CHAPTER 9: *DAEs*

- DAE theory is more recent than ODE theory.

  $\rightarrow$ There are similarities, but differences too!

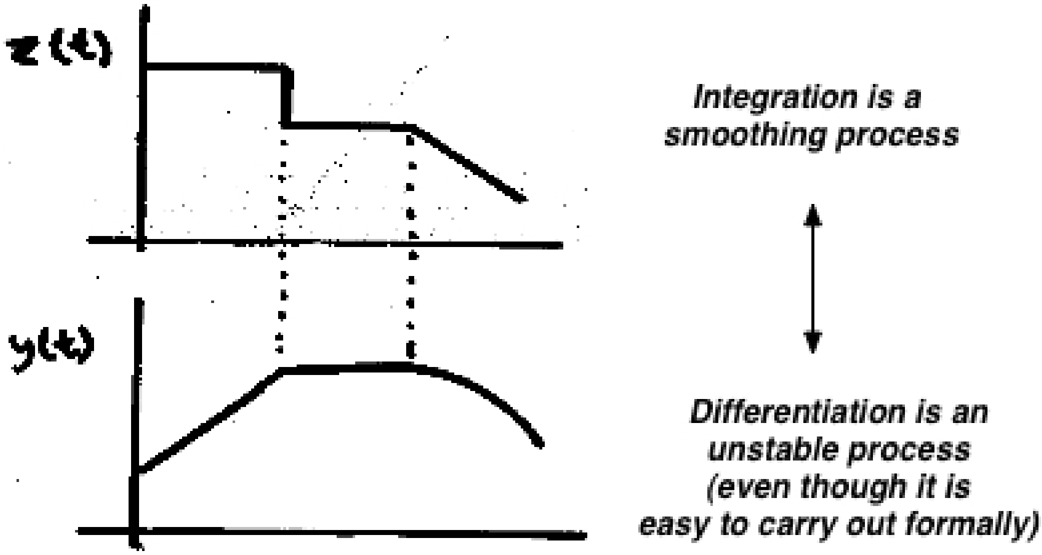
Consider two functions $y(t), z(t)$ related by

$$\dot{y}(t) = z(t), \qquad 0 \leq t \leq t_f.$$

To recover $z(t)$ from $y(t)$,
we must differentiate (an automated process).

To recover $y(t)$ from $z(t)$,
we must integrate (tricky! also need $\underbrace{y(0)}_{\text{side condition}}$).

Note 1. $y(t)$ *is smoother than* $z(t)$.



To solve a DE, you must integrate
$\leftrightarrow$ smoothing.

If $\mathbf{y}$ satisfies $\dot{\mathbf{y}} = \mathbf{A}\mathbf{y} + \mathbf{q}(t)$,
then $\mathbf{y}$ is one derivative smoother than $\mathbf{q}(t)$.

A DAE involves both integration and differentiation.

Because ODEs are a subset of DAEs, we can see the process of integration.

However, these processes may be intertwined!
(That's where the fun begins.)

9.1 Index

One might hope that (repeated) differentiation of constraints might reduce a DAE to an ODE.

It turns out it is possible to do this unless the problem is singular.

The number of differentiations required is called the (differential) index of the DAE.

Thus ODEs have index 0.

Note 2. *Higher-index DAEs are harder to solve.*

Example 1. *Let $q(t)$ be a given smooth function.*

- *The scalar equation*

$$y = q(t)$$

  is a (trivial) index-1 DAE.
  (With a differentiation, you obtain an ODE for y.)

- *Consider*

$$\begin{aligned} y_1 &= q(t) \\ y_2 &= \dot{y}_1 \end{aligned}$$

Take $\frac{d}{dt}$ *of first equation:* $\dot{y}_1 = \dot{q}(t) = y_2$
Now take $\frac{d}{dt}$ *of this equation:*

$$\dot{y}_2 = \ddot{q}(t)$$

This is an index-2 DAE (constraint differentiated twice to get ODE for y_2*).*

- *The system*

$$\begin{aligned} u &= q(t) \\ y_3 &= \ddot{u} \end{aligned}$$

is index 3. (verify!)

Note 3. *Before, when we had a system of* m *first-order ODEs, we needed* m *(independent) initial/boundary conditions.*

Now the DAEs are determined by the right-hand sides!
$\rightarrow$ No extra conditions are needed!

In general, a DAE will have l degrees of freedom where $0 \leq l \leq m$.

- It may not be obvious what l pieces of information are needed to specify the solution of the DAE! Often, all m components may have initial values prescribed.

