Verified Interval Orbit Propagation in Satellite Collision Avoidance

Bart A. Römgens*, Erwin Mooij† and Marc C. Naeije‡

Delft University of Technology, Delft, 2600 AA, The Netherlands

Verified interval integration methods enclose a solution set corresponding to interval initial values and parameters, and bound integration and rounding errors. Verified methods suffer from overestimation of the solution, i.e., non-solutions are also included in the solution enclosure. Two verified integration methods, interval Taylor-series and Taylor-model based methods and their implementation in VNODE-LP and VSPODE, are used to reduce overestimation in verified satellite orbit propagation. Furthermore, two orbital state models based on integration constants, the modified equinoctial elements (MEE) and unified state model (USM), are used to reduce overestimation. Earth-satellite trajectories propagated using VSPODE have 2-3 times less overestimation than those propagated using VNODE-LP. Using the USM and MEE state models, overestimation is further reduced by a factor 4 to 10, depending on initial and parameter uncertainties. It is demonstrated that verified collision detection is feasible and may contribute to prevent satellite collisions to reduce future space debris.

I. Introduction

Conventional numerical satellite-trajectory propagation methods are based on fixed-point initial values and parameters, and give approximate fixed-point solutions at a fixed-point time. These methods are approximations because they include truncation (integration) errors that propagate and accumulate throughout the integration, possibly causing large deviations from the true solution for unstable systems, like the along-track instability of satellites. The size of these errors is estimated at best, but often ignored and assumed to be insignificant.

In reality, initial values come from uncertain measurements and we may want to know the solution over a time interval. Furthermore, we want to place bounds on truncation errors so that we can be certain that our simulation translates to reality. Verified integration methods allow us to deal with both uncertain initial values and parameters, represented by an interval, and the numerical integration errors, which can be bounded by finite upper and lower bounds.

Several verified integration methods to solve ordinary differential equations have been developed by Makino and Berz, Lin and Stadtherr, Nedialkov and others. Neher, Nedialkov and Hoefkens give an overview of the different methods.

Solutions created using verified integration can be represented by an interval valid over a time interval. Applied to satellite orbits, we can guarantee by mathematical proof that a satellite will be within a certain position interval, during a certain time interval. The only assumption we have to make is that the satellite starts within a certain position and velocity interval.

Berz and Makino were the first to apply verified integration to a satellite orbit example in 2001. Hoefkens did the same in 2001 and 2003, comparing two methods. Alessi et al. used interval integration to integrate the orbit of asteroid Apophis in 2007. The major problem with verified integration is the overestimation of the solution set, i.e., the inclusion of non-solutions in the solution enclosures. This means that verified integration methods can guarantee a solution to be within certain bounds, but not all points within the bounds are true solutions. Significant decreases in overestimation have to be made before verified integration can be useful in real non-linear problems like orbit propagation.

---

* Graduate Student, Faculty of Aerospace Engineering, bart.romgens@gmail.com.
† Assistant Professor, Faculty of Aerospace Engineering, e.mooij@tudelft.nl, Senior member AIAA.
‡ Assistant Professor, Faculty of Aerospace Engineering, m.c.naeije@tudelft.nl.
In this paper, we will try to reduce the overestimation by using and comparing VSPODE and VNODE-LP for verified integration, and two perturbation-based state models to model the motion of Earth satellites. These models are based on constants or linear changes and may therefore suffer less from overestimation.

The feasibility of the application of verified interval integration to Earth-satellite collision detection is investigated. Space debris is an increasing problem and a single collision can greatly increase the amount of debris in certain orbit ranges. It is therefore vital that we can guarantee that satellites will not collide, and adjust orbits when this cannot be guaranteed. Verified integration could provide this guarantee. The collision between Iridium-33 and Cosmos-2251 is reconstructed to investigate the feasibility of verified integration and interval analysis to detect satellite collision.

We start with an introduction to interval analysis (Section II). Next, verified integration of ordinary differential equations (ODEs) is introduced in Section III. In Section IV, the different orbital state models that should reduce some disadvantages of verified integration will be discussed. An application of verified integration to satellite-orbit propagation and satellite-collision detection is presented in Section V. Section VI, finally, will conclude this paper with some final remarks and recommendations.

II. Interval Analysis

Interval analysis (also called interval arithmetic or interval computing) was developed by Moore in the early 1960s. It can be seen as an extension of real numbers, an interval, represented by a pair of real numbers, the lower and upper bound. This interval contains all real numbers between its minimum and maximum. This section gives a brief overview of the most important rules of interval analysis. The reader is referred to the works of Moore for more information about interval analysis.

In mathematical terms, an interval is a closed bounded set of real numbers

\[ [a, b] = \{ x : a \leq x \leq b \} \] (1)

We denote intervals by \([x]\), and the endpoints of the interval by \(x\) and \(\bar{x}\). Thus, \([x] = [x, \bar{x}]\).

II.A. Interval Arithmetic

Interval arithmetic describes the rules of elementary operations, like addition and multiplication, for intervals. The rules for addition and subtraction are,

\[ [x, \bar{x}] + [y, \bar{y}] = [x + y, \bar{x} + \bar{y}] \] (2)
\[ [x, \bar{x}] - [y, \bar{y}] = [x - \bar{y}, \bar{x} - y] \] (3)

The rules for multiplication,

\[ [x] \cdot [y] = [\min(xy, x\bar{y}, \bar{x}y, \bar{x}\bar{y}), \max(xy, x\bar{y}, \bar{x}y, \bar{x}\bar{y})] \] (4)

When \([x]\) is an interval not containing the number 0, the multiplicative inverse (reciprocal) is defined as,

\[ \frac{1}{[x]} = \left[ \frac{1}{\bar{x}}, \frac{1}{x} \right] \] (5)

However, when \([x]\) contains 0, the set is unbounded and cannot be represented as an interval whose endpoints are real numbers; both bounds go to infinity. This causes the loss of valuable information that cannot be contained in a single interval. However, splitting the interval in two parts, \([-\infty, 1/y_1]\) and \([1/y_2, \infty]\), and using two intervals can contain the information.

The division of two intervals is defined by

\[ \frac{[x]}{[y]} = [x] \cdot \frac{1}{[y]} \] (6)

and can therefore be computed using Eq. (4) and (5).

