Optimal control of Goursat-Volterra systems.

S. A. Belbas
Mathematics Department
University of Alabama
Tuscaloosa, AL 35487-0350. USA.

e-mail: SBELBAS@BAMA.UA.EDU

Abstract. We analyze an optimal control problem for systems of integral equations of Volterra type with two independent variables. These systems generalize both, the hyperbolic control problems for systems of Goursat-Darboux type, and the optimal control of ordinary (i.e. with one independent variable) Volterra integral equations. We prove extremal principles akin to Pontryagin's maximum principle.

AMS Subject Classification: 49B34, 45D05, 45G15.
1. Introduction.

The optimal control of hyperbolic equations has received a lot of attention in the research literature. This is due to the large variety of applications of hyperbolic equations to physical and engineering problems.

One important direction of research in the optimal control of hyperbolic equations is the study of necessary conditions for optimality using characteristic coordinates for a controlled system governed by quasilinear second order hyperbolic equations. This has led to a family of optimal control problems for hyperbolic equations in characteristic coordinates, also known as Goursat-Darboux systems.

From a mathematical point of view, the place of optimal control for Goursat-Volterra systems in the overall framework of calculus of variations and optimal control theory can be illustrated in the following table:

| One independent variable | Several independent variables |
|--------------------------|------------------------------|
| (1). Calculus of variations for single integrals. | (4). Calculus of variations for multiple integrals. |
| (2). Optimal control of ordinary differential equations. | (5). Optimal control of hyperbolic systems, in particular Roesser-type systems and Goursat-Darboux systems. |
| (3). Optimal control of Volterra integral equations. | (6). Optimal control of Goursat-Volterra systems (for the first time in the present paper). |

Table 1.1. Calculus of variations and optimal control in one and in several independent variables.

The topic of optimal control of Goursat-Volterra systems is being treated for the first time in the present paper. The other topics shown in the table above have been previously treated by several authors. It is well known that in all those other topics, # (1) through (5) in the table above, there are new non-trivial developments as we move either horizontally or vertically in this table. The extremum principle of Pontryagin is generally acknowledged as the separation step of optimal control from the calculus of variations; the optimal control of Volterra integral equations has the additional feature that the Hamiltonian is a functional of the co-state, and not merely a function; the theory of calculus of variations for multiple integrals, beginning in the works of Gauss and further developed in the works of Hilbert, Carathéodory, Haar, Lichtenstein, Prange, Boerner, Kötzler, et al., has many new features that do not exist in calculus of variations for single...
integrals, such as Haar's lemma, Carathéodory's development of a counterpart of the Legendre transformation for multiple integral problems, etc. The transition to optimal control for hyperbolic systems presents, among other things, the new feature that, without strong differentiability assumptions on an optimal control, the Hamiltonian equations cannot be written in differential form, but only in the form of integral equations. The results in the present paper, for Goursat-Volterra systems, have the further features that (i) in addition to the case of Goursat-Darboux problems, the Hamiltonian is a functional (not just a function) of the co-state, and (ii) an extremum principle of Pontryagin's type is generally not possible, but a partial extremum principle is valid, and an extremum principle holds only in certain particular cases. These features will be made precise later in this paper.

The optimal control of Goursat-Darboux partial differential equations has been treated in, among other publications, [B1-B4, BDMO, VST]; the optimal control of ordinary Volterra equations (i.e. one independent variable) has been treated in, among other places, [B5, S, V]. The theory of optimal control of Goursat-Volterra equations is not simply a combination of Goursat-Darboux problems and ordinary Volterra equations, but rather it requires a separate treatment ab initio; the optimality conditions for Goursat-Volterra control problems cannot be simply guessed on the basis of Goursat-Darboux and ordinary Volterra control problems.
2. The basic Goursat-Volterra optimal control problem.

In this section, we shall define a certain type of hereditary systems of Goursat-Darboux type and the associated optimal control problems, and we shall examine the simplest properties of such problems.

We consider a system with partitioned state \( y(s, t) = \text{col}[y_1(s, t): 1 \leq i \leq N] \) where "col" signifies a column vector obtained by placing the column vectors \( y_1(s, t), y_2(s, t), \ldots \) one underneath the other in a partitioned column arrangement, and each \( y_i(s, t) \) is a column vector in \( \mathbb{R}^{n_i} \). We set

\[
y_{(i)}(s, t) := \text{col}[y_{j(s, t): 1 \leq j \leq i - 1}], \text{ for } 2 \leq i \leq N;
\]

\[
y_{[i]}(s, t) := \text{col}[y_{j(s, t): 1 \leq j \leq i}], \text{ for } 1 \leq i \leq N
\]

--- (2.1)

The controlled system will have 3 control functions \( u_1(s), u_2(t), u_{12}(s, t) \). The system has cascaded form with respect to the components (vector components, i.e. parts of a partitioned column vector) of the state:

\[
y_1(s, t) = f_{0,1}(s, t, y_{(i)}(s, t), u_1(s), u_2(t), u_{12}(s, t)) +
\]

\[
+ \int_0^s f_{1,i}(s, t, \sigma, y_{[i]}(\sigma, t), u_1(\sigma), u_2(t), u_{12}(\sigma, t))d\sigma +
\]

\[
+ \int_0^t f_{2,i}(s, t, \tau, y_{[i]}(s, \tau), u_1(s), u_2(\tau), u_{12}(s, \tau))d\tau +
\]

\[
+ \int_0^s \int_0^t f_{12,i}(s, t, \sigma, \tau, y_{[i]}(\sigma, \tau), u_1(\sigma), u_2(\tau), u_{12}(\sigma, \tau))d\tau d\sigma
\]

--- (2.2)

The cost functional associated with (2.2) is

\[
J := F_0(y(A, B), u_1(A), u_2(B), u_{12}(A, B)) +
\]

\[
+ \int_0^A F_1(s, y(s, B), u_1(s), u_2(B), u_{12}(s, B))ds +
\]

\[
+ \int_0^B F_2(t, y(A, t), u_1(A), u_2(t), u_{12}(A, t))dt +
\]

\[
+ \int_0^A \int_0^B F_{12}(s, t, y(s, t), u_1(s), u_2(t), u_{12}(s, t))dt ds
\]

--- (2.3)

The model (2.2-2.3) encompasses the general form of controlled (non-hereditary) Goursat-Darboux systems.
The general problem of optimal control for systems governed by partial differential equations of Goursat-Darboux type has the following form of state dynamics:

\[ \frac{\partial^2 x(s,t)}{\partial s \partial t} = \varphi(s,t,x(s,t),x_s(s,t),x_t(s,t),v(s,t)); \]
\[ x(s,0) = x_1(s); \quad x(0,t) = x_2(t); \]
\[ \frac{dx_1(s)}{ds} = \varphi_1(s,x_1(s),v_1(s)); \quad \frac{dx_2(t)}{dt} = \varphi_2(t,x_2(t),v_2(t)); \quad x_1(0) = x_2(0) = x_0 \]

--- (2.4)

The cost functional to be minimized for (2.3) is

\[ J_1 := \int_0^A \int_0^B \Phi(s,t,x(s,t),x_s(s,t),x_t(s,t),v(s,t),v_1(s),v_2(s))dtds + \]
\[ + \int_0^A \Phi_1(s,x(s,B),x_s(s,B),v_1(s))ds + \int_0^B \Phi_2(t,x(A,t),x_t(A,t),v_2(t))dt + \]
\[ + \Phi_0(x(A,B)) \]

--- (2.5)

We write the equations of the group (2.4) in integral form:

\[ x(s,t) = x_1(s) + x_2(t) - x_0 + \]
\[ + \int_0^s \int_0^t \varphi(\sigma,\tau,x(\sigma,\tau),x_s(\sigma,\tau),x_t(\sigma,\tau),v(\sigma,\tau))d\tau d\sigma; \]
\[ x_1(s) = x_0 + \int_0^s \varphi_1(\sigma,x_1(\sigma),v_1(\sigma))d\sigma; \quad x_2(t) = x_0 + \int_0^t \varphi_2(\tau,x_2(\tau),v_2(\tau))d\tau \]

--- (2.6)

We set

\[ p(s,t) := x_s(s,t), \quad q(s,t) := x_t(s,t) \]

--- (2.7)

Then the first equation in (2.6) becomes

\[ x(s,t) = x_1(s) + x_2(t) - x_0 + \int_0^s \int_0^t \varphi(\sigma,\tau,x(\sigma,\tau),p(\sigma,\tau),q(\sigma,\tau),v(\sigma,\tau))d\tau d\sigma \]

--- (2.8)

Differentiation of (2.8) with respect to s and t yields
\[ p(s, t) = \varphi_1(s, x_1(s), v_1(s)) + \int_0^t \varphi(s, \tau, x(s, \tau), p(s, \tau), q(s, \tau), v(s, \tau)) d\tau; \]
\[ q(s, t) = \varphi_2(t, x_2(t), v_2(t)) + \int_0^s \varphi(\sigma, t, x(\sigma, t), p(\sigma, t), q(\sigma, t), v(\sigma, t)) d\sigma \]
--- (2.9)

We define
\[ y_1(s, t) := \begin{bmatrix} x_1(s) \\ x_2(t) \end{bmatrix}, \quad y_2(s, t) := \begin{bmatrix} x(s, t) \\ p(s, t) \\ q(s, t) \end{bmatrix} \]
--- (2.10)

Then the system of (2.6) through (2.9) is seen to be a particular case of (2.2). Also, the functional \( J_1 \) becomes a particular case of \( J \).

Many of the features of the system consisting of (2.2) and (2.3) are contained in the simpler problem consisting of the state dynamics equations in the simplified form (2.11) below and the problem of minimizing the same functional \( J \) as in (2.3) but with the new simpler equations of state:
\[ y(s, t) = f_0(s, t, u_1(s), u_2(t), u_{12}(s, t)) + \]
\[ + \int_0^s f_1(s, t, \sigma, y(\sigma, t), u_1(\sigma), u_2(t), u_{12}(\sigma, t)) d\sigma + \]
\[ + \int_0^t f_2(s, t, \tau, y(s, \tau), u_1(s), u_2(\tau), u_{12}(s, \tau)) d\tau + \]
\[ + \int_0^s \int_0^t f_{12}(s, t, \sigma, \tau, y(\sigma, \tau), u_1(\sigma), u_2(\tau), u_{12}(\sigma, \tau)) d\sigma d\tau \]
--- (2.11)

where now the state \( y(.) \) takes values in \( \mathbb{R}^n \).

In some cases, we shall give complete proofs only for the model (2.11).

Next, we outline the existence and uniqueness theory for integral systems of the form (2.11).

Exclusively in the context of discussing existence and uniqueness questions, we make the following assumptions concerning the control functions and the functions \( f_0, f_1, f_2, f_{12} \):
(i). $u_1(.,u_2(.,u_{12}(.,\cdot)\text{ are bounded measurable functions taking values in the compact sets } U_1, U_2, U_{12}, \text{ respectively.}$

(ii). The functions $f_0, f_1, f_2, f_{12}$ are bounded measurable functions with respect to $s, t, \sigma, \tau$ for all $(s, \sigma, t, \tau)$ such that $0 \leq \sigma \leq s \leq a, 0 \leq \tau \leq t \leq b$, and continuous with respect to $y, u_1, u_2, u_{12}$.

(iii). The functions $f_1, f_2, f_{12}$ are Lipschitz with respect to $y$, with Lipschitz constants $L_1, L_2, L_{12}$, respectively. ///

Under these conditions, the system (2.11) has a unique solution in $L^\infty([0,A] \times [0,B])$. This is proved by a standard contraction mapping argument, similar to the case of single Volterra equations. For completeness, we briefly outline one proving that. We define an operator $S$ on $L^\infty([0,A] \times [0,B])$ by

$$(S\mathbf{z})(s,t) = f_0(s,t,u_1(s),u_2(t),u_{12}(s,t)) +$$

$$+ \int_0^s f_1(s,t,\sigma,z(\sigma,t),u_1(\sigma),u_2(t),u_{12}(\sigma,t))d\sigma +$$

$$+ \int_0^t f_2(s,t,\tau,z(s,\tau),u_1(s),u_2(\tau),u_{12}(s,\tau))d\tau +$$

$$+ \int_0^s \int_0^t f_{12}(s,t,\sigma,\tau,z(\sigma,\tau),u_1(\sigma),u_2(\tau),u_{12}(\sigma,\tau))d\tau d\sigma.$$

Then the problem of solving (2.11) amounts to finding a fixed point of $S$.

Using the weighted norm (which is equivalent to the standard norm on $L^\infty([0,a] \times [0,b])$)

$$\|z\|_\mu := \text{ess sup} \{\exp(-\mu(s+t)) \cdot |z(s,t)| : (s,t) \in [0,a] \times [0,b]\},$$

with $\mu > 0$, we have, for every two functions $z_1, z_2$
The expression
\[ \exp(-\mu(s + t)) \mid (Sz_1)(s,t) - (Sz_2)(s,t) \mid \leq \]
\[ \leq \exp(-\mu(s + t)) \left[ L_1 \int_0^s \mid z_1(\sigma, t) - z_2(\sigma, t) \mid d\sigma + L_2 \int_0^t \mid z_1(s, \tau) - z_2(s, \tau) \mid d\tau + \right. \]
\[ + L_{12} \int_0^s \int_0^t \mid z_1(\sigma, \tau) - z_2(\sigma, \tau) \mid d\tau d\sigma \leq \]
\[ \leq \exp(-\mu(s + t)) \left[ L_1 \int_0^s \exp(\mu\sigma) \mid z_1 - z_2 \mid_{\mu} d\sigma + L_2 \int_0^t \exp(\mu\tau) \mid z_1 - z_2 \mid_{\mu} d\tau + \right. \]
\[ + L_{12} \int_0^s \int_0^t \exp(\mu(\sigma + \tau)) \mid z_1 - z_2 \mid_{\mu} d\tau d\sigma \right] = \]
\[ = \{\mu^{-1}[\exp(-\mu t) - \exp(-\mu(s + t))]L_1 + \mu^{-1}[\exp(-\mu s) - \exp(-\mu(s + t))]L_2 + \]
\[ + \mu^{-2}[1 - \exp(-\mu t) - \exp(-\mu s) + \exp(-\mu(s + t))]L_{12}\} \mid z_1 - z_2 \mid_{\mu} \leq \]
\[ \leq \{\mu^{-1}[1 - \exp(-\mu(a + b))]L_1 + \mu^{-1}[1 - \exp(-\mu(a + b))] + \]
\[ + \mu^{-2}[2 - \exp(-\mu a) - \exp(-\mu b)]\} \mid z_1 - z_2 \mid_{\mu} . \]

The expression
\[ \{\mu^{-1}[1 - \exp(-\mu(A + B))]L_1 + \mu^{-1}[1 - \exp(-\mu(A + B))] + \mu^{-2}[2 - \exp(-\mu A) - \exp(-\mu B)]\} \]
go to 0 as \( \mu \to \infty \), thus it can be made \( \leq q \) for \( \mu \) sufficiently large, if \( q \) is any predetermined number in \([0, 1)\), and this makes the operator \( S \) a contraction on \( L^\infty([0, A] \times [0, B]) \) with respect to the norm \( \| \cdot \|_{\mu} \), and then existence and uniqueness of a fixed point follows from the standard contraction mapping theorem.

We note that related existence results, for less general problems and without considering controlled systems, have been obtained by a number of authors, for example [SU].

Although the above conditions yield the existence and uniqueness of a bounded measurable solution \( y \), they are not enough to make all terms in the cost functional \( J \) meaningful. If the terms \( u_1(A), u_2(B), u_{12}(A,t), u_{12}(s,B), u_{12}(A,B) \) are absent from \( J \), then the stated conditions suffice to make \( J \) well defined. If these terms are included in \( J \), then we need, additionally, the following set of conditions:

(iv). The limits \( \lim_{s \to A^-} u_{12}(s,\cdot) \equiv u_{12}(A,\cdot) \) and \( \lim_{t \to B^-} u_{12}(\cdot,t) \equiv u_{12}(\cdot,B) \) exist in \( L^\infty([0,B]) \), respectively; the limits \( \lim_{s \to A^-} u_1(s) \equiv u_1(A) \) and \( \lim_{t \to B^-} u_2(t) \equiv u_2(B) \) exist; the following equality involving limits holds:
An alternative, and more restrictive, set of conditions is provided below. It is possible, judging by analogy with the case of controlled ordinary differential equations, that, in cases in which optimal controls can actually be constructed, they may turn out to be piecewise continuous functions. It is, therefore, of practical interest to have optimality conditions formulated for piecewise continuous controls. The conditions below make exact the concept of "piecewise continuous controls" for the problem under consideration.

(iv*). There exist finitely many piecewise \( C^1 \) curves \((\gamma_j)\) in \( G := [0, A] \times [0, B] \), such that the subdomains \( G_i \) determined by the requirements \( G_i \subseteq \text{int} \ G \ \forall \ i \), each \( G_i \) does not contain any part of any \((\gamma_j)\), and each \( \partial G_i \) is a subset of \( \bigcup_j (\gamma_j) \), are open sets. The function \( u_{12}(\cdot, \cdot) \) is continuous on each \( G_i \), and the restriction of \( u_{12}(\cdot, \cdot) \) to \( G_i \) can be extended to a function that is continuous in \( \overline{G_i} \). The limits \( \lim_{s \to A^-} u_{12}(s,t) = u_{12}(A,t) \) and \( \lim_{t \to B^-} u_{12}(s,t) = u_{12}(s,B) \) exist for all except possibly finitely many values of \( t \), respectively finitely many values of \( s \), and they exist uniformly in \( t \), respectively uniformly in \( s \); the functions \( u_{12}(\cdot,A) \) and \( u_{12}(\cdot,B) \) are piecewise continuous (in the usual sense for functions of one variable, i.e. they can have at most a finite number of jump discontinuities with finite limits from the left and from the right at each point of discontinuity) on \([0, B]\), respectively \([0, A]\); the following holds:

\[
\begin{align*}
\lim_{(s,t) \to (A^-,B^-)} u_{12}(s,t) &= u_{12}(A,t) = \lim_{t \to B^-} u_{12}(s,B) = u_{12}(A,B) ; \\
\lim_{s \to A^-} u_1(s) &= u_1(A), \quad \lim_{t \to B^-} u_2(t) = u_2(B) \quad \text{exist.} \quad ///
\end{align*}
\]

We conclude this section with a few remarks on applied scenarios that can lead to optimal control problems for Goursat-Volterra systems of the type we have defined above.

