Numerical Integration - 1st draft

Thomas Grønneløv

April 15, 2008

Contents

List of Figures

1 Introduction

In this text I will go over a few subjects related to integration and its relevance for simulation. I try to focus on the areas where I had trouble when I originally learned about these things, and I will generally focus less on the things I found to be easy. The subjects I touch upon are...

- Very brief overview of what it means to integrate.
- Introduction to ordinary differential equations and their use.
- Example of using analytical integration to solve an initial value problem.
- Introduction to numerical integration and Taylor series as well as truncation error.
- A description of phase space and its use.
- Example of using Explicit Euler to solve the problem previously solved analytically.
- Example of using Midpoint to solve the same problem.
- Example of using Runge Kutta 4 to solve the same problem.
- Example of using Implicit Euler to solve the same problem.

Early on I will present a physical problem in form of an initial value problem in 1 which is then solved analytically by integration, with Taylor series and numerically using three different methods.

2 Integration recapped

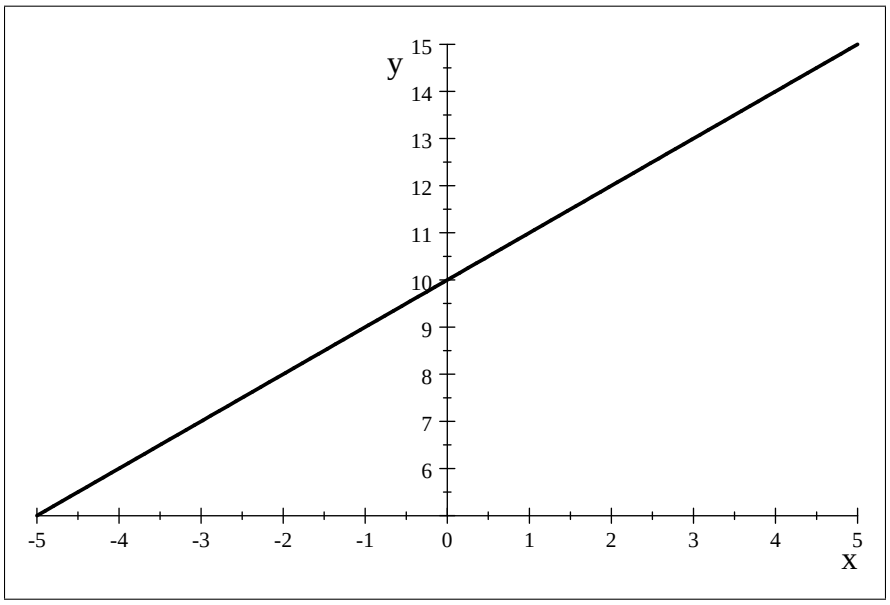
I would like to take a step back and look at what integration really means. I will however not get into how you perform integration by hand. I assume that the reader knows about this, I just want to point out what integration actually means. If this is too simplistic for you, just skip over this section.

What does it mean when you say that you integrate $f(t) = at$ and get $F(t) = \frac{1}{2}at^2$? It means that you have some function f of t which in this case returns a single real value and you would like to translate it into another function of t which returns the sum of all the function values multiplied by the distance between each sample $\sum_{t=-\infty}^{\infty} f(t)dt$. This means that you sample the function with some interval dt and multiply the function value with dt each time. When you make the interval small enough, you get the exact value of the integral. You

can visualize this by observing "the area under the function curve" for t in some interval. An example for a sinus function is shown in figure 4. In that figure we look at $0 \leqq t \leqq 4$. The sum of all $f(x)dx$ when dx approaches zero is in fact the area under the curve which is $F(4) - F(0)$.

The unit of the integral is the unit of the function multiplied by the unit of the variable over which we integrate. If you for example have velocity $v(t) = 10m/s$ and distance $s = \int v \cdot dt = vt$ then the unit is meters per second times seconds which is meters as one should expect for a distance. You essentially multiply all the function values with all the steps of the variable over which you integrate.

Another way to look at integration is that you multiply the average value of the function over some interval and then multiply that average with the length of the interval. Imagine the function $f(x) = x + 10$. The average value of the function between 0 and 5 is $12\frac{1}{2}$. Multiply $12\frac{1}{2}$ with 5 and you get 62.5. Integrating the function you get $F(x) = \frac{1}{2}x^2 + 10x$. The definite integral from 0 to 5 is then $F(5) - F(0) = (\frac{1}{2} \cdot 5^2 + 10 \cdot 5) - 0 = 62.5$.



3 Ordinary differential equations

An ordinary differential equation ODE presents a problem in which you are looking for an unknown which in itself is an equation. The differential equation can for example be stated as $\frac{dy}{dx} = -kx$ which only says something about how the function value should change when changing x, and you now need to find an equation which satisfies that. Finding that equation is called to solve the differential equation. It is a problem where the result is not a value but an equation!

This means that when you differentiate that solution you will end up with the differential equation again. One solution to $\frac{dy}{dx} = -kx$ could be $y = -\frac{1}{2}kx^2 + C$ which is found by simply integrating both sides of the equation. It is not a unique solution however since $y = -\frac{1}{2}kx^2 + 10$ and $y = -\frac{1}{2}kx^2 + 20$ are both solutions and are both equally valid. Figure 1 shows five such solutions for $k = 10$.