Interval arithmetic has algebraic properties similar to real numbers. Addition and multiplication are associative, commutative, but not always distributive. For example

\[ [1, 2] \cdot (1 - 1) = [0, 0] = 0 \] (7)
However

\[ [1, 2] \cdot 1 - [1, 2] \cdot 1 = [-1, 1] \neq 0 \quad (8) \]

This means that interval functions are not uniquely defined; the resulting bounds depend on the form of the function. However, the following algebraic property always holds

\[ [x](y + z) \subseteq [x][y] + [x][z] \quad (9) \]

This property is called subdistributivity and implies that if we rearrange an interval expression, we may obtain tighter bounds, i.e., a smaller interval.\(^{14}\)

Furthermore, the width of an interval, \(w([x])\), is defined as the difference between the upper and lower bounds,

\[ w([x]) = \bar{x} - \underline{x} \quad (10) \]

II.B. Inclusion Monotonic Interval Extensions of Functions

An interval extension of a function \(f(x_1, \ldots, x_n)\) is an interval-valued function \(F([x_1], \ldots, [x_n])\) with the property

\[ f(x_1, \ldots, x_n) = F(x_1, \ldots, x_n) \quad (11) \]

So an interval extension of \(f\) is an interval valued function, which has real values when the arguments are all real and coincides with \(f\).

\(F\) is an inclusion monotonic interval extension of \(f\) if

\[ f([x_1], \ldots, [x_n]) \subseteq F([x_1], \ldots, [x_n]) \quad (12) \]

This means that an inclusion monotonic interval extension \(F\) gives bounds on the function when the function variables are within the intervals \([x_1], \ldots, [x_n]\).

For example (from Moore\(^{13}\)), consider the polynomial

\[ p(x) = 1 - 5x + \frac{1}{3}x^3 \quad (13) \]

and suppose we want the range of all values of \(p(x)\) when \(x\) is a number in the range \([2, 3]\), i.e., \(x \subseteq [2, 3]\).

A natural interval extension of \(p\) is

\[ P([x]) = 1 - 5[x] + \frac{1}{3}[x] \cdot [x] \cdot [x] \quad (14) \]

Computing \(P([2,3])\) using interval arithmetic

\[ P([2, 3]) = 1 - 5[2, 3] + \frac{1}{3}[8, 27] = [-\frac{34}{3}, 0] \quad (15) \]

All rational functions (the function can be expressed as a ratio of two polynomials) have natural inclusion monotonic interval extensions; we only have to replace the real variables with interval variables and real arithmetic operations with interval arithmetic operations. Moreover, one can construct inclusion monotonic interval extensions of all commonly used non-rational functions.

II.C. Overestimation

The bounds placed on functions by evaluating them using interval analysis are generally larger than the true bounds of the function. This so-called overestimation, including values that are not in the range of the function, is mainly due to two reasons, the dependency problem and the wrapping effect.\(^{13}\) Methods to reduce this overestimation are discussed in Section III.B.
The Dependency Problem

Interval analysis cannot detect multiple occurrences of the same variable. This can lead to overestimation as we demonstrate with the following example. In fixed-point arithmetic, $x - x = 0$. But if we evaluate this expression using interval arithmetic for all $x \in [1, 2]$, the result is $[-1, 1]$. The true solution, $[0, 0]$, is contained in the result, but contains a large overestimation. This cause of overestimation is called the dependency problem.

Writing an interval function in a different form can change the enclosure. It may be useful to rewrite the function such that the overestimation is reduced, but removing the dependency is not always possible. The dependency problem can be the cause of significant overestimation and should always be kept in mind when using interval arithmetic.\(^8,13\)

The Wrapping Effect

Enclosing a 3-dimensional set (shape) that is not a rectangular cuboid (or parallelepiped, in the general sense) in an interval, always includes points that are outside the set we try to enclose. In more practical terms: you cannot fit a ball in a box without leaving some empty space. This is visualized in Fig. 1 where the 3-dimensional set (red) is enclosed in a three dimensional interval box (green). There are many points inside the green box that are not part of the original set; wrapping the set in a new set introduced elements that were not part of the original set. This cause of overestimation is called the wrapping effect. It becomes a significant problem when used in iterative methods like numerical integration, where rotated intervals are enclosed by new intervals.

III. Verified Integration of Ordinary Differential Equations

A wide range of conventional numerical integration methods is available to compute solutions to differential equations. However, these methods only yield approximate solutions due to truncation errors that are at best approximated, but not placed between guaranteed bounds. Moreover, traditional methods can only solve for fixed-point initial values, instead of interval-valued initial values. A traditional solution is a single point in $n$-dimensional space. Also, traditional solutions are valid at a single moment in time and not over a time interval.

Verified (also called validated or guaranteed) integration methods are part of a group of integration methods that provide guaranteed bounds, by mathematical proof, on solutions. These methods bound truncation errors and can handle interval initial values and parameters. This guarantee is particularly interesting for critical dynamical systems with high safety requirements such as chemical reactions,\(^2\) but can also be used for verified global optimization, sensitivity analysis or to verify results of conventional methods.
The first idea of verified integration came from Moore in 1968 when he worked on interval analysis, a method to bound functions for a range (interval) of variables. His method is based on bounding truncated Taylor series expansions by using interval arithmetic to evaluate the Taylor series at each integration step, and using these bounds as initial values for the next step. Development of interval-analysis based methods was continued by Eijgenraam, Löhner, Corliss, Rihm and others. These methods, based on interval analysis and Taylor series, are called Interval Taylor Series (ITS) methods.

A second group of verified integration methods was developed by Makino and Berz in the 1990s, initially to analyse the stability of particles in particle accelerators. Their method uses Taylor models (a Taylor polynomial with an interval remainder bound) to propagate higher-order information about the dependency of the solution on initial values. A good overview of Taylor models and a theoretical comparison with ITS methods is given by Neher et al.

A new hybrid method combining ITS with TM methods was created by Lin and Stadtherr in 2007. It uses the traditional first steps of ITS methods to verify existence and uniqueness. However, where ITS methods evaluate the coefficients of the Taylor series using interval analysis, their method uses Taylor models for the coefficient to express the dependency on the interval initial values and parameters, reducing overestimation. They implemented their method in VSPODE.

