

Cálculo de cotas para el tiempo de permanencia en politopos para sistemas no lineales

Alexis J. Vallarella¹, Felipe Cinto¹ y Hernan Haimovich¹

¹Centro Internacional Franco-Argentino de Ciencias de la Información y de Sistemas (CIFASIS), CONICET-UNR, Rosario, Argentina,
{vallarella, cinto, haimovich}@cifasis-conicet.gov.ar

Resumen Diversas estrategias de control requieren determinar el tiempo durante el cual el estado de un sistema permanece dentro de una región considerada. Esto ocurre a menudo en estrategias de control por conmutación, donde se desea seleccionar un modo actual que asegure que el estado permanezca durante mayor tiempo posible dentro de la región. En general, las ecuaciones de los sistemas no lineales no tienen solución analítica y sólo se dispone de una solución aproximada, usualmente derivada mediante métodos de integración numérica. En este trabajo se presenta un método para garantizar que, a partir del conocimiento de soluciones aproximadas, la solución exacta permanezca dentro de un conjunto convexo deseado a lo largo de un intervalo de tiempo. Mostramos que, cuando el conjunto en cuestión es politópico, los cálculos requeridos se simplifican. El método se aplica a un sistema no lineal para ilustrar su funcionamiento.

Palabras clave: Tiempo de permanencia, sistemas commutados, simulación.

Computation of bounds for permanence time in polytopes for nonlinear systems

Abstract Diverse control strategies require determining how long the state of a system remains within a considered region. This often arises in switching control strategies, where the goal is to select the current mode to ensure that the state stays inside the region for as long as possible. In general, nonlinear system equations do not admit analytical closed-form solutions; thus, only approximate solutions are available, usually obtained through numerical integration methods. This work presents a method to guarantee that, based on approximate solutions, the exact solution remains within a desired convex set over a time interval. We show that when the set in question is polytopic, the required computations become simpler. The method is applied to a nonlinear system to illustrate its operation.

Keywords: Permanence time, switched systems, simulation.

1 Introducción

Un objetivo primordial en los sistemas de control es mantener la evolución del estado del sistema dentro de una región deseada. Las estrategias de control que persiguen este objetivo pueden requerir predecir la evolución del estado futuro para determinar qué acción de control aplicar en el presente instante. La determinación del período de tiempo máximo durante el cual la trayectoria permanece dentro de una región deseada es un problema de gran relevancia práctica. Éste cobra especial interés en sistemas controlados mediante leyes de conmutación, para los que se tienen restricciones de tiempo mínimo de permanencia para cada modo (Liberzon, 2003) y en los cuales puede no estar permitido comutar de un modo a otro antes de que un período mínimo de tiempo haya transcurrido.

La predicción de la evolución del estado a menudo se basa en métodos de integración numérica o en soluciones aproximadas que, inevitablemente, incurren en algún grado de error y que por lo general sólo permiten calcular la trayectoria en instantes discretos de tiempo (ver Griffiths and Higham, 2010; Grüne and Pannek, 2018; LeVeque, 2007 y las referencias allí citadas). El hecho de que la trayectoria permanezca dentro de una región deseada en ciertos instantes discretos de tiempo es una condición necesaria pero no suficiente para garantizar que lo haga en todos los tiempos intermedios. La determinación del período de tiempo en que la trayectoria permanece en la región deseada no es trivial.

Diferente del objetivo de encontrar una región invariante para sistemas conmutados (Julius & van der Schaft, 2002; Perez et al., 2022, 2023) en algunos sistemas controlados por conmutación es de interés determinar si una solución permanece dentro de una región durante el denominado tiempo de permanencia. Existen métodos que requieren calcular la solución en un instante futuro para determinar la ley de control en el instante presente según diversos objetivos (Russo et al., 2022). Por ejemplo, el control por conmutación puede requerir seleccionar el modo que haga que el estado permanezca dentro de un conjunto durante el mayor tiempo posible. Tal objetivo requiere algún método para predecir o aproximar la trayectoria futura del estado, ya que la solución exacta usualmente no se conoce o para calcularla se incurre en errores computacionales. Las soluciones aproximadas obtenidas mediante métodos de integración numérica consisten en una secuencia de valores discretos con garantías numéricas dadas habitualmente como cotas de incertidumbre alrededor de cada punto calculado. Para un método dado, la cota puede hacerse más pequeña reduciendo el paso de integración, es decir, disminuyendo el tiempo entre los instantes consecutivos. El hecho de que una solución aproximada dada, contemplando sus cotas de incertidumbre, esté completamente contenida en una región no implica necesariamente que el estado verdadero haya permanecido dentro de la región en todos los tiempos intermedios entre los instantes discretos de cálculo, como se mostrará en el ejemplo de la sección 2.2.

En este contexto, la principal contribución de este trabajo es desarrollar un método *ad-hoc* para proporcionar tiempos de permanencia con garantías del estado dentro de una región convexa dada. El método utiliza propiedades de convexidad de la aproximación de Euler para asegurar que la solución permanece

en la región en cuestión. Al mismo tiempo permite el uso de métodos de orden superior más precisos para el cálculo del avance de la solución aproximada. En el caso en que la región convexa sea un politopo, la verificación de que la solución aproximada, con su correspondiente error, se encuentra en dicha región se vuelve sencilla y completamente computable. Ilustramos la aplicación de los resultados sobre un modelo no lineal del tipo Lotka-Volterra, donde aseguramos que el estado permanece dentro de una región politópica dada.