The point is that both solutions will, when differentiated, result in the original differential equation since the constant term C becomes zero.

An initial value problem IVP is formulated as a differential equation as well as some value which the resulting equation should take for some specific input value , or values. An example could be 1 which has only one solution which is found by solving as before by integration, substituting x with the initial value of x and then isolating the unknown constant C. That is $y(0) = -\frac{1}{2}k0^2 + C = 10$ $=> C = 10 => y(x) = -\frac{1}{2}kx^2 + 10$.

The problem with analytical solutions as the above is that many problems are far more complex than that and then there are no easy solutions. You may also want to make a computer program that solves differential equations in general and then you would have to program the analytical integration methods yourself, which is not that easy a task. Yet another problem could be a problem stated as $\frac{dy}{dx} = blackbox(x)$, where you have a black-box function of x but you do not have the function itself. You can call $blackbox(10)$ and get perhaps 17 back, but you do not know *how* $blackbox(x)$ works and therefore you can not integrate it.

The nice thing about analytical solutions, when you can get them, is that you get a function which tells what you want to know in constant time for any input variable. If you for example know how position develops over time, as we will derive in the next section, then you can find the position at any time wether it is a million years in the future or in a second from now, and it will be absolutely accurate, within the limits of the model you are using. After all, nature can never be described with absolute accuracy, so the differential equations are just describing some model. A numerical method would instead need to take a lot of small steps or fewer huge steps to get from now to then and every step will increase the error of the solution.

$$\begin{aligned} \frac{dy}{dx} &= -kx \\ y(0) &= 10 \\ y(x) &= -\frac{1}{2}k^2 + 10 \end{aligned} \tag{1}$$

3.1 Deriving an expression for movement

In this section I will try to show a physical problem which can be solved analytically by integrating differential equations.

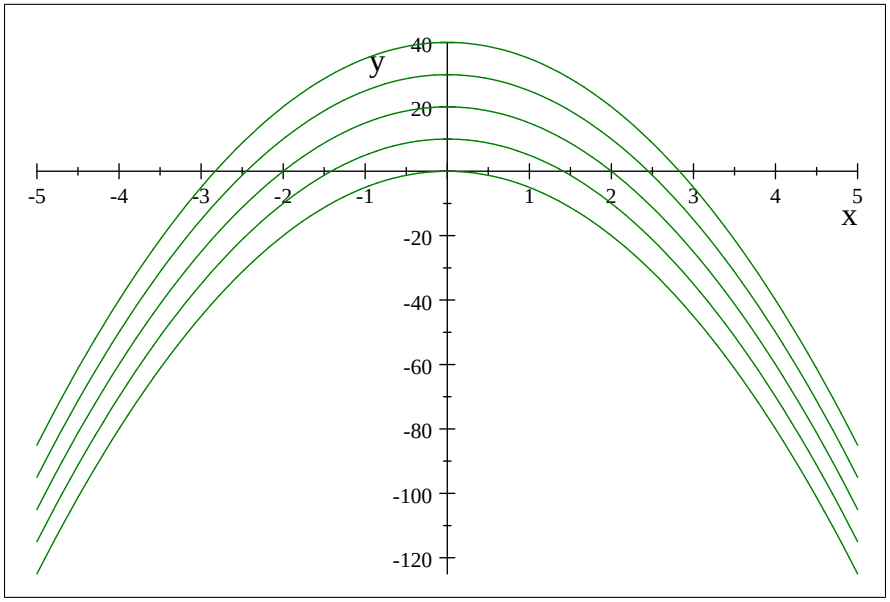


Figure 1: Five different solutions to the differential equation $\frac{dy}{dx} = -kx$

An object has position $s(t)$, velocity $v(t)$ and acceleration $a(t)$. All three elements are functions of time. At time $t = 0$ the object has a well defined position $s(0) = 5m$ and velocity $v(0) = 10m/s$. Acceleration is defined to be $a(t) = \sin t + 1$.

We now want to describe its state at any time based on our knowledge of the initial state and how that state changes over time. Note that you actually have differential equations here.

Acceleration is the rate at which velocity changes over time or $dv/dt = a$ and velocity is defined as change in position over time or $ds/dt = v$. The acceleration at any time is a defined function. If we integrate both sides of the differential equations we end up with $s = \int_0^t v\ dt$ and $v = \int_0^t a\ dt$. The position, velocity and acceleration for $t = 0$ were given, so it is an initial value problem which is stated in 2.

We solve it from the bottom up by first integrating $\frac{dv}{dt} = a$ over t on both sides, which gives us $v(t) = \int_0^t a\ dt = \int_0^t (\sin t + 1)dt = (-\cos t + 1t) - (-\cos 0 + 0t) + C = -\cos t + t + 1 + C$. C is actually the initial velocity $v(0)$, which we can see by setting $t = 0$ and evaluating $v(0) = -\cos 0 + 0 + 1 + C = 10m/s$. By isolating C we get $C = 10m/s$.

Now that we have $v(t)$ we can integrate that part to get $s(t)$ using the same method.

