Departamento de Matemáticas

Ecuaciones Diferenciales Ordinarias

Jesús Antonio Bueno Linares

Estos apuntes corresponden a la asignatura Ecuaciones Diferenciales Ordinarias, impartida en el segundo curso del Grado en Matemáticas y los dobles Grados en Matemáticas e Informática/Física. Lo que comenzó como unas notas personales para preparar la asignatura diariamente ha ido tomando forma hasta adquirir este resultado, el cual está lejos de ser su versión final. Los apuntes irán cambiando añadiendo nuevos contenidos que, aunque no se cursen en la asignatura según la guía docente aprobada, creo que son interesantes y necesarios para tener la visión más amplia posible de esta asignatura. Cuando encontréis alguna errata, fallo o imprecisión, os agradecería que contactaseis conmigo indicando el error para su corrección.

Índice general

Capítulo 1

Introducción a las EDOs

Las Ecuaciones Diferenciales son una de las herramientas más usadas tanto en Matemáticas como en otras disciplinas como la Física, la Biología, las ingenierías o la Economía, por mencionar algunas. Ayudan a modelizar y resolver problemas que involucran la tasa de cambio de una magnitud, expresada de manera funcional, con respecto de algún parámetro continuo, generalmente el tiempo. El concepto matemático que se encarga de medir el cambio de una función respecto de una variable independiente que la define es la derivada, y es por ello que la palabra "Diferencial"no debería sorprendernos. La otra palabra clave es la de "Ecuación". Una ecuación no es más que una igualdad entre dos expresiones matemáticas, en las que aparecen elementos conocidos e incógnitas.

En las ecuaciones polinómicas que se han estudiado hasta ahora, la incógnita, generalmente denotada por x, es un número (posiblemente complejo!) y la ecuación establece una relación entre algunas potencias de x; por ejemplo

$$3x^3 - 2x - 1 = 0,$$

que admite la solución real $x = 1$ y las soluciones complejas $x = \frac{-1}{2} \pm i\frac{\sqrt{3}}{6}$. Obviamente existen ecuaciones más complejas que las polinómicas, siendo incluso probable que tal ecuación tenga una solución pero no seamos capaces de calcularla explícitamente. Un ejemplo es

$$e^x + \log x = 0, \qquad x > 0.$$

En este punto de los estudios de Grado, es un ejercicio bastante trivial comprobar que, en efecto, la ecuación anterior tiene una única solución para cierto $x_0 \in (0, 1)$.

En las ecuaciones que vamos a tratar este curso, la incógnita es una función $x \equiv x(t)$, dependiendo de la variable independiente t, y los términos de la ecuación podrán involucrar ciertos elementos conocidos, a la propia función $x(t)$ y alguna de sus derivadas.

1.1. Definición de EDO

Comenzamos dando la definición del objeto central de estudio de esta asignatura.

Definición 1.1.1. Dado $n \in \mathbb{N}$, sea $\Omega \subset \mathbb{R}^{n+2}$ abierto y conexo y $\Phi : \Omega \to \mathbb{R}$ una función continua. Una *ecuación diferencial ordinaria (EDO) de orden* n es una expresión de la forma

$$\Phi(t, x(t), x'(t), ..., x^{n)}(t)) = 0, \tag{1.1.1}$$

donde $x : I \subset \mathbb{R} \to \mathbb{R}$ es una función n veces derivable e I un intervalo abierto. La EDO se dice autónoma si Φ no tiene dependencia explícita de t,

$$\Phi(x(t), x'(t), ..., x^{n)}(t)) = 0.$$

Si la derivada de mayor orden de (1.1.1) podemos despejarla, encontramos el siguiente tipo de EDOs.

Definición 1.1.2. La EDO (1.1.1) se dice que se encuentra en *forma normal* si la derivada de mayor orden se puede despejar, esto es

$$x^{n)}(t) = F(t, x(t), ..., x^{n-1)}(t)), \tag{1.1.2}$$

con $F : \Omega \subset \mathbb{R}^{n+1} \to \mathbb{R}$ una función continua.

Aunque algunas de las definiciones que daremos serán para EDOs de la forma (1.1.1), en general a lo largo de este curso supondremos que la EDO siempre se encuentra en forma normal. A la hora de intentar expresar una EDO en forma normal, debemos tener cuidado al despejar la derivada de mayor orden. Por ejemplo,

$$x'(t)^2 + x(t)^2 = 1$$

es una EDO de la forma (1.1.1). Para escribirla de forma normal despejamos su derivada, lo cual nos lleva a la expresión multivaluada

$$x'(t) = \pm\sqrt{1 - x(t)^2}.$$

Cualquier función que resuelva una de estas dos EDOs resuelve la primera, pero el recíproco no es cierto. En general, escribir EDOs en forma normal requiere de despejar la incógnita $x'(t)$ y en última instancia tomar funciones inversas. Hay que ser cautos en cómo realizar estos despejes y considerar todos los casos posibles.

El objetivo primario de una EDO es encontrar una función $x(t)$ que la resuelva. Aunque intuitivo, esto todavía está lejos de ser una definición formal del concepto de solución de una EDO. A continuación damos precisión a esta idea intuitiva.

Definición 1.1.3. Dado un intervalo abierto $I \subset \mathbb{R}$, una función $x : I \to \mathbb{R}$ es una solución de la EDO (1.1.1) si se cumple:

1. $x \in C^n(I)$,
2. para todo $t \in I$ se tiene $(t, x(t), x'(t), ..., x^{n)}(t)) \in \Omega$,
3. la función x satisface $\Phi(t, x(t), x'(t), ..., x^{n)}(t)) = 0$ para todo $t \in I$.

Unos comentarios acerca de esta definición.

- No es de extrañar que se pida $x \in C^n$, puesto que para poder evaluar $x(t)$ en la expresión de la EDO debemos derivar, al menos, n veces. La continuidad de esta última derivada es algo estándar en la teoría y no restringe los resultados obtenidos.
- La segunda condición nos dice que dada $x(t)$, la curva $t \mapsto (t, x(t), x'(t), ..., x^{n)}(t)) \in \Omega$ y en particular tiene sentido aplicar Φ sobre esta curva. Además, por ser Ω conexo y $t \in I$ siendo I un intervalo, la curva $(t, x(t), x'(t), ..., x^{n)}(t))$ no *salta* entre posibles componentes conexas.
- La tercera condición es quizás la que menos análisis merece; simplemente nos dice que x, efectivamente, satisface la ecuación diferencial.

Ejemplo 1.1. Uno de los ejemplos más sencillos de EDO es

$$x'(t) = h(t),$$

siendo $h : I \subset \mathbb{R} \to \mathbb{R}$ una función continua. Sabemos por el Teorema Fundamental del Cálculo que si denotamos por $H(t)$ a cualquier primitiva de $h(t)$ en I, entonces $H'(t) = h(t)$ y por tanto $x(t) = H(t)$ resuelve la EDO. Más aún, dados $t_0 \in I$ y $x_0 \in \mathbb{R}$ la función

$$H(t) = \int_{t_0}^{t} h(s)\, ds + x_0$$

es la única primitiva de $h(t)$ tal que $H(t_0) = x_0$. El Teorema Fundamental del Cálculo nos asegura que todas las posibles soluciones se construyen variando x_0. No todas las EDOs de esta forma se pueden resolver explícitamente. Por ejemplo,

$$x'(t) = e^{-t^2}$$

no tiene solución explícita en términos de funciones elementales, aunque de nuevo el Teorema Fundamental del Cálculo nos asegura la existencia de una primitiva de e^{-t^2}.

Ejemplo 1.2. La primera EDO que no se reduce al cálculo de primitivas y que es la más sencilla de tratar es

$$x'(t) = x(t).$$

Buscamos una función cuya derivada sea ella misma. No es difícil comprobar que las funciones

$$x(t) = \lambda e^t, \qquad \lambda \in \mathbb{R},$$

son soluciones. ¿Habrá más soluciones? Si estas son las únicas, ¿cómo se demuestra?

Ejemplo 1.3. Consideremos la EDO

$$x'(t)^2 + x(t)^2 = 1.$$

Esta EDO no se encuentra en forma normal y hemos visto que para despejar la derivada de mayor orden tenemos que llegar a una expresión multivaluada. Otra opción es derivar la ecuación, obteniendo

$$x'(t)x''(t) + x(t)x'(t) = 0.$$

Si $x'(t) \neq 0$ (o lo que es lo mismo, $x(t)$ no es constante), entonces obtenemos la EDO de segundo orden

$$x''(t) + x(t) = 0.$$

1. ¿Es toda solución de $x''(t) + x(t) = 0$ solución de $x'(t)^2 + x(t)^2 = 1$?
2. ¿Es toda solución de $x'(t)^2 + x(t)^2 = 1$ solución de $x''(t) + x(t) = 0$?
3. ¿Es toda solución no constante de $x'(t)^2 + x(t)^2 = 1$ solución de $x''(t) + x(t) = 0$?
4. Dada $x(t)$ solución de $x''(t)+x(t) = 0$, ¿podrías construir una EDO para la expresión $x'(t)^2 + x(t)^2$ que tenga a $x(t)$ como solución?

Ejemplo 1.4. Si consideramos directamente la EDO

$$x''(t) + x(t) = 0,$$

nos damos cuenta que las funciones $x(t) = \sin t,\ x(t) = \cos t$ son soluciones de esta EDO. En general,

$$x(t) = A \sin t + B \cos t,\ A, B \in \mathbb{R}.$$

Fijémonos ahora que describimos una serie de funciones que cumplen esta ecuación en términos de dos parámetros, A, B. ¿Tendrá algo que ver con que la EDO es de segundo orden? ¿Habrá más soluciones? ¿Son estas soluciones de $x'(t)^2 + x(t)^2 = 1$?

Ejemplo 1.5. Sea la EDO

$$x'(t) = -x(t)^2.$$

Una solución es $x(t) = 0$, definida en todo $\mathbb{R}$. Otra función que resuelve la EDO es $x(t) = 1/t$, pero su dominio de definición es $\mathbb{R} - \{0\}$ que no es un intervalo. En consecuencia, si consideramos $I_1 = \{x < 0\}$ e $I_2 = \{x > 0\}$, entonces $x_1(t) = x(t)_{|I_1}$ y $x_2(t) = x(t)_{|I_2}$ son soluciones distintas de la EDO, puesto que sus intervalos de definición son distintos.

En particular, incluso siendo $\Phi(t, u, v) = v + u^2$ una función continua definida en todo $\mathbb{R}^3$ y la EDO asociada autónoma, algunas soluciones no están definidas en toda la recta real. Esto nos dice que no siempre podemos garantizar que las soluciones estén definidas en todo $\mathbb{R}$, incluso cuando Φ sí está definida en todo $\mathbb{R}^n$.

Ejemplo 1.6. Consideremos ahora la EDO

$$x'(t) = \frac{1}{x(t)}.$$

Si vemos esta EDO en forma normal (1.1.2), la función que la define es $F(t, x) = 1/x$. Esta función es continua y está definida en $\Omega_1 \cup \Omega_2$, donde

$$\Omega_- = \{(t, x) \in \mathbb{R}^2 \colon x < 0\}, \qquad \Omega_+ = \{(t, x) \in \mathbb{R}^2 \colon x > 0\}.$$

El dominio de F no es conexo y por tanto a la hora de considerar soluciones de esta EDO lo tenemos que hacer restringiéndonos a uno de los dos conjuntos Ω_i. Por ejemplo, la función

$$x : (-\frac{c}{2}, \infty) \to \mathbb{R}, \qquad x(t) = \sqrt{2t + c},$$

cumple que es solución de la EDO en Ω_+. Fijémonos que no podemos extender $x(t)$ a $t = -c/2$ porque $x(t)$ no es derivable en tal punto. De igual forma, la función

$$x : (-\frac{c}{2}, \infty) \to \mathbb{R}, \qquad x(t) = -\sqrt{2t + c},$$

es solución en Ω_-. Las soluciones se comportan como deben, ya que ambas cumplen $x(t) \neq 0$ para todo t, siendo solo $x(t) = 0$ para $t = -c/2$, que no está en el dominio de las soluciones, y donde su derivada tiende a infinito.

La existencia o no de soluciones a una EDO, así como su número, es muy variado y depende fuertemente de la estructura de la propia EDO. A continuación, destacamos algunos casos.

1. Hay EDOs que no tienen solución, como por ejemplo
$$e^{x'(t)} + x(t)^2 = 0.$$

2. Hay EDOs que tienen una única solución, como por ejemplo
$$x'(t)^2 + x(t)^2 = 0,$$
que tiene $x(t) = 0$ como única solución.

3. Hay EDOs que tienen infinitas soluciones, como las que hemos visto dependiendo de parámetros. O por ejemplo,
$$x'(t)^2 + x(t)^2 = 1$$
tiene como soluciones $x(t) = 1$, $x(t) = \cos t$ y $x(t) = \sin t$. Incluso
$$x(t) = \begin{cases} 1 & \text{si} \quad t < 0 \\ \cos t & \text{si} \quad t \geq 0 \end{cases}$$
es solución.

1.2. Ecuaciones escalares y vectoriales. Reducción de orden

Recordemos que una de las motivaciones de las EDOs provienen de la mecánica clásica y en concreto en entender el comportamiento de una cierta partícula, por ejemplo plana, cuando sobre ella actúan unas determinadas fuerzas. Es por ello por lo que surge de forma natural el generalizar la ecuación (1.1.1) para poder considerar las siguientes soluciones.

Definición 1.2.1. Una *curva parametrizada* es una aplicación
$$x : I \subset \mathbb{R} \to \mathbb{R}^k, \quad x(t) = (x_1(t), ..., x_n(t)) \quad k \in \mathbb{N}$$
La curva se dice *regular* si $||x'(t)|| \neq 0$ para todo $t \in I$.

Para el caso de curvas que sean solución de una cierta EDO, la ecuación (1.1.1) tiene la forma
$$\Phi : \Omega \subset \mathbb{R} \times \overbrace{\mathbb{R}^k \times \mathbb{R}^k \cdots \mathbb{R}^k}^{n+1)} \to \mathbb{R}, \qquad \Phi(t, x(t), x'(t), ..., x^{n)}(t)) = 0, \tag{1.2.1}$$
y nos referiremos a estas EDOs como vectoriales. El concepto de solución para (1.2.1) es similar al dado para funciones escalares. Por tanto, cualquier solución de (1.2.1) es una curva parametrizada de clase C^n y que resuelve la ecuación diferencial. Aquí debemos remarcar que en una EDO vectorial necesitamos resolver k incógnitas: las coordenadas de la curva $x(t)$.

Ejemplo 1.7. Un caso de especial interés surge al tomar el codominio como $\mathbb{R}^2$ y considerar una curva plana,

$$\alpha : I \subset \mathbb{R} \to \mathbb{R}^2, \qquad \alpha(t) = (x(t), y(t)).$$

Si α es regular, se define su curvatura como

$$\kappa_\alpha(t) = \frac{\langle \alpha''(t), J\alpha'(t) \rangle}{||\alpha'(t)||^3}, \tag{1.2.2}$$

donde $J : \mathbb{R}^2 \to \mathbb{R}^2$ es el giro de ángulo $\pi/2$ dado por $J(x, y) = (-y, x)$. Por ejemplo, un problema que se abordará en cursos posteriores es determinar las curvas $\alpha(t)$ en $\mathbb{R}^2$ con curvatura constante, $\kappa \in \mathbb{R}$. La función que define la EDO es

$$\Phi(t, \alpha(t), \alpha'(t), \alpha''(t)) = \frac{\langle \alpha''(t), J\alpha'(t) \rangle}{||\alpha'(t)||^3} - \kappa, \qquad \kappa \in \mathbb{R}.$$

¿Podrías determinar cuáles son las curvas con curvatura nula?

En general, si $\varphi : I \times D \subset \mathbb{R} \times \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ podemos considerar el problema de *predeterminar la curvatura* de α de la siguiente forma

$$\kappa_\alpha(t) = \varphi(t, \alpha(t), \alpha'(t)), \qquad \forall t \in I,$$

lo cual nos lleva a la función

$$\Phi(t, \alpha(t), \alpha'(t), \alpha''(t)) = \frac{\langle \alpha''(t), J\alpha'(t) \rangle}{||\alpha'(t)||^3} - \varphi(t, \alpha(t), \alpha'(t)).$$

Es decir, podemos considerar el problema de predeterminar la curvatura de α en términos del tiempo, o de forma más interesante en términos de la propia α y su derivada. Un problema realmente interesante, incluso en la investigación actual, tiene que ver con curvas planas satisfaciendo la relación

$$\kappa_\alpha(t) = \langle J\alpha'(t), e_2 \rangle, \qquad e_2 = (0, 1).$$

Ya hemos visto que al considerar EDOs vectoriales tenemos k incógnitas. Tal y como sucede en las ecuaciones algebraicas que llevamos estudiando *desde siempre*, lo natural es imponer más de una ecuación para poder obtener la solución de forma satisfactoria. Dados $k, n, m \in \mathbb{N}^3$, un sistema de m EDOs de orden n es un sistema diferencial

$$\begin{cases} \Phi_1(t, x(t), x'(t), ..., x^{n)}(t)) &= 0, \\ \Phi_2(t, x(t), x'(t), ..., x^{n)}(t)) &= 0, \\ \qquad\qquad \vdots & \\ \Phi_m(t, x(t), x'(t), ..., x^{n)}(t)) &= 0, \end{cases}$$

donde $x : I \subset \mathbb{R} \to \mathbb{R}^k$ y $\Phi_i : \Omega \subset \mathbb{R} \times \mathbb{R}^{k(n+1)} \to \mathbb{R}$. Se dice que x es solución del sistema si es solución de cada una de las m EDOs. Esto se puede resumir de forma vectorial definiendo la función

$$\Phi : \Omega \to \mathbb{R}^m, \qquad \Phi = (\Phi_1, ..., \Phi_m).$$

Unos cuantos comentarios.

1. Las funciones Φ_i podrían a priori estar definidas en dominios distintos, $\Omega_i \subset \mathbb{R} \times \mathbb{R}^{k(n+1)}$. No obstante, como para toda solución x se tiene que cumplir $t \mapsto (t, x(t), ..., x^{n)}(t)) \in \Omega_i$, se debe tomar un dominio común donde esta condición se cumpla para todo i. Salvo intersección finita, se considera Ω el mismo dominio.

2. Cada función que define una ecuación podría tener un orden de derivación distinto. No obstante, por ser un número finito de ecuaciones, nos podemos quedar con el orden máximo y escribir de forma general que todas las EDOs tienen dependencia de una derivada máxima n-ésima.

3. Debemos recordar que la incógnita es la curva $x(t) = (x_1(t), ..., x_k(t))$, es decir tenemos k incógnitas que resolver. Por razones que serán claras más adelante, sólo consideraremos sistemas que contengan el mismo número de ecuaciones que de incógnitas ($m = k$), de tal forma que las ecuaciones sean independientes entre sí, es decir, ninguna de ellas se deduzca de las demás.

Ejemplo 1.8. Dadas $x, y : I \to \mathbb{R}$, un ejemplo sencillo de sistema de ecuaciones diferenciales es el siguiente

$$\begin{cases} x'(t) = y(t), \\ y'(t) = -x(t). \end{cases}$$

¿Podrías identificar este sistema con alguna EDO escalar?

Ejemplo 1.9. Supongamos que un objeto orbita alrededor de la Tierra y suponemos que tal movimiento se realiza en un plano conteniendo al centro de la tierra. Si el objeto se parametriza por $\alpha(t) = (x(t), y(t))$ y el centro de la Tierra se supone con coordenadas $(0, 0)$, la EDO satisfecha por las coordenadas de α es

$$\begin{cases} x''(t) = -\dfrac{gR^2 x(t)}{(x(t)^2 + y(t)^2)^{3/2}}, \\ \\ y''(t) = -\dfrac{gR^2 y(t)}{(x(t)^2 + y(t)^2)^{3/2}}, \end{cases}$$

donde R es el radio de la Tierra y g la constante de gravitación. Si $\alpha(t)$ orbita alrededor de la Tierra entonces podemos suponer $x(t)^2 + y(t)^2 = r^2$, para cierto $r > 0$ y por tanto se tiene

$$\begin{cases} x''(t) = -\dfrac{gR^2 x(t)}{r^3}, \\ \\ y''(t) = -\dfrac{gR^2 y(t)}{r^3}. \end{cases}$$

Asumiendo que las funciones $x(t) = r\cos(\omega t),\ y(t) = r\sin(\omega t),\ \omega \in \mathbb{R}$ son soluciones, determina el valor de ω.

El siguiente resultado prueba que toda EDO, con soluciones escalares o vectoriales, de orden n de la forma (1.1.1) se puede transformar en un sistema equivalente de primer orden con n ecuaciones.

Proposición 1.2.2. *Sea la EDO (1.1.1), esto es*

$$\Phi(t, x(t), x'(t), ..., x^{n)}(t)) = 0,$$

con $\Phi : \Omega \subset \mathbb{R} \times \mathbb{R}^{k(n+1)} \to \mathbb{R}^m$ *continua y* $x : I \subset \mathbb{R} \to \mathbb{R}^k$ *de clase* C^n*, y definamos*

$$y_k(t) = x^{k-1)}(t), \qquad k = 1, ..., n,$$

entendiendo $x^{0)}(t) = x(t)$*. Entonces, la curva* $(y_1(t), ..., y_n(t))$ *es solución del sistema de orden 1 con* n *ecuaciones,*

$$\begin{cases} y_1'(t) = y_2(t), \\ y_2'(t) = y_3(t), \\ \vdots \\ y_{n-1}'(t) = y_n(t), \\ \Phi(t, y_1(t), ..., y_n(t), y_n'(t)) = 0. \end{cases} \tag{1.2.3}$$

Recíprocamente, si la curva $(y_1(t), ..., y_n(t))$ *es solución del sistema* (1.2.3)*, entonces* $x(t) = y_1(t)$ *es solución de la EDO* (1.1.1) *y además cada* $y_k(t)$ *es la derivada* $(k-1)$*-ésima de* $x(t)$*.*

En particular, si consideramos la EDO en forma normal,

$$x^{n)}(t) = F(t, x(t), x'(t), ..., x^{n-1)}(t)),$$

con $F : \Omega \subset \mathbb{R} \times \mathbb{R}^{kn} \to \mathbb{R}^k$ continua, y los cambios de variable $y_k(t) = x^{k-1)}(t),\ k = 1, ..., n$, podemos reducirla a la ecuación vectorial de primer orden

$$\begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}' = \begin{pmatrix} y_2(t) \\ y_3(t) \\ \vdots \\ F(t, y_1(t), ..., y_n(t)) \end{pmatrix}$$

Ejemplo 1.10. Cuando tenemos EDOs de orden 2, es usual abusar la notación y denotar $y_1(t) \equiv x(t)$ e $y_2(t) \equiv y(t)$. En tal caso, si la EDO es de la forma

$$\Phi(t, x(t), x'(t), x''(t)) = 0,$$

obtendríamos el sistema equivalente

$$\begin{cases} x'(t) = y(t), \\ \Phi(t, x(t), y(t), y'(t)) = 0. \end{cases}$$

Si la EDO es normal, entonces tendríamos

$$\begin{cases} x'(t) = y(t), \\ y'(t) = F(t, x(t), y(t)). \end{cases}$$

Por ejemplo, si consideramos la EDO de segundo orden

$$x''(t) + x(t) = 0$$

y el cambio de variable $y(t) = x'(t)$, entonces esta EDO se transforma en

$$\begin{cases} x'(t) = y(t), \\ y'(t) = -x(t). \end{cases}$$

1.3. Problema de Cauchy. Ecuaciones diferenciales e integrales

Como hemos visto en la EDO más sencilla posible $x'(t) = h(t)$, este problema tiene infinitas soluciones dadas por $x(t) = H(t) + \lambda,\ \lambda \in \mathbb{R}$, donde $H(t)$ es *cualquier* primitiva de $h(t)$. No obstante, dados $t_0, x_0 \in \mathbb{R}$ con $t_0 \in \operatorname{dom}(h)$, si imponemos la condición $x(t_0) = x_0$ entonces este problema tiene existencia y unicidad dada por la función $x(t) = H(t) + x_0 - H(t_0)$. La mera existencia de estos problemas está lejos de ser trivial. Por ejemplo, la EDO $x'(t)^2 + x(t)^2 = 1$ tiene infinitas soluciones, pero no existe ninguna tal que $x(0) = 2$. Otro ejemplo: la EDO $x''(t)+x(t) = 0$ tiene infinitas soluciones, así como el problema con la condición $x(0) = 2$, pero si imponemos además la condición $x'(0) = 2$, entonces sólo hay una función que cumpla este problema.

La siguiente definición introduce uno de los objetos clave de estudio en la teoría.

Definición 1.3.1. Un *problema de Cauchy*, llamado también problema de valores iniciales, es una EDO de orden n junto con n condiciones iniciales de la forma

$$\begin{cases} \Phi(t, x(t), x'(t), ..., x^{n)}(t)) = 0, \\ x(t_0) = x_0, \\ x'(t_0) = x_1, \\ \vdots \\ x^{n-1)}(t_0) = x_{n-1}, \end{cases} \tag{1.3.1}$$

siendo $(t_0, x_0, x_1, ..., x_{n-1}) \in \Omega$.

A las n restricciones $x^{i)}(t_0) = x_i,\ i = 0, ..., n-1$ se les llama *condiciones iniciales*. Una solución del problema de Cauchy (1.3.1) es una solución de la EDO que cumple las condiciones iniciales. Es importante destacar que el instante t_0 para el cual se fijan las n condiciones es el mismo. Si consideramos instantes distintos estaríamos trabajando con un *problema de contorno*, el cual escapa los contenidos de este curso introductorio a las EDOs.

En general, para resolver un problema de Cauchy primero se resuelve la EDO asociada y se obtienen soluciones dependientes de una serie de parámetros. Estos parámetros se determinarán mediante las condiciones iniciales.

Observación 1.1. Un par de observaciones sobre el problema de Cauchy asociado a una EDO.

- No todo problema de Cauchy tiene solución. Por ejemplo,

$$x'(t)^2 + x(t)^2 = 0, \qquad x(0) = 1$$

  no tiene solución puesto que la única solución de la EDO es $x(t) = 0$. En general, el problema de Cauchy

$$x'(t)^2 + x(t)^2 = a, \qquad x(0) = b, \qquad a, b \in \mathbb{R},$$

  solo tiene solución para $a \geq b$.

- Hay problemas de Cauchy que tienen más de una solución. Por ejemplo, consideremos el problema

$$x'(t) = \sqrt{|x(t)|}, \qquad x(0) = 0.$$

Claramente $x(t) = 0$ es solución pero $x(t) = t^2/4$ también lo es si $t > 0$, al igual que $-t^2/4$ si $t < 0$. En general, si $c_1, c_2 \in \mathbb{R}$ con $-\infty \leq c_1 \leq 0 \leq c_2 \leq \infty$ se tiene que

$$x(t) = \begin{cases} \dfrac{-(t-c_1)^2}{4} & \text{si} \quad t < c_1 \\ 0 & \text{si} \quad c_1 \leq t \leq c_2 \\ \dfrac{(t-c_2)^2}{4} & \text{si} \quad t > c_2 \end{cases}$$

resuelve el problema de Cauchy. Aquí debemos entender que si $c_2 = \infty$ (resp. $c_1 = -\infty$) entonces $x(t) = 0$ para todo $t \geq c_1$ (resp. $t \leq c_2$).

Ejemplo 1.11. Dado el problema de Cauchy

$$x''(t) + x(t) = 0, \qquad x(0) = 1, \quad x'(0) = 2,$$

la única función de la forma $x(t) = A\sin t + B\cos t$ que es solución de este problema es $x(t) = \cos t + 2\sin t$. ¿Habrá más soluciones?

1.3.1. Integral de Volterra

Supongamos que tenemos una EDO en forma normal ya reducida a primer orden,

$$x'(t) = F(t, x(t)), \qquad x : I \subset \mathbb{R} \to \mathbb{R}^k,\ k \in \mathbb{N},$$

con $F : \Omega \subset \mathbb{R}^{k+1} \to \mathbb{R}^k$ continua, y consideramos el problema de Cauchy con condición inicial $x(t_0) = x_0,\ t_0 \in I,\ x_0 \in \mathbb{R}^k$. El siguiente resultado nos da una caracterización de las soluciones de este problema de Cauchy en términos de una ecuación integral.

Teorema 1.3.2. *Sean $F : \Omega \subset \mathbb{R} \times \mathbb{R}^k \to \mathbb{R}^k$ una función continua e $I \subset \mathbb{R}$ un intervalo. Consideremos el problema de Cauchy*

$$\begin{cases} x'(t) = F(t, x(t)), \\ x(t_0) = x_0. \end{cases}$$

y sea $x : I \to \mathbb{R}^k$. Equivalen:

1. *$x(t)$ es solución del problema de Cauchy.*

2. *La función $x(t)$ es continua, $(t, x(t)) \in \Omega$ y se cumple*

$$x(t) = x(t_0) + \int_{t_0}^{t} F(s, x(s))\, ds.$$

La ecuación integral anterior se conoce como *ecuación de Volterra* y nos aparecerá en temas posteriores.

Demostración. Si $x(t)$ es solución del problema de Cauchy entonces $x \in C^1$ y $(t, x(t)) \in \Omega$ por la propia definición. Además, por ser solución del problema de Cauchy se tiene

$$x'(t) = F(t, x(t)), \qquad x(t_0) = x_0.$$

Integrando entre t_0 y t y aplicando la regla de Barrow,

$$x(t) = x(t_0) + \int_{t_0}^{t} F(s, x(s))\, ds,$$

es decir se cumple la integral de Volterra.

Recíprocamente, si $x(t)$ es continua y $(t, x(t)) \in \Omega$ entonces $t \mapsto F(t, x(t))$ es continua. Por hipótesis

$$x(t) = x(t_0) + \int_{t_0}^{t} F(s, x(s))\, ds,$$

y por tanto el Teorema Fundamental del Cálculo asegura que $x \in C^1(I)$. Si derivamos esta expresión, obtenemos $x'(t) = F(t, x(t))$. Además, $x(t_0) = x_0$ y por tanto $x(t)$ es solución del problema de Cauchy. $\square$

1.4. Ecuaciones diferenciales a través de funciones implícitas

Por lo que hemos visto hasta ahora, sabemos que algunas funciones dadas a priori son soluciones de algunas EDOs concretas. Por ejemplo, toda función constante tiene derivada cero, lo que se traduce en que las funciones $x(t) = k,\ k \in \mathbb{R}$, son soluciones de la EDO $x'(t) = 0$. O por ejemplo, ya que la función exponencial cumple que es su propia derivada también sabemos que las funciones de la forma $x(t) = \lambda e^t,\ \lambda \in \mathbb{R}$, son soluciones de la EDO $x'(t) = x(t)$. A continuación vamos a generalizar esta idea a un contexto algo más complejo.

Sea $H\colon \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^1, donde en Ω consideramos coordenadas (t, x). Dado $(t_0, x_0) \in \Omega$, nos preguntamos por la existencia de una función $x : I \subset \mathbb{R} \to \mathbb{R}$ de clase C^1 tal que $x(t_0) = x_0$, se tenga $t \mapsto (t, x(t)) \in \Omega$ y además $H(t, x(t)) = 0$ para todo $t \in I$. En otras palabras, queremos expresar localmente la curva $H(t, x) = 0$ alrededor del punto (t_0, x_0) como un grafo $x = x(t)$.

Ejemplo 1.12. Dado $r > 0$, consideramos la función $H : \mathbb{R}^2 \to \mathbb{R}$ dada por $H(t, x) = t^2 + x^2 - r^2$. Sabemos que la curva $H(t, x) = 0$ define la circunferencia de centro el origen y radio r, $C((0, 0), r)$. Si queremos expresar esta circunferencia como un grafo $x = x(t)$, basta despejar la variable x obteniendo

$$x = \pm\sqrt{r^2 - t^2},$$

lo cual define las dos funciones $x(t) = \pm\sqrt{r^2 - t^2}$, cada una de ellas correspondiente a una semicircunferencia abierta. Esta función es de clase C^1 en todos los puntos salvo en $t = \pm r$, que es precisamente donde los puntos de la curva $H(t, x) = 0$ tienen tangente vertical. No obstante, para cualquier $(t_0, x_0) \in C((0, 0), r)$ tal que $t_0 \neq \pm r$, existe exactamente una única función (dependiendo de si $x_0 > 0$ ó $x_0 < 0$) tal que la circunferencia se expresa localmente alrededor de (t_0, x_0) como el grafo de la función $x(t) = \mathrm{signo}(x_0)\sqrt{r^2 - t^2}$.

El caso de la circunferencia es excepcional porque podemos resolver explícitamente la ecuación $H(t, x) = 0$ en función de t, pero este no es el caso en general; basta considerar por ejemplo $H(t, x) = e^x + x + t$. La existencia y unicidad de tal función $x(t)$ está garantizada por el Teorema de la función implícita, el cual enunciamos a continuación para una función de dos variables.

Teorema 1.4.1 (Teorema de la función implícita)**.** *Sea $\Omega \subset \mathbb{R}^2$ abierto y $H : \Omega \to \mathbb{R}$ de clase C^1. Supongamos que $(t_0, x_0) \in \Omega$ y $\dfrac{\partial H}{\partial x}(t_0, x_0) \neq 0$. Entonces existe $I \subset \mathbb{R}$ intervalo abierto con $t_0 \in I$, y una única función $x : I \to \mathbb{R}$ de clase C^1 cumpliendo $x(t_0) = x_0$, $(t, x(t)) \in \Omega$ y $H(t, x(t)) = 0$ para todo $t \in I$.*

La unicidad debe entenderse del siguiente modo: si $z : J \subset \mathbb{R} \to \mathbb{R}$ es tal que $z(t_0) = x_0$ y $H(t, z(t)) = 0$ para todo $t \in J$, entonces en particular $I \cap J \neq \varnothing$ por ser I, J intervalos abiertos con $t_0 \in I \cap J$. En tal caso, $x(t) = z(t)$ para todo $I \cap J$. En particular, si $I \neq J$ la solución se puede extender de la siguiente forma

$$\hat{x}(t) = \begin{cases} x(t) & \text{si } t \in I, \\ z(t) & \text{si } t \in J. \end{cases}$$

En tal caso, seguiremos llamando $x(t)$ a la solución extendida por simplificar la notación.

Podemos decir un poco más de la función $x(t)$. Si derivamos respecto de t la relación $H(t, x(t)) = 0$ y aplicamos la regla de la cadena, obtenemos

$$\frac{d}{dt}\left(H(t, x(t))\right) = \frac{\partial H}{\partial t}(t, x(t)) + \frac{\partial H}{\partial x}(t, x(t))x'(t) = 0.$$

Por tanto, toda función definida implícitamente por la relación $H(t, x(t)) = 0$ es solución de la EDO

$$\frac{\partial H}{\partial t}(t, x(t)) + \frac{\partial H}{\partial x}(t, x(t))x'(t) = 0.$$

Además, fijada la condición inicial $x(t_0) = x_0$ existe una única solución del problema de Cauchy asociado. Recíprocamente, supongamos que $x : I \subset \mathbb{R} \to \mathbb{R}$ es solución de la EDO

$$\frac{\partial H}{\partial t}(t, x(t)) + \frac{\partial H}{\partial x}(t, x(t))x'(t) = 0,$$

siendo $H : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^1. Entonces $\dfrac{d}{dt}\left(H(t, x(t))\right) = 0$ y por tanto existe $\lambda \in \mathbb{R}$ tal que

$$H(t, x(t)) = \lambda, \qquad \forall t \in I.$$

Aquí debemos incidir en que si la solución $x(t)$ cambia, entonces también cambiará el valor de λ. Si tomamos la condición inicial $x(t_0) = x_0$ entonces se tiene

$$H(t, x(t)) = H(t_0, x_0), \qquad \forall t \in I.$$

Hemos probado el siguiente resultado.

Teorema 1.4.2. *Sea $H : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^1, con $(t, x) \in \Omega$. Entonces, son equivalentes:*

1. *Dado $(t_0, x_0) \in \Omega$ con $H(t_0, x_0) = 0$ y $\dfrac{\partial H}{\partial x}(t_0, x_0) \neq 0$, existe una única función $x : I \subset \mathbb{R} \to \mathbb{R}$ con $t_0 \in I$ que cumple $x(t_0) = x_0$, $(t, x(t)) \in \Omega$ y $H(t, x(t)) = 0$ para todo $t \in I$. Además, $x(t)$ es solución de la EDO*

$$\frac{\partial H}{\partial t}(t, x(t)) + \frac{\partial H}{\partial x}(t, x(t))x'(t) = 0.$$

2. *Sea $x(t)$ una solución de la EDO*

$$\frac{\partial H}{\partial t}(t,x(t)) + \frac{\partial H}{\partial x}(t,x(t))x'(t) = 0.$$

Entonces, $x(t)$ está definida implícitamente por la ecuación $H(t,x(t)) = \lambda$, para algún $\lambda \in \mathbb{R}$. Más aún, dado $(t_0,x_0) \in \Omega$, la solución del problema de Cauchy con condición inicial $x(t_0) = x_0$ está definida implícitamente sin ambigüedad por la única función que satisface la ecuación implícita

$$H(t,x(t)) = H(t_0,x_0).$$

Ejemplo 1.13. Volvamos al ejemplo de la circunferencia, $H(t,x) = t^2 + x^2 - r^2$. Entonces

$$\frac{\partial H}{\partial x} = 2x,$$

la cual se anula si y sólo si $x = 0$ y en tal caso $t = \pm r$. Sea $x_0 \neq 0$ y tomamos (t_0,x_0) tal que $H(t_0,x_0) = 0$. Sabemos que existe $x(t)$ definida en $I \subset \mathbb{R}$ que contiene a t_0 tal que $H(t,x(t)) = 0$ para todo $t \in I$. Además, $x(t)$ es solución de la EDO

$$x(t)x'(t) = -t.$$

Ya sabemos que la función $x(t)$ se puede calcular explícitamente como

$$x(t) = \pm\sqrt{r^2 - t^2}.$$

Ejemplo 1.14. Consideramos la EDO en forma normal

$$x'(t) = \frac{1}{e^{x(t)} + 1}.$$

Si definimos $H(t,x) = e^x + x - t$, entonces es fácil comprobar que las soluciones de la EDO cumplen $H(t,x(t)) = \lambda,\ \lambda \in \mathbb{R}$. El problema ahora es que $x(t)$ no se puede despejar. No obstante, $\frac{\partial H}{\partial t}(t,x) = -1$ y el Teorema de la función implícita nos asegura la existencia de una función $t : J \to \mathbb{R}$ tal que $H(t(x),x) = \lambda$ para todo $x \in J$. En este caso $t(x)$ se puede despejar explícitamente como $t(x) = e^x + x - \lambda$, y bastaría imponer una condición inicial para determinar el parámetro λ.

1.4.1. Familias de curvas como soluciones de EDOs

Otro caso donde aparece la derivación implícita es el siguiente. Sabemos que dada una EDO, sus soluciones suelen depender de parámetros. Por ejemplo, $x'(t) = x(t)$ tiene como familia de soluciones $x(t) = \lambda e^t,\ \lambda \in \mathbb{R}$. O por ejemplo, $x''(t) + x(t) = 0$ tiene como familia de soluciones $x(t) = A\sin t + B\cos t,\ A,B \in \mathbb{R}$. En general, las soluciones de una EDO de primer orden vienen dadas por una relación de la forma

$$H(t,x(t),\lambda) = 0,$$

donde las distintas soluciones se obtienen variando el parámetro λ. En este punto, es lógico plantearnos el problema inverso. Sea $\Omega \subset \mathbb{R}^3$ con coordenadas (t,x,λ) y $H : \Omega \to \mathbb{R}$ una función de clase C^1. Para cada $\lambda \in \mathbb{R}$, consideramos el subconjunto

$$\Gamma_{H,\lambda} = \{(t,x) \in \mathbb{R}^2 \colon H(t,x,\lambda) = 0\}.$$

Nos preguntamos si existe una EDO de forma que sus soluciones sean precisamente las curvas de la familia $\Gamma_{H,\lambda}$. Aquí debemos hacer el siguiente inciso importante. La familia de curvas depende de un parámetro λ, pero estamos buscando una EDO cuyas soluciones definan *todas* las curvas de la familia. De nuevo, es bastante obvio que debemos aplicar el Teorema de la función implícita. Fijado $\lambda \in \mathbb{R}$ y dado $(t_0, x_0) \in \Gamma_{H,\lambda}$, si $\frac{\partial H}{\partial x}(t_0, x_0, \lambda) \neq 0$, el Teorema de la función implícita asegura la existencia de cierta $x \in C^1(I)$ tal que $x(t_0) = x_0$ y $H(t, x(t), \lambda) = 0$, $\forall t \in I$, o de forma equivalente $(t, x(t)) \in \Gamma_{H,\lambda}$. Para intentar obtener la EDO satisfecha por $x(t)$ derivamos implícitamente esta expresión y obtenemos

$$\frac{\partial H}{\partial t}(t, x(t), \lambda) + \frac{\partial H}{\partial x}(t, x(t), \lambda)x'(t) = 0.$$

Esto todavía no es una EDO para $x(t)$ por la dependencia de la variable λ. Para eliminarla podemos considerar el sistema

$$\begin{cases} H(t, x(t), \lambda) = 0 \\ \dfrac{\partial H}{\partial t}(t, x(t), \lambda) + \dfrac{\partial H}{\partial x}(t, x(t), \lambda)x'(t) = 0, \end{cases}$$

y ayudarnos de ambas ecuaciones para eliminar esta dependencia. Si existe más de un parámetro habrá que derivar tantas veces como el número de parámetros que tengamos para intentar despejarlos.

Observación 1.2. En el caso particular en el que H sea de la forma

$$H(t, x, \lambda) = G(t, x) - g(\lambda),$$

la dependencia del parámetro λ siempre queda descartada. En efecto,

$$\frac{\partial H}{\partial t}(t, x(t), \lambda) + \frac{\partial H}{\partial x}(t, x(t), \lambda)x'(t) = \frac{\partial G}{\partial t}(t, x(t)) + \frac{\partial G}{\partial x}(t, x(t))x'(t) = 0,$$

donde la dependencia de λ ha desaparecido. Este es el caso por ejemplo de la familia de circunferencias $H(t, x, r) = t^2 + x^2 - r^2$, donde la EDO satisfecha por estas circunferencias es $2t + 2xx' = 0$.

Ejemplo 1.15. Se considera la familia de curvas

$$H(t, x, \lambda) := e^{\lambda x} - \frac{t^2}{x} = 0, \qquad x > 0.$$

La condición $x > 0$ se toma porque de lo contrario esta ecuación no tiene solución. Sea $(t_0, x_0) \in \mathbb{R}^2$, con $x_0 > 0$, tal que $H(t_0, x_0, \lambda) = 0$. Como

$$\frac{\partial H}{\partial x}(t_0, x_0, \lambda) = \lambda e^{\lambda x_0} + \frac{t^2}{x^2} = \frac{t_0^2}{x_0}(\lambda + t_0),$$

si $\lambda \neq -t_0$ garantizamos la existencia de una función $x(t)$ cumpliendo $e^{\lambda x(t)} = t^2/x(t)$ para todo $t \in I$ con $t_0 \in I \subset \mathbb{R}$. ¿Qué ecuación diferencial satisface x? Si derivamos implícitamente,

$$\lambda e^{\lambda x(t)} x'(t) - \frac{2t}{x(t)} + \frac{t^2}{x(t)^2}x'(t) = 0,$$

o de forma equivalente

$$x'(t) = \frac{2tx(t)}{\lambda x(t)^2 e^{\lambda x(t)} + t^2}.$$

Esto todavía no es una EDO para $x(t)$ pues depende de λ. Despejemos las expresiones que acompañen a λ en términos de $t, x(t)$. Por una parte, $e^{\lambda x(t)} = t^2/x(t)$. De esta igualdad despejamos λ, obteniendo

$$\lambda = \frac{1}{x(t)} \log \frac{t^2}{x(t)}.$$

Finalmente, concluimos que x es solución de la EDO

$$x'(t) = \frac{2x(t)}{t \log \dfrac{t^2}{x(t)} + t}.$$

1.5. Problemas geométricos

Las Ecuaciones Diferenciales nos permiten describir numerosos problemas que tienen una interpretación o motivación geométrica o física. Hacer un estudio exhaustivo de esta materia se correspondería con una asignatura propia, así que vamos a exponer un par de ejemplos que creemos interesantes.

Un problema geométrico que surge de forma natural es el siguiente. Sea

$$H : \Omega \subset \mathbb{R}^3 \to \mathbb{R}.$$

Dado $\lambda \in \mathbb{R}$, supongamos que tenemos definida una familia de curvas de forma implícita,

$$\Gamma_{H,\lambda} = \{(t,x) \in \mathbb{R}^2 \colon H(t,x,\lambda) = 0\}.$$

Sabemos por el Teorema de la función implícita que dado $(t_0, x_0) \in \Gamma_{H,\lambda}$, si se tiene

$$\frac{\partial H}{\partial x}(t_0, x_0, \lambda) \neq 0,$$

existe un entorno $I \subset \mathbb{R}$ de t_0 y una función $x : I \to \mathbb{R}$ tal que

$$H(t, x(t), \lambda) = 0,$$

es decir que $x(t)$ es solución local de esta familia de curvas pasando por (t_0, x_0). Además, sabemos un método para deducir la EDO satisfecha por $x(t)$: basta despejar λ del sistema

$$\begin{cases} H(t, x(t), \lambda) = 0 \\ \dfrac{\partial H}{\partial t}(t, x(t), \lambda) + \dfrac{\partial H}{\partial x}(t, x(t), \lambda) x'(t) = 0, \end{cases}$$

siempre que se pueda. Nos preguntamos por la existencia de una familia de curvas $\Psi(t, x, \mu) = 0$ cumpliendo la siguiente propiedad: *si $\gamma_1 \in \Gamma_{H,\lambda}$ y $\gamma_2 \in \Gamma_{\Psi,\mu}$, entonces $\gamma_!$ y γ_2 siempre se cortan de forma ortogonal, y eso para todo $\lambda, \mu \in \mathbb{R}$*. A la familia definida por la relación $\Psi(t, x, \lambda) = 0$ se le conoce como las trayectorias ortogonales a la familia $H(t, x, \lambda) = 0$. A continuación, presentamos un método para calcular Ψ.

1. De la familia $H(t, x, \lambda) = 0$, obtenemos la EDO satisfecha por $x(t)$. Supongamos pues

$$x'(t) = F(t, x(t)).$$

2. Geométricamente, la pendiente de la curva $(t, x(t))$ en cada instante t coincide con $x'(t)$. Por tanto, la curva ortogonal pasando por el punto $(t, x(t))$ tendrá pendiente $-\frac{1}{x'(t)}$, en aquellos puntos donde $x'(t) \neq 0$. Consideramos pues la ecuación diferencial
$$x'(t) = -\frac{1}{F(t, x(t))},$$
en aquellos puntos que tenga sentido (esto es, donde $F(t, x(t)) \neq 0$).

3. Resolvemos esta última ecuación diferencial, obteniendo la familia de soluciones
$$\Psi(t, x(t), \mu) = 0,$$
que definen las trayectorias ortogonales a la familia original $H = 0$.

Observación 1.3. Es importante resaltar que *cada* curva de la familia Ψ es ortogonal a *cada* curva de la familia H. Es decir, no es que la curva $H(t, x, \lambda) = 0$ sea ortogonal únicamente a la curva $\Psi(t, x, \lambda) = 0$, es que para todo $\lambda, \mu \in \mathbb{R}$ se tiene que $H(t, x, \lambda) = 0$ es ortogonal a $\Psi(t, x, \mu) = 0$.

Ejemplo 1.16. Nos planteamos encontrar las trayectorias ortogonales a la a familia de parábolas $x(t) = \lambda t^2$. Para ello, seguimos los pasos anteriormente expuestos.