Notación. \mathbb{R} , $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$, \mathbb{N} y \mathbb{N}_0 denotan, respectivamente, los números reales, reales no negativos, reales positivos, naturales y naturales no negativos. $\bar{1}$ denota un vector con componentes todas iguales a 1. Para $X, Y \in \mathbb{R}^{r \times n}$, $X \preceq Y$ denota el conjunto de desigualdades $X_{i,j} \leq Y_{i,j}$, para $i = 1, \dots, r$, $j = 1, \dots, n$. La norma euclídea e infinito se denotan como $|\cdot|$ y $\|\cdot\|_\infty$, respectivamente. Dada $A \in \mathbb{R}^{n \times m}$, $|A|$ denota la norma matricial inducida por la norma euclídea. Para $x \in \mathbb{R}^n$, x^\top representa su transpuesta. Dado $x \in \mathbb{R}^n$ y $\varepsilon > 0$, definimos la bola cerrada $\mathbb{B}_\varepsilon(x) := \{y \in \mathbb{R}^n : |x - y| \leq \varepsilon\}$.

2 Formulación del Problema y Motivación

2.1 Formulación del Problema

Consideramos sistemas no lineales de tiempo continuo de la forma

$$\dot{x} = f(x), \quad (1)$$

donde $x \in \mathbb{R}^n$ y $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisface la siguiente hipótesis de continuidad de Lipschitz dentro de un conjunto convexo dado \mathcal{C} .

Suposición 1 Existe $L \geq 0$ tal que, para todo $x, y \in \mathcal{C}$,

$$|f(x) - f(y)| \leq L|x - y|. \quad (2)$$

La Suposición 1 garantiza la existencia y unicidad de las soluciones mientras las mismas se encuentren dentro del conjunto \mathcal{C} . Una solución de (1) con condición inicial $x_0 \in \mathbb{R}^n$ se denota como $\phi(t, x_0)$, y mientras exista satisface

$$\phi(t, x_0) := x_0 + \int_0^t f(\phi(s, x_0)) ds. \quad (3)$$

Dado que (1) generalmente no posee una solución analítica explícita, determinar si una solución permanece dentro de \mathcal{C} durante un intervalo de tiempo requiere considerar aproximaciones numéricas de la solución. Además, incluso en el caso de existir una solución explícita, la verificación de que la solución continua permanece en \mathcal{C} implica evaluarla en todo el intervalo, lo cual resulta impráctico desde el punto de vista computacional.

El problema a abordar consiste entonces en determinar si es posible garantizar que una solución permanece dentro de un conjunto convexo \mathcal{C} durante un intervalo de tiempo evaluando únicamente una aproximación de la misma en un conjunto finito de instantes de tiempo.

2.2 Limitaciones de las Aproximaciones Numéricas

Los métodos de integración numérica, como los métodos de Runge-Kutta explícitos e implícitos, aproximan soluciones de ecuaciones diferenciales con cotas de error garantizadas, usualmente expresadas en términos de la diferencia entre la solución exacta y la aproximada en instantes de tiempo específicos (Nešić et al., 1999; Vallarella et al., 2021). Aunque dichas cotas pueden utilizarse para verificar si una solución está dentro de \mathcal{C} en dichos instantes de tiempo, garantizar la permanencia dentro del conjunto para todos los tiempos intermedios no es sencillo, considerando que sólo se desea imponer la Suposición 1 sobre el sistema.

Una propiedad clave de los métodos numéricos mencionados es que el error producido por la discretización en cada paso puede reducirse disminuyendo el tamaño del paso, asegurando así que las soluciones aproximadas y exactas permanezcan cercanas (Cellier and Kofman, 2006; Stuart and Humphries, 1998). Por lo tanto podría esperarse que una respuesta afirmativa al problema presentado anteriormente se obtenga simplemente disminuyendo suficientemente el tamaño del paso. Sin embargo, como se ilustra en el siguiente ejemplo, incluso utilizando aproximaciones numéricas arbitrariamente precisas —ya sea refinando el paso o utilizando métodos de orden superior— existen casos en los que la solución aproximada no es suficiente para determinar si la solución exacta permanece dentro de la región convexa \mathcal{C} .

Ejemplo 1. Consideremos el sistema lineal

$$\dot{x}(t) = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_A x(t), \quad x_0 := x(0) \quad (4)$$

y el conjunto convexo $\mathcal{C} := \{x \in \mathbb{R}^2 : \|x\|_\infty \leq 1\}$, que define un cuadrado centrado en el origen. Para cualquier condición inicial $x_0 \in \mathcal{C}$, la solución exacta de (4) está dada por

$$\phi(t, x_0) = \begin{bmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{bmatrix} x_0, \quad (5)$$

la cual describe trayectorias circulares en sentido horario con radio $|x_0|$ y centradas en el origen. Por lo tanto toda solución que parte de una condición inicial $x_0 \in \mathcal{C}$ con $|x_0| > 1$ saldrá de \mathcal{C} , aunque volverá a ingresar luego, como se ilustra con la línea roja en la Figura 1.

Para este caso lineal sencillo, una solución aproximada de orden $N \in \mathbb{N}$ puede obtenerse truncando la expansión en serie de Taylor de la solución exacta como

$$\hat{\phi}_N(t, x) = \sum_{k=0}^N \frac{(At)^k}{k!} x. \quad (6)$$

Definimos con $h > 0$ al paso utilizado para la aproximación. A continuación mostraremos que incluso si la precisión de la solución aproximada $\hat{\phi}_N(h, x_0)$ es