One category of classical applications of Goursat-Darboux and Roesser problems concerns phenomena with exchange of a quantity (e.g. mass or energy) between a moving medium and a stationary medium. Such phenomena include, among other things, gas chromatography, drying processes, air sparging methods of remediation of contaminated groundwater sites, and "moving bed operations" in chemical industry. In some of these problems, certain memoryless constitutive laws can be replaced with laws
with memory, and the corresponding equations become Goursat-Volterra equations. For example, consider the problem of gas chromatography with one-dimensional flow of gas in a porous medium. If \( v(s, t) \) is the speed of the flow at position \( s \) at time \( t \), \( \varphi(s, t) \) is the mass density of flowing gas, \( a(s, t) \) is the mass density of absorbed gas, and \( y(s, t) \) is the mass density of gas in chemical equilibrium with the absorbed gas, then the mass balance equations can be written as

\[
v(s, t) \frac{\partial \varphi(s, t)}{\partial s} + \frac{\partial}{\partial t} (a(s, t) + \varphi(s, t)) = 0; \quad \frac{\partial a(s, t)}{\partial t} = \beta (\varphi(s, t) - y(s, t))
\]

where \( \beta > 0 \) is an absorption coefficient. The system of two mass balance equations with 3 unknowns \( \varphi, a, \) and \( y \), is supplemented by a constitutive law, which here we take as a law with memory,

\[a(s, t) = \int_0^t K(s, t, \tau) y(s, \tau) d\tau\]

(In the case of a constitutive law without memory, the corresponding equation is the so-called adsorption isotherm.)

It follows from the hereditary constitutive law that

\[a_1(s, t) = K(s, t, t) y(s, t) + \int_0^t K_1(s, t, \tau) y(s, \tau) d\tau\]

Then we can eliminate \( a_1(s, t) \) between the above integral equation and the second equation out of the two mass balance equations:

\[\beta + K(s, t, t) y(s, t) = \beta \varphi(s, t) + \int_0^t K_1(s, t, \tau) y(s, \tau) d\tau\]

Assuming \( K(s, t, t) \geq 0 \), the above equation can be written in the form

\[y(s, t) = \ell_0(s, t) \varphi(s, t) + \int_0^t \frac{K_1(s, t, \tau)}{\beta + K(s, t, t)} y(s, \tau) d\tau; \quad \ell_0(s, t) := \frac{\beta}{\beta + K(s, t, t)}\]

and, by using the resolvent kernel for a linear Volterra equation, we can represent the solution \( y(s, t) \) as

\[y(s, t) = \ell_0(s, t) \varphi(s, t) + \int_0^t \ell_1(s, t, \tau) \varphi(s, \tau) d\tau\]

from which it follows that
\[ y_t(s, t) = \ell_{0,t}(s, t)\varphi(s, t) + \ell_0(s, t)\varphi_t(s, t) + \ell_1(s, t)\varphi(s, t) + \int_0^t \ell_{1,t}(s, \tau)\varphi(s, \tau)\,d\tau \]  
--- (2.13)

We differentiate eq. (2.12) with respect to \( t \) to obtain

\[ v_t(s, t)\varphi_s(s, t) + v(s, t)\varphi_s(t, t) + \varphi_u(s, t) + \beta(\varphi_t(s, t) - y_t(s, t)) = 0 \]

Under the condition that the second derivative of \( \varphi \) with respect to time is negligible compared to the other terms, and using the expression (2.13) for \( y_t(s, t) \), we can write the following equation:

\[ \frac{\partial^2 \varphi(s, t)}{\partial s \partial t} = a_0(s, t)\varphi(s, t) + a_1(s, t)\varphi_s(s, t) + \varphi_t(s, t) + \varphi(s, t) \]

\[ + a_3(s, t)\int_0^t \ell_{1,t}(s, \tau)\varphi(s, \tau)\,d\tau \]  
--- (2.14)

We write (2.14) in integral form:

\[ \varphi(s, t) = \varphi(s, 0) + \varphi(0, t) - \varphi(0, 0) + \int_0^s \int_0^t \{a_0(\sigma, \tau)\varphi(\sigma, \tau) + a_1(\sigma, \tau)\varphi_s(\sigma, \tau) + \varphi_t(\sigma, \tau)\}\,d\tau\,d\sigma + \int_0^s \int_0^t \int_0^{\tau_1} a_3(\sigma, \tau_1)\ell_{1,\tau_1}(\sigma, \tau, \tau_1)\varphi(\sigma, \tau)\,d\tau\,d\tau_1\,d\sigma \]  
--- (2.15)

We define \( \tilde{a}_3(t, \sigma, \tau) := \int_\tau^t a_3(\sigma, \tau_1)\ell_{1,\tau_1}(\sigma, \tau, \tau_1)\,d\tau_1 \), and then eq. (2.15) can be written as

\[ \varphi(s, t) = \varphi(s, 0) + \varphi(0, t) - \varphi(0, 0) + \int_0^s \int_0^t \{a_0(\sigma, \tau)\varphi(\sigma, \tau) + a_1(\sigma, \tau)\varphi_s(\sigma, \tau) + \varphi_t(\sigma, \tau) + \tilde{a}_3(t, \sigma, \tau)\varphi(\sigma, \tau)\}\,d\tau\,d\sigma \]  
--- (2.16)

Then, by differentiating (2.16) with respect to each of \( s, t \), we find
\[ \varphi_s(s, t) = \varphi_s(s, 0) + \int_0^t \{ a_0(s, \tau) \varphi(s, \tau) + a_1(s, \tau) \varphi_s(s, \tau) + a_2(s, \tau) \varphi_t(s, \tau) + \\
+ \tilde{a}_3(t, s, \tau) \varphi(s, \tau) \} \, d\tau; \]

\[ \varphi_t(s, t) = \varphi_t(0, t) + \int_0^s \{ a_0(\sigma, t) \varphi(\sigma, t) + a_1(\sigma, t) \varphi_\sigma(\sigma, t) + a_2(\sigma, t) \varphi_t(\sigma, t) + \\
+ \tilde{a}_3(t, \sigma, t) \varphi(\sigma, t) \} \, d\tau + \int_0^t \int_0^s \tilde{a}_3(t, \sigma, \tau) \varphi(\sigma, \tau) \, d\tau \, d\sigma \]

--- (2.17)

and it is seen that the system of (2.17) is a Goursat-Volterra system, according to our definition, in the vector-valued unknown function \([\varphi(s, t) \quad \varphi_s(s, t) \quad \varphi_t(s, t)]^T\).

An optimal control problem may be formulated by taking the velocity \(v(s, t)\) as the control, and asking for the minimization of a functional \(J\) given by

\[ J := \int_0^A \int_0^B \{ (\varphi(s, t) - \varphi_0(s, t))^2 + \mu(v(s, t))^2 \} \, dt \, ds \]

i.e. we want the state \(\varphi\) to be as close as possible, in the mean square sense, to a given profile \(\varphi_0\), without letting the velocity become too large. This can be interpreted also as an identification problem, if the velocity field is to be estimated on the basis of observations \(\varphi_0\) of the state \(\varphi\).

It should be emphasized that this is just one example, among many possible applications, and it is not intended as the motivation for the type of control problems we consider in the present paper.
3. Formal derivation of the Hamiltonian equations.

The formal derivation of Hamiltonian equations relies on introducing formal penalty terms and the associated Lagrange (functional) multipliers. This formal derivation is not proof of the validity of the Hamiltonian equations, but only establishes the form that we expect the Hamiltonian equations to have. The proofs will be carried out in subsequent sections.

For the problem consisting of (2.2) and (2.3), we introduce the following penalty terms and co-state functions:

\[ P_0 := \sum_i \psi_{0,i}(-y_i(A,B) + f_{0,i}(A,B,y_{(i)}(A,B),u_1(A),u_2(B),u_{12}(A,B)) + 
\]
\[ + \int_0^A f_{1,i}(A,B,\sigma,y_{[i]}(\sigma,B),u_1(\sigma),u_2(B),u_{12}(\sigma,B))d\sigma + 
\]
\[ + \int_0^B f_{2,i}(A,B,\tau,y_{[i]}(A,\tau),u_1(A),u_2(\tau),u_{12}(A,\tau))d\tau + 
\]
\[ + \int_0^A \int_0^B f_{12,i}(A,B,\sigma,\tau,y(\sigma,\tau),u_1(\sigma),u_2(\tau),u_{12}(\sigma,\tau))d\tau d\sigma; \]

\[ P_1 := \sum_i \int_0^A \psi_{1,i}(s)[-y_i(s,B) + f_{0,i}(s,B,y_{(i)}(s,B),u_1(s),u_2(B),u_{12}(s,B)) + 
\]
\[ + \int_0^s f_{1,i}(s,B,\sigma,y_{[i]}(\sigma,B),u_1(\sigma),u_2(B),u_{12}(\sigma,B))d\sigma + 
\]
\[ + \int_0^B f_{2,i}(s,B,\tau,y_{[i]}(s,\tau),u_1(s),u_2(\tau),u_{12}(s,\tau))d\tau + 
\]
\[ + \int_0^s \int_0^B f_{12,i}(s,B,\sigma,\tau,y(\sigma,\tau),u_1(\sigma),u_2(\tau),u_{12}(\sigma,\tau))d\tau d\sigma]ds; \]

\[ P_2 := \sum_i \int_0^B \psi_{2,i}(t)[-y_i(A,t) + f_{0,i}(A,t,y_{(i)}(A,t),u_1(A),u_2(t),u_{12}(A,t)) + 
\]
\[ + \int_0^A f_{1,i}(A,t,\sigma,y_{[i]}(\sigma,t),u_1(\sigma),u_2(t),u_{12}(\sigma,t))d\sigma + 
\]
\[ + \int_0^t f_{2,i}(A,t,\tau,y_{[i]}(A,\tau),u_1(A),u_2(\tau),u_{12}(A,\tau))d\tau + 
\]
\[ + \int_0^A \int_0^t f_{12,i}(A,t,\sigma,\tau,y(\sigma,\tau),u_1(\sigma),u_2(\tau),u_{12}(\sigma,\tau))d\tau d\sigma]dt; \]
\[ P_{12} := \sum_i \int_0^A \int_0^B \psi_{12,i}(s,t)[-y_i(s,t) + f_{0,i}(s,t,y_{(i)}(s,t),u_1(s),u_2(t),u_{12}(s,t)) + \\
+ \int_0^s f_{1,i}(s,t,\sigma,y_{[i]}(\sigma,t),u_1(\sigma),u_2(t),u_{12}(\sigma,t))d\sigma + \\
+ \int_0^t f_{2,i}(s,t,\tau,y_{[i]}(s,\tau),u_1(s),u_2(\tau),u_{12}(s,\tau))d\tau + \\
+ \int_0^s \int_0^t f_{12,i}(s,\tau,\sigma,y_{[i]}(\sigma,\tau),u_1(\sigma),u_2(\tau),u_{12}(\sigma,\tau))d\sigma d\tau]dtds \quad (3.1) \]

The Lagrangian functional associated with (2.2) and (2.3) is

\[ L := J + P_0 + P_1 + P_2 + P_{12} \quad (3.2) \]

We denote by \( \delta L \) the variation of \( L \) obtained by using only variations \( \delta y \) of the state \( y \), while treating \( y \) as if it were independent from \( u_1, u_2, u_{12} \).

We use the following notation: for any \( \mathbb{R}^N \)-valued function \( g(z,w,...) \), where \( z \) takes values in \( \mathbb{R}^M \), \( w \) takes values in \( \mathbb{R}^Q \), etc., we denote by \( \nabla_z g, \nabla_w g, ... \) the Jacobians of \( g \) with respect to \( z, w, ... \), i.e. the matrices \( \nabla_z g = \left[ \frac{\partial g_i}{\partial z_j} \right]_{1 \leq i \leq N, 1 \leq j \leq M} \), \( \nabla_w g = \left[ \frac{\partial g_i}{\partial w_j} \right]_{1 \leq i \leq N, 1 \leq j \leq Q} \), etc. In these formulae, the subscript \( i \) denotes rows and \( j \) denotes columns of the indicated matrices. According to this notation, the gradient of a real-valued function is a row vector.

Then we have
\( \delta^{-} P_0 = \sum_i \left( -\psi_{0,i} \delta^{-} y_i(A,B) + \sum_{i,j} \psi_{0,i} \nabla y_j f_{0,i}(A,B,...) \delta^{-} y_j(A,B) + \right. \\
+ \sum_{i,j} \int_0^A \psi_{0,i} \nabla y_j f_{1,i}(A,B,s,...) \delta^{-} y_j(s,B)ds + \\
+ \sum_{i,j} \int_0^B \psi_{0,i} \nabla y_j f_{2,i}(A,B,t,...) \delta^{-} y_j(A,t)dt + \\
+ \sum_{i,j} \int_0^A \int_0^B \psi_{0,i} \nabla y_j f_{12,i}(A,B,s,t,...) \delta^{-} y_j(s,t)dt ds \right) \\
\tag{3.3} \\

\delta^{-} P_1 = \sum_i \left\{ \int_0^A \psi_{1,i}(s) \left[ -\delta^{-} y_i(s,B) + \sum_{j:j<i} \nabla y_j f_{0,i}(s,B,...) \delta^{-} y_j(s,B) \right] ds + \\
+ \sum_{j:j<i} \int_0^A \int_0^s \psi_{1,i}(s) \nabla y_j f_{1,i}(s,B,\sigma,...) \delta^{-} y_j(\sigma,B)d\sigma ds + \\
+ \sum_{j:j<i} \int_0^A \int_0^B \psi_{1,i}(s) \nabla y_j f_{2,i}(s,B,\tau,...) \delta^{-} y_j(s,\tau)d\tau ds + \\
+ \sum_{j:j<i} \int_0^A \int_0^s \int_0^B \psi_{1,i}(s) \nabla y_j f_{12,i}(s,B,\sigma,\tau,...) \delta^{-} y_j(\sigma,\tau)d\tau d\sigma ds \right\} \\
\tag{3.4} \\

\delta^{-} P_2 = \sum_i \left\{ \int_0^B \psi_{2,i}(t) \left[ \delta^{-} y_i(A,t) + \sum_{j:j<i} \nabla y_j f_{0,i}(A,t,...) \delta^{-} y_j(A,t) \right] dt + \\
+ \sum_{j:j<i} \int_0^A \int_0^B \psi_{2,i}(t) \nabla y_j f_{1,i}(A,t,\sigma,...) \delta^{-} y_j(\sigma,t)d\sigma dt + \\
+ \sum_{j:j<i} \int_0^B \int_0^t \psi_{2,i}(t) \nabla y_j f_{2,i}(A,t,\tau,...) \delta^{-} y_j(A,\tau)d\tau dt + \\
+ \sum_{j:j<i} \int_0^A \int_0^t \int_0^B \psi_{2,i}(t) \nabla y_j f_{12,i}(A,t,\sigma,\tau,...) \delta^{-} y_j(\sigma,\tau)d\tau d\sigma \right\} \\
\tag{3.5}
By setting the variation of $L$ equal to zero, and taking into account the above formulae for the variations of the individual terms that make up $L$, and after performing a few changes in the order of integration and renaming certain variables, we obtain the following integral equations for the co-state $[\psi_0 \; \psi_1(s) \; \psi_2(t) \; \psi_{12}(s,t)]$:
\[ \psi_{12,i}(s,t) = \nabla Y_i f_{12}(s,...) + \sum_{j:i} \psi_{12,j}(s,t)\nabla Y_i f_{0,j}(s,...) + \]
\[ + \sum_{j:i} \left\{ \psi_{1,j}(s)\nabla Y_i f_{2,j}(s,B,t,...) + \psi_{2,j}(t)\nabla Y_i f_{1,j}(A,t,s,...) + \psi_{0,j}\nabla Y_i f_{12}(A,B,s,t,...) + \right. \]
\[ + \int_{s}^{A} [\psi_{12,j}(S,t)\nabla Y_i f_{1,j}(S,t,s,...) + \psi_{1,j}(S)\nabla Y_i f_{12,j}(S,B,s,t,...)]d\sigma + \]
\[ + \int_{t}^{B} [\psi_{12,j}(S,\tau)\nabla Y_i f_{2,j}(S,\tau,t,...) + \psi_{2,j}(\tau)\nabla Y_i f_{12,j}(S,\tau,s,t,...)]d\tau + \]
\[ + \int_{s}^{A} \int_{t}^{B} \psi_{12,j}(S,\tau)\nabla Y_i f_{12}(S,\tau,s,t,...)d\tau d\sigma ] \]
\[ \text{--- (3.8)} \]

\[ \psi_{1,j}(s) = \nabla Y_i f_{1}(s,...) + \sum_{j:i} \psi_{1,j}(s)\nabla Y_i f_{0,j}(s,B,...) + \]
\[ + \sum_{j:i} \left\{ \psi_{0,j}\nabla Y_i f_{1,j}(A,B,s,...) + \int_{s}^{A} \psi_{1,j}(S)\nabla Y_i f_{1,j}(S,B,s,...)d\sigma \right\} \]
\[ \text{--- (3.9)} \]

\[ \psi_{2,i}(t) = \nabla Y_i f_{2}(t,...) + \sum_{j:i} \psi_{2,j}(t)\nabla Y_i f_{0,j}(A,t,...) + \]
\[ + \sum_{j:i} \left\{ \psi_{0,j}\nabla Y_i f_{2,j}(A,B,t,...) + \int_{t}^{B} \psi_{2,j}(\tau)\nabla Y_i f_{2,j}(A,\tau,t,...)d\tau \right\} \]
\[ \text{--- (3.10)} \]

\[ \psi_{0,j} = \nabla Y_i f_{0}(y(A,B),...) + \sum_{j:i} \psi_{0,j}f_{0,j}(A,B,...) \]
\[ \text{--- (3.11)} \]

This leads us to define 4 Hamiltonians \( H_0, H_1, H_2, H_{12} \). In addition to the variables \( u_1, u_2, u_{12} \), these Hamiltonians also depend on the additional variables \( u_{1,1}, u_{2,2}, u_{12,1}, u_{12,2}, u_{12,12} \), where \( u_{1,1} \) plays the role of \( u_1(A) \), \( u_{2,2} \) that of
\begin{align*}
  u_2(B), \ u_{12,1} \text{ that of } u_{12}(A, t), \ u_{12,2} \text{ that of } u_{12}(s, B), \text{ and } u_{12,12} \text{ that of } u_{12}(A, B).
\end{align*}

Also, these Hamiltonians depend on the variables \(y, y_{\{1\}}, y_{\{2\}}, y_{\{12\}}\), where \(y_{\{1\}}\) plays the role of \(y(A, t)\), \(y_{\{2\}}\) that of \(y(s, B)\), and \(y_{\{12\}}\) that of \(y(A, B)\). The symbols \(y_{\{1\}}, y_{\{2\}}, y_{\{1\}j}, y_{\{2\}j}, \text{ etc.}\) will have the analogous meaning with \(y_j, y_{(j)}, y_{[j]}\), for example \(y_{\{1\}}(j)\) plays the role of \(y_{(j)}(A, t)\), and so forth.

