Some theorems about the solutions of Stieltjes differential equations and the existence of optimal Stieltjes controllers

TECHNOLOGICAL UNIVERSITY EINDHOVEN

Department of Mathematics

STATISTICS AND OPERATIONS RESEARCH GROUP

Memorandum COSOR 75-05

Some Theorems about the Solutions of Stieltjes Differential Equations and the Existence of Optimal Stieltjes Controllers

by

A.J.E.M. Janssen

Eindhoven, April 1975

Some Theorems about the Solutions of Stieltjes Differential Equations and the Existence of Optimal Stieltjes Controllers

by

A.J.E.M. Janssen

0. Introduction.

Consider a differential equation of the following type

$$(1)\qquad \begin{aligned} \dot{x}(t) &= A(x(t),t)u(t) \qquad (t \in [0,T]) \\ x(0) &= x_0 \end{aligned}$$

where $x(t) \in \mathbb{R}^n$ $(t \in [0,T])$ $x_0 \in \mathbb{R}^n$ and $T > 0$ (given constants), the controller u a non-negative integrable real function defined on $[0,T]$, $A : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$.
In this paper we consider what we call Stieltjes differential equations

$$(2)\qquad x(t) = x_0 + \int_{[0,t)} A(x(\tau),\tau)d\lambda(\tau) \qquad (t \in [0,T])$$

where λ is a finite measure on ($[0,T]$, $\mathcal{B}$) ($\mathcal{B}$ is the set of Borelsets in $[0,T]$).
We may compare (1) and (2) by interpreting λ as a controller having the character of a δ-function. Such controllers will be called Stieltjes controllers.
In section 1 of this paper we prove a number of lemmas. These lemmas will be used to find sufficient conditions (restrictions on A) for the existence and uniqueness of solutions of (2) (theorem E I). It should be remarked that (2) has , in general, not a (unique) solution if we replace the integration interval $[0,t)$ by $[0,t]$.
In section 2 we shall also deal with systems of the form

$$(3)\qquad x(t) = x_0 + \int_{[0,t)} A(x(\tau),\tau)d\lambda(\tau) + \int_{[0,t)} B(x(\tau),\tau) \qquad (t \in [0,T])$$

where $B: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$.
It is proved that (3) can be reduced to a system of the type (2).
In section 3 some theorems of the theory of ordinary differential equations will be generalized to theorems about Stieltjes differential equations.
In the last section of this paper, section 4, we shall discuss the existence of optimal Stieltjes controllers. We consider the case that λ is an element

of the class W of measures μ satisfying $\mu([0,T]) \leq M$ (where M is a given positive constant, something like the "total energy").
The notion of optimality is not quite obvious here; we use the following interpretation of optimal. Let $\lambda \in W$ and let x be the solution of (3). Define

$$y_\lambda(T) := x_0 + \int_{[0,T]} A(x(t),t)d\lambda(t) + \int_{[0,T]} B(x(t),t)dt.$$

We prove, using the theorems of section 3, that the set

$$V := \left\{ y_\lambda(T) \mid \lambda \in W \right\}$$

is closed and bounded if A en B satisfy certain conditions (theorem E II).
Let now $h : \mathbb{R}^n \to \mathbb{R}$ be a continuous mapping. We shall call the Stieltjes controller λ <u>optimal</u> for the 6-tuple (x_0,T,A,B,M,h) if $h(y_\lambda(T))$ is maximal.

1. <u>Notation and some lemmas</u>.

1.1. <u>Remarks about the notation used</u>. Let $T > 0$. The collection of Borelsets of $[0,T]$ is denoted by $\mathcal{B}$. If we use the word measurable then this should be interpreted as $\mathcal{B}$-measurable. In this section ν is a finite measure on $([0,T], \mathcal{B})$. We consider $\mathbb{R}^n (n \in \mathbb{N})$ with the usual inner product norm

$$|x| := \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} \qquad (x \in \mathbb{R}^n).$$

Furthermore we use the lambda-notation of Church: if a function f is defined on a set X and its value is given by an expression E(x) for every $x \in X$, we write $f = \underset{x \in X}{\Psi} E(x)$ (or, if it is obvious that $x \in X$, $\underset{x}{\Psi} E(x)$).

1.2. <u>Lemma</u> : Let $y : [0,T] \to \mathbb{R}$ be a bounded, measurable function. Define

$$f(t) := \int_{[0,t)} y(\tau)d\nu(\tau) \qquad (t \in [0,T]).$$

Then f is of bounded variation and continuous from the left in $[0,T]$.
<u>Proof</u>: We can split $y(t) = y^+(t) - y^-(t)$ into a non-negative and a non-positive part on $[0,T]$. We have

$$f(t) = \int_{[0,t)} y^+(\tau)d\nu(\tau) - \int_{[0,t)} y^-(\tau)d\nu(\tau) \qquad (t \in [0,T]).$$

Now f is written as the difference of two bounded, non-decreasing functions on [0,T]. It follows that f is of bounded variation in [0,T].
Now let $t \in [0,T]$. For every s in $[0,t)$ we have

$$\left| \int_{[0,t)} y(\tau)d\nu(\tau) - \int_{[0,s)} y(\tau)d\nu(\tau) \right| \le \int_{[s,t)} K d\nu(\tau) = K\nu([s,t))$$

where K is an upper bound for $|y(\tau)|$ $(\tau \in [0,T])$. Using

$$\lim_{s \uparrow t} \nu([s,t)) = 0$$

we obtain that f is continuous from the left in [0,T]. □

1.3. For the proof of theorem E I we shall need some estimates of the Stieltjes-Lebesgue integrals

$$\int_{(\tau,1]} g(F(t))dF(t) \quad \text{resp.} \int_{(\tau,1]} g(F(t-0))dF(t) \quad (\tau \in [0,1]),$$

where F is a non-decreasing, non-negative bounded function, continuous from the right on [0,1], and g is a non-decreasing, continuous function on $\mathbb{R}$.

