# The Initial Value Problem For First Order ODE

Richard Palais

Math 32a

Brandeis, Fall, 2003

# Part I

# The General Theory of Initial Value Problems

# Predicting the Future

What we are going to study in this course can be looked at as a generalization of the following interesting and important question:

# Predicting the Future

What we are going to study in this course can be looked at as a generalization of the following interesting and important question:

Suppose that we know the wind velocity at every point of space and at each moment of time.

# Predicting the Future

What we are going to study in this course can be looked at as a generalization of the following interesting and important question:

Suppose that we know the wind velocity at every point of space and at each moment of time.

Suppose also that at a particular moment $t_0$ we see a tiny puff of smoke pass by, with its center located at a point $x_0$.

# Predicting the Future

What we are going to study in this course can be looked at as a generalization of the following interesting and important question:

Suppose that we know the wind velocity at every point of space and at each moment of time.

Suppose also that at a particular moment $t_0$ we see a tiny puff of smoke pass by, with its center located at a point $x_0$.

Can we then predict the position $x(t)$ of the smoke at times $t$ close to $t_0$?

# Mathematical Model Building

Translating this kind of vague question concerning the real world into a precisely stated question about rigorously defined mathematical objects is called building a mathematical model, or simply "model building".

# Mathematical Model Building

Translating this kind of vague question concerning the real world into a precisely stated question about rigorously defined mathematical objects is called building a mathematical model, or simply "model building".

We will see that in the present case it leads us to what is called an initial value problem (IVP) for a time-dependent ordinary differential equation (ODE).

# Mathematical Model Building

Translating this kind of vague question concerning the real world into a precisely stated question about rigorously defined mathematical objects is called building a mathematical model, or simply "model building".

We will see that in the present case it leads us to what is called an initial value problem (IVP) for a time-dependent ordinary differential equation (ODE).

Let's see how this works.

# Space and Time

In our model, we will represent "time" by $\mathbf{R}$.

# Space and Time

In our model, we will represent "time" by $\mathbf{R}$.

It is natural to represent "space" by $\mathbf{R}^3$, but nothing essential changes if we generalize to the case that space is $\mathbf{R}^k$, for some positive integer $k$, or even by an arbitrary finite dimensional real vector space $V$.

# Space and Time

In our model, we will represent "time" by $\mathbf{R}$.

It is natural to represent "space" by $\mathbf{R}^3$, but nothing essential changes if we generalize to the case that space is $\mathbf{R}^k$, for some positive integer $k$, or even by an arbitrary finite dimensional real vector space $V$.

Moreover—as we shall soon see—this extra generality proves **very** useful, so in what follows we identify space with a finite dimensional orthogonal vector space $V$, that you may think of $\mathbf{R}^3$ if that helps your intuition.

# Space and Time

In our model, we will represent "time" by $\mathbf{R}$.

It is natural to represent "space" by $\mathbf{R}^3$, but nothing essential changes if we generalize to the case that space is $\mathbf{R}^k$, for some positive integer $k$, or even by an arbitrary finite dimensional real vector space $V$.

Moreover—as we shall soon see—this extra generality proves **very** useful, so in what follows we identify space with a finite dimensional orthogonal vector space $V$, that you may think of $\mathbf{R}^3$ if that helps your intuition.

Later, we will also consider cases where $V$ is an infinite dimensional vector space and we will see that this will lead us to consider initial value problems for partial differential equations (PDE).)

## Wind is a Time-Dependent Vector Field

Now the wind velocity at a point $x$ of $V$ at time $t$ is a vector in $V$ that we will denote by $f(x, t)$.

## Wind is a Time-Dependent Vector Field

Now the wind velocity at a point $x$ of $V$ at time $t$ is a vector in $V$ that we will denote by $f(x,t)$.

A function $f : V \times \mathbf{R} \to V$ is called a *time-dependent vector field on* $V$, so saying that we know the wind velocity at all points of space and all instants of time just amounts to saying that we are given such a function $f$.

## Wind is a Time-Dependent Vector Field

Now the wind velocity at a point $x$ of $V$ at time $t$ is a vector in $V$ that we will denote by $f(x,t)$.

A function $f : V \times \mathbf{R} \to V$ is called a *time-dependent vector field on* $V$, so saying that we know the wind velocity at all points of space and all instants of time just amounts to saying that we are given such a function $f$.

(We will always assume that $f$ is at least continuous, but to prove theorems we will actually assume more than that.)

## The Path $x(t)$ of the "Puff of Smoke"

We select a smoke particle that at time $t_0$ (which we will call the *initial time*) is located at the center of the puff of smoke, and we identify the position $x(t)$ of this particle at time $t$ with the position of the puff at time $t$.

# The Path $x(t)$ of the "Puff of Smoke"

We select a smoke particle that at time $t_0$ (which we will call the *initial time*) is located at the center of the puff of smoke, and we identify the position $x(t)$ of this particle at time $t$ with the position of the puff at time $t$.

We will assume that the position $x(t)$ is defined for all times $t$ sufficiently close to $t_0$, so we can think of $t \mapsto x(t)$ as a function defined in some open interval $I \subseteq \mathbf{R}$ containing $t_0$ and having values in $V$.

# The Path $x(t)$ of the "Puff of Smoke"

We select a smoke particle that at time $t_0$ (which we will call the *initial time*) is located at the center of the puff of smoke, and we identify the position $x(t)$ of this particle at time $t$ with the position of the puff at time $t$.

We will assume that the position $x(t)$ is defined for all times $t$ sufficiently close to $t_0$, so we can think of $t \mapsto x(t)$ as a function defined in some open interval $I \subseteq \mathbf{R}$ containing $t_0$ and having values in $V$.

Note that by definition, $x(t_0) = x_0$, and we will call $x_0$ the *initial position*.

# Going With the Flow—the ODE

The characteristic property of a smoke particle is that it "goes with the flow", i.e., its velocity at any time $t$ is the same as the wind velocity at its position, $x(t)$, at time $t$.

# Going With the Flow—the ODE

The characteristic property of a smoke particle is that it "goes with the flow", i.e., its velocity at any time $t$ is the same as the wind velocity at its position, $x(t)$, at time $t$.

Now the velocity of the smoke particle (at time $t$) is just the tangent vector to its path $x(t)$ at this time—namely $x'(t)$.

# Going With the Flow—the ODE

The characteristic property of a smoke particle is that it "goes with the flow", i.e., its velocity at any time $t$ is the same as the wind velocity at its position, $x(t)$, at time $t$.

Now the velocity of the smoke particle (at time $t$) is just the tangent vector to its path $x(t)$ at this time—namely $x'(t)$.

This means that the path $x(t)$ satisfies the relation $x'(t) = f(x(t), t)$ !

# Going With the Flow—the ODE

The characteristic property of a smoke particle is that it "goes with the flow", i.e., its velocity at any time $t$ is the same as the wind velocity at its position, $x(t)$, at time $t$.

Now the velocity of the smoke particle (at time $t$) is just the tangent vector to its path $x(t)$ at this time—namely $x'(t)$.

This means that the path $x(t)$ satisfies the relation $x'(t) = f(x(t), t)$ !

Such a relation is called a *time-dependent ordinary differential equation*.

# Initial Value Problems

**Definition.** Let $V$ be a finite dimensional real vector space and let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field on $V$. Given an initial time $t_0 \in \mathbf{R}$ and an initial position $x_0$ in $V$, we associate a so-called Initial Value Problem (IVP)

# Initial Value Problems

**Definition.** Let $V$ be a finite dimensional real vector space and let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field on $V$. Given an initial time $t_0 \in \mathbf{R}$ and an initial position $x_0$ in $V$, we associate a so-called Initial Value Problem (IVP)

$\frac{dx}{dt} = f(x,t)$ (The ODE)

# Initial Value Problems

**Definition.** Let $V$ be a finite dimensional real vector space and let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field on $V$. Given an initial time $t_0 \in \mathbf{R}$ and an initial position $x_0$ in $V$, we associate a so-called Initial Value Problem (IVP)

$\frac{dx}{dt} = f(x,t)$ (The ODE)

$x(t_0) = x_0$ (The Initial Condition)

# Initial Value Problems

**Definition.** Let $V$ be a finite dimensional real vector space and let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field on $V$. Given an initial time $t_0 \in \mathbf{R}$ and an initial position $x_0$ in $V$, we associate a so-called Initial Value Problem (IVP)

$$\frac{dx}{dt} = f(x,t) \qquad \text{(The ODE)}$$

$$x(t_0) = x_0 \qquad \text{(The Initial Condition)}$$

and we define a solution to this IVP to be any differentiable path $x : I \to V$ defined in some interval $I$ containing $t_0$ and satisfying $x(t_0) = x_0$ and $x'(t) = f(x(t),t)$ for all $t \in I$.

# Initial Value Problems

**Definition.** Let $V$ be a finite dimensional real vector space and let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field on $V$. Given an initial time $t_0 \in \mathbf{R}$ and an initial position $x_0$ in $V$, we associate a so-called Initial Value Problem (IVP)

$$\frac{dx}{dt} = f(x,t) \qquad \text{(The ODE)}$$

$$x(t_0) = x_0 \qquad \text{(The Initial Condition)}$$

and we define a solution to this IVP to be any differentiable path $x : I \to V$ defined in some interval $I$ containing $t_0$ and satisfying $x(t_0) = x_0$ and $x'(t) = f(x(t),t)$ for all $t \in I$.

Our original problem can be paraphrased in terms of this model as: Can we always solve such an IVP?

