

Politecnico di Milano - Bovisa  
via La Masa 34, 20156 Milano

Dottorato di Ricerca  
Corso di “dinamica di sistemi multicorpo”

# **Analisi di sistemi di equazioni algebrico-differenziali**

Paolo Mantegazza, Pierangelo Masarati

8 gennaio 2010

# 1 Problemi meccanici in forma algebrico-differenziale

La dinamica dei sistemi liberi è descritta da un sistema di equazioni differenziali; la sua soluzione non presenta particolari problemi. La dinamica dei sistemi vincolati affianca alle equazioni dei sistemi liberi un insieme di equazioni di vincolo di tipo algebrico. La soluzione di questi problemi può essere affrontata in modi molto diversi, la cui convenienza è opportuno valutare caso per caso.

## 1.1 Sistemi liberi

Un tipico sistema meccanico in generale ha la forma differenziale

$$\mathbf{M}(\mathbf{q}) \ddot{\mathbf{q}} = \mathbf{f}(\dot{\mathbf{q}}, \mathbf{q}, t)$$

che descrive la dinamica di un insieme di corpi liberi soggetti a forze dipendenti dalla loro configurazione,  $\mathbf{q}$ , e dalle sue derivate. Esso esprime anche la dinamica del continuo quando la sua configurazione sia descritta in forma discreta.

## 1.2 Sistemi vincolati

Esistono due categorie di vincoli; essi sono detti *olonomi* quando esprimono un legame tra i gradi di libertà del sistema; sono detti *anonomi* quando esprimono un legame tra le derivate dei gradi di libertà del sistema. La presenza di vincoli olonomi comporta l'aggiunta di equazioni algebriche di vincolo

$$\mathbf{0} = \phi(\mathbf{q}, t)$$

a cui corrispondono reazioni vincolari; la presenza di vincoli anonomi comporta l'aggiunta di equazioni differenziali di vincolo

$$\mathbf{0} = \mathbf{A}(\mathbf{q}, t) \dot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, t),$$

tipicamente lineari nelle derivate delle variabili cinematiche, a cui corrispondono, allo stesso modo, reazioni vincolari. Si noti che le equazioni di vincolo, pur essendo differenziali nei gradi di libertà del problema meccanico, rendono comunque il problema algebrico-differenziale, perché non vi compaiono le derivate delle incognite algebriche aggiuntive, che esprimono le reazioni vincolari. Si consideri il caso di vincoli olonomi; con il metodo dei moltiplicatori di Lagrange,

$$\boldsymbol{\lambda} \cdot \phi(\mathbf{q}, t)$$

$$\delta(\boldsymbol{\lambda} \cdot \phi(\mathbf{q}, t)) = \delta\boldsymbol{\lambda} \cdot \phi(\mathbf{q}, t) + \delta\mathbf{q} \cdot \phi_{/\mathbf{q}}^T \boldsymbol{\lambda}$$

si giunge a scrivere

$$\begin{aligned} \mathbf{M}(\mathbf{q}) \ddot{\mathbf{q}} &= \mathbf{f}(\dot{\mathbf{q}}, \mathbf{q}, t) + \phi_{/\mathbf{q}}^T \boldsymbol{\lambda} \\ \mathbf{0} &= \phi(\mathbf{q}, t) \end{aligned}$$

analogamente, per problemi anonomi, si ottiene:

$$\begin{aligned} \mathbf{M}(\mathbf{q}) \ddot{\mathbf{q}} &= \mathbf{f}(\dot{\mathbf{q}}, \mathbf{q}, t) + \mathbf{A}^T \boldsymbol{\lambda} \\ \mathbf{0} &= \mathbf{A}(\mathbf{q}, t) \dot{\mathbf{q}} + \mathbf{b}(\mathbf{q}, t) \end{aligned}$$

### 1.3 Sistemi multicampo

Si giunge spesso alla scrittura di sistemi algebrico-differenziali quando si scrivono problemi multicampo, ovvero che legano tra loro modelli fisici diversi. Si pensi, ad esempio, ad un problema accoppiato, in cui ad una parte meccanica si affianchi un impianto idraulico che consente l'attuazione del dispositivo meccanico, una rete elettrica che descrive il controllo dell'impianto idraulico, il tutto nell'ambito di un sistema aeroelastico in cui la configurazione del sistema meccanico determina le condizioni al contorno del problema aerodinamico. Le reti elettriche ed idrauliche sono spesso scritte in forma di equazioni di bilancio ai nodi (di corrente, di portata) a cui si affiancano i legami costitutivi dei componenti del circuito (capacità, resistenze, induttanze, generatori di corrente e di tensione, o tubi, orifici, accumulatori, generatori di pressione, attuatori, servovalvole). L'interfaccia tra diversi modelli fisici, per ragioni di generalità, può essere più comodamente espressa in termini di equazioni di congruenza tra grandezze omogenee, attraverso vincoli algebrici.

Nell'impostare questi problemi risulta agevole rinunciare a priori a criteri di *causalità*, ovvero a scrivere le relazioni costitutive dei componenti in termini di ingresso-uscita; conviene anzi lasciare le relazioni in forma implicita, imponendo le relazioni più appropriate in termini di vincoli algebrici.

Il risultato è un sistema di equazioni algebrico-differenziali (detto a *tableau sparso*) formalmente analogo al problema della dinamica di sistemi vincolati.

### 1.4 Filosofie di soluzione

Esistono fondamentalmente due filoni filosofici sulle modalità di soluzione dei problemi di questo tipo:

- riduzione ad equazioni differenziali pure;
- soluzione diretta delle equazioni algebrico-differenziali.

Le due modalità sono profondamente differenti, eppure presentano numerosi punti di contatto; la convenienza di una tecnica nei confronti dell'altra deve essere soppesata con cura, anche se nelle applicazioni pratiche la seconda sta prendendo il sopravvento. Queste note, oltre a descrivere le metodologie di soluzione dei problemi, vogliono aiutare a formare la capacità critica di scegliere le tecniche di soluzione più adatte al problema che si deve affrontare.

### 1.5 Eliminazione diretta dei moltiplicatori

La soluzione esatta del problema vincolato comporta la derivazione del vincolo fino al secondo ordine:

$$\begin{aligned} \mathbf{0} &= \dot{\boldsymbol{\phi}}(\mathbf{q}, t) = \boldsymbol{\phi}_{/q} \dot{\mathbf{q}} + \boldsymbol{\phi}_{/t} &= \boldsymbol{\phi}_{/q} \dot{\mathbf{q}} + \mathbf{b} \\ \mathbf{0} &= \ddot{\boldsymbol{\phi}}(\mathbf{q}, t) = \boldsymbol{\phi}_{/q} \ddot{\mathbf{q}} + (\boldsymbol{\phi}_{/q} \dot{\mathbf{q}} + \boldsymbol{\phi}_{/t})_{/q} \dot{\mathbf{q}} + (\boldsymbol{\phi}_{/q} \dot{\mathbf{q}} + \boldsymbol{\phi}_{/t})_{/t} &= \boldsymbol{\phi}_{/q} \ddot{\mathbf{q}} + \mathbf{b}' \end{aligned}$$

si noti che il caso di vincoli anolonomi equivale a quello di vincoli olonomi in cui si parta direttamente dalla derivata dell'equazione di vincolo:

$$\mathbf{0} = \mathbf{A}\ddot{\mathbf{q}} + (\mathbf{A}\dot{\mathbf{q}} - \mathbf{b})_{/q} \dot{\mathbf{q}} + (\mathbf{A}\dot{\mathbf{q}} - \mathbf{b})_{/t} = \mathbf{A}\ddot{\mathbf{q}} + \mathbf{b}'$$

A questo punto le derivate seconde dell'incognita possono essere sostituite nell'equazione di vincolo, e quindi si possono esplicitare i moltiplicatori:

$$\lambda = -(\phi_{/q} \mathbf{M}^{-1} \phi_{/q}^T)^{-1} (\phi_{/q} \mathbf{M}^{-1} \mathbf{f}(\dot{\mathbf{q}}, \mathbf{q}, t) + \mathbf{b}')$$

Sostituendo i moltiplicatori nell'equazione differenziale della dinamica si ottiene un problema puramente differenziale e intrinsecamente rispettoso dei vincoli (Gauss [1], Udwadia & Kalaba [2]):

$$\mathbf{M}(\mathbf{q}) \ddot{\mathbf{q}} = \left( \mathbf{I} - \phi_{/q}^T (\phi_{/q} \mathbf{M}^{-1} \phi_{/q}^T)^{-1} \phi_{/q} \mathbf{M}^{-1} \right) \mathbf{f}(\dot{\mathbf{q}}, \mathbf{q}, t) - \phi_{/q}^T (\phi_{/q} \mathbf{M}^{-1} \phi_{/q}^T)^{-1} \mathbf{b}'$$

Il problema è più complesso nel caso di vincoli non-ideali, ovvero vincoli le cui forze di reazione compiono lavoro per uno spostamento virtuale.

## 1.6 Riduzione del sistema

In seguito a questa operazione al problema libero viene sostituito un problema formalmente delle stesse dimensioni, ma la cui soluzione deve rimanere confinata in un sottospazio dello spazio originario. A questo punto è lecita la domanda: è possibile restringere il problema alla dimensione del sottospazio effettivamente libero? Si giunge a questo risultato utilizzando le equazioni di vincolo per esprimere i gradi di libertà vincolati in funzione di quelli liberi. Si immagini di poter partizionare i gradi di libertà del sistema in liberi e vincolati:

$$\mathbf{q} = \left\{ \begin{array}{l} \mathbf{q}_a \\ \mathbf{q}_o \end{array} \right\} \quad \begin{array}{l} \text{g.d.l. usati nell'analisi} \\ \text{g.d.l. omessi} \end{array}$$

dove, in analogia con la notazione usata in NASTRAN, si sono indicati con *o* i gradi di libertà *omessi*, e con *a* i gradi di libertà mantenuti nell'*analisi*. Si espliciti la derivata prima del vincolo in funzione dei gradi di libertà vincolati:

$$\dot{\mathbf{q}}_o = -\phi_{/q_o}^{-1} (\phi_{/q_a} \dot{\mathbf{q}}_a + \phi_{/t})$$

questo è possibile a patto che la matrice  $\phi_{/q_o}$  sia quadrata (lo deve essere per definizione, in quanto i gradi di libertà che si eliminano sono pari al numero delle equazioni di vincolo) e non-singolare<sup>1</sup>. Quest'ultima condizione indica che i vincoli non devono essere sovrabbondanti e che la matrice di vincolo non deve esprimere una condizione di impuntamento.

---

<sup>1</sup>Se la matrice è singolare i vincoli sono insufficienti o sovrabbondanti. Entrambi i casi hanno un preciso significato fisico, ma l'indeterminazione e/o la sovradeterminazione possono essere eliminate dal punto di vista formale. A volte, tuttavia, dal punto di vista computazionale può essere conveniente non eliminarla

## 1.7 Equazioni di Maggi

Le derivate dei gradi di libertà del sistema possono essere espresse in forma compatta:

$$\dot{\mathbf{q}} = \begin{bmatrix} \mathbf{I} \\ -\phi_{/\mathbf{q}_o}^{-1} \phi_{/\mathbf{q}_a} \end{bmatrix} \mathbf{e} + \begin{Bmatrix} \mathbf{0} \\ -\phi_{/\mathbf{q}_o}^{-1} \phi_{/t} \end{Bmatrix}$$

dove l'ultimo termine a secondo membro contiene il termine noto dell'equazione di vincolo nelle derivate, mentre  $\mathbf{e} = \dot{\mathbf{q}}_a$ . Le derivate seconde delle coordinate libere diventano

$$\ddot{\mathbf{q}} = \mathbf{T}\ddot{\mathbf{e}} + \mathbf{h}(\dot{\mathbf{q}}, \mathbf{q}, t)$$

in cui nel vettore  $\mathbf{h}(\dot{\mathbf{q}}, \mathbf{q}, t)$  è stato confinato tutto quello che non contiene  $\ddot{\mathbf{e}}$ , mentre in  $\mathbf{T}$  è stata racchiusa la parte lineare della trasformazione.

