Mathematische Grundlagen der dynamischen Simulation

Dynamische Systeme sind Systeme, die sich verändern. Es geht dabei um eine zeitliche Entwicklung und wie immer in der Informatik betrachten wir dabei nicht die Entwicklung einer realen Welt, sondern eines Abbildes dieser realen Welt in einem Modell.

Beschreibung der Welt durch Größen

Das Modell der realen Welt wird dabei beschrieben durch Größen und ihre Wechselwirkungen. Die Gesamtheit aller Größen beschreibt dann mit ihren Werten einen Zustand des Systems und wenn es dynamisch ist, verändern sich diese Größen zeitlich.

Veränderung von Größen

Prinzipiell gibt es an Veränderungen nur die beiden Möglichkeiten Zunahme und Abnahme[1]. Interessant wird es daher nur bei der Frage, durch welche Wechselwirkung zwischen Größen welche Art von Veränderungen auftreten und eben auch, wie sich diese Veränderungen selbst zeitlich entwickeln.

Analytische Verfahren

Sehen wir aber davon zunächst ab, dann gibt es eine Größe und eine Änderungsrate dieser Größe. Dies ist uns mathematisch aus der Analysis bekannt und manche Systeme lassen sich analytisch beschreiben und lösen.
Man hat dann

- eine Funktion, welche die Werte der Größe beschreibt
- eine andere Funktion , die Ableitung, welche die Werte der Änderung der Größe beschreibt
- eine weitere Funktion , die zweite Ableitung, welche die Werte der Änderung der Änderung der Größe beschreibt
- usw.

Das System wird so durch ein System von Gleichungen beschrieben, das diese Größen enthält, ein Differentialgleichungssystem.
Für einige Differentialgleichungssysteme gibt es Lösungen:
Die Gleichung $f'(x) = k \cdot f(x)$ mit $f(0)=c$ [2] lässt sich beispielsweise durch eine Funktion mit der Funktionsgleichung $f(x) = c \cdot e^{k \cdot x}$ lösen.

Numerische Verfahren

Bei unserem Verfahren der dynamischen Simulation geht es nicht um diese Fälle. Wir überlassen diese allein den Mathematikern. Wir betrachten (prinzipiell) Fälle, die sich auf analytischem Wege nicht lösen lassen, obwohl es manchmal natürlich interessant sein kann, die entwickelten Methoden gerade auf die analytisch lösbaren anzuwenden, weil man die Güte seiner Simulationsverfahren daran testen kann.

Näherungslösungen

Wir betrachten Fälle, bei denen wir zu Näherungen einer richtigen Lösung dadurch kommen, dass wir iterativ ausgehend von Startwerten unter Berücksichtigung der

1 Wenn wir einmal von solchem Schwachsinn wie „Nullwachstum“ absehen, was ja nichts anderes als „keine Veränderung“ bedeutet.
2 Der Zusatz f(0)=c definiert einen Anfangswert der Größe f.

Änderungsraten mit diskreten Zeitschritten zu nachfolgenden Werten weiter rechnen.

Ein triviales Beispiel

Startwert Zeitschritt Änderungsrate

Jemand hat auf seinem Konto 35.000,-€ und zahlt jeden Monat 400,-€ ein.
Hieraus eine Tabelle der Werte zu bestimmen, fällt sicher jedem leicht. Ein passendes Werkzeug dazu wäre eine Tabellenkalkulation.

diskret

Ein besonderes Merkmal des eben betrachteten Beispiels ist, dass es ein diskretes Problem ist. Damit ist nicht unsere Diskretion über den Umfang des Kapitals zu schweigen gemeint, sondern die Tatsache, dass eine Unterteilung des Monates in kleinere Abschnitte nicht sinnvoll ist, da die Änderung zu genau einem Zeitpunkt mit genau einem Wert erfolgt.
Ganz anders sieht es aus, wenn die betrachteten Größen selbst aber kontinuierlich (oder quasi-kontinuierlich) sind, wie im folgenden Beispiel:
Jemand befindet sich auf der Startlinie einer Laufbahn, startet und läuft bis zur Ziellinie. Auch hier gibt es einen Startwert und eine Änderung, was man sehr leicht daran erkennen kann, dass er ja irgendwann am Ziel ankommt. Seine Position ändert sich aber nicht in kleinen Positionen, da er während der Bewegung an jedem Ort gewesen sein muss. Es gibt daher keine vorgegebenen Zeitschritte und selbst wenn die Bewegung gleichmäßig wäre, hinge die Größe der Änderung von der Länge des Zeitschrittes ab. Trotzdem geht man genau so vor wie vorher:

Wieder ein Beispiel:

Die ***Größe*** ist die Strecke s mit dem Startwert 0, die ***Änderung*** der Strecke ist Δs[1], der ***Zeitschritt*** Δt. Bei einem Δs von 0,1 zum Simulationszeitschritt Δt hat unsere Strecke s dann den Wert 0,1. Bleibt die Änderung konstant, kann man leicht die folgenden Werte direkt errechnen.

