A Direct Multiple Shooting Method for Initial Satellite Orbit Determination

Simon M. Lenz, Hans Georg Bock, and Johannes P. Schlöder

University of Heidelberg, 69120 Heidelberg, Germany

Ekaterina A. Kostina

University of Marburg, 35032 Marburg, Germany

Gottlob Gienger, Gerald Ziegler

European Space Operations Center, 64293 Darmstadt, Germany

After a satellite has been launched, the determination of its initial orbit is the primary question to be answered by mission control in order to prepare for high precision orbit determination and orbital transfers. For this purpose, tracking measurements performed by ground stations are used and parameters are determined by solving a parameter estimation problem. We suggest to use the direct multiple shooting approach for the solution of this problem. We develop a tailored method called "analytical projection" for the generation of reasonable initial guesses. As a test scenario for our method, we use operational tracking data of the ARTEMIS satellite, where the actual satellite trajectory differed significantly from the intended trajectory. A comparison to the widely employed direct single shooting method shows a significant improvement of the convergence behaviour.

I. Introduction

Traditionally it is distinguished between two different kinds of solution methods for orbit determination problems. On the one hand there are methods for the direct computation of six orbital elements from a corresponding number of measurements. Methods of this kind were developed already by Gauss, by Laplace and by Lambert and Euler. Those methods and variants of them are given, for example, in the book by Escobal, and were also discussed later in the survey by Marsden. These methods assume unperturbed Kepler orbits and are restricted to particular measurement types, although in more recent work by Allgower and Georg this approach has also been made applicable to situations with arbitrary tracking data.

In contrast to the direct computation of orbital elements there are also estimation methods that are typically based on large amounts of available data. Here, it is further distinguished between batch methods, where one tries to find a solution for the whole data set at once, and sequential methods, where one measurement at a time is processed. Both batch and sequential methods are iterative and require sufficiently good initial guesses.
Batch methods, which we want to consider in this article, typically rely on the formulation of a least-squares parameter estimation problem, which is constrained by a system of ordinary differential equations (ODE) that describes the satellite dynamics. In the standard single shooting approach, the unknown position and velocity vectors (state vector) at a reference time (epoch) are used as optimization variables. Since satellites are tracked regularly, the necessity of having good initial guesses for the optimization process is usually not a problem. However, shortly after the launch when the satellite’s orbit has to be determined for the first time the available a priori knowledge on the orbit may be poor, in particular in cases of launcher injection errors. Montenbruck and Gill state that in such situations a preliminary orbit determination using the direct computation methods may be required.

In this article we present a particular batch method for the solution of the orbit determination problems that is far less dependent on good initial guesses for the satellites’ state vector at the epoch. In particular, we resume the direct multiple shooting method for the solution of ODE-constrained parameter estimation problems and discuss the important feature of having the opportunity to exploit knowledge about the solution, which is given by the available measurement data. For this purpose we develop a projection algorithm that generates good initial guesses for the optimization process even if the a priori knowledge on the state vector at the epoch is very poor.

In order to test the performance of the method we exploit actual tracking data of the launch and early orbit phase of the ARTEMIS satellite. As this satellite was, due to a severe launcher underperformance, into significantly different orbit than expected, this scenario is an excellent study object to compare the convergence behavior of our method compared to the standard single shooting approach.

The article is structured as follows: In the next section, we briefly recall the theory of satellite motion and describe which perturbations to Kepler orbits we take into account for our computations. In section III, we discuss the measurement types typically used for initial orbit determination as well as the measurement functions used for modelling the outcome of measurements for a given description of the dynamics. In section IV, we formulate a least-squares parameter estimation problem that is subject to a system of ordinary differential equations. The direct single shooting and the direct multiple shooting approaches as possible parametrizations of the ODE-restriction to the parameter estimation problem are presented in section V and the important advantages of the latter approach are pointed out. In section VI we describe the generalized Gauss-Newton method as an iterative procedure to solve the parametrized optimization problems arising from the single shooting and multiple shooting approaches. Section VII is dedicated to the description of the “analytical projection”, which allows a generation of suitable initial guesses and is thus crucial for the success of multiple shooting for initial satellite orbit determination. Numerical results for tracking data of the ARTEMIS satellite are given in section VIII. Finally, we give a conclusion in section IX.
II. Satellite Dynamics

A. Keplerian Motion

As a first approximation, we consider the motion of a satellite around a central body as a two-body-problem. Further we will assume that the central body is at rest at the center of our coordinate system. This is well justified if the mass of the central body is by several orders of magnitude greater than that of the satellite, which is true for Earth-orbiting satellites. Under these conditions the motion of the satellite can be described by the second order differential equation:

$$\ddot{r}(t) = -\frac{GM_\oplus}{|r(t)|^2} r(t).$$

(1)

Herein, $r(t)$ is the position vector given in an appropriate inertial reference frame, as for example the J2000.0 system.\footnote{25} Further, $G$ is the gravitational constant and $M_\oplus$ is the mass of the central body. Finally, we have the acceleration $\ddot{r}(t)$ of the satellite, where the two dots indicate the second time derivative. We have used bold notation for vectors here and will do this throughout the whole document.

For the solution of equation (1), the three Kepler laws hold, which were published by Johannes Kepler in his two famous books "Astronomia Nova"\textsuperscript{17} and "Harmonices Mundi"\textsuperscript{18}:

- The orbit of the satellite is a conic section. The conic section equation

$$|r(t)|^2 = \frac{p}{1 + e \cos(\nu(t))}$$

holds, with $e$ and $\nu(t)$ being the Kepler elements eccentricity and true anomaly and $p$ being the semi-latus rectum. In particular, one has $0 \leq e < 1$ for bound orbits, with $e = 0$ corresponding to a circle.

- The position vector $r(t)$ sweeps over equal areas $\Delta A$ in equal time intervals $\Delta t$:

$$\frac{\Delta A}{\Delta t} = const.$$  

(3)

- The orbital period $T$ squared is proportional to the third power of the semi-major axis $a$:

$$T^2 = 4\pi^2 \frac{a^3}{GM_\oplus}.$$  

(4)

B. Perturbation Effects

Keplerian orbits are only a first approximation to satellite motion because several forces are not included in the model (1). In particular, for Earth orbiting satellites additional force terms arise from gravitational forces from Sun
and Moon and differences of the Earth gravitational potential from that of a point-like mass as implicitly assumed in Keplerian motion. Further, there are surface forces like air drag (for satellites at low altitudes) and solar radiation pressure. All forces arising from these perturbation effects are usually at least three orders of magnitude smaller than the gravitation of a point mass $M_0$ in the center of the Earth as assumed in the two-body problem (1) (cf. calculations performed by Montenbruck and Gill[23]). However, modern high precision applications like the Global Positioning System (GPS) require a detailed knowledge and accurate modelling of these perturbations up to relative precisions of $10^{-12}$. To obtain these very impressive accuracies, even relativistic effects need to be taken into account.3

In contrast to that, this article deals with initial satellite orbit determination where medium accuracies suffice. Accordingly, we do not go into the details of the models for the perturbation effects, and we only give an overview on the effects included in our calculations:

- **Air Drag**: The acceleration due to drag can be modeled by the expression:

$$\ddot{r}_{\text{drag}}(t) = -\frac{1}{2} C_D \frac{A_D}{m_{\text{sat}}} \rho_{\text{atm}} |\dot{r}_e(t)| \dot{r}_e(t),$$

with $\rho_{\text{atm}}$ being the density of the atmosphere, $A_D$ being the cross-section of the satellite in flight direction, $m_{\text{sat}}$ being the satellite mass and $C_D$ being the drag coefficient, which accounts for the satellite material and surface-atmosphere interaction. Further, $\dot{r}_e(t)$ denotes the velocity of the satellite relative to the atmosphere.

In order to actually compute the air drag, we therefore need to provide values for $A_D$, $C_D$ and $\rho_{\text{atm}}$. For the density $\rho_{\text{atm}}$, we used the MSIS-77 (Mass Spectrometer and Incoherent Scatter) model,15 which relies on satellite measurements performed in the 1970’s. The parameter $C_D$ and the area $A_D$ were provided by ESA.

- **Geopotential**: The deviation of the Earth’s gravitational field from that of a point-like mass in its center of mass is computed by using the Goddard Earth Model from Tracking Data22 (GEM-T1). The corresponding acceleration term is denoted as $\ddot{r}_{\text{geo}}(t)$.

- **Sun and Moon Gravitation**: Gravitational forces by Sun and Moon are computed by using the planetary ephemerides provided by the Jet Propulsion Laboratory (JPL)a. The employed model is called DE200, with DE standing for “development ephemerides”. We end up with acceleration terms that we call $\ddot{r}_{\text{Sun}}$ and $\ddot{r}_{\text{Moon}}$.

