_0^1 \sin(st) dt \right\| < 0.46$.
2. $\gamma_\zeta = \gamma(F, DF(\zeta), \zeta) \leq \frac{1}{2} \frac{\|D^2 F(\zeta)\|}{1 - \left\| \int_0^1 \sin(st) \zeta(t) dt \right\|} < 0.54$

For $\bar{s} = 0.23$, the convergence of Hald's sequence is quadratic for all $x_0 \in B(\zeta, 0.43)$.

Proof. To determine $DF(\zeta)^{-1} z(t)$ we need to solve $DF(\zeta)y(s) = z(s)$, i.e.

$$y(s) - \int_0^1 \sin(st) \zeta(t) y(t) dt = z(s).$$

Hence if $DF(\zeta)^{-1}$ exists it satisfies

$$z(s) + \int_0^1 \sin(st) \zeta(t) DF(\zeta)^{-1} z(t) dt = DF(\zeta)^{-1} z(s).$$

A numerical computation gives $\|I - DF(\zeta)^{-1}\| = \left\| \int_0^1 \sin(st) \zeta(t) dt \right\| < 0.58\dots$ Hence $DF(\zeta)$ is invertible and

$$\|DF(\zeta)^{-1}\| = \frac{1}{1 - \left\| \int_0^1 \sin(st) \zeta(t) dt \right\|} < 2.34$$

Moreover $\|D^2 F(\zeta)\| = \left\| \int_0^1 \sin(st) \right\| < 0.46$. Then

$$\gamma(F, DF(\zeta), \zeta) < \frac{1}{2} \frac{\|D^2 F(\zeta)\|}{1 - \left\| \int_0^1 \sin(st) \zeta(t) dt \right\|} < 0.54.$$

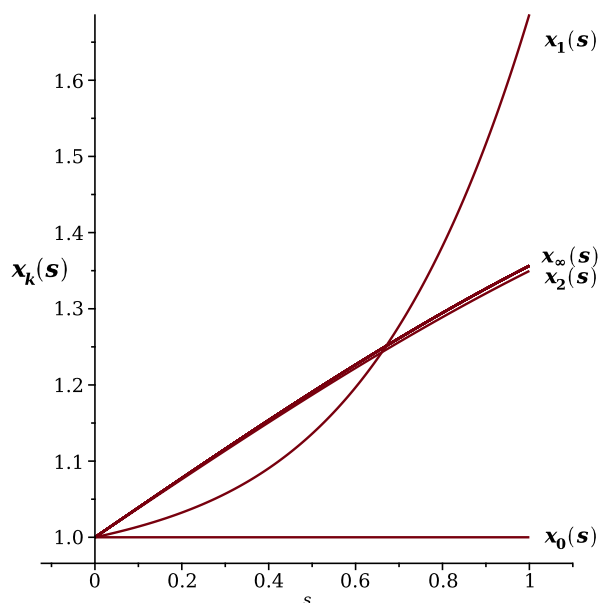


Figure 7: Iterated curves from $x_{1,0}$.

Then, letting $\bar{s} = 0.23$ we get $r = \frac{\bar{s}}{\gamma_\zeta} = 0.43$ and from Theorem 1.3 we deduce that the Hald sequence converges quadratically for all $x_0 \in B(\zeta, 0.43)$. \square

Table 2 corroborates the quadratic convergence of Hald sequence and Figures 7 and 8 show the iterated curves corresponding to Hald sequence initialized respectively to $x_{1,0}$ and $x_{2,0}$.

k	1	2	3	4	5	6	7
$x_0=x_{1,0}, \ x_{k+1}-x_k\ $	0.22	0.78	0.028	2×10^{-5}	5×10^{-11}	2×10^{-22}	2×10^{-45}
$x_0=x_{2,0}, \ x_{k+1}-x_k\ $	0.08	0.33	2×10^{-3}	2×10^{-7}	5×10^{-15}	2×10^{-30}	3×10^{-61}

Table 2: Quadratic convergence of Hald's sequence.

10 Concluding remark

We first have given a new condition of existence of a zero of an analytic function. Next, we have studied the local behaviour of numerical Hald's method. An application of the results presented here is to find by homotopy method a point satisfying the assumptions of our results. Classically this is realized using Newton's method. It could be interesting to study

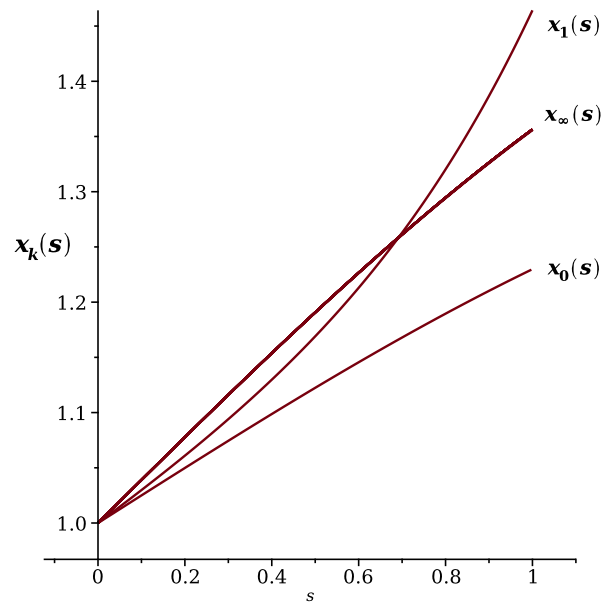


Figure 8: Iterated curves from $x_{2,0}$.

the use of homotopy coupled with Hald's method, for instance in the multiscale methods, to numerically solve integral equations.

References

- [1] **Blum, E., Cucker, F., Shub, M., and Smale, S.** : *Complexity and Real Computation*. Springer, 1998
- [2] **Ezquerro, J. A., and Hernández, M. A.** : *The Ulm method under mild differentiability conditions*. *Numerische Mathematik*, 109-2:193–207, Springer, 2008
- [3] **Giusti, M., Lecerf, G., Salvy, B., and Yakoubsohn, J.-C.** : *On location and approximation of clusters of zeros of analytic functions*. *Foundations of Computational Mathematics*, 5:257–311, Springer, 2005
- [4] **Gutiérrez, J. M., Hernández, M. A., and Romero, N.** : *A note on a modification of Moser's method*. *Journal of Complexity*, 24-2:185–197, 2008
- [5] **Gutiérrez, J. M., and Hernández, M. A.** : *On the convergence of Newton-Moser method from a data at one point*. In: *Understanding Banach Spaces*, Daniel González Sánchez eds., Nova Science Publishers, 2020
- [6] **Gutiérrez, J. M., Hernández, M. A., and Romero, N.** : *α -theory for Newton-Moser method*. *Monografías Matemáticas García de Galdeano*, 35:155–162, 2010

- [7] **Hald, O. H.** : *On a Newton-Moser type method.* Numer. Math., 23:411-426, 1975
- [8] **Kantorovich, L. V.** : *On Newton method for functional equations.* Dokl. Acad. N. USSR, 59(7): 1237, 1948
- [9] **Kantorovich, L. V., and Akilov, G. P.** : *Functional analysis.* Elsevier, 2016
- [10] **Lairez, P.** : *A deterministic algorithm to compute approximate roots of polynomial systems in polynomial average time.* Foundations of computational mathematics, 17-5:1265–1292, Springer, 2017
- [11] **Moser, J.** : *Stable and random motions in dynamical systems. With special emphasis on celestial mechanics.* Herman Weyl Lectures, Annals of Mathematics Studies, no. 77. Princeton, New Jersey: Princeton University Press, 1973
- [12] **Smale, S.** : *Complexity theory and numerical analysis.* Act. Numer., 6: 523, 1997
- [13] **Schulz, G.** : *Iterative Berechnung der reziproken Matrix.* ZAMM-Journal of Applied Mathematics and Mechanics, 13-1:57-59, 1933
- [14] **Ulm, S.** : *On iterative methods with successive approximation of the inverse operator (Russian).* Izv. Akad Nauk Est. SSR, 16:403–411, 1967
- [15] **Wang, Xinghua, and Han, Danfu** : *On the dominating sequence method in the point estimates and Smale's theorem.* Science in China, 135–144, 1990
- [16] **Deren, Wang, and Fengguan, Zhao** : *The theory of Smale's point estimation and its applications.* J. Comput. Appl. Math., 60:253–269, 1995

received: November 5, 2024

Authors:

Jean-luc Voléry
Toulouse School of Economics,
Département de Mathématiques,
Université de Toulouse-Capitole,
1, Esplanade de l'Université,
31000 Toulouse Cedex, France

e-mail: jean-luc.volery@tse-fr.eu

Jean-claude Yakoubsohn
Institut de Mathématiques de Toulouse,
Université Paul Sabatier,
118, route de Narbonne,
31062 Toulouse Cédex,
France

e-mail: jean-claude.yakoubsohn@math.univ-toulouse.fr

COLIN BUTTCHEREIT

Monotones Iterationsschema in Risikomodellen

1 Einführung

Partielle Differentialgleichungen haben häufig einen naturwissenschaftlichen Hintergrund. Beispielsweise sind in dem Werk [2] von W. Arendt und K. Urban die Transportgleichung, die Wärmeleitungsgleichung und die Wellengleichung zu finden. Allerdings gibt es auch partielle Differentialgleichungen aus dem Bereich der Finanzwirtschaft. Der US-amerikanische Wirtschaftswissenschaftler Fischer Sheffey Black und der kanadische Wirtschaftswissenschaftler Myron Samuel Scholes veröffentlichten im Jahr 1973 das bekannte Black-Scholes-Modell, in dem es im Wesentlichen darum geht, den fairen Preis $U(S, t)$ eines Derivats für eine Risikoanlage S (z.B. eine Aktie) zu einem gewissen Zeitpunkt t zu bestimmen. Es ist üblich, dass für Modelle gewisse Annahmen zur Vereinfachung getroffen werden. In diesem Artikel wird das „klassische“ Black-Scholes-Modell um das Ausfallrisiko des Ausgebers B und der Gegenpartei C sowie um die Finanzierungskosten erweitert, was insgesamt zu einer nichtlinearen partiellen Differentialgleichung vom Black-Scholes-Typ führt. Diese Gleichung hat folgende Gestalt:

$$\frac{\partial}{\partial t}U + \mathcal{A}_t U - rU = -(1 - R_B) \lambda_B U^- + (1 - R_C) \lambda_C U^+ + s_F U^+ \quad (1)$$

für $(S, t) \in (0, \infty) \times (0, T)$.

Für die genaue Herleitung der Gleichung (1) sei auf C. Burgard und M. Kjaer in [6] (Kapitel 3) verwiesen.

Der zeitabhängige Black-Scholes-Operator \mathcal{A}_t ist definiert durch

$$\mathcal{A}_t U := \frac{1}{2} [\sigma(t)]^2 S^2 \frac{\partial^2}{\partial S^2} U + [q_S(t) - \gamma_S(t)] S \frac{\partial}{\partial S} U \quad (2)$$

für $U : (0, \infty) \times (0, T) \rightarrow \mathbb{R}$.

Mit $r > 0$ bezeichnen wir den risikolosen Zinssatz, mit $r_B > 0$ bzw. $r_C > 0$ die Rendite der Anleihen von B bzw. C und mit r_F die Finanzierungsrate. Außerdem definieren wir die nichtnegativen Konstanten $\lambda_B := r_B - r$, $\lambda_C := r_C - r$ und $s_F := r_F - r$. Des Weiteren ist

$\gamma_S(t) > 0$ die Dividenden-Rendite der Aktien-Beteiligung, $q_S(t) > 0$ die Finanzierungskosten-Rate der Aktien-Beteiligung und $\sigma(t) > 0$ die Volatilität. Zuletzt sind $R_B \in [0, 1]$ bzw. $R_C \in [0, 1]$ die Erholungsraten im Falle des Ausfalls von B bzw. C . Außerdem ist $\varphi^+ = \max(\varphi, 0)$ der Positivteil und $\varphi^- = \max(-\varphi, 0)$ der Negativteil der Funktion φ .

Die nichtlineare parabolische Gleichung (1) wird durch den Endwert

$$U(S, T) = H(S) \quad (3)$$

für $S \in (0, \infty)$ ergänzt.

Dieser Endwert wird durch die Art der Option festgelegt. Anhand einer Aktie werden nun kurz die Optionen und ihre dazugehörigen Auszahlungsfunktionen vorgestellt, die wir in diesem Artikel vorrangig betrachten. Dabei betrachten wir die selben Optionen wie F. Baustian, J. Pospíšil und V. Švígler in [5]:

1. Call-Option
2. Forward
3. Gap-Option

Eine **Call-Option** gibt dem Ausgeber B das Recht, jedoch nicht die Pflicht, eine Aktie zu der Zeit T für einen festgelegten Strike-Preis $K > 0$ zu kaufen. Die Auszahlungsfunktion der Call-Option ist daher durch

$$H_{\text{call}}(S) = (S - K)^+ \quad (4)$$

gegeben (siehe Abbildung 1).

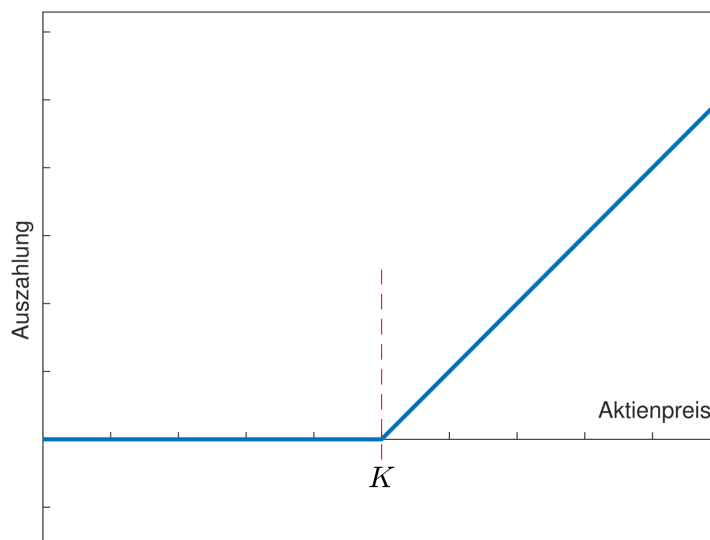


Abbildung 1: Auszahlungsfunktion einer Call-Option

Bei einem **Forward** hat der Ausgeber B zu dem Zeitpunkt T die Pflicht, eine Aktie zu dem festgelegten Strike-Preis $K > 0$ zu kaufen. Die Auszahlungsfunktion eines Forwards ist daher durch

$$H_{\text{fwd}}(S) = S - K \quad (5)$$

gegeben (siehe Abbildung 2).

