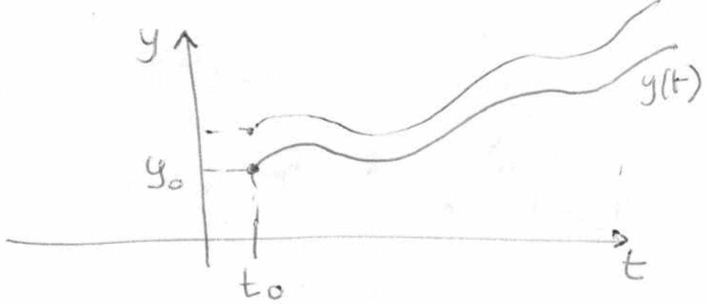
ORDINARY DIFFERENTIAL EQUATIONS (ODE)

$I \subseteq \mathbb{R}$ interval , $f: I \times \mathbb{R} \to \mathbb{R}$
$t_0 \in I$ e $y_0 \in \mathbb{R}$

? $y: I \to \mathbb{R}$ solution of the 1st order Cauchy problem

$$(1) \quad \begin{cases} y'(t) = f(t, y(t)) & t > t_0 ,\ t \in I \\ y(t_0) = y_0 \end{cases}$$



<u>theorem</u>: f continuous w.r.t. both t and y,
f lipsch on the second argument, uniformly w.r.t. to t, i.e. $\exists L > 0 : \forall y_1, y_2 \in \mathbb{R}$

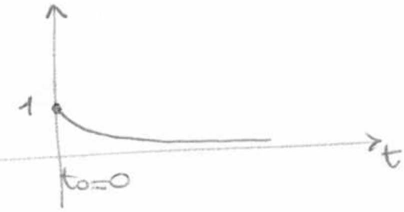
$$|f(t, y_1) - f(t, y_2)| \le L\, |y_1 - y_2| \quad \forall t \in I$$

$\Rightarrow \exists!\, y: I \to \mathbb{R}$, $y \in C^1(I)$ solution of (1)

From the numerical point of view we are not interested in looking for the ~~ex~~ mathematical expression of $y(t)$.

Ex: $\begin{cases} y'(t) = 3y(t) \\ y(0) = 1 \end{cases}$ the exact solution is

$y(t) = e^{3t}$



We are interested in approximating the exact solution in some points in $(t_0, T]$.

1. We discretize $(t_0, T]$ the time interval).

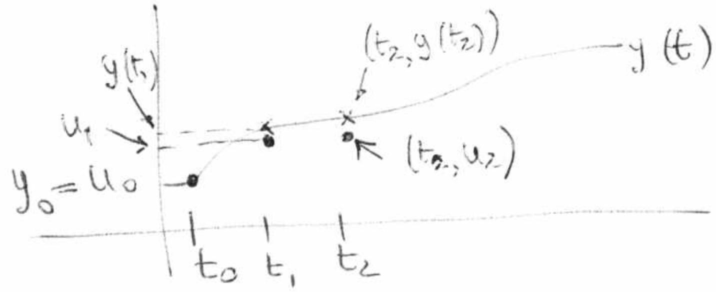
Given $h > 0$: $t_n = t_0 + n \cdot h$ $n = 0, \dots, N_h$:

$t_{N_h} = T$.

$t_{n+1} = t_n + h$

? $\{u_0, u_1, \dots, u_{N_h}\}$: $\underset{\substack{\uparrow \\ \text{appx sol.}}}{u_n} \simeq \underbrace{y(t_n)}_{\substack{\uparrow \\ \text{exact sol}}}$

We set $u_0 = y_0$ then we compute u_1 by appx y'
Once u_1 is known, we compute u_2, and so on



How to build $\{u_0, u_1, \dots, u_{N_h}\}$?
There are a lot of possibility depending on the way the derivative is approximated by.