Die Änderung ändert sich

Schwieriger wird es, wenn die Änderung z.B. vom aktuellen Wert abhängt. Bei der Abkühlung einer Flüssigkeit beispielsweise hängt die Abkühlung, also die Änderung der Temperatur, vom Unterschied zwischen eben dieser Temperatur und der Umgebungs-temperatur ab.
Nehmen wir einmal an, die Ausgangstemperatur beträgt 80°C, die Umgebungstemperatur sei 20°C und die Abnahme im betrachteten Zeitintervall sei das 0,05-fache der Temperaturdifferenz.
Dann erhalten wir folgende Rechnungen:

$$80-0{,}05\cdot(80-20) = 80-0{,}05\cdot 60 = 80-3 = 77$$
$$77-0{,}05\cdot(77-20)=77-0{,}05\cdot\underline{57}=77-\underline{2{,}85}=74{,}15$$

usw.

Da sich die Größe selbst geändert hat, ändern sich nun auch die Änderungen!

1 Unsere Simulationssoftware verwendet die Bezeichnung ds

Bedeutung der Schrittweite

An dieser Stelle macht sich nun bemerkbar, dass die Änderung nicht linear ist, bei der o.a. Berechnung aber so getan wird, als wäre sie das zumindest in dem betrachteten Intervall! Die Folge ist, dass die Abnahme – da wir sie jeweils an der vorigen Stelle bestimmen – geringfügig größer ist, als sie tatsächlich ist. Statt einer glatten Kurve arbeiten wir nämlich mit einer Kette von Tangentenstücken, was sich leicht zu größeren Fehlern verstärken kann.

Verbesserung durch Verkleinerung der Schrittweite?

Man kann den Fehler, den man dabei macht, zwar verkleinern, indem man die Schrittweite der Zeitintervalle verkleinert. Das hat aber wiederum zwei Nachteile:

1. die Zahl der durchzuführenden Berechnungen verdoppelt sich, wenn man die Schrittweite halbiert.
2. Bei jeder realen Berechnung[1] wird man Rundungsfehler machen müssen. Je mehr Berechnungen man macht, desto größer wird auch die mögliche Abweichung durch Rundungsfehler

Man wird also irgendeinen Kompromiss finden müssen.

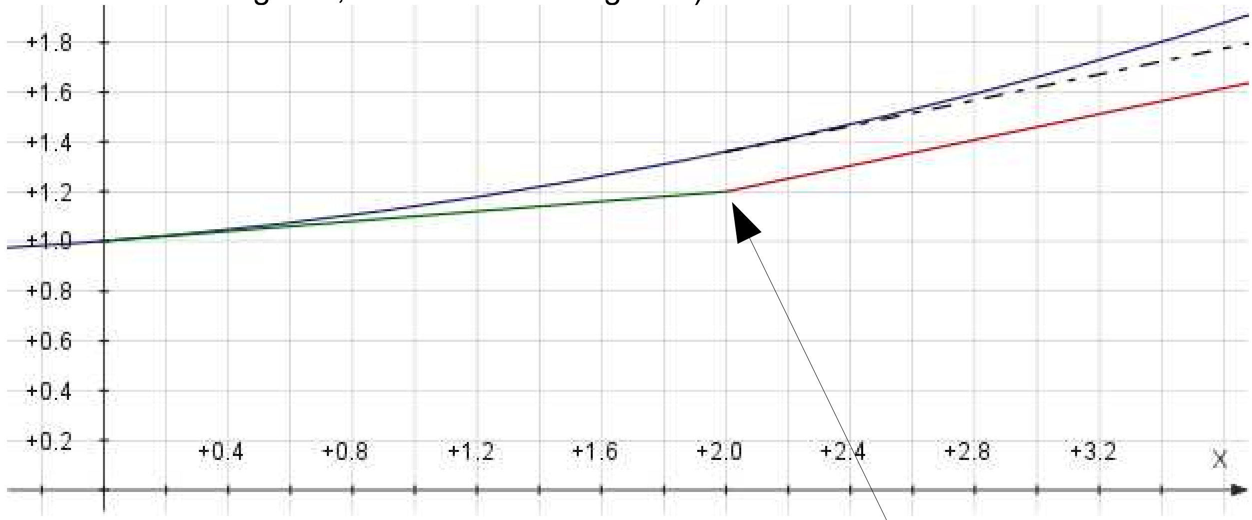
Die von Dynasys verwendeten Verfahren

Dynasys bietet zwei Rechenverfahren an:

Euler-Cauchy[2]:

$$\boldsymbol{B(t+\Delta t) = B(t) + m \cdot \Delta t}$$

(B ist die Bestandsgröße, m ist die Änderungsrate)



Vereinfacht gesagt, macht dieses Verfahren den Fehler „aus der Kurve zu rutschen“. Man trifft dadurch nicht den wirklichen Verlauf der Funktion, was dann bei den nachfolgenden Punkten auch die Folge hat, dass die Änderung auf der Basis von nicht ganz richtigen Werten der Bestandsgröße berechnet wird.