$\rightarrow$ Too many conditions!

$\rightarrow$ They must be consistent.

e.g., if from $y(t) = q(t)$ we get the ODE $\dot{y}(t) = \dot{q}(t)$, we must specify $y(0) = q(0)$.

e.g., from $y_1 = q(t)$, we obtained $\dot{y}_2 = \ddot{q}(t)$

$y_2 = \dot{y}_1(t)$ $\qquad\qquad$ $\dot{y}_1 = y_2(t)$,

obviously we must have $\quad y_1(0) = q(0)$.

But we must also satisfy the hidden constraint $y_2 = \dot{q}(t)$ by specifying $\quad y_2(0) = \dot{q}(0)$.

This is an important difference between index-1 DAEs and <u>higher-index</u> DAEs (index > 1).

$\rightarrow$ Higher-index DAEs have hidden constraints!

$\updownarrow$

derivatives of explicitly stated constraints

Index $n \leftrightarrow (n-1)$ hidden derivatives of explicit constraints.

For the index-3 example, the hidden constraints are

$$\dot{u} = \dot{q}(t) \quad \text{and} \quad y_3 = \ddot{q}(t). \qquad \text{(verify)}$$

The most general form of a DAE is

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

where $\frac{\partial \mathbf{F}}{\partial \dot{\mathbf{y}}}$ may be singular.

An important special case is the
semi-explicit DAE $\leftrightarrow$ ODE with constraints:

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

Index 1 if $\frac{\partial \mathbf{g}}{\partial \mathbf{z}}$ is nonsingular. (verify)

For the semi-explicit index-1 DAE,
we can distinguish between differential variables $\mathbf{x}$ and algebraic variables $\mathbf{z}$.

Algebraic variables may be one derivative less smooth than differential variables.

$\rightarrow$ Algebraic variables may not be differentiable!

In general, a DAE variable $\mathbf{y}$ can have differential <u>and</u> algebraic components.
$\rightarrow$ This makes the numerical solution of DAEs harder when index > 1.

Semi-explicit form is nice because differential / algebraic variables are decoupled.

Note 4. *Any DAE can be written in semi-explicit form with index increased by one as follows:*
For $\mathbf{F}(t, \mathbf{y}, \dot{\mathbf{y}}) = \mathbf{0}$, *let* $\quad \dot{\mathbf{y}} = \mathbf{z}$.

$$\text{Then} \qquad \begin{aligned} \dot{\mathbf{y}} &= \mathbf{z}, \\ \mathbf{0} &= \mathbf{F}(t, \mathbf{y}, \mathbf{z}). \end{aligned}$$

Note 5. *This re-writing doesn't make a given DAE any easier to solve!*

It is also possible to take a semi-explicit index-2 DAE

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

to a fully implicit, index-1 form: Let $\dot{\mathbf{w}} = \mathbf{z}$.

$$\text{Then} \qquad \begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \dot{\mathbf{w}}), \\ \mathbf{0} &= \mathbf{g}(t, \mathbf{x}, \dot{\mathbf{w}}). \end{aligned}$$

$\rightarrow$ These problem classes are equivalent!

Note 6. *Index depends on the solution, not just on what the DAE looks like!*

Example 2.

$$\begin{aligned} \dot{y}_1 &= y_3, \\ 0 &= y_2(1 - y_2), \\ 0 &= y_1 y_2 + y_3(1 - y_2) - t. \end{aligned}$$

Then $\quad y_2(t) = 0 \quad$ *or* $\quad y_2(t) = 1$.
(Suppose $y_2(t)$ is continuous so it does not jump arbitrarily between 0 and 1.)

$$\text{If} \quad y_2(t) = 0: \quad \begin{aligned} y_3(t) &= t, && \text{(verify)} \\ y_1(t) &= \frac{1}{2}t^2 + y_1(0). && \text{(verify)} \end{aligned}$$

→ *index-1 DAE in semi-explicit form.*

$$\text{If} \quad y_2(t) = 1: \quad \begin{aligned} y_1(t) &= t, && \text{(verify)} \\ y_3(t) &= 1. && \text{(verify)} \end{aligned}$$

→ *This is an index-2 DAE.*

Note 7. *No initial values required in this case!*

What if we replace $0 = y_2(1 - y_2)$ with its derivative?