For example if $y'(t_n) \simeq \frac{y_n - y_{n-1}}{h}$ backward finite difference

$(y_n = y(t_n))$

from the Cauchy problem:

$$y'(t) = f(t, y(t)) \quad \forall t \in (t_0, T)$$

$$\Rightarrow \quad y'(t_n) = f(t_n, y_n) \quad (\text{since } t_n \in (t_0, T))$$

$$\simeq \frac{y_n - y_{n-1}}{h} \quad \Rightarrow \quad \frac{y_n - y_{n-1}}{h} \simeq f(t_n, y_n)$$

this is not =

We define $\{u_1, u_2, \dots u_{N_h}\}$ the real numbers that satisfy the equality $\frac{u_n - u_{n-1}}{h} = f(t_n, u_n)$.

We explicite u_n : $\begin{cases} u_n = u_{n-1} + h f(t_n, u_n) & n \geq 1 \\ u_0 = y_0 \end{cases}$

Backward Euler method

$$\begin{cases} u_0 = y_0 \\ u_{n+1} = u_n + h\, f(t_{n+1}, u_{n+1}) & n \geq 0 \end{cases}$$

If we approximate $y'(t_n) \simeq \dfrac{y_{n+1}-y_n}{h}$

FORWARD FINITE DIFF

$$\Rightarrow \left[\begin{array}{l} y'(t_n) = f(t_n, y_n) \\ \simeq \dfrac{y_{n+1}-y_n}{h} \end{array}\right.$$

$$\frac{y_{n+1}-y_n}{h} \simeq f(t_n, y_n)$$

$$\frac{u_{n+1}-u_n}{h} = f(t_n, u_n)$$

$$\boxed{\begin{cases} u_{n+1} = u_n + h f(t_n, u_n) & n \geq 0 \\ u_0 = y_0 \end{cases} \qquad \text{FORWARD EULER.}}$$

Another approach:

by the fundamental theorem of integral calculus

$$\begin{cases} y'(t) = f(t, y(t)) & t \in (t_0, T] \\ y(t_0) = y_0 \end{cases} \iff y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

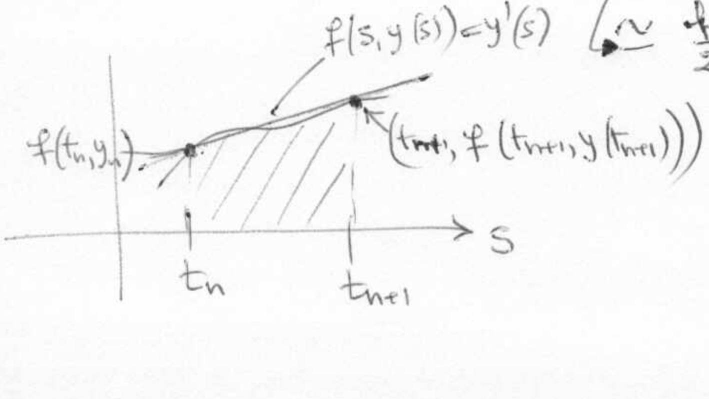
differential form $\qquad\qquad$ integral form

Here we approximate y' $\qquad\qquad$ here we approximate $\int$

the CRANK-NICOLSON Method:

$\forall t_n \in (t_0, T]$ and $t_{n+1} = t_n + h$

$$y(t_{n+1}) = y(t_n) + \underbrace{\int_{t_n}^{t_{n+1}} f(s, y(s))\, ds}_{\simeq \frac{h}{2}\left[f(t_n, y(t_n)) + f(t_{n+1}, y(t_{n+1}))\right]}$$



the $\int$ is approximated by the area of the trapezoid

$$\Rightarrow \quad y_{n+1} \simeq y_n + \frac{h}{2}\left(f(t_n, y_n) + f(t_{n+1}, y_{n+1})\right)$$

we define $\{u_1, u_2, \ldots, u_{N_h}\}$ the set of real numbers

s.t.

$$\begin{cases} u_{n+1} = u_n + \frac{h}{2}\left[f(t_n, u_n) + f(t_{n+1}, u_{n+1})\right] & n \geq 0 \\ u_0 = y_0 \end{cases}$$

CRANK NICOLSON METHOD -

Each of these 3 methods yields u_{n+1} once u_n is known.

