

Bang-Bang principle for controls in a simplex

Alexander Gudev

Definition. Let $I \subset \mathbb{R}$ be an interval, $A \in L_1(I, M_{n \times n})$. *Fundamental matrix* of the linear system

$$\dot{x}(t) = A(t)x(t) \quad (1)$$

is any matrix function $\Phi : I \rightarrow M_n$, whose columns are linearly independent solutions to the system.

Remark 1. All solutions are expressible in terms of Φ ,

$$x(t) = \sum_i \alpha_i \Phi_i(t) = \Phi(t)\alpha, \quad \alpha \in \mathbb{R}^n.$$

Provided $\det \Phi(t) \neq 0$ for all $t \in I$, each column of $\Phi(t)$ separately is a solution to (1) iff $\dot{\Phi}(t) = A(t)\Phi(t)$. That is, the fundamental matrices of (1) are determined by

$$\dot{\Phi}(t) = A(t)\Phi(t).$$

Everywhere below we assume $\Phi(t_0) = E$; then the vector α above is actually the initial condition $\alpha = x(t_0)$.

In this sense the matrix $\Phi(t)$ „walks“ the solution *forward* as time passes from t_0 to t .

Remark 2. It is natural to expect $\Phi^{-1}(t)$ to walk the solution *backward*, i.e. to solve the system with $(-A)$ (opposite velocity \dot{x}). Indeed, differentiating $\Phi^{-1}\Phi = E$ we get $\frac{d\Phi^{-1}}{dt}\Phi + \Phi^{-1}\dot{\Phi} = 0$, from where $\frac{d\Phi^{-1}}{dt}\Phi = -\Phi^{-1}A\Phi$ and

$$\frac{d\Phi^{-1}}{dt} = -\Phi^{-1}A. \quad (2)$$

Remark 3. Below at each point in time $t \in I$ the state of the system is a vector $x(t) \in \mathbb{R}^n$, and the control is a vector $u(t) \in \mathbb{R}^m$.

Problem. Let $A \in L_1(I, M_{n \times n})$ and $B \in L_1(I, M_{n \times m})$. Solve the system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t)$$

in terms of the fundamental matrix $\Phi(t)$ of $\dot{x}(t) = A(t)x(t)$.

Solution (wiki books: Control Systems/Linear System Solutions). Rewrite $\dot{x} - Ax = Bu$, multiply by $\Phi^{-1}(t)$, and integrate:

$$\begin{aligned} \Phi^{-1}\dot{x} + \Phi^{-1}(-A)x &= \Phi^{-1}Bu \\ \frac{d}{dt}(\Phi^{-1}x) &= \Phi^{-1}Bu \quad (\text{by (2)}) \\ \Phi^{-1}x &= x(t_0) + \int_{t_0}^t \Phi^{-1}(s)B(s)u(s) ds, \end{aligned}$$

from where

$$x(t) = \Phi(t)x(t_0) + \Phi(t) \int_{t_0}^t \Phi^{-1}(s)B(s)u(s) ds.$$

Below for convenience we will assume $x(t_0) = 0$, in which case $x(t) = \int_{t_0}^t \Phi^{-1}(s)B(s)u(s) ds$.

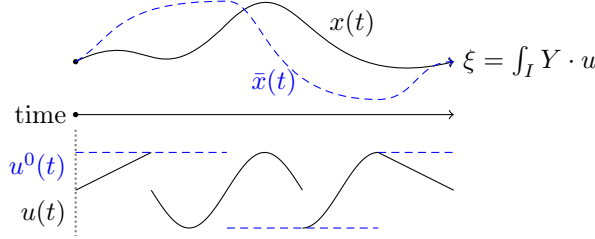


Figure 1: An extremal control u^0 leading to the same ξ as the u control.

Theorem (*The Bang-Bang principle*). *Let $C \subset \mathbb{R}^m$ be a convex compact (of allowed control values). For the system*

$$\begin{aligned} \dot{x} &= A(t)x(t) + B(t)u(t) \\ x(0) &= \vec{0} \end{aligned}$$

every point accessible with some control $u : I \rightarrow C$ in time t is accessible with an extremal control $u^0 : I \rightarrow \text{Ext } C$ as well.