Si scriva ora l'equazione di equilibrio in termini di  $\ddot{\mathbf{e}}$ :

$$\mathbf{M}\mathbf{T}\ddot{\mathbf{e}} = \mathbf{f}(\dot{\mathbf{q}}, \mathbf{q}, t) + \mathbf{h}(\dot{\mathbf{q}}, \mathbf{q}, t) + \phi_{/\mathbf{q}}^T \boldsymbol{\lambda}$$

Il problema viene quindi ridotto all'insieme delle incognite libere (*minimal set*) premoltiplicandolo per la trasposta della matrice  $\mathbf{T}$ ; si noti che tale matrice è intrinsecamente ortogonale alla matrice dei moltiplicatori,  $\phi_{/\mathbf{q}}^T$ :

$$\begin{bmatrix} \mathbf{I} \\ -\phi_{/\mathbf{q}_o}^{-1} \phi_{/\mathbf{q}_a} \end{bmatrix}^T \begin{bmatrix} \phi_{/\mathbf{q}_a}^T \\ \phi_{/\mathbf{q}_o}^T \end{bmatrix} = \phi_{/\mathbf{q}_a}^T - \left( \phi_{/\mathbf{q}_o}^{-1} \phi_{/\mathbf{q}_a} \right)^T \phi_{/\mathbf{q}_o}^T = \mathbf{0}$$

Queste sono le cosiddette *equazioni di Maggi* [3], dal nome del matematico italiano che per primo le ha formulate all'inizio del secolo scorso, più comunemente note come *equazioni di Kane* dal nome del matematico che le ha indipendentemente riscoperte nella seconda metà del XX° secolo. Sono ottimamente illustrate in [4].

Il sistema che si ottiene è costituito dalle equazioni differenziali del secondo ordine nelle coordinate libere, e dalle derivate delle equazioni di vincolo nelle coordinate vincolate:

$$\begin{aligned} \hat{\mathbf{M}}\ddot{\mathbf{q}}_a &= \hat{\mathbf{f}}(\dot{\mathbf{q}}_a, \dot{\mathbf{q}}_o, \mathbf{q}_a, \mathbf{q}_o, t) \\ \dot{\mathbf{q}}_o &= -\phi_{/\mathbf{q}_o}^{-1} (\phi_{/\mathbf{q}_a} \dot{\mathbf{q}}_a + \phi_{/t}) \end{aligned}$$

con  $\hat{\mathbf{M}} = \mathbf{T}^T \mathbf{M} \mathbf{T}$  e  $\hat{\mathbf{f}} = \mathbf{T}^T \mathbf{f}$  più i termini non lineari della trasformazione. Le coordinate vincolate, in realtà, devono essere risolte a partire dalla equazione di vincolo e non dalla sua derivata, per evitare il fenomeno noto come *deriva*; ne consegue che i valori ottenuti per integrazione debbano poi essere corretti risolvendo il problema algebrico.

## 1.8 Minimal set e ortogonalità

Il concetto fondamentale che sta alla base di questa formulazione consiste nel trovare un sottospazio dello spazio delle incognite che sia ortogonale al vincolo, ovvero, con la notazione precedente, tale che  $\mathbf{T}^T \phi_{/\mathbf{q}}^T = \mathbf{0}$ . Il formalismo che porta al calcolo della matrice  $\mathbf{T}$  non è l'unico praticabile; anzi, spesso la scelta dei gradi di libertà da eliminare non è semplice e può portare ad algoritmi poco robusti.

Un metodo robusto ed efficiente si basa sulla decomposizione QR, la quale esprime una generica matrice rettangolare come prodotto di una matrice quadrata *ortogonale*,  $\mathbf{Q}$ , ed una matrice rettangolare,  $\mathbf{R}$ , in forma triangolare superiore. Nel caso in esame, si consideri la matrice rettangolare dei vincoli, che, trasposta, esprime l'effetto delle reazioni vincolari sulle equazioni della dinamica. Applicando la decomposizione QR alla matrice  $\phi_{/q}^T$ , di dimensioni  $n \times m$ , con  $n > m$ , la matrice ortogonale ha dimensione  $n \times n$ , mentre la matrice triangolare è  $n \times m$ , ovvero riporta nella parte superiore,  $\mathbf{R}_1$ , di dimensioni  $m \times m$ , una matrice triangolare, mentre nella parte inferiore,  $\mathbf{R}_2$ , contiene coefficienti nulli:

$$\phi_{/q}^T = \mathbf{Q}\mathbf{R} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \end{bmatrix}$$

Solo la sottomatrice  $\mathbf{Q}_1$ , di dimensioni  $n \times m$ , è effettivamente usata, perché moltiplica la sottomatrice  $\mathbf{R}_1$ , ovvero la parte di  $\mathbf{R}$  non nulla; quindi

$$\phi_{/q}^T = \mathbf{Q}\mathbf{R} = \mathbf{Q}_1\mathbf{R}_1$$

A questo punto è evidente che la sottomatrice  $\mathbf{Q}_2$  rappresenta una conveniente trasformazione che esprime tutti i gradi di libertà in funzione di una sottobase ortogonale ai vincoli; infatti, essendo la matrice  $\mathbf{Q}$  ortogonale, deve valere  $\mathbf{Q}_2^T\mathbf{Q}_1 = \mathbf{0}$ ; ne risulta la scelta ottimale di  $\mathbf{T} = \mathbf{Q}_2$ .

Ad un risultato analogo si arriva attraverso la decomposizione ai valori singolari, o SVD:

$$\phi_{/q}^T = \mathbf{U}\mathbf{S}\mathbf{V}^T = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{bmatrix} \mathbf{V}^T$$

dove  $\mathbf{U}$  e  $\mathbf{V}$  sono matrici ortogonali, di dimensioni rispettivamente  $n \times n$  e  $m \times m$ . La matrice  $\mathbf{\Sigma}$  è diagonale, e contiene i valori singolari della matrice  $\phi_{/q}$ . Essi sono maggiori di 0 se i vincoli sono ben definiti, quindi il rango di  $\phi_{/q}$  è pari a  $m$ . La matrice  $\mathbf{U}_2$  svolge lo stesso ruolo di  $\mathbf{Q}_2$  nel caso precedente.

## 1.9 Riduzione ed eliminazione simultanee

Un approccio che in un certo senso sintetizza i due precedenti è descritto da Finzi [5]. Si partizionino anche in questo caso i gradi di libertà in un set che viene preservato,  $\mathbf{q}_a$ , ed in uno che viene omesso,  $\mathbf{q}_o$ ; le equazioni della dinamica diventano:

$$\begin{aligned} \mathbf{M}_{oo}\ddot{\mathbf{q}}_o + \mathbf{M}_{oa}\ddot{\mathbf{q}}_a &= \mathbf{f}_o + \phi_{/q_o}^T \boldsymbol{\lambda} \\ \mathbf{M}_{ao}\ddot{\mathbf{q}}_o + \mathbf{M}_{aa}\ddot{\mathbf{q}}_a &= \mathbf{f}_a + \phi_{/q_a}^T \boldsymbol{\lambda} \end{aligned}$$

Dalla derivata seconda delle equazioni di vincolo si esplicitino le derivate seconde delle incognite da omettere:

$$\ddot{\mathbf{q}}_o = -\phi_{/q_o}^{-1} \phi_{/q_a} \ddot{\mathbf{q}}_a + \mathbf{h}$$

Dalla equazione della dinamica relativa alle coordinate da omettere si ricavano i moltiplicatori  $\lambda$  attraverso l'inversione della matrice  $\phi_{/q_o}^T$ ; li si sostituisca quindi nell'altra equazione, assieme alle coordinate da omettere ricavate dalla derivata seconda dell'equazione di vincolo:

$$\left( M_{aa} - M_{ao} \phi_{/q_o}^{-1} \phi_{/q_a} - \phi_{/q_a}^T \phi_{/q_o}^{-T} M_{oa} + \phi_{/q_a}^T \phi_{/q_o}^{-T} M_{oo} \phi_{/q_o}^{-1} \phi_{/q_a} \right) \ddot{q}_a = \hat{f}_a$$

Si ottiene così direttamente un'equazione nelle coordinate preservate, avendo eliminato i moltiplicatori. Si noti che la matrice a primo membro, con la notazione usata in precedenza, non è altro che  $\mathbf{T}^T \mathbf{M} \mathbf{T}$ .

## 2 Equazioni differenziali ordinarie

Un sistema di *equazioni differenziali ordinarie* (Ordinary Differential Equations, ODE) in forma esplicita è dato da:

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t)$$

Lo studio delle proprietà di questo tipo di sistemi rappresenta un problema sicuramente interessante, ma, fatto salvo il caso di sistemi lineari a coefficienti costanti, risulta difficile se non impossibile senza ricorrere a metodi numerici. I metodi numerici, d'altra parte, tipicamente non consentono lo studio diretto del sistema, ma semplicemente la soluzione di un problema specifico, dato dall'unione del sistema e da un insieme di condizioni iniziali. Si ottiene così un *problema differenziale ordinario ai valori iniziali* (Initial Value problem, IV):

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t), \quad \mathbf{y}(t_0) = \mathbf{y}_0$$

La sua soluzione è data dalla funzione  $\mathbf{y}(t)$ .

### 2.1 Approssimazione della soluzione

La soluzione di un problema di equazioni differenziali alle condizioni iniziali si dice anche integrazione. Quando non è possibile integrare un sistema di equazioni differenziali in forma chiusa occorre utilizzare metodi numerici di approssimazione della soluzione. Ne esistono numerosi, con proprietà molto diverse. A seconda delle applicazioni è necessario conoscere i criteri in base ai quali operare la scelta del metodo più opportuno. Esistono numerose classificazioni dei metodi; le proprietà fondamentali sono:

- implicito/esplicito
- condizionatamente/incondizionatamente stabile
- a passo singolo/a passo multiplo

Si consideri a titolo di esempio un semplice problema meccanico costituito da un sistema non smorzato e non forzato (i più critici), nella forma

$$m\ddot{x} + kx = 0$$

Si immagini di applicare un metodo di approssimazione alle differenze, in cui

$$\ddot{x} = \frac{x_{k+1} - 2x_k + x_{k-1}}{h^2}$$

Si hanno:

- le differenze centrate se la derivata numerica viene valutata al centro dell'intervallo, ovvero  $\ddot{x} = \ddot{x}_k$ ;
- le differenze in avanti se la derivata numerica viene valutata al passo corrente, ovvero  $\ddot{x} = \ddot{x}_{k+1}$ .

### 2.1.1 Differenze Centrate

Nel primo caso ( $\ddot{x} = \ddot{x}_k$ ) la predizione è immediata:

$$\ddot{x}_k = \frac{x_{k+1} - 2x_k + x_{k-1}}{h^2}$$

da cui

$$m \frac{x_{k+1} - 2x_k + x_{k-1}}{h^2} + kx_k = 0$$

e quindi

$$x_{k+1} = 2x_k - x_{k-1} - h^2 \frac{k}{m} x_k$$

### 2.1.2 Differenze in avanti

Nel secondo caso ( $\ddot{x} = \ddot{x}_{k+1}$ ) l'incognita si trova sia nella derivata che nello stato, quindi si ottiene una formula implicita:

$$\ddot{x}_{k+1} = \frac{x_{k+1} - 2x_k + x_{k-1}}{h^2}$$

da cui

$$m \frac{x_{k+1} - 2x_k + x_{k-1}}{h^2} + kx_{k+1} = 0$$

e quindi

$$\left( \frac{m}{h^2} + k \right) x_{k+1} = \frac{m}{h^2} (2x_k - x_{k-1})$$

È semplice verificare che mentre la prima formula offre una maggiore accuratezza, la sua stabilità è condizionata alla scelta di un opportuno passo inferiore ad un certo valore che dipende da  $m$  e  $k$ . Al contrario la seconda formula, implicita, sebbene meno accurata offre caratteristiche di stabilità che non dipendono dalla dimensione del passo di integrazione.



