

Obyčejné diferenciální rovnice

J. Schmidt

Datum poslední aktualizace: 6. února 2024

Text je zatím jen souhrn poznámek a nemá dostatečnou koherenci. Některé odstavce jsou plně popsány, jinde jsou rovnice uvedeny bez poznámek a jinde chybí třeba doplnit obrázky. Časem budou nejasné odstavce přepsány a do nutných míst doplněny obrázky nebo text.

Obsah

| | | |
|---|---------------------------------|----|
| 1 | Dopředný a zpětný Euler | 3 |
| 2 | Verletova metoda | 9 |
| 3 | Rungeova-Kuttova metoda | 12 |
| 4 | Exponenciální řešiče | 15 |
| 5 | Newmarkova metoda | 15 |
| 6 | Motivační příklad | 15 |
| 7 | Diskrétní Lagrangeova mechanika | 17 |
| 8 | Konkrétní integrátory | 18 |

V předcházející kapitole byl představen diferenciální způsob popisu fyzikálních dějů a představen způsob, jakým lze nalézt analytické přesné řešení. Ne vždy je to ovšem možné, zejména pokud popisujeme nelineární děje. V tom případě nastupují na scénu numerické metody. Pojd'me si nejprve představit ty nejzákladnější integrátory spolu s analýzou jejich vlastností.

1 Dopředný a zpětný Euler

Fyzikální systémy mohou mít různý tvar, ale ukazuje se, že v praktických úlohách lze velmi často problémy převést na soustavu diferenciálních rovnic 1. řádu. V předcházející kapitole jsme se omezovali zejména na soustavy lineárních homogenních rovnic. Takové omezení je nyní již bezpředmětné a uvažujeme proto obyčejnou diferenciální rovnici s obecnou pravou stranou ve tvaru

$$\frac{dy(t)}{dt} = f(y, t). \quad (1)$$

Opět budeme zkoumat nejprve jedinou rovnici. K soustavám přejdeme záhy. Pojd'me tedy zkoumat numerický způsob řešení takového systému. Předpokládejme prvně, že vyšetřovaný časový interval $\langle 0, T \rangle$ diskretizujeme pomocí konstantního časového kroku Δt na soubor po sobě jdoucích diskrétních časových okamžiků $0 = t_0, t_1, \dots, t_n = T$, kde pro každé i platí $t_{i+1} - t_i = \Delta t$. V takovém případě lze postupovat aproximací derivace na levé straně konečnou diferencí, tedy derivaci můžeme aproximovat například jako

$$\frac{dy(t_i)}{dt} \approx \frac{y_{i+1} - y_i}{\Delta t}. \quad (2)$$

Dodejme, že funkce f_i s indexem i je aproximací funkce $f(t_i)$ vyčíslené v čase t_i . Pokud rovnici (1) vyčíslíme v čase t_i a dosadíme výše zmíněnou aproximaci, potom dostáváme

$$\frac{y_{i+1} - y_i}{\Delta t} = f(y_i, t_i), \quad (3)$$

což lze případně přepsat do tvaru

$$y_{i+1} = y_i + \Delta t f(y_i, t_i). \quad (4)$$

Této formě říkáme numerický integrátor, protože se jedná o numerický způsob, jak integrovat onu diferenciální rovnici, v tomto případě v čase. Jedná se o tzv. **Eulerovu dopřednou metodu**. Název pochází z faktu, že derivaci aproximujeme pomocí aktuálního času a času o jeden krok vzdálenějším. Aproximaci lze ale volit více způsoby, další velmi přirozenou aproximací je definice na základě zpětné hodnoty, tedy

$$\frac{dy(t_{i+1})}{dt} \approx \frac{y_{i+1} - y_i}{\Delta t}, \quad (5)$$

čímž dostáváme další přirozený integrátor, tzv. **Eulerovu zpětnou metodu**. Po vyčíslení rovnice (1) v čase t_{i+1} a dosazení výše uvedeného vztahu dostáváme

$$y_{i+1} = y_i + \Delta t f(y_{i+1}, t_{i+1}). \quad (6)$$

Jedná se o implicitní rovnici, jejímž vyřešením dostáváme odhad y_{i+1} . Z tohoto důvodu označujeme tento integrátor za implicitní metodu. Řešení rovnice může být triviální, ale může také vyžadovat nutnost využít další numerickou metodu pro její řešení. Vše závisí na složitosti funkce $f(y, t)$. [TODO: Přidat ukázkový příklad na obě metody a ukázat jejich numerické chování. Tím motivovat další vývoj textu.]

Numerických metod lze odvodit výrazně více různými modifikacemi. Pojd'me ale nejprve zkoumat vlastnosti výše uvedených. Další metody budou uvedeny později.

Uvedené integrátory zůstávají v platnost bez ohledu jak divoká je funkce $f(y, t)$. Problematictější je zkoumání jejich hlubších vlastností jako je stabilita nebo vývoj chyby v čase, tzv. numerický řád integrátoru. Z toho důvodu se opět omezíme na lineární diferenciální rovnici a pro tuto chvíli také vyloučíme explicitní závislost na čase. Dostáváme tedy zjednodušenou verzi rovnice

$$\frac{dy(t)}{dt} = \zeta y(t). \quad (7)$$

Pokud aplikujeme na tento problém dopřednou Eulerovu metodu, dostáváme

$$y_{i+1} = y_i + \Delta t \zeta y_i = (1 + \Delta t \zeta) y_i = A y_i. \quad (8)$$

Jak je vidět, odhad funkce y_{i+1} v čase t_{i+1} lze spočítat jako A násobek hodnoty y_i . Z toho je okamžitě vidět, že lze postupně rekurzivně získávat postupně odhad celé funkce až do požadovaného času. Nastává zde ovšem jeden zásadní problém. Pokud je absolutní hodnota parametru A větší jak jedna, potom velikost řešení v každém kroku narůstá až nade všechny meze. Říkáme tomu **ztráta numerické stability**. Je dobré připomenout, že $A > 1$ nemusí být vždy nežádoucí efekt, protože existují lineární rovnice 1. řádu, jejichž přesné řešení roste nade všechny meze. Zde je ale problematické to, že dochází k růstu i za předpokladu, že se jedná o stabilní diferenciální rovnici. Jak je vidět, parametr $A(\Delta t)$ je funkcí časového kroku a při zvětšování časového kroku dochází i ke zvětšování parametru A . Ztrátě numerické stability lze tak zamezit volbou dostatečně jemného kroku. Jak je vidět, pro zachování stability musí platit

$$|1 + \Delta t \zeta| < 1. \quad (9)$$

Je ihned vidět, že pro $\zeta > 0$ je metoda nestabilní pro libovolně velký krok. To je ovšem přirozené i vzhledem k charakteru původní úlohy. Pro $\zeta < 0$ očekáváme stabilní řešení, které v se v dlouhém čase blíží k nule. To dostáváme ovšem pouze pro krok $\Delta t > \frac{-2}{\zeta}$. V opačném případě dostáváme kvalitativně špatnou odezvu rostoucí nade všechny meze.

| | Integrátor | Stabilita |
|----------------|--|---|
| Dopředný Euler | $y_{i+1} = y_i + \Delta t \zeta y_i$ | $\Delta t \zeta < 0 \wedge \Delta t \zeta > -2$ |
| Zpětný Euler | $y_{i+1} = y_i + \Delta t \zeta y_{i+1}$ | $\Delta t \zeta < 0 \vee \Delta t \zeta > 2$ |

Tabulka 1: Řešení epidemiologické úlohy dopřednou a zpětnou Eulerovou metodou.

