# Module 4F3: Nonlinear and Predictive Control
# Solutions 2009

J.M. Maciejowski

13 May 2009

1. **(a)**

   **(i)** $x_e$ is an equilibrium point if $f(x_e) = 0$.

   **(ii)** An equilibrium $x_e$ is stable if $\forall \epsilon > 0\ \exists \delta > 0$ s.t. $\|x(0) - x_e\| < \delta$ implies $\|x(0) - x_e\| < \epsilon\ \ \forall\ t \geq 0$.

   **(iii)** An equilibrium $x_e$ is asymptotically stable if it is stable and $\exists \delta > 0$ s.t. $\|x(0) - x_e\| < \delta$ implies $\lim_{t\to\infty} x(t) = x_e$ .

   **(iv)** Domain of attraction of an asymptotically stable equilibrium point $x_e$ is the set $S$ of initial conditions s.t. if $x(0) \in S$ then $x(t) \to x_e$ as $t \to \infty$ .

   **(b)**

   **(i)** First verify that $x = 0$ is an equilibrium point. Note that this is unique for $|x_i| < \alpha$, $i = 1, 2, 3$. It is clear that $V(0) = 0$.

   Note that, providing that $|x_i| < \alpha$ for $i = 1, 2, 3$, each of the integrals appearing in $V$ is nonnegative, and positive if $|x_i| > 0$. Thus $V > 0$ in some neighbourhood of 0.

$$
\begin{aligned}
\dot{V} = \nabla V(x)\dot{x} &= \frac{\partial V}{\partial x_1}\dot{x}_1 + \frac{\partial V}{\partial x_2}\dot{x}_2 + \frac{\partial V}{\partial x_3}\dot{x}_3 \\
&= f(x_1)[-x_1 + h(x_3)] + g(x_2)[-h(x_3)] + h(x_3)[-f(x_1) + g(x_2) - h(x_3)] \\
&= -x_1 f(x_1) - h^2(x_3) \leq 0
\end{aligned}
$$

   Choose $c > 0$ sufficiently small s.t. $S := \{x : V(x) \leq c\} \subset \{x : |x_i| < \alpha \text{ for } i = 1, 2, 3\}$. $S$ is a closed and bounded invariant set, hence from Lasalle's theorem for any $x(0) \in S$, $x(t) \to M$ as $t \to \infty$, where $M$ is the largest invariant set in $S$ included in $\{x : \dot{V}(x) = 0\}$.

$$\dot{V}(t) = 0 \Rightarrow x_1(t) = 0, x_3(t) = 0$$

   If $x_2(t) \neq 0$ then $\exists \tau$ s.t. $x_3(t + \tau) \neq 0$ hence $M$ only includes the origin.

   **(ii)** The Jacobian is given by

$$
A = \begin{bmatrix} -1 & 0 & \frac{\partial h}{\partial x_3} \\ 0 & 0 & -\frac{\partial h}{\partial x_3} \\ -\frac{\partial f}{\partial x_1} & \frac{\partial g}{\partial x_2} & -\frac{\partial h}{\partial x_3} \end{bmatrix}
$$

   where the derivatives are evaluated at the equilibrium point. The eigenvalues $\lambda$ of $A$ are the solutions of $|\lambda I - A| = 0$ i.e.

$$
\begin{vmatrix} \lambda + 1 & 0 & -\frac{\partial h}{\partial x_3} \\ 0 & \lambda & \frac{\partial h}{\partial x_3} \\ \frac{\partial f}{\partial x_1} & -\frac{\partial g}{\partial x_2} & \lambda + \frac{\partial h}{\partial x_3} \end{vmatrix} = 0
$$

which gives at the origin $(\lambda+1)\lambda(\lambda+\frac{\partial h}{\partial x_3})=0$ (also follows by inspection of $A$ since $\partial f/\partial x_1$ and $\partial g/\partial x_2$ are both zero at the origin in this case). There is hence at least one pole on the imaginary axis $\Rightarrow$ linearization inconclusive.

**(iii)** Is the origin also globally asymptotically stable?

Not necessarily, e.g. other equilibria could be present for $x_i > \alpha$ depending on the form of the functions $f, g, h$.