1. Derivando y despejando λ obtenemos
$$x'(t) = \frac{2x(t)}{t}.$$

2. La familia de trayectorias ortogonales tiene como EDO
$$x'(t) = -\frac{t}{2x(t)};$$

3. cuya solución es
$$x(t)^2 + \frac{t^2}{2} = \lambda, \qquad \lambda > 0.$$
Geométricamente son elipses centradas en $(0, 0)$ y con semiejes de longitud 2 y 1, respectivamente.

Otro tipo de problemas de interés son los relacionados con la geometría de curvas planas regulares. Existen numerosos problemas que involucran propiedades geométricas de estas curvas y pretender hacer una mínima recopilación daría para un tema de una asignatura de Grado, como poco. Nos limitamos a detallar uno de los primeros que se plantean en esta temática.

Ejemplo 1.17. Hallar, si existe, una curva que cumpla que sus rectas normales pasan todas por un mismo punto.

No es difícil convencerse de que, salvo una traslación en el plano, podemos suponer que tal punto es el origen de coordenadas. Sea pues $\alpha(t) = (x(t), y(t))$. En cada instante $t \in I$ la recta normal en $\alpha(t)$ tendrá como dirección un cierto vector unitario $n_\alpha(t)$. De hecho, $n_\alpha(t)$ se puede tomar como
$$n_\alpha(t) = \frac{J\alpha'(t)}{||J\alpha'(t)||}$$

y por tanto $n_\alpha(t)$ es diferenciable. Para que la recta normal de $\alpha(t)$ pase por $(0,0)$, existe $\lambda(t)$ tal que

$$\alpha(t) - \lambda(t) n_\alpha(t) = (0,0).$$

Primero veamos que así definida, $\lambda(t)$ es diferenciable. En efecto,

$$\lambda(t) = \langle \alpha(t), n_\alpha(t) \rangle,$$

y como $\alpha \in C^2$ concluimos que $\lambda \in C^1$. De hecho,

$$|\lambda(t)| = ||\alpha(t)|| = \sqrt{\langle \alpha(t), \alpha(t) \rangle}.$$

Es más, podemos suponer que existe $t_0 \in I$ tal que $\lambda(t_0) \neq 0$ ya que de lo contrario $\alpha(t)$ se reduciría al origen. Restringimos I a I_0 conteniendo a t_0 y salvo renombrar la función suponemos $\lambda(t) > 0$ para todo $t \in I_0$. Ahora bien,

$$\lambda'(t) = \frac{\langle \alpha'(t), \alpha(t) \rangle}{||\alpha(t)||} = 0$$

por ser $\alpha(t) = \lambda(t) n_\alpha(t)$ y en consecuencia $\lambda(t)$ es constante. Esto nos dice que $||\alpha(t)||$ es constante y por tanto $\alpha(t)$ está a distancia constante del origen, lo cual prueba que la imagen $\alpha(I_0)$ está contenida en una circunferencia. Por continuidad, esta propiedad es cierta para todo $t \in I$.

1.6. Isoclinas y campos de direcciones

Las ecuaciones escalares de primer orden expresadas en forma normal $x'(t) = F(t, x(t))$ tienen la siguiente interpretación geométrica. Si $x(t)$ es solución de la EDO, entonces podemos representar la gráfica de $x(t)$ como $t \mapsto (t, x(t))$. Dado $t_0 \in I$, por ser $x \in C^1$ existe la recta tangente a la gráfica de $x(t)$ en $t = t_0$, y su pendiente vale exactamente $x'(t_0) = F(t_0, x(t_0))$. En general, fijado $K \in \mathbb{R}$ podemos considerar el subconjunto

$$M_K = \{(t,x) \in \mathbb{R}^2 \colon F(t,x) = K\}.$$

Puede ser que $M_K = \varnothing$, o que incluso M_K se reduzca a un único punto (por ejemplo, si $F(t,x) = t^2 + x^2$ y $K \leq 0$ se tienen los dos casos mencionados). Supongamos que K es tal que M_K no se reduce a un punto. En tal caso, si $x(t)$ es solución de la EDO y $(t_0, x(t_0)) \in M_K$, la pendiente de la gráfica de la solución $x(t)$ tiene pendiente K en t_0. Variando el parámetro K podemos hacernos una idea del comportamiento de las soluciones y hacer una representación gráfica aproximada de las mismas.

Definición 1.6.1. Con la notación anterior, las curvas que forman el conjunto M_K se les llama *isoclinas*.

Ejemplo 1.18. Estudiemos las isoclinas de las siguientes EDOs.

1. $x'(t) = k,\ k \in \mathbb{R}$.

   La función $F(t,x) = k$ y por tanto

   $$M_K = \{(t,x) \in \mathbb{R}^2 \colon F(t,x) = K\}.$$

   Si $K \neq k$ se tiene $M_K = \varnothing$, y si $K = k$ entonces $M_k = \mathbb{R}^2$. Todas las soluciones son rectas de pendiente k.

2. $x'(t) = x(t)$.

   Ahora $F(t,x) = x$ y por tanto
$$M_K = \{(t,x) \in \mathbb{R}^2 \colon F(t,x) = K\}.$$
   Distinguimos casos.

   - Si $K = 0$ entonces $F(t,x) = 0$ implica $x = 0$, es decir M_0 es el eje t.
   - Si $K > 0$ entonces $F(t,x) = K$ solo tiene sentido para $x > 0$ y su solución es la recta horizontal $x = K$.
   - Si $K < 0$ entonces $F(t,x) = K$ solo tiene sentido para $x < 0$ y su solución es la recta horizontal $x = K$.

   De aquí podemos ver que si $x > 0$ las soluciones son crecientes, mientras que si $x < 0$ las soluciones son decrecientes.

3. $x'(t) = t^2 + x(t)^2$.

   Por último, $F(t,x) = t^2 + x^2$. Observamos que si $K < 0$ entonces $M_K = \varnothing$ y si $K = 0$ entonces $M_K = \{(0,0)\}$. Por tanto suponemos $K > 0$ y se tiene
$$M_K = \{(t,x) \in \mathbb{R}^2 \colon t^2 + x^2 = K\},$$
   esto es la circunferencia centrada en el origen y de radio $\sqrt{K}$. Variando K obtenemos una foliación de $\mathbb{R}^2$ de forma que isoclinas alejadas del origen se corresponden con puntos de derivada cada vez mayores.

El concepto de isoclina consta de fijar un valor $K \in \mathbb{R}$ y preguntarnos por los puntos de la curva $(t, x(t))$ tales que $x'(t) = K$. Un problema de cierta forma inverso nos introduce el concepto de campo de direcciones.

Definición 1.6.2. Sea $x : I \to \mathbb{R}^2$ una solución de la EDO $x'(t) = F(t, x(t))$. El campo de vectores
$$V(t) = (1, x'(t)) = (1, F(t, x(t))$$
se conoce como *campo de direcciones* de la EDO.

El campo de direcciones tiene la siguiente interpretación geométrica: si $x(t)$ es solución de la EDO $x'(t) = F(t, x(t))$, entonces sabemos que la gráfica de $x(t)$ es la curva en el plano $\alpha(t) = (t, x(t))$. La derivada de esta curva plana es $\alpha'(t) = (1, x'(t)) = (1, F(t, x(t))) = V(t)$. Este campo pues nos indica en cada instante cuál es el vector tangente de la gráfica de $x(t)$ y obviamente en particular el valor de $x'(t)$.

Los campos de direcciones son de gran utilidad a la hora de esbozar gráficamente las soluciones de una EDO. En el caso más general en el que tenemos un sistema autónomo plano de orden 1 dado por
$$\begin{cases} x'(t) = f(x(t), y(t)), \\ y'(t) = g(x(t), y(t)), \end{cases}$$
donde $f, g : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ son continuas, la aplicación $t \mapsto (x(t), y(t))$ define una curva parametrizada en $\mathbb{R}^2$, cuyo campo de direcciones se define de manera análoga por $V(t) = (x'(t), y'(t)) = (f(x(t), y(t)), g(x(t), y(t))$. Estudiar las propiedades de este campo de direcciones suele ser una herramienta poderosísima que nos ayuda a deducir las propiedades cualitativas de las soluciones de estos sistemas autónomos de orden 1.

1.7. Sistemas planos y ecuación de las órbitas

Consideremos un sistema plano autónomo

$$\begin{cases} x'(t) = f(x(t), y(t)) \\ y'(t) = g(x(t), y(t)), \end{cases} \tag{1.7.1}$$

donde $f, g \colon \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ son funciones continuas. El problema de Cauchy se tiene sin más que fijar una condición inicial $(x(t_0), y(t_0)) = (x_0, y_0) \in \Omega$. La interpretación geométrica de este sistema es la siguiente: si definimos $F(x, y) = (f(x, y), g(x, y))$ y $\gamma(t) = (x(t), y(t))$, entonces el sistema (1.7.1) se escribe como

$$\gamma'(t) = F(\gamma(t)).$$

Podemos interpretar F como un campo de velocidades o de fuerzas, en el que en cada punto $F(x, y)$ se mide una magnitud vectorial. La ecuación $\gamma'(t) = F(\gamma(t))$ modela una curva cuya tangente en cada punto de la curva coincide exactamente con el campo de velocidades actuando sobre tal punto. Es decir, $\gamma(t)$ modela el movimiento de una partícula cuando sobre ella actúa el campo F.

Definición 1.7.1. Llamaremos *órbita* a toda solución del sistema (1.7.1), y la denotaremos comúnmente por $\gamma(t) = (x(t), y(t))$.

Por ejemplo, la EDO $x''(t) + x(t) = 0$ nos define el sistema

$$\begin{cases} x'(t) = y(t) \\ y'(t) = -x(t). \end{cases}$$

Si consideramos una órbita $\gamma(t) = (x(t), y(t))$ entonces se tiene

$$\frac{d}{dt}\left(x(t)^2 + y(t)^2\right) = 0$$

y por tanto $x(t)^2 + y(t)^2 = k$, siendo $k \in \mathbb{R}$. Obviamente $k \geq 0$ y $k = 0$ si y sólo si $x(t) = y(t) = 0$. Para cada $k > 0$ las órbitas son circunferencias centradas en el origen y de radio $\sqrt{k}$. En particular, las solución $x(t)$ de la EDO original $x''(t) + x(t)$ es una función periódica.

Definición 1.7.2. Se llama *plano de fases* de (1.7.1) al conjunto de todas sus órbitas.

Obviamente el plano de fases estará definido en el mismo abierto en el que estén definidas las funciones f, g. Denotaremos comunmente al plano de fases por Θ.

Dada una órbita $\gamma(t) = (x(t), y(t))$, se puede ver como una curva descrita en paramétricas sobre el plano xy. Un problema que surge es preguntarnos si existe una ecuación diferencial satisfecha por las coordenadas de las órbitas, de manera que una de las funciones, digamos $x(t)$, actúe como variable independiente $x = x(t)$ y la otra función sea una función $y(x)$. Esta pregunta pasa por resolver la cuestión previa de si dado $t_0 \in I$, la órbita $\gamma(t)$ se puede expresar localmente alrededor de $(x_0, y_0) = (x(t_0), y(t_0))$ como un grafo $y(x)$. La respuesta a esta cuestión nos la da el Teorema de la función inversa.

Supongamos $x'(t_0) \neq 0$, lo cual se traduce en que $f(x_0, y_0) \neq 0$. Por ser $x'(t_0) \neq 0$ el Teorema de la función inversa nos asegura que existe $I_0 \subset I$ conteniendo a t_0 tal que la función $x : I_0 \to J$

es invertible. Si denotamos por $\varphi = x^{-1}$ y $s = x(t)$, entonces $\varphi(s) = x^{-1}(x(t))$ es de clase C^1. Además,

$$\varphi'(s) = \frac{1}{x'(\varphi^{-1}(s))} = \frac{1}{x'(t)},$$

para todo $s \in J$. Definimos $\hat{y}(s) = y(\varphi(s))$. Por ser $y(t)$ solución y dado que $t = \varphi(s)$,

$$\begin{aligned}\frac{d}{ds}\hat{y}(s) &= \frac{d}{ds}y(\varphi(s)) = y'(\varphi(s))\varphi'(s) = \frac{y'(\varphi(s))}{x'(\varphi(s))}\\ &= \frac{g(x(\varphi(s)), y(\varphi(s)))}{f(x(\varphi(s)), y(\varphi(s)))}\\ &= \frac{g(s, \hat{y}(s))}{f(s, \hat{y}(s))}.\end{aligned}$$

Sin más que renombrar la variable s por x. Concluimos que $\hat{y} : J \to \mathbb{R}$ es solución de la EDO

$$y'(x) = \frac{g(x, y(x))}{f(x, y(x))}.$$

Aquí se ve perfectamente la necesidad de que $f(x_0, y_0) \neq 0$ ya que por continuidad $f(x, y) \neq 0$ en un entorno de este punto.

1.7.1. El plano de fases de un sistema autónomo

El estudio de sistemas planos es uno de los pilares de la *teoría cualitativa de las Ecuaciones Diferenciales*. Esta disciplina estudia el comportamiento de las soluciones sin conocerlas explícitamente, mediante el análisis de las órbitas. Supongamos que escribimos el sistema plano (1.7.1) como la ecuación vectorial

$$\gamma'(t) = F(\gamma(t)),$$

Por ejemplo, sin necesidad de conocer explícitamente la solución $\gamma(t)$, su movimiento en cada instante viene determinada por el campo de direcciones $V(t) = F(\gamma(t))$ y en última instancia por el signo de las funciones $f(x(t), y(t))$ y $g(x(t), y(t))$. El conocimiento del comportamiento cualitativo nos puede permitir inferir propiedades analíticas de las componentes de $\gamma(t)$.

Definición 1.7.3. Dado el plano de fases Θ de (1.7.1), llamamos *región crítica* al subconjunto del plano de fases

$$\mathcal{R} = \{(x, y) \in \Theta \colon f(x, y)g(x, y) = 0\}.$$

Llamamos *puntos de equilibrio* a aquellos $(x_0, y_0) \in \Theta$ tales que $f(x_0, y_0) = g(x_0, y_0) = 0$.

Los puntos de equilibrio se caracterizan por ser soluciones constantes de (1.7.1). Por ser f, g funciones continuas, la región crítica $\mathcal{R}$ es un conjunto cerrado en Θ y por tanto $\Theta - \mathcal{R}$ es un conjunto abierto, posiblemente disconexo. En cada componente conexa de $\Theta - \mathcal{R}$ las coordenadas de una órbita son funciones estrictamente monótonas y por tanto el movimiento de una órbita está unívocamente determinado.

Definición 1.7.4. Llamaremos *regiones de monotonía* a las componentes conexas de $\Theta - \mathcal{R}$.

Ejemplo 1.19. En el sistema plano

$$\begin{cases} x'(t) = y(t) \\ y'(t) = -x(t) \end{cases}$$

se tiene un único punto de equilibrio: el origen $(0,0)$. Además, la región crítica son los ejes coordenados y por tanto las regiones de monotonía son los cuadrantes abiertos.

Todavía no hemos hecho un estudio exhaustivo de la existencia y unicidad de un problema de Cauchy general de la forma $\gamma'(t) = F(\gamma(t))$, estudio que se llevará a cabo de forma sistemática en el tema 4. Sin embargo, si damos por buena esta existencia y unicidad llegamos al siguiente resultado.

Teorema 1.7.5. *Supongamos que el problema de Cauchy*

$$\begin{cases} \gamma'(t) = F(\gamma(t)), \\ \gamma(t_0) = (x_0, y_0) \end{cases}$$

tiene existencia y unicidad para todo $t_0 \in I$ y $(x_0, y_0) \in \Theta$. Entonces,

1. *Las órbitas definen una foliación de Θ, esto es, dado cualquier punto $p_0 \in \Theta$ existe una órbita pasando por p_0.*

2. *Dos órbitas distintas no pueden intersecarse. Esto es, si γ_1, γ_2 son dos órbitas, $\gamma_i : I_i \to \mathbb{R}^2,\ i = 1, 2$, y existe $t_0 \in I_1 \cap I_2$ tal que $\gamma_1(t_0) = \gamma_2(t_0)$, entonces $\gamma_1(t) = \gamma_2(t)$ para todo $t \in I_1 = I_2$.*

1.8. Ecuaciones en Derivadas Parciales

En este curso únicamente veremos ecuaciones cuya incógnita es una función dependiendo de una sola variable y donde aparecen la propia función y alguna de sus derivadas (de cualquier orden). Otro problema, cuya dificultad y abordaje se escapa a los contenidos de este curso, es el de considerar una ecuación cuya incógnita es una función de varias variables y donde ahora pueden aparecer las derivadas parciales de tal función.

Antes de nada, necesitamos introducir notación. Sean $n \in \mathbb{N}$ y $\alpha = (\alpha_1, ..., \alpha_n),\ \alpha_k \in \mathbb{N} \cup \{0\}$ y definimos $|\alpha| = \sum_k \alpha_k$. Sean $\Omega \subset \mathbb{R}^n$, con coordenadas $x = (x_1, ..., x_n)$, y $u \in C^m(\Omega)$. Definimos

$$\partial^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, \qquad D^k u = (\partial^\alpha u)_{|\alpha|=k}, \tag{1.8.1}$$

donde $D^k u$ es un vector formado por las distintas derivadas $\partial^\alpha u$ con $|\alpha| = k$. En particular,

$$Du = (\frac{\partial u}{\partial x_1}, ..., \frac{\partial u}{\partial x_n})$$

no es más que el gradiente de u.

Definición 1.8.1. Una Ecuación en Derivadas Parciales (EDP) de orden $k \in \mathbb{N}$ es una expresión de la forma

$$\Phi(x, u(x), D^1 u(x), ..., D^k u(x)) = 0,$$

donde $\Phi : \Omega \subset \mathbb{R}^p \to \mathbb{R}$ es una función continua y $u(x)$ es una función de clase C^k.

Las EDPs para funciones de dos variables, $u(x,y)$, aparecen de forma natural en numerosos problemas de índole geométrica. Como localmente toda superficie puede describirse como un grafo $(x,y,u(x,y))$, las propiedades geométricas de la superficie describirán una EDP satisfecha por la función u. Análogamente, toda solución u de una EDP describe el grafo de una superficie cuyas propiedades geométricas se pueden inferir a través de las propiedades analíticas de la EDP. Esto hace que exista una relación muy fuerte entre la Geometría Diferencial y el Análisis de Ecuaciones en Derivadas Parciales. Esta rama de las Matemáticas se conoce como Análisis Geométrico.

Las EDPs de orden 2 satisfechas por funciones $u(x,y)$ tienen un interés especial dentro del Análisis Geométrico. La notación que introducimos a continuación tiene su interés propio y será útil para la definición posterior. Dada $u \in C^2(\Omega)$ con $\Omega \subset \mathbb{R}^2$, definimos

$$D_1 u = \frac{\partial u}{\partial x}, \qquad D_2 u = \frac{\partial u}{\partial y}, \qquad D_{ij} u = \partial^{(i,j)} u = \frac{\partial^2 u}{\partial x^i \partial y^j}, \quad i+j=2.$$

No confundir con la notación de la Ecuación (1.8.1), para la cual se tiene $D^1 u = Du = (D_1 u, D_2 u)$.

Definición 1.8.2. Supongamos $u : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ y consideremos EDPs de orden 2. Definimos los siguientes tipos de EDPs.

1. La EDP es lineal si se escribe como
$$\sum_{i+j=2} a_{ij}(x,y) D_{ij} u + \sum_{i=1}^{2} a_i(x,y) D_i u + a_0(x,y) u + b(x,y) = 0.$$

2. La EDP es semi-lineal si se escribe como
$$\sum_{i+j=2} a_{ij}(x,y) D_{ij} u + b(x,y,u,Du) = 0.$$

3. La EDP es cuasi-lineal si se escribe como
$$\sum_{i+j=2} a_{ij}(x,y,u,Du) D_{ij} u + b(x,y,u,Du) = 0.$$

4. La EDP es totalmente no lineal si se escribe como
$$\Phi(x,y,u,Du,D^2 u) = 0,$$
es decir no tiene ningún tipo de linealidad previamente definidas.

Ejemplo 1.20. Supongamos $u = u(t,x)$. Algunos ejemplos de EDPs interesantes son los siguientes.

1. La ecuación de ondas. Describe el comportamiento ondulatorio de las ondas físicas, ya sean sonoras, mecánicas o electromagnéticas.
$$u_{tt} = c^2 u_{xx}.$$

2. La ecuación del calor. Describe la distribución de la temperatura según la posición y el tiempo.
$$u_t = c^2 u_{xx}$$

3. Ecuación de transporte. Modela la concentración de una sustancia que viaja en un fluido a velocidad constante.
$$u_t + cu_x = 0.$$

4. Una EDP geométrica:
$$\frac{u_{xx}u_{yy} - u_{xy}^2}{(1 + u_x^2 + u_y^2)^2} = K, \qquad K \in \mathbb{R}.$$
Describe superficies gráficas $(x, y, u(x, y))$ con *curvatura constante*. Esta EDP es totalmente no lineal.

5. La siguiente EDP modela las superficies que tienen localmente área mínima entre todas las superficies con su mismo borde:
$$\mathrm{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right) = 0,$$
donde div es el operador divergencia usual en $\mathbb{R}^2$ y $Du = (u_x, u_y)$. En general, si $H_0 \in \mathbb{R}$ las superficies que localmente cumplen la EDP
$$\mathrm{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right) = 2H_0,$$
son superficies que minimizan el área encerrando un volumen constante. Estas EDPs son cuasilineales.

La resolución explícita de EDPs es muchísimo más difícil que en el caso de EDOs.

1.9. Breve introducción al Cálculo de Variaciones

El Cálculo de Variaciones es una disciplina que busca optimizar funcionales; esto es, operadores definidos en espacios de funciones y con valores reales. Son una vasta generalización del cálculo elemental de máximos y mínimos de funciones reales de una variable real. Sus aplicaciones abarcan la física teórica, la física matemática, la biología, la aerodinámica o la ingeniería, entre otras tantas disciplinas.

Su estudio especializado formal tuvo lugar durante el siglo XVIII principalmente desarrollado por Euler y Lagrange, aunque su motivación original data del siglo XVII cuando Johann Bernoulli propuso en 1696 el *problema de la braquistocrona*: dados dos puntos A, B en un plano vertical, estudiar la curva que traza un punto que va de A a B en el menor tiempo posible, cuando sobre éste actúa únicamente la fuerza de la gravedad. A continuación, damos un breve recorrido sobre este problema y su resolución.

Tanto Johann como su hermano Jakob obtuvieron la misma solución, pero ésta era incompleta: Johann tuvo errores en la dedución y Jakob simplificó el problema considerando menos constantes que las que realmente existen. Johann dio un plazo de seis meses para su resolución, y sin recibir ninguna demostración satisfactoria amplió el plazo un año y medio más. En enero de 1697, Isaac Newton encontró en su buzón una carta de Johann planteándole el problema de la braquistocrona; lo resolvió esa misma noche usando su Cálculo Diferencial y envió la solución de forma anónima. No obstante, al recibir la demostración, Johann reconoció inmediatamente el estilo de Newton afirmando *se conoce al león por su garra*.

El primer matemático que desarrolló formalmente una teoría de Cálculo de Variaciones fue Euler en 1733 en su *Elementa Calculi Variationum*, lo cual dio nombre a la disciplina. Por otra parte, en 1755 y contando con sólo 19 años de edad, Lagrange envió una carta a Euler sobre un problema de isoperimetría mediante una técnica que involucraba el Cálculo de Variaciones. Euler reconoció la genialidad del método, el cual puso a Lagrange en primera línea entre los matemáticos de su época.

A continuación, introducimos la piedra angular de esta disciplina: la ecuación de Euler-Lagrange. Sea $X \subset C^1([a,b])$ y $L : [a,b] \times TX \to \mathbb{R}$, una función dos veces diferenciable, donde TX denota el fibrado tangente de X. Definimos el funcional

$$\mathcal{L} : X \to \mathbb{R}, \qquad \mathcal{L}[f] = \int_a^b L(t, f(t), f'(t))\, dt.$$

Denotamos comúnmente las variables de L por $L(t,p,q)$. La interpretación y motivación física es la siguiente. La función L se llama *lagrangiano* y suele definirse mediante las magnitudes físicas que actúan sobre una partícula $t \mapsto (t, f(t))$ dependiendo de su posición $(t, f(t))$ y de su velocidad $f'(t)$.

La pregunta central de esta teoría es: ¿existe alguna función $f \in X$ tal que $\mathcal{L}[f] \leq \mathcal{L}[g]$ para cualquier otra función $g \in X$? Es muy común imponer ciertas condiciones de contorno sobre una eventual solución. Es decir, que junto con la minimización del funcional $\mathcal{L}$ se impone que la solución f satisfaga $f(a) = A$ y $f(b) = B$, por ejemplo. En el caso 1-dimensional, sabemos que si una función diferenciable f tiene un extremo relativo en algún punto interior x_0, entonces necesariamente $f'(x_0) = 0$.

El siguiente objetivo es dar una caracterización necesaria de una función f que optimice $\mathcal{L}$ con las condiciones de contorno anteriores. Redefinimos X imponiendo además que para toda $g \in X$ se satisface $g(a) = A$ y $g(b)B = B$ y sea $f \in X$ un extremo de $\mathcal{L}$. Dada η una función diferenciable tal que $\eta(a) = \eta(b) = 0$ y $\epsilon > 0$, la función $f + \epsilon\eta \in X$ y por tanto la función real de variable real

$$\varphi(\epsilon) = \mathcal{L}[f + \epsilon\eta] = \int_a^b L(t, (f(t) + \epsilon\eta(t), f'(t) + \epsilon\eta'(t))\, dt,$$

está bien definida. Si f es un extremo de $\mathcal{L}$ entonces φ tiene un extremo relativo en 0 si ϵ es suficientemente pequeño y por tanto $\varphi'(0) = 0$. Vamos a calcular esta derivada real.

$$\begin{aligned}
\varphi'(\epsilon) = \frac{d\mathcal{L}}{d\epsilon} &= \frac{d}{d\epsilon}\int_a^b L(t, f(t) + \epsilon\eta(t), f'(t) + \epsilon\eta'(t))\, dt \\
&= \int_a^b \frac{d}{d\epsilon} L(t, f(t) + \epsilon\eta(t), f'(t) + \epsilon\eta'(t))\, dt \\
&= \int_a^b \eta(t)\frac{\partial L}{\partial p}(t, f(t) + \epsilon\eta(t), f'(t) + \epsilon\eta'(t)) + \eta'(t)\frac{\partial L}{\partial q}(t, f(t) + \epsilon\eta(t), f'(t) + \epsilon\eta'(t))\, dt.
\end{aligned}$$

Si hacemos $\epsilon = 0$ tenemos

$$\varphi'(0) = \int_a^b \eta(t)\frac{\partial L}{\partial p}(t, f(t), f'(t)) + \eta'(t)\frac{\partial L}{\partial q}(t, f(t), f'(t))\, dt = 0.$$

El problema es que aparece la función η y su derivada y nos gustaría tener todo dependiendo únicamente de η. Para tal fin, hacemos uso de la integración por partes. En primer lugar,

$$\frac{d}{dt}\left(\eta(t)\frac{\partial L}{\partial q}(t, f(t), f'(t))\right) = \eta'(t)\frac{\partial L}{\partial q}(t, f(t), f'(t)) + \eta(t)\frac{d}{dt}\frac{\partial L}{\partial q}(t, f(t), f'(t)).$$

Por otra parte,

$$\int_a^b \frac{d}{dt}\left(\eta(t)\frac{\partial L}{\partial q}(t, f(t), f'(t))\right) dt = 0$$

ya que $\eta(a) = \eta(b) = 0$. Sustituyendo en la expresión de $\varphi'(0)$ tenemos

$$\varphi'(0) = \int_a^b \eta(t)\left(\frac{\partial L}{\partial p}(t, f(t), f'(t)) - \frac{d}{dt}\frac{\partial L}{\partial q}(t, f(t), f'(t))\right) dt. \tag{1.9.1}$$

Y esto se cumple para toda $\eta \in C^1([a,b])$ tal que $\eta(a) = \eta(b) = 0$. Llegados a este punto, hacemos uso del *lema fundamental del Cálculo de Variaciones*.

Teorema 1.9.1 (Lema fundamental del Cálculo de Variaciones)**.** *Sea $f : (a,b) \to \mathbb{R}$ continua tal que*

$$\int_a^b f(t)h(t)\, dt = 0,$$

para toda función $h : (a,b) \to \mathbb{R}$ con soporte compacto. Entonces, $f(t) = 0$ para todo $t \in (a,b)$.

Demostración. Supongamos por reducción al absurdo que f no es idénticamente nula. Existe $t_0 \in (a,b)$ tal que $f(t_0) \neq 0$. Por continuidad, existe $\epsilon > 0$ tal que $f(t)$ tiene el mismo signo que $f(t_0)$ en $(t_0 - \epsilon, t_0 + \epsilon)$. Sea h tal que $h(t) = 0$ para todo $t \in (a,b)\backslash(t_0 - \epsilon, t_0 + \epsilon)$ y h tiene el mismo signo que f en $(t_0 - \epsilon, t_0 + \epsilon)$. Entonces,

$$0 = \int_a^b f(t)h(t)\, dt = \int_{t_0-\epsilon}^{t_0+\epsilon} f(t)h(t)\, dt > 0,$$

llegando a contradicción. □

Volviendo a la ecuación (1.9.1) y aplicando el lema fundamental del Cálculo de Variaciones concluimos que una condición necesaria para que f sea extremo de $\mathcal{L}$ es que se cumpla la siguiente ecuación en derivadas parciales

$$\frac{\partial L}{\partial p}(t, f(t), f'(t)) - \frac{d}{dt}\frac{\partial L}{\partial q}(t, f(t), f'(t)) = 0. \tag{1.9.2}$$

Definición 1.9.2. La ecuación (1.9.2) se conoce como *ecuación de Euler-Lagrange*.

Este resultado es fácilmente generalizable, siguiendo las mismas técnicas, a los siguientes casos de lagrangianos más generales. Por agilizar la notación, se omitirá la dependencia explícita de las coordenadas de L.

1. El lagrangiano L involucra una función y derivadas de orden superior. Si

$$L = (t, f(t), f'(t), ..., f^{n)}(t) = L(t, p, q_1, ..., q_n),$$

entonces integando por partes n veces se obtiene la ecuación de Euler-Lagrange

$$\frac{\partial L}{\partial p} - \frac{d}{dt}\frac{\partial L}{\partial q_1} + \frac{d^2}{dt^2}\frac{\partial L}{\partial q_2} + \cdots + (-1)^n\frac{d^n}{dt^n}\frac{\partial L}{\partial q_n} = 0.$$

2. El lagrangiano L involucra varias funciones con derivadas de primer orden. Si

$$L = (t, f_1(t), ..., f_n(t), f_1'(t), ..., f_n'(t)) = L(t, p_1, ..., p_n, q_1, ..., q_n),$$

entonces obtenemos un sistema de n ecuaciones de Euler-Lagrange,

$$\frac{\partial L}{\partial p_k} - \frac{d}{dt}\frac{\partial L}{\partial q_k} = 0, \qquad k = 1, ..., n.$$

3. El lagrangiano L involucra una función de varias variables, con derivadas parciales de primer orden. Por agilizar la notación, escribimos $\mathbf{t} = (t_1, ..., t_n)$. Si

$$L(\mathbf{t}, f(\mathbf{t}), f_1(\mathbf{t}), ..., f_n(\mathbf{t})) = L(\mathbf{t}, p, q_1, ..., q_n), \quad f_k(\mathbf{t}) = \frac{\partial f}{\partial t_k}(\mathbf{t}),$$

entonces obtenemos una ecuación en derivadas parciales

$$\frac{\partial L}{\partial p} - \sum_{k=1}^{n} \frac{\partial}{\partial t_k}\frac{\partial L}{\partial q_k} = 0.$$

El caso $n = 2$ es de especial interés por modelizar problemas que satisfacen localmente superficies inmersas en $\mathbb{R}^3$. Por ejemplo, sea $\Omega \subset \mathbb{R}^2$ un abierto y $u : \Omega \to \mathbb{R}$ una función. Ahora estamos imponiendo $u = \varphi$ en $\partial\Omega$, donde $\varphi : \partial\Omega \to \mathbb{R}$ es continua. Si consideramos la superficie $M_u = (x, y, u(x, y))$ dada por el grafo de u, su área viene dada por

$$A(M_u) = \int_\Omega \sqrt{1 + |\nabla u|^2}.$$

Si queremos estudiar qué superficies tienen menor área teniendo como contorno fijo la curva $(x, y, \varphi(x, y)),\ (x, y) \in \partial\Omega$, defiinimos el funcional

$$\mathcal{L}[u] = \int_\Omega \sqrt{1 + |\nabla u|^2}.$$

Todo punto crítico de $\mathcal{L}$ es una función cuya gráfica define una superficie conocida como *superficie mínima*. En este caso, la ecuación de Euler-Lagrange es

$$\frac{\partial}{\partial x}\left(\frac{u_x}{\sqrt{1 + u_x^2 + u_y^2}}\right) + \frac{\partial}{\partial y}\left(\frac{u_y}{\sqrt{1 + u_x^2 + u_y^2}}\right) = \operatorname{div}\left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}}\right) = 0,$$

siendo div el operador divergencia.

1.9.1. Aplicación al cálculo de geodésicas en una superficie

Un problema clásico en Geometría Diferencial es el de calcular curvas en una superficie dada que minimicen, al menos de forma local, la distancia entre dos puntos fijos. Una tal curva se conoce como *geodésica* y viene caracterizada de diversas formas, todas ellas equivalentes. Aunque se necesitan de conocimientos que se adquirirán en cursos posteriores para poder entender y justificar todo el planteamiento teórico que expondremos a continuación, nos planteamos deducir la expresión local del sistema de EDOs que tiene que satisfacer una geodésica en una superficie de $\mathbb{R}^3$ haciendo uso de las ecuaciones de Euler-Lagrange.

Sea M una superficie en $\mathbb{R}^3$, consideramos $\psi : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^3$ una parametrización, $\psi = \psi(u, v)$, y $\gamma : I \subset \mathbb{R} \to M$ una curva. Más aún, suponemos que $\gamma(t) \in \psi(\Omega)$ para todo $t \in I$ y por tanto podemos expresar localmente γ de la forma

$$\gamma(t) = \psi(u(t), v(t)), \quad (u(t), v(t)) \in \Omega,$$

para ciertas funciones $u(t), v(t)$. La derivada de γ en coordenadas es

$$\gamma'(t) = u'(t)\psi_u(u(t), v(t)) + v'(t)\psi_v(u(t), v(t)).$$

Por agilizar la notación, a partir de ahora omitiremos las variables en las funciones involucradas, entendiendo sobre qué variable actúa cada función. Si denotamos

$$g_{11} = \langle \psi_u, \psi_u \rangle, \qquad g_{12} = g_{21} = \langle \psi_u, \psi_v \rangle, \qquad g_{22} = \langle \psi_v, \psi_v \rangle, \tag{1.9.3}$$

se tiene que el elemento de longitud, ds^2, en M viene dado por

$$ds^2 = g_{11}du^2 + 2g_{12}du\, dv + g_{22}dv^2.$$

En coordenadas, este elemento de longitud es

$$ds^2 = \langle \gamma', \gamma' \rangle = u'^2 g_{11}(u, v) + 2u'v' g_{12}(u, v) + v'^2 g_{22}(u, v).$$

Sean $[a, b] \subset I$ y $A = \gamma(a), B = \gamma(b) \in M$. La longitud del segmento de γ que une A con B es

$$l(\gamma) = \int_\gamma ds = \int_a^b \sqrt{u'^2 g_{11}(u, v) + 2u'v' g_{12}(u, v) + v'^2 g_{22}(u, v)}.$$

En tal caso, podemos considerar el problema de obtener la curva cuya longitud uniendo los puntos A y B es mínima. Como optimizar una función positiva y su raíz es equivalente, podemos tomar como lagrangiano

$$L(t, u, v, u', v') = u'^2 g_{11}(u, v) + 2u'v' g_{12}(u, v) + v'^2 g_{22}(u, v).$$

En este punto, debemos incidir en que las incógnitas son las funciones $u(t), v(t)$. Las ecuaciones de Euler-Lagrange se traducen en el sistema de EDOs

$$\begin{cases} \dfrac{\partial L}{\partial u} - \dfrac{d}{dt}\dfrac{\partial L}{\partial u'} = 0, \\ \dfrac{\partial L}{\partial v} - \dfrac{d}{dt}\dfrac{\partial L}{\partial v'} = 0. \end{cases} \tag{1.9.4}$$

Vamos a estudiar cada uno de los términos de las ecuaciones. Por una parte, las derivadas más sencillas son $\partial L/\partial u'$, $\partial L/\partial v'$, que se traducen en

$$\frac{\partial L}{\partial u'} = 2u' g_{11}(u, v) + 2v' g_{12}(u, v), \qquad \frac{\partial L}{\partial v'} = 2v' g_{22}(u, v) + 2u' g_{12}(u, v).$$

Para derivar L respecto de u y v, recordemos que los coeficientes de la métrica g_{ij} vienen dados por (1.9.3).

Llegados a este punto, debemos expresar las derivadas $\psi_{uu}, \psi_{uv}, \psi_{vv}$ en la base de $\mathbb{R}^3$ dada por $\{\psi_u, \psi_v, N\}$, donde N es un campo de vectores ortogonal a la superficie M localmente en $\psi(\Omega)$; basta definir por ejemplo

$$N = \frac{\psi_u \times \psi_v}{|\psi_u \times \psi_v|}.$$

En tal caso,

$$\begin{cases} \psi_{uu} = \Gamma_{11}^1 \psi_u + \Gamma_{11}^2 \psi_v + b_{11} N, \\ \psi_{uv} = \Gamma_{12}^1 \psi_u + \Gamma_{12}^2 \psi_v + b_{12} N, \\ \psi_{vv} = \Gamma_{22}^1 \psi_u + \Gamma_{22}^2 \psi_v + b_{22} N. \end{cases}$$

Como $\psi_{uv} = \psi_{vu}$ se tiene $\Gamma_{12}^k = \Gamma_{21}^k$ y $b_{12} = b_{21}$. Los coeficientes Γ_{ij}^k se conocen como *símbolos de Christoffel* de M y nos dicen cómo se curva M de forma intrínseca. Los coeficientes b_{ij} se conocen como los *coeficientes de la segunda forma fundamental* y nos dicen cómo se curva M cuando la *vemos* en el espacio ambiente $\mathbb{R}^3$. El lector perspicaz habrá percibido una posible contradicción: hemos definido los símbolos de Christoffel mediante derivadas en el ambiente, pero afirmamos que los mismos nos dicen cómo se curva M de forma intrínseca. Esto es porque tales funciones sólo dependen de la métrica g_{ij} de M y no de cómo M se *encaja* en el ambiente.

Ahora podemos derivar los coeficientes g_{ij} respecto de u, v, obteniendo

$$\begin{aligned}
\frac{\partial g_{11}}{\partial u} &= 2\langle \psi_{uu}, \psi_u \rangle = 2(\Gamma_{11}^1 g_{11} + \Gamma_{11}^2 g_{12}), \\
\frac{\partial g_{12}}{\partial u} &= \langle \psi_{uu}, \psi_v \rangle + \langle \psi_{uv}, \psi_u \rangle = \Gamma_{11}^1 g_{12} + \Gamma_{11}^2 g_{22} + \Gamma_{12}^1 g_{11} + \Gamma_{12}^2 g_{12}, \\
\frac{\partial g_{22}}{\partial u} &= 2\langle \psi_{uv}, \psi_v \rangle = 2(\Gamma_{12}^1 g_{12} + \Gamma_{12}^2 g_{22}), \\
\frac{\partial g_{11}}{\partial v} &= 2\langle \psi_{uv}, \psi_u \rangle = 2(\Gamma_{12}^1 g_{11} + \Gamma_{12}^2 g_{12}), \\
\frac{\partial g_{12}}{\partial v} &= \langle \psi_{uv}, \psi_v \rangle + \langle \psi_{vv}, \psi_u \rangle = \Gamma_{12}^1 g_{12} + \Gamma_{12}^2 g_{22} + \Gamma_{22}^1 g_{11} + \Gamma_{22}^2 g_{12}, \\
\frac{\partial g_{22}}{\partial v} &= 2\langle \psi_{vv}, \psi_v \rangle = 2(\Gamma_{22}^1 g_{12} + \Gamma_{22}^2 g_{22}).
\end{aligned}$$

De aquí podemos justificar que los símbolos de Christoffel sólo dependen de la métrica g_{ij}. Por ejemplo, sustituyendo la cuarta ecuación en la segunda y junto con la primera se tiene el sistema

$$\begin{cases} \dfrac{(g_{11})_u}{2} = \Gamma_{11}^1 g_{11} + \Gamma_{11}^2 g_{12}, \\ (g_{12})_u - \dfrac{(g_{11})_v}{2} = \Gamma_{11}^1 g_{12} + \Gamma_{11}^2 g_{22}, \end{cases}$$

o matricialmente

$$\begin{pmatrix} \dfrac{(g_{11})_u}{2} \\ (g_{12})_u - \dfrac{(g_{11})_v}{2} \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix} \begin{pmatrix} \Gamma_{11}^1 \\ \Gamma_{11}^2 \end{pmatrix}.$$

Aquí los subíndices $(\cdot)_u, (\cdot)_v$ denotan la derivada parcial respecto de u, v respectivamente. Como los coeficientes g_{ij} definen una métrica en una superficie, se tiene $\det(g_{ij}) > 0$ y en particular podemos despejar los coeficientes Γ_{11}^k en función de la métrica y sus derivadas.

Ya estamos en condiciones para poder derivar L respecto de u, obteniendo

$$\begin{aligned}\frac{\partial L}{\partial u} &= u'^2\frac{\partial g_{11}}{\partial u} + 2u'v'\frac{\partial g_{12}}{\partial u} + v'^2\frac{\partial g_{22}}{\partial u}\\ &= 2u'^2(\Gamma_{11}^1 g_{11} + \Gamma_{11}^2 g_{12}) + 2u'v'(\Gamma_{11}^1 g_{12} + \Gamma_{11}^2 g_{22} + \Gamma_{12}^1 g_{11} + \Gamma_{12}^2 g_{12}) + 2v'^2(\Gamma_{12}^1 g_{12} + \Gamma_{12}^2 g_{22}),\\ &= 2(\Gamma_{11}^1(u'^2 g_{11} + u'v'g_{12}) + \Gamma_{11}^2(u'^2 g_{12} + u'v'g_{22}) + \Gamma_{12}^1(u'v'g_{11} + v'^2 g_{12}) + \Gamma_{12}^2(u'v'g_{12} + v'^2 g_{22})).\end{aligned}$$

Un razonamiento similar nos permite concluir

$$\frac{\partial L}{\partial v} = 2(\Gamma_{12}^1(u'^2 g_{11}+u'v'g_{12})+\Gamma_{12}^2(u'^2 g_{12}+u'v'g_{22})+\Gamma_{22}^1(u'v'g_{11}+v'^2 g_{12})+\Gamma_{22}^2(u'v'g_{12}+v'^2 g_{22})).$$

Notemos que las derivadas parciales coinciden salvo en los símbolos de Christoffel.

Por otra parte, debemos derivar respecto de t las derivadas parciales $\partial L/\partial u'$, $\partial L/\partial v'$ y eventualmente derivar respecto de t los coeficientes $g_{ij}(u(t), v(t))$. Podemos intentar derivar todos estos coeficientes *a la vez*, usando la siguiente notación. Si identificamos simbólicamente $u = 1$ y $v = 2$, entonces $\psi_1 \equiv \psi_u$, $\psi_2 \equiv \psi_v$ y así sucesivamente. Por tanto,

$$\begin{aligned}\frac{d}{dt}g_{ij}(u(t), v(t)) &= \frac{d}{dt}\langle\psi_i, \psi_j\rangle\\ &= \langle u'\psi_{1i} + v'\psi_{2i}, \psi_j\rangle + \langle u'\psi_{1j} + v'\psi_{2j}, \psi_i\rangle\\ &= u'(\Gamma_{1i}^1 g_{1j} + \Gamma_{1i}^2 g_{2j}) + v'(\Gamma_{2i}^1 g_{1j} + \Gamma_{2i}^2 g_{2j}) + u'(\Gamma_{1j}^1 g_{1i} + \Gamma_{1j}^2 g_{2i}) + v'(\Gamma_{2j}^1 g_{1i} + \Gamma_{2j}^2 g_{2i})\\ &= u'(\Gamma_{1i}^1 g_{1j} + \Gamma_{1i}^2 g_{2j} + \Gamma_{1j}^1 g_{1i} + \Gamma_{1j}^2 g_{2i}) + v'(\Gamma_{2i}^1 g_{1j} + \Gamma_{2i}^2 g_{2j} + \Gamma_{2j}^1 g_{1i} + \Gamma_{2j}^2 g_{2i}).\end{aligned}$$

Ahora podemos calcular las derivadas respecto de t de las parciales de L respecto de u', v'.

$$\begin{aligned}\frac{d}{dt}\frac{\partial L}{\partial u'} &= \frac{d}{dt}(2u'g_{11} + 2v'g_{12}) = 2(u''g_{11} + u'\frac{d}{dt}g_{11} + v''g_{12} + v'\frac{d}{dt}g_{12}) = 2(u''g_{11} + v''g_{12}\\ &+ u'(u'((\Gamma_{11}^1 g_{11} + \Gamma_{11}^2 g_{21} + \Gamma_{11}^1 g_{11} + \Gamma_{1j}^2 g_{21})) + v'(\Gamma_{21}^1 g_{11} + \Gamma_{21}^2 g_{21} + \Gamma_{21}^1 g_{11} + \Gamma_{21}^2 g_{21}))\\ &+ v'(u'((\Gamma_{11}^1 g_{12} + \Gamma_{11}^2 g_{22} + \Gamma_{12}^1 g_{11} + \Gamma_{12}^2 g_{21})) + v'(\Gamma_{21}^1 g_{12} + \Gamma_{21}^2 g_{22} + \Gamma_{22}^1 g_{11} + \Gamma_{22}^2 g_{21}))),\\ \frac{d}{dt}\frac{\partial L}{\partial v'} &= \frac{d}{dt}(2v'g_{22} + 2u'g_{12}) = 2(v''g_{22} + v'\frac{d}{dt}g_{22} + u''g_{12} + u'\frac{d}{dt}g_{12}) = 2(v''g_{22} + u''g_{12}\\ &+ v'(u'((\Gamma_{12}^1 g_{12} + \Gamma_{12}^2 g_{22} + \Gamma_{12}^1 g_{12} + \Gamma_{12}^2 g_{22})) + v'(\Gamma_{22}^1 g_{12} + \Gamma_{22}^2 g_{22} + \Gamma_{22}^1 g_{12} + \Gamma_{22}^2 g_{22}))\\ &+ u'(u'((\Gamma_{11}^1 g_{12} + \Gamma_{11}^2 g_{22} + \Gamma_{12}^1 g_{11} + \Gamma_{12}^2 g_{21})) + v'(\Gamma_{21}^1 g_{12} + \Gamma_{21}^2 g_{22} + \Gamma_{22}^1 g_{11} + \Gamma_{22}^2 g_{2i}))).\end{aligned}$$

Finalmente, podemos calcular el sistema de ecuaciones de Euler-Lagrange (1.9.4) sin más que sustituir cada una de las expresiones calculadas. Un pequeño comentario: llegados a este punto, lo normal es usar software especializado para que éste calcule por nosotros lo que es cierto que es una cuenta larga y tediosa. No obstante, merece la pena hacer la simplificación a mano, al menos una vez en la vida, sólo por el gustazo de ver cómo se cancelan los términos y comprobar que, aunque a priori parezca que el resultado final será una expresión muy complicada y extensa, todo se acaba simplificando y se llega al siguiente resultado.