We define
\[ H_0(y_{(1,2)}, u_{1,1}, u_{2,2}, u_{12,12}, \psi_0) := \\
= F_0(y_{(1,2)}, u_{1,1}, u_{2,2}, u_{12,12}) + \sum_j \psi_{0,j} f_{0,j}(A, B, y_{(1,2)}, (j), u_{1,1}, u_{2,2}, u_{12,12}); \]
\[ H_1(s, y_{(2)}, u_{1}, u_{2,2}, u_{12,12}, \psi_0, \psi_1(\cdot)) := F_1(s, y_{(2)}, u_{1}, u_{2,2}, u_{12,12}) + \]
\[ + \sum_j \left[ \psi_{1,j}(s) f_{1,j}(s, B, y_{(2)}, (j), u_{1}, u_{2,2}, u_{12,12}) + \psi_{0,j} f_{1,j}(A, B, s, y_{(2)}, (j), u_{1}, u_{2,2}, u_{12,12}) + \right. \]
\[ + \left. \int_s^A \psi_{1,j}(\sigma) f_{1,j}(\sigma, B, s, y_{(2)}, (j), u_{1}, u_{2,2}, u_{12,12}) d\sigma \right]; \]
\[ H_2(t, y_{(1)}, u_{1,1}, u_{2,1}, \psi_0, \psi_2(\cdot)) := F_2(t, y_{(1)}, u_{1,1}, u_{2,1}) + \]
\[ + \sum_j \left[ \psi_{2,j}(t) f_{2,j}(A, t, y_{(1)}, (j), u_{1,1}, u_{2,1}) + \psi_{0,j} f_{2,j}(A, B, t, y_{(1)}, (j), u_{1,1}, u_{2,1}) + \right. \]
\[ + \left. \int_t^B \psi_{2,j}(\tau) f_{2,j}(A, \tau, t, y_{(1)}, (j), u_{1,1}, u_{2,1}) d\tau \right]; \]
\[ H_{12}(s, t, y, u_1, u_2, u_{12}, \psi_0, \psi_1(.), \psi_2(.), \psi_{12}(..)) := F_{12}(s, t, y, u_1, u_2, u_{12}) + \]
\[ + \sum_j \left[ \psi_{0,j} f_{12,j}(A, B, s, t, y_{[i]}, u_{1}, u_{2}, u_{12}) + \psi_{1,j}(s) f_{2,j}(s, B, t, y_{[i]}, u_{1}, u_{2}, u_{12}) + \right. \]
\[ + \left. \psi_{2,j}(t) f_{1,j}(A, t, s, y_{[i]}, u_{1}, u_{2}, u_{12}) + \psi_{12,j}(s, t) f_{0,j}(s, t, y_{(j)}, u_{1}, u_{2}, u_{12}) + \right. \]
\[ + \left. \int_s^A \left[ \psi_{12,j}(\sigma, t) f_{1,j}(\sigma, t, s, y_{[i]}, u_{1}, u_{2}, u_{12}) + \psi_{1,j}(\sigma) f_{12,j}(\sigma, B, s, t, y_{[i]}, u_{1}, u_{2}, u_{12}) \right] d\sigma + \right. \]
\[ + \left. \int_t^B \left[ \psi_{12,j}(s, \tau) f_{2,j}(s, \tau, t, y_{[i]}, u_{1}, u_{2}, u_{12}) + \psi_{2,j}(\tau) f_{12,j}(A, t, s, y_{[i]}, u_{1}, u_{2}, u_{12}) \right] d\tau + \right. \]
\[ + \left. \int_s^A \int_t^B \psi_{12,j}(\sigma, \tau) f_{12,j}(\sigma, \tau, s, t, y_{[i]}, u_{1}, u_{2}, u_{12}) d\tau d\sigma \right] \]

--- (3.12)

The Hamiltonian equations for the co-state can now be written in the form
\[ \psi_{0,i} = \nabla_y H_0(y_{[12]}, u_{1,1}, u_{2,2}, u_{12,12}, \psi_0); \]
\[ \psi_{1,i}(s) = \nabla_y H_1(s, y_{[2]}, u_{1,1}, u_{2,2}, u_{12,12}, \psi_0, \psi_1(\cdot)); \]
\[ \psi_{2,i}(t) = \nabla_y H_2(t, y_{[1]}, u_{1,1}, u_{2,2}, u_{12,12}, \psi_0, \psi_2(\cdot)); \]
\[ \psi_{12,i}(s, t) = \nabla_y H_{12}(s, t, y_{[1]}, u_{1,1}, u_{2,2}, u_{12,12}, \psi_0, \psi_1(\cdot), \psi_2(\cdot), \psi_{12}(\cdot)) \]

--- (3.13)

For later reference, we also record below the Hamiltonian equations for the simplified model (2.11):

\[ \psi_{12}(s, t) = \nabla_y F_{12}(s, \ldots) + \psi_1(s) \nabla_y f_2(s, B, \ldots) + \psi_2(t) \nabla_y f_1(A, t, s, \ldots) + \]
\[ + \psi_0 \nabla_y f_{12}(A, B, s, t, \ldots) + \int_s^A \left[ \psi_{12}(\sigma, t) \nabla_y f_1(\sigma, t, s, \ldots) + \right. \]
\[ + \psi_1(\sigma) \nabla_y f_{12}(\sigma, B, s, t, \ldots) \right] d\sigma + \]
\[ + \int_t^B \left[ \psi_{12}(s, \tau) \nabla_y f_2(s, \tau, t, \ldots) + \psi_2(\tau) \nabla_y f_{12}(A, \tau, s, t, \ldots) \right] d\tau + \]
\[ + \int_s^A \int_t^B \psi_{12}(\sigma, \tau) \nabla_y f_{12}(\sigma, \tau, s, t, \ldots) d\tau d\sigma; \]
\[ \psi_1(s) = \nabla_y F_1(s, \ldots) + \psi_0 \nabla_y f_1(A, B, s, \ldots) + \int_s^A \psi_1(\sigma) \nabla_y f_1(\sigma, B, s, \ldots) d\sigma; \]
\[ \psi_2(t) = \nabla_y F_2(t, \ldots) + \psi_0 \nabla_y f_2(A, B, t, \ldots) + \int_t^B \psi_2(\tau) \nabla_y f_2(A, \tau, t, \ldots) d\tau; \]
\[ \psi_0 = \nabla_y F_0(y(A, B), \ldots) \]

--- (3.14)

The Hamiltonians for the simplified model (2.11) are
\[ h_0(y_{12}, u_{1,1}, u_{2,2}, u_{12,12}) := F_0(y_{12}, u_{1,1}, u_{2,2}, u_{12,12}); \]
\[ h_1(s, y_{2}, u_{1,1}, u_{2,2}, u_{12,12}, \psi_0, \psi_1(\cdot)) := F_1(s, y_{2}, u_{1,1}, u_{2,2}, u_{12,12}) + \]
\[ + \psi_0 f_1(A, B, s, y_{2}, u_{1,1}, u_{2,2}, u_{12,12}) + \]
\[ + \int_A \psi_1(\sigma) f_1(\sigma, B, s, y_{2}, u_{1,1}, u_{2,2}, u_{12,12}) d\sigma; \]
\[ h_2(t, y_{11}, u_{1,1,1}, u_{2,2,2}, u_{12,12,1}) := F_2(t, y_{11}, u_{1,1,1}, u_{2,2,2}, u_{12,12,1}) + \]
\[ + \psi_0 f_2(A, B, t, y_{11}, u_{1,1,1}, u_{2,2,2}, u_{12,12,1}) + \int_B \psi_2(\tau) f_2(\tau, t, y_{11}, u_{1,1,1}, u_{2,2,2}, u_{12,12,1}) d\tau; \]
\[ h_{12}(s, t, y, u_{1,1,2}, u_{12,12}, \psi_0, \psi_1(\cdot), \psi_2(\cdot), \psi_{12}(\cdot)) := \]
\[ = F_{12}(s, t, y, u_{1,1,2}, u_{12,12}) + \psi_0 f_{12}(A, B, s, t, y, u_{1,1,2}, u_{12,12}) + \]
\[ + \psi_1(s) f_{2}(s, B, t, y, u_{1,1,2}, u_{12,12}) + \psi_2(t) f_1(A, t, s, y, u_{1,1,2}, u_{12,12}) + \]
\[ + \int_A \int_B \psi_{12}(\sigma, \tau) f_{12}(\sigma, \tau, t, s, y, u_{1,1,2}, u_{12,12}) d\sigma d\tau = \]
\[ = \int_A \int_B \int_s t \psi_{12}(\sigma, \tau) f_{12}(\sigma, \tau, t, s, y, u_{1,1,2}, u_{12,12}) d\tau d\sigma \]

--- (3.15)
4. The variational equations.

In this section we shall carry out the exact calculation of the variation of the functional $J$ of section 2. In order to somewhat reduce the notational complexity, we shall examine in detail the simplified model.

Assuming adequate continuous differentiability of all functions involved, we calculate the variation of the functional $J$ under an admissible variation $\delta u$ of the control function $u$. We are using the standard definition of admissible variations, i.e. a variation $\delta u(\cdot)$ is said to be admissible relative to an admissible control function $u(\cdot)$ if there exists an $\varepsilon_0 > 0$ such that, for all $\varepsilon \in [0, \varepsilon_0]$, the function $u(\cdot) + \varepsilon \delta u(\cdot)$ is an admissible control function. The admissible control functions were defined in section 2. The set of admissible variations $\delta u(\cdot)$ relative to an admissible control function $u(\cdot)$ will be denoted by $A(u(\cdot))$. The variation $\delta y(s, t) = \delta y(s, t; u(\cdot), \delta u(\cdot))$ of the state is defined as

$$\delta y(s, t) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \left( y(s, t; u(\cdot) + \varepsilon \delta u(\cdot)) - y(s, t; u(\cdot)) \right)$$

--- (4.1)

Similarly, the variation $\delta J = \delta J(u(\cdot), \delta u(\cdot))$ of the functional $J$ is defined as

$$\delta J = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon} \left( J(u(\cdot) + \varepsilon \delta u(\cdot)) - J(u(\cdot)) \right)$$

--- (4.2)

Then we have

$$\delta y(s, t) = \int_0^s \left\{ (\nabla_y f_1(s, t, \sigma, y(\sigma, t), u(\sigma, t))) \delta y(\sigma, t) + 
+ (\nabla_u f_1(s, t, \sigma, y(\sigma, t), u(\sigma, t))) \delta u(\sigma, t) \right\} d\sigma + 
+ \int_0^t \left\{ (\nabla_y f_2(s, t, \tau, y(s, \tau), u(s, \tau))) \delta y(s, \tau) + 
+ (\nabla_u f_2(s, t, \tau, y(s, \tau), u(s, \tau))) \delta u(s, \tau) \right\} d\tau + 
+ \int_0^s \int_0^t \left\{ (\nabla_y f_{12}(s, t, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau))) \delta y(\sigma, \tau) + 
+ (\nabla_u f_{12}(s, t, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau))) \delta u(\sigma, \tau) \right\} d\tau d\sigma$$

--- (4.3)

and
\[ \delta J = (\nabla_y F_0(y(A, B)))\delta y(A, B) + \]
\[ + \int_0^A \{(\nabla_y F_1(s, y(s, B), u(s, B)))\delta y(s, B) + \]
\[ + (\nabla_u F_1(s, y(s, B), u(s, B)))\delta u(s, B)\}ds + \]
\[ + \int_0^B \{(\nabla_y F_2(t, y(A, t), u(A, t)))\delta y(A, t) + \]
\[ + (\nabla_u F_2(t, y(A, t), u(A, t)))\delta u(A, t)\}dt + \]
\[ + \int_0^A \int_0^B \{(\nabla_y F_{12}(s, t, y(s, t), u(s, t)))\delta y(s, t) + \]
\[ + \{(\nabla_y F_{12}(s, t, y(s, t), u(s, t)))\delta y(s, t)\}dtds \]

--- (4.4)

As it is well known, the necessary conditions for optimality are

\[ \delta J(u^*(.), \delta u(.)) \geq 0 \quad \forall \delta u(.) \in \mathbf{A}(u^*(.)) \]

--- (4.5)

In order to proceed further, we need a theory of resolvent kernels and duality for linear Goursat-Volterra equations of the type of (4.3). This theory is developed in the next section.
5. Linear Goursat-Volterra equations.

We consider a linear Goursat-Volterra equation

\[
z(s, t) = z_0(s, t) + \int_0^s K_1(s, t, \sigma)z(\sigma, t)d\sigma + \int_0^t K_2(s, t, \tau)z(s, \tau)d\tau + \\
+ \int_0^s \int_0^t K_{12}(s, t, \sigma, \tau)z(\sigma, \tau)d\tau d\sigma
\]

--- (5.1)

For the purpose of developing a theory of resolvent kernels, we denote by \( K \) the triple \((K_1, K_2, K_{12})\), and by \( K \) the integral operator

\[
(Kz)(s, t) := \int_0^s K_1(s, t, \sigma)z(\sigma, t)d\sigma + \int_0^t K_2(s, t, \tau)z(s, \tau)d\tau + \\
+ \int_0^s \int_0^t K_{12}(s, t, \sigma, \tau)z(\sigma, \tau)d\tau d\sigma
\]

--- (5.2)

We also consider a second operator \( L \) of the same type as \( K \), and the triple \( L = (L_1, L_2, L_{12}) \) associated with the operator \( L \):

\[
(Lw)(s, t) := \int_0^s L_1(s, t, \sigma)w(\sigma, t)d\sigma + \int_0^t L_2(s, t, \tau)w(s, \tau)d\tau + \\
+ \int_0^s \int_0^t L_{12}(s, t, \sigma, \tau)w(\sigma, \tau)d\tau d\sigma
\]

--- (5.3)

We are interested in the composition \( M := L \circ K \). As it will be shown, this composition is again a linear Goursat-Volterra operator, and it has the corresponding triple of kernels \( M = (M_1, M_2, M_{12}) \).

We have
\[(L \circ K)z(s, t) = \int_0^s \int_0^{\sigma_1} L_1(s, t, \sigma_1)K_1(\sigma_1, t, \sigma)z(\sigma, t)d\sigma d\sigma_1 + \]
\[+ \int_0^s \int_0^{\sigma_1} L_1(s, t, \sigma_1)K_2(\sigma_1, t, \tau)z(\sigma_1, \tau)d\tau d\sigma_1 + \]
\[+ \int_0^s \int_0^{\sigma_1} L_1(s, t, \sigma_1)K_{12}(\sigma_1, t, \sigma, \tau)z(\sigma, \tau)d\tau d\sigma d\sigma_1 + \]
\[+ \int_0^t \int_0^s L_2(s, t, \tau_1)K_1(s, \tau_1, \sigma)z(\sigma, \tau_1)d\sigma d\tau_1 + \]
\[+ \int_0^t \int_0^s \int_0^{\tau_1} L_2(s, t, \tau_1)K_{12}(s, \tau_1, \sigma, \tau)z(\sigma, \tau)d\tau d\sigma d\tau_1 + \]
\[+ \int_0^t \int_0^s \int_0^{\tau_1} L_{12}(s, t, \sigma_1, \tau_1)K_1(\sigma_1, \tau_1, \sigma)z(\sigma, \tau_1)d\sigma d\tau_1 d\sigma_1 + \]
\[+ \int_0^t \int_0^s \int_0^{\tau_1} \int_0^{\tau_1} L_{12}(s, t, \sigma_1, \tau_1)K_{12}(\sigma_1, \tau_1, \sigma, \tau)z(\sigma, \tau)d\tau d\sigma d\tau_1 d\sigma_1 \]

\[--- (5.4) \]

We define the 9 convolutions \( \otimes_{\alpha, \beta} \), where each of \( \alpha, \beta \) takes values in \( \{1, 2, 12\} \), by

\[
(L_1 \otimes_{1,1} K_1)(s, t, \sigma) := \int_0^s L_1(s, t, \sigma_1)K_1(\sigma_1, t, \sigma)d\sigma_1 ;
\]
\[
(L_2 \otimes_{2,2} K_2)(s, t, \tau) := \int_\tau^t L_2(s, t, \tau_1)K_2(s, \tau_1, \tau)d\tau_1 ;
\]
\[
(L_1 \otimes_{1,2} K_2)(s, t, \sigma) := L_1(s, t, \sigma)K_2(s, t, \sigma) ;
\]
\[
(L_2 \otimes_{2,1} K_1)(s, t, \tau) := L_2(s, t, \tau)K_1(s, t, \tau) ;
\]
\[
(L_1 \otimes_{1,12} K_{12})(s, t, \sigma, \tau) := \int_0^s L_1(s, t, \sigma_1)K_{12}(\sigma_1, t, \sigma, \tau)d\sigma_1 ;
\]
\[
(L_{12} \otimes_{12,1} K_1)(s, t, \sigma, \tau) := \int_\sigma^\tau L_{12}(s, t, \sigma_1, \tau)K_1(\sigma_1, t, \sigma, \tau)d\sigma_1 ;
\]
\[
(L_2 \otimes_{2,12} K_{12})(s, t, \sigma, \tau) := \int_\tau L_2(s, t, \tau_1)K_{12}(s, \tau_1, \sigma, \tau)d\tau_1 ;
\]
\[
(L_{12} \otimes_{12,2} K_2)(s, t, \sigma, \tau) := \int_\tau L_{12}(s, t, \sigma_1, \tau_1)K_2(\sigma_1, \tau_1, \sigma, \tau)d\tau_1 ;
\]
\[
(L_{12} \otimes_{12,12} K_{12})(s, t, \sigma, \tau) := \int_\sigma \int_\tau L_{12}(s, t, \sigma_1, \tau_1)K_{12}(\sigma_1, \tau_1, \sigma, \tau)d\tau_1 d\sigma_1 \]

\[--- (5.5) \]

It follows from (5.4) and (5.5) that
\[ M_1 = L_1 \otimes_{1,1} K_1, \quad M_2 = L_2 \otimes_{2,2} K_2, \quad M_{12} = \sum_{\alpha, \beta: (1,1) \neq (\alpha, \beta) \neq (2,2)} L_\alpha \otimes_{\alpha, \beta} K_\beta \]  \quad \text{--- (5.6)}

We shall write
\[ M = L \otimes K \]  \quad \text{--- (5.7)}

This convolution is associative but, in general, noncommutative. We set
\[ K^{\otimes N} := K \otimes K \otimes \cdots \otimes K \]  \quad \text{N times}  \quad \text{--- (5.8)}

For consistency of notation, we define the convolutions
\[
\begin{align*}
(K_1 \otimes_{1,0} z)(s,t) &:= \int_0^s K_1(s,t,\sigma)z(\sigma,t)d\sigma; \\
(K_2 \otimes_{2,0} z)(s,t) &:= \int_0^t K_2(s,t,\tau)z(s,\tau)d\tau; \\
(K_{12} \otimes_{12,0} z)(s,t) &:= \int_0^s \int_0^t K_{12}(s,t,\sigma,\tau)z(\sigma,\tau)d\tau d\sigma; \\
(K \otimes_0 z)(s,t) &:= (K_1 \otimes_{1,0} z)(s,t) + (K_2 \otimes_{2,0} z)(s,t) + (K_{12} \otimes_{12,0} z)(s,t)
\end{align*}
\]  \quad \text{--- (5.9)}