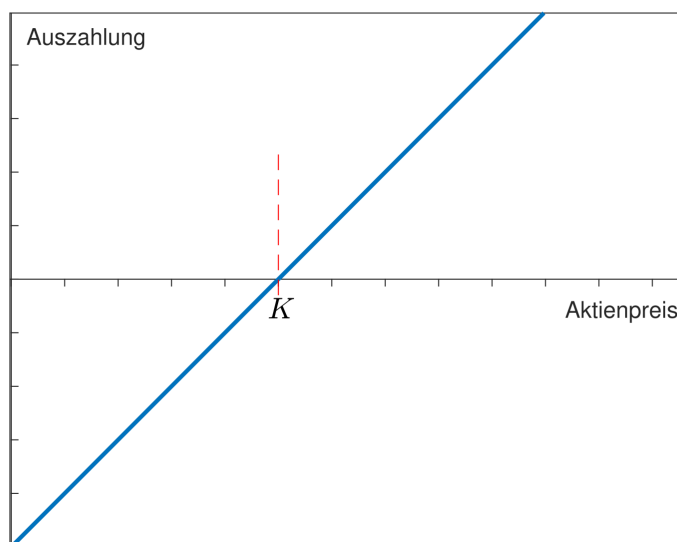


Abbildung 2: Auszahlungsfunktion eines Forwards

Bei der **Gap-Option** gibt es zusätzlich zum Strike-Preis K_s noch einen Trigger-Preis K_t , wobei $K_s > K_t > 0$ gelten soll. Es ist maßgeblich, ob der Aktienpreis zum Zeitpunkt T mindestens so groß ist wie der Trigger-Preis. Die Auszahlungsfunktion der Gap-Option ist durch

$$H_{\text{gap}}(S) = (S - K_s) \mathbb{1}_{\{S \geq K_t\}} \quad (6)$$

gegeben (siehe Abbildung 3).

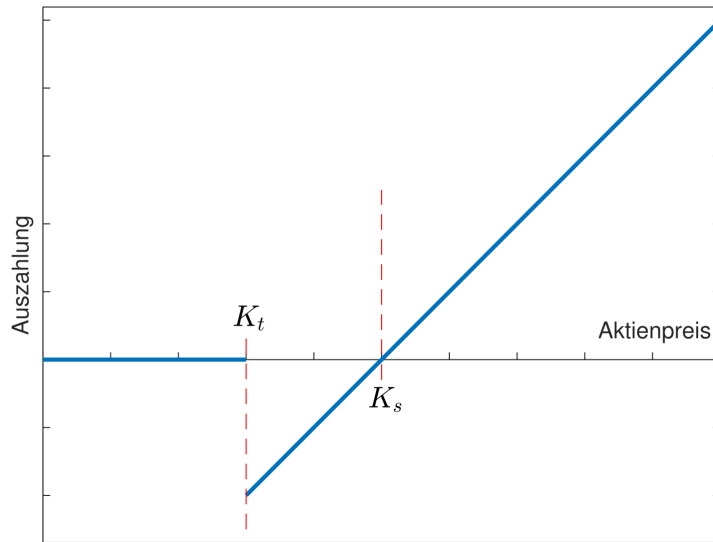


Abbildung 3: Auszahlungsfunktion einer Gap-Option

Es ist anzumerken, dass die Call-Option immer eine nichtnegative Auszahlung besitzt, während die beiden anderen Verträge auch negative Auszahlungen besitzen können.

Der Artikel ist folgendermaßen aufgebaut. Zunächst wird im zweiten Kapitel das nichtlineare Problem umgeformt, damit sich aus dem Endwertproblem (1), (3) ein Anfangswertproblem ergibt. Des Weiteren betrachten wir, in welchen Räumen sich die betrachteten Funktionen befinden. Dabei werden die hierfür relevanten gewichteten Lebesgue-Räume eingeführt.

Im dritten Kapitel wird das Prinzip von Ober- und Unterlösungen erläutert. Daraufhin wird eine Folge von Funktionen konstruiert, welche Lösungen linearer partieller Differentialgleichungen sind. Diese Folge konvergiert aufgrund dieser Konstruktion gegen die Lösung des ursprünglichen nichtlinearen Problems. Zuletzt leiten wir mithilfe eines „Variation der Konstanten“-Ansatzes eine explizite Lösungsformel für das lineare Problem her.

Im vierten Kapitel wird die Methode der finiten Differenzen auf das monotone Iterationschema angewendet, indem konkret einige Iterationen für festgelegte Parameter durchgeführt und Ergebnisse grafisch dargestellt werden. So kann eingeordnet werden, ob die Anwendung der Finiten-Differenzen-Methode für das monotone Iterationschema sinnvolle Ergebnisse liefert.

Abschließend wird im fünften Kapitel ein Fazit gezogen, inwieweit das monotone Iterations-

schema zur Behandlung des betrachteten Risiko-Modells geeignet ist und wie aussagekräftig die Ergebnisse der numerischen Behandlung sind.

Während im zweiten und dritten Kapitel die analytischen Aspekte im Vordergrund stehen, welche beispielsweise von B. Alziary und P. Takáč in [1] erforscht wurden, konzentriert sich das vierte Kapitel auf die numerische Behandlung des Iterationsschemas. Auf diese Weise werden die bisherigen analytischen Ergebnisse um interessante numerische Resultate erweitert.

2 Vorbereitung

Bevor in Kapitel 3 das monotone Iterationsschema definiert werden kann, formen wir das nichtlineare Problem (1), (3) zunächst um. Außerdem sind wir daran interessiert, in welchen Räumen sich der Anfangswert und die Lösung des Problems befinden.

2.1 Umformung des nichtlinearen Problems

Die Umformungen basieren auf [1] (Kapitel 2) und [5] (Kapitel 2, Kapitel 3.1).

Wir betrachten das nichtlineare Problem (1), (3) und setzen $\tau = T - t$, $x = \ln S$ sowie $u(x, \tau) = U(S, t)$. Aus dem Endwert (3) ergibt sich mit $h(x) = H(S)$ dann der Anfangswert

$$u(x, 0) = h(x) = H(e^x) \quad (7)$$

für $x \in \mathbb{R}$.

Aus der Gleichung (1) ergibt sich in verkürzter Schreibweise

$$\frac{\partial}{\partial \tau} u - \mathcal{A}(\tau)u + ru = f(u) \quad (8)$$

für $(x, \tau) \in \mathbb{R} \times (0, T)$.

Dabei ist nun der zeitabhängige Black-Scholes-Operator $\mathcal{A}(\tau)$ definiert durch

$$\mathcal{A}(\tau)u := \frac{1}{2} [\sigma(\tau)]^2 \frac{\partial^2}{\partial x^2} u + \left[q_S(\tau) - \gamma_S(\tau) - \frac{1}{2} [\sigma(\tau)]^2 \right] \frac{\partial}{\partial x} u \quad (9)$$

für $u : \mathbb{R} \times (0, T) \rightarrow \mathbb{R}$.

An dieser Stelle sei angemerkt, dass alle linearen Operatoren $\mathcal{A}(\tau)$, für $\tau \in (0, T)$, paarweise kommutativ sind.

Die Nichtlinearität $f(\cdot, x, \tau) : \mathbb{R} \rightarrow \mathbb{R}$ auf der rechten Seite von (8) definieren wir durch

$$f(u, x, \tau) := (1 - R_B) \lambda_B u^- - (1 - R_C) \lambda_C u^+ - s_F u^+ \quad (10)$$

für $u \in \mathbb{R}$ und $(x, \tau) \in \mathbb{R} \times (0, T)$.

Addiert man bei der Gleichung (8) auf beiden Seiten den Term $L_f u$, wobei die Konstante $L_f \in [0, \infty)$ definiert ist durch

$$L_f := \max \{(1 - R_B)\lambda_B, (1 - R_C)\lambda_C + s_F\}, \quad (11)$$

so erhält man insgesamt also das äquivalente nichtlineare Problem

$$\begin{cases} \frac{\partial}{\partial \tau} u - \mathcal{A}(\tau)u + r_g u = g(u) & \text{für } (x, \tau) \in \mathbb{R} \times (0, T) \\ u(x, 0) = h(x) = H(e^x) & \text{für } x \in \mathbb{R}. \end{cases} \quad (12a)$$

$$(12b)$$

Hierbei haben wir $r_g := r + L_f$ gesetzt und die Nichtlinearität $g(\cdot, x, \tau) : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$g(u, x, \tau) := f(u, x, \tau) + L_f u \quad (13)$$

für $u \in \mathbb{R}$ und $(x, \tau) \in \mathbb{R} \times (0, T)$.

Das nichtlineare Problem (12a), (12b) wird im Folgenden zugrunde gelegt. Der nachfolgende Satz zeigt, warum die Addition des Terms $L_f u$ sinnvoll ist. Wir gehen an dieser Stelle etwas präziser auf den Beweis ein als in [1].

Satz 1 Für alle $(x, \tau) \in \mathbb{R} \times (0, T)$ ist die Funktion $g(\cdot, x, \tau) : u \mapsto g(u, x, \tau) : \mathbb{R} \rightarrow \mathbb{R}$ monoton wachsend, d.h. für alle $u_1, u_2 \in \mathbb{R}$ gilt die Implikation

$$u_1 \leq u_2 \implies g(u_1, x, \tau) \leq g(u_2, x, \tau). \quad (14)$$

Beweis. Für alle $(x, \tau) \in \mathbb{R} \times (0, T)$ gilt

$$\begin{aligned} g(u, x, \tau) &= f(u, x, \tau) + L_f u \\ &= (1 - R_B)\lambda_B u^- - (1 - R_C)\lambda_C u^+ - s_F u^+ + L_f (u^+ - u^-) \\ &= (-L_f + (1 - R_B)\lambda_B) u^- + (L_f - (1 - R_C)\lambda_C - s_F) u^+ \\ &= -(L_f - (1 - R_B)\lambda_B) u^- + (L_f - ((1 - R_C)\lambda_C + s_F)) u^+ \\ &= \begin{cases} (L_f - (1 - R_B)\lambda_B) u & , \text{ falls } u < 0 \\ (L_f - ((1 - R_C)\lambda_C + s_F)) u & , \text{ falls } u > 0 \\ 0 & , \text{ falls } u = 0. \end{cases} \end{aligned} \quad (15)$$

Also gilt

$$\frac{\partial}{\partial u} g = \begin{cases} L_f - (1 - R_B)\lambda_B & , \text{ falls } u < 0 \\ L_f - ((1 - R_C)\lambda_C + s_F) & , \text{ falls } u > 0. \end{cases} \quad (16)$$

Weiter gilt

$$L_f = \max \{(1 - R_B)\lambda_B, (1 - R_C)\lambda_C + s_F\} \geq (1 - R_B)\lambda_B \quad (17)$$

sowie

$$L_f = \max \{(1 - R_B)\lambda_B, (1 - R_C)\lambda_C + s_F\} \geq (1 - R_C)\lambda_C + s_F. \quad (18)$$

Damit folgt

$$0 \leq \frac{\partial}{\partial u} g \leq L_f \quad (19)$$

für alle $u \in \mathbb{R} \setminus \{0\}$ und $(x, \tau) \in \mathbb{R} \times (0, T)$.

Außerdem gilt wegen (15) für alle $(x, \tau) \in \mathbb{R} \times (0, T)$

$$\begin{aligned} g(u, x, \tau) &\leq 0, \text{ falls } u < 0, \\ g(u, x, \tau) &= 0, \text{ falls } u = 0, \\ g(u, x, \tau) &\geq 0, \text{ falls } u > 0, \end{aligned} \quad (20)$$

sodass die Behauptung mit (19) und (20) folgt. \square

In [1] (Kapitel 2) werden ganz allgemein fünf Forderungen (G1 - G5) an die Nichtlinearität $g : \mathbb{R} \times \mathbb{R} \times (0, T) \rightarrow \mathbb{R}$ gestellt. Neben der gerade bewiesenen Monotonie (G3) erfüllt die konkret betrachtete Nichtlinearität (13) alle diese Forderungen, denn:

Für jedes feste $u \in \mathbb{R}$ ist die Funktion $g(u, \cdot, \cdot) : (x, \tau) \mapsto g(u, x, \tau) : \mathbb{R} \times (0, T) \rightarrow \mathbb{R}$ Lebesgue-messbar (G1). Die exponentielle Wachstumsbeschränkung (G4) der Funktion $g(0, \cdot, \tau) : x \mapsto g(0, x, \tau) : \mathbb{R} \rightarrow \mathbb{R}$ ist gegeben, da $g(0, x, \tau) = 0$ gilt. Auch die Hölder-Stetigkeit (G5) der Funktion $g(u, x, \cdot) : \tau \mapsto g(u, x, \tau) : (0, T) \rightarrow \mathbb{R}$ liegt vor, da g nicht explizit von τ abhängt. Zuletzt erhalten wir auch die gleichmäßige Lipschitz-Stetigkeit (G2) der Funktion $g(\cdot, x, \tau) : u \mapsto g(u, x, \tau) : \mathbb{R} \rightarrow \mathbb{R}$ durch (19) zusammen mit der Tatsache, dass $g(0, x, \tau) = 0$ gilt.

2.2 Die betrachteten Räume

In diesem Abschnitt wird dargelegt, in welchen Räumen sich die Lösung $u : \mathbb{R} \times (0, T) \rightarrow \mathbb{R}$ des nichtlinearen Problems (12a), (12b) bzw. die Funktionen in dem später vorgestellten Iterationsschema, welche Lösungen linearer partieller Differentialgleichungen sind, befinden. Dies hängt wesentlich davon ab, in welchem Raum sich der Anfangswert (12b) befindet. Häufig werden σ , q_S und γ_S als positive Konstanten angenommen. Es genügen allerdings auch schwächere Annahmen, die wir nun gemäß [1] (Kapitel 2) aufzählen:

Annahme 1 $\sigma : [0, T] \rightarrow (0, \infty)$ ist eine Hölder-stetige Funktion, d.h. es existieren von $\tau \in [0, T]$ unabhängige Konstanten $C_\sigma \in [0, \infty)$ und $\vartheta_\sigma \in (0, 1)$, sodass

$$|\sigma(\tau_1) - \sigma(\tau_2)| \leq C_\sigma |\tau_1 - \tau_2|^{\vartheta_\sigma} \quad (21)$$

für alle $\tau_1, \tau_2 \in [0, T]$ gilt.

Annahme 2 $q_S, \gamma_S : [0, T] \rightarrow (0, \infty)$ sind Hölder-stetige Funktion, d.h. es existieren von $\tau \in [0, T]$ unabhängige Konstanten $C_q, C_\gamma \in [0, \infty)$ und $\vartheta_q, \vartheta_\gamma \in (0, 1)$, sodass

$$|q_S(\tau_1) - q_S(\tau_2)| \leq C_q |\tau_1 - \tau_2|^{\vartheta_q} \quad (22)$$

und

$$|\gamma_S(\tau_1) - \gamma_S(\tau_2)| \leq C_\gamma |\tau_1 - \tau_2|^{\vartheta_\gamma} \quad (23)$$

für alle $\tau_1, \tau_2 \in [0, T]$ gilt.