1.4. <u>Lemma E I</u> : Let F and g be as described in 1.3. Then we have

$$\int_{(\tau,1]} g(F(t-0))dF(t) \le \int_{F(\tau)}^{F(1)} g(u)du \le \int_{(\tau,1]} g(F(t))dF(t).$$

<u>Proof</u>. We give the proof in three steps.
1) First assume that F is continuous on $(\tau,1)$ We define

$$\bar{F}(t) := \begin{cases} F(\tau+0) & \text{at } t = \tau \\ F(t) & \text{for } \tau < t < 1 \\ F(1-0) & \text{at } t = 1. \end{cases}$$

Then $\bar{F}$ is uniformly continuous on $[\tau,1]$. Moreover we have

$$\int_{(\tau,1)} g(F(t))dF(t) = \int_{[\tau,1]} g(\bar{F}(t))d\bar{F}(t) = \int_{(\tau,1)} g(F(t-0))dF(t).$$

Given $\varepsilon > 0$, there is a partition $\tau = \lambda_1 < \ldots < \lambda_k = 1$ such that $\bar{F}(\lambda_i) - \bar{F}(\lambda_{i-1}) < \varepsilon$ $(i=2,\ldots,k)$. From the fact that F and g are non-decreasing we infer

$$L := \sum_{i=1}^{k-1} g(\bar{F}(\lambda_i))\,[\bar{F}(\lambda_{i+1}) - \bar{F}(\lambda_i)] \le \int_{[\tau,1]} g(\bar{F}(t))\,d\bar{F}(t)$$

$$\le \sum_{i=1}^{k-1} g(\bar{F}(\lambda_{i+1}))\,[\bar{F}(\lambda_{i+1}) - \bar{F}(\lambda_i)] =: R,$$

and for $i=1,\ldots,k-1$

$$g(\bar{F}(\lambda_i))\left[\bar{F}(\lambda_{i+1})-\bar{F}(\lambda_i)\right] \le \int_{\bar{F}(\lambda_i)}^{\bar{F}(\lambda_{i+1})} g(u)du \le g(\bar{F}(\lambda_{i+1}))\left[\bar{F}(\lambda_{i+1})-\bar{F}(\lambda_i)\right].$$

So we may conclude

$$L \le \int_{[\tau,1]} g(\bar{F}(t))d\bar{F}(t) \le R \; ; \; L \le \int_{\bar{F}(\tau)}^{\bar{F}(1)} g(u)du \le R.$$

Finally we have

$$0 \le R-L \le \sum_{i=1}^{k-1} \left[g(\bar{F}(\lambda_{i+1})) - g(\bar{F}(\lambda_i))\right].\ (\bar{F}(\lambda_{i+1}) - \bar{F}(\lambda_i))$$

$$\le \varepsilon \sum_{i=1}^{k-1} \left[g(\bar{F}(\lambda_{i+1})) - g(\bar{F}(\lambda_i))\right]$$

$$= \varepsilon\left[g(\bar{F}(1)) - g(\bar{F}(0))\right].$$

We can take $\varepsilon > 0$ arbitrarily small, thus proving

$$\int_{[\tau,1]} g(\bar{F}(t))d\bar{F}(t) = \int_{\bar{F}(\tau)}^{\bar{F}(1)} g(u)du.$$

And from this it follows that

$$\int_{(\tau,1)} g(F(t))dF(t) = \int_{F(\tau+0)}^{F(1-0)} g(u)du = \int_{(\tau,1)} g(F(t-0))dF(t).$$

2) Now assume that F is continuous on $(\tau,1]$ with the exception of the points of discontinuity $a_1 < \ldots < a_m$ and define $a_0 = \tau$ and (only in the case that $a_m \neq 1$) $a_{m+1} = 1$. Applying the result of 1) and using the fact that F is continuous from the right we obtain

$$\int_{(\tau,1]} g(F(t))dF(t) = \sum_{i=0}^{m} \left\{ \int_{(a_i,a_{i+1})} g(F(t))dF(t) + \int_{\{a_{i+1}\}} g(F(t))dF(t) \right\}$$

$$= \sum_{i=0}^{m} \left\{ \int_{F(a_i)}^{F(a_{i+1}-0)} g(u)du + g(F(a_{i+1}))[F(a_{i+1}) - F(a_{i+1}-0)] \right\}$$

$$\geq \sum_{i=0}^{m} \left\{ \int_{F(a_i)}^{F(a_{i+1}-0)} g(u)du + \int_{F(a_{i+1}-0)}^{F(a_{i+1})} g(u)du \right\}$$

$$= \int_{F(\tau)}^{F(1)} g(u)du.$$

Here we again used that F and g are non-decreasing on $(\tau,1]$. It is easily seen that we can use the same argument to prove

$$\int_{(\tau,1]} g(F(t-0))dF(t) \leq \int_{F(\tau)}^{F(1)} g(u)du.$$

3) Now let $(a_m)_{m \in \mathbb{N}}$ be an enumeration of the discontinuities of F in $(\tau,1]$. We define for $m \in \mathbb{N}$

$$F_m(t) := F(t) - \sum_{\substack{a_k \leq t, \\ k>m}} [F(a_k) - F(a_k - 0)] \qquad (t \in [\tau,1]).$$

Notice that every F_m is continuous from the right, non-decreasing, and, moreover, has a finite number of discontinuities. Furthermore we observe that

$$\sum_{\substack{a_k \leq t, \\ k>m}} [F(a_k) - F(a_k - 0)] \qquad (t \in [\tau,1]).$$

is non-decreasing and continuous from the right. So we find

$$\left| \int_{(\tau,1]} g(F(t\pm 0))d(F(t) - F_m(t)) \right| \leq$$

$$\int_{(\tau,1]} M_0 d\Big(\sum_{\substack{a_k \leq t, \\ k>m}} [F(a_k) - F(a_k - 0)] \Big) =$$

$$M_0 \sum_{\substack{a_k \leq 1, \\ k>m}} [F(a_k) - F(a_k - 0)] \to 0 \qquad (m\to\infty).$$