Teorema 1.9.3. *Sea M una superficie en $\mathbb{R}^3$ y $\psi : \Omega \to \mathbb{R}^3$ una parametrización. Una curva $\gamma(t) = \psi(u(t), v(t))$ es geodésica en M si y sólo si sus componentes $u(t), v(t)$ resuelven el sistema*

$$\begin{cases} u'' + \Gamma_{11}^1 u'^2 + 2\Gamma_{12}^1 u'v' + \Gamma_{22}^1 v'^2 = 0,\\ v'' + \Gamma_{11}^2 u'^2 + 2\Gamma_{12}^2 u'v' + \Gamma_{22}^2 v'^2 = 0,\end{cases}$$

donde Γ_{ij}^k son los símbolos de Christoffel de M.

Por ejemplo, un plano en $\mathbb{R}^3$ tiene todos sus símbolos de Christoffel idénticamente nulos y por tanto toda curva plana que sea geodésica cumple $u''(t) = v''(t) = 0$ para todo t. Esto nos dice que

$$u(t) = A + tv_1, \quad v(t) = B + tv_2, \qquad A, B, v_1, v_2 \in \mathbb{R}.$$

Supongamos sin perder generalidad que el plano es $z = 0$ y una parametrización global es simplemente $\psi(u, v) = (u, v, 0)$. En consecuencia, toda geodésica tiene la forma

$$\gamma(t) = \psi(u(t), v(t)) = (A, B, 0) + t(v_1, v_2, 0).$$

Esta no es más que la fórmula de una recta pasando por el punto $(A, B, 0)$ y con vector director $(v_1, v_2, 0)$.

Ejemplo 1.21. Un ejemplo más sofisticado es la esfera unidad, parametrizada por

$$\psi(u, v) = (\sin u \cos v, \sin u \sin v, \cos u), \qquad u \in (-\pi/2, \pi/2),\ v \in (0, 2\pi).$$

Se llama polo norte al punto $N = (0, 0, 1)$ y polo sur al punto $S = (0, 0, -1)$. Para esta parametrización, los símbolos de Christoffel se anulan todos salvo $\Gamma^1_{22} = -\sin u \cos u,\ \Gamma^2_{12} = \cot u$. En consecuencia, el sistema de ecuaciones de Euler-Lagrange que satisface cualquier geodésica de la esfera es

$$\begin{cases} u'' - v'^2 \sin u \cos u = 0, \\ v'' + 2u'v' \cot u = 0. \end{cases}$$

Una primera consecuencia inmediata es que las curvas con $v(t) = v_0$ constante resuelven la segunda ecuación y *casi* la primera ecuación. Si $u(t) = at + b$ con $a, b \in \mathbb{R}$, se tiene que las curvas $t \mapsto \psi(at+b, v_0)$ son geodésicas. Estas curvas parametrizan meridianos que van del polo norte al polo sur. Si $u(t) = u_0$ es constante entonces $v(t) = at + b$ resuelve la segunda ecuación. Si tomamos además $u_0 = 0$, entonces la curva $\gamma(t) = \psi(0, at + b)$ es también geodésica. Aquí estamos parametrizando el ecuador de la esfera.

Podemos decir un poco más. De la segunda ecuación se tiene

$$\frac{v''}{v'} = -2 \cot u\, u',$$

e integrando llegamos a la conocida como *relación de Clairaut*

$$v' = \frac{c}{\sin^2 u}, \qquad c \in \mathbb{R}.$$

Si u es constante ya sabemos que es necesariamente el ecuador, así que suponemos que u no es constante en un intervalo y trabajamos ahí. Multiplicando la primera ecuación por u', sustituyendo v' por $c/\sin^2 u$ e integrando explícitamente se llega a

$$u'^2 = -\frac{c^2}{\sin^2 u} + d, \qquad c, d \in \mathbb{R}.$$

De aquí se observa que $d > 0$ necesariamente. Despejando concluimos

$$u' = \pm\sqrt{d - \frac{c^2}{\sin^2 u}}.$$

Aunque esta EDO se puede integrar explícitamente, no podemos resolver $u(t)$ como una función de t. Incluso aunque se pudiera, luego tendríamos que integrar $v' = c/\sin^2 u$.

Ejercicios tema 1

1. Considera la EDO
$$x'(t) = x(t), \qquad t \in I \subset \mathbb{R}.$$

   - Supongamos que $x(t_0) = 0$ para algún $t_0 \in \mathbb{R}$. Usando el Teorema del Valor Medio, probar que $x(t) = 0$ para todo $t \leq t_0$.
   - Probar que $x(t) = 0$ para todo $t \geq t_0$. *Indicación: usar la función $f(t) = x(t) - \lambda e^t$ para λ adecuado.*
   - Deducir que si $x(t)$ es solución de la EDO, o bien $x(t) \neq 0$ para todo t, o bien $x(t) = 0$ para todo t.
   - Sean $x(t), y(t)$ soluciones de la EDO con $y(t) \neq 0$. Probar que existe $\lambda \in \mathbb{R}$ tal que $x(t) = \lambda y(t)$.
   - Concluir que las únicas soluciones de esta EDO son de la forma $x(t) = \lambda e^t,\ \lambda \in \mathbb{R}$. En particular, $x(t) = x_0 e^t$ es la única solución del problema de Cauchy
$$\begin{cases} x'(t) = x(t), \\ x(t_0) = x_0. \end{cases}$$

2. Probar que si $x(t)$ es una solución de la EDO
$$x''(t) + x(t) = 0,$$
entonces es también solución de
$$x'(t)^2 + x(t)^2 = k,$$
para algún $k \in \mathbb{R}$. Encontrar una solución de $x'(t)^2 + x(t)^2 = k$ que no sea solución de $x''(t) + x(t) = 0$. Qué tiene que cumplir una solución de $x'(t)^2 + x(t)^2 = k$ para ser solución de $x''(t) + x(t) = 0$?

3. Se considera el problema de Cauchy
$$\begin{cases} x'(t)^2 + x(t)^2 = a^2, \\ x(0) = b. \end{cases} \qquad a \geq 0,\ b \in \mathbb{R}.$$

   - Probar que $x(t) = a \sin t$ es solución de la EDO.
   - Probar que si $a < |b|$ entonces el problema anterior no tiene solución.
   - Usando que $x(t) = a \sin t$ es solución de la EDO, encontrar una solución del problema de Cauchy cuando $a \geq |b|$.

4. Reducir las siguientes EDOs a un sistema equivalente de primer orden
$$a) \quad x''(t) = \frac{1 - x'(t)^2}{x(t)} - 2H\sqrt{1 - x'(t)^2},\ H \in \mathbb{R}, \qquad b) \quad x''(t) = f(x'(t))x(t),$$
$$c) \quad \sin x'''(t) + \cos x''(t) + \tan x'(t) = x^2(t) - 1.$$

5. Se define
$$x(t) = x(t_0) + \int_{t_0}^{t} tf(s, x(s))\, ds,$$
siendo f una función continua. ¿Qué EDO cumple $x(t)$?

6. Encontrar una EDO que admita como soluciones a las siguientes familias de funciones.
$$a)\quad x(t) = Ae^t + Be^{-t}, \qquad b)\quad x(t) = Ae^t + Bte^t, \qquad c)\quad Ae^t \cos t + Be^t \sin t.$$

7. Se considera la EDO
$$x'(t) = t + x^3.$$
Se pide:
   - Representar de forma aproximada las isoclinas.
   - Representar la curva en $\mathbb{R}^2$ donde las soluciones de la EDO alcanzan un punto crítico.
   - Estudiar el número de ceros de una solución $x(t)$ en función de la condición inicial $x(0) = x_0$.

8. Dada la curva
$$C = \{(t, x) \in \mathbb{R}^2 \colon t^2 + 2x^2 + 2t + 2x = 1\},$$
se pide:
   - Estudiar cuántas componentes de C se pueden extraer que se puedan expresar como un grafo $(t, x(t))$.
   - Encontrar una EDO de la forma $x'(t) = f(t, x(t))$ que admita como soluciones las funciones del apartado anterior.
   - Usando la expresión de C, reducir la EDO anterior a una autónoma.

9. Sea g la constante gravitacional y $L > 0$. Consideremos la EDO de segundo orden
$$x''(t) + \frac{g}{L} \sin x(t) = 0,$$
donde $x(t)$ describe la amplitud de un péndulo de longitud L que oscila al actuar sobre él únicamente la fuerza de la gravedad.
   - Escribir esta EDO como un sistema de primer orden equivalente. Calcular el campo de direcciones de este sistema y deducir el comportamiento de cualquier solución del mismo.
   - Probar que si $(x(t), y(t))$ es solución del sistema, entonces $(x(-t), -y(-t))$ es igualmente solución.
   - Probar que si $(x(t), y(t))$ es solución del sistema, entonces $(-x(-t), y(-t))$ es igualmente solución.
   - Asumiendo que por cada punto del plano pasa una única solución y que éstas están definidas para todo valor del parámetro, deducir que las soluciones del sistema son curvas periódicas encerrando al origen en el interior del dominio que acotan.

10. Encontrar la familia de trayectorias ortogonales a las siguientes familias de curvas

$$a) \quad x = e^{\lambda t}, \qquad b) \quad tx = \lambda, \qquad c) \quad x = \lambda t^4.$$

11. Se considera la EDO de segundo orden

$$x''(t) = \frac{1 - x'(t)^2}{x(t)} - 2H\sqrt{1 - x'(t)^2}, \qquad x > 0,\ x'(t) \in [-1, 1],\ H > 0.$$

- Probar que el punto $(1/(2H), 0)$ es solución constante del sistema.
- Estudiar la curva tal que el campo de direcciones es horizontal (tiene segunda componente nula).
- Representar de forma aproximada el campo de direcciones del sistema de primer orden equivalente.

12. Expresar cada enunciado como una ecuación diferencial.

- El grafo de $x(t)$ verifica que la pendiente de la recta tangente en un punto es el cuadrado de la distancia del punto al origen.
- El grafo de $x(t)$ verifica en cada punto que la distancia del origen al punto de corte de la recta tangente con el eje x coincide con la distancia del origen al punto de corte de la recta normal con el eje t.
- El grafo de $x(t)$ verifica que la recta normal en cada punto pasa por el origen de coordenadas.

13. Supongamos que está nevando con intensidad constante. Una máquina quitanieves sale a las 12 del medio día y recorre 2km en la primera hora, y posteriormente 1km en la segunda hora. ¿A qué hora empezó a nevar? *Indicaciones: 1) la altura de la nieve es directamente proporcional al tiempo transcurrido; 2) la velocidad de la máquina es inversamente proporcional a la altura de la nieve*.

14. Una persona, partiendo del origen, se mueve en la dirección del eje x positivo tirando de una cuerda de longitud L atada a una piedra. Supongamos que la cuerda se mantiene tensa en todo momento y que inicialmente la piedra se encuentra en la posición $(0, L)$. La trayectoria que describe la piedra cuando la persona anda se llama *tractriz* y es una curva clásica. Encontrar la ecuación diferencial satisfecha por la piedra.

- Escribir la trayectoria como un grafo $x(t)$, suponiendo que la cuerda es tangente a la trayectoria de la piedra en todo momento.
- Resolver la EDO obtenida.

Capítulo 2

Algunas EDOs integrables. Aplicaciones a distintos ámbitos

A lo largo de este tema y en lo que queda de curso, omitiremos la dependencia de la variable t de la función incógnita $x(t)$ de nuestras EDOs. Esta misma notación se seguirá para las derivadas sucesivas de x, entendiéndose $x' = x'(t),\ x'' = x''(t),\ \cdots,\ x^{n)} = x^{n)}(t)$. No obstante, mantendremos la dependencia de t de las funciones dadas en la definición de las EDOs, así como en la función $x(t)$ en algunos casos concretos donde pueda llevar a confusión su omisión.

El principal objetivo de este tema es dar la solución explícita de ciertas EDOs de primer orden y dar algunas aplicaciones en diferentes ámbitos en los que estas EDOs aparecen. La mayoría de EDOs que vamos a integrar y los métodos empleados para ello, eran ya bien conocidos hacia 1740, si bien su origen se remonta a la propia fundación del Cálculo Diferencial. El interés por estas EDOs surgió de la necesidad de encontrar nuevos métodos para abordar problemas geométricos que involucraban el estudio de curvas planas con ciertas propiedades notables.

La forma de entender este tema ha de ser doble. Por una parte, se exponen métodos de resolución con rigor matemático, el cual a veces nos hace perder en intuición y capacidad de cálculo Por otra parte, en la práctica suele ser recomendable aplicar los métodos básicos de integración sin preocuparse del rigor matemático en todos y cada uno de los pasos (aunque siempre hay que tener cuidado!) y posteriormente se comprueban que las *soluciones* obtenidas cumplen con la EDO que queremos resolver. Esta pérdida de rigor matemático que nos permite obtener soluciones explícitas de nuestras EDOs, tiene como contraparte la duda de si hemos encontrado todas las soluciones de nuestro problema o si hemos dejado alguna por el camino. No será hasta el tema 4 cuando podamos dedicarnos a fondo en este problema.

2.1. Ecuaciones lineales

Sean $a_0, a_1, b : I \subset \mathbb{R} \to \mathbb{R}$, con $a_0(t) \neq 0$ para todo $t \in I$, funciones continuas. Se define la *ecuación diferencial lineal* como la EDO de primer orden

$$a_0(t)x' = a_1(t)x + b(t).$$

Dividiendo por $a_0(t)$ y renombrando las funciones podemos suponer que toda ecuación lineal se expresa como

$$x' = a(t)x + b(t). \tag{2.1.1}$$

Si $b(t) = 0$ la ecuación se dice *homogénea* y en caso contrario *completa*. Antes de resolver la ecuación (2.1.1) en toda su generalidad, vamos a pararnos un poco a estudiar la ecuación homogénea.

2.2. La ecuación homogénea

Consideremos la ecuación lineal homogénea,

$$x' = a(t)x,$$

con $a : I \subset \mathbb{R} \to \mathbb{R}$ continua. El caso más sencillo no trivial es cuando $a(t) = 1$ para todo t. Si nos restringimos a este caso y consideramos

$$x' = x,$$

sabemos del tema pasado que las funciones $x(t) = \lambda e^t$ son soluciones de esta ecuación. En concreto, para $\lambda = 0$ recuperamos la función constantemente cero. El siguiente resultado es una generalización de este hecho para cuando la función $a(t)$ es arbitraria.

Teorema 2.2.1. *Dada $a : I \subset \mathbb{R} \to \mathbb{R}$ una función continua, consideremos la EDO lineal homogénea*

$$x' = a(t)x.$$

1. *Si $x(t_0) = 0$, entonces $x(t) = 0$ para todo $t \in I$.*
2. *Las soluciones forman un espacio vectorial real de dimensión 1.*
3. *Las soluciones de la EDO son de la forma $x(t) = \lambda e^{\int a(t)\,dt},\ \lambda \in \mathbb{R}$.*
4. *Dados $t_0 \in I,\ x_0 \in \mathbb{R}$, el problema de Cauchy*

$$\begin{cases} x' = a(t)x, \\ x(t_0) = x_0, \end{cases}$$

tiene a $x(t) = x_0 e^{\int_{t_0}^t a(s)\,ds}$ como única solución.

Demostración. Vamos a probar todos los puntos de este teorema.

1. Si $x(t_0) = 0$, supongamos que existe $t_1 > t_0$ tal que $x(t_1) \neq 0$; podemos suponer sin perder generalidad $x(t_1) > 0$. Sea
$$t_* = \sup\{t < t_1 \colon x(t) = 0\}.$$
Este t_* es de hecho un máximo por continuidad de x y se tiene $x(t) > 0$ para todo $t \in (t_*, t_1]$. En consecuencia existe φ tal que
$$x(t) = e^{\varphi(t)}$$
para todo $t \in (t_*, t_1]$. Por definición, φ es de clase C^1 al serlo también x. Más aún, como x es solución de la EDO lineal homogénea, φ satisface la EDO $\varphi'(t) = a(t)$, de donde
$$\varphi(t) = \int a(t)\,dt + \lambda.$$
Por tanto, $x(t) = e^{\int a(t)\,dt + \lambda}$, pero $x(t_*) = 0$ lo cual es imposible. En consecuencia, $x(t) = 0$ para todo $t \in I$. Aquí es fundamental resaltar que $t_* \in I$ y $a(t)$ es continua en I, por lo que

la integral $\int a(t)$ es finita en un entorno de t_*. Dicho de otra forma, si $A(t)$ es una primitiva de $a(t)$ en I, entonces no puede ocurrir que $\lim_{t\to t_*} A(t) = -\infty$, haciendo que $x(t_*) = 0$. En efecto, si t_* es un punto interior entonces existe $\epsilon > 0$ tal que $t_* - \epsilon \in I$ y podemos hacer este argumento en el compacto $[t_* - \epsilon, t_1]$, donde $a(t)$ es continua y alcanza su máximo y mínimo. Si t_* no es interior entonces es frontera y como $t_* \in I$ necesariamente $I = [t_*, t_\infty\rangle$, con $t_\infty \leq \infty$ y denotamos por $\rangle$ al hecho de que el intervalo puede ser abierto o cerrado en t_∞ (si $t_\infty = \infty$ entonces es abierto necesariamente). De cualquier forma, $[t_*, t_1] \subset I$ y $a(t)$ vuelve a alcanzar máximo y mínimo en tal compacto.

2. Sean x, y soluciones de la EDO homogénea. Entonces es trivial comprobar que $x + y$ y λx con $\lambda \in \mathbb{R}$ son también soluciones y en particular la función nula es solución. Para ver que la dimensión es 1, sean x, y soluciones con $y(t) \neq 0$ para todo $t \in I$. Entonces,

$$\left(\frac{x}{y}\right)' = \frac{x'y - xy'}{y^2} = \frac{a(t)xy - a(t)xy}{y^2} = 0,$$

y por tanto existe $\lambda \in \mathbb{R}$ tal que

$$x(t) = \lambda y(t), \qquad \forall t \in I.$$

Esto nos dice que dos soluciones diferentes difieren de una constante multiplicativa y por tanto el espacio vectorial se define mediante una solución no nula, probando que la dimensión es 1.

3. Como es inmediato comprobar que la función $e^{\int a(t)\,dt}$ es solución, para cualquier otra solución $x(t)$ la estructura de espacio vectorial asegura la existencia de $\lambda \in \mathbb{R}$ tal que

$$x(t) = \lambda e^{\int a(t)\,dt}.$$

4. La función $x(t) = \lambda e^{\int a(t)\,dt}$ es solución del problema de Cauchy si y sólo si $x(t_0) = x_0$. En la definición de $x(t)$ hemos tomado integrales indefinidas, pero nada nos impide tomar la integral definida

$$x(t) = \lambda e^{\int_{t_0}^t a(s)\,ds}.$$

Para esta elección, si $x(t_0) = x_0$ necesariamente $\lambda = x_0$, lo que concluye el resultado.

□

Observación 2.1. En el resultado anterior, la hipótesis de continuidad de $a(t)$ es esencial. En efecto, si definimos

$$a(t) = \begin{cases} 2 & \text{si} \quad t \leq 1 \\ 1 & \text{si} \quad t > 1, \end{cases}$$

el problema de Cauchy

$$\begin{cases} x' = a(t)x, \\ x(1) = e^2 \end{cases}$$

no tiene ninguna solución definida en $\mathbb{R}$. En efecto, si $x : \mathbb{R} \to \mathbb{R}$ es solución, en particular lo será en el intervalo $(-\infty, 1]$. Como en este intervalo la ecuación es $x' = 2x$, la solución del problema de

Cauchy es $x(t) = e^{2t}$. Si ahora $t > 1$ la ecuación es $x' = x$ y la solución del problema de Cauchy es $x(t) = e^{t+1}$. No obstante, $x(t)$ definida por

$$x(t) = \begin{cases} e^{2t} & \text{si} \quad t \leq 1, \\ e^{t+1} & \text{si} \quad t > 1, \end{cases}$$

no es derivable en $t = 1$ ya que sus derivadas laterales no coinciden. De hecho, si

$$a(t) = \begin{cases} 2 & \text{si} \quad t = 1 \\ 1 & \text{si} \quad t \neq 1, \end{cases}$$

el problema de Cauchy anterior no tiene *ninguna* solución. En efecto, si x es solución entonces para todo $t \neq 1$ se tendría $x(t) = e^t$ y por continuidad $x(t) = e^t$ para todo $t \in \mathbb{R}$. Pero la EDO $x'(t) = a(t)x(t)$ para $t = 1$ cumple $x'(1) = a(1)x(1) = 2x(1)$, lo cual es falso para la función $x(t)$.

2.3. La ecuación completa

Una vez estudiadas en detalle las primeras propiedades de la ecuación lineal homogénea, así como sus soluciones, nos disponemos a estudiar de forma similar la ecuación completa

$$x' = a(t)x + b(t).$$

Para tal fin, nos ayudaremos de la teoría de espacios vectoriales y afines; a continuación recordamos la definición de los últimos.

Definición 2.3.1. Sea V un espacio vectorial. Un conjunto no vacío $\mathbb{A}$ es un *espacio vectorial afín sobre* V si existe una aplicación

$$\varphi : \mathbb{A} \times \mathbb{A} \to V,$$

satisfaciendo:

1. Para cada $a \in \mathbb{A}$, la aplicación $\varphi_a : \mathbb{A} \to V,\ \varphi_a(b) = \varphi(a, b)$ es biyectiva.
2. Se cumple la relación $\varphi(a, c) = \varphi(a, b) + \varphi(b, c)$

A los elementos de $\mathbb{A}$ se les llama *puntos*, y la aplicación φ asocia a dos puntos un vector de V. Es posible que dados puntos $a_k, b_k \in \mathbb{A},\ k = 1, 2$, se cumpla $\varphi(a_1, b_1) = \varphi(a_2, b_2)$. No obstante, la biyectividad de φ_a nos permite tomar un punto de $\mathbb{A}$ como *origen*. En efecto, fijado $\mathbf{o} \in \mathbb{A}$ se dice que el vector $\varphi_{\mathbf{o}}(a)$ tiene origen $\mathbf{o}$ y extremo a. Por ejemplo, $\mathbb{R}^2$ como conjunto de puntos tiene estructura de espacio afín sobre el espacio vectorial $\mathbb{R}^2$ sin más que definir la aplicación

$$\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, \qquad \varphi(a, b) = b - a.$$

Es inmediato comprobar que, así definida, φ cumple las dos propiedades necesarias que dotan a $\mathbb{R}^2$ de estructura de espacio afín. Si fijamos $\mathbf{o} = (x_0, y_0)$, entonces la aplicación $\varphi_{\mathbf{o}}(x, y) = (x, y) - (x_0, y_0)$ define el vector con extremo (x, y) y origen $\mathbf{o}$.

Teorema 2.3.2. *El espacio de soluciones de la ecuación* (2.1.1) *es un espacio vectorial afín de dimensión 1, cuyo espacio vectorial base es el espacio de soluciones de la ecuación homogénea* $x' = a(t)x$. *En particular, las soluciones de* (2.1.1) *vienen dadas por*

$$x(t) = \lambda e^{\int a(t)\,dt} + x_p(t), \qquad \lambda \in \mathbb{R},$$

donde $x_p(t)$ es una solución particular de (2.1.1). Finalmente, el problema de Cauchy

$$\begin{cases} x' = a(t)x + b(t), \\ x(t_0) = x_0, \end{cases}$$

tiene como única solución $x(t) = x_p(t) + (x_0 - x_p(t_0))e^{\int_{t_0}^{t} a(s)\,ds}$.

Demostración. Sean $\mathcal{S}_h, \mathcal{S}_c$ el espacio de las soluciones de la ecuación homogénea y completa, respectivamente. Entonces,

$$\varphi : \mathcal{S}_c \times \mathcal{S}_c \to \mathcal{S}_h, \qquad \varphi(x,y) = y - x,$$

está bien definida y dota a $\mathcal{S}_c$ de estructura de espacio afín sobre $\mathcal{S}_h$. En efecto, que φ está bien definida es trivial puesto que dadas $x, y \in \mathcal{S}_c$ se tiene

$$(x-y)' = x' - y' = a(t)x + b(t) - a(t)y - b(t) = a(t)(x-y),$$

y por tanto $x - y \in \mathcal{S}_h$. Por otra parte, probemos que dada $x \in \mathcal{S}_c$ se tiene que la aplicación $\varphi_x(y) = y - x$ es biyectiva. En efecto, si $\varphi_x(y) = \varphi_x(z)$ entonces $y = z$ de la propia definición de φ. Por otra parte, si $z \in \mathcal{S}_h$ entonces $x + z \in \mathcal{S}_c$ y por tanto $\varphi_x(x+z) = z$, probando la sobreyectividad y finalmente la biyectividad. La segunda propiedad se demuestra de igual forma.

Esta estructura de espacio afín nos permite concluir lo siguiente: sea $x_p \in \mathcal{S}_c$ *cualquier* solución particular. Entonces, $\varphi_{x_p} : \mathcal{S}_c \to \mathcal{S}_h$ es biyectiva y en conclusión para cualquier otra $x \in \mathcal{S}_c$ existe una única $x_h \in \mathcal{S}_h$ tal que $x = x_h + x_p$. En consecuencia, fijada $x_p \in \mathcal{S}_c$, se tiene

$$\mathcal{S}_c = \{x_h + x_p \colon x_h \in \mathcal{S}_h\}.$$

Como conocemos las soluciones de la ecuación homogénea, a saber $x_h(t) = \lambda e^{\int a(t)\,dt}$ con $\lambda \in \mathbb{R}$, las soluciones de (2.1.1) vendrán dadas por

$$x(t) = \lambda e^{\int a(t)\,dt} + x_p(t), \qquad \lambda \in \mathbb{R}.$$

Referente al problema de Cauchy, sustituyendo $x(t_0)$ se tiene que $\lambda = x_0 - x_p(t_0)$, lo que finaliza el resultado. $\square$

Como consecuencia de este resultado, para resolver la ecuación (2.1.1) basta conocer una solución particular suya, $x_p(t)$. Una forma de buscar una solución particular $x_p(t)$ es imponer que sea de la forma

$$x_p(t) = \lambda(t)e^{\int a(t)\,dt}.$$

Es decir, estamos considerando una solución de la ecuación homogénea, pero permitiendo que la constante λ que la acompañaba sea ahora una función arbitraria. Este método es conocido como *método de variación de las constantes* por motivos obvios y tendrá un análisis teórico en mayor profundidad en futuros temas. Si imponemos que así definida $x_p(t)$ sea solución de (2.1.1), obtenemos

$$x_p'(t) = \lambda'(t)e^{\int a(t)\,dt} + \lambda(t)a(t)e^{\int a(t)\,dt} = a(t)x_p(t) + b(t) = a(t)\lambda(t)e^{\int a(t)\,dt} + b(t),$$

lo cual nos lleva a que

$$\lambda'(t) = b(t)e^{-\int a(t)\,dt},$$

e integrando concluimos

$$\lambda(t) = \int b(t)e^{-\int a(t)\,dt}\,dt.$$

Aquí no hace falta añadir la constante de integración, ya que nos basta obtener *una* solución particular. En consecuencia,

$$x_p(t) = e^{\int a(t)\,dt}\int b(t)e^{-\int a(t)\,dt}\,dt,$$

y por tanto la solución de la ecuación lineal (2.1.1) viene dada por

$$x(t) = e^{\int a(t)\,dt}\left(\int e^{-\int a(t)\,dt}b(t)\,dt + \lambda\right), \qquad \lambda \in \mathbb{R}. \tag{2.3.1}$$

Con todo este desarrollo, tenemos el resultado de existencia y unicidad del problema de Cauchy para la EDO completa.

Teorema 2.3.3. *La ecuación lineal completa* (2.1.1) *tiene como solución la familia 1-paramétrica dada por* (2.3.1). *Más aún, el problema de Cauchy*

$$\begin{cases} x' = a(t)x + b(t), \\ x(t_0) = x_0 \end{cases}$$

tiene como única solución

$$x(t) = e^{\int_{t_0}^t a(s)\,ds}\left(\int_{t_0}^t b(s)e^{-\int_{t_0}^s a(u)\,du}\,ds + x_0\right).$$

Demostración. Que la ecuación lineal (2.1.1) tiene como solución (2.3.1) se deduce del Teorema 2.3.2, que asegura que el espacio de soluciones de la ecuación completa tiene estructura de espacio afín sobre el espacio vectorial del conjunto de soluciones de la ecuación homogénea, y por tanto toda solución $x(t)$ de (2.1.1) es de la forma

$$x(t) = \lambda e^{\int a(t)\,dt} + x_p(t), \qquad \lambda \in \mathbb{R},$$

con $x_p(t)$ una solución particular. Tomando la calculada anteriormente,

$$x_p(t) = e^{\int a(t)\,dt}\int b(t)e^{-\int a(t)\,dt}\,dt,$$

concluimos que cualquier solución de (2.1.1) es de la forma (2.3.1). Por otra parte, las integrales involucradas en (2.3.1) son todas indefinidas, pero nada nos impide elegir las primitivas concretas

$$x(t) = e^{\int_{t_0}^t a(s)\,ds}\left(\int_{t_0}^t e^{-\int_{t_0}^s a(u)\,du}b(s)\,ds + \lambda\right).$$

Con estas elecciones, para que se dé $x(t_0) = x_0$ necesariamente $\lambda = x_0$. Esto concluye la prueba. □

Observación 2.2. Un par de comentarios acerca de este resultado.

1. Las soluciones obtenidas están definidas en el mismo intervalo I donde están definidas las funciones $a(t), b(t)$. Este resultado de existencia de soluciones es *global*, lo cual es algo que está lejos de ser trivial. Por ejemplo, la EDO $x' = -x^2$ tiene como soluciones las funciones $x(t) = 1/(t + c)$, que no están definidas para $t = -c$.

2. Que se tenga una expresión explícita de las soluciones no significa que éstas puedan expresarse en términos de funciones elementales. Por ejemplo, la ecuación lineal

$$x' = 2tx + 1$$

tiene como solución

$$x(t) = e^{t^2} \int e^{-t^2}\, dt,$$

y es bien sabido que la primitiva de e^{-t^2} no se puede expresar como una función elemental.

A continuación, presentamos dos métodos alternativos para calcular soluciones particulares de (2.1.1) y en última instancia abordar la resolución explícita de esta EDO. Estos métodos tendrán una generalización y estudio en profundidad tanto en este tema como en otros más adelante, pero creemos conveniente introducirlos ya para que el lector se vaya haciendo a ellos desde un planteamiento más sencillo.

El primer método es el conocido como *método de los coeficientes indeterminados* y se puede aplicar en los casos en que $a(t) = a \in \mathbb{R}$ y $b(t)$ es de la forma

$$b(t) = e^{\alpha t}\left(p_n(t)\cos(\beta t) + q_m(t)\sin(\beta t)\right),$$

donde $p_n(t), q_m(t)$ son polinomios de grado n, m, respectivamente. El método, cuyo fundamento teórico veremos en temas posteriores, nos plantea buscar una solución particular de la forma

$$x_p(t) = t^{\delta_{a\alpha}} e^{\alpha t}\left(r_k(t)\cos(\beta t) + s_k(t)\sin(\beta t)\right),$$

donde $\delta_{a\alpha} = 1$ si $a = \alpha$ y 0 en caso contrario, y $r(t), s(t)$ son polinomios de grado $k = \max\{n, m\}$ cuyos *coeficientes indeterminados* debemos determinar imponiendo que $x_p(t)$ sea solución particular de la EDO. Este método se suele usar junto con el conocido como *principio de superposición de soluciones*.

Proposición 2.3.4 (Método de superposición)**.** *Sean $a, b_1, ..., b_n : I \subset \mathbb{R} \to \mathbb{R}$ funciones continuas. Para cada $k = 1, .., n$, sea x_{p_k} una solución particular de la EDO lineal completa $x' = a(t)x + b_k(t)$. Entonces, $x_p = x_{p_1} + \cdots + x_{p_n}$ es solución particular de la EDO $x' = a(t)x + b_1(t) + \cdots + b_n(t)$.*

El segundo método consiste en multiplicar (2.1.1) por una función $\mu(t)$ no nula de forma que el miembro de la izquierda se escriba como la derivada de $\mu(t)x$ y poder integrar para eventualmente obtener la solución x. Esto es,

$$\mu(t)(x' - a(t)x) = (\mu(t)x)' = \mu'(t)x + \mu(t)x'.$$

En consecuencia,

$$-\mu(t)a(t)x = \mu'(t)x,$$

y despejando se llega a $\mu'(t) = -a(t)\mu(t)$. Como suponemos $\mu(t) \neq 0$, podemos suponer sin perder generalidad $\mu(t) > 0$, lo cual nos permite dividir por $\mu(t)$ y llegamos a

$$\frac{\mu'(t)}{\mu(t)} = -a(t) \Longrightarrow \log \mu(t) = -\int a(t)\, dt,$$

o de forma equivalente

$$\mu(t) = e^{-\int a(t)\, dt}.$$

Por tanto, (2.1.1) se expresa como

$$(e^{-\int a(t)\,dt}x)' = \mu(t)b(t),$$

e integrando y despejando se llega a

$$x(t) = e^{\int a(t)\,dt}\left(\int b(t)e^{-\int a(t)\,dt}\,dt + \lambda\right), \qquad \lambda \in \mathbb{R},$$

que coincide con (2.3.1).

Observación 2.3. Realmente, podemos probar un poco más. Si $z(t)$ es solución de (2.1.1) en un cierto $J \subset I$, entonces necesariamente z es la restricción a J de alguna de las soluciones dadas en (2.3.1). La prueba de este hecho tan intuitivo presenta más problemas formales que conceptuales, y para tal fin debemos introducir alguna notación.

Denotemos por $\tilde{a}(t), \tilde{b}(t)$ a las restricciones de $a(t), b(t)$ a J, respectivamente. Si $A(t) = \int a(t)\,dt$ es una primitiva de $a(t)$ y consideramos la restricción $\tilde{A}(t)$ de $A(t)$ a J, entonces $\tilde{A}(t)$ es una primitiva de $\tilde{a}(t)$. Evidentemente, si $\mu(t) = e^{A(t)}$ y $\tilde{\mu}(t) = e^{\tilde{A}(t)}$, entonces $\tilde{\mu}(t)$ es la restricción de $\mu(t)$ a J. Por último, si $B(t) = \int \mu(t)b(t)\,dt$ es una primitiva de $\mu(t)b(t)$ y consideramos su restricción $\tilde{B}(t)$ a J, entonces $\tilde{B}(t)$ es una primitiva de $\tilde{\mu}(t)\tilde{b}(t)$.

Sea pues $z(t)$ una solución de (2.1.1) en $J \subset I$. Entonces, $z(t)$ es de la forma

$$z(t) = \frac{1}{\tilde{\mu}(t)}(\tilde{B}(t) + \lambda_0),$$

para algún $\lambda_0 \in \mathbb{R}$. Si ahora definimos

$$x(t) = \frac{1}{\mu(t)}(B(t) + \lambda_0),$$

entonces obviamente x es solución de (2.1.1) y además $z(t)$ es la restricción de $x(t)$ a J.

Ejemplo 2.1. Queremos resolver la EDO

$$x' = 5x + t^3e^t.$$

Vamos a hacerlo por los dos métodos anteriormente descritos.

- Primero, busquemos $\mu(t)$ adecuado para que $\mu(t)(x' - 5x) = (\mu(t)x)'$, lo cual nos lleva a que

  $$\mu(t) = e^{-5t}.$$

  En consecuencia, la solución es

  $$x(t) = e^{5t}\left(\int t^3e^{-4t}\,dt + \lambda\right), \qquad \lambda \in \mathbb{R}.$$

  Esta solución no es difícil de calcular explícitamente, pero debemos efectuar la integración de

  $$\int t^3e^{-4t}\,dt,$$

que sin ser la integral más complicada a la que nos vayamos a enfrentar no deja de ser algo tediosa de calcular. De hecho, su solución es

$$\int t^3 e^{-4t}\, dt = -\frac{e^{-4t}}{128}\left(3 + 12t + 24t^2 + 32t^3\right).$$

Esta integral es la misma que obtenemos si intentamos el método de variación de las constantes, suponiendo la solución particular de la forma $x_p(t) = \lambda(t)e^{5t}$, e imponiendo que sea solución de $x' = 5x + t^3e^t$, lo que implica

$$\lambda'(t) = t^3 e^{-4t}.$$

- Ahora aplicamos el método de los coeficientes indeterminados. Por una parte, $a(t) = a = 5$ y por otra parte de $b(t) = t^3e^t$ deducimos $\alpha = 1, \beta = 0$ y $n = 3$ y $m = 0$; en consecuencia $k = 3$. Como $\alpha = 1$ y $a = 5$, buscamos la solución particular de la forma

$$x_p(t) = e^t\left(c_1 + c_2t + c_3t^2 + c_4t^3\right).$$

Sustituyendo en la EDO $x_p' = 5x_p + t^3e^t$ tenemos

$$e^t\left(c_1 + c_2t + c_3t^2 + c_4t^3\right) + e^t\left(c_2 + 2c_3t + 3c_4t^2\right) = 5e^t\left(c_1 + c_2t + c_3t^2 + c_4t^3\right) + t^3e^t.$$

De esta igualdad polinómica concluimos que los coeficientes son solución del sistema

$$\begin{cases} 4c_1 = c_2, \\ 4c_2 = 2c_3, \\ 4c_3 = 3c_4, \\ 4c_4 = -1. \end{cases}$$

De aquí obtenemos los mismos coeficientes que por el método anterior.

2.3.1. Aplicación a la biología. Datación por carbono-14

El método de datación de acontecimientos por el isótopo de carbono 14 es una técnica ideada en la década de 1940 por el químico estadounidense Willard Frank Libby y que le valió para ganar el Nobel de Química en 1960. Este método se sustenta en el fenómeno de la radiactividad. Existen ciertos tipos de átomos que tienen en su núcleo un número elevado de neutrones y esto los hace inestables. Los átomos se desintegran, expulsando neutrones sobrantes y convergiendo a una configuración más estable. Experimentalmente se ha observado que la velocidad de desintegración de una sustancia radiactiva es proporcional a la cantidad de sustancia existente. Esto es la conocida como *Ley de desintegración radiactiva*. De forma analítica, si $x(t)$ es la cantidad de sustancia en tiempo t, entonces

$$x' = -kx,$$

donde $k > 0$ es una constante que depende de la sustancia en sí. Por tanto, si x_0 es la cantidad de sustancia que se tiene al comienzo (con $t = 0$), se tiene

$$x(t) = x_0 e^{-kt}.$$

Un parámetro importante que aparece en el fenómeno de la desintegración radiactiva es la *vida media*, denotada por $t_{1/2}$, y que indica el tiempo que debe transcurrir hasta que la sustancia se desintegra hasta la mitad. En concreto, si se comienza con una cantidad x_0, la vida media vendrá dada por

$$\frac{x_0}{2} = x(t_{1/2}) = x_0 e^{-kt_{1/2}},$$

de donde concluimos

$$t_{1/2} = \frac{\log 2}{k}$$

y por tanto $t_{1/2}$ depende exclusivamente de la constante de desintegración k y no de la cantidad de materia existente. A su vez, la constante de desintegración puede calcularse sabiendo la cantidad de materia en dos instantes distintos. Si $x_1 = x(t_1)$ y $x_2 = x(t_2)$, entonces

$$x_1 = x_0 e^{-kt_1}, \qquad x_2 = x_0 e^{-kt_2},$$

y el valor de k lo calculamos como

$$k = \frac{\log x_1 - \log x_2}{t_2 - t_1}.$$

Generalmente es imposible saber la cantidad de materia existente en dos instantes distintos, así que otro método para estimar el tiempo que una sustancia lleva desintegrándose es hacer uso de la velocidad de desintegración, la cual se puede calcular de forma instantanea. Si denotamos por $r(t) = x'(t) = -kx_0 e^{-kt}$ a la velocidad de desintegración de la sustancia y conocemos las velocidades $r_1 = r(t_1)$ y $r_2 = r(t_2)$ en dos tiempos distintos, se tiene

$$k = \frac{\log r_1 - \log r_2}{t_2 - t_1}.$$

Supongamos que tenemos un trozo de madera cuya antigüedad queremos estimar. Sea $t = 0$ el instante en el que se cortó el trozo de madera y $t_1 > 0$ el instante actual en el que se estudia el trozo. Denotemos por r_0 la velocidad de desintegración en $t = 0$ (que cortando un trozo similar se puede calcular) y por r_1 a la velocidad en $t = t_1$. Entonces,

$$t_1 = \frac{1}{k} \log \frac{r_0}{r_1}.$$

Por otra parte, la vida media del carbono-14 es de 5568 años y por tanto $k \approx 1'245 \cdot 10^{-4}$, con lo que obtenemos el tiempo t_1 que ha pasado desde que se cortó la madera.

Por ejemplo, unos exploradores afirman haber encontrado el Santo Grial en una incursión en un antiguo templo que habitaron los Caballeros Templarios. Según las leyendas, el Grial está tallado en la misma madera que hoy día se encuentra en el Monte de los Olivos. Las mediciones mostraron que la madera de la copa encontrada se desintegraba a una velocidad de $5'6$ átomos por gramo y minuto, mientras que la velocidad de desintegración de un trozo de madera similar era de $6'68$ átomos por gramo y minuto. En consecuencia, $r_0 = 6'68$ y $r_1 = 5{,}6$ y como $k \approx 1'245 \cdot 10^{-4}$ basta calcular t_1 obteniendo

$$t_1 \approx 1416'48$$

Es decir, han pasado unos 1416 años y medio desde que se cortó la madera para construir la copa, haciendo imposible que sea el Santo Grial que afirmaban los Templarios.

2.3.2. Aplicación a medicina forense. Determinar hora de fallecimiento

La ley de enfriamiento de Newton establece que en situaciones en las que la temperatura del medio permanece aproximadamente constante, la temperatura superficial de un objeto varía proporcionalmente a la diferencia entre la temperatura del ambiente y la del objeto. En otras palabras, si $x(t)$ es la temperatura en el instante t y T_{amb} es la temperatura del medio (supuesta constante), entonces

$$x' = k(T_{amb} - x),$$

donde $k > 0$ es una cierta constante de proporcionalidad que sólo depende de la naturaleza del objeto. La EDO que describe la ley del enfriamiento de Newton es lineal y completa, con $a(t) = -k$ y $b(t) = kT_{amb}$, ambas constantes. En efecto, la solución de la EDO lineal completa que describe la ley de enfriamiento de Newton es

$$x(t) = T_{amb} + (T_0 - T_{amb})e^{-k(t-t_0)},$$

para una cierta condición inicial $x(t_0) = T_0$.

Supongamos el siguiente problema que surge en la medicina forense. Un cuerpo sin vida se encuentra en una habitación cerrada y se encuentra en tiempo $t = 0$. El objetivo es determinar el tiempo de fallecimiento, $t_m < 0$. Si hemos encontrado el tiempo a una temperatura $T_0 = x(0)$, la solución de la EDO vendrá dada por

$$x(t) = T_{amb} + (T_0 - T_{amb})e^{-kt}.$$

Lo primero que debemos hacer para conocer explícitamente esta función es determinar el valor del parámetro k. Para ello, medimos la temperatura en otro tiempo $t_1 > 0$, obteniendo $T_1 = x(t_1)$, y podemos despejar la constante k como

$$k = \frac{1}{t_1} \log \frac{T_0 - T_{amb}}{T_1 - T_{amb}}.$$

Si intuimos que conocemos la temperatura de su muerte T_m (por ejemplo, no es descabellado suponer que es aproximadamente 37ºC), entonces podemos estimar su hora de fallecimiento como

$$t_m = \frac{1}{k} \log \frac{T_0 - T_{amb}}{T_m - T_{amb}} = t_1 \frac{\log \dfrac{T_0 - T_{amb}}{T_m - T_{amb}}}{\log \dfrac{T_0 - T_{amb}}{T_1 - T_{amb}}}.$$

Por ejemplo, supongamos que el cadáver se descubre en $t = 0$ para una temperatura inicial de $T_0 = 29C$ y una hora y media después ($t_1 = 1'5$) tiene una temperatura de $T_1 = 25C$. Si la habitación se encuentra a $T_{amb} = 20$ y supuesta que esta temperatura de la habitación ha permanecido constante, estimamos

$$k = \frac{1}{1'5} \log \frac{29 - 20}{25 - 20} \approx 0'39, \qquad t_m = \frac{1}{k} \log \frac{29 - 20}{37 - 20} \approx -1'62h,$$

lo cual equivale a que la persona falleció 1h y 37’ aproximadamente antes de ser encontrada.

2.4. La Ecuación de Bernoulli

A continuación vamos a estudiar cómo ciertas ecuaciones no lineales de primer orden pueden reducirse a una ecuación lineal tras una serie de manipulaciones convenientes, las cuales suelen ser cambios de variable. La primera ecuación de este tipo que vamos a analizar es la *ecuación de Bernoulli*.

Definición 2.4.1. Una ecuación diferencial se dice de Bernoulli si es de la forma

$$x' = a(t)x + b(t)x^{\alpha}, \tag{2.4.1}$$

donde $a, b : I \subset \mathbb{R} \to \mathbb{R}$ son funciones continuas y $\alpha \in \mathbb{R}$.

Si $\alpha = 0$ tenemos una ecuación lineal completa, y si $\alpha = 1$ la ecuación lineal obtenida es homogénea. Supondremos pues $\alpha \neq 0, 1$ para evitar estos casos ya analizados.

El método usado para estudiar sistemáticamente las soluciones de esta ecuación es realizar un cambio de variable adecuado. En paralelo, realizaremos el estudio del problema de Cauchy asociado.

Teorema 2.4.2. *Sea $x : I \subset \mathbb{R} \to \mathbb{R}$ una solución de (2.4.1) y supongamos que existe $J \subset I$ tal que $x(t) > 0$ para todo $t \in J$. Entonces, $u(t) = x(t)^{1-\alpha}$ es solución de la EDO lineal*

$$u' = (1-\alpha)a(t)u + (1-\alpha)b(t). \tag{2.4.2}$$

Recíprocamente, si $u : J \to (0, \infty)$ es solución de (2.4.2), entonces $x(t) = u(t)^{\frac{1}{1-\alpha}}$ es solución de (2.4.1). En particular, dados $t_0 \in I$ y $x_0 > 0$, el problema de Cauchy

$$\begin{cases} x' = a(t)x + b(t)x^{\alpha}, \\ x(t_0) = x_0, \end{cases} \tag{2.4.3}$$

tiene existencia y unicidad.

Demostración. Sea $x : I \subset \mathbb{R} \to \mathbb{R}$ una solución de (2.4.1) y supongamos que existe $J \subset I$ tal que $x(t) > 0$ para todo $t \in J$. Entonces, la función $u : J \to \mathbb{R},\ u(t) = x(t)^{1-\alpha}$, está bien definida independientemente del valor de α por ser $x(t) > 0$ para todo $t \in J$; es más, es C^1 por serlo x. Si derivamos u, obtenemos

$$\begin{aligned} u' &= (1-\alpha)x^{-\alpha}x' \\ &= (1-\alpha)u^{\frac{-\alpha}{1-\alpha}}(a(t)x + b(t)x^{\alpha}) \\ &= (1-\alpha)u^{\frac{-\alpha}{1-\alpha}}(a(t)u^{\frac{1}{1-\alpha}} + b(t)u^{\frac{\alpha}{1-\alpha}}) \\ &= (1-\alpha)(a(t)u + b(t)). \end{aligned}$$

Es decir, u es solución de la EDO lineal (2.4.2). Recíprocamente, si $u : J \subset \mathbb{R} \to (0, \infty)$ es solución de (2.4.2), la función $x(t) = u(t)^{\frac{1}{1-\alpha}}$ está bien definida independientemente de α por ser $u(t) > 0$ para todo t y es solución de la EDO de Bernoulli (2.4.1). En efecto,

$$\begin{aligned} x' &= \frac{1}{1-\alpha}u^{\frac{\alpha}{1-\alpha}}u' \\ &= \frac{1}{1-\alpha}u^{\frac{\alpha}{1-\alpha}}(1-\alpha)(a(t)u + b(t)) \\ &= a(t)u^{\frac{1}{1-\alpha}} + b(t)u^{\frac{\alpha}{1-\alpha}} \\ &= a(t)x + b(t)x^{\alpha}, \end{aligned}$$

lo que concluye la prueba. En cuanto al problema de Cauchy (2.4.3), consideremos el problema de Cauchy

$$\begin{cases} u' = (1-\alpha)a(t)u + (1-\alpha)b(t) \\ u(t_0) = x_0^{1-\alpha}. \end{cases}$$

Aquí remarcamos que la condición inicial $x_0^{1-\alpha}$ está bien definida independientemente del valor de α por ser $x_0 > 0$. Sea $u : I \subset \mathbb{R} \to \mathbb{R}$ la solución única de este problema de Cauchy, donde I es el intervalo de definición de las funciones $a(t), b(t)$. Definimos

$$J = \{t \in I \colon u(t) \neq 0\},$$

el cual es una unión disjunta, posiblemente infinita, de intervalos abiertos. Llamemos $J_0 \subset J$ al intervalo que contiene a t_0 y restringimos $u : J_0 \to (0, \infty)$. Si definimos $x(t) = u(t)^{\frac{1}{1-\alpha}}$, $t \in J_0$, entonces x está bien definida, es de clase C^1 y es solución de la EDO de Bernoulli (2.4.1). Más aún, $x(t_0) = x_0$ por lo que resuelve el problema de Cauchy.