We won't prove it in full generality, but only in the special case when the allowed control values form the unit ball (Holmes) or a *simplex* (Hermes).

Theorem 1 (Holmes). *Let $Y \in L_1^{loc}(I, M_{n \times m})$ have columns Y_i . For each control $u \in \Psi$ where*

$$\Psi := L^\infty(I, \{\|x\|_\infty \leq 1\}) = (\Psi')^m$$

$$\Psi' := L^\infty(I, [-1, 1]),$$

there exists $v \in \Psi$ with $\text{Im } v \subseteq \{-1, 1\}^m$ s.t. $\int_I Y \cdot u = \int_I Y \cdot v$ (here $Y \cdot u$ denotes matrix-vector multiplication).

Proof. We want $\int_I Y \cdot u = \int_I Y \cdot v$, but with $|v_j(t)| \stackrel{\text{a.e.}}{=} 1$. For $j = 1, \dots, m$ we are looking for a partition $I = B_j \cup \overline{B}_j$ with $v_j|_{B_j} \equiv +1$, $v_j|_{\overline{B}_j} \equiv -1$. Equivalently,

$$\begin{aligned} \int_I u_j Y_j &= \int_I v_j Y_j = \int_{B_j} Y_j - \int_{\overline{B}_j} Y_j \\ &= \int_{B_j} Y_j - \left(\int_I Y_j - \int_{B_j} Y_j \right) \\ &= 2 \int_{B_j} Y_j - \int_I Y_j, \end{aligned}$$

that is,

$$\int_{B_j} Y_j = \frac{1}{2} \int_I (u_j + 1) Y_j. \quad (3)$$

We want to prove that to each $u_j \in \Psi'$ there corresponds a measurable $B_j \subseteq I$, satisfying the last equality. To express more clearly the relationship between the given u_j and the B_j looked for, we introduce the mappings, defining their part in the respective sides in the equality.

- For B_j this is the vector measure

$$\vec{\mu}_j(E) := \int_E Y_j, \quad E \in \mathcal{L}_I,$$

where $\mathcal{L}_I = \{E \subset I : E \text{ is measurable}\}$.

- For u_j this is the affine operator $T_j : \Psi' \rightarrow \mathbb{R}^n$,

$$T_j(y) := \frac{1}{2} \int_I (y(t) + 1) d\vec{\mu}_j, \quad y \in \Psi'.$$

T_j is w^* -continuous, since for its linear part

$$L_\infty \ni y \mapsto \int_I y Y_j \leq \|y\|_\infty \vec{\mu}_j(I) \xrightarrow{\|y\| \rightarrow \infty} 0.$$

Thus the equality (3) can be expressed with the formula $\vec{\mu}_j(B_j) = T u_j$, and the statement we want to prove becomes the inclusion

$$T_j \Psi' \subset \vec{\mu}_j(\mathcal{L}_I).$$

By Lyapunov's theorem the RHS $\vec{\mu}_j(\mathcal{L}_I)$ is a convex compact, and since Ψ' is obviously convex and T_j is linear, $T_j \Psi'$ is convex as well. Then it suffices to show $T_j(\text{Ext } \Psi') \subset \vec{\mu}_j(\mathcal{L}_I)$.

Let $u \in \text{Ext } \Psi'$, then $u = \chi_E$ for some $E \subset I$ and

$$\begin{aligned} T_j u &= \frac{1}{2} \int_I (\chi_E + 1) Y_j \\ &= \int_E Y_j + \frac{1}{2} \int_{\bar{E}} Y_j \\ &= \vec{\mu}_j(E) + \frac{1}{2} \vec{\mu}_j(\bar{E}). \end{aligned}$$

Let $\bar{E}' \subset \bar{E}$ have $\vec{\mu}_j(\bar{E}') = \frac{1}{2} \vec{\mu}_j(\bar{E})$. Then

$$T_j u = \vec{\mu}_j(E \cup \bar{E}').$$

□

Definition 1. From here on we fix a closed bounded interval I , the convex compact

$$\mathcal{C} := \{x \in \mathbb{R}_{\geq 0}^m : \sum_i x_i = 1\},$$

as well as a matrix-valued function $Y \in L_1(I, M_{n \times m})$ (in the system above, $Y(s) := \Phi^{-1}(s)B(s)$).