Here M_0 is an upper bound for $|g(x)|$ $(x\in[0,F(1)])$. It is not difficult to show that $F_m(t\pm 0)\to F(t\pm 0)$ $(m\to\infty)$ uniformly on $(\tau,1]$ and, using the fact that g is uniformly continuous on $[0,F(1)]$, we find that $g(F_m(t\pm 0)) \to g(F(t\pm 0))$ uniformly on $(\tau,1]$. We conclude

$$\lim_{m\to\infty} \int_{(\tau,1]} | g(F_m(t\pm 0)) - g(F(t\pm 0))| dF_m(t) = 0.$$

Now, using the result of 2), we obtain

$$\int_{(\tau,1]} g(F(t))dF(t) = \lim_{m\to\infty} \int_{(\tau,1]} g(F_m(t))dF_m(t)$$

$$\geq \lim_{m\to\infty} \int_{F_m(\tau)}^{F_m(1)} g(u)du$$

$$= \int_{F(\tau)}^{F(1)} g(u)du.$$

In the same way we prove that

$$\int_{(\tau,1]} g(F(t-0))dF(t) \leq \int_{F(\tau)}^{F(1)} g(u)du.$$

□

1.5. Remark: The inequalities of 1.4. are no longer valid if we replace the interval of integration by $[\tau,1]$.

Example: Let $g : \mathbb{R} \to \mathbb{R}$ be a continuous non-decreasing function and define

$$F(t) = \begin{cases} 0 \text{ if } t \in [0,\frac{1}{2}) \\ 1 \text{ if } t \in [\frac{1}{2},1]. \end{cases}$$

If we choose $\tau = \frac{1}{2}$ we find successively

$$\int_{[\frac{1}{2},1]} g(F(t))dF(t) = g(1)$$

$$\int_{[\frac{1}{2},1]} g(F(t-0))dF(t) = g(0)$$

$$\int_{F(\frac{1}{2})}^{F(1)} g(u)du = 0$$

By making suitable choices for g, it is easily seen that both inequalities can be violated.

2. A theorem about the existence and uniqueness of the solution of Stieltjes differential equations (Theorem E I).

2.1. We consider the following Stieltjes differential equation

$$(*) \qquad x(t) = x_0 + \int_{[0,t)} H(x(\tau),\tau)d\lambda(\tau) \qquad (t \in [0,T]),$$

where $x(t) \in \mathbb{R}^n$ $(t \in [0,T])$, $H : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$, λ a finite measure on $([0,T], \mathcal{B})$, and $x_0 \in \mathbb{R}^n$ and $T > 0$ given constants. In section 4 we shall consider Stieltjes differential equations of the form

$$(**) \qquad x(t) = x_0 + \sum_{i=1}^{k} \int_{[0,t)} A_i(x(\tau),\tau)d\lambda_i(\tau) \quad (t \in [0,T]),$$

where $A_i : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ $(i = 1,\ldots,k)$ and λ_i a finite measure on $([0,T], \mathcal{B})$ $(i = 1,\ldots,k)$.

2.2. Lemma: It is possible to reduce system (**) to system (*).

Proof: Define a new measure λ on $([0,T],\mathcal{B})$ by

$$\lambda(A) := \sum_{i=1}^{k} \lambda_i(A) \qquad (A\in\mathcal{B}).$$

Notice that every λ_i is absolutely continuous with respect to λ. Therefore we can apply the Radon Nikodym theorem (see [SG] page 189). For $i=1,\dots,k$ there exists a measurable function f_i, non-negative and λ-summable, such that for each $A\in\mathcal{B}$

$$\lambda_i(A) = \int_A f_i d\lambda.$$

Furthermore it follows that for each $A\in\mathcal{B}$

$$\lambda(A) = \sum_{i=1}^{k} \int_A f_i d\lambda = \int_A \sum_{i=1}^{k} f_i d\lambda ,$$

so that $\sum_{i=1}^{k} f_i = 1$ holds almost everywhere in the sense of λ. By choosing $f_i = 0$ on the set where $\sum_{i=1}^{k} f_i \neq 1$ we achieve that $0\le f_i \le 1$ on $[0,T]$ $(i=1,\dots,k)$. Now we are able to write (**) formally as a system of the form (*):

$$x(t) = x_0 + \int_{[0,t)} \sum_{i=1}^{k} A_i(x(\tau),\tau) f_i(\tau) d\lambda(\tau) \qquad (t\in[0,T]). \quad \square$$

2.3. From now on we consider the system (*).

Definition: For $t\in[0,T]$ we have the λ-generated function

$$F(t) := \lambda([0,t]).$$

It is easily seen that F is non-decreasing, continuous from the right and non-negative. Therefore we can write the integral

$$\int_A g(\tau)d\lambda(\tau)$$

as a Stieltjes-Lebesgue integral viz.

$$\int_A g(\tau)dF(\tau).$$

We also replace λ in (*) by F.

2.4. Remark: In the case that λ is a signed measure, its generated function F will be continuous from the right and of bounded variation. We can split F into a non-decreasing part G_1 and a non-increasing part G_2:

$$F(t) = G_1(t) + G_2(t) \qquad (t \in [0,T]).$$

If we define

$$\bar{G}_1(t) := \lim_{\tau \downarrow t} G_1(\tau), \; \bar{G}_2(t) = \lim_{\tau \downarrow t} G_2(\tau) \qquad (t \in [0,T)).$$

$$\bar{G}_1(T) = G_1(T), \; \bar{G}_2(T) = G_2(T)$$

then $\bar{G}_1$ and $\bar{G}_2$ are continuous from the right and

$$F(t) = \bar{G}_1(t) + \bar{G}_2(t) \qquad (t \in [0,T]).$$

We can consider $\bar{G}_1$ and $\bar{G}_2$ as generated functions of the measures λ_1 and λ_2 defined by $(0 \le t_0 \le t_1 \le T)$

$$\lambda_1(\{0\}) = \bar{G}_1(0), \lambda_1((t_0,t_1]) = \bar{G}_1(t_1) - \bar{G}_1(t_0);$$

$$\lambda_2(\{0\}) = \bar{G}_2(0), \lambda_2((t_0,t_1]) = \bar{G}_2(t_1) - \bar{G}_2(t_0).$$

So we have written a system with a signed measure as a system (**).