# Questions Concerning the IVP

Here are a few of the interesting questions associated to the IVP that we will consider.

# Questions Concerning the IVP

Here are a few of the interesting questions associated to the IVP that we will consider.

• Given an interval $I$ containing $t_0$, when can we be sure there exists a solution of the IVP defined in $I$?

# Questions Concerning the IVP

Here are a few of the interesting questions associated to the IVP that we will consider.

- Given an interval $I$ containing $t_0$, when can we be sure there exists a solution of the IVP defined in $I$?

- If a solution does exist in $I$, is it necessarily unique?

# Questions Concerning the IVP

Here are a few of the interesting questions associated to the IVP that we will consider.

• Given an interval $I$ containing $t_0$, when can we be sure there exists a solution of the IVP defined in $I$?

• If a solution does exist in $I$, is it necessarily unique?

• What can we say about the dependence of a solution on the initial time and initial position.

# Questions Concerning the IVP

Here are a few of the interesting questions associated to the IVP that we will consider.

• Given an interval $I$ containing $t_0$, when can we be sure there exists a solution of the IVP defined in $I$?

• If a solution does exist in $I$, is it necessarily unique?

• What can we say about the dependence of a solution on the initial time and initial position.

• Assuming that a solution does exist, can we find efficient algorithms for actually computing it numerically with any desired accuracy?

# Questions Concerning the IVP

Here are a few of the interesting questions associated to the IVP that we will consider.

• Given an interval $I$ containing $t_0$, when can we be sure there exists a solution of the IVP defined in $I$?

• If a solution does exist in $I$, is it necessarily unique?

• What can we say about the dependence of a solution on the initial time and initial position.

• Assuming that a solution does exist, can we find efficient algorithms for actually computing it numerically with any desired accuracy?

• What are good methods for visualizing solutions in low dimensions.

Show examples using 3DXM:1D2ndOrder Forced Oscillator and UserDefined, higher dimensional examples.

# A Reformulation of the IVP

Let $J$ be a closed and bounded interval of real numbers containing $t_0$ and let $C(J, V)$ denote the space of all continuous maps of $J$ into $V$.

# A Reformulation of the IVP

Let $J$ be a closed and bounded interval of real numbers containing $t_0$ and let $C(J, V)$ denote the space of all continuous maps of $J$ into $V$.

We associate to a time-dependent vector field $f$ on $V$ and $x_0 \in V$ a mapping $F$ of $C(J, V)$ to itself as follows: if $x : J \to V$ is continuous, $F(x) : J \to V$ is defined by:

$$F(x)(t) := x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

# A Reformulation of the IVP

Let $J$ be a closed and bounded interval of real numbers containing $t_0$ and let $C(J, V)$ denote the space of all continuous maps of $J$ into $V$.

We associate to a time-dependent vector field $f$ on $V$ and $x_0 \in V$ a mapping $F$ of $C(J, V)$ to itself as follows: if $x : J \to V$ is continuous, $F(x) : J \to V$ is defined by:

$$F(x)(t) := x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

**Proposition.** *A mapping $x : J \to V$ solves the IVP $\frac{dx}{dt} = f(x, t)$, $x(t_0) = x_0$ if and only if $x$ is a fixed point of $F$, i.e., if and only if $x(t) = x_0 + \int_{t_0}^{t} f(x(s), s)\, ds$.*

# A Reformulation of the IVP

Let $J$ be a closed and bounded interval of real numbers containing $t_0$ and let $C(J, V)$ denote the space of all continuous maps of $J$ into $V$.

We associate to a time-dependent vector field $f$ on $V$ and $x_0 \in V$ a mapping $F$ of $C(J, V)$ to itself as follows: if $x : J \to V$ is continuous, $F(x) : J \to V$ is defined by:

$$F(x)(t) := x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

**Proposition.** *A mapping $x : J \to V$ solves the IVP $\frac{dx}{dt} = f(x, t)$, $x(t_0) = x_0$ if and only if $x$ is a fixed point of $F$, i.e., if and only if $x(t) = x_0 + \int_{t_0}^{t} f(x(s), s)\, ds$.*

Proof. Trivial.

# Successive Approximations

Suppose that $X$ is a metric space and $F$ is a continuous map of $X$ to itself. The "Method of Successive Approximations" is a technique for locating fixed points of $F$. It works as follows. Define $F^n : X \to X$ by composing $F$ with itself $n$ times. If $x$ is any element of $X$ we call the sequence $\{F^n(x)\}$ the sequence of successive approximations defined by $x$.

# Successive Approximations

Suppose that $X$ is a metric space and $F$ is a continuous map of $X$ to itself. The "Method of Successive Approximations" is a technique for locating fixed points of $F$. It works as follows. Define $F^n : X \to X$ by composing $F$ with itself $n$ times. If $x$ is any element of $X$ we call the sequence $\{F^n(x)\}$ the sequence of successive approximations defined by $x$.

**Proposition.** *If a sequence of successive approximations $\{F^n(x)\}$ converges to a limit $p$, then $p$ is a fixed point of $F$.*

# Successive Approximations

Suppose that $X$ is a metric space and $F$ is a continuous map of $X$ to itself. The "Method of Successive Approximations" is a technique for locating fixed points of $F$. It works as follows. Define $F^n : X \to X$ by composing $F$ with itself $n$ times. If $x$ is any element of $X$ we call the sequence $\{F^n(x)\}$ the sequence of successive approximations defined by $x$.

**Proposition.** *If a sequence of successive approximations $\{F^n(x)\}$ converges to a limit $p$, then $p$ is a fixed point of $F$.*

Proof. Since $F$ is continuous, $F(p) = F(\lim F^n(x)) = \lim F^{n+1}(x) = p$.

# Solving IVPs by Successive Approximations

As above, let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field, $J$ a closed, bounded interval, and define $F : C(J, V) \to C(J, V)$ by:

$$F(x)(t) := x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

# Solving IVPs by Successive Approximations

As above, let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field, $J$ a closed, bounded interval, and define $F : C(J, V) \to C(J, V)$ by:

$$F(x)(t) := x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

Let's try to use Sucessive Approximations to solve IVPs for a couple of special classes of vector fields, $f$.

# Solving IVPs by Successive Approximations

As above, let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field, $J$ a closed, bounded interval, and define $F : C(J, V) \to C(J, V)$ by:

$$F(x)(t) := x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

Let's try to use Sucessive Approximations to solve IVPs for a couple of special classes of vector fields, $f$. Perhaps the simplest kind of time-dependent vector fields is one that is constant in space, i.e., of the form $f(x, t) = \phi(t)$, In this case the mapping $F$ is clearly the constant map with value $\Phi \in C(J, V)$ defined by $\Phi(t) = x_0 + \int_{t_0}^{t} \phi(s)\, ds$, i.e., the antiderivative of $\phi$.

# Solving IVPs by Successive Approximations

As above, let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field, $J$ a closed, bounded interval, and define $F : C(J, V) \to C(J, V)$ by:

$$F(x)(t) := x_0 + \int_{t_0}^{t} f(x(s), s)\, ds.$$

Let's try to use Sucessive Approximations to solve IVPs for a couple of special classes of vector fields, $f$. Perhaps the simplest kind of time-dependent vector fields is one that is constant in space, i.e., of the form $f(x, t) = \phi(t)$, In this case the mapping $F$ is clearly the constant map with value $\Phi \in C(J, V)$ defined by $\Phi(t) = x_0 + \int_{t_0}^{t} \phi(s)\, ds$, i.e., the antiderivative of $\phi$.

This is clearly a fixed point of $F$ and also the solution of the IVP, and we see that Successive approximations works in this case.

# Linear ODE

Let us denote by $L(V)$ the vector space of linear maps of $V$ to itself. Any $T \in L(V)$, defines a time-independent vector field: $f(x,t) = Tx$.

# Linear ODE

Let us denote by $L(V)$ the vector space of linear maps of $V$ to itself. Any $T \in L(V)$, defines a time-independent vector field: $f(x,t) = Tx$.

Recall how the exponential of $T$ is defined using a power series, namely $\sum_{k=0}^{\infty} \frac{1}{k!} T^k$. The series is clearly absolutely convergent and so defines an element $\exp(T)$ in $L(V)$.

# Linear ODE

Let us denote by $L(V)$ the vector space of linear maps of $V$ to itself. Any $T \in L(V)$, defines a time-independent vector field: $f(x,t) = Tx$.

Recall how the exponential of $T$ is defined using a power series, namely $\sum_{k=0}^{\infty} \frac{1}{k!} T^k$. The series is clearly absolutely convergent and so defines an element $\exp(T)$ in $L(V)$.

By absolute convergence, we can differentiate the series for $\exp(tT)$ termwise, and we see that $\frac{d}{dt} \exp(tT) = T \exp(tT)$.

# Linear ODE

Let us denote by $L(V)$ the vector space of linear maps of $V$ to itself. Any $T \in L(V)$, defines a time-independent vector field: $f(x,t) = Tx$.

Recall how the exponential of $T$ is defined using a power series, namely $\sum_{k=0}^{\infty} \frac{1}{k!} T^k$. The series is clearly absolutely convergent and so defines an element $\exp(T)$ in $L(V)$.

By absolute convergence, we can differentiate the series for $\exp(tT)$ termwise, and we see that $\frac{d}{dt} \exp(tT) = T \exp(tT)$.

It follows that the solution of the IVP for the vector field $f$ with initial data $t_0, x_0$ is $\exp((t-t_0)T)x_0$.