$\rightarrow$

$$\begin{aligned} \dot{y}_1 &= y_3, \\ \dot{y}_2 &= 0, \qquad \longleftarrow \quad \text{verify} \\ 0 &= y_1 y_2 + y_3(1 - y_2) - t. \end{aligned}$$

Now if $y_2(0) = 0$, index = 1;
($y_2(t) = 0$ is the solution case 1.)
if $y_2(0) = 1$, index = 2.
($y_2(t) = 1$ is the solution case 2.)

$\rightarrow$ The index depends on the IC!

- Formal definition of (differential) index:
  For the DAE $\mathbf{F}(t, \mathbf{y}, \dot{\mathbf{y}}) = \mathbf{0}$, the index along solution $\mathbf{y}(t)$ is the minimum number of differentiations required to uniquely solve for $\dot{\mathbf{y}}$ in terms of $\mathbf{y}$ and t;
  i.e., define an ODE for $\mathbf{y}$.

Note 8. *These differentiations are often not done in practice!*

But index helps you determine the difficulty level of the problem and to choose appropriate software.

Note 9. *Recall for IVPs, we had nice theory guaranteeing existence, uniqueness, and continuous dependence on ICs.*
$\rightarrow$ *No such theory held for BVPs.*
$\rightarrow$ *We have no such theory for IV DAEs.*

9.1.1 Special DAE Forms

The general DAE $\mathbf{F}(t, \mathbf{y}, \dot{\mathbf{y}}) = \mathbf{0}$ can include problems that are not well-defined mathematically or cannot be discretized directly (i.e., without reformulation).

Fortunately, DAEs that arise in practice often appear as ODEs with constraints.

$\rightarrow$ Differential and algebraic variables can be identified and treated appropriately.

$\rightarrow$ Algebraic variables can be eliminated (in principle!) with the same number of differentiations.

$\rightarrow$ These are called Hessenberg forms.

- Hessenberg index-1:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{z}), \\ \mathbf{0} &= \mathbf{g}(t, \mathbf{x}, \mathbf{z}), \qquad \frac{\partial \mathbf{g}}{\partial \mathbf{z}} \quad \text{nonsingular} \quad \forall t.\end{aligned}$$

  Also called semi-explicit index-1 DAE.
  $\rightarrow$ Very closely related to implicit ODEs because we can solve (in principle) for $\mathbf{z}$ in terms of $\mathbf{x}, t$ (implicit function theorem).

- Hessenberg index-2:

$$\begin{aligned}\dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{z}), \\ \mathbf{0} &= \mathbf{g}(t, \mathbf{x}), \qquad \frac{\partial \mathbf{g}}{\partial \mathbf{x}}\frac{\partial \mathbf{f}}{\partial \mathbf{z}} \quad \text{nonsingular} \quad \forall t.\end{aligned}$$

Note 10. $\mathbf{g}$ *is independent of* $\mathbf{z}$.

Also called pure index-2 DAE (algebraic variables are index-2 only, not a mixture of index 1 and 2).

Example 3. *Incompressible fluid flow considered as an index-2 DAE:*

$$\begin{aligned}\frac{\partial \mathbf{u}}{\partial t} &= \nu \nabla^2 \mathbf{u} - (\mathbf{u} \cdot \boldsymbol{\nabla})\mathbf{u} - \boldsymbol{\nabla} p, \\ \boldsymbol{\nabla} \cdot \mathbf{u} &= 0. \text{ \textit{(incompressibility condition constraint)}}\end{aligned}$$

$\mathbf{u} = \mathbf{u}(x, y, z, t) =$ *vector of fluid velocities* *differential*
$p = p(x, y, z, t) =$ *pressure* *algebraic*

Applying method of lines yields

$$\begin{aligned}\mathbf{M}\dot{\mathbf{u}} + (\mathbf{K} + \mathbf{N}(\mathbf{u}))\mathbf{u} + \mathbf{C}\mathbf{p} &= \mathbf{f}, \\ \mathbf{C}^T \mathbf{u} &= \mathbf{0}, \\ \mathbf{C}^T \mathbf{M}^{-1} \mathbf{C} \text{ \textit{nonsingular}}&,\end{aligned}$$

where $\mathbf{u}$, $\mathbf{p}$ *are the vectors of the nodal values.*
$\rightarrow$ *An index-2 DAE in Hessenberg form.*

Note 11. *Index-2 variables can be viewed as Lagrange multipliers. Here,* p *forces* $\mathbf{u}$ *onto* $\boldsymbol{\nabla} \cdot \mathbf{u} = 0$.