2.5. Definition: Let L be the set of n-vector functions that are measurable and λ-summable on $[0,T]$.

2.6. Definition: Let $\alpha \ge 0$. For every $f \in L$ we define

$$\|f\|_\alpha := \int_{[0,T]} \exp(-\alpha F(t)) \cdot |f(t)| \, dF(t).$$

It is easily seen that $\| \ \|_\alpha$ is well defined on L. Furthermore it is not hard to prove that $\| \ \|_\alpha$ is a semi-norm in L.

2.7. For every $\alpha > 0$ and every $f \in L$ we have

$$\exp(-\alpha F(T)).\ \| f \|_0 \leq \| f \|_\alpha \leq \| f \|_0 .$$

So $\| \ \|_\alpha$ and $\| \ \|_0$ are equivalent on L for $\alpha \geq 0$.

2.8. From now on we identify elements $f, g \in L$ if $\| f-g \|_0 = 0$. Then, using 2.7, we obtain $\| f-g \|_\alpha = 0$ $(\alpha \geq 0)$. We shall call the space constructed this way $(\mathcal{L}, \| \ \|_\alpha)$ on which $\| \ \|_\alpha$ is a positive norm.

2.9. <u>Theorem</u>: For every $\alpha \geq 0$ $(\mathcal{L}, \| \ \|_\alpha)$ is a Banach space.

<u>Proof</u>: We remark that $(\mathcal{L}, \| \ \|_0)$ is a Banach space and according to 2.7. $\| \ \|_0$ and $\| \ \|_\alpha$ are equivalent on $\mathcal{L}$. □

2.10. Let $H : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ be a function satisfying the following conditions

1) $\exists_{L > 0}\ \forall_{x,y \in \mathbb{R}^n}\ \forall_{t \in [0,T]} [|H(x,t) - H(y,t)| \leq L|x-y|]$

2) $\forall_{x \in \mathbb{R}^n} [\ \Psi_{t \in [0,T]} H(x,t) \in \mathcal{L}]$

3) $\Psi_{t \in [0,T]} H(0,t)$ is bounded.

2.11. <u>Lemma</u>: If H satisfies the conditions 2.10. 1), 2) and 3) we have

i) there exists a positive number M such that for every $x_0 \in \mathbb{R}^n$

$$H(x_0,t) \leq M + L|x_0| \qquad (t \in [0,T])$$

ii) if $f : [0,T] \to \mathbb{R}^n$ is measurable then so is $\Psi_{t \in [0,T]} H(f(t),t)$

iii) if $f \in \mathcal{L}$ then $\Psi_{t \in [0,T]} H(f(t),t) \in \mathcal{L}$.

<u>Proof</u>: (i) For every $x_0 \in \mathbb{R}^n$ and $t \in [0,T]$ we have

$$|H(x_0,t) - H(0,t)| \leq L|x_0|$$

therefore

$$|H(x_0,t)| \leq M + L|x_0|$$

where M is an upper bound for $|H(0,t)|$ $(t \in [0,T])$.

ii) Let $f: [0,T] \to \mathbb{R}^n$ be a measurable function. There exists a sequence of elementary functions

$$f_k = \sum_{i=1}^{\infty} \alpha_i^{(k)} A_i^{(k)}$$

for which $f_k \to f$ holds pointwise. The sets $(A_i^{(k)})_{i \in \mathbb{N}}$ are, for every $k \in \mathbb{N}$, measurable and mutually disjoint. Using the fact that H is continuous with respect to x we have

$$H(f_k(t),t) \to H(f(t),t) \qquad (t \in [0,T]).$$

For every $k \in \mathbb{N}$ we can write

$$H(f_k(t),t) = H(\sum_{i=1}^{\infty} \alpha_i^{(k)} A_i^{(k)}(t),t)$$

$$= \sum_{i=1}^{\infty} H(\alpha_i^{(k)},t)\ A_i^{(k)}(t) \qquad (t \in [0,T]).$$

So $\Psi_{t \in [0,T]} H(f_k(t),t)$ is measurable $(k \in \mathbb{N})$ and therefore $\Psi_{t \in [0,T]} H(f(t),t)$.

iii) This follows immediately from i) and ii). □

2.12. <u>Remark</u>: If $H(x,t) = \sum_{i=1}^{k} A_i(x,t)\ f_i(t)$ as in 2.2., then H satisfies the conditions 2.10. 1), 2) and 3) if every A_i does. This follows from the fact that f_i is measurable and bounded for $i = 1,\dots,k$.

2.13. Now we define on $\mathcal{L}$ the following operator

$$\varphi := \Psi_{x \in \mathcal{L}} \ \Psi_{t \in [0,T]} (x_0 + \int_{[0,t)} H(x(\tau),\tau)dF(\tau)).$$

We observe that φ maps $\mathcal{L}$ into $\mathcal{L}$.

2.14. <u>Theorem</u>: If $\alpha > L$ then φ is a contraction of $(\mathcal{L}, \| \ \|_\alpha)$.

<u>Proof</u>: Let $y,z \in \mathcal{L}$. For every $t \in [0,T]$ we have

$$|(\varphi y)(t) - (\varphi z)(t)| \le L \int_{[0,t)} |y(\tau) - z(\tau)| dF(\tau)$$

hence

$$\| \varphi y - \varphi z \|_\alpha = \int_{[0,T]} \exp(-\alpha F(t)) \; |(\varphi y)(t) - (\varphi z)(t)| \; dF(t)$$

$$\le L \int_{[0,T]} \exp(-\alpha F(t)) \; [\int_{[0,t)} |y(\tau) - z(\tau)| \; dF(\tau)] dF(t).$$

Now we apply Fubini's theorem and find

$$\| \varphi y - \varphi z \|_\alpha \le L \int_{[0,T]} \{ \int_{(\tau,T]} |y(\tau) - z(\tau)| \exp(-\alpha F(t)) dF(t) \} dF(\tau)$$

$$= L \int_{[0,T]} |y(\tau) - z(\tau)| \{ \int_{(\tau,T]} \exp(-\alpha F(t)) dF(t) \} dF(\tau).$$

Lemma E I (1.4.) applied to the inner integral with $g = \Psi_{x \in \mathbb{R}} -\exp(-\alpha x)$ gives

$$\int_{(\tau,T]} \exp(-\alpha F(t)) dF(t) \le \alpha^{-1} [\exp(-\alpha F(\tau)) - \exp(-\alpha F(T))]$$

$$\le \alpha^{-1} \exp(-\alpha F(\tau)).$$

Therefore we conclude

$$\| \varphi y - \varphi z \|_\alpha \le \alpha^{-1} L \int_{[0,T]} |y(\tau) - z(\tau)| \exp(-\alpha F(\tau)) dF(\tau)$$

$$= \alpha^{-1} L \| y - z \|_\alpha .$$

□

2.15. Theorem E I:

i) There exists an in λ-sense unique solution of (*).

ii) The solution of (*) is pointwise unique.