# Linear ODE by Successive Approximation

Let's try to solve the linear ODE $\frac{dx}{dt} = Tx$ with the initial condition $x(t_0) = x_0$ by successive approximation, choosing as the initial approximation the costant curve $x_1(t) = x_0$ for all $t$ in $J$.

# Linear ODE by Successive Approximation

Let's try to solve the linear ODE $\frac{dx}{dt} = Tx$ with the initial condition $x(t_0) = x_0$ by successive approximation, choosing as the initial approximation the costant curve $x_1(t) = x_0$ for all $t$ in $J$.
The next approximation is:
$x_2(t) := x_0 + \int_{t_0}^{t} T(x_1(s))\, ds = (I + (t - t_0)T)(x_0)$,

# Linear ODE by Successive Approximation

Let's try to solve the linear ODE $\frac{dx}{dt} = Tx$ with the initial condition $x(t_0) = x_0$ by successive approximation, choosing as the initial approximation the costant curve $x_1(t) = x_0$ for all $t$ in $J$.
The next approximation is:

$$x_2(t) := x_0 + \int_{t_0}^{t} T(x_1(s))\, ds = (I + (t - t_0)T)(x_0),$$

and an easy induction shows that the $n+1$-st successive approximation is:

$$x_{n+1}(t) := x_0 + \int_{t_0}^{t} T(x_n(s))\, ds = \left(\sum_{k=0}^{n} \frac{(t-t_0)^k}{k!} T^k\right)(x_0)$$

# Linear ODE by Successive Approximation

Let's try to solve the linear ODE $\frac{dx}{dt} = Tx$ with the initial condition $x(t_0) = x_0$ by successive approximation, choosing as the initial approximation the costant curve $x_1(t) = x_0$ for all $t$ in $J$.
The next approximation is:
$x_2(t) := x_0 + \int_{t_0}^{t} T(x_1(s))\, ds = (I + (t - t_0)T)(x_0)$,
and an easy induction shows that the $n+1$-st successive approximation is:
$x_{n+1}(t) := x_0 + \int_{t_0}^{t} T(x_n(s))\, ds = (\sum_{k=0}^{n} \frac{(t-t_0)^k}{k!} T^k)(x_0)$.
Since this converges to $\exp((t - t_0)T)(x_0)$, the solution of the IVP, we see that the method of successive approximations solves the IVP for the case of linear ODE also.

## Locally Lipschitz Vector Fields

There is a simple answer to the existence question for an IVP. It turns out that just as long as the time-dependent vector field $f$ is continuous, a solution will always exist on some interval containing $t_0$.

## Locally Lipschitz Vector Fields

There is a simple answer to the existence question for an IVP. It turns out that just as long as the time-dependent vector field $f$ is continuous, a solution will always exist on some interval containing $t_0$.

However to get a satisfactory theory for the IVP, a somewhat more stringent condition than just continuity is required.

# Locally Lipschitz Vector Fields

There is a simple answer to the existence question for an IVP. It turns out that just as long as the time-dependent vector field $f$ is continuous, a solution will always exist on some interval containing $t_0$.

However to get a satisfactory theory for the IVP, a somewhat more stringent condition than just continuity is required.

**Definition.** A time-dependent vector field on $V$, $f : V \times \mathbf{R} \to V$ is called **locally Lipschitz** if for each $(x_0, t_0) \in V \times \mathbf{R}$ there exists a $K > 0$ such that $\|f(x_1, t) - f(x_2, t)\| < K \|x_1 - x_2\|$ for all $x_1$ and $x_2$ sufficiently close to $x_0$ and all $t$ sufficiently close to $t_0$.

# Existence and Uniqueness Theorem for Locally Lipschitz Vector Fields

**Theorem.** *Let $V$ be a finite dimensional orthogonal vector space and let $f : V \times \mathbf{R} \to V$ be a time-dependent locally Lipschitz vector field in $V$. Given any $T \in \mathbf{R}$ and $P \in V$ there exist positive real numbers $\delta$ and $\epsilon$ such that if $|T - t_0| < \delta$ and $\|P - x_0\| < \delta$ then the IVP:*

$\frac{dx}{dt} = f(x, t)$
$x(t_0) = x_0$
*has a unique solution $x_{x_0,t_0}(t)$ on the interval $I = (t_0 - \epsilon, t_0 + \epsilon)$. Moreover, this solution is contiuously diferrerentiable in $t$ and is Lipschitz in the initial data $x_0$ and $t_0$.*

We will sketch the proof below after a few remarks.

# $C^1$ Implies Locally Lipschitz

Let $V$ and $W$ be orthogonal vector spaces, $U$ a convex open set in $V$, $f : U \to W$ a $C^1$ map, and $Df_p$, the differential of $f$ at $p \in U$.

# $C^1$ Implies Locally Lipschitz

Let $V$ and $W$ be orthogonal vector spaces, $U$ a convex open set in $V$, $f : U \to W$ a $C^1$ map, and $Df_p$, the differential of $f$ at $p \in U$.

If $p, q \in U$ and $\sigma(t) = p + t(q - p)$ is the line joining them, the so-called "finite difference formula" says: $f(q) - f(p) = \int_0^1 Df_{\sigma(t)}(q-p)\, dt$, and it follows that:

$$\|f(q) - f(p)\| \leq \left(\int_0^1 \left\|Df_{\sigma(t)}\right\| \, dt\right) \|(q - p)\|,$$

# $C^1$ Implies Locally Lipschitz

Let $V$ and $W$ be orthogonal vector spaces, $U$ a convex open set in $V$, $f : U \to W$ a $C^1$ map, and $Df_p$, the differential of $f$ at $p \in U$.

If $p, q \in U$ and $\sigma(t) = p + t(q - p)$ is the line joining them, the so-called "finite difference formula" says: $f(q) - f(p) = \int_0^1 Df_{\sigma(t)}(q-p)\, dt$, and it follows that:

$$\|f(q) - f(p)\| \leq \left(\int_0^1 \left\|Df_{\sigma(t)}\right\| \, dt\right) \|(q - p)\|,$$

Consequently, the supremum of $\|Df_p\|$ for $p$ in $U$ is a Lipschitz constant for $f$. (In fact, the smallest one.)

# $C^1$ Implies Locally Lipschitz

Let $V$ and $W$ be orthogonal vector spaces, $U$ a convex open set in $V$, $f : U \to W$ a $C^1$ map, and $Df_p$, the differential of $f$ at $p \in U$.

If $p, q \in U$ and $\sigma(t) = p + t(q - p)$ is the line joining them, the so-called "finite difference formula" says: $f(q) - f(p) = \int_0^1 Df_{\sigma(t)}(q-p)\, dt$, and it follows that:

$$\|f(q) - f(p)\| \leq \left(\int_0^1 \left\|Df_{\sigma(t)}\right\| \, dt\right) \|(q - p)\|,$$

Consequently, the supremum of $\|Df_p\|$ for $p$ in $U$ is a Lipschitz constant for $f$. (In fact, the smallest one.)

In particular it follows that a $C^1$ time-dependent vector field is locally Lipschitz and so satisfies the Local Existence and Uniqueness Theorem.

# A Counter-Example

**Exercise.** Show that continuity of $V$ is **not** sufficient to guarantee uniqueness for an IVP.

# A Counter-Example

**Exercise.** Show that continuity of $V$ is **not** sufficient to guarantee uniqueness for an IVP.

Hint: the classic example (with $n = 1$) is the initial value problem $\frac{dx}{dt} = \sqrt{x}$, and $x(0) = 0$.

# A Counter-Example

**Exercise.** Show that continuity of $V$ is **not** sufficient to guarantee uniqueness for an IVP.

Hint: the classic example (with $n = 1$) is the initial value problem $\frac{dx}{dt} = \sqrt{x}$, and $x(0) = 0$.

Show that for each $T > 0$, we get a distinct solution $x_{_T}(t)$ of this IVP by defining $x_{_T}(t) = 0$ for $t < T$ and $x_{_T}(t) = \frac{1}{4}(t - T)^2$ for $t \geq T$.

# Contracting Maps

A mapping $F$ of a metric space $X$ to itself is called a contracting map (or a contraction) if it satisfies a Lipschitz condition with constant $K < 1$.

# Contracting Maps

A mapping $F$ of a metric space $X$ to itself is called a contracting map (or a contraction) if it satisfies a Lipschitz condition with constant $K < 1$.

**Fundamental Contraction Inequality.**
*If $F : X \to X$ is a contraction, and if $K < 1$ is a Lipschitz constant for $F$, then for all $x_1$ and $x_2$ in $X$,*

$$\rho(x_1, x_2) \leq \frac{1}{1-K}\Big(\rho(x_1, F(x_1)) + \rho(x_2, F(x_2))\Big).$$

Proof. Exercise.

# Contracting Maps

A mapping $F$ of a metric space $X$ to itself is called a contracting map (or a contraction) if it satisfies a Lipschitz condition with constant $K < 1$.

**Fundamental Contraction Inequality.**
*If $F : X \to X$ is a contraction, and if $K < 1$ is a Lipschitz constant for $F$, then for all $x_1$ and $x_2$ in $X$,*

$$\rho(x_1, x_2) \leq \frac{1}{1-K}\Big(\rho(x_1, F(x_1)) + \rho(x_2, F(x_2))\Big).$$

Proof. Exercise.

**Corollary.** *A contraction mapping can have at most one fixed point.*

Proof. Assuming that $x_1$ and $x_2$ are fixed points we deduce immediately that $\rho(x_1.x_2)$ must be zero.