We have that $\frac{ds}{dt} = v(t)$ which we integrate on both sides to get $s(t) = \int_0^t v\ dt = \int_0^\tau -\cos t + t + 11 = (-\sin t + \frac{1}{2}t^2 + 11t) - (-\sin 0 + \frac{1}{2}0^2 + 11 \cdot 0) + C =$

$-\sin t + \frac{1}{2}t^2 + 11t + C$. As before C is just the initial value which we can find (though we already know it) by setting $t = 0$ and evaluating the new function $s(t)$. We then get $s(0) = -\sin 0 + \frac{1}{2}0^2 + 11 \cdot 0 + C = 5m$. Isolating C gives us $C = 5m$.

The conclusion is that by integrating the differential equations we went from a general description of how the system changed over time to useful functions which tells us exactly how the system is at any time. We now know that $s(t) = 5m + (11m/s)t - \sin t + \frac{1}{2}t^2$ and we know that $v(t) = 11m/s + t - \cos t$. We know that those functions are solutions to the problem since $s(0) = 5m$ and $v(0) = 10m/s$ and we can see that if we differentiate the functions we go back to the original problem. In figure 2 the three functions are plotted together. You can see that acceleration is a sinus going through a full period over the duration of 2π. Velocity is generally increasing, but when acceleration goes to zero, then velocity remains constant. Position is always growing since the velocity is always greater than zero.

$$\begin{aligned} \frac{ds}{dt} &= v \\ \frac{dv}{dt} &= a \\ a(t) &= \sin t + 1 \\ s(0) &= 5m \\ v(0) &= 10m/s \end{aligned} \tag{2}$$

The point of this exercise was to show why you would want to do integration. It is often easier to define how something changes than to define the result of the change. That is certainly the common way of describing physical systems. In the previous example you can easily see how velocity relates to acceleration and how position relates to velocity. It was however not as easy to see how you would go from acceleration directly to change in position. You may have noted that if the problem was somewhat more complex, it would have been a true pain to solve like I did.

4 Numerical integration

Often you can not readily obtain an analytical solution to a differential equation. This is when you must use numerical integration.

Numerical integration works by doing something quite simple a lot of times. The more you do it, the better the solution generally gets. It is easy to make a computer do something simple a lot of times, so this sounds like a nice method to use. Assuming you have $f(x) = \sin x + 1$ and you need to integrate it from 0 to 4, you can approach this in several ways. One method, which is known as integration by left boxes, or Reimann Left, is to define a number of

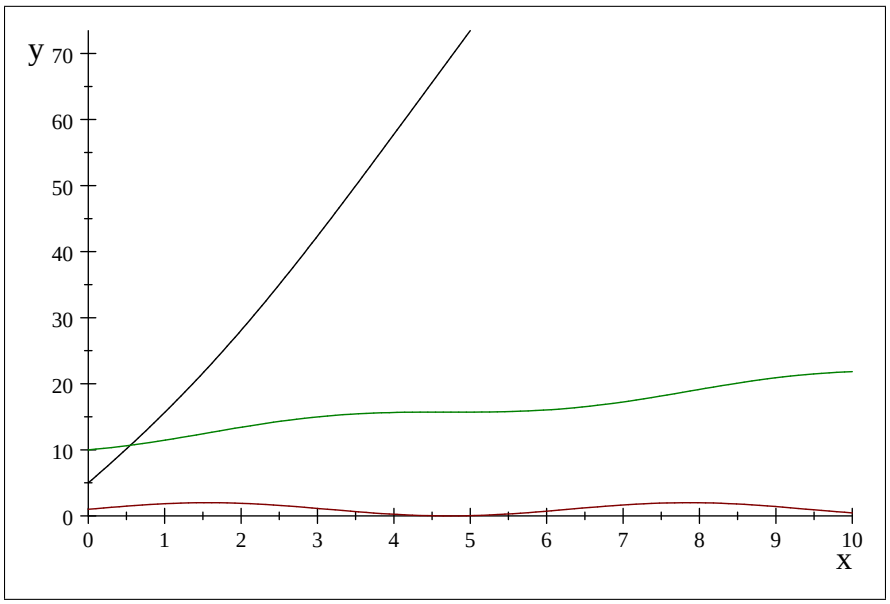


Figure 2: Plot of a(t), v(t) and s(t)

columns under the function which have a left height equal to the function value at the point and width equal to some dx defining the distance between function samplings. The integral is then the sum of the area of all the boxes. Reimann Left is shown in figure 4 and figure 5. The first is a coarse sampling and the later is a fine sampling with thin boxes.

The column building is what I call the simple part and the thinner the columns the more times you have to do that simple part. It is easy to see that a fine sampling means more work but also means that the result is more accurate. In essence this "boxing" works under the assumption that $f(x)$ is constant over dx. If it was, which is certainly is not in this example, then the boxing method would be absolutely accurate, but the fact of the matter is that when dx approaches zero then that assumption becomes true.

As stated previously the left boxes method is very simple and not that accurate. When I say it is not very accurate I mean that is takes a lot of steps to make it accurate. If we make dx small enough it is accurate, but we would like to make dx as large as possible and still have some decent accuracy. I will describe three methods for making it more accurate. They are explicit Euler, Midpoint and Runge Kutta4. Before going into details with them I need to touch on another subject which is Taylor series.