**Verifica:** la soluzione di una equazione lineare a coefficienti costanti omogenea alle differenze ha la forma  $x_k = \rho x_{k-1}$ , da cui discende  $x_k = \rho^k x_0$ . Perché sia asintoticamente stabile, deve valere la condizione  $|\rho| < 1$ .

Nel caso delle differenze centrate si ottiene:

$$\frac{m}{h^2} \rho^2 = \frac{m}{h^2} (2\rho - 1) - k\rho$$

ovvero

$$\rho = \left(1 - \frac{kh^2}{2m}\right) \pm j\sqrt{1 - \left(1 - \frac{kh^2}{2m}\right)^2}$$

da cui è evidente che, finché la soluzione si mantiene complessa coniugata, ovvero

$$1 - \left(1 - \frac{kh^2}{2m}\right)^2 > 0,$$

$|\rho|$  è unitario, e la precisione è massima: la soluzione non decade in ampiezza, come è giusto che sia in assenza di smorzamento, mentre è presente un certo errore sul periodo, dal momento che  $\rho$  non è reale.

La soluzione cessa di essere complessa per  $h = 2\sqrt{m/k}$ ; a questo punto una delle due radici reali rende  $\rho > 1$  e quindi la soluzione numerica di un problema stabile diventa instabile: la stabilità del metodo di soluzione è quindi condizionata dal passo di integrazione utilizzato.

Nel caso delle differenze in avanti, l'equazione diventa

$$\left(\frac{m}{h^2} + k\right) \rho^2 - 2\frac{m}{h^2} \rho + \frac{m}{h^2} = 0$$

ovvero

$$\rho = \frac{1 \pm j\sqrt{\frac{k}{m}h^2}}{\left(1 + \frac{k}{m}h^2\right)}$$

Si noti che ora la soluzione ha sempre forma complessa coniugata, tranne che per  $h = 0$ ; il modulo di  $\rho$  è sempre minore di 1 tranne che per  $h = 0$ ; quindi, anche la soluzione omogenea di un sistema stabile, ovvero non smorzato, sarà convergente al valore nullo. Questa caratteristica non deve essere vista come assolutamente negativa. Infatti, ciò che tipicamente interessa della soluzione di sistemi dinamici è la soluzione dovuta ad una forzante. Di conseguenza, un metodo di integrazione incondizionatamente stabile che introduca dissipazione algoritmica si limita a cancellare la risposta del sistema a perturbazioni delle condizioni iniziali, mentre il suo effetto sulla soluzione forzata è molto più limitato, e lo è tanto più quanto più il passo di integrazione è piccolo e quindi tanto più  $\rho$  è vicino a 1.

Ciò che rende le differenze in avanti poco appetibili per l'integrazione delle equazioni del moto di sistemi meccanici è essenzialmente il fatto che, per  $h > 0$ , esse introducono una dissipazione anche in sistemi a smorzamento nullo, in quanto  $|\rho| < 1$ .

## 2.2 Metodi espliciti/impliciti

Un metodo è *esplicito* quando la determinazione della soluzione ad un dato istante richiede la valutazione della funzione integranda  $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t)$  solo ad istanti precedenti; altrimenti lo si definisce *implicito*.

Esempio: Eulero esplicito

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h\mathbf{f}(\mathbf{y}_k, t_k)$$

Esempio: Eulero implicito

$$\mathbf{y}_{k+1} = \mathbf{y}_k + h\mathbf{f}(\mathbf{y}_{k+1}, t_{k+1})$$

Si noti come nel secondo caso l'incognita  $\mathbf{y}_{k+1}$  compaia a primo e a secondo membro, e quindi non possa essere valutata direttamente.

## 2.3 Metodi condizionatamente/incondizionatamente stabili

Un metodo è *condizionatamente stabile* se la sua stabilità è garantita solo per passi temporali appartenenti ad un certo intervallo (a partire da zero, come enunciato in seguito); è *incondizionatamente stabile* se la sua stabilità è garantita indipendentemente dalla lunghezza del passo di integrazione. In seguito verranno formulati criteri di stabilità rigorosi che consentono di valutare le proprietà dei metodi di approssimazione della soluzione.

## 2.4 Metodi a passo singolo/multiplo

In genere i metodi si classificano in base al numero di passi all'indietro ai quali il valore della funzione incognita viene utilizzato per calcolarne il valore al passo corrente. I metodi a passo singolo partono dalla conoscenza della funzione all'ultimo passo risolto (o dalle condizioni iniziali se si tratta del primo passo); quelli a passo multiplo richiedono la conoscenza della funzione a più intervalli. Si nota subito che i metodi a passo singolo sono autopartenti, mentre quelli a passo multiplo richiedono un criterio di avviamento.

## 2.5 Metodi lineari a passo multiplo

I tipici metodi lineari a passo multiplo hanno la forma

$$\sum_{j=0,k} \alpha_j \mathbf{y}_{n-j} = h \sum_{j=0,k} \beta_j \mathbf{f}(\mathbf{y}_{n-j}, t_{n-j})$$

con la condizione  $\alpha_0 = 1$  e, per i metodi espliciti,  $\beta_0 = 0$ . Se li si esplicita rispetto al valore dell'incognita al passo corrente:

$$\mathbf{y}_n = - \sum_{j=1,k} \alpha_j \mathbf{y}_{n-j} + h \sum_{j=0,k} \beta_j \mathbf{f}(\mathbf{y}_{n-j}, t_{n-j})$$

In pratica l'incognita viene espressa come combinazione lineare del suo valore e della funzione integranda ai passi precedenti. Varianti interessanti dei metodi a passo multiplo sono i cosiddetti metodi *one-leg* (letteralmente, a “singola gamba”), ovvero:

$$\sum_{j=0,k} \alpha_j \mathbf{y}_{n-j} = h \mathbf{f} \left( \sum_{j=0,k} \beta_j \mathbf{y}_{n-j}, t_n \right)$$

in cui la funzione viene valutata una volta sola nell'istante corrente, utilizzando una combinazione lineare dei valori dell'incognita ai passi precedenti (se si include il valore incognito, ovvero  $\beta_0 \neq 0$ , il metodo risulta implicito).

Un metodo implicito di grande interesse è la regola dei trapezi (metodo di Crank-Nicolson):

$$\mathbf{y}_n = \mathbf{y}_{n-1} + h \frac{\mathbf{f}(\mathbf{y}_n, t_n) + \mathbf{f}(\mathbf{y}_{n-1}, t_{n-1})}{2}$$

che, in versione one-leg, diventa:

$$\mathbf{y}_n = \mathbf{y}_{n-1} + h \mathbf{f} \left( \frac{\mathbf{y}_n + \mathbf{y}_{n-1}}{2}, t_n \right)$$

mantenendo le stesse proprietà di stabilità ed accuratezza. Essa in pratica valuta la funzione  $\mathbf{y}$  al passo  $n$  come il suo valore al passo  $n - 1$  più la proiezione in avanti con pendenza pari alla media delle pendenze ai due estremi del passo.

## 2.6 Metodi a passo singolo

I tipici metodi a passo singolo (oltre ai metodi a passo multiplo con  $k = 1$ , che costituiscono l'anello di collegamento tra le due tipologie) sono i metodi alla Runge-Kutta, che consistono nel definire un certo numero di punti intermedi all'interno del passo in cui la funzione viene calcolata mediante interpolazione di pseudo-valori della derivata, a loro volta calcolati interpolando i valori stimati della funzione. L'approssimazione della soluzione si ottiene da una interpolazione del tipo:

$$\begin{aligned} \mathbf{y}_n &= \mathbf{y}_{n-1} + h \sum_{i=1,s} b_i \mathbf{Y}_i \\ \mathbf{Y}_i &= \mathbf{f} \left( \mathbf{y}_{n-1} + h \sum_{j=1,s} a_{ij} \mathbf{Y}_j, t_{n-1} + c_i h \right) \end{aligned}$$

soggetta alle condizioni  $c_i = \sum_j a_{ij}$ . Si noti che la definizione delle pseudo-derivate è ricorsiva; inoltre:

- se la matrice dei coefficienti  $a_{ij}$  è sub-triangolare inferiore ( $a_{ij} = 0$  per  $j \geq i$ ) il metodo è *esplicito*;
- se è triangolare inferiore ( $a_{ij} = 0$  per  $j > i$ ), il metodo è *diagonalmente implicito* (Diagonally Implicit Runge-Kutta, DIRK); infine,

- se i coefficienti della diagonale sono uguali tra loro, il metodo si dice *singularmente implicito* (Singly Implicit Runge-Kutta, SIRK)<sup>2</sup>.

Nel caso esplicito, le pseudo-derivate ad un dato punto intermedio dipendono solo da quelle valutate a punti precedenti, mentre nei casi impliciti dipendono anche da se stesse e da quelle a punti successivi. Nei casi Diagonally- e Singly-, tuttavia, non c'è dipendenza tra il valore ad un punto intermedio ed i valori in punti successivi; questo porta alcuni vantaggi computazionali. I metodi alla Runge-Kutta si differenziano per i valori che i diversi coefficienti assumono, da cui risultano diverse caratteristiche di stabilità e di accuratezza. Nel caso vengano applicati a sistemi di equazioni algebrico-differenziali, poi, occorre rispettare alcuni vincoli per ottenere un metodo stabile.

I coefficienti dei metodi alla Runge-Kutta di solito vengono rappresentati mediante la cosiddetta matrice di Butcher:

$$\begin{array}{c|cccc}
 c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\
 c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\
 \vdots & \vdots & & \ddots & \vdots \\
 c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\
 \hline
 & b_1 & b_2 & \cdots & b_s
 \end{array}$$

I metodi a passo singolo, come illustrato, richiedono comunque la valutazione della funzione  $\mathbf{f}$  ad istanti intermedi. Se il metodo è esplicito, ne è semplicemente richiesta la valutazione in cascata. Se il metodo è implicito, la soluzione può richiedere la soluzione simultanea del problema a più istanti temporali. Solo nel caso di un algoritmo DIRK, la soluzione, ancorché implicita, avviene a cascata, e quindi la dimensione del singolo problema implicito che viene risolto è pari alla dimensione di  $\mathbf{f}$ .

Un algoritmo a passo singolo può quindi essere più oneroso di uno a passo multiplo. Il fatto di consentire una certa libertà nella scelta dei punti intermedi e dei pesi consente tuttavia una maggiore flessibilità nel progettare le caratteristiche di stabilità ed accuratezza.

## 2.7 Sintesi delle due classi

Si può dare una interpretazione interessante dei metodi alla Runge-Kutta in termini di integrazione discreta per collocazione di una funzione approssimata mediante interpolazione di funzioni elementari, nello spirito del metodo degli elementi finiti. Si esprima la soluzione approssimata mediante interpolazione:

$$\hat{\mathbf{y}}(t) = \sum_{i=0,s} m_i(t) \mathbf{y}_{k-i} + h \sum_{i=0,s} n_i(t) \dot{\mathbf{y}}_{k-i}$$

La soluzione ad un generico passo è data dal teorema fondamentale del calcolo integrale applicato alla funzione integranda:

$$\mathbf{y}_k = \mathbf{y}_{k-s} + \int_{k-s}^k \mathbf{f}(\mathbf{y}, t) dt$$

---

<sup>2</sup> La nomenclatura non è univoca; spesso in letteratura si indicano come SIRK i metodi diagonalmente impliciti in genere, senza distinzione tra i due tipi.