Verified integration has two important disadvantages. First, it often overestimates the solution set which is mainly the result of the dependency problem and the wrapping effect. Most research focuses on reducing this overestimation. Second, verified integration is computationally more expensive than conventional integration. For the more complex algorithms and when using high orders, verified integration can be a few orders of magnitude slower.

### III.A. Bounding Taylor Series

Most verified methods use Taylor-series expansions of state variables with respect to time

$$x(t_0 + h) = x(t_0) + hf(x(t_0)) + \frac{1}{2!} h^2 f'(x(t_0)) + ... + \frac{1}{n!} h^n f^{(n-1)}(x(t_0)) + R_n(t)$$  \hspace{1cm} (16)

and then find bounds on this Taylor series by evaluating it using either interval analysis or Taylor models, and bounding the remainder term, $R_n(t)$, which represents all the higher-order terms. This is the truncation error in numerical integration.

#### Lagrange Remainder Term

The remainder term, $R_n$, can be enclosed by several expressions, the most commonly used is the Lagrange form of the remainder term,

$$R_n(t) = \frac{f^n(x(\xi))}{(n+1)!} h^{n+1}$$  \hspace{1cm} (17)

where we first need to enclose the solution $x(t)$ on the interval $t \in [t_i, t_{i+1}]$, which is what we already try to do by bounding the remainder term $R_n(t)$; we need a-priori bounds on $x(t)$ to create tighter bounds on $x(t)$.

A crude a-priori enclosure of the solution is found by validating existence and uniqueness of a solution within a certain interval, $[\hat{x}]$, on a certain time interval $h = (t_1 - t_0)$.

Validating the existence and uniqueness is usually done using the Picard-Lindelöf theorem and Picard iterator, which are results of the Banach fixed-point theorem. These theorems are modified for the interval case and require an initial guess for the enclosure, $[\hat{x}]$, and stepsize, $h$, which can be checked for existence and uniqueness using the Picard iterator in interval arithmetic.

With a known a-priori enclosure, $[\hat{x}]$, we can write the remainder term in Eq. (17) as

$$[R_n(t)] = \frac{f^n([\hat{x}])}{(n+1)!} h^{n+1}$$  \hspace{1cm} (18)
Taylor Coefficients

The remainder part of the Taylor series expansion is now enclosed by an interval \([ R_n(t) ]\) using a rough solution enclosure. The Taylor series’ coefficients, up to order \(n\), can be bound by evaluating the series using interval arithmetic (ITS methods), making use of the fact that a Taylor series is an inclusion monotonic function. Or by writing the coefficients as Taylor models that contain higher-order information about the dependency on the initial values (TM methods).

A Taylor model consists of a polynomial, \(P\), and interval, \([R]\), part,

\[
TM = P + [R]
\]

also denoted as, \(TM = (P, [R])\). Just as with intervals, there is an arithmetic for computing with Taylor models; the reader is referred to the work of Makino and Berz,\(^{21}\) and Neher\(^{4}\) for more information about Taylor models and Taylor-model arithmetic.

The polynomial part is the Taylor-series expansion of finite order \(n\), and the interval part contains truncation and rounding errors and depends on the algorithm used to compute it. Any function that can be represented on a computer and is continuously differentiable can be modeled (bounded) by a Taylor model.

The Taylor-model methods do not result in an interval representation, but a Taylor-model representation of the solution set. This Taylor model representation is generally tighter than an interval representation, which is a cuboid or parallelepiped, and can also be evaluated using interval analysis (wrapped in an interval) to produce interval solutions.

Taylor-series coefficients are computed in a symbolic-numeric way using automatic differentiation, a way to accurately compute Taylor coefficients up to an arbitrary order using the chain-rule.

III.B. Reducing Overestimation

Verified integration of ODEs is an iterative process where a new Taylor-series expansion of the solution is created for each integration step. Each step uses the solution set of the previous step as initial value set. These initial values are represented by intervals for ITS methods and Taylor models for TM methods.

Overestimation, the inclusion of non-solutions in the solution enclosure, occurs in both methods since they all use some form of interval arithmetic and therefore suffer from both the dependency problem and the wrapping effect. Because the integration is an iterative process, the overestimated solution is used as initial value for the next step, where the solution is again wrapped in an interval including even more non-solutions. This iterative wrapping can quickly cause the solution to grow out of bounds, rendering it useless; i.e., the solution has exploded.

A wide range of methods has been created to reduce the overestimation. Taylor models instead of direct interval evaluation of Taylor series were invented to reduce the huge overestimation of interval methods, especially for non-linear differential equations.\(^{22}\)

There are several methods to reduce overestimation for the different types of verified integration methods. The major strategies are subdivision of the intervals, coordinate transformations, shrink wrapping and preconditioning.

Subdivision divides the problem into multiple problems with smaller intervals.\(^{8}\) This is a simple and effective method to reduce the wrapping effect, but suffers from the curse of dimensionality.

Coordinate transformations reduce the wrapping effect. Instead of using an interval enclosure, \([x]\), defined by the unit vectors of the original coordinate system, we can also represent the solution as a single point with a locally transformed interval box around it\(^{23}\)

\[
[x] = m([x]) + A[r]
\]

where \(A\) is the transformation matrix and \([r]\) the interval vector. Different coordinate transformations exist, the most commonly used are QR factorization (orthogonal) and the parallelepiped (non-orthogonal) method.\(^{3}\) This coordinate transformation using an orthogonal transformation is visualized in Fig. 2 where the left graph shows the interval enclosure (grey) of a true solution set (white) in the original coordinate system. The right graph shows the coordinate transformation, and the new smaller interval enclosure of the same solution set.

Shrink wrapping and preconditioning are methods created by Berz and Makino\(^{24}\) to reduce overestimation in their TM-based method. Shrink wrapping absorbs a part of the interval remainder term in the symbolic
polynomial part of the Taylor Model, $P$, by modifying the polynomial’s coefficients. Preconditioning is a coordinate transformation that makes the solutions sets less prone to wrapping, not very different from QR-factorization methods. The reader is referred to the works by Berz and Makino, Neher and Hoefkens for more information about shrink wrapping and preconditioning.