- **Solar Radiation Pressure**: For computing the acceleration due to solar radiation pressure the equation

$$\ddot{r}_{\text{srp}}(t) = - P_\odot C_R \frac{A_R}{m_{\text{sat}}} \frac{\dot{r}_\odot}{|\dot{r}_\odot|} \frac{a_\odot^2}{2}$$

(6)

can be used. Therein, we have introduced the average solar radiation pressure $P_\odot$ at a distance of 1 AU from the Sun, the radiation pressure coefficient $C_R$, the cross-sectional area of the satellite $A_R$ exposed to solar radiation,

[a]Available at http://ssd.jpl.nasa.gov/?ephemerides#planets
the vector $\mathbf{r}_\odot$ from the satellite to the Sun and the semi-major axis of the Earth orbit around the Sun $a_\odot$. For all our computations we use guesses for $C_R$ and $A_R$ provided by ESA.

By summing up all additional acceleration terms into the function

$$pert(t, \mathbf{r}(t), \mathbf{v}(t), \mathbf{q}) = \mathbf{r}_{\text{drag}}(t) + \mathbf{r}_{\text{geo}}(t) + \mathbf{r}_{\text{Sun}}(t) + \mathbf{r}_{\text{Moon}}(t) + \mathbf{r}_{\text{srp}}(t)$$

we can write the equation of motion for the satellite in the enhanced force model as

$$\ddot{\mathbf{r}}(t) = -\frac{GM_\odot}{|\mathbf{r}(t)|^2} \mathbf{r}(t) + pert(t, \mathbf{r}(t), \mathbf{v}(t), \mathbf{q}).$$

(8)

In this equation, the vector $\mathbf{q}$ sums up all additional parameters. By standard transformation techniques, equation (8) can also be written as first order ordinary differential equation with a $\mathbf{q}$-dependent right hand side:

$$\dot{\mathbf{y}}(t) = f(t, \mathbf{y}(t), \mathbf{q}).$$

(9)

The state vector $\mathbf{y}(t)$ includes both position and velocity of the satellite:

$$\mathbf{y}(t) = \begin{pmatrix} \mathbf{r}(t) \\ \mathbf{v}(t) \end{pmatrix}.$$  

(10)

An orbit is uniquely determined, if the state vector $\mathbf{y}(t^*)$ is known for a time $t^*$.

### III. Tracking Data

For the determination of the satellite orbit we need to have tracking measurements. In addition, we need measurement functions that describe the relation between the variables appearing in the differential equation (time $t$, differential state vector $\mathbf{y}(t)$, parameters $\mathbf{q}$) and the observables of the system. The difference between a measurement $\eta_{i,j}$ of observable $j$ taken at a time $t_i$ and the evaluation of the corresponding measurement function $m_j(t_i, \hat{y}(t_i), \hat{q})$ for the unknown true parameter values $\hat{q}$ and the unknown true state vector at the given measurement time $\hat{y}(t_i)$ is the measurement error $\epsilon_{i,j}$:

$$\epsilon_{i,j} = \eta_{i,j} - m_j(t_i, \hat{y}(t_i), \hat{q})$$

(11)

In this section, we will describe the three most common measurement types, namely range, range rate and angular measurements and introduce the measurement functions.
A. Two-Way Range Measurements

Range measurements rely on the travel time $\tau$ of a signal. In particular, two-way range measurements are performed by sending out a signal, and detecting the phase shift $\Delta \Phi_{\text{meas}}$ between the outgoing and incoming signal. Then the signal travel time can be computed as

$$\tau_{\text{meas}} = \frac{\Delta \Phi_{\text{meas}}}{2\pi f_0},$$

with $f_0$ being the frequency of the radio wave. From this the measured two-way range can be computed as

$$\rho_{\text{meas}} = c\tau_{\text{meas}},$$

with the speed of light $c$. In the preceding two equations the subscript $\text{meas}$ has been introduced to indicate that these are measured quantities or quantities that are computed from measured quantities using exact analytical relations. The quantity $\rho_{\text{meas}}$ corresponds to a measurement $\eta_{i,\text{range}}$ in equation (11).

The ambiguity of the measurement $\Delta \rho = \frac{\Delta \rho}{2f_0}$ may be removed by adding sub-harmonic minor tones to the carrier wave with frequency $f_0$.

For the mathematical model of the range measurement, we first introduce the auxiliary function

$$\rho(t) := 2 \cdot |r(t) - r_{\text{stat}}(t)|_2,$$

which delivers twice the distance between the position of the satellite $r(t)$ and the position of the station $r_{\text{stat}}(t)$ at a given time $t$. Then the range measurement function is defined as

$$m_{\text{range}}(t_M, y(t_M)) = \rho(t_M) + \frac{\rho(t_M)}{2c} \frac{\partial \rho(t)}{\partial t} \bigg|_{t=t_M}$$

with $t_M$ being the time where the measurement was started, that is it denotes the time of the emission of the ranging signal.

For the motivation of the second term it is pointed out that the measured range $\rho_{\text{meas}}$ does not correspond to the distance between station and satellite at the nominal measurement time $t_M$, but rather to

$$|r(t_M + \tau_u) - r_{\text{stat}}(t_M)| + |r(t_M + \tau_u + \tau_d) - r_{\text{stat}}(t_M + \tau_u + \tau_d)|,$$

because the signal does not instantly arrive at the satellite, but needs certain travel times $\tau_u$ and $\tau_d$ to get up to the
satellite and down to the station again. These uplink and downlink light travel times are only implicitly known as

\[ \tau_u = \left\| \mathbf{r}(t_M + \tau_u) - \mathbf{r}_{\text{stat}}(t_M) \right\| / c \]  
\[ \tau_d = \left\| \mathbf{r}(t_M + \tau_u) - \mathbf{r}_{\text{stat}}(t_M + \tau_u + \tau_d) \right\| / c. \]  

The travel times \( \tau_u \) and \( \tau_d \) can be calculated by a fixed-point iteration up to machine precision. However, for Earth orbiting satellites a simple approximation is sufficient. Therefore it is assumed that the expression \((*)\) can be approximated by twice the distance at the time \( t_M + \frac{\tau}{2} \), that is \( \rho(t_M + \frac{\tau}{2}) \), where \( \tau = \tau_u + \tau_d \).

This distance is then approximated by the first order of a Taylor expansion

\[ \rho \left( t_M + \frac{\tau}{2} \right) = \rho(t_M) + \frac{\tau}{2} \frac{\partial \rho(t)}{\partial t} \bigg|_{t_M}, \]  

delivering exactly equation (15), if \( \frac{\tau}{2} \) is replaced by \( \frac{\mathbf{r}(t_M) - \mathbf{r}_{\text{stat}}(t_M)}{c} \).

Note, that for the time derivative of \( \rho \) appearing in (15) not only the motion of the satellite, but also the motion of the station due to the Earth’s rotation must be taken into account.

In the discussion of the range measurement function we have excluded several technical effects such as timing biases, transponder delays and corrections due to atmospheric perturbations in order to keep the description short by focussing on the main points. The presentation of the full range measurement function including additional terms for the correction of these errors is therefore shifted to the appendix.

B. Two-Way Range Rate Measurements

Two-way range rate measurements rely on the Doppler shift of frequencies. A ground station emits a signal with a certain emission frequency \( f_e \). The frequency \( f_r \) of the signal retransmitted by the satellite is then greater than \( f_e \), if the satellite is approaching the station, and lower, if the distance increases. The received signal \( f_r \) also underlies the Doppler effect because the station itself is in motion as well.

Unfortunately it is not possible in practice to directly detect the frequency shift \( f_r - f_e \). Instead, a so-called integrated Doppler measurement is used, where the zero crossings \( N_{\text{meas}} \) of a wave with frequency \( f_r - f_e \) are counted over a count time interval \( t_c \). The measured two-way range rate \( \dot{\rho}_{\text{meas}} \) is then obtained from

\[ \dot{\rho}_{\text{meas}} = -\frac{1}{2} \frac{c N_{\text{meas}}}{f_e t_c} \]  

This quantity corresponds to a measurement \( \eta_{i,\text{rangerate}} \) at a time \( t_i \) in equation (11).

The range rate measurement function that models the range rate for a given state vector can be computed by a
difference quotient of range measurement functions:

\[ m_{\text{range rate}}(t_M, y(t_M)) = \frac{m_{\text{range}}(t_M + t_c, y(t_M + t_c)) - m_{\text{range}}(t_M, y(t_M))}{t_c} \]  

(19)

The evaluations are done at times \( t_M \) and \( t_M + t_c \), thus at the beginning and at the end of the count time interval of length \( t_c \).