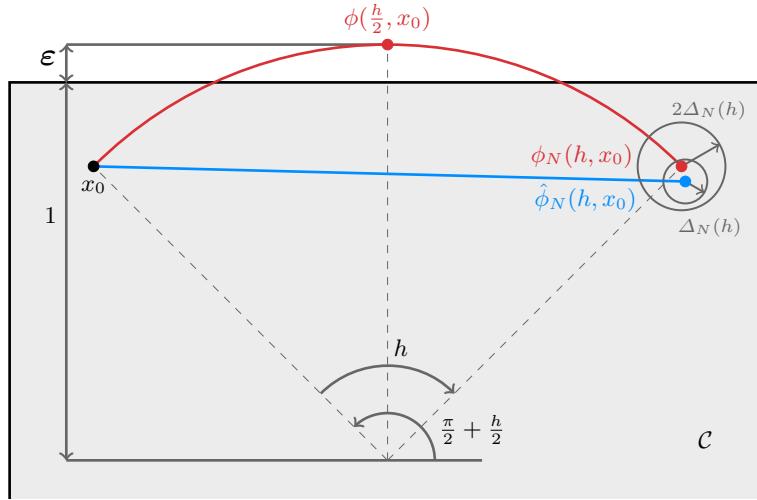


Fig. 1: Solución exacta del sistema (4) (línea roja) y la solución aproximada (6) en los instantes $t = 0$ y $t = h > 0$ (conectados por una línea azul). Ambas parten desde la condición inicial $x_0 = (1 + \varepsilon) \left[\cos\left(\frac{\pi}{2} + \frac{h}{2}\right) \sin\left(\frac{\pi}{2} + \frac{h}{2}\right) \right]^\top$.

suficientemente alta como para garantizar que la solución exacta en el tiempo $t = h$ se encuentra dentro de \mathcal{C} , es decir, $\phi(h, x_0) \in \mathcal{C}$, esto aún puede no ser suficiente para asegurar que la solución exacta permanece dentro de \mathcal{C} en puntos intermedios, por ejemplo $t = h/2$.

La solución aproximada (6) satisface la cota de error

$$|\phi(h, x) - \hat{\phi}_N(h, x)| \leq \frac{(|A|h)^{N+1}}{(N+1)!} e^{|A|h} = \frac{h^{N+1}}{(N+1)!} e^h =: \Delta_N(h), \quad (7)$$

donde se usó el hecho de que $|A| = 1$ para la matriz A en (4).

Definimos $\varepsilon := \frac{1-2\Delta_N(h)-\cos(\frac{h}{2})}{\cos(\frac{h}{2})}$. Para $N \geq 2$, puede demostrarse que $\frac{d\varepsilon}{dh}|_{h=0} = 0$ y que $\frac{d^2\varepsilon}{dh^2}|_{h=0} > 0$, por lo que para $h > 0$ suficientemente pequeño, se tiene $\varepsilon > 0$. Para la condición inicial

$$x_0 = (1 + \varepsilon) \left[\cos\left(\frac{\pi}{2} + \frac{h}{2}\right) \sin\left(\frac{\pi}{2} + \frac{h}{2}\right) \right]^\top, \quad (8)$$

la solución exacta satisface, considerando (5), lo siguiente

$$\begin{aligned} \phi(h, x_0) &= (1 + \varepsilon) \left[\sin\left(\frac{h}{2}\right) \cos\left(\frac{h}{2}\right) \right]^\top \\ &= (1 - 2\Delta_N(h)) \left[\tan\left(\frac{h}{2}\right) 1 \right]^\top \preceq \bar{1}. \end{aligned} \quad (9)$$

Por lo tanto, $\phi(h, x_0) \in \mathcal{C}$.

Además a partir de (7) y (9), y debido a la cota $|\hat{\phi}_N(h, x_0)| \leq |\phi(h, x_0)| + \Delta_N(h)$, la solución aproximada también satisface $\hat{\phi}_N(h, x_0) \in \mathcal{C}$. A su vez la bola

de radio $\Delta_N(h)$ centrada en la solución aproximada $\hat{\phi}_N(h, x_0)$ está contenida en \mathcal{C} . Sin embargo, puede notarse que $\phi\left(\frac{h}{2}, x_0\right) = [0 \ 1 + \varepsilon]^\top \notin \mathcal{C}$, como se ilustra en la Figura 1. •

El Ejemplo 1 muestra que, independientemente del orden de aproximación $N \geq 2$, pueden existir casos en los que no se puede garantizar que la solución exacta permanezca dentro de \mathcal{C} en todo el intervalo de tiempo $t \in [0, h]$ basándose únicamente en la solución aproximada, ya que la solución exacta puede salir brevemente y luego reingresar. Un método numérico de alta precisión debe captar esta breve ventana de salida y reentrada. Si los instantes de tiempo seleccionados no caen dentro de este intervalo, el método podría sugerir erróneamente que la trayectoria permanece dentro de \mathcal{C} , conduciendo a un cálculo incorrecto del tiempo de permanencia.

En la siguiente sección, desarrollamos un método para determinar un tiempo de permanencia garantizado aprovechando propiedades de convexidad.

3 Resultados principales

Los resultados principales se presentan en dos subsecciones. En la Sección 3.1 se brindan resultados teóricos que permiten reducir la complejidad de determinar si la solución exacta permanece dentro de una región convexa. Luego, éstos se emplean en la Sección 3.2, donde se presenta un procedimiento que permite calcular una cota inferior del tiempo de permanencia con garantías.

3.1 Propiedades de convexidad

La aproximación de Euler de la solución del sistema (1) partiendo de un estado $x \in \mathbb{R}^n$ y luego de un paso de longitud $h \in \mathbb{R}_{\geq 0}$, se define como

$$\hat{\phi}^E(h, x) := x + hf(x). \quad (10)$$

Para un $x \in \mathbb{R}^n$ fijo, la función $\hat{\phi}^E(h, x) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^n$ satisface la siguiente propiedad de convexidad, como puede comprobarse fácilmente reemplazando (10) en (11).

Propiedad 1. Para todo $h \geq 0$ y $x \in \mathcal{C}$ se cumple que

$$\hat{\phi}^E(\alpha h, x) = (1 - \alpha)\hat{\phi}^E(0, x) + \alpha\hat{\phi}^E(h, x), \quad \forall \alpha \in [0, 1]. \quad (11)$$

Es decir, las soluciones para tiempos intermedios de la aproximación de Euler $\hat{\phi}(\tau, x)$ con $\tau \in [0, h]$ pertenecen al segmento de recta que une las soluciones en los instantes extremos $t = 0$ y $t = h$.