Proof:

i) Let $\alpha > L$ and apply the Banach theorem to the operator φ in

$(\mathcal{L}, \|\ \|_\alpha)$; remark that $\|\ \|_\alpha$ and $\|\ \|_0$ are equivalent on $\mathcal{L}$ (2.7.).
ii) Assume that (for $t \in [0,T]$) $\varphi x = x$ and $\varphi y = y$:

$$x(t) = x_0 + \int_{[0,t)} H(x(\tau),\tau)dF(\tau)$$

$$y(t) = x_0 + \int_{[0,t)} H(y(\tau),\tau)dF(\tau).$$

Then x and y are equal in the sense of λ, thus for $t \in [0,T]$

$$\int_{[0,t)} H(x(\tau),\tau)dF(\tau) = \int_{[0,t)} H(y(\tau),\tau)dF(\tau).$$

This means that $x = y$ pointwise in $[0,T]$. □

3. Some theorems about the solutions of (*).

3.1. The theorems proved in this section will be applied in section 4 with the exception of theorem 3.7. . This last theorem is included here because it can serve as a point of departure to extend the theory of optimal Stieltjes controllers (e.g. finding necessary conditions for optimal Stieltjes controllers).

3.2. Theorem: Let x be the solution of (*). There exist constants C_1 and C_2 (only depending on H, x_0 and F(T)) such that

1) $\|x\|_0 \le C_1$

2) $|x(t)| \le C_2 \quad (t \in [0,T])$.

Proof:
1) Let $\alpha = 2L$ and let φ be the operator as defined in 2.13.. Then φ is a contraction of $(\mathcal{L},\|\ \|_\alpha)$ with contraction constant $\le \frac{1}{2}$. Therefore we have

$$\|x\|_\alpha = \|\varphi x\|_\alpha \le \|\varphi x - \varphi(\underset{t}{\Psi} x_0)\|_\alpha + \|\varphi(\underset{t}{\Psi} x_0)\|_\alpha$$

$$\le \tfrac{1}{2}\|x - \underset{t}{\Psi} x_0\|_\alpha + \|\varphi(\underset{t}{\Psi} x_0)\|_\alpha$$

$$\leq \tfrac{1}{2}\|x\|_\alpha + \tfrac{1}{2}\|\underset{t}{\Psi} x_0\|_\alpha + \|\varphi(\underset{t}{\Psi} x_0)\|_\alpha.$$

Now using 2.11. we find

$$\|x\|_\alpha \leq \|\underset{t}{\Psi} x_0\|_\alpha + 2\|\varphi(\underset{t}{\Psi} x_0)\|_\alpha$$

$$\leq F(T)\,|x_0| + 2\|\underset{t}{\Psi}(x_0 + L\,|x_0|\,F(T) + MF(T))\|_\alpha$$

$$\leq 3F(T)|x_0| + 2\,L|x_0|F^2(T) + 2MF^2(T).$$

The first part of the theorem now follows from 2.7..

2) For every $t \in [0,T]$ we have, again using 2.11.,

$$|x(t)| \leq |x_0| + \int_{[0,t)} |H(x(\tau),\tau)|\,dF(\tau)$$

$$\leq |x_0| + L\,\|x\|_0 + MF(T).$$

The theorem follows immediately from 1). □

3.3. <u>Theorem</u>: Let x_0, $y_0 \in \mathbb{R}^n$. The solutions of (*) corresponding to x_0 resp. y_0 are indicated by x resp. y. There exist numbers M_1, $M_2 \geq 0$ (only depending on H and F(T)) such that

1) $\|x - y\|_0 \leq M_1\,|x_0 - y_0|$

2) $|x(t) - y(t)| \leq M_2|x_0 - y_0|$ $\qquad (t \in [0,T])$.

<u>Proof</u>:

1) Let φ_1 resp. φ_2 be the operator as defined in 2.13. corresponding to x_0 resp. y_0. We choose $\alpha = 2L$ and find

$$\|x - y\|_\alpha = \|\varphi_1 x - \varphi_2 y\|_\alpha \leq \|\varphi_1 x - \varphi_1 y\|_\alpha + \|(\varphi_1 - \varphi_2)y\|_\alpha$$

$$\leq \tfrac{1}{2}\|x - y\|_\alpha + \|\underset{t}{\Psi}(x_0 - y_0)\|_\alpha.$$

We may complete the proof in a manner similar to 3.2.1).

2)For every $t \in [0,T]$ we have according to 2.10.

$$|x(t) - y(t)| \leq |x_0 - y_0| + L\int_{[0,t)} |x(\tau) - y(\tau)|\,dF(\tau)$$

$$\leq |x_0 - y_0| + L\,\|x - y\|_0.$$

Application of 1) gives the required inequality immediately. □

3.4. <u>Remark</u> : Theorem 3.3. states that the solution of (*) depends continuously on the initial value x_0. This theorem also occurs in the theory of ordinary differential equations.

3.5. <u>Theorem</u>: Let $x_0 \in \mathbb{R}^n$. There exists a constant $M > 0$ (only depending on H, x_0 and F(T)) such that the total variation of the solution of (*) is at most M.