Again, the full measurement function accounting for technical details and including atmospheric correction is given in the appendix.

C. Angular Measurements

For the discussion of the angular measurement function we first introduce the local tangent coordinate system, the origin of which is located in the observing ground station. The three coordinate axes are then given as follows: The \( x \)-axis is aligned with the local latitude, pointing toward the east. The \( y \)-axis is aligned with the local longitude, pointing toward the north. And finally, the \( z \)-axis points toward the zenith. An illustration is given in figure 1a.

In an angular measurement, the azimuth and elevation angle are measured. For their definition, see figure 1b: As it displays, the azimuth angle lies in the local horizontal plane and is counted positively from north to east. The elevation angle specifies the angle between the observation light ray and the horizontal plane and is counted positively to the zenith.

In practice, angular measurements are performed using parabolic antennas in S-, X- or Ka-band that scan for the satellite in a conical area.

For the definition of the measurement function we introduce the auxiliary functions

\[ A(t) := \arctan2(r_E(t), r_N(t)) \]  

\[ E(t) := \arctan2(r_Z(t), \sqrt{r_E(t)^2 + r_N(t)^2}) \]  

(20a)

(20b)

which give the azimuth and elevation angle for a satellite in dependence of the local tangent coordinates east \( r_E \), north \( r_N \) and zenith \( r_Z \) under the assumption that the signal travel time is zero. The function \( \arctan2(y, x) \) in equations (20) is defined such that it gives the angle whose tangent is \( \frac{y}{x} \) and which is in the correct quadrant.

However, since the signal travel time is of course finite, the angular measurement function contains an appropriate
Figure 1: Illustration of local tangent coordinate system and measured angles
(a) The local tangent coordinate system: The $x$-axis is aligned with the local latitude, the $y$-axis is aligned with the local longitude. The $z$-axis points toward the zenith of the horizontal plane created by the $x$ and $y$ axes. Image from http://standards.iso.org/ittf/PubliclyAvailableStandards/C030811e_FILES/MAIN_C030811e/text/ISOIEC_18026E_SRF.HTM:
(b) Measured angles: The observer (light gray) is located in the origin of the local tangent coordinate system. The azimuth angle lies in the horizontal plane and is counted positively from the $y$-axis (north) to east. The elevation angle is measured from the horizontal plane towards the zenith.

The correction term:

$$m_{\text{ang}}(t_M, y(t_M)) = \begin{pmatrix} m_{\text{azimuth}}(t_M, y(t_M)) \\ m_{\text{elevation}}(t_M, y(t_M)) \end{pmatrix} = \begin{pmatrix} A(t) + \frac{\dot{\rho}(t_M)}{2c} \cdot \frac{\partial A(t)}{\partial t} \\ E(t) + \frac{\dot{\rho}(t_M)}{2c} \cdot \frac{\partial E(t)}{\partial t} \end{pmatrix} \cdot t_M$$

(21)

For the full measurement function, please see the appendix.

**IV. Parameter Estimation**

Given the equation of motion in the enhanced force model (8), tracking measurements of the types range, range rate and angles, and further the corresponding measurement functions for these measurement types (15), (19) and (21), we can now formulate a parameter estimation problem. Therefore we first write down a general, constrained least-squares parameter estimation problem with underlying system of ordinary differential equations:
Problem IV.1 Determine a parameter vector \( q \) and a solution \( y(t) \) such that

\[
\min_{y(t), q} \left| r_1(y_0, \ldots, y_1, q) \right|^2_2 \\
r_2(y_0, \ldots, y_1, q) = 0
\]

where \( y_j - y(t_j), \ j = 0, \ldots, l \) and \( y(t) \) fulfills the system of ordinary differential equations

\[
\dot{y}(t) = f(t, y(t), q) \\
y : [t_a, t_b] \rightarrow \mathbb{R}^{nd} \\
q \in \mathbb{R}^{nq}
\]

The \( t_j \in [t_a, t_b] \) are known and

\[
r_1 \in C^2(D, \mathbb{R}^{n_1}), \quad \text{on the domain } D \subset \mathbb{R}^{(l+1)n_d+n_q}.
\]

The objective function of the minimization problem IV.1 is usually chosen as a least-squares sum of the differences between measurements \( \eta_{i,j} \) and model predictions \( m_j(t_i, y(t_i), q) \), weighted with factors \( w_{i,j} \). In the case of satellite orbit determination, this means that

\[
\left| r_1(y_0, \ldots, y_1, q) \right|^2_2 = \sum_{j \in \{ \text{range} \}} \sum_{i=1}^{n_{j,\text{meas}}} (w_{i,j} \cdot (\eta_{i,j} - m_j(t_i, y(t_i), q)))^2 \\
+ \sum_{i=1}^{n_{\text{ang,meas}}} \left( w_{i,\text{azim}} \cdot (\eta_{i,\text{azim}} - m_{\text{azim}}(t_i, y(t_i), q)) \right)^2 + w_{i,\text{elev}} \cdot (\eta_{i,\text{elev}} - m_{\text{elev}}(t_i, y(t_i), q)) \right)^2_2, \quad (22)
\]

where we have written the residuals of the angular measurements in a separate term because of the two dimensional measurement function and measurements, resp. Further we have used \( n_{j,\text{meas}} \) for the total number of measurements of type \( j \) and have abbreviated azimuth to "azim" and elevation to "elev".

For satellite orbit determination, the ordinary differential equation in problem IV.1 is given by the equation of motion (8).

Given independent and normally distributed measurement errors with zero mean

\[
e_{i,j} - \eta_{i,j} - m_j(t_i, y(t_i), q) \in \mathcal{N}(0, \sigma^2_{i,j}), \ j \in \{ \text{range, rangerate, azim, elev} \} \quad (23)
\]
and choosing \( w_{i,j} = \frac{1}{\sigma_{i,j}} \), the solution of the constrained least-squares problem provides a maximum likelihood estimate for the unknown parameter vector.

V. Direct Single Shooting and Direct Multiple Shooting

The problem IV.1 stated in the previous section is an optimization problem, which is coupled to a differential equation. It is infinite-dimensional, because one of the optimization variables is the continuous function \( y(t) \). So-called direct optimization methods rely on the formulation of a finite dimensional optimization problem, whereas indirect optimization methods are based on applying the maximum principle, what leads to a multi-point boundary value problem. In the following, we will only consider direct optimization methods. For an overview of direct and indirect optimization methods we refer to Binder et al.\(^4\)

The conventional way for the solution of least-squares parameter estimation is to use the direct single shooting approach. Here, problem IV.1 is reduced to finite dimension by using the initial value \( y_{a} = y(t_{a}) \) as optimization variable. The state vectors for the times \( y(t_{j}), j = 0, \ldots, l \) are obtained by solving an initial value problem.

The resulting parameterized constrained least-squares parameter estimation problem with its optimization variables \((y_{a}, q)\) can then be solved by an iterative procedure, for example by the generalized Gauss-Newton method described in the next section. Note that in each of these iterations the initial value problem (IVP)

\[
\dot{y}(t) = f(t, y(t), q) \\
y(t_{a}) = y_{a}
\]

has to be solved on \([t_{a}, t_{b}]\) in order to evaluate the measurement functions.

The single shooting method has the advantage of being easily implemented, but it may be difficult or even impossible to determine the solution of the IVP if the initial guess for the unknowns \((y_{a}, q)\) is poor. If one uses single shooting, it is practically impossible to incorporate information about the solution trajectory - e.g. information given by the measurement data or expert knowledge - in the initial guesses. Unstable trajectories, which may occur due to poor initial guesses in the satellite orbit determination problem, can also not be treated by this approach.

An alternative, which does not have the mentioned disadvantages, is to use the direct multiple shooting method. In the following we briefly describe how to use this method for the solution of parameter estimation problems.\(^5\)–\(^7,9\)

The idea of multiple shooting is to choose a grid of so called multiple shooting nodes \(\tau_{j}\)

\[
t_{a} = \tau_{0} < \tau_{1} < \ldots < \tau_{m} = t_{b}
\]

where \([t_{a}, t_{b}]\) is the interval, in which the measurements are given.
At each grid point we introduce variables $s_j \in \mathbb{R}^n$ and then solve $m$ initial value problems

\[
\dot{y}(t) = f(t, y(t), q)
\]

\[
y(\tau_j) = s_j
\]

on the subintervals $I_j := [\tau_j, \tau_{j+1}]$. The solutions are denoted as $y(t; \tau_j, s_j, q)$ for $t \in I_j$. Solutions generated in this way are in general not continuous at the multiple shooting nodes $\tau_j, j = 1, \ldots, m$. Hence, it is necessary to formulate matching conditions $h_j$ which assure that the solution is eventually continuous on the whole interval $[t_a, t_b]$.

\[
h_j(s_j, s_{j+1}, q) := y(\tau_{j+1}; \tau_j, s_j, q) - s_{j+1} = 0, \quad j = 0, \ldots, m - 1
\]

(24)

Using this parameterisation the parameter estimation problem is reformulated in terms of the augmented variable vector $(s_0, \ldots, s_m, q)^T$.