Remark. \mathcal{C} is actually a **simplex**, since for $f_j \in (\mathbb{R}^m)^\star, j = 0, \dots, m$,

$$\begin{aligned} f_0(x_1, \dots, x_m) &= \sum_i x_i \\ f_i(x_1, \dots, x_m) &= x_i, \quad i = 1, \dots, m \end{aligned}$$

we can express $\mathcal{C} = \bigcap_{i=1}^m f_i^{-1}[0, \infty) \cap f_0^{-1}(1)$, that is, \mathcal{C} is an intersection of closed half-spaces.

Since \mathcal{C} is a simplex, $\text{Ext } \mathcal{C}$ consists of m **corners** we denote by r_1, \dots, r_m .

Definition. Just to clarify, I will use the term **non-null set** in the sense of a set of strictly positive measure.

In the proof of the Bang-bang principle later we'll refer to the following separability proposition:

Lemma 1. *Let $E \subset \mathbb{R}$ be a non-null set and $u(t) \in \mathbb{R}^m \setminus \mathcal{C}$ a.e. in E . Then there exists a vector $\eta \in \mathbb{R}^m$, separating strictly \mathcal{C} from $u(t)$ a.e. in a non-null $E' \subset E$, that is,*

$$\exists \varepsilon, B > 0, \eta \in \mathbb{R}^m : \begin{aligned} \forall x \in \mathcal{C} : \quad & \langle \eta, x \rangle \geq B \\ \forall t \in E' : \quad & \langle \eta, u(t) \rangle \leq B - \varepsilon \end{aligned}$$

Proof. According to the previous remark, we can represent with a countable union

$$\begin{aligned} \mathbb{R}^m \setminus \mathcal{C} &= \bigcup_{i=1}^{m+2} g_i^{-1}(-\infty, a_i) \\ &= \bigcup_{i=1}^{m+2} \bigcup_{k \in \mathbb{N}} g_i^{-1}(-\infty, a_i - 1/k], \end{aligned}$$

where $g_i \in (\mathbb{R}^m)^\star, a_i \in \mathbb{R}, i = 1, \dots, m+2$. Then the image under u of some non-null E' must lie in $g_i^{-1}(-\infty, a_i - \frac{1}{k}]$ for some $k \in \mathbb{N}$. Let $\eta \in \mathbb{R}^m$ be the respective (by Riesz' theorem) vector to the functional g_i (for which $\langle \eta, \cdot \rangle = g_i(\cdot)$). Then

$$\begin{aligned} \forall x \in \mathcal{C} : \quad & \langle \eta, x \rangle = g_i(x) \geq a_i \\ \forall t \in E' : \quad & \langle \eta, u(t) \rangle = g_i(u(t)) \leq a_i - \frac{1}{k}. \end{aligned}$$

□

Definition. We denote the sets of the *allowed* and the *extremal* controls, respectively:

$$\begin{aligned} \Psi &:= L_\infty(I, \mathcal{C}) \subset L_\infty(I, \mathbb{R}^m) \\ \Psi_e &:= L_\infty(I, \text{Ext } \mathcal{C}) \subseteq \Psi. \end{aligned}$$

Remark. From here on we'll talk about the space $L_\infty(I, \mathbb{R}^m)$, which, taken as the dual of $L_1(I, \mathbb{R}^m)$ by the action

$$f(\varphi) := \int \langle \varphi(t), f(t) \rangle dt, \quad \text{for all } \varphi \in L_1(I, \mathbb{R}^m), f \in L_\infty(I, \mathbb{R}^m),$$

is equipped with the w^\star topology.

Lemma 2. $\Psi := L_\infty(I, \mathcal{C})$ is a convex w^* -compact. It is essential that here \mathcal{C} is the simplex defined earlier.

