1 Introduction

In this paper we present [1], where a strategy for mesh adaption for the numerical solution of our usual problem by collocation is described. After a short introduction to the problem, the collocation method as underlying numerical scheme and the error estimation, we will have a look at the error equidistribution principle, which is the important part of the mesh selection algorithm. Since we have an asymptotically correct estimation for the global error, we can adopt this strategy to equidistribute the global error. We will present the method for mesh selection step-by-step and try to explain the underlying ideas.

1.1 Collocation

There will be a talk about collocation methods, but we give a short definition. Consider a first-order system of ODE's

\[ z'(t) = f(t, z(t)), \quad a < t < b \]
\[ g(z(a), z(b)) = 0 \] (1.1)

**Definition 1.** For a given mesh \( \Delta = \{\tau_i : i = 0, \cdots N\} \) and points \( 0 \leq \rho_1 < \cdots < \rho_k \leq 1, \)

define the collocation points

\[ \tau_{ij} = \tau_i + h_i \rho_j, \quad 1 \leq j \leq k, \quad 1 \leq i \leq N - 1. \]

A collocation solution to (1.1) is a continuous function \( z_{\Delta} \) which reduces to a polynomial of order \( k + 1 \) on each mesh subinterval and satisfies

\[ z'_{\Delta}(\tau_{ij}) = f(\tau_{ij}, z_{\Delta}(\tau_{ij})), \quad 1 \leq j \leq k, \quad 1 \leq i \leq N - 1. \]
\[ g(z_{\Delta}(a), z_{\Delta}(b)) = 0 \]

i.e., the differential equation holds in the collocation points and the boundary condition is fulfilled.
1.2 Error estimation

Another talk will be about this topic, but since we need it here we give a short introduction. We are interested in an asymptotically correct a posteriori estimate for the global error of the numerical solution obtained by the collocation method. This estimate is based on the defect correction principle: a ‘neighboring problem’ to (1.1) is constructed, for which the exact solution is known. The numerical solution of (1.1), obtained by the collocation, is interpolated by a continuous piecewise polynomial $q$. We construct the problem

$$
\begin{align*}
y'(t) &= f(t, y(t)) + d(t), \quad a < t < b \\
g(z(a), z(b)) &= 0
\end{align*}
$$

where $d(t)$ is called defect:

$$
d(t) := q'(t) - f(t, q(t)).
$$

By construction, $q$ is the exact solution to (1.2). Now, if the error of the collocation solution is small, then we expect the problems (1.1) and (1.2) and their solutions to be closely related. (1.1) and (1.2) are then solved by a low order scheme, e.g., the backward Euler scheme at the mesh points, which yields the solutions $\xi$ and $\pi$. Then $\pi - \xi$ is used as an estimate for the global error of the collocation solution of the original problem (1.1). This approach is further extended by using the information computed at the collocation points and by using a modified defect term.

1.3 The Singular BVP

Consider the boundary value problem

$$
\begin{align*}
z'(t) &= \frac{M(t)}{t} z(t) + f(t, z(t)), \quad t \in (0, 1] \\
B_a z(0) + B_b z(1) &= \beta \\
z &\in C[0, 1]
\end{align*}
$$

where $z$ is a d-dimensional real function, $M$ is a smooth $n \times n$ matrix and $f$ is an $n$-dimensional smooth function. $B_a$ and $B_b$ are constant $r \times n$ matrices, where $r < n$. We solve (1.3) on a given mesh $\Delta := \{\tau_i : i = 0 \ldots N\}$ where $\tau_0 = 0$ and $\tau_N = 1$ by collocation at an even number $m$ of collocation points in the interior of every mesh interval, i.e.,

$$
\Delta^m := \{t_{i,j} = \tau_i + \rho_j (\tau_{i+1} - \tau_i), i = 0, \ldots, N - 1, j = 0, \ldots, m + 1\}
$$

with $\rho_j := j/(m + 1)$. This leads to equidistantly spaced collocation points $t_{i,j}$ in the interior of every mesh interval. The collocation solution is a continuous piecewise polynomial function

$$
p(t) := p_i(t), \quad t \in J_i := [\tau_i, \tau_{i+1}], \quad i = 0, \ldots, N - 1
$$
where every $p_i$ is a polynom of order $m + 1$, which satisfies

$$p_i'(t_{i,j}) = \frac{M(t_{i,j})}{t_{i,j}} p_i(t_{i,j}) + f(t_{i,j}, p_i(t_{i,j}))$$

$$p_i(\tau_i) = p_{i-1}(\tau_i)$$

$$B a p_0(0) + B b p_{N-1}(1) = \beta.$$ 

Throughout this paper we will use $h_i := \tau_{i+1} - \tau_i$, with $h_{\text{max}} := \max_i h_i$ and $h_{\text{min}} := \min_i h_i$. If we speak of an element $K$ of the mesh $\Delta = \{\tau_i : i = 0, \ldots, N\}$, we mean an interval of the form $K = [\tau_i, \tau_{i+1}]$. See [3] for details about the solution of this collocation method.