The construction of the convolution \( \otimes \) implies
\[ L \otimes_0 (K \otimes_0 z) = (L \otimes K) \otimes_0 z \]  \quad \text{--- (5.10)}

The solution of the integral equation (5.1) by the method of Picard iterations (as in section 2) leads therefore to the following formula for the N-th iteration:
\[ z(N)(s,t) = z_0(s,t) + \sum_{k=1}^N (K^{\otimes k} \otimes_0 z)(s,t) \]  \quad \text{--- (5.11)}

We set
\[ R(s, t, \sigma, \tau) = (R_1(s, t, \sigma), R_2(s, t, \tau), R_{12}(s, t, \sigma, \tau)) := \]
\[ = \left( \sum_{k=1}^{\infty} K_{1k}(s, t, \sigma), \sum_{k=1}^{\infty} K_{2k}(s, t, \tau), \sum_{k=1}^{\infty} K_{12k}(s, t, \sigma, \tau) \right) \]

--- (5.12)

Under the condition that all kernels \( K_1, K_2, K_{12} \) are measurable and bounded for all relevant values of \( s, t, \sigma, \tau \), we can prove the convergence of the 3 series in (5.12). We set

\[ \mathcal{D} := \{(s, t, \sigma, \tau) : 0 \leq \sigma \leq s \leq A, 0 \leq \tau \leq t \leq B\} \]

--- (5.13)

and we denote by \( C \) a common bound on \( K_1, K_2, K_{12} \):

\[ |K_1(s, t, \sigma)| \leq C, |K_2(s, t, \tau)| \leq C, |K_{12}(s, t, \sigma, \tau)| \leq C, \forall (s, t, \sigma, \tau) \in \mathcal{D} \]

--- (5.14)

We have:

**Lemma 5.1.** The terms \((K^{\otimes k})_1, (K^{\otimes k})_2 \) satisfy

\[ |(K^{\otimes k})_1(s, t, \sigma)| \leq C_k \frac{(s-\sigma)^{k-1}}{(k-1)!}, |(K^{\otimes k})_2(s, t, \tau)| \leq C_k \frac{(t-\tau)^{k-1}}{(k-1)!}, \forall (s, t, \sigma, \tau) \in \mathcal{D} \]

--- (5.15)

**Proof:** According to (5.5) and (5.6), we have

\[ (K^{\otimes k})_1 = (K_1)^{\otimes k} \equiv \underbrace{K_1 \otimes K_1 \otimes \ldots \otimes K_1}_{k \text{ times}} \]

--- (5.16)

and consequently the estimate \(|(K^{\otimes k})_1(s, t, \sigma)| \leq C_k \frac{(s-\sigma)^{k-1}}{(k-1)!} \) follows by the same method as for standard Volterra equations (equations with single integrals). The estimate for \((K^{\otimes k})_2 \) is, of course, justified in the same way. ///

**Lemma 5.2.** We have the estimate
\[ |(K^{\otimes k})_{12}(s,t,\sigma,\tau)| \leq C^k \sum_{\lambda=1}^{N_k} \varphi_{k,\lambda}(s-\sigma,t-\tau); \quad N_k := 3^k - 2 \]  
\hfill (5.17)

where each \( \varphi_{k,\lambda} \) is a term of the form

\[ \varphi_{k,\lambda}(s-\sigma,t-\tau) = \frac{(s-\sigma)^{\alpha_{k,\lambda}}(t-\tau)^{\beta_{k,\lambda}}}{(\alpha_{k,\lambda})!(\beta_{k,\lambda})!} \]  
\hfill (5.18)

with \( \alpha_{k,\lambda}, \beta_{k,\lambda} \) nonnegative integers satisfying

\[ \alpha_{k,\lambda} + \beta_{k,\lambda} \geq k - 2, \quad \max(\alpha_{k,\lambda}, \beta_{k,\lambda}) \leq k - 1 \]  
\hfill (5.19)

(The terms \( \varphi_{k,\lambda} \) appearing in the summation in (5.17) are not necessarily distinct: the summation may include repeated terms, but each term is of the form specified in (5.18) and (5.19). The set of pairs \( (\alpha_{k,\lambda}, \beta_{k,\lambda}) \) that actually appear in the terms in the summation in (5.17) does not necessarily comprise all possible pairs of nonnegative integers that satisfy (5.19).)

**Proof:** The numbers \( N_k \) are solutions of the finite-difference equation \( N_{k+1} = 4 + 3N_k \) with initial condition \( N_1 = 1 \). The justification for these numbers comes from the number of terms (each term being of the form \( C^k \varphi_{k,\lambda}(s-\sigma,t-\tau) \)) involved in the estimation of \( (K^{\otimes(k+1)})_{12} \) in each new iteration, assuming \( K^{\otimes k} \) satisfies the asserted estimates. The kernel \( (K^{\otimes(k+1)})_{12} \) is expressed as

\[ (K^{\otimes(k+1)})_{12} = (K_1 \otimes_1 k)K_2 + (K_2 \otimes_2 k)K_1 + (K_1 \otimes_1 k) \otimes_1,12 K_{12} + \]  
\hfill (5.20)

\[ + (K_2 \otimes_2 k) \otimes_2,12 K_{12} + (K^{\otimes k})_{12} \otimes_1,12 K_{12} + (K^{\otimes k})_{12} \otimes_2,12 K_{12} \]

Now, the right-hand side of (5.20) contains the four terms

\( (K_1 \otimes_1 k)K_2, (K_2 \otimes_2 k)K_1, (K_1 \otimes_1 k) \otimes_1,12 K_{12}, (K_2 \otimes_2 k) \otimes_2,12 K_{12} \).
which contribute, respectively, terms of the form
\[
C^{k+1} \frac{(s-\sigma)^{k-1}}{(k-1)!}, \ C^{k+1} \frac{(t-\tau)^{k-1}}{(k-1)!}, \ C^{k+1} \frac{(s-\sigma)^k}{k!}, \ C^{k+1} \frac{(t-\tau)^k}{k!}
\]
to the estimation of \(|(K^{\otimes(k+1)})_{12}(s,t,\sigma,\tau)|\); each of the remaining terms
\((K^{\otimes k})_{12} \otimes_{12,1} K_1, (K^{\otimes k})_{12} \otimes_{12,2} K_2, (K^{\otimes k})_{12} \otimes_{12,12} K_{12}\)
contributes \(N_k\) terms to the estimation of \(|(K^{\otimes(k+1)})_{12}(s,t,\sigma,\tau)|\); thus \(N_{k+1} = 4 + 3N_k\). For each term of the form
\[
C^k \frac{(s-\sigma)^{\alpha_{k,\lambda}}(t-\tau)^{\beta_{k,\lambda}}}{(\alpha_{k,\lambda})!(\beta_{k,\lambda})!}
\]
of the estimate for \(|(K^{\otimes k})_{12}(s,t,\sigma,\tau)|\), the three terms
\((K^{\otimes k})_{12} \otimes_{12,1} K_1, (K^{\otimes k})_{12} \otimes_{12,2} K_2, (K^{\otimes k})_{12} \otimes_{12,12} K_{12}\)
contribute, respectively, the terms
\[
C^{k+1} \frac{(s-\sigma)^{\alpha_{k,\lambda}+1}(t-\tau)^{\beta_{k,\lambda}}}{(\alpha_{k,\lambda}+1)!(\beta_{k,\lambda})!}, \ C^{k+1} \frac{(s-\sigma)^{\alpha_{k,\lambda}}(t-\tau)^{\beta_{k,\lambda}+1}}{(\alpha_{k,\lambda})!(\beta_{k,\lambda}+1)!},
\]
\[
C^{k+1} \frac{(s-\sigma)^{\alpha_{k,\lambda}+1}(t-\tau)^{\beta_{k,\lambda}+1}}{(\alpha_{k,\lambda}+1)!(\beta_{k,\lambda}+1)!}
\]
to the estimation of \(|(K^{\otimes(k+1)})_{12}(s,t,\sigma,\tau)|\). The pairs \((\alpha_{k+1,\lambda_i}, \beta_{k+1,\lambda_i})\), \(i = 1, \ldots, 7\), defined by specifying that they take the values
\((k-1,0),(0,k-1),(k,0),(0,k),(\alpha_{k,\lambda_i}+1,\beta_{k,\lambda_i}),(\alpha_{k,\lambda_i},\beta_{k,\lambda_i}+1),(\alpha_{k,\lambda_i}+1,\beta_{k,\lambda_i}+1)\),
satisfy the conditions \(\alpha_{k+1,\lambda_i}+\beta_{k+1,\lambda_i} \geq k-1\), \(\max(\alpha_{k+1,\lambda_i}, \beta_{k+1,\lambda_i}) \leq k\) provided
\(\alpha_{k,\lambda_i}+\beta_{k,\lambda_i} \geq k-2\), \(\max(\alpha_{k,\lambda_i}, \beta_{k,\lambda_i}) \leq k-1\). This is the inductive step in proving (5.19) for all \(k\); for \(k=1\) and \(k=2\), the validity of the wanted inequalities is verifiable by direct calculation. ///

**Theorem 5.1.** We set
\[
Q := \max(A,B,1), \ m_k = \left\lfloor \frac{k}{2} - 1 \right\rfloor
\]
--- (5.21)

(where \(\lfloor w \rfloor\) signifies the greatest among all integers that do not exceed the real number \(w\), in other words \(\lfloor w \rfloor\) is the "floor function" of \(w\)).

Then we have
\[
|(K^{\otimes k})_{12}(s,t,\sigma,\tau)| \leq \frac{(3C)^k}{[(m_k)!][k-2-m_k]!} Q^{2(k-1)}
\]
--- (5.22)
Proof: Among all pairs \((\alpha_{k,\lambda}, \beta_{k,\lambda})\) that satisfy \(\alpha_{k,\lambda} + \beta_{k,\lambda} = m\), the maximum of 

\[
\frac{1}{(\alpha_{k,\lambda})!(\beta_{k,\lambda})!}
\]

is achieved at \(\alpha_{k,\lambda} = \alpha_m := \left\lfloor \frac{m}{2} \right\rfloor\), and that maximum value is equal to

\[
\frac{1}{[(\alpha_m!)![m - \alpha_m!]]!}
\]

(This is tantamount to the well-known result that, in the expansion of \((x + y)^m\), the maximum of the coefficients is in the middle if \(m\) is even, and it is achieved at the two terms in the middle if \(m\) is odd.) The smallest possible value of \(m\), in the context of the estimates for \(|K_{12}(s,t,\sigma,\tau)|\), is \(k-2\), and the corresponding maximum over \(m\) is therefore achieved at \(m_k^*\). The sum of the exponents \(\alpha_{k,\lambda} + \beta_{k,\lambda}\) does not exceed \(2k-2\), and consequently the powers \((s-\sigma)^{\alpha_{k,\lambda}}(t-\tau)^{\beta_{k,\lambda}}\) do not exceed \(Q^{2(k-1)}\). The number of terms, \(N_k\), in the estimate (5.17), does not exceed \(3^k\). Thus the assertion of this theorem has been proved. ///

Since the estimates on \((K_{12}^k)_1, (K_{12}^k)_2, (K_{12}^k)_{12}\) are all terms of absolutely convergent series, we have the following:

**Theorem 5.2.** Under the conditions stated before lemma 5.1, the 3 series defining the resolvent kernel \(R\) in (5.12) converge uniformly in \(D\). ///

The construction of \(R\) and the representation of \(z_{(N)}\) in (5.11) imply the following:

**Theorem 5.3.** The solution \(z(s, t)\) of (5.1) is given by

\[
z(s, t) = z_0(s, t) + (R \otimes z_0)(s, t)
\]

--- (5.23)

The resolvent kernel \(R\) satisfies

\[
K \otimes R = R \otimes K = R - K
\]

--- (5.24)

///

We now define and analyze an equation of a type adjoint to (5.1):
\[ \zeta(s, t) = \zeta_0(s, t) + \int_A^B \zeta(\sigma, t)K_1(\sigma, t, s)d\sigma + \int_A^B \zeta(s, \tau)K_2(s, \tau, t)d\tau + \]
\[ + \int_s^t \int_s^t \zeta(\sigma, \tau)K_{12}(\sigma, \tau, s, t)d\tau d\sigma \]

--- (5.25)

where the values of \( \zeta(s, t) \) are \( n \)-dimensional row vectors.

We shall prove:

**Theorem 5.4.** For initial function \( \zeta_0 \) that is continuous on \( D \), and under the conditions of this section on the kernel \( K \), the solution of (5.25) is given by

\[ \zeta(s, t) = \zeta_0(s, t) + \int_A^B \zeta_0(\sigma, t)R_1(\sigma, t, s)d\sigma + \int_A^B \zeta_0(s, \tau)R_2(s, \tau, t)d\tau + \]
\[ + \int_s^t \int_s^t \zeta_0(\sigma, \tau)R_{12}(\sigma, \tau, s, t)d\tau d\sigma \]

--- (5.26)

**Proof:** Suppose \( \zeta \) is given by (5.26); we shall show that \( \zeta \) solves (5.25).

For convenience, we define the convolution \( \otimes \) by

\[ (\zeta_0 \otimes M)(s, t) = \int_A^B \zeta_0(\sigma, t)M_1(\sigma, t, s)d\sigma + \int_A^B \zeta_0(s, \tau)M_2(s, \tau, t)d\tau + \]
\[ + \int_s^t \int_s^t \zeta_0(\sigma, \tau)M_{12}(\sigma, \tau, s, t)d\sigma d\tau \]

--- (5.27)

with the operations corresponding to the 3 terms on the right-hand side of (5.27) denoted by \( \otimes_{1,1}, \otimes_{0,2}, \otimes_{0,12} \), respectively, i.e.
With $\zeta$ given by (5.26), we calculate

\[
\begin{align*}
(\zeta_0 \otimes_{0,1} M_1)(s,t) &:= \int_{s}^{A} \int_{t}^{B} \zeta_0(\sigma,t)M_1(\sigma,t,s)d\sigma; \\
(\zeta_0 \otimes_{0,2} M_2)(s,t) &:= \int_{t}^{B} \zeta_0(s,\tau)M_2(s,\tau,t)d\tau; \\
(\zeta_0 \otimes_{0,12} M_{12})(s,t) &:= \int_{s}^{A} \int_{t}^{B} \zeta_0(\sigma,\tau)M_{12}(\sigma,\tau,s,t)d\sigma d\tau
\end{align*}
\] --- (5.28)

With $\zeta$ given by (5.26), we calculate
It is seen from (5.29) that, if \( \zeta \) is given by (5.26), then
Indeed, those terms in (5.29) that contain \( R_1, R_2, R_{12} \) make up \( \zeta_0 \otimes_0 (R \otimes K) \). We verify that one of those terms of (5.29) is one of the terms that make up \( \zeta_0 \otimes_0 (R \otimes K) \):

\[
\begin{align*}
\int_A^B \int_T^S \int_\sigma^\tau \zeta_0(\sigma_1, \tau)R_1(\sigma_1, \tau, \sigma)K_{12}(\sigma, \tau, s, t)d\sigma_1d\tau d\sigma = \\
= \int_A^B \int_T^S \int_\sigma^\tau \zeta_0(\sigma_1, \tau)(R_1 \otimes_{1,12} K_{12})(\sigma_1, \tau, s, t)d\tau d\sigma_1 = \\
= (\zeta_0 \otimes_{0,12} (R_1 \otimes_{1,12} K_{12}))(s, t)
\end{align*}
\]

--- (5.31)

Of course, all the relevant terms of (5.29) can be handled similarly. On the basis of the identity \( R \otimes K = R - K \), we have

\[
\zeta_0 \otimes_0 K + \zeta_0 \otimes_0 (R \otimes K) = \zeta_0 \otimes_0 R
\]

--- (5.32)

Therefore, if \( \zeta \) is given by (5.26), we have

\[
\zeta_0 + \zeta \otimes_0 K = \zeta_0 + \zeta_0 \otimes_0 K + \zeta_0 \otimes_0 (R \otimes K) = \zeta_0 + \zeta_0 \otimes_0 R = \zeta
\]

--- (5.33)

which is another way of writing (5.25). ///
6. Representation of the co-state.

For the simplified model (2.11), we define the co-state $\psi := [\psi_0 \psi_1(,) \psi_2(,) \psi_{12}(,)]$ by

$$\psi_0 := \nabla_y F_0(y(A,B),...)$$;

$$\psi_1(s) := \nabla_y F_1(s,...) + \psi_0 \nabla_y f_1(A,B,s,...) + \int_s^A \{ \nabla_y F_1(\sigma,...) + \psi_0 \nabla_y f_1(A,B,\sigma,...) \} R_1(\sigma,B,s) \, d\sigma$$;

$$\psi_2(t) := \nabla_y F_2(t,...) + \psi_0 \nabla_y f_2(A,B,t,...) + \int_t^B \{ \nabla_y F_2(\tau,...) + \psi_0 \nabla_y f_2(A,B,\tau,...) \} R_1(\tau,A,t) \, d\tau$$;

$$\psi_{12}(s,t) := \nabla_y F_{12}(s,t,...) + \psi_1(s) \nabla_y f_2(s,B,t,...) + \psi_2(t) \nabla_y f_1(A,t,s,...) +$$

$$\int_s^A \{ \nabla_y F_{12}(\sigma,t,...) + \psi_1(\sigma) \nabla_y f_2(\sigma,B,t,...) + \psi_2(t) \nabla_y f_1(A,t,\sigma,...) \} \, d\sigma +$$

$$\int_t^B \{ \nabla_y F_{12}(s,\tau,...) + \psi_1(s) \nabla_y f_2(s,B,\tau,...) + \psi_2(\tau) \nabla_y f_1(A,\tau,s,...) \} \, d\tau$$

$$+ \psi_0 \nabla_y f_{12}(A,B,s,t,...) \, d\sigma +$$

$$+ \psi_0 \nabla_y f_{12}(A,B,\sigma,t,...) R_2(s,\tau,t) \, d\tau + \psi_0 \nabla_y f_{12}(A,B,\sigma,t) R_1(\sigma,t,s) \, d\sigma +$$

$$+ \int_s^A \int_t^B \{ \nabla_y F_{12}(\sigma,\tau,...) + \psi_1(\sigma) \nabla_y f_2(\sigma,B,\tau,...) + \psi_2(\tau) \nabla_y f_1(A,\tau,\sigma,...) \} \, d\sigma \, d\tau$$

$$+ \int_s^A \int_t^B \{ \psi_1(\sigma) \nabla_y f_{12}(\sigma_1,B,\sigma,t,...) R_1(\sigma,t,s) \, d\sigma_1 \, d\sigma +$$

$$+ \int_s^A \int_t^B \{ \psi_2(\tau) \nabla_y f_{12}(A,\tau,\sigma,t,...) R_1(\sigma,t,s) \, d\tau \, d\sigma +$$

$$+ \psi_0 \nabla_y f_{12}(A,B,s,t,...) R_2(s,\tau,t) \, d\tau \, d\sigma +$$
\[
\begin{align*}
B & B \\
+ \int_{t}^{A} \int_{\tau}^{B} \psi_2(\tau_1) V_y f_{12}(A, \tau_1, s, \tau, ...) R_2(s, \tau, t) d\tau_1 d\tau + \\
& + \int_{s}^{A} \int_{t}^{B} \left\{ \int_{\sigma}^{A} \psi_1(\sigma_1) V_y f_{12}(\sigma_1, B, \sigma, \tau, ...) d\sigma_1 \right\} R_{12}(\sigma, \tau, s, t) d\tau d\sigma \\
\end{align*}
\]  
\[--- (6.1)\]

where \( R = (R_1, R_2, R_{12}) \) is the resolvent kernel associated with

\[
K(s, t, \sigma, \tau) = (K_1(s, t, \sigma), K_2(s, t, \tau), K_{12}(s, t, \sigma, \tau)) = (V_y f_1(s, t, \sigma, ...), V_y f_2(s, t, \tau, ...), V_y f_{12}(s, t, \sigma, \tau, ...))
\]  
\[--- (6.2)\]

The Hamiltonians \( h_0, h_1, h_2, h_{12} \) are given by (3.15), but now with \( \psi_0, \psi_1, \psi_2, \psi_{12} \) defined as in (6.1). It now follows that, when \( (\psi_0, \psi_1, \psi_2, \psi_{12}) \) are defined as in (6.1), then they satisfy the system of Hamiltonian equations (3.14), on the basis of the results on duality of linear Goursat-Volterra problems that were proved in section 5.