L'integrale può essere approssimato per collocazione:

$$\mathbf{y}_k \cong \mathbf{y}_{k-s} + \sum_{j=1,n} w_j \mathbf{f}(\mathbf{y}(t_j), t_j)$$

Utilizzando l'approssimazione mediante interpolazione della soluzione nel valutare la funzione integranda, si ottiene:

$$\mathbf{y}_k \cong \mathbf{y}_{k-s} + \sum_{j=1,n} w_j \mathbf{f}(\hat{\mathbf{y}}(t_j), t_j)$$

ovvero una forma analoga a quella alla Runge-Kutta ma sviluppata su più passi. I vari metodi noti in letteratura, e tutta una serie di nuovi metodi con caratteristiche intermedie, si ottengono attraverso una opportuna scelta delle funzioni di forma e della tecnica di integrazione numerica.

Questo approccio mostra come i coefficienti dei metodi alla Runge-Kutta possano essere interpretati come il risultato dell'unione di un criterio di integrazione per collocazione e di un criterio di interpolazione della soluzione, ovvero dell'operazione fondamentale che sta alla base del metodo degli elementi finiti.

All'integrazione per collocazione si può sostituire una integrazione alla Galerkin in senso generalizzato, di cui la collocazione rappresenta un caso particolare; in questo modo si ottiene un criterio ancora più generale per il progetto del metodo di integrazione, che fa capo ai metodi dei residui pesati in interpretazione estesa, a cui sono riconducibili i metodi del tipo *elementi finiti nel tempo* (Finite Elements Time Integration, FETI) anche discontinui. Si noti tuttavia che, nelle applicazioni pratiche, l'integrazione analitica delle formule alla Galerkin in senso generalizzato non è quasi mai possibile, per cui si ricorre di nuovo ad integrazioni numeriche riconducibili a tecniche di collocazione.

## 2.8 Accuratezza

Una caratteristica importante dei metodi di approssimazione della soluzione è l'accuratezza. Una definizione intuitiva dell'accuratezza dei metodi lineari è data dall'ordine del polinomio che sono in grado di integrare esattamente. Una misura dell'errore locale commesso durante l'integrazione, a sua volta, è data dal coefficiente del primo termine del resto. Una definizione più rigorosa è data dal limite, al tendere a zero del passo, del rapporto dell'errore compiuto al dimezzarsi del passo stesso, ovvero

$$\lim_{h \rightarrow 0} \log_2 \left( \frac{E(2h)}{E(h)} \right)$$

La convergenza in smorzamento e fase di un insieme di metodi è illustrata in Figura 1, mentre l'errore in smorzamento e fase è illustrato in Figura 2.

## 3 Stabilità lineare

La stabilità lineare di un metodo di approssimazione della soluzione si studia applicando il metodo al semplice problema differenziale lineare scalare:

$$\dot{y} = \lambda y$$

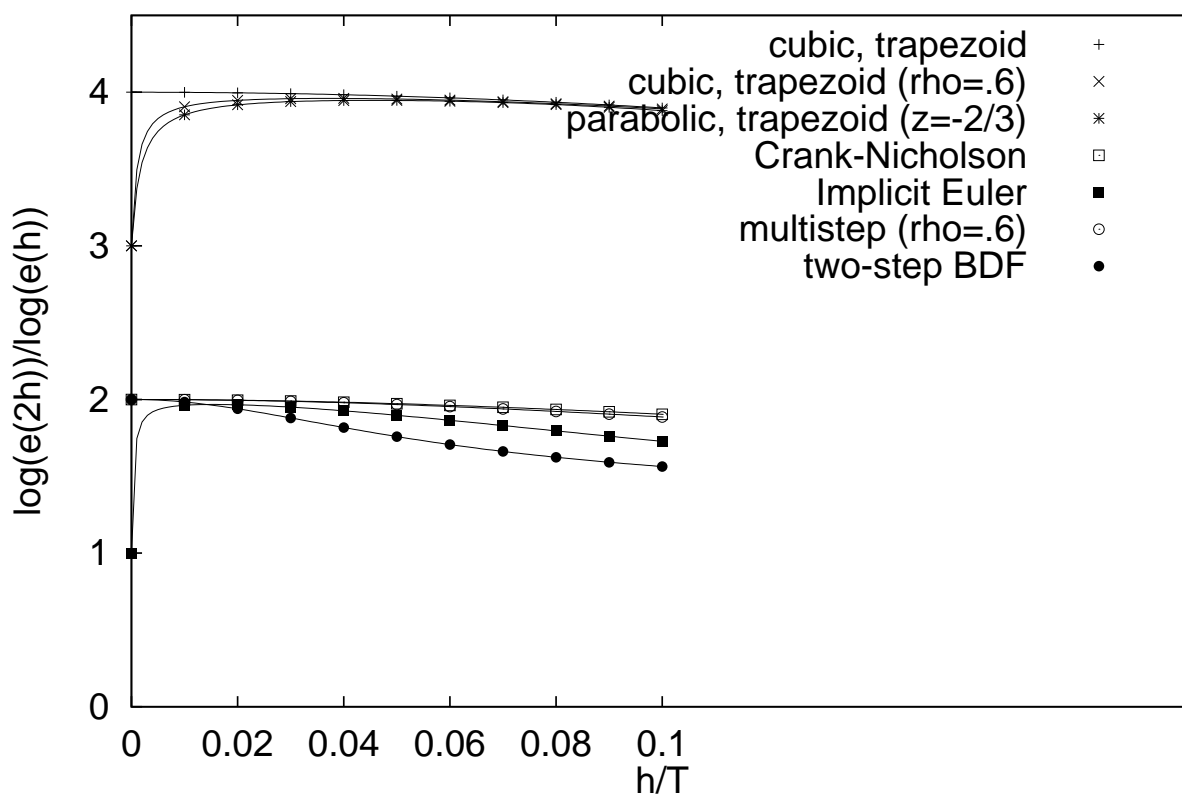
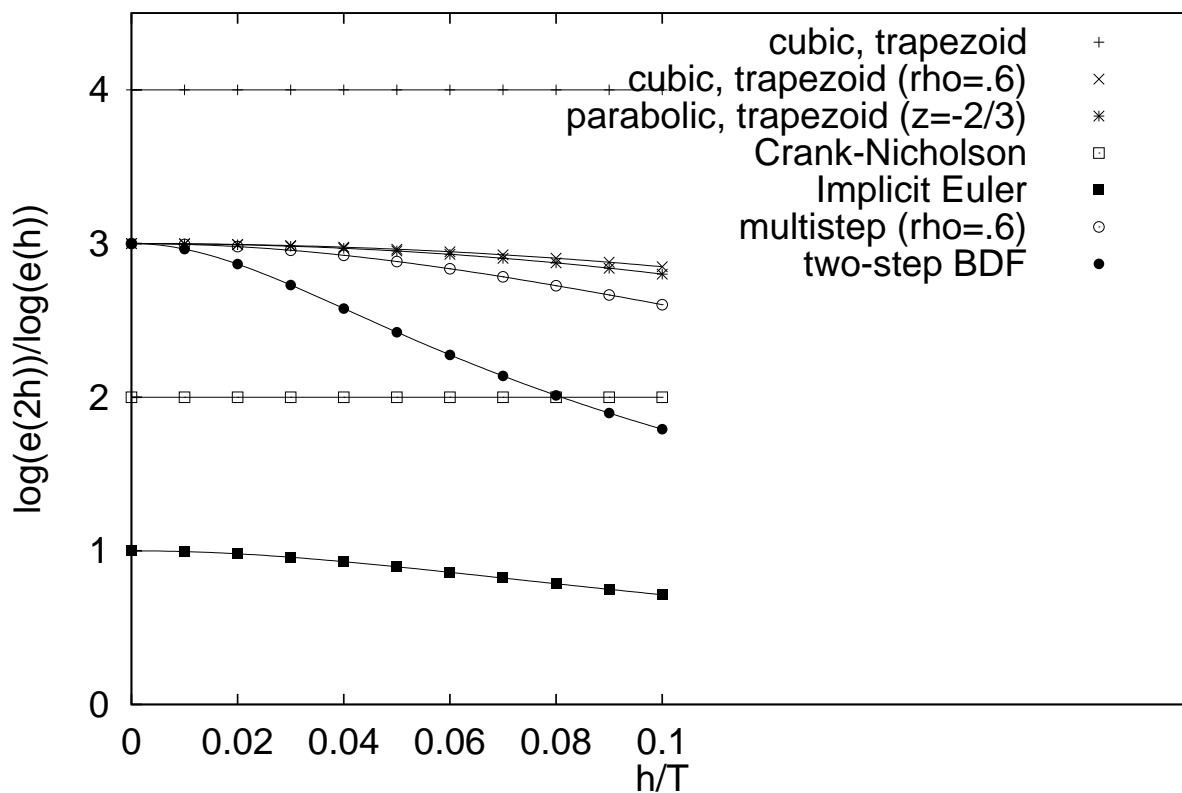