Podobným způsobem lze analyzovat taktéž zpětnou Eulerovu metodu, pro kterou dotáváme integrátor

$$y_{i+1} = y_i + \Delta t \lambda y_{i+1}, \quad (10)$$

tedy

$$y_{i+1} - \Delta t \lambda y_{i+1} = y_{i+1} (1 - \Delta t \lambda) = y_i, \quad (11)$$

a nakonec

$$y_{i+1} = (1 - \Delta t \lambda)^{-1} y_i. \quad (12)$$

V tomto jednoduchém případě jsme dokázali analyticky vyřešit implicitní rovnici, tedy nová hodnota je rovna původní krát nějaký parametr závislý na časovém kroku. Opět platí, že nemá-li dojít ke ztrátě numerické stability, musí být tento parametr větší jak jedna, tedy pro zachování stability musí platit následující nerovnost

$$|(1 - \Delta t \lambda)^{-1}| < 1. \quad (13)$$

Výsledky stabilitní analýzy pro tyto dva integrátory jsou shrnuty v tabulce 1. Pro kladnou hodnotu ζ při uvažování dopředného Eulerova je vidět, že musí platit $\Delta t \zeta \in (-2, 0)$, nepočítáme-li vyšetřování rovnice zpětně v čase, potom pro $\Delta t > 0$ neexistuje délka časového kroku, pro kterou by výpočet nerostl nade všechny meze. To není ovšem dáno jen nestabilitou řešiče, ale zejména tím, že pro $\zeta > 0$ roste nade všechny meze i analytické řešení. Nemůžeme tedy od numerického řešení požadovat něco jiného. Z toho plyne, že podmínku pro numerickou stabilitu má smysl vyšetřovat jen pro případy, kdy analytické řešení zůstává s rostoucím časem omezené, tedy pro $\zeta < 0$. To je důvod proč říkáme o dopředné Eulerově metodě, že je **podmíněně stabilní**, protože stabilita nastává jen pro $\Delta t < -2/\zeta$. Podobnou podmíněnost dostáváme i pro zpětnou Eulerovu metodu, nicméně pokud se omezíme na $\zeta < 0$, potom je metoda stabilní pro libovolnou hodnotu Δt . Z toho důvodu mluvíme o metodě **podmíněně stabilní**. Zajímavé je, že pro $\zeta > 0$ je samotná diferenciální rovnice a její analytické řešení nestabilní - roste nade všechny meze, přestože numerický integrátor má naopak nadměrný útlum a řešení konverguje k nule.

Pojďme se nyní z jedné rovnice přesunout na soustavy obyčejných diferenciálních rovnic 1. řádu. Pokud budeme opět uvažovat lineární systém, potom lze popsat tuto soustavu maticově jako

$$\dot{\mathbf{y}} = \mathbf{A}\mathbf{y}, \quad (14)$$

kde \mathbf{y} je vektor neznámých funkcí a A je matice soustavy. Opět lze zapsat numerické řešení pomocí dopředné a zpětné Eulerovi metody jako

$$\text{Dopředný Euler: } \mathbf{y}_{i+1} = \mathbf{y}_i + \Delta t A \mathbf{y}_i, \quad (15)$$

$$\text{Zpětný Euler: } \mathbf{y}_{i+1} = \mathbf{y}_i + \Delta t A \mathbf{y}_{i+1}. \quad (16)$$

Obě tyto metody jdou opět přepsat do explicitního vyjádření

$$\text{Dopředný Euler: } \mathbf{y}_{i+1} = (I + \Delta t A) \mathbf{y}_i, \quad (17)$$

$$\text{Zpětný Euler: } \mathbf{y}_{i+1} = (I - \Delta t A)^{-1} \mathbf{y}_i, \quad (18)$$

kde I je jednotková matice. Zde platí podobný závěr jako u jedné rovnice. Aby byla metoda stabilní, nesmí se vektor \mathbf{y}_i zvětšovat nade všechny meze. Pokud maticí násobíme nějaký vektor, lze to vnímat jako zobrazení, kdy vektoru \mathbf{y}_i přiřazujeme nový vektor $A\mathbf{y}_i$. Jak rozpoznat, zda je nový vektor větší než předchozí? Lze to rozpoznat např. z vlastních čísel. Pokud jsou všechna vlastní čísla menší než 1, potom s jistotou víme, že nový vektor je v jistém smyslu menší než ten předchozí. Pokud je ovšem nějaké vlastní číslo větší než jedna, potom se může stát (ale nemusí), že nový vektor má větší velikost. Opakovanou aplikací tohoto zobrazení může růst řešení nade všechny meze a není tedy stabilní. Veličinu, která popisuje takovou maticovou kvantitu nazýváme spektrální poloměr a značíme ji $\rho(A)$. Ta je rovna právě největšímu vlastnímu číslu matice. Pro stabilní řešení tedy platí [TODO: Doplnit ukázkou toho, že množina tvořící kruh na kterou aplikujeme maticové násobení se stane elipsou. Pro jistotu, že je nové řešení menší musí mít elipsa poloměry menší než jedna. Totéž platí i pro komplexní čísla.]

$$\text{Dopředný Euler: } \rho(I + \Delta t A) < 1, \quad (19)$$

$$\text{Zpětný Euler: } \rho\left((I - \Delta t A)^{-1}\right) < 1. \quad (20)$$

[TODO: Přidat příklad na akademickou úlohu pro FE a BE a ukázat, že i BE někdy selhává, není stále "stabilní".] Pojdme se ještě podívat na chybu generovanou těmito integrátory. Při řešení problému na počítači vzniká chyb vícero druhů. Počítač neumí některá čísla ukládat s libovolnou přesností a z kapacitních důvodů je zaokrouhluje. My nyní nebudeme vyšetřovat zaokrouhlovací chybu, ale budeme se soustředit pouze na distribuci chyby vzniklé naší aproximací derivace. Odvození tzv. lokální chyby není tolik obtížné. V případě dopředného Eulerova jsme využili následující aproximaci dalšího kroku hledané funkce

$$y_{i+1} = y_i + \Delta t f(y_i) \quad (21)$$

Předpokládejme nyní pro jednoduchost, že funkční hodnotu v čase t_i známe přesně a zajímá nás, jak velká vznikne chyba onou aproximací v čase t_{i+1} . Pro funkční hodnotu v čase t_{i+1} můžeme využít rozvoj do Taylorovy řady. Pro ten dostáváme

$$y(t_{i+1}) = y(t_i + \Delta t) = y(t_i) + y'(t_i)\Delta t + \frac{y''(t_i)}{2}\Delta t^2 + \dots \quad (22)$$