**IV. Orbital State-Propagation Models**

The state of an orbiting body, and the related equations of motion, are often described in Cartesian coordinates. These Cartesian coordinates have a physical meaning (position and velocity) and the orthogonality of the system results in simple equations of motion and transformations. It is therefore the preferred state model in many celestial-mechanics applications.

However, an unperturbed two-body problem is an orbit, which shape and orientation can be described by five constants, and the position along the orbit by one variable. The Cartesian model does not use this information, causing a larger error or a smaller stepsize requirement. The two-body orbit-shape and orientation description is the basis for the traditional Kepler elements. Such a state model, including the equations of motion, is often more efficient in numerical integration.

When used with verified integration, the state variables and the "shape" of the solution in the state space have a great influence on the wrapping effect. Enclosing a linear flow does not suffer as much from the wrapping effect as a non-linear flow. It is for this reason that using a state model that is based on small variations of constants and/or linearly changing state variables, is expected to reduce overestimation. This, in turn, results in smaller interval enclosures, a longer time to solution explosion (when the size of the uncertainty box exceeds a certain value) and thus a more useful verified solution.

A large number of orbital state models and related variational equations have been created over the years. The influence on overestimation of two such models, compared to traditional Cartesian coordinates, will be investigated.

The first model is the modified equinoctial element (MEE) set. This set is based on the classical Kepler elements and is modified to remove singularities in the inclination and eccentricity. The MEE set was chosen as it was tested best for traditional orbit integration in a survey of orbit elements by Hintz.

The second model is the unified state model (USM), first proposed by Samuel P. Altman in 1972. It uses quaternions for the orientation of the orbit and the velocity to describe the shape of the orbit, resulting in seven state variables. This model is chosen, because it is based on different integration constants than Kepler-like state models such as the MEE, and has only recently been successfully applied to a few example problems by Vittaldev et al.

The equations of motion of both models are expressed in terms of accelerations in the following orthogonal directions. A radial component, $e_1$, a component in the orbital plane perpendicular to the radius vector, $e_2$, and a component perpendicular to the orbital plane, $e_3$, completing the right-handed coordinate system. The left graph of Fig. 3 shows the direction of $e_1$ and $e_2$. 

![Figure 2. Visualization of local coordinate transformation to reduce overestimation of interval enclosures, using Löhner's QR-factorization method. Source: Nedialkov](image)
IV.A. Modified Equinoctial Elements

The modified equinoctial elements are a modified version of the traditional Kepler elements to remove mathematical singularities. MEE were first introduced by Walker in 1985.²⁹ The relation between the MEE and Kepler elements is as follows,

\[ p = a(1 - e^2) \]
\[ f = e \cos(\omega + \Omega) \]
\[ g = e \sin(\omega + \Omega) \]
\[ h = \tan \frac{i}{2} \cos \Omega \]
\[ k = \tan \frac{i}{2} \sin \Omega \]
\[ L = \Omega + \omega + \theta \]

where \( a \) is the semi-major axis, \( e \) the eccentricity, \( \omega \) the argument of periapsis, \( \Omega \) right ascension of the ascending node, \( i \) the inclination, and \( \theta \) the true anomaly. All but \( L \) are constant for an unperturbed two-body problem. When perturbation forces act on the body, the following differential equations describe the change of the MEE due to these forces³⁰

\[
\frac{dp}{dt} = \sqrt{\frac{p}{\mu}} \frac{2p}{w} a_{c2}
\]
\[
\frac{df}{dt} = \sqrt{\frac{p}{\mu}} \left( \sin(L) a_{c1} + ((w + 1) \cos L + f) \frac{a_{c2}}{w} - \left( h \sin L - k \cos L \right) \frac{a_{c3}}{w} \right)
\]
\[
\frac{dg}{dt} = \sqrt{\frac{p}{\mu}} \left( -\cos(L) a_{c1} + ((w + 1) \sin L + g) \frac{a_{c2}}{w} + \left( h \sin L - k \cos L \right) \frac{f}{w} a_{c3} \right)
\]
\[
\frac{dh}{dt} = \sqrt{\frac{p}{\mu}} \frac{s^2}{2w} \cos(L) a_{c3}
\]
\[
\frac{dk}{dt} = \sqrt{\frac{p}{\mu}} \frac{s^2}{2w} \sin(L) a_{c3}
\]
\[
\frac{dL}{dt} = \sqrt{\frac{p}{\mu}} \frac{w^2}{p^2} + \frac{1}{w} \sqrt{\frac{p}{\mu}} (h \sin L - k \cos L) a_{c3}
\]

where

\[ s^2 = 1 + h^2 + k^2 \]
\[ w = 1 + f \cos L + g \sin L \]
\[ r = \frac{p}{w} \]

IV.B. Unified State Model

The unified state model uses seven state variables to describe the state of an orbiting body. Three to define the shape of the orbit in the velocity phase space and four quaternions to define the orientation with respect to the inertial reference frame.

\( C, R_{f1} \) and \( R_{f2} \) are functions of the radial and angular momentum that determine the energy of the satellite and the size and position of the orbit in velocity phase space (the velocity hodograph), see the right graph of Fig. 3. \( R_{f1} \) and \( R_{f2} \) are the components of \( R \) in a convenient intermediate reference frame that is rotated about the line of nodes.²⁸

\( R \) and \( C \) are velocity components, and have been defined in Fig. 3. \( R \) has a constant direction, and points in the direction of the velocity at periapsis. \( C \) is always perpendicular to the radius vector. The satellite’s velocity, \( \mathbf{v} \), is the sum of both

\[ \mathbf{v} = \mathbf{R} + \mathbf{C} \]  

(23)

For an unperturbed two-body problem, the magnitude of both \( \mathbf{R} \) and \( \mathbf{C} \) is constant. Since the orbit in velocity phase space is constant for an unperturbed two-body problem, these variables are also constant for the unperturbed case.
Where the MEE use the shape and size of an orbit (ellipse) in the position phase space, the USM uses \( C \) and \( R \) to define the orbit (circle) in the velocity phase space.

The USM uses a quaternion to keep track of the orientation of the orbital plane. \( q_4, q_1, q_2, q_3 \) are the four quaternions

\[
q = q_4 + q_1 i + q_2 j + q_3 k
\]

and are also called Euler parameters. Since four parameters are used to define three degrees of freedom, the following constraint on the quaternions needs to be introduced

\[
q_4^2 + q_1^2 + q_2^2 + q_3^2 = 1
\]

Quaternions have two major advantages over Euler angles. No singular conditions occur in a rotation, the degrees of freedom are never lost i.e., there is no gimbal-lock, and all angular functions are algebraic instead of trigonometric.