Figura 1: Convergenza in smorzamento e fase di un insieme di metodi

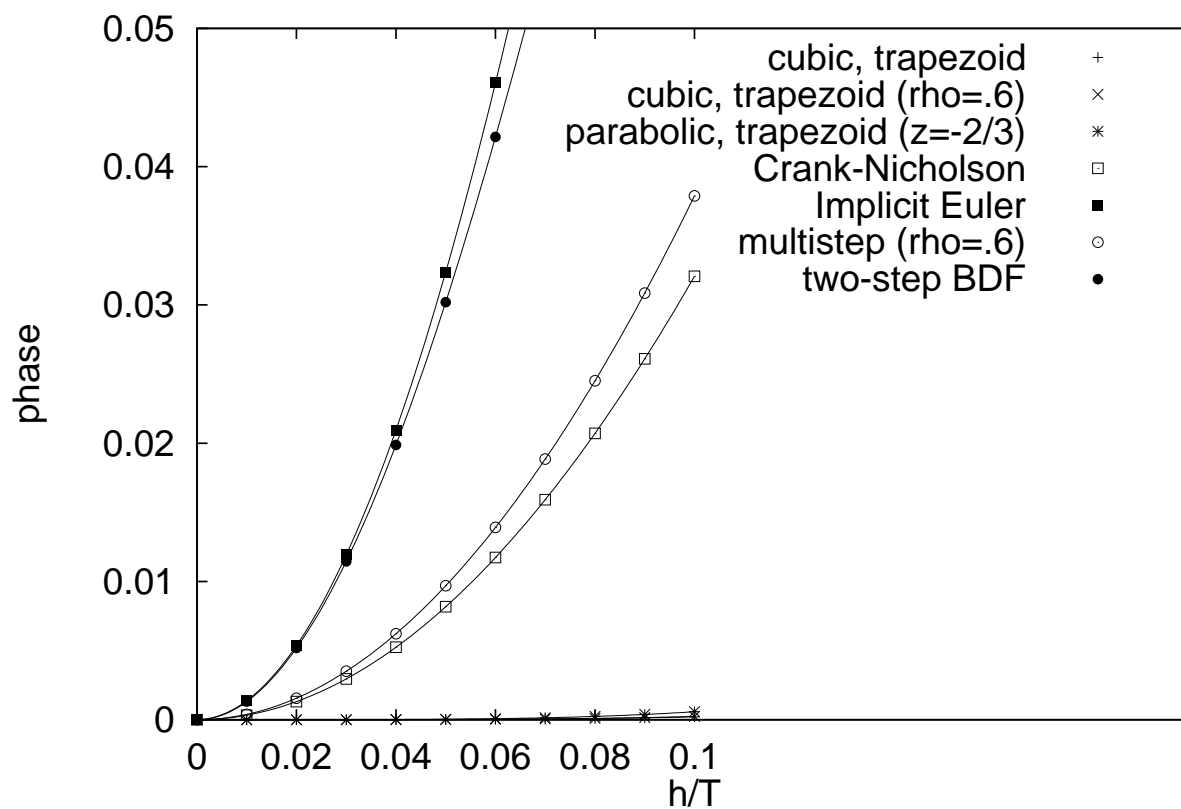
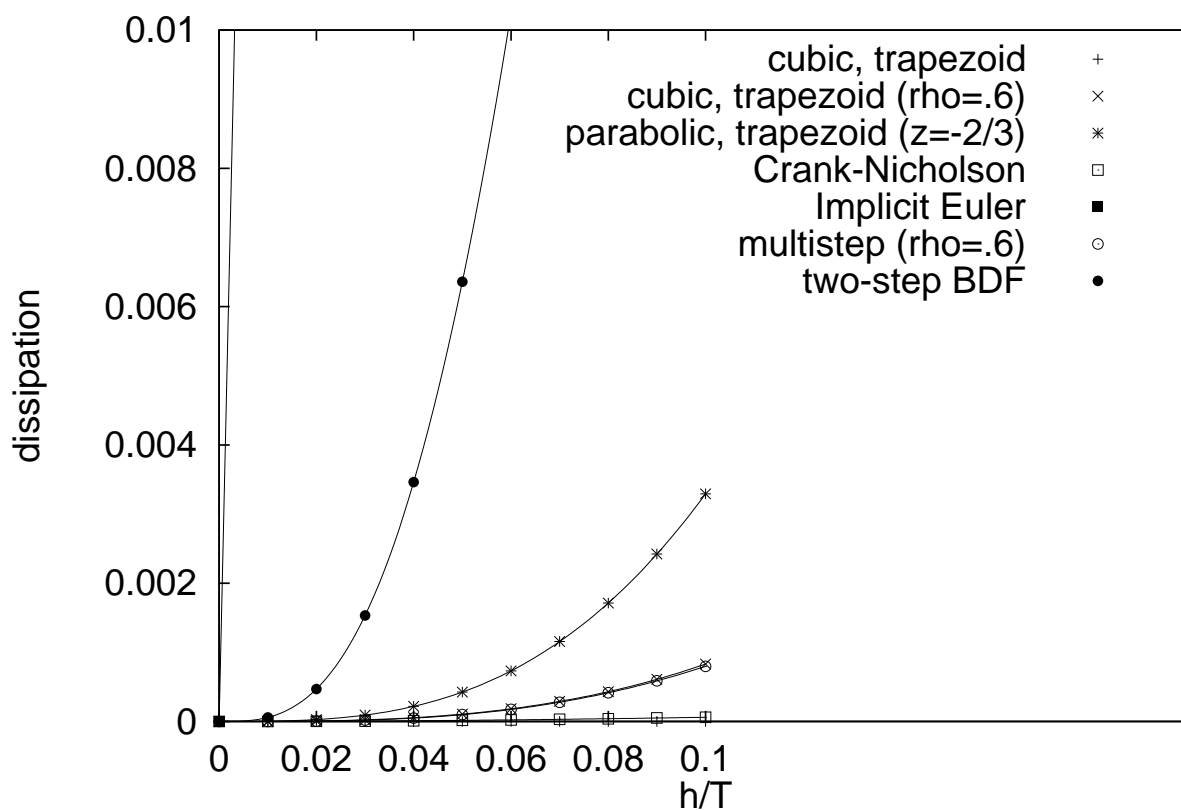


Figura 2: Errore in smorzamento e fase di un insieme di metodi

che rappresenta la linearizzazione locale di un generico problema non-lineare e quindi contiene informazioni sulla stabilità del problema non-lineare in un intorno della soluzione. In questo senso la stabilità lineare è detta anche stabilità “in piccolo”, in quanto i suoi risultati hanno valore generale solo per sistemi lineari. L’applicazione di un generico metodo di soluzione dà luogo, nel dominio del tempo discreto, ad un problema alle differenze

$$y_{k+1} = \rho y_k$$

in cui  $\rho = \rho(\lambda h)$ ; la *regione di stabilità* del metodo è data dall’insieme dei valori di  $\lambda$  e di  $h$  per cui vale la condizione  $|\rho| < 1$ , ovvero l’integrazione del semplice problema lineare dà una soluzione convergente o al più oscillante.

Quando si considerano metodi a passo multiplo, in genere si ottiene una relazione polinomiale in  $\rho$  di ordine pari al numero di passi; in questi casi si hanno radici multiple, di cui quella più grande in modulo caratterizza il comportamento lineare del metodo.

Vengono ora date le definizioni di stabilità di maggiore interesse nella trattazione di sistemi ODE e DAE.

### 3.1 Zero-stabilità

Un metodo si dice Zero-stabile quando lo scostamento tra due soluzioni approssimate si mantiene limitato ad ogni passo se a sua volta il passo di integrazione si mantiene limitato (ovvero  $h \in [0, h_0]$ ).

La Zero-stabilità è un requisito minimo che un metodo di approssimazione deve soddisfare per poter essere utilizzato; esprime rigorosamente il concetto di *stabilità condizionata*.

### 3.2 Primo Teorema di Dahlquist

Il teorema di Dahlquist (1956), noto anche come *prima barriera di Dahlquist*, afferma che un metodo di integrazione lineare a passo multiplo, a  $k$  passi, non può avere accuratezza superiore a  $k + 1$  se  $k$  è dispari, o a  $k + 2$  se  $k$  è pari, senza perdere la proprietà di Zero-stabilità.

Nei metodi di uso pratico per applicazioni di dinamica strutturale, ovvero per problemi con autovalori complessi e smorzamento molto basso, si usano metodi con  $k = 2$ ; questo significa che un metodo condizionatamente stabile non può avere accuratezza superiore a 4.

### 3.3 A-stabilità

Un metodo si dice A-stabile (Dahlquist) quando la sua regione di stabilità, per  $\Re(\lambda) < 0$ , non dipende dal valore del passo temporale, ovvero la soluzione converge a zero al crescere del numero di passi. L’A-stabilità esprime rigorosamente il concetto di *stabilità incondizionata*.



### 3.4 Seconda barriera di Dahlquist

Un ulteriore teorema di Dahlquist (1963), noto anche come *seconda barriera di Dahlquist*, afferma che un metodo di integrazione esplicito non può essere A-stabile; l'ordine di accuratezza di un metodo di integrazione a passo multiplo A-stabile non può essere superiore a 2; tra questi, quello con il minore errore è la regola dei trapezi (detta anche regola di Simpson o metodo di Crank-Nicolson).

Esistono metodi definiti *stiffly stable*, ovvero stabili quando applicati a problemi cosiddetti *stiff*, o rigidi (Gear). Questa nozione di rigidità non deve essere confusa con la nozione intuitiva di rigidità dei sistemi meccanici, associata alla presenza di frequenze elevate. I sistemi *stiff*, infatti, sono caratterizzati da autovalori molto grandi che stanno nel semipiano sinistro del piano complesso, ma in prossimità dell'asse reale, e corrispondono quindi a sistemi con smorzamento subcritico prossimo a quello critico.

Sono stati formulati metodi a passo multiplo incondizionatamente stabili per problemi di questo tipo che consentono di ottenere ordini di accuratezza superiori a 2 in violazione della seconda barriera di Dahlquist. Tuttavia questi metodi falliscono nel caso di problemi con autovalori complessi coniugati che si avvicinano all'asse immaginario, tipici dei sistemi meccanici. Di conseguenza non possono essere considerati A-stabili.

### 3.5 L-stabilità

Un metodo si dice L-stabile (Ehle, Axelsson) quando è A-stabile (quindi, se  $\Re(\lambda) < 0$ ,  $|\rho| < 1$  qualunque sia  $\lambda h$ ) e, per  $\Re(\lambda h) \rightarrow -\infty$ , vale la relazione  $y_{k+1} = \rho y_k = 0$ , ovvero il raggio spettrale si contrae nell'origine del cerchio unitario al crescere del passo  $h$ .

Questo criterio di stabilità è molto importante perché suggerisce un modo per eliminare dalla soluzione le oscillazioni legate ad autovalori particolarmente grandi, con  $\Re(\lambda h) \rightarrow -\infty$ , conservando al tempo stesso l'accuratezza del metodo per i termini della soluzione i cui autovalori siano dell'ordine della frequenza di integrazione,  $|\angle(\lambda h)| \ll \pi$ .

La Figura 3 illustra il raggio spettrale di un insieme di metodi al variare del passo di integrazione.

### 3.6 Stabilità non-lineare (in grande)

Un interessante capitolo riguarda la stabilità “in grande” di sistemi non-lineari. Il sistema lineare utilizzato per valutare la stabilità in piccolo dei metodi di approssimazione della soluzione può essere interpretato come una linearizzazione locale di un generico problema non-lineare. Le considerazioni svolte in quel caso si basavano sull'analisi della costante  $\lambda$ , associabile agli autovalori del sistema. Nel caso lineare tale informazione caratterizza tutta l'evoluzione del sistema. Nel caso non-lineare, invece, l'analisi degli autovalori locali per una configurazione del sistema, per forza di cose, non può dare informazioni globali. Al contrario, in letteratura si trovano esempi di problemi globalmente stabili i cui autovalori locali sono strutturalmente nel semipiano destro del piano complesso, e viceversa. Senza entrare nel dettaglio, si constata come una misura più appropriata della stabilità globale di un sistema, sempre nell'intorno di una soluzione, è data dalla nozione di *contrattività*, che esprime la tendenza di una soluzione perturbata ad avvicinarsi o

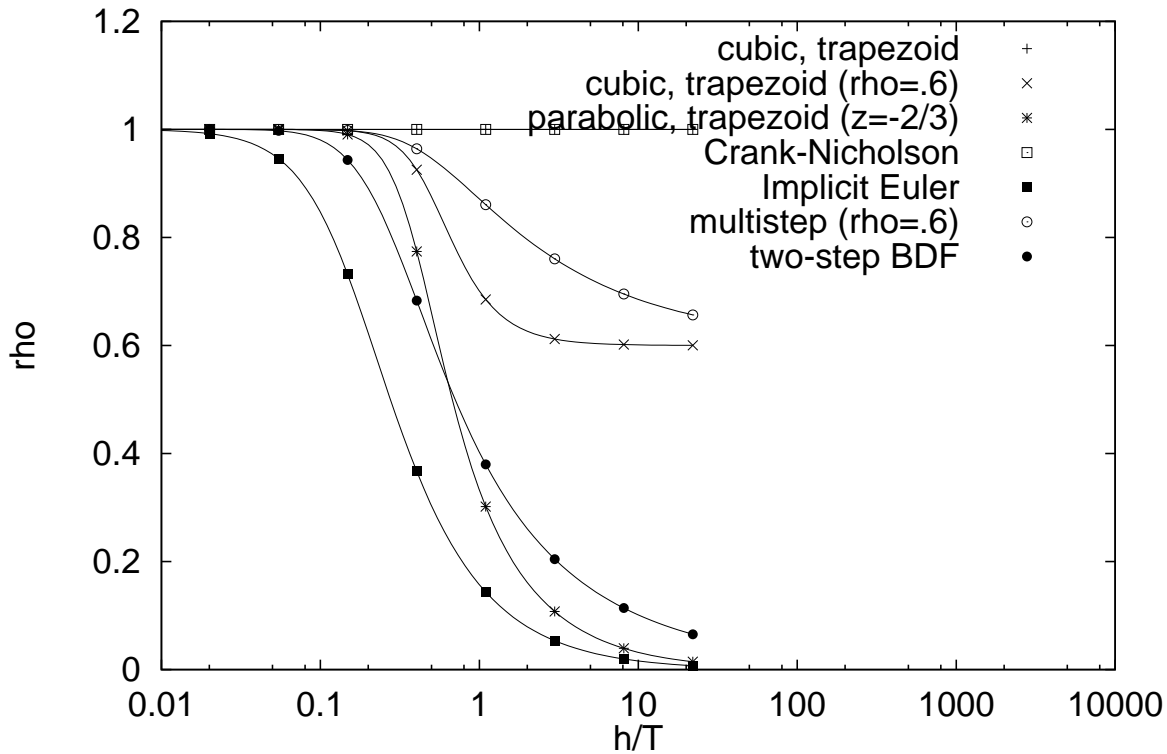


Figura 3: Raggio spettrale di un insieme di metodi

meno alla corrispondente soluzione imperturbata su un determinato intervallo di tempo. Questo fenomeno, evidente nel caso lineare in cui  $y = y_0 e^{\lambda t}$ , per cui la soluzione perturbata  $\hat{y} = (y_0 + \delta) e^{\lambda t}$  si avvicina a quella originaria a condizione che  $\Re(\lambda) < 0$ , può essere analizzata per via numerica nel caso non-lineare, previa definizione di una opportuna norma della perturbazione e di opportune condizioni di contrazione. Il risultato è una definizione di stabilità, quando sia ottenibile, che impiega direttamente i criteri di stabilità alla Liapunov.