<u>Proof</u>: If x is the solution of (*) and $0 = t_0 < t_1 < \ldots < t_k = T$ is a partition of $[0,T]$, we have

$$\sum_{j=1}^{k} |x(t_j) - x(t_{j-1})| = \sum_{j=1}^{k} \Big| \int_{[t_{j-1},t_j)} H(x(\tau),\tau)dF(\tau) \Big|$$

$$\leq \sum_{j=1}^{k} \int_{[t_{j-1},t_j)} |H(x(\tau),\tau)| dF(\tau)$$

$$= \int_{[0,T)} |H(x(\tau),\tau)| dF(\tau).$$

By using 2.11. and 3.2. the theorem is easily proved. □

3.6. We now choose in (*)

$$H(x,\tau) = A(\tau)x \qquad (x \in \mathbb{R}^n, t \in [0,T]).$$

Here A is a bounded, measurable, matrix valued function on $[0,T]$.
It is easily seen that H satisfies the conditions 2.10. and therefore we have for every $x_0 \in \mathbb{R}^n$ a unique solution of (*) which will be indicated by $\xi(t,x_0)$ $(t \in [0,T])$. We define (just as in the theory of ordinary differential equations) a <u>fundamental solution</u> φ_A^F by

$$\varphi_A^F(t) = I + \int_{[0,t)} A(\tau)\varphi_A^F(\tau)dF(\tau) \qquad (t \in [0,T]).$$

For every $x_0 \in \mathbb{R}^n$ we have

$$\xi(t,x_0) = \varphi_A^F(t)x_0 \qquad (t \in [0,T]).$$

3.7. We now assume that H in (*) has a continuous derivative with respect to x (notation for this derivative $\nabla_x H(c,t)$)and furher we assume

(i) $\forall_{c\in\mathbb{R}^n} [\underset{t\in[0,T]}{\Psi} \nabla_x H(c,t)$ is measurable]

(ii) $\exists_{M,L>0}\forall_{c\in\mathbb{R}^n}\forall_{t\in[0,T]}[\ |\nabla_x H(c,t)| \le M + L|c|\ ]$.

It is a matter of routine (see 2.11.) to show that

$$\underset{t\in[0,T]}{\Psi} \nabla_x H(f(t),t)$$

is measurable and λ-summable if f is. We denote the solution of (*) with initial value x_0 by $\xi(t,x_0)$ $(t \in [0,T])$ and we are interested in the derivative of $\xi(t,x_0)$ with respect to x_0 (notation: $\nabla_{x_0}\xi(t,c)$). We prove the following

<u>Theorem</u>: $\xi(t,x_0)$ $(t \in [0,T])$ has a continuous derivative with respect to x_0 and

$$\nabla_{x_0}\xi(t,c) = \varphi_A^F(t) \qquad (t \in [0,T]).$$

Where φ_A^F satisfies

$$\varphi_A^F(t) = I + \int_{[0,t)} \nabla_x H(\xi(\tau,c),\tau)\ \varphi_A^F(\tau)dF(\tau) \qquad (t \in [0,T]).$$

<u>Proof</u>: Take a fixed $c \in \mathbb{R}^n$ and let $x_0 \in \mathbb{R}^n$. Consider

$$v(t) := \xi(t,x_0) - \xi(t,c) \qquad (t \in [0,T]).$$

and

$$A(\tau) := \nabla_x H(\xi(\tau,c),\tau) \qquad (t \in [0,T]).$$

We have for every $t \in [0,T]$

$$v(t) = x_0 - c + \int_{[0,t)} A(\tau)v(\tau)dF(\tau) +$$

$$(1) \qquad + \int_{[0,t)} \{[H(\xi(\tau,x_0),\tau) - H(\xi(\tau,c),\tau)\ ] - A(\tau)v(\tau)\}dF(\tau).$$

We wish to compare v with the function z that satisfies

$$(2) \qquad z(t) = x_0 - c + \int_{[0,t)} A(\tau)z(\tau)dF(\tau) \qquad (t \in [0,T]),$$

and therefore we have to estimate the third term of (1). Let K be a bounded set in $\mathbb{R}^n$ and $x,y \in K$. For $\tau \in [0,T]$ we consider

$$R: =[\ H(x,\tau) - H(y,\tau)] - \nabla_x H(y,\tau)\ (x - y)$$

and define for $s \in [0,1]$

$$h(s)\ :=\ [H(x,\tau) - H(x + s(y - x),\tau)] - s\nabla_x H(y,\tau)\ (x - y).$$

We observe that $h(0) = 0$ and $h(1) = R$. Now we have

$$|R| = |h(1)| \leq \int_0^1 |\dot{h}(s)|ds.$$

Using the fact that $\nabla_x H(c,t)$ is continuous with respect to c we find

$$\dot{h}(s) = - \nabla_x H(x+s(y - x),\tau)\ (y - x) + \nabla_x H(y,\tau)\ (y - x)$$

$$= [\nabla_x H(y,\tau) - \nabla_x H(x + s(y - x),\tau)]\ (y - x)$$

$$= o(|y - x|) \qquad (y - x \to 0)$$

uniformly in $x,y \in K$.

From the theorems 3.2. and 3.3. we know that

$$|\xi(\tau,x_0) - \xi(\tau,c)| = 0\ (|x_0 - c|)$$

uniformly on $[0,T]$, and therefore we have for every $\tau \in [0,T]$

$$|[H(\xi(\tau,x_0),\tau) - H(\xi(\tau,c),\tau)] - A(\tau)v(\tau)| = o(|x_0 - c|)(x_0 - c \to 0).$$

Applying the theorem of Lebesgue on dominated convergence we finally find

$$v(t) = x_0 - c + \int_{[0,t)} A(\tau)v(\tau)dF(\tau) + o(|x_0 - c|) \qquad (x_0 - c \to 0)$$

uniformly in $[0,T]$.