Falta probar la unicidad y primero debemos aclarar qué vamos a entender por *unicidad de solución*: si $y : \tilde{J} \to \mathbb{R}$ es otra solución del problema de Cauchy (2.4.3), entonces $x(t) = y(t)$ para todo $t \in J_0 \cap \tilde{J}$. Supongamos que existe $y : \tilde{J} \to \mathbb{R}$ solución del problema de Cauchy (2.4.3) distinta a x. Sea $J_* = J_0 \cap \tilde{J}$ y supongamos pues que existe $t_1 \in J_*$ tal que $x(t_1) \neq y(t_1)$; digamos $t_1 > t_0$. Sabemos que $x(t_1) > 0$ pero podría ser $y(t_1) \leq 0$. No obstante, si este fuese el caso podemos definir

$$t_* = \text{mín}\{t > t_0 \colon y(t) = 0\}.$$

En tal caso, existe $\epsilon > 0$ suficientemente pequeño tal que $0 < y(t_* - \epsilon) < x(t_0)$ y basta redefinir t_1 como este $t_* - \epsilon$. Haciendo pues $\tilde{J}$ más pequeño si es necesario, existe $t_1 \in J_0 \cap \tilde{J}$ tal que $x(t_1) \neq y(t_1)$ y además la restricción de y a $\tilde{J}$ es positiva. Tiene pues sentido definir la función $v(t) = y(t)^{1-\alpha}$, la cual está bien definida, es solución de la EDO (2.4.2) y cumple $v(t_0) = x_0^{1-\alpha}$. Por una parte, la unicidad del problema de Cauchy para la EDO lineal completa asegura que $u(t) = v(t)$ para todo $t \in J_0 \cap \tilde{J}$, pero por otra parte $u(t_1) \neq v(t_1)$, lo cual es contradictorio. Esta contradicción nos permite asegurar que $x(t) = y(t)$ para todo $t \in J_0 \cap J$. □

Observación 2.4. El cambio propuesto $u(t) = x(t)^{1-\alpha}$ puede llevar a problemas para determinados valores de α si la función $x(t)$ no es positiva. Esto ocurre, por ejemplo, si $\alpha = 1/2$. Es por eso por lo que inicialmente pedimos que $x(t)$ y $u(t)$ sean funciones positivas. No obstante, se puede ser menos restrictivos en algunos casos, como veremos a continuación. Por ejemplo, si α es tal que $1 - \alpha = \frac{1}{2n+1}$, con $n \in \mathbb{N}$, entonces $u = x^{1-\alpha}$ está bien definida incluso si x es negativa o cero. En tal caso tendríamos garantizada la existencia de soluciones, pero podríamos encontrarnos algún problema con la unicidad del problema de Cauchy para condiciones iniciales que sean nulas, tal y como veremos a continuación.

Ejemplo 2.2. Resolvamos el problema de Cauchy

$$\begin{cases} x' = x^2, \\ x(0) = 1. \end{cases}$$

Se trata de una ecuación de Bernoulli con $a(t) = 0$, $b(t) = 1$ y $\alpha = 2$. El cambio $u = 1/x$ nos lleva a considerar el problema de Cauchy equivalente

$$\begin{cases} u' = -1, \\ u(0) = 1. \end{cases}$$

Esta ecuación lineal se reduce al cálculo de una primitiva, obteniendo $u(t) = 1 - t$ como solución única del problema de Cauchy. Obsérvese que $u : \mathbb{R} \to \mathbb{R}$, pero debemos restringir $u : (-\infty, 1) \to \mathbb{R}$ para que $u(t) > 0$. Deshaciendo el cambio obtenemos

$$x(t) = \frac{1}{1-t},$$

cuyo intervalo de definición es $t \in (-\infty, 1)$, lo que hace $x(t) > 0$. Además, por el Teorema 2.4.2 sabemos que esta es la única solución del problema de Cauchy.

Ejemplo 2.3. Consideremos el problema de Cauchy similar

$$\begin{cases} x' = x^2, \\ x(0) = x_0, \end{cases}$$

con $x_0 \neq 0$. Sabemos por el Teorema 2.4.2 que si $x_0 > 0$ entonces tenemos existencia y unicidad. Si $x_0 < 0$, razonamos de manera similar al ejemplo anterior. En efecto, con el mismo cambio $u = 1/x$ llegamos al problema de Cauchy equivalente para la EDO lineal

$$\begin{cases} u' = -1, \\ u(0) = 1/x_0. \end{cases}$$

Este problema tiene como solución única $u(t) = -t + \dfrac{1}{x_0}$, y deshaciendo el cambio concluimos

$$x(t) = \frac{x_0}{1 - x_0\, t},$$

lo cual está definido para todo $t \neq 1/x_0$. Como las soluciones deben estar definidas en intervalos, tomamos el que contenga a $t = 0$, que es el instante donde planteamos el problema de Cauchy. Tenemos

$$x(t) = \frac{x_0}{1 - x_0\, t}, \quad t \in (-\infty, \frac{1}{x_0}) \text{ si } x_0 > 0, \qquad x(t) = \frac{x_0}{1 - x_0\, t}, \quad t \in (\frac{1}{x_0}, \infty) \text{ si } x_0 < 0.$$

En cuanto a la condición inicial $x_0 = 0$, sabemos que la función idénticamente nula es solución del problema de Cauchy correspondiente. Si existiese otra solución no nula, entonces $x(t) \neq 0$ en algún intervalo J. Ahora el cambio $u = x^{1-\alpha}$ nos diría que u es solución de la EDO lineal $u' = -1$ y por tanto

$$x(t) = \frac{1}{\lambda - t}, \qquad \lambda \in \mathbb{R}.$$

Para esta elección es imposible $x(0) = 0$ y por tanto si $x_0 = 0$ el problema de Cauchy tiene como solución única la función idénticamente nula.

Ejemplo 2.4. Consideremos ahora el problema de Cauchy

$$\begin{cases} x' = 3x^{2/3}, \\ x(1) = 1. \end{cases}$$

Ahora $u(t) = x(t)^{1/3}$ es solución del problema de Cauchy para la EDO lineal

$$\begin{cases} u' = 1, \\ u(1) = 1, \end{cases}$$

por lo que $u(t) = t$ y $x(t) = t^3$, la cual está definida en todo $\mathbb{R}$ aunque el método solo garantiza a priori la existencia y unicidad cuando $x(t) > 0$, esto es cuando $t > 0$. No obstante, podemos comprobar directamente que $x(t) = t^3$ es solución de la EDO $x' = 3x^{2/3}$ para todo $t \in \mathbb{R}$. En cuanto a la unicidad, el Teorema 2.4.2 nos la garantiza para cualquier otra solución del problema de Cauchy que sea positiva.

En este caso concreto podemos decir un poco más: la función $x(t) = t^3$ es la única solución del problema de Cauchy

$$\begin{cases} x' = 3x^{2/3}, \\ x(t_0) = t_0^3, \end{cases}$$

siempre y cuando $t_0 \neq 0$; incluso para $t_0 < 0$. En efecto, dado $t_0 \neq 0$ consideremos el problema de Cauchy

$$\begin{cases} u' = 1, \\ u(t_0) = t_0, \end{cases}$$

que tiene como única solución $u(t) = t$ definida en $\mathbb{R}$. Por ser $t_0 \neq 0$ necesariamente $t_0 > 0$ ó $t_0 < 0$. En el primer caso restringimos $t \in (0, \infty)$; en el segundo caso $t \in (-\infty, 0)$. En cualquier caso $u(t) \neq 0$ y la función $x(t) = t^3$ es la única solución del problema de Cauchy en cada uno de estos casos. En efecto, si $y(t)$ fuese otra solución del problema de Cauchy, entonces en un entorno de t_0 necesariamente $y(t_0) \neq 0$ y el cambio $y(t) = v(t)^{1/3}$ nos llevaría localmente al problema de Cauchy

$$\begin{cases} v' = 1, \\ v(t_0) = t_0, \end{cases}$$

cuya solución vuelve a ser $v(t) = t$ y por tanto $x(t) = y(t)$.

Ahora bien, si $t_0 = 0$ sabemos que $x(t) = t^3$ es solución, el cambio $u(t) = x(t)^{1/3} = t$ es C^1 y cumple $u'(t) = 1,\ u(0) = 0$. Pero si $y(t)$ es otra solución distinta tal que $y(0) = 0$, el cambio $v(t) = y(t)^{1/3}$ es C^1 siempre y cuando $y(t) \neq 0$. En particular, en ningún intervalo que contenga a $t_0 = 0$ podemos garantizar unicidad del problema (y de hecho no se puede garantizar en ningún intervalo donde $y(t)$ se anule). De hecho, existen infinitas soluciones de este problema dadas por

$$y(t) = \begin{cases} (t - c_1)^3 & \text{si} \quad t < c_1, \\ 0 & \text{si} \quad c_1 \leq t \leq c_2, \\ (t - c_2)^3 & \text{si} \quad t > c_2, \end{cases}$$

con $-\infty \leq c_1 \leq 0 \leq c_2 \leq \infty$.

Como se puso de manifiesto al comienzo de este tema, en la práctica se aplica despreocupadamente el cambio de variable y a posteriori se verifica que las funciones obtenidas aplicando el método son, de hecho, soluciones de la EDO. Este enfoque tiene la ventaja de poderse aplicar a situaciones menos restrictivas que las pedidas en las hipótesis del Teorema 2.4.2. Por ejemplo, consideremos la EDO

$$(x^2)' = x^2 \cos t - x^3 \sin t,$$

la cual está lejos de ser una ecuación de Bernoulli. Derivando el miembro de la izquierda tendríamos

$$2xx' = x^2 \cos t - x^3 \sin t.$$

Vemos que la función $x(t) = 0$ es solución, así que suponemos que x no es idénticamente nula. Restringimos entonces x a un cierto intervalo J donde $x(t) \neq 0$ para todo $t \in J$ (sin preocuparnos por si J es único o si existen varios, incluso infinitos, intervalos de este tipo) simplificamos y llegamos a

$$x' = \frac{1}{2}x\cos t - \frac{1}{2}x^2 \sin t,$$

que ahora sí es una EDO de Bernoulli. Haciendo el cambio $u = 1/x$ llegaríamos a la EDO lineal asociada. El precio que hay que pagar es que no podemos asegurar que hayamos obtenido todas las posibles soluciones. Veamos otro ejemplo donde realizaremos los cálculos pertinentes sin pararnos en el rigor matemático.

Ejemplo 2.5. Resolvamos la EDO

$$x' = -x + tx^3.$$

Tras el cambio $u = x^{-2}$ llegamos a que u es solución de la EDO lineal

$$u' = 2u - 2t.$$

La solución de la EDO homogénea es $u(t) = \lambda e^{2t}$, mientras que una solución particular $u_p(t)$ se puede buscar como $u_p(t) = \lambda(t)e^{2t}$, con lo que concluimos $\lambda(t) = (t + 1/2)e^{-2t}$ y por tanto

$$u_p(t) = t + \frac{1}{2}.$$

En consecuencia,

$$u(t) = \lambda e^{2t} + t + \frac{1}{2}, \qquad \lambda \in \mathbb{R},$$

y deshaciendo el cambio concluimos

$$x(t) = \frac{1}{(\lambda e^{2t} + t + \frac{1}{2})^{1/2}}, \qquad \lambda \in \mathbb{R}.$$

A continuación estudiamos estas funciones en un intervalo adecuado para que estén bien definidas y para tal fin definimos la función $z(t) = \lambda e^{2t} + t + \frac{1}{2}$. Notemos que no podemos resolver explícitamente la inecuación $z(t) > 0$ en términos de t y por eso tenemos que hace un análisis en mayor profundidad.

- Si $\lambda \geq 0$ entonces $z'(t) = 2\lambda e^{2t} + 1$. Además, $\lim_{t\to-\infty} z(t) = -\infty$ y $\lim_{t\to\infty} z(t) = \infty$ y por tanto existe un único t_λ tal que $z(t_\lambda) = 0$. En consecuencia, $x(t)$ está definida en (t_λ, ∞).
- Si $\lambda < 0$ entonces $\lim_{t\to-\infty} z(t) = -\infty$. Además, $z'(t) = 0$ para $t_\lambda = -1/2\log(-2\lambda)$, donde $z(t)$ alcanza un máximo absoluto. Por otra parte, $z(t_\lambda) = t_\lambda$ y $z(t)$ tomará valores positivos si y sólo si $t_\lambda > 0$, lo que equivale a $\lambda \in (-1/2, 0)$. Llamando t_1, t_2 a las soluciones de $z(t) = 0$, concluimos que $x(t)$ estará definida en (t_1, t_2), donde $z(t) > 0$.

2.4.1. El caso $\alpha = 2$

A continuación estudiaremos con mayor profundidad el caso $\alpha = 2$ por ser de importancia en la próxima sección, y este estudio se realizará en paralelo al del problema de Cauchy

$$\begin{cases} x' = a(t)x + b(t)x^2, \\ x(t_0) = x_0. \end{cases} \tag{2.4.4}$$

Como es usual, $a, b : I \subset \mathbb{R} \to \mathbb{R}$ son funciones continuas. Primero observemos que si $x_0 = 0$ entonces la función $x(t) = 0$ es solución de (2.4.4). De hecho, vamos a demostrar que es su única solución.

Proposición 2.4.3. *Si x es solución del Problema de Cauchy* (2.4.4) *y* $x(t_0) = 0$*, entonces* $x(t) = 0$ *para todo* $t \in I$*.*

Demostración. Sea $x(t)$ solución de (2.4.4). Por reducción al absurdo, supongamos que x no es idénticamente nula, esto es existe $t_1 \neq t_0$ tal que $x(t_1) \neq 0$. Podemos suponer sin perder generalidad que $t_1 > t_0$, tomar $\varepsilon > 0$ tal que $x(t_1 + \varepsilon)$ tiene el mismo signo que $x(t_1)$ y redefinir t_0 como

$$t_0 = \max\{t < t_1 \colon x(t) = 0\}.$$

Todo esto asegura que $x(t) \neq 0$ para todo $t \in (t_0, t_1 + \varepsilon)$. Consideramos ahora el problema de Cauchy

$$\begin{cases} u' = -a(t)u - b(t), \\ u(t_1) = 1/x(t_1), \end{cases}$$

que tiene existencia y unicidad dada por $u : I \to \mathbb{R}$ definida en todo el intervalo I de definición de las funciones $a(t), b(t)$. Como quiera que $v(t) = 1/x(t)$ para $t \in (t_0, t_1 + \varepsilon)$ es solución igualmente de este problema de Cauchy, necesariamente

$$u(t) = \frac{1}{x(t)}, \qquad \forall t \in (t_0, t_1 + \varepsilon).$$

Por una parte $u(t_0)$ está bien definida, pero

$$u(t_0) = \lim_{t \to t_0^+} u(t) = \lim_{t \to t_0^+} \frac{1}{x(t)} = \pm\infty,$$

dependiendo del signo de $x(t_1)$. Esta contradicción surge de suponer $x(t_1) \neq 0$ y por tanto $x(t) = 0$ para todo $t \in \mathbb{R}$.

$\square$

Observación 2.5. Un argumento similar se puede usar para probar que el problema de Cauchy

$$\begin{cases} x' = a(t)x + b(t)x^\alpha, \\ x(t_0) = 0, \end{cases}$$

con $\alpha > 1$, tiene como única solución a la función constantemente nula. Aquí la clave es que para $\alpha > 1$, el cambio de variable $u = x^{1-\alpha} = 1/x^{\alpha-1}$ por una parte debe estar definida en todo I pero por otra parte en t_0 tiene límite infinito.

Estamos en disposición de probar el resultado de existencia y unicidad del problema de Cauchy (2.4.4).

Teorema 2.4.4. *El problema de Cauchy* (2.4.4) *tiene existencia y unicidad.*

Demostración. Si $x_0 = 0$ sabemos que la función nula es la única solución de este problema de Cauchy. Suponemos $x_0 \neq 0$ y consideremos el problema de Cauchy para la EDO lineal

$$\begin{cases} u' = -a(t)u - b(t), \\ u(t_0) = 1/x_0, \end{cases}$$

para cierto $t_0 \in I$. Remarcamos que la condición inicial está bien definida por ser $x_0 \neq 0$. Sabemos que existe una única solución de este problema de Cauchy, $u : I \subset \mathbb{R} \to \mathbb{R}$, definida en el mismo intervalo I que las funciones $a(t), b(t)$. Ahora bien, consideramos

$$J = \{t \in I \colon u(t) \neq 0\},$$

que por continuidad de u es una unión disjunta, posiblemente infinita, de intervalos abiertos, y denotemos por $J_0 \subset J$ al intervalo que contiene a t_0. Entonces, la función $x(t) = 1/u(t)$ para $t \in J$ está bien definida y es solución de la EDO de Bernoulli

$$x' = a(t)x + b(t)x^2.$$

Más aún, $x : J_0 \to \mathbb{R}$ es la única solución del problema de Cauchy (2.4.4). □

2.5. La ecuación de Riccati

Una vez analizadas las EDOs de Bernoulli, a continuación estudiaremos las EDOs de Riccati.

Definición 2.5.1. Una ecuación diferencial se dice de Riccati si es de la forma

$$x' = a(t)x^2 + b(t)x + c(t), \tag{2.5.1}$$

donde $a, b, c : I \subset \mathbb{R} \to \mathbb{R}$ son funciones continuas.

Si $c(t) \equiv 0$ obtenemos una ecuación de Bernoulli con $\alpha = 2$. Si $a(t) \equiv 0$ obtenemos una ecuación lineal.

La forma de proceder para resolver la ecuación de Riccati es distinta a la de Bernoulli, pues para tal fin necesitamos conocer previamente una solución particular $x_p(t)$. Desafortunadamente, no existe ningún método general para obtener tal solución particular y deberemos proceder por tanteo o intuición. En el caso en que las funciones a, b, c sean constantes, una solución particular es alguna de las raíces reales (si existen) del polinomio $ax^2 + bx + c$.

Teorema 2.5.2. *Sean x, x_p soluciones de la EDO de Riccati* (2.5.1)*. Entonces, $z = x - x_p$ es solución de la EDO de Bernoulli*

$$z' = (2x_p(t)a(t) + b(t))z + a(t)z^2. \tag{2.5.2}$$

Recíprocamente, si x_p es solución particular de (2.5.1) *y z solución de* (2.5.2)*, entonces $x = x_p + z$ es solución de* (2.5.1)*. En particular, el problema de Cauchy*

$$\begin{cases} x' = a(t)x(t)^2 + b(t)x(t) + c(t), \\ x(t_0) = x_0, \end{cases} \tag{2.5.3}$$

tiene existencia y unicidad.

Antes de probar este resultado hay que remarcar que x_p es conocida y por tanto el término que acompaña a z es una función determinada.

Demostración. Sea x_p una solución particular de (2.5.1), tomemos x una solución arbitraria de (2.5.1) y definamos $z(t) = x(t) - x_p(t)$. Entonces,

$$\begin{aligned} z' &= x' - x_p' \\ &= a(t)(x^2 - x_p^2) + b(t)(x - x_p) \\ &= a(t)(x - x_p)(x + x_p) + b(t)(x - x_p) \\ &= a(t)(x - x_p)(x - x_p + 2x_p) + b(t)(x - x_p) \\ &= a(t)z(z + 2x_p) + b(t)z \\ &= a(t)z^2 + (2x_p a(t) + b(t))z. \end{aligned}$$

En consecuencia, z es solución de (2.5.2). Recíprocamente, si z es solución de la EDO (2.5.2) y x_p es solución de (2.5.1), entonces es inmediato comprobar que $x = x_p + z$ es solución de (2.5.1).

En cuanto al problema de Cauchy (2.5.3), sea x_p solución particular de (2.5.1). Si $x_p(t_0) = x_0$ hemos terminado. De lo contrario, definimos $x_1 = x_p(t_0)$ y en particular $x_1 \neq x_0$. Sabemos por el Teorema 2.4.4 que el problema de Cauchy para la EDO de Bernoulli

$$\begin{cases} z' = (2x_p(t)a(t) + b(t))z + a(t)z^2, \\ z(t_0) = x_0 - x_1, \end{cases} \tag{2.5.4}$$

tiene existencia y unicidad. De hecho, este problema de Cauchy es equivalente al problema para la EDO lineal

$$\begin{cases} u' = -(2x_p(t)a(t) + b(t))u - a(t), \\ u(t_0) = \frac{1}{x_0 - x_1}. \end{cases}$$

Aquí el hecho de que $x_0 \neq x_1$ es fundamental. Si u es solución de este último problema de Cauchy, entonces $u(t_0) \neq 0$ y en particular existe J_0 tal que $t_0 \in J_0$ y $u : J_0 \to \mathbb{R}$ cumple $u(t) \neq 0$ para todo $t \in J_0$. Definiendo $z(t) = 1/u(t)$, que está perfectamente definida en J_0 tenemos que z es solución del problema de Cauchy (2.5.4) y finalmente $x = z + x_p$ es la solución buscada del problema de Cauchy (2.5.3), lo que concluye la demostración. □

Observación 2.6. De la demostración del teorema anterior podemos deducir que si x_p es solución particular de (2.5.1) y x es tal que $x(t) \neq x_p(t)$ para todo $t \in J$, entonces x es solución de (2.5.1) si y sólo si $u(t) = \frac{1}{x(t) - x_p(t)}$, con $t \in J$, es solución de la EDO lineal completa $u' = -(2x_p a(t) + b(t))u - a(t)$.

Ejemplo 2.6. Considérese la ecuación diferencial

$$x' = t^3 x^2 + \frac{x}{t} - t^5.$$

Una opción en este tipo de ecuaciones es buscar soluciones polinómicas. Sea pues $x(t) = t^k$ y forcemos a que sea solución de la EDO.

$$kt^{k-1} = t^{k-1} + t^{2k+3} - t^5.$$

Si $k = 1$ esta igualdad se da y por tanto $x_p(t) = t$ es solución particular. Definiendo $z(t) = x(t) - x_p(t)$, sabemos que z es solución de la EDO de Bernoulli

$$z' = (2t^4 + \frac{1}{t})z + t^3 z^2.$$

Realizando el cambio $u = 1/z$ se tiene que u es solución de

$$u' = -(2t^4 + \frac{1}{t})u - t^3,$$

la cual es ya una EDO lineal que sabemos integrar. En efecto, la EDO homogénea

$$u' = -(2t^4 + \frac{1}{t})u$$

tiene como solución

$$u(t) = \lambda \frac{1}{t} e^{-\frac{2}{5}t^5}, \qquad \lambda \in \mathbb{R}.$$

Si buscamos una solución particular de la forma

$$u_p(t) = \lambda(t) \frac{1}{t} e^{-\frac{2}{5}t^5},$$

tenemos que $\lambda(t)$ viene dada por $\lambda(t) = \dfrac{-1}{2} e^{\frac{2}{5}t^5}$ y en consecuencia $u_p(t) = \frac{-1}{2t}$. La solución de la EDO lineal es

$$u(t) = \lambda \frac{1}{t} e^{-\frac{2}{5}t^5} - \frac{1}{2t}, \qquad \lambda \in \mathbb{R}.$$

Deshaciendo el primer cambio $u = 1/z(t)$ llegamos a

$$z(t) = \frac{2te^{\frac{2t^5}{5}}}{\lambda - e^{\frac{2t^5}{5}}}, \qquad \lambda \in \mathbb{R}.$$

Por último deshacemos el cambio final $x = z + x_p$, obteniendo

$$x(t) = \frac{2te^{\frac{2t^5}{5}}}{\lambda - e^{\frac{2t^5}{5}}} + t, \qquad \lambda \in \mathbb{R}.$$

Quedaría por ver cuándo estas funciones están bien definidas y para ello llamemos $z(t) = \lambda - e^{\frac{2t^5}{5}}$. Si $\lambda \leq 0$, se tiene que $z(t) < 0$ para todo $t \in \mathbb{R}$. Si $\lambda > 0$, entonces $z(t) = 0$ para

$$t_* = \left(\frac{5}{2}\log\lambda\right)^{1/5}.$$

Por tanto, $x(t)$ está bien definida para todo $t \in \mathbb{R} - \{t_*\}$.

2.5.1. Aplicación al crecimiento de poblaciones. Modelo de Verhulst

El modelo de Verhulst, llamado también modelo logístico, es una formulación matemática que describe el crecimiento de una población cuando existe un conocido como *límite ambiental* máximo sostenible, también conocido como capacidad de carga. Por ejemplo, en el desarrollo de un embrión el óvulo fecundado comienza a crecer doblando en cada iteración el número de células, lo cual da inicialmente un crecimiento exponencial. No obstante, este crecimiento está obviamente acotado por un tamaño que el útero materno pueda soportar, lo cual hace que la tasa de crecimiento disminuya con el tiempo; aunque sigue creciendo, éste lo hace cada vez menos. Introducimos a continuación el modelo en cuestión.

Sea $P(t)$ la cantidad de una determinada población en tiempo t, sea $r > 0$ la tasa intrínseca de crecimiento y $K > 0$ el límite ambiental o capacidad de carga del medio. Por ejemplo. Entonces, la tasa de crecimiento de la población se modela por la EDO

$$P' = rP\left(1 - \frac{P}{K}\right).$$

Esta EDO es de tipo Riccati y por tanto podemos usar toda la teoría que hemos desarrollado para deducir la expresión explícita de su solución, así como algunas propiedades de unicidad. Generalmente, las constantes r, K se ajustan mediante algún argumento de *interpolación*, usando datos pasados de crecimiento para intentar ajustar lo máximo posible la solución y poder deducir comportamientos futuros. Algunos comentarios sobre esta EDO antes de proceder a resolverla explícitamente.

- Si $P(t)$ es pequeño en comparación con K, esto es se tiene $P(t)/K \approx 0$, entonces $P' \approx rP$ y el crecimiento es casi exponencial. No obstante, si $P(t) \to K$ se tiene $P'(t) \approx 0$ y el crecimiento se frena.
- Si $P(t)$ comienza siendo mayor que K entonces $P'(t) < 0$ y la población decrecería hasta converger a su valor límite K.
- Supongamos que existe t_0 tal que $P(t_0) = K$. Entonces, $P(t) = K$ para todo t; esto es, o bien $P(t)$ nunca llega a su valor límite en tiempo finito, o bien la solución es constantemente la capacidad de carga. Para probarlo, observemos que la función constante K es trivialmente solución de la EDO y en particular del problema de Cauchy

$$\begin{cases} P' = rP\left(1 - \dfrac{P}{K}\right), \\ P(t_0) = K. \end{cases} \tag{2.5.5}$$

  Este problema de Cauchy tiene existencia y unicidad, como probamos en el Teorema 2.5.2, y por tanto si $P(t_0) = K$ necesariamente $P(t) = K$ para todo t.

Una vez realizadas estas observaciones, resolvemos explícitamente el problema de Cauchy. Supongamos que la población inicial es $P(0) = P_0 \neq K$ y escribamos la EDO de la forma usual

$$P' = -\frac{r}{K}P^2 + rP.$$

Es decir, con la notación de la ecuación (2.5.1), se tiene $a(t) = -r/K$, $b(t) = r$ y $c(t) = 0$. Sabemos que una solución particular de esta EDO es $x_p(t) = K$ y por tanto mediante el cambio $z(t) = P(t) - K$ podemos considerar el problema de Cauchy para la EDO de Bernoulli

$$\begin{cases} z' = -rz - \frac{r}{K}z^2, \\ z(t_0) = P_0 - K. \end{cases}$$

Mediante el cambio de variable $u(t) = 1/z(t)$, este problema de Cauchy se reduce al problema para la EDO lineal

$$\begin{cases} u' = ru + \frac{r}{K}, \\ u(t_0) = \frac{1}{P_0 - K}. \end{cases}$$

La solución de este problema de Cauchy es

$$u(t) = \frac{P_0}{K(P_0 - K)} e^{r(t-t_0)} - \frac{1}{K},$$

y por tanto

$$z(t) = \frac{K(P_0 - K)}{K + P_0(e^{r(t-t_0)} - 1)}$$

es solución del problema de Cauchy para la EDO de Bernoulli. Finalmente, concluimos que la solución del problema de Cauchy (2.5.5) es

$$P(t) = \frac{K P_0 e^{r(t-t_0)}}{K + P_0(e^{r(t-t_0)} - 1)}.$$

2.6. Ecuaciones en variables separadas

Las ecuaciones que hemos visto hasta ahora son todas reducibles a lineales, pero este no es el escenario lógico (ni deseable, ¡de lo contrario la asignatura carecería de interés!). Las ecuaciones diferenciales más sencillas que no son reducibles a lineales, al menos de forma general, son las de *variables separadas*.

Definición 2.6.1. Una ecuación diferencial se dice de variables separadas si se puede expresar de la forma

$$x'(t) = f(t)g(x(t)), \tag{2.6.1}$$

con $f : I \subset \mathbb{R} \to \mathbb{R},\ g : J \subset \mathbb{R} \to \mathbb{R}$ funciones continuas.

Para resolver esta ecuación, una idea preliminar que hemos empleado *alegremente* en el primer tema de esta asignatura, y como veremos totalmente acertada, es dividir por $g(x(t))$ e integrar a ambos lados, usando en el de la izquierda la regla de la cadena. No obstante, debemos cuidar la formalidad de estos pasos y para ello necesitamos introducir notación. Definimos

$$J_0 = \{x \in J \colon g(x) \neq 0\}.$$

Por ser g continua en J, se tiene que J_0 es un conjunto abierto en J. Es más, J_0 no puede tener puntos aislados y por tanto si $x_0 \in J_0$ entonces existe $\varepsilon > 0$ tal que $(x_0 - \varepsilon, x_0 + \varepsilon) \subset J_0$ (si x_0 es frontera entonces se tiene $[x_0, x_0 + \epsilon) \subset J_0$ ó $(x_0 - \epsilon, x_0] \subset J_0$). En consecuencia, J_0 es unión disjunta (tal vez infinita) de intervalos abiertos, donde por abiertos se entiende en la topología relativa de J_0. Tiene pues sentido considerar una primitiva de $1/g(x)$ en J_0, a la cual denotaremos por $G(x)$. Igualmente, denotamos por $F(t)$ a una primitiva de f. Por último, definimos

$$H : I \times J_0 \to \mathbb{R}, \qquad H(t,x) = G(x) - F(t). \tag{2.6.2}$$

El resultado principal de esta sección es el siguiente.

Teorema 2.6.2. *Sea $x : I_0 \subset I \to \mathbb{R}$ de clase C^1 y tal que $x(I_0) \subset J_0$. Entonces, x es solución de la EDO en variables separadas (2.6.1) si y sólo si está definida implícitamente por la ecuación*

$$H(t, x(t)) = \lambda, \qquad \lambda \in \mathbb{R}, \tag{2.6.3}$$

donde $H(t,x)$ viene dada por (2.6.2). Más aún, dados $t_0 \in I$ y $x_0 \in J_0$, el problema de Cauchy

$$\begin{cases} x'(t) = f(t)g(x(t)), \\ x(t_0) = x_0, \end{cases} \tag{2.6.4}$$

tiene existencia y unicidad dada por la única solución implícita de

$$H(t, x(t)) = H(t_0, x_0).$$

Demostración. Sea $I_0 \subset I$ un intervalo y $x \in C^1(I_0)$ tal que $x(I_0) \subset J_0$. En particular $g(x(t)) \neq 0$ para todo $t \in I_0$ y tiene sentido considerar en (2.6.1) el cociente

$$\frac{x'(t)}{g(x(t))} = f(t).$$

Integrando en ambas partes respecto de t y usando la regla de la cadena en el miembro de la izquierda se tiene

$$G(x(t)) = F(t) + \lambda, \qquad \lambda \in \mathbb{R},$$

En otras palabras, si definimos la función $H(t,x)$ por (2.6.2), entonces $x(t)$ es solución de (2.6.1) si y sólo si x es solución de la ecuación implícita (2.6.3) para algún $\lambda \in \mathbb{R}$. Por otra parte, si x viene definida por la relación implícita (2.6.2), derivando a ambos lados concluimos que x es solución de la EDO (2.6.1).

Vamos a probar existencia y unicidad de cada problema de Cauchy. Tomemos $t_0 \in I$ y $x_0 \in J_0$. Como J_0 es una unión disjunta de intervalos abiertos, denotamos por J_0^* a la componente conexa de J_0 que contiene a x_0. Suponemos $g(x)$ restringida a J_0^* y por tanto $G(x)$ es una primitiva de $1/g(x)$ una vez restringida a J_0^*. Consideremos de nuevo $H(t,x)$ dada por (2.6.2) y definamos

$$\widehat{H} : I \times J_0^* \to \mathbb{R}, \qquad \widehat{H}(t,x) = H(t,x) - H(t_0,x_0).$$

Observemos que se cumple

$$\widehat{H}(t_0,x_0) = 0, \qquad \frac{\partial \widehat{H}}{\partial x}(t_0,x_0) = \frac{1}{g(x_0)} \neq 0,$$

y por tanto el Teorema de la función implícita garantiza la existencia de abiertos $U \subset I \times J_0^*$ e $I_0 \subset I$ tales que $(t_0,x_0) \in U$, $t_0 \in I_0$, y de una función $x : I_0 \to \mathbb{R}$ de clase C^1 satisfaciendo $x(t_0) = x_0$ y $(t, x(t)) \in U$ y además

$$\widehat{H}(t, x(t)) = 0, \qquad \forall t \in I_0.$$

Nótese que como $U \subset I \times J_0^*$ y $(t, x(t)) \in U$ en particular $x : I_0 \to J_0^*$ y por tanto $1/g(x(t))$ está perfectamente definida. Como quiera que $\widehat{H}(t,x(t)) = H(t,x(t)) - H(t_0,x_0)$ concluimos que la solución $x(t)$ dada por el Teorema de la función implícita está caracterizada por cumplir la relación

$$H(t, x(t)) = H(t_0, x_0).$$

Finalmente, la unicidad del Teorema de la función implícita garantiza la unicidad del problema de Cauchy. $\square$

Observación 2.7. Sea $x_0 \in J - J_0$, esto es $g(x_0) = 0$. Entonces, la función $x(t) = x_0$ es solución de (2.6.1). En relación con el problema de Cauchy (2.6.4) para la condición inicial $x(t_0) = x_0$, tenemos garantizada la existencia pero no podemos asegurar la unicidad. Por ejemplo, consideremos los problemas de Cauchy

$$\begin{cases} x' = x^2, \\ x(t_0) = 0, \end{cases} \qquad \begin{cases} x' = x^{2/3}, \\ x(t_0) = 0. \end{cases}$$

Sabemos por el Teorema 2.4.4 que $x(t) = 0$ es la única solución del primero, mientras que el segundo problema de Cauchy tiene infinitas soluciones dadas por las funciones

$$\begin{cases} (t - c_1)^3 & \text{si} \quad t < c_1, \\ \quad 0 & \text{si} \quad c_1 \leq t \leq c_2, \\ (t - c_2)^3 & \text{si} \quad t > c_2, \end{cases}$$

con $-\infty \leq c_1 \leq t_0 \leq c_2 \leq \infty$.

Ejemplo 2.7. Resolvamos

$$\begin{cases} x' = \dfrac{x \cos t}{1 + 2x^2}, \\ x(0) = 1. \end{cases}$$

Con la notación anterior, $f(t) = \cos t$, $g(x) = \frac{x}{1+2x^2}$ y

$$J_0 = \{x \in \mathbb{R} \colon g(x) \neq 0\} = \mathbb{R} - \{0\}.$$

Como la condición inicial es $x(0) = 1$, nos quedamos con la componente conexa $J_0^* \subset J_0$ que contiene al valor 1. En consecuencia, $J_0^* = (0, \infty)$, restringimos $g : J_0^* \to \mathbb{R}$ y una primitiva de $1/g(x)$ en J_0^* viene dada por

$$G(x) = \int \frac{1 + 2x^2}{x}\, dx = \log x + x^2.$$

Como primitiva de $f(t) = \cos t$ tomamos $F(t) = \sin t$. Sabemos entonces que x es solución de la EDO si y sólo si x es solución implícita de $H(t, x(t)) = G(x(t)) - F(t) = \lambda,\ \lambda \in \mathbb{R}$, donde

$$H(t, x) = \log x + x^2 - \sin t,$$

lo cual implica

$$\log x(t) + x(t)^2 = \sin t + \lambda, \qquad \lambda \in \mathbb{R}.$$

Finalmente, la solución del problema de Cauchy se obtiene al tomar $\lambda = H(0, 1) = 1$,

$$\log x(t) + x(t)^2 = \sin t + 1.$$

Notemos que $\log x(t)$ está perfectamente definido puesto que hemos restringido a $J_0^* = (0, \infty)$ y por tanto $x(t) > 0$.

Observación 2.8. Supongamos la ecuación implícita

$$\log x(t) + x(t)^2 = \sin t + 1, \tag{2.6.5}$$

definida como $H(t, x(t)) = 0$, con $H(t, x) = \log x + x^2 - \sin t - 1$. El Teorema de la función implícita asegura la existencia de una solución local $x(t)$ alrededor del punto $(0, 1)$, ya que $H(0, 1) = 0$ y

$$\frac{\partial H}{\partial x}(0,1) = 3 \neq 0.$$

Del tema pasado sabemos que $x(t)$ es solución de la EDO obtenida al derivar la ecuación (2.6.5) respecto de t y aplicar la regla de la cadena, obteniendo

$$\frac{x'}{x} + 2xx' = \cos t,$$

y despejando x' se tiene

$$x' = \frac{x \cos t}{1 + 2x^2},$$

obteniendo (como no podía ser de otra forma) la EDO original.

2.7. Ecuaciones homogéneas

Un tipo de ecuaciones que son siempre reducibles a variables separadas son las denominadas *homogéneas*. Recordemos que una función $f : \mathbb{R}^2 \to \mathbb{R}$ se dice homogénea de grado $\alpha \in \mathbb{R}$ si

$$f(\lambda x, \lambda y) = \lambda^\alpha f(x, y).$$

Definición 2.7.1. Una ecuación diferencial se dice homogénea si se puede escribir como

$$x' = \frac{M(t,x)}{N(t,x)}, \tag{2.7.1}$$

con M, N funciones homogéneas del mismo grado

La resolución de (2.7.1) pasa por realizar un cambio de variable. Concretamente, tenemos el siguiente resultado

Teorema 2.7.2. *Sea x una solución de la EDO homogénea (2.7.1). Entonces, $u = x/t$, con $t \neq 0$, es solución de la EDO en variables separadas*

$$u' = \frac{1}{t}\left(\frac{M(1,u)}{N(1,u)} - u\right). \tag{2.7.2}$$

Recíprocamente, si u es solución de (2.7.2) y definimos $x = tu$, entonces x es solución de la EDO homogénea (2.7.1). En particular, dados $t_0 \neq 0$ y x_0 tales que $N(t_0, x_0) \neq 0$ y

$$\frac{M(t_0,x_0)}{N(t_0,x_0)} \neq \frac{x_0}{t_0}, \tag{2.7.3}$$

el problema de Cauchy

$$\begin{cases} x' = \dfrac{M(t,x)}{N(t,x)}, \\ x(t_0) = x_0, \end{cases} \tag{2.7.4}$$

tiene existencia y unicidad.

Demostración. Sea x solución de (2.7.1) y consideremos $u = x/t$ con $t \neq 0$. Derivando,

$$u' = \frac{x'}{t} - \frac{x}{t^2} = \frac{1}{t}\left(\frac{M(t,x)}{N(t,x)} - \frac{x}{t}\right) = \frac{1}{t}\left(\frac{M(1,u)}{N(1,u)} - u\right),$$

donde en la última igualdad se ha usado que M y N son homogéneas del mismo grado. Análogamente, si u es solución de (2.7.2) y $x = tu$, entonces

$$x' = u + tu' = u + \frac{M(1,u)}{N(1,u)} - u = \frac{t^\alpha M(1,u)}{t^\alpha N(1,u)} = \frac{M(t,tu)}{N(t,tu)} = \frac{M(t,x)}{N(t,x)},$$

lo que concluye la prueba. Referente la existencia y unicidad de (2.7.4), consideremos el problema de Cauchy

$$\begin{cases} u' = \dfrac{1}{t}\left(\dfrac{M(1,u)}{N(1,u)} - u\right) \\ u(t_0) = \dfrac{x_0}{t_0}. \end{cases}$$

Si definimos $g(u) = \frac{M(1,u)}{N(1,u)} - u$, entonces este problema tiene existencia y unicidad siempre que $g(u_0) \neq 0$ en virtud del Teorema 2.6.2, lo que se traduce a que (2.7.3) se cumpla. Ahora definimos $x = tu$, que cumple (2.7.1) y la condición inicial $x(t_0) = x_0$ y por tanto es solución del problema de Cauchy (2.7.4). □

Observación 2.9. Si $t_0 \neq 0$ y x_0 son tales que se cumple

$$\frac{M(t_0,x_0)}{N(t_0,x_0)} = \frac{x_0}{t_0},$$

entonces de manera similar al caso de las ecuaciones en variables separadas podemos garantizar existencia del problema de Cauchy pero no unicidad. En efecto, es inmediato comprobar que la función $x(t) = tx_0/t_0$ es solución de la EDO homogénea (2.7.1).

Ejemplo 2.8. Resolvamos la EDO

$$x' = \frac{t+x}{t-x}.$$

El numerador y denominador son funciones homogéneas de grado 1 y por tanto podemos hacer el cambio de variable $x(t) = tu(t)$, siempre que $t \neq 0$, obteniendo la EDO en variables separadas

$$u' = \frac{1}{t}\left(\frac{1+u}{1-u} - u\right) = \frac{1}{t}\frac{1+u^2}{1-u}.$$

Integrando se tiene

$$\arctan u - \frac{1}{2}\log(1+u^2) = \log|t| + \lambda, \qquad \lambda \in \mathbb{R},$$

y deshaciendo el cambio $u = x/t$ obtenemos la solución implícita

$$\arctan\frac{x(t)}{t} - \frac{1}{2}\log(1 + \frac{x(t)^2}{t^2}) = \log|t| + \lambda, \qquad \lambda \in \mathbb{R},$$

o de forma equivalente

$$\arctan\frac{x(t)}{t} - \frac{1}{2}\log(t^2 + x^2) = \lambda, \qquad \lambda \in \mathbb{R},$$

2.8. Ecuaciones exactas

Sea $\Omega \subset \mathbb{R}^2$ un abierto y $H : \Omega \to \mathbb{R}$ una función de clase C^1. Supongamos que $x(t)$ es una función definida implícitamente por $H(t, x(t)) = \lambda$ para un cierto $\lambda \in \mathbb{R}$. Sabemos que x es solución de la EDO obtenida al derivar la relación $H(t, x(t))$ respecto de t, obteniendo

$$\frac{\partial H}{\partial t}(t,x) + \frac{\partial H}{\partial x}(t,x)x' = 0.$$

Recíprocamente, consideremos la EDO

$$M(t,x) + N(t,x)x' = 0, \tag{2.8.1}$$

donde $M, N : \Omega \to \mathbb{R}$ son continuas, y supongamos que existe $H = H(t,x) \in C^1(\Omega)$ tal que

$$\frac{\partial H}{\partial t}(t,x) = M(t,x), \qquad \frac{\partial H}{\partial x}(t,x) = N(t,x). \tag{2.8.2}$$

Si $x : I \to \mathbb{R}$ es solución de (2.8.1), entonces es inmediato que $\frac{d}{dt}H(t,x(t)) = 0$ y por tanto existe $\lambda \in \mathbb{R}$ de manera que $H(t,x(t)) = \lambda$ para todo $t \in I$. En otras palabras, x está definida implícitamente por la expresión $H(t,x(t)) = \lambda$.

Definición 2.8.1. Una EDO se dice exacta si es de la forma (2.8.1) y existe H de clase C^1 cumpliendo el criterio de exactitud (2.8.2).

Es común escribir el criterio de exactitud (2.8.2) en términos del gradiente de H. En efecto, si definimos $F : \Omega \to \mathbb{R},\ F(t,x) = (M(t,x), N(t,x))$, entonces (2.8.2) se escribe como la ecuación vectorial

$$\nabla H(t,x) = F(t,x).$$

En la literatura, se suelen llamar a los campos F satisfaciendo esta ecuación *campos gradientes* y si su dominio de definición es conexo entonces se habla de *campos conservativos*. En ambos casos, a la función H se le suele llamar *potencial*. El siguiente paso lógico es dar un criterio que nos permita asegurar cuándo una EDO de la forma (2.8.1) es exacta y en tal caso cómo calcular la función H cumpliendo (2.8.2).

Teorema 2.8.2. *Supongamos que M, N admiten derivadas parciales continuas en $I \times J$, con I, J intervalos abiertos. Entonces, son equivalentes en $I \times J$:*

1. *La ecuación (2.8.1) es exacta;*
2. $\frac{\partial M}{\partial x}(t,x) = \frac{\partial N}{\partial t}(t,x)$.

En tal caso, fijados $t_0 \in I,\ x_0 \in J$ y definiendo

$$H(t,x) = \int_{t_0}^{t} M(s,x)\,ds + \int_{x_0}^{x} N(t_0,u)\,du,$$

la solución de (2.8.1) viene dada por

$$H(t,x(t)) = \lambda, \qquad \lambda \in \mathbb{R}.$$

Además, si $N(t_0, x_0) \neq 0$*, el problema de Cauchy*

$$\begin{cases} M(t,x) + N(t,x)x' = 0, \\ x(t_0) = x_0, \end{cases}$$

tiene solución única dada de forma implícita por

$$H(t, x(t)) = 0.$$

Demostración. Supongamos que se cumple *1*. Entonces, existe H de clase C^1 cumpliendo (2.8.2). De hecho, H es de clase C^2 al ser M, N de clase C^1. Por el lema de Schwarz, las derivadas segundas cruzadas de H coinciden, esto es

$$\frac{\partial^2 H}{\partial t\, \partial x}(t,x) = \frac{\partial^2 H}{\partial x\, \partial t}(t,x).$$

Así pues, $\dfrac{\partial M}{\partial x}(t,x) = \dfrac{\partial N}{\partial t}(t,x)$ y hemos probado que *1*. implica *2*.