# The Banach Contraction Principle

Let $F : X \to X$ and $F^n$ its $n$-fold composition with itself. If $F$ satisfies a Lipschitz condition with constant $K$, by an easy induction $F^n$ satisfies a Lipschitz condition with constant $K^n$, so by the Fundamental Contraction Mapping Inequality, if $K < 1$ then

$$\rho(F^n(x), F^m(x)) \leq \frac{K^n + K^m}{1 - K}\Big(\rho(x, F(x))\Big).$$

# The Banach Contraction Principle

Let $F : X \to X$ and $F^n$ its $n$-fold composition with itself. If $F$ satisfies a Lipschitz condition with constant $K$, by an easy induction $F^n$ satisfies a Lipschitz condition with constant $K^n$, so by the Fundamental Contraction Mapping Inequality, if $K < 1$ then

$$\rho(F^n(x), F^m(x)) \leq \frac{K^n + K^m}{1 - K}\Big(\rho(x, F(x))\Big).$$

In particular the successive approximation sequence $\{F^n(x)\}$ is a Cauchy sequence.

# The Banach Contraction Principle

Let $F : X \to X$ and $F^n$ its $n$-fold composition with itself. If $F$ satisfies a Lipschitz condition with constant $K$, by an easy induction $F^n$ satisfies a Lipschitz condition with constant $K^n$, so by the Fundamental Contraction Mapping Inequality, if $K < 1$ then

$$\rho(F^n(x), F^m(x)) \leq \frac{K^n + K^m}{1 - K}\Big(\rho(x, F(x))\Big).$$

In particular the successive approximation sequence $\{F^n(x)\}$ is a Cauchy sequence. Hence:

**Banach Contraction Principle.** *If $X$ is a complete metric space and if $F : X \to X$ is a contraction mapping, then $F$ has a unique fixed point $p$ in $X$ and for any $x \in X$ the successive approximation sequence $\{F^n(x)\}$ converges to $p$.*

# A Stopping Rule

When do we stop iterating and accept the current approximation? Suppose an "error" of $\epsilon$ is acceptable, and we start our iteration at $x \in X$. The Fundamental Inequality, with $x_1 = f^N(x)$ and $x_2 = p$ gives:

# A Stopping Rule

When do we stop iterating and accept the current approximation? Suppose an "error" of $\epsilon$ is acceptable, and we start our iteration at $x \in X$. The Fundamental Inequality, with $x_1 = f^N(x)$ and $x_2 = p$ gives:

$$\begin{aligned}\rho(f^N(x), p) &\leq \tfrac{1}{1-K}\rho(f^N(x), f^N(f(x))) \\ &\leq \tfrac{K^N}{1-K}\rho(x, f(x)).\end{aligned}$$

# A Stopping Rule

When do we stop iterating and accept the current approximation? Suppose an "error" of $\epsilon$ is acceptable, and we start our iteration at $x \in X$. The Fundamental Inequality, with $x_1 = f^N(x)$ and $x_2 = p$ gives:

$$\begin{aligned}\rho(f^N(x), p) &\leq \tfrac{1}{1-K}\rho(f^N(x), f^N(f(x))) \\ &\leq \tfrac{K^N}{1-K}\rho(x, f(x)).\end{aligned}$$

To insure $\rho(f^N(x), p) \leq \epsilon$, we must choose $N$ so large that $\frac{K^N}{1-K}\rho(x, f(x)) < \epsilon$. We can compute $\rho(x, f(x))$ after the first iteration and then find $N$ by solving the above inequality for $N$:

# A Stopping Rule

When do we stop iterating and accept the current approximation? Suppose an "error" of $\epsilon$ is acceptable, and we start our iteration at $x \in X$. The Fundamental Inequality, with $x_1 = f^N(x)$ and $x_2 = p$ gives:

$$\begin{aligned}\rho(f^N(x), p) &\leq \tfrac{1}{1-K}\rho(f^N(x), f^N(f(x))) \\ &\leq \tfrac{K^N}{1-K}\rho(x, f(x)).\end{aligned}$$

To insure $\rho(f^N(x), p) \leq \epsilon$, we must choose $N$ so large that $\frac{K^N}{1-K}\rho(x, f(x)) < \epsilon$. We can compute $\rho(x, f(x))$ after the first iteration and then find $N$ by solving the above inequality for $N$:

**Stopping Rule.** If $N > \frac{\log(\epsilon)+\log(1-K)-\log(\rho(x,f(x)))}{\log(K)}$ then $\rho(f^N(x), p) < \epsilon$.

# Speed of Convergence

Suppose we take $\epsilon = 10^{-m}$ in our stopping rule inequality. What we see is that the growth of $N$ with $m$ is a constant plus $m/|\log(K)|$, or in other words, to get one more decimal digit of precision we have to do (roughly) $1/|\log(K)|$ more iteration steps. So if we need $N$ iterative steps to get $m$ decimal digits of precision, then we need another $N$ to double the precision to $2m$ digits.

# Existence and Uniqueness Proof

We use the "sup" norm $\|\sigma\| = \sup_{t \in J} \|\sigma(t)\|$ to make $C(J, V)$ into a normed space. It is well-known that it is complete. (This is just the theorem that a uniformly Cauchy sequence is uniformly convergent.)

# Existence and Uniqueness Proof

We use the "sup" norm $\|\sigma\| = \sup_{t \in J} \|\sigma(t)\|$ to make $C(J, V)$ into a normed space. It is well-known that it is complete. (This is just the theorem that a uniformly Cauchy sequence is uniformly convergent.)

Given $\epsilon > 0$, let $M$ be Lipschitz constant for $f$ on the set of $(x, t) \in V \times \mathbf{R}$ with $\|x - p\| \leq 2\epsilon$ and $|t - t_0| \leq \epsilon$, and let $B$ be the maximum of $f(x, t)$ on this set.

# Existence and Uniqueness Proof

We use the "sup" norm $\|\sigma\| = \sup_{t \in J} \|\sigma(t)\|$ to make $C(J, V)$ into a normed space. It is well-known that it is complete. (This is just the theorem that a uniformly Cauchy sequence is uniformly convergent.)

Given $\epsilon > 0$, let $M$ be Lipschitz constant for $f$ on the set of $(x, t) \in V \times \mathbf{R}$ with $\|x - p\| \leq 2\epsilon$ and $|t - t_0| \leq \epsilon$, and let $B$ be the maximum of $f(x, t)$ on this set.

Choose $\delta > 0$ so that $K := M\delta < 1$ and $B\delta < \epsilon$, and define $X$ to be the set of $\sigma$ in $C(J, V)$ such that $\|\sigma(t) - p\| \leq 2\epsilon$ for all $|t| \leq \delta$. It is easy to see that $X$ is closed in $C(J, V)$ and so a complete metric space.

# Existence and Uniqueness Proof (Cont.)

The theorem will follow from the Banach Contraction Principle if we can show that for $\|v_0\| < \epsilon$, $F : C(J, V) \to C(J, V)$ maps $X$ into itself and has $K$ as a Lipschitz bound.

If $\sigma \in X$ then

$$\begin{aligned}\|F(\sigma)(t) - p\| &\leq \|v_0 - p\| + \int_0^t \|f(\sigma(s), s)\| \, ds \\ &\leq \epsilon + \delta B \leq 2\epsilon,\end{aligned}$$

so $F$ maps $X$ to itself.

## Existence and Uniqueness Proof (Cont.)

And if $\sigma, \tau \in X$ then

$\|f(\sigma(t), t) - f(\tau(t), t)\| \leq M \|\sigma(t) - \tau(t)\|$, so

$$\begin{aligned}
\|F(\sigma)(t) - F(\tau)(t)\| &\leq \int_0^t \|f(\sigma(s), s) - f(\tau(s), s)\| \, ds \\
&\leq \int_0^t M \|\sigma(s) - \tau(s)\| \, ds \\
&\leq \int_0^t M \rho(\sigma, \tau) \, ds \\
&\leq \delta M \rho(\sigma, \tau) \leq K \rho(\sigma, \tau),
\end{aligned}$$

and it follows that $\rho(F(\sigma), F(\tau) \leq K \rho(\sigma, \tau)$. ∎

# Maximal Solutions of the IVP

The Existence and Uniqueness Theorem is a central result in the theory of ODE with a great many important consequences. We next consider one easy corollary. A solution $\sigma : J \to V$ of the IVP is called the **maximal solution** for given initial data $t_0$ and $x_0$ if any other solution $x : I \to V$ with the same initial data, is a restriction of $\sigma$ to a subinterval $I$ of $J$.

# Maximal Solutions of the IVP

The Existence and Uniqueness Theorem is a central result in the theory of ODE with a great many important consequences. We next consider one easy corollary. A solution $\sigma : J \to V$ of the IVP is called the **maximal solution** for given initial data $t_0$ and $x_0$ if any other solution $x : I \to V$ with the same initial data, is a restriction of $\sigma$ to a subinterval $I$ of $J$.

**Proposition.** *If $f : V \times \mathbf{R} \to V$ is a locally Lipschitz time-dependent vector field, then for any initial data $t_0$ and $x_0$, the maximal solution of the IVP exists.*

# Maximal Solutions of the IVP

The Existence and Uniqueness Theorem is a central result in the theory of ODE with a great many important consequences. We next consider one easy corollary. A solution $\sigma : J \to V$ of the IVP is called the **maximal solution** for given initial data $t_0$ and $x_0$ if any other solution $x : I \to V$ with the same initial data, is a restriction of $\sigma$ to a subinterval $I$ of $J$.