To compute u_{n+1} only the previous value u_n is needed. ~~more precisely the last value u_n~~ -

We say that all these methods are ONE-STEP methods since only the last known step is required to compute u_{n+1} -

For example the scheme

$$u_{n+1} = u_n + \frac{h}{2}\left[3f(t_n, u_n) - f(t_{n-1}, u_{n-1})\right]$$

is a TWO-STEPS methods since 2 previous steps are needed to compute u_n -

BE and CN are <u>implicit</u> methods, while FE is <u>explicit</u> -

As a matter of fact in FE, u_{n+1} is explicit while in both BE and CN u_{n+1} depends on itself through $f(t_{n+1}, u_{n+1})$ -

Explicit methods are cheaper

implicit methods are very expensive since to pass from u_n to u_{n+1} we need to solve a non-linear system of equations (which is equivalent to about 6-10 linear systems)

Each method is characterized by:

1 - n° of steps ✓

2 - explicit or implicit type ✓

3 - order of convergence

4 - stability properties.

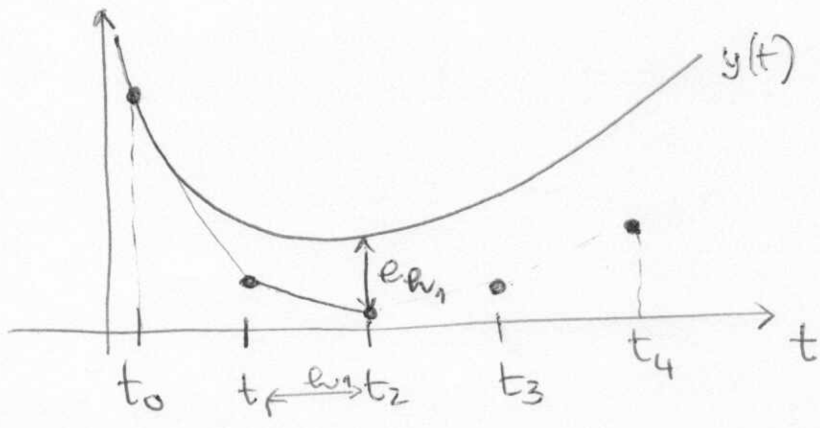
CONVERGENCE

Def. A numerical scheme ~~to~~ approximating the solution of an ODE is CONVERGENT if $\exists$ a function $C = C(h)$: $\lim_{h \to 0} C(h) = 0$ s.t.

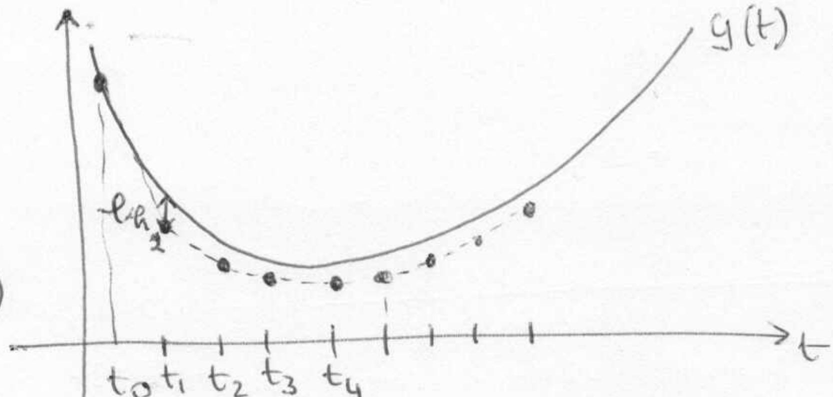
$$e_h = \max_{0 \le n \le N_h} |y(t_n) - u_n| \le C(h)$$

moreover, if $\exists p \in \mathbb{N}$: $C(h) \simeq h^p$ when $h \to 0$
$\Rightarrow$ the scheme has order of convergence $= p$.