Supongamos ahora que *2*. es cierto. Queremos probar que (2.8.1) es exacta y para tal fin debemos encontrar $H(t,x)$ de clase C^1 tal que $\dfrac{\partial H}{\partial t}(t,x) = M(t,x)$. Sean $t_0 \in I$ y $x_0 \in J$. Si integramos entre t_0 y t obtenemos

$$H(t,x) = \int_{t_0}^{t} M(s,x)\, ds + H(t_0, x). \tag{2.8.3}$$

Para deducir la expresión de $\phi(x) = H(t_0, x)$ derivamos ahora respecto de x, aplicamos los teoremas de derivación bajo el signo integral y usamos que $\dfrac{\partial H}{\partial x}(t,x) = N(t,x)$, obteniendo

$$\frac{\partial H}{\partial x}(t,x) = \int_{t_0}^{t} \frac{\partial M}{\partial x}(s,x)\, ds + \phi'(x) = \int_{t_0}^{t} \frac{\partial N}{\partial s}(s,x)\, ds + \phi'(x),$$

de donde deducimos

$$\phi'(x) = N(t_0, x).$$

Integrando entre x_0 y x se tiene

$$H(t_0, x) = \int_{x_0}^{x} N(t_0, u)\, du + H(t_0, x_0).$$

Sustituyendo en (2.8.3) se tiene

$$H(t,x) = \int_{t_0}^{t} M(s,x)\, ds + \int_{x_0}^{x} N(t_0, u)\, du + H(t_0, x_0).$$

Esta función es de clase C^2 y cumple el criterio de exactitud (2.8.2); que $\dfrac{\partial H}{\partial t}(t,x) = M(t,x)$ es trivial y por otro lado

$$\frac{\partial H}{\partial x}(t,x) = \int_{t_0}^{t} \frac{\partial M}{\partial x}(s,x)\, ds + N(t_0, x) = \int_{t_0}^{t} \frac{\partial N}{\partial s}(s,x)\, ds + N(t_0, x) = N(t,x),$$

concluyendo la veracidad del punto *1*. Más aún, como $H(t,x)$ está definida módulo una constante aditiva, podemos tomar como potencial de la EDO exacta la función

$$H(t,x) = \int_{t_0}^{t} M(s,x)\,ds + \int_{x_0}^{x} N(t_0,u)\,du.$$

Además, por ser exacta, las soluciones de la EDO (2.8.1) vienen dadas implícitamente por la relación

$$H(t,x(t)) = \lambda, \qquad \lambda \in \mathbb{R},$$

donde se vuelve a poner de manifiesto la independencia de la constante elegida en la definición de $H(t,x)$. Por último, en relación con el problema de Cauchy, como $H(t_0,x_0) = 0$ y $\dfrac{\partial H}{\partial x}(t_0,x_0) = N(t_0,x_0) \neq 0$, la afirmación es consecuencia del Teorema de la función implícita. □

El Teorema 2.8.2 se ha enunciado para funciones definidas sobre rectángulos, y esta consideración es de vital importancia para poder definir correctamente las integrales y no *salirnos* del dominio de integración. Este resultado se podría probar para dominios más generales como son los dominios *estrellados*, o de forma más general aún los *simplemente conexos* (véanse los ejercicios 16, 17 y 18 de la relación).

Teorema 2.8.3. *Sea $\Omega \subset \mathbb{R}^2$ un dominio simplemente conexo y sean $M, N \colon \Omega \to \mathbb{R}$ de clase C^1 tales que $\dfrac{\partial M}{\partial x}(t,x) = \dfrac{\partial N}{\partial t}(t,x)$. Entonces, la EDO (2.8.1) es exacta, esto es existe $H \in C^2(\Omega)$ tal que $\dfrac{\partial H}{\partial t}(t,x) = M(t,x)$, $\dfrac{\partial H}{\partial x}(t,x) = N(t,x)$.*

A continuación, probaremos que el resultado no es cierto para un dominio no simplemente conexo. Considérense las funciones

$$M(t,x) = \frac{-x}{t^2+x^2}, \qquad N(t,x) = \frac{t}{t^2+x^2},$$

definidas en $\mathbb{R}^2 - \{(0,0)\}$. Por una parte,

$$\frac{\partial M(t,x)}{\partial x} = \frac{\partial N(t,x)}{\partial t} = \frac{x^2-t^2}{(t^2+x^2)^2}.$$

Sin embargo, la EDO

$$M(t,x(t)) + N(t,x(t))x'(t) = 0$$

no es exacta; es decir, no existe ninguna H de clase C^1 tal que $\dfrac{\partial H}{\partial t} = M$ y $\dfrac{\partial H}{\partial x} = N$. Para convencernos de esta afirmación, definimos

$$\theta(t,x) = \begin{cases} \arctan \dfrac{x}{t} & \text{si} \quad t > 0, \\ \pi + \arctan \dfrac{x}{t} & \text{si} \quad t < 0. \end{cases}$$

Es inmediato comprobar que $\dfrac{\partial \theta}{\partial t}(t,x) = M(t,x)$ y $\dfrac{\partial \theta}{\partial x}(t,x) = N(t,x)$ y por tanto

$$\frac{\partial (H-\theta)}{\partial t}(t,x) = 0, \qquad \frac{\partial (H-\theta)}{\partial x}(t,x) = 0,$$

para todo (t,x) con $t \neq 0$. Esto nos dice que la diferencia $H - \theta$ es constante en cada componente conexa de su dominio. Existen pues constantes λ_+, λ_- tales que

$$\begin{aligned} \theta(t,x) &= H(t,x) + \lambda_+, \qquad t > 0, \\ \theta(t,x) &= H(t,x) + \lambda_-, \qquad t < 0. \end{aligned}$$

Ahora bien, como $\theta(t,x)$ puede extenderse por continuidad a la semirrecta $(0,x),\ x > 0$, tomando el valor $\theta(0,x) := \pi/2$ y ya que $H(t,x)$ es continua en $\mathbb{R}^2 - \{0\}$, necesariamente $\lambda_+ = \lambda_- = \lambda$. Esto nos diría que $\theta(t,x)$ se puede extender por continuidad a toda la semirrecta $(0,x)$ con $x < 0$, pero esto es imposible. En efecto, si $x_0 < 0$ y $t_n \to 0$,

$$\begin{aligned} \theta(t_n, x_0) &\to -\frac{\pi}{2} \quad \text{si} \quad t_n > 0, \\ \theta(t_n, x_0) &\to \frac{3\pi}{2} \quad \text{si} \quad t_n < 0. \end{aligned}$$

2.9. Factor integrante

Realmente, el concepto de ecuación diferencial exacta es ambiguo. Podría ocurrir que la EDO (2.8.1) no sea exacta, pero que exista $\mu(t,x) \neq 0$ continua tal que

$$\mu(t,x)M(t,x) + \mu(t,x)N(t,x)x' = 0$$

sí que lo sea. Advertimos que estamos hablando de la misma ecuación debido a que $\mu(t,x)$ no se anula. En este punto, está más que justificado el estudio de funciones $\mu(t,x)$ que hagan exactas ecuaciones que originalmente no lo son.

Definición 2.9.1. Dada la EDO (2.8.1), se dice que una función $\mu(t,x)$ no nula es un factor integrante si al multiplicar (2.8.1) por $\mu(t,x)$, ésta se convierte en exacta.

En lo que sigue, supondremos adicionalmente que $\mu(t,x)$ es de clase C^1. En general, encontrar factores integrantes es una labor difícil. Una primera aproximación es imponer la condición de que $\mu(t,x)$ sea factor integrante, obteniendo

$$\mu(t,x)\frac{\partial M}{\partial x}(t,x) + M(t,x)\frac{\partial \mu}{\partial x}(t,x) = \mu(t,x)\frac{\partial N}{\partial t}(t,x) + N(t,x)\frac{\partial \mu}{\partial t}(t,x).$$

En consecuencia, ha de satisfacerse la *ecuación en derivadas parciales*

$$N(t,x)\frac{\partial \mu}{\partial t}(t,x) - M(t,x)\frac{\partial \mu}{\partial x}(t,x) = \mu(t,x)\left(\frac{\partial M}{\partial x}(t,x) - \frac{\partial N}{\partial t}(t,x)\right).$$

La teoría de EDPs asegura que esta ecuación siempre tiene existencia local y por tanto toda EDO (2.8.1) es exacta salvo multiplicar por un factor integrante. Otro problema radicalmente distinto es encontrar tal factor integrante, tarea que puede ser incluso más complicada que resolver la ecuación original.

A continuación, daremos algunos tipos concretos de factores integrales que serán con los que se tendrán que probar antes de nada.

1. Supongamos que $N(t,x) \neq 0$ y que $\mu(t,x)$ solo depende de la variable t. Entonces, $\partial\mu/\partial x = 0$ y se tiene
$$\frac{\mu'(t)}{\mu(t)} = \frac{\dfrac{\partial M}{\partial x}(t,x) - \dfrac{\partial N}{\partial t}(t,x)}{N(t,x)},$$
la cual es una EDO en variables separadas. Una forma de verificar si es admisible tal factor integrante es comprobar a priori si la función
$$\frac{\dfrac{\partial M}{\partial x}(t,x) - \dfrac{\partial N}{\partial t}(t,x)}{N(t,x)}$$
sólo depende de la variable t.

2. Análogamente, si $M(t,x) \neq 0$ podemos buscar un factor integrante $\mu(t,x) = \mu(x)$. En efecto, tal factor integrante ha de satisfacer la EDO
$$\frac{\mu'(x)}{\mu(x)} = \frac{\dfrac{\partial N}{\partial t}(t,x) - \dfrac{\partial M}{\partial x}(t,x)}{M(t,x)}.$$
La admisibilidad de tal factor integrante es similar, comprobando si
$$\frac{\dfrac{\partial N}{\partial t}(t,x) - \dfrac{\partial M}{\partial x}(t,x)}{M(t,x)}$$
sólo depende de x.

2.9.1. Aplicación al Cálculo de Variaciones. La braquistocrona

El problema de la braquistocrona ya ha aparecido en estas notas, concretamente en la introducción al Cálculo de Variaciones al final del Capítulo 1. La formulación, de nuevo, es la siguiente: dados dos puntos $A, B \in \mathbb{R}^2$, encontrar la trayectoria que une A y B de forma que el descenso de una partícula sobre tal trayectoria se realiza en el menor tiempo posible, asumiendo que sobre la partícula sólo actúa la fuerza de la gravedad.

Sean $A = (0, x_0)$ y $B = (t_0, 0)$ con $t_0, x_0 > 0$. La naturaleza del problema invita a pensar que la trayectoria, si existe, se podrá expresar como un grafo vertical y horizontal al mismo tiempo. Por otra parte, lo natural es pensar que inicialmente la partícula desciende de forma vertical para que adquiera inicialmente la mayor velocidad posible. En consecuencia, si denotamos por $(t, x(t))$ a tal grafo se tendría $x'(0) = -\infty$ y por este motivo es por el que es mejor expresar la solución como una función $t = t(x)$. El problema pues es *determinar una función* $t : [0, x_0] \to \mathbb{R}$ *tal que* $t(0) = t_0,\ t(x_0) = 0$ *y que el tiempo de recorrido sea mínimo*. La primera cuestión es determinar cómo se calcula el tiempo recorrido por una partícula. Por una parte, el espacio que la partícula recorre viene dado por el elemento de longitud, $\sqrt{1 + t'(x)^2}\, dx$. Por otra parte, de la conservación de la energía sabemos que la partícula tiene inicialmente una energía potencial $E_p = mgx_0$. En cada altura $t(x)$, la energía total viene dada por
$$mgx_0 = E_T = E_c + E_p = \frac{1}{2}mv^2 + mg(x_0 - x),$$

y en consecuencia $v = \sqrt{2g(x_0 - x)}$. Dada una función $t(x)$ con $t(0) = t_0$ y $t(x_0) = 0$, el tiempo transcurrido por una partícula recorriendo el grafo $(t(x), x)$ por la acción de la gravedad viene dado por

$$T[t] = \int_0^{x_0} \sqrt{\frac{1 + t'(x)^2}{2g(x_0 - x)}}.$$

El lagrangiano de este problema variacional es $L(x, t(x), t'(x)) = \sqrt{\frac{1+t'(x)^2}{2g(x_0-x)}}$, cuya ecuación de Euler-Lagrange es

$$\frac{\partial L}{\partial t} - \frac{d}{dx}\frac{\partial L}{\partial t'} = -\frac{d}{dx}\frac{\partial L}{\partial t'} = 0,$$

y en consecuencia

$$\frac{\partial L}{\partial t'} = \frac{t'(x)}{\sqrt{1 + t'(x)^2}} \frac{1}{\sqrt{2g(x_0 - x)}} = k, \qquad k \in \mathbb{R}.$$

Observemos que el sistema de referencia escogido hace que $t'(x) < 0$ localmente alrededor de $x = 0$, en consecuencia $k < 0$ y por tanto $t'(x) < 0$ para todo $x \in [0, x_0]$. Denotando $\lambda = \frac{1}{2gk^2}$ y despejando se tiene

$$t'(x) = -\sqrt{\frac{x_0 - x}{\lambda - x_0 + x}}.$$

Finalmente, integrando concluimos

$$t(x) = \sqrt{x - x_0 + \lambda}\sqrt{x_0 - x} - \lambda \arctan \frac{\sqrt{x_0 - x}}{\sqrt{\lambda - (x_0 - x)}} + c,$$

donde c es tal que $t(0) = t_0$.

Ejercicios tema 2

1. Se considera la EDO
$$x' = f(t, x),$$
con $f : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ continua. Supongamos que el conjunto de todas las soluciones de esta EDO es un espacio vectorial real. Probar que la EDO es lineal homogénea.

2. Sean $a, b : \mathbb{R} \to \mathbb{R}$ funciones continuas tales que existe $t_0 \in \mathbb{R}$ con $a(t) \geq c > 0$ para todo $t \geq t_0$. Supongamos además que
$$\lim_{t\to\infty} b(t) = 0.$$
Probar que las soluciones de la EDO $x' + a(t)x = b(t)$ tienden a cero cuando $t \to \infty$.

3. Considérese la ecuación
$$x' = f(t, x),$$
siendo $f \in C^\infty(\Omega)$ definida en cierto abierto $\Omega \subset \mathbb{R}^2$. Probar que sus soluciones son también C^∞.

4. Razonar las siguientes preguntas.

   a) ¿Puede la función $\varphi(t) = t^2$, para todo $t \in \mathbb{R}$, resolver una EDO lineal de primer orden homogénea? ¿Y una EDO lineal de primer orden completa?

   b) ¿Pueden las funciones $\varphi(t) = e^t, \psi(t) = e^{-t}$, para todo $t \in \mathbb{R}$, resolver una misma EDO lineal de primer orden homogénea? ¿Y una EDO lineal de primer orden completa?

   c) Sea $x : [1, 2] \to \mathbb{R}$ una solución del problema de Cauchy
$$\begin{cases} x' = t + x^2 \\ x(1) = 2. \end{cases}$$
¿Puede ser $x(2) = 1$?

   d) ¿Existe alguna constante $\mu \in \mathbb{R}$ tal que la EDO $x' = (\mu + \cos^2 t)x$ tenga soluciones π-periódicas no triviales?

   e) Sea $a : \mathbb{R} \to \mathbb{R}$ una función continua y supongamos que existe $t_0 \in \mathbb{R}$ tal que $a(t) \leq 1/t^2$ para todo $t > t_0$. Probar que cualquier solución de la EDO $x' = a(t)x$ cumple $\lim_{t\to\infty} x(t) < \infty$. ¿Es cierto si se impone $a(t) \leq 1/t$?

5. Dadas f, g funciones derivables, en general es falso que se cumple la identidad
$$(f(t)g(t))' = f'(t)g'(t).$$
Fijando $f(t) = e^{t^3+2t}$, determinar las funciones $g(t)$ que satisfacen tal identidad.

6. Resolver las siguientes EDOs usando el método que convenga en cada caso.

   a) $3tx' - 2x = t^3/x^2$

   b) $x' = e^t x^7 + 2x$

   c) $x' + x/t = \log t/tx^2$

   d) $x' = x^2 - tx + 1$

   e) $x' = 1/t^4 - x^2$, con $x_p = 1/t - 1/t^2$

   f) $x' = -1/t^2 - x/t + x^2$

g) $x' = x/t + t^3x^2 - t^5$

h) $x' = 1 - 1/2tx + 1/2x^2$

i) $2tx + 3x + (t^2 + 3t)x' = 0$

7. Encontrar las soluciones de las siguientes EDOs buscando (si procede) factores integrantes de la forma que se indica.

   a) $\sin(tx) + tx\cos(tx) + t^2\cos(tx)x' = 0$.

   b) $3tx^2 - 4x + (3t - 4t^2x)x' = 0$, con $\mu(t,x) = t^n x^m$.

   c) $(t+1)^2 + (1+t^2)x' = 0$, con $\mu(t,x) = \mu(t+x)$.

   d) $t + x^2 + 2(x^2 + x + t - 1)x' = 0$, con $\mu(t,x) = \mu(e^{at+bx})$.

   e) $2txx' = t^2 + x^2 + 1$, con $\mu(t,x) = \mu(t^2 - x^2)$.

   f) $x^2 + x - 2tx + (t^2 + t - 2tx)x' = 0$, con $\mu(t,x) = \mu(t+x)$.

   g) $3x^2 - t + (2x^3 - 6tx)x' = 0$, con $\mu(t,x) = \mu(t+x^2)$.

8. Dada $h : (a,b) \to \mathbb{R}$ una función continua, con $-\infty \leq a < b \leq \infty$, se considera la EDO

$$x' = h\left(\frac{x}{t}\right), \qquad t \neq 0.$$

   a) Probar que

$$x' = \frac{M(t,x)}{N(t,x)},$$

   con M, N homogéneas del mismo grado, son un caso particular de esta EDO.

   b) Dado $(t,x) \in \mathbb{R}^2$, obtener la relación entre t y x para que $f(t,x) = h(x/t)$ esté bien definida y sea continua. Determinar el dominio de la EDO $x' = h(x/t)$.

   c) Probar que el cambio $u = x/t$ con $t \neq 0$ transforma esta EDO en una de variables separadas.

   d) Probar que si $-\infty < a < b < \infty$ y $h : [a,b] \to \mathbb{R}$ no tiene puntos fijos, entonces el problema de Cauchy

$$\begin{cases} x' = h\left(\frac{x}{t}\right), \\ x(t_0) = x_0, \end{cases}$$

   tiene existencia y unicidad.

   e) Se considera el sistema de ecuaciones

$$\begin{cases} at + bx + c = 0, \\ At + Bx + C = 0. \end{cases}$$

   Determinar la existencia y unicidad de soluciones en función de a, b, A, B.

   f) Si existe una única solución (t_0, x_0), probar que el cambio de variable

$$\begin{cases} t = s + t_0 \\ y(s) = x(s + t_0) - x_0 \end{cases}$$

   transforma la EDO

$$x' = h\left(\frac{at + bx + c}{At + Bx + C}\right), \qquad a, b, c, A, B, C \in \mathbb{R}$$

   en una homogénea.

g) Si no existe solución, probar que el cambio de variable

$$y(t) = at + bx(t),$$

transforma la EDO

$$x' = h\left(\frac{at+bx+c}{At+Bx+C}\right), \qquad a,b,c,A,B,C \in \mathbb{R}$$

en una de variables separadas.

h) Aplicar lo anterior para resolver la EDO

$$x' = \frac{3t+x-2}{1-x}.$$

9. Sean $t_0, x_0 \in \mathbb{R}$ con $t_0 \neq x_0$.

a) Resolver el problema de Cauchy

$$\begin{cases} x' = \dfrac{t+x}{t-x}, \\ x(t_0) = x_0. \end{cases}$$

¿Cuándo se puede garantizar la unicidad de soluciones?

b) Realizar un esbozo de la solución con condición inicial $x(0) = 1$. ¿Se puede extender la solución de forma C^1 en $t = 0$?

10. Encontrar un factor integrante de la forma $\mu(t^2 + x^2)$ para la EDO

$$x + t + (x-t)x' = 0,$$

y resolverla. Comparar con la solución obtenida en el ejercicio anterior. ¿Existe algún tipo de contradicción?

11. Sea x solución del problema de Cauchy

$$\begin{cases} x' = e^{t(x+1)} - \cos t, \\ x(0) = 0. \end{cases}$$

Probar que x tiene un mínimo relativo en $t = 0$.

12. Probar que el problema de Cauchy

$$\begin{cases} tx' = tx^2 + x, \\ x(0) = 0. \end{cases}$$

tiene infinitas soluciones distintas definidas en $\mathbb{R}$.

13. Dadas $f : I \subset \mathbb{R} \to \mathbb{R}$ y $g : J \subset \mathbb{R} \to \mathbb{R}$ funciones continuas, supongamos que $|g(x)| \leq |x|$ en un entorno de $x = 0$. Probar que dado $t_0 \in I$, la única solución del problema de Cauchy

$$\begin{cases} x' = f(t)g(x) \\ x(t_0) = 0, \end{cases}$$

es $x(t) = 0$ para todo $t \in I$.

14. Consideremos la EDO $M(t,x)+N(t,x)x'=0$.

    a) Probar que si existe un factor integrante $\mu(t,x)$ tal que $\mu(t,x)^n$ es también factor integrante, entonces la EDO original es exacta.

    b) Dar una relación entre M, N y μ de tal forma que si la EDO original es exacta, μ sea factor integrante.

    c) Probar que si μ, ξ son factores integrantes, entonces $\mu+\xi$ es también factor integrante.

    d) Probar que si μ, ξ son factores integrantes, entonces $H(t,x)=\frac{\mu(t,x)}{\xi(t,x)}$ es constante a lo largo de las soluciones de la EDO.

15. Probar que no existe ninguna función $x: I\subset\mathbb{R}\to\mathbb{R}$ de clase C^1 tal que su gráfica $(t,x(t))$ pase por el origen y las rectas normales a cada uno de los puntos de la gráfica también pasen por el origen.

16. Sea $\alpha:[a,b]\to\mathbb{R}^2$ una curva C^1.

    a) Probar que $\varphi:[a,b]\to[a,b],\ \varphi(s)=-s+a+b$ es un difeomorfismo.

    b) Probar que
    $$\beta:[a,b]\to\mathbb{R}, \qquad \beta(s)=\alpha(\varphi(s))$$
    es una curva C^1 cuya imagen $\beta([a,b])$ (también llamada *traza*) coincide con la traza de α. Intuitivamente, la curva β recorre los mismos puntos en $\mathbb{R}^2$ que la curva α pero en *sentido contrario*, y se llama *reparametrización de α recorrida hacia atrás*.

17. *Integración curvilínea.* Sea $\Omega\subset\mathbb{R}^2$ abierto y conexo y $F\colon\Omega\to\mathbb{R}^2$, $F(t,x)=(M(t,x),N(t,x))$, un campo vectorial. Dada $\gamma:[a,b]\to\mathbb{R}$ una curva con coordenadas $\gamma(s)=(t(s),x(s))$ y tal que $\gamma([a,b])\subset\Omega$, se define
    $$\int_\gamma F=\int_a^b\langle F(\gamma(s)),\gamma'(s)\rangle.$$

    a) Probar que si α es una reparametrización de γ *recorrida hacia atrás*, se tiene
    $$\int_\gamma F=-\int_\alpha F.$$

    b) Probar que si existe $H\in C^1(\Omega)$ tal que $\partial H/\partial t=M,\ \partial H/\partial x=N$, entonces
    $$\int_\gamma F=H(\gamma(b))-H(\gamma(a)).$$

    c) Concluir que si existe una curva cerrada γ_0 en Ω (esto es, $\gamma(a)=\gamma(b)$) tal que
    $$\int_{\gamma_0}F\neq 0,$$
    entonces no puede existir $H\in C^1(\Omega)$ tal que $\partial H/\partial t=M,\ \partial H/\partial x=N$.

18. Sea $\Omega \subset \mathbb{R}^2$ un dominio *estrellado*, esto es existe $p_0 = (t_0, x_0) \in \Omega$ tal que para todo $p = (t, x) \in \Omega$ el segmento que une p_0 con p está incluido en Ω. Fijado $p_0 = (t_0, x_0)$, denotemos por $\overline{p_0\, p}$ a tal segmento. Sean $M, N \in C^1(\Omega)$ tales que $\partial M/\partial x = \partial N/\partial t$. Probar que la función

$$H(p) = \int_{\overline{p_0\, p}} F(t,x), \qquad F(t,x) = (M(t,x), N(t,x)), \quad p \in (t,x) \in \Omega,$$

cumple el criterio de exactitud

$$\frac{\partial H}{\partial t} = M, \qquad \frac{\partial H}{\partial x} = N.$$

19. Sea Ω abierto y conexo y $F\colon \Omega \subset \mathbb{R}^2 \to \mathbb{R}^2$, $F(t,x) = (M(t,x), N(t,x))$, un campo vectorial. Supongamos que existe $p_0 \in \Omega$ con la siguiente propiedad: toda curva cerrada γ conteniendo a p_0 cumple

$$\int_\gamma F = 0.$$

a) Probar que la función

$$H(p) = \int_{\gamma_p} F, \qquad p = (t,x) \in \Omega,$$

donde γ_p es *cualquier* curva uniendo p_0 con p está bien definida. En otras palabras, si γ_p^1, γ_p^2 son dos curvas uniendo p_0 y p, entonces

$$\int_{\gamma_p^1} F = \int_{\gamma_p^2} F.$$

b) Si $\partial M/\partial x = \partial N/\partial t$, probar que H cumple el criterio de exactitud

$$\frac{\partial H}{\partial t} = M, \qquad \frac{\partial H}{\partial x} = N.$$

Capítulo 3

Ecuaciones lineales

En el tema pasado nos centramos en el estudio de EDOs que se podían integrar para obtener una expresión, en general implícita, de las soluciones. Incluso sin poder obtener una fórmula explícita en términos de funciones elementales, se pueden deducir propiedades cualitativas de las soluciones a través de la derivación implícita y el uso de técnicas analíticas elementales.

Aunque en el tema pasado nos dedicamos exclusivamente a EDOs de orden 1, esto es, la derivada de mayor orden es 1, este está lejos de ser el escenario real. Ya en la que posiblemente es la obra científica más importante de la historia, el famoso *Philosophiæ naturalis principia mathematica* de Isaac Newton (1686-1687), se trata, entre otras muchas, la EDO de segundo orden

$$x'' + k^2 x = 0, \tag{3.0.1}$$

cuando se estudia el movimiento de un punto atraído por otro con una fuerza proporcional a la distancia. Esta ecuación aparece igualmente en la ley de Hooke del movimiento de un muelle o resorte elástico, bajo la hipótesis de rozamiento nulo. Newton también observó que dada $f : I \subset \mathbb{R} \to \mathbb{R}$ continua, las soluciones de la ecuación

$$x^{n)} = f(t),$$

dependen de n parámetros arbitrarios. Otra EDO similar a (3.0.1) es

$$x'' + k^2 \sin x = 0, \tag{3.0.2}$$

que describe la amplitud angular respecto del reposo de un péndulo (de nuevo, sin rozamiento). Para amplitudes pequeñas, se suele realizar la aproximación $\sin x \approx x$ y volvemos a la EDO (3.0.1). Llegados a este punto, es importante resaltar una diferencia fundamental entre las ecuaciones (3.0.1) y (3.0.2); las soluciones de la primera forman un espacio vectorial real, al contrario que las segundas.

Estos pocos ejemplos ponen de manifiesto el interés natural de estudiar EDOs de orden superior. Una forma de abordar este estudio, vista en el tema 1, es reducir una EDO escalar de orden n a un sistema equivalente de orden 1 con n ecuaciones. En efecto, si tenemos la EDO en forma normal

$$x^{n)} = F(t, x, x', ..., x^{n-1)}), \qquad F : \Omega \subset \mathbb{R}^{n+1} \to \mathbb{R}, \tag{3.0.3}$$

introduciendo las funciones $y_k = x^{k-1)}$ para $k = 1, ..., n$, se tiene que la EDO anterior es equivalente al sistema

$$\begin{cases} y_1' = y_2, \\ \qquad \vdots \\ y_{n-1}' = y_n, \\ y_n' = F(t, y_0, y_1, ..., y_n). \end{cases}$$

Este hecho motiva el estudio de los sistemas de orden 1, ya que toda solución $y = (y_1, ..., y_n)$ cumple que la función $x = y_1$ es solución de (3.0.3). Con el mismo esfuerzo se puede estudiar un sistema de EDOs de la forma

$$\begin{cases} x_1' = f_1(t,x), \\ \quad\vdots \\ x_n' = f_n(t,x), \end{cases} \qquad x = (x_1, ..., x_n), \tag{3.0.4}$$

que son equivalentes al estudio de EDOs vectoriales de primer orden de la forma $x' = f(t,x)$ con $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n,\ f = (f_1, ..., f_n)$. En el caso particular pero de suma importancia en que la función $f(t,x)$ es lineal en x, esto es $f(t,x) = A(t)x + b(t)$, con $A : I \to \mathcal{M}_n(\mathbb{R})$ y $b : I \to \mathbb{R}^n$ continuas, el sistema (3.0.4) se escribe matricialmente

$$\begin{cases} x_1' = a_{11}(t)x_1 + \cdots + a_{1n}(t)x_n + b_1(t), \\ \qquad\qquad\qquad\vdots \\ x_n' = a_{n1}(t)x_1 + \cdots + a_{nn}(t)x_n + b_n(t), \end{cases} \qquad A(t) = (a_{ij}(t))_{i,j}, \quad b(t) = (b_1(t), ..., b_n(t))^T.$$

Este será el objeto de estudio principal en el presente tema.

3.1. Ecuación lineal de orden uno n-dimensional

En esta sección trabajaremos con sistemas lineales n-dimensionales expresados de la forma

$$x' = A(t)x + b(t),$$

donde $A(t)$ es una matriz $n \times n$, $b(t)$ es un vector (columna) de n componentes y las aplicaciones $t \mapsto A(t)$ y $t \mapsto b(t)$, con $t \in I \subset \mathbb{R}$, son continuas. Esto último equivale a que cada una de las componentes $a_{ij}(t)$ y $b_i(t)$ de $A(t)$ y $b(t)$, respectivamente, sean funciones continuas. Dado un vector $x_0 \in \mathbb{R}^n$, el problema de Cauchy con condición inicial $x_0 \in \mathbb{R}^n$ en el instante t_0 no es más que el sistema

$$\begin{cases} x' = A(t)x + b(t), \\ x(t_0) = x_0. \end{cases} \tag{3.1.1}$$

Antes de enunciar el resultado fundamental de existencia y unicidad de este tema, necesitamos recordar unas definiciones que serán de utilidad.

Definición 3.1.1. Dado un intervalo $I \subset \mathbb{R}$ y una función $f : I \to \mathbb{R}^n$, definimos

$$||f||_\infty = \sup_{t \in I} ||f(t)||,$$

donde la segunda norma $|| \cdot ||$ es *cualquier* norma en $\mathbb{R}^n$.

Definición 3.1.2. Una sucesión de funciones $f_n : I \to \mathbb{R}^n$ se dice que *converge uniformemente a* $f : I \to \mathbb{R}^n$ si se cumple

$$||f_n - f||_\infty \to 0.$$

En otras palabras, para cada $\epsilon > 0$ existe $n_0 \in \mathbb{N}$ sólo dependiendo de ϵ tal que para todo $n \geq n_0$ y para todo $t \in I$ se tiene

$$||f_n(t) - f(t)|| < \epsilon.$$

La convergencia uniforme de funciones tiene algunas propiedades muy interesantes, enunciadas a continuación.

1. Convergencia puntual no implica uniforme, incluso cuando la sucesión está definida en un compacto. Por ejemplo, $f_n(t) = t^n,\ t \in [0,1]$ converge puntualmente a la función $f(t) = 0$ si $t \in [0,1)$ y $f(1) = 1$. Sin embargo, esta convergencia no puede ser jamás uniforme porque $f_n(t)$ es tan cercano a 1 como queramos si t es suficientemente próximo a $t = 1$.

2. Si $f_n : I \to \mathbb{R}^n$ es continua para cada n y $f_n \underset{c.u.}{\to} f$, entonces f es continua en I. Observemos que no hay hipótesis adicionales sobre la topología de I.

3. Si $[a,b] \subset I$ y $f_n \underset{c.u.}{\to} f$ en I, entonces
$$\int_a^b f_n \to \int_a^b f.$$
Esto se puede probar, por ejemplo, usando el Teorema de la convergencia dominada.

Ojo! No va tan bien con las derivadas. Se puede probar que
$$f_n(t) = \frac{\sin(nt)}{n} \underset{c.u.}{\to} 0,$$
pero la sucesión de las derivadas, $f_n'(t)$, no puede converger ni siquiera puntualmente a una función ya que $f_n'(\pi) = \cos(n\pi)$ no tiene límite.

Definición 3.1.3. Dada una matriz $A \in \mathcal{M}_n(\mathbb{R})$, se define la norma del supremo como
$$||A|| = \sup\left\{\frac{||Av||}{||v||} : v \in \mathbb{R}^n,\ v \neq 0\right\} = \text{máx}\{||Av|| : v \in \mathbb{R}^n,\ ||v|| = 1\}.$$

3.1.1. Existencia y unicidad del problema de Cauchy

El resultado principal de esta sección es probar el resultado de existencia y unicidad del problema de Cauchy (3.1.1).

Teorema 3.1.4. *El problema de Cauchy* (3.1.1) *tiene una única solución definida en todo* I.

Demostración. Comenzaremos observando que x es una solución del problema de Cauchy (3.1.1) si y sólo si se cumple
$$x(t) = x_0 + \int_{t_0}^t A(s)x(s) + b(s)\,ds, \tag{3.1.2}$$
donde la integral se debe entender componente a componente,
$$\int x(t)\,dt = \left(\int x_1(t)\,dt, \int x_2(t)\,dt, \ldots, \int x_n(t)\,dt\right).$$
Análogamente al caso $n = 1$, la integral (3.1.2) es conocida como *integral de Volterra*. Procedemos a definir por recurrencia la sucesión de funciones
$$x_0(t) = x_0,$$
$$x_{k+1}(t) = x_0 + \int_{t_0}^t A(s)x_k(s) + b(s)\,ds.$$

La sucesión $\{x_k\}$ se conoce en la literatura como *iterantes de Picard*. La idea detrás de esta definición es comprobar que la sucesión x_k converge a una función que será de hecho solución de la EDO. Fijemos $[a,b] \subset I$ tal que $t_0 \in (a,b)$ y sean

$$||A||_\infty = \max_{t\in[a,b]} ||A(t)||, \qquad ||b||_\infty = \max_{t\in[a,b]} ||b(t)||$$

$$c = \max_{t\in[a,b]} \int_{t_0}^{t} ||A(s)||\,||x_0|| + ||b(s)||\,ds \geq \max_{t\in[a,b]} ||x_1(t) - x_0(t)||,$$

donde los máximos anteriores están bien definidos por ser $A(t), x_k(t)$ continuas y $[a,b]$ compacto. Entonces, para $t \in [a,b]$ se tiene

$$||x_2(t) - x_1(t)|| = ||\int_{t_0}^{t} A(s)(x_1(s) - x_0(s))\,ds|| \leq ||A||_\infty c|t - t_0|.$$

Análogamente,

$$||x_3(t) - x_2(t)|| = ||\int_{t_0}^{t} A(s)(x_2(s) - x_1(s))\,ds|| \leq \frac{||A||_\infty^2 c}{2}|t - t_0|^2.$$

Por inducción,

$$||x_{k+1}(t) - x_k(t)|| \leq \frac{||A||_\infty^k c}{k!}|t - t_0|^k,$$

y en cualquier caso

$$\max_{t\in[a,b]} ||x_{k+1}(t) - x_k(t)|| \leq \frac{||A||_\infty^k c}{k!}|t - t_0|^k.$$

Consideremos la sucesión de funciones $a_0(t) = x_0$ y $a_k(t) = x_k(t) - x_{k-1}(t)$ para $k > 0$. Como la serie formada por los términos $\frac{||A||_\infty^k c}{k!}|t - t_0|^k$ converge, por ejemplo aplicando el criterio del cociente, la serie funcional $\sum_{k=0}^{\infty} a_k(t)$ converge absoluta y uniformemente en $[a,b]$, sin más que aplicar el criterio de Weierstrass para series de funciones vectoriales. Ahora bien, una serie funcional converge uniformemente si la sucesión de sus sumas parciales

$$S_k(t) = \sum_{i=0}^{k} a_i(t),$$

converge uniformemente. Como las sumas parciales cumplen $S_k(t) = x_k(t)$, la propia sucesión $\{x_k\}$ converge de manera uniforme en $[a,b]$ a una cierta función $x : [a,b] \to \mathbb{R}$. Es más, por ser $x(t)$ límite uniforme de la sucesión $\{x_k\}$ y siendo cada x_k continua, tenemos que $x(t)$ es continua para cada $t \in [a,b]$. Como el intervalo $[a,b]$ se cogió en I de forma arbitraria, sin más que suponer $t_0 \in [a,b]$, tenemos que $x(t)$ es continua para todo $t \in I$. Más aún, el buen comportamiento de la convergencia uniforme respecto de la integral nos permite deducir lo siguiente

$$\begin{aligned} x(t) &= \lim_{k\to\infty} x_{k+1}(t) \\ &= \lim_{k\to\infty} \left(x_0 + \int_{t_0}^{t} A(s)x_k(s) + b(s)\,ds \right) ds \\ &= x_0 + \int_{t_0}^{t} \lim_{k\to\infty} \Big(A(s)x_k(s) + b(s) \Big)\,ds \\ &= x_0 + \int_{t_0}^{t} A(s)x(s) + b(s)\,ds, \end{aligned}$$

y esto para todo $t \in I$. En consecuencia, $x(t)$ es de clase C^1 por el Teorema Fundamental del Cálculo y es solución de la EDO (3.1.1). Como además $x(t_0) = x_0$ hemos probado que el problema de Cauchy tiene una solución en $[a, b]$.

Ahora probamos que $x(t)$ es la única solución en $[a, b]$. Supongamos que existe otra solución $y(t)$ del problema de Cauchy (3.1.1) en $[a, b]$ y demostraremos que $x(t) = y(t)$ para todo $t \in [a, b]$. Denotemos

$$c = \max_{t\in[a,b]} ||x(t) - y(t)||.$$

Como $y(t)$ satisface la ecuación integral de Volterra (3.1.2) por ser solución, se tiene la siguiente estimación para cada $t \in [a, b]$,

$$||x(t) - y(t)|| \leq ||A||_\infty \int_{t_0}^{t} ||x(s) - y(s)||\, ds, \tag{3.1.3}$$

y en particular

$$||x(t) - y(t)|| \leq ||A||_\infty c|t - t_0|.$$

Si se sustituye la estimación anterior en (3.1.3), se tiene

$$||x(t) - y(t)|| \leq K \int_{t_0}^{t} ||x(s) - y(s)||\, ds \leq ||A||_\infty^2 c \int_{t_0}^{t} |s - t_0|\, ds = \frac{||A||_\infty^2 c}{2} |t - t_0|^2.$$

Repitiendo el proceso n veces concluimos

$$||x(t) - y(t)|| \leq \frac{c||A||_\infty^n}{n!} |t - t_0|^n.$$

Como la sucesión $\frac{c||A||_\infty^n}{n!}|t - t_0|^n$ tiende a cero cuando $n \to \infty$ independientemente de t, concluimos que $x(t) = y(t)$ para cada $t \in [a, b]$.

Ojo! Hemos probado que existencia y unicidad de solución en cada intervalo de la forma $[a, b] \subset I$. Para extender la existencia y unicidad a todo el intervalo I, consideramos una exhaución $I_n = [a_n, b_n],\ n \geq 1$, de I por intervalos compactos con $t_0 \in I_1$; esto es, $I_n \subset I_{n+1}$ y $\bigcup_{n=1}^{\infty} I_n = I$.

Probemos primero la unicidad en I. Sean $x, y : I \to \mathbb{R}^n$ dos soluciones de (3.1.1) consideremos $t_* \in I$ arbitrario. Entonces existe n_0 tal que $t_* \in I_{n_0}$. Como la restricción de x e y a I_{n_0} son soluciones de (3.1.1), se tiene $x(t_*) = y(t_*)$. Ya que t_* era arbitrario, se tiene $x(t) = y(t)$ para cada $t \in I$.

Para finalizar, probaremos que existe una solución en I. Si aplicamos el resultado de existencia sobre cada I_n se tendrá una solución x_n del problema de Cauchy (3.1.1). Además, por unicidad si $n < m$ se tiene $x_n(t) = x_m(t)$ para cada $t \in I_n$. En concreto podemos definir la función

$$x : I \to \mathbb{R}^n, \qquad x(t) = x_n(t), \text{ si } t \in I_n.$$

Esta función está bien definida, de nuevo por unicidad. Dado $t \in I$ podemos encontrar n grande tal que $t \in I_n$, por tanto $x = x_n$ en I_n y en particular en un entorno de t. Entonces, x es derivable en t y se tiene

$$x'(t) = x_n'(t) = A(t)x_n(t) + b(t) = A(t)x(t) + b(t).$$

Deducimos que $x \in C^1(I)$ y cumple (3.1.2), es decir x es solución del problema de Cauchy (3.1.1) en I. $\square$

3.1.2. Ecuación homogénea. Estructura de espacio vectorial y matrices fundamentales

Tal y como hicimos en el tema 2, comenzaremos estudiando el sistema lineal homogéneo

$$x' = A(t)x, \tag{3.1.4}$$

donde $t \in I \subset \mathbb{R}$ y $A : I \to \mathcal{M}_n(\mathbb{R})$ es continua. El primer resultado de esta sección es probar que las soluciones de (3.1.4) forman un espacio vectorial, lo cual no debería ni sorprendernos ni presentarnos dificultad en su prueba. En cuanto a la dimensión, sí que tendremos que hacer un desarrollo algo más elaborado. Antes de enunciar el resultado debemos dar una definición.

Definición 3.1.5. Se dice que $x_1, ..., x_n : I \to \mathbb{R}^k$ son linealmente independientes en I si dados $\lambda_1, ..., \lambda_n \in \mathbb{R}$ tales que

$$\lambda_1 x_1(t) + \cdots + \lambda_n x_n(t) = 0$$

para todo $t \in I$, se tiene $\lambda_1 = ... = \lambda_n = 0$.

Debemos remarcar que en la definición anterior los escalares $\lambda_1, ..., \lambda_n$ son fijos; no varían cuando t se mueve en I. Ahora sí podemos enunciar el resultado deseado.

Teorema 3.1.6. *El conjunto de las soluciones de* (3.1.4)*, S_h, es un espacio vectorial de dimensión n.*

Demostración. Que el conjunto de las soluciones tiene estructura de espacio vectorial es trivial y se deja como ejercicio. Para deducir la dimensión de este espacio vectorial, fijemos $v_1, ...v_n$ una base de $\mathbb{R}^n$ y $t_0 \in I$. Sean $x_1, ..., x_n$ las únicas soluciones de (3.1.4), en virtud del Teorema 3.1.4, con condiciones iniciales $x_k(t_0) = v_0$ para cada $k = 1, ..., n$. Si probamos que las funciones $x_1, ..., x_n$ forman una base de soluciones, habremos concluido.

Para empezar, sean $\lambda_1, ..., \lambda_n \in \mathbb{R}$ tales que

$$\lambda_1 x_1 + \cdots \lambda_n x_n = 0.$$

Sustituyendo en t_0 se tiene

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0,$$

y como $v_1, ..., v_n$ forman una base de $\mathbb{R}^n$ concluimos que $\lambda_1 = \cdots = \lambda_n = 0$, por lo que las funciones $x_1, ..., x_n$ son linealmente independientes. Sea ahora z una solución de (3.1.4) y denotemos $z(t_0) = v$. Existen $\lambda_1, ..., \lambda_n \in \mathbb{R}$ tales que

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n.$$

Definimos $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$, que cumple $x(t_0) = v$. Por la unicidad del problema de Cauchy para la condición inicial $x(t_0) = v$ concluimos que $z = x$ y hemos probado que $x_1, ..., x_n$ forman un sistema generador. □

Definición 3.1.7. Un sistema fundamental de (3.1.4) es cualquier base de su espacio de soluciones. En tal caso, si $x_1, ..., x_n$ forman un sistema fundamental de (3.1.4) y definimos la matriz

$$\Phi(t) = \begin{pmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \end{pmatrix}$$

cuyas columnas son las funciones $x_1, ..., x_n$, diremos que Φ es una *matriz fundamental* de (3.1.4).

Antes de avanzar, nos parece interesante hacer un par de aclaraciones que serán útiles en lo sucesivo. Dado un sistema fundamental $\{x_1, ..., x_n\}$, denotaremos sus funciones coordenadas con el siguiente convenio,

$$x_1 = \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix}, \qquad \cdots \qquad x_k = \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{nk} \end{pmatrix}, \qquad \cdots \qquad x_n = \begin{pmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{pmatrix},$$

y así la matriz fundamental $\Phi(t)$ en coordenadas es

$$\Phi(t) = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nn} \end{pmatrix}.$$

Sea ahora $\lambda = (\lambda_1, ..., \lambda_n)^T$. Dado que el producto de una matriz por un vector es hacer la combinación lineal de las columnas de la matriz según las coordenadas del vector, el producto $\Phi(t)\lambda$ no es más que el vector

$$\begin{pmatrix} x_{11}\lambda_1 + x_{12}\lambda_2 + \cdots + x_{1n}\lambda_n \\ \vdots \\ x_{k1}\lambda_1 + x_{k2}\lambda_2 + \cdots + x_{kn}\lambda_n \\ \vdots \\ x_{n1}\lambda_1 + x_{n2}\lambda_2 + \cdots + x_{nn}\lambda_n \end{pmatrix} = \lambda_1 \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix} + \cdots + \lambda_n \begin{pmatrix} x_{1n} \\ x_{2n} \\ \vdots \\ x_{nn} \end{pmatrix} = \lambda_1 x_1 + \cdots + \lambda_n x_n.$$

Si expresamos $A(t)$ en coordenadas

$$A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1n}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nn}(t) \end{pmatrix},$$

el hecho de que cada $x_k = (x_{1k}, ..., x_{nk})^T$ sea solución del sistema homogéneo (3.1.4), esto es $x_k' = A(t)x_k$, se expresa como el sistema de ecuaciones diferenciales

$$\begin{cases} x_{1k}' = a_{11}x_{1k} + \cdots + a_{1n}x_{nk} = \sum_{i=1}^n a_{1i}x_{ik}, \\ \vdots \\ x_{jk}' = a_{j1}x_{1k} + \cdots + a_{jn}x_{nk} = \sum_{i=1}^n a_{ji}x_{ik}, \\ \vdots \\ x_{nk}' = a_{n1}x_{1k} + \cdots + a_{nn}x_{nk} = \sum_{i=1}^n a_{ni}x_{ik}. \end{cases}$$

Esto se resume matricialmente de la forma

$$\Phi'(t) = A(t)\Phi(t).$$

Sabemos que si $\{x_1, ..., x_k\}$ es sistema fundamental de (3.1.4), entonces por definición estas funciones son linealmente independientes. Es por tanto esperable que toda matriz fundamental tenga determinante no nulo, en virtud de la independencia lineal de sus columnas en todo instante t, pero sería también deseable que este hecho caracterizase a las matrices fundamentales. En otras palabras, que dada una matriz cuyas columnas sean solución de (3.1.4) y además tenga determinante distinto de cero, tal matriz nos proporcione automáticamente un sistema fundamental de (3.1.4). El siguiente resultado da respuesta a estas cuestiones.

Proposición 3.1.8. *Sea $\Phi(t)$ una matriz cuyas columnas son soluciones de* (3.1.4). *Equivalen:*

1. $\Phi(t)$ *es una matriz fundamental.*
2. *Para todo $t \in I$ se tiene* $\det \Phi(t) \neq 0$*;*
3. *Existe $t_0 \in I$ tal que* $\det \Phi(t_0) \neq 0$.

Demostración. Denotemos por $x_1, ..., x_n$ a las columnas de $\Phi(t)$.

Primero probemos $1 \Rightarrow 2$. Sean $x_1, ..., x_n$ las columnas de $\Phi(t)$, las cuales son una base del espacio vectorial de soluciones de (3.1.4) y en particular linealmente independientes. Por reducción al absurdo, supongamos que existe $t_0 \in I$ tal que $\det \Phi(t_0) = 0$. En tal caso, existen $\lambda_1, ..., \lambda_n \in \mathbb{R}$, no todos nulos, tales que

$$\lambda_1 x_1(t_0) + \cdots + \lambda_n x_n(t_0) = 0.$$

Definimos la función $x = \lambda_1 x_1 + \cdots + \lambda_n x_n$. Obviamente x es solución de (3.1.4) y cumple $x(t_0) = 0$. Por la unicidad del problema de Cauchy se tiene que x es idénticamente nula, lo que contradice la independencia lineal de las funciones $x_1, ..., x_n$.