- Hessenberg index-3:

$$
\begin{aligned}
\dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}, \mathbf{y}, \mathbf{z}), \\
\dot{\mathbf{y}} &= \mathbf{g}(t, \mathbf{x}, \mathbf{y}), \\
\mathbf{0} &= \mathbf{h}(t, \mathbf{y}), \qquad \left(\frac{\partial \mathbf{h}}{\partial \mathbf{y}}\frac{\partial \mathbf{g}}{\partial \mathbf{x}}\frac{\partial \mathbf{f}}{\partial \mathbf{z}}\right) \quad \text{nonsingular} \quad \forall t.
\end{aligned}
$$

Example 4. *Mechanical systems with holonomic constraints:* $\rightarrow$ *Second-order ODEs (Newton's law,* $\mathbf{M}\ddot{\mathbf{x}} = \mathbf{F}$*) subject to constraints (e.g., pendulum arm has fixed length).*

$$
\begin{aligned}
\dot{q}_1 &= v_1, \\
\dot{q}_2 &= v_2, \\
\dot{v}_1 &= -\lambda q_1, \\
\dot{v}_2 &= -\lambda q_2 - g, \\
0 &= q_1^2 + q_2^2 - 1.
\end{aligned}
$$

(verify index 3)

Some important observations about DAEs:

1. In practice, DAEs often arise as limits of singular perturbations of ODEs, e.g.,

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

   as $\epsilon \to 0$.

   The solution of the DAE is then called the reduced solution of the ODE.

2. A higher-index DAE can be the result of a "simplified" lower-index DAE (or ODE) and hence "simpler" to solve.

3. A given DAE can be "close" to another DAE with a different index. So generally one needs more than just the concept of index to quantify stability.

9.1.2 DAE Stability

The definition of index suggests it is a local quantity, subject to the solution relative to which it is defined.

For a rigorous stability analysis, we consider perturbations of linear DAEs and their relationship to index and stability bounds.

(As usual, for stability of nonlinear problems, we form a variational problem.)

We wish to analyze the behaviour of the solution in terms of the perturbations to the data.

Consider the linear ODE

$$\dot{\mathbf{y}} = \mathbf{A}(t)\mathbf{y} + \mathbf{q}(t), \quad 0 < t < 1,$$

where we have normalized t_f to 1 and subject to homogeneous initial or boundary conditions.

Then

$$\|\mathbf{y}\| := \max_{t\in[0,1]} |\mathbf{y}(t)| \leq \kappa \int_0^1 |\mathbf{q}(s)|\, ds = \kappa \|\mathbf{q}\|_1.$$

So for the trivial index-1 DAE $\mathbf{y}(t) = \mathbf{q}(t)$, we have

$$\|\mathbf{y}\| \leq \|\mathbf{q}\|.$$

For the semi-explicit index-1 DAE

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} + \mathbf{q}_1(t) \\ \mathbf{0} &= \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{z} + \mathbf{q}_2(t), \end{aligned}$$

where $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ are bounded functions of t and $\mathbf{D}$ has a bounded inverse, we similarly get

$$\|\mathbf{y}\| \leq \kappa \|\mathbf{q}\|,$$

where $\mathbf{y}^T = (\mathbf{x}^T, \mathbf{z}^T)$ and $\mathbf{q}^T = (\mathbf{q}_1^T, \mathbf{q}_2^T)$.

The (generic) stability constant κ involves bounds on $\mathbf{D}^{-1}$ and the stability constant of the underlying ODE for $\mathbf{x}$ after substituting for $\mathbf{z}$.

This bound can actually be tightened to

$$\|\mathbf{z}\| \leq \kappa \|\mathbf{q}\|, \quad \|\mathbf{x}\| \leq \kappa \|\mathbf{q}\|_1.$$

For the general linear index-1 DAE

$$\mathbf{E}(t)\dot{\mathbf{y}} = \mathbf{A}(t)\mathbf{y} + \mathbf{q}(t)$$

subject to homogeneous initial or boundary conditions, we can write

$$\mathbf{E}(t) = \mathbf{S}(t) \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{T}^{-1}(t),$$

where $\mathbf{S}(t)$, $\mathbf{T}(t)$ are nonsingular matrices with uniformly bounded condition numbers.

Then a change of variables

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{z} \end{pmatrix} = \mathbf{T}^{-1}\mathbf{y}$$

yields a semi-explicit system (see above) (verify!)
and (as before) a stability estimate

$$\|\mathbf{y}\| \leq \kappa \|\mathbf{q}\|,$$

where κ contains condition numbers of the transformations $\mathbf{S}$ and $\mathbf{T}$.