Jelikož předpokládáme, že funkci v čase t_i známe přesně, tedy $y_i = y(t_i)$ a navíc platí diferenciální vztah $y' = f(y)$, potom můžeme odečtením přesného řešení (22) a přibližného řešení (21) získat vyjádření chyby

$$y(t_i + \Delta t) - y_{i+1} = \frac{y''(t_i)}{2} \Delta t^2 + \frac{y'''(t_i)}{6} \Delta t^3 + \dots \quad (23)$$

Vzhledem k danému předpokladu se jedná pouze o tzv. **lokální chybu** vznikající při přechodu z jednoho času do dalšího. Vystupuje zde Δt umocněné na různé mocniny, nicméně člen Δt^3 je při malém Δt výrazně převyšující a proto hovoříme lokálně o chybě druhého řádu a značíme ji $\mathcal{O}(\Delta t^2)$. Značíme tím, že pokud zmenšíme krok desetkrát, potom se lokální chyba zmenší stokrát. Ve výsledku nás ovšem zajímá tzv. **globální chyba**, která vyjadřuje přímo rozdíl analytického a numerického řešení. Vyjadřuje fakt, že lokální chyba vzniklá v jednom kroku se v dalším kroku více zvětšuje a kumuluje se. Výpočet globální chyby již není tolik triviální záležitost, proto se v tomto místě uspokojíme s inženýrskou úvahou. Pokud v jednom časovém kroku vzniká chyba úměrná Δt^2 a počet kroků je úměrný Δt (Když krok zmenšíme desetkrát, počet nutných kroků do téhož času se zdesetinásobí), potom by bylo přirozené odhadovat globální chybu řádu $\mathcal{O}(\Delta t)$. Neuvádíme zde důkaz, pouze konstatování, že tomu tak opravdu je a rigorózní rozbor potvrzuje naši intuici.

Pojďme, už v rychlosti, ukázat, že chyba zpětné Eulerovi metody je stejná. Nyní se budeme dívat o krok zpět, v tom případě pro tuto metodu dostáváme aproximaci $y_i = y_{i-1} + \Delta t f(y_i)$, neboli $y_{i-1} = y_i - \Delta t f(y_i)$. Nyní je ale potřeba vyčíslit Taylorův polynom příslušející takovému zpětnému pohledu, tedy

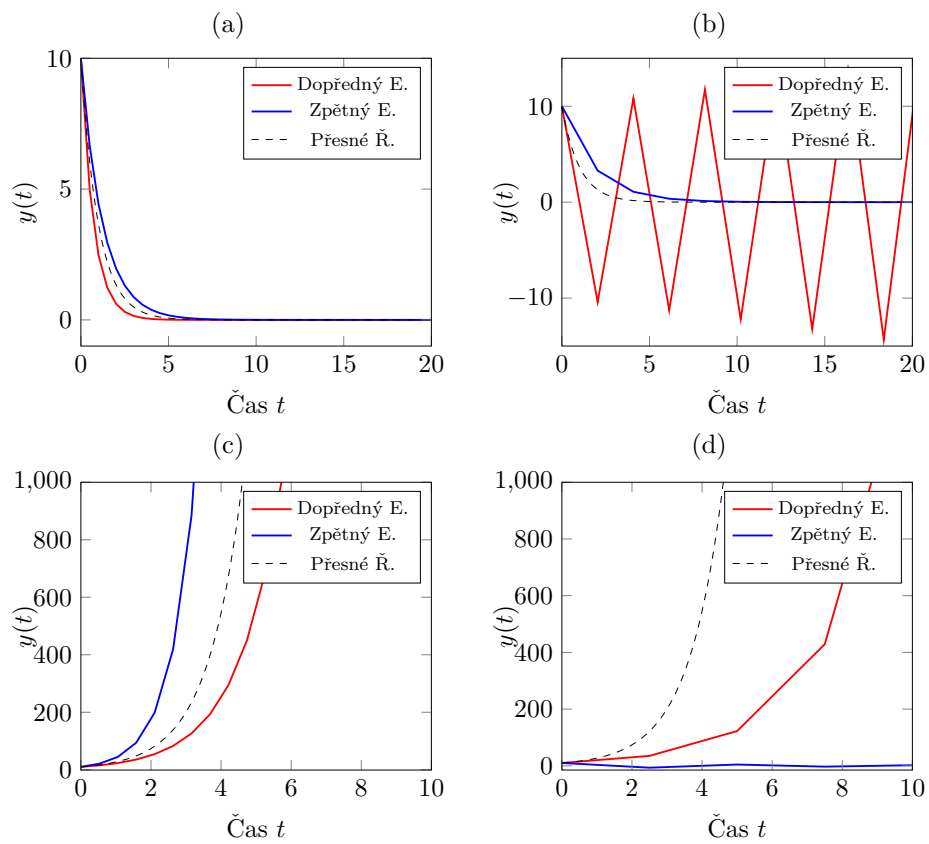
$$y(t_{i-1}) = y(t_i - \Delta t) = y(t_i) - y'(t_i)\Delta t + \frac{y''(t_i)}{2} \Delta t^2 - \dots \quad (24)$$

Nyní stejnou úvahou jako předtím a odečtením aproximace a přesného řešení dostáváme

$$y(t_i - \Delta t) - y_{i-1} = \frac{y''(t_i)}{2} \Delta t^2 - \frac{y'''(t_i)}{6} \Delta t^3 + \dots, \quad (25)$$

tedy můžeme prohlásit, že i zpětná Eulerova metoda je lokálně řádu $\mathcal{O}(\Delta t^2)$ a globálně řádu $\mathcal{O}(\Delta t)$.