$h_1 = \vdash\!\!\dashv$



$h_2 = \vdash\!\dashv$
(~~the~~ one half of the previous one)



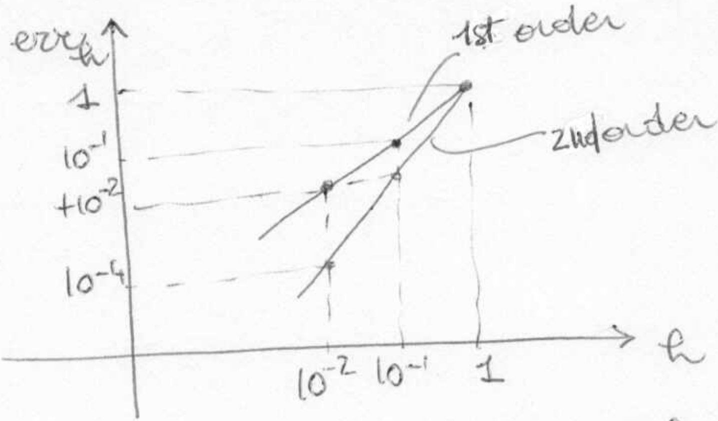
By taking $h_3 < h_2$ we obtain $e_{h_3} < e_{h_2}$
and when $h \to 0$ $e_h \to 0$. All the values $\{u_1, \dots u_{N_h}\}$ come close to the exact $y(t)$.

BE and FE are convergent methods with order $p=1$

CN is convergent with order $p=2$.

the higher the order is and the more appealing the method is,

By fixing h, the second-order method will provide a more accurate (precise) solution than the first-order method does.



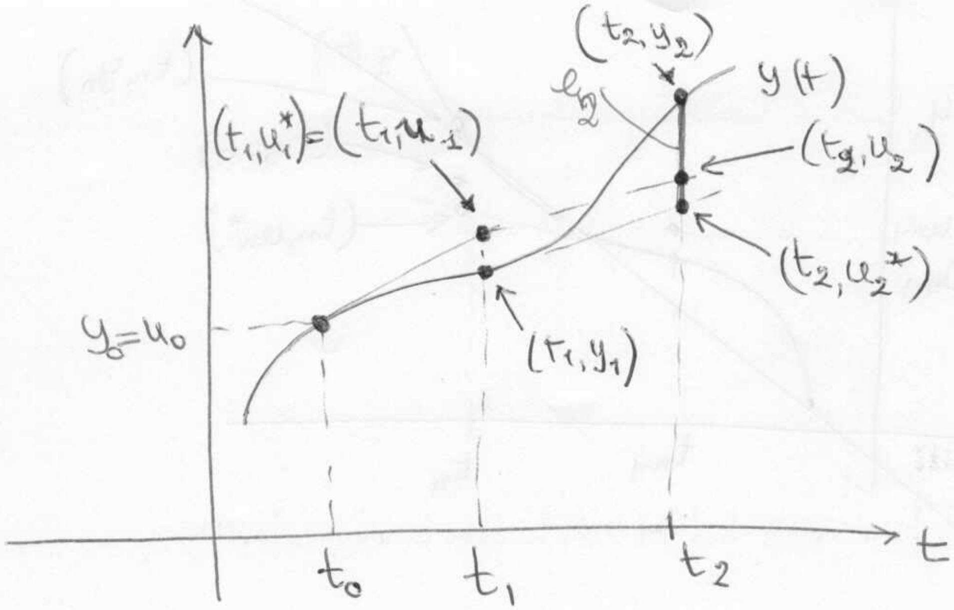
1st order method: by reducing h of one order (from 1 to 10^{-1}) the error reduces of one order

2nd-order method: by reducing h of one order the error reduces of two orders.