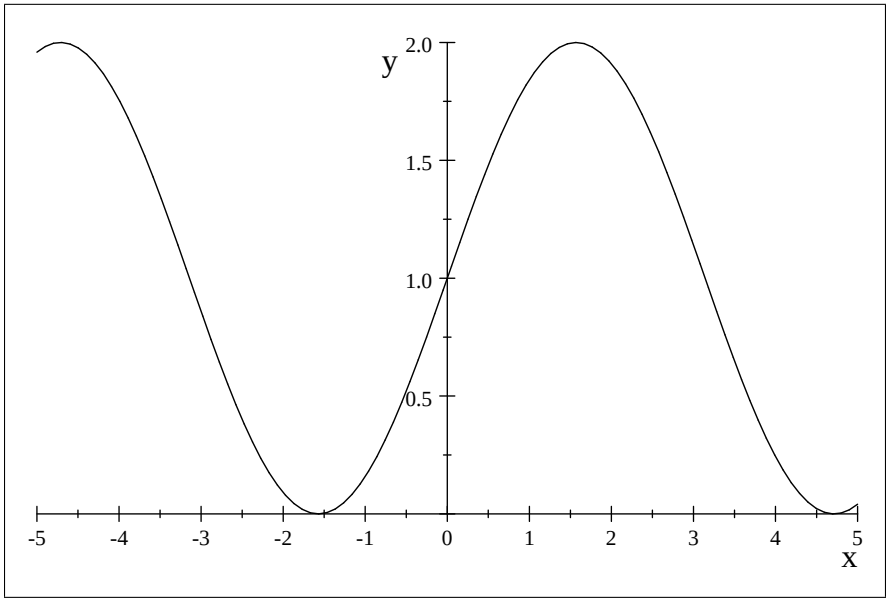


Figure 3: $f(x) = \sin x + 1$

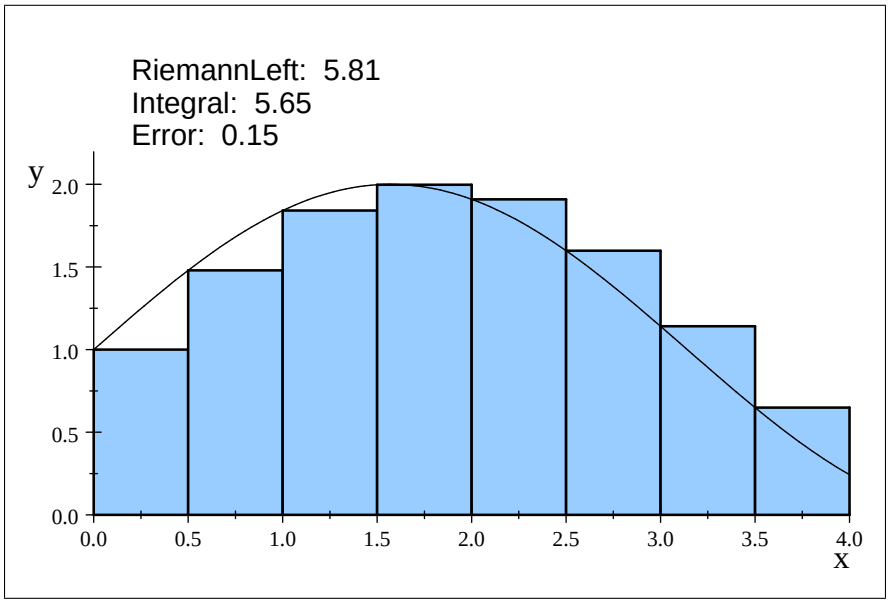


Figure 4: $\sin x + 1$ integrated by left boxes with $dx = \frac{1}{2}$

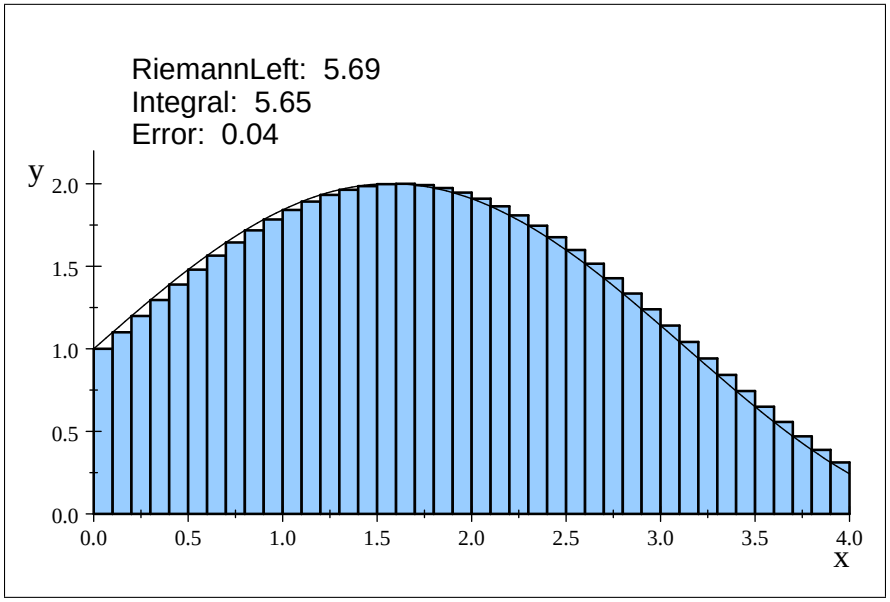


Figure 5: $\sin x + 1$ integrated by left boxes with $dx = 0.1$

4.1 Taylor series

If you have a function $f(t)$ then you can evaluate it for any t and get a value. If you do not have the function definition but only the *value* of the function and its derivatives at some point then you are seemingly stuck. As an example, if you know $f(t) = x^2$ then you can easily find $f(10) = 10^2 = 100$.The problem would be if you only know that $f(0) = 0$ and $f'(0) = 0$ and $f''(0) = 2$ and you still want to find $f(10)$. You can not do that right away.