La continuidad de Lipschitz de la Suposición 1 es suficiente para garantizar la existencia de una cota para el error local de la aproximación de Euler tras un paso de longitud h . Consideremos la solución exacta de (1), $\phi(h, x)$, y la solución obtenida mediante el método de Euler, $\hat{\phi}^E(h, y)$, que parten de condiciones iniciales $x, y \in \mathbb{R}^n$, respectivamente. Aplicando la desigualdad triangular,

el error entre ambas soluciones puede descomponerse en dos términos: el primero, $|\phi(h, x) - \phi(h, y)|$, corresponde a la diferencia entre soluciones exactas que parten de condiciones iniciales distintas; el segundo, $|\phi(h, y) - \hat{\phi}^E(h, y)|$, representa la diferencia entre la solución exacta y la de Euler con la misma condición inicial. Acotando ambos términos por separado mediante la Suposición 1 y la desigualdad de Gronwall, se obtiene una cota explícita del error en un paso dada por

$$|\phi(h, x) - \hat{\phi}^E(h, y)| \leq e^{Lh} \left(r + \frac{Lh^2}{2} |f(y)| \right) =: \Delta^E(r, y, h), \quad (12)$$

donde el parámetro r representa una cota superior de la diferencia entre las condiciones iniciales x de la solución exacta e y de la aproximación, i.e. $r \geq |x - y|$. Para r e y dados, la función $\Delta^E(r, y, h)$ resulta convexa en la variable h .

Considerando la propiedad de convexidad dada en (11) y la cota en un paso (12) presentamos a continuación un teorema que asegura que la solución exacta pertenece, a lo largo de todo el intervalo de tiempo continuo $t \in [0, h]$, a una región convexa específica determinada por la solución de Euler y la cota (12).

Teorema 1. *Defina $\mathcal{B}_{r,y}^E(h) := \mathbb{B}_{\Delta^E(r,y,h)}(\hat{\phi}^E(h, y))$. Sea $\mathcal{C} \subseteq \mathbb{R}^n$ un conjunto convexo y sea $h \geq 0$. Suponga $x, y \in \mathcal{C}$ tales que*

- i) $\phi(0, x) \in \mathcal{B}_{r,y}^E(0) \subseteq \mathcal{C}$,
- ii) $\mathcal{B}_{r,y}^E(h) \subseteq \mathcal{C}$.

donde $r \geq |x - y|$. Entonces,

$$\phi(t, x) \in \text{conv}(\mathcal{B}_{r,y}^E(0) \cup \mathcal{B}_{r,y}^E(h)) \subseteq \mathcal{C}. \quad (13)$$

para todo $t \in [0, h]$.

Demostración. Para simplificar la notación, dados $x, y \in \mathcal{C}$ y $r := |x - y|$, definimos

$$\Delta(h) := \Delta^E(r, y, h) \quad (14)$$

$$\mathcal{B}(\alpha) := \mathcal{B}_{r,y}^E(\alpha h) = \mathbb{B}_{\Delta^E(r,y,\alpha h)}(\hat{\phi}^E(\alpha h, y)) \quad \forall \alpha \in [0, 1]. \quad (15)$$

Dado que \mathcal{C} es convexo, por los ítems i) y ii) se tiene que $\text{conv}(\mathcal{B}(0) \cup \mathcal{B}(1)) \subseteq \mathcal{C}$. Además, por la cota (12), la solución $\phi(\alpha h, x)$ satisface $\phi(\alpha h, x) \in \mathcal{B}(\alpha)$ para todo $\alpha \in [0, 1]$. En consecuencia, existe un vector $d \in \mathbb{R}^n$ dentro de la bola de radio uno $\mathbb{B}_1(0)$ tal que $\phi(\alpha h, x)$ puede escribirse como

$$\begin{aligned} \phi(\alpha h, x) &= \hat{\phi}^E(\alpha h, y) + \Delta(\alpha h)d \\ &= (1 - \alpha)\hat{\phi}^E(0, y) + \alpha\hat{\phi}^E(h, y) + \Delta(\alpha h)d, \end{aligned} \quad (16)$$

donde se utilizó la Propiedad 1. Definimos $d_\alpha := d \frac{\Delta(\alpha h)}{(1 - \alpha)\Delta(0) + \alpha\Delta(h)}$. Como Δ es convexa, se tiene que $|d_\alpha| = |d| \frac{\Delta(\alpha h)}{(1 - \alpha)\Delta(0) + \alpha\Delta(h)} \leq |d| \leq 1$. A partir de (16) y la definición de d_α , obtenemos

$$\phi(\alpha h, x) = (1 - \alpha)(\hat{\phi}^E(0, y) + \Delta(0)d_\alpha) + \alpha \left(\hat{\phi}^E(h, y) + \Delta(h)d_\alpha \right).$$

Por lo tanto, $\phi(\alpha h, x)$ puede expresarse como una combinación convexa entre un punto en $\mathcal{B}(0)$ y un punto en $\mathcal{B}(1)$. En consecuencia, $\phi(\alpha h, x) \in \text{conv}(\mathcal{B}(0) \cup \mathcal{B}(1)) \subseteq \mathcal{C}$ para todo $\alpha \in [0, 1]$, lo que establece el resultado. \square

El Teorema 1 muestra que si una bola de radio $r \geq |x - y|$ centrada en $\hat{\phi}^E(0, y) = y$ está contenida en \mathcal{C} , entonces sólo se requiere que una bola de radio igual a $\Delta^E(r, y, h)$ centrada en $\hat{\phi}^E(h, y)$ esté contenida en \mathcal{C} para garantizar que la solución exacta $\phi(t, x)$ permanece dentro de la región \mathcal{C} durante todo el intervalo. La Figura 2 ilustra la idea detrás del resultado.

