Let \( p \) be a polynomial of degree at most \( N \) and let 
\[-1 < \tau_1 < \tau_2 < \ldots < \tau_N < 1\]
be the Gauss quadrature points. Suppose that \( p(-1) = 0 \) and \( |p'(\tau_i)| \leq 1 \) for all \( 1 \leq i \leq N \). Show that \( |p(\tau_i)| \leq 2 \) for all \( 1 \leq i \leq N \).

This gives \( \|D^{-1}\|_\infty \leq 2 \): Let \( p \in \mathcal{P}_N \) with \( p(-1) = 0 \). Let \( \mathbf{p} \) be the vector with components \( p_i = p(\tau_i) \), and let \( \dot{\mathbf{p}} \) be the vector with components \( \dot{p}_i = \dot{p}(\tau_i) \). Since \( D\mathbf{p} = \dot{\mathbf{p}} \), we have \( \mathbf{p} = D^{-1}\dot{\mathbf{p}} \). Since

\[
\|D^{-1}\|_\infty = \max\{\|D^{-1}\mathbf{y}\|_\infty : -1 \leq \mathbf{y} \leq 1\},
\]

it follows that \( \|D^{-1}\|_\infty \) is the maximum value \( |p(\tau_i)| \) over all polynomials for which \( p(-1) = 0 \) and \( |\dot{p}(\tau_i)| \leq 1 \).

NOTE: the polynomial that achieves the maximum value for \( |p(\tau_i)| \) is \( p(\tau) = 1 + \tau \).
minimize \[ C(x(1)) \] 

subject to \[ \dot{x}(t) = f(x(t), u(t)), \quad u(t) \in \mathcal{U}, \quad t \in \Omega_0, \]

\[ x(0) = x_0. \]

- \( \Omega_0 = [0, 1], \) \( x_0 \) given, \( x(t) \in \mathbb{R}^n, \)
- \( \mathcal{U} \subset \mathbb{R}^m \) closed and convex,
- \( f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \) and \( C : \mathbb{R}^n \to \mathbb{R} \)
Partition time domain $\Omega_0$ into $K$ subintervals and use different polynomial to approximate state on each subinterval. Require continuity across subintervals. Change variables so that $x_k(\tau), \tau \in [-1, 1]$ corresponds to $x(t), t \in [t_{k-1}, t_k]$.

\[
h = 1/2K = \frac{t_k-t_{k-1}}{2} \quad \tau \in [-1, 1]
\]

Continuous: $x_k(\tau) = x(t_{k-1/2} + h\tau)$

Discrete: $x_k \in P_N$
Radau Collocation Points

- Radau quadrature points (zeros of $P_{N-1}^{1,0}$ plus $\tau_N = 1$):
  
  $$-1 < \tau_1 < \tau_2 < \ldots < \tau_N = +1.$$  

  Additional points in analysis: $\tau_0 = -1$

- Quadrature weights:
  
  $$\omega_i = \frac{2(1 + \tau_i)}{[(1 - \tau_i^2)P_{N-1}^{1,0}(\tau_i)]^2}, \quad 1 \leq i \leq N - 1, \quad \omega_N = \frac{2}{N^2}.$$  

  For every $p \in P_{2N-2}$:
  
  $$\int_{-1}^{1} p(\tau) d\tau = \sum_{i=1}^{N} \omega_i p(\tau_i).$$  

N = 20 Radau Points
Continuous and Discrete Problems on $K$ Mesh Intervals

Continuous:

minimize $C(x_K(1))$

subject to $\dot{x}_k(\tau) = hf(x_k(\tau), u_k(\tau)), \quad u_k(\tau) \in U, \quad \tau \in \Omega,$

$x_k(-1) = x_{k-1}(1), \quad 1 \leq k \leq K.$

Discrete:

minimize $C(x_K(1))$

subject to $\dot{x}_k(\tau_i) = hf(x_k(\tau_i), u_{ki}), \quad u_{ki} \in U, \quad 1 \leq i \leq N,$

$x_k(-1) = x_{k-1}(1), \quad x_k \in \mathcal{P}_N^n, \quad 1 \leq k \leq K.$
Lagrange interpolating polynomials: For $0 \leq j \leq N$,

$$
\Phi_i(\tau) = \prod_{j=0}^{N} \frac{\tau - \tau_j}{\tau_i - \tau_j}, \quad \Phi_i(\tau_j) = \begin{cases} 1 & \text{for } j = i \\ 0 & \text{otherwise.} \end{cases}
$$

If $x_k \in \mathcal{P}_N$, then $x_k(\tau) = \sum_{j=0}^{N} X_{kj}(\tau_j)\Phi_j(\tau), \quad X_{kj} = x_k(\tau_j)$.