Un approccio completamente diverso è quello usato nelle formulazioni cosiddette a *conservazione/dissipazione di energia*. In questo caso il metodo di integrazione è strettamente connesso alla formulazione del problema, di cui sfrutta proprietà intrinseche di conservazione/dissipazione (energia, momento della quantità di moto, ecc.). Se da una parte questi metodi garantiscono la stabilità globale per lo meno nei riguardi della grandezza sulla quale sono basati, dall'altra richiedono una formulazione ad hoc del problema e non si prestano ad applicazioni generali. Inoltre, la forma conservativa di questi metodi, interessante dal punto di vista teorico, non trova applicazione pratica se non in rari casi, perché la conservazione globale dell'energia in sistemi decisamente rigidi non impedisce il suo trasferimento dai modi a bassa frequenza, più fisici, a quelli ad alta frequenza, tipicamente di origine numerica o comunque legati alla discretizzazione spaziale, facendo degenerare ben presto la soluzione in modo caotico.

## 4 Metodi di approssimazione

### 4.1 A passo multiplo

Tra i metodi multi-passo significativi per l'integrazione di problemi meccanici, oltre alla già citata regola dei trapezi, un posto particolare è occupato dalle cosiddette BDF (Backward Differentiation Formulas). La formula è generale ed ha la forma:

$$\sum_{j=0,s} \alpha_j \mathbf{y}_{k-j} = h\beta_0 \mathbf{f}(\mathbf{y}_k, t_k)$$

con i coefficienti  $\alpha_i$  legati da alcune relazioni particolari. Le BDF sono Zero-stabili fino al sesto ordine, mentre, per la seconda barriera di Dahlquist, solo fino al secondo ordine ammettono A-stabilità; in particolare, le BDF fino al secondo ordine godono di L-stabilità (le BDF di primo ordine corrispondono al già citato metodo di Eulero implicito). La formula del secondo ordine è:

$$\mathbf{y}_k = \frac{4}{3}\mathbf{y}_{k-1} - \frac{1}{3}\mathbf{y}_{k-2} + \frac{2}{3}h\mathbf{f}(\mathbf{y}_k, t_k)$$

Una formula interessante, nell'ambito del formalismo descritto in precedenza, si ottiene utilizzando una interpolazione parabolica su due passi solo sugli stati ( $n_i = 0$ ) e integrando con la regola dei trapezi a tre punti, in modo da avere accuratezza del secondo ordine con una indeterminazione  $\delta$  sul valore dei pesi:

$$\mathbf{y}_k = \mathbf{y}_{k-2} + h \left( \left( \frac{1}{2} + \delta \right) \mathbf{f}(\mathbf{y}_k, t_k) + (1 - 2\delta) \mathbf{f}(\mathbf{y}_{k-1}, t_{k-1}) + \left( \frac{1}{2} + \delta \right) \mathbf{f}(\mathbf{y}_{k-2}, t_{k-2}) \right)$$

Alla forma risultante si aggiunga un'integrazione con la regola dei trapezi sui due passi più lontani, pesata da un opportuno coefficiente  $(\beta - 1)$ :

$$\mathbf{0} = (\beta - 1) \left( -\mathbf{y}_{k-1} + \mathbf{y}_{k-2} + \frac{h}{2} (\mathbf{f}(\mathbf{y}_{k-1}, t_{k-1}) + \mathbf{f}(\mathbf{y}_{k-2}, t_{k-2})) \right)$$

Siccome la soluzione ai passi  $k-1$ ,  $k-2$  è nota, questa relazione è una identità; si ottiene:

$$\begin{aligned} \mathbf{y}_k &= (1 - \beta) \mathbf{y}_{k-1} + \beta \mathbf{y}_{k-2} \\ &+ h \left( \left( \frac{1}{2} + \delta \right) \mathbf{f}(\mathbf{y}_k, t_k) + \left( \frac{1}{2} + \frac{\beta}{2} - 2\delta \right) \mathbf{f}(\mathbf{y}_{k-1}, t_{k-1}) + \left( \frac{\beta}{2} + \delta \right) \mathbf{f}(\mathbf{y}_{k-2}, t_{k-2}) \right) \end{aligned}$$

Il metodo risultante è intrinsecamente del secondo ordine; una opportuna scelta dei parametri, negli intervalli  $\beta = 1 \rightarrow -1/3$  e  $\delta = 0 \rightarrow 1/6$ , consente di passare con regolarità dalla A- alla L-stabilità (nel qual caso ovviamente si riottengono le BDF). Se si impone la coincidenza asintotica delle due radici del raggio spettrale, si può usare quest'ultimo come unico parametro per la caratterizzazione del metodo.

## 4.2 A passo singolo

Esistono innumerevoli metodi a passo singolo. Tipicamente si usano metodi di ordine elevato, a volte variabile in modo adattativo, per risolvere problemi con andamento molto regolare. In tali casi l'ordine elevato consente di procedere con passi elevati senza perdere in accuratezza. Nel caso dei problemi meccanici, in genere, non si supera il quarto ordine. Si consideri ad esempio un metodo il cui diagramma di Butcher è

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1/2 & 5/24 & 1/3 & -1/24 \\ 1 & 1/6 & 2/3 & 1/6 \\ \hline & 1/6 & 2/3 & 1/6 \end{array}$$

Questo metodo è noto come Lobatto IIIA del quarto ordine (il metodo di Crank-Nicolson viene anche classificato come Lobatto IIIA del secondo ordine). Lo si ottiene con la procedura proposta considerando una interpolazione Hermitiana cubica e una integrazione per collocazione con la regola dei trapezi a tre punti. Questo metodo è del quarto ordine, A-stabile. A partire da questo metodo, spostando in modo opportuno il punto di collocazione mediano verso la fine del passo, si ottiene un interessante metodo con caratteristiche di dissipazione algoritmica intermedie tra la A- e la L-stabilità.

Si consideri ora un metodo il cui diagramma di Butcher è:

$$\begin{array}{c|cc} 1/3 & 5/12 & -1/12 \\ 1 & 3/4 & 1/4 \\ \hline & 3/4 & 1/4 \end{array}$$

Questo metodo, noto come Radau IIA, è del terzo ordine, ed è L-stabile. Lo si ottiene con la procedura proposta utilizzando come pesi  $3/4$  e  $1/4$  e interpolando con polinomi quadratici con la funzione  $n_1 = 0$  sulla derivata all'istante iniziale.

## 5 Sistemi di equazioni algebrico-differenziali

Si consideri un sistema di equazioni differenziali in forma implicita

$$\mathbf{s}(\dot{\mathbf{y}}, \mathbf{y}, t) = 0$$

Esso si dice algebrico-differenziale (Differential-Algebraic Equations, DAE) quando la sua derivata parziale rispetto alla derivata dell'incognita è strutturalmente singolare:

$$\det(\mathbf{s}_{/\dot{\mathbf{y}}}) = 0$$

ma quando il *pencil* costituito da

$$\mathbf{s}_{/\dot{\mathbf{y}}} + \lambda \mathbf{s}_{/\mathbf{y}}$$

è singolare al più per un numero finito di valori di  $\lambda \neq 0$  (corrispondenti agli autovalori del problema generalizzato costituito dalle matrici  $\mathbf{s}_{/\dot{\mathbf{y}}}$  e  $\mathbf{s}_{/\mathbf{y}}$ ).

In tale caso, operando un opportuno partizionamento delle incognite in differenziali,  $\mathbf{y}$ , ed algebriche,  $\mathbf{z}$ , il sistema può essere riscritto come:

$$\begin{aligned}\mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{z}, t) &= \mathbf{0} \\ \mathbf{g}(\mathbf{x}, \mathbf{z}, t) &= \mathbf{0}\end{aligned}$$

o, nel caso particolare in cui le equazioni differenziali siano esplicitabili:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{z}, t) \\ \mathbf{0} &= \mathbf{g}(\mathbf{x}, \mathbf{z}, t)\end{aligned}$$

Si hanno diversi tipi di DAE a seconda che la parte differenziale  $\mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{z}, t)$  sia effettivamente esplicitabile o meno rispetto alla derivata dell'incognita differenziale  $\dot{\mathbf{x}}$ , e che le incognite algebriche  $\mathbf{z}$  siano o meno contenute nella parte algebrica del sistema  $\mathbf{g}(\mathbf{x}, \mathbf{z}, t)$ .

## 5.1 Perturbazioni singolari

Spesso si ottiene un sistema di DAE quando si considerano sistemi in cui alcune derivate dipendono da parametri che possono assumere valori molto piccoli; si consideri ad esempio il sistema

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{z}, t) \\ \epsilon \dot{\mathbf{z}} &= \mathbf{g}(\mathbf{x}, \mathbf{z}, t)\end{aligned}$$

con  $0 < \epsilon \ll 1$ . Questa situazione in genere si verifica qualora la dinamica di una parte del problema sia molto più veloce di quella più significativa, descritta dalle incognite  $x$ . In questi casi, generalmente, un sistema si dice *stiff*, ovvero rigido (anche se di solito con questa accezione ci si riferisce a problemi i cui autovalori giacciono lungo l'asse reale).