At each step t_n the error between the exact and the numerical solution $e_n = y(t_n) - u_n$ can be decomposed in the sum of two terms:

- the first one is the error introduced at step n when the exact derivative is approximated;
- the second one is due to the accumulation of the errors introduced at the previous steps.

Both of them have to be infinitesimal to guarantee the convergence



Analysis for FE $\quad u_{n+1} = u_n + h f(t_n, u_n)$

$t_0 \;:\; y_0 - u_0 = 0$

$t_1 \;:\; y_1 - u_1 = c h^2 \qquad$ (c: constant > 0)

$$\left(y(t_{n+1}) = y(t_n) + h\, y'(t_n) + \frac{h^2}{2} y''(\xi_n) \right)$$

Taylor expansion

$t_2 \;:\; y_2 - u_2 = (y_2 - u_2^*) + (u_2^* - u_2)$

where u_2^* would be the numerical solution if we start from the exact value y_1

(u_2^* is a fictitious variable)

$(y_2 - u_2^*)$ is due to the appx error introduced at step 2 when we appx the derivative

$(u_2^* - u_2)$ is due to the accumulation of previous errors.

<u>Def</u> $\forall n$ $\quad \tau_n(h) = \dfrac{y_n - u_n^*}{h}$ is called LOCAL TRUNCATION ERROR

and it is the error that would be generated by forcing the exact solution to satisfy the numerical scheme

EX: FE $\quad y_{n+1} = y_n + h \underbrace{y'(t_n)}_{f(t_n, y(t_n))} + \frac{h^2}{2} y''(\xi_n)$ (Taylor)

$$u_{n+1} = u_n + h f(t_n, u_n)$$

when we force y_{n+1} to satisfy FE we neglect the remainder of the taylor expansion

$\left(\frac{h^2}{2} y''(\xi_n)\right) \Rightarrow \tau_n(h) = \dfrac{y_n - u_n^*}{h} = \frac{1}{h}\frac{h^2}{2} y''(\xi_n) = h \dfrac{y''(\xi_n)}{2}$

If y'' is bounded $\Rightarrow \tau_n(h) \simeq C_n h$ and it is infinitesimal when $h \to 0$

<u>Def</u> $\tau(h) = \max\limits_n \tau_n(h)$ GLOBAL TRUNCATION ERROR

<u>Def</u> A numerical scheme for approximating ODE is CONSISTENT if $\tau(h)$ is infinitesimal when $h \to 0$.

EX: FE: $\tau(h) = Ch$; BE: $\tau(h) = Ch$

CN: $\tau(h) = Ch^2$

All these scheme are consistent.

Let us examine $(u_n^* - u_n)$, it is due to the propagation / accumulation of previous local truncation errors.

Local truncation errors can be interpreted as perturbations on the data introduced at each step.

~~We~~ ZERO STABILITY

Roughly speaking, we say that a scheme is zero-stable for the numerical solution of ~~the~~ Cauchy problem on a <u>bounded interval</u> if small perturbations on the data yield small perturbations on the solution when $h \to 0$.

More precisely if

$\exists C > 0, \exists h_0 > 0, \exists \varepsilon_0 > 0$:

$$\forall h \in (0, h_0], \forall \varepsilon \in (0, \varepsilon_0]$$

$$|z_n - u_n| \leq C\varepsilon \qquad \forall n = 0, \dots, N_h$$

where: C is a const indep of h, n, ε (it can depend on the interval (t_0, T))

- ε is the maximum size of the perturbations
- u_n is the unperturbed solution
- z_n is the perturbed solution.

therefore, if our errors on the data are $\tau_n(h)$ (which are infinitesimal) and the scheme is

zero-stable $\Rightarrow |u_n - u_n^*| \leq C \cdot \underbrace{\varphi_n(h)}$

that depends on $\tau_n(h)$ and is infinitesimal when $h \to 0$.