**Proposition.** *If $f : V \times \mathbf{R} \to V$ is a locally Lipschitz time-dependent vector field, then for any initial data $t_0$ and $x_0$, the maximal solution of the IVP exists.*

Proof. Exercise. Hint:If $x_1 : I_1 \to V$ and $x_2 : I_2 \to V$ are solutions of the IVP with the same initial data, show that the set of $t \in \mathbf{R}$ with $x_1(t) = x_2(t)$ is a non-empty open and closed subset of $I_1 \cap I_2$.

# Maximal Solutions (Cont.)

**Exercise.** Show that if $\sigma : (a, b) \to V$ is a maximal solution of an IVP, then either $b = \infty$ or $\|\sigma(t)\| \to \infty$ as $t \to b$. Similarly, either $a = -\infty$ or $\|\sigma(t)\| \to \infty$ as $t \to a$. Hint: If $b < \infty$ and $\|\sigma(t)\| \not\to \infty$ as $t \to b$, there is a sequence $\{t_n\}$ converging to $b$ with$\{\sigma(t_n)\}$ converging to $p \in V$. Use the Existence and Uniqueness Theorem with $t_0 = b$ and $x_0 = p$ to show that the solution $\sigma$ could be extended to $(a, b + \epsilon)$ with $\epsilon > 0$, contradicting maximality of $\sigma$.

# Maximal Solutions (Cont.)

**Exercise.** Show that if $\sigma : (a, b) \to V$ is a maximal solution of an IVP, then either $b = \infty$ or $\|\sigma(t)\| \to \infty$ as $t \to b$. Similarly, either $a = -\infty$ or $\|\sigma(t)\| \to \infty$ as $t \to a$. Hint: If $b < \infty$ and $\|\sigma(t)\| \not\to \infty$ as $t \to b$, there is a sequence $\{t_n\}$ converging to $b$ with$\{\sigma(t_n)\}$ converging to $p \in V$. Use the Existence and Uniqueness Theorem with $t_0 = b$ and $x_0 = p$ to show that the solution $\sigma$ could be extended to $(a, b + \epsilon)$ with $\epsilon > 0$, contradicting maximality of $\sigma$.

**Exercise.** Suppose the vector field $f$ is bounded, or more generally satisfies $\int_1^\infty \frac{dr}{B(r)} = \infty$ where $B(r) = \sup_{\|x\|<r} \|f(x, t)\|$. Show that each maximal solution is defined on all of $\mathbf{R}$. Hint: How long does it take a solution to get outside a ball of radius $R$?

## Global Existence vs. Finite Time Blowup

If for a particular initial condition the maximal solution is defined on the entire real line then we say we have global existence for that initial condition, otherwise we say that there is finite-time blowup.

## Global Existence vs. Finite Time Blowup

If for a particular initial condition the maximal solution is defined on the entire real line then we say we have global existence for that initial condition, otherwise we say that there is finite-time blowup.

**Exercise.** On $\mathbf{R}$, consider the time-independent ODE $\frac{dx}{dt} = x^2$ with the initial condition $x(0) = x_0$. Show that in this case the maximal solution is $x(t) = \frac{x_0}{1-x_0 t}$ with the interval of definition is $(-\infty, \frac{1}{x_0})$ if $x_0 > 0$ and $(\frac{1}{x_0}, \infty)$ if $x_0 < 0$—in other words we have finite-time blowup at time $T = \frac{1}{x_0}$.

# Autonomous vs. Non-Autonomous ODE

A time-independent vector field $f$ on $V$ is also called **autonomous**. An obvious and characteristic property of autonomous ODEs $\frac{dx}{dt} = f(x)$ is that if $x(t)$ is a solution defined on $(a, b)$ then $x(t + c)$ is a solution defined on $(a - c, b - c)$. In particular, if the maximal solution for the initial condition $x(0) = p$ is $\sigma_p : (a, b) \to V$ then the maximal solution for the initial condition $x(t_0) = p$ is just $\sigma_p(t - t_0)$, defined on $(a + t_0, b + t_0)$.

# Autonomous vs. Non-Autonomous ODE

A time-independent vector field $f$ on $V$ is also called **autonomous**. An obvious and characteristic property of autonomous ODEs $\frac{dx}{dt} = f(x)$ is that if $x(t)$ is a solution defined on $(a, b)$ then $x(t + c)$ is a solution defined on $(a - c, b - c)$. In particular, if the maximal solution for the initial condition $x(0) = p$ is $\sigma_p : (a, b) \to V$ then the maximal solution for the initial condition $x(t_0) = p$ is just $\sigma_p(t - t_0)$, defined on $(a + t_0, b + t_0)$.

**Exercise.** We call $f$ **complete** if $\sigma_p$ has domain $\mathbf{R}$ for all $p \in V$. In this case we define the map $\phi_t : V \to V$ for each $t \in \mathbf{R}$ by $\phi_t(p) = \sigma_p(t)$. Show that $t \mapsto \phi_t$ is a homomorphism of $\mathbf{R}$ into the group of diffeomorphisms of $V$ (i.e., $\phi_{t_1+t_2} = \phi_{t_1} \circ \phi_{t_2}$).

# Reduction Theorems

We remarked earlier that even if one is interested only in solving the IVP for time-dependent vector fields in $\mathbf{R}^3$, there are still good reasons to consider the problem in more general vector spaces.

# Reduction Theorems

We remarked earlier that even if one is interested only in solving the IVP for time-dependent vector fields in $\mathbf{R}^3$, there are still good reasons to consider the problem in more general vector spaces.

We illustrate this by showing how to:

1) reduce an IVP for a non-autonomous vector field in $V$ to an IVP for a time-independent vector field in $V \times \mathbf{R}$, and

# Reduction Theorems

We remarked earlier that even if one is interested only in solving the IVP for time-dependent vector fields in $\mathbf{R}^3$, there are still good reasons to consider the problem in more general vector spaces.

We illustrate this by showing how to:

1) reduce an IVP for a non-autonomous vector field in $V$ to an IVP for a time-independent vector field in $V \times \mathbf{R}$, and

2) reduce an IVP for higher order ODE in $V$ to an IVP for a vector field in a product of copies of $V$.

# Reduction Theorems

We remarked earlier that even if one is interested only in solving the IVP for time-dependent vector fields in $\mathbf{R}^3$, there are still good reasons to consider the problem in more general vector spaces.

We illustrate this by showing how to:

1) reduce an IVP for a non-autonomous vector field in $V$ to an IVP for a time-independent vector field in $V \times \mathbf{R}$, and

2) reduce an IVP for higher order ODE in $V$ to an IVP for a vector field in a product of copies of $V$.

The proofs will be left as exercises.

## Time-Dependent → Time-Independent

As we have seen, autonomous ODE have a number of simplifying features, and this makes the following reduction quite useful.

# Time-Dependent → Time-Independent

As we have seen, autonomous ODE have a number of simplifying features, and this makes the following reduction quite useful.

**Exercise.** Let $f : V \times \mathbf{R} \to V$ be a time-dependent vector field in $V$, and define an associated time independent vector field $\tilde{f}$ in $V \times \mathbf{R}$ by $\tilde{f}(x, z) = (f(x, z), 1)$. Show that $y(t) = (x(t), z(t))$ is a solution of the differential equation $\frac{dy}{dt} = \tilde{f}(y)$ if and only if $z(t) = t + c$ and $x(t)$ is a solution of $\frac{dx}{dt} = f(x, t + c)$. Deduce that if $y(t) = (x(t), z(t))$ solves the IVP $\frac{dy}{dt} = \tilde{f}(y)$, $y(t_0) = (x_0, t_0)$, then $x(t)$ is a solution of the IVP $\frac{dx}{dt} = f(x, t)$, $x(t_0) = x_0$.

# Second Order $\rightarrow$ First Order

A curve $x(t)$ in $V$ is a solution of the second ODE $\frac{d^2x}{dt^2} = f(x, \frac{dx}{dt}, t)$ in $V$ if $x''(t) = f(x(t), x'(t), t)$. (Here of course $f$ is a map $V \times V \times \mathbf{R} \rightarrow V$.)

# Second Order → First Order

A curve $x(t)$ in $V$ is a solution of the second ODE $\frac{d^2x}{dt^2} = f(x, \frac{dx}{dt}, t)$ in $V$ if $x''(t) = f(x(t), x'(t), t)$. (Here of course $f$ is a map $V \times V \times \mathbf{R} \to V$.)

Define an associated time-dependent vector field $\tilde{f}$ on $V \times V$ by $\tilde{f}(x, v, t) = (v, f(x, v, t))$—so the associated first order ODE in $V \times V$ is $\frac{dx}{dt} = v$, $\frac{dv}{dt} = f(x, v, t)$.

# Second Order → First Order

A curve $x(t)$ in $V$ is a solution of the second ODE $\frac{d^2x}{dt^2} = f(x, \frac{dx}{dt}, t)$ in $V$ if $x''(t) = f(x(t), x'(t), t)$. (Here of course $f$ is a map $V \times V \times \mathbf{R} \to V$.)

Define an associated time-dependent vector field $\tilde{f}$ on $V \times V$ by $\tilde{f}(x, v, t) = (v, f(x, v, t))$—so the associated first order ODE in $V \times V$ is $\frac{dx}{dt} = v$, $\frac{dv}{dt} = f(x, v, t)$.

**Exercise.** Define the IVP for for the above second order ODE, and analyze the relation of this IVP and the IVP for the time-dependent vector field $\tilde{f}$ on $V \times V$. Use this to formulate and prove an Existence and Uniqueness Theorem for second order ODE. Now, generalize this to $m$-th order ODE in $V$

# Gronwall's Inequality.

The following estimate plays a very important role in ODE theory.