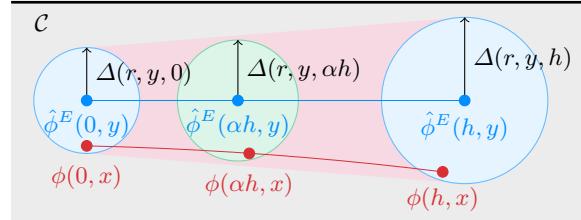


Fig. 2: Ilustración del Teorema 1. Dado que $\phi(0, x) = x$ se encuentra dentro de una bola de radio $\Delta^E(r, y, 0) = r \geq |x - y|$ centrada en $\hat{\phi}^E(0, y) = y$, la cual a su vez está contenida en \mathcal{C} , y que la bola de radio $\Delta(r, y, h)$ centrada en $\hat{\phi}^E(h, y)$ permanece dentro del conjunto convexo \mathcal{C} , se garantiza que $\phi(t, x)$ (línea roja) permanece dentro de la envoltura convexa de dichas bolas a lo largo de $[0, h]$.

3.2 Cómputo del tiempo de permanencia

Retomamos el problema planteado al final de la Sección 2.1, suponiendo que la región \mathcal{C} es cerrada y convexa. Consideraremos datos una condición inicial $x_0 \in \mathcal{C}$, un paso de tiempo mínimo $h_{\min} > 0$ y un tiempo máximo de cómputo $T_{\max} \geq h_{\min}$. El problema específico a resolver es el siguiente:

Problema 1. Determinar computacionalmente un tiempo $T \in [h_{\min}, T_{\max}]$ tal que la solución exacta de (1) satisfaga $\phi(t, x_0) \in \mathcal{C}$ para todo $t \in [0, T]$, evaluando soluciones aproximadas en instantes separados por al menos $h_{\min} > 0$.

Como solución a este problema, proponemos el Algoritmo 1, donde asumimos la Suposición 1 para un conjunto cerrado y convexo $\mathcal{C} \subset \mathbb{R}^n$ y, dados $T_{\max} \geq h_{\min} > 0$, utilizamos la definición $\mathcal{B}_{r,y}^E(h) := \mathbb{B}_{\Delta^E(r,y,h)}(\hat{\phi}^E(h, x))$ del Teorema 1 y la siguiente definición

$$\bar{h}_k(T) := \max\{h \in [h_{\min}, T_{\max} - T] : \mathcal{B}_{r_k, y_k}^E(h) \subseteq \mathcal{C}\}. \quad (17)$$

La lógica del algoritmo se ilustra en la Fig. 3. Si el Algoritmo 1 retorna $T > 0$, entonces T es solución del problema planteado. El algoritmo requiere la existencia de una aproximación del sistema no lineal $\hat{\phi}(\cdot, y)$ y una cota para el error en un paso $\Delta_{r,y}$ asociada que cumpla

$$|\phi(t, x) - \hat{\phi}(t, y)| \leq \Delta_{r,y}(t). \quad (18)$$

Algoritmo 1: Tiempo de permanencia garantizado**Entradas:** \mathcal{C} , f , x_0 , h_{\min} , and T_{\max}

```

1  $k \leftarrow 0$ 
2  $T \leftarrow 0$ 
3  $y_0 \leftarrow x_0$ 
4  $r_0 \leftarrow |x_0 - y_0|$ 
5 while  $T_{\max} - T \geq h_{\min}$  and  $\mathcal{B}_{r_k, y_k}^E(0) \subseteq \mathcal{C}$  do
6   if  $\mathcal{B}_{r_k, y_k}^E(h_{\min}) \subseteq \mathcal{C}$  then
7     Usar  $r_k$ ,  $y_k$  para computar  $\bar{h}_k := \bar{h}_k(T)$  como en (17)
8      $T \leftarrow T + \bar{h}_k$ 
9      $y_{k+1} \leftarrow \hat{\phi}(\bar{h}_k, y_k)$ 
10     $r_{k+1} \leftarrow \Delta_{r_k, y_k}(\bar{h}_k)$ 
11     $k \leftarrow k + 1$ 
12  else
13    break
14 return  $T$ 

```

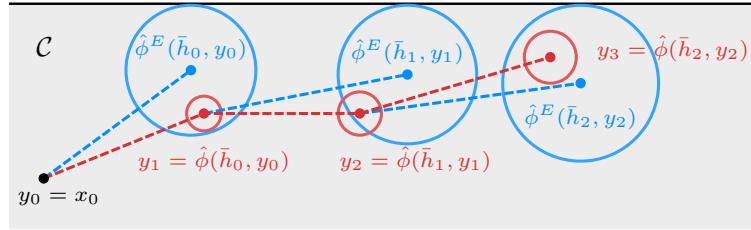


Fig. 3: Ilustración de la lógica del Algoritmo 1. En cada paso, el tiempo \bar{h}_k se calcula como el valor que maximiza el tiempo de permanencia garantizado proporcionado por el Teorema 1. Las bolas que representan la aproximación de Euler se muestran en azul. Al mismo tiempo se calcula una aproximación de orden mayor $\hat{\phi}$ para refinar la cota del error (bolas ilustradas en rojo).

En el algoritmo se verifican en cada ciclo las condiciones del Teorema 1 para calcular un período de tiempo garantizado hasta el cual una solución aproximada dada por $\hat{\phi}$ puede computarse asegurando que la solución continua se mantiene en el conjunto \mathcal{C} durante todo dicho intervalo.