Finally: consistency guarantees that $\tau_n(h) \underset{h\to 0}{\to} 0 \quad \forall n$
and $|y_n - u_n^*| \to 0$ when $h \to 0$
zero-stab guarantees that
$|u_n^* - u_n| \to 0$ when $h \to 0$

then $\boxed{\text{CONS} + \text{ZEROSTAB} \Rightarrow |y_n - u_n| \underset{h\to 0}{\to} 0 \Leftrightarrow \text{CONVERGENCE}}$

Actually CONVERGENCE $\Leftrightarrow$ (CONSISTENCY + ZERO STABILITY).

BE, FE, CN are zero-stab, then are convergent.

When the integration interval (t_0, T) is unbounded the constant $C \to +\infty$ and the zero-stab is insufficient to guarantee convergence, we have to control ~~something~~ a stronger concept of stability: the ABSOLUTE STABILITY.

ABSOLUTE STABILITY

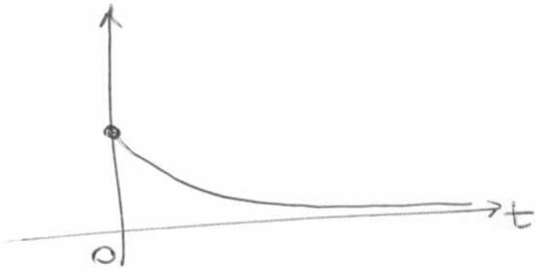
We study this property on a very simple problem but the concept can be extended to every ODE.

Let us consider the linear model Cauchy prob

$$(1)\quad \begin{cases} y' = \lambda y & t>0 \\ y(0)=1 \end{cases} \qquad \text{where } \lambda \in \mathbb{C} : Re(\lambda)<0$$

If $\lambda \in \mathbb{R}^- \Rightarrow y(t) = e^{\lambda t}$

and $y(t) \to 0,\ t \to \infty$



If $\lambda \in \mathbb{C} \Rightarrow y(t) = e^{\lambda t} = e^{Re(\lambda)t} \cdot \left(\cos(Im(\lambda)t) + i \sin(Im(\lambda)t)\right)$

and, since $Re(\lambda)<0 \Rightarrow \begin{matrix} Re(y(t)) \xrightarrow[t\to\infty]{} 0 \\ Im(y(t)) \to 0 \end{matrix}$

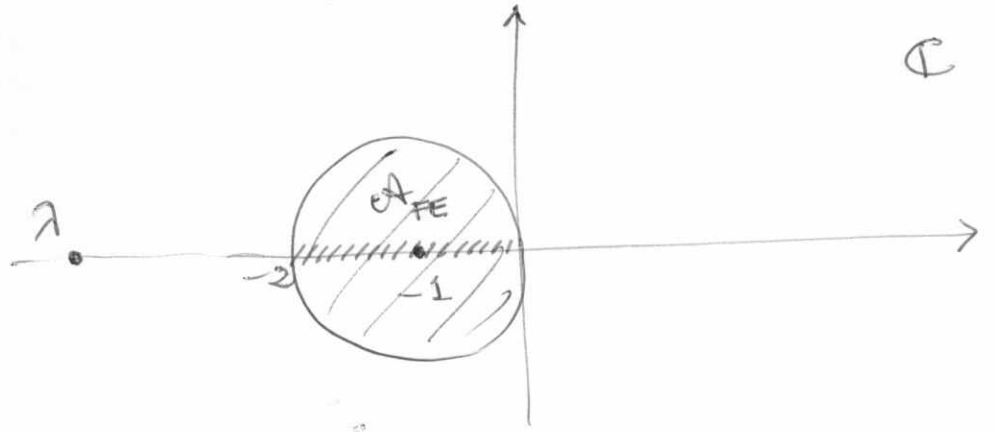
We say that a numerical scheme is abs stable if the numerical sol (approximating the solution of (1)) $u_n \to 0$ when $t \to \infty$, that is it behaves like the exact sol when $t \to \infty$.