Na fyzikálně smysluplný příklad se vrhneme až v následující kapitole, nyní se pojdme podívat na akademickou úlohu $\dot{y}(t) = \pm y(t)$, neboli předpokládáme $\zeta = \pm 1$. Navíc uvažujeme okrajovou podmínku $y(0) = 10$. V obrázku 2 lze potom nalézt odezvy pro dva časové kroky $\Delta t = 0.5$ a $\Delta t = 2.04$ obou těchto systémů. Horní grafy představují systém $\dot{y}(t) = -y(t)$, zatímco spodní jsou simulací systému $\dot{y}(t) = y(t)$. tyto časové kroky nejsou voleny náhodně, protože pro naši úlohu dotáváme kritickou hodnotu časového kroku jako $\Delta t_{\text{crit}} = 2$. Tedy dopředná Eulerova metoda by měla být pro časové kroky $\Delta t \leq \Delta t_{\text{crit}}$ stabilní a pro ostatní časy nestabilní. Opravdu tomu tak je, jak ukazují grafy, nicméně naše stabilitní analýza není tak úplně přesná. Pro rovnici se stabilním analytickým řešením opravdu dostáváme kvalitativně správnou předpověď jen



Obrázek 1: Odezva dopředné a zpětné Eulerovi metody při simulaci rovnice $\dot{y} = -y$ (a,b) a $\dot{y} = y$ (c,d) dva časový krok $\Delta t = 0.5$ (a,c) a $\Delta t = 2$ (b,d).

pro $\Delta t = 0.5$ a pro $\Delta t = 2.04$ se nám řešení rozpadá a začíná divergovat. Pro druhý případ je ale nestabilní předpověď ta správná a dopředná Eulerova metoda zde dává kvalitativně správnou předpověď pro libovolné časové kroky. Oproti tomu zpětný Euler je pro první rovnici stabilní vždy, zatímco pro druhý případ je zde opět správná kvalitativní předpověď jen do $\Delta t \leq \Delta t_{\text{crit}}$. Proč se tedy jedna z metod označuje za podmíněčně stabilní a druhá jako nepodmínečně stabilní? Jde čistě jen o to, že drtivá většina fyzikálních systémů, které chceme vyšetřovat nevykazují exponenciální růst a nevzniká tedy nutnost simulovat i je. Většinou se tedy omezujeme na systémy jako je $\dot{y}(t) = -y(t)$.

2 Verletova metoda

Přesnosti těchto metod tedy nejsou příliš vysoké, pojďme se proto podívat na nějakou numerickou metodu, která dosahuje lepších výsledků. Numerických integrátorů existuje celá škála, představme si alespoň 2 nejnámější a to Rungeova-Kuttova metoda a Verletovu integraci. Začneme s druhou zmíněnou.

Tato metoda se používá pro popis dynamických systémů, které závisejí na rychlostech i zrychleních. Metoda se dá použít i pro širší škálu úloh, nicméně z didaktických důvodů se přidržme těchto systémů. Analýza bude opět provedena nejprve na systému s jedním stupněm volnosti, tedy systémem, který lze formálně zapsat jako

$$\ddot{y}(t) = f(y), \quad (26)$$

kde přirozeně dvě tečky značí druhou derivaci podle času. Základ metody spočívá v diskretizaci členu na levé straně. Využijeme zde diskretizaci tzv. centrálními diferencemi, která říká

$$\frac{d^2 y(t_i)}{dt^2} \approx \frac{y_{i+1} - 2y_i + y_{i-1}}{\Delta t^2}. \quad (27)$$

Tuto aproximaci můžeme rovnou dosadit do naší rovnice vyčíslené v čase t_i a vyjádřit y_{i+1} , čímž dostáváme výsledné schéma, kterému říkáme **Verletova metoda**

$$y_{i+1} = 2y_i - y_{i-1} + \Delta t^2 f(y_i). \quad (28)$$

V tomto přístupu tedy již nestačí znám předchozí krok, ale pracujeme s dvěmi kroky zpětně. Za tuto cenu dostáváme metodu s větší přesností. To lze přímo ukázat hned ze součtu Taylorova rozvoje pro $y(t_i + \Delta t)$ a $y(t_i - \Delta t)$, tedy sečtením rovnic (23) a (23). Dostáváme

$$y(t_i + \Delta t) + y(t_i - \Delta t) = 2y_i + y''(t_i)\Delta t^2 + \frac{y^{(4)}(t_i)}{12}\Delta t^4 + \dots, \quad (29)$$

což vyplývá z toho, že se členy s lichou mocninou odečtou. Pokud opět předpokládáme, že funkční hodnoty v čase t_{i-1} a t_i známe přesně, tedy že $y(t_i - \Delta t) = y_{i-1}$,

$y(t_i) = y_i$ a zároveň víme, že druhá derivace v čase t_i je rovna $f(y_i)$, potom dostáváme

$$y(t_i + \Delta t) - y_{i+1} = \frac{y^{(4)}(t_i)}{12} \Delta t^4 + \dots \quad (30)$$

Naše předpoklady nejsou sice platné, ale jsou dostačující na odhad chyby vznikající v jednom časovém kroku. Představená metoda je tedy lokálně řádu $\mathcal{O}(\Delta t^4)$. Mohlo by se zdát, že globálně tedy dostáváme opět o jeden řád nižší chybu, ale pozor. Tím, že aproximujeme druhou derivaci, tak se řád sníží dvakrát a metoda je ve skutečnosti globálně řádu $\mathcal{O}(\Delta t^2)$.

Tato metoda je známá i v jiné úpravě pod jiným jménem. Pokud rovnici (28) lehce přeskupíme, jednoduše dostaneme

$$\frac{y_{i+1} - y_i}{\Delta t} = \frac{y_i - y_{i-1}}{\Delta t} + f(y_i) \Delta t. \quad (31)$$

Úprava do tohoto tvaru není náhodná, protože ve zlomcích lze nyní vidět aproximaci časové derivace funkce y . Pokud budeme předpokládat, že se jedná o aproximaci zpětným Eulerem a tuto aproximovanou veličinu nazveme rychlostí, potom lze tedy uvažovat

$$v(t_i + \Delta t) \approx v_{i+1} = \frac{y_{i+1} - y_i}{\Delta t}. \quad (32)$$

Pokud tuto úvahu aplikujeme na Verletův systém, potom na něj lze nahlížet jako soustavu dvou numerických rovnic pro dvě neznáme y a v , tedy

$$v_{i+1} = v_i + f(y_i) \Delta t, \quad (33)$$

$$y_{i+1} = y_i + v_{i+1} \Delta t. \quad (34)$$

Tomuto schématu se potom říká **semi-implicitní Eulerova metoda**, nebo také v nějakých zdrojích **symplektická Eulerova metoda**.