The unified state variables \( C, R_{f1} \) and \( R_{f2} \) are constant for the unperturbed two-body problem. If perturbing forces act on the body, the equations of motion in terms of the perturbing force components in the \( e_1, e_2 \) and \( e_3 \) direction were first derived by Altman\(^{27}\) and later corrected by Chodas\(^{32}\) and Vittaldev\(^{28}\)

\[
\frac{d}{dt} \begin{bmatrix} C \\ R_{f1} \\ R_{f2} \end{bmatrix} = \begin{bmatrix} 0 & -p & 0 \\ \cos \lambda & -(1 + p) \sin \lambda & -\gamma R_{f2}/v_{e2} \\ \sin \lambda & (1 + p) \cos \lambda & \gamma R_{f1}/v_{e2} \end{bmatrix} \begin{bmatrix} a_{e1} \\ a_{e2} \\ a_{e3} \end{bmatrix}
\]

\[
\frac{d}{dt} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & \omega_3 & 0 & \omega_1 \\ -\omega_3 & 0 & \omega_1 & 0 \\ 0 & -\omega_1 & 0 & \omega_3 \\ -\omega_1 & 0 & -\omega_3 & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix}
\]

where

\[
\begin{bmatrix} \sin \lambda \\ \cos \lambda \end{bmatrix} = \frac{1}{(q_3^2 + q_4^2)} \begin{bmatrix} 2q_3 q_4 \\ q_4^2 - q_3^2 \end{bmatrix}
\]

which becomes singular for \( q_3 = q_4 = 0 \). This corresponds with retrograde equatorial orbits \( (i = 180^\circ) \). This means that the USM breaks down for retrograde equatorial orbits. This orbit can be avoided by reformulating it with a negative orbital velocity and zero inclination.
The body-fixed velocity components \(v_{e1}\) and \(v_{e2}\) are given by

\[
\begin{bmatrix}
v_{e1} \\
v_{e2}
\end{bmatrix} = \begin{bmatrix} 0 \\ C \end{bmatrix} + \begin{bmatrix}
\cos \lambda & \sin \lambda \\
-\sin \lambda & \cos \lambda
\end{bmatrix} \begin{bmatrix}
R_{f1} \\
R_{f2}
\end{bmatrix}
\]

(27)

The parameter \(\gamma\) in Eq. (24) is used for compactness, and is given by

\[
\gamma = \frac{q_3 q_4 - q_2 q_1}{q_3^2 + q_4^2}
\]

(28)

This term also becomes singular for \(q_3 = q_4 = 0\), which are retrograde equatorial orbits as explained above.

The variable \(p\) is the ratio between \(C\) and the velocity perpendicular to the radius vector

\[
p = \frac{C}{v_{e2}}
\]

(29)

and \(\omega_i\) are the body-fixed angular velocity components

\[
\omega_1 = \frac{a_{e3}}{v_{e2}}, \quad \omega_2 = 0, \quad \omega_3 = \frac{C v_{e2}^2}{\mu}
\]

(30)

The angular velocity, \(\omega_2\), about the \(e_2\) axis is zero, because there is no velocity component out of the local instantaneous orbital plane.

V. Applications and Results

This section presents the results of verified interval integration applied to satellite orbit propagation with interval initial values and parameters (forces). The equations of motion of a satellite from Section IV are integrated using verified integration discussed in Section III to create guaranteed satellite orbits. These are essentially 4-dimensional (space and time) trajectory enclosures.

First, we present the performance of verified orbit propagation using different verified integration methods, orbital state-propagation models and interval initial value widths. Second, verified interval integration is applied to detect a collision between two satellites.

V.A. Propagation

We first compare two modern verified integration software tools, VNODE-LP\textsuperscript{10} and VSPODE,\textsuperscript{9} for tightness of the solution enclosures and computational cost. Second, we compare the three different orbital state models, Cartesian, MEE and USM, when used in verified orbit propagation, to see how they influence solution tightness, and how the orbit type and initial interval widths influence solution enclosures.

The orbit propagations in this section use an interval initial state and two interval parameters. They also include three perturbation forces, \(i.e.,\) atmospheric drag, \(J_2\) gravity perturbation and an unknown force modeled as interval. The uncertainty in the atmospheric drag is modeled as an interval with upper and lower bounds of 50% above and below the nominal value. The third, unknown force is an interval with a width of \(2 \times 10^{-8}\) km s\(^{-2}\). This models the uncertainty in all other perturbing forces.

A third, state-of-the-art verified TM-based integrator, COSY Infinity,\textsuperscript{33} could not be used due to pending patent claims.

VNODE-LP and VSPODE

As described previously, there are verified integration methods based on interval or Taylor-model evaluation of a Taylor series expansion of the solution. In simple example problems, and applications from literature,\textsuperscript{5} the Taylor-model based integrators create tighter solution enclosures. This comes with the disadvantage of 10 to 100 higher computational cost.

The most modern ITS-based verified integration library is VNODE-LP.\textsuperscript{10,34} We will compare VNODE-LP with the TM based verified integrator VSPODE.\textsuperscript{9,35} VSPODE was originally created to simulate bioreactors.\textsuperscript{36}
Both literature and simple example simulations suggest that VSPODE creates tighter enclosures than VNODE-LP. Although this advantage comes with a high computational cost, the tightness of the solution is more important, because solution overestimation is the main problem when applying verified integration and interval analysis to real-world problems.

The volume of the solution enclosure as a function of time for a circular, 45°-inclined low Earth orbit (LEO) at 800 km, computed using the different state models, is shown in Fig. 4. VSPODE produces smaller volumes and is able to propagate about twice as long for all tested cases. This confirms that the Taylor-model based integrator VSPODE produces better solutions, albeit at the cost of more CPU time. Anticipating the discussion about the quality of the results obtained with the different orbit-state models later on in this section, it is obvious from Fig. 4 that there is a significant difference in quality of the results. Both the MEE and the USM seem to perform much better than the Cartesian model.