Específicamente, el algoritmo comienza estableciendo $r_0 = 0$ dado que al comienzo la solución de la aproximación coincide con la condición inicial $y_0 := x_0$. Se fija $k = 0$. Luego se ingresa al bucle principal, el cual primero verifica que todavía se puede avanzar una cantidad superior a h_{\min} sin superar T_{\max} a partir del instante actual T y si la solución aproximada actual y_k se encuentra, dado el radio actual r_k asegurado, dentro del conjunto \mathcal{C} . En caso afirmativo, se verifica que la bola dada por la solución de Euler luego de un paso h_{\min} se encuentre también dentro de \mathcal{C} , de otro modo la solución aproximada $\hat{\phi}$ no podría avanzar asegurando el cumplimiento del Teorema 1. Si este es el caso, el estado actual de

la solución aproximada y_k y el radio asegurado r_k se utilizan en (17) para computar \bar{h}_k . Por definición, este intervalo de tiempo garantiza que $\mathcal{B}_{r_k, y_k}^E(\bar{h}_k) \subseteq \mathcal{C}$, resultando \bar{h}_k una longitud de paso admitida. Con esta información se actualiza el instante de tiempo en la línea 8, se computa el siguiente estado de la solución y_{k+1} en la línea 9 y se actualiza el radio de la bola r_{k+1} dada por la cota en (18). El ciclo vuelve a repetirse pero esta vez partiendo del nuevo estado y_{k+1} y del radio asegurado r_{k+1} . De esta forma, en cada ciclo del algoritmo, se garantiza por el Teorema 1 que la solución continua se encuentra dentro del conjunto convexo (13), i.e. $\phi(t, y_k) \in \text{conv}(\mathcal{B}_{r_k, y_k}^E(0) \cup \mathcal{B}_{r_{k+1}, y_{k+1}}^E(h)) \subseteq \mathcal{C}$.

La Figura 3 ilustra el resultado de la aplicación de tres ciclos del algoritmo.

Observación 1. La función $\hat{\phi}$ utilizada en el Algoritmo 1 representa cualquier solución aproximada, siempre que exista una cota del error local correspondiente $\Delta_{r,y}$ que satisfaga (18).

Observación 2. En cada iteración del algoritmo, el tiempo de permanencia garantizado T aumenta al menos en h_{\min} . Dado que la condición del ciclo while asegura que el algoritmo se ejecuta solamente mientras $T_{\max} - T \geq h_{\min}$, el número total de iteraciones está acotado por arriba por $\frac{T_{\max}}{h_{\min}}$, garantizando la terminación en un número finito de pasos.

Observación 3. En algunos casos el algoritmo puede devolver $T = 0$ a pesar de que el tiempo real de permanencia de la solución es positivo. Este resultado puede mejorarse reduciendo el valor elegido de h_{\min} .

Observación 4. Resolver (17) implica típicamente resolver un problema de optimización. Surgen dos dificultades principales: verificar si la bola está contenida en \mathcal{C} , y calcular el máximo correspondiente. Estos problemas pueden abordarse considerando lo siguiente:

- si la región \mathcal{C} tiene una estructura geométrica simple, por ejemplo, es politópica, verificar que la bola está incluida en \mathcal{C} se vuelve sencillo y directo,
- el procedimiento sigue siendo válido si en lugar de calcular \bar{h}_k se calcula h_k con $h_{\min} \leq h_k \leq \bar{h}_k$. Esto puede implementarse de manera eficiente mediante un algoritmo de búsqueda, por ejemplo, el algoritmo de búsqueda binaria.

4 Ejemplos

Aplicamos el Algoritmo 1 a dos ejemplos: primero retomamos el Ejemplo 1 y luego consideramos un sistema no lineal de segundo orden. En ambos casos, h_k se determina mediante búsqueda binaria en cada paso, se consideran regiones politópicas, y la solución aproximada $\hat{\phi}$ de (1) se calcula utilizando el método de Runge-Kutta explícito de cuarto orden (RK4).

Definimos $k_1 := f(x)$, $k_2 := f(x + \frac{h}{2}k_1)$, $k_3 := f(x + \frac{h}{2}k_2)$ y $k_4 := f(x + hk_3)$. Entonces, la solución RK4 en el siguiente instante de tiempo está dada por

$$\hat{\phi}(h, x) := x + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4). \quad (19)$$

Para obtener una cota $\Delta_{r,y}(h)$ para esta aproximación, aplicamos la desigualdad triangular, obteniendo $|\phi(h, x) - \hat{\phi}(h, y)| \leq |\phi(h, x) - \hat{\phi}(h, x)| + |\hat{\phi}(h, x) - \hat{\phi}(h, y)|$.

Existen resultados para derivar una cota superior del error local de truncamiento del método RK4 en un solo paso (Lotkin, 1951). Para un sistema de segundo orden $\dot{x} = f(x)$ con $f(x) = [f_1(x), f_2(x)]^\top$, para derivar la cota se requiere que existan constantes $M, P > 0$ tales que

$$M \geq \sup_{x \in \mathcal{C}, \ell \in \{1,2\}} |f_\ell(x)| \quad (20)$$

$$\frac{P^{q+r}}{M^{q+r-1}} \geq \sup_{x \in \mathcal{C}, \ell \in \{1,2\}} \left| \frac{\partial^{q+r} f_\ell(x)}{\partial x_1^q x_2^r} \right| \quad (21)$$

para todo $1 \leq q + r \leq p$, donde en este caso $p = 4$ es el orden del método. Si se cumplen (20)–(21), entonces se asegura que

$$|\phi(h, x) - \hat{\phi}(h, x)| \leq \sqrt{2} \frac{973}{720} P^4 M h^5. \quad (22)$$

Además, utilizando la Suposición 1 y la (19), se puede obtener

$$\sup_{|x-y| \leq r} |\hat{\phi}(t, x) - \hat{\phi}(t, y)| \leq r \left(1 + \frac{Lh}{2} + \frac{L^2 h^2}{3} + \frac{L^3 h^3}{8} + \frac{L^4 h^4}{24} \right). \quad (23)$$

A partir de (22) y (23), se tiene

$$\begin{aligned} |\phi(h, x) - \hat{\phi}(h, y)| &\leq \sqrt{2} \frac{973}{720} P^4 M h^5 + \left(1 + \frac{Lh}{2} + \frac{L^2 h^2}{3} + \frac{L^3 h^3}{8} + \frac{L^4 h^4}{24} \right) \\ &=: \Delta_{r,y}(h). \end{aligned} \quad (24)$$

4.1 Sistema del Ejemplo 1

Para el sistema lineal (4), la constante de Lipschitz es $L = 1$. Observando que $\sup_{x \in \mathcal{C}, \ell \in \{1,2\}} |f_\ell(x)| \leq 1 =: M$ y definiendo $P := L$, entonces se satisfacen (20) y (21). Consideramos la condición inicial

$$x_0 = 1.0005 [\cos(\pi/2 + 0.43) \sin(\pi/2 + 0.43)]^\top, \quad (25)$$

y los parámetros $h_{\min} = 0.1$ y $T_{\max} = 10$.