Pojďme se ještě podívat na stabilitu - tu budeme vyšetřovat pro ten nejjednodušší lineární systém a to $\ddot{y}(t) = Ay$, kde $A \in \mathbb{R}$. V takovém případě je naše soustava dána jako

$$v_{i+1} = v_i + A \Delta t, \quad (35)$$

$$y_{i+1} = y_i + v_{i+1} \Delta t. \quad (36)$$

Pro komfortní analýzu nejprve rovnice upravíme do lepšího tvaru. Vhodné je do druhé rovnice za v_{i+1} dosadit první rovnici. Po úpravě potom dostáváme maticové vyjádření

$$\begin{pmatrix} y_{i+1} \\ v_{i+1} \end{pmatrix} = \begin{pmatrix} 1 + A \Delta t^2 & \Delta t \\ A \Delta t & 1 \end{pmatrix} \begin{pmatrix} y_i \\ v_i \end{pmatrix}. \quad (37)$$

Analýza stability zůstává stejná. Musíme vyjádřit vlastní čísla a ptát se na jejich velikost. Pro vlastní čísla dostáváme vztah

$$\lambda_{1,2} = \frac{2 + A \Delta t^2}{2} \pm \sqrt{\frac{(2 + A \Delta t^2)^2}{4} - 1}. \quad (38)$$

Tento výsledek žádá drobnou analýzu a diskuzi. Jedná se o výraz typu $\xi \pm \sqrt{\xi^2 - 1}$. Pokud platí, že $|\xi| > 1$, potom je celý výraz pod odmocninou kladný a dostáváme reálná vlastní čísla. Nicméně v takovém případě je jedno z těchto dvou čísel vždy větší než jedna a nemáme tedy zaručené vhodné chování, v tom případě se jedná o nestabilní chování. Naproti tomu pro $|\xi| < 1$ je výraz pod odmocninou záporný a dostáváme tedy dvojici komplexně sdružených kořenů $\xi \pm i\sqrt{1 - \xi^2}$. Ohledně stability je nutné vyřešit jejich velikost. Pokud analýzu vztáhneme zpět na náš případ, potom pro velikost dostáváme vztah

$$|\lambda_1| = |\lambda_2| = \sqrt{\frac{(2 + A\Delta t^2)^2}{4} + 1 - \frac{(2 + A\Delta t^2)^2}{4}} = 1. \quad (39)$$

Pro $\left| \frac{2 + A\Delta t^2}{2} \right| < 1$ tedy dostáváme stabilní řešení. Taková podmínka je splněna právě tehdy když

$$0 \geq \Delta t \geq \sqrt{-\frac{4}{A}}. \quad (40)$$

Nenechme se zmást záporným znaménkem pod odmocninou, pokud je totiž parametr A záporný také, potom pod odmocninou dostáváme bezproblémový výraz, který lze přímo odmocnit. V běžném reálném dynamickém systému se objevuje právě záporný parametr A , jak je vidět na následujícím příkladu.

Řešený příklad

Pojďme analyzovat případ našeho lineárního kyvadla odvozeného v sekci ???. Pro jednotlivé integrátory ihned dostáváme vztahy mezi $\{\theta_i, \xi_i\}$ a $\{\theta_{i+1}, \xi_{i+1}\}$ jako

$$\text{Dopředný Euler: } \xi_{i+1} = \xi_i - \Delta t \frac{g}{l} \theta_i, \quad \theta_{i+1} = \theta_i + \Delta t \xi_i. \quad (41)$$

$$\text{Zpětný Euler: } \xi_{i+1} = \xi_i - \Delta t \frac{g}{l} \theta_{i+1}, \quad \theta_{i+1} = \theta_i + \Delta t \xi_{i+1}. \quad (42)$$

$$\text{Symplektický Euler: } \xi_{i+1} = \xi_i - \Delta t \frac{g}{l} \theta_i, \quad \theta_{i+1} = \theta_i + \Delta t \xi_{i+1}. \quad (43)$$

Opět lze pro obě metody odvodit maticové zobrazení mezi aktuálním a následujícím časem. Pro přehlednost pojďme nejprve analyzovat dopřednou metodu. Pro ní získáváme toto zobrazení maticovým vztahem

$$\begin{pmatrix} \theta_{i+1} \\ \xi_{i+1} \end{pmatrix} = \begin{pmatrix} 1 & \Delta t \\ -\Delta t g/l & 1 \end{pmatrix} \begin{pmatrix} \theta_i \\ \xi_i \end{pmatrix}. \quad (44)$$

Pro vyšetření stability musíme pouze vyšetřit vlastní čísla matice soustavy. Z matice dostáváme charakteristickou rovnici

$$(1 - \lambda)^2 + \Delta t^2 g/l = 0, \quad (45)$$

která má řešení

$$\lambda = 1 \pm i\Delta t\sqrt{g/l}, \quad (46)$$

z kterého lze přímo vypočítat velikosti vlastních čísel. Ty jsou shodné a jsou rovny

$$|\lambda| = \sqrt{1 + \Delta t^2 g/l}. \quad (47)$$

Pokud tedy uvažujeme základní fyzikální konstanty l, g jako kladné, potom je velikost vlastních čísel větší než jedna pro libovolnou volbu délky časového kroku $\Delta t > 0$. Jak bylo uvedeno analytické řešení matematického kyvadla je periodická funkce, která nemění svoji amplitudu. Řešení dopřednou Eulerovou metodou neumí toto chování předpovídat a amplitudu v každé periodě systematicky zvětšuje a postupně diverguje, jedná se tedy o nestabilní chování. Duální chování má zpětná Eulerova metoda. Ta je dána zobrazením

$$\begin{pmatrix} \theta_{i+1} \\ \xi_{i+1} \end{pmatrix} = \begin{pmatrix} 1 & -\Delta t \\ \Delta t g/l & 1 \end{pmatrix}^{-1} \begin{pmatrix} \theta_i \\ \xi_i \end{pmatrix} = \frac{1}{1 + \Delta t^2 g/l} \begin{pmatrix} 1 & \Delta t \\ -\Delta t g/l & 1 \end{pmatrix} \begin{pmatrix} \theta_i \\ \xi_i \end{pmatrix}. \quad (48)$$

Matice vpravo je identická s výše uvedenou a její vlastní čísla jsou taktéž

$$\lambda = 1 \pm i\Delta t\sqrt{g/l}, \quad (49)$$

a tedy se zohledněním faktoru před maticí dostáváme pro velikost vlastních čísel vztah

$$|\lambda| = \frac{\sqrt{1 + \Delta t^2 g/l}}{1 + \Delta t^2 g/l} = \frac{1}{\sqrt{1 + \Delta t^2 g/l}}. \quad (50)$$

Zde tedy platí naopak, že pro libovolnou volbu délky časového kroku $\Delta t > 0$ je metoda z výpočetního pohledu stabilní - nedochází k divergenci. Nicméně už od pohledu je vidět, že pro jinou délku kroku než nulovou se amplituda řešení bude zmenšovat, což opět není v souladu s analytickým vztahem.