Zuletzt muss noch eine Annahme über den Anfangswert (12b) getroffen werden. Am Beispiel des Forwards sieht man, dass die Funktion $h_{\text{fwd}} : \mathbb{R} \rightarrow \mathbb{R}$ mit $h_{\text{fwd}}(x) = H_{\text{fwd}}(e^x) = e^x - K$ unbeschränkt ist. Jedoch gilt immerhin $h_{\text{fwd}}(x) \leq e^x$ für alle $x \in \mathbb{R}$. Allgemein fordern wir:

Annahme 3 Die Funktion $h : \mathbb{R} \rightarrow \mathbb{R}$ ist Lebesgue-messbar und es existiert eine Konstante $C_h \in [0, \infty)$, sodass für fast alle $x \in \mathbb{R}$ die Ungleichung

$$|h(x)| = |H(e^x)| \leq C_h e^{|x|} \quad (24)$$

gilt.

Es sei angemerkt, dass die drei konkret betrachteten Verträge (Call-Option, Forward und Gap-Option) jeweils die Annahme 3 erfüllen.

Durch diese Annahme motiviert, betrachten wir zunächst den Vektorraum

$$\mathcal{L}^2(\mathbb{R}, w) := \left\{ f : \mathbb{R} \rightarrow \mathbb{C} \mid f \text{ ist Lebesgue-messbar und } \int_{\mathbb{R}} |f(x)|^2 w(x) dx < \infty \right\}, \quad (25)$$

wobei $w(x) = e^{-\mu|x|}$ eine Gewichtsfunktion mit einer Konstanten $\mu \in (2, \infty)$ ist.

Zunächst erhalten wir durch die Abbildung

$$\|\cdot\|_{\mathcal{L}^2(\mathbb{R}, w)} : \mathcal{L}^2(\mathbb{R}, w) \rightarrow [0, \infty), \quad f \mapsto \left(\int_{\mathbb{R}} |f(x)|^2 w(x) dx \right)^{\frac{1}{2}} \quad (26)$$

eine Seminorm. Um eine Norm zu erhalten, definieren wir

$$\mathcal{N} := \left\{ f \in \mathcal{L}^2(\mathbb{R}, w) : \|f\|_{\mathcal{L}^2(\mathbb{R}, w)} = 0 \right\} \quad (27)$$

und definieren den **gewichteten Lebesgue-Raum** durch

$$L^2(\mathbb{R}, w) := \mathcal{L}^2(\mathbb{R}, w) / \mathcal{N}. \quad (28)$$

Dieser Raum entsteht aus $\mathcal{L}^2(\mathbb{R}, w)$ durch Identifizieren von Funktionen, die fast überall gleich sind. Wir verweisen an dieser Stelle auf F. Baustian, K. Filipová und J.Pospíšil in

[4] (Kapitel 2.1) und O. Forster in [9] (Kapitel 12) für weitere Details über Lebesgue- bzw. gewichtete Lebesgue-Räume.

Es ist bekannt, dass $H_{\mathbb{C}} := L^2(\mathbb{R}, w)$ ein Hilbertraum mit der Norm

$$\|f\|_H := \left(\int_{\mathbb{R}} |f(x)|^2 w(x) dx \right)^{\frac{1}{2}} < \infty \quad (29)$$

ist. Diese Norm wird von dem Skalarprodukt

$$\langle f, g \rangle_H := \int_{\mathbb{R}} f(x) \overline{g(x)} w(x) dx \quad (30)$$

für $f, g \in H_{\mathbb{C}}$ induziert.

Wir bezeichnen den abgeschlossenen Unterraum aller reellwertigen Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$ von $H_{\mathbb{C}}$ mit $H_{\mathbb{R}}$.

Da die Lösung $u(x, \tau)$ für das nichtlineare Problem (12a), (12b) reellwertig sein wird, werden wir insgesamt $u(\cdot, \tau) \in H_{\mathbb{R}}$ für alle $\tau \in (0, T)$ erhalten.

Zuletzt geben wir ein für das monotone Iterationsschema wichtiges Existenz- und Eindeutigkeitsresultat an. Dafür betrachten wir das lineare Problem

$$\begin{cases} \frac{\partial}{\partial \tau} u - \mathcal{A}(\tau)u + r_g u = F(x, \tau) & \text{für } (x, \tau) \in \mathbb{R} \times (0, T) \\ u(x, 0) = h(x) = H(e^x) & \text{für } x \in \mathbb{R} \end{cases} \quad (31a)$$

$$(31b)$$

in Anlehnung an das nichtlineare Problem (12a), (12b).

Dabei ist die Inhomogenität $F : [0, T] \rightarrow H_{\mathbb{R}}$ eine stetige Funktion, das heißt $F \in C([0, T] \rightarrow H_{\mathbb{R}})$, wobei $F : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ die Bedingung $F(\tau) \equiv F(\cdot, \tau) \in H_{\mathbb{R}}$ für alle $\tau \in [0, T]$ erfüllt. Die Ergebnisse in [1] (Kapitel 2) und von L. C. Evans in [8] (Kapitel 7.1) liefern, dass das parabolische lineare Problem (31a), (31b) genau eine schwache Lösung $u : [0, T] \rightarrow H_{\mathbb{R}}$ besitzt, sofern die Anfangsbedingung $h \in H_{\mathbb{R}}$ die Bedingung (24) erfüllt. Diese schwache Lösung u ist stetig als $H_{\mathbb{R}}$ -wertige Funktion von $\tau \in [0, T]$, das heißt $u \in C([0, T] \rightarrow H_{\mathbb{R}})$.

Es sei an dieser Stelle angemerkt, dass diese schwache Lösung $u : [0, T] \rightarrow H_{\mathbb{R}}$ auch eine klassische Lösung ist, falls $F : [0, T] \rightarrow H_{\mathbb{R}}$ Hölder-stetig zum Hölder-Exponenten $\vartheta_F \in (0, 1)$ ist. Für Details sei auf [1] (Kapitel 2, Hypothesis (f1')) verwiesen. Wir verwenden daher im Weiteren häufig die Notation der (schwachen) Lösung.

3 Monotones Iterationsschema

In diesem Kapitel behandeln wir nun das monotone Iterationsschema. Wir konstruieren eine Folge von Ober- bzw. Unterlösungen, welche selbst Lösungen von linearen Problemen sind. Diese Folge von Funktionen konvergiert dann monoton gegen die Lösung des nichtlinearen Problems (12a), (12b).

3.1 Ober- und Unterlösungen

Als Erstes müssen die Begriffe Oberlösung und Unterlösung definiert werden. Dieser Abschnitt basiert auf [1] (Kapitel 3.1) und [5] (Kapitel 3.1).

Wir definieren zunächst die Begriffe im Sinne von klassischen Lösungen.

Definition 1 Die Funktion $\bar{u} : [0, T] \rightarrow H_{\mathbb{R}}$ ist eine **Oberlösung** des linearen Problems (31a), (31b), wenn gilt:

1. $\bar{u}(x, 0) \geq h(x)$ für fast alle $x \in \mathbb{R}$.
2. Für das Lösungs-Residuum

$$\mathcal{D}(\bar{u}) := \frac{\partial}{\partial \tau} \bar{u}(x, \tau) - \mathcal{A}(\tau) \bar{u}(x, \tau) + r_g \bar{u}(x, \tau) - F(x, \tau) \quad (32)$$

gilt: $\mathcal{D}(\bar{u}) \geq 0$ für fast alle $(x, \tau) \in \mathbb{R} \times (0, T)$.

Analog lässt sich die Unterlösung definieren.

Definition 2 Die Funktion $\underline{u} : [0, T] \rightarrow H_{\mathbb{R}}$ ist eine **Unterlösung** des linearen Problems (31a), (31b), wenn gilt:

1. $\underline{u}(x, 0) \leq h(x)$ für fast alle $x \in \mathbb{R}$.
2. Für das Lösungs-Residuum

$$\mathcal{D}(\underline{u}) := \frac{\partial}{\partial \tau} \underline{u}(x, \tau) - \mathcal{A}(\tau) \underline{u}(x, \tau) + r_g \underline{u}(x, \tau) - F(x, \tau) \quad (33)$$

gilt: $\mathcal{D}(\underline{u}) \leq 0$ für fast alle $(x, \tau) \in \mathbb{R} \times (0, T)$.

An dieser Stelle sei angemerkt, dass für die schwache Lösung von (31a), (31b) der Begriff der schwachen Ober- bzw. Unterlösung (Definition 1 und Definition 2) im Sinne einer Approximation durch eine Folge von klassischen Lösung zu verstehen ist. Dafür sei erneut auf [1] (Kapitel 3.1) verwiesen.

Um das monotone Iterationsschema definieren zu können, wird zunächst eine anfängliche Oberlösung \bar{u}_0 beziehungsweise eine anfängliche Unterlösung \underline{u}_0 des nichtlinearen Problems (12a), (12b) benötigt. Der Begriff der Ober- und Unterlösung ist dabei analog zu Definition 1 und 2 definiert, indem man $F(x, \tau)$ durch $g(\bar{u})$ beziehungsweise durch $g(\underline{u})$ ersetzt.

Für die drei oben genannten Verträge (Call-Option, Forward und Gap-Option) kann jeweils eine anfängliche Oberlösung \bar{u}_0 und eine anfängliche Unterlösung \underline{u}_0 des nichtlinearen Problems (12a), (12b) angegeben werden.

Satz 2 Für die Call-Option, den Forward und die Gap-Option ist jeweils eine anfängliche Oberlösung \bar{u}_0 des nichtlinearen Problems (12a), (12b) gegeben durch

$$\bar{u}_0(x, \tau) = \exp(x + \bar{\lambda}\tau) \text{ für } (x, \tau) \in \mathbb{R} \times [0, T], \quad (34)$$

wobei $\bar{\lambda} \in \mathbb{R}$ eine Konstante ist, die

$$\bar{\lambda} \geq \max_{\tau \in [0, T]} [q_S(\tau) - \gamma_S(\tau)] - r - (1 - R_C)\lambda_C - s_F \quad (35)$$

erfüllt.

Beweis. Zunächst stellt man folgende Äquivalenz fest:

$$S \geq K_t \iff x \geq \ln K_t \quad (36)$$

Weiter gilt für alle $x \in \mathbb{R}$

$$\begin{aligned} \bar{u}_0(x, 0) &= e^x \geq (e^x - K)^+ = H_{\text{call}}(e^x) = h_{\text{call}}(x), \\ \bar{u}_0(x, 0) &= e^x \geq e^x - K = H_{\text{fwd}}(e^x) = h_{\text{fwd}}(x), \\ \bar{u}_0(x, 0) &= e^x \geq (e^x - K_s) \mathbb{1}_{\{x \geq \ln K_t\}} = H_{\text{gap}}(e^x) = h_{\text{gap}}(x), \end{aligned} \quad (37)$$

sodass Bedingung (1.) von Definition 1 erfüllt ist.

Außerdem gilt

$$\begin{aligned} \frac{\partial}{\partial \tau} \bar{u}_0 &= \bar{\lambda} \bar{u}_0, \\ \frac{\partial^2}{\partial x^2} \bar{u}_0 &= \frac{\partial}{\partial x} \bar{u}_0 = \bar{u}_0. \end{aligned} \quad (38)$$

Damit gilt also

$$\begin{aligned} \mathcal{A}(\tau) \bar{u}_0 &= \frac{1}{2} [\sigma(\tau)]^2 \bar{u}_0 + \left[q_S(\tau) - \gamma_S(\tau) - \frac{1}{2} [\sigma(\tau)]^2 \right] \bar{u}_0 = [q_S(\tau) - \gamma_S(\tau)] \bar{u}_0, \\ r_g \bar{u}_0 &= (r + L_f) \bar{u}_0, \\ g(\bar{u}_0) &= f(\bar{u}_0) + L_f \bar{u}_0 = -(1 - R_C) \lambda_C \bar{u}_0 - s_F \bar{u}_0 + L_f \bar{u}_0 \\ &= (L_f - (1 - R_C) \lambda_C - s_F) \bar{u}_0. \end{aligned} \quad (39)$$

Somit gilt für das Lösungs-Residuum

$$\begin{aligned} \mathcal{D}(\bar{u}_0) &= \frac{\partial}{\partial \tau} \bar{u}_0 - \mathcal{A}(\tau) \bar{u}_0 + r_g \bar{u}_0 - g(\bar{u}_0) \\ &= (\bar{\lambda} - [q_S(\tau) - \gamma_S(\tau)] + r + (1 - R_C) \lambda_C + s_F) \bar{u}_0 \\ &= (\bar{\lambda} - [q_S(\tau) - \gamma_S(\tau)] + r + (1 - R_C) \lambda_C + s_F) \exp(x + \bar{\lambda}\tau). \end{aligned} \quad (40)$$

Wegen $\exp(x + \bar{\lambda}\tau) > 0$ führt die Wahl von $\bar{\lambda} \geq \max_{\tau \in [0, T]} [q_S(\tau) - \gamma_S(\tau)] - r - (1 - R_C) \lambda_C - s_F$ dazu, dass $\mathcal{D}(\bar{u}_0) \geq 0$, sodass also auch die Bedingung (2.) aus Definition 1 erfüllt ist. \square

Ähnlich lässt sich für die drei Verträge auch eine anfängliche Unterlösung \underline{u}_0 des nichtlinearen Problems (12a), (12b) konstruieren.

Satz 3 *Eine anfängliche Unterlösung \underline{u}_0 des nichtlinearen Problems (12a), (12b) ist gegeben durch*

$$\underline{u}_0(x, \tau) = -\underline{C} \exp(\lambda\tau) \text{ für } (x, \tau) \in \mathbb{R} \times [0, T] \quad (41)$$

mit $\underline{C} \geq 0$ und $\lambda \in \mathbb{R}$, wobei wir für die Call-Option $\underline{C} = 0$, für den Forward $\underline{C} = K$ und für die Gap-Option $\underline{C} = K_s - K_t$ setzen. Zusätzlich muss bei dem Forward und der Gap-Option noch $\lambda \geq (1 - R_B) \lambda_B - r$ gewählt werden.

Beweis. Für die Call-Option ist $\underline{u}_0(x, \tau) \equiv 0$. Also gilt für alle $x \in \mathbb{R}$

$$\underline{u}_0(x, 0) = 0 \leq (e^x - K)^+ = H_{\text{call}}(e^x) = h_{\text{call}}(x), \quad (42)$$

sodass Bedingung (1.) von Definition 2 erfüllt ist.

Außerdem gilt für das Lösungs-Residuum $\mathcal{D}(\underline{u}_0) = 0$, sodass Bedingung (2.) aus Definition 2 auch erfüllt ist.