**Problem V.1** Determine values for $s_0, \ldots, s_m, q$ such that

\[
\min_{s_0, \ldots, s_m, q} \|r_1(s_0, \ldots, s_m, q)\|_2^2
\]

s.t. \qquad $r_2(s_0, \ldots, s_m, q) = 0$

and that the matching conditions

\[
h_j(s_j, s_{j+1}, q) := y(\tau_{j+1}; \tau_j, s_j, q) - s_{j+1} = 0, \quad j = 0, \ldots, m - 1
\]
hold and \( y(t; \tau_j, s_j, q) \) for \( t \in [\tau_j, \tau_{j+1}] \) is given as the solution of the initial value problem

\[
\begin{align*}
\dot{y}(t) &= f(t, y(t), q) \\
y(\tau_j) &= s_j
\end{align*}
\]

for \( j = 0, \ldots, m - 1 \).

The direct single shooting method can be considered as a special direct multiple shooting method with \( m - 1 \).

VI. Generalized Gauss-Newton Method

For the solution of the discretized least-squares parameter estimation problem \( \text{V.1} \) we use the generalized Gauss-Newton method. The basic algorithm, numerical aspects and a statistical assessment of the solution are given in this section.

In order to simplify the notation, we rewrite problem \( \text{V.1} \) using the abbreviation \( x = (s_0^T, \ldots, s_m^T, q^T)^T \) for the optimization variables. Furthermore, we combine the equality constraint function \( r_2 \) and the matching conditions \((24)\) into a single function \( F_2 \) and rename \( r_1 \) as \( F_1 \). By this we obtain the following nonlinear constrained least-squares problem:

**Problem VI.1** Determine \( x \) such that

\[
\begin{align*}
\min_x |F_1(x)|_2^2 \\
s.t. \quad F_2(x) = 0
\end{align*}
\]

A. Basic Algorithm and Convergence Properties

The generalized Gauss-Newton method is an iterative method. Given an initial guess \( x^0 \), the iterates are given by

\[
x^{k+1} = x^k + t^k \Delta x^k,
\]

with increments \( \Delta x^k \) that are determined as the solution of the constrained linearized problem

\[
\begin{align*}
\min_{\Delta x^k} |F_1(x^k) + J_1(x^k)\Delta x^k|_2^2 \\
s.t. \quad F_2(x^k) + J_2(x^k)\Delta x^k = 0
\end{align*}
\]

where \( J_i(x^k) = \frac{\partial F_i}{\partial x}(x^k) \in \mathbb{R}^{n_i \times m} \) are the Jacobian matrices of the functions \( F_i \) that are assumed to be three times continuously differentiable, \( F_i(x) \in C^3 \). The stepsize \( t^k \in (0, 1] \) in equation \((25)\) is determined following a
globalization strategy.

The functions $F_1$ and $F_2$ and the corresponding Jacobians are summed up in

$$F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \quad \text{and} \quad \mathcal{J} = \begin{pmatrix} \mathcal{J}_1 \\ \mathcal{J}_2 \end{pmatrix}. \quad (27)$$

For the solution of the constrained linearized problem there exists a linear solution operator $\mathcal{J}^+$:

$$\Delta x^k = -\mathcal{J}^+(x^k)F(x^k) \quad (28)$$

that has the following explicit form:

$$\mathcal{J}^+ = \begin{pmatrix} I & 0 \\ J_c^T J_c & 0 \end{pmatrix}^{-1} \begin{pmatrix} J_1^T & 0 \\ 0 & I \end{pmatrix}. \quad (29)$$

The solution operator $\mathcal{J}^+$ has the property of being a generalized inverse

$$\mathcal{J}^+ = \mathcal{J}^+ \mathcal{J} \mathcal{J}^+. \quad (30)$$

Extensions of the generalized Gauss-Newton method for additional inequality constraints $F_3(x) \geq 0$ in the problem formulation VI.1 using an active set strategy exist.\textsuperscript{7} Local convergence of the method can be shown under relatively mild assumptions.\textsuperscript{5–7} For globalization of convergence we employ the restrictive monotonicity test\textsuperscript{7,10} (RMT), which provides appropriate values for the choice of the stepsize $t^k$. The iteration (25) is stopped, if the scaled increment $[S \Delta x^k]$, $S$ being a diagonal scaling matrix, falls below a given threshold.
B. Numerical Aspects

Condensing

The multiple shooting parameterization leads to large, but structured Jacobian matrices. Splitting up the function $F_2$ into its compartments $r_2$ and $h_j$, $j = 0, \ldots, m - 1$, we can investigate this structure in detail:

$$J(x) = \begin{pmatrix} D_1^0 & D_1^1 & \ldots & \ldots & D_1^m & D_1^q \\ D_2^0 & D_2^1 & \ldots & \ldots & D_2^m & D_2^q \\ G_L^0 & -I & 0 & 0 & 0 & G_q^0 \\ 0 & 0 & G_L^1 & -I & 0 & 0 & G_q^1 \\ 0 & 0 & 0 & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & \ddots & \ddots & 0 & \vdots \\ 0 & 0 & 0 & 0 & G_L^{m-1} & -I & G_q^{m-1} \end{pmatrix}, \quad F(x) = \begin{pmatrix} r_1 \\ r_2 \\ h_0 \\ h_1 \\ \vdots \\ \vdots \\ h_{m-1} \end{pmatrix}. \quad (31)$$

Herein, we have used the abbreviations

$$D_i^j := \frac{\partial r_1}{\partial s_j} \in \mathbb{R}^{n_d \times n_d}, \quad j = 0, \ldots, m; \quad i = 1, 2 \quad (32)$$

$$D_i^q := \frac{\partial r_1}{\partial q} \in \mathbb{R}^{n_i \times n_q}, \quad i = 1, 2 \quad (33)$$

$$G_L^i := \frac{\partial h_i(s_i, s_{i+1}, q)}{\partial s_i} \in \mathbb{R}^{n_d \times n_d}, \quad i = 0, \ldots, m - 1. \quad (34)$$

$$G_q^i := \frac{\partial h_i(s_i, s_{i+1}, q)}{\partial q} \in \mathbb{R}^{n_i \times n_q}, \quad i = 0, \ldots, m - 1. \quad (35)$$

Techniques for the exploitation of the special structure of $J(x)$ have been developed, which lead to a "condensed" equality constrained linearized problem of form (26). The number of columns in the condensed problem is then equal to the number $n_d$ of differential states plus the number of parameters $n_q$. Examples for condensing algorithms are the backward recursion\(^7\) and the forward recursion.\(^9\) After the solution of the condensed problem all other variables can be determined recursively. Following the condensing algorithms, the computational costs needed in each iteration are quite low.

Internal Numerical Differentiation

Since the generalized Gauss-Newton method is derivative based, derivatives of the solution of the ordinary differential equation with respect to initial values and parameters need to be computed. Such derivatives appear for example in the matrices $G_L^i$ (34), but also in the matrices $D_i^j$, because the functions $r_1$ depend on the differential states at the measurement times, which may be located anywhere on the given multiple shooting interval.
In order to compute these derivatives with high accuracies and low computational costs, we apply the internal numerical differentiation. In contrast to external numerical differentiation, where every directional derivative requires an additional integrator call, the derivatives are calculated together with the nominal trajectory. All adaptive components of the integrator are fixed for the derivative computation, thus reducing the computational effort. Moreover, we obtain the exact derivative for the discretization scheme, which makes the procedure stable also for low accuracies.

VII. Generation of Initial Guesses

One of the big advantages of multiple shooting is the possibility to incorporate knowledge about the solution in the initial guesses for the intermediate values and not only for the initial guesses of the parameter values and the initial values of the differential equation. This often accelerates the solution process for parameter estimation problems. If the differential states of the ordinary differential equation are accessible by measurements, these measurements can directly be used as initial guesses for the intermediate state variables \( s_j \). However, for satellite orbit determination problems the measurement functions \( m_{\text{range}}, m_{\text{rangerate}} \) and \( m_{\text{ang}} \) are more complicated and a method for the generation of reasonable initial guesses is needed.