Si consideri ad esempio un problema costituito da due masse,  $m_1$  e  $m_2$ , di cui la prima messa a terra da una molla,  $k_1$ , e collegate da un'altra molla  $k_2$ :

$$\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \begin{Bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{Bmatrix} + \begin{bmatrix} k_1 + k_2 & -k_2 \\ -k_2 & k_2 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} = \begin{Bmatrix} F_1 \\ F_2 \end{Bmatrix}$$

Si considerino quindi gli autovalori del problema omogeneo associato:

$$s_{1,2}^2 = -\frac{1}{2} \left( \frac{k_2}{m_2} + \frac{k_1 + k_2}{m_1} \right) \pm \frac{1}{2} \sqrt{\left( \frac{k_2}{m_2} + \frac{k_1 + k_2}{m_1} \right)^2 - 4 \frac{k_1}{m_1} \frac{k_2}{m_2}}$$

Al tendere all'infinito della rigidità del supporto,  $k_1 \rightarrow \infty$ , la prima equazione del sistema,

$$\frac{m_1}{k_1} \ddot{x}_1 + \left( 1 + \frac{k_2}{k_1} \right) x_1 - \frac{k_2}{k_1} x_2 = \frac{F_1}{k_1}$$

degenera nell'equazione algebrica

$$x_1 = 0$$

mentre gli autovalori diventano

$$s_1^2 = \infty \quad \text{e} \quad s_2^2 = \frac{k_2}{m_2}$$

## 5.2 Soluzione diretta

Per capire il significato di questo tipo di equazioni si può pensare di ridurle al set minimo di incognite differenziali esplicitando la parte algebrica rispetto alle incognite algebriche, nell'ipotesi che questo sia localmente possibile, in modo da ottenere

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{z}, t) \\ \mathbf{z} &= \hat{\mathbf{g}}(\mathbf{x}, t)\end{aligned}$$

a questo punto le incognite algebriche possono essere sostituite nella parte differenziale, in modo da ottenere un set di equazioni ordinarie ODE nelle sole  $\mathbf{x}$ :

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \hat{\mathbf{g}}(\mathbf{x}, t), t)$$

La motivazione che spinge verso la soluzione diretta delle DAE anziché ridurle ad ODE è legata al fatto che spesso si studiano problemi di grandi dimensioni, ma molto sparsi (ovvero le matrici hanno moltissimi coefficienti nulli); la riduzione ad ODE le renderebbe piene. Inoltre, se si adottano particolari tecniche di soluzione di tipo iterativo anziché diretto, non è neppure necessario generare ed assemblare le matrici; questo vantaggio può essere sfruttato solo se si preserva la sparsità delle matrici.

## 5.3 Concetto di indice

A partire dal sistema DAE in forma esplicita, si immagini di volerlo portare in forma differenziale. Assumendo che la funzione  $\mathbf{g}$  sia localmente invertibile, la si derivi rispetto al tempo; il sistema diventa:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{z}, t) \\ \mathbf{0} &= \mathbf{g}_{/\mathbf{x}}\dot{\mathbf{x}} + \mathbf{g}_{/\mathbf{z}}\dot{\mathbf{z}} + \mathbf{g}_{/t}\end{aligned}$$

quindi, sostituendo la prima equazione nella seconda, si ottiene:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, \mathbf{z}, t) \\ \dot{\mathbf{z}} &= -\mathbf{g}_{/\mathbf{z}}^{-1}(\mathbf{g}_{/\mathbf{x}}\mathbf{f}(\mathbf{x}, \mathbf{z}, t) + \mathbf{g}_{/t})\end{aligned}$$

il problema viene ricondotto in forma puramente differenziale in entrambi gli insiemi di incognite. Si definisce indice di un sistema DAE il numero di differenziazioni richieste per portare il sistema in forma puramente differenziale. Nota: questo non significa che l'operazione debba essere effettivamente svolta; al contrario, un metodo di approssimazione della soluzione basato su questa sequenza di differenziazioni risulterebbe in un sistema di equazioni sostanzialmente diverso da quello iniziale e darebbe risultati profondamente errati. Infatti, se si risolve un sistema costituito dalla parte differenziale del sistema originario e dalla derivata dell'equazione di vincolo, non è più garantito il soddisfacimento dei vincoli algebrici originari, dato che, poiché la soluzione è calcolata per via numerica, si può incorrere nel fenomeno della deriva.

## 5.4 Equazioni di indice 1

Le equazioni di indice 1 sono quelle descritte in precedenza, nella forma

$$\begin{aligned}\mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{z}, t) &= \mathbf{0} \\ \mathbf{g}(\mathbf{x}, \mathbf{z}, t) &= \mathbf{0}\end{aligned}$$

Un tipico esempio è dato dalle equazioni in cui la parte algebrica rappresenta la definizione di un termine che viene utilizzato nella parte differenziale; ad esempio:

$$\begin{aligned}\dot{\mathbf{u}} &= \mathbf{M}^{-1}\mathbf{v} \\ \dot{\mathbf{v}} &= -\mathbf{r} + \mathbf{f}(t) \\ \mathbf{r} &= \mathbf{k}(\mathbf{u})\end{aligned}$$

in cui l'equazione algebrica contiene la definizione della forza elastica non lineare,  $\mathbf{R}$ . La soluzione di queste equazioni in genere non comporta problemi, in quanto la teoria per DAE di indice 1 è completa.

## 5.5 Equazioni di indice 2

Le equazioni di indice 2 hanno la forma

$$\begin{aligned}\mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, \mathbf{z}, t) &= \mathbf{0} \\ \mathbf{g}(\mathbf{x}, t) &= \mathbf{0}\end{aligned}$$

ovvero la parte algebrica non dipende direttamente dalle incognite algebriche. In questo caso sono richieste due differenziazioni per portarsi in forma differenziale, ovvero è necessario derivare anche la parte differenziale.

## 5.6 Equazioni di indice 3

Esiste un tipo di equazioni di indice 3 particolarmente importante perché ad esso sono riconducibili le equazioni della dinamica di sistemi meccanici con vincoli olonomi; questa forma è detta di Hessemberg:

$$\begin{aligned}\dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{z}, t) \\ \dot{\mathbf{x}}_2 &= \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2, t) \\ \mathbf{0} &= \mathbf{g}(\mathbf{x}_2, t)\end{aligned}$$

Come si può notare, le equazioni algebriche dipendono solo da incognite differenziali le cui equazioni non contengono le incognite algebriche stesse; sono quindi necessarie tre differenziazioni per ricondursi alla forma puramente differenziale. In questo caso, le incognite algebriche giocano il ruolo delle reazioni vincolari, spesso esprimibili mediante la tecnica dei moltiplicatori di Lagrange. Modificando l'esempio meccanico precedente si ottiene:

$$\begin{aligned}\dot{\mathbf{u}} &= \mathbf{M}^{-1}\mathbf{v} \\ \dot{\mathbf{v}} &= -\mathbf{k}(\mathbf{u}) + \phi_{\mathbf{u}}^T \boldsymbol{\lambda} + \mathbf{f}(t) \\ \mathbf{0} &= \phi(\mathbf{u}, t)\end{aligned}$$

La soluzione di queste equazioni comporta difficoltà numeriche non trascurabili. La teoria delle DAE di indice superiore a 2 è tutt'altro che consolidata. Si noti che il metodo di eliminazione dei moltiplicatori illustrato inizialmente, applicato all'equazione differenziale di secondo grado, richiedeva due derivazioni dell'equazione di vincolo; questo illustra come la nozione di indice sia legata naturalmente al numero di derivazioni necessarie per ridurre il problema a differenziale (la terza derivazione è legata alla scrittura del problema al secondo ordine).

## 5.7 Soluzione

La soluzione di sistemi DAE presenta problemi non banali. Innanzitutto, se si considera la parte algebrica in termini di parametri singolari, il sistema, in generale non lineare, localmente presenta autovalori molto grandi (in confronto a quelli della parte puramente differenziale) che, al tendere a zero dei parametri, vanno all'infinito. Questo significa che un metodo di integrazione esplicito non può assolutamente essere usato, in quanto non ne è garantita la Zero-stabilità. Si rende necessario il ricorso a metodi impliciti, ovvero A-stabili.

Tuttavia anche la A-stabilità non è un requisito sufficiente, perché in tale caso le oscillazioni corrispondenti agli autovalori legati alla parte algebrica non sarebbero attenuate e, in assenza di qualche forma di dissipazione, il calcolo tenderebbe a divergere con oscillazioni puramente numeriche legate al passo di integrazione, come ampiamente verificato in letteratura. Si rende quindi necessario il ricorso a metodi L-stabili, o comunque in grado di introdurre dissipazione algoritmica.

Tra i metodi utilizzati in passato, vi erano tecniche basate sull'eliminazione dei gradi di libertà algebrici attraverso la decomposizione spettrale del problema. Tra i principali svantaggi la difficoltà di ottenere buone decomposizioni spettrali e la perdita della sparsità della matrice dopo la decomposizione.

Un approccio investigato in passato è il cosiddetto alla Baumgarte [6], in cui le equazioni algebriche vengono derivate (due volte, in quanto il problema è tipicamente formulato in termini di equazioni differenziali del secondo ordine) e, per evitare il fenomeno della deriva, all'equazione algebrica vengono aggiunte le sue derivate prima e seconda, pesate da opportuni coefficienti:

$$\begin{aligned} \mathbf{M}\ddot{\mathbf{x}} &= \mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, t) + \mathbf{g}_{/\mathbf{x}}^T \boldsymbol{\lambda} & \rightarrow & \quad \mathbf{M}\ddot{\mathbf{x}} = \mathbf{f}(\dot{\mathbf{x}}, \mathbf{x}, t) + \mathbf{g}_{/\mathbf{x}}^T \boldsymbol{\lambda} \\ \mathbf{0} &= \mathbf{g}(\mathbf{x}, t) & & \quad \mathbf{0} = \ddot{\mathbf{g}}(\mathbf{x}, t) + 2\alpha\dot{\mathbf{g}}(\mathbf{x}, t) + \beta^2\mathbf{g}(\mathbf{x}, t) \end{aligned}$$

L'inconveniente di questo metodo, oltre alla necessità di calcolare le derivate seconde del vincolo, sta nel fatto che il problema che si integra ha autovalori (e quindi dinamica) diversi da quello reale per effetto delle matrici aggiunte; inoltre, in caso di condizioni iniziali non consistenti con i vincoli o in caso di forti gradienti o discontinuità di velocità, si possono avere transitori non fisici, in quanto anziché imporre il soddisfacimento delle equazioni di vincolo  $\mathbf{g}(\mathbf{x}, t) = \mathbf{0}$ , in realtà si impone il soddisfacimento di un'equazione omogenea nelle equazioni di vincolo. Infine, la scelta dei coefficienti di peso  $\alpha$  e  $\beta$  presenta un grado di aleatorietà spesso fuorviante.

Oggi è comunemente riconosciuta la possibilità di integrare direttamente le equazioni algebrico-differenziali fino ad indice 2 con opportuni metodi A-stabili, e fino ad indice 3



e superiori con metodi che presentino dissipazione algoritmica ai limiti della L-stabilità, a patto di accettare, in casi estremi, la violazione (nei limiti dell'accuratezza del metodo) delle derivate delle equazioni di vincolo, e una riduzione di ordine di accuratezza sulle reazioni vincolari.

### 5.7.1 Soddisfacimento delle equazioni di vincolo

Il problema algebrico-differenziale di partenza esprime una condizione di vincolo sulle incognite di posizione del sistema; dal punto di vista matematico, una condizione del tipo

$$\mathbf{g}(\mathbf{x}, t) = \mathbf{0}$$

implica anche

$$\frac{d^n}{dt^n} \mathbf{g}(\mathbf{x}, t) = \mathbf{0}$$

Durante la soluzione numerica del problema, tuttavia, se viene imposto soltanto il soddisfacimento della condizione di vincolo originaria, niente assicura che le condizioni sulle derivate del vincolo siano soddisfatte. Tipicamente i metodi di integrazione multi-passo impongono il soddisfacimento delle equazioni di vincolo solo negli istanti corrispondenti ai passi; i metodi a passo singolo spesso utilizzano anche punti intermedi, o tecniche di pesatura, che in ogni caso alla fine si riducono a valutazioni collocate. Questo significa che tutte le soluzioni che soddisfano i vincoli nei soli punti di integrazione sono accettabili, quindi la soluzione del problema numerico non è unica anche se la soluzione esatta lo sarebbe. Questo vale sempre per soluzioni numeriche; tuttavia, nel caso di problemi algebrico-differenziali, diventa in special modo critico.