In summary, a linear index-1 DAE is stable if

- it can be transformed into a semi-explicit system and then into an ODE by eliminating the algebraic variables (no differentiations!)
- the transformations are well-conditioned
- the underlying ODE is stable.

For higher-index problems, we must differentiate some of the constraints, in general p differentiations for an index-$(p+1)$ DAE to obtain an index-1 DAE.

Hence, the best case for a stability bound is

$$\|\mathbf{y}\| \leq \kappa \sum_{j=1}^{p+1} \|\mathbf{q}^{(j-1)}\|.$$

Having the bound depend on (high-order) derivatives of $\mathbf{q}$ is not generally desirable.

Fortunately for DAEs in Hessenberg form, this bound can be improved:

For index-2 DAEs in Hessenberg form ($\mathbf{D} \equiv \mathbf{0}$ and $\mathbf{CB}$ nonsingular in the semi-explicit index-1 form), we have

$$\|\mathbf{x}\| \leq \kappa\|\mathbf{q}\|, \quad \|\mathbf{z}\| \leq \kappa\|\dot{\mathbf{q}}\|.$$

Hence, direct discretization of higher-index DAEs other than Hessenberg index-2 may be problematic.

9.2 Index Reduction and Stabilization

Often the best way to solve a high-index DAE is to reduce its index via (analytic) differentiation(s).

The DAE is viewed as an ODE with an invariant.

Recall an index-$(p+1)$ DAE in Hessenberg form (with m ODEs and l constraints) requires p differentiations to eliminate the AEs (and obtain an ODE of size m).

The AEs and their first $p-1$ derivatives form an invariant set (define a constraint manifold) of size pl.

The solution of the ODE is required to remain on (or close to) this set, so the problem really has only $m-pl$ degrees of freedom.

Alternatively, one can imagine using the AEs to eliminate pl of the m unknowns in the original ODE to obtain an ODE with reduced size $m-pl$.

We expand on these ideas in the following subsections.

9.2.1 Reformulation of Higher-Index DAEs

Recall the DAEs that model the motion of mechanical systems:

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{v}, \\ \mathbf{M}(\mathbf{q})\dot{\mathbf{v}} &= \mathbf{f}(\mathbf{q}, \mathbf{v}) - \mathbf{G}^T(\mathbf{q})\boldsymbol{\lambda}, \\ \mathbf{0} &= \mathbf{g}(\mathbf{q}), \end{aligned}$$

where

- $\mathbf{q}$ are generalized body positions,
- $\mathbf{v}$ are generalized velocities,
- $\boldsymbol{\lambda}$ are Lagrange multiplier functions,
- $\mathbf{g}(\mathbf{q})$ are the (holonomic) constraints,
- $\mathbf{G} = \partial \mathbf{g}/\partial \mathbf{q}$ has full row rank for all t,
- $\mathbf{M}$ is a (symmetric) positive definite mass matrix,
- $\mathbf{f}$ are the applied forces.

Let $\mathbf{x}^T = (\mathbf{q}^T, \mathbf{v}^T)$.

Differentiate the constraint twice, first to yield constraints on the velocity level

$$\mathbf{0} = \dot{\mathbf{g}} = \mathbf{G}\mathbf{v}$$

and then on the acceleration level

$$\mathbf{0} = \ddot{\mathbf{g}} = \mathbf{G}\dot{\mathbf{v}} + \frac{\partial(\mathbf{G}\mathbf{v})}{\partial \mathbf{q}}\mathbf{v}.$$

Solving for $\boldsymbol{\lambda}$,

$$\boldsymbol{\lambda} = (\mathbf{G}\mathbf{M}^{-1}\mathbf{G}^T)^{-1}\left(\mathbf{G}\mathbf{M}^{-1}\mathbf{f} + \frac{\partial(\mathbf{G}\mathbf{v})}{\partial \mathbf{q}}\mathbf{v}\right) \qquad \text{(verify!)}$$

we can obtain an ODE for $\mathbf{x} = (\mathbf{q}^T, \mathbf{v}^T)^T$

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{v}, \\ \mathbf{M}\dot{\mathbf{v}} &= \mathbf{f} - \mathbf{G}^T(\mathbf{G}\mathbf{M}^{-1}\mathbf{G}^T)^{-1}\left(\mathbf{G}\mathbf{M}^{-1}\mathbf{f} + \frac{\partial(\mathbf{G}\mathbf{v})}{\partial \mathbf{q}}\mathbf{v}\right). \end{aligned}$$

(verify!)