Que $2 \Rightarrow 3$ es trivial. Finalmente para probar que $3 \Rightarrow 1$ debemos demostrar que las columnas de $\Phi(t)$ son una base del espacio de soluciones. Si definimos

$$v_1 = x_1(t_0), \quad v_2 = x_2(t_0), \quad \cdots \quad v_n = x_n(t_0),$$

entonces $\{v_1, ..., v_n\}$ es base de $\mathbb{R}^n$ por ser $\det \Phi(t_0) \neq 0$. Que $x_1, ..., x_n$ son linealmente independientes se prueba tomando una combinación lineal $\lambda_1 x_1 + \cdots + \lambda_n x_n = 0$, sustituyendo en $t = t_0$ y usando que $v_1, ..., v_n$ es base de $\mathbb{R}^n$. Sea ahora z una solución de (3.1.4) y sea $v = z(t_0)$. Existen $\lambda_1, ..., \lambda_n$ tales que $v = \lambda_1 v_1 + \cdots \lambda_n v_n$. Por unicidad del problema de Cauchy en $t = t_0$,

$$z = \lambda_1 x_1 + \cdots + \lambda_n x_n,$$

lo que prueba que $x_1, ..., x_n$ generan el espacio de soluciones y por tanto $\Phi(t)$ es matriz fundamental. □

Ejemplo 3.1. ¿Pueden ser las funciones

$$x_1(t) = \begin{pmatrix} 1 \\ t \end{pmatrix}, \qquad x_2(t) = \begin{pmatrix} \cos t \\ 0 \end{pmatrix},$$

soluciones linealmente independientes de algún sistema con coeficientes continuos en $\mathbb{R}$? ¿Y en un intervalo más pequeño? En caso afirmativo, deducir el sistema del que x_1, x_2 son soluciones.

Si lo fueran,

$$\Phi(t) = \begin{pmatrix} 1 & \cos t \\ t & 0 \end{pmatrix}$$

sería una matriz fundamental, pero su determinante es $\det\varphi(t) = -t\cos t$, el cual se anula en $t = 0$ y para todo $t = \frac{\pi}{2} + k\pi$, con $k \in \mathbb{Z}$. En consecuencia, el sistema diferencial no puede estar definido en todo $\mathbb{R}$. Si nos restringimos a intervalos que no contengan estos puntos, entonces pueden ser solución de un sistema diferencial, el cual viene caracterizado por

$$\Phi'(t) = A(t)\Phi(t).$$

Ahora bastaría despejar $A(t)$, obteniendo

$$A(t) = \Phi'(t)\Phi^{-1}(t) = \begin{pmatrix} -\tan t & \dfrac{\tan t}{t} \\ & \\ 0 & \dfrac{1}{t} \end{pmatrix}.$$

Nótese que si $\Phi(t)$ es una matriz fundamental, la solución general de (3.1.4) viene dada por

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n = \Phi(t)\lambda, \qquad \lambda = (\lambda_1, ..., \lambda_n)^T \in \mathbb{R}^n.$$

Dicho de otra forma, para generar el espacio de soluciones del sistema homogéneo (3.1.4) basta multiplicar cualquier matriz fundamental $\Phi(t)$ por vectores de $\mathbb{R}^n$. Esto generaliza de cierta forma el caso de la EDO lineal homogénea de orden 1; dada una solución, cualquier otra es proporcional. El siguiente resultado relaciona las matrices fundamentales y nos dice que, en esencia, todas son un cierto múltiplo matricial de una matriz fundamental dada.

Teorema 3.1.9. *Sea $\Phi(t)$ una matriz fundamental de* (3.1.4)*. Entonces, $\Psi(t)$ es otra matriz fundamental de* (3.1.4) *si y sólo si existe una matric $C \in \mathcal{M}_n(\mathbb{R})$ con $\det C \neq 0$ tal que $\Psi(t) = \Phi(t)C$.*

Demostración. Comenzando recordando un resultado básico de la teoría de espacios vectoriales. Si V es un espacio vectorial de dimensión n, tomamos una base $v_1, ...v_n$ de V, consideramos vectores $w_1, ..., w_n$ y los expresamos respecto de la base $v_1, ..., v_n$ como $w_i = \lambda_{1i}v_1 + \cdots + \lambda_{ni}v_n$, entonces $w_1, ..., w_n$ forman base de V si y sólo si

$$\det\begin{pmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \end{pmatrix} \neq 0.$$

Sea pues $C = (c_{ij})$ una matriz con $\det C \neq 0$ y definamos $\Psi(t) = \Phi(t)C$. Entonces si las columnas de $\Psi(t)$ son $z_1, ..., z_n$, es claro que $z_k = c_{1k}x_1 + \cdots + c_{nk}x_n$ y cada z_k es solución de (3.1.4). Como $\det C \neq 0$ las funciones $z_1, ..., z_n$ vuelven a formar una base del espacio de soluciones de (3.1.4) y hemos probado que $\Psi(t) = \Phi(t)C$ es matriz fundamental.

Recíprocamente, si $\Psi(t)$ es una matriz fundamental y sus columnas son $z_1, ..., z_n$ entonces para todo $k = 1, ..., n$ se tiene $z_k = c_{1k}x_1 + \cdots + c_{nk}x_n$ para ciertos $c_{jk},\ j = 1, ..., n$ (por ser $x_1, ..., x_n$ base). Matricialmente se escribe como $\Psi(t) = \Phi(t)C$, donde $C = (c_{ij})$. Ahora bien, como $\Psi(t)$ es matriz fundamental entonces $\det\Psi(t) \neq 0$ lo que necesariamente implica $\det C \neq 0$. □

Si $\Phi(t)$ es una matriz fundamental de (3.1.4) el Teorema 3.1.9 garantiza que $\Phi(t)\Phi^{-1}(t_0)$ es de nuevo matriz fundamental para cualquier $t_0 \in I$. De hecho, esta matriz fundamental es la identidad en el instante t_0. Tenemos pues el siguiente corolario que concierne al problema de Cauchy.

Corolario 3.1.10. *El problema de Cauchy*

$$\begin{cases} x' = A(t)x, \\ x(t_0) = x_0. \end{cases}$$

tiene como solución

$$x = \Phi(t)\Phi^{-1}(t_0)x_0.$$

La dificultad de resolver el problema de Cauchy radica en obtener n funciones independientes de su espacio de soluciones para formar su matriz fundamental asociada.

Ejemplo 3.2. Consideremos el problema de Cauchy

$$\begin{cases} x' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x, \\ \\ x(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{cases}$$

Si denotamos por $x = (x_1, x_2)$, entonces el sistema anterior es $x_1' = x_2,\ x_2' = -x_1$, lo cual es equivalente sin más que derivar la primera ecuación a

$$x_1'' + x_1 = 0.$$

Esta ecuación es lineal homogénea de segundo orden y por tanto su espacio de soluciones tiene dimensión 2. Como las funciones $\sin t, \cos t$ son linealmente independientes (¡pruébese!), se tiene que forman un sistema fundamental de soluciones. En consecuencia, usando que $x_2 = x_1'$ construimos la matriz fundamental

$$\Phi(t) = \begin{pmatrix} \sin t & \cos t \\ \cos t & -\sin t \end{pmatrix}.$$

Por tanto, dado $\lambda = (\lambda_1, \lambda_2)^T \in \mathbb{R}^2$, la solución de la ecuación homogénea es

$$x(t) = \Phi(t)\lambda = \lambda_1 \begin{pmatrix} \sin t \\ \cos t \end{pmatrix} + \lambda_2 \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}.$$

Para calcular la solución del problema de Cauchy, calculamos

$$\Phi^{-1}(1) = \begin{pmatrix} \sin 1 & \cos 1 \\ \cos 1 & -\sin 1 \end{pmatrix}.$$

La solución pues del problema de Cauchy es

$$x(t) = \Phi(t)\Phi^{-1}(1) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \cos(t-1) + \sin(t-1) \\ -\sin(t-1) + \cos(t-1) \end{pmatrix}.$$

3.1.3. Ecuación completa. Método de variación de las constantes

En esta sección nos dedicamos al estudio de las soluciones de la ecuación completa

$$x' = A(t)x + b(t), \tag{3.1.5}$$

donde ahora el vector $b(t) = (b_1(t), ..., b_n(t))^T$ no es idénticamente nulo. Al igual que en el caso de dimensión 1, las soluciones del sistema homogéneo (3.1.4) y del completo (3.1.5) guardan una estrecha relación. La demostración de este resultado es totalmente análoga al Teorema 2.3.2 dada en el Tema 2 y por eso se omite.

Teorema 3.1.11. *Sean S_h y S_c los conjuntos de soluciones de las EDOs homogénea* (3.1.4) *y completa* (3.1.5)*, respectivamente. Entonces, S_c tiene estrcutrua de espacio afín sobre el espacio vectorial S_h.*

De manera similar al caso $n = 1$, basta conocer una solución particular x_p del sistema completo para concluir que cualquier otra solución x es de la forma

$$x = x_p + \lambda_1 x_1 + \cdots + \lambda_n x_n,$$

con $\lambda_1, ..., \lambda_n \in \mathbb{R}$ y $x_1, ..., x_n$ una base del espacio de soluciones del sistema homogéneo. En términos de matrices fundamentales, si $\Phi(t)$ es matriz fundamental de (3.1.4) con columnas el sistema fundamental $x_1, ..., x_n$, las soluciones de (3.1.4) vienen dadas por $\Phi(t)\lambda$, con $\lambda = (\lambda_1, ..., \lambda_n)^T$, y por tanto las soluciones de (3.1.5) son de la forma

$$x = x_p + \Phi(t)\lambda, \qquad \lambda \in \mathbb{R}^n.$$

La obtención de una solución particular de (3.1.5) tiene su inspiración en la resolución de la ecuación lineal completa para $n = 1$, tal y como vimos en el pasado tema. Para tal fin, debemos conocer una matriz fundamental Φ de (3.1.4), cosa que como estamos incidiendo está lejos de ser una cuestión trivial (o incluso fácil...). Supongamos conocida una matriz fundamental y busquemos una solución particular x_p de (3.1.5). La forma de proceder está inspirada por el método de variación de las constantes introducido para el caso $n = 1$, y por tanto suponemos x_p de la forma

$$x_p = \Phi(t)\lambda(t),$$

donde ahora $\lambda(t)$ es un vector variable. Por una parte,

$$x_p' = \Phi'(t)\lambda(t) + \Phi(t)\lambda'(t) = A(t)\Phi(t)\lambda(t) + \Phi(t)\lambda'(t),$$

y si imponemos que x_p sea solución particular se tiene

$$x_p' = A(t)x_p + b(t) = A(t)\Phi(t)\lambda(t) + b(t).$$

Igualando ambos términos concluimos

$$\lambda(t) = \int \Phi^{-1}(t)b(t)\, dt.$$

Así pues,

$$x_p = \Phi(t) \int \Phi^{-1}(t)b(t)\, dt.$$

Por último, la solución general del sistema completo (3.1.5) viene dada por

$$x = \Phi(t)\lambda + \Phi(t) \int \Phi^{-1}(t)b(t)\, dt, \qquad \lambda \in \mathbb{R}^n.$$

En relación al problema de Cauchy, obtenemos el siguiente corolario.

Corolario 3.1.12. *El problema de Cauchy*

$$\begin{cases} x' = A(t)x + b(t), \\ x(t_0) = x_0, \end{cases}$$

tiene como solución

$$x = \Phi(t)\Phi^{-1}(t_0)x_0 + \Phi(t) \int_{t_0}^{t} \Phi^{-1}(s)b(s)\, ds.$$

Ejemplo 3.3. Resolvamos el problema de Cauchy

$$\begin{cases} x' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}, \\ x(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{cases}$$

Sabemos que una matriz fundamental es

$$\Phi(t) = \begin{pmatrix} \cos t & \sin t) \\ -\sin t) & \cos t \end{pmatrix}.$$

La matriz inversa es

$$\Phi^{-1}(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix},$$

y el producto $\Phi^{-1}(s)b(s)$ no es más que

$$\Phi^{-1}(s)b(s) = \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} \begin{pmatrix} \sin s \\ \cos s \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

con primitiva $(0, t-1)^T$. La solución es pues

$$x = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} \cos 1 & -\sin 1 \\ \sin 1 & \cos 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \begin{pmatrix} 0 \\ t-1 \end{pmatrix}$$

Llegados a este punto, debemos advertir que no existe un método general para calcular una matriz fundamental del sistema homogéneo. A continuación, destacamos algunos casos concretos en los que este cálculo sí que es posible y de cierta forma efectivo.

1. La matriz $A(t)$ es diagonal. En efecto, si $a_{ij}(t) = 0$ para todo $i \neq j$ entonces el sistema $x' = A(t)x$ se expresa como n EDOs lineales homogéneas

$$x_k' = a_{kk}(t)x_k, \qquad k = 1, ..., n,$$

cuya solución general para cada k es $x_k = c_k e^{\int a_{kk}(t)\,dt}$, siendo $c_k \in \mathbb{R}$. En particular, tenemos n soluciones linealmente independientes,

$$x_k = \begin{pmatrix} 0 \\ \vdots \\ x_{kk} = e^{\int a_{kk}(t)\,dt} \\ \vdots \\ 0 \end{pmatrix},$$

y una matriz fundamental es

$$\Phi(t) = \begin{pmatrix} e^{\int a_{11}(t)\,dt} & 0 & \cdots & 0 \\ 0 & e^{\int a_{22}(t)\,dt} & \ddots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & e^{\int a_{nn}(t)\,dt} \end{pmatrix}$$

2. La matriz $A(t)$ es triangular. Por ejemplo, triangular superior

$$A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) & \cdots & \cdots & a_{1n}(t) \\ 0 & a_{22}(t) & \cdots & \cdots & a_{2n}(t) \\ 0 & 0 & a_{33}(t) & \cdots & a_{3n}(t) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn}(t) \end{pmatrix}$$

En este caso podemos integrar la última ecuación

$$x_n' = a_{nn}(t)x_n,$$

obteniendo $x_n = e^{\int a_{nn}(t)\,dt}$. Ahora subimos de fila y consideramos la ecuación

$$x_{n-1}' = a_{n-1,n-1}(t)x_{n-1} + a_{n-1,n}(t)x_n.$$

Como la solución x_n es conocida, podemos resolver la ecuación lineal completa anterior obteniendo x_{n-1}. Iteramos este proceso y vamos obteniendo soluciones x_k hasta llegar a la primera ecuación

$$x_1' = a_{11}(t)x_1 + a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n,$$

donde las funciones $x_2, ..., x_n$ son conocidas por los pasos previos. Ahora solo restaría resolver la ecuación completa $x_1' = a_{11}(t)x_1 + b(t)$, con $b(t) = a_{12}(t)x_2 + \cdots + a_{1n}(t)x_n$ y tendríamos resueltas todas las ecuaciones escalares.

Notemos que para cada $k = 1, ..., n$, la aplicación vectorial

$$t \mapsto \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_k(t) \\ x_{k+1}(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

es solución de la EDO, y son todas linealmente independientes. Por tanto, una matriz fundamental viene dada por

$$\Phi(t) = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ x_2 & x_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ x_n & x_n & \cdots & x_n \end{pmatrix}$$

3. La matriz $A(t)$ es de coeficientes constantes. Esto lo veremos en la siguiente sección

3.2. Ecuación lineal de orden n uno-dimensional

Sean $a_0, a_1, ..., a_{n-1}, b \in C^0(I)$. La EDO lineal de orden n es la EDO

$$x^{n)} + a_{n-1}(t)x^{n-1)} + ... + a_1(t)x' + a_0(t)x = b(t). \tag{3.2.1}$$

Si $b \equiv 0$ la ecuación se dice homogénea, y si $b \not\equiv 0$ se dice completa.

Todos los resultados teóricos de resolución de EDOs desarrollados en el tema 2 han sido para el caso de orden uno. Llegados a este punto, podemos hacer uso del resultado teórico del pasado tema de que toda EDO de orden n se puede reducir a un sistema equivalente de n ecuaciones y orden 1. El cambio de variable usual $x_k = x^{k-1)},\ k = 1, ..., n$, transforma la EDO (3.2.1) en el sistema de primer orden equivalente

$$\begin{cases} x_1' = x_2, \\ x_2' = x_3, \\ \quad \vdots \\ x_{n-1}' = x_n, \\ x_n' = -a_{n-1}(t)x_n - ... - a_1(t)x_2 - a_0(t)x_1 + b(t). \end{cases}$$

Ahora bien, este sistema se puede escribir matricialmente como $x' = A(t)x + b(t)$, con $x = (x_1, ..., x_n)^T$ y

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_{j,j+1} = 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \cdots & -a_{n-2}(t) & -a_{n-1}(t) \end{pmatrix}, \qquad b(t) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}. \tag{3.2.2}$$

Aquí hemos abusado de la notación e identificado la función $b(t)$ con el vector cuyas $n-1$ componentes son nulas y la última es precisamente $b(t)$. El siguiente resultado es consecuencia de la reducción de orden dada por (3.2.2) y nos permite relacionar soluciones de las EDOs escalar y matricial.

Teorema 3.2.1. *Sea S_h^n el espacio de soluciones de la EDO homogénea (3.2.1). Entonces, S_h^n tiene estructura de espacio vectorial. Más aún, si denotamos por S_h' al espacio vectorial de soluciones del sistema $x' = A(t)x$ para $A(t)$ dada por (3.2.2) y consideremos el operador lineal*

$$T : C^n(I, \mathbb{R}) \to C^1(I, \mathbb{R}^n), \qquad Tx = (x, x', ..., x^{n-1)})^T, \tag{3.2.3}$$

entonces $T : S_h^n \to S_h'$ es un isomorfismo de espacios vectoriales y se tiene $\dim S = n$.

Demostración. Que el conjunto S_h^n de soluciones de la EDO homogénea es un espacio vectorial es algo trivial. Dado el operador lineal T definido en (3.2.3), consideremos la restricción $T : S_h^n \to S_h'$ y probemos que es un isomorfismo de espacios vectoriales. Para probar que T es inyectivo basta observar que $\ker T = \{0\}$. En efecto, sea $x \in S$ tal que $Tx = 0$. Por la propia definición $Tx = (x, x', ..., x^{n-1)})$ y por ser esta función vectorial nula en particular su primera componente es nula y por tanto $x = 0$. Para probar que T es sobreyectivo, tomemos $(x_1, ..., x_n) \in S_h'$. Por ser $(x_1, ..., x_n)$

solución de $x' = A(t)x$, para $A(t)$ definida por (3.2.2), se tiene

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}' = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 1 \\ -a_0(t) & -a_1(t) & -a_2(t) & \cdots & -a_{n-2}(t) & -a_{n-1}(t) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

o expresesado en coordenadas,

$$x_k' = x_{k+1}, \quad k = 0, ..., n-1, \qquad x_n' = -a_0(t)x_1 - a_1(t)x_1' - \cdots - a_{n-1}x_1^{n-1)}.$$

Ahora bien, definiendo $x = x_1$ concluimos $x_k = x^{k-1)}$ para $k = 1, ..., n$ y por tanto $Tx_1 = (x_1, ..., x_n)$. Para concluir, debemos probar que efectivamente $x \in S$, pero esto es inmediato sin más que derivar x_n y usar que $x_n = x^{n-1)}$,

$$x_n' = (x^{n-1)})' = x^{n)} = -a_0(t)x_1 - a_1(t)x_1' - \cdots - a_{n-1}(t)x_1^{n-1)}.$$

Esto implica que $x \in S_h^n$, probando la sobreyectividad de T. □

Una consecuencia de este resultado es la estructura de espacio afín de la ecuación completa.

Corolario 3.2.2. *El espacio de soluciones, S_c^n, de la EDO completa (3.2.1) tiene estructura de espacio afín sobre el espacio vectorial S_h^n de soluciones de la EDO homogénea. En particular, dada x_p solución particular y $\{x_1, ...x_n\}$ una base de S_h^n, cualquier solución de la EDO completa es de la forma*

$$x = x_p + c_1x_1 + \cdots + c_nx_n.$$

Más aún si T es el operador (3.2.3) y denotamos por S_c' al conjunto de soluciones de la EDO matricial completa $x' = A(t)x + b(t)$, con $A(t), b(t)$ dados por (3.2.2), entonces $T : S_c^n \to S_c'$ es una biyección.

Demostración. Sea S_c^n el espacio de soluciones de la EDO completa (3.2.1). Entonces la aplicación

$$\varphi : S_c^n \times S_c^n \to S_h^n, \qquad \varphi(x, y) = x - y$$

está bien definida e induce una estructura de espacio afín en S_c^n. En efecto, si $x, y \in S_c^n$ entonces la linealidad de la derivada y de la propia EDO nos permite concluir que $x - y \in S_h^n$, por lo que φ está bien definida. Que φ cumple las propiedades de aplicación que induce estructura de espacio afín es trivial.

Consideremos la restricción $T : S_c^n \to S_c'$. Primero veamos que T está bien definido, es decir dada $x \in S_c^n$ probemos que $Tx \in S_c'$. Por una parte,

$$(Tx)' = (x', x'', ..., x^{n)})^T = (x', x'', ..., -a_0(t)x - a_1(t)x' - \cdots - a_{n-1}(t)x^{n-1)} + b(t))^T.$$

Por otra parte,

$$A(t)Tx + b(t) = \begin{pmatrix} x' \\ x'' \\ \vdots \\ -a_0(t)x - a_1(t)x' - \cdots - a_{n-1}(t)x^{n-1)} + b(t) \end{pmatrix},$$

lo que prueba que $Tx \in S'_c$. Que T es biyectivo se prueba de la misma forma que en el Teorema 3.2.1. □

Del corolario anterior deducimos que si $(x_1, ..., x_n) \in S'_c$ entonces $x = x_1 \in S^n_c$. Esto tiene como consecuencia la existencia y unicidad del problema de Cauchy para la EDO (3.2.1).

Corolario 3.2.3. *El problema de Cauchy*

$$\begin{cases} x^{n)} + a_{n-1}(t)x^{n-1)} + ... + a_1(t)x' + a_0(t)x = b(t) \\ x(t_0) = u_1, \\ x'(t_0) = u_2, \\ \quad\vdots \\ x^{n-1)}(t_0) = u_n, \end{cases}$$

con $t_0 \in I$ *y* $u_k \in \mathbb{R}$ *para* $k = 1, ..., n$*, tiene existencia y unicidad.*

Demostración. Basta definir el problema de Cauchy análogo

$$\begin{cases} x' = A(t)x + b(t), \\ x(t_0) = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \end{cases} \qquad x = (x_1, ..., x_n)^T,$$

el cual sabemos por el Teorema 3.1.4 que tiene existencia y unicidad. En vista del operador T dado por (3.2.3) podemos afirmar que $x = x_1$ es solución de (3.2.1) y cumple $x^{k-1)}(t_0) = x_k(t_0) = u_k$ para todo $k = 1, ..., n$. □

Por analogía con todo lo visto hasta ahora, llamaremos sistema fundamental de soluciones de (3.2.1) a cualquier base $\{x_1, ..., x_n\}$ de su espacio de soluciones. Ahora bien, dado un sistema fundamental de (3.2.1), las funciones vectoriales dadas por el isomorfismo T,

$$Tx_1 = \begin{pmatrix} x_1 \\ x_1' \\ \vdots \\ x_1^{n-1)} \end{pmatrix}, \quad Tx_2 = \begin{pmatrix} x_2 \\ x_2' \\ \vdots \\ x_2^{n-1)} \end{pmatrix}, \quad \cdots, \quad Tx_n = \begin{pmatrix} x_n \\ x_n' \\ \vdots \\ x_n^{n-1)} \end{pmatrix},$$

forman un sistema fundamental del sistema $x' = A(t)x$, lo cual equivale a que la matriz

$$\Phi(t) = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1' & x_2' & \cdots & x_n' \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{n-1)} & x_2^{n-1)} & \cdots & x_n^{n-1)} \end{pmatrix}$$

sea una matriz fundamental. Sabemos por el Teorema 3.1.9 que las matrices fundamentales tienen determinante no nulo y de hecho esta propiedad las caracteriza. Esto motiva la siguiente definición.

Definición 3.2.4. Dadas $x_1, ..., x_n \in C^{n-1}(I)$, se define el wronskiano como la función

$$W(x_1, ..., x_n)(t) = \det \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1' & x_2' & \cdots & x_n' \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{n-1)} & x_2^{n-1)} & \cdots & x_n^{n-1)} \end{pmatrix}.$$

Ahora la Proposición 3.1.8 se enuncia como sigue.

Teorema 3.2.5. *Sean $x_1, ..., x_n$ soluciones de (3.2.1). Las siguientes afirmaciones son equivalentes.*

1. *$x_1, ..., x_n$ es un sistema fundamental de (3.2.1); esto es, $x_1, ..., x_n$ son linealmente independientes.*
2. *Existe $t_0 \in I$ tal que $W(x_1, ..., x_n)(t_0) \neq 0$.*
3. *Para todo $t \in I$ se tiene $W(x_1, ..., x_n)(t) \neq 0$.*

Observación 3.1. Si $x_1, ..., x_n$ son funciones linealmente dependientes, su wronskiano es idénticamente nulo ya que encontramos columnas linealmente dependientes entre ellas. Ahora bien, el recíproco no es cierto (salvo que impongamos además que $x_1, ..., x_n$ son soluciones de la misma ecuación lineal homogénea con coeficientes continuos). En efecto, sean

$$\begin{aligned} &x_1(t) \neq 0, \quad \text{si} \quad t \in [0, \tfrac{1}{2}), \qquad &x_1(t) = 0, \quad \text{si} \quad t \in [\tfrac{1}{2}, 1], \\ &x_2(t) = 0, \quad \text{si} \quad t \in [0, \tfrac{1}{2}), \qquad &x_2(t) \neq 0, \quad \text{si} \quad t \in [\tfrac{1}{2}, 1], \end{aligned}$$

donde se impone que las funciones sean C^1 (de hecho se pueden definir para que sean C^∞). Claramente ambas funciones son linealmente independientes, pero sin embargo su wronskiano es idénticamente nulo.

3.2.1. Ecuación completa. Método de variación de las constantes

Para finalizar, estudiamos cómo resolver la ecuación completa

$$x^{n)} + a_{n-1}(t)x^{n-1)} + ... + a_1(t)x' + a_0(t)x = b(t).$$

Supongamos conocido un sistema fundamental $x_1, ..., x_n$ de la ecuación homogénea. Entonces, toda solución de la ecuación homogénea es de la forma

$$x = c_1x_1 + \cdots + c_nx_n, \qquad c_k \in \mathbb{R}.$$

De la misma forma que en el caso de orden 1 y para sistemas diferenciales completos, bastará conocer una solución particular x_p de la ecuación completa. Para encontrar una solución particular usamos la idea de la variación de las constantes, es decir buscamos x_p de la forma

$$x_p = c_1(t)x_1 + \cdots + c_n(t)x_n,$$

con $c_k : I \to \mathbb{R}$ de clase C^1. Sabemos que si $x_p \in S_c$ entonces $Tx_p \in S_c'$ y viceversa, y por tanto bastará encontrar una solución vectorial particular del sistema matricial $x' = A(t)x + b(t)$ para posteriormente quedarnos con su primera coordenada como solución particular de (3.2.1). Sea pues la matriz fundamental

$$\Phi(t) = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_1' & x_2' & \cdots & x_n' \\ \vdots & \vdots & \cdots & \vdots \\ x_1^{n-1)} & x_2^{n-1)} & \cdots & x_n^{n-1)} \end{pmatrix}.$$

Sabemos que una solución particular de la forma $\Phi(t)c(t)$, con $c(t) = (c_1(t), ..., c_n(t))^T$, satisface $\Phi(t)c'(t) = b(t)$. El sistema de ecuaciones lineales obtenido es

$$\begin{cases} c_1'(t)x_1 + c_2'(t)x_2 & + \cdots + & c_n'(t)x_n & = & 0, \\ c_1'(t)x_1' + c_2'(t)x_2' & + \cdots + & c_n'(t)x_n' & = & 0, \\ \vdots & \cdots & \vdots & & \vdots \\ c_1'(t)x_1^{n-1)} + c_2'(t)x_2^{n-1)} & + \cdots + & c_n'(t)x_n^{n-1)} & = & b(t), \end{cases}$$

donde debemos incidir en que las incógnitas son las funciones $c_k(t)$. Como la matriz de coeficientes $\Phi(t)$ tiene determinante no nulo al ser una matriz fundamental, la regla de Cramer nos da la solución

$$c_k'(t) = \frac{\det \Phi_k(t)}{\det \Phi(t)} b(t),$$

donde $\Phi_k(t)$ es la matriz obtenida al sustituir en $\Phi(t)$ la columna k-ésima por el vector $(0, ..., 0, 1)$. En términos del wronskiano,

$$c_k'(t) = \frac{W_k(t)}{W(t)} b(t),$$

donde $W(t) = W(x_1, ..., x_n)(t)$ y $W_k(t) = \det \Phi_k(t)$. Finalmente, integrando $c_k'(t)$ tenemos que una solución particular es

$$x_p = \int \frac{W_1(t)}{W(t)} b(t)\, dt\, x_1 + \cdots + \int \frac{W_n(t)}{W(t)} b(t)\, dt\, x_n.$$

Para concluir, la solución de la ecuación completa (3.2.1) es

$$x(t) = c_1x_1 + \cdots + c_nx_n + x_p, \qquad (c_1, ..., c_n) \in \mathbb{R}^n.$$

Ejemplo 3.4. Consideremos la ecuación

$$x'' - \frac{3}{t}x' + \frac{3}{t^2}x = t^2 + t, \qquad t \in (0, \infty).$$

Para resolver esta ecuación, primero necesitamos un sistema fundamental de soluciones. Como la EDO es de segundo orden, necesitaremos dos funciones linealmente independientes que resuelvan la EDO homogénea. Además, por la naturaleza de la ecuación no parece descabellado intuir que posibles soluciones sean polinómicas. Supongamos pues $x(t) = t^k$ y sustituimos en la EDO homogénea, obteniendo

$$k(k-1)t^{k-2} - 3kt^{k-2} + 3t^{k-2} = 0,$$

de donde concluimos

$$k^2 - 4k + 3 = 0 \Longrightarrow k = 1, 3.$$

Por tanto, las funciones $x_1 = t$, $x_2 = t^3$ son soluciones de la EDO homogénea y claramente son linealmente independientes. En consecuencia,

$$\Phi(t) = \begin{pmatrix} t & t^3 \\ 1 & 3t^2 \end{pmatrix}$$

es matriz fundamental y la solución de la EDO homogénea es

$$x_h = c_1 t + c_2 t^3, \qquad c_1, c_2 \in \mathbb{R}.$$

Para resolver la EDO completa, primero calculamos

$$W(t) = \det \Phi(t) = 2t^3, \qquad W_1(t) = -t^3, \qquad W_2(t) = t.$$

Si buscamos $x_p = c_1(t)t + c_2(t)t^3$, sabemos que cada $c_k(t)$ es solución de la EDO

$$c_1'(t) = \frac{W_1(t)}{W(t)} b(t) = -\frac{1}{2}(t^2 + t), \qquad c_2'(t) = \frac{W_2(t)}{W(t)} b(t) = \frac{1}{2t^2}(t^2 + t),$$

de donde concluimos

$$c_1(t) = -\frac{t^3}{6} - \frac{t^2}{4}, \qquad c_2(t) = \frac{t}{2} + \frac{1}{2}\log t.$$

Tras sustituir y simplificar, concluimos que la solución general de la EDO es de la forma

$$x(t) = \frac{t^4}{3} - \frac{t^3}{4} + \frac{t^3}{2}\log t + c_1 t + c_2 t^3, \qquad c_1, c_2 \in \mathbb{R}.$$

Finalmente, si se nos diesen condiciones iniciales $x(t_0) = x_0$ y $x'(t_0) = x_1$, bastaría sustituir en la expresión anterior y obtener los valores c_1, c_2 adecuados.

3.3. Ecuación lineal de orden uno n-dimensional con coeficientes constantes

Hasta ahora el poder resolver la ecuación (3.1.5) depende fuertemente de poder encontrar un sistema fundamental de la ecuación homogénea $x' = A(t)x$. A final de la sección 3.1 vimos que un caso concreto de matriz A para la que podíamos resolver el sistema homogéneo correspondiente es el caso en que A es una matriz diagonal,

$$A(t) = \operatorname{diag}(a_1(t), ..., a_n(t)).$$

En tal caso, $x_k' = a_k(t)x_k$ tiene como solución $x_k = e^{\int a_k(t)\,dt}$. Una matriz fundamental es por tanto

$$\Phi(t) = \operatorname{diag}(e^{\int a_1(t)\,dt}, ..., e^{\int a_n(t)\,dt}).$$

Podríamos definir de forma simbólica, sin ningún tipo de rigor matemático por ahora, la exponencial de una matriz diagonal $A = \operatorname{diag}(a_1, ..., a_n)$ como

$$e^A = \exp(\operatorname{diag}(a_1, ..., a_n)) = \operatorname{diag}(e^{a_1}, ..., e^{a_n}).$$

Es decir, la exponencial de una matriz diagonal la definiríamos como la matriz diagonal formada por la exponencial de las entradas en la diagonal principal. Con esta notación, insistimos informal, tendríamos que una matriz fundamental del sistema $x' = A(t)x$ para $A(t)$ diagonal sería

$$\Phi(t) = e^{\int A(t)\,dt}.$$

Notemos que toda matriz fundamental cumple $\Phi'(t) = A(t)\Phi(t)$ y en este caso concreto es como si la propia exponencial de la matriz recuperase esta propiedad al derivarla de manera informal.

Supongamos que la matriz $A(t)$ es constante. Si A fuese diagonalizable, existiría P invertible tal que $A = PDP^{-1} = P\mathrm{diag}(\lambda_1, ..., \lambda_n)P^{-1}$. Si llamamos $y = P^{-1}x$ el problema original es

$$y' = P^{-1}x' = P^{-1}PDP^{-1}x = Dy.$$

Este problema ya sabemos resolverlo, obteniendo la solución general

$$y(t) = \mathrm{diag}(e^{\lambda_1 t}, ..., e^{\lambda_n t})c, \qquad c \in \mathbb{R}^n.$$

Ahora basta deshacer el cambio, obteniendo

$$x = P\,\mathrm{diag}(e^{\lambda_1 t}, ..., e^{\lambda_n t})c, \qquad c \in \mathbb{R}^n.$$

Como P es regular en particular se puede ver como un isomorfismo de $\mathbb{R}^n$ en sí mismo y por tanto podemos sustituir c por $P^{-1}c$, reescribiendo la solución como

$$x = P\,\mathrm{diag}(e^{\lambda_1 t}, ..., e^{\lambda_n t})P^{-1}c, \qquad c \in \mathbb{R}^n$$

Observemos que esta solución solo depende de cierta forma de la parte semejante diagonal de A y no tanto de las matrices de paso P, P^{-1}. Con la notación informal $e^{tD} = \mathrm{diag}(e^{\lambda_1 t}, ..., e^{\lambda_n t})$, la solución sería

$$x = Pe^{tD}P^{-1}c, \qquad c \in \mathbb{R}^n.$$

Como quiera que $tA = P(tD)P^{-1}$, es como si la solución *rompiese* la semejanza y solo se quedase con la parte diagonal. Si finalmente definiésemos de manera informal

$$e^{At} = Pe^{tD}P^{-1} = P\,\mathrm{diag}(e^{\lambda_1 t}, ..., e^{\lambda_n t})P^{-1}, \tag{3.3.1}$$

tendríamos que la solución del sistema $x' = Ax$ es

$$x(t) = e^{At}c, \qquad c \in \mathbb{R}^n.$$

En particular, la solución del problema de Cauchy $x(t_0) = x_0$, que viene dada por

$$x(t) = P\,\mathrm{diag}(e^{\lambda_1(t-t_0)}, ..., e^{\lambda_n(t-t_0)})P^{-1}x_0,$$

se escribiría con esta notación informal como

$$x(t) = e^{A(t-t_0)}x_0.$$

Esto no debería extrañarnos, puesto que en el caso de dimensión 1 estudiado en el pasado tema, la solución del problema de Cauchy $x' = ax,\ x(t_0) = x_0$, con $a(t) = a \in \mathbb{R}$, es $x = e^{a(t-t_0)}x_0$. En lo que queda de tema daremos formalidad a esta idea de *exponencial de una matriz*, proporcionando una forma de calcular soluciones del sistema homogéneo cuando A es de coeficientes constantes y que, al menos, satisfaga que la exponencial de una matriz diagonal es la matriz diagonal formada por la exponencial de sus coeficientes diagonales.

3.3.1. Solución general. Exponencial de una matriz

Sea $A \in \mathcal{M}_n(\mathbb{R})$. Fijada una norma en $\mathbb{R}^n$, por ejemplo la usual, recordemos que se define la *norma uniforme* de A como

$$||A|| = \text{máx}\{||Ax||\colon ||x|| \leq 1\} = \text{máx}\{||Ax||\colon ||x|| = 1\}.$$

Así definida, la norma uniforme cumple todas las propiedades que debe satisfacer cualquier norma. Más aún, se tienen las siguientes propiedades específicas.

Proposición 3.3.1. *Sean $A, B \in \mathcal{M}_n(\mathbb{R})$ y $x \in \mathbb{R}^n$. La norma uniforme satisface las siguientes propiedades.*

1. $||Ax|| \leq ||A||\,||x||$.
2. $||AB|| \leq ||A||\,||B||$.
3. $||A^k|| \leq ||A||^k$*, para todo* $k \in \mathbb{N}$.

Demostración.

1. Si $x = 0$ no hay nada que probar. Si $x \neq 0$ entonces consideramos el vector unitario $x/||x||$. Por la definición de norma uniforme,

$$||A\frac{x}{||x||}|| = \frac{1}{||x||}||Ax|| \leq ||A|| \Longrightarrow ||Ax|| \leq ||A||\,||x||.$$

2. Dado $x \in \mathbb{R}^n$ con $||x|| \leq 1$, en virtud del apartado anterior se tiene

$$||ABx|| \leq ||A||\,||Bx|| \leq ||A||\,||B||\,||x|| \leq ||A||\,||B||.$$

   Ahora se concluye tomando el máximo entre todos los x con $||x|| \leq 1$.

3. Si $k = 0, 1$, es trivial. Si $k > 1$, es una consecuencia del apartado anterior sin más que aplicarlo k veces.

□

Estamos en condiciones de introducir el concepto de exponencial de una matriz.

Definición 3.3.2. Dada $A \in \mathcal{M}_n(\mathbb{R})$, se define la exponencial de A y se denota por e^A o $\exp(A)$ como

$$e^A = \exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!},$$

donde la suma se entiende componente a componente y A^0 es la matriz identidad I_n de orden n.

Vamos a probar que la exponencial de una matriz está bien definida.

Teorema 3.3.3. *Dada $A \in \mathcal{M}_n(\mathbb{R})$, la matriz e^A está bien definida. Más aún, la serie*

$$\sum_{k=0}^{\infty} \frac{A^k}{k!}$$

converge absolutamente.

Demostración. Recordemos que dada una sucesión $\{a_k\}_{k\geq 1}$ en un espacio normado $(X, ||\cdot||)$, se dice que la serie $\sum_{k=1}^{\infty} a_k$ es absolutamente convergente si la serie $\sum_{k=1}^{\infty} ||a_k||$ es convergente. Por la propiedad *3.* de la Proposición 3.3.1 se tiene

$$\frac{||A^k||}{k!} \leq \frac{||A||^k}{k!}.$$

Como la serie $\sum_{k=0}^{\infty} \frac{||A||^k}{k!}$ es convergente (precisamente converge a $e^{||A||}$), el criterio de comparación prueba la convergencia absoluta de la serie que define e^A. □

Observación 3.2. De la demostración anterior deducimos

$$||\sum_{k=0}^{n} \frac{A^k}{k!}|| \leq \sum_{k=0}^{n} \frac{||A^k||}{k!} \leq e^{||A||}.$$

En consecuencia, tomando límite cuando $n \to \infty$ en el miembro de la izquierda se tiene

$$||e^A|| \leq e^{||A||}.$$

El siguiente resultado es un lema técnico que es análogo a la fórmula del producto de Cauchy para series numéricas.

Lema 3.3.4. *Sean $A = \sum_{k=0}^{\infty} A_k$ y $B = \sum_{j=0}^{\infty} B_j$ dos series de matrices absolutamente convergentes. Para cada $l \geq 0$ se define*

$$C_l = \sum_{k+j=l} A_k B_j.$$

Entonces, la serie $C = \sum_{l=0}^{\infty} C_l$ es absolutamente convergente y $C = AB$.

Una vez conocido el lema que concierne al producto de Cauchy para series numéricas, enunciamos el resultado que recopila las propiedades más importantes de la exponencial de una matriz.

Teorema 3.3.5. *Dadas $A, B \in \mathcal{M}_n(\mathbb{R})$, se tienen las siguientes propiedades.*

1. *$e^{0_n} = I_n$, donde 0_n denota la matriz cuadrada nula.*

2. *Si $AB = BA$, entonces $e^{A+B} = e^A e^B$.*

3. *$e^{-A} = (e^A)^{-1}$. En particular, e^A es invertible.*

4. *Si P es invertible y $B = PAP^{-1}$, entonces $e^B = Pe^A P^{-1}$.*

5. *Sean $p, q \in \mathbb{N}$ con $p + q = n$ y $A_1 \in \mathcal{M}_p(\mathbb{R})$ y $A_2 \in \mathcal{M}_q(\mathbb{R})$. Se define $A = \operatorname{diag}\{A_1, A_2\}$ como la matriz cuyas primeras $p \times p$ entradas forman la matriz A_1, cuyas últimas $q \times q$ son la matriz A_2 y el resto de entradas son cero. Entonces*

$$e^A = \operatorname{diag}\{e^{A_1}, e^{A_2}\}.$$

6. *Si $v \in \mathbb{R}^n$ es vector propio de A con valor propio λ, entonces v es vector propio de e^A correspondiente al valor propio e^{λ}.*

Demostración. La propiedad *1.* es consecuencia de la propia definición de exponencial de una matriz, mientras que *3.* es consecuencia inmediata de *2.* y del hecho de que A y $-A$ conmutan. Probemos la propiedad *2.* La conmutatividad de A y B aseguran que se cumple la fórmula del binomio de Newton,

$$(A+B)^l = \sum_{k=0}^{l} \binom{l}{k} A^k B^{l-k} = l! \sum_{k+j=l} \frac{A^k}{k!}\frac{B^j}{j!}.$$

Ahora bien,

$$e^{A+B} = \sum_{l=0}^{\infty} \frac{(A+B)^l}{l!} = \sum_{l=0}^{\infty} \sum_{k+j=l} \frac{A^k}{k!}\frac{B^j}{j!},$$

y en virtud del Lema 3.3.4,

$$e^{A+B} = \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} \right) \left(\sum_{j=0}^{\infty} \frac{B^j}{j!} \right) = e^A e^B.$$

La propiedad *4.* es consecuencia de las identidades evidentes $P(A+B)P^{-1} = PAP^{-1} + PBP^{-1}$ y $(PAP^{-1})^k = PA^kP^{-1}$, de las que se deduce

$$\sum_{k=0}^{n} \frac{(tB)^k}{k!} = \sum_{k=0}^{n} \frac{(tPAP^{-1})^k}{k!} = P \left(\sum_{k=0}^{n} \frac{(tA)^k}{k!} \right) P^{-1}.$$

Tomando límites cuando $n \to \infty$, se tiene *4.*

Referente a la propiedad *5.*, podemos ver A como la suma directa

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} = A_1 \oplus A_2,$$

y usamos la propiedad $(A_1 \oplus A_2)^k = A_1^k \oplus A_2^k$, obteniendo

$$e^A = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(A_1 \oplus A_2)^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{A_1^k \oplus A_2^k}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1}{k!} \begin{pmatrix} A_1^k & 0 \\ 0 & A_2^k \end{pmatrix} = \begin{pmatrix} e^{A_1} & 0 \\ 0 & e^{A_2} \end{pmatrix} = \mathrm{diag}\{e^{A_1}, e^{A_2}\}.$$

Para finalizar con la prueba de la propiedad *6.*, sea $v \in \mathbb{R}^n$ vector propio de A con valor propio λ, esto es $Av = \lambda v$. Como quiera que $A^k v = \lambda^k v$,

$$e^A v = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{A^k v}{k!} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{\lambda^k v}{k!} = e^\lambda v,$$

concluyendo la prueba. □

Observación 3.3. El resultado anterior nos facilita el cálculo de la exponencial de una matriz de la siguiente forma.

1. La propiedad *4.* nos asegura que basta expresar A respecto de una base adecuada. En concreto, si $A = PDP^{-1}$ con D diagonal, su exponencial es

   $$e^A = P\mathrm{diag}\{e^{\lambda_1}, ..., e^{\lambda_n}\}P^{-1},$$

   lo cual concuerda con la definición informal (3.3.1) que nos proporcionaba una matriz fundamental del sistema $x' = Ax$.

2. La propiedad *5.* nos asegura que si A no es diagonalizable pero tiene una expresión $A = PJP^{-1}$ siendo
$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix}$$
una matriz con bloques J_i cuyas exponenciales e^{J_i} son conocidas o calculables, entonces la exponencial e^A es inmediata de calcular.

Hasta ahora hemos definido la exponencial de una matriz de coeficientes constantes, pero de igual forma que sucede con la exponencial real podemos definirla para una matriz con coeficientes variables.

Definición 3.3.6. Sea $A : I \subset \mathbb{R} \to \mathcal{M}_n(\mathbb{R})$. Se define la exponencial de $A(t)$ como
$$e^{A(t)} = \sum_{k=0}^{\infty} \frac{A(t)^k}{k!}.$$

Observamos que para cada $t \in I$ fijo, la serie $\sum_{k=0}^{\infty} \frac{A(t)^k}{k!}$ converge puntualmente a $e^{A(t)}$. Más aún, sea $K \subset I$ compacto y definimos
$$M = \max_{t \in K} ||A(t)||,$$
donde $||A(t)||$ es la norma uniforme de cada matriz $A(t)$. Si denotamos $a_k(t) = \frac{A(t)^k}{k!}$, para cada $t \in K$ se tiene
$$||a_k(t)|| = ||\frac{A(t)^k||}{k!} \leq \frac{||A(t)||^k}{k!} \leq \frac{M^k}{k!}.$$
Como la serie $\sum_{k=0}^{\infty} M^k/k!$ converge a e^M, se tiene que la serie $\sum_{k=0}^{\infty} \frac{A(t)^k}{k!}$ converge absoluta y uniformemente en K a una cierta función matricial límite, que llamamos $e^{A(t)}$, sin más que aplicar el criterio de Weierstrass. Más aún, para todo $t \in K$ se tiene
$$||\sum_{k=0}^{n} \frac{A(t)^k}{k!}|| \leq \sum_{k=0}^{n} \frac{||A(t)^k||}{k!} \leq \sum_{k=0}^{n} \frac{||A(t)||^k}{k!} \leq \sum_{k=0}^{n} \frac{M^k}{k!} \leq e^M.$$
Tomando límite $n \to \infty$ concluimos $||e^{A(t)}|| \leq e^M$.

El siguiente paso es estudiar cómo se comporta la exponencial de una matriz respecto de la derivación. Es esperable que reproduzca de cierta forma las propiedades de la derivada exponencial para funciones reales. No obstante, impondremos una hipótesis de conmutatividad a priori muy fuerte. En la prueba del resultado, necesitaremos el siguiente lema.