## 2 Error Equidistribution

Numerical Methods to solve BVPs on a computer involve a discretization of the differential equations on a mesh. A natural question which arises is how the mesh should be chosen so that the solution is sufficiently accurate and can be computed as inexpensively as possible. We can make this precise: we want to find a mesh with as few elements $N$ as possible. In practice it is usually too expensive, if not impossible to find a optimal mesh, i.e., a mesh which yields a minimal mesh size for a given error tolerance. Thus we search for good meshes, i.e., meshes with a size not much larger than the optimal one. The appropriateness of the discretized system is directly determined by the mesh, so the mesh must be fine in regions where the solution changes rapidly, but can be coarse elsewhere. We discuss here an adaptive type of mesh selection, this means that the mesh and the approximate solution are repeatedly updated until a prescribed tolerance is satisfied. Therefore we have to deal with the following two basic questions of adaptivity:

1. we want to ensure that the error is smaller than a given tolerance $TOL$.

2. conversely to the first question, we want to minimize the work, i.e., the number of collocation points.

As we stated above, usually one does not seek for an optimal mesh, but looks for a good one. The error equidistribution principle (EP) is a commonly used device which deals with the two questions stated above. The idea is, that a mesh $\Delta$ is chosen in a way that a error measure $e_i$ on the ith element of the mesh is equidistributed, i.e., every element of the mesh yields the same error. Imho it is more a heuristic principle, and really not an obvious way to go. Nevertheless there are problems, where it can easily be shown that the solutions do indeed fulfill the EP, so that the mesh should be chosen this
way. For this purpose we will state such a problem.

Given a mesh $\Delta$ on $[a,b]$ and $f \in C([a,b])$, define the Operator

$$I_{\Delta} : C([a,b]) \to S^{0,0}(\Delta)$$

$$f \mapsto (I_{\Delta}f)|_K := 1/2(\max_K f + \min_K f)$$

where

$$S^{0,0}(\Delta) := \{u|\forall K, K \text{ element in } \Delta : u|_K = \text{const}\}$$

i.e., $S^{0,0}$ consists of step functions, which are constant on every element of the mesh. Consider the following question: Given $N \in \mathbb{N}$, find a mesh $\Delta$ with $N$ elements such that

$$|f - I_{\Delta}f|_{\infty} = \min_{\Delta' \text{ has } N \text{ elements}} |f - I_{\Delta'}f|_{\infty} \tag{2.1}$$

**Theorem 1.** If $\Delta$ is a solution of (2.1), then $\Delta$ fulfills the EP, i.e., $\exists C \forall K, K \text{ element in } \Delta : |f - I_{\Delta}f|_{L^\infty(K)} = C$

**Proof.** Let $e_k := \min_{\lambda \in \mathbb{R}} ||f - \lambda||_{L^\infty(K)} = 1/2(\max_K f - \min_K f)$ and $e_{\Delta} := ||f - v||_{\infty}$. $f \in C([a,b])$ implies that the error $e_{\Delta}$ depends continuous on the knots of the element $K$, furthermore this dependence is monotone, i.e., $K \subset K'$ implies $e_{K} \leq e_{K'}$.

Now we assume that we have a mesh $\Delta$ and we show that, if $\Delta$ is a solution, it has to fulfill the EP. To that end, assume $\Delta$ is a solution of (2.1), but it does not fulfill the EP.

$$M_{\Delta} := |\{K'| K' \text{ element in } \Delta, e_{K'} = \max_K e_K\}|$$

i.e., $M$ is the number of elements on which the error reaches its maximum. Note that $M = N$ means that the EP is fulfilled. We assumed $M < N$, then there exist adjacent elements $K, K'$ with $e_{K'} < e_K$. Take the knot $x$ which connects $K$ and $K'$ and move it in the direction of $K$. Because of the continuous dependence on the knots we can achieve $e_K = e_{K'}$. This transformation gives us a new mesh $\Delta$ with $N$ elements, and since $\Delta$ is a solution, $e_{\Delta} \leq e_{\Delta'}$ holds. Therefore

$$M_{\Delta} \leq M_{\Delta} - 1$$

has to hold. This transformation shows that, after less then $N - 2$ steps, we obtain a mesh $\Delta'$ with

$$M_{\Delta'} = 1$$

and the same transformation as before yields a mesh $\Delta''$ with $e_{\Delta''} < e_{\Delta}$. This is a contradiction. $\square$
One can think about related questions, e.g., given the number of elements, what is the achievable error, and what are the properties of the mesh? The equidistribution of the error is a desirable property in the above example. Applied to other problems, the EP might be plausible, but as we said, is more a heuristic approach. Take the explicit euler scheme as a simple example:

\[ y_i = y_{i-1} + h_i f(t_{i-1}, y_{i-1}) \]