## 5.8 Condizioni iniziali consistenti con le derivate del vincolo

Uno dei punti fondamentali della soluzione di DAE di indice 3 è che la parte algebrica  $\mathbf{0} = \mathbf{g}(\mathbf{x}_2, t)$  implica che anche la condizione  $\mathbf{0} = \mathbf{g}_{/\mathbf{x}_2} \dot{\mathbf{x}}_2 + \mathbf{g}_{/t}$  sia soddisfatta; ne risulta

$$\mathbf{0} = \mathbf{g}_{/\mathbf{x}_2} \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2, t) + \mathbf{g}_{/t}$$

ovvero una ulteriore equazione puramente algebrica nelle sole incognite differenziali. Questa equazione ci dice che le derivate di una parte delle incognite differenziali devono sottostare ad un ulteriore set di condizioni per essere consistenti; dal punto di vista cinematico, non solo i gradi di libertà di spostamento devono soddisfare l'equazione di vincolo, ma l'atto di moto legato alle loro derivate deve essere consistente con le derivate dei vincoli, ovvero deve svolgersi in modo da non violarli (si dice che l'atto di moto deve essere tangente alla *varietà*, o *manifold*, definita dalle equazioni di vincolo). Questa condizione è importante all'avvio del calcolo, mentre nella maggior parte dei casi una soluzione sufficientemente accurata che adotti una formula di integrazione più che A-stabile fa sì che la violazione delle derivate dei vincoli sia limitata e contenuta entro l'accuratezza del metodo di approssimazione, risultando in una perdita di accuratezza (tipicamente un ordine) solo sulle derivate seconde del moto (le accelerazioni) e sulle incognite algebriche.

La DAE di indice 3 descritta in precedenza, aumentata dalla derivata dell'equazione di vincolo e mediante l'introduzione di un ulteriore set di moltiplicatori  $\mathbf{w}$  sulla equazione priva di incognite algebriche, dà luogo ad una DAE di indice 2:

$$\begin{aligned}\dot{\mathbf{x}}_1 &= \mathbf{f}_1(\mathbf{x}_1, \mathbf{x}_2, \mathbf{z}, t) \\ \dot{\mathbf{x}}_2 &= \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2, t) + \mathbf{w} \\ 0 &= \mathbf{g}(\mathbf{x}_2, t) \\ 0 &= \mathbf{g}_{/\mathbf{x}_2} \mathbf{f}_2(\mathbf{x}_1, \mathbf{x}_2, t) + \mathbf{g}_{/t}\end{aligned}$$

Il significato del nuovo set di moltiplicatori  $\mathbf{w}$  sta nel fatto che l'approssimazione legata all'integrazione numerica, per consentire il soddisfacimento delle derivate delle equazioni di vincolo, deve rilassare il soddisfacimento delle equazioni differenziali che originariamente non dipendevano dalle incognite algebriche. Quanto più l'integrazione è accurata, tanto più i moltiplicatori  $\mathbf{w}$  tendono a zero.

Questa forma, inizialmente suggerita da Gear, Gupta e Leimkuhler [7], consente la riduzione dell'indice a 2, ed è detta *formulazione ad indice 2 stabilizzata*; questa forma è importante perché la teoria delle DAE di indice 2 è più completa rispetto a quella per DAE di indice superiore; ovviamente la stabilizzazione delle equazioni avviene al costo di un aumento delle dimensioni del problema e con la complessità legata alla derivazione delle equazioni di vincolo.

## 6 Aspetti numerici

### 6.1 Problema algebrico non-lineare

I problemi differenziali e algebrico-differenziali risultano sempre nella risoluzione di un problema algebrico non-lineare; infatti, quando si passa alla linearizzazione dell'equazione implicita

$$\mathbf{f}(\dot{\mathbf{y}}, \mathbf{y}, t) = \mathbf{0}$$

si ottiene:

$$\mathbf{f}_{/\dot{\mathbf{y}}} \Delta \dot{\mathbf{y}} + \mathbf{f}_{/\mathbf{y}} \Delta \mathbf{y} = -\mathbf{f}$$

e, a partire dall'equazione del metodo numerico utilizzato, si ottiene:

$$\Delta \dot{\mathbf{y}} = c \Delta \mathbf{y}$$

o la relazione inversa, dove il parametro  $c$  dipende dal metodo e contiene il passo di integrazione. Quindi il problema da risolvere è del tipo

$$(c\mathbf{f}_{/\dot{\mathbf{y}}} + \mathbf{f}_{/\mathbf{y}}) \Delta \mathbf{y} = -\mathbf{f}$$

e la matrice risolvibile è sicuramente strutturalmente non-singolare, a meno di occasionali impuntamenti o labilità.

## 6.2 Applicazione a problemi impliciti

Si consideri l'applicazione di un metodo alla Runge-Kutta alla soluzione del problema implicito precedente; in accordo con la descrizione del metodo vista in precedenza, l'incognita e la sua derivata in un generico punto sono descritte da:

$$\begin{aligned}\dot{\mathbf{y}}(t_i) &= \mathbf{Y}_i \\ \mathbf{y}(t_i) &= \mathbf{y}_{n-1} + h \sum_{j=1,s} a_{ij} \mathbf{Y}_j\end{aligned}$$

La loro perturbazione dà:

$$\begin{aligned}\Delta \dot{\mathbf{y}}(t_i) &= \Delta \mathbf{Y}_i \\ \Delta \mathbf{y}(t_i) &= h \sum_{j=1,s} a_{ij} \Delta \mathbf{Y}_j\end{aligned}$$

La funzione  $\mathbf{f}$  va valutata in tutti i punti di collocazione; a tal fine se ne consideri direttamente la linearizzazione nel punto  $i$ -esimo:

$$\mathbf{f}_{/\dot{\mathbf{y}}} \Delta \mathbf{Y}_i + \mathbf{f}_{/\mathbf{y}} h \sum_{j=1,s} a_{ij} \Delta \mathbf{Y}_j = -\mathbf{f}$$

L'assemblaggio di tutte le valutazioni porta alla risoluzione di un problema algebrico implicito di ordine pari al numero di incognite  $\mathbf{y}$  per il numero di punti di collocazione  $s$ :

$$\begin{bmatrix} \mathbf{f}_{/\dot{\mathbf{y}}} + h a_{11} \mathbf{f}_{/\mathbf{y}} & \cdots & h a_{1s} \mathbf{f}_{/\mathbf{y}} \\ \vdots & \ddots & \vdots \\ h a_{s1} \mathbf{f}_{/\mathbf{y}} & \cdots & \mathbf{f}_{/\dot{\mathbf{y}}} + h a_{ss} \mathbf{f}_{/\mathbf{y}} \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{Y}_1 \\ \vdots \\ \Delta \mathbf{Y}_s \end{Bmatrix} = - \begin{Bmatrix} \mathbf{f}(t_{n-1} + c_1 h) \\ \vdots \\ \mathbf{f}(t_{n-1} + c_s h) \end{Bmatrix}$$

Un problema algebrico-differenziale non può essere risolto con un metodo esplicito, perché in tale caso la matrice risolvibile sarebbe strutturalmente singolare.

La soluzione diventa interessante qualora si consideri un metodo diagonalmente implicito:

$$\begin{bmatrix} \mathbf{f}_{/\dot{\mathbf{y}}} + h a_{11} \mathbf{f}_{/\mathbf{y}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ h a_{s1} \mathbf{f}_{/\mathbf{y}} & \cdots & \mathbf{f}_{/\dot{\mathbf{y}}} + h a_{ss} \mathbf{f}_{/\mathbf{y}} \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{Y}_1 \\ \vdots \\ \Delta \mathbf{Y}_s \end{Bmatrix} = - \begin{Bmatrix} \mathbf{f}(t_{n-1} + c_1 h) \\ \vdots \\ \mathbf{f}(t_{n-1} + c_s h) \end{Bmatrix}$$

in questo caso occorre risolvere gli  $s$  sottoproblemi dati dalle matrici sulla diagonale principale. Queste sottomatrici sono diverse e quindi devono essere fattorizzate separatamente a meno che non si consideri un metodo singolarmente implicito, in cui tutti i coefficienti diagonalmente sono uguali,  $a_{ii} = d$ :

$$\begin{bmatrix} \mathbf{f}_{/\dot{\mathbf{y}}} + h d \mathbf{f}_{/\mathbf{y}} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ h a_{s1} \mathbf{f}_{/\mathbf{y}} & \cdots & \mathbf{f}_{/\dot{\mathbf{y}}} + h d \mathbf{f}_{/\mathbf{y}} \end{bmatrix} \begin{Bmatrix} \Delta \mathbf{Y}_1 \\ \vdots \\ \Delta \mathbf{Y}_s \end{Bmatrix} = - \begin{Bmatrix} \mathbf{f}(t_{n-1} + c_1 h) \\ \vdots \\ \mathbf{f}(t_{n-1} + c_s h) \end{Bmatrix}$$

in questo caso è richiesta una sola fattorizzazione, e la soluzione avviene poi per sostituzione con i termini extra-diagonali. Questo a rigore vale solo nel caso in cui il problema

sia a coefficienti costanti, ovvero le matrici  $\mathbf{f}_{/y}$  e  $\mathbf{f}_{/y}$  non devono dipendere né dal tempo né dallo stato. L'inconveniente principale è comunque dato dal fatto che il risparmio in termini computazionali che si ottiene viene pagato con una riduzione dell'ordine del metodo. Ad esempio, mentre è possibile disegnare metodi di Runge-Kutta L-stabili del terzo ordine con un punto intermedio, quando li si rende diagonalmente o singolarmente impliciti l'accuratezza cade al secondo ordine, se gli si impone anche la condizione di essere *accurati su problemi rigidi* (stiffly accurate), ovvero  $c_s = 1$ , che implica che l'ultimo punto intermedio coincida con la fine del passo stesso.

## Riferimenti bibliografici

- [1] C. F. Gauss, “Ueber ein neues allgemeines Grundgesetz der Mechanik,” *J. fuer die reine und angewandte Mathematik*, vol. 4, pp. 232–235, 1829.
- [2] F. E. Udvardi and R. E. Kalaba, *Analytical Dynamics*. New York: Cambridge University Press, 1996.
- [3] G. A. Maggi, *Principii di stereodinamica: Corso sulla formazione, l'interpretazione e l'integrazione delle equazioni del movimento dei solidi*. Milano: Hoepli, 1903. In Italian.
- [4] T. Levi-Civita and U. Amaldi, *Lezioni di Meccanica Razionale — Vol. II: Dinamica dei Sistemi con un numero finito di gradi di libertà, parte II*. Bologna: Zanichelli, 1974. In Italian.
- [5] B. Finzi, *Meccanica razionale. Volume secondo: dinamica*. Bologna: Zanichelli, 1975. In Italian.
- [6] J. Baumgarte, “Stabilization of constraints and integrals of motion in dynamical systems,” *Computer Methods in Applied Mechanics*, vol. 1, pp. 1–36, 1972.
- [7] C. W. Gear, G. K. Gupta, and B. Leimkuhler, “Automatic integration of euler-lagrange equations with constraints,” *J. Comp. Appl. Math.*, vol. 12&13, pp. 77–90, 1985.
- [8] K. E. Brenan, S. L. V. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*. New York: North-Holland, 1989.
- [9] E. Hairer, C. Lubich, and M. Roche, *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*. Lecture Notes in Mathematics, Berlin Heidelberg, Germany: Springer-Verlag, 1989.
- [10] J. D. Lambert, *Numerical Methods for Ordinary Differential Systems*. Chichester, England: John Wiley & Sons Ltd., 1991.
- [11] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations*, vol. II. Berlin Heidelberg, Germany: Springer-Verlag, 1996. 2<sup>nd</sup> rev. ed.

- [12] P. Masarati, M. Lanz, and P. Mantegazza, “Multistep integration of ordinary, stiff and differential-algebraic problems for multibody dynamics applications,” in *XVI Congresso Nazionale AIDAA*, (Palermo), pp. 71.1–10, 24–28 Settembre 2001.