Lema 3.3.7. *Sean $f_n : [a,b] \to \mathbb{R}^n$ aplicaciones C^1 que convergen puntualmente a $f : [a,b] \to \mathbb{R}^n$. Si $\{f_n'\} \to g$ uniformemente, entonces f es derivable y $f' = g$.*

Demostración. Sea $f : [a,b] \to \mathbb{R}^n$ la función límite puntual. Por ser f_n de clase C^1, el Teorema Fundamental del Cálculo asegura
$$f_n(t) = f_n(t_0) + \int_{t_0}^{t} f_n'(s)\, ds.$$

Además, como f_n' son continuas, la convergencia uniforme asegura la continuidad de g. Tomando límite cuando $n \to \infty$ en la igualdad anterior y usando que la convergencia uniforme de f_n' a g nos permite intercambiar derivada y límite, se tiene

$$f(t) = f(t_0) + \int_{t_0}^{t} g(s)\, ds.$$

Esto nos dice que f es de clase C^1 por ser g continua, y además $f'(t) = g(t)$. $\square$

Ahora sí, enunciamos y demostramos el resultado de derivación de la exponencial de una matriz.

Teorema 3.3.8. *Sea $A : I \to \mathcal{M}_n(\mathbb{R})$. Si $A(t)$ es de clase C^1 en I, entonces $e^{A(t)}$ es de clase C^1 en I. Más aún, si $A(t)$ y $A'(t)$ conmutan, entonces*

$$(e^{A(t)})' = A'(t)e^{A(t)}.$$

Demostración. Como la derivada es una cuestión local, fijamos t_0 y trabajamos en un intervalo compacto $I_0 = [t_0 - \varepsilon, t_0 + \varepsilon] \subset I$ (suponemos $[t_0, t_0 + \varepsilon]$ o $[t_0 - \varepsilon, t_0]$ si t_0 es uno de los extremos del intervalo I). El objetivo es demostrar que la serie

$$\sum_{k=0}^{\infty} \left(\frac{A(t)^k}{k!} \right)' \tag{3.3.2}$$

converge absolutamente. Hay que destacar que esta suma realmente comienza para $k = 1$, porque para $k = 0$ se tiene $A(t)^0 = I_n$ y la derivada de la identidad es nula al ser una matriz constante. Primero, observemos que se tiene

$$\left(\frac{A(t)^k}{k!} \right)' = \sum_{i=1}^{k} \frac{A(t) \cdots \overset{i)}{A'(t)} \cdots A(t)}{k!},$$

donde estamos indicando que derivamos el término i-ésimo en el producto $A(t)^k$ (recordamos que todavía $A(t), A(t)'$ no se han supuesto conmutativas). Sean ahora

$$M = \max_{t \in I_0} ||A(t)||, \qquad N = \max_{t \in I_0} ||A'(t)||,$$

los cuales existen por ser $A(t)$ de clase C^1 en I. Ahora bien, para cada $n \in \mathbb{N}$,

$$\sum_{k=1}^{n} \left|\left| \left(\frac{A(t)^k}{k!} \right)' \right|\right| \leq \sum_{k=1}^{n} \sum_{i=1}^{k} \frac{||A(t)|| \cdots ||\overset{i)}{A'(t)}|| \cdots ||A(t)||}{k!} \leq \sum_{k=1}^{n} \frac{kM^{k-1}N}{k!} = \sum_{k=0}^{\infty} \frac{M^k N}{k!} = Ne^M,$$

y por tanto la serie (3.3.2) converge absolutamente; de hecho, por ser I_0 compacto la convergencia es también uniforme.

Ahora podemos aplicar el Lema 3.3.7 a la sucesión de sumas parciales de la serie (3.3.2) y concluir que la función $e^{A(t)}$ es derivable y su derivada es precisamente el valor de la suma

$$(e^{A(t)})' = \sum_{k=0}^{\infty} \left(\frac{A(t)^k}{k!} \right)'.$$

Finalmente, si $A(t), A'(t)$ conmutan entonces la suma anterior puede evaluarse como

$$\begin{aligned}(e^{A(t)})' &= \sum_{k=0}^{\infty}\left(\frac{A(t)^k}{k!}\right)' = \sum_{k=1}^{\infty}\frac{kA'(t)A(t)^{k-1}}{k!}\\ &= \sum_{k=1}^{\infty}\frac{A'(t)A(t)^{k-1}}{(k-1)!} = \sum_{k=0}^{\infty}\frac{A'(t)A(t)^{k}}{k!} = A'(t)e^{A(t)}.\end{aligned}$$

□

Estamos ya en condiciones de poder aplicar los resultados obtenidos a la resolución de sistemas de ecuaciones lineales. El siguiente resultado es consecuencia de todo lo estudiado hasta ahora.

Teorema 3.3.9. *Consideremos el sistema lineal homogéneo*

$$x' = A(t)x,$$

y sea $\int A(t)\,dt$ *una primitiva de* $A(t)$. *Supongamos además que* $A(t)$ *y* $\int A(t)\,dt$ *conmutan. Entonces,*

$$\Phi(t) = e^{\int A(t)\,dt}$$

es una matriz fundamental del sistema $x' = A(t)x$ *y la solución viene dada por* $x(t) = e^{\int A(t)\,dt}c$, $c \in \mathbb{R}^n$.

En particular, si A *es constante entonces* e^{tA} *es matriz fundamental del sistema homogéneo* $x' = Ax$ *y* $x(t) = e^{tA}c$, $c \in \mathbb{R}^n$, *es la solución general. Más aún, la solución del problema de Cauchy*

$$\begin{cases} x' = Ax, \\ x(t_0) = x_0, \end{cases}$$

viene dada por $x(t) = e^{(t-t_0)A}x_0$.

El problema de resolver un sistema matricial de la forma

$$x' = A(t)x + b(t)$$

se reduce pues al cálculo de la exponencial de la matriz $A(t)$. En la siguiente sección nos dedicaremos a esta empresa cuando A es constante; es decir, calcularemos e^{tA}. No obstante, podemos calcular matrices fundamentales de sistemas más generales siempre y cuando podamos calcular $e^{A(t)}$. Por ejemplo, considérese el sistema $x' = A(t)x$, donde

$$A(t) = \begin{pmatrix} -1 & 0 \\ t & -1 \end{pmatrix}.$$

Una primitiva de $A(t)$ es

$$B(t) = \begin{pmatrix} -t & 0 \\ \frac{t^2}{2} & -t \end{pmatrix},$$

la cual satisface $A(t)B(t) = B(t)A(t)$. Vamos a calcular $e^{B(t)}$ y así anticiparemos algunos métodos para el cálculo de la exponencial de matrices constantes. Para empezar, obsérvese que

$$B(t) = \begin{pmatrix} -t & 0 \\ 0 & -t \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \frac{t^2}{2} & 0 \end{pmatrix}.$$

Como ambas matrices conmutan, se tiene

$$e^{B(t)} = \exp\begin{pmatrix} -t & 0 \\ 0 & -t \end{pmatrix} \cdot \exp\begin{pmatrix} 0 & 0 \\ \frac{t^2}{2} & 0 \end{pmatrix}.$$

La primera es la exponencial de una matriz diagonal y por tanto

$$\exp\begin{pmatrix} -t & 0 \\ 0 & -t \end{pmatrix} = \begin{pmatrix} e^{-t} & 0 \\ 0 & e^{-t} \end{pmatrix}.$$

Por otra parte, de la definición de exponencial se tiene

$$\exp\begin{pmatrix} 0 & 0 \\ \frac{t^2}{2} & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \frac{t^2}{2} & 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} + \cdots = \begin{pmatrix} 1 & 0 \\ \frac{t^2}{2} & 1 \end{pmatrix}.$$

En resumen,

$$e^{B(t)} = \begin{pmatrix} e^{-t} & 0 \\ 0 & e^{-t} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ \frac{t^2}{2} & 1 \end{pmatrix} = \begin{pmatrix} e^{-t} & 0 \\ \frac{t^2}{2}e^{-t} & e^{-t} \end{pmatrix}$$

es una matriz fundamental del sistema

$$x' = \begin{pmatrix} -1 & 0 \\ t & -1 \end{pmatrix} x.$$

Es inmediato comprobar que, en efecto, las funciones

$$x_1(t) = \begin{pmatrix} e^{-t} \\ \frac{t^2}{2}e^{-t} \end{pmatrix}, \qquad x_2(t) = \begin{pmatrix} 0 \\ e^{-t} \end{pmatrix}$$

son soluciones de este sistema lineal homogéneo y que ambas son linealmente independientes, por lo que definen un sistema fundamental.

3.3.2. Cálculo la matriz exponencial en algunos casos concretos

A continuación, detallamos el cálculo explícito de la exponencial de una matriz $A \in \mathcal{M}_n(\mathbb{R})$ para algunos casos concretos. El caso general no obstante será el objeto de estudio de la siguiente sección.

1. **Caso A diagonalizable en $\mathbb{R}$.** Si A es diagonalizable en $\mathbb{R}$ entonces existen $\lambda_1, ..., \lambda_n \in \mathbb{R}$ valores propios de A y $\mathcal{B} = \{v_1, ..., v_n\} \subset \mathbb{R}^n$ una base de vectores propios, $Av_k = \lambda_k v_k,\ k = 1, ..., n$. En tal caso, $x_k(t) = e^{\lambda_k t} v_k$ es solución de $x' = Ax$. En efecto,

$$x_k' = \lambda_k e^{\lambda_k t} v_k = e^{\lambda_k t}\lambda_k v_k = e^{\lambda_k t} A v_k = A x_k.$$

Además, $\{x_1, .., x_n\}$ son funciones linealmente independientes, puesto que

$$\det(e^{\lambda_1 t} v_1, ..., e^{\lambda_n t} v_n) = e^{(\lambda_1 + \cdots + \lambda_n)t} \det(v_1, ..., v_n) \neq 0,$$

por ser $v_1, ..., v_n$ elementos de una base. En consecuencia, la matriz

$$\Phi(t) = \begin{pmatrix} e^{\lambda_1 t} v_1 & \cdots & e^{\lambda_n t} v_n \end{pmatrix},$$

con columnas formadas por $x_1, ..., x_k$, es matriz fundamental del sistema $x' = Ax$. Como sabemos que e^{tA} es también matriz fundamental y dos matrices fundamentales se diferencian por el producto de una matriz regular C, se tiene $e^{tA} = \Phi(t)C$ y por tanto si denotamos por P a la matriz del cambio de base de $\mathcal{B}$ a la base canónica, se tiene $C = \Phi(0)^{-1} = P^{-1}$, o de forma equivalente,

$$e^{tA} = \begin{pmatrix} e^{\lambda_1 t} v_1 & \cdots & e^{\lambda_n t} v_n \end{pmatrix} P^{-1}.$$

Finalmente, observemos que

$$\Phi(t) = \begin{pmatrix} e^{\lambda_1 t} v_1 & \cdots & e^{\lambda_n t} v_n \end{pmatrix} = P \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix},$$

con lo que concluimos

$$e^{tA} = P \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} P^{-1}.$$

Esto se podría haber razonado equivalentemente a través de la propiedad *4.* del Teorema 3.3.5. En particular, si A es diagonal entonces la propia base canónica es una base de vectores propios y $P = P^{-1} = I_n$, obteniendo

$$e^{tA} = \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix}.$$

2. **Caso A diagonalizable en $\mathbb{C}$.** Supongamos que A es diagonalizable pero tiene valores propios reales y complejos

$$\{\lambda_1, ..., \lambda_k, \mu_1, ..., \mu_r, \overline{\mu_1} ..., \overline{\mu_r}\},$$

con $k+2r = n$. En tal caso será conveniente considerar soluciones complejas, $x : I \to \mathbb{C}^n$, del sistema $x' = A(t)x$. Ahora bien, si llamamos $u, w : I \to \mathbb{R}^n$ a las partes reales e imaginarias de $x = u + iw$, se tiene

$$u' + iw' = x' = A(t)x = A(t)u + iA(t)w,$$

es decir las partes reales e imaginarias de una solución compleja son soluciones reales.

Sea pues μ valor propio complejo y $z \in \mathbb{C}^n$ un vector propio, el cual se descompone como $z = \Re z + i\Im z,\ \Re z, \Im z \in \mathbb{R}^n$. Razonando como en el caso real se tiene que $x = e^{\mu t} z$ es solución compleja. Además, por ser A real entonces $\overline{\mu}$ es igualmente valor propio con vector propio $\overline{z}$ y la función $\overline{x} = e^{\overline{\mu} t}\overline{z}$ es solución. Si expresamos $\mu = a + bi$ tenemos

$$x = e^{\mu t} z = e^{at} (\cos(bt)\Re z - \sin(bt)\Im z) + ie^{at}(\sin(bt)\Re z + \cos(bt)\Im z).$$

En consecuencia, cada valor propio complejo y su conjugado definen las soluciones reales

$$u(t) = e^{at} (\cos(bt)\Re z - \sin(bt)\Im z), \qquad w(t) = e^{at}(\sin(bt)\Re z + \cos(bt)\Im z).$$

Finalmente, bastaría comprobar que estas soluciones son linealmente independientes. Para ello, observemos que los vectores $\Re z, \Im z$ son linealmente independientes. De lo contrario, si existiese una constante $\alpha \in \mathbb{R}$ tal que $\Im z = \alpha \Re z$, de la condición $Az = \mu z$ tendríamos

$$Az = A(1 + i\alpha)\Re z = \mu(1 + i\alpha)\Re z \Longrightarrow (A - \mu I_n)(1 + i\alpha)\Re z = 0,$$

de lo que concluimos $A\Re z = \mu \Re z$, lo cual es imposible por ser A real. Ahora la independencia lineal de las funciones u, w se concluye sin más que tomar una combinación lineal de éstas y usar la independencia lineal de $\Re z, \Im z$. Omitimos los detalles.

Con todo esto hemos probado lo siguiente: si A es diagonalizable en $\mathbb{C}$ entonces existe una base de $\mathbb{C}^n$ de la forma

$$\{v_1, ..., v_k, z_1, ..., z_r, \overline{z_1}, ..., \overline{z_r}\},$$

con $v_j \in \mathbb{R}^n$ vectores propios de valores propios reales y $z_j \in \mathbb{C}^n$ vectores propios de valores propios complejos. En consecuencia, las siguientes funciones son soluciones reales

$$x_j = e^{\lambda_j t} v_j, \qquad u_j = e^{at}\Big((\cos(bt)\Re(z_j) - \sin(bt)\Im(z_j))\Big), \quad w_j = e^{at}\Big(\sin(bt)\Re(z_j) + \cos(bt)\Im(z_j))\Big),$$

y por tanto podemos considerar la matriz solución

$$\Phi(t) = \begin{pmatrix} x_1 & \cdots & x_k & u_1 & \cdots & u_r & w_1 & \cdots & w_r \end{pmatrix}.$$

Para que esta matriz sea fundamental es suficiente probar que $\det \Phi(0) \neq 0$. Pero justo en $t = 0$ se tiene

$$\det \Phi(0) = \det(v_1, ..., v_k, \Re(z_1), ..., \Re(z_r), \Im(z_1), ..., \Im(z_r)) \neq 0.$$

En esta última igualdad se ha usado que si $\{v_1, ..., v_k, z_1, ..., z_r, \overline{z_1}, ..., \overline{z_r}\}$ es base de $\mathbb{C}^n$, entonces $\{v_1, ..., v_k, \Re(z_1), ..., \Re(z_r), \Im(z_1), ..., \Im(z_r)\}$ es base de $\mathbb{R}^n$, cuya prueba es inmediata y se deja como ejercicio.

Por último, es claro que $Av_j = \lambda_j v_j$ y por otra parte con la descomposición usual $z_j = \Re(z_j) + i\Im(z_j)$ se tiene

$$Az_j = \mu z_j = (a + bi)z_j \Longrightarrow A\Re(z_j) + iA\Im(z_j) = a\Re(z_j) - b\Im(z_j) + i(b\Re(z_j) + a\Im(z_j)).$$

Es decir,

$$A\Re(z_j) = a\Re(z_j) - b\Im(z_j), \qquad A\Im(z_j) = b\Re(z_j) + a\Im(z_j).$$

Si finalmente consideramos la base ordenada $\mathcal{B} = \{v_1, ..., v_k, \Im(z_1), \Re(z_1), ..., \Im(z_r), \Re(z_r)\}$, la matriz de A en $\mathcal{B}$ es

$$D = \begin{pmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \lambda_k & & & & & & \\ & & & \Re\mu_1 & -\Im\mu_1 & & & & \\ & & & \Im\mu_1 & \Re\mu_1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & \Re\mu_r & -\Im\mu_r \\ & & & & & & \Im\mu_r & \Re\mu_r \end{pmatrix}.$$

Observemos que D no es diagonal como tal, pero es diagonal por bloques, donde cada valor propio complejo y su conjugado generan un bloque 2×2. Si ahora definimos la matriz fundamental ordenada de la forma

$$\Phi(t) = \begin{pmatrix} x_1 & \cdots & x_k & w_1 & u_1 & \cdots & w_r & \cdots & u_r \end{pmatrix},$$

y llamamos P a la matriz cuyas columnas son los vectores de $\mathcal{B}$, concluimos como en el caso real

$$e^{tA} = \Phi(t)P^{-1}.$$

Ejemplo 3.5. Consideremos el sistema $x' = Ax$, donde

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Los valores propios son $\mu = a \pm ib = \pm i$ y una base de vectores propios en $\mathbb{C}^2$ es $\{(i,1), (-i,1)\}$. Estudiamos únicamente $\mu = i$ y por tanto tomamos $a = 0, b = 1$. Si descomponemos $z = (i,1) = (0,1) + i(1,0)$ y consideramos la base ordenada $\{\Im z, \Re z\} = \{(1,0), (0,1)\}$ se tiene que $P = I_2$. Por otra parte, las funciones

$$w = (\cos t, \sin t), \qquad u = (-\sin t, \cos t)$$

son soluciones y forman un sistema fundamental. La matriz pues

$$\Phi(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

es matriz fundamental y se concluye

$$e^{tA} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}.$$

Ejemplo 3.6. Sea ahora el sistema $x' = Ax$, donde

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 1 & 2 & -1 \end{pmatrix}.$$

El polinomio característico es

$$p(\lambda) = 1 - \lambda^3 = (1-\lambda)(\lambda^2 + \lambda + 1),$$

que tiene como raíces

$$\lambda = 1, \qquad \lambda = -\frac{1}{2} \pm i\frac{\sqrt{3}}{2}.$$

Vamos a distinguir casos real y complejo.

- Consideremos $\lambda = 1$. Como $m_a(\lambda) = 1$, necesariamente $m_g(\lambda) = 1$ y una base del subespacio propio $V_1 = \ker(A - I_3)$ es $v = (4, -1, 1)$. En consecuencia, una solución del sistema viene dada por

$$x(t) = e^t \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix}.$$

- Sea ahora $\mu = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$. En particular, $a = -1/2$ y $b = \sqrt{3}/2$. Ahora calculamos una base compleja de $\ker(A - \mu I_3)$, obteniendo $z = (-1, 1, 1/2 - i\sqrt{3}/2)$. Ahora bien, teniendo en cuenta
$$\Re z = \begin{pmatrix} -1 \\ 1 \\ \frac{1}{2} \end{pmatrix}, \qquad \Im z = \begin{pmatrix} 0 \\ 0 \\ -\frac{\sqrt{3}}{2} \end{pmatrix},$$
se tiene que las funciones
$$u = e^{-t/2}\Big(\cos\frac{\sqrt{3}t}{2}\Re z - \sin\frac{\sqrt{3}t}{2}\Im z\Big), \qquad w = e^{-t/2}\Big(\sin\frac{\sqrt{3}t}{2}\Re z + \cos\frac{\sqrt{3}t}{2}\Im z\Big),$$
son soluciones reales. Por último, la matriz en columnas
$$\Phi(t) = (x(t), w(t), u(t))$$
es matriz fundamental y definiendo la base $\mathcal{B} = \{v, \Im z, \Re z\}$ y la matriz en columnas $P = (v, \Im z, \Re z)$ concluimos
$$e^{tA} = \Phi(t)P^{-1}.$$
Finalmente, la expresión de A en la base $\mathcal{B}$ es la matriz
$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

3.4. La forma canónica de Jordan

Hemos calculado la exponencial de $A \in \mathcal{M}_n(\mathbb{R})$ y una matriz fundamental del sistema $x' = Ax$ para los casos en que A es diagonalizable tanto real como complejo, pero este está lejos de ser el escenario real. Incluso en el caso diagonalizable complejo, no está del todo claro cómo calcular la exponencial de e^{At} sin conocer un sistema fundamental de soluciones. En efecto, en el ejemplo 3.6 se tenía
$$A = PDP^{-1}, \qquad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$
Si queremos calcular e^{tA} entonces sabemos que basta calcular e^{tD}, pero la exponencial de esta matriz no es calculable usando las propiedades básicas de exponenciación, al menos de forma trivial.

El objetivo de esta sección es explicitar el cálculo efectivo de e^{tA} sea cual sea la matriz A. La idea básica es intentar expresar A de forma adecuada respecto de una base y usar las propiedades de la exponenciación matricial para simplificar el cálculo de e^{tA} al cálculo de exponenciales de matrices más sencillas, como las diagonales. Antes de nada, necesitamos definir algunas matrices que serán de importancia, y por la estructura que éstas tienen a lo largo de esta sección vamos a tomar el siguiente convenio: si en una matriz aparece alguna entrada vacía, se entiende que se completa con un cero.

Definición 3.4.1. Sean $n \in \mathbb{N}$, $\lambda \in \mathbb{R}$ y $\mu = a + bi \in \mathbb{C}$. Se definen

$$D_n(\lambda) = \begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{pmatrix}, \quad E_1(\mu) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad E_n(\mu) = \begin{pmatrix} E_1(\mu) & & & & \\ I_2 & E_1(\mu) & & & \\ & I_2 & E_1(\mu) & & \\ & & \ddots & \ddots & \\ & & & I_2 & E_1(\mu) \end{pmatrix}.$$

Se entiende que $D_1(\lambda) = \lambda$ y que cada entrada de $E_n(\mu)$ es un bloque 2×2 dado por $E_1(\mu)$, la matriz identidad I_2, o la matriz nula 0_2. Así definida, $E_n(\mu) \in \mathcal{M}_{2n}(\mathbb{R})$.

Obsérvese que $D_n(\lambda)$ se puede escribir como $D_n(\lambda) = \lambda I_n + N_n$, donde

$$N_n = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix} = (N_{ij}) = (\delta_{i,j+1}), \qquad i, j = 1, ..., n. \tag{3.4.1}$$

Igualmente, podemos expresar de forma simbólica

$$E_n(\mu) = E_1(\mu) I_n + I_2 N_n,$$

donde se debe entender que cada una de las entradas no nulas de la matriz de orden $n \times n$ se sustituye por el bloque 2×2, haciendo eventualmente la matriz resultante de orden $2n \times 2n$.

A continuación, introducimos uno de los conceptos principales alrededor del cual gira el desarrollo de esta sección.

Definición 3.4.2. Una matriz $A \in \mathcal{M}_n(\mathbb{R})$ se dice *nilpotente* si existe $k \in \mathbb{N}$ tal que $A^k = 0_n$. Al menor $k \in \mathbb{N}$ que cumple esta propiedad se le llama índice de nilpotencia y se tiene

$$A^j \neq 0_n, \quad \forall j < k, \ A^k = 0_n.$$

A la vista de cuándo se ha introducido la definición anterior, el siguiente resultado no debería sorprendernos.

Proposición 3.4.3. *La matriz N_n dada por* (3.4.1) *es nilpotente, con índice de nilpotencia n, y la llamaremos* bloque nilpotente elemental*.*

Demostración. Vamos a estudiar las potencias sucesivas de N_n. Si denotamos en coordenadas $N_n = (N_{ij}) = (\delta_{i,j+1})$ y para cada $k \in \mathbb{N}$, $N_n^k = N_{ij}^k$, entonces

$$N_{ij}^2 = \sum_{k=1}^{n} N_{ik} N_{kj} = \sum_{k=1}^{n} \delta_{i,k+1} \delta_{k,j+1}.$$

De aquí se infiere que $i = k + 1, \ k = j + 1$ para que esta suma no sea idénticamente nula, lo que implica $i = j + 2$ y en consecuencia $N_{ij}^2 = \delta_{i,j+2}$. Por una parte es claro que si $i \leq j$ entonces $N_{ij}^2 = 0$ y por tanto la estructura triangular superior se mantiene nula. Más aún, la primera diagonal inferior $N^2_{j+1,j}$ es ahora también nula y los únicos términos que aparecen son los de la segunda diagonal inferior, esto es $N^2_{j+2,j} = 1$. Lo que hemos hecho es desplazar la diagonal no nula *hacia abajo*. Por inducción es ya fácil ver que $N_n^k \neq 0_n$ para todo $k < n$, habiendo desplazado en cada potencia la diagonal no nula hacia abajo de forma que $N_{ij}^k = \delta_{i,j+k}$ y desapareciendo exactamente para $k = n$, es decir $N_n^n = 0_n$. $\square$

Estamos en condiciones de definir el concepto de matriz de Jordan.

Definición 3.4.4. Dada $J \in \mathcal{M}_n(\mathbb{R})$, diremos que J es una *matriz de Jordan* si

$$J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{pmatrix},$$

donde para cada $m = 1, ..., k$ existen $r_m \in \mathbb{N}$ y o bien $\lambda_m \in \mathbb{R}$ o bien $\mu_m \in \mathbb{C}$ tales que $J_m = D_{r_m}(\lambda_m)$ o $J_m = E_{r_m}(\mu_m)$. Llamaremos *bloques elementales de Jordan* a cada una de las submatrices J_m dadas por $D_{r_m}(\lambda)$ o $E_{r_m}(\mu)$.

Observación 3.4. Un par de comentarios.

1. En la definición anterior es posible que el mismo valor propio aparezca en distintos bloques elementales de Jordan. Por ejemplo, bastaría considerar la matriz

$$\begin{pmatrix} \lambda & & & \\ 1 & \lambda & & \\ & & \lambda & \\ & & 1 & \lambda \end{pmatrix}$$

para concluir que tiene dos bloques de Jordan de la forma $D_2(\lambda)$.

2. En el importante caso 2×2, las matrices de Jordan han de ser de los siguientes tipos:

$$\begin{pmatrix} D_1(\lambda) & 0 \\ 0 & D_1(\theta) \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \theta \end{pmatrix}, \qquad D_2(\lambda) = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}, \qquad E_1(a+bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

con $\lambda, \theta, a, b \in \mathbb{R}$. En el primer caso podría ser $\lambda = \theta$.

El siguiente resultado es uno de los resultados centrales del Álgebra Lineal y nos asegura que toda matriz es semajante a una matriz de Jordan.

Teorema 3.4.5. *Toda matriz $A \in \mathcal{M}_n(\mathbb{R})$ es semejante a una matriz de Jordan, J, llamada* forma canónica de Jordan de *A, la cual es única salvo reordenación de los bloques elementales que la conforman. Concretamente, si $\lambda_1, ..., \lambda_k \in \mathbb{R}$ son los valores propios reales y $\mu_1, ..., \mu_l, \overline{\mu}_1, ..., \overline{\mu}_l \in \mathbb{C}$ los valores propios complejos de A, entonces*

$$J = \begin{pmatrix} J(\lambda_1) & & & & & \\ & \ddots & & & & \\ & & J(\lambda_k) & & & \\ & & & J(\mu_1) & & \\ & & & & \ddots & \\ & & & & & J(\mu_l) \end{pmatrix},$$

donde

$$J(\lambda_j) = \begin{pmatrix} D_{r_1}(\lambda_j) & & \\ & \ddots & \\ & & D_{r_{n(\lambda_j)}}(\lambda_j) \end{pmatrix}, \qquad J(\mu_j) = \begin{pmatrix} E_{r_1}(\mu_j) & & \\ & \ddots & \\ & & E_{r_{n(\mu_j)}}(\mu_j) \end{pmatrix}.$$

Más aún, si δ es valor propio de A, real o complejo, y denotamos por $m_a(\delta)$ y $m_g(\delta)$ a las multiplicidades algebraica y geométrica, respectivamente, se tiene

$$m_a(\delta) = r_1 + \cdots + r_{n(\delta)}, \qquad m_g(\delta) = n(\delta).$$

Es decir, el número de bloques de Jordan para un valor propio coincide con su multiplicidad geométrica, y la suma de los órdenes de los bloques de Jordan es la multiplicidad algebraica.

Observación 3.5. Como A es real, si μ es valor propio de A entonces necesariamente $\overline{\mu}$ también lo será. No obstante, en el Teorema 3.4.5 únicamente se hace mención a una mitad de los valores propios complejos, sin referirnos a sus conjugados. En lo que sigue, siempre que tengamos que calcular algún bloque elemental de Jordan para algún valor propio complejo μ, supondremos $\Im\mu > 0$.

3.4.1. La exponencial de un bloque elemental de Jordan

Como quiera que

$$J = J(\lambda_1) \oplus \cdots J(\lambda_k) \oplus J(\mu_1) \oplus \cdots \oplus J(\mu_l),$$

y en vista de las propiedades *4.* y *5.* del Teorema 3.3.5, para el cálculo de e^{tA} bastará expresar A en su forma canónica de Jordan y calcular la exponencial de cada uno de los bloques elementales de Jordan, $tD_{r_j}(\lambda_j)$ o $tE_{r_j}(\mu_j)$. El siguiente resultado muestra la sencillez del cálculo de la matriz exponencial de un bloque de Jordan.

Teorema 3.4.6. *Con la notación anterior, se tienen las siguientes propiedades.*

1. *La exponencial de la matriz $tD_n(\lambda)$ viene dada por*

$$e^{tD_n(\lambda)} = e^{t\lambda} \begin{pmatrix} 1 & & & & \\ t & 1 & & & \\ \frac{t^2}{2} & t & 1 & & \\ \vdots & \ddots & \ddots & \ddots & \\ \frac{t^{n-1}}{(n-1)!} & \cdots & \frac{t^2}{2} & t & 1 \end{pmatrix}.$$

2. *La exponencial de la matriz $tE_n(\mu)$, con $\mu = a + bi$, viene dada por*

$$e^{tE_n(\mu)} = e^{ta} \begin{pmatrix} \cos(tb) & -\sin(tb) \\ \sin(tb) & \cos(tb) \end{pmatrix} \begin{pmatrix} I_2 & & & & \\ tI_2 & I_2 & & & \\ \frac{t^2}{2} I_2 & tI_2 & I_2 & & \\ \vdots & \ddots & \ddots & \ddots & \\ \frac{t^{n-1}}{(n-1)!} I_2 & \cdots & \frac{t^2}{2} I_2 & tI_2 & I_2 \end{pmatrix},$$

donde el producto de las matrices se debe entender en el sentido de que la primera se debe multiplicar por cada una de las cajas 2×2 que conforman la segunda.

Demostración. Para verificar *1.*, usamos que $D_n(\lambda) = \lambda I_n + N_n$. Ya que $t\lambda I_n$ y tN_n conmutan, se tiene

$$e^{tD_n(\lambda)} = e^{t\lambda I_n + tN_n} = e^{t\lambda I_n} e^{tN_n} = e^{t\lambda} I_n e^{tN_n},$$

y basta evaluar la exponencial de la matriz e^{tN_n}. Como hemos probado que N_n es nilpotente también es inmediato que tN_n lo es y en consecuencia la serie que define la exponencial de tN_n es finita. Habiendo deducido igualmente la estructura triangular de N_n^k para todo $k \leq n$, se tiene

$$e^{tN_n} = \sum_{k=0}^{n-1} t^k \frac{N_n^k}{k!} = \begin{pmatrix} 1 & & & & \\ t & 1 & & & \\ \frac{t^2}{2} & t & 1 & & \\ \vdots & \ddots & \ddots & \ddots & \\ \frac{t^{n-1}}{(n-1)!} & \cdots & \frac{t^2}{2} & t & 1 \end{pmatrix}.$$

De aquí concluimos *1*.

La comprobación de *2*. es algo más laboriosa. Comenzamos con el caso $n = 1$ y para ello consideramos el subconjunto de $\mathcal{M}_2(\mathbb{R})$ de las matrices de la forma

$$\mathcal{M}_2(a, b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \qquad a, b \in \mathbb{R},$$

el cual es un subespacio vectorial de $\mathcal{M}_2(\mathbb{R})$ y es cerrado para el producto de matrices; de hecho tiene estructura de cuerpo puesto que la única matriz en $\mathcal{M}_2(a, b)$ con determinante nulo es la matriz nula. Definimos la biyección natural entre $\mathcal{M}_2(a, b)$ y el cuerpo de los números complejos $\mathbb{C}$ dada por

$$\Psi : \mathcal{M}_2(a, b) \to \mathbb{C}, \qquad \Psi \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a + bi.$$

Esta aplicación es un isomorfismo de cuerpos, donde el producto en $\mathcal{M}_2(a, b)$ es el producto usual de matrices y en $\mathbb{C}$ es el producto complejo. De hecho, su inversa es

$$E_1(a + bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

introducida en la Definición 3.4.4. Por ser $\mathcal{M}_2(a, b)$ un subespacio vectorial de $\mathcal{M}_2(\mathbb{R})$ en particular es cerrado[1] y por tanto contiene los límites de sus sucesiones. En particular, $e^{tE_1(\mu)} \in \mathcal{M}_2(a, b)$ y podemos aplicarle Ψ sin ningún tipo de problema. Usando que Ψ es continua y homomorfismo de

[1]Esto es porque todo subespacio vectorial de un espacio vectorial es el núcleo de una aplicación lineal. Como toda aplicación lineal entre espacios vectoriales finito dimensionales es continua, su núcleo es siempre cerrado.

cuerpos se tiene

$$\begin{aligned}
\Psi(e^{tE_1(a+bi)}) &= \Psi\left(\sum_{k=0}^{\infty}\frac{t^k}{k!}E_1(a+bi)^k\right)\\
&= \Psi\left(\lim_{m\to\infty}\sum_{k=0}^{m}\frac{t^k}{k!}E_1(a+bi)^k\right)\\
&= \lim_{m\to\infty}\Psi\left(\sum_{k=0}^{m}\frac{t^k}{k!}E_1(a+bi)^k\right)\\
&= \lim_{m\to\infty}\sum_{k=0}^{m}\Psi\left(\frac{t^k}{k!}E_1(a+bi)^k\right)\\
&= \lim_{m\to\infty}\sum_{k=0}^{m}\frac{t^k}{k!}\Psi\left(E_1(a+bi)\right)^k\\
&= \sum_{k=0}^{\infty}\frac{t^k}{k!}\Psi\left(E_1(a+bi)\right)^k\\
&= e^{t\Psi(E_1(a+bi))} = e^{t(a+bi)} = e^{ta}(\cos(tb)+i\sin(tb)).
\end{aligned}$$

La inyectividad de Ψ asegura

$$e^{tE_1(\mu)} = e^{ta}\begin{pmatrix}\cos(tb) & -\sin(tb)\\ \sin(tb) & \cos(tb)\end{pmatrix}.$$

Para el caso general, descomponemos $E_n(\mu) = E_1(\mu)I_n + E_n(0)$ donde $E_n(0) = I_2N_n$ es la matriz $2n\times 2n$ por bloques

$$E_n(0) = \begin{pmatrix}0_2 & & & & \\ I_2 & 0_2 & & & \\ & I_2 & 0_2 & & \\ & & \ddots & \ddots & \\ & & & I_2 & 0_2\end{pmatrix}.$$

De nuevo $tE_1(\mu)I_n$ y $tE_n(0)$ conmutan y en consecuencia

$$\begin{aligned}
e^{tE_n(\mu)} &= e^{tE_1(\mu)I_n}e^{tE_n(0)}\\
&= \begin{pmatrix}e^{tE_1(\mu)} & & & \\ & e^{tE_1(\mu)} & & \\ & & \ddots & \\ & & & e^{tE_1(\mu)}\end{pmatrix}\sum_{k=0}^{\infty}\frac{t^k}{k!}E_n(0)^k\\
&= e^{ta}\begin{pmatrix}\cos(tb) & -\sin(tb)\\ \sin(tb) & \cos(tb)\end{pmatrix}\sum_{k=0}^{\infty}\frac{t^k}{k!}E_n(0)^k
\end{aligned}$$

La prueba finaliza tras comprobar por inducción que si $1\le k<n$ con $k\in\mathbb{N}$ entonces las entradas de $E_n(0)^k$ son los bloques identidad I_2 desplazados a los bloques diagonales $k-1$-ésimos inferiores. En consecuencia, $E_n(0)^n = 0_{2n}$, la matriz $E_n(0)$ es nilpotente y su estructura y la de sus potencias es similar a N_n definida en (3.4.1). $\square$

Finalmente, enunciamos el siguiente resultado acerca de la estructura de las soluciones de un sistema lineal homogéneo con coeficientes constantes. Su demostración es consecuencia del Teorema 3.4.5, que asegura que toda matriz es semejante a su forma canónica de Jordan, y del Teorema 3.4.6 mediante el cual se calcula la exponencial de cada uno de los bloques elementales de Jordan.

Corolario 3.4.7. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ y x una solución del sistema homogéneo $x' = Ax$. Entonces, cada una de las coordenadas de x es una combinación lineal de funciones de la forma*

$$t^k e^{at}\cos(bt), \qquad t^k e^{at}\sin(bt),$$

donde $a + bi$ recorre el conjunto de valores propios de A y $0 \leq k < n$ (admitimos $b = 0$ para contemplar el caso real). Más aún, para cada valor propio $a+bi$, el exponente k debe ser menor que r, siendo r el tamaño del mayor bloque elemental de Jordan $D_r(a)$ ó $E_r(a+bi)$ correspondiente a $a+bi$ en la forma canónica de Jordan de A.

3.5. Cálculo de la forma canónica de Jordan

Dada $A \in \mathcal{M}_n(\mathbb{R})$, el camino que hemos seguir para el cálculo de e^{tA} está bastante claro: si identificamos A con el endomorfismo $A : \mathbb{R}^n \to \mathbb{R}^n$ tal que su matriz en la base canónica de $\mathbb{R}^n$ es precisamente la matriz A, debemos encontrar una base $\mathcal{B}$ tal que la matriz del endomorfismo A en $\mathcal{B}$ sea su forma canónica de Jordan, J. Llamando P a la matriz del cambio de base de $\mathcal{B}$ a la base canónica, entonces $A = PJP^{-1}$ y por la propiedad *4.* del Teorema 3.3.5 se tendrá

$$e^{tA} = Pe^{tJ}P^{-1}.$$

Por otra parte, si $J = J_1 \oplus \cdots \oplus J_r$ entonces la propiedad *5.* del Teorema 3.3.5 nos permite expresar e^{tJ} como

$$e^{tJ} = \mathrm{diag}\{e^{tJ_1}, ..., e^{tJ_r}\},$$

donde cada una de las exponenciales e^{tJ_k} son conocidas y calculables en virtud del Teorema 3.4.6. Finalmente, recordamos que a la hora de calcular la matriz fundamental

$$\Phi(t) = e^{tA} = Pe^{tJ}P^{-1},$$

no hace falta calcular la inversa P^{-1}. Esto es porque si $\Phi(t)$ es matriz fundamental, entonces $\Phi(t)C$ sigue siendo matriz fundamental para cualquier matriz C con $\det C \neq 0$. En particular, se puede tomar como matriz fundamental

$$\Phi(t) = e^{tA} = Pe^{tJ}$$

y obviar el (casi siempre) tedioso cálculo de P^{-1}.

3.5.1. Valor propio real

Supongamos $\lambda \in \mathbb{R}$ valor propio de A y sean $m_a(\lambda), m_g(\lambda)$ las multiplicidades aritmética y geométrica de λ. El objetivo es encontrar vectores linealmente independientes $\{v_1, ..., v_{m_a(\lambda)}\}$ tales que al completar a una base $\mathcal{B}$ de todo $\mathbb{R}^n$, la matriz de A en $\mathcal{B}$ sea de la forma

$$\begin{pmatrix} J(\lambda) & \\ & A' \end{pmatrix},$$

siendo A' una matriz cuadrada de orden $n - m_a(\lambda)$ que tiene a todos los demás valores propios de A. Sabemos por el Teorema de Jordan que $J(\lambda)$ está formada por $m_g(\lambda)$ cajas de la forma $D_r(\lambda)$, es decir

$$J(\lambda) = \begin{pmatrix} D_{r_1}(\lambda) & & \\ & \ddots & \\ & & D_{r_{m_g(\lambda)}}(\lambda) \end{pmatrix}, \tag{3.5.1}$$

con $r_1 + \cdots + r_{m_g(\lambda)} = m_a(\lambda)$. La construcción de esta base se basa en el estudio del operador $A - \lambda I_n$. La discusión que sigue es más propia de un curso de Álgebra Lineal, pero su inclusión en estas notas se cree importante y necesaria para justificar el algoritmo del cálculo de la base $\mathcal{B}$.

Consideremos la siguiente sucesión de espacios vectoriales de $\mathbb{R}^n$,

$$\{0\} \subset \ker(A - \lambda I_n) \subset \ker(A - \lambda I_n)^2 \subset \cdots \subset \ker(A - \lambda I_n)^k \subset \cdots \tag{3.5.2}$$

Como $\mathbb{R}^n$ es finito dimensional, la torre anterior no puede crecer indefinidamente y por tanto existe un menor $p \in \mathbb{N}$ tal que $\ker(A - \lambda I_n)^p = \ker(A - \lambda I_n)^{p+1}$. De hecho, una vez que la torre se estabiliza en una etapa esta no vuelve a crecer. En efecto, supongamos que se tiene

$$\ker(A - \lambda I_n)^p = \ker(A - \lambda I_n)^{p+1} \subset \ker(A - \lambda I_n)^{p+2},$$

y sea $v \in \ker(A - \lambda I_n)^{p+2}$. Entonces, $0 = (A - \lambda I_n)^{p+2} v = (A - \lambda I_n)^{p+1}(A - \lambda I_n)v$, en consecuencia $(A - \lambda I_n)v \in \ker(A - \lambda I_n)^{p+1}$ y por hipótesis $(A - \lambda I_n)v \in \ker(A - \lambda I_n)^p$, esto es $v \in \ker(A - \lambda I_n)^{p+1} = \ker(A - \lambda I_n)^p$.

Definición 3.5.1. Sea $p \in \mathbb{N}$ el valor mínimo para el cual se estabiliza la torre (3.5.2). Se define el *subespacio propio generalizado correspondiente a* λ como el subsepacio vectorial de $\mathbb{R}^n$

$$E(A, \lambda) = \ker(A - \lambda I_n)^p.$$

En particular, el operador $A - \lambda I_n$ restringido a $E(A, \lambda)$ es nilpotente con índice de nilpotencia exactamente p.

Una de las propiedades clave en la demostración del Teorema de Jordan es la descomposición que admite $E(A, \lambda)$ como suma directa de subespacios muy sencillos de tratar, que introducimos a continuación.

Definición 3.5.2. Sea V espacio vectorial y $N : V \to V$ aplicación nilpotente con índice de nilpotencia p. Un subespacio W de V se dice *cíclico* para N si existe $v \in W$ tal que W está generado por los vectores $N^k v,\ k = 0, ..., p$. En tal caso, denotaremos a W por $Z(v)$.

Nótese que la longitud de $N^k v$ es a lo sumo p, puesto que $N^p = 0$, y que en cualquier caso depende de v. Por ejemplo, si $v \in \ker N$ entonces $Nv = 0$ y $Z(v) = \mathrm{span}(v)$. Por tanto, para cada $v \in V$ existe un entero minimal denotado por $\mathrm{nil}(v)$ tal que $N^{\mathrm{nil}(v)} v = 0$ y $N^k v \neq 0$ para $k < \mathrm{nil}(v)$. Los subespacios cíclicos para una aplicación nilpotente se describen a partir de la siguiente proposición.

Proposición 3.5.3. *Sea $N : V \to V$ una aplicación nilpotente y $v \in V$. Entonces,* $\dim Z(v) = \mathrm{nil}(v)$ *y*

$$\{v, Nv, ..., N^{\mathrm{nil}(v)-1} v\}$$

es una base de $Z(v)$.

Demostración. Que los vectores $v, Nv, ..., N^{\mathrm{nil}(v)-1}v$ generan $Z(v)$ es trivial por la propia definición de $Z(v)$. Si fuesen linealmente dependientes existirían $\alpha_0, ..., \alpha_{\mathrm{nil}(v)-1}$ no todos nulos tales que

$$\sum_{k=0}^{\mathrm{nil}(v)-1} \alpha_k N^k v = 0.$$

Sea j el menor índice tal que $\alpha_j \neq 0$. Entonces,

$$0 = N^{\mathrm{nil}(v)-j-1} \sum_{k=0}^{\mathrm{nil}(v)-1} \alpha_k N^k v = \alpha_j N^{\mathrm{nil}(v)-1} v,$$

lo cual es contradictorio. □

Notemos que si restringimos N a $Z(v)$, la representación de N en la base $\{v, Nv, ..., N^{\mathrm{nil}(v)-1}v\}$ viene dada por

$$\begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}, \tag{3.5.3}$$

y en particular $N^{\mathrm{nil}(v)-1}v \in \ker N$. Estos son los unos que aparecerán bajo la diagonal principal en la forma canónica de Jordan.

El siguiente resultado es uno de los pilares en la demostración del Teorema de Jordan.

Lema 3.5.4. *Sea $N : V \to V$ una aplicación lineal nilpotente. Entonces existen $v_1, ..., v_r \in V$ tal que*

$$V = Z(v_1) \oplus \cdots \oplus Z(v_r).$$

Además, tal descomposición es esencialmente única, es decir si existen $w_1, ..., w_s$ tales que

$$V = Z(w_1) \oplus \cdots \oplus Z(w_s),$$

entonces $r = s$ y los vectores w_k se pueden reordenar de forma que $\mathrm{nil}(v_k) = \mathrm{nil}(w_k)$ *para todo* $k = 1, ..., r$.

Demostración. Primero observemos que podemos suponer $p > 1$, porque de lo contrario $Nv = 0$ para todo $v \in V$ y N sería el operador nulo. Bastaría entonces tomar $\{v_1, ..., v_n\}$ una base de V, definir $Z(v_k) = \mathrm{span}(v_k)$ y se tendría trivialmente $V = Z(v_1) \oplus \cdots \oplus Z(v_n)$. Supondremos pues $p > 1$.

Por ser N nilpotente, existe $p > 1$ su índice de nilpotencia tal que $N^p v = 0$ para todo $v \in V$. Ahora bien, por ser p el índice de nilpotencia existe $v_1 \in V$ tal que $N^{p-1}v_1 \neq 0$. En tal caso, la Proposición 3.5.3 asegura que el conjunto $\{v_1, Nv_1, ..., N^{p-1}v_1\}$ es un conjunto linealmente independiente; más aún, es una base del subespacio cíclico $V_1 = Z(v_1)$. Vamos a probar el siguiente hecho: V_1 *tiene un complemento N-invariante en V.*

Supongamos que lo tenemos probado y llamemos W_1 a tal complemento N-invariante de V_1, es decir $V = V_1 \oplus W_1 = Z(v_1) \oplus W_1$. En este punto, la restricción $N_{|W_1}$ es nilpotente con índice de nilpotencia $p_1 \leq p$. En particular, existe $v_2 \in W_2$ tal que $N^{p_1-1}v_2 \neq 0$ y argumentando como en el párrafo anterior el subespacio $V_2 = Z(v_2)$ es cíclico y tendrá un complemento N invariante en W_1, llamémoslo W_2. Por tanto, $V = V_1 \oplus V_2 \oplus W_2 = Z(v_1) \oplus Z(v_2) \oplus W_2$. Como V es finito-dimensional, repitiendo este argumento un número finito de pasos llegaremos a obtener una descomposición de V como suma finita de subespacios cíclicos, como queremos probar.