Note 12. *In practice, the matrix $\frac{\partial(\mathbf{G}\mathbf{v})}{\partial \mathbf{q}}$ is not computed explicitly, rather only its product with $\mathbf{v}$.*

Note 13. *The ODE system has size m and is the result of unstabilized index reduction.*

The constraints on the position and velocity levels define an invariant set of dimension $2l$

$$\mathbf{h}(\mathbf{x}) := \begin{pmatrix} \mathbf{g}(\mathbf{q}) \\ \mathbf{G}(\mathbf{q})\mathbf{v} \end{pmatrix} = \mathbf{0}.$$

The exact solution of the ODE system with consistent ICs $\mathbf{h}(\mathbf{x}(0)) = \mathbf{0}$ satisfies $\mathbf{h}(\mathbf{x}(t)) = \mathbf{0}$ for all $t > 0$.

For the mechanical system, the constraint Jacobian

$$\mathbf{H} = \frac{\partial \mathbf{h}}{\partial \mathbf{x}} = \begin{pmatrix} \mathbf{G} & \mathbf{0} \\ \frac{\partial(\mathbf{G}\mathbf{v})}{\partial \mathbf{q}}\mathbf{v} & \mathbf{G} \end{pmatrix}$$

has full row rank $2l$; so on the constraint manifold, the ODE has dimension $m - 2l$.

Example 5. *For the simple pendulum,*

$$-\lambda = q_2 g - v_1^2 - v_2^2; \qquad \text{(verify!)}$$

so the ODEs to be satisfied are

$$\begin{aligned}
\dot{q}_1 &= v_1,\\
\dot{q}_2 &= v_2,\\
\dot{v}_1 &= (q_2 g - v_1^2 - v_2^2) q_1,\\
\dot{v}_2 &= (q_2 g - v_1^2 - v_2^2) q_2 - g,
\end{aligned}$$

with invariants

$$\begin{aligned}
0 &= q_1^2 + q_2^2 - 1,\\
0 &= q_1 v_1 + q_2 v_2.
\end{aligned}$$

(verify!)

9.2.2 ODEs with Invariants

ODEs with invariants arise in other ways besides just from index reduction of DAEs; e.g., conservation of mass, energy, momentum, charge, etc.

We have seen how DAEs lead to ODEs with invariants.

The relationship goes both ways: the ODE with invariant system

$$\dot{\mathbf{x}} = \hat{\mathbf{f}}(\mathbf{x}),$$
$$\mathbf{h}(\mathbf{x}) = \mathbf{0},$$

is equivalent to the index-2 Hessenberg DAE

$$\begin{aligned} \dot{\mathbf{x}} &= \hat{\mathbf{f}}(\mathbf{x}) - \mathbf{D}(\mathbf{x})\mathbf{z}, \\ \mathbf{0} &= \mathbf{h}(\mathbf{x}), \end{aligned}$$

where $\mathbf{D}(\mathbf{x})$ is any bounded matrix function such that $\mathbf{HD}$ has a bounded inverse for all t.

These two systems share the same exact solution, specifically $\mathbf{z}(t) \equiv \mathbf{0}$, but a numerical solution will generally not satisfy this (unless it is exact).

But this does not mean that the Hessenberg index-2 DAE is (precisely) the same as the DAE that may have led to the ODE with invariant.

Note 14. $\mathbf{D}$ *defines the direction of the projection onto the constraint manifold.*

A common choice is orthogonal projection, $\mathbf{D} = \mathbf{H}^T$*.*

Sometimes, it may be OK in practice to simply integrate the ODE and have it turn out that $\mathbf{h}(\mathbf{x}) \approx \mathbf{0}$.

But other times it is not, e.g., if the problem is not stable off the manifold or $\mathbf{h}(\mathbf{x})$ must be zero to within roundoff errors not just to within truncation errors.

It turns out instability off the manifold is a typical consequence of index reduction.

To see this, consider a (nonsingular) change of variables

$$\mathbf{q} \to \begin{pmatrix} \tilde{\mathbf{q}}_1 \\ \tilde{\mathbf{q}}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{g}(\mathbf{q}) \\ \mathbf{g}^{\perp}(\mathbf{q}) \end{pmatrix},$$

such that $\partial \mathbf{g}^{\perp}/\partial \mathbf{q}$ is orthogonal to $\mathbf{G}^T$; i.e., the new variables are the constraints themselves and their orthogonal complement.