In this section, we describe a sequential procedure that makes use of the geometrical information contained in the measurements in order to perform projections onto subsets of the three dimensional space where the measurements are fulfilled. Since we will use simplified measurement functions, an analytical computation of the projected position with elementary methods of linear algebra is possible, thus suggesting the name "analytical projection" for the method described here.

A. The Analytical Projection

For the generation of initial guesses for the variables \( s_i, i = 1, \ldots, m \) we suggest the following procedure, for which only an initial guess for the state vector at the first multiple shooting node \( s_0 = y(t_0) \) and the parameters \( q \) is needed. Note that in general not every measurement time needs to be a multiple shooting node or vice versa. However, for simplicity we assume that the grid of measurement times \( t_i \) is a superset of the grid of multiple shooting nodes \( \tau_i \). The algorithm can easily be modified if this is not the case.

1. Start with \( i = 0, l = 0 \) and the initial guess \( s_0 = y(t_0) - y_0 \).

2. Solve the initial value problem on interval \([t_i, t_{i+1}]\),

\[
\begin{align*}
\dot{y}(t) &= f(t, y(t), q) \\
y(t_i) &= y_i
\end{align*}
\]

where \( t_i \) and \( t_{i+1} \) are both measurement times. The solution of the ODE at time \( t_{i+1} \) is denoted as \( y_{\text{int}}(t_{i+1}; t_i, y_i, q) \).
3. Determine $y_{i+1}$ by performing a projection according to the measurement type. Detailed descriptions are given below.

4. If the measurement time $t_{i+1}$ is not a multiple shooting node, set $i = i+1$ and continue with 2.

5. If the measurement time $t_{i+1}$ is the multiple shooting node $\tau_{i+1}$, then set $s_{i+1} = y_{i+1}$. If $\tau_{i+1} = t_b$, stop the algorithm. Otherwise, set $l = l + 1$ and continue with 2.

Let us now consider the projections in detail.

**Range Measurement**

If a two-way range measurement was performed at time $t_{i+1}$, we apply the following steps for the computation of the projected state vector:

- Decompose the state vector obtained from the integration into $y_{\text{int}}(t_{i+1}; t_i, y_i, q) = (r_{\text{int}}^T, v_{\text{int}}^T)^T$ and transform the position part $r_{\text{int}}$ into the local tangent coordinate system.

- Compute the unit vector

\[ e_{r_{\text{int}}} = \frac{r_{\text{int}}(t_{i+1})}{\|r_{\text{int}}(t_{i+1})\|} \]  

(36)

- Compute the projected position with the measured range $\eta_{i+1,\text{range}}$

\[ r(t_{i+1}) = \frac{\eta_{i+1,\text{range}}}{2}, e_{r_{\text{int}}} \]  

(37)

The factor two comes from the fact that we always measure the two-way range.

- Transform the projected position $r(t_{i+1})$ into the J2000.0 inertial coordinate system and combine the projected position and the unchanged velocity vector to $y_{i+1} = (r(t_{i+1})^T, v_{\text{int}}^T)^T$.

**Angular Measurement**

If the measurement at $t_{i+1}$ is an angular measurement the following procedure is applied:

- Decompose the state vector obtained from the integration into $y_{\text{int}}(t_{i+1}; t_i, y_i, q) = (r_{\text{int}}^T, v_{\text{int}}^T)^T$ and transform the position part $r_{\text{int}}$ into the local tangent coordinate system.
• Compute the unit vector in the observation direction from the measured azimuth and elevation angle:

\[ \mathbf{e}_{\text{obs}} = \begin{pmatrix} \sin(\eta_{i+1, \text{azim}}) \cos(\eta_{i+1, \text{elev}}) \\ \cos(\eta_{i+1, \text{azim}}) \cos(\eta_{i+1, \text{elev}}) \\ \sin(\eta_{i+1, \text{elev}}) \end{pmatrix} \]  

(38)

• Compute the scalar product:

\[ \alpha_{\text{proj}} = \mathbf{r}_{\text{int}}(t_{i+1}) \cdot \mathbf{e}_{\text{obs}} \]  

(39)

• Compute the projected position as

\[ \mathbf{r}(t_{i+1}) - \alpha_{\text{proj}} \cdot \mathbf{e}_{\text{obs}}. \]  

(40)

• Transform the projected position \( \mathbf{r}(t_{i+1}) \) into the J2000 coordinate system and combine the projected position and the unchanged velocity vector to \( \mathbf{y}_{i+1} = \left( \mathbf{r}(t_{i+1})^T, \mathbf{\dot{r}}_{\text{int}}^T \right)^T. \)

Range Rate Measurement

Finally, if a range rate measurement is performed we can use the following projection:

• Decompose the state vector obtained from the integration into \( \mathbf{y}_{\text{int}}(t_{i+1}; \mathbf{r}, \mathbf{v}, \mathbf{q}) - \left( \mathbf{r}_{\text{int}}^T, \mathbf{\dot{r}}_{\text{int}}^T \right)^T \) and transform both the position part \( \mathbf{r}_{\text{int}} \) and the velocity part \( \mathbf{\dot{r}}_{\text{int}}(t_{i+1}) \) into the local tangent coordinate system.

• Decompose the velocity into two components: one that is parallel and one that is perpendicular to \( \mathbf{\dot{r}}_{\text{int}}(t_{i+1}) \):

\[ \mathbf{\dot{r}}_{\text{int}}(t_{i+1}) = \mathbf{\dot{r}}_{\text{int, \parallel}}(t_{i+1}) + \mathbf{\dot{r}}_{\text{int, \perp}}(t_{i+1}). \]  

(41)

• Correct the parallel part with the measured range rate

\[ \mathbf{\dot{r}}_{\parallel}(t_{i+1}) = \frac{1}{2} \eta_{i+1, \text{rangerate}} \cdot \frac{\mathbf{\dot{r}}_{\text{int}}(t_{i+1})}{\left| \mathbf{\dot{r}}_{\text{int}}(t_{i+1}) \right|} \]  

(42)

Again, the factor \( \frac{1}{2} \) is needed because the measurements are two-way range rate measurements.

• For the projected velocity the unchanged perpendicular component and the corrected parallel part is used:

\[ \mathbf{\dot{r}}(t_{i+1}) = \mathbf{\dot{r}}_{\parallel}(t_{i+1}) + \mathbf{\dot{r}}_{\text{int, \perp}}(t_{i+1}). \]  

(43)
The projections described here are a slight modification to the ones suggested by Gienger. An illustration of the projections for all measurement types is given in figure 3.

It shall be mentioned that the analytical projection is also applicable to particular combinations of measurements, because the gathered measurement information is complementary for different measurement types. However, the projections do not commute. Therefore it is suggested to perform the projections in the order "angle - range - range rate", because only in this order the value of the measurement function for the previously performed projections is left unchanged.

![Figure 3: Projection types](image)

Figure 3: Projection types: (a) Range projection: All points that fulfill the range measurement are located on the shaded sphere centered at the ground station (white circle). The position obtained from the integration (star shaped) is projected to a position on the sphere (black circle), the velocity vector is unchanged. (b) Angle projection: All points that fulfill the angular measurement are on the half-line originating from the ground station. Again, the position vector obtained from the integration (star shaped) is projected to that point on the half-line that has minimum distance to the position obtained from the integration. (c) Range rate projection: The velocity vector obtained from the integration (solid gray arrow) can be decomposed into a part that is parallel to the connecting line between station and satellite and another part, which is perpendicular to it (dashed gray arrows). The parallel part is then corrected using the measurement information (dashed black). Together with the unchanged perpendicular part the projected velocity vector (dotted black arrow) is obtained.

**B. Alternatives**

Instead of using the analytical formulas for the projection as described above, it is also possible to determine the solutions of local minimization problems of the following form:

**Problem VII.1** Determine $y_{i+1}$ as a solution of the constrained multistage least-squares problem

$$
\min_{y_{i+1}} \|S(y_{i+1} - y_{\text{int}}(t_{i+1}))\|_2^2
$$

subject to

$$
\min_{y_{i+1}} \|g(y_{i+1}) - z_{i+1}\|_2^2
$$

$$
h(y_{i+1}) \leq 0,
$$

American Institute of Aeronautics and Astronautics
where $S \in \mathbb{R}^{n_d \times n_d}$ is a diagonal scaling matrix. Further, $z_{i+1}$ represents expert knowledge, typically a vector of available measurement data. Correspondingly, $g$ is the vector of the corresponding measurement functions. Finally, $h$ represents additional inequality constraints.