**Gronwall's Inequality.** *Let* $u : [0, T) \to [0, \infty)$ *be a continuous, non-negative, real-valued function and assume that* $u(t) \leq U(t) := C + K \int_0^t u(s)\, ds$ *for certain constants* $C \geq 0$ *and* $K > 0$. *Then* $u(t) \leq Ce^{Kt}$.

# Gronwall's Inequality.

The following estimate plays a very important role in ODE theory.

**Gronwall's Inequality.** *Let $u : [0, T) \to [0, \infty)$ be a continuous, non-negative, real-valued function and assume that $u(t) \leq U(t) := C + K \int_0^t u(s)\, ds$ for certain constants $C \geq 0$ and $K > 0$. Then $u(t) \leq Ce^{Kt}$.*

**Exercise.** Prove Gronwall's Inequality. Hint: Since $u \leq U$, it is enough to show that $U(t) \leq Ce^{Kt}$, or equivalently that $e^{-Kt}U(t) \leq C$, and since $U(0) = C$, it will suffice to show that $e^{-Kt}U(t)$ is non-increasing, i.e., that $(e^{-Kt}U(t))' \leq 0$. But, since $(e^{-Kt}U(t))' = e^{-Kt}(U'(t) - KU)$ and $U' = Ku$, this just says that $Ke^{-Kt}(u - U) \leq 0$.

# Continuity w.r.t Initial Conditions.

**Theorem.** *If $f$ is a $C^1$ vector field on $V$ and $\sigma_p(t)$ the maximal solution curve of $\frac{dx}{dt} = f(x)$ with initial condition $p$, then as $q$ tends to $p$, $\sigma_q(t)$ approaches $\sigma_p(t)$, uniformly for $t$ in a bounded interval $I$.*

# Continuity w.r.t Initial Conditions.

**Theorem.** *If $f$ is a $C^1$ vector field on $V$ and $\sigma_p(t)$ the maximal solution curve of $\frac{dx}{dt} = f(x)$ with initial condition $p$, then as $q$ tends to $p$, $\sigma_q(t)$ approaches $\sigma_p(t)$, uniformly for $t$ in a bounded interval $I$.*

**Proof.** Since $\sigma_p(t) = p + \int_0^t f(\sigma_p(s), s)\, ds$,
$\|\sigma_p(t) - \sigma_q(t)\| \leq$
$\|p - q\| + \int_0^t \|f(\sigma_p(s), s) - f(\sigma_q(s), s)\|\, ds.$
But on any bounded set (so on some neighborhood of $\sigma_p(I) \times I$), $f$ satisfies a Lipschitz condition:
$\|f(x,t) - f(y,t)\| \leq K\|x - y\|$, so $\|\sigma_p(t) - \sigma_q(t)\| \leq$
$\|p - q\| + K\int_{t_0}^t \|\sigma_p(s) - \sigma_q(s)\|\, ds$, and by Gronwall's Inequality, $\|\sigma_p(t) - \sigma_q(t)\| \leq \|p - q\|\, e^{Kt}$. ∎

# The IVP for Inhomogeneous Linear ODE

For a linear ODE $\frac{dx}{dt} = Ax$ with initial condition $x(t_0) = x_0$ we saw that the solution is $\exp((t-t_0)A)x_0$. If $g : \mathbf{R} \to V$ is a smooth function, then we can add it to the right hand side of the ODE, getting a so-called inhomogeneous linear ODE, and it turns out that the IVP for such equations can be solved in a fairly explicit form by a formula that for historical reasons is called "The Variation of Parameters Formula".

# The IVP for Inhomogeneous Linear ODE

For a linear ODE $\frac{dx}{dt} = Ax$ with initial condition $x(t_0) = x_0$ we saw that the solution is $\exp((t-t_0)A)x_0$. If $g : \mathbf{R} \to V$ is a smooth function, then we can add it to the right hand side of the ODE, getting a so-called inhomogeneous linear ODE, and it turns out that the IVP for such equations can be solved in a fairly explicit form by a formula that for historical reasons is called "The Variation of Parameters Formula".

**Exercise.** Show (by direct verification) that the solution of the IVP for the inhomogeneous linear differential equation $\frac{dx}{dt} = Ax + g(t)$ with initial condition $x(0) = x_0$ is given by:

$$x(t) = \exp(tA)x_0 + \int_0^t \exp((t-s)A)g(s)\,ds.$$

# Existence of a Periodic Orbit

**Exercise.** Assume that the linear operator $A$ is what is called "asymptotically stable"—namely that all of its eigenvalues have negative real part—and also that the forcing term $g(t)$ is periodic with period $T > 0$. Show that there is a point $p \in V$ for which the solution $x(t)$ with initial value $x(0) = p$ is periodic with period $T$.

# Existence of a Periodic Orbit

**Exercise.** Assume that the linear operator $A$ is what is called "asymptotically stable"—namely that all of its eigenvalues have negative real part—and also that the forcing term $g(t)$ is periodic with period $T > 0$. Show that there is a point $p \in V$ for which the solution $x(t)$ with initial value $x(0) = p$ is periodic with period $T$.

Hint: Since $x$ is given by the above variation of parameters formula, the condition that it be periodic of period $T$ is that $p = e^{TA}p + \int_0^T e^{(T-s)A} g(s)\, ds$, or that $p = (I - e^{TA})^{-1} \int_0^T e^{(T-s)A} g(s)\, ds$. Why is the operator $(I - e^{tA})$ invertible?

# Part II

# Numerical Solutions of Initial Value Problems

# Numerical Algorithms for Solving IVPs

Very few initial value problems admit explicit closed form solutions, so in general we must construct solutions numerically with the aid of a computer. But what algorithm should we use?

# Numerical Algorithms for Solving IVPs

Very few initial value problems admit explicit closed form solutions, so in general we must construct solutions numerically with the aid of a computer. But what algorithm should we use?

The natural first guess is successive approximations. But while that is a powerful theoretical tool for studying general properties of solutions (in particular, existence and uniqueness), it is not an efficient method for constructing numerical solutions.

# Numerical Algorithms for Solving IVPs

Very few initial value problems admit explicit closed form solutions, so in general we must construct solutions numerically with the aid of a computer. But what algorithm should we use?

The natural first guess is successive approximations. But while that is a powerful theoretical tool for studying general properties of solutions (in particular, existence and uniqueness), it is not an efficient method for constructing numerical solutions.

In fact there is no one simple answer to this question, for there is no one algorithm that is "best" in all situations. Below we will look at just two methods from the numerical analyst's extensive toolbox for solving initial value problems, Euler and Runge-Kutta.

# The General Approach

In what follows we will suppose that $f$ is a $C^1$ time-dependent vector field on $\mathbf{R}^n$, and given $t_0$ in $\mathbf{R}$ and $x_0$ in $\mathbf{R}^n$ we will denote by $x(t)$ or $\sigma(f, x_0, t_0, t)$ the maximal solution of the IVP $\frac{dx}{dt} = f(x, t)$ with initial condition $x(t_0) = x_0$.

# The General Approach

In what follows we will suppose that $f$ is a $C^1$ time-dependent vector field on $\mathbf{R}^n$, and given $t_0$ in $\mathbf{R}$ and $x_0$ in $\mathbf{R}^n$ we will denote by $x(t)$ or $\sigma(f, x_0, t_0, t)$ the maximal solution of the IVP $\frac{dx}{dt} = f(x, t)$ with initial condition $x(t_0) = x_0$.

The goal in the numerical integration of ODE is to devise efficient methods for approximating the solution $x(t)$ on an interval $I = [t_0, T]$.

# The General Approach

In what follows we will suppose that $f$ is a $C^1$ time-dependent vector field on $\mathbf{R}^n$, and given $t_0$ in $\mathbf{R}$ and $x_0$ in $\mathbf{R}^n$ we will denote by $x(t)$ or $\sigma(f, x_0, t_0, t)$ the maximal solution of the IVP $\frac{dx}{dt} = f(x, t)$ with initial condition $x(t_0) = x_0$.

The goal in the numerical integration of ODE is to devise efficient methods for approximating the solution $x(t)$ on an interval $I = [t_0, T]$.

The basic strategy is to interpolate $N$ equally spaced gridpoints $t_1, \ldots t_N$ in the interval $I$ (defined by $t_k := t_0 + k\Delta t$ with $\Delta t = \frac{T - t_0}{N}$), and use some algorithm to define values $x_1, \ldots, x_N$ in $\mathbf{R}^n$, in such a way that when $N$ is large each $x_k$ is close to the corresponding $x(t_k)$.

# The General Approach (Cont.)

The quantity $\max_{1 \leq k \leq N} \|x_k - x(t_k)\|$ is called the **global error** of the algorithm, and if it converges to zero as $N$ tends to infinity (for every choice of $f$, $t_0$, $x_0$, and $T$), then we say that we have a **convergent algorithm**.

# The General Approach (Cont.)

The quantity $\max_{1 \leq k \leq N} \|x_k - x(t_k)\|$ is called the **global error** of the algorithm, and if it converges to zero as $N$ tends to infinity (for every choice of $f$, $t_0$, $x_0$, and $T$), then we say that we have a **convergent algorithm**.

Even if a algorithm is convergent, that does not necessarily mean that it will provide an adequate method for solving initial value problems in all situations; other considerations such as stability and rate of convergence are important. However, if an algorithm is not at least convergent, that is sufficient reason to reject it as a tool for solving IVPs numerically.