The variation \( \delta y(s, t) \), obtained in (4.3), can be expressed, in view of the results of section 5, as
\[
\begin{align*}
\delta y(s,t) &= \int_0^s \int_0^t \nabla_u f_1(s,t,\sigma,...) \delta u(\sigma, t) d\sigma + \int_0^t \nabla_u f_2(s,t,\tau,...) \delta u(s, \tau) d\tau + \\
&+ \int_0^s \int_0^t \nabla_u f_{12}(s,t,\sigma,\tau,...) \delta u(\sigma, \tau) d\tau d\sigma + \\
&+ \int_0^s \int_0^t R_1(s,t,\sigma) \nabla_u f_1(s,t,\sigma_1,...) \delta u(\sigma_1, t) d\sigma_1 d\sigma + \\
&+ \int_0^s \int_0^t R_1(s,t,\sigma) \nabla_u f_2(s,t,\tau,...) \delta u(\sigma, \tau) d\tau d\sigma + \\
&+ \int_0^s \int_0^t \int_0^t R_1(s,t,\sigma) \nabla_u f_{12}(s,t,\sigma_1,\tau,...) \delta u(\sigma_1, \tau) d\tau d\sigma_1 d\sigma + \\
&+ \int_0^t \int_0^t R_2(s,t,\tau) \nabla_u f_1(s,t,\sigma,...) \delta u(\sigma, \tau) d\sigma d\tau + \\
&+ \int_0^t \int_0^t R_2(s,t,\tau) \nabla_u f_2(s,t,\tau_1,...) \delta u(s, \tau_1) d\tau_1 d\tau + \\
&+ \int_0^t \int_0^t \int_0^t R_2(s,t,\tau) \nabla_u f_{12}(s,t,\sigma,\tau,\tau_1,...) \delta u(\sigma, \tau_1) d\tau_1 d\sigma d\tau + \\
&+ \int_0^t \int_0^t \int_0^t R_2(s,t,\tau) \nabla_u f_{12}(s,t,\sigma,\tau,\tau_1,...) \delta u(\sigma, \tau_1) d\tau_1 d\sigma d\tau + \\
&+ \int_0^t \int_0^t \int_0^t \int_0^t R_2(s,t,\tau) \nabla_u f_{12}(s,t,\sigma,\tau,\tau_1,...) \delta u(\sigma, \tau_1) d\tau_1 d\sigma d\tau d\sigma \\
\end{align*}
\]

In order to simplify the presentation in this part of the paper, we write down the terms for the representation of the variation of the state.

Terms in the expression for \(\delta y(s,t)\):
\[ V_u f_1(s, t, \sigma, ...) \delta u(\sigma, t) \quad \{\sigma\} \]
\[ V_u f_2(s, t, \tau, ...) \delta u(s, \tau) \quad \{\tau\} \]
\[ V_u f_{12}(s, t, \sigma, \tau, ...) \delta u(\sigma, \tau) \quad \{\sigma, \tau\} \]

\[ R_1(s, t, \sigma) \]
\[ \left\{ \begin{array}{l}
    V_u f_1(s, t, \sigma_1, ...) \delta u(\sigma_1, t) \quad \{\sigma_1, \sigma\} \\
    V_u f_2(s, t, \tau, ...) \delta u(s, \tau) \quad \{\tau, \sigma\} \\
    V_u f_{12}(s, t, \sigma_1, \tau, ...) \delta u(\sigma_1, \tau) \quad \{\tau, \sigma_1, \sigma\}
\end{array} \right. \]

\[ R_2(s, t, \tau) \]
\[ \left\{ \begin{array}{l}
    V_u f_1(s, \tau, \sigma, ...) \delta u(\sigma, \tau) \quad \{\sigma, \tau\} \\
    V_u f_2(s, \tau, \tau_1, ...) \delta u(s, \tau_1) \quad \{\tau_1, \tau, \sigma\} \\
    V_u f_{12}(s, \tau, \sigma_1, \tau_1, ...) \delta u(\sigma_1, \tau_1) \quad \{\tau_1, \sigma_1, \tau, \sigma\}
\end{array} \right. \]

Terms in the expression for \( \delta y(s, B) \):

\[ V_u f_1(s, B, \sigma, ...) \delta u(\sigma, B) \quad \{\sigma\} \]
\[ V_u f_2(s, B, \tau, ...) \delta u(s, \tau) \quad \{\tau\} \]
\[ V_u f_{12}(s, B, \sigma, \tau, ...) \delta u(\sigma, \tau) \quad \{\sigma, \tau\} \]
Terms in the expression for $\delta y(A, t)$:

\[
\begin{align*}
\nabla_u f_1(A, t, \sigma, ...) & \delta u(\sigma, t) \quad \{\sigma\} \\
\nabla_u f_2(A, t, \tau, ...) & \delta u(A, \tau) \quad \{\tau\} \\
\nabla_u f_{12}(A, t, \sigma, \tau, ...) & \delta u(\sigma, \tau) \quad \{\sigma, \tau\}
\end{align*}
\]

\[
\begin{align*}
R_1(A, t, \sigma) & \quad \{\sigma_1, \sigma\} \\
R_2(A, t, \tau) & \quad \{\tau_1, \tau\} \\
R_{12}(A, t, \sigma, \tau) & \quad \{\tau_1, \sigma, \tau\}
\end{align*}
\]

Terms in the expression for $\delta y(A, B)$:

\[
\begin{align*}
\nabla_u f_1(A, B, \sigma, ...) & \delta u(\sigma, B) \quad \{\sigma\} \\
\nabla_u f_2(A, B, \tau, ...) & \delta u(A, \tau) \quad \{\tau\} \\
\nabla_u f_{12}(A, B, \sigma, \tau, ...) & \delta u(\sigma, \tau) \quad \{\sigma, \tau\}
\end{align*}
\]
Next, we introduce the following notation:

\[ u_{[0]} := \text{col}(u_{1,1}, u_{2,2}, u_{12,12}); \]
\[ u_{[1]} := \text{col}(u_{1,1}, u_{2,2}, u_{12,2}); \]
\[ u_{[2]} := \text{col}(u_{1,1}, u_{2,2}, u_{12,2}); \]
\[ u_{[12]} := \text{col}(u_{1,1}, u_{2,2}, u_{12}) \]

--- (6.4)

We claim that the variation \( \delta J \) can be represented as

\[
\delta J = \int_0^A \int_0^B \nabla u_{[12]} h_{12}(s, t, y(s, t), u_1(s), u_2(t), u_{12}(s, t), \psi_0, \psi_1(\cdot), \psi_2(\cdot), \psi_{12}(\cdot)) \cdot \delta u_{[12]}(s, t)dt \, ds +
\]
\[
+ \int_0^A \nabla u_{[1]} h_{1}(s, y(s, B), u_1(s), u_2(B), u_{12}(s, B), \psi_0, \psi_1(\cdot)) \delta u_{[1]}(s)ds +
\]
\[
+ \int_0^B \nabla u_{[2]} h_{2}(t, y(A, t), u_1(A), u_2(t), u_{12}(A, t), \psi_0, \psi_2(\cdot)) \delta u_{[2]}(t)dt +
\]
\[
+ \nabla u_{[0]} h_0(y(A, B), u_1(A), u_2(B), u_{12}(A, B), \psi_0) \delta u_{[0]}
\]

--- (6.5)

The proof of (6.5) relies on taking inventory of all the terms that appear in the expression for \( \delta J \) in (4.4), using the expressions for \( \delta y \) in tables 6.1 through 6.4 above, and comparing with the analogous inventory for the terms on the right-hand side of (6.5). Because of the large number of terms involved, we shall illustrate this comparison for only a few of the relevant terms.
The term \( \int \int (\nabla_y F_{12}(s,t,y(s,t),u(s,t))\delta y(s,t)) dt ds \) out of (4.4) involves all the terms for \( \delta y(s,t) \). Out of those terms, we look at these 3 terms:

\[
R_{12}(s,t,\tau,\sigma) = \begin{pmatrix}
\nabla_u f_1(\sigma,\tau,\sigma_1,\ldots)\delta u(\sigma_1,\tau) & \{\sigma_1,\tau,\sigma\} \\
\nabla_u f_2(\sigma,\tau,\tau_1,\ldots)\delta u(\sigma,\tau_1) & \{\tau_1,\tau,\sigma\} \\
\nabla_u f_{12}(\sigma,\tau,\sigma_1,\tau_1,\ldots)\delta u(\sigma_1,\tau_1) & \{\tau_1,\sigma_1,\tau,\sigma\}
\end{pmatrix}
\]

The term \( R_{12}(s,t,\sigma,\tau) \nabla_u f_{12}(\sigma,\tau,\sigma_1,\tau_1,\ldots)\delta u(\sigma_1,\tau_1) \), multiplied by \( \nabla_y F_{12}(s,t,\ldots) \), is matched with the term \( \nabla_y F_{12}(\sigma,\tau,\ldots)R_{12}(\sigma,\tau,s,t,\ldots) \), inside the integral \( \int \int \) in the definition (6.1) of \( \psi_{12} \), multiplied by \( f_{12} \) in the definition of \( h_{12} \) and then differentiated with respect to \( u \). The term \( R_{12}(s,t,\sigma,\tau)\nabla_u f_1(\sigma,\tau,\sigma_1,\ldots)\delta u(\sigma_1,\tau) \), multiplied by \( \nabla_y F_{12}(s,t,\ldots) \), is matched with the term \( \nabla_y F_{12}(\sigma,\tau,\ldots)R_{12}(\sigma,\tau,s,t,\ldots) \), inside the integral \( \int \int \) in the definition (6.1) of \( \psi_{12} \), multiplied by \( f_1 \) to produce another one of the terms in the definition of \( h_{12} \), and then differentiated with respect to \( u \). The term \( R_{12}(s,t,\sigma,\tau)\nabla_u f_{12}(\sigma,\tau,\tau_1,\ldots)\delta u(\sigma,\tau_1) \) is analogous to the term

\( R_{12}(s,t,\sigma,\tau)\nabla_u f_1(\sigma,\tau,\sigma_1,\ldots)\delta u(\sigma_1,\tau) \).

Other terms are matched in a similar way.

Therefore, we have proved:

**Theorem 6.1.** For the simplified model (2.11) and under the conditions stated in this section, the co-state \( (\psi_0,\psi_1,\psi_2,\psi_{12}) \) satisfies equation (3.14), and the variation of the cost functional \( J \) is given by (6.5). ///
7. Partial extremal principles.

The optimal control problem with state equation (2.11), and, \textit{a fortiori}, the more general model (2.1), does not generally admit the derivation of necessary optimality conditions of the type of an extremal principle akin to the famous maximum principle of Pontryagin et al. [PBMG]. This is due to the fact that the controls that depend on a one-dimensional variable (only \(s\) or only \(t\)) appear both in single integrals and in a double integral, in the definition of the cost functional \(J\), as well as in the variational equation (6.5); as a consequence, the integral expression (6.5) cannot be simultaneously reduced to pointwise form in the entire triple \((u_1(s), u_2(t), u_{12}(s, t))\). If we were to take, simultaneously, needle-shaped variations in all three controls, the passage to pointwise form would require division by \(\delta s\) or \(\delta t\) in the single integrals, and passage to the limit as \(\delta s \to 0^+\) or \(\delta t \to 0^+\), separately for each increment \(\delta s\) or \(\delta t\), but it would require division by the product \((\delta s)(\delta t)\) and passage to the limit as \((\delta s, \delta t) \to (0^+, 0^+)\) simultaneously for the Hamiltonian \(h_{12}\) inside the double integral. The incompatibility of the two types of operations makes it impossible to obtain a pointwise form of the variational equation (6.5) jointly in all components of the control.

The only case in which it is possible to have an extremum principle akin to Pontryagin's maximum principle is the case in which only the control \(u_{12}\) is present.

In general, we can have only a weaker statement of an extremum principle; we shall call this result the partial extremum principle. This result states that, if

\[
(u_1^*(s), u_2^*(t), u_{12}^*(s, t), u_1^*(A), u_1^*(B), u_{12}^*(s, B), u_{12}^*(A, t), u_{12}^*(A, B))
\]

is an optimal control, \((y^*(s, B), y^*(A, t), y^*(A, B))\) is the corresponding optimal trajectory (solution of the state equation with control

\[
(u_1^*(s), u_2^*(t), u_{12}^*(s, t), u_1^*(A), u_1^*(B), u_{12}^*(s, B), u_{12}^*(A, t), u_{12}^*(A, B))
\]

and

\[
(\psi_0^*, \psi_1^*(s, B), \psi_2^*(t), \psi_{12}^*(s, t))
\]

is the associated co-state, then, at points of continuity of

\[
(u_1^*(s), u_2^*(t), u_{12}^*(s, t), u_1^*(A), u_1^*(B), u_{12}^*(s, B), u_{12}^*(A, t), u_{12}^*(A, B))
\]

if \(f_0, f_2\) are independent of \(u_1\), then \(u_1^*(s)\) minimizes, over the variable \(u_1\), the function

\[
\int_B h_1(s, y^*(s, B), u_1, u_2^*(B), u_{12}^*(s, B), \psi_0^*, \psi_1^*(.) + \\
\int_0^{B} h_{12}(s, t, y^*(s, t), u_1, u_2^*(t), u_{12}^*(s, t), \psi_0^*, \psi_1^*(.), \psi_2^*(.), \psi_{12}^*(., .))dt
\]

if \(f_0, f_1\) are independent of \(u_2\), then \(u_2^*(t)\) minimizes, over the variable \(u_2\), the function

\[
\int_B h_1(s, y^*(s, B), u_1, u_{12}^*(s, B), \psi_0^*, \psi_1^*(.) + \\
\int_0^{B} h_{12}(s, t, y^*(s, t), u_1, u_{12}^*(t), u_{12}^*(s, t), \psi_0^*, \psi_1^*(.), \psi_2^*(.), \psi_{12}^*(., .))dt
\]
\[ h_2(t,y^*(A,t),u_1^*(A),u_2,u_{12}^*(A,t),\psi_0^*,\psi_2^*(\cdot)) + \]
\[ + \int_0^{A} h_{12}(s,t,y^*(s,t),u_1^*(s),u_2,u_{12}^*(s,t),\psi_0^*,\psi_1^*,\psi_2^*(\cdot),\psi_{12}^*(\cdot))ds \]

if \( f_0, f_1, f_2 \) are independent of \( u_{12} \), then \( u_{12}^*(s,t) \) minimizes, over the variable \( u_{12} \), the function \( h_{12}(s,t,y^*(s,t),u_1^*(s),u_2^*(t),u_{12},\psi_0^*,\psi_1^*,\psi_2^*(\cdot),\psi_{12}^*(\cdot)) \) ;

if \( f_0 \) is independent of \( u_{12} \), then \( u_{12}^*(A,t) \) minimizes, over the variable \( u_{12,1} \), the function \( h_2(t,y^*(A,t),u_1^*(A),u_2^*(t),u_{12,1},\psi_0^*,\psi_2^*(\cdot)) \) ;

if \( f_0 \) is independent of \( u_{12} \), then \( u_{12}^*(s,B) \) minimizes, over the variable \( u_{12,2} \), the function \( h_1(s,y^*(s,B),u_1^*(s),u_2^*(B),u_{12,2},\psi_0^*,\psi_1^*(\cdot)) \) ;

\( u_1^*(A) \) minimizes, over the variable \( u_{1,1} \), the function \( h_0(y^*(A,B),u_{1,1},u_2^*(B),u_{12}^*(A,B)) + \)
\[ + \int_0^{B} F_2(t,y^*(A,t),u_{1,1},u_2^*(t),u_{12}^*(A,t))dt \]

\( u_2^*(B) \) minimizes, over the variable \( u_{2,2} \), the function \( h_0(y^*(A,B),u_1^*(A),u_{2,2},u_{12}^*(A,B)) + \)
\[ + \int_0^{A} F_1(s,y^*(s,B),u_1^*(s),u_{2,2},u_{12}^*(s,B))ds \]

\( u_{12}^*(A,B) \) minimizes, over the variable \( u_{12,12} \), the function \( h_0(y^*(A,B),u_1^*(A),u_2^*(B),u_{12,12}) \).
We shall prove one of the above cases of a partial extremum principle, the proofs in the remaining cases being similar. We prove that, under the stated conditions, \( u_{12}^* (s,t) \) minimizes, over the variable \( u_{12} \), the function

\[
h_{12}(s, t, y^*(s, t), u_1^*(s), u_2^*(t), u_{12}, \psi_0^*, \psi_1^*, \psi_2^*), \psi_{12}^*(\ldots)) .
\]
8. Proof of the extremum principle for the control $u_{12}(s,t)$.

We structure this proof so that it shows the significance of the condition that $f_0, f_1, f_2$ are independent of $u_{12}$. Thus, we start without that condition, and we proceed up to a certain point where that condition needs to be introduced. We make use of some ideas from [GK]; we denote that [GK] treats optimal control of ordinary differential equations, and there is no straightforward way to extend those methods to Goursat-Volterra systems.