Nun betrachten wir den Forward und die Gap-Option mit $\underline{u}_0(x, \tau) = -\underline{C} \exp(\lambda\tau)$. Für alle $x \in \mathbb{R}$ gilt für den Forward

$$\underline{u}_0(x, 0) = -K \leq e^x - K = H_{\text{fwd}}(e^x) = h_{\text{fwd}}(x) \quad (43)$$

und für die Gap-Option

$$\begin{aligned} \underline{u}_0(x, 0) &= K_t - K_s \leq (e^x - K_s) \mathbb{1}_{\{e^x \geq K_t\}} = (e^x - K_s) \mathbb{1}_{\{x \geq \ln K_t\}} \\ &= H_{\text{gap}}(e^x) = h_{\text{gap}}(x), \end{aligned} \quad (44)$$

wobei hier noch einmal angemerkt sei, dass $K_t - K_s < 0$. Damit ist Bedingung (1.) von Definition 2 erfüllt.

Nun gilt weiter

$$\begin{aligned} \frac{\partial}{\partial \tau} \underline{u}_0 &= \lambda \underline{u}_0, \\ \frac{\partial^2}{\partial x^2} \underline{u}_0 &= \frac{\partial}{\partial x} \underline{u}_0 = 0. \end{aligned} \quad (45)$$

Damit gilt

$$\begin{aligned} \mathcal{A}(\tau) \underline{u}_0 &= 0, \\ r_g \underline{u}_0 &= (r + L_f) \underline{u}_0, \\ g(\underline{u}_0) &= f(\underline{u}_0) + L_f \underline{u}_0 = (1 - R_B) \lambda_B \underline{u}_0 + L_f \underline{u}_0. \end{aligned} \quad (46)$$

Somit gilt für das Lösungs-Residuum

$$\begin{aligned} \mathcal{D}(\underline{u}_0) &= \frac{\partial}{\partial \tau} \underline{u}_0 - \mathcal{A}(\tau) \underline{u}_0 + r_g \underline{u}_0 - g(\underline{u}_0) = (\underline{\lambda} + r - (1 - R_B) \lambda_B) \underline{u}_0 \\ &= -\underline{C} (\underline{\lambda} + r - (1 - R_B) \lambda_B) \exp(\underline{\lambda} \tau). \end{aligned} \quad (47)$$

Da $\underline{C} \geq 0$ und $\exp(\underline{\lambda} \tau) > 0$, führt die Wahl von $\underline{\lambda} \geq (1 - R_B) \lambda_B - r$ dazu, dass $\mathcal{D}(\underline{u}_0) \leq 0$, sodass also auch die Bedingung (2.) aus Definition 2 erfüllt ist. \square

Wir schließen das Unterkapitel 3.1 mit einem sehr wichtigen Satz ab, der auf dem Maximumprinzip basiert.

Das schwache Maximumprinzip für die schwache Lösung u von (31a), (31b) besagt laut [1] (Kapitel 3.1) Folgendes:

Wenn $u(x, 0) \geq 0$ und $F(x, \tau) \geq 0$ für fast alle $(x, \tau) \in \mathbb{R} \times (0, T)$ gilt, dann gilt auch $u(x, \tau) \geq 0$ für fast alle $(x, \tau) \in \mathbb{R} \times (0, T)$.

Folgender Satz hilft für die anschließende Konstruktion der monotonen Iterationen weiter:

Satz 4 Seien $\bar{u}, \underline{u} : \mathbb{R} \times (0, T) \rightarrow \mathbb{R}$ eine (schwache) Ober- bzw. Unterlösung des linearen Problems (31a), (31b), sodass $\underline{u}(x, 0) \leq \bar{u}(x, 0)$ für fast alle $x \in \mathbb{R}$ gilt. Dann gilt für jedes $\tau \in [0, T)$ auch $\underline{u}(x, \tau) \leq \bar{u}(x, \tau)$ für fast alle $x \in \mathbb{R}$.

Beweis. Anwendung des schwachen Maximumprinzips auf die Differenz $v = \bar{u} - \underline{u}$. Siehe Lemma 3.1 (Weak comparison) in [1] (Kapitel 3.1) sowie Erkenntnisse von A. Friedman in [10] (Kapitel 2). \square

3.2 Konstruktion der Iterationen

Wir definieren nun das monotone Iterationsschema für die Oberlösungen mithilfe der bisherigen Erkenntnisse. Der Abschnitt basiert auf [1] (Kapitel 3.2). Dort wird eine allgemeinere anfängliche Ober- bzw. Unterlösung angegeben. Wir beschränken uns nun auf die konkret betrachteten Verträge (Call-Option, Forward, Gap-Option) und deren in Satz 2 und Satz 3 benannten anfänglichen Ober- bzw. Unterlösungen.

Wir betrachten zunächst die anfängliche Oberlösung \bar{u}_0 aus Satz 2. Nun sei $\bar{u}_1 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ die (schwache) Lösung von

$$\begin{cases} \frac{\partial}{\partial \tau} \bar{u}_1 - \mathcal{A}(\tau) \bar{u}_1 + r_g \bar{u}_1 = g(\bar{u}_0) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \bar{u}_1(x, 0) = h(x) = H(e^x) & \text{für } x \in \mathbb{R}. \end{cases} \quad (48a)$$

$$(48b)$$

Aufgrund der bisherigen Erkenntnisse wissen wir, dass \bar{u}_1 existiert und eindeutig ist. Außerdem gilt $\bar{u}_1 \in C([0, T] \rightarrow H_{\mathbb{R}})$.

Da \bar{u}_1 eine (schwache) Lösung von (48a), (48b) ist, ist \bar{u}_1 gleichzeitig eine (schwache) Oberlösung und eine (schwache) Unterlösung von (48a), (48b). Da \bar{u}_0 eine Oberlösung von (12a), (12b) ist, ist \bar{u}_0 auch eine (schwache) Oberlösung von (48a), (48b). Somit kann man mit Satz 4 folgern, dass $\bar{u}_0(x, \tau) \geq \bar{u}_1(x, \tau)$ für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ gilt.

Analog gilt dann auch $\bar{u}_1(x, \tau) \geq \underline{u}_0(x, \tau)$ für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$, da \underline{u}_0 (gemäß Satz 3) eine Unterlösung von (12a), (12b) ist. Insgesamt folgt dann also

$$\bar{u}_0(x, \tau) \geq \bar{u}_1(x, \tau) \geq \underline{u}_0(x, \tau) \quad (49)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$.

Nun nehmen wir an (Induktionsvoraussetzung), dass wir die Funktionen $\bar{u}_0, \bar{u}_1, \dots, \bar{u}_n : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ so konstruiert haben, dass gilt:

1. Für alle $i = 0, \dots, n$ ist $\bar{u}_i : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ Lebesgue-messbar und es gilt $\bar{u}_i \in C([0, T] \rightarrow H_{\mathbb{R}})$.
2. Für alle $i = 1, \dots, n$ gilt

$$\bar{u}_0(x, \tau) \geq \bar{u}_{i-1}(x, \tau) \geq \bar{u}_i(x, \tau) \geq \underline{u}_0(x, \tau) \quad (50)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$.

3. Für alle $i = 1, \dots, n$ ist $\bar{u}_i : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ die (schwache) Lösung von

$$\begin{cases} \frac{\partial}{\partial \tau} \bar{u}_i - \mathcal{A}(\tau) \bar{u}_i + r_g \bar{u}_i = g(\bar{u}_{i-1}) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \bar{u}_i(x, 0) = h(x) = H(e^x) & \text{für } x \in \mathbb{R}. \end{cases} \quad (51a)$$

$$\quad \quad \quad (51b)$$

Zuletzt muss nun noch die Funktion $\bar{u}_{n+1} : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ als (schwache) Lösung von

$$\begin{cases} \frac{\partial}{\partial \tau} \bar{u}_{n+1} - \mathcal{A}(\tau) \bar{u}_{n+1} + r_g \bar{u}_{n+1} = g(\bar{u}_n) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \bar{u}_{n+1}(x, 0) = h(x) = H(e^x) & \text{für } x \in \mathbb{R} \end{cases} \quad (52a)$$

$$\quad \quad \quad (52b)$$

konstruiert werden. Da wir durch (50) insbesondere wissen, dass

$$\bar{u}_{n-1}(x, \tau) \geq \bar{u}_n(x, \tau) \quad (53)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ gilt, können wir mit Satz 1 folgern, dass

$$g(\bar{u}_{n-1}(x, \tau)) \geq g(\bar{u}_n(x, \tau)) \quad (54)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ gilt.

Dann kann man wiederum mit Satz 4 folgern, dass auch

$$\bar{u}_n(x, \tau) \geq \bar{u}_{n+1}(x, \tau) \quad (55)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ gilt.

Analog zur Konstruktion von \overline{u}_1 mithilfe von \overline{u}_0 erhalten wir nun die Existenz von \overline{u}_{n+1} und insgesamt

$$\overline{u}_0(x, \tau) \geq \overline{u}_n(x, \tau) \geq \overline{u}_{n+1}(x, \tau) \geq \underline{u}_0(x, \tau) \quad (56)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ sowie $\overline{u}_{n+1} \in C([0, T] \rightarrow H_{\mathbb{R}})$.

Mit dieser Induktion ist die Konstruktion der monotonen Iterationen abgeschlossen.

Durch obige Konstruktion erhalten wir also eine Folge von (schwachen) Oberlösungen $\overline{u}_0, \overline{u}_1, \overline{u}_2, \dots$ für das nichtlineare Problem (12a), (12b).

Mit dem Satz von der monotonen Konvergenz bzw. mit dem Satz von der majorisierten Konvergenz (siehe beispielsweise [9] (Kapitel 5)) erhält man zum einen die monotone punktweise Konvergenz

$$\overline{u}_n(x, \tau) \searrow u^*(x, \tau), \text{ wenn } n \rightarrow \infty \quad (57)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ und zum anderen die Konvergenz

$$\|\overline{u}_n(\tau) - u^*(\tau)\|_H \rightarrow 0, \text{ wenn } n \rightarrow \infty \quad (58)$$

für fast alle $\tau \in (0, T)$, wobei $u^*(\tau) \equiv u^*(\cdot, \tau) \in H_{\mathbb{R}}$ eine Funktion ist, welche die Ungleichung

$$\overline{u}_0(x, \tau) \geq \overline{u}_1(x, \tau) \geq \overline{u}_2(x, \tau) \geq \dots \geq u^*(x, \tau) \geq \underline{u}_0(x, \tau) \quad (59)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ erfüllt.

Betrachtet man den Lebesgue-Raum $L^2([0, T] \rightarrow H_{\mathbb{R}})$, erhält man des Weiteren auch die Konvergenz

$$\|\overline{u}_n - u^*\|_{L^2([0, T] \rightarrow H_{\mathbb{R}})} := \left(\int_0^T \|\overline{u}_n(\tau) - u^*(\tau)\|_H^2 d\tau \right)^{\frac{1}{2}} \rightarrow 0 \text{ für } n \rightarrow \infty. \quad (60)$$

Die Resultate von A. Pazy in [12] (Kapitel 6, Theorem 1.2) liefern zusammen mit [8] (Kapitel 7.1) zum einen, dass $u^* \in C([0, T] \rightarrow H_{\mathbb{R}})$ und zum anderen, dass u^* eine (schwache) Lösung des nichtlinearen Problems (12a), (12b) ist.

Analog konstruiert man die Folge von (schwachen) Unterlösungen, indem man mit der in Satz 3 beschriebenen anfänglichen Unterlösung \underline{u}_0 beginnt und die Iterierten so berechnet wie in (48a), (48b) (ersetze im Schema jeweils \overline{u}_i durch \underline{u}_i). Auf diese Weise erhält man eine Folge von (schwachen) Unterlösungen $\underline{u}_0, \underline{u}_1, \underline{u}_2, \dots$, für das nichtlineare Problem (12a), (12b).

Analog erhält man die monotone punktweise Konvergenz

$$\underline{u}_n(x, \tau) \nearrow u^*(x, \tau), \text{ wenn } n \rightarrow \infty \quad (61)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$, wobei

$$\underline{u}_0(x, \tau) \leq \underline{u}_1(x, \tau) \leq \underline{u}_2(x, \tau) \leq \dots \leq u^*(x, \tau) \leq \overline{u}_0(x, \tau) \quad (62)$$

für fast alle $(x, \tau) \in \mathbb{R} \times [0, T]$ gilt.

Die Konvergenz (58) sowie die Konvergenz (60) gelten analog.

3.3 Lösungsformel für das lineare Problem

In diesem Abschnitt wird eine Lösung des linearen Problems (31a), (31b) hergeleitet. Mithilfe bekannter Resultate über den Kern der Diffusionsgleichung kann dann sogar eine explizite Lösungsformel angegeben werden. Wir verwenden hierfür [1] (Kapitel 5) und [12] (Kapitel 5.6) sowie die Ergebnisse für die Diffusionsgleichung in [2] (Kapitel 3.4.4).

Wir beginnen mit der homogenen abstrakten Differentialgleichung

$$\begin{cases} \frac{\partial}{\partial \tau} u - \mathcal{A}(\tau)u = 0 \text{ in } H_{\mathbb{R}} \text{ für } \tau \in (0, T) \\ u(0) = h \in H_{\mathbb{R}}. \end{cases} \quad (63a)$$

Um eine Lösung für (63a), (63b) anzugeben, teilen wir zunächst den Black-Scholes-Operator $\mathcal{A}(\tau) : D_{\mathbb{C}} \subset H_{\mathbb{C}} \rightarrow H_{\mathbb{C}}$ in zwei Operatoren auf. Dabei besteht der Definitionsbereich $D(\mathcal{A}(\tau)) = D_{\mathbb{C}}$ aus allen Funktionen $f \in H_{\mathbb{C}}$, deren schwachen Ableitungen $f' = \frac{d}{dx}f$ und $f'' = \frac{d^2}{dx^2}f$ ebenfalls in $H_{\mathbb{C}}$ liegen.