By this approach an arbitrary number of measurements at a time can be treated, especially problems where an exact fulfillment of the in principle arbitrary large number of measurements ($g(y_{i+1}) - z_{i+1}$) is impossible and only a least-squares minimization can be done. Further, additional expert knowledge can be used for the projection by defining appropriate inequality constraints, e.g. that the orbit is bounded $\epsilon \leq 1 - \epsilon$ with $\epsilon$ small.

Problem VII.1 is a nonlinear constrained least squares problem itself, so for its solution basically the same methods are required as for the original problem. But since no dynamical process is involved, it is not as complicated as the original problem and has the advantage that a natural initial guess of the variables is given by the result of the integration, $y_{\text{int}}(t_{i+1})$.

However, multiple measurements at a time are very rare in practice, and for simple combinations of measurements we have already discussed that the analytical approach is still suitable. In particular, for the test scenario considered in the next section no multiple measurements occurred and we can therefore use the analytical projection with its geometrically intuitive solutions.

Finally, let us discuss which special minimization problems are solved by the analytical projection:

- For range and angular measurements with $\alpha_{\text{proj}} > 0$ the solution of problem VII.1 and the analytical solutions are the same, if the diagonal elements of $S$ that correspond to the position vector components are all the same, $S_{1,1} = S_{2,2} = S_{3,3}$, and if nominal measurement functions are employed:

$$n_{\text{range}}^\text{nominal}(t, y(t)) = 2|r(t) - r_{\text{stat}}(t)|_2$$

and

$$n_{\text{ang}}^\text{nominal}(t, y(t)) = \begin{pmatrix} A(t) \\ E(t) \end{pmatrix}.$$ (45)

- For angular measurements with $\alpha_{\text{proj}} \leq 0$ the analytical approach projects to a position on the extended line, whereas the minimization problem has no well-defined solution because all feasible points are on the open half-line.

- The solution of the minimization problem for range rate measurements depend on the scaling matrix $S$, because fulfilling the measurement can also be achieved by varying the observation direction instead of varying the velocity vector. The result of the analytical projection corresponds to infinitely high penalty factors for the changes in the position vector $S_{1,1} = S_{2,2} = S_{3,3} = \infty$ and to the usage of the nominal range rate measurement
function:

\[
m_{\text{nominal}}^\text{rangerate}(t, y(t)) = 2 \frac{d}{dt} \| \mathbf{r}(t) - \mathbf{r}_{\text{stat}}(t) \|_2
\]

(46)

VIII. Numerical Investigations

For a comparison of the direct single shooting and the direct multiple shooting approach we use operational tracking data from the ARTEMIS launch and early orbit phase on July 12th in 2001\(^1\) provided by ESOC\(^b\) Flight Dynamics. Because of severe underperformance of the Ariane 5G launcher, the actual orbit of the satellite after the separation from the launch vehicle was significantly lower than expected (apogee 8500\(km\) lower than for geostationary transfer orbit). The ARTEMIS scenario therefore constitutes a very challenging test case.

A. Software

The parameter estimation software we use for the solution of initial satellite orbit determination problems is PARFITBAHN. The core of this software is the program PARFIT,\(^5–7\) which is in an implementation of the multiple shooting parameterization and the generalized Gauss-Newton method for parameter estimation problems. Initial value problems are solved in PARFIT using the integrator DIFSYS, which relies on the explicit mid-point rule developed by Bulirsch and Stoer.\(^11\)

For the solution of initial satellite orbit determination problems the problem specific environment was implemented for PARFIT. This problem specific environment contains subroutines for data processing of input files, the generation of suitable initial guesses for the multiple shooting variables with the analytical projection, computation of the right hand side of the differential equation, the least squares terms and trajectory computation and data output to text files. A lot of (partly modified) routines of the orbit determination software BAHN\(^24\) provided by ESA were used. This motivates the name for the whole multiple shooting software package for initial satellite orbit determination: PARFITBAHN. The latest version is version 1.2.\(^8,19\)

Since single shooting can be regarded as a special case of multiple shooting, PARFITBAHN can also be used in single shooting mode. The comparison between the two approaches in this section relies on calculations where all other features of the numerical methods are kept fixed in order to assure that the observed difference of the convergence behavior is only due to the application of multiple shooting and the generation of suitable initial values.

However, we will also additionally compare the convergence behavior of PARFITBAHN to that of the software BAHN used by ESA. Without going into the details we simply mention that BAHN works with a single shooting method, with increments determined by the solution of the well-known normal equations for unconstrained minimiza-
tion problems:

$$J^T(x)J(x)\Delta x = -J^T(x)F(x). \quad (47)$$

No step size control for the iterations is done. The integrator relies on the predictor-corrector Adams-Bashforth-Moulton method. BAHN exits successful if the relative change of the least squares sum $$(F_1(x^{k-1}) - F_1(x^k))/F_1(x^{k-1})$$ is lower than a given criterion. It exits with an error message whenever the estimated orbit has become hyperbolic or ballistic.

B. Implementation of the Parameter Estimation Problem

Formulation of the ODE

All perturbations to Kepler orbits are taken into account as discussed in section II if the satellite trajectory does not lead to a crash on Earth. If the satellite is on a ballistic trajectory, meaning that for the perigee of the orbit the equation

$$r_{min} - \frac{p}{1+e} \leq R_0, \quad (48)$$

holds, then the simplified right hand side (1) without perturbation is used.

Some general satellite properties that need to be known in order to compute the perturbations are given in table 1. The values were provided by ESA.

<table>
<thead>
<tr>
<th></th>
<th>$m_{sat}[kg]$</th>
<th>$C_D$</th>
<th>$C_R$</th>
<th>$A_D[m^2]$</th>
<th>$A_R[m^2]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARTEMIS</td>
<td>3098.0</td>
<td>1.0</td>
<td>1.4</td>
<td>7.75</td>
<td>14.61</td>
</tr>
</tbody>
</table>

Table 1: General satellite properties: Masses, air drag coefficient $C_D$, solar radiation pressure coefficient $C_R$ and areas $A_D$ and $A_R$ exposed to air drag and solar radiation.

Measurement Functions

The measurement functions are implemented as given in the equations (53), (54) and (55) (see appendix), which are up to minor correction terms identical to the ones given in the equations (15), (19) and (21).

For the evaluation of the range rate measurement function, two evaluations of the range measurement function at times $t_M$ and $t_M + t_c$ are necessary. Considering the fact that the count time intervals $t_c$ are usually small (typically a few seconds), a step with Heun’s method\textsuperscript{16} provides sufficient accuracy.

The ambiguity of measurements of the azimuth angle $\angle \rightarrow \angle + 2\pi$ is addressed by adding or subtracting $2\pi$ if the residual of the azimuth angle is not in $[-\pi, \pi]$. 

American Institute of Aeronautics and Astronautics
Choice of Weights

For a full description of a parameter estimation problem we have to define weights for the measurements of different type. Different choices for the weights result in different parameter estimation problems, which usually also have a different solution. As discussed in section IV, the weights should be chosen as the reciprocal of the standard deviation of the measurement error for measurements of a specific type in order to obtain a maximum likelihood estimate for the optimization variables. However, standard deviations for the applied measurement types are not known and we therefore use the values given in table 2.

<table>
<thead>
<tr>
<th>Measurement Type</th>
<th>(1/w)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range ([km])</td>
<td>0.02</td>
</tr>
<tr>
<td>Velocity ([km/s])</td>
<td>0.00003</td>
</tr>
<tr>
<td>Azimuth angle ([deg])</td>
<td>0.02</td>
</tr>
<tr>
<td>Elevation angle ([deg])</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Table 2: Choice of weighting factors for different measurement types. Please note that the weights for the angles are given in degree here. In radiant, the reciprocal weights \(1/w\) for azimuth and elevation angle would be approximately \(3.49 \cdot 10^{-3}\).

Given the weights, the solutions obtained for the parameter estimation problem is the state vector given in table 4. Additional parameters are not estimated.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(x[km])</td>
<td>5046.07</td>
</tr>
<tr>
<td>(y[km])</td>
<td>-7018.14</td>
</tr>
<tr>
<td>(z[km])</td>
<td>-263.990</td>
</tr>
<tr>
<td>(v_x[km/s])</td>
<td>7.87936</td>
</tr>
<tr>
<td>(v_y[km/s])</td>
<td>2.02064</td>
</tr>
<tr>
<td>(v_z[km/s])</td>
<td>-0.401513</td>
</tr>
</tbody>
</table>

Figure 4: Actual ARTEMIS trajectory Solution state \(s_{0,sol}\) for the ARTEMIS scenario for the choice of weights given in table 2. The state vectors are given in the J2000 reference frame.