*Note: even if the conditions specified for $f, g, h$ hold for all $|y| > 0$ the set $S$ would not necessarily be bounded for all $c > 0$, hence the analysis above would not be sufficient to conclude that $S$ is included in the domain of attraction of the origin for all $c > 0$ (the latter would be the case if e.g. $f, g, h$ are additionally non decreasing functions).*

---

2. **(a)** Consider $e = E\sin\theta$. If $E \le \delta$, $f(e) = e/\delta$, hence $N_1(E) = 1/\delta$. If $E > \delta$

$$N_1(E) = \frac{U_1 + jV_1}{E}$$

$V_1 = 0$ since $f(e)$ is an odd function.

$$\begin{aligned}
U_1 &= \frac{1}{\pi}\int_0^{2\pi} f(E\sin\theta)\sin\theta d\theta = \frac{4}{\pi}\int_0^{\pi/2} f(E\sin\theta)\sin\theta d\theta \\
&= \frac{4}{\pi}\int_0^{\sin^{-1}(\delta/E)} \frac{1}{\delta}E\sin^2\theta d\theta + \frac{4}{\pi}\int_{\sin^{-1}(\delta/E)}^{\pi/2} \sin\theta d\theta \\
&= \frac{4E}{\pi\delta}\int_0^{\sin^{-1}(\delta/E)} \frac{1-\cos(2\theta)}{2} d\theta - \frac{4}{\pi}[\cos\theta]_{\sin^{-1}(\delta/E)}^{\pi/2} \\
&= \frac{2E}{\pi\delta}\left[\theta - \frac{\sin(2\theta)}{2}\right]_0^{\sin^{-1}(\delta/E)} + \frac{4}{\pi}\cos(\sin^{-1}(\delta/E)) \\
&= \frac{2E}{\pi\delta}\left[\sin^{-1}(\delta/E) - \frac{1}{2}\sin(2\sin^{-1}(\delta/E))\right] + \frac{4}{\pi}\cos(\sin^{-1}(\delta/E)) \\
&= \frac{2E}{\pi\delta}\left[\sin^{-1}(\delta/E) - \sin(\sin^{-1}(\delta/E))\cos(\sin^{-1}(\delta/E))\right] + \frac{4}{\pi}\cos(\sin^{-1}(\delta/E)) \\
&= \frac{2E}{\pi\delta}\sin^{-1}(\delta/E) + \frac{2}{\pi}\cos(\sin^{-1}(\delta/E)) \\
&= \frac{2E}{\pi\delta}\sin^{-1}(\delta/E) + \frac{2}{\pi}\sqrt{1-(\delta/E)^2}
\end{aligned}$$

Hence $N_1(E)$ is as required.

**(b)** For $\delta = 1$ we have $g(e) = e - f(e)$. So $N_2(E) = 1 - N_1(E)$ , i.e.

$$N_2(E) = \begin{cases} 0, & \text{if } E \le 1 \\ 1 - \frac{2}{\pi}\left[\sin^{-1}\left(\frac{1}{E}\right) + \frac{1}{E}\sqrt{1-\left(\frac{1}{E}\right)^2}\right] & \text{if } E > 1 \end{cases}$$

**(c)** $f$ is an odd function. Therefore

$$N_1(E) = \frac{U_1}{E} = \frac{1}{\pi}\int_0^{2\pi} \frac{f(E\sin\theta)}{E}\sin\theta d\theta \ge 0$$

Also, since $f(E\sin\theta) \le E\sin\theta/\delta$, we have

$$N_1(E) \le \frac{1}{\pi}\int_0^{2\pi}\frac{1}{\delta}\sin^2\theta d\theta = \frac{1}{\pi\delta}\int_0^{2\pi}\frac{1-\cos(2\theta)}{2}d\theta = \frac{1}{\pi\delta}\left[\frac{\theta}{2} - \frac{\sin(2\theta)}{4}\right]_0^{2\pi} = \frac{1}{\delta}$$

*Alternatively, a less formal argument invoking the concept of 'equivalent linear gain' is acceptable here: from the form of the nonlinearity, it is clear that $N_1(E)$ does not increase with $E$. Hence its largest value is that for $E \leq \delta$. Also an argument based on showing that $d\{N_1(E)\}/dE \leq 0$ could be used.*

**(d)**

$$G(j\omega) = \frac{k}{(j\omega+1)^2} = k\frac{(1-j\omega)^2}{(\omega^2+1)^2} = k\frac{1-\omega^2}{(1+\omega^2)^2} - k\frac{2\omega j}{(1+\omega^2)^2}$$

**(i)** $\Im[G(j\omega)] = 0$ for $\omega = 0$ or $\omega \to \infty$, hence no intersections with negative real axis, therefore no limit cycle from describing function method (using (c)).