Some schemes are abs stable $\forall h>0$, others only for $h \in A \subset \mathbb{R}^+$

Def $\mathcal{A} = \{ z \in \mathbb{C} : z = h\lambda \text{ s.t. } u_n \to 0,\ t_n \to \infty \}$
is called absolute stability region.

Given λ (it is a problem datum), $\mathcal{A}$ says us which values of h guarantee abs. stab.

EX: FE $\quad \mathcal{A}_{FE} = \{ z \in \mathbb{C} : |1+z| < 1 \}$



EX 1 $\begin{cases} y' = -3y \\ y(0) = 1 \end{cases}$ $\qquad \lambda = -3$

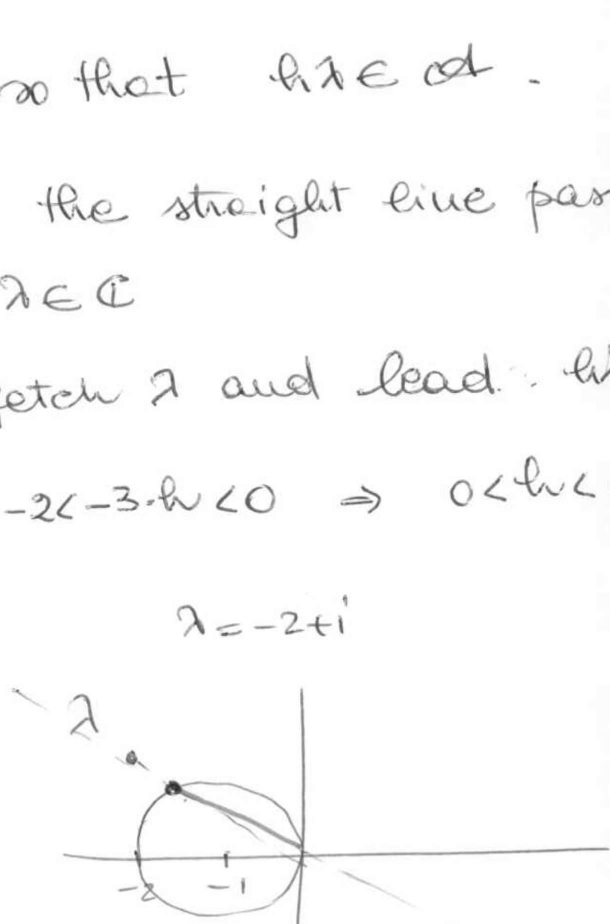
h has to be chosen so that $h\lambda \in \mathcal{A}$.

$z = h\lambda$ ~~must~~ belong to the straight line passing from $(0,0)$ and $\lambda \in \mathbb{C}$

It's up to h to fetch λ and lead $h\lambda \in \mathcal{A}$.

FE is abs stab if $-2 < -3 \cdot h < 0 \Rightarrow 0 < h < \frac{2}{3}$

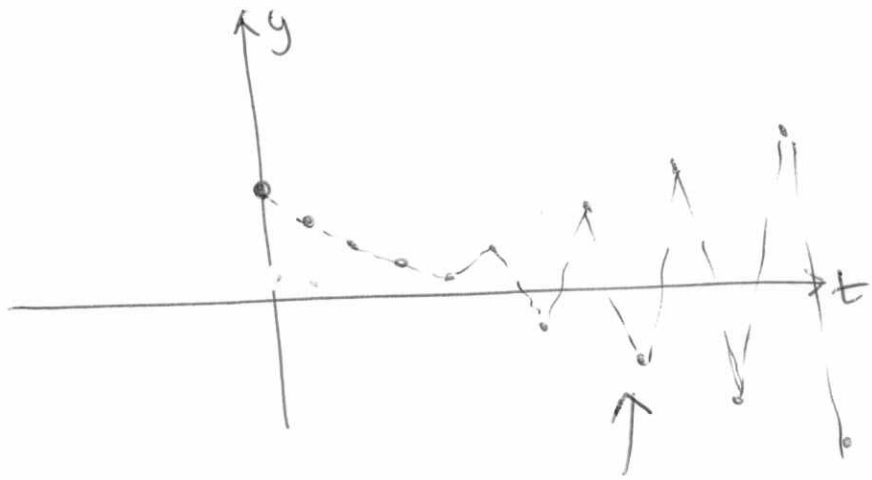
EX 2 $\quad y' = (-2+i)y \qquad \lambda = -2+i$