Proof. Obviously Ψ is convex, since for $u_1, u_2 \in \Psi$, $\lambda u_1 + (1 - \lambda)u_2$ is again a L_∞ function, and by that again \mathcal{C} -valued due to convexity of \mathcal{C} . Moreover, Ψ is bounded (in norm) with

$$\|u\|_\infty = \sup_{t \in I} u(t) \leq \max_{y \in \mathcal{C}} \|y\| \quad \text{for all } u \in \Psi.$$

For Ψ to be a w^* -compact it remains only to check it is w^* -closed, that is¹, any $u^0 \in L_\infty(I, \mathbb{R}^m) \setminus \Psi$ is w^* -separable from Ψ . In order to w^* -separate u^0 from Ψ , we need a suitable functional $\varphi \in L_1 \hookrightarrow (L_\infty)^*$, with which

$$\varphi(u^0) \leq B - \varepsilon \quad \text{and} \quad \forall u \in \Psi : \varphi(u) \geq B.$$

Since $u^0 \notin \Psi$, then in a non-null $E \subseteq I$, $u^0(t) \notin \mathcal{C}$. Then by lemma 1 there exist $\varepsilon > 0, B > 0, \eta \in \mathbb{R}^m$ and a non-null $E' \subset E$ with the property

$$\begin{aligned} \langle \eta, x \rangle &\geq B && \text{for all } x \in \mathcal{C} \\ \langle \eta, u^0(t) \rangle &\leq B - \varepsilon && \text{for all } t \in E' \end{aligned}$$

We set $\varphi(t) = \begin{cases} \frac{1}{\mu(E')} \eta, & t \in E' \\ 0, & \text{else} \end{cases}$, and then $\varphi \in L_1$ and fulfills the required conditions:

$$\begin{aligned} \varphi(u^0) &= \int_I \langle \varphi(t), u^0(t) \rangle dt = \frac{1}{\mu(E')} \int_{E'} \langle \eta, u^0(t) \rangle dt \\ &\leq \frac{1}{\mu(E')} \int_{E'} (B - \varepsilon) = B - \varepsilon \end{aligned}$$

and analogously $\varphi(u) = \int_I \langle \varphi, u \rangle = \frac{1}{\mu(E')} \int_{E'} \langle \eta, u \rangle \geq B$ for all $u \in \Psi$.

Since Ψ is w^* -separated by any $u^0 \notin \Psi$, then Ψ is w^* -closed, and since it is $\|\cdot\|_\infty$ -bounded – it is a w^* -compact. \square

Remark. For $u \in L_\infty(I, \mathbb{R}^m)$, the product $Y \cdot u$ is a function from $L_1(I, \mathbb{R}^m)$, because $Y_{ij} \in L_1$ and

$$(Y \cdot u)_i = \sum_{j=1}^m Y_{ij}(s) u_j(s) \leq \sum_{j=1}^m Y_{ij}(s) \|u\|_\infty.$$

Then it makes sense to talk about the integral $\int_I Y \cdot u$ that we will look at from here on.

Definition. We say that the control $u \in \Psi$ leads to the point $\xi = Tu$, where $T : L_\infty(I, \mathbb{R}^m) \rightarrow \mathbb{R}^n$ is the linear operator

$$Tu := \int_I Y(s) \cdot u(s) ds \quad \text{for all } u \in L_\infty(I, \mathbb{R}^m).$$

¹In Hermes's book the existence of a $u^0 \in \partial\Psi \setminus \Psi$ is assumed, and a contradiction is derived later with the same argument. I find this style unnecessarily complicated, so I just prove openness of the complement.

Remark. T is continuous (in norm), since for $\varepsilon > 0$, with $\delta := \frac{\varepsilon}{\mu(I) \int_I Y}$ and $\|h\|_\infty < \delta$ we have

$$|T(u+h) - Tu| = \left| \int_I Y(s) \cdot h(s) \, ds \right| \leq \mu(I) \|h\|_\infty \left| \int_I Y \right| < \varepsilon.$$

Moreover, it is w^* -continuous, since its coordinate functionals

$$T_i : L_\infty(I, \mathbb{R}^m) \rightarrow \mathbb{R}, i = 1, \dots, n$$

are defined by $T_i(u) = \sum_j \int_I Y_{ij} u_j$ where Y_{ij} are L_1 -functions.