**(ii)** To deduce stability from circle criterion need $\Re[G(j\omega)] > -\delta \ \forall\omega$.

$$\frac{\partial\Re[G(j\omega)]}{\partial\omega^2} = \frac{-(1+\omega^2)^2 - 2(1+\omega^2)(1-\omega^2)}{(1+\omega^2)^4} = 0 \Rightarrow -(1+\omega^2) - 2(1-\omega^2) = 0 \Rightarrow \omega^2 = 3$$

So $\min_\omega \Re[G(j\omega)] = -2k/16 = -k/8$ . $\Re[G(j\omega)]$ is maximized when $\omega^2$ is minimized, i.e. $\omega = 0$. So $-\frac{k}{8} \leq \Re[G(j\omega)] \leq k$. So need $-\delta < k < 8\delta$ .

---

3. (a) If a linear model is used (as in the standard formulation of MPC), then linear inequality constraints of the form $MX \leq m$, applied to the predicted states, transform into linear inequality constraints on the predicted inputs, which are the decision variables of the optimization problem that is solved in MPC. If a convex optimization criterion is used, such as a quadratic cost (which is the standard formulation) or a linear cost (absolute values or peak values), then the resulting optimization problem is convex. Since the optimization problem has to be solved on-line, it is important to solve a convex problem if possible, since that guarantees that a solution will be found if a 'descent' search strategy is used, and that this solution will be a global optimum of the problem.

(b) A constraint of the form $|x^i| \leq \ell_i$ can be written as two linear inequalities:

$$x^i \leq \ell_i \quad \text{and} \quad -x^i \leq \ell_i \tag{1}$$

which can be written as

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} x^i \leq \begin{bmatrix} \ell_i \\ \ell_i \end{bmatrix} \tag{2}$$

This is now written for every predicted state in the prediction horizon:

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} x^i_s \leq \begin{bmatrix} \ell_i \\ \ell_i \end{bmatrix} \tag{3}$$

for $s = 1, 2, \ldots, N$. Since $x^i_s$ appears in the vector $X$, the inequalities (3) can be included in the set of inequalities $MX \leq m$ by inserting the coefficients on the left and right hand sides of (3) in the appropriate entries of $M$and $m$.

(c) Following the above, the inequality $|\dot{z}| \leq 0.01$ is written as

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} \dot{z} \leq \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix} \tag{4}$$

and the inequality $|z| \leq 0.1$ is written as

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix} z \leq \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} \tag{5}$$

Now applying these constraints over the prediction horizon, and writing them in terms of the complete predicted state vector, gives:

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ \dot{y}_1 \\ z_1 \\ \dot{z}_1 \\ y_2 \\ \dot{y}_2 \\ z_2 \\ \dot{z}_2 \end{bmatrix} \leq \begin{bmatrix} 0.1 \\ 0.1 \\ 0.01 \\ 0.01 \\ 0.1 \\ 0.1 \\ 0.01 \\ 0.01 \end{bmatrix} \tag{6}$$

(d) If $x_0$ is the latest measurement of the state vector, we have predictions (since $N = 2$):

$$x_1 = Ax_0 + Bu_0 \tag{7}$$
$$x_2 = Ax_1 + Bu_1 \tag{8}$$
$$= A^2x_0 + ABu_0 + Bu_1 \tag{9}$$

which can be written as

$$X = \begin{bmatrix} A \\ A^2 \end{bmatrix} x_0 + \begin{bmatrix} B & 0 \\ AB & B \end{bmatrix} U \tag{10}$$

where $U = [u_0^T, u_1^T]^T$. Hence the inequalities $MX \leq m$ can be expressed as

$$M\left(\begin{bmatrix} A \\ A^2 \end{bmatrix} x_0 + \begin{bmatrix} B & 0 \\ AB & B \end{bmatrix} U\right) \leq m \tag{11}$$

---

4. (a) The question asks for the principle rather than details, so the essential ingredients that should be mentioned are

- An internal model, used for prediction,
- A cost function, which is minimised at each step,
- Constraints which should not be violated,
- The receding horizon idea,
- New measurements bringing in feedback action.

*Benefits:*

- Constraints can be considered explicitly,
- Easy to understand,
- Deals easily with time delays,
- Allows operation close to constraints,
- Adaptation easily implemented, eg by changing the model.