Nun definieren wir den **Diffusions-Operator** $\mathcal{A}_1(\tau) : D_{\mathbb{C}} \subset H_{\mathbb{C}} \rightarrow H_{\mathbb{C}}$ durch

$$(\mathcal{A}_1(\tau)u)(x) := \frac{1}{2} [\sigma(\tau)]^2 \frac{\partial^2}{\partial x^2} u \quad \text{für } x \in \mathbb{R} \quad (64)$$

und den **Drift-Operator** $\mathcal{A}_2(\tau) : D_{\mathbb{C}} \subset H_{\mathbb{C}} \rightarrow H_{\mathbb{C}}$ durch

$$(\mathcal{A}_2(\tau)u)(x) := \left(q_S(\tau) - \gamma_S(\tau) - \frac{1}{2} [\sigma(\tau)]^2 \right) \frac{\partial}{\partial x} u \quad \text{für } x \in \mathbb{R}. \quad (65)$$

Somit erhalten wir $\mathcal{A}(\tau) = \mathcal{A}_1(\tau) + \mathcal{A}_2(\tau)$ als Summe zweier kommutativer Differential-Operatoren. Wir definieren nun noch die Hilfsfunktionen

$$\mathcal{S}(\tau) := \frac{1}{2} \int_0^\tau [\sigma(t)]^2 dt \quad \text{und} \quad \mathcal{R}(\tau) := \int_0^\tau \left(q_S(t) - \gamma_S(t) - \frac{1}{2} [\sigma(t)]^2 \right) dt \quad (66)$$

für alle $\tau \in [0, T]$ und erhalten durch die Ergebnisse von [1] (Kapitel 5) und [12] (Kapitel 5.6, Theorem 6.8), dass die eindeutige klassische Lösung $u \in C([0, T] \rightarrow H_{\mathbb{R}})$ des homogenen Problems (63a), (63b) gegeben ist durch

$$u(\tau) = \mathcal{U}_{1,2}(\tau, 0)h \quad \text{in } H_{\mathbb{R}} \text{ für } \tau \in [0, T], \quad (67)$$

wobei die Evolutionsfamilie $\mathcal{U}_{1,2}(\tau, s) = \{\mathcal{U}_{1,2}(\tau, s) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}} : 0 \leq s \leq \tau \leq T\}$ von beschränkten linearen Operatoren gegeben ist durch

$$\begin{aligned} \mathcal{U}_{1,2}(\tau, s) &:= \mathcal{U}_1(\tau, s)\mathcal{U}_2(\tau, s) \\ &:= \exp\left([\mathcal{S}(\tau) - \mathcal{S}(s)] \frac{\partial^2}{\partial x^2}\right) \exp\left([\mathcal{R}(\tau) - \mathcal{R}(s)] \frac{\partial}{\partial x}\right) \end{aligned} \quad (68)$$

für $0 \leq s \leq \tau \leq T$.

Nun betrachten wir die „vollständige“ homogene Differentialgleichung

$$\begin{cases} \frac{\partial}{\partial \tau} u - \mathcal{A}(\tau)u + r_g u = 0 \text{ in } H_{\mathbb{R}} \text{ für } \tau \in (0, T) \\ u(0) = h \in H_{\mathbb{R}}. \end{cases} \quad (69a)$$

$$(69b)$$

und definieren nun die Evolutionsfamilie $\mathcal{U}(\tau, s) = \{\mathcal{U}(\tau, s) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}} : 0 \leq s \leq \tau \leq T\}$ durch

$$\mathcal{U}(\tau, s) := \exp(-r_g(\tau - s)) \cdot \mathcal{U}_{1,2}(\tau, s) \quad (70)$$

für $0 \leq s \leq \tau \leq T$ und erhalten die eindeutige klassische Lösung $u \in C([0, T] \rightarrow H_{\mathbb{R}})$ des homogenen Problems (69a), (69b) durch

$$u(\tau) = \mathcal{U}(\tau, 0)h \text{ in } H_{\mathbb{R}} \text{ für } \tau \in [0, T]. \quad (71)$$

Zuletzt betrachten wir nun die inhomogene abstrakte Differentialgleichung

$$\begin{cases} \frac{\partial}{\partial \tau} u - \mathcal{A}(\tau)u + r_g u = F(\tau) \text{ in } H_{\mathbb{R}} \text{ für } \tau \in (0, T) \\ u(0) = h \in H_{\mathbb{R}}. \end{cases} \quad (72a)$$

$$(72b)$$

An dieser Stelle erinnern wir daran, dass $F : [0, T] \rightarrow H_{\mathbb{R}}$ eine Hölder-stetige Funktion zum Hölder-Exponenten $\vartheta_F \in (0, 1)$ ist. Mithilfe von [12] (Kapitel 5.6, Theorem 7.1) können wir die eindeutige klassische Lösung $u \in C([0, T] \rightarrow H_{\mathbb{R}})$ des inhomogenen abstrakten Problems (72a), (72b) durch die folgende Variation-der-Konstanten-Formel angeben:

$$u(\tau) = \mathcal{U}(\tau, 0)h + \int_0^{\tau} \mathcal{U}(\tau, t)F(t)dt \text{ in } H_{\mathbb{R}} \text{ für } \tau \in [0, T], \quad (73)$$

wobei die Evolutionsfamilie $\mathcal{U}(\tau, s) = \{\mathcal{U}(\tau, s) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}} : 0 \leq s \leq \tau \leq T\}$ durch (70) definiert ist.

Jede Abbildung $\mathcal{U}(\tau, s) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}}$ ($0 \leq s < \tau \leq T$) ist ein Integral-Operator mit dem Kern $\mathcal{K}(x, y; \tau, s)$, für den wir eine explizite Formel bestimmen möchten. Dafür nutzen wir den bekannten **Gauß-Kern**

$$\mathcal{G}(x; \tau) := \frac{1}{\sqrt{4\pi\tau}} \exp\left(-\frac{x^2}{4\tau}\right) \text{ für } (x, \tau) \in \mathbb{R} \times (0, \infty), \quad (74)$$

für den wir auf [2] (Kapitel 3.4.4) verweisen. Außerdem betrachten wir in Anlehnung an (69a), (69b) nun für $s \in [0, T]$ das folgende Problem auf dem Intervall $[s, T]$:

$$\begin{cases} \frac{\partial}{\partial \tau} u - \mathcal{A}(\tau)u + r_g u = 0 \text{ in } H_{\mathbb{R}} \text{ für } \tau \in (s, T) & (75a) \\ u(s) = h_s \in H_{\mathbb{R}}. & (75b) \end{cases}$$

Das Problem (75a), (75b) besitzt die klassische Lösung $u \in C([s, T] \rightarrow H_{\mathbb{R}})$, welche gegeben ist durch

$$u(\tau) = \mathcal{U}(\tau, s)h_s \text{ in } H_{\mathbb{R}} \text{ für } \tau \in [s, T]. \quad (76)$$

Die Diffusions-Halbgruppe $\mathcal{U}_1 = \{\mathcal{U}_1(\tau) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}} : 0 \leq \tau < \infty\}$ besteht aus Integral-Operatoren $\mathcal{U}_1(\tau) = \exp\left(\tau \frac{\partial^2}{\partial x^2}\right) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}}$ ($0 \leq \tau < \infty$) mit dem Kern $\mathcal{G}(x - y; \tau)$, d.h.

$$[\mathcal{U}_1(\tau)h](x) = \int_{\mathbb{R}} \mathcal{G}(x - y; \tau)h(y)dy \quad (77)$$

für alle $x \in \mathbb{R}$, $\tau \in (0, \infty)$ und $h \in H_{\mathbb{R}}$.

Damit erhalten wir für alle $0 \leq s < \tau \leq T$ den Integraloperator

$$[\mathcal{U}_1(\tau, s)h_s](x) = \int_{\mathbb{R}} \mathcal{G}(x - y; \mathcal{S}(\tau) - \mathcal{S}(s))h_s(y)dy \quad (78)$$

für alle $x \in \mathbb{R}$ und $h_s \in H_{\mathbb{R}}$.

Die Translations-Gruppe $\mathcal{U}_2 = \{\mathcal{U}_2(\tau) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}} : \tau \in \mathbb{R}\}$ besteht aus Translations-Operatoren $\mathcal{U}_2(\tau) = \exp\left(\tau \frac{\partial}{\partial x}\right) : H_{\mathbb{R}} \rightarrow H_{\mathbb{R}}$ ($-\infty < \tau < \infty$) mit

$$[\mathcal{U}_2(\tau)h](x) = h(x + \tau) \quad (79)$$

für alle $x \in \mathbb{R}$, $\tau \in \mathbb{R}$ und $h \in H_{\mathbb{R}}$.

Damit erhalten wir für alle $0 \leq s \leq \tau \leq T$ den Translations-Operator

$$[\mathcal{U}_2(\tau, s)h_s](x) = h_s(x + \mathcal{R}(\tau) - \mathcal{R}(s)) \quad (80)$$

für alle $x \in \mathbb{R}$ und $h_s \in H_{\mathbb{R}}$.

Somit erhalten wir also für $0 \leq s < \tau \leq T$ den Integraloperator

$$[\mathcal{U}_{1,2}(\tau, s)h_s](x) = \int_{\mathbb{R}} \mathcal{G}(x - y + \mathcal{R}(\tau) - \mathcal{R}(s); \mathcal{S}(\tau) - \mathcal{S}(s))h_s(y)dy \quad (81)$$

für alle $x \in \mathbb{R}$ und $h_s \in H_{\mathbb{R}}$.

Insgesamt erhalten wir also für das Problem (75a), (75b) die klassische Lösung $u \in C([s, T] \rightarrow H_{\mathbb{R}})$ durch

$$u(\tau) = [\mathcal{U}(\tau, s)h_s](x) = \int_{\mathbb{R}} \mathcal{K}(x, y; \tau, s)h_s(y)dy \quad (82)$$

für alle $x \in \mathbb{R}$, $0 \leq s < \tau \leq T$ und $h_s \in H_{\mathbb{R}}$, wobei der Kern $\mathcal{K}(x, y; \tau, s)$ gegeben ist durch

$$\mathcal{K}(x, y; \tau, s) := \exp(-r_g(\tau - s)) \mathcal{G}(x - y + \mathcal{R}(\tau) - \mathcal{R}(s); \mathcal{S}(\tau) - \mathcal{S}(s)) \quad (83)$$

für alle $x, y \in \mathbb{R}$ und $0 \leq s < \tau \leq T$.

4 Numerische Behandlung des Iterationsschemas

In diesem Abschnitt wird die Finite-Differenzen-Methode auf das monotone Iterationsschema angewendet. Dafür werden die Resultate von C. Munz und T. Westermann aus [11] sowie von F. Baustian, M. Fencl, J. Pospíšil und V. Švígler aus [3] benutzt. Wir nehmen ab jetzt an, dass σ , q_S und γ_S positive Konstanten sind. Damit ist der bisher zeitabhängige Black-Scholes-Operator $\mathcal{A}(\tau)$ nun ein zeitunabhängiger Operator \mathcal{A} .

4.1 Finite-Differenzen-Methode

Für die erste Iteration betrachten wir in Anlehnung an das lineare Problem (48a), (48b) mit gegebener anfänglicher Ober- bzw. Unterlösung u_0 folgendes Problem:

$$\begin{cases} \frac{\partial}{\partial \tau} u - \mathcal{A}u + r_g u = g(u_0) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ u(x, 0) = h(x) = H(e^x) & \text{für } x \in \mathbb{R}. \end{cases} \quad (84a)$$

$$(84b)$$

Dafür betrachten wir zum einen das Intervall $[x_{\min}, x_{\max}]$ für x , wobei $x_{\min}, x_{\max} \in \mathbb{R}$ mit $x_{\min} < x_{\max}$ und zum anderen das Intervall $[0, T]$ für τ . Man unterteile die Intervalle äquidistant, sodass

$$x_i = x_{\min} + i \cdot \delta_x, \quad i = 0, \dots, N, \quad (85)$$

wobei $N \in \mathbb{N}$ und die konstante Ortsschrittweite $\delta_x = \frac{x_{\max} - x_{\min}}{N} = x_{i+1} - x_i > 0$ für $i = 0, \dots, N-1$ gegeben ist. Es gilt also insbesondere $x_0 = x_{\min}$ und $x_N = x_{\max}$. Analog unterteile man das Intervall $[0, T]$, sodass

$$\tau_k = k \cdot \delta_\tau, \quad k = 0, \dots, M \quad (86)$$

gilt, wobei $M \in \mathbb{N}$ und die konstante Zeitschrittweite $\delta_\tau = \frac{T}{M} = \tau_{k+1} - \tau_k > 0$ für $k = 0, \dots, M-1$ gegeben ist. Es gilt also insbesondere $\tau_0 = 0$ und $\tau_M = T$.

Nun sei $U_{i,k}$ der Näherungswert der Lösung von (84a), (84b) im Punkt (x_i, τ_k) und es sei $g_{i,k} = g(u_0(x_i, \tau_k))$. Die partiellen Ableitungen werden für $i = 1, \dots, N-1$, $k = 0, \dots, M-1$ durch folgende finite Differenzen approximiert:

$$\frac{\partial}{\partial \tau} u(x_i, \tau_k) \approx \frac{U_{i,k+1} - U_{i,k}}{\delta_\tau}, \quad (87)$$

$$\frac{\partial}{\partial x}u(x_i, \tau_k) \approx \frac{U_{i+1,k} - U_{i-1,k}}{2\delta_x}, \quad (88)$$

$$\frac{\partial^2}{\partial x^2}u(x_i, \tau_k) \approx \frac{U_{i-1,k} - 2U_{i,k} + U_{i+1,k}}{\delta_x^2}. \quad (89)$$

Mithilfe der finiten Differenzen ergibt sich aus (84a) die Gleichung

$$\begin{aligned} \frac{U_{i,k+1} - U_{i,k}}{\delta_\tau} - \frac{1}{2}\sigma^2 \frac{U_{i-1,k} - 2U_{i,k} + U_{i+1,k}}{\delta_x^2} \\ - \left(q_S - \gamma_S - \frac{1}{2}\sigma^2 \right) \frac{U_{i+1,k} - U_{i-1,k}}{2\delta_x} + r_g U_{i,k} = g_{i,k} \end{aligned} \quad (90)$$

für $i = 1, \dots, N-1$, $k = 0, \dots, M-1$.

Die Gleichung (90) lässt sich explizit nach $U_{i,k+1}$ auflösen. Man erhält

$$\begin{aligned} U_{i,k+1} = U_{i,k} + \delta_\tau \left[g_{i,k} + \frac{1}{2}\sigma^2 \frac{U_{i-1,k} - 2U_{i,k} + U_{i+1,k}}{\delta_x^2} \right. \\ \left. + \left(q_S - \gamma_S - \frac{1}{2}\sigma^2 \right) \frac{U_{i+1,k} - U_{i-1,k}}{2\delta_x} - r_g U_{i,k} \right] \end{aligned} \quad (91)$$

für $i = 1, \dots, N-1$, $k = 0, \dots, M-1$.

Durch die Anfangsbedingung (84b) erhält man

$$U_{i,0} = h(x_i), \quad i = 0, \dots, N. \quad (92)$$

4.2 Randbedingungen

Zuletzt werden noch Randbedingungen benötigt, um auch Werte für $U_{0,k}$ und $U_{N,k}$ ($k = 1, \dots, M$) zu erhalten. Da diese nicht vorgegeben sind, müssen wir sie uns sinnvoll überlegen. Wir beginnen mit der Randbedingung bei x_{\min} . Dafür betrachten wir zunächst in Anlehnung an [3] (Kapitel 3.3) die ursprüngliche partielle Differentialgleichung (1), dividieren durch S^2 und betrachten den Grenzwert für $S \rightarrow \infty$. Dadurch erhält man

$$\lim_{S \rightarrow \infty} \frac{\partial^2}{\partial S^2} U(S, t) = 0 \quad (93)$$

unter der Voraussetzung, dass das Wachstum von $U(S, t)$ in S kleiner als quadratisch im asymptotischen Verhalten ist.