For the local discretization error

\[ l_i := \frac{y(t_i) - (y(t_{i-1}) + h_i f(t_{i-1}, y(t_{i-1})))}{h_i} \]

a simple Taylor expansion yields

\[ l_i = \frac{1}{2} h_i y''(t_{i-1}) + O(h_i^2). \]

For a m-th order method, this representation is of the general form

\[ l_i = h_i^m T(t_i) + O(h_i^{m+1}) \quad (2.2) \]

where the function \( T \) depends on certain derivatives of the true solution. This error representation depends on the mesh, but not in a crucial way, so this dependence is usually ignored.

### 2.1 Error Model

The representation (2.2) gives us a error model of the form

\[ h_i^m \|T(t)\|. \quad (2.3) \]

To achieve equidistribution we claim \( h_i^m \|T(t)\| \approx TOL \), i.e., the error is equidistributed and fulfills a tolerance. It turns out that it is convenient to consider a corresponding error measure which only depends linearly on \( h_i \), i.e.,

\[ h_i \|T(t_i)\|^{1/m} \approx TOL^{1/m} \quad (2.4) \]

Since \( T(t) \) is not available precisely, a local error estimate \( \bar{T}(t) \) must be available. The discrete requirement (2.4) is then replaced by the more practicable requirement

\[ \int_{t_{i-1}}^{t_i} \|\bar{T}(t)\|^{1/m} \approx TOL^{1/m} = \frac{1}{N} \int_0^1 \|\bar{T}(t)\|^{1/m} dt \quad (2.5) \]

which gives the number of grid points \( N \) required to meet the local error tolerance \( TOL \). In the following, we will use the same ideas for equidistribution of the global error.
3 Mesh Characterization

Suppose we are given a mesh $\Delta = \{\tau_i : i = 0, \ldots, N\}$. We will now characterize this mesh in terms of a function, its mesh density. This also gives us a good starting point for the following computations. First we give the formal definition: We define a continuous and piecewise affine function $g$ which satisfies

$$g(\tau_i) := \frac{i}{N}, \quad i = 0, \ldots, N.$$ 

The function $g$ is monotone increasing with values in $[0, 1]$ and has a piecewise constant derivative

$$\rho(t) := g'(t) = \frac{1}{Nh_i}, \quad i = 0, \ldots, N - 1.$$ 

It is easy to see that

$$\int_0^1 \rho(t)dt = 1. \quad (3.1)$$

If a positive function $\rho$ has the property (3.1), we say that $\rho$ is a mesh density. For a better understanding, consider the following figure.

On the vertical axis we have an equidistant mesh which is scaled to the grid $\Delta$ on the horizontal axis by the function $g$. By means of the figure it is easy to see that a bigger mesh density leads to a steeper function $g$, and consequently the mesh gets finer. Conversely, for a given positive step function $\rho$
(or any positive function) which fulfills (3.1) and a number of mesh points \( N \), the corresponding mesh \( \Delta \) can be reconstructed by integrating \( \rho \):

\[
g(t) := \int_0^t \rho(s) ds \tag{3.2a}
\]

\[
\tau_i := g^{-1} \left( \frac{i}{N} \right) \tag{3.2b}
\]

Note that in terms of a mesh density \( \rho \), the equidistribution of the local error writes as

\[
\rho(t) = \frac{|T(t)|^{1/m}}{TOL^{1/m} N}.
\]

If one’s goal is not error equidistribution but only to meet the error tolerance, it is also possible to have a fixed mesh density and just increase the number of mesh points. This method could be applied if knowledge of the shape of the solution is available a priori.

4 Mesh refinement

Now we consider the error model (2.3) for the global error and set

\[
e(t) := h_i^m |T(t)|. \tag{4.1}
\]

It is plausible to have error representation like that. The global error estimate which was described in §1.2 leads to a representation of this type. Our goal is to find \( N \in \mathbb{N} \) and a mesh \( \Delta \) with \( N + 1 \) knots, which leads to

1. equidistribution of the global error
2. fulfilled tolerance of the global error

A first approach would be to adapt the techniques for error equidistribution described in §2. The mesh (the step sizes) is determined by the mesh density, so we want to find a density function \( \rho \), which is of the form

\[
\rho(t) = \frac{1}{Nh(t)} \tag{4.2}
\]

where \( h \) is the stepwidth. The density should be chosen according to the equidistribution of the error model (2.3), this actually means

\[
\rho(t) := \frac{|T(t)|^{1/m}}{TOL^{1/m} N}. \tag{4.3}
\]
In order to obtain a mesh on the interval $[0, 1]$ we have to ensure that the function $\rho$ is a mesh density according to our definition. This means that $N$ has to be chosen such that $\int_0^1 \rho(t) dt = 1$, so this reads
\[
\frac{1}{N} TOL^{-1/m} \int_0^1 |T(t)|^{1/m} dt = 1
\]
Note that this is similar to the requirement (2.5).