*Disadvantages:*

- On-line computational complexity,
- Lack of transparency of behaviour,

(b) i. From the definition of $V$, we have

$$\begin{aligned}
V(Ax_0 + Bu_0^*, 0) &= (Ax_0 + Bu_0^*)^T Q(Ax_0 + Bu_0^*) + 0 + \\
&\quad +(A^2x_0 + ABu_0^*)^T P(A^2x_0 + ABu_0^*) && (12)\\
&= (x_0^T A^T Q A x_0 + 2x_0^T A^T Q B u_0^* + u_0^{*T} B^T Q B u_0^*) + \\
&\quad +(x_0^T A^{2T} P A^2 x_0 + 2x_0^T A^{2T} P A B u_0^* + u_0^{*T} B^T A^T P A B u_0^*) && (13)\\
&= x_0^T A^T (Q + A^T P A) A x_0 + 2x_0^T A^T (Q + A^T P A) B u_0^* + \\
&\quad +u_0^{*T} B^T (Q + A^T P A) B u_0^* && (14)\\
&= x_0^T A^T P A x_0 + 2x_0^T A^T P B u_0^* + u_0^{*T} B^T P B u_0^* && (15)
\end{aligned}$$

where in the last line we have used the fact that $P = A^T P A + Q$.
But

$$\begin{aligned}
V^*(x_0) &= x_0^T Q x_0 + u_0^{*T} R u_0^* + (Ax_0 + Bu_0^*)^T P (Ax_0 + Bu_0^*) && (16)\\
&= x_0^T Q x_0 + u_0^{*T} R u_0^* + (x_0^T A^T P A x_0 + 2x_0^T A^T P B u_0^* + u_0^{*T} B^T P B u_0^*) && (17)\\
&= x_0^T (Q + A^T P A) x_0 + 2x_0^T A^T P B u_0^* + u_0^{*T} (R + B^T P B) u_0^* && (18)\\
&= x_0^T P x_0 + 2x_0^T A^T P B u_0^* + u_0^{*T} (R + B^T P B) u_0^* && (19)
\end{aligned}$$

where in the last line we have again used the fact that $P = A^T P A + Q$.
Now comparing (15) and (19) we see that

$$\begin{aligned}
V^*(x_0) &= V(Ax_0 + Bu_0^*, 0) + x_0^T (P - A^T P A) x_0 + u_0^{*T} R u_0^* && (20)\\
&= V(Ax_0 + Bu_0^*, 0) + x_0^T Q x_0 + u_0^{*T} R u_0^* && (21)\\
&> V(Ax_0 + Bu_0^*, 0) && (22)
\end{aligned}$$

if $x_0 \neq 0$, since $Q > 0$ and $R > 0$.
But

$$V^*(Ax_0 + Bu_0^*) = \min_u V(Ax_0 + Bu^*, u) \leq V(Ax_0 + Bu_0^*, 0) \qquad (23)$$

so $V^*(Ax_0 + Bu_0^*) < V^*(x_0)$. QED

ii. The idea is that we can use $V^*$ (the *value function*) as a Lyapunov function. In discrete-time systems a Lyapunov function is one which decreases at each step, and has a minimum at an equilibrium. We have just shown that $V^*$ has the decreasing property. Clearly we have $V(0,0) = 0$, $V(x,u) > 0$ if $x \neq 0$ or $u \neq 0$, and $(x = 0, u = 0)$ is an equilibrium of the system. Hence $V^*(0) = 0$ and $V^*(x) > 0$ if $x \neq 0$. The other condition that needs to be established is the continuity of $V^*$ — this is harder, and not covered in the course.

iii. Consider $V(k) = x(k)^T P x(k)$. For the system $x(k+1) = Ax(k)$ we have

$$\begin{aligned}
V(k+1) - V(k) &= x(k+1)^T P x(k+1) - x(k)^T P x(k) && (24)\\
&= x(k)^T (A^T P A - P) x(k) && (25)\\
&= -x(k)^T Q x(k) && (26)\\
&< 0 && (27)
\end{aligned}$$

so $V$ is a (discrete-time) Lyapunov function for the open-loop system, and so the open-loop system must be stable.

---

# Module 4F3: Nonlinear and Predictive Control
## Answers to 2009 exam.

1. (b)(iii) Not necessarily.

2. (b) $N_2(E) = 1 - N_1(E)$. (d)(i) No. (d)(ii) $-\delta < k < 8\delta$.