Somit ist es sinnvoll, für ein hinreichend großes $S_{\max} > 0$ anzunehmen, dass

$$\frac{\partial^2}{\partial S^2} U(S_{\max}, t) = 0 \quad (94)$$

für alle $t \in [0, T]$ gilt. Daher wählen wir als Randbedingung

$$\frac{\partial^2}{\partial x^2} u(x_{\max}, \tau) = 0 \quad (95)$$

für alle $\tau \in [0, T]$. Um dies umzusetzen, nehme man an, dass der Punkt $x_{N+1} = x_N + \delta_x$ existiere. Mithilfe der finiten Differenz (89) erhält man dann für $i = N$ und $k = 1, \dots, M$

$$\frac{U_{N-1,k} - 2U_{N,k} + U_{N+1,k}}{\delta_x^2} = 0 \quad (96)$$

und damit insbesondere

$$U_{N+1,k} = 2U_{N,k} - U_{N-1,k}. \quad (97)$$

Betrachtet man nun (91) für $i = N$ und $k = 0, \dots, M - 1$ und setzt (96) und (97) ein, so erhält man

$$U_{N,k+1} = U_{N,k} + \delta_\tau \left[g_{N,k} + \left(q_S - \gamma_S - \frac{1}{2}\sigma^2 \right) \frac{U_{N,k} - U_{N-1,k}}{\delta_x} - r_g U_{N,k} \right] \quad (98)$$

für $k = 0, \dots, M - 1$.

Die Randbedingung bei x_{\min} legen wir fest, indem wir annehmen, dass die Funktion dort konstant der Anfangsbedingung entspricht, d.h.

$$u(x_{\min}, \tau) = u(x_{\min}, 0) = h(x_{\min}) \quad (99)$$

für alle $\tau \in [0, T]$. Somit setzen wir also

$$U_{0,k} = U_{0,0} \quad (100)$$

für $k = 1, \dots, M$.

4.3 Berechnungen für die unterschiedlichen Verträge

Wir fixieren nun die Parameter ähnlich wie in [5] (Kapitel 4):

Parameter	Wert	Bemerkung
q_S	0.1	
γ_S	0.02	
σ	0.2	
r	0.03	
λ_B	0.01	
λ_C	0.02	
R_B	0.7	
R_C	0.8	
s_F	0.01	
K	15	für die Call-Option und den Forward
K_s	15	für die Gap-Option
K_t	12	für die Gap-Option

Tabelle 3.1: Wahl der Parameter

Für alle folgenden Berechnungen mithilfe der finiten Differenzen setzen wir außerdem $T = 2$, sodass wir also $\tau \in [0, 2]$ betrachten. Darüber hinaus setzen wir $x \in [x_{\min}, x_{\max}] = [1, 4]$ und führen die Berechnungen mit $N = 100$ und $M = 2000$ durch. Anschließend stellen wir die Ergebnisse auf dem Intervall $S \in [K - 5, K + 5] = [10, 20]$ bzw. $x \in [\ln(10), \ln(20)]$ dar. Für die Gap-Option betrachten wir anschließend $x \in [\ln(7), \ln(20)]$. Alle Berechnungen wurden mit MATLAB durchgeführt.

4.3.1 Call-Option Bei der Call-Option ist eine anfängliche Unterlösung laut Satz 3 durch

$$\underline{u}_0(x, \tau) \equiv 0 \text{ für } (x, \tau) \in \mathbb{R} \times [0, T] \quad (101)$$

gegeben. Für die erste Iterierte $\underline{u}_1 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ betrachten wir also

$$\begin{cases} \frac{\partial}{\partial \tau} \underline{u}_1 - \mathcal{A}\underline{u}_1 + r_g \underline{u}_1 = 0 & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \underline{u}_1(x, 0) = (e^x - K)^+ & \text{für } x \in \mathbb{R}. \end{cases} \quad (102a)$$

$$(102b)$$

Zusammen mit den in Kapitel 4.2 erwähnten Randbedingungen kann nun die Methode der finiten Differenzen angewendet werden, um eine Näherungslösung für \underline{u}_1 zu erhalten. Diese ist in Abbildung 5 dargestellt.

Für die zweite Iterierte $\underline{u}_2 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ betrachten wir

$$\begin{cases} \frac{\partial}{\partial \tau} \underline{u}_2 - \mathcal{A}\underline{u}_2 + r_g \underline{u}_2 = g(\underline{u}_1) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \underline{u}_2(x, 0) = (e^x - K)^+ & \text{für } x \in \mathbb{R}. \end{cases} \quad (103a)$$

$$(103b)$$

Mithilfe der eben berechneten Näherungslösung für \underline{u}_1 könnte man nun mit denselben Randbedingungen erneut die Methode der finiten Differenzen anwenden. Dies würde aber wieder zu derselben Näherungslösung führen, denn:

Sei $U_{i,k}$ der Näherungswert für \underline{u}_1 im Punkt (x_i, τ_k) . Es gilt $U_{i,k} \geq 0$ für $i = 0, \dots, N$ und $k = 0, \dots, M$. Für die neue Iteration wäre nun $g_{i,k} = g(U_{i,k})$. Wegen der spezifischen Wahl der Parameter gilt hier

$$L_f = \max \{(1 - R_B)\lambda_B, (1 - R_C)\lambda_C + s_F\} = (1 - R_C)\lambda_C + s_F \quad (104)$$

und damit

$$\begin{aligned} g(U_{i,k}) &= (1 - R_B)\lambda_B U_{i,k}^- - (1 - R_C)\lambda_C U_{i,k}^+ - s_F U_{i,k}^+ + L_f U_{i,k} \\ &= (1 - R_B)\lambda_B U_{i,k}^- - (1 - R_C)\lambda_C U_{i,k}^+ - s_F U_{i,k}^+ + ((1 - R_C)\lambda_C + s_F) U_{i,k} \\ &= -(1 - R_C)\lambda_C U_{i,k}^+ - s_F U_{i,k}^+ + ((1 - R_C)\lambda_C + s_F) U_{i,k}^+ \\ &= 0. \end{aligned} \quad (105)$$

Also ergibt sich mithilfe der finiten Differenzen für alle weiteren Iterationen genau dieselbe Rechnung wie bei der ersten Iteration und damit stimmen alle weiteren Näherungslösungen mit der Näherungslösung von \underline{u}_1 überein.

Analog betrachten wir nun die anfängliche Oberlösung \overline{u}_0 aus Satz 2

$$\overline{u}_0(x, \tau) = \exp(x + \bar{\lambda}\tau) \quad \text{für } (x, \tau) \in \mathbb{R} \times [0, T], \quad (106)$$

wobei wir hier $\bar{\lambda} = q_S - \gamma_S - r - (1 - R_C)\lambda_C - s_F = 0.036$ wählen. In Abbildung 4 ist diese anfängliche Oberlösung abgebildet. Es sei an dieser Stelle noch einmal angemerkt, dass diese für die drei verschiedenen Verträge dieselbe ist.

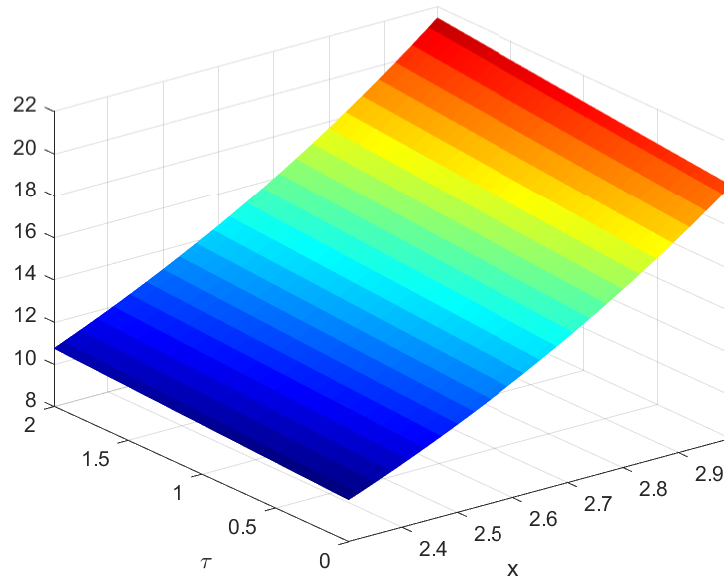


Abbildung 4: Anfängliche Oberlösung \overline{u}_0

Dann betrachten wir für die erste Iterierte $\overline{u}_1 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ nun

$$\begin{cases} \frac{\partial}{\partial \tau} \overline{u}_1 - \mathcal{A}\overline{u}_1 + r_g \overline{u}_1 = g(\overline{u}_0) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \overline{u}_1(x, 0) = (e^x - K)^+ & \text{für } x \in \mathbb{R}. \end{cases} \quad (107a)$$

$$(107b)$$

Wir wenden nun die Methode der finiten Differenzen mit denselben Randbedingungen wie bei der Unterlösung an. Allerdings gilt analog zur Rechnung in (105) auch $g(\overline{u}_0) = 0$, da $\overline{u}_0(x, \tau) > 0$ für alle $(x, \tau) \in \mathbb{R} \times [0, T]$. Daher stimmt die Näherungslösung für \overline{u}_1 mit der eben errechneten Näherungslösung für \underline{u}_1 überein.

Ganz analog zu der Argumentation bei den Unterlösungen stimmen alle weiteren Näherungslösungen mit der Näherungslösung von \overline{u}_1 überein. Folglich ist es nicht sinnvoll, an dieser Stelle weitere Iterationen durchzuführen.

Da die Call-Option eine nichtnegative Auszahlung besitzt, kann in diesem Fall eine analytische Lösung für das Problem (12a), (12b) angegeben werden. In [5] (Kapitel 3.2) ist eine Formel für die analytische Lösung gegeben. In Abbildung 6 ist diese dargestellt. Im Vergleich zu der errechneten Näherungslösung (Abbildung 5, Abbildung 6) sind nur leichte Unterschiede zu erkennen.

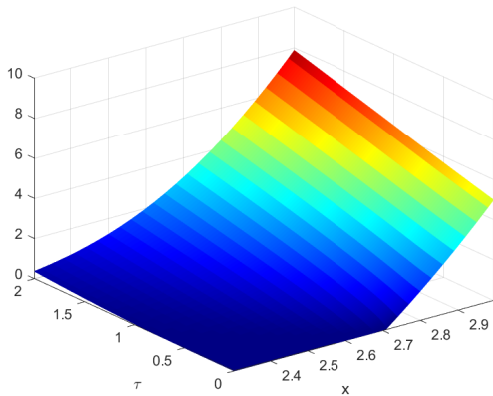


Abbildung 5: Näherungslösung für \underline{u}_1

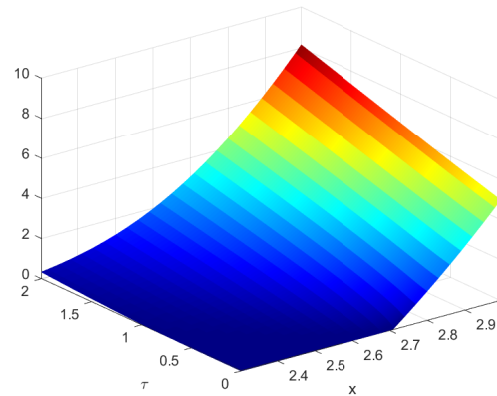


Abbildung 6: Analytische Lösung

Wir betrachten die Differenz von der analytischen Lösung und der Näherungslösung (Abbildung 7) und stellen fest, dass insbesondere am rechten Rand der Fehler relativ groß ist. Ebenso ist ein kleiner Fehler an der Stelle zu erkennen, an der die Auszahlung der Call-Option positiv wird.

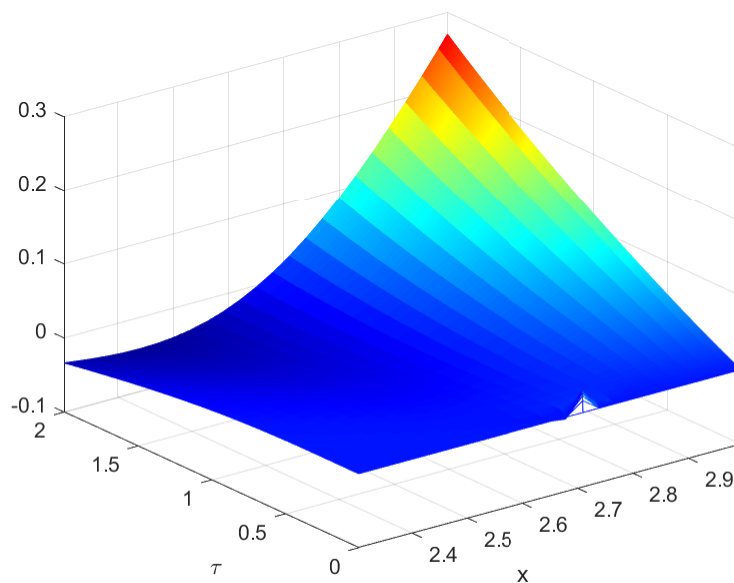


Abbildung 7: Differenz der analytischen Lösung und der Näherungslösung von \underline{u}_1

4.3.2 Forward Bei dem Forward ist eine anfängliche Unterlösung laut Satz 3 durch

$$\underline{u}_0(x, \tau) = -K \exp(\underline{\lambda}\tau) \text{ für } (x, \tau) \in \mathbb{R} \times [0, T] \quad (108)$$

gegeben, wobei wir $\underline{\lambda} = (1 - R_B) \lambda_B - r = -0.027$ setzen. In Abbildung 8 ist diese anfängliche Unterlösung dargestellt.

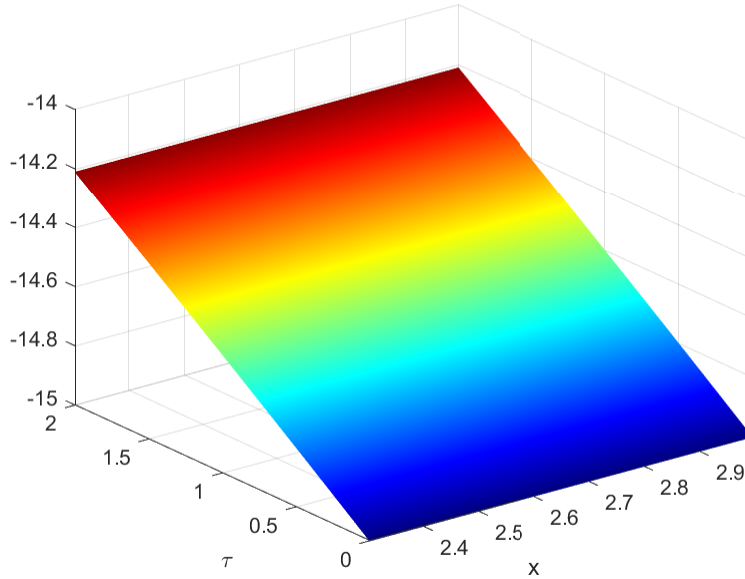


Abbildung 8: Anfängliche Unterlösung \underline{u}_0 beim Forward

Für die erste Iterierte $\underline{u}_1 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ betrachten wir

$$\begin{cases} \frac{\partial}{\partial \tau} \underline{u}_1 - \mathcal{A} \underline{u}_1 + r_g \underline{u}_1 = g(\underline{u}_0) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \underline{u}_1(x, 0) = e^x - K & \text{für } x \in \mathbb{R} \end{cases} \quad (109a)$$

$$\quad (109b)$$

und wenden die Methode der finiten Differenzen an. Mit der erhaltenen Näherungslösung kann dann die nächste Iterierte errechnet werden usw. Die ersten beiden Iterierten sind in Abbildung 9 und Abbildung 10 dargestellt. Es ist festzustellen, dass diese beiden sich nicht signifikant voneinander unterscheiden.