When all other controls $u_1(s), u_2(t), u_1(A), u_2(B), u_{12}(s,B), u_{12}(A,t), u_{12}(A,B)$ are set to optimal values and only the control $u_{12}(s,t), 0 < s < A, 0 < t < B$ is to be determined, we rewrite the state dynamics, the cost functional, and the Hamiltonian equations in relevant form, i.e. we suppress the dependence on controls that have been set to optimal values, and we show only the dependence on the control $u_{12}(s,t)$. In this section, we shall also denote $u_{12}(s,t)$ by simply $u(s,t)$. Thus we have the state dynamics

$$y(s,t) = g_0(s,t,u(s,t)) + \int_0^s g_1(s,t,\sigma,y(\sigma,t),u(\sigma,t))d\sigma +$$

$$\quad + \int_0^t g_2(s,t,\tau,y(\tau,t),u(\tau,t))d\tau + \int_0^s \int_0^t g_{12}(s,t,\sigma,\tau,y(\sigma,\tau),u(\sigma,\tau))d\tau d\sigma$$

--- (8.1)

The cost functional is now expressed as

$$J := G_0(y(A,B)) + \int_0^A G_1(s,y(s,B))ds + \int_0^B G_2(t,y(A,t))dt +$$

$$\quad + \int_0^A \int_0^B G_{12}(s,t,y(s,t),u(s,t))dt ds$$

--- (8.2)

The Hamiltonian equations are
\[
\psi_{12}(s, t) = \nabla_y G_{12}(s, t, y(s, t), u(s, t)) + \psi_1(s)\nabla_y g_2(s, B, t, y(s, t), u(s, t)) + \\
+ \psi_2(t)\nabla_y g_1(A, t, s, y(s, t), u(s, t)) + \psi_0\nabla_y g_{12}(A, B, s, t, y(s, t), u(s, t)) + \\
\int_A^B \int_s^t \left[ \psi_{12}(\sigma, t)\nabla_y g_1(\sigma, t, s, y(s, t), u(s, t)) + \psi_1(\sigma)\nabla_y g_{12}(\sigma, B, s, t, y(s, t), u(s, t)) \right] d\sigma + \\
\int_B^t \int_s^t \left[ \psi_{12}(s, \tau)\nabla_y g_2(s, \tau, t, y(s, t), u(s, t)) + \psi_2(\tau)\nabla_y g_{12}(A, \tau, s, t, y(s, t), u(s, t)) \right] d\tau + \\
\int_s^t \int_A^B \psi_{12}(\sigma, \tau)\nabla_y g_{12}(\sigma, \tau, s, t, y(s, t), u(s, t)) d\tau d\sigma; \\
\psi_1(s) = \nabla_y G_1(s, y(s, B)) + \psi_0\nabla_y g_1(A, B, s, y(s, B), u(s, B)) + \\
\int_A^B \psi_1(\sigma)\nabla_y g_1(\sigma, B, s, y(s, B), u(s, B)) d\sigma; \\
\psi_2(t) = \nabla_y G_2(t, y(A, t)) + \psi_0\nabla_y g_2(A, B, t, y(A, t), u(A, t)) + \\
\int_B^t \psi_2(\tau)\nabla_y g_2(A, \tau, t, y(A, t), u(A, t)) d\tau; \\
\psi_0 = \nabla_y G_0(y(A, B)) \\
\text{--- (8.3)}
\]

The corresponding Hamiltonians are denoted, in this section, by \( \eta_0, \eta_1, \eta_2, \eta_{12} \), thus
\[ \eta_0 = G_0(y(A,B)); \]
\[ \eta_1(s,y,\psi_0,\psi_1(\cdot)) := G_1(s,y) + \psi_0 g_1(A,B,s,y,u^*(s,B)) + \int_0^A \psi_1(\sigma)g_1(\sigma,B,s,y,u^*(s,B))d\sigma; \]
\[ \eta_2(t,y,\psi_0,\psi_2(\cdot)) := G_2(t,y) + \psi_0 g_2(A,B,t,y,u^*(A,t)) + \int_t^B \psi_2(\tau)g_2(\tau,B,s,y,u^*(A,t))d\tau; \]
\[ \eta_{12}(s,t,y,\psi_0,\psi_1(\cdot),\psi_2(\cdot),\psi_{12}(\cdot),u) := \]
\[ = G_{12}(s,t,y,u) + \psi_1(s)g_2(s,B,t,y,u) + \psi_2(t)g_1(A,t,s,y,u) + \psi_0 g_{12}(A,B,s,t,y,u) + \]
\[ + \int_0^A \left[ \psi_{12}(\sigma,t)g_1(\sigma,t,s,y,u) + \psi_1(\sigma)g_{12}(\sigma,B,s,t,y,u) \right]d\sigma + \]
\[ + \int_t^B \left[ \psi_{12}(\sigma,\tau)g_2(\sigma,\tau,t,y,u) + \psi_2(\tau)g_{12}(A,\tau,s,t,y,u) \right]d\tau + \]
\[ + \int_s^t \int_0^\sigma \psi_{12}(\sigma,\tau)g_{12}(\sigma,\tau,s,t,y,u)dtd\sigma; \]

--- (8.4)

The proof consists of the following parts:

**Part I:** properties of the finite variations of the state, the cost functional, and the Hamiltonian;

**Part II:** construction of "thin" variations of the control;

**Part III:** construction of multi-exponential functions, and proof of a two-dimensional version of the Gronwall inequality;

**Part IV:** passing to the limit and completion of the proof of the extremum principle.

**Part I.**

We assume adequate differentiability of all functions involved.

We denote by \((u, y), (u^-, y^-)\) two pairs of state and control.

We show explicitly the state and the control in the functional \(J\). Thus, for any continuous function \(z(\cdot, \cdot)\), and any piecewise continuous function \(v(\cdot, \cdot)\), where \(z(\cdot, \cdot)\) is not necessarily solution of the state dynamics with control \(v(\cdot, \cdot)\), we set

\[ J(z, v) := G_0(z(A,B)) + \int_0^A G_1(s,z(s,B))ds + \int_0^B G_2(t,z(A,t))dt + \]
\[ + \int_0^A \int_0^B G_{12}(s,t,z(s,t),v(s,t))dt ds \]
We conventionally designate the pair \((y, u)\) as the reference pair, and the pair \((\tilde{y}, \tilde{u})\) as the comparison pair.

The finite variation of the functional \(J\) is denoted by \(\Delta J(y, u)\) and is defined by

\[
\Delta J(y, u) := J(\tilde{y}, \tilde{u}) - J(y, u)
\]

The partial finite variations are defined by

\[
\Delta_y J(y, u) := J(\tilde{y}, u) - J(y, u); \quad \Delta_u J(y, u) := J(y, \tilde{u}) - J(y, u)
\]

The same notation applies to other functions or functionals. For example, for the Hamiltonian \(\eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u)\), we have

\[
\begin{align*}
\Delta \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u) := \\
= \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u) - \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u);
\end{align*}
\]

\[
\begin{align*}
\Delta_y \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u) := \\
= \eta_{12}(s, t, \tilde{y}, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u) - \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u);
\end{align*}
\]

\[
\begin{align*}
\Delta_u \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u) := \\
= \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), \tilde{u}) - \eta_{12}(s, t, y, \psi_0, \psi_1(,), \psi_2(,), \psi_{12}(,,), u).
\end{align*}
\]

We shall use the following formula for the variation of a differentiable (with respect to \(y\)) function \(\varphi(s, t, y(s, t), u(s, t))\):

\[
\begin{align*}
\Delta \varphi(s, t, y(s, t), u(s, t)) &= \Delta_u \varphi(s, t, y(s, t), u(s, t)) + \\
+ [\nabla_y \varphi(s, t, y(s, t), u(s, t)) + \nabla_y \Delta_u \varphi(s, t, y(s, t), u(s, t))] \Delta y(s, t) + o_p(||\Delta y(s, t)\||)
\end{align*}
\]

This is proved as follows:
\[ \Delta \varphi(s, t, y(s, t), u(s, t)) = \varphi(s, t, y^- (s, t), u^- (s, t)) - \varphi(s, t, y(s, t), u(s, t)) = \\
\quad = \varphi(s, t, y^- (s, t), u^- (s, t)) - \varphi(s, t, y^- (s, t), u(s, t)) + \varphi(s, t, y^- (s, t), u(s, t)) - \\
\quad - \varphi(s, t, y(s, t), u(s, t)) = \\
\quad = \Delta_u - \varphi(s, t, y^- (s, t), u(s, t)) + \Delta_y - \varphi(s, t, y(s, t), u(s, t)) = \\
\quad = \Delta_u - \varphi(s, t, y(s, t), u(s, t)) + \nabla_y \Delta_u - \varphi(s, t, y(s, t), u(s, t)) \Delta y(s, t) + o_{\varphi, 1}(|| \Delta y(s, t) ||) + \\
\quad + \nabla_y \varphi(s, t, y(s, t), u(s, t)) \Delta y(s, t) + o_{\varphi, 2}(|| \Delta y(s, t) ||) \\
\quad \text{--- (8.10)} \]

thus we obtain (8.9) with \( o_{\varphi}(|| \Delta y(s, t) ||) = o_{\varphi, 1}(|| \Delta y(s, t) ||) + o_{\varphi, 2}(|| \Delta y(s, t) ||) \). (The symbols \( o_j(\xi) \), for different expressions substituted in lieu of the subscript \( j \), signify different quantities that are \( o(\xi) \) as \( \xi \to 0^+ \), where \( o \) stands for Landau's little-oh symbol.)

Now, we claim that the finite variation of the functional \( J \) can be represented as follows:

\[ \Delta J(y, u) = \int_A^B \int_0^1 \{ \Delta_u \eta_{12}(s, t, y(s, t), \psi_0, \psi_1, \psi_2, \ldots, u(s, t)) + \\
\quad + \nabla_y \Delta_u \eta_{12}(s, t, y(s, t), \psi_0, \psi_1, \psi_2, \ldots, u(s, t)) \Delta y(s, t) + o_{1,12}(|| \Delta y(s, t) ||) \} \, dt \, ds + \\
\quad + \int_0^B o_{1,1}(|| \Delta y(s, B) ||) \, ds + \int_0^A o_{1,2}(|| \Delta y(A, t) ||) \, dt + o_{0,0}(|| \Delta y(A, B) ||) \\
\quad \text{--- (8.11)} \]

To prove (8.11), we first calculate, in accordance with (8.9),

\[ \Delta J(y, u) = \nabla y G_0(y(A, B)) \Delta y(A, B) + o_{1,0,1}(|| \Delta y(A, B) ||) + \\
\quad + \int_0^A \{ \nabla y G_1(s, y(s, B)) \Delta y(s, B) + o_{1,1,1}(|| \Delta y(s, B) ||) \} \, ds + \\
\quad + \int_0^B \{ \nabla y G_2(t, y(A, t)) \Delta y(A, t) + o_{1,2,1}(|| \Delta y(A, t) ||) \} \, dt + \\
\quad + \int_0^A \int_0^B \{ \Delta_u \eta_{12}(s, t, y(s, t), u(s, t)) + \nabla y \eta_{12}(s, t, y(s, t), u(s, t)) + \\
\quad + \nabla_y \Delta_u \eta_{12}(s, t, y(s, t), u(s, t)) \Delta y(s, t) + o_{1,2,1}(|| \Delta y(s, t) ||) \} \, dt \, ds \]
The quantities $\Delta y(s,t), \Delta y(s,B), \Delta y(A,t), \Delta y(A,B)$ are expressed, on the basis of (8.1) and (8.9), as follows:

\[
\Delta y(s,t) = \Delta_u - g_0(s, t, u(s,t)) + \int_0^s \{\Delta_u - g_1(s, t, \sigma, y(\sigma, t), u(\sigma, t)) + \\
+ [V_y g_1(s, t, \sigma, y(\sigma, t), u(\sigma, t)) + V_y \Delta_u - g_1(s, t, \sigma, y(\sigma, t), u(\sigma, t))]\Delta y(\sigma, t) + \\
+ o_{y,12,1}(||\Delta y(\sigma, t)||)\} d\sigma + \\
+ \int_0^t \{\Delta_u - g_2(s, t, \tau, y(s, \tau), u(s, \tau)) + [V_y g_2(s, t, \tau, y(s, \tau), u(s, \tau)) + \\
+ V_y \Delta_u - g_2(s, t, \tau, y(s, \tau), u(s, \tau))]\Delta y(s, \tau) + o_{y,12,2}(||\Delta y(s, \tau)||)\} d\tau + \\
+ \int_0^s \int_0^t \{\Delta_u - g_{12}(s, t, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau)) + [V_y g_{12}(s, t, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau)) + \\
+ V_y \Delta_u - g_{12}(s, t, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau))]\Delta y(\sigma, \tau) + o_{y,12,12}(||\Delta y(\sigma, \tau)||)\} d\tau d\sigma
\]

\[--- (8.13)\]

\[
\Delta y(s,B) = \int_0^s \{V g_1(s, B, y(\sigma, B), u^*(\sigma, B)) + o_{y,1,1}(||\Delta y(\sigma, B)||)\} d\sigma + \\
+ \int_0^B \{\Delta_u - g_2(s, B, \tau, y(s, \tau), u(s, \tau)) + [V_y g_2(s, B, \tau, y(s, \tau), u(s, \tau)) + \\
+ V_y \Delta_u - g_2(s, B, \tau, y(s, \tau), u(s, \tau))]\Delta y(s, \tau) + o_{y,1,2}(||\Delta y(s, \tau)||)\} d\tau + \\
+ \int_0^s \int_0^B \{\Delta_u - g_{12}(s, B, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau)) + [V_y g_{12}(s, B, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau)) + \\
+ V_y \Delta_u - g_{12}(s, B, \sigma, \tau, y(\sigma, \tau), u(\sigma, \tau))]\Delta y(\sigma, \tau) + o_{y,1,12}(||\Delta y(\sigma, \tau)||)\} d\tau d\sigma
\]

\[--- (8.14)\]
On the basis of formulae (8.13) through (8.16), we evaluate the expression

$$E := \int_{A}^{B} \psi_{1}(s,t)\Delta y(s,t)dt\,ds + \int_{0}^{A} \psi_{2}(t)\Delta y(A,t)dt + \psi_{0}\Delta y(A,B)$$

--- (8.17)

The quantity $E$ plays a role similar to that of $\int_{t_{0}}^{t_{1}} \psi(t)\Delta x(t)dt$ in the theory of Pontryagin's maximum principle for controlled ordinary differential equations, as developed in [GK]; $E$ is also related, but not exactly analogous, to the energy of a mechanical system that is described by the equations of Hamiltonian dynamics. The properties of $E$ are established next.
We have

\[
\int_0^B \int_A^{B_t} \psi_{12}(s,t) \Delta y(s,t) \, dt \, ds = \int_0^B \int_A^{B_s} \psi_{12}(s,t) \Delta_{u - g_0}(s,t,u(s,t)) \, ds \, dt + \int_0^B \int_A^{B_s} \psi_{12}(s,t) \Delta_{u - g_1}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \, ds \, dt + \left[ \nabla_y g_1(s,t,\delta, y(\sigma, t), u(\sigma, t)) + \nabla_y \Delta_{u - g_1}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \right] \Delta y(\sigma, t) + \mathbf{o}_{y,12,1}(\| \Delta y(\sigma, t) \|) \, d\sigma \, dt \, ds + \int_1^B \int_A^{B_t} \psi_{12}(s,t) \Delta_{u - g_2}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \, ds \, dt + \int_1^B \int_A^{B_t} \psi_{12}(s,t) \Delta_{u - g_2}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \, ds \, dt + \left[ \nabla_y g_2(s,t,\delta, y(\sigma, t), u(\sigma, t)) + \nabla_y \Delta_{u - g_2}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \right] \Delta y(\sigma, t) + \mathbf{o}_{y,12,2}(\| \Delta y(\sigma, t) \|) \, dt \, d\sigma \, dt \, ds \quad \text{(8.18)}
\]

After some changes in the order of integration, we get

\[
\int_0^B \int_A^{B_t} \psi_{12}(s,t) \Delta y(s,t) \, dt \, ds = \int_0^B \int_A^{B_s} \psi_{12}(s,t) \Delta_{u - g_0}(s,t,u(s,t)) \, ds \, dt + \int_0^B \int_A^{B_s} \psi_{12}(s,t) \Delta_{u - g_1}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \, ds \, dt + \left[ \nabla_y g_1(s,t,\delta, y(\sigma, t), u(\sigma, t)) + \nabla_y \Delta_{u - g_1}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \right] \Delta y(\sigma, t) + \mathbf{o}_{y,12,1}(\| \Delta y(\sigma, t) \|) \, d\sigma \, dt \, ds + \int_1^B \int_A^{B_t} \psi_{12}(s,t) \Delta_{u - g_2}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \, ds \, dt + \int_1^B \int_A^{B_t} \psi_{12}(s,t) \Delta_{u - g_2}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \, ds \, dt + \left[ \nabla_y g_2(s,t,\delta, y(\sigma, t), u(\sigma, t)) + \nabla_y \Delta_{u - g_2}(s,t,\delta, y(\sigma, t), u(\sigma, t)) \right] \Delta y(\sigma, t) + \mathbf{o}_{y,12,2}(\| \Delta y(\sigma, t) \|) \, dt \, d\sigma \, dt \, ds \quad \text{(8.19)}
\]
Similarly, we find

\[
\begin{align*}
&\int_0^s \psi_1(s) \Delta y(s, B) \, ds = \int_0^s \int_0^s \psi_1(s) \{ \nabla g_1(s, B, y(\sigma, B), u^*(\sigma, B)) + o_{y,1,1}(\| \Delta y(\sigma, B) \|) \} d\sigma \, ds + \\
&+ \int_0^s \int_0^s \psi_1(s) \{ \Delta u_2 \cdot g_2(s, B, \tau, y(s, \tau), u(s, \tau)) + [\nabla y g_2(s, B, \tau, y(s, \tau), u(s, \tau)) + \\
&+ \nabla y \Delta u_2 \cdot g_2(s, B, \tau, y(s, \tau), u(s, \tau))] \Delta y(s, \tau) + o_{y,1,2}(\| \Delta y(s, \tau) \|) \} d\tau \, ds + \\
&= \int_0^s \int_0^s \psi_1(\sigma) \{ \nabla g_1(\sigma, B, y(s, B), u^*(s, B)) + o_{y,1,1}(\| \Delta y(s, B) \|) \} d\sigma \, ds + \\
&+ \int_0^s \int_0^s \psi_1(s) \{ \Delta u_2 \cdot g_2(s, B, t, y(s, t), u(s, t)) + [\nabla y g_2(s, B, t, y(s, t), u(s, t)) + \\
&+ \nabla y \Delta u_2 \cdot g_2(s, B, t, y(s, t), u(s, t))] \Delta y(s, t) + o_{y,1,2}(\| \Delta y(s, t) \|) \} dt \, ds + \\
&+ \int_0^s \int_0^s \psi_1(\sigma) \{ \Delta u_2 \cdot g_1(\sigma, B, s, t, y(s, t), u(s, t)) + [\nabla y g_1(\sigma, B, s, t, y(s, t), u(s, t)) + \\
&+ \nabla y \Delta u_2 \cdot g_1(\sigma, B, s, t, y(s, t), u(s, t))] \Delta y(s, t) + o_{y,1,12}(\| \Delta y(s, t) \|) \} d\sigma \, dt \, ds
\end{align*}
\]