Constraints

No additional equality or inequality constraints were formulated for the parameter estimation problem. It is thus an unconstrained optimization problem in the single shooting mode, but becomes an equality-constrained problem in multiple shooting mode because of the matching conditions (24).
C. Results for Artemis Scenario

Regarded Measurements and Multiple Shooting Grid

We consider a tracking arc of 21478.4s of flight (\(\approx 6h\)) starting at the time where the satellite separated from the rocket (epoch). 103 measurements were performed during that time, 76 of which are measurements of azimuth elevation angle pairs and 27 are range measurements. Range rate measurements were not performed within the considered tracking arc.

For the application of multiple shooting we use all measurement times plus the epoch as multiple shooting nodes, meaning that we have altogether 104 multiple shooting nodes. Their distribution is given in figure 5.

Test Problems

In order to compare single shooting to multiple shooting for the solution of the nonlinear least squares problem, a set of initial guesses was generated using a homotopy from the solution to the original initial guess for the ARTEMIS scenario:

\[
\begin{align*}
    s^0_0(\beta) - (1 - \beta) \cdot s_{0,\text{sol}} + \beta \cdot s_{0,\text{expected}} \\
    = (1 - \beta) \cdot \begin{pmatrix} 5046.07 \\ -7018.14 \\ -263.990 \\ 7.87936 \\ 2.02064 \\ -0.401513 \end{pmatrix} + \beta \cdot \begin{pmatrix} 4368.15 \\ -7141.46 \\ -209.580 \\ 8.696263 \\ 1.836859 \\ -0.279152 \end{pmatrix}
\end{align*}
\]

where for the choice \(\beta = 0\) the solution of the problem is used as initial guess and for the choice \(\beta = 1\) the original ARTEMIS initial guess provided by ESA, that is the expected trajectory, is used. We tested

\[
\beta \in \{0.0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0, 3.0, 5.0\},
\]

meaning that we also considered values for \(\beta\) outside the natural range \([0, 1]\) of the homotopy. A subset of the trajectories obtained with these single shooting initial guesses are illustrated in figure 6.
Figure 6: Illustration of initial values. The initial positions of all three trajectories are located in the dotted circle at the bottom. With time, the satellite moves in the direction of the black arrow (1), i.e. counterclockwise. The true trajectory (solution) is the inner black line. It corresponds to $\beta = 0.0$ and appears as a closed ellipse because the revolution time of the ARTEMIS satellite is less than the considered time horizon of 6 hours. The black arrow (2) points to the position where Artemis was after 6 hours. The ESA initial guess is displayed in gray ($\beta = 1.0$), further the trajectories corresponding to homotopy parameters $\beta = 0.4$ (black, dashed) and $\beta = 2.0$ (gray, dashed) are displayed.

**Generation of Multiple Shooting Initial Guesses with the Analytical Projection**

Whether or not the projection strategy for the generation of initial guesses described in section VII is reasonable or not can only be decided by comparing the convergence behaviour of the multiple shooting method to that of the single shooting method. Nevertheless, we also want to visually compare the projected positions to the unprojected positions and the true positions of the satellite.

Therefore we consider figure 7, which shows three trajectories: The true trajectory of the ARTEMIS satellite, a single shooting initial guess obtained by choosing the homotopy parameter as $\beta = 0.4$, and a multiple shooting initial guess that is obtained by applying the analytical projection. We find that in the beginning the projected positions (dashed gray) are so close to the true trajectory (black) that the two curves overlay. However, during the time where no measurements were done we can see a significant error propagation (starting at (1)). After the time without measurements we need several projections until the projected positions are close to the true positions again (see (2)). Note that the single shooting initial guess leads to a completely different trajectory: After six hours it has not yet completed one revolution and is, compared to the true position of ARTEMIS after that time, on the other side of the Earth at a distance of more than 25000 km.

**Comparison of Single Shooting and Multiple Shooting**

The visual comparison is only a first, though strong, indication that the projection strategy and multiple shooting are useful for initial satellite orbit determination. In this paragraph we now show that this approach does indeed also lead to an improvement of the convergence behaviour of the Gauss-Newton method.
Figure 7: Illustration of the true trajectory of the ARTEMIS satellite (black), the single shooting initial guess for $\beta = 0.4$ (gray) and the multiple shooting initial guess generated by the analytical projection (dashed gray). Again, the initial positions of all three trajectories are located in the dotted circle. The time interval without measurements, cf. figure 5, corresponds to the curve between points (1) and (2). After the considered tracking arc of 6 hours, multiple shooting initial guess and true position are close again (3).

Therefore we have a look at table 3. Therein the number of iterations are listed until convergence was achieved.

As a criterion for convergence we demand that the scaled norm of the increment fulfills $|S\Delta x| \leq 10^{-5}$. Again, we emphasize that for the comparison between the single shooting and the multiple shooting mode of PARFIT all additional parameters concerning the termination criterion, scaling and so on were kept fixed; the only difference are the number of shooting nodes and the usage of the projection method.

As we can see, the ESA software BAHN and PARFIT in single shooting mode converge only for very good initial guesses. Already if the homotopy parameter $\beta$ is chosen to be 0.2, no convergence can be obtained with the single shooting approach. In contrast to this, the multiple shooting approach leads to considerably better results: Convergence was achieved not only for the original ESA initial guess ($\beta = 1$), but even for initial guesses much worse than that. Even for highly hyperbolic initial guesses ($\beta = 5.0$, corresponds to an eccentricity $e = 1.82$) the solution was obtained.

The failures of the single shooting mode in PARFIT are due to the fact that the maximum number of iterations was reached (set to 100 for all tests). The ESA software BAHN did moreover terminate whenever the estimated orbit became hyperbolic.

D. Further Verification of the Benefit of Multiple Shooting

Random Initial Guesses

More tests were performed in order to evaluate the advantage of multiple shooting even better. Therefore, random initial guesses were tried out, in order to check whether the good performance of multiple shooting observed in the
Table 3: Comparison of single shooting and multiple shooting: Iterations needed until convergence was obtained. For $\beta = 3.0$ and $\beta = 5.0$ the initial orbit is hyperbolic, and hyperbolic initial guesses are not allowed as input in BAHN.

last paragraph is only due to the specific choice of initial guesses.

The test set of initial guesses was obtained using a generator of Gaussian random numbers. All components $i$ of the solution state vector $s_{0,\text{sol}}$ were disturbed by errors

$$\epsilon_i \in \mathcal{N}(0, (\sigma_g \cdot (s_{0,\text{sol}})_i)^2),$$

that is we generate Gaussian random number with a relative standard deviation $\sigma_g$ and add them to the corresponding component of the $s_{0,\text{sol}}$. Altogether, 21 single shooting initial guesses were obtained following this strategy, three state vectors for each of the following values of relative standard deviations:

$$\sigma_g = \{0.025, 0.05, 0.10, 0.20, 0.30, 0.50, 0.75\}.$$  

The results obtained for the random initial guesses are listed in table 4. Again we observe a significant progress with the multiple shooting method: While single shooting converges only for seven of the 21 generated initial guesses, the multiple shooting method (initialized with the results of the projection method) successfully converges to the solution for all test cases.

Note, that among the tested set of random initial guesses are already highly hyperbolic (eccentricity $e > 3$) and ballistic trajectories, that is extremely poor initial guesses. But still the multiple shooting approach reliably leads to a convergence of the generalized Gauss-Newton method to the solution of the problem. We therefore conclude that if the measurement information is sufficient for a successful application of the projection strategy, the parameter estimation problem can be solved with the multiple shooting method for virtually every practically relevant initial state vector.
Table 4: Comparison of single shooting and multiple shooting: Iterations needed until convergence was obtained. The word "hyperbolic" indicates that this initial guess could not be supplied to BAHN because it is a hyperbolic orbit.