La Figura 4 ilustra la aplicación del Algoritmo 1, el cual proporciona un tiempo garantizado de permanencia $T = 0.2761$. En contraste, si se considera únicamente la solución de orden superior RK4 con el valor fijo $h_{\min} = 0.1$, se sugiere incorrectamente que la trayectoria permanece dentro de \mathcal{C} hasta $\tilde{T} \approx 1.97$.

4.2 Sistema no lineal

Consideraremos el modelo de presa depredador de Lotka–Volterra

$$\dot{x} = f(x) = \begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix} = \begin{bmatrix} ax_1 - bx_1 x_2 \\ dx_1 x_2 - cx_2 \end{bmatrix} \quad (26)$$

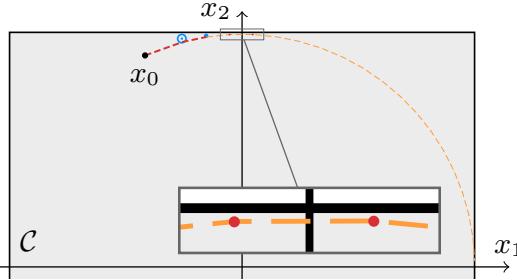


Fig. 4: Trayectoria simulada del sistema (4) desde la condición inicial $x_0 = 1.0005 [\cos(\pi/2 + 0.43) \sin(\pi/2 + 0.43)]^\top$ usando el Algoritmo 1 con RK4 como aproximación de orden superior y parámetros $h_{\min} = 0.1$ y $T_{\max} = 10$. Los pasos de la aproximación RK4 hasta la terminación del algoritmo en $T = 0.2761$ están conectados por una línea roja discontinua. Las bolas $\mathcal{B}_{r_k, y_k}^E(\bar{h}_k)$ en cada paso del algoritmo están ilustradas en azul. Las bolas rojas conectadas por una línea discontinua naranja representan la continuación de la aproximación RK4 con paso fijo h_{\min} . Observando sólo estas bolas de la solución aproximada, como se aprecia en la ampliación en la figura, sugiere incorrectamente que la solución exacta permanece en \mathcal{C} , continuando luego hasta $\tilde{T} \approx 1.97$.

con parámetros $a = 1$, $b = 4$, $c = 0.5$ y $d = 1$ y la región politópica dada por $\mathcal{C} := \{x \in \mathbb{R}^n : x_1 \in [0, 2], x_2 \in [0, 2]\}$. Aunque para el modelo Lotka–Volterra las trayectorias en el espacio de estados cumplen $H = dx_1 - c \ln(x_1) + bx_2 - a \ln(x_2)$, donde H es una constante, no hay una solución cerrada general para despejar un estado a partir del otro. Utilizaremos el Algoritmo 1 para determinar el tiempo de permanencia en la región \mathcal{C} partiendo desde una condición inicial $x_0 = [1.77025 \ 0.13142]^\top$ elegida específicamente para mostrar el funcionamiento del algoritmo. Una cota superior para la constante de Lipschitz en la región \mathcal{C} se puede calcular como $L := 11 \geq \max_{x \in \mathcal{C}} \|J(x)\|_2 = \max_{x \in \mathcal{C}} \left\| \begin{bmatrix} 1 - 4x_2 & -4x_1 \\ x_2 & x_1 - 0.5 \end{bmatrix} \right\|_2$.

Por otro lado es fácil calcular la constante M en (20) como $M := 14 \geq \max \{ \sup_{x \in \mathcal{C}} |f_1(x)|, \sup_{x \in \mathcal{C}} |f_2(x)| \} = \max\{14, 3\}$. Considerando que $\frac{\partial^2 f_1}{\partial x_1^2} = 0$, $\frac{\partial^2 f_1}{\partial x_1 \partial x_2} = -4$, $\frac{\partial^2 f_1}{\partial x_2^2} = 0$, $\frac{\partial^2 f_2}{\partial x_1^2} = 0$, $\frac{\partial^2 f_2}{\partial x_1 \partial x_2} = 1$ y $\frac{\partial^2 f_2}{\partial x_2^2} = 0$, y que las derivadas de grado superior son nulas, se puede calcular la constante P en (21) como $P := \max\{4, 1\} = 4$. Con estos valores queda determinada la cota (24) para el presente sistema.