Differentiating the constraints twice, we obtain

$$\ddot{\tilde{\mathbf{q}}}_1 = \mathbf{0}.$$

This leads to a mild instability known as drift.

If $\tilde{\mathbf{q}}_1(0) = \boldsymbol{\epsilon}_1$ and $\dot{\tilde{\mathbf{q}}}_1(0) = \boldsymbol{\epsilon}_2$, then

$$\tilde{\mathbf{q}}_1(t) = \boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2 t,$$
$$\dot{\tilde{\mathbf{q}}}_1(t) = \boldsymbol{\epsilon}_2.$$

Theorem 1. *For a given time t_n, the numerical solution $\tilde{\mathbf{q}}_n$ obtained from a numerical method of order p applied to the index-1 formulation of an index-3 DAE and consistent initial conditions satisfies*

$$\|\tilde{\mathbf{q}}_n\| \leq (\Delta t)^p (A t_n + B t_n^2), \quad \|\dot{\tilde{\mathbf{q}}}_n\| \leq (\Delta t)^p C t_n,$$

where A, B, and C are constants and Δt is the maximum step size used.

We see violation of the original constraint grow quadratically in time.

Drift is the result of differentiation!

It is not present in the original DAE (and hence also not in the equivalent ODE restricted to the manifold).

Proof:

Let $\tilde{\mathbf{q}}(t; \tilde{\mathbf{q}}_j)$ be the exact solution at time t to the index-1 problem with initial value $\tilde{\mathbf{q}}_j$ at time $t = t_j$.

Let $\tilde{\mathbf{q}}_n(\tilde{\mathbf{q}}_j)$ be the numerical solution at time t_n to the index-1 problem with initial value $\tilde{\mathbf{q}}_j$ at time $t = t_j$.

Let $\tilde{\mathbf{q}}_0 = \mathbf{0}$ at the initial time $t = t_0$.

The local error satisfies

$$\tilde{\mathbf{q}}(t_n; \tilde{\mathbf{q}}_{n-1}) - \tilde{\mathbf{q}}_n(\tilde{\mathbf{q}}_{n-1}) = \mathcal{O}((\Delta t)^{p+1}).$$

Because $\ddot{\tilde{\mathbf{q}}} = \mathbf{0}$,

$$\begin{aligned} &\|\tilde{\mathbf{q}}(t_n; \tilde{\mathbf{q}}_{j+1}) - \tilde{\mathbf{q}}(t_n; \tilde{\mathbf{q}}_j)\| \\ &\qquad\qquad \leq (\Delta t_j)^{p+1}(A + 2B(t_n - t_{j+1})) \end{aligned}$$

Adding up these inequalities for $j = 0$ to $n - 1$ gives the desired bound on $\|\tilde{\mathbf{q}}_n\|$.

Rather than converting the ODE to a DAE, we consider stabilizing the ODE with respect to the invariant set

$$\mathcal{M} = \{\mathbf{x} : \mathbf{h}(\mathbf{x}) = \mathbf{0}\}.$$

On $\mathcal{M}$ (i.e., when $\mathbf{h}(\mathbf{x}) = \mathbf{0}$), the ODE

$$\dot{\mathbf{x}} = \hat{\mathbf{f}}(\mathbf{x}) - \gamma \mathbf{F}(\mathbf{x})\mathbf{h}(\mathbf{x})$$

has the same solution as

$$\dot{\mathbf{x}} = \hat{\mathbf{f}}(\mathbf{x}).$$

$\mathcal{M}$ will also be attracting if the matrix $\mathbf{HF}$ is positive definite and $\gamma > 0$ is sufficiently large.

The stabilized system may not even be stiff!

This leads to the potential of using non-stiff solvers instead of being condemned to using stiff ones when solving a DAE.

Example 6. *Consider the simple pendulum subject to the ICs* $\mathbf{q}(0) = (1, 0)^T$ *and* $\mathbf{v}(0) = (0, -5)^T$*.*