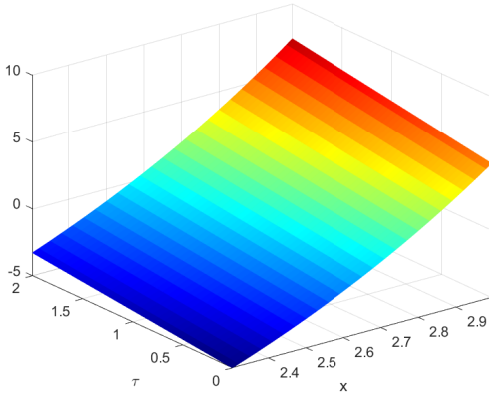
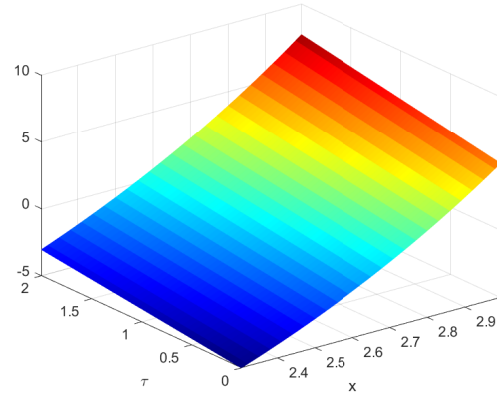
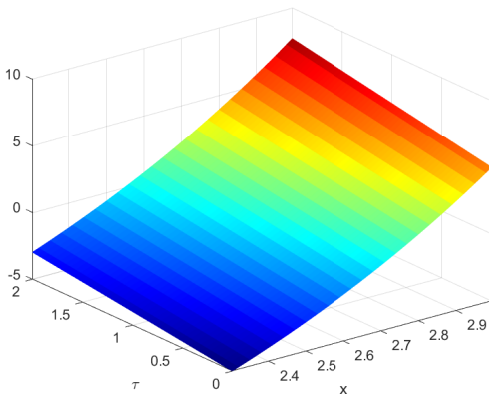
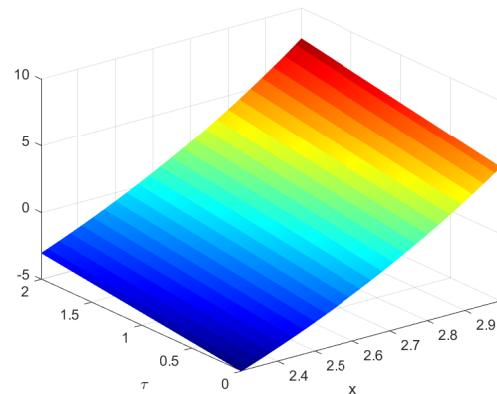
Analog betrachten wir bei den Oberlösungen die erste Iterierte $\overline{u}_1 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$

$$\begin{cases} \frac{\partial}{\partial \tau} \overline{u}_1 - \mathcal{A} \overline{u}_1 + r_g \overline{u}_1 = g(\overline{u}_0) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \overline{u}_1(x, 0) = e^x - K & \text{für } x \in \mathbb{R}, \end{cases} \quad (110a)$$

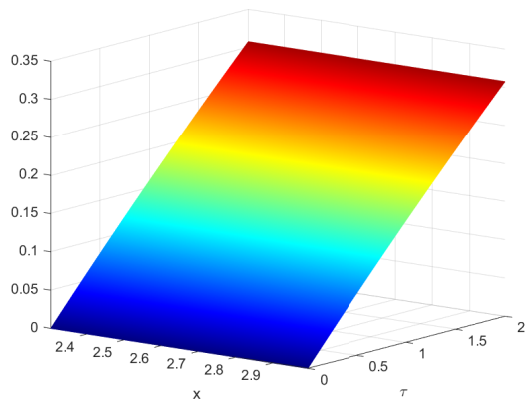
$$\quad (110b)$$

wobei die anfängliche Oberlösung \overline{u}_0 aus Satz 2 für alle betrachteten Verträge dieselbe ist. Die Randwerte werden auch genau so gewählt wie bei den Unterlösungen, sodass die Methode der

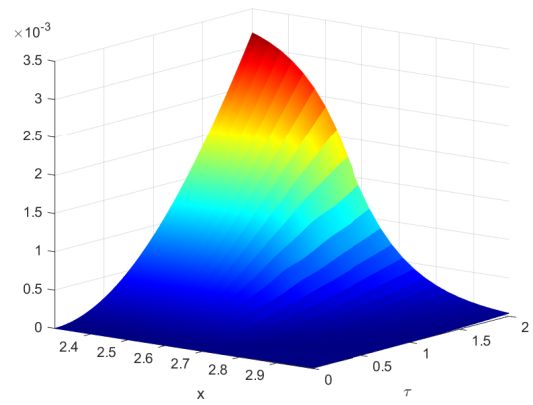
finiten Differenzen angewendet werden kann. Analog kann dann wieder die nächste Iterierte berechnet werden usw. In Abbildung 11 und Abbildung 12 sind diese dargestellt. Beim Vergleich aller berechneten Näherungslösungen (Abbildung 9 - Abbildung 12) fällt der sehr ähnliche Verlauf auf.

Abbildung 9: Näherungslösung für \underline{u}_1 Abbildung 10: Näherungslösung für \underline{u}_2 Abbildung 11: Näherungslösung für \overline{u}_1 Abbildung 12: Näherungslösung für \overline{u}_2

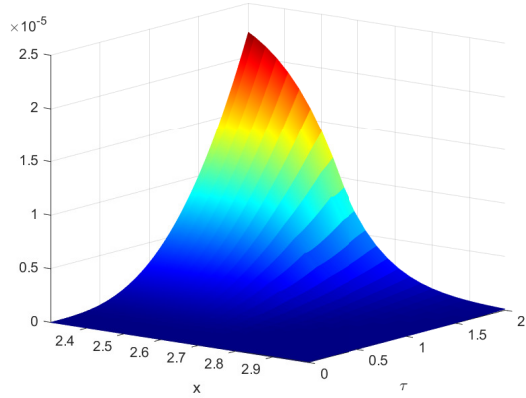
Da nun keine explizite analytische Lösung angegeben werden kann, betrachten wir stattdessen nach jedem Iterationsschritt die Differenz der Näherungslösung der Oberlösung und Unterlösung. Die Ergebnisse sind für die ersten fünf Iterationen in Abbildung 13 dargestellt. Zunächst kann man feststellen, dass die Differenzen immer nichtnegativ sind, sodass es nicht dazu kommt, dass die errechnete Oberlösung unterhalb der errechneten Unterlösung liegt. Des Weiteren stellt man in jedem Iterationsschritt fest, dass die Differenz kleiner wird, was nahelegt, dass Ober- und Unterlösung auch tatsächlich gegen dieselbe Funktion konvergieren. Jedoch kann man so keine verlässliche Aussage darüber treffen, inwieweit die errechneten Näherungslösungen die tatsächliche Lösung des nichtlinearen Problems widerspiegeln.



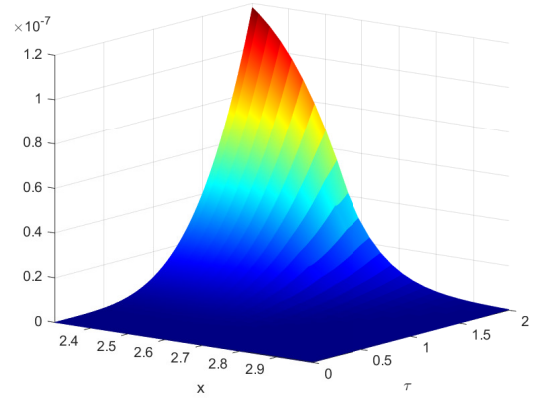
(a)



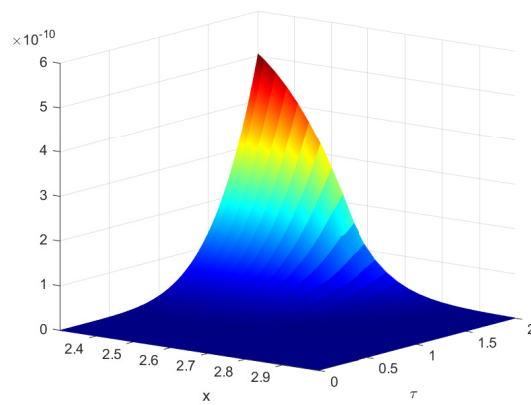
(b)



(c)



(d)



(e)

Abbildung 13: Differenz der Näherungslösungen von Ober- und Unterlösung nach (a) einer Iteration, (b) zwei Iterationen, (c) drei Iterationen, (d) vier Iterationen, (e) fünf Iterationen

4.3.3 Gap-Option Bei der Gap-Option ist eine anfängliche Unterlösung laut Satz 3 durch

$$\underline{u}_0(x, \tau) = (K_t - K_s) \exp(\underline{\lambda}\tau) \text{ für } (x, \tau) \in \mathbb{R} \times [0, T] \quad (111)$$

gegeben, wobei wir wieder $\underline{\lambda} = (1 - R_B) \lambda_B - r = -0.027$ setzen. Diese anfängliche Unterlösung ist in Abbildung 14 dargestellt.

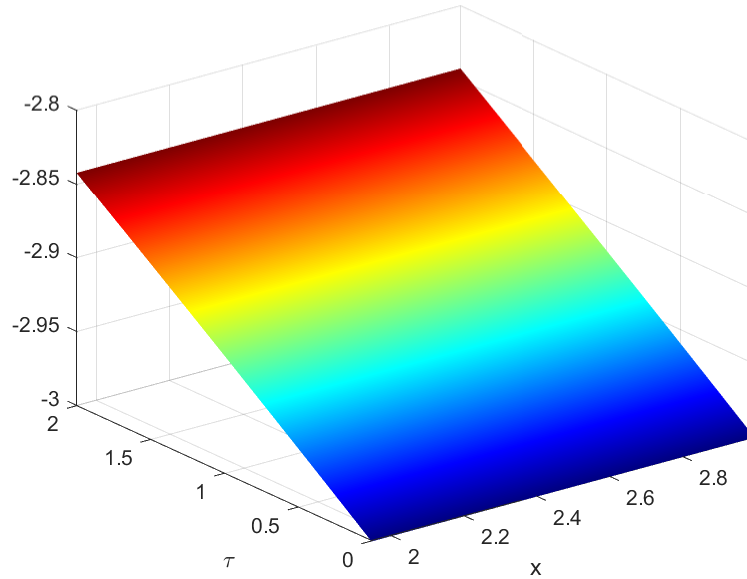


Abbildung 14: Anfängliche Unterlösung \underline{u}_0 bei der Gap-Option

Für die erste Iterierte $\underline{u}_1 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$ betrachten wir

$$\begin{cases} \frac{\partial}{\partial \tau} \underline{u}_1 - \mathcal{A} \underline{u}_1 + r_g \underline{u}_1 = g(\underline{u}_0) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \underline{u}_1(x, 0) = (e^x - K_s) \mathbb{1}_{\{x \geq \ln K_t\}} & \text{für } x \in \mathbb{R}. \end{cases} \quad (112a)$$

$$(112b)$$

Mit den gewohnten Randbedingungen kann die Methode der finiten Differenzen angewendet werden. Mit der erhaltenen Näherungslösung kann erneut die nächste Iterierte errechnet werden usw. Die ersten beiden Iterierten sind erneut abgebildet (Abbildung 15 und Abbildung 16).

Analog betrachten wir bei den Oberlösungen die erste Iterierte $\overline{u}_1 : \mathbb{R} \times [0, T] \rightarrow \mathbb{R}$

$$\begin{cases} \frac{\partial}{\partial \tau} \overline{u}_1 - \mathcal{A} \overline{u}_1 + r_g \overline{u}_1 = g(\overline{u}_0) & \text{für } (x, \tau) \in \mathbb{R} \times [0, T] \\ \overline{u}_1(x, 0) = (e^x - K_s) \mathbb{1}_{\{x \geq \ln K_t\}} & \text{für } x \in \mathbb{R}, \end{cases} \quad (113a)$$

$$(113b)$$

wobei die anfängliche Oberlösung \overline{u}_0 aus Satz 2 für alle betrachteten Verträge dieselbe ist. Die Randwerte werden auch wieder so gewählt wie bei den Unterlösungen, sodass die Methode der finiten Differenzen angewendet werden kann. Analog kann dann wieder die nächste

Iterierte berechnet werden usw. Auch hier stellen wir die ersten beiden Iterierten dar (Abbildung 17 und Abbildung 18). Die verschiedenen Ergebnisse weisen erneut keine signifikanten Unterschiede auf.