Aplicamos el algoritmo con parámetros $h_{\min} = 0.014$ y $T_{\max} = 1.5$, obteniendo una cota inferior del tiempo de permanencia de $T = 0.419$. La Figura 5 muestra la aproximación RK4 en líneas rojas punteadas hasta la terminación del algoritmo. La continuación de las aproximaciones RK4 con paso $h_{\min} = 0.014$ luego de la finalización del algoritmo se muestra en naranja. Las sucesivas ampliaciones de la trayectoria muestran que basándose únicamente en las soluciones RK4 calculadas con paso h_{\min} el tiempo de permanencia sería mucho mayor dado

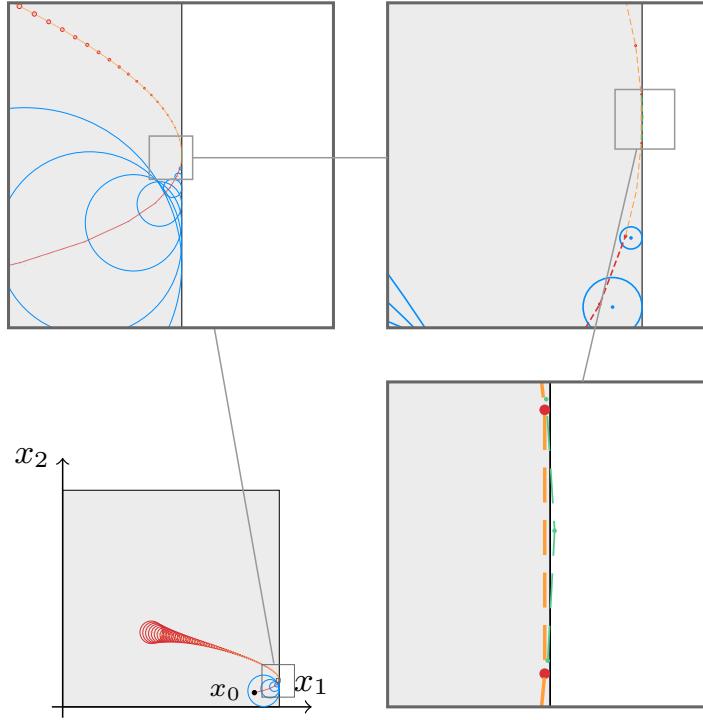


Fig. 5: Trayectorias simuladas del sistema (26) desde la condición inicial $x_0 = [1.77025 \ 0.13142]^\top$ mediante el Algoritmo 1 con RK4 como aproximación de orden superior y parámetros $h_{\min} = 0.014$ y $T_{\max} = 1.5$. La aproximación RK4 hasta la finalización del algoritmo en $T = 0.419$ está ilustrada mediante una línea roja discontinua. Las bolas $B_{r_k, y_k}^E(\bar{h}_k)$ calculadas en cada paso están representadas en azul. Las bolas rojas conectadas por líneas discontinuas naranjas representan la continuación de RK4 tras la finalización del algoritmo, calculada con paso fijo h_{\min} . Las bolas verdes conectadas por una línea discontinua representan la solución RK4 con paso $\frac{h_{\min}}{2} = 0.007$, y muestran que la solución exacta escapa de \mathcal{C} poco después del tiempo de permanencia computado T .

que, como se muestra en la ampliación de abajo a la derecha, las soluciones calculadas con su respectivo error se encuentran dentro del politopo. Sin embargo, simulando la trayectoria de RK4 con paso $\frac{h_{\min}}{2}$ podemos observar que la trayectoria exacta se escapa del politopo poco después del tiempo de permanencia asegurado por el algoritmo.

5 Conclusiones

En este trabajo mostramos que utilizar únicamente la información de las aproximaciones numéricas para verificar el tiempo de permanencia del estado dentro de un conjunto puede ser problemático. A través del uso de las propiedades de convexidad del método de Euler en combinación con aproximaciones de orden

superior, se presentó un método que proporciona un tiempo de permanencia garantizado para sistemas no lineales dentro de una región politópica convexa dada. Los trabajos futuros se concentrarán en añadir robustez al esquema actual y aplicar el algoritmo propuesto para implementar leyes de control en sistemas conmutados, en particular en escenarios donde el objetivo es minimizar la frecuencia de conmutación y maximizar el tiempo de permanencia de cada modo dentro de la región deseada.

Referencias

Cellier, F. E., & Kofman, E. (2006). *Continuous system simulation*. Springer Science & Business Media.

Griffiths, D. F., & Higham, D. J. (2010). *Numerical methods for ordinary differential equations*. Springer London.

Grüne, L., & Pannek, J. (2018). *Nonlinear model predictive control: Theory and algorithms*. Springer Cham.

Julius, A., & van der Schaft, A. (2002). The maximal controlled invariant set of switched linear systems. *Proceedings of the 41st IEEE Conference on Decision and Control*, 2002., 3, 3174–3179 vol.3.

LeVeque, R. (2007). *Finite difference methods for ordinary and partial differential equations: Steady-state and time-dependent problems*. Society for Industrial; Applied Mathematics.

Liberzon, D. (2003). *Switching in systems and control*. Birkhäuser.

Lotkin, M. (1951). On the accuracy of Runge-Kutta's method. *Mathematical tables and other aids to computation*, 5(35), 128–133.

Nesić, D., Teel, A. R., & Kokotović, P. V. (1999). Sufficient conditions for stabilization of sampled-data nonlinear systems via discrete-time approximations. *Systems & Control Letters*, 38(4-5), 259–270.

Perez, M., Anderson, A., Abuin, P., Hernandez-Vargas, E. A., Gonzalez, A. H., & Actis, M. (2023). Computing control invariant sets for waiting-time switched systems: A study case of glucose regulation. *IEEE Control Systems Letters*, 7, 3114–3119.

Perez, M., Anderson, A., Hernandez-Vargas, E. A., González, A. H., & Actis, M. (2022). Permanence regions for switched linear systems under waiting-time constraints [18th IFAC Workshop on Control Applications of Optimization CAO 2022]. *IFAC-PapersOnLine*, 55(16), 184–189.

Russo, A., Paolo Incremona, G., Cavallo, A., & Colaneri, P. (2022). State dependent switching control of affine linear systems with dwell time: Application to power converters. *2022 American Control Conference (ACC)*, 3807–3813.

Stuart, A., & Humphries, A. R. (1998). *Dynamical systems and numerical analysis* (Vol. 2). Cambridge University Press.

Vallarella, A. J., Cardone, P., & Haimovich, H. (2021). Semiglobal exponential input-to-state stability of sampled-data systems based on approximate discrete-time models. *Automatica*, 131, 109742.