Enter Taylor series expansion. This method lets you rediscover your function as a possibly infinite sum of its derivatives as seen in 3. If you rewrite the above problem to a Taylor series then you have 4 which is called the Taylor series expansion around the point 0, since all you information comes from that point. In this case the series can be easily simplified all the way back to the original hidden function. You may also note that since the third and following derivative of the function are 0 then the series is not infinite but only consists of three terms. If the Taylor series is complete, you do not neglect any terms, then its result is exact.

$$f(x+h) = f(x) + \frac{f'(x)h}{1!} + \frac{f''(x)h^2}{2!} + \frac{f''(x)h^3}{3!} + ... \quad (3)$$

Where h is the sample distance from the known position

$$
\begin{aligned}
f(0+h) &= f(0) + \frac{f'(0)h}{1!} + \frac{f''(0)h^2}{2!} + \frac{f''(0)h^3}{3!} + ... \qquad (4)\\
f(t) &= 0 + \frac{0t}{1!} + \frac{2t^2}{2!} + 0\\
f(t) &= t^2
\end{aligned}
$$

Truncation error If you do not evaluate *all* terms of the Taylor series then you do not get the exact function. This is the result of "truncation error" since you truncate the series. If for example $f(x) = \sin x + 1$ then the series is infinite. A Taylor series with six terms for this function will show the effect of truncation error. As can clearly be seen in figure 6, the truncated series 5 is exact at $x = 0$ which is the point it it evaluated around. This makes sense since the series at that point reduces to the function value in that point. As you get farther away from zero, the series begin to diverge since it really just is a polynomial of fifth degree which is unable to maintain the oscillation around $y = 1$ which $\sin x + 1$ can.

$$
\begin{aligned}
f(x) &= \sin x + 1 \qquad (5)\\
T(x) &= f(0) + \frac{\cos(0)x}{1} + \frac{-\sin(0)x^2}{2!} + \frac{-\cos(0)x^3}{3!} + \frac{\sin(0)x^4}{4!} + \frac{\cos(0)x^5}{5!} + O(t^6)
\end{aligned}
$$

A function with an infinite Taylor series was approximated with six terms and it resulted in a large error when you moved $2\frac{1}{2}$ away from the point around which the series was generated. The error is a function of the distance from the evaluation point. In the truncated series 5 we dropped the seventh term which was $\frac{-\sin(0)x^6}{6!}$ as well as all following terms. The seventh term has a size related to the value of x. If you double x^6 then you double the terms value and vice verse. This is means that the order of that term is $O(x^6)$. That term is the first and largest in an infinite chain and the order of such a chain is the largest order. This means that the order of the entire cut of chain in this case is $O(x^6)$. In fact the seventh term reduces to zero, so the largest one is the eight one, but that is besides the point right now.

If this notation is unfamiliar to you then you can look at

`http://en.wikipedia.org/wiki/Big_O_notation` to read the details or just think of it as meaning that there is a linear relationship between x^6 and the size of that term, which is a somewhat simplified view. To be more specific it actually means that there is a point (here an $x$value) after which kx^6 is forever thereafter greater than the error, for some constant k. The point is that you do not know exactly what the error is, based on $O(x^6)$ but you know how the error develops over x.

Let me show one last example of Taylor series. The initial value problem from 2 can be solved with such a series as well. You want to know $s(t)$ and you know only $s(0)$, $v(0)$ and $a(0)$. Since $s'(0) = v(0)$ and $s''(0) = a$, you have the