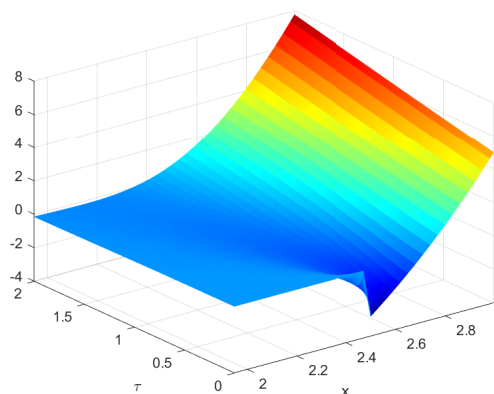


Abbildung 15: Näherungslösung für u_1

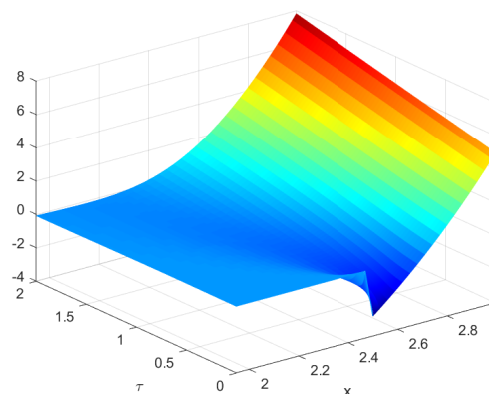


Abbildung 16: Näherungslösung für u_2

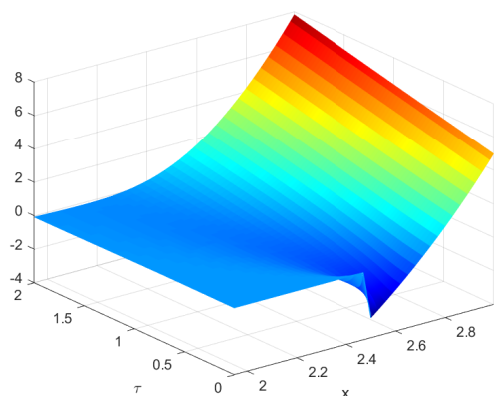


Abbildung 17: Näherungslösung für \bar{u}_1

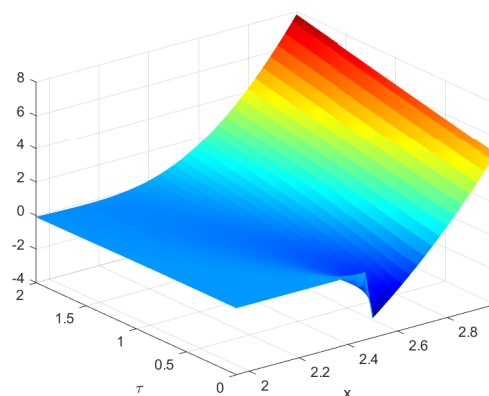
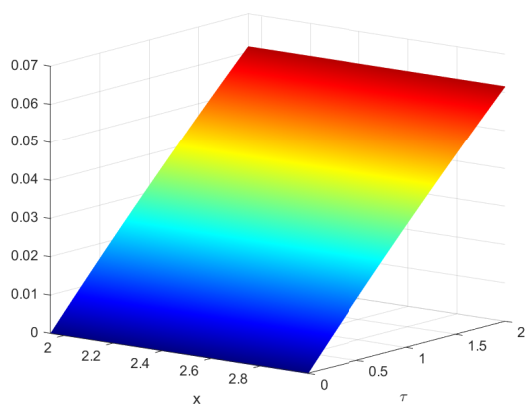


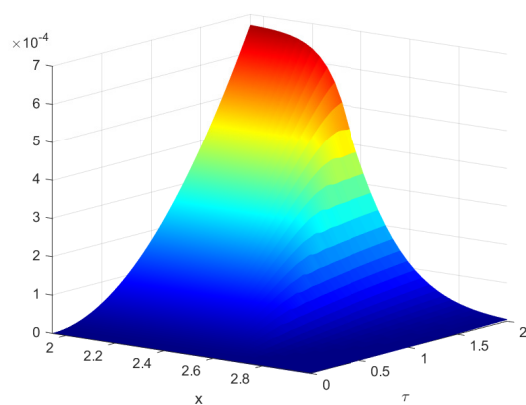
Abbildung 18: Näherungslösung für \bar{u}_2

Erneut betrachten wir nach jedem Iterationsschritt die Differenz der Näherungslösung der Oberlösung und Unterlösung (Abbildung 19). Die Ergebnisse lassen sich ähnlich wie beim Forward interpretieren.

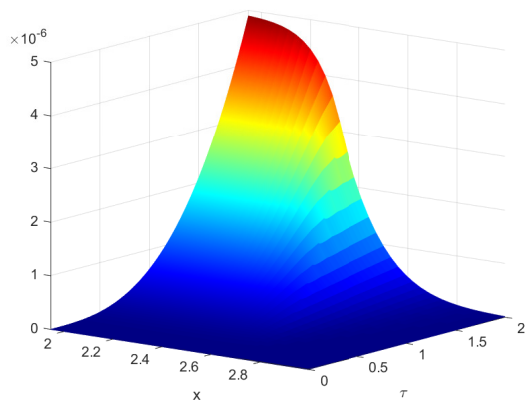
Die Differenzen sind auch hier immer nichtnegativ, sodass es nicht dazu kommt, dass die errechnete Oberlösung unterhalb der errechneten Unterlösung liegt. Ebenso wird auch hier die Differenz in jedem Iterationsschritt kleiner, was nahelegt, dass Ober- und Unterlösung auch tatsächlich gegen dieselbe Funktion konvergieren. Da aber auch hier keine analytische Lösung zum Vergleich vorliegt, kann keine verlässliche Aussage darüber getroffen werden, inwieweit die errechneten Näherungslösungen die tatsächliche Lösung des nichtlinearen Problems widerspiegeln.



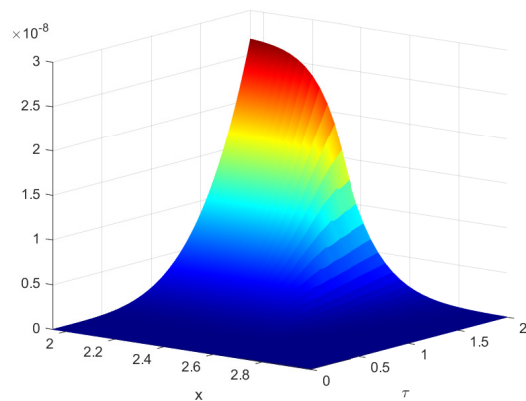
(a)



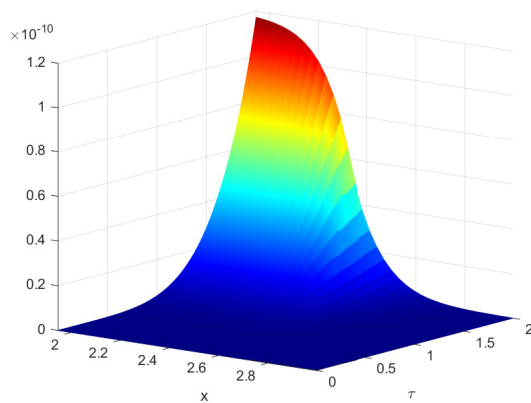
(b)



(c)



(d)



(e)

Abbildung 19: Differenz der Näherungslösungen von Ober- und Unterlösung nach (a) einer Iteration, (b) zwei Iterationen, (c) drei Iterationen, (d) vier Iterationen, (e) fünf Iterationen

5 Fazit

In diesem Artikel wurde zunächst das betrachtete Risiko-Modell vorgestellt. Da eine nicht-lineare partielle Differentialgleichung vorliegt, betrachten wir das monotone Iterationsschema, welches eine Folge von Ober- bzw. Unterlösungen konstruiert, die gegen die Lösung des nichtlinearen Problems konvergiert. Insbesondere liefert das monotone Iterationsschema den Vorteil, dass jeweils die vorherige Iterierte genutzt wird, damit das in jedem Iterationsschritt zu lösende Problem linear ist. Daher ist es insbesondere für ursprünglich nichtlineare Probleme (wie das hier betrachtete Risiko-Modell vom Black-Scholes-Typ) sehr sinnvoll, dieses Schema anzuwenden. Es wurde gezeigt, dass man für das lineare Problem sogar eine Lösungsformel angeben kann.

In der Praxis werden aber häufig numerische Verfahren verwendet, um die linearen Differentialgleichungen zu lösen. Daher wurde im Anschluss für drei Verträge (Call-Option, Forward und Gap-Option) versucht, das monotone Iterationsschema mithilfe der Finiten-Differenzen-Methode numerisch zu behandeln. Dafür mussten Parameter fixiert werden sowie sinnvolle Randbedingungen gefunden werden, sodass die Berechnungen durchgeführt werden können. Da die Call-Option eine nichtnegative Auszahlung besitzt (und daher eine analytische Lösung), konnte die errechnete Näherungslösung mit der analytischen Lösung verglichen werden. Die Ergebnisse zeigen zum einen, dass eine sinnvolle Näherung gefunden wurde, zum anderen aber auch, dass der Fehler am rechten Rand relativ groß ist. Dies legt die Vermutung nahe, dass die Randbedingungen (insbesondere die am rechten Rand) für Abweichungen sorgen.

Beim Forward und bei der Gap-Option haben wir die Differenz der Ober- und Unterlösung betrachtet, um die Qualität der Lösungen bewerten zu können. Die Ergebnisse legen nahe, dass beide gegen dieselbe Funktion konvergieren, allerdings lässt sich keine verlässliche Aussage darüber treffen, ob diese Funktion auch die Lösung des nichtlinearen Problems widerspiegelt.

Jedoch weisen alle errechneten Näherungslösungen einen ähnlichen qualitativen Verlauf wie die Ergebnisse in [5] (Kapitel 4) auf. Dort wurde eine Integralformel hergeleitet und mithilfe der Gauß-Quadratur die Näherungslösung nach fünf Iterationen berechnet. Da die Parameterwahl dort ähnlich ist und dieselben Verträge betrachtet wurden, sind die Ergebnisse vergleichbar.

Somit ist insgesamt anzunehmen, dass die numerischen Ergebnisse dieser Arbeit sinnvoll sind und somit die Finite-Differenzen-Methode (mit den festgelegten Randbedingungen) für das Problem geeignet ist.

Es ist möglich, dass durch andere Randbedingungen oder aber auch durch andere nume-

rische Methoden zur Behandlung partieller Differentialgleichungen, zum Beispiel die Finite-Elemente-Methode, neue Ergebnisse erzielt werden können.

Literaturverzeichnis

- [1] **Alziary, B.**, and **Takáč, P.** : *Monotone methods in counterparty risk models with non-linear Black-Scholes-type equations*. SeMA Journal, 2022
- [2] **Arendt, W.**, and **Urban, K.** : *Partielle Differenzialgleichungen*. Spektrum Akademischer Verlag, 2010
- [3] **Baustian, F.**, **Fencl, M.**, **Pospíšil, J.**, and **Švígler, V.** : *A Note on a PDE Approach to Option Pricing Under xVA*. WILMOTT magazine, 60-69, 2022
- [4] **Baustian, F.**, **Filipová, K.**, and **Pospíšil, J.** : *Solution of option pricing equations using orthogonal polynomial expansion*. Application of Mathematics, 66 (2021), No. 4, 553-582, 2021
- [5] **Baustian, F.**, **Pospíšil, J.**, and **Švígler, V.** : *Monotone iteration scheme for nonlinear PDEs in risk models*. arXiv:2306.17320 [math.AP], 2023
- [6] **Burgard, C.**, and **Kjaer, M.** : *PDE representation of derivatives with bilateral counterparty risk and funding costs*. J. Credit Risk 7(3), 75-93, 2011
- [7] **Drábek, P.**, and **Holubová, G.** : *Elements of Partial Differential Equations*. de Gruyter, 2014
- [8] **Evans, L.C.** : *Partial Differential Equations*. American Mathematical Society, 2010
- [9] **Forster, O.** : *Analysis 3*. Springer Spektrum, 2017
- [10] **Friedman, A.** : *Partial Differential Equations of Parabolic Type*. Dover Publications, 2008
- [11] **Munz, C.-D.**, and **Westermann, T.** : *Numerische Behandlung gewöhnlicher und partieller Differenzialgleichungen*. Springer Vieweg, 2012
- [12] **Pazy, A.** : *Semigroups of Linear Operators and Application to Partial Differential Equations*. Springer, 1983

received: November 5, 2024

Author:

Colin Buttchereit
Institut für Mathematik,
Universität Rostock,
Ulmenstraße 69, Haus 3,
18057 Rostock,
Germany

e-mail: colin.buttchereit2@uni-rostock.de

Hints for Authors

Rostock. Math. Kolloq. appears once or twice per year.

Submission

Papers should be sent by e-mail to

romako@uni-rostock.de

or by mail to

Universität Rostock
Institut für Mathematik
D-18051 Rostock

We will only consider original contributions which are related to research areas at the University of Rostock. All papers will be reviewed.

AMS-Subject-Classification

Please add one or two AMS-classification numbers which describe the content of your paper.

Manuscript Format

Papers should be written in German or English. Please send your article as a Latex-file and a pdf-file or a ps-file. The Latex-file should not contain self defined commands. Textwidth should be 165 mm and font size 12pt.

Authors Adresses

Please add the current complete addresses of all authors including first name / surname / institution / department / street / house number / postal code / place / country / e-mail-address.

Bibliography

Current numbers within the text ([3], [4]; [7, 8, 9]) refer to the corresponding item in the bibliography at the end of the article, which has the headline References. Please follow the structure of the examples:

- [3] **Zariski, O.**, and **Samuel, P.**: *Commutative Algebra*. Princeton 1958
- [4] **Steinitz, E.**: *Algebraische Theorie der Körper*. J. Reine Angew. Math. **137**, 167-309 (1920)
- [8] **Gnedenko, B.W.**: *Über die Arbeiten von C.F. Gauß zur Wahrscheinlichkeitsrechnung*. In: Reichardt, H. (Ed.): C.F. Gauß, Gedenkband anlässlich des 100. Todestages. S. 193-204, Leipzig 1957

Each citation should be written in the original language. Cyrillic letters must be transliterated as it is usual in libraries.

Hinweise für Autoren

Das Rostock. Math. Kolloq. erscheint ein- bis zweimal pro Jahr.

Einreichen

Senden Sie bitte Ihre Arbeiten per e-mail an

`romako@uni-rostock.de`

oder per Post an

Universität Rostock
Institut für Mathematik
D-18051 Rostock

Wir berücksichtigen nur Originalarbeiten, die in Bezug zu einem der Rostocker Arbeitsgebiete stehen. Alle Arbeiten werden begutachtet.

AMS-Subject-Klassifikation

Bitte geben Sie zur inhaltlichen Einordnung der Arbeit ein bis zwei Klassifizierungsnummern (AMS-Subject-Classification) an.

Manuskript

Manuskripte sollen in Deutsch oder Englisch abgefasst sein. Bitte senden Sie uns Ihre Arbeit als \LaTeX - und als PDF-Datei bzw. PS-Datei. In der \LaTeX -Datei sollten selbst definierte Befehle vermieden werden. Die Textbreite sollte 165 mm betragen und die Schriftgröße 12pt.

Adressen der Autoren

Die aktuelle, vollständige Adresse des Autors sollte enthalten: Vornamen Name / Institution / Struktureinheit / Straße Hausnummer / Postleitzahl Ort / Land / e-mail-Adresse.

Literaturzitate

Literaturzitate sind im Text durch laufende Nummern (vgl. [3], [4]; [7, 8, 10]) zu kennzeichnen und am Schluss der Arbeit unter der Zwischenüberschrift **Literatur** zusammenzustellen. Hierbei ist die durch die nachfolgenden Beispiele veranschaulichte Form einzuhalten.

- [3] **Zariski, O.**, and **Samuel, P.**: *Commutative Algebra*. Princeton 1958
- [4] **Steinitz, E.**: *Algebraische Theorie der Körper*. J. Reine Angew. Math. **137**, 167-309 (1920)
- [8] **Gnedenko, B.W.**: *Über die Arbeiten von C.F. Gauß zur Wahrscheinlichkeitsrechnung*. In: Reichardt, H. (Ed.): C.F. Gauß, Gedenkband anlässlich des 100. Todestages. S. 193-204, Leipzig 1957

Die Angaben erfolgen in Originalsprache; bei kyrillischen Buchstaben sollte die (bibliothekarische) Transliteration verwendet werden.