If one thinks about an iterative solution process it is natural to think of (4.1) as an error model, which comes from an existing mesh $\Delta_{old}$ with density $\rho_{old}$ and a corresponding collocation solution $p_{old}$. Based on these preceding approximate solutions a new mesh $\Delta$ is computed. The sought mesh density $\rho$, given by (4.2), has then to fulfill (4.3), which can then be written as
\[
\rho(t) = \frac{1}{Nh_i} |T(t)|^{1/m} h_i TOL^{-1/m}
\]
where $\rho$ occurs on both sides of the equation. Now, if we replace $\rho$ on the right side of the equation by the available $\rho_{old}$, we obtain
\[
\rho(t) = \rho_{old}(t) |T(t)|^{1/m} h_i TOL^{-1/m}.
\]
One should think about $|T(t)|^{1/m} h_i$ as an error measure which is independent of $h_i$, therefore we define
\[
\Theta(t) := |T(t)|^{1/m} h_i.
\]
If we think of $\Theta$ as an error measure, (4.4) is very intuitive, because we take the old density $\rho_{old}$ and multiply it with an error measure, so that the density gets bigger where the error is bigger and consequently the mesh gets finer. We have to ensure that we obtain a density function by
\[
I := \int_0^1 \rho(t) dt
\]
and hence, to satisfy the tolerance, we must set the new number of elements
\[
\widetilde{N} := NI
\]
The new mesh $\widetilde{\Delta} = (\widetilde{\tau}_0, \ldots, \widetilde{\tau}_{\widetilde{N}})$ is then computed according to
\[
\widetilde{g}'(t) := \widetilde{\rho}(t),
\]
\[
\widetilde{\tau}_j := \widetilde{g}^{-1}\left(\frac{j}{\widetilde{N}}\right)
\]
Note that this leads to
\[ \tilde{g}(0) = 0, \quad \tilde{g}(1) = 1, \]
\[ \tilde{g}(\tau_j) = \frac{j}{N} \implies \frac{1}{N} = \tilde{g}(\tau_j) - \tilde{g}(\tau_{j-1}) = \int_{\tau_{j-1}}^{\tau_j} \tilde{\rho}(t)\,dt = \frac{1}{I} \int_{\tau_{j-1}}^{\tau_j} \rho(t)\,dt \]

A short outline of the algorithm reads as:

1: choose a starting mesh, compute collocation solution and error model (measure)
2: compute new density by (4.4), (4.5), (4.6)
3: compute collocation solution and error model
4: if satisfied, stop. if not, goto 2:

4.1 Further Comments

In certain contexts such as acceleration techniques it can be desirable to restrict the variation of step sizes
\[ \frac{h_{\text{max}}}{h_{\text{min}}} \leq K \quad (4.7) \]
with \( K \geq 1 \), i.e. to have quasuniform meshes. Another reason is that a high variation in the step sizes makes the solution process unstable and deteriorates the efficiency (see [1]). Therefore we change (4.4) of the algorithm to
\[ \tilde{\rho}(t) := \max \left\{ \rho(t)\Theta(t)TOL^{-1/m}, \, \frac{\max_{\tau \in [0,1]} \rho(t)\Theta(t)TOL^{-1/m}}{K} \right\}. \]

Actually this means nothing else than that the new mesh density is bounded from below, with respect to its maximum. It is easy to check that this adjustment restricts the step size variation. Let \( M := \max_{\tau \in [0,1]} \rho(t)\Theta(t)TOL^{-1/m} \), then
\[ \frac{1}{I} K^{-1} \tilde{h}_{j-1} \leq \int_{\tau_{j-1}}^{\tau_j} \tilde{\rho}(t)\,dt = \int_{\tau_{k-1}}^{\tau_k} \tilde{\rho}(t)\,dt \leq \frac{1}{I} M \tilde{h}_{k-1} \]
which results in (4.7). In the algorithm presented in [1], (4.6) is changed to
\[ \tilde{N} := \max \{1.5N, (1 + \delta)IN\}. \]
where \( \delta \) is a safety margin.
References