Interval Solution Enclosure Verification

Fig. 5 shows the interval position and Monte-Carlo fixed-point solution set computed using MEE, for a LEO, after 10 hours. The green cuboid represents the interval enclosure, valid both in space and time. The Monte-Carlo solution set lies completely within the interval enclosure. This is evidence for the correct implementation of MEE transformations, equations of motion and VSPODE integrator.

This figure also shows the overestimation that occurs when enclosing a non-convex three dimensional shape in 3 dimensional interval described by only 6 coordinates. The Monte Carlo solution set can be considered to be close to the true solution shape, although it does contain numerical integration errors. All other positions inside the green interval box are non-solutions, i.e., extra information we do not need and want. Although the true solution set will still be inside the interval, practical application of verified integration becomes useless when there are many non-solutions inside the solution enclosure, e.g., the knowledge that a satellite is within a huge cuboid enclosing the Earth and all LEO space around it, is not useful in any application.

Orbital State Models

The advantage of using perturbation state models like MEE and USM over Cartesian coordinates lies in the almost constant or linearly changing variables of the perturbation models. The major gravity force of the central body is removed from the differential equations and therefore has no contribution to overestimation. However, the differential equations of the MEE and USM are more complex and may therefore suffer more from the dependency problem, and are also computationally more expensive.

To compare the effect of the different state models, we have to start with comparable initial intervals, representing a similar uncertainty in position and velocity. Table 1 and 2 show the initial interval widths

![Figure 4. Comparison between verified integrators VNODE-LP and VSPODE.](image-url)
Figure 5. Left: The nominal orbit (blue) with the solution enclosure (small box) after 10 hour, valid over 5 seconds, computed using the MEE. Right: A close-up of the interval solution box (green) shown on the left, now with Monte Carlo solution space plotted as red dots and the nominal orbit in blue.

Table 1. Interval Monte Carlo coordinate transformation from Cartesian coordinates to MEE. The first two columns give the width of the Cartesian velocity and position components. The last six columns give the resulting widths of the MEE.

<table>
<thead>
<tr>
<th>Cartesian</th>
<th>MEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w([x]) ) [m]</td>
<td>( w([\dot{x}]) ) [m/s]</td>
</tr>
<tr>
<td>1.0</td>
<td>0.01</td>
</tr>
<tr>
<td>10.0</td>
<td>0.10</td>
</tr>
<tr>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>1000.0</td>
<td>10.00</td>
</tr>
</tbody>
</table>

Table 2. Interval Monte Carlo coordinate transformation from Cartesian coordinates to USM variables. The first two columns give the width of the Cartesian velocity and position components. The last six columns give the resulting widths of the USM variables.

<table>
<thead>
<tr>
<th>Cartesian</th>
<th>USM</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w([x]) ) [m]</td>
<td>( w([\dot{x}]) ) [m/s]</td>
</tr>
<tr>
<td>1.0</td>
<td>0.01</td>
</tr>
<tr>
<td>10.0</td>
<td>0.10</td>
</tr>
<tr>
<td>100.0</td>
<td>1.00</td>
</tr>
<tr>
<td>1000.0</td>
<td>10.00</td>
</tr>
</tbody>
</table>
(uncertainties) of MEE and USM variables corresponding to the interval widths of Cartesian coordinates shown in the first two columns. These are computed using 100,000 Monte Carlo transformations by randomly picking points from within a Cartesian interval and converting these to MEE or USM variables using fixed-point transformations. The corresponding transformation equations can be found in the works by Vittaldev et al. 28 for USM and Betts et al. for MEE. 30 Monte-Carlo transformations are used, because interval analysis suffers too much from the dependency problem, which makes a fair comparison between the state models impossible. Taylor model based transformations can probably solve this problem in future research.

A visualization of an interval trajectory is shown in Fig. 6. The cuboids enclose the solutions over a time interval of 2 seconds. Only half of all intervals are shown in this figure to prevent cluttering, but when we would show all cuboids, they create a tube of overlapping cuboids enclosing the trajectory over the full integration time interval, including all points within the time interval and thus not only at single moments in time.

What happens when solution enclosures grow without bounds is shown in Fig. 7. In this example case, the solution grows exponentially and becomes too large for the integration to continue, the solution explodes. This is the result of accumulation and propagated overestimation of individual integration steps, and shows the largest problem of using verified integration.

The evolution of the volume of the position enclosure, valid over a 2-second time interval, is shown for different initial uncertainties in position and velocity, and different orbital-state models in Fig. 8. The example orbit is a circular 45° inclined LEO satellite at 400 km.

This figure clearly shows that both the MEE and USM are able to integrate for about 5 times longer than the Cartesian model before the solution explodes and becomes useless.

As expected, a smaller uncertainty in the initial values results in smaller position enclosure volumes and a longer time to solution explosion. The difference between the smallest two initial uncertainties is very small for the USM, this shows that a change in initial interval has no significant effect on both overestimation and true solution set size when the initial intervals are sufficiently small, i.e., 10 m and 0.1 m/s for the USM. This insignificance of small changes in small initial values indicates that the only other uncertainties, the interval perturbation forces, are the main contributor to the growth of the position intervals. The effect of the size of initial intervals and interval perturbations is investigated next.

The relation between the uncertainty in initial interval position and velocity widths, and the time to solution explosion is shown in Fig. 9. The time of solution explosion is defined as the moment when the interval position volume is larger than $1 \times 10^9$ km$^3$. This corresponds to cube with sides of 1000 km, although the actual enclosure can be any parallelepiped. A larger interval initial value width results in a smaller time-to-solution explosion for all three state models. Reducing the interval widths even further than 10 m and 0.1 m/s has little effect on the time-to-solution explosion as can be seen from the lines in the graph having a smaller slope.

The MEE and USM are able to integrate about 4 times longer than the Cartesian model. USM and MEE
Figure 7. Visualization of interval solution explosion for a LEO satellite propagated using Cartesian coordinates.