--- (8.20)
B \int \psi_2(t)\Delta y(A, t)\, dt = \int_0^B \int_A^B \psi_2(t)\{|\Delta_u g_1(A, t, y(\sigma, t), u(\sigma, t)) + \\
+ [\nabla y g_1(A, t, \sigma, y(\sigma, t), u(\sigma, t)) + \nabla y \Delta_u g_1(A, t, \sigma, y(\sigma, t), u(\sigma, t))|\Delta y(\sigma, t) + \\
+ o_{y,2,1} (|| \Delta y(\sigma, t) ||)|\} \, d\sigma \, dt + \\
+ \int_0^B \int_A^B \psi_2(t)\{|\Delta u g_1(A, t, \sigma, y(\sigma, t), u(\sigma, t)) + [\nabla y g_1(A, t, \sigma, y(\sigma, t), u(\sigma, t)) + \\
+ \nabla y \Delta u g_1(A, t, \sigma, y(\sigma, t), u(\sigma, t))|\Delta y(\sigma, t) + \\
+ o_{y,2,1} (|| \Delta y(\sigma, t) ||)|\} \, d\sigma \, dt = \\
= \int_0^B \int_A^B \psi_2(t)\{|\Delta u g_1(A, t, s, y(s, t), u(s, t)) + \\
+ [\nabla y g_1(A, t, s, y(s, t), u(s, t)) + \nabla y \Delta u g_1(A, t, s, y(s, t), u(s, t))|\Delta y(s, t) + \\
+ o_{y,2,1} (|| \Delta y(s, t) ||)|\} \, ds \, dt + \\
+ \int_0^B \int_A^B \psi_2(t)\{|\Delta u g_1(A, t, s, y(s, t), u(s, t)) + [\nabla y g_1(A, t, s, y(s, t), u(s, t)) + \\
+ \nabla y \Delta u g_1(A, t, s, y(s, t), u(s, t))|\Delta y(s, t) + \\
+ o_{y,2,1} (|| \Delta y(s, t) ||)|\} \, ds \, dt \\
--- (8.21)

\psi_0 \Delta y(A, B) = \int_0^B \psi_0 \{|\nabla y g_1(A, B, s, y(s, B), u^*(s, B))|\Delta y(s, B) + o_{y,0,1} (|| \Delta y(s, B) ||)|\} \, ds + \\
+ \int_0^B \psi_0 \{|\nabla y g_2(A, B, s, y(A, t), u^*(A, t))|\Delta y(A, t) + o_{y,0,2} (|| \Delta y(A, t) ||)|\} \, dt + \\
+ \int_0^B \int_A^B \psi_0 \{|\Delta u g_1(A, B, s, t, y(s, t), u(s, t)) + [\nabla y g_1(A, B, s, t, y(s, t), u(s, t)) + \\
+ \nabla y \Delta u g_1(A, B, s, t, y(s, t), u(s, t))|\Delta y(s, t) + \\
+ o_{y,0,1} (|| \Delta y(s, t) ||)|\} \, ds \, dt \\
--- (8.22)
We now use the expression (8.12) for $\Delta J$ and the Hamiltonian equations (8.3) and (8.4). Then we can verify that

$$E + \Delta J = E + \int_{0}^{A} \int_{0}^{B} \left\{ \Delta_{u} - \eta_{12}(s, t, y(s, t), \psi_{0}, \psi_{1}(.), \psi_{2}(.), \psi_{12}(.), u(s, t)) + \right.$$  

$$+ \nabla_{y} \Delta_{u} - \eta_{12}(s, t, y(s, t), \psi_{0}, \psi_{1}(.), \psi_{2}(.), \psi_{12}(.), u(s, t)) \Delta y(s, t) + \mathbf{o}_{J,12}(\| \Delta y(s, t) \|) \right\} dt ds +$$  

$$+ \int_{0}^{A} \mathbf{o}_{J,1}(\| \Delta y(s, B) \|) ds + \int_{0}^{B} \mathbf{o}_{J,2}(\| \Delta y(A, t) \|) dt + \mathbf{o}_{J,0}(\| \Delta y(A, B) \|)$$

--- (8.23)

which is tantamount to (8.11).

Indeed,
\[ E + \Delta J = \int \int_{A} \int_{B} \psi_{12} (s, t) \Delta u - g_{0} (s, t, u(s, t)) \, ds \, dt + \]
\[ = \int \int_{A} \int_{B} \int_{A} \psi_{12} (\sigma, t) \{ \Delta u - g_{1} (\sigma, t, s, y(s, t), u(s, t)) + + [\nabla_{y} g_{1} (\sigma, t, s, y(s, t), u(s, t)) + \nabla_{y} \Delta u - g_{1} (\sigma, t, s, y(s, t), u(s, t))] \Delta y(s, t) + + o_{y,1,2,1} (|| \Delta y(s, t) ||) \} \, d\sigma \, dt \, ds + \]
\[ + \int \int_{A} \int_{B} \int_{A} \psi_{12} (s, \tau) \{ \Delta u - g_{2} (s, \tau, t, y(s, t), u(s, t)) + + [\nabla_{y} g_{2} (s, \tau, t, y(s, t), u(s, t)) + + \nabla_{y} \Delta u - g_{2} (s, \tau, t, y(s, t), u(s, t))] \Delta y(s, t) + + o_{y,2,1,2} (|| \Delta y(s, t) ||) \} \, d\tau \, dt \, ds + \]
\[ + \int \int_{A} \int_{B} \int_{A} \psi_{12} (\sigma, \tau) \{ \Delta u - g_{12} (\sigma, \tau, s, t, y(s, t), u(s, t)) + + [\nabla_{y} g_{12} (\sigma, \tau, s, t, y(s, t), u(s, t)) + + \nabla_{y} \Delta u - g_{12} (\sigma, \tau, s, t, y(s, t), u(s, t))] \Delta y(s, t) + + o_{y,1,2,12} (|| \Delta y(s, t) ||) \} \, d\sigma \, dt \, ds + \]
\[ + \int \int_{A} \int_{B} \int_{A} \psi_{1} (\sigma) \{ \nabla_{y} g_{1} (\sigma, B, y(s, B), u^{*} (s, B)) + + o_{y,1,1} (|| \Delta y(s, B) ||) \} \, d\sigma \, ds + \]
\[ + \int \int_{A} \int_{B} \int_{A} \psi_{1} (s) \{ \Delta u - g_{2} (s, B, t, y(s, t), u(s, t)) + + [\nabla_{y} g_{2} (s, B, t, y(s, t), u(s, t)) + + \nabla_{y} \Delta u - g_{2} (s, B, t, y(s, t), u(s, t))] \Delta y(s, t) + + o_{y,1,2} (|| \Delta y(s, t) ||) \} \, dt \, ds + \]
\[ + \int \int_{A} \int_{B} \int_{A} \psi_{1} (\sigma) \{ \Delta u - g_{12} (\sigma, B, s, t, y(s, t), u(s, t)) + + [\nabla_{y} g_{12} (\sigma, B, s, t, y(s, t), u(s, t)) + + \nabla_{y} \Delta u - g_{12} (\sigma, B, s, t, y(s, t), u(s, t))] \Delta y(s, t) + + o_{y,1,12} (|| \Delta y(s, t) ||) \} \, d\sigma \, ds + \]
\[ + \int \int_{A} \int_{B} \int_{A} \psi_{2} (t) \{ \Delta u - g_{1} (A, t, s, y(s, t), u(s, t)) + + [\nabla_{y} g_{1} (A, t, s, y(s, t), u(s, t)) + + \nabla_{y} \Delta u - g_{1} (A, t, s, y(s, t), u(s, t))] \Delta y(s, t) + + o_{y,2,1} (|| \Delta y(s, t) ||) \} \, dt \, ds + \]
\[ + \int \int_{B} \int_{A} \int_{B} \int_{A} \psi_{2} (\tau) \{ \nabla_{y} g_{2} (A, \tau, t, y(A, t), u^{*} (A, t)) \Delta y(A, t) + + o_{y,2,2} (|| \Delta y(A, t) ||) \} \, d\tau \, dt + \]
\[ + \int \int_{B} \int_{A} \int_{B} \int_{A} \psi_{2} (\tau) \{ \Delta u - g_{12} (A, \tau, s, t, y(s, t), u(s, t)) + + [\nabla_{y} g_{12} (A, \tau, s, t, y(s, t), u(s, t)) + + \nabla_{y} \Delta u - g_{12} (A, \tau, s, t, y(s, t), u(s, t))] \Delta y(s, t) + + o_{y,2,12} (|| \Delta y(s, t) ||) \} \, d\tau \, ds \, dt + \]
\[\int_0^B \psi_0 \left\{ \nabla_y g_1(A, B, s, y(s, B), u^*(s, B)) \Delta y(s, B) + o_{y,0,1}(\| \Delta y(s, B) \|) \right\} ds + \]

\[\int_0^A \int_0^B \psi_0 \left\{ \Delta_u - g_{12}(A, B, s, t, y(s, t), u(s, t)) + [\nabla_y g_{12}(A, B, s, t, y(s, t), u(s, t)) + \nabla_y \Delta_u - g_{12}(A, B, s, t, y(s, t), u(s, t))] \Delta y(s, t) + o_{y,0,12}(\| \Delta y(s, t) \|) \right\} dt ds +\]

\[\nabla_y G_0(y(A, B)) \Delta y(A, B) + o_{J,0,1}(\| \Delta y(A, B) \|) + \]

\[\int_0^A \{ \nabla_y G_1(s, y(s, B)) \Delta y(s, B) + o_{J,1,1}(\| \Delta y(s, B) \|) \} ds + \]

\[\int_0^B \{ \nabla_y G_2(t, y(A, t)) \Delta y(A, t) + o_{J,2,1}(\| \Delta y(A, t) \|) \} dt + \]

\[\int_0^A \int_0^B \{ \Delta_u - G_{12}(s, t, y(s, t), u(s, t)) + [\nabla_y G_{12}(s, t, y(s, t), u(s, t)) + \nabla_y \Delta_u - G_{12}(s, t, y(s, t), u(s, t))] \Delta y(s, t) + o_{J,12,1}(\| \Delta y(s, t) \|) \} dt ds \]

--- (8.24)

Because of the Hamiltonian equations (8.3), some of the terms on the right-hand side of (8.24) make up E. This is seen from the following equations:
\[
\int_0^A \int_0^B \nabla_y G_{12}(s, t, y(s, t), u(s, t)) \Delta y(s, t) \, ds \, ds + \\
\int_0^A \int_0^B \nabla_y g_2(s, B, t, y(s, t), u(s, t)) + \psi_2(t) \nabla_y g_1(A, t, s, y(s, t), u(s, t)) + \\
\psi_0 \nabla_y g_{12}(A, B, s, t, y(s, t), u(s, t)) \Delta y(s, t) \, dt \, ds + \\
\int_0^A \int_0^B \int_0^s \left[ \psi_1(s) \nabla_y g_2(s, B, t, y(s, t), u(s, t)) + \psi_2(t) \nabla_y g_1(A, t, s, y(s, t), u(s, t)) + \\
\psi_0 \nabla_y g_{12}(A, B, s, t, y(s, t), u(s, t)) \right] \Delta y(s, t) \, dt \, ds + \\
\int_0^A \int_0^B \int_0^t \left[ \psi_1(s) \nabla_y g_2(s, \tau, t, y(s, t), u(s, t)) + \psi_2(\tau) \nabla_y g_{12}(A, \tau, s, t, y(s, t), u(s, t)) \right] \Delta y(s, t) \, dt \, ds + \\
\int_0^A \int_0^B \int_0^s \int_0^t \psi_1(\sigma, \tau) \nabla_y g_{12}(\sigma, \tau, s, t, y(s, t), u(s, t)) \Delta y(s, t) \, dt \, ds \, ds = \\
\int_0^A \int_0^B \psi_1(s) \Delta y(s, t) \, ds \\
--- (8.25)
\]

\[
\int_0^A \left[ \nabla_y G_1(s, y(s, B)) + \psi_0 \nabla_y g_1(A, B, s, y(s, B), u(s, B)) \right] \Delta y(s, B) \, ds + \\
\int_0^A \int_0^A \psi_1(\sigma) \nabla_y g_1(\sigma, B, s, y(s, B), u(s, B)) \Delta y(s, B) \, d\sigma \, ds = \\
\int_0^A \psi_1(s) \Delta y(s, B) \, ds \\
--- (8.26)
\]

\[
\int_0^B \left[ \nabla_y G_2(t, y(A, t)) + \psi_0 \nabla_y g_2(A, B, t, y(A, t), u(A, t)) \right] \Delta y(A, t) \, dt + \\
\int_0^B \int_0^B \psi_2(\tau) \nabla_y g_2(A, \tau, t, y(A, t), u(A, t)) \Delta y(A, t) \, d\tau \, dt = \\
\int_0^B \psi_2(t) \Delta y(A, t) \, dt \\
--- (8.27)
\]
\( V_y G_0 (y(A, B)) \Delta y(A, B) = \psi_0 \Delta y(A, B) \)  
\[\text{--- (8.28)}\]

The terms on the right-hand sides of (8.25) through (8.28) make up \( E \). On the other hand, the terms on the right-hand side of (8.24), other than those that appear on the left-hand sides of (8.25) through (8.28), and other than those that are \( o(\| \Delta y \|) \), make up

\[
\int \int \{ \Delta_u - \eta_{12} (s, t, y(s, t), \psi_0, \psi_1, \psi_2, \psi_{12}, u(s, t)) + \\
+ V_y \Delta_u - \eta_{12} (s, t, y(s, t), \psi_0, \psi_1, \psi_2, \psi_{12}, u(s, t)) \Delta y(s, t) \} dt ds .
\]

This proves (8.23), and thus also (8.11). This concludes Part I of the proof of the extremum principle for the control \( u_{12} \).

Part II.

If \( u_{12}^* \) is an optimal control, we denote by \( y^* \) the corresponding state function, and by \( (\psi_0^*, \psi_1^*, \psi_2^*, \psi_{12}^*) \) the corresponding solution of the Hamiltonian equations. If \( (s_1, t_1) \in (0, A) \times (0, B) \) is a point of continuity of \( u_{12}^* \), if \( \hat{u} \) is an admissible value of the control \( u_{12} \), we define a control \( \tilde{u}_{\varepsilon \delta} \), for every \( \varepsilon \) and \( \delta \) positive and sufficiently small, by

\[
\tilde{u}_{\varepsilon \delta} (s, t) = \hat{u} , \text{ if } (s, t) \in (s_1, s_1 + \delta) \times (t_1, t_1 + \varepsilon) ; \\
\tilde{u}_{\varepsilon \delta} (s, t) = u_{12}^* (s, t) , \text{ if } (s, t) \notin (s_1, s_1 + \delta) \times (t_1, t_1 + \varepsilon)
\]
\[\text{--- (8.29)}\]

We set
We denote by \( \tilde{y}_{\varepsilon \delta} \) the solution of the state dynamics (8.1) with control \( \tilde{u}_{\varepsilon \delta} \). Then \( \tilde{y}_{\varepsilon \delta} \) solves the following problem:
\[ y_{\delta}(s,t) = y^*(s,t), \text{ for } (s,t) \notin D_1; \]

\[
y_{\delta}(s,t) = g_0(s,t,u(s,t)) + \int_0^{s_1} \left( \int_0^{t_1} g_1(s,t,\sigma, y^*(\sigma,t), u^*(\sigma,t)) d\sigma \right) + \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \left( \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \left( \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \right) \]

\[
y_{\delta}(s,t) = g_0(s,t,u^*(s,t)) + \int_0^{s_1} \left( \int_0^{t_1} g_1(s,t,\sigma, y^*(\sigma,t), u^*(\sigma,t)) d\sigma \right) + \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \left( \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \int_0^s \left( \int_0^t \int_0^{s_1} g_{12}(s,t,\sigma,\tau, y^*(\sigma,\tau), u^*(\sigma,\tau)) d\sigma d\tau \right) d\sigma + \right) \]

for \( (s,t) \in D_{\delta}^1 \);
\[ y_{\varepsilon \delta}(s, t) = g_0(s, t, u(s, t)) + \int_0^{s_1} g_1(s, t, \sigma, y^*(\sigma, t), u^*(\sigma, t))d\sigma + \]
\[ + \int_0^{t_1} g_2(s, t, \tau, y^*(\tau), u^*(\tau))d\tau + \int_0^{s_1} \int_0^{t_1} g_{12}(s, t, \sigma, \tau, y^{\prime}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} g_{12}(s, t, \sigma, y_{\varepsilon \delta}(\sigma, t), u^*(\sigma, t))d\sigma + \int_{s_1}^{t_{1+\varepsilon}} g_2(s, t, \tau, y_{\varepsilon \delta}(\sigma, \tau), u^*(\sigma, \tau))d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, \tau, y_{\varepsilon \delta}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, \tau, y_{\varepsilon \delta}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau + \]
\[ y_{\varepsilon \delta}(s, t) = g_0(s, t, u^*(s, t)) + \int_0^{s_1} g_1(s, t, \sigma, y^*(\sigma, t), u^*(\sigma, t))d\sigma + \]
\[ + \int_0^{t_1} g_2(s, t, \tau, y^*(\tau), u^*(\tau))d\tau + \int_0^{s_1} \int_0^{t_1} g_{12}(s, t, \sigma, \tau, y^{\prime}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, y_{\varepsilon \delta}(\sigma, t), u^*(\sigma, t))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, \tau, y_{\varepsilon \delta}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, \tau, y_{\varepsilon \delta}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, \tau, y_{\varepsilon \delta}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau , \text{ for } (s, t) \in D_{\varepsilon \delta}^2 ; \]
\[ y_{\varepsilon \delta}(s, t) = g_0(s, t, u^*(s, t)) + \int_0^{s_1} g_1(s, t, \sigma, y^*(\sigma, t), u^*(\sigma, t))d\sigma + \]
\[ + \int_0^{t_1} g_2(s, t, \tau, y^*(\tau), u^*(\tau))d\tau + \int_0^{s_1} \int_0^{t_1} g_{12}(s, t, \sigma, \tau, y^{\prime}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, y_{\varepsilon \delta}(\sigma, t), u^*(\sigma, t))d\sigma d\tau + \]
\[ + \int_{s_1}^{t} \int_{s_1}^{t_{1+\varepsilon}} g_{12}(s, t, \sigma, \tau, y_{\varepsilon \delta}(\sigma, \tau), u^*(\sigma, \tau))d\sigma d\tau , \text{ for } (s, t) \in D_{\varepsilon \delta}^3 \]
--- (8.31)
Consequently, for the controls $u_{\tilde{\phi}}$ introduced above, the difference 
\[ \Delta y(s, t) = y_{\tilde{\phi}}(s, t) - y^*(s, t) \] can be non-zero only for $(s, t) \in D_1$. It is thus seen that \( \Delta y(s, t) \) satisfies an equation

\[
\Delta y(s, t) = \Delta_{u_{\tilde{\phi}}} g_0(s, t, u^*(s, t)) + \int_{s_1}^{s} \Delta g_1(s, t, \sigma, y^*(\sigma, t), u^*(\sigma, t)) d\sigma + \\
+ \int_{\tau_1}^{t} \Delta g_2(s, t, \tau, y^*(s, \tau), u^*(s, \tau)) d\tau + \int_{s_1}^{s} \int_{\tau_1}^{t} \Delta g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u^*(\sigma, \tau)) d\tau d\sigma
\]