# Stepping Methods

A common way to construct the algorithm that produces the values $x_1, \ldots, x_N$ uses a recursion based on a so-called **stepping procedure**, namely a function, $\Sigma(f, x_0, t_0, \Delta t)$, having as inputs:

**1)** a time-dependent vector field $f$ on $\mathbf{R}^n$,
**2)** an initial condition $x_0$ in $\mathbf{R}^n$,
**3)** an initial time $t_0$ in $\mathbf{R}$, and
**4)** a "time-step" $\Delta t$ in $\mathbf{R}$,
and with output a point of $\mathbf{R}^n$ that approximates $\sigma(f, x_0, t_0, t_0 + \Delta t)$ well when $\Delta t$ is small.

# Stepping Methods

A common way to construct the algorithm that produces the values $x_1, \ldots, x_N$ uses a recursion based on a so-called **stepping procedure**, namely a function, $\Sigma(f, x_0, t_0, \Delta t)$, having as inputs:

**1)** a time-dependent vector field $f$ on $\mathbf{R}^n$,
**2)** an initial condition $x_0$ in $\mathbf{R}^n$,
**3)** an initial time $t_0$ in $\mathbf{R}$, and
**4)** a "time-step" $\Delta t$ in $\mathbf{R}$,
and with output a point of $\mathbf{R}^n$ that approximates $\sigma(f, x_0, t_0, t_0 + \Delta t)$ well when $\Delta t$ is small.

More precisely, the so-called **local truncation error**, defined by $\|\sigma(f, x_0, t_0, t_0 + \Delta t) - \Sigma(f, x_0, t_0, \Delta t)\|$, should approach zero at least quadratically in the time-step $\Delta t$.

# Stepping Methods (Cont.)

Given such a stepping procedure, the approximations $x_k$ of the $x(t_k)$ are defined recursively by $x_{k+1} := \Sigma(f, x_k, t_k, \Delta t)$. Numerical integration methods that follow this general pattern are referred to as **finite difference methods**.

# Stepping Methods (Cont.)

Given such a stepping procedure, the approximations $x_k$ of the $x(t_k)$ are defined recursively by $x_{k+1} := \Sigma(f, x_k, t_k, \Delta t)$. Numerical integration methods that follow this general pattern are referred to as **finite difference methods**.

There are two main sources contributing to the global error, $\|x_k - x(t_k)\|$. At each step there will be an additional local truncation error, and after the first step, there will be an error because the recursion uses $\Sigma(f, x_k, t_k, \Delta t)$ rather than $\Sigma(f, x(t_k), t_k, \Delta t)$. In practice there is a third source of error, namely round-off error from using floating-point arithmetic, We will ignore this, pretending that our computers do precise real arithmetic, but there are situations where it is important to take round-off error into consideration.

# Euler's Method

Euler's Method is defined by the particularly simple and natural stepping procedure: :

**Euler Step:** $\Sigma^E(f, x_0, t_0, \Delta t) := x_0 + \Delta t\, f(x_0, t_0)$.

# Euler's Method

Euler's Method is defined by the particularly simple and natural stepping procedure: :

**Euler Step:** $\Sigma^E(f, x_0, t_0, \Delta t) := x_0 + \Delta t\, f(x_0, t_0)$.

It is not hard to see why this is a good choice. If as above we denote $\sigma(f, x_0, t_0, t)$ by $x(t)$, then by Taylor's Theorem:

$$\begin{aligned} x(t_0 + \Delta t) &= x(t_0) + \Delta t\, x'(t_0) + O(\Delta t^2) \\ &= x_0 + \Delta t\, f(x_0, t_0) + O(\Delta t^2) \\ &= \Sigma^E(f, x_0, t_0, \Delta t) + O(\Delta t^2), \end{aligned}$$

# Euler's Method

Euler's Method is defined by the particularly simple and natural stepping procedure: :

**Euler Step:** $\Sigma^E(f, x_0, t_0, \Delta t) := x_0 + \Delta t\, f(x_0, t_0)$.

It is not hard to see why this is a good choice. If as above we denote $\sigma(f, x_0, t_0, t)$ by $x(t)$, then by Taylor's Theorem:

$$\begin{aligned} x(t_0 + \Delta t) &= x(t_0) + \Delta t\, x'(t_0) + O(\Delta t^2) \\ &= x_0 + \Delta t\, f(x_0, t_0) + O(\Delta t^2) \\ &= \Sigma^E(f, x_0, t_0, \Delta t) + O(\Delta t^2), \end{aligned}$$

so that $\|\sigma(f, x_0, t_0, t_0 + \Delta t) - \Sigma(f, x_0, t_0, \Delta t)\|$, the local truncation error for Euler's Method, does go to zero quadratically in $\Delta t$.

# Euler's Method (Cont.)

When we partition $[0, T]$ into $N$ equal parts, $\Delta t = \frac{T - t_0}{N}$, each step in the recursion for computing $x_k$ will contribute a local truncation error that is $O(\Delta t^2) = O(\frac{1}{N^2})$, and since there are $N$ steps in the recursion, this suggests that the global error will be $O(\frac{1}{N})$, and hence will go to zero as $N$ tends to infinity. Thus we expect Euler's Method to be a convergent algorithm. We will give a rigorous argument below.

# Euler's Method (Cont.)

When we partition $[0, T]$ into $N$ equal parts, $\Delta t = \frac{T-t_0}{N}$, each step in the recursion for computing $x_k$ will contribute a local truncation error that is $O(\Delta t^2) = O(\frac{1}{N^2})$, and since there are $N$ steps in the recursion, this suggests that the global error will be $O(\frac{1}{N})$, and hence will go to zero as $N$ tends to infinity. Thus we expect Euler's Method to be a convergent algorithm. We will give a rigorous argument below.

**Exercise.** Show that Euler's Method applied to the initial value problem $\frac{dx}{dt} = x$ with $x(0) = 1$ gives $\lim_{N\to\infty}(1 + \frac{t}{N})^N = e^t$.

# Error Estimate for Euler Method

Assume the vector field $f$ satisfies the local Lipschitz bound $\|f(p,t) - f(q,t)\| \leq L \|p - q\|$. We use an argument of Hermann Karcher to estimate the error in Euler's method.

# Error Estimate for Euler Method

Assume the vector field $f$ satisfies the local Lipschitz bound $\|f(p,t) - f(q,t)\| \leq L\,\|p - q\|$. We use an argument of Hermann Karcher to estimate the error in Euler's method.

Recall that Euler's Method approximates $x(t)$ at points $t_n := t_{n-1} + \Delta T = t_0 + n\,\Delta T$, where $\Delta T = \frac{T - t_0}{N}$, and the approximations $e(t_n)$ are defined inductively by $e(t_0) := x_0$, and $e(t_{n+1}) := e(t_n) + \Delta T\, f(e(t_n), t_n)$.

# Error Estimate for Euler Method

Assume the vector field $f$ satisfies the local Lipschitz bound $\|f(p,t) - f(q,t)\| \leq L\,\|p - q\|$. We use an argument of Hermann Karcher to estimate the error in Euler's method.

Recall that Euler's Method approximates $x(t)$ at points $t_n := t_{n-1} + \Delta T = t_0 + n\,\Delta T$, where $\Delta T = \frac{T - t_0}{N}$, and the approximations $e(t_n)$ are defined inductively by $e(t_0) := x_0$, and $e(t_{n+1}) := e(t_n) + \Delta T\, f(e(t_n), t_n)$.

We interpolate the discrete approximations $e(t_n)$ with a piecewise-linear curve $e(t)$ defined on $[t_n, t_{n+1}]$ by $e(t) := e(t_n) + (t - t_n)\, f(e(t_n), t_n)$. Think of this as considering the time difference $\Delta t := t - t_n$ as a variable, allowing us to estimate the difference or "error", $\mathrm{Err}(t) := \|x(t) - e(t)\|$ by a Gronwall-like argument.

# Error Estimate for Euler Method (Cont.)

We first estimate the error for a single time-step, i.e., on the interval $t_0 \leq t \leq t_0 + \Delta T$. From the definition of $e$, $\dot{e}(t) = f(x_0, t_0)$, $t_0 \leq t \leq t_0 + \Delta T$, so $\ddot{e} = 0$. It follows that $\dot{x}(t) - \dot{e}(t) = f(x(t), t) - f(x_0, t_0) = f(x(t), t) - f(e(t), t) + f(e(t), t) - f(x_0, t_0)$, so

$$\|\dot{x}(t) - \dot{e}(t)\| \leq L \|x(t) - e(t)\| + \|f(e(t), t) - f(x_0, t_0)\|.$$

# Error Estimate for Euler Method (Cont.)

We first estimate the error for a single time-step, i.e., on the interval $t_0 \le t \le t_0 + \Delta T$. From the definition of $e$, $\dot{e}(t) = f(x_0, t_0)$, $t_0 \le t \le t_0 + \Delta T$, so $\ddot{e} = 0$. It follows that $\dot{x}(t) - \dot{e}(t) = f(x(t), t) - f(x_0, t_0) = f(x(t), t) - f(e(t), t) + f(e(t), t) - f(x_0, t_0)$, so

$$\|\dot{x}(t) - \dot{e}(t)\| \le L \|x(t) - e(t)\| + \|f(e(t), t) - f(x_0, t_0)\|.$$