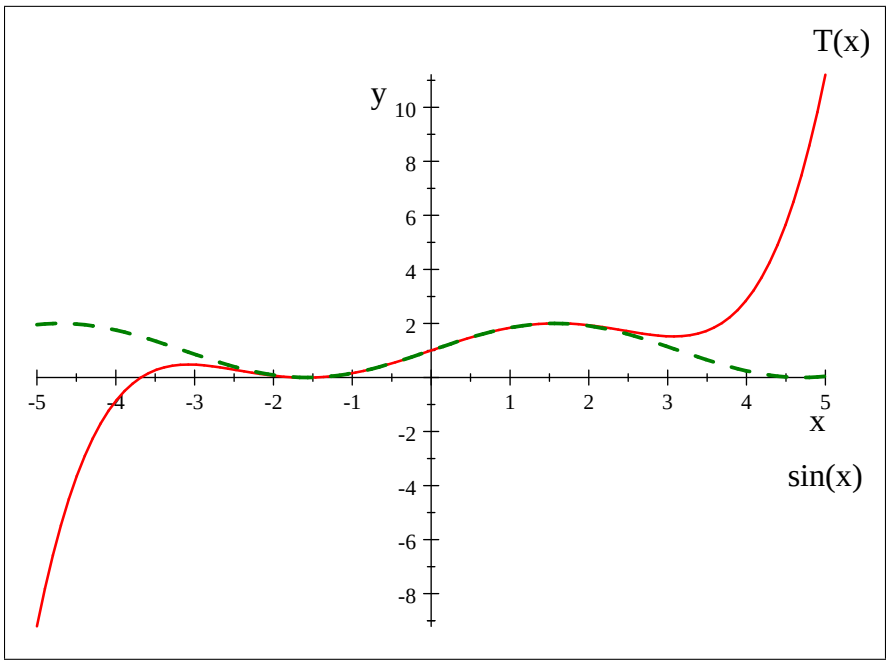


Figure 6: $\sin x$ and its six term taylor series T arround 0

basis for a series with three terms consisting of $s(0)$, $s'(0)$ and $s''(0)$. That is not a lot though. Such a series would assume that acceleration is constant and that it always has the value of $a(0) = 1$, which is not the case at all. Therefor we will add $s'''(t) = a'(t) = \cos t$ which at least throws *that* silly assumption (constant acceleration) out the window. This gives the series shown in 6. Note however that the series *is* truncated now since $a(t)$ has an infinite number of derivatives, and not just a single one, which are left out.

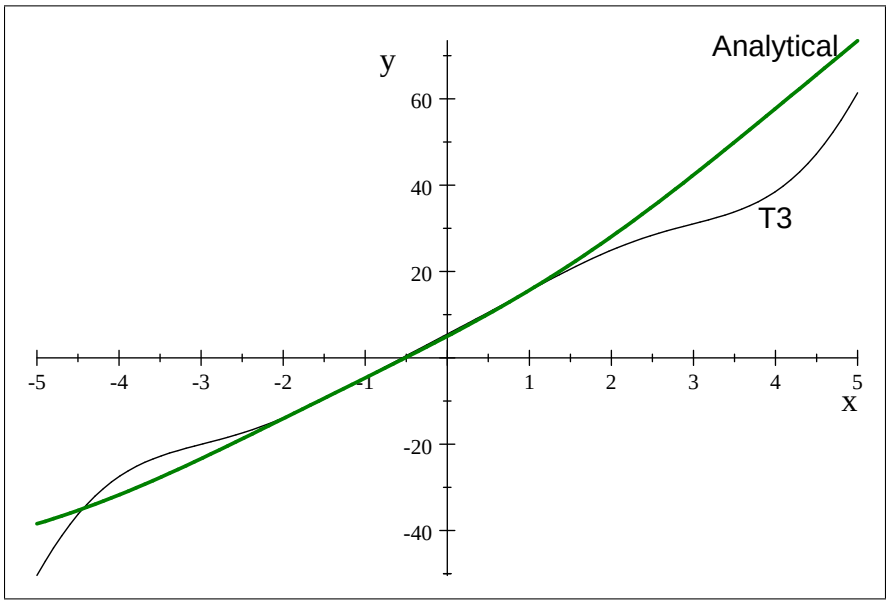
You clearly see that the Taylor series is fairly accurate near zero, but it breaks off around 1.

$$s(t) = s(0) + \frac{s'(0)t}{1} + \frac{s''(0)t^2}{2!} + \frac{s'''(0)t^3}{3!} + ... \tag{6}$$

$$s(t) = s(0) + v(0)t + \frac{a(0)t^2}{2} + \frac{(\cos t)t^3}{3!} + O(t^4)$$

$$s(t) = 5m + (10m/s)t + \frac{1}{2}m/s^2 + \frac{t^3}{6} + O(t^4) \tag{7}$$

Just to show the accuracy, or lack thereof, of the Taylor series, figure **??** shows both the analytical solution for $s(t)$ and the Taylor series approximation.



The point of going around truncation error and Taylor series was to explain why a large "step size" in numerical integration is bad. We have previously in section 4 seen how you can integrate a function by assuming that it is constant over some interval and that the smaller interval you use the better that method finds the correct result. Here we actually had the function definition and therefore did not really need to fool around like that, but in physics you rarely have

that much information. You generally know some variable and some of its derivatives at specific points and then you can use Taylor series to approximate the variable value at some other point. The farther away that other point is, the farther away from the true value your result might be. In the simple example in section 4 we used only the function value itself and none of its derivatives. That is a Taylor series with only one term, which only makes sense when the function value is a constant. A better method would seem to be to use the function value and at least one of its derivatives, which is what we will do later on.

5 Phase space

Phase space is a space in which everything about a system can be described by a point in that space. If you consider a system with a weight connected to a spring so the weight can move in 1D as shown in figure 7, then you see a system with two variables. They are one dimensional position and one dimensional velocity. The spring constant and the weights mass are constants. This means that any state of this system can be defined with just two numbers. If you now consider a 2D space (a plane) where we call one axis position and the other velocity, then you can see how any point in this space will corespond to two values - the points coordinates - which are position and velocity. This means that a system is now a point in 2D or you could consider it a 2D vector, which in essense is the same.