Theorem (*Bang-bang* principle for controls in a simplex, Hermes).

$$T(\Psi) = T(\Psi_e).$$

Moreover, the two sets are convex compacts.

Proof. $T\Psi_e \subseteq T\Psi$ is obvious, we need to prove the reverse inclusion. The proof goes in the following steps:

1. According to the last remark, T is w^* -continuous, hence the image $T\Psi \subset \mathbb{R}^n$ of the w^* -compact Ψ is again a convex compact (the second statement in this theorem).
2. T is a continuous map between compacts, then the possible controls $T^{-1}(\xi) \cap \Psi$ leading to any point $\xi \in \mathbb{R}^n$ form a convex w^* -compact, necessarily having an extreme point $u \in T^{-1}(\xi) \subset \Psi$ (Krein-Milman).
3. It turns out that this extremum u is a *bang-bang* control.

If we assume the opposite, then in some non-null set $E \subset I$ the controls are far from $\text{Ext } \mathcal{C}$ and from the lemma (3) below it follows u is a midpoint of a proper line segment of controls $[u-h, u+h] \subset \Psi$, leading to the same ξ , contradicting extremality of u .

□

Lemma 3. Let $n \in \mathbb{N}$, $Y \in L_1(I, M_{n \times m})$, $E \subset I$ is non-null, and the control $u \in \Psi$ leads to $\xi := Tu$ and satisfies $\text{dist}(u(E), \text{Ext } \mathcal{C}) > \varepsilon$ for some $\varepsilon > 0$.

Then there exists a displacement $h \in L_\infty(I, \mathbb{R}^m)$ with the following properties:

1. $u(t) \pm h(t) \in \mathcal{C}$ for all $t \in I$, where $h(t) \neq 0$ for all $t \in E$ (u lies in a proper line segment);
2. $\int_I Y \cdot h = 0$ (that is, $u \pm h$ leads again to ξ).

In other words, $u \in [u-h, u+h] \subset T^{-1}(\xi) \cap \Psi$.

Proof. The proof goes in the following steps:

1. For each $t \in E$, according to the next lemma (4), we can embed $u(t)$ in a segment with endpoints $u(t) \pm \tilde{h}(t) \in \mathcal{C}$ so that $\|\tilde{h}(t)\|_\infty > \varepsilon^2$. Then $\|\tilde{h}\|_\infty > \varepsilon^2$ too. Unfortunately, however, $\int_I Y \cdot \tilde{h} \neq 0$ in the general case.

2. From the non-atomicity of the Lebesgue measure we can partition for any $k \in \mathbb{N}^+$ (the suitable k will be determined later) $E = \bigcup_{j=1}^k E_k$, so that $E_i \cap E_j = \emptyset$ for all $i \neq j$ and $\mu(E_j) \geq 0$ for all j .
3. Now in each $E_j, j = 1, \dots, k$ we will weight the restriction of \tilde{h} with a suitable $\alpha_j \in \mathbb{R}$ so that $\int_I Y \cdot h = 0$:

$$h(t) := \begin{cases} \alpha_j \tilde{h}(t), & t \in E_j \\ 0, & \text{otherwise} \end{cases}.$$

The condition $\int_I Y \cdot h = 0$ is a homogenous linear system with n equations and k unknowns

$$\int_I Y \cdot h = \sum_{j=1}^k \alpha_j \int_{E_j} Y \cdot \tilde{h} = 0.$$

A non-trivial solution for the coefficients α_j always exists whenever $k \geq n + 1$ – fix one with $|\alpha_j| < 1$ for all j .