Sea $w \in W$ arbitrario y supongamos, por reducción al absurdo, que $Nw \in Z(v_1)$. Si $Nw = 0 \in W$ hemos terminado, así que supongamos $Nw \neq 0$. Existen por tanto $\alpha_1, ..., \alpha_n$ no todos nulos tales que

$$Nw = \alpha_1 v_1 + \alpha_2 N v_1 + \cdots + \alpha_p N^{p-1} v_1.$$

Afirmamos que $\alpha_1 = 0$. De lo contrario, aplicando a la ecuación anterior N^{p-1} a ambos lados se tiene

$$N^p w = \alpha_1 N^{p-1} v_1 \neq 0,$$

y por tanto el índice de nilpotencia de N sería mayor que p, contradicción. En consecuencia,

$$Nw = \alpha_2 N v_1 + \cdots + \alpha_p N^{p-1} v_1 = N(\alpha_2 v_1 + \alpha_3 N v_1 + \cdots + \alpha_p N^{p-2} v_1)$$

Como se suponía $p > 1$ el último miembro $N^{p-2}v_1$ está bien definido. Llegados a este punto, recordamos el siguiente resultado que es propio de un primer curso de Álgebra Lineal y cuya demostración pondremos por hacer estas notas lo más autocontenidas posibles:

Si $W \subset V$ es subespacio vectorial y $T : V \to V$ es aplicación lineal, entonces $T^{-1}(T(W)) = W + \ker T$.

> Por ser W subespacio vectorial, $T(W)$ es también subespacio y de nuevo $T^{-1}(T(W))$ es subespacio; todas estas conclusiones son triviales. Ahora bien, si W es subespacio de V, entonces $T^{-1}(T(W))$ contiene a W y también contiene a $\ker T$. En efecto, dado $w \in W$, si llamamos $v = Tw$ se tiene $w \in T^{-1}(v)$ y por tanto $W \subset T^{-1}(T(W))$. Que contiene a $\ker T$ se prueba de forma similar. Como $T(W)$ es subespacio vectorial en particular $0 \in T(W)$ y por tanto $\ker T = T^{-1}(0)$; en cualquier caso, $\ker T \subset T^{-1}(T(W))$ y por ser $T^{-1}(T(W))$ subsepacio vectorial se tiene $W + \ker T \subset T^{-1}(T(W))$. La implicación contraria se prueba de forma similar. Dado $u \in T^{-1}(T(W))$, existe $v \in T(W)$ tal que $Tu = v$. Pero $v \in T(W)$ significa a su vez que existe $w \in W$ tal que $v = Tw$ y concatenando igualdades concluimos $Tu = Tw$. Por linealidad $T(u - w) = 0$ y se tiene $u - w \in \ker T$. Ahora es claro que podemos escribir $u = w + (u - w)$, donde el primero pertenece a W y el segundo a $\ker T$, completando la prueba.

Volvamos a la prueba de nuestro lema, y en concreto al estudio de la igualdad

$$Nw = N(\alpha_2 v_1 + \alpha_3 N v_1 + \cdots + \alpha_p N^{p-2} v_1).$$

Si llamamos $v = \alpha_2 v_1 + \alpha_3 N v_1 + \cdots + \alpha_p N^{p-2} v_1 \in Z(v_1)$, se tiene $v \neq 0$ (porque se suponía $Nw \neq 0$), $Nw = Nv$ y en consecuencia $w = N^{-1}(Nv)$, por lo que $w \in N^{-1}(N(Z(v_1)) = Z(v_1) + \ker N$. Como se suponía $Nw \neq 0$ se tiene $w \notin \ker N$ y necesariamente $w \in Z(v_1)$, lo cual es contradictorio. Esto nos dice que $N(W) \subset W$ y W es invariante por N. Repitiendo este proceso obtenemos la descomposición deseada,

$$V = Z(v_1) \oplus \cdots \oplus Z(v_r).$$

Que la descomposición es única salvo reordenación es consecuencia de que esta descomposición está determinada por el índice de nilpotencia de N cuando se restringe a cada subespacio invariante, y por tanto sólo depende del vector v_k con $\mathrm{nil}(v_k)$ igual al índice de nilpotencia de N. $\square$

Vamos a aplicar esta discusión al operador $A - \lambda I_n$, el cual es trivialmente nilpotente una vez restringido a $E(A,\lambda)$. En concreto, entender el siguiente resultado nos permitirá extraer las ideas principales para el cálculo de una base adecuada respecto de la cual la matriz de A una vez restringida a $E(A,\lambda)$ sea de Jordan.

Proposición 3.5.5. *Existen $v_1, ..., v_r \in E(A,\lambda)$ tales que*

$$E(A,\lambda) = Z(v_1) \oplus \cdots \oplus Z(v_r).$$

Además, dada $\{v_k, ..., (A-\lambda I_n)^{\mathrm{nil}(v_k)-1} v_k\}$ una base de $Z(v_k)$, el conjunto

$$\mathcal{B} = \{v_1, ..., (A-\lambda I_n)^{\mathrm{nil}(v_1)-1} v_1, ..., v_r, ..., (A-\lambda I_n)^{\mathrm{nil}(v_r)-1} v_r\},$$

es una base de $E(A,\lambda)$. Más aún:

1. *El subespacio $E(A,\lambda)$ es invariante por A, esto es $A(E(A,\lambda)) \subset E(A,\lambda)$. En particular, podemos considerar la restricción $A : E(A,\lambda) \to E(A,\lambda)$.*

2. *Si denotamos por P a la matriz cuyas columnas son los vectores de la base $\mathcal{B}$, se tiene $A_{|_{E(A,\lambda)}} = PJP^{-1}$, donde J viene dada por los bloques diagonales*

$$J(\lambda) = \mathrm{diag}\{D_{\mathrm{nil}(v_1)}(\lambda), ..., D_{\mathrm{nil}(v_r)}(\lambda)\},$$

y cada $D_{\mathrm{nil}(v_k)}(\lambda)$ es un bloque elemental de Jordan definido en 3.4.1. Matricialmente,

$$J(\lambda) = \begin{pmatrix} \underbrace{\begin{matrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{matrix}}_{\mathrm{nil}(v_1)\times\mathrm{nil}(v_1)} & & \\ & \ddots & \\ & & \underbrace{\begin{matrix} \lambda & & & \\ 1 & \lambda & & \\ & \ddots & \ddots & \\ & & 1 & \lambda \end{matrix}}_{\mathrm{nil}(v_r)\times\mathrm{nil}(v_r)} \end{pmatrix}.$$

3. *Si $m_g(\lambda)$ denota la multiplicidad geométrica de λ, se tiene $m_g(\lambda) = r$. Si $m_a(\lambda)$ denota la multiplicidad algebraica de λ, se tiene $m_a(\lambda) = \mathrm{nil}(v_1) + \cdots \mathrm{nil}(v_r)$.*

4. *El índice de nilpotencia de $A - \lambda I_n$ restringido a $E(A,\lambda)$ es $n_\lambda = \mathrm{máx}\{\mathrm{nil}(v_k)\}$.*

5. *La diferencia* $\dim\ker(A-\lambda I_n)^k - \dim\ker(A-\lambda I_n)^{k-1}$ *coincide con el número de bloques elementales de Jordan de orden al menos* k*. En consecuencia, el número de bloques de Jordan de orden exactamente* k *es*

$$2\dim\ker(A-\lambda I_n)^k - \dim\ker(A-\lambda I_n)^{k+1} - \dim\ker(A-\lambda I_n)^{k-1}.$$

Demostración. 1. Inmediato porque A conmuta con cada potencia $(A-\lambda I_n)^k$.

2. Consideramos la base $\{v_k, ..., (A-\lambda I_n)^{\mathrm{nil}(v_k)-1}v_k\}$ de cada espacio cíclico $Z(v_k)$ y por comodidad la enumeramos como $\{v_1, ...v_{\mathrm{nil}(v_k)}\}$, de donde se tiene $v_{j+1} = (A-\lambda I_n)v_j$ para $j = 1, ..., \mathrm{nil}(v_k) - 1$. Entonces,

$$Av_j = (A-\lambda I_n)v_j + \lambda v_j = v_{j+1} + \lambda v_j, \quad j = 1, ..., \mathrm{nil}(v_k) - 1, \quad Av_{\mathrm{nil}(v_k)} = \lambda v_{\mathrm{nil}(v_k)},$$

y la representación de A una vez restringida a $Z(v_k)$ en la base $\{v_k, ..., (A-\lambda I_n)^{\mathrm{nil}(v_k)-1}v_k\}$ es $D_{\mathrm{nil}(v_k)}(\lambda)$. Como $E(A,\lambda)$ es suma directa de los subespacios cíclicos, la representación de A en $\mathcal{B}$ es la concatenación diagonal de las representaciones matriciales de A en cada $Z(v_k)$.

3. Llamemos $d = \dim E(A,\lambda)$, descompongamos $\mathbb{R}^n = E(A,\lambda) \oplus E(A,\lambda)^\perp$ y sea $\mathcal{B}'$ una base de $\mathbb{R}^n$ obtenida al concatenar $\mathcal{B}$ con una base $\mathcal{B}_0$ de $E(A,\lambda)^\perp$. La matriz de A en $\mathcal{B}'$ viene dada por los bloques cuadrados diagonales

$$\begin{pmatrix} J(\lambda) & B \\ & C \end{pmatrix},$$

donde la matriz B no nos interesa y $C \in \mathcal{M}_{n-d}(\mathbb{R})$. El polinomio característico de A es $p(x) = \det(A - xI_n) = \det(J(\lambda) - xI_d)\det(C - xI_{n-d}) = (\lambda - x)^d q(x)$, para cierto polinomio $q(x)$. De aquí se deduce $m_a(\lambda) \geq d$. Supongamos $m_a(\lambda) > d$ y por tanto $q(\lambda) = 0$. Esto nos dice que λ es valor propio de C y existe $v \in E(A,\lambda)^\perp$ no nulo tal que $Cv = \lambda v$. Ahora bien, definimos el vector $\tilde{v} = (0, \overset{d)}{\ldots}, 0, v) \in \mathbb{R}^n$, el cual es no nulo y satisface $A\tilde{v} = Cv = \lambda v = \lambda\tilde{v}$, y por tanto $\tilde{v}$ es vector propio de A para el valor propio λ. En consecuencia, $\tilde{v} \in \ker(A - \lambda I_n) \subset E(A,\lambda)$, lo cual es imposible ya que $v \notin E(A,\lambda)$ y se tiene $m_a(\lambda) = d$. Como la dimensión de $E(A,\lambda)$ coincide con la suma $\mathrm{nil}(v_1) + \cdots \mathrm{nil}(v_r)$ de los subespacios cíclicos que lo generan, se tiene $m_a(\lambda) = \mathrm{nil}(v_1) + \cdots \mathrm{nil}(v_r)$. Además, $m_g(\lambda) = \dim\ker(A - \lambda I_n)$ coincide exactamente con $\dim\ker(J(\lambda) - I_{m_a(\lambda)})$ de nuevo por no poder ser λ valor propio de C. Como en $J(\lambda) - \lambda I_{m_a(\lambda)}$ existen exactamente r columnas nulas, una por cada bloque elemental de Jordan, concluimos $m_g(\lambda) = r$.

4. Observemos que la matriz de $A - \lambda I_n$ en $\mathcal{B}$ está formada por r bloques nilpotentes elementales de la forma (3.5.3). Cada potencia de $A - \lambda I_n$ baja en una las subdiagonales de cada uno de estos bloques elementales. El último en quedar nulo será el de mayor orden, y este será precisamente el índice de nilpotencia de $A - \lambda I_n$.

5. La demostración de este resultado no radica en su dificultad, sino en entender (e intentar transmitir) la estructura del operador $A - \lambda I_n$ y sus potencias. Recordemos que $(A - \lambda I_n)^{n_\lambda}$ es el operador nulo una vez restringido a $E(A,\lambda)$ y $(A - \lambda I_n)^k \neq 0$ para $k < n_\lambda$. Como la representación de $A - \lambda I_n$ en $\mathcal{B}$ es una diagonal por bloques, siendo cada uno un bloque nilpotente elemental (3.5.3), cada potencia de $A - \lambda I_n$ baja una unidad la subdiagonal de cada

uno de estos bloques nilpotentes elementales. Esto se repite hasta llegar a la potencia $n_\lambda - 1$, donde todos los bloques de Jordan de órdenes menores que n_λ han desaparecido y sólo quedan los bloques de orden máximo n_λ. Finalmente, en la potencia n_λ desaparecen exactamente los bloques de orden máximo.

Comenzamos estudiando la diferencia de dimensiones de $\ker(A-\lambda I_n)^{n_\lambda}$ y $\ker(A-\lambda I_n)^{n_\lambda-1}$ y para ello recordemos que $(A-\lambda I_n)^{n_\lambda}$ es el operador nulo y que por tanto la representación matricial de $(A-\lambda I_n)^{n_\lambda-1}$ consta de todos los bloques de Jordan de órdenes menores que n_λ siendo ya nulos, y quedando exactamente un único valor 1 en la posición $j_{n_\lambda,1}$ en cada uno de los bloques de Jordan de orden máximo n_λ. En consecuencia, la diferencia $\dim\ker(A-\lambda I_n)^{n_\lambda} - \dim\ker(A-\lambda I_n)^{n_\lambda-1}$ proviene exactamente de descender una unidad el último elemento de los bloques de Jordan de orden n_λ, haciéndolos desaparecer y en consecuencia el número de bloques de Jordan de orden máximo n_λ es esta diferencia de dimensiones.

Sea ahora $k < n_\lambda$ y estudiemos la potencia $(A-\lambda I_n)^{k-1}$. Aquí los bloques de Jordan de órdenes menores que k ya han desaparecido y sólo quedan bloques de órdenes al menos k. En consecuencia, la diferencia de dimensiones $\dim\ker(A-\lambda I_n)^k - \dim\ker(A-\lambda I_n)^{k-1}$ proviene de hacer descender una unidad la subdiagonal de los bloques de Jordan de orden al menos k y por tanto el número de bloques de Jordan de orden al menos k es $\dim\ker(A-\lambda I_n)^k - \dim\ker(A-\lambda I_n)^{k-1}$.

Finalmente, como el número de bloques elementales de Jordan de orden exactamente k se calcula como el número de orden al menos k menos el número de orden al menos $k+1$, concluimos que el número de bloques de orden exactamente k es

$$2\dim\ker(A-\lambda I_n)^k - \dim\ker(A-\lambda I_n)^{k+1} - \dim\ker(A-\lambda I_n)^{k-1}.$$

$\square$

A la vista de la Proposición 3.5.5, los pasos a seguir para calcular la base $\mathcal{B}$ del subespacio generalizado deberían estar ahora claros. Partimos del subespacio generalizado

$$E(A,\lambda) = \ker(A-\lambda I_n)^{n_\lambda},$$

donde el exponente n_λ lo podemos deducir del hecho de que $\dim E(A,\lambda) = m_a(\lambda)$, o bien de que $E(A,\lambda)$ se estabiliza por primera vez en la potencia n_λ. Sea $d_1 = \ker(A-\lambda I_n)^{n_\lambda} - \ker(A-\lambda I_n)^{n_\lambda-1}$, el cual nos dice el número de bloques de Jordan de orden máximo n_λ. Tomamos $v_1, ...v_{d_1}$ vectores linealmente independientes en $\ker(A-\lambda I_n)^{n_\lambda}\setminus\ker(A-\lambda I_n)^{n_\lambda-1}$ y a cada uno de estos vectores le aplicamos el operador $A-\lambda I_n$, obteniendo d_1 vectores en $\ker(A-\lambda I_n)^{n_\lambda-1}$. Estos vectores son linealmente independientes entre sí en virtud del Lema 3.5.4 y de la Proposición 3.5.5. Repetimos este proceso y aplicamos $A-\lambda I_n$ un total de $n_\lambda - 1$ veces a cada $v_j \in \ker(A-\lambda I_n)^{n_\lambda}$, obteniendo.

$$\begin{array}{ccccc}
\ker(A-\lambda I_n) & \ker(A-\lambda I_n)^2 & \cdots & \ker(A-\lambda I_n)^{n_\lambda-1} & \ker(A-\lambda I_n)^{n_\lambda} \\
 & & & & \\
(A-\lambda I_n)^{n_\lambda-1}v_1 & (A-\lambda I_n)^{n_\lambda-2}v_1 & \cdots & (A-\lambda I_n)v_1 & v_1 \\
(A-\lambda I_n)^{n_\lambda-1}v_2 & (A-\lambda I_n)^{n_\lambda-2}v_2 & \cdots & (A-\lambda I_n)v_2 & v_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
(A-\lambda I_n)^{n_\lambda-1}v_{d_1} & (A-\lambda I_n)^{n_\lambda-2}v_{d_1} & \cdots & (A-\lambda I_n)v_{d_1} & v_{d_1}
\end{array}$$

Todos estos vectores $(A-\lambda I_n)^k v_j$ para $k=1,...,n_\lambda-1$ y $j=1,...,d_1$ formarán parte de la base $\mathcal{B}$; de hecho cada cadena

$$\{v_j,(A-\lambda I_n)v_j,...,(A-\lambda I_n)^{n_\lambda-1}v_j\}, \qquad j=1,...,d_1,$$

define un conjunto linealmente independiente respecto del cual la matriz de $A-\lambda I_n$ es un bloque de Jordan de orden máximo n_λ; equivalentemente, son la base del subespacio cíclico $Z(v_j)$ que tendrá longitud máxima n_λ.

Pasamos ahora a estudiar la diferencia $d_2=\dim\ker(A-\lambda I_n)^{n_\lambda-1}-\dim\ker(A-\lambda I_n)^{n_\lambda-2}$. Esta diferencia nos dice el número de bloques elementales de Jordan de orden al menos $n_\lambda-1$. Como ya sabemos que existen d_1 bloques elementales de orden n_λ, existen dos posibilidades: o bien $d_2>d_1$ o bien $d_2=d_1$. Supongamos primero que $d_2>d_1$ y definimos $\epsilon_2=d_2-d_1$. Sabemos por el punto 5. de la Proposición 3.5.5 que ϵ_2 coincide con el número de bloques elementales de Jordan de orden exactamente $n_\lambda-1$. Tomemos vectores $w_1,...,w_{\epsilon_2}$ en $\ker(A-\lambda I_n)^{n_\lambda-1}\setminus\ker(A-\lambda I_n)^{n_\lambda-2}$ linealmente independientes con los previamente calculados $(A-\lambda I_n)v_k$.

$$\begin{array}{ccccc}
\ker(A-\lambda I_n) & \ker(A-\lambda I_n)^2 & \cdots & \ker(A-\lambda I_n)^{n_\lambda-1} & \ker(A-\lambda I_n)^{n_\lambda} \\
\\
(A-\lambda I_n)^{n_\lambda-1}v_1 & (A-\lambda I_n)^{n_\lambda-2}v_1 & \cdots & (A-\lambda I_n)v_1 & v_1 \\
(A-\lambda I_n)^{n_\lambda-1}v_2 & (A-\lambda I_n)^{n_\lambda-2}v_2 & \cdots & (A-\lambda I_n)v_2 & v_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
(A-\lambda I_n)^{n_\lambda-1}v_{d_1} & (A-\lambda I_n)^{n_\lambda-2}v_{d_1} & \cdots & (A-\lambda I_n)v_{d_1} & v_{d_1} \\
\\
 & & & w_1 & \\
 & & & w_2 & \\
 & & & \vdots & \\
 & & & w_{\epsilon_2} &
\end{array}$$

Ahora aplicamos $A-\lambda I_n$ a cada uno de estos vectores w_j un total de $n_\lambda-2$ veces, hasta llegar a elementos de $\ker(A-\lambda I_n)$.

$$\begin{array}{ccccc}
\ker(A-\lambda I_n) & \ker(A-\lambda I_n)^2 & \cdots & \ker(A-\lambda I_n)^{n_\lambda-1} & \ker(A-\lambda I_n)^{n_\lambda} \\
\\
(A-\lambda I_n)^{n_\lambda-1}v_1 & (A-\lambda I_n)^{n_\lambda-2}v_1 & \cdots & (A-\lambda I_n)v_1 & v_1 \\
(A-\lambda I_n)^{n_\lambda-1}v_2 & (A-\lambda I_n)^{n_\lambda-2}v_2 & \cdots & (A-\lambda I_n)v_2 & v_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
(A-\lambda I_n)^{n_\lambda-1}v_{d_1} & (A-\lambda I_n)^{n_\lambda-2}v_{d_1} & \cdots & (A-\lambda I_n)v_{d_1} & v_{d_1} \\
\\
(A-\lambda I_n)^{n_\lambda-2}w_1 & (A-\lambda I_n)^{n_\lambda-3}w_1 & \cdots & w_1 & \\
(A-\lambda I_n)^{n_\lambda-2}w_2 & (A-\lambda I_n)^{n_\lambda-3}w_2 & \cdots & w_2 & \\
\vdots & \vdots & \vdots & \vdots & \\
(A-\lambda I_n)^{n_\lambda-2}w_{\epsilon_2} & (A-\lambda I_n)^{n_\lambda-3}w_{\epsilon_2} & \cdots & w_{\epsilon_2} &
\end{array}$$

Igualmente, los vectores $(A-\lambda I_n)^k w_j$ con $k=1,...,n_\lambda-2$ y $j=1,...,\epsilon_2$ formarán parte de la base $\mathcal{B}$. De hecho, cada conjunto linealmente independiente,

$$\{w_j,(A-\lambda I_n)w_j,...,(A-\lambda I_n)^{k-1}w_j\}, \qquad j=1,...,\epsilon_2,$$

es la base de un subespacio cíclico respecto del cual la matriz de $A - \lambda I_n$ es un bloque de Jordan de orden k.

Si $d_2 = d_1$ entonces la cantidad de bloques de orden al menos $n_\lambda - 1$ coincide con la cantidad de bloques de orden n_λ y no estamos añadiendo ningún bloque de orden $n_\lambda - 1$. Sea $k < n_\lambda$ el mayor natural tal que $d_k = \dim\ker(A - \lambda I_n)^k - \dim\ker(A - \lambda I_n)^{k-1}$ cumple $d_k > d_{k-1} = \cdots = d_1$ y definimos $\epsilon_k = d_k - d_1$ Ahora tomaríamos $w_1, ..., w_{\epsilon_k}$ en $\ker(A - \lambda I_n)^k \backslash \ker(A - \lambda I_n)^{k-1}$ linealmente independientes con los previamente calculados $(A - \lambda I_n)v_k$ y aplicaríamos $A - \lambda I_n$ a cada uno de estos w_j un total de $k - 1$ veces hasta llegar a $\ker(A - \lambda I_n)$.

$$
\begin{array}{cccccccc}
\ker(A-\lambda I_n) & \ker(A-\lambda I_n)^2 & \cdots & \ker(A-\lambda I_n)^{k-1} & \ker(A-\lambda I_n)^{k} & \cdots & \ker(A-\lambda I_n)^{n_\lambda-1} & \ker(A-\lambda I_n)^{n_\lambda} \\
 & & & & \cdots & & & \\
(A-\lambda I_n)^{n_\lambda-1}v_1 & (A-\lambda I_n)^{n_\lambda-2}v_1 & & \cdots & \cdots & & (A-\lambda I_n)v_1 & v_1 \\
(A-\lambda I_n)^{n_\lambda-1}v_2 & (A-\lambda I_n)^{n_\lambda-2}v_2 & & \cdots & \cdots & & (A-\lambda I_n)v_2 & v_2 \\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\
(A-\lambda I_n)^{n_\lambda-1}v_{d_1} & (A-\lambda I_n)^{n_\lambda-2}v_{d_1} & & \cdots & \cdots & & (A-\lambda I_n)v_{d_1} & v_{d_1} \\
 & & & & \cdots & & & \\
(A-\lambda I_n)^{k-1}w_1 & (A-\lambda I_n)^{k-2}w_1 & & (A-\lambda I_n)w_1 & w_1 & & & \\
(A-\lambda I_n)^{k-1}w_2 & (A-\lambda I_n)^{k-2}w_2 & & (A-\lambda I_n)w_2 & w_2 & & & \\
\vdots & \vdots & & \vdots & \vdots & & & \\
(A-\lambda I_n)^{k-1}w_{\epsilon_2} & (A-\lambda I_n)^{k-2}w_{\epsilon_2} & & (A-\lambda I_n)w_{\epsilon_2} & w_{\epsilon_2} & & &
\end{array}
$$

En general, cada fila que sale del subespacio $\ker(A - \lambda I_n)^k$ desde un vector v_j nos define el subespacio cíclico $Z(v_j)$ de dimensión exactamente k y por tanto un bloque elemental de Jordan de orden k. Repitiendo este proceso hasta llegar finalmente a $\ker(A - \lambda I_n)$, acabaríamos por completar la base $\mathcal{B}$. Aquí surgen tres preguntas que responderemos tras formularlas.

1. ¿Cómo sabemos que hemos acabado? Pues existen varias maneras, una de ellas por ejemplo es recordar que $\dim\ker(A - \lambda I_n) = m_g(\lambda)$ y por tanto necesariamente el proceso acabará cuando tengamos $m_g(\lambda)$ filas en este diagrama.

2. ¿Qué pasa si al finalizar el proceso se tiene un número de filas menor que $m_g(\lambda)$? Basta completar a una base de $\ker(A - \lambda I_n)$ con vectores propios de A que no sean los que previamente hemos obtenido.

3. ¿Cómo se construye la base $\mathcal{B}$ para obtener la forma de Jordan? Se colocan en orden de derecha a izquierda y de arriba a abajo. Es decir,

$$\mathcal{B} = \{v_1, (A - \lambda I_n)v_1, ..., (A - \lambda I_n)^{n_\lambda - 1}v_1, w_1, (A - \lambda I_n)w_1, ...\}.$$

Una vez se entienden las propiedades del subespacio generalizado y la estructura nilpotente de $A - \lambda I_n$ cuando se restringe a $E(A, \lambda)$, estamos listos para demostrar el Teorema de Jordan 3.4.5. Antes, un lema previo.

Lema 3.5.6. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ y supongamos que sus valores propios son todos reales. Existe un polinomio $m_A(t)$ mónico con las siguientes propiedades:*

1. $m_A(t)$ es el polinomio de menor grado tal que $m_A(A) = 0$.

2. *Si λ es valor propio de A, $m_A(\lambda)=0$.*

3. *Si j_λ es el orden del mayor bloque de Jordan correspondiente la valor propio λ, entonces*

$$m_A(t)=\prod_{i=1}^{r}(t-\lambda_i)^{j_{\lambda_i}},$$

donde $\lambda_1,...,\lambda_r\in\mathbb{R}$ son todos los valores propios de A y cada j_λ es el orden del mayor bloque de Jordan correspondiente a cada valor propio λ.

Demostración. Dado en general un polinomio

$$p(t)=t^n+a_{n-1}t^{n-1}+\cdots+a_1t+a_0,\qquad a_k\in\mathbb{R},$$

se define $p(A)$ como el operador $p(A):\mathbb{R}^n\to\mathbb{R}^n$ dado por

$$p(A)v=A^nv+a_{n-1}A^{n-1}v+\cdots+a_1Av+a_0v.$$

Demostremos punto por punto.

1. Sea $R[t]$ el anillo de los polinomios en la variable t y consideremos $I=\{p(t)\in R[t]\colon p(A)=0\}$. Entonces, $I\neq\varnothing$ y es cerrado para suma y multiplicación por cualquier polinomio y por tanto I es un ideal de $R[t]$. Como $R[t]$ es un Dominio Euclídeo, en particular es un Dominio de Ideales Principales y por tanto I está generado por un único elemento, el cual se puede suponer mónico sin perder generalidad,

$$I=(m_A(t))=\{m_A(t)q(t)\colon q(t)\in R[t]\}.$$

Que $m_A(t)$ tiene grado mínimo es consecuencia de que genera tal ideal.

2. Sea λ un valor propio de A y sea v un vector propio no nulo. Como quiera que $A^kv=\lambda^kv$, sustituyendo en la expresión de $p(A)$ se tiene $p(A)v=p(\lambda)v$. En particular, $0=m_A(A)v=m_A(\lambda)v$ y como v es no nulo, necesariamente $m_A(\lambda)=0$.

3. Del punto anterior se tiene que si λ es valor propio, entonces existe $j_\lambda\in\mathbb{N}$ tal que $m_A(t)=(t-\lambda)^{j_\lambda}q(t)$, con $q(\lambda)\neq 0$. Más aún, como $(t-\lambda)^{j_\lambda}$ y $q(t)$ son coprimos, la identidad de Bezout asegura la existencia de polinomios $a(t),b(t)\in R[t]$ tales que

$$a(t)(t-\lambda)^{t_\lambda}+b(t)q(t)=1,$$

y en particular

$$a(A)(A-\lambda I_n)^{t_\lambda}+b(A)q(A)=I_n.$$

Si aplicamos la igualdad anterior a $v\in E(A,\lambda)$ se tiene $b(A)q(A)v=v$. Si $q(A)v=0$ entonces se concluye $v=0$ y por tanto si v es no cero, necesariamente $q(A)v\neq 0$. En consecuencia, para todo $v\in E(A,\lambda)$ se tiene $(A-\lambda I_n)^{j_\lambda}v=0$, y esto sucede exactamente cuando j_λ coincide con el orden del mayor bloque elemental de Jordan definido por el subespacio cíclico en $E(A,\lambda)$ de mayor longitud. Repitiendo este proceso en un número finito de pasos acabaríamos de descomponer

$$m_A(t)=\prod_{i=1}^{r}(t-\lambda_i)^{j_{\lambda_i}}.$$

□

Ahora sí, damos el resultado con el que concluye la prueba del Teorema 3.4.5.

Teorema 3.5.7. *Sea $A \in \mathcal{M}_n(\mathbb{R})$ y sean $\lambda_1, ..., \lambda_r \in \mathbb{R}$ sus valores propios. Entonces,*

$$\mathbb{R}^n = E(A, \lambda_1) \oplus \cdots \oplus E(A, \lambda_r).$$

Demostración. Fijemos $\lambda = \lambda_1 \in \mathbb{R}$ valor propio de A y sea $m_A(t)$ el polinomio mínimo de A. Sabemos por el Lema 3.5.6 que $m_A(t) = (t - \lambda)^{j_\lambda}$ y que se cumple

$$a(t)(t - \lambda)^{j_\lambda} + b(t)q(t) = 1,$$

para ciertos $a(t), b(t)$ y siendo $q(t)$ tal que $q(\lambda) \neq 0$. Definiendo $p(t) = b(t)q(t)$ se tiene

$$I_n - p(A) = a(A)(A - \lambda I_n)^{j_\lambda}.$$

Veamos que así definido, el operador $p(A)$ es una *proyección*, esto es $p(A)^2 = p(A)$. Para ello, calculemos $p(t)^2 - p(t)$.

$$p(t)^2 - p(t) = p(t)(p(t) - 1) = p(t)(-a(t)(t - \lambda)^{j_\lambda}).$$

Por otra parte,

$$(t - \lambda)^{j_\lambda} p(t) = (t - \lambda)^{j_\lambda} b(t)q(t) = b(t)m_A(t),$$

y por tanto

$$p(t)^2 - p(t) = -a(t)b(t)m_A(t).$$

Sustituyendo finalmente $t = A$ se tiene $p(A)^2 = p(A)$ y por tanto $p(A)$ es una proyección.

Aunque es conocido el hecho de que para toda proyección se tiene la descomposición

$$\mathbb{R}^n = \ker p(A) \oplus \operatorname{im} p(A),$$

lo probamos por completitud de las notas. Por una parte, por ser $p(A)$ proyección se tiene que $I_n - p(A)$ es también proyección. Para cualquier $v \in \mathbb{R}^n$ se tiene $v = p(A)v +‘ (I_n - p(A))v$ y en particular $p(A)v \in \operatorname{im} p(A)$. Por otra parte, $p(A)(I_n - p(A))v = 0$ y por tanto $(I_n - p(A))v \in \ker p(A)$. En consecuencia, $\mathbb{R}^n = \ker p(A) + \operatorname{im} p(A)$. Falta probar que la suma es directa y para ello tomemos $v \in \ker p(A) \cap \operatorname{im} p(A)$. Por una parte, existe $w \in \mathbb{R}^n$ tal que $v = p(A)w$. Como $v \in \ker p(A)$ se tiene

$$0 = p(A)v = p(A)^2 w = p(A)w = v,$$

lo que implica $v = 0$.

El paso final es estudiar los subsespacios $\ker p(A)$ e $\operatorname{im} p(A)$. Lo primero es observar que como $p(A)$ es un polinomio en A se tiene que $p(A)A = Ap(A)$ y por tanto $\ker p(A)$ es A-invariante. En cuanto a la imagen, veamos que $\operatorname{im} p(A) = E(A, \lambda)$. De la factorización $m_A(t) = (t - \lambda)^{j_\lambda} q(t)$ se tiene

$$(t - \lambda)^{j_\lambda} p(t) = (t - \lambda)^{j_\lambda} b(t)q(t) = b(t)(t - \lambda)^{j_\lambda} q(t) = b(t)m_A(t).$$

Sustituyendo en A y usando que $m_A(A) = 0$ llegamos a $(A - \lambda I_n)^{j_\lambda} p(A) = 0$. En consecuencia, para cada $v \in \mathbb{R}^n$ se tiene $p(A)v \in \ker(A - \lambda I_n)^{j_\lambda}$ y por tanto $\operatorname{im} p(A) \subset E(A, \lambda)$. Como por otra parte $p(A)$ es la identidad en $E(A, \lambda)$ concluimos que $\operatorname{im} p(A) = E(A, \lambda)$.

Con todo esto, tenemos la descomposición

$$\mathbb{R}^n = E(A,\lambda) \oplus V,$$

siendo $V = \ker p(A)$ un subespacio A-invariante. Sea ahora el operador $A_{|_V} : V \to V$. Su polinomio mínimo es precisamente

$$m_{A_{|_V}}(t) = \prod_{i=2}^{r} (t-\lambda_i)^{j_{\lambda_i}}.$$

Fijando $\lambda = \lambda_2$, descompondríamos $V = E(A,\lambda_2) \oplus V_1$, siendo otra vez V_1 un subespacio A-invariante. Repitiendo este proceso llegamos a la descomposición final

$$\mathbb{R}^n = E(A,\lambda_1) \oplus \cdots \oplus E(A,\lambda_r).$$

$\square$

La forma ahora de obtener la base $\mathcal{B}$ de $\mathbb{R}^n$ está clara. Cada subespacio $E(A,\lambda_k)$ nos proporciona una base $\mathcal{B}_k$ y tras concatenarlas obtendríamos una base de $\mathbb{R}^n$ respecto de la cual la matriz de A en tal base es su forma canónica de Jordan.

Ejemplo 3.7. Vamos a calcular la forma canónica de Jordan de

$$A = \begin{pmatrix} -1 & 1 & -2 \\ 0 & -1 & 4 \\ 0 & 0 & 1 \end{pmatrix}.$$

Su ecuación característica es $(\lambda+1)^2(1-\lambda) = 0$, lo que nos proporciona valores propios $\lambda = 1$ simple y $\lambda = -1$ doble. Empezamos con el valor propio simple, puesto que $m_g(1) \leq m_a(1) = 1$ y por tanto coinciden. En particular, solo hay un bloque de Jordan correspondiente a $\lambda = 1$. Calculamos un vector propio para $\lambda = 1$ y para ello estudiemos $\ker(A - I_3)$.

$$\ker(A - I_3) = \ker \begin{pmatrix} -2 & 1 & -2 \\ 0 & -2 & 4 \\ 0 & 0 & 0 \end{pmatrix},$$

de lo que deducimos $\ker(A - I_3) = \{(0, 2z, z) \colon z \in \mathbb{R}\}$. Un vector propio pues es $v_1 = (0,2,1)$. Para el valor propio $\lambda = -1$ estudiamos análogamente $\ker(A + I_3)$.

$$\ker(A + I_3) = \ker \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 4 \\ 0 & 0 & 2 \end{pmatrix},$$

de lo que deducimos $\ker(A + I_3) = \{(x,0,0) \colon x \in \mathbb{R}\}$ y en particular $\dim \ker(A + I_3) = 1 < 2 = m_a(-1)$. En consecuencia hay un bloque de Jordan correspondiente a $\lambda = -1$, necesariamente de orden 2. Para hallar una base de vectores propios generalizados, comenzamos calculando $(A + I_3)^2$.

$$(A + I_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 8 \\ 0 & 0 & 4 \end{pmatrix},$$

y se tiene $\ker(A+I_3)^2 = \{z=0\}$. Tomamos $v_2 \in \ker(A+I_3)^2 \setminus \ker(A+I_3)$, por ejemplo $v_2 = (0,1,0)$ y posteriormente $v_3 = (A+I_3)v_2 = (1,0,0)$. En particular, v_3 es vector propio para $\lambda = -1$. Obtenemos pues la base $\{v_1, v_2, v_3\}$ de $\mathbb{R}^3$, respecto de la cual la matriz A vista como endomorfismo satisface

$$Av_1 = v_1, \quad Av_2 = -v_2 + v_3, \quad Av_3 = -v_3.$$

Si tomamos P como la matriz cuyas columnas son v_1, v_2, v_3, concluimos

$$A = PJP^{-1}, \qquad J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}, \qquad P = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Si queríamos resolver el sistema $x' = Ax$, sabemos que $\Phi(t) = e^{tA}$ es matriz fundamental del sistema. En otras palabras,

$$\Phi(t) = Pe^{tJ}P^{-1}$$

es matriz fundamental, donde sabemos que podemos obviar el cálculo de P^{-1}. Para calcular e^{tJ}, calculamos la exponencial de cada uno de sus bloques elementales de Jordan. El primer bloque de Jordan consiste únicamente de t, obteniendo e^t. En cuanto a la exponencial del segundo bloque elemental de Jordan,

$$\exp\begin{pmatrix} -t & 0 \\ t & -t \end{pmatrix} = \exp\begin{pmatrix} -t & 0 \\ 0 & -t \end{pmatrix} \exp\begin{pmatrix} 0 & 0 \\ t & 0 \end{pmatrix} = \begin{pmatrix} e^{-t} & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} = \begin{pmatrix} e^{-t} & 0 \\ te^{-t} & e^{-t} \end{pmatrix}.$$

Concluimos que una matriz fundamental del sistema es

$$\Phi(t) = P \begin{pmatrix} e^t & 0 & 0 \\ 0 & e^{-t} & 0 \\ 0 & te^{-t} & e^{-t} \end{pmatrix} = \begin{pmatrix} 0 & te^{-t} & e^{-t} \\ 2e^t & e^{-t} & 0 \\ e^t & 0 & 0 \end{pmatrix},$$

y por tanto las columnas forman un sistema fundamental de soluciones.

Ejemplo 3.8. Sea ahora

$$A = \begin{pmatrix} 0 & -1 & 1 \\ 3 & 4 & -2 \\ -1 & 0 & 2 \end{pmatrix}.$$

Se tiene $p(\lambda) = \det(A - \lambda I_3) = \lambda^3 - 6\lambda^2 + 12\lambda - 8 = (\lambda-2)^3$ y por tanto la única raíz real es $\lambda = 2$ con $m_a(\lambda) = 3$. Ahora bien, si calculamos $\ker(A - 2I_3)$ se tiene

$$\begin{pmatrix} -2 & -1 & 1 \\ 3 & 2 & -2 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Longrightarrow x = 0, \quad y = z.$$

En consecuencia, $\ker(A - 2I_3) = \operatorname{span}(0,1,1)$ y $m_g(2) = 1$. Existe por tanto un único bloque de Jordan de tamaño máximo 3 y la torre de subespacios toma la forma

$$\ker(A - 2I_3) \subset \ker(A - 2I_3)^2 \subset \ker(A - 2I_3)^3 = \mathbb{R}^3.$$

La matriz $(A-2I_3)^2$ es

$$(A-2I_3)^2 = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 1 & -1 \\ 2 & 1 & -1 \end{pmatrix},$$

y se tiene $\ker(A-2I_3)^2 = \operatorname{span}((0,1,1),(1,0,2))$. Tomamos $v_1 \in \ker(A-2I_3)^3 \setminus \ker(A-2I_3)^2$, por ejemplo $v_1 = (1,0,0)$ y definimos

$$v_2 = (A-2I_3)v_1 = (-2,3,1), \qquad v_3 = (A-2I_3)^2 v_1 = (A-2I_3)v_2 = (0,2,2).$$

Si definimos la base $\mathcal{B} = \{v_1, v_2, v_3\}$ y la matriz por columnas $P = (v_1, v_2, v_3)$, se tiene $A = PJP^{-1}$, donde

$$P = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix}, \qquad J = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

Finalmente, el sistema diferencial $x' = Ax$ tiene como matriz fundamental

$$\Phi(t) = Pe^{tJ} = P \begin{pmatrix} e^{2t} & 0 & 0 \\ te^{2t} & e^{2t} & 0 \\ \frac{t^2}{2}e^{2t} & te^{2t} & e^{2t} \end{pmatrix}.$$

3.5.2. Valor propio complejo

A lo largo de esta sección suponemos que $\mu = a + ib \in \mathbb{C}$ es solución de $\det(A - \mu I_n) = 0$, con $b \neq 0$ para no caer en el caso real anteriormente estudiado. Además, al igual que en el caso real, consideraremos indistintamente A como matriz y a su vez como el endomorfismo tal que su representación en la base canónica de $\mathbb{R}^n$ es precisamente la matriz A. Como $A \in \mathcal{M}_n(\mathbb{R})$, su polinomio característico es de coeficientes reales y por tanto $\overline{\mu}$ es igualmente solución de $\det(A - \mu I_n) = 0$. Ahora bien, ¿qué significa que $\mu \in \mathbb{C}$ sea valor propio de una matriz A de coeficientes reales? Obviamente no puede significar $Aw = \mu w$ para algún vector $w \in \mathbb{R}^n$, porque esta ecuación carece de sentido ya que $Aw \in \mathbb{R}^n$, pero $\mu w \in \mathbb{C}^n$. La interpretación de que μ sea valor propio de A puede entenderse extendiendo de cierta forma el endomorfismo A a un operador lineal $A_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n$, de manera que dotamos de sentido a la relación $A_{\mathbb{C}} w = \mu w$. Como todo vector $w \in \mathbb{C}^n$ se puede descomponer como $w = u + iv$, con $u, v \in \mathbb{R}^n$, una forma de extender A es la siguiente,

$$A_{\mathbb{C}} : \mathbb{C}^n \to \mathbb{C}^n, \qquad A_{\mathbb{C}} w = Au + iAv, \qquad w = u + iv \in \mathbb{C}^n,$$

el cual es un operador $\mathbb{C}$-lineal sobre $\mathbb{C}^n$. Todas estas ideas se pueden abstraer mediante el concepto del *complexificado* de un espacio vectorial real E y un operador $\mathbb{R}$-lineal en E.

Definición 3.5.8. Dado E un espacio vectorial sobre $\mathbb{R}$, definimos su *complexificado* como el espacio vectorial

$$E_{\mathbb{C}} = \left\{ \sum_{k=1}^{n} \alpha_k v_k \colon n \in \mathbb{N},\ \alpha_k \in \mathbb{C},\ v_k \in E \right\}.$$

En esta definición estamos considerando *sumas formales*, puesto que a priori no tiene sentido matemático el producto αv con $\alpha \in \mathbb{C}$ y $v \in E$. Esto se puede reescribir como

$$E_{\mathbb{C}} = \{u + iv \colon u, v \in E\},$$

siendo $i \in \mathbb{C}$ la unidad imaginaria. Así definido se tiene que $E_{\mathbb{C}}$ es un espacio vectorial sobre $\mathbb{C}$. Además, si $\mathcal{B}$ es una base de E, entonces $\mathcal{B}$ es una base de $E_{\mathbb{C}}$ como espacio vectorial complejo. En efecto, si $\mathcal{B} = \{v_1, ..., v_n\}$ con $v_k \in E$ y tomamos $w = u+iv$, entonces existen $\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_n$ tales que

$$u = \sum_{k=1}^{n} \alpha_k v_k, \qquad v = \sum_{k=1}^{n} \beta_k v_k.$$

Por tanto,

$$w = u + iv = \sum_{k=1}^{n} \alpha_k v_k + i \sum_{k=1}^{n} \beta_k v_k = \sum_{k=1}^{n} \alpha_k v_k + \sum_{k=1}^{n} i\beta_k v_k = \sum_{k=1}^{n} (\alpha_k + i\beta_k) v_k = \sum_{k=1}^{n} \mu_k v_k,$$

lo que prueba que $\mathcal{B}$ es sistema de generadores de $E_{\mathbb{C}}$. Como son linealmente independientes en E también lo son trivialmente en $E_{\mathbb{C}}$, lo que concluye el resultado. Como consecuencia, E y $E_{\mathbb{C}}$ tienen la misma dimensión pero con un matiz: el primero como espacio vectorial real y el segundo como espacio vectorial complejo. Obsérvese igualmente que E es un subconjunto de $E_{\mathbb{C}}$, pero no es subespacio vectorial (visto $E_{\mathbb{C}}$ como espacio vectorial complejo; sí que lo es si $E_{\mathbb{C}}$ es visto como espacio vectorial sobre $\mathbb{R}$). Para el caso particular $E = \mathbb{R}^n$, el complexificado $\mathbb{R}^n_{\mathbb{C}}$ no es más que $\mathbb{C}^n = \{u + iv \colon u, v \in \mathbb{R}^n\}$.

Definición 3.5.9. Dado un operador lineal $T : E \to E$, se define su complexificado $T_{\mathbb{C}} : E_{\mathbb{C}} \to E_{\mathbb{C}}$ como el operador dado de las dos formas equivalentes siguientes,

$$T_{\mathbb{C}} \left(\sum_{k=1}^{n} \alpha_k v_k \right) = \sum_{k=1}^{n} \alpha_k T v_k, \quad \alpha_k \in \mathbb{C},\ v_k \in E, \qquad T_{\mathbb{C}}(u + iv) = Tu + iTv, \quad u, v \in E.$$

Así definido, $T_{\mathbb{C}}$ es un operador $\mathbb{C}$-lineal. Además, se cumple la siguiente propiedad fundamental. Si $\mathcal{B}$ una base de E (e igualmente de $E_{\mathbb{C}}$) y denotamos por $A \in \mathcal{M}_n(\mathbb{R})$ a la representación matricial de T en la base $\mathcal{B}$, entonces la matriz de $T_{\mathbb{C}}$ en $\mathcal{B}$ coincide con A. Esto es inmediato al ser $\mathcal{B} = \{v_1, ..., v_n\}$ con $v_k \in E$ y por tanto $T_{\mathbb{C}} v_k = T v_k$. Ahora bien, sea $p(\mu) = \det(T - \mu I)$ el polinomio característico del operador real T. Entonces, $p(\mu)$ es igualmente el polinomio característico de $T_{\mathbb{C}}$, ya que el polinomio característico se obtiene a través de una representación matricial de un operador y en nuestro caso A es representación matricial de T y $T_{\mathbb{C}}$. Como μ es valor propio de $T_{\mathbb{C}}$ si y sólo si μ es raíz del polinomio característico de $T_{\mathbb{C}}$, esta discusión justifica la siguiente definición.

Definición 3.5.10. Decimos que $\mu \in \mathbb{C}$ es valor propio de $T : E \to E$ si y sólo si μ es valor propio del operador complexificado $T_{\mathbb{C}}$, es decir existe $w \in \mathbb{C}^n$ tal que $T_{\mathbb{C}} w = \mu w$.

Un problema de interés es determinar cuándo un subespacio de $E_{\mathbb{C}}$ es el complexificado de un subespacio de E. Es decir, sea W un subespacio vectorial de $E_{\mathbb{C}}$. ¿Existe V subespacio vectorial de E tal que $V_{\mathbb{C}} = W$? La misma cuestión la planteamos para operadores. Dado $L : E_{\mathbb{C}} \to E_{\mathbb{C}}$ un operador $\mathbb{C}$-lineal, dar condiciones que aseguren que $L = T_{\mathbb{C}}$, siendo $T : E \to E$ lineal. Para tal fin, definimos la aplicación conjugación

$$\sigma : E_{\mathbb{C}} \to E_{\mathbb{C}}, \qquad \sigma(u + iv) = u - iv, \qquad u, v \in E.$$

Es inmediato comprobar que para cualesquiera $w \in E_{\mathbb{C}}$ y $\mu \in \mathbb{C}$ se tiene $\sigma