Figure 8. Time evolution of the volume of the position enclosure, for a circular LEO at 400 km, using the four different orbit models, for four different sizes of initial position and velocity widths (the initial uncertainty). Enclosures are valid over 2 seconds.
are about equally good, although the MEE is slightly better for initial interval widths smaller than 400 m and 0.4 m/s, and the USM for larger widths. However, this result depends on the orbit type and interval perturbations and cannot be generalized to all orbits with all interval perturbation force widths. An unexpected peak can be seen in the Cartesian line in Fig. 9. This peak is unexpected as smaller intervals always lay within the larger intervals and should therefore create smaller or equally sized solution enclosures. The cause of this peak could not be determined conclusively and is most likely the result of the a-priori stepsize selection within VSPODE. The relation between the uncertainty in the perturbation force, and the time to solution explosion, is shown in Fig. 10. This figure shows that the USM is able to integrate for a longer time than MEE for small initial intervals, while MEE is better for shorter intervals. However, the time-to-solution explosion is not the only indication of quality interval orbit propagation, the volume of the position enclosure can also be important. Example applications are verified collision detection, where smaller sizes create more guaranteed safe trajectories, and global optimization, where the solution volume influences how much of the search space can be discarded. Both MEE and USM have singularities for inclinations of 180° (retrograde equatorial orbits). Furthermore, the inclination determines the \( J_2 \) perturbation force and the relation between inclination and time to solution explosion. Fig. 11 shows the relation between the inclination and solution explosion. It is clear that the inclination has a large influence on both the MEE and USM, but especially on the MEE model. This makes the MEE better than the USM for equatorial orbits (zero inclination), while it is worse for inclined orbits with an inclination between 60° and 120°. This suggests that a change in size or direction of the \( J_2 \) perturbation force has a larger influence in the MEE model. Using the right model or combination of models for a specific inclination can therefore improve the interval integration significantly. The USM is only affected by the singularity when the initial inclination interval is chosen within half a degree from 180° while the MEE model shows an abrupt halving of the time to solution explosion in the inclination range from 177° to 183°. This shows that the USM is the more reliable and preferred state model for inclinations around 180°.

V.B. Collision Detection

Verified collision-detection methods try to guarantee that no collisions will occur, based on assumptions on bounds of measurement uncertainties. This in contrast to conventional collision detection methods, which
Figure 10. Time to solution explosion as function of the width of the uncertain interval force acting on the satellite, for a circular LEO at 800 km with a 28.6 deg inclination.

Figure 11. Time to solution explosion as function of inclination for a circular LEO.
estimate a collision probability. Verified methods either guarantee no collision (a collision probability of 0%), or a possible collision with unknown probability. Future applications may combine verified integration with probability-based methods, such that a verified collision probability distribution can be given.

Verified methods use verified orbit integration to create a guaranteed satellite-trajectory enclosure in the form of adjacent 3-dimensional boxes. Each box encloses the satellite position during a certain time interval. A collision cannot be ruled out when the position enclosures of two satellites intersect during the same time interval. In other words, if the satellite trajectories intersect both in space and time, their paths may intersect. It should be emphasized that due to overestimation verified methods only guarantee a safe trajectory for satellites, i.e., a collision-free trajectory having interval trajectories that do not intersect.

Verified collision-detection methods include any possible integration errors, uncertain initial conditions, uncertain constants and uncertain perturbation forces. Furthermore, collisions are detected over a continuous time interval, and not based on the approximated time of closest approach that conventional methods use.

The major disadvantages of current verified methods are the lack of probability distribution propagation, the overestimation of the position enclosures introducing non-solutions, and its computational cost. Probability information could theoretically be propagated using TM methods although they need major modifications. Overestimation can be reduced by decreasing the integration stepsize and thereby the time interval over which solutions are valid, by using different state models, using higher integration order, subdivision and applying methods as discussed in Section III.B.

**Interval Intersections**

The intersection of two, 3-dimensional intervals is the set of points that are within both intervals. This intersection is a new 3-dimensional interval. 3-dimensional interval vectors can intersect in five different ways, as shown in Fig. 12.

Finding the intersection of two boxes requires only the position of the corners of both boxes. Here we use the knowledge that both boxes have the same attitude because they are both defined in the same orthogonal (Cartesian) coordinates. The position of the corners is given by the 3-dimensional intervals of both boxes. A computer can quickly find interval intersections of n-dimensional intervals.
For collisions between objects, we have to find the intersection of 4-dimensional intervals, three space components and one time, because objects only collide when they are in the same place at the same time.

**Hybrid Interval Trajectories**

Interval intersections allow us to combine the best of multiple state models and integrators by creating a new interval trajectory from the intersection of all individual trajectories. For example, if we have two interval trajectories computed using the USM and MEE, only the points that are inside both trajectory enclosures can be part of the true solutions. This means that the intersection of both trajectories is a new verified interval trajectory, which we will call a hybrid trajectory.

This hybrid trajectory will have a smaller volume than the individual trajectories if, and only if, one of the trajectories is not completely included in the other. All positions that are not inside both trajectories are discarded. The evolution of the position volume for a circular LEO at 800 km is shown in Fig. 13. The red line in the graph represents the volume of the hybrid trajectory. This volume is, as expected, always equal or smaller than the individual USM and MEE trajectory volumes. During some time periods, e.g., between 0.5 and 1 hour, the intersection of the two interval trajectories is significantly smaller than the individual trajectories.

The hybrid method can cause a large or no reduction in the volume (and everything in between), depending on the satellite orbit, initial interval widths and propagation time. Finding the intersection of two interval trajectories comes with a low computational cost compared to the interval integration and is therefore always advised when multiple interval trajectories are available.

**Iridium-33/Cosmos-2251 Collision**

Few major accidental collisions have been detected since the launch of the first artificial satellite, Sputnik, in 1957. One between two intact satellites, two between an operational satellite and large space debris, and one between two rocket bodies. Collisions between inactive satellites are hard to detect and can only be reconstructed from observed debris. Smaller collisions occur more often, but are also more difficult to detect.

The largest and most notable collision in Earth orbit was the collision between two intact satellites, the operational communication satellite, Iridium 33, and the inactive navigation satellite, Cosmos 2251, that occurred at 16:56 UTC on February 10, 2009, at 789 km above the Taymyr Peninsula in Siberia. Both satellites were completely destroyed. The incident was observed by the U.S. Space Surveillance Network, above Siberia at an altitude of 790 km, and later tracked two large clouds of debris.\(^{37}\)

Precise state information of both satellites is not openly available. The only available information are the TLE from Space-Track.\(^{38}\) These provide no accuracy information and accuracy is often in the km range. A
realistic verified collision detection simulation using real measurements uncertainties is therefore not possible.