1 also dann, wenn die Werte nicht so glatt gewählt wurden wie im Beispiel
2 Zu prüfen wäre, ob Dynasys nicht das verbesserte Euler-Cauchy – Verfahren verwendet.

Man erkennt an diesem Beispiel sehr schön eines der grundlegenden Probleme der numerischen Lösungen von Anfangswertproblemen: Sie neigen dazu zu divergieren und sich immer mehr von den richtigen Werten zu entfernen.
Warum nimmt man denn dann nicht einfach den Wert der Änderung am Ende des Intervalls und berechnet aus diesem und dem zum Startwert die Steigung? Ganz einfach: Man hat den Wert ja nicht! Hätte man ihn, bräuchte man gar nicht zu rechnen.

Euler-Cauchy verbessern

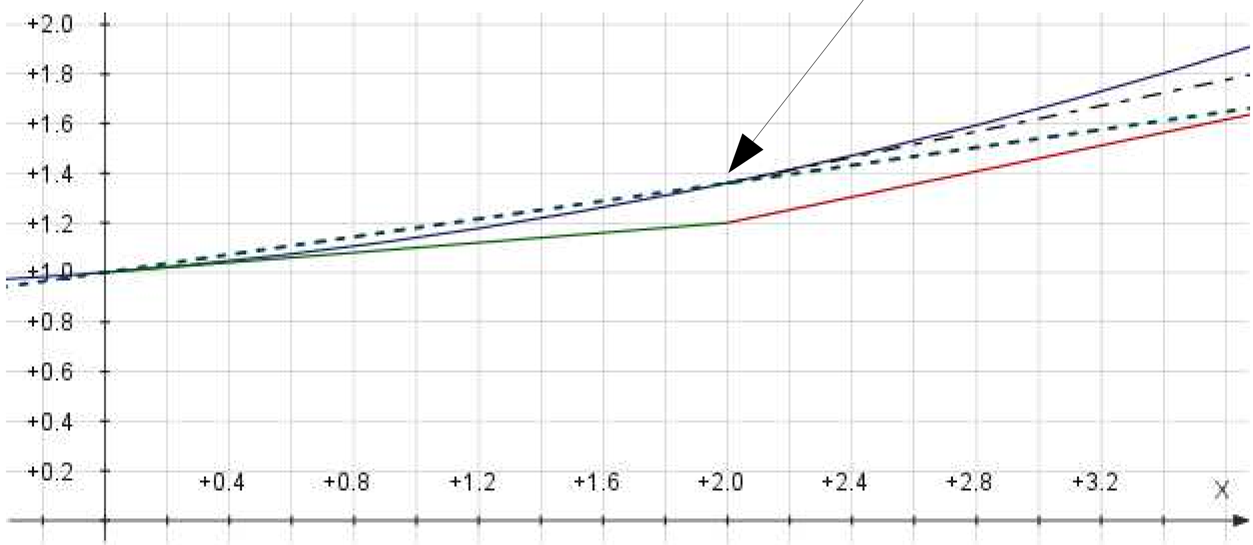
Man kann das Verfahren von Euler-Cauchy verbessern, indem man man den oben formulierten Gedanken sinnvoll aufnimmt und vorausschauend arbeitet. Dazu verwendet man zwar zunächst wiederum die Steigung am Anfang und damit den Folgewert am Ende des Intervalls. Dieser wird aber nun nicht als nächster Bestandswert verwendet, sondern nur benutzt, um die Änderung an diesem Punkt zu berechnen. Aus den beiden so berechneten Änderungsraten (Steigungen) berechnet man den Mittelwert. Man erhofft sich davon, eine bessere Information über das weitere Verhalten der Funktion nach der Anfangsstelle zu bekommen.

m_1 ist die Änderungsrate am Startwert des Intervalls.

m_2 ist die Änderungsrate, die man bekäme, wenn man mit m_1 bis zum Ende des Intervalls gehen würde. Die tatsächlich verwendete Änderungsrate ist der Mittelwert dieser beiden Änderungsraten.

$$m = \frac{m_1 + m_2}{2}$$

Das folgende Bild zeigt, weshalb man in der Regel bessere Ergebnisse erzielt. [Die gepunktete Linie trifft in dem gezeichneten Beispiel sogar exakt die Kurve, da ich eine quadratische Funktion für den Graphen verwendet habe.]