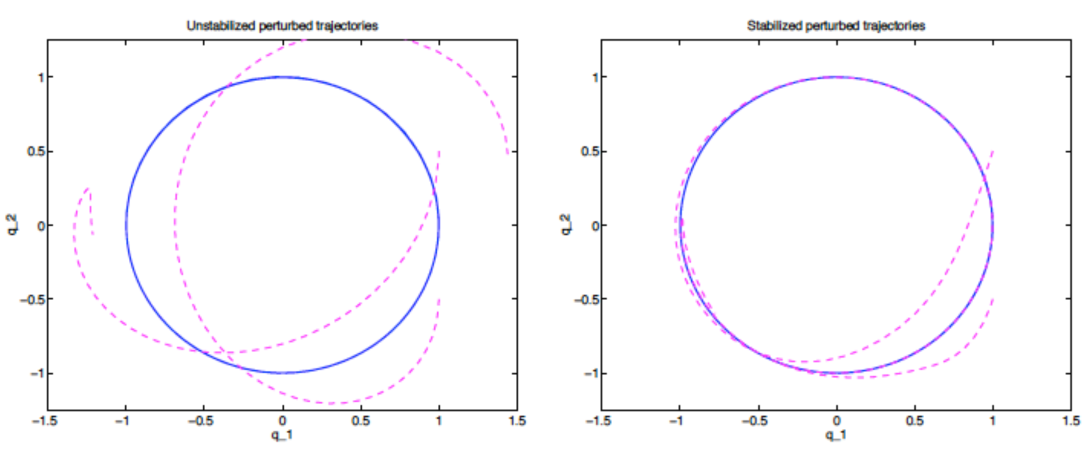
This problem is simple enough that a solution obtained with `ode45` *remains on the unit circle in* $\mathbf{q}$*-space for one orbit to 4 significant figures.*

If we perturb the IC for $\mathbf{q}$ *to* $\mathbf{q}(0) = (1, \pm 0.5)^T$ *and repeat the calculation, we see the manifold is unstable.*

If however we use $\gamma = 10$ *and*

$$\mathbf{D}^T = \mathbf{H} = \begin{pmatrix} 2q_1 & 2q_2 & 0 & 0 \\ v_1 & v_2 & q_1 & q_2 \end{pmatrix},$$

we see the manifold is stable.



(a) Unstabilized pendulum equations

(b) Stabilized pendulum equations, $\gamma = 10$

One of the first stabilization methods is by Baumgarte.

Baumgarte stabilization writes

$$\mathbf{0} = \ddot{\mathbf{g}} + \gamma_1 \dot{\mathbf{g}} + \gamma_2 \mathbf{g},$$

then tries to choose parameters γ_1, γ_2 such that

$$x^2 + \gamma_1 x + \gamma_2 = 0$$

has roots with negative real parts, making the ODE for $\mathbf{g}$ stable, and thus stabilizing the invariant set.

Unfortunately, the choice of γ_1, γ_2 is tricky in practice.

9.2.3 State-Space Formulation

Differentiating the constraints of an index-$(p+1)$ DAE yields an ODE with a larger size (higher dimension).

The number of degrees of freedom remains $m-pl$, but we have m ODEs and pl AEs.

An alternative to invariant stabilization is to use the AEs to define a reduced ODE of size $m-pl$.

The advantages include the reduced ODE system size as well as the guarantee that there is no drift (the constraints are satisfied automatically).

The drawbacks center around the process of eliminating the constraints.

Constraints often contain strong nonlinearities.

This not only makes the process complex and expensive, it must be updated (the same elimination procedure may not always work) and monitored (it may become singular) as the solution proceeds.

9.3 Modelling with DAEs

Many researchers are awakening to the fact that their favourite mathematical models are (better) described as DAEs instead of ODEs.

This revelation can lead to new and fruitful perspectives on an otherwise old problem.

But there is a tradeoff in the formulation, analysis, and solution of the DAE formulations
$\rightarrow$ now less familiar / straightforward.

High-index DAEs are unstable, so direct discretization generally leads to disaster
$\rightarrow$ some reformulation is required (can be costly).

It is important to keep in mind that for a semi-explicit DAE, the ODE is solved to within truncation errors whereas the constraints are solved to roundoff errors.

$\rightarrow$ more importance on the constraints (invariants)!

Sometimes this makes sense, but not always.

Example 7. *Consider the solution of a PDE by the method of lines.*

Normally, the mesh used in the method of lines to convert the PDE to a system of coupled ODEs is fixed.

Of course, we appreciate the limitations of using a fixed mesh, so one can imagine a method of lines for which the mesh used for the spatial discretization automatically adapts to concentrate in the areas where it is most needed as the solution evolves in time.

Moving mesh methods achieve this by equidistributing some measure of the solution (e.g., arc length) on each subinterval of the mesh.

However, strictly enforcing this equidistribution as an algebraic constraint does not guarantee solution accuracy, so in some sense a lot of effort is expended on a non-essential goal
$\rightarrow$ *satisfying such a constraint only approximately often more than suffices.*