Now imagine that you plot the system at time $t = 0$ and then at $t = 1, 2, 3..n$ and you connectthe dots. Yoy will end up with a curve in phase space. Figure 8 shows such a curve for $t = 0..10$ plotted for a system as the one in figure 7 with spring constant $k = 1$ (and restlength=0), mass $m = 1$, initial position $x(0) = 1$, initial velocity $v(0) = 0$ and $\frac{1}{3}$ dampning each cycle. The little arrows define the phase space velocity for the points. For any system state, which is a point, there is a next state based on the timestep between states and the current position and velocity and acceleration. The horizontal axis is position and the vertical is velocity. You could consider a more simple system which is just a point moving along the x-axis with a constant velocity of 1. If your current state is $(0, 1)$ then the next state one second into the future *is* $(1, 1)$. The same is true for the spring system.

We see the curve starting at $(1, 0)$ which means that the position is 1 and the velocity is 0. Then position starts becomming smaller and moving towards 0 while velocity goes towards an increasingly larger negative velocity. After the position passes 0, the velocity starts becomming smaller and eventually it is 0 at which point the position starts growing again. After a full cycle, the weight is again as far to the "right" as it can come (this time one tird of last distance) and things start over. The dampning eventually causes the phase space position to become $(0, 0)$ meaning that it is positioned at 0 and not moving.

Actually the described systsem is not 2D but 3D since time is a variable as well. This is called the extended phase space. Had any of the systems

Figure 7: A weight connected to a spring with a fixed endpoint. The weight can only move in one dimension left/right.

derivatived depended on time, it would be absolulely nessesary to include time - they did not in this example. If they do not, it should still be included since it is more correct. It is however also more difficult to visualize. In the extended phase space a point is not just describing the system but the system at some time. Has the spring system from before not been damped then the curve would have been a circle. In that case, any point on the circle would describe n infinite number of identical systems. The point $(1, 0)$ would describe the systems with position 1 and velocity 0, but that would be the initial system and all other systems at $t = periodLength \cdot n$ for any integer n. The point is that the coordinate $(1, 0)$ is not one system but many systems. Had we included time, and defined the time axis to go from the paper/screen out towards us, then the curve could have been a spiral moving towards us. Any point on that spiral would be a unique system at one point in time.

That was a basic system which could be visualized easily. This is easily extended to a system of any complexity. If the system is for example a 2D simulation of two planets circling arround eachother due to gravitational pull, then the system will consist of 8 variables plus time. Those are position and velocity for both objects as shown in 8 as well as time. Phase space for this system is thererefore a 9 dimensional space. Any state of such a system can be defined with 9 variables which can be considered a point in this 9D space.

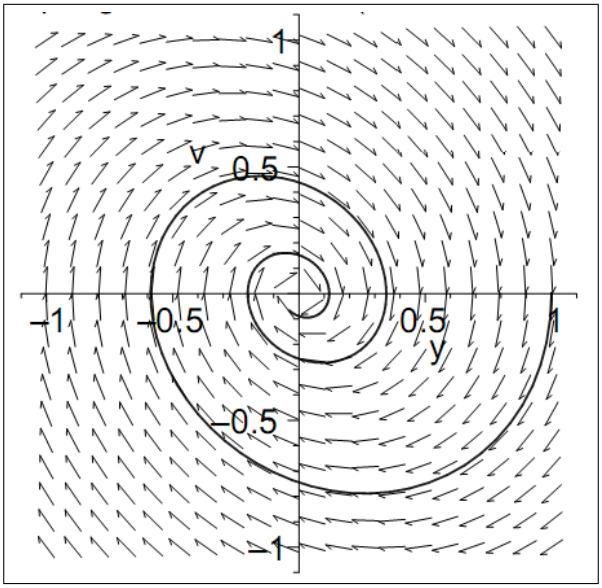


Figure 8: The direction field and a curve through phase space for a 1D weighted spring. The horizontal axis is position and the vertical is velocity.

$$
\begin{aligned}
&\vec{object_0}position_x \\
&\vec{object_0}position_y \\
&\vec{object_0}velocity_x \\
&\vec{object_0}velocity_y \\
&\vec{object_1}position_x \\
&\vec{object_1}position_y \\
&\vec{object_1}velocity_x \\
&\vec{object_1}velocity_y \\
&time
\end{aligned}
\tag{8}
$$

The point of all this is that the mathematical formulation of the numerical integration methods use notation which sees the systems as points/vectors in phase space. If we want to define a gradient function of the system in 8, we could write it as $f(x) : \mathbb{R}^9 \to \mathbb{R}^9$ where x would be a 9 dimensional vector. To put it another way; you have a function which takes a system and returns a gradient. When writing Taylor series you need the current value (the current state of the system) and its derivatives (the current systems gradient). I remember this as being something which confused me. Now I wonder why, because it is very simple. Describe your system from a phase space point of view as a single vector.