Runge-Kutta [vierter Ordnung]

Dieses Verfahren setzt den Gedanken der Verbesserung fort und arbeitet mit vier Stützstellen, an denen dann jeweils die Änderungsraten berechnet werden und mit ihrer Hilfe dann erst der Folgewert berechnet wird. Dies Verfahren ist daher sehr viel besser in

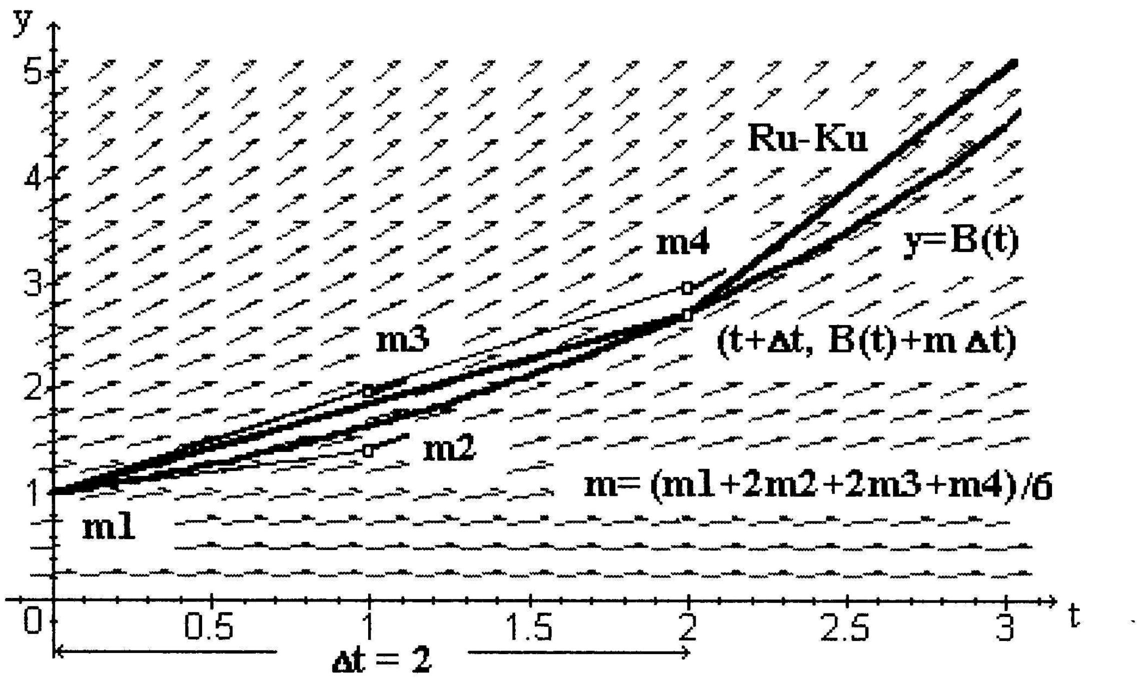
der Lage, die vorhandene Krümmung einer Kurve zu berücksichtigen.
In der nachfolgend angegebenen Gleichung werden vier Änderungsraten verwendet:

- m_1 ist die Änderungsrate am Startwert des Intervalls.
- m_2 ist die Änderungsrate, die man bekäme, wenn man mit m_1 bis zur Mitte des Intervalls gehen würde.
- m_3 ist die Änderungsrate, die man bekäme, wenn man mit m_2 bis zur Mitte des Intervalls gehen würde.
- m_4 ist die Änderungsrate, die man bekäme, wenn man mit m_3 bis zum Ende des Intervalls gehen würde.

Aus diesen vier Werten berechnet man nun durch einen gewichteten Mittelwert, bei dem die mittleren Werte stärker berücksichtigt werden, den tatsächlich verwendeten Wert der Änderungsrate

$$m = \frac{m_1 + 2 \cdot m_2 + 2 \cdot m_3 + m_4}{6}$$

und rechnet damit nun wie beim Euler-Cauchy – Verfahren den nächsten Wert der Bestandsgröße aus.
Das Bild aus dem Buch zur Simulation[1] zeigt den möglichen Vorteil:



Es gelingt auf diesem Wege die Änderung der Änderung stärker zu berücksichtigen und man kommt in der Regel zu besseren Simulationsergebnissen.

1 Bild aus „Simulation dynamischer Vorgänge“ von Klett. [Leider nicht mehr lieferbar.]