--- (8.32)

where

\[
\Delta g_1(s, t, \sigma, y^*(\sigma, t), u^*(\sigma, t)) := g_1(s, t, \sigma, y_{\tilde{\phi}}(\sigma, t), u_{\tilde{\phi}}(\sigma, t)) - g_1(s, t, \sigma, y^*(\sigma, t), u^*(\sigma, t)) ;
\]

\[
\Delta g_2(s, t, \tau, y^*(s, \tau), u^*(s, \tau)) := g_2(s, t, \tau, y_{\tilde{\phi}}(s, \tau), u_{\tilde{\phi}}(s, \tau)) - g_2(s, t, \tau, y^*(s, \tau), u^*(s, \tau)) ;
\]

\[
\Delta g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u^*(\sigma, \tau)) :=
= g_{12}(s, t, \sigma, \tau, y_{\tilde{\phi}}(\sigma, \tau), u_{\tilde{\phi}}(\sigma, \tau)) - g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u^*(\sigma, \tau))
\]

--- (8.33)

Each of the finite variations above can be expressed according to the equation

\[
\Delta g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u^*(\sigma, \tau)) =
= \Delta_{y_{\tilde{\phi}}} g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u_{\tilde{\phi}}(\sigma, \tau)) + \Delta_{u_{\tilde{\phi}}} g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u^*(\sigma, \tau))
\]

--- (8.34)

which, of course, also applies to the functions $g_1$ and $g_2$, by replacing \( \{(s, t, \sigma, \tau), (\sigma, \tau)\} \) by \( \{(s, t, \sigma), (\sigma, t)\}, \{(s, t, \tau), (s, \tau)\} \), respectively. Also, we note that

\[ \Delta_{u_{\tilde{\phi}}} g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u^*(\sigma, \tau)) \] can be non-zero only on the set \( D_1^{0} \).

If $B_1, B_2, B_{12}$ are Lipschitz constants for the functions $g_1, g_2, g_{12}$, respectively, with respect to the variable $y$, then we have
\[ \| \Delta_y g_1(s, t, \sigma, y^*(\sigma, t), u_{\tilde{\epsilon}\delta}(\sigma, t)) \| \leq B_1 \| \Delta y(\sigma, t) \| ; \]
\[ \| \Delta_y g_2(s, t, \tau, y^*(\sigma, \tau), u_{\tilde{\epsilon}\delta}(\sigma, \tau)) \| \leq B_2 \| \Delta y(\sigma, \tau) \| ; \]
\[ \| \Delta_y g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u_{\tilde{\epsilon}\delta}(\sigma, \tau)) \| \leq B_{12} \| \Delta y(\sigma, \tau) \| \]
--- (8.35)

Consequently \( \Delta y \) satisfies the integral inequality
\[
\| \Delta y(s, t) \| \leq a(s, t) + B_1 \int_s^t \| \Delta y(\sigma, t) \| d\sigma + B_2 \int_{t_1}^t \| \Delta y(\sigma, \tau) \| d\tau + \\
B_{12} \int_{t_1}^t \| \Delta y(\sigma, \tau) \| d\tau d\sigma ;
\]
\[ a(s, t) := \left[ \Delta_y g_0(s, t, u^*(s, t)) \chi((s, t) \in D_{\epsilon\delta}^0) + \\
+ \int_{s_1}^s \Delta_y g_1(s, t, \sigma, y^*(\sigma, t), u^*(\sigma, t)) d\sigma + \int_{t_1}^{t+\epsilon} \Delta_y g_2(s, t, \tau, y^*(s, \tau), u^*(s, \tau)) d\tau + \\
+ \int_{s_1}^s \int_{t_1}^{t+\epsilon} \Delta_y g_{12}(s, t, \sigma, \tau, y^*(\sigma, \tau), u^*(\sigma, \tau)) d\tau d\sigma \right] \]
--- (8.36)

Here, \( \chi \) is the truth function, i.e. if \( P \) is a logical statement, then \( \chi(P) = 1 \) if \( P \) is true, and \( \chi(P) = 0 \) if \( P \) is false. The symbol \( \land \) denotes the operation of taking the minimum of two real numbers.
This concludes Part II.

Part III.
First we prove an inequality of Gronwall's type. Even though there exist, in the research literature, several extensions of the standard Gronwall inequality, we shall prove directly a result that we need for the present paper.
If \( z(s, t), \varphi_0(s, t), \varphi_1(s, t, \sigma), \varphi_2(s, t, \tau), \varphi_{12}(s, t, \sigma, \tau) \) are piecewise continuous nonnegative functions and \( z \) satisfies
\[ z(s, t) \leq \varphi_0(s, t) + \int_{s_1}^{s} \varphi_1(s, t, \sigma) z(\sigma, t) \, d\sigma + \int_{t_1}^{t} \varphi_2(s, t, \tau) z(s, \tau) \, d\tau + \]
\[ + \int_{s_1}^{s} \int_{t_1}^{t} \varphi_{12}(s, t, \sigma, \tau) z(\sigma, \tau) \, d\sigma \, d\tau \]
--- (8.37)

and if \( \zeta \) is the solution of

\[ \zeta(s, t) = \varphi_0(s, t) + \int_{s_1}^{s} \varphi_1(s, t, \sigma) \zeta(\sigma, t) \, d\sigma + \int_{t_1}^{t} \varphi_2(s, t, \tau) \zeta(s, \tau) \, d\tau + \]
\[ + \int_{s_1}^{s} \int_{t_1}^{t} \varphi_{12}(s, t, \sigma, \tau) \zeta(\sigma, \tau) \, d\sigma \, d\tau \]
--- (8.38)

then

\[ z(s, t) \leq \zeta(s, t) \text{ on } [s_1, A] \times [t_1, B] \]
--- (8.39)

We shall use the convolution notation of section 5. (The fact that we now have integrations with starting point \((s_1, t_1)\), instead of \((0, 0)\) that we had in section 5, does not affect the validity of the results.) We set \( \varphi := (\varphi_1, \varphi_2, \varphi_{12}) \), and we write (8.37) in the form

\[ z \leq \varphi_0 + \varphi \otimes_0 z \]
--- (8.40)

Because of the nonnegativity of all functions involved, we can iterate (8.40) to get

\[ z \leq \varphi_0 + \varphi \otimes_0 (\varphi_0 + \varphi \otimes_0 z) = \varphi_0 + \varphi \otimes_0 \varphi_0 + \varphi \otimes^2 \otimes_0 z \]
--- (8.41)

and, inductively,
\[ z \leq \varphi_0 + \sum_{k=1}^{n} \varphi^{\otimes k} \otimes_0 \varphi_0 + \varphi^{\otimes(n+1)} \otimes_0 z \]  
--- (8.42)

The convergence analysis of section 5 shows that \( \varphi^{\otimes(n+1)} \otimes_0 z \to 0 \) as \( n \to \infty \) in the uniform (sup) norm over \([s_1, A] \times [t_1, B]\). By passing to the limit in (8.42), we obtain

\[ z \leq \varphi_0 + \sum_{k=1}^{\infty} \varphi^{\otimes k} \otimes_0 \varphi_0 \]  
--- (8.43)

Now, the solution of (8.38) is

\[ \zeta = \varphi_0 + \sum_{k=1}^{\infty} \varphi^{\otimes k} \otimes_0 \varphi_0 \]  
--- (8.44)

and consequently (8.43) is tantamount to the wanted inequality (8.40).

Next, we obtain more explicit information about the solution of (8.39) in the case in which all functions \( \varphi_0(s, t), \varphi_1(s, t, \sigma), \varphi_2(s, t, \tau), \varphi_{12}(s, t, \sigma, \tau) \) are constants. Thus we consider the equation

\[ \zeta(s, t) = A + B_1 \int_{s_1}^{s} \zeta(\sigma, t) d\sigma + B_2 \int_{t_1}^{t} \zeta(s, \tau) d\tau + B_{12} \int_{s_1}^{s} \int_{t_1}^{t} \zeta(\sigma, \tau) d\sigma d\tau \]  
--- (8.41)

We seek a solution in the form

\[ \zeta(s, t) = A \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} C_{k\ell} (s-s_1)^k (t-t_1)^\ell \]  
--- (8.42)

By substituting (8.42) into (8.41) and equating the coefficients of the same powers of \((s-s_1), (t-t_1)\), we find
\[ C_{k0} = \frac{B_1^k}{k!}, \quad C_{0\ell} = \frac{B_2^\ell}{\ell!}, \]

\[ C_{k+1,\ell+1} = \frac{B_1}{k+1} C_{k,\ell+1} + \frac{B_2}{\ell+1} C_{k+1,\ell} + \frac{B_{12}}{(k+1)(\ell+1)} C_{k\ell} \quad \text{--- (8.43)} \]

We set

\[ B := \max \left( \left| B_1 \right|, \left| B_2 \right|, \sqrt{\frac{|B_{12}|}{3}} \right) \quad \text{--- (8.44)} \]

We claim that

\[ |\zeta(s, t)| \leq A |\exp(3B(s - s_1)(t - t_1))| \quad \text{--- (8.45)} \]

In order to prove (8.45), it suffices to show that

\[ |C_{k\ell}| \leq \frac{(3B)^{k+\ell}}{(k!)(\ell!)} \quad \text{--- (8.46)} \]

The inequality (8.46) is clearly true for \( k = 0 \) or \( \ell = 0 \). For other values of \((k, \ell)\), we use double induction on \((k, \ell)\), and we invoke (8.43):

\[ |C_{k+1,\ell+1}| \leq \frac{|B_1|}{k+1} |C_{k,\ell+1}| + \frac{|B_2|}{\ell+1} |C_{k+1,\ell}| + \frac{|B_{12}|}{(k+1)(\ell+1)} |C_{k\ell}| \leq \]

\[ \leq \frac{|B_1|}{k+1} \left(3B\right)^{k+\ell+1} \frac{1}{(k!(\ell+1)!)} + \frac{|B_2|}{\ell+1} \left(3B\right)^{k+\ell+1} \frac{1}{((k+1)!(\ell)!)} + \frac{|B_{12}|}{(k+1)(\ell+1)} \left(3B\right)^{k+\ell} \leq \]

\[ \leq 3B \left(3B\right)^{k+\ell+1} \frac{1}{((k+1)!(\ell+1)!)} = \left(3B\right)^{k+\ell+2} \frac{1}{((k+1)!(\ell+1)!)} \quad \text{--- (8.47)} \]

This proves (8.46) (and thus (8.45)), and (8.46) also proves the uniform convergence of the series (8.42).

This concludes Part III.
Part IV.

In the case in which \( g_0, g_1, g_2 \) are independent of \( u \), the inequality (8.36) becomes

\[
\| \Delta y(s, t) \| \leq \int_{s_1}^{s} \int_{t_1}^{t} \Delta \eta_{12}(s, t, \Delta, u, y^*(s, \tau), u^*(\sigma, \tau)) d\tau d\sigma + B_{12} \int_{s_1}^{s} \int_{t_1}^{t} \| \Delta y(\sigma, \tau) \| d\tau d\sigma
\]

--- (8.48)

If \( C_0 \) is a bound on \( g_{12}(s, t, \Delta, y^*(s, \tau), u) \), uniform with respect to \( s, t, \Delta, u \), then the first double integral term in (8.48) is, in absolute value, \( \leq 2C_0\varepsilon \delta \). Consequently, by the results established in part III, we have

\[
\| \Delta y(s, t) \| \leq 2C_0\varepsilon \delta \exp(3M\varepsilon \delta) = O_1(\varepsilon \delta)
\]

--- (8.49)

And similarly

\[
\| \Delta y(A, t) \| = O_2(\varepsilon \delta), \| \Delta y(s, B) \| = O_3(\varepsilon \delta), \| \Delta y(A, B) \| = O_4(\varepsilon \delta)
\]

--- (8.50)

Now (8.11) becomes

\[
\Delta J(y, u) = \int_{D_0} \left\{ \eta_{12}(s, t, y(s, t), \psi_0, \psi_1, \psi_2, \psi_{12}, \ldots, u(s, t)) + \right.
\]

\[
+ \nabla_y \Delta \eta_{12}(s, t, y(s, t), \psi_0, \psi_1, \psi_2, \psi_{12}, \ldots, u(s, t)) \Delta y(s, t) \right\} dt ds +
\]

\[
+ \int_{D_1} \left\{ o_{J,1}(\| \Delta y(s, t) \|) \right\} dt ds +
\]

\[
+ \int_{s_1}^{A} o_{J,1}(\| \Delta y(s, B) \|) ds + \int_{t_1}^{B} o_{J,2}(\| \Delta y(A, t) \|) dt + o_{J,0}(\| \Delta y(A, B) \|)
\]

--- (8.51)

In view of (8.49) and (8.50), we obtain from (8.51)
\[ \Delta J(y, u) = \int_{D_{\varepsilon \delta}} \{ \Delta_{u_{\varepsilon \delta}} \eta_{12}(s, t, y(s, t), \psi_0, \psi_1, \psi_2, \psi_{12}, u(s, t)) \} dt \, ds + o(\varepsilon \delta) \]  

--- (8.52)

For an optimal control \( u^* \) and the corresponding state trajectory \( y^* \), we have

\[ \Delta J(y^*, u^*) \geq 0, \]

and, consequently,

\[ \int_{D_{\varepsilon \delta}} \{ \Delta_{u_{\varepsilon \delta}} \eta_{12}(s, t, y^*(s, t), \psi_0^*, \psi_1^*, \psi_2^*, \psi_{12}^*, u^*(s, t)) \} dt \, ds + o(\varepsilon \delta) \geq 0 \]

--- (8.53)

By taking \( \delta = q \varepsilon \), for a fixed positive number \( q \), we obtain

\[ \int_{D_{\varepsilon, q \varepsilon}} \{ \Delta_{u_{\varepsilon, q \varepsilon}} \eta_{12}(s, t, y^*(s, t), \psi_0^*, \psi_1^*, \psi_2^*, \psi_{12}^*, u^*(s, t)) \} dt \, ds + o(\varepsilon^2) \geq 0 \]

--- (8.54)

By invoking standard results on differentiability of integrals, we have, for almost all points \( (s_1, t_1) \),

\[ \int_{D_{\varepsilon, q \varepsilon}} \{ \Delta_{u_{\varepsilon, q \varepsilon}} \eta_{12}(s, t, y^*(s, t), \psi_0^*, \psi_1^*, \psi_2^*, \psi_{12}^*, u^*(s, t)) \} dt \, ds = \]

\[ = q \varepsilon^2 \Delta_{u_{\varepsilon, q \varepsilon}} \eta_{12}(s_1, t_1, y^*(s, t), \psi_0^*, \psi_1^*, \psi_2^*, \psi_{12}^*, u^*(s_1, t_1)) + o(\varepsilon^2) \]

--- (8.55)

For points \( (s_1, t_1) \) for which (8.55) holds, by dividing both sides of (8.54) by \( q \varepsilon^2 \) and taking the limit as \( \varepsilon \to 0^+ \), we obtain
\[ \eta_{12}(s_1, t_1, y^*(s, t), \psi_0^*, \psi_1^*(.), \psi_2^*(.), \psi_{12}^*(.), u^*(s_1, t_1)) - 
\eta_{12}(s_1, t_1, y^*(s, t), \psi_0^*, \psi_1^*(.), \psi_2^*(.), \psi_{12}^*(.), u_1) \geq 0 \quad \text{--- (8.55)} \]

which is tantamount to the partial extremal principle for \( u_{12} \) that we stated in section 7 above. ///
References

[B1]. S. A. Belbas, "Optimal control of 2-D systems with nonlocal terms", International Multiconference on Circuits, Systems, Communications and Computers, Athens, 2000, CD proceedings, 6 pp.

[B2]. S. A. Belbas, "Dynamic programming approach to the optimal control of systems governed by Goursat-Darboux equations", International Journal of Control, Vol. 51, 1990, pp. 1279-1294.

[B3]. S. A. Belbas, "Dynamic programming and maximum principle for discrete Goursat systems", Journal of Mathematical Analysis and Applications, Vol. 161, 1991, pp. 57-77.

[B4]. S. A. Belbas, "Optimal control of Goursat-Darboux systems with discontinuous co-state", Applied Mathematics and Computation, Vol. 186, 2007, pp. 101-116.

[B5]. S. A. Belbas, "A new method for optimal control of Volterra integral equations", Applied Mathematics and Computation, Vol. 189, 2007, pp. 1902-1915.

[BDMO]. E. P. Bokmelder, V. A. Dykhta, A. I. Moskalenko, N. A. Ovsyanikova, Conditions of extremum and constructive methods of solution in problems of optimization of hyperbolic systems (in Russian), Nauka, Novosibirsk, 1993.

[GK]. R. Gabasov, F. M. Kirillova, The maximum principle in the theory of optimal control (in Russian), Nauka i Tekhnika, Minsk, 1974.

[S]. W. H. Schmidt, "Durch Integralgleichungen beschriebenene optimale Prozesse mit Nebenbedingungen in Banachräumen - notwendige Optimalitätsbedingungen", Zeitschrift für angewandte Mathematik und Mechanik, Vol. 62, 1982, pp. 65-75.

[SU]. M. B. Suryanarayana, On multidimensional integral equations of Volterra type, University of Michigan Mathematics Department Technical Report #12, Ann Arbor, MI, 1970.

[VST]. O. V. Vasil'ev, V. A. Srochko, V. I. Terletski, Methods of optimization and their applications, vol. 2: optimal control (in Russian), Nauka, Novosibirsk, 1990.

[V]. V. R. Vinokurov, "Optimal control of processes described by integral equations, Parts I, II, and III", SIAM Journal of Control, Vol. 7, 1969, pp. 324-355.