Nevertheless, this case is a good test to see if this collision can be detected, i.e., not ruled out, by verified collision detection. The initial interval widths used are 10 m in all position components and 0.1 m/s in the velocity components. Drag and $J_2$ are the perturbations that are explicitly modelled. The uncertainty in the density is enclosed in an interval with a width of $2 \times 10^{-6}$ kg/m$^3$ and an uncertainty in the perturbation acceleration is modelled by an interval with a width of $2 \times 10^{-9}$ km/s$^2$.

The nominal initial positions as computed from the TLE are input to SPG4 to propagate them to 00:00 UTC on February 10, 2009, i.e., 16 hour and 56 minutes before the actual collision. These nominal Cartesian states are provided in Table 3.

The corresponding nominal orbits of both satellites are shown in Fig. 14 where the orbit of Iridium 33 is shown in green and Cosmos 2251 in blue. The intersection of the two orbits on the Northern hemisphere is the location of the actual collision.

Both orbits are propagated using verified interval integration. VSPODE is used with a TM-order of 3, ITS-order of 12 and initial stepsize of 10 seconds. Two possible collisions are found per revolution for the initial integration with a stepsize of 10 seconds. One at the location where the actual collision occurred and one at the close approach area of both orbits on the Southern hemisphere.

When the stepsize is reduced to 1.25 seconds, only possible collisions remain around the location of the actual collision and all close approaches on the Southern hemisphere are determined safe. The first possible

![Figure 14. Nominal orbits of Iridium-33 and Cosmos-2251 and their projections on the three planes.](image)

Table 3. Cosmos-33 and Iridium-2251 Cartesian state in an Earth-centered inertial reference frame at 00:00 UTC on February 10, 2009, 16 hour and 56 minutes before the collision between the two satellites.

<table>
<thead>
<tr>
<th>Satellite</th>
<th>x [km]</th>
<th>y [km]</th>
<th>z [km]</th>
<th>$v_x$ [km/s]</th>
<th>$v_y$ [km/s]</th>
<th>$v_z$ [km/s]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cosmos 2251</td>
<td>2689.06</td>
<td>2726.88</td>
<td>6049.935</td>
<td>-6.53545</td>
<td>-1.16479</td>
<td>3.40109</td>
</tr>
<tr>
<td>Iridium 33</td>
<td>-3240.99</td>
<td>4759.60</td>
<td>4246.546</td>
<td>2.01107</td>
<td>-3.99140</td>
<td>5.98213</td>
</tr>
</tbody>
</table>
collision is detected at 15:15, the second between 16:55:48.75 and 16:55:51.25. This second collision is the exact time of the real collision. Reducing the stepsize to 0.3125 seconds removes the possible collision at 15:15. The real collision can still not be ruled out, as we expect since the satellites did collide. Further reduction of the stepsize, changing state model and using a hybrid trajectory cannot rule out the real collision.

The interval trajectories and the detected collision interval computed using a stepsize of 5 seconds are shown in Fig. 15. The blue and green line indicate the nominal trajectory of Cosmos 2251 and Iridium 33 respectively. The time dimension of the interval trajectory is indicated by the opacity of the intervals where a darker interval indicates a later moment in time. Projections of the interval trajectory and the red collision interval are shown on all three planes. This shows that a real collision can be successfully detected by verified collision detection.

The interval trajectories, even for the hybrid-trajectory with time intervals (stepsizes) of a fraction of a second, still have interval sides of a few km, indicating large overestimation. It is therefore likely that although this collision would have been correctly detected, collisions would also have been detected for cases where they did not actually collide.

Socrates\textsuperscript{39} predicted a closest approach of 117 m between the two satellites.\textsuperscript{40} This is far smaller than the interval enclosures of a few km computed using verified integration. Approaches within 5 km occur about a thousand times per week for one of the 66 active Iridium satellites with other tracked satellites. Interval enclosure size thus has to decrease for practical applications.

VI. Conclusions and Recommendations

Verified interval integration is successfully applied to satellite orbit propagation to create guaranteed solution enclosures on satellite trajectories, valid over a time interval and taking uncertainties in initial state and perturbation forces into account by enclosing them in interval. Also bounding truncation (numerical integration) and rounding errors are included. Modern verified integration methods VNODE-LP and VSPODE are applied in combination with orbital state models based on integration constants to try to create tight solution enclosures and reduce overestimation, the largest problem of verified interval integration.
Overestimation in verified interval integration of satellite orbits can be greatly reduced by using the Taylor model based software implementation, VPSODE, combined with perturbation orbital state models like the modified equinoctial elements or the unified state model. VPSODE is able to propagate orbits for at least twice as long as VNODE-LP.

Earth-orbiting satellites can be verified propagated for 15-20 orbits before the interval trajectory grows too large and solution enclosures become useless, depending on the size of interval initial values and parameters (forces). The example cases include an uncertainty of 10 m in all position components and 0.1 m/s in the velocity components, an uncertainty in the atmospheric density, \( J_2 \)-gravity term, and an uncertain perturbation acceleration of \( 1 \times 10^{-8} \) km/s\(^2\). Although 15-20 orbits is still not much for many applications, previous methods using interval Taylor series and Cartesian coordinates were only able to integrate for 2-4 orbits before breaking down.

The MEE produce the smallest solution enclosures for most orbit types, about 4-10 times smaller than the Cartesian model and twice as small as the USM. However, the USM produces more consistent results for different orbit types. A hybrid method produced by combining interval trajectories computed using MEE and USM combines the best characteristics of both methods.

Verified interval trajectories can be used to detect collisions between satellites, as has been demonstrated by the reconstruction and successful detection of the collision between the Iridium-33 and Cosmos-2251 satellite in 2009. It can guarantee that no collisions will occur when certain bounds on measurement uncertainties are assumed.

Further research is necessary to determine the performance of a verified collision detection system that detects collisions between a large number of satellites, including debris. Interval analysis has been proven to be too crude for orbit propagation and Taylor model based computations should ideally be used in every step of the verified integration and collision detection.

To make verified collision detection practically possible in the long term, it may be necessary to combine it with conventional probability-based methods to be able to provide a verified collision probability. This makes it possible to prioritize possible collision avoidance maneuvers.

References