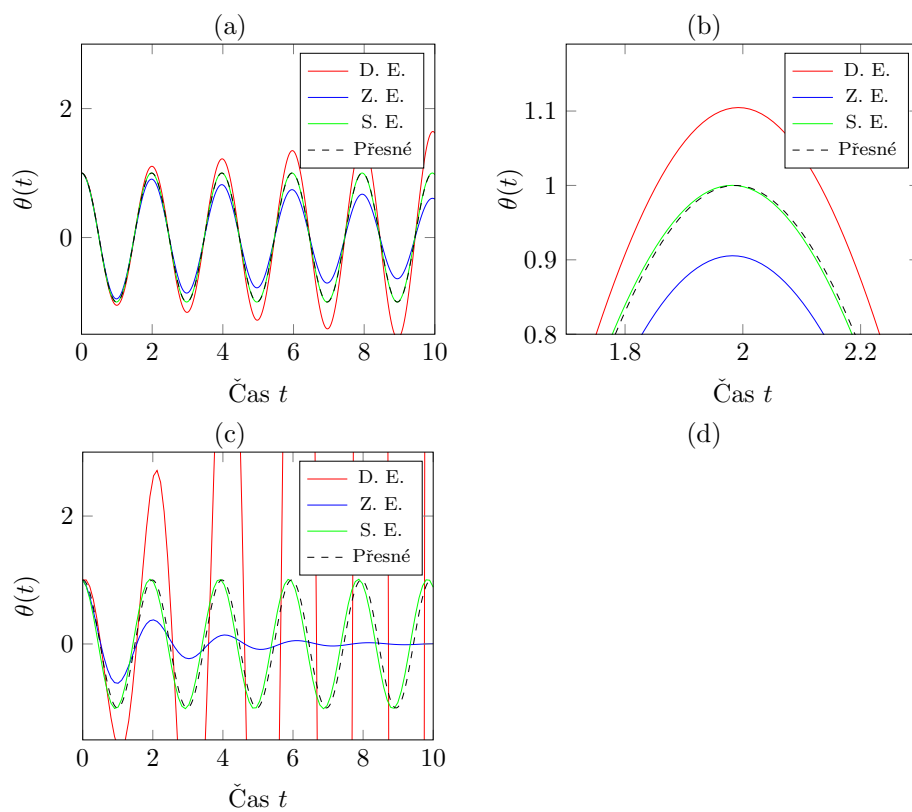
Pojďme ještě vyřešit tzv. symplektickou Eulerovu (Verletovu) metodu, která nabízí kompromis mezi oběma výše zmíněnými metodami. Zde dostáváme po lehké algebraické úpravě následující maticové zobrazení

$$\begin{pmatrix} \theta_{i+1} \\ \xi_{i+1} \end{pmatrix} = \begin{pmatrix} 1 - \Delta t^2 g/l & \Delta t \\ -\Delta t g/l & 1 \end{pmatrix} \begin{pmatrix} \theta_i \\ \xi_i \end{pmatrix}. \quad (51)$$

Stabilitu řešiče můžeme opět vyšetřit analýzou velikosti vlastních čísel této matice. Případně můžeme využít už odvozený vztah (40), kde je pro náš případ $A = -g/l$. Stabilní časový krok je tedy ten, pro který platí $\Delta t < 2\sqrt{l/g}$.

3 Rungeova-Kuttova metoda

Celou dobu se bavíme o diferenciálních rovnicích typu $\dot{y}(t) = f(y(t), t)$. Tedy rovnicích, jejichž funkční hodnota v každém čase je přímo rovna časové derivaci.



Obrázek 2: Odezva dopředné, zpětné a symplektické Eulerovi metody při simulaci matematického lineárního kyvadla pro dva časové kroky $\Delta t = 0.5$ (a,c) a $\Delta t = 2$ (b,d).

Dopřednou Eulerovu metodu lze také interpretovat tak, že v čase t_i vyčíslením $f(y_i, t_i)$ zjistíme přesně velikost derivace a krok z t_i do t_{i+1} linearizujeme, dostaneme tentýž integrátor. Podobnou úvahu lze použít na zpětnou Eulerovu metodu s tím rozdílem, že sklon funkce, neboli časovou derivaci, nezjišťujeme v t_i , ale v t_{i+1} . Takový přístup lze volně interpretovat tak, že známe y_i a hledáme takové y_{i+1} , abychom linearizací funkce y v okolí času t_{i+1} dostali lineární funkci, která protne bod y_i . Jednodušeji řečeno jdeme zpět v čase.

Tuto interpretaci tady zmiňujeme zejména proto, abychom si uvědomili, že lze tuto myšlenku zobecnit a pracovat s funkčními hodnotami funkce $f(y, t)$ mnohem více pro zpřesnění odhadu sklonu křivku.

Náaledující úvaha je tedy legitimní: Pokud funkce $f(y_i, t_i)$ říká, jaký sklon má tečna ke grafu $y(t_i)$, potom $f(y_{i+1}, t_{i+1})$ říká totéž akorát pro následující čas. Jelikož se přesouváme z času t_i do času t_{i+1} je legitimní odhadovat průměrný sklon jako $\frac{1}{2}(f(y_i, t_i) + f(y_{i+1}, t_{i+1}))$. Díky tomu lze integrátor sestavit jako

$$y_{i+1} = y_i + \frac{\Delta t}{2} (f(y_i, t_i) + f(y_{i+1}, t_{i+1})). \quad (52)$$

Takovému schématu říkáme **lichoběžníkové pravidlo**. Tuto metodu lze opět analyzovat podobným způsobem jako předchozí. Všechny principy už byly představeny, proto už zde nebudeme vypisovat každý krok, pouze konstatujeme, že lichoběžníková metoda má globálně chybu $\mathcal{O}(\Delta t^2)$ a je nepodmínečně stabilní.

Podobnou úvahou lze rovnou navrhnout následující integrátor

$$y_{i+1} = y_i + \Delta t f\left(y_i + \frac{\Delta t}{2} f(y_i, t_i), t_i + \frac{\Delta t}{2}\right), \quad (53)$$

který se v literatuře označuje jako tzv. **Midpoint metoda** (nebudeme se zde pouštět do českého překladu tohoto názvu). Metoda je opět řádu $\mathcal{O}(\Delta t^2)$. Stabilita už je obtížnější, zde se omezíme pouze na konstatování, že se jedná o podmíněně stabilní metodu. Je také vidět, že funkce f je nejprve vyčíslena v jednom bodě a posléze na základě této hodnoty v dalším bodě. To je také důvod, proč této metodě říkáme **dvoukroková**. Takové vícekrokové schéma je dobré vhodně interpretovat. V naše případě nejprve vyčísleme $f(y_i, t_i)$, tedy zjistíme sklon tečny v čase t_i . Posléze pomocí této linearizace nalezneme řešení $y_{i+1/2}$ v polovičním čase $t_{i+1/2} = \Delta t/2$ a posléze odhadneme sklon tečny v čase $t_{i+1/2}$. S tímto sklonem nakonec vypočteme funkční hodnotu f_{i+1} . [TODO: Lépe vysvětlit]