Continuity $\iff X_{k-1,N} = X_{k0}$.

Differentiation matrix $D \in \mathbb{R}^{N \times (N+1)}$

$$
\dot{x}(\tau_i) = \sum_{j=0}^{N} \dot{\Phi}_j(\tau_i)x(\tau_j) = \sum_{j=0}^{N} D_{ij}x(\tau_j), \quad D_{ij} = \dot{\Phi}_j(\tau_i)
$$
Discrete Control Problem

Hence, in terms of the state and control values $X_{kj}$ and $U_{kj}$ at the $N$ collocation points on each of the $K$ intervals, the discrete control problem can be formulated as

$$\text{minimize} \quad C(X_{KN})$$

subject to

$$\sum_{j=0}^{N} D_{ij} X_{kj} = h_f(X_{ki}, U_{ki}), \quad U_{ki} \in \mathcal{U}, \quad 1 \leq i \leq N,$$

$$X_{k0} = X_{k-1,N}, \quad 1 \leq k \leq K,$$

where $X_{0N} = X_0$ is the starting condition.
\[
\mathcal{L}(\lambda, X, U) = C(X_{KN}) + \\
\sum_{k=1}^{K} \sum_{i=1}^{N} \left< \lambda_{ki}, hf(X_{ki}, U_{ki}) - \sum_{j=0}^{N} D_{ij} X_{kj} \right> + \\
\sum_{k=1}^{K} \left< \lambda_{k0}, (X_{k-1,N} - X_{k0}) \right>.
\]

Differentiating with respect to each of the variables:

\(X_{k0} \Rightarrow \sum_{i=1}^{N} D_{i0} \lambda_{ki} = -\lambda_{k0},\)

\(X_{kj} \Rightarrow \sum_{i=1}^{N} D_{ij} \lambda_{ki} = h \nabla_x H(X_{kj}, U_{kj}, \lambda_{kj}), \quad 1 \leq j < N,\)

\(X_{kN} \Rightarrow \sum_{i=1}^{N} D_{iN} \lambda_{ki} = h \nabla_x H(X_{kN}, U_{kN}, \lambda_{kN}) + \lambda_{k+1,0},\)

\(\lambda_{K+1,0} := \nabla C(X_{KN}),\)

\(U_{ki} \Rightarrow -\nabla_u H(X_{ki}, U_{ki}, \lambda_{ki}) \in N_U(U_{ki}).\)
Lemma. The multipliers $\lambda_k \in \mathbb{R}^{Nn}$, $1 \leq k \leq K$, satisfy the KKT conditions if and only if the polynomial $\lambda_k \in \mathcal{P}_{n-1}^n$ given by $\lambda_k(\tau_i) = \frac{\lambda_{ki}}{\omega_i}$, $1 \leq i \leq N$, satisfies the following conditions and $\lambda_{k0} = \lambda_k(-1)$.

$$\dot{\lambda}_k(\tau_i) = -h \nabla_x H(x_k(\tau_i), u_{ki}, \lambda_k(\tau_i)), \quad 1 \leq i < N,$$

$$\dot{\lambda}_k(1) = -h \nabla_x H(x_k(1), u_{kN}, \lambda_k(1)) + (\lambda_k(1) - \lambda_{k+1}(-1)) / \omega_N,$$

where $\lambda_{K+1}(-1) := \nabla C(x_K(1))$

$$\mathcal{N}(u_{ki}) \ni -\nabla_u H(x_k(\tau_i), u_{ki}, \lambda_k(\tau_i)), \quad 1 \leq i \leq N.$$

NOTE: The discrete $\lambda$ is typically discontinuous at the mesh points.
The differentiation matrix $D^\dagger$ for the space of polynomials of degree $N - 1$ evaluated at $\tau_i$, $1 \leq i \leq N$, is given by

$$D_{NN}^\dagger = -D_{NN} + \frac{1}{\omega_N} \quad \text{and} \quad D_{ij}^\dagger = -\frac{\omega_j}{\omega_i} D_{ji} \quad \text{otherwise.}$$

In other words, if $p$ is a polynomial of degree at most $N - 1$ and if $p \in \mathbb{R}^N$ is the vector with $i$-th component $p_i = p(\tau_i)$, then

$$(D^\dagger p)_i = \dot{p}(\tau_i), \quad 1 \leq i \leq N.$$
Proof

If \( p \in \mathcal{P}_N \) and \( q \in \mathcal{P}_{N-1} \) with \( p(-1) = 0 \), then integration by parts gives

\[
\int_{-1}^{1} \dot{p}(\tau)q(\tau)d\tau = p(1)q(1) - \int_{-1}^{1} p(\tau)\dot{q}(\tau)d\tau.
\]