The second term on the right is not yet in a form to apply a Gronwall argument. However, if we define $K := \max_{t_0 \le t \le t_0 + \Delta t} \left\| \frac{d}{dt} (f(e(t), t) - f(x_0, t_0) \right\|$ then we obtain the differential inequality

$$\|\dot{x}(t) - \dot{e}(t)\| \le L \|\mathrm{Err}(t)\| + K(t - t_0).$$

# Error Estimate for Euler Method (Cont.)

Since
$\mathrm{Err}(t) = \left\| \int_{t_0}^{t} (\dot{x}(t) - \dot{e}(t))\, dt \right\| \leq \int_{t_0}^{t} \|\dot{x}(t) - \dot{e}(t)\|\, dt,$
we see that $\|\mathrm{Err}(t)\| \leq \psi(t)$, where $\psi$ is the differentiable function:

$$\psi(t) := \|\mathrm{Err}(t_0)\| + L \int_{t_0}^{t} \|\mathrm{Err}(t)\|\, dt + K \int_{t_0}^{t} (t - t_0)\, dt.$$

# Error Estimate for Euler Method (Cont.)

Since
$\mathrm{Err}(t) = \left\| \int_{t_0}^{t} (\dot{x}(t) - \dot{e}(t))\, dt \right\| \leq \int_{t_0}^{t} \|\dot{x}(t) - \dot{e}(t)\|\, dt$,
we see that $\|\mathrm{Err}(t)\| \leq \psi(t)$, where $\psi$ is the differentiable function:

$$\psi(t) := \|\mathrm{Err}(t_0)\| + L \int_{t_0}^{t} \|\mathrm{Err}(t)\|\, dt + K \int_{t_0}^{t} (t - t_0)\, dt.$$

Since $\dot{\psi} = L\,\|\mathrm{Err}(t)\| + K(t - t_0)$ and $\|\mathrm{Err}(t)\| \leq \psi$, we have the differential inequality $\dot{\psi} \leq L\psi + K(t - t_0)$ that we now use for a Gronwall argument.

# Error Estimate for Euler Method (Cont.)

Compute the derivative of the function

$$\left(\psi + \frac{K}{L^2} + \frac{K}{L}(t - t_0)\right) \cdot e^{-L \cdot (t - t_0)}.$$

# Error Estimate for Euler Method (Cont.)

Compute the derivative of the function

$$\left(\psi + \frac{K}{L^2} + \frac{K}{L}(t - t_0)\right) \cdot e^{-L\cdot(t-t_0)}.$$

$$\frac{d}{dt}\left(\left(\psi + \frac{K}{L^2} + \frac{K}{L}(t - t_0)\right) \cdot e^{-L\cdot(t-t_0)}\right) =$$

$$\left(\dot{\psi} + \frac{K}{L} - L\left(\psi + \frac{K}{L^2} + \frac{K}{L}(t - t_0)\right)\right) \cdot e^{-L\cdot(t-t_0)}.$$

# Error Estimate for Euler Method (Cont.)

Compute the derivative of the function

$$\left(\psi + \frac{K}{L^2} + \frac{K}{L}(t - t_0)\right) \cdot e^{-L\cdot(t-t_0)}.$$

$$\frac{d}{dt}\left(\left(\psi + \frac{K}{L^2} + \frac{K}{L}(t - t_0)\right) \cdot e^{-L\cdot(t-t_0)}\right) =$$

$$\left(\dot{\psi} + \frac{K}{L} - L\left(\psi + \frac{K}{L^2} + \frac{K}{L}(t - t_0)\right)\right) \cdot e^{-L\cdot(t-t_0)}.$$

By the differential inequality for $\psi$, this function has a non-positive derivative, so all of its values are less than its "initial" value at $t_0$, namely $\mathrm{Err}(t_0) + \frac{K}{L^2}$, and solving for $\psi$ we obtain the desired error estimate:

$$\mathrm{Err}(t) \leq \psi(t) \leq \left(\mathrm{Err}(t_0) + \frac{K}{2}(t - t_0)^2\right) e^{L\cdot(t-t_0)}$$

for $t_0 \leq t \leq t_0 + \Delta T$.

# Error Estimate for Euler Method (Cont.)

To iterate this estimate we define the starting point for the second time step as $x_1 := e(t_0 + \Delta T)$, so that we have the initial error bound

$$|x(t_0 + \Delta t) - x_1| \leq$$

$$\mathrm{Err}(t_1) := \left(\mathrm{Err}(t_0) + \frac{K}{2}(t - t_0)^2\right) e^{L \cdot \Delta T},$$

## Error Estimate for Euler Method (Cont.)

To iterate this estimate we define the starting point for the second time step as $x_1 := e(t_0 + \Delta T)$, so that we have the initial error bound

$$|x(t_0 + \Delta t) - x_1| \leq$$

$$\mathrm{Err}(t_1) := \left(\mathrm{Err}(t_0) + \frac{K}{2}(t - t_0)^2\right) e^{L \cdot \Delta T},$$

and after the second time step we have

$$\mathrm{Err}(t) \leq \psi(t) \leq \left(\mathrm{Err}(t_1) + \frac{K}{2}(t - t_1)^2\right) e^{L \cdot (t - t_1)}$$

for $t_1 \leq t \leq t_1 + \Delta T$.

## Error Estimate for Euler Method (Cont.)

To reach the fixed time $T$ one needs $N$ time steps of size $\Delta t := (T - t_0)/N$ and the $N$-fold iteration of the error estimate gives (replace $N \cdot \Delta T$ by $(T - t_0)$, recall $\mathrm{Err}(t_0) = 0$ and use the sum of the geometric series):

$$\begin{aligned}
\mathrm{Err}(t) &\leq \mathrm{Err}(t_0) \cdot e^{L(T-t_0)} + \frac{K}{2}\Delta T^2 \cdot \sum_{k=1}^{N} e^{kL\cdot\Delta T} \\
&\leq \frac{K}{2}\Delta T^2 \cdot (e^{L(T-t_0+\Delta T)} - 1)/(e^{L\Delta t} - 1) \\
&\leq \frac{K}{2L}\Delta T \cdot (e^{L(T-t_0+\Delta T)} - 1).
\end{aligned}$$

This proves in particular that when $\Delta T \to 0$ the iterated Euler curves converge uniformly to the exact solution, or in other words that Euler's Method is a convergent algorithm.

# Runge-Kutta

Despite what we have just proved, perhaps the only positive thing that can be said about the Euler method for solving an IVP is that it is intuitive and easy to program. Beyond that there is little to recommend it as a practical method for solving real-world problems. It requires very small time steps to get reasonable accuracy, making it very slow, and in fact it is rarely used except for pedagogical purposes.

# Runge-Kutta

Despite what we have just proved, perhaps the only positive thing that can be said about the Euler method for solving an IVP is that it is intuitive and easy to program. Beyond that there is little to recommend it as a practical method for solving real-world problems. It requires very small time steps to get reasonable accuracy, making it very slow, and in fact it is rarely used except for pedagogical purposes.

A general purpose finite difference method for solving IVPs that is the most useful (and the only other one that we will consider) is Runge-Kutta, or more precisely the fourth order Runge-Kutta Method, as there is a whole family of Runge-Kutta methods. It is in fact one of the most implemented an useful pieces of numerical software for any purpose.

# Runge-Kutta Step

The stepping procedure for fourth order Runge-Kutta is a **lot** less transparent than that for Euler. It is given by the following formula:

## Runge-Kutta Step

$$\Sigma^{RK^4}(f, x_0, t_0, \Delta t) := x_0 + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),$$

# Runge-Kutta Step

The stepping procedure for fourth order Runge-Kutta is a **lot** less transparent than that for Euler. It is given by the following formula:

## Runge-Kutta Step

$\Sigma^{RK^4}(f, x_0, t_0, \Delta t) := x_0 + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4)$, where:

$k_1 = \Delta t\, f(x_0, t_0)$

$k_2 = \Delta t\, f(x_0 + \frac{1}{2}k_1, t_0 + \frac{\Delta t}{2})$

$k_3 = \Delta t\, f(x_0 + \frac{1}{2}k_2, t_0 + \frac{\Delta t}{2})$

$k_4 = \Delta t\, f(x_0 + k_3, t_0 + \Delta t)$

# Runge-Kutta Pseudo Code

```
function RungeKutta4(f,x0,t0,h,N):vector;
f: function(v:vector;s:real):vector;
x0:vector ;t0,h:real; N:integer;
var  j:integer; t:real; x,k1,k2,k3,k4:vector;
begin
   t := t0;    x := x0;
   for j := 1 to N do
      begin
         k1 := h f(x,t);
         k2 := h f(x+ k1/2,t+h/2);
         k3 := h f(x+ k2/2,t+h/2);
         k4 := h f(x+ k3,t+h);
         x := x + (k1+2(k2+k3)+k4)/6;
         t := t + h;
      end;
   RungeKutta4 := x;
end;
```

# Runge-Kutta (Cont.)

It is of course a fair question to ask where such a strange formula comes from. If you are familiar with Simpson's Rule for evaluating the definite integral of a function $\phi(t)$, then the above should not look unreasonable, and indeed if $f(x,t) = \phi(t)$ then recall that the solution of the IVP reduces to the integral of $\phi$ and in this case the Runge-Kutta formula reduces precisely to Simpson's Rule. And like Simpson's Rule, Runge-Kutta is fourth order, meaning that the local truncation error goes to zero as the fifth power of the step-size, and the global error as the fourth power. So if for a fixed step-size we have attained an accuracy of 0.1, then with one-tenth the step-size (and so ten times the number of steps and ten times the time) we can expect an accuracy of 0.00001, whereas with the Euler method, ten times the time would only increase accuracy from 0.1 to 0.01.