4. Let us convince ourselves that h satisfies property 1 (from the lemma statement). Obviously $h(t) \neq 0$ for all $t \in E$, and $u(t) \pm h(t)$ lies in \mathcal{C} , since it lies (for all $t \in E$) in the line segment with endpoints $u(t) \pm \tilde{h}(t) \in \mathcal{C}$ (due to $|\alpha_j| < 1$), and \mathcal{C} is convex.

Therefore $u \pm h \in \Psi = L_\infty(I, \mathcal{C})$ and $u \in [u - h, u + h]$ from convexity of Ψ .

□

Lemma 4. *Let $\xi \in \mathcal{C}$ satisfy $\text{dist}(\xi, \text{Ext } \mathcal{C}) > \varepsilon$ for some $\varepsilon > 0$. Then ξ can be put in a line segment $[\xi - h, \xi + h] \subset \mathcal{C}$ with $\|h\|_\infty \geq \varepsilon^2$.*

Moreover, the mapping $\xi \mapsto h$ is measurable.

(interestingly, we cannot state $\|h\|_\infty \geq \varepsilon$, the largest lower bound is $\frac{\varepsilon^2}{1-\varepsilon}$)

Proof. Geometrically it is natural to choose h parallel to $r_j - \xi$, where $r_j \in \text{Ext } \mathcal{C}$ is the closest to ξ corner with smallest index j . We put $h := \beta(r_j - \xi)$. Clearly $\xi \pm h \in \text{affine span } \mathcal{C}$. We will show that for sufficiently small $\beta > 0$ even $\xi \pm h \in \mathcal{C}$.

$$\xi \pm h = \xi \pm \beta(r_j - \xi) = (1 \mp \beta)\xi \pm \beta r_j.$$

Since $\xi \in \mathcal{C}$, then $\xi = \sum_i \beta_i r_i$ with some $\beta_i \geq 0$ and $\sum_i \beta_i = 1$. Then

$$\begin{aligned} \xi \pm h &= (1 \mp \beta) \sum_i \beta_i r_i \pm \beta r_j \\ &= \underbrace{(\pm\beta + (1 \mp \beta)\beta_j)}_{\beta'_j} r_j + \sum_{i \neq j} \underbrace{(1 \mp \beta)\beta_i}_{\beta'_i} r_i \end{aligned}$$

and we have to check that the RHS is a convex combination of $\{r_i\}_{i=1}^m$. $\sum_i \beta'_i = 1$ is equivalent to $\xi \pm h \in \text{affine span } \mathcal{C}$, so it remains to check $0 < \beta'_i$ for all i .

If $|\beta| \leq 1$ then obviously $\beta'_i \geq 0$ for all $i \neq j$, and

$$\beta'_j = \pm(1 - \beta_j)\beta + \beta_j$$

and hence $\beta'_j \geq 0$ holds always if $\beta \leq \frac{\beta_j}{1-\beta_j}$. We choose namely the maximal $\beta = \beta_j/(1-\beta_j)$.

Since $\text{dist}(\xi, \text{Ext } \mathcal{C}) > \varepsilon$ and r_j is a closest corner to ξ , then $\beta_j > \varepsilon$, $\beta > \frac{\varepsilon}{1-\varepsilon}$ and

$$\|h\|_\infty = \beta \|r_j - \xi\|_\infty > \beta \varepsilon = \frac{\varepsilon^2}{1-\varepsilon} > \varepsilon^2.$$

For measurability of the mapping $\xi \mapsto h$ it suffices to observe that $\xi \mapsto j$ is constant on sets

$$F_j = \left\{ \xi \in \mathcal{C} : \begin{array}{l} \text{d}(\xi, r_j) = \text{dist}(\xi, \text{Ext } \mathcal{C}) \\ \text{and } \text{d}(\xi, r_j) > \text{d}(\xi, r_i) \text{ for all } i < j \end{array} \right\}, \quad j = 1, \dots, m,$$

which partition \mathcal{C} and are measurable due to continuity of the metric $\text{d}(\cdot, \cdot)$. \square

Sources:

- Functional analysis and time optimal control, Volume 56 (Mathematics in Science and Engineering) by Hermes (Editor)
- Geometric Functional Analysis and its Applications Richard B. Holmes