Since \( \dot{pq} \) and \( pq \) are polynomials of degree at most \( 2N - 2 \), Radau quadrature is exact and we have

\[
\sum_{j=1}^{N} w_j \dot{p}_j q_j = p_N q_N - \sum_{j=1}^{N} w_j p_j \dot{q}_j,
\]

where \( p_j = p(\tau_j) \) and \( \dot{p}_j = \dot{p}(\tau_j) \). This yields

\[
(W\dot{p})^T q = p_N q_N - (Wp)^T \dot{q}.
\]

Substituting \( \dot{p} = D_{1:N}p \) and \( \dot{q} = D^\dagger q \) gives

\[
p^T D_{1:N}^T Wq = p_N q_N - p^T WD^\dagger q.
\]
Rearrange into the following form:

$$p^T(D^{T}_{1:N}W + WD^\dagger - e_Ne_N^T)q = 0,$$

where $e_N$ is the last column of $I$. Since this identity must be satisfied for all choices of $p$ and $q$, we deduce that

$$D^\dagger = W^{-1}e_Ne_N^T - W^{-1}D^{T}_{1:N}W.$$
Define $\Lambda_{ki} = \lambda_{ki}/\omega_i$ for $1 \leq i \leq N$, $\Lambda_{k0} = \lambda_{k0}$, and $D^\ddagger_{ij} = -\omega_j D_{ji}/\omega_i$. These substitutions in the KKT conditions yield

\[
\sum_{j=1}^{N} D^\ddagger_{ij} \Lambda_{kj} = -h \nabla_x H(X_{ki}, U_{ki}, \Lambda_{ki}), \quad 1 \leq i < N,
\]

\[
\sum_{i=1}^{N} D^\ddagger_{Ni} \Lambda_{ki} = -[h \nabla_x H(X_{kN}, U_{kN}, \Lambda_{kN}) + \Lambda_{k+1,0}/\omega_N], \quad (1)
\]

\[
N_{\mathcal{U}}(U_{ki}) \ni -\nabla_u H(X_{ki}, U_{ki}, \Lambda_{ki}), \quad 1 \leq i \leq N.
\]
Since the polynomial that is identically equal to 1 has derivative 0, we have \( D1 = 0 \), which implies that \( D0 = -\sum_{j=1}^{N} D_j \), where \( D_j \) is the \( j \)-th column of \( D \). Hence,

\[
\Lambda_{k0} = -\sum_{i=1}^{N} \lambda_{ki} D_{i0} = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{ki} D_{ij} = \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_j \left( \frac{\lambda_{ki}}{\omega_i} \right) (\omega_i D_{ij}/\omega_j) \\
= -\sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i D_{ij}^\dagger \Lambda_{kj} \tag{\dagger}
\]

\[
= \Lambda_{k+1,0} + h \sum_{i=1}^{N} \omega_i \nabla_{x} H(X_{ki}, U_{ki}, \Lambda_{ki}),
\]

where the discrete costate equations are used in the last step.
Let $\lambda_k \in P_{N-1}^n$ satisfy $\lambda_k \tau_i = \Lambda_{ki}$, $1 \leq i \leq N$.

$$\sum_{j=1}^{N} D_{ij}^{+} \Lambda_{kj} = \dot{\lambda}_k(\tau_i), \quad 1 \leq i < N, \quad \text{and}$$

$$\sum_{j=1}^{N} D_{Nj}^{+} \Lambda_{kj} = \dot{\lambda}_k(1) - \dot{\lambda}_k(1)/\omega_N. \quad (2)$$

This substitution in (‡) yields

$$\Lambda_{k0} = \lambda_k(1) - \sum_{i=1}^{N} \omega_i \dot{\lambda}_k(\tau_i).$$

Since $\dot{\lambda}_k \in P_{N-2}^n$ and $N$-point Radau quadrature is exact for these polynomial, we have

$$\sum_{i=1}^{N} \omega_i \dot{\lambda}_k(\tau_i) = \int_{-1}^{1} \dot{\lambda}_k(\tau) d\tau = \lambda_k(1) - \lambda_k(-1).$$
Combine last two equations to obtain

\[ \Lambda_{k0} = \lambda_k(-1) \quad ((3)) \]

(1) + (2) + (3) yield the formula for \( \dot{\lambda}_k(1) \).
For Gauss, the entries of the matrices $D$ and $D^\dagger$ satisfy

$$D_{ij} = -D^\dagger_{N+1-i,N+1-j}, \quad 1 \leq i \leq N, \quad 1 \leq j \leq N.$$  

In other words, $D_{1:N} = -JD_{1:N}^\dagger$ where $J$ is the exchange matrix with ones on its counterdiagonal and zeros elsewhere. Hence, $\|D^{-1}\|_\infty = \|(D^\dagger)^{-1}\|_\infty$.