$h\lambda \in \mathcal{A}_{FE}$
iff $\quad 0 < h < \dfrac{-2\,\mathrm{Re}(\lambda)}{|\lambda|^2}$

(recall that $\mathrm{Re}(\lambda) < 0$)

When $h\lambda \notin \mathcal{A} \Rightarrow$ oscillations arise



unstable solution.

Both BE and CN are absolutely stable $\forall h > 0$.

and $\mathcal{A}_{BE} = \{ z \in \mathbb{C} : |z-1| > 1 \}$

$$\mathcal{A}_{CN} = \{ z \in \mathbb{C} : Re\, \lambda < 0 \}$$

If a scheme is absolutely stable $\Rightarrow$
it is zero-stable, too

Absolute stability is [illegible]
than zero-stab.

Systems of ODE
and ODE of order $p > 1$.

? $x = x(t)$

Ex 1: $(P)\begin{cases} 3x'' - 5x' + 6x = \cos t \\ x(0) = 1 \\ x'(0) = 3 \end{cases}$ order $p=2$, $t_0 = 0$

Rem 1 we need p initial conditions on $x, x', \dots, x^{(p-1)}$ in t_0.

We convert the second-order equation in a system of first-order eqs:

$\underline{y}(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}$ with $y_1(t) = x(t)$, $y_2(t) = x'(t) = y_1'(t)$

$\Rightarrow x''(t) = y_2'(t)$

and problem (P) becomes

$$\begin{cases} 3y_2'(t) - 5y_2(t) + 6y_1(t) = \cos t \\ y_1'(t) = y_2(t) \\ y_1(0) = 1 \\ y_2(0) = 3 \end{cases} \qquad \begin{cases} y_1'(t) = y_2(t) \\ y_2'(t) = -2y_1(t) + \frac{5}{3}y_2(t) + \cos t \\ y_1(0) = 1 \\ y_2(0) = 3 \end{cases}$$

$$\begin{cases} \underline{y}'(t) = \underline{F}(t, \underline{y}(t)) = \begin{bmatrix} y_2(t) \\ -2y_1(t) + \frac{5}{3} y_2(t) + \cos t \end{bmatrix} \\ \underline{y}(0) = \underline{y}_0 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \end{cases}$$

We can apply BE, FE, CN to the vectorial problem

$$\begin{cases} \underline{y}'(t) = \underline{F}(t, \underline{y}(t)) & t > t_0 \\ \underline{y}(t_0) = \underline{y}_0 \end{cases}$$

FE: $\underline{u}_{n+1} = \underline{u}_n + h\, \underline{F}(t_n, \underline{u}_n)$

BE $\underline{u}_{n+1} = \underline{u}_n + h\, \underline{F}(t_n, \underline{u}_{n+1})$

CN $\underline{u}_{n+1} = \underline{u}_n + \frac{h}{2} \left[\underline{F}(t_n, \underline{u}_n) + F(t_{n+1}, \underline{u}_{n+1}) \right]$

---

the vectorial counterpart of $y' = \lambda y$ is $\underline{y}' = A \cdot \underline{y}$ and $\lambda_i(A)$ eigenvalues of A play the role of λ in the study of abs stab.

In order to guarantee abs stab, the conditions $h \lambda_i(A) \in \mathcal{A}$ have to be satisfied $\forall\, i = 1, \dots, P$ (P = dimension of the system).