Now let φ be the operator corresponding to (2) as defined in 2.13. and choose

α such that φ is a contraction of $(\mathcal{L}, \|\ \|_\alpha)$ with contraction constant $\leq \frac{1}{2}$. We can conclude

$$\|v - z\|_\alpha = \|v - \varphi z\|_\alpha \leq \|\varphi v - \varphi z\|_\alpha + \|v - \varphi v\|_\alpha$$

$$\leq \tfrac{1}{2}\|v - z\|_\alpha + o(|x_0 - c|) \qquad (x_0 \to c),$$

and by 2.7. we obtain

$$\|v - z\|_0 = o(|x_0 - c|) \qquad (x_0 \to c).$$

By the boundedness of A we have as a final result

$$\xi(t,x_0) - \xi(t,c) - \varphi_A^F(t)\,(x_0 - c) = v(t) - z(t) =$$

$$= \int_{[0,t)} A(\tau)[v(\tau) - z(\tau)]dF(\tau) + o(|x_0 - c|)$$

$$= o(|x_0 - c|) \qquad (x_0 \to c).$$

uniformly in $t \in [0,T]$. This completes the proof of the first part of the theorem. We still have to prove that $\nabla_{x_0} \xi(t,c)$ is continuous with respect to x_0.
The proof is analogous to the proof of 3.3. and we shall omit it. □

4. A theorem on the existence of optimal Stieltjes controllers.

4.1. Let $x_0 \in \mathbb{R}^n$ and $M, T > 0$ be given. In this section we consider systems of the following type

$$\text{(I)} \begin{cases} x(t) = x_0 + \int_{[0,t)} A(x(\tau),\tau)d\lambda(\tau) + \int_{[0,t)} B(x(\tau),\tau)d\tau \quad (t \in [0,T]) \\ \lambda([0,T]) \leq M. \end{cases}$$

We assume that A and B satisfy the conditions 2.10. so that, according to 2.12. and 2.15., system (I) has a unique solution. Notice that the theorem 3.2., 3.3. and 3.5. are applicable to system (I). We again define

$$F(t) := \lambda([0,t]) \qquad (t \in [0,T]).$$

4.2. We now restrict ourselves to functions A for which there exists a $k \in \mathbb{N}$, $1 \leq k \leq n$ such that

(1) the first k components of $A(x,\tau)$ are identically zero for $x \in \mathbb{R}^n$, $\tau \in [0,T]$

(2) the last n - k components of $A(x,\tau)$ only depend on the first k components of x for $x \in \mathbb{R}^n$, $\tau \in [0,T]$

(3) $A(x,\tau)$ is a continuous function of $x \in \mathbb{R}^n, \tau \in [0,T]$.

4.3. Definition: Let W be the set of functions defined on $[0,T]$ that are continuous from the right, non-decreasing, and for which

$$0 \le F(t) \le M \qquad (t \in [0,T]).$$

Furthermore we define for $F \in W$

$$y_F := x_0 + \int_{[0,T]} A(x(\tau),\tau)dF(\tau) + \int_{[0,T]} B(x(\tau),\tau)d\tau$$

where x is the solution of (I) corresponding to F, and finally we define

$$V := \{y_F \mid F \in W\}.$$

4.4. For the proof of theorem E II we need Helly's selection principle (see [N] page 222) :

If $(g_m)_{m \in \mathbb{N}}$ is a sequence of functions, uniformly bounded and of uniformly bounded variation on $[0,T]$, then there exists a subsequence of $(g_m)_{m \in \mathbb{N}}$ that converges pointwise to a function g of bounded variation on $[0,T]$.

If $(g_m)_{m \in \mathbb{N}}$ are continuous from the left resp. on the right, it is easily verified that there is a limit function $\bar{g}$, continuous from the left resp. from the right, such that the subsequence converges to $\bar{g}$ in the points of continuity of $\bar{g}$ (it only requires the definition of $\bar{g}(t_0) = \lim_{t \uparrow t_0} g(t)$ $(t_0 \in (0,T])$ $\bar{g}(0) = g(0)$ resp. $\bar{g}(t_0) = \lim_{t \downarrow t_0} g(t) (t_0 \in [0,T)), \bar{g}(T) = g(T)).$

A second theorem we use is the theorem of Arzelà (see [HS] page 226) :

If $(h_m)_{m \in \mathbb{N}}$ is a sequence of functions that are equicontinuous and uniformly bounded on $[0,T]$, then there exists a subsequence that converges to a continuous function g uniformly in $[0,T]$.

The third theorem we use is a special form of the theorem of Helly-Bray (see [N] page 233):

If F and $(F_m)_{m \in \mathbb{N}}$ are non-decreasing, non-negative, uniformly bounded functions which are continuous from the right on $[0,T]$ and $F_m \to F$ on a set that is dense in $[0,T]$ then, for every continuous function g,

$$\lim_{i\to\infty} \int_{[0,t]} g(\tau)dF_i(\tau) = \int_{[0,t]} g(\tau)dF(\tau)$$

in every point $t \in [0,T]$ for which $\lim_{i\to\infty} F_i(t) = F(t)$.

4.5. We are able now to prove the following

Theorem: If A satisfies the conditions 4.3. then V is closed and bounded.

Proof: Let y be a limitpoint of V. Hence there exists a sequence $(F_i)_{i \in \mathbb{N}}$ in W and a corresponding sequence of solutions $(x^{(i)})_{i \in \mathbb{N}}$ of (I) such that $\lim_{i\to\infty} y_{F_i} = y$.

From the form of A it is easy to see that the sequence $((x_1^{(i)},\ldots,x_k^{(i)}))_{i \in \mathbb{N}}$ is equicontinuous on $[0,T]$. Therefore there exists a subsequence (again denoted by $((x_1^{(i)},\ldots,x_k^{(i)}))_{i \in \mathbb{N}}$) and a continuous function $(x_1,\ldots,x_k)$ such that

$$\lim_{i\to\infty} (x_1^{(i)},\ldots,x_k^{(i)}) = (x_1,\ldots.x_k) \text{ uniformly in } [0,T].$$

On the other hand we may apply theorem 3.5. to the sequence $((x_1^{(i)},\ldots,x_n^{(i)}))_{i\in\mathbb{N}}$ and we find that it is of uniformly bounded variation. We conclude that there exists a subsequence (again denoted by $((x_1^{(i)},\ldots,x_n^{(i)}))_{i\in\mathbb{N}}$.) and a function $(x_1,\ldots,x_n)$ continuous from the left in $[0,T]$ such that

$$\lim_{i\to\infty} (x_1^{(i)},\ldots,x_n^{(i)}) = (x_1,\ldots,x_n) =: x$$

in the points of continuity of x and in $t = 0$.