20,000 Yen Prize Problem: Show that $\|(D^\dagger)^{-1}\|_\infty \leq 2$ and the rows of $[W^{1/2}D^\dagger]^{-1}$ have Euclidean length bounded by $\sqrt{2}$.

In the $hp$-scheme, only need to have $K$ is large enough, or equivalently $h$ is small enough, that $2hd_1 < 1$ and $2hd_2 < 1$, where

$$d_1 = \sup_{t \in \Omega_0} \|\nabla_x f(x^*(t), u^*(t))\|_\infty$$

$$d_2 = \sup_{t \in \Omega_0} \|\nabla_x f(x^*(t), u^*(t))^T\|_\infty$$
Discrete and continuous vectors

\[ X_{kj}^* = x_k^*(\tau_j), \quad 0 \leq j \leq N, \quad 1 \leq k \leq K, \]
\[ U_{kj}^* = u_k^*(\tau_j), \quad 1 \leq j \leq N, \quad 1 \leq k \leq K, \]
\[ \Lambda_{kj}^* = \lambda_k^*(\tau_j), \quad 0 \leq j \leq N, \quad 1 \leq k \leq K. \]
\[ X_{kj}^N = x_k^N(\tau_j), \quad 0 \leq j \leq N, \quad 1 \leq k \leq K, \]
\[ U_{kj}^N = u_k^N, \quad 1 \leq j \leq N, \quad 1 \leq k \leq K, \]
\[ \Lambda_{kj}^N = \lambda_k^N(\tau_j), \quad 0 \leq j \leq N, \quad 1 \leq k \leq K. \]

**NOTE:** \( \lambda^*_k(1) = \lambda^*_{k+1}(-1) \) but \( \lambda^N_k(1) \neq \lambda^N_{k+1}(-1) \).
Theorem. Suppose \((x^*, u^*)\) is a local minimizer for the continuous problem with \((x^*, \lambda^*) \in \mathcal{P}\mathcal{H}^\eta(\Omega_0; \mathbb{R}^n)\) for some \(\eta \geq 2\). If some smoothness and local convexity conditions hold, then for \(N\) sufficiently large or for \(h\) sufficiently small with \(N \geq 2\), the discrete problem has a local minimizer \(x_k^N \in \mathcal{P}_N^\eta\) and \(u_k^N \in \mathbb{R}^{mN}\), and an associated multiplier \(\lambda_k^N \in \mathcal{P}_{N-1}^\eta\), \(1 \leq k \leq K\); moreover, there exists a constant \(c\), independent of \(h, N\) and \(\eta\), such that

\[
\max \left\{ \|X^N - X^*\|_\infty, \|U^N - U^*\|_\infty, \|\Lambda^N - \Lambda^*\|_\infty \right\} \\
\leq h^{p-1} \left( \frac{c}{N} \right)^{p-1} |x^*|_{\mathcal{P}\mathcal{H}^p(\Omega_0)} + h^{q-1} \left( \frac{c}{N} \right)^{q-1.5} |\lambda^*|_{\mathcal{P}\mathcal{H}^q(\Omega_0)},
\]

where \(p = \min(\eta, N + 1)\) and \(q = \min(\eta, N)\).
Example

minimize \[ \frac{1}{2} \int_0^1 [x^2(t) + u^2(t)] \, dt \]

subject to \[ \dot{x}(t) = u(t), \quad u(t) \leq 1, \quad x(t) \leq \frac{2\sqrt{e}}{1-e}, \quad t \in [0, 1], \]

\[ x(0) = \frac{5e + 3}{4(1-e)}. \]

Exact solution:

\[ 0 \leq t \leq \frac{1}{4} : \quad x^*(t) = t - \frac{1}{4} + \frac{1+e}{1-e}, \quad u^*(t) = 1, \]

\[ \frac{1}{4} \leq t \leq \frac{3}{4} : \quad x^*(t) = \frac{e^{t-\frac{1}{4}}}{1-e} (1 + e^{\frac{3}{2}-2t}), \quad u^*(t) = \frac{e^{t-\frac{1}{4}}}{1-e} (1 - e^{\frac{3}{2}-2t}), \]

\[ \frac{3}{4} \leq t \leq 1 : \quad x^*(t) = \frac{2\sqrt{e}}{1-e}, \quad u^*(t) = 0. \]
Error in Solution

\[ \log_{10}(\text{error}) \]

state
control

N
2 3 4 5 6 7 8