S pomocí představeného konceptu lze vytvořit velmi mnoho rozličných integrátorů. Pro úplnost zde uvedeme alespoň ten nejznámější a tím je Runge-Kutta metoda řádu 4, často označovaná jako **RK4** nebo jako klasická Rungeova-Kuttova metoda. Integrátor vypadá následovně

$$y_{i+1} = y_i + \frac{\Delta t}{6} (k_1 + 2k_2 + 2k_3 + k_4), \quad (54)$$

kde

$$k_1 = f(t_i, y_i), \quad (55)$$

$$k_2 = f\left(t_i + \frac{\Delta t}{2}, y_i + \Delta t \frac{k_1}{2}\right), \quad (56)$$

$$k_3 = f\left(t_i + \frac{\Delta t}{2}, y_i + \Delta t \frac{k_2}{2}\right), \quad (57)$$

$$k_4 = f(t_i + \Delta t, y_i + \Delta t k_3). \quad (58)$$

Opět se můžeme pokusit o interpretaci: Člen k_1 označuje sklon v čase t_i , zatímco k_2 označuje sklon v čase $t_{i+1/2}$ s odhadem $y_{i+1/2}$ pomocí k_1 . Sklon v čase $t_{i+1/2}$ je poté odhadnut ještě jednou, tentokrát ale s pomocí sklonu k_2 . Nakonec je za pomocí sklonu k_3 dopočtena hodnota y_{i+1} . Otázka zůstává, který z těchto členů je nejpravděpodobnější, z toho důvodu je ze všech čtyř vytvořen vážený průměr a takto je spočten skutečný odhad y_{i+1} . Chyba metody je lokálně $\mathcal{O}(\Delta t^5)$ zatímco globálně $\mathcal{O}(\Delta t^4)$.

4 Exponenciální řešiče

5 Newmarkova metoda

[Odtud dále jen nakopírováno z jiné kapitoly, je nutné opět vsadit do pořádného kontextu.]

6 Motivační příklad

Případ obyčejného kyvadla je krásný příklad toho, kdy Lagrangián má krásný jednoduchý tvar, taktéž řídicí diferenciální rovnice a přesto ty nejjednodušší interátory selhávají.

Pro sestavení Lagrangiánu nám stačí sestavit kinetickou a potenciální energii. Ty lze odvodit přímo z geometrie problému a mají tvar

$$\mathcal{K}(\dot{\theta}) = \frac{1}{2}ml^2\dot{\theta}^2, \quad \mathcal{U}(\theta) = mgl(1 - \cos \theta), \quad (59)$$

tedy výsledný Lagrangián je roven

$$\mathcal{L}(\theta, \dot{\theta}) = \frac{1}{2}ml^2\dot{\theta}^2 - mgl(1 - \cos \theta). \quad (60)$$

Je vhodné zmínit, že z důvodu zvyklostí je v tomto příkladu zobecněná souřadnice q nahrazena symbolem pro pootočení θ . Pro odvození řídicích diferenciálních rovnic stačí dosadit do Euler-Lagrangeovy rovnice (??) čímž dostáváme

$$mgl \sin \theta + ml^2\ddot{\theta} = 0. \quad (61)$$

Pro potřeby následujících úvah zavedeme ještě zobecněnou hybnost jako

$$p_\theta = ml^2\dot{\theta}. \quad (62)$$

Za použití této hybnosti můžeme diferenciální rovnici 2. řádu převést na soustavu dvou rovnic 1. řádu, tedy

$$\dot{p} = -mgl \sin \theta, \quad (63)$$

$$\dot{\theta} = \frac{p}{ml^2}. \quad (64)$$

Případně lze soustavu dvou rovnic 1. řádu odvodit jednoduchým zavedením nové funkce $\xi = \dot{\theta}$, v této variantě dostáváme

$$\dot{\xi} = -\frac{g}{l} \sin \theta, \quad (65)$$

$$\dot{\theta} = \xi. \quad (66)$$

Oba způsoby uvádíme proto, aby byla jasná fyzikální interpretace jednotlivých veličin a lze tím také nahlédnout, že obě soustavy jsou ekvivalentní. Liší se jen škálováním, ale principiálně jsou tytéž. Uveďme ještě pro úplnost, že ξ má fyzikální význam úhlové rychlosti.

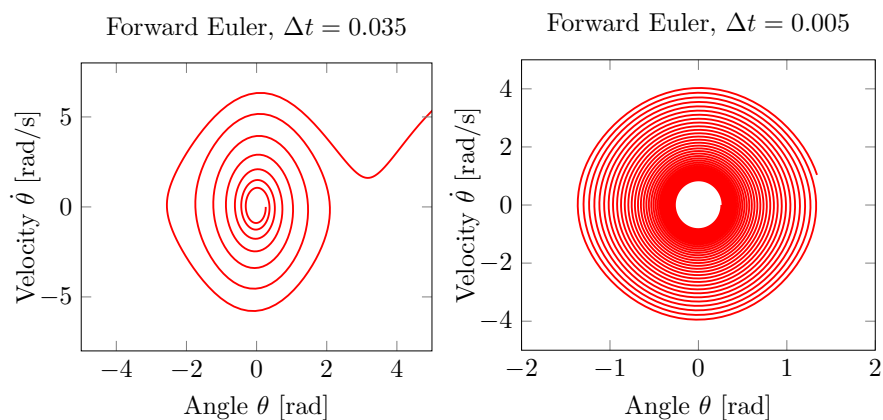
Pokusme se nyní tento problém vyřešit. Přesně řešení není vždy možné odvodit a proto ke slovu přicházejí numerické metody, které nám umožní vyšetřit přibližné řešení v diskrétních časových okamžicích t_n . Říkáme jim také integrátory, protože umožňují numericky zintegrovat diferenciální rovnice a získat přibližné řešení. Integrátorů existuje celá řada. Z didaktických důvodů začneme vůbec tím nejjednodušším, kterým je dopředná Eulerova metoda. Ta říká, že pokud máme soustavu diferenciálních rovnic ve tvaru $\dot{y}(t) = f(t)$, potom odhad řešení $y_n \approx y(t_n)$ v čase t_n lze spočítat rekurzivně jako $y_{n+1} = y_n + \Delta t f(t_n)$, kde Δt je zvolený časový krok. Pro případ našeho kyvadla tedy dostáváme

$$\xi_{n+1} = \xi_n - \Delta t \frac{g}{l} \sin \theta_n, \quad (67)$$

$$\theta_{n+1} = \theta_n + \Delta t \xi_n. \quad (68)$$

Tento příklad slouží pouze jako motivace, ale pro přesnost uveďme parametry úlohy a její počáteční podmínky. Uvažujeme $g = 10, l = 1, m = 1$ a dále $\theta_0 = 0.26$ a $\xi_0 = 0$. Pro zhodnocení výkonnosti integrátoru je lepší se dívat na tzv. fázový portrét radši než na vývoj výsledku v čase. Jedná se o graf ukazující vzájemný vztah zobecněné souřadnice a zobecněné rychlosti. Takový fázový portrét pro 2 časové kroky $\Delta t = 0.035$ a $\Delta t = 0.005$ je zobrazen v grafu 3.

