\[ \bar{J} = \Phi[x(N)] - \lambda^T(N) x(N) + \sum_{i=1}^{N-1} [H^i - \lambda^T(i) x(i)] + H^o. \] (2.6.7)

Now consider differential changes in \( \bar{J} \) due to differential changes in \( u(i) \):

\[
d\bar{J} = \left[ \frac{\partial \Phi}{\partial x(N)} - \lambda^T(N) \right] dx(N) + \sum_{i=1}^{N-1} \left\{ \left[ \frac{\partial H^i}{\partial x(i)} - \lambda^T(i) \right] dx(i) + \frac{\partial H^i}{\partial u(i)} du(i) \right\} + \frac{\partial H^o}{\partial x(0)} dx(0) + \frac{\partial H^o}{\partial u(0)} du(0). \tag{2.6.8}
\]

The coefficients multiplying \( dx(i)(i = 0, \ldots, n) \) vanish if we choose the multiplier sequence \( \lambda(i) \) so that we have

\[
\lambda^T(i) - \frac{\partial H^i}{\partial x(i)} = 0, \tag{2.6.9}
\]

or

\[
\lambda^T(i) = \frac{\partial L^i}{\partial x(i)} + \lambda^T(i + 1) \frac{\partial f^i}{\partial x(i)}, \quad i = 0, \ldots, N - 1, \tag{2.6.9a}
\]

with boundary conditions

\[
\lambda^T(N) = \frac{\partial \Phi}{\partial x(N)}, \tag{2.6.10}
\]

or

\[
\lambda^T(N) = \frac{\partial \phi}{\partial x(N)} + \nu^T \frac{\partial \psi}{\partial x(N)}. \tag{2.6.10a}
\]

Equation (2.6.8) then becomes

\[
d\bar{J} = \lambda^T(0) dx(0) + \sum_{i=0}^{N-1} \frac{\partial H^i}{\partial u(i)} du(i). \tag{2.6.11}
\]

Thus \( \partial H^i/\partial u(i) \) is the gradient of \( \bar{J} \) with respect to \( u(i) \) while holding \( x(0) \) constant and satisfying (2.6.2), and \( \lambda^T(0) \) is the gradient of \( \bar{J} \) with respect to \( x(0) \) while holding \( u(i) \) constant and satisfying (2.6.2). If \( x(0) \) is given, we have \( dx(0) = 0 \).

For a stationary value of \( \bar{J} \), \( d\bar{J} \) must be zero for admissible \( du(i) \). If \( u(i) \) is unconstrained and \( H^i \) is differentiable with respect to \( u(i) \) and the problem is "normal", this can happen only if:

\[
\frac{\partial H^i}{\partial u(i)} = 0. \tag{2.6.12}
\]

†See Sections 5.3 and 6.3 for the argument concerning "normality" which relates to the existence of neighboring optimal paths.
or
\[ \frac{\partial L}{\partial u(i)} + \lambda^T(i + 1) \frac{\partial f}{\partial u(i)} = 0, \quad i = 0, \ldots, N - 1. \quad (2.6.12a) \]

In summary, to find a control-vector sequence \( u(i) \) that produces a stationary value of the performance index \( J \), we must solve the “two-point boundary-value problem” defined by (2.6.2), (2.6.3), (2.6.9), (2.6.10), and (2.6.12).

These equations constitute \((2n + m) N + n + p\) equations for as many unknowns: \( x(0), \ldots, x(N) \) where \( x \) is an \( n \)-vector; \( u(0), \ldots, u(N - 1) \), where \( u \) is an \( m \)-vector; \( \lambda(0), \lambda(1), \ldots, \lambda(N) \), where \( \lambda \) is an \( n \)-vector; and \( \nu \), a \( p \)-vector.

To solve (2.6.2) and (2.6.9a) together sequentially in the forward direction, using (2.6.12a) to determine \( u(i) \), it is necessary to solve (2.6.9a) for \( \lambda(i + 1) \) in terms of \( \lambda(i) \) and \( x(i) \):
\[ \lambda^T(i + 1) = \left[ \lambda^T(i) - \frac{\partial L}{\partial x(i)} \right] \left[ \frac{\partial f}{\partial x(i)} \right]^{-1}. \quad (2.6.13) \]

The inverse of \( \frac{\partial f}{\partial x(i)} \) exists since it is, essentially, the linearized transition-matrix; however, the computation of this inverse is time-consuming. The alternative of sequential backward solution offers no improvement since (2.6.2), (2.6.9a), and (2.6.12a) would have to be viewed as a set of implicit equations for \( x(i), \lambda(i) \), and \( u(i) \), given \( x(i + 1), \lambda(i + 1), u(i + 1) \).

2.7 Continuous systems; some state variables specified at an unspecified terminal time (including minimum-time problems)

This is almost the same set of problems as in Section 2.4, with the important difference that the terminal time, \( t_f \), is not specified. It is convenient to regard \( t_f \) as a control parameter to be chosen in addition to the control functions, \( u(t) \), so as to minimize the performance index and satisfy the constraints. We shall show that the same necessary conditions apply as in Section 2.4, but, in addition, the following condition must be satisfied by the optimal choice of the terminal time, \( t_f \):
\[ \left( \frac{\partial \phi}{\partial t} + \lambda^T f + L \right)_{t=t_f} = 0. \]

As in Section 2.3, we adjoin the system differential equations to the performance index, as follows:

†See Appendix A3.
‡The computation of the inverse is circumvented in the algorithm given in Section 7.7.