Create a function which takes such a system and returns a vector of equal size containing the derivatives of each vector element as seen in 9. I will try to make this more concrete when explaining the Explicit Euler integration method, but when moving on to Midpoint and RK4, I will assume that it has become clear.

$$\begin{aligned} &\vec{object_0}velocity_x \\ &\vec{object_0}velocity_y \\ &\vec{object_0}acceleration_x \\ &\vec{object_0}acceleration_y \\ &\vec{object_1}velocity_x \\ &\vec{object_1}velocity_y \\ &\vec{object_1}acceleration_x \\ &\vec{object_1}acceleration_y \\ &timestepsize\ dt \end{aligned} \tag{9}$$

6 Explicit Euler

Explicit Euler (10) is one of the simplest numerical integration methods. It is a good first step even though it is quite unstable. It is in essense a Taylor series with two terms. Only the first derivative is included. For systems with only one derivative, the method is accurate, but it is unlikely that you would want to use numerical integration on such systems. The methods local error is $O(h^2)$.

The idea is simple: you come from here to there by assuming that the first derivative is constant over the duration of the step.

To illustrate the method, we will again solve the initial value problem 2 which is repeated below. It is a system of coupled differential equations. There is one equation describing each system variable, here s and v. As explained section 5 we can describe the system as one vector containing all variables, here s, v and t, and we can describe the derivative of the system using one vector containing s', v' and t', which are just v, a and 1. The time derivative is 1 since it is always true that $\frac{dt}{dt} = 1$.

Lets define a function $f(x)$ which takes a vector3 describing the current system state, its phase space position, and returns a vector3 describing the derivative of the system. I will use C# syntax and assume that there exist a datatype called Vector3 which implements 3D vectors. Note that x is a *system state* - a vector - and not just a single number.

$$f(x+h) = f(x) + f\,'(x)h \tag{10}$$

$$\begin{aligned}
\frac{ds}{dt} &= v \\
\frac{dv}{dt} &= a \\
a(t) &= \sin t + 1 \\
s(0) &= 5m \\
v(0) &= 10m/s
\end{aligned}$$

```
Vector3 f(Vector3 stateVector)
{
  //this function takes a current state and returns its first derivatives
  Vector3 derivative = new Vector3();
  derivative[0] = stateVector[1];          //  ds/dt is v
  derivative[1] = sin(stateVector[2])+1;   //  dv/dt is a which is sin(t)+1
  derivative[2] = 1;                       //  dt/dt is always 1
  return derivative;
}
```

Given the derivative function it is trivial to integrate over time using explicit Euler. We have a current state which is the initial state. We now evaluate the derivative for that state and take a step in time size h to get the next state. This is repeated as much as needed - step by step.

```
Vector3 ExplicitEuler(Vector3 stateVector, double h)
{
  //this function takes a state x(t) and calculates the next state x(t+h)
  return stateVector + f(stateVector)*h;
}
```

7 Midpoint

The midpoint method has a local error of $O(h^3)$. It uses the derivative at $t + \frac{1}{2}h$ and not at t as the explicit Euler does. Right now I will just show its definition and how to use it. In a later version of this text I may explain how the method was originally obtained. Untill then, you can look at `http://en.wikipedia.org/wiki/Midpoint_method.`

Midpoint is defined as seen in (**??**).

The function $f(x)$ used in explicit euler will work fine for this method as well - and for RK4 for that matter, since it simply evaluates the derivatives of all variables at a given point in phase space. When using $f(x)$ as before, we have

the folowing code for implementing Midpoint. Note that $k1$ is also a 3D vector which is this case is the change in x over a timestep size h, or $f(x)$ integrated over h. This means that $x + \frac{1}{2}k1$ is the state of the system halfway between time t and $t + h$, or in other words, it is the midpoint.

$$\begin{aligned} k1 &= f(x_0)h \\ x_1 &= x_0 + f(x_0 + \frac{1}{2}k1) \end{aligned}$$

```
Vector3 Midpoint(Vector3 stateVector, double h)
{
  //this function takes a state x(t) and calculates the next state x(t+h)
  Vector3 k1 = f(stateVector)*h;
  return stateVector + f(stateVector+\U{bd}k1)*h;
}
```

8 RK4

Runge Kutta 4 is more seemingly more complex than Explicit Euler and Midpoint, but it is really just more of the same. Its local error is $O(h^4)$ which is yet another magnitude better than Explicit Euler, but it comes at a higher computational cost. Rather than evaluating $f(x)$ and $f(x+\frac{1}{2}h)$ this method evaluates the midpoint one more time and also uses the endpoint. Its deriviation is also not explained here (yet), but you can look at

`http://en.wikipedia.org/wiki/Runge-Kutta` if you want the finer details. The method is defined as seen in (11).

$$\begin{aligned} k1 &= f(x_0)h \\ k2 &= f(x_0 + \frac{1}{2}k1)h \\ k3 &= f(x_0 + \frac{1}{2}k2)h \\ k4 &= f(x_0 + k3)h \\ x_1 &= x_0 + \frac{k1 + 2(k2 + k3) + k4}{6} \end{aligned} \tag{11}$$

Again the function $f(x)$ can be used without changes, and the stepping is done as folowing.

```
Vector3 RK4(Vector3 stateVector, double h)
{
  //this function takes a state x(t) and calculates the next state x(t+h)
  Vector3 k1 = f(stateVector)*h;
  Vector3 k2 = f(stateVector+0.5*k1)*h;
```

```
  Vector3 k3 = f(stateVector+0.5*k2)*h;
  Vector3 k4 = f(stateVector+k3)*h;
  return stateVector + (k1+2*(k2+k3)+k4)/6;
}
```

9 Implicit Euler

TODO: well... do it.

10 Conclusion