In the same way we find that there exists a subsequence of $(F_i)_{i \in \mathbb{N}}$ (again denoted by $(F_i)_{i \in \mathbb{N}}$) and a function F which is non-decreasing and continuous from the right such that

$$\lim_{i\to\infty} F_i = F$$

in the points of continuity of F, and, according to 4.4., $\lim_{i\to\infty} F_i(T) = F(T)$;

this F satisfies $F(T) \leq M$, so $F \in W$.
We shall prove the following statements

(a) $$x(t) = x_0 + \int_{[0,t)} A(x(\tau),\tau)dF(\tau) + \int_{[0,t)} B(x(\tau),\tau)d\tau \quad (t \in [0,T])$$

(b) $$y = \lim_{i\to\infty} y_{F_i} = x_0 + \int_{[0,T]} A(x(\tau),\tau)dF(\tau) + \int_{[0,T]} B(x(\tau),\tau)d\tau.$$

It is easy to see that $x_i \to x$ dominated in $[0,T]$ with the exception of a set of measure zero.
Therefore we conclude that

$$\lim_{i\to\infty} \int_{[0,T]} B(x^{(i)}(\tau),\tau)d\tau = \int_{[0,T]} B(x(\tau),\tau)d\tau$$

and

$$\lim_{i\to\infty} \int_{[0,t)} B(x^{(i)}(\tau),\tau)d\tau = \int_{[0,t)} B(x(\tau),\tau)d\tau \quad (t \in [0,T]).$$

Using the fact that A only depends on the first k components of x we find

(1) $$\lim_{i\to\infty} A(x^{(i)}(\tau),\tau) = A(x(\tau),\tau)$$

uniformly in $\tau \in [0,T]$. Moreover we have in the points of continuity of F and of the F_i's, using the theorem of Helly-Bray,

(2) $$\lim_{i\to\infty} \int_{[0,t)} A(x(\tau),\tau)d(F_i(\tau) - F(\tau)) = \lim_{i\to\infty} \int_{[0,t]} A(x(\tau),\tau)d(F_i(\tau)-F(\tau))=0.$$

Combining (1) and (2) and recalling that $(F_i)_{i \in \mathbb{N}}$ is uniformly bounded on $[0,T]$ we find

$$\lim_{i\to\infty} \int_{[0,T]} A(x^{(i)}(\tau),\tau)dF_i(\tau) = \int_{[0,T]} A(x(\tau),\tau)dF(\tau)$$

$$\lim_{i\to\infty} \int_{[0,t)} A(x^{(i)}(\tau),\tau)dF_i(\tau) = \int_{[0,t)} A(x(\tau),\tau)dF(\tau)$$

in the continuity points of F and of the F_i's. This proves (b) and furthermore we can conclude that (a) holds in a set that is dense in [0,T]. We observe however, that both sides of equation (a) are continuous from the left and therefore statement (a) holds everywhere in [0,T].
This proves the first part of the theorem: V is closed. It is trivial (by 3.2.) that V is bounded. □

4.6. Remark: In 4.5. we have to take y_F instead of x(T). It is easily verified that the set $\{x(T) \mid F \in W\}$ is not necessarily closed.

4.7. We have restricted ourselves in theorem 4.5. to functions satisfying the conditions 4.2.. The necessity of the restriction is illustrated in the following

Example: Consider the system

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \int_{[0,t)} \begin{bmatrix} f(\tau)\ y(\tau) \\ 1 \end{bmatrix} dF(\tau) \qquad (t \in [0,1])$$

maximize

$$\int_{[0,1]} f(\tau)y(\tau)dF(\tau)$$

under the condition $0 \leq F(t) \leq 1 \qquad (t \in [0,1])$.

Here f is a continuous, non-negative real function on [0,1] for which

$$f(t) < f(1) = 1 \qquad (t \in [0,1], t \neq 1).$$

Using lemma E I (1.4.) we conclude

$$\int_{[0,1]} f(\tau)y(\tau)dF(\tau) = \int_{[0,1]} f(\tau)F(\tau - 0)dF(\tau) \leq \tfrac{1}{2}.$$

Consider the following sequence of controllers ($n \in \mathbb{N}$)

$$F_n = \begin{cases} 0 & (0 \leq x \leq 1 - \frac{1}{n}) \\ n(x - 1 + \frac{1}{n}) & (1 - \frac{1}{n} \leq x \leq 1) \end{cases}$$

From lemma E I (1.4.) it easily follows that

$$\lim_{n\to\infty} \int_{[0,1]} f(\tau)y_n(\tau)dF_n(\tau) = \tfrac{1}{2}.$$

Now suppose we have a controller F for which

$$\int_{[0,1]} f(\tau)y(\tau)dF(\tau) = \tfrac{1}{2}.$$

From the fact that the maximum of f is uniquely achieved in t = 1 we conclude that F = 0 on [0,1) and F(1) = 1. But then we have

$$\int_{[0,1]} f(\tau)y(\tau)dF(\tau) = 0$$

Hence the set V is not closed.

Acknowledgement

This paper is a result of a senior project at the Technological University of Eindhoven. Therefore the author owes many thanks to the supervisor of this project Prof. Dr. Ir. M.L.J. Hautus for his inspiring support and remarks.

References

[SG] : Shilov, G.E. and Gurevich, B.L. ; Integral Measure and Derivative: A Unified Approach; Prentice-Hall, Inc., Englewood Cliffs, N.J.; Revised English Edition, 1966.

[N] : Natanson, I.P. ; Theory of Functions of a Real Variable (Chapter I to IX) ; Frederick Ungar Publishing Co., New York; American Edition, 1955.

[HS] : Heider, L.J. and Simpson, J.E.; Theoretical Analysis; W.B. Saunders Company, Philadelphia-London; 1st edition, 1967.