Jak je vidět integrátor neodvádí úplně dobrou práci, protože předpovídá postupné zvětšování jak úhlu natočení kyvadla, tak i zvětšování jeho rychlosti. Řešení lze samozřejmě porovnávat s přesným řešením, ale ani to není potřeba. Z intuice víme, že kyvadlo s počátečním vychýlením θ_0 se nemůže vychýlit o více než $\pm\theta_0$ z nulové polohy. Daný integrátor vytvořit tedy jakési umělé perpetuumobile a systém rozhodně není v rovnováze. Do systému přitéká numericky



Obrázek 3: Fázový portrét získaný dopřednou Eulerovou metodou zobrazeno pro 2 různé časové kroky.

energie. Co když se tedy pokusíme upravit náš integrátor? Udělejme velmi jednoduchou úpravu - nejprve aktualizujeme rychlost ξ a tuto aktualizovanou hodnotu ihned využijme pro aktualizaci úhlu θ , tedy

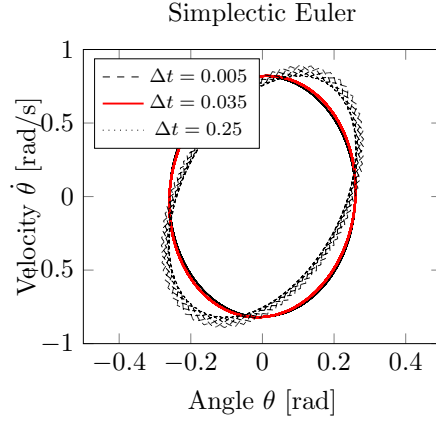
$$\xi_{n+1} = \xi_n - \Delta t \frac{g}{l} \sin \theta_n, \quad (69)$$

$$\theta_{n+1} = \theta_n + \Delta t \xi_{n+1}. \quad (70)$$

Tento integrátor dává o poznání lepší výsledky, viz obrázek 4. Samozřejmě, že chyba daná zmenšováním časového kroku se musí někde projevit. Na grafech je vidět, že fázový portrét se pootočí a lehče zploští. Nicméně charakter řešení není poškozen a nedochází k disipaci nebo přebytku energie. Nicméně otázka zůstává, zda je toto tvrzení opravdu pravdivé. Nemůže dojít v dlouhých časech k téměř negativnímu efektu? A jak vlastně upravovat integrátory nebo jak je vybírat tak, aby byly nějaké vlastnosti řešení zachovány. Na tuto otázku odpovídá následující kapitola, kde je představeno odvozování integrátorů založených na Lagrangeově mechanice.

7 Diskrétní Lagrangeova mechanika

Numerické metody založené na Lagrangeově mechanice nejsou založeny na diskretizaci řídicích diferenciálních rovnic, ale diskretizován je řídicí integrál akce. V první řadě zmiňme, že zkoumaný časový interval $(0, T)$ je diskretizován na jednotlivé časové okamžiky $t_0 = 0, t_1, \dots, t_{n-1}, t_n = T$. Předpokládejme v následujících, že toto dělení je ekvidistantní, tedy že $\Delta t = t_{i+1} - t_i$ je konstantní pro všechna i . Pro získání variačního integrálu je nejdříve nutné aproximovat



Obrázek 4: Fázový portrét získaný symplektickou Eulerovou metodou zobrazeno pro 3 různé časové kroky.

integrál akce na daném intervalu $\langle t_i, t_{i+1} \rangle$ následovně

$$\int_{t_i}^{t_{i+1}} \mathcal{L} dt \approx \mathcal{L}_d(q_i, q_{i+1}, t_i, t_{i+1}). \quad (71)$$

Představená funkce \mathcal{L}_d je diskrétní verze lagrangiánu, která určuje vlastnosti a typ daného integrátoru. Jednotlivé příklady budou představeny později. Lagrangián nyní není funkcí, ale pouze funkcí skalárů q_i a t_i . Minimalizace je tedy získána obyčejnou derivací podle jednotlivých souřadnic. Diskrétní Euler-lagrangeova rovnice nabývá tvaru

$$\frac{d\mathcal{L}_d(q_i, q_{i+1}, t_i, t_{i+1})}{dq_i} + \frac{d\mathcal{L}_d(q_{i-1}, q_i, t_{i-1}, t_i)}{dq_i} = 0. \quad (72)$$

8 Konkrétní integrátory

Centrální diference. Zmávané centrální diference ze získat za předpokladu uvažování následujícího diskrétního lagrangiánu

$$\mathcal{L}_d(q_i, q_{i+1}, t_i, t_{i+1}) = \Delta t \mathcal{L}\left(q_i, \frac{q_{i+1} - q_i}{\Delta t}, t_i\right). \quad (73)$$

Po dosazení dostáváme diskrétní Euler-Lagrangeovu rovnici

$$\Delta t \left(m \frac{q_{i+1} - 2q_i + q_{i-1}}{\Delta t^2} + \mathcal{U}(q_i) \right) \quad (74)$$

Mid-point schéma. Lze získat za předpokladu uvažování následujícího diskrétního lagrangiánu

$$\mathcal{L}_d(q_i, q_{i+1}, t_i, t_{i+1}) = \Delta t \mathcal{L}\left(\frac{q_i + q_{i+1}}{2}, \frac{q_{i+1} - q_i}{\Delta t}, t_i\right). \quad (75)$$

Po dosazení dostáváme diskrétní Euler-Lagrangeovu rovnici

$$? \tag{76}$$

Lichoběžníkové pravidlo. Lze získat za předpokladu uvažování následujícího diskrétního lagrangiánu

$$\mathcal{L}_d(q_i, q_{i+1}, t_i, t_{i+1}) = \frac{\Delta t}{2} m \left(\frac{q_{i+1} - q_i}{\Delta t} \right)^2 - \frac{\Delta t}{2} (\mathcal{U}(q_i) + \mathcal{U}(q_{i+1})). \tag{77}$$

Po dosazení dostáváme diskrétní Euler-Lagrangeovu rovnici

$$? \tag{78}$$