<table>
<thead>
<tr>
<th>$\sigma_q$</th>
<th>BAHN (single shooting)</th>
<th>PARFITBAHN (single shooting)</th>
<th>PARFITBAHN (multiple shooting)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.025</td>
<td>no convergence</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>0.025</td>
<td>no convergence</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>0.025</td>
<td>no convergence</td>
<td>no convergence</td>
<td>17</td>
</tr>
<tr>
<td>0.05</td>
<td>no convergence</td>
<td>no convergence</td>
<td>5</td>
</tr>
<tr>
<td>0.05</td>
<td>8</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>0.10</td>
<td>no convergence</td>
<td>25</td>
<td>11</td>
</tr>
<tr>
<td>0.10</td>
<td>no convergence</td>
<td>no convergence</td>
<td>11</td>
</tr>
<tr>
<td>0.20</td>
<td>no convergence</td>
<td>no convergence</td>
<td>10</td>
</tr>
<tr>
<td>0.20</td>
<td>no convergence</td>
<td>no convergence</td>
<td>5</td>
</tr>
<tr>
<td>0.20</td>
<td>hyperbolic</td>
<td>no convergence</td>
<td>32</td>
</tr>
<tr>
<td>0.30</td>
<td>no convergence</td>
<td>87</td>
<td>15</td>
</tr>
<tr>
<td>0.30</td>
<td>no convergence</td>
<td>no convergence</td>
<td>10</td>
</tr>
<tr>
<td>0.30</td>
<td>hyperbolic</td>
<td>no convergence</td>
<td>27</td>
</tr>
<tr>
<td>0.50</td>
<td>no convergence</td>
<td>56</td>
<td>20</td>
</tr>
<tr>
<td>0.50</td>
<td>hyperbolic</td>
<td>no convergence</td>
<td>29</td>
</tr>
<tr>
<td>0.50</td>
<td>no convergence</td>
<td>no convergence</td>
<td>16</td>
</tr>
<tr>
<td>0.75</td>
<td>hyperbolic</td>
<td>no convergence</td>
<td>25</td>
</tr>
<tr>
<td>0.75</td>
<td>hyperbolic</td>
<td>no convergence</td>
<td>19</td>
</tr>
<tr>
<td>0.75</td>
<td>no convergence</td>
<td>no convergence</td>
<td>63</td>
</tr>
</tbody>
</table>

Benefit of Range Rate Measurements

The third common measurement type for initial orbit determination is the range rate measurement. However, the first range rate measurements in the ARTEMIS scenario were done after more than 17 hours of flight and already 177 performed range or angular measurements. This is considered to be too much in order to investigate whether the exploitation of additional range rate measurements leads to further improvements of the convergence behaviour of the multiple shooting method.

Table 5: Benefit of range rate measurements

<table>
<thead>
<tr>
<th>Exploitation of measurement information of given type in the analytical projection</th>
<th>Problems solved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angle</td>
<td>Range</td>
</tr>
<tr>
<td>Single Shooting</td>
<td>no</td>
</tr>
<tr>
<td>Multiple Shooting</td>
<td>no</td>
</tr>
<tr>
<td>Multiple Shooting</td>
<td>yes</td>
</tr>
<tr>
<td>Multiple Shooting</td>
<td>yes</td>
</tr>
</tbody>
</table>

We therefore had a look at the tracking data of the launch and early orbit phase of the second pair of CLUSTER-II
satellites, launched on August 9th 2001 from Baikonur. Due to a underperformance of the Soyuz first stage the apogee height was approximately 1000 km lower than planned. Tracking data were also provided by ESA/ESOC.

The considered tracking arc was chosen as the first 2910.4 s after the separation. During this time 195 measurements were performed, 76 of which are range rate measurements, i.e. more than one third. The remaining measurements are divided in angular measurements (51) and range measurements (68).

For this scenario we investigated the effect of exploiting the range rate measurements for the generation of multiple shooting initial guesses under the two guiding questions: First, if only range rate measurements are used in the analytical projection, does this already lead to a improvements compared to single shooting. And second, in a comparison of two multiple shooting methods, where in one case only angle and range measurements are used in the analytical projection and in the other case measurements of all three measurement types are used, does the latter one benefit significantly from the additional information in range rate measurements?

For the investigation, 21 random initial guesses were generated in the same manner as done in the ARTEMIS scenario, that is using each of the seven relative standard deviations as listed in (51) for the generation of three Gaussian disturbed initial state vectors to the solution

\[ s_{0, sol} = \begin{pmatrix} -5847.78, -350.767, 4502.82, 2.79892, -4.51364, 7.34956 \end{pmatrix}^T. \] (52)

The epoch of this state vector is 2000 – 08 – 09, 12 : 43 : 25.651 UTC.

Concerning the two guiding questions the following answers were found: Already the exploiting of the range rate measurements leads to a better convergence behaviour of the multiple shooting method compared to single shooting method. The number of runs where no convergence was achieved was reduced from 5 to 3. However, for the comparison of the two multiple shooting methods we found that in both cases all 21 problems were solved. The results are summarized in table 5.

**IX. Conclusions**

A reliable determination of the initial orbit of a satellite is of utmost importance. It needs to be done as soon as possible after the separation of the satellite from the launcher for operations planning. However, in case of a significant misperformance of the launcher, the nominal initial state vector is very often insufficient for obtaining convergence with the well-known and commonly used direct single shooting method.

In this article, we have introduced the direct multiple shooting method for initial satellite orbit determination. We have presented results of a comparison between our method and direct single shooting using operational ARTEMIS and CLUSTER-II tracking data.

It was clearly demonstrated that multiple shooting combined with an effective strategy for the determination of
initial guesses leads to dramatic improvements concerning the convergence behavior of the generalized Gauss-Newton method for the solution of the parameter estimation problem. Successful convergence was obtained even for extremely poor initial guesses such as those corresponding to ballistic or escape trajectories.

**Appendix: Measurement Functions**

The full measurement functions for range, range rate and angular measurements are given as follows:

\[
m_{\text{range}}(t_M, y(t_M)) = \rho(t_M) + \left( \frac{\rho(t_M)}{2c} - \Delta_{tb} \right) \frac{\ddot{\rho}(t)}{\ddot{t}} \bigg|_{t=t_M} + \Delta_{td} \cdot c + \Delta_{cb} + \Delta_{atm}\]  \(53\)

\[
m_{\text{range rate}}(t_M, y(t_M)) = \frac{\dot{m}_{\text{range}}(t_M + t_c, y(t_M + t_c)) - \dot{m}_{\text{range}}(t_M, y(t_M))}{t_c} + \Delta_{cb} + \Delta_{atm}.\]  \(54\)

\[
m_{\text{ang}}(t_M, y(t_M)) = \begin{pmatrix} m_{\text{azimuth}}(t_M, y(t_M)) \\ m_{\text{elevation}}(t_M, y(t_M)) \end{pmatrix} - \begin{pmatrix} A(t) + \left( \frac{\Delta_{td}}{2c} - \Delta_{tb} \right) \frac{\ddot{A}(t)}{\ddot{t}} \bigg|_{t=t_M} + \Delta_{cb} \\ E(t) + \left( \frac{\Delta_{td}}{2c} - \Delta_{tb} \right) \frac{\ddot{E}(t)}{\ddot{t}} \bigg|_{t=t_M} + \Delta_{cb} + \Delta_{atm} \end{pmatrix}.\]  \(55\)

In equation (54) we have used the auxiliary range measurement function

\[
\dot{m}_{\text{range}}(\hat{t}, y(\hat{t})) = \rho(\hat{t}) + \left( \frac{\rho(\hat{t})}{2c} - \Delta_{tb} \right) \frac{\ddot{\rho}(\hat{t})}{\ddot{t}} \bigg|_{t=\hat{t}}.\]  \(56\)

Let us now discuss the correction terms in detail: In all measurement functions the station timing bias \(\Delta_{tb}\) occurs, which is the same for all measurement types, but different for every ground station. It accounts for the fact that between the nominal beginning of a measurement and the actual beginning there can be small deviations.

The transponder delay \(\Delta_{td}\) takes into account the difference between signal reception and transmission at the satellite. It only effects the range measurement, because only the range measurement is based on the signal travel time.

The observation component bias \(\Delta_{cb}\) is different for every measurement type, although for simplicity the same symbol was used in all equations. It accounts for a general offset to the result of a measurement.

All the systematic measurement errors \(\Delta_{tb}, \Delta_{td}\) and \(\Delta_{cb}\) are known from the calibration of the measurement system.

Finally, an atmospheric correction term \(\Delta_{atm}\) has to be taken into account for all measurements except for the azimuth angle. However, it should be mentioned that for the atmospheric correction term \(\Delta_{atm}\), only the influence of the troposphere is taken into account by using the model described in Marini.\(^{20}\) Corrections of the ionosphere are not included. For details concerning the calculation of atmospheric disturbances we refer to the literature.\(^{14}\)
Acknowledgments

The members of the Simulation and Optimization working group at the Interdisciplinary Center for Scientific computing in Heidelberg, H. G. Bock, J. P. Schlöder and S. M. Lenz, and E. A. Kostina from the University of Marburg, thank the European Space Agency (ESA) for the financial support of this work within ESOC study contract 14802/00/D/IM.

References


24. S. Pallaschke. *Introduction to the Multi-Project Orbit Determination Subsystem, Short Description of Program BAHN*. European Space Operations Center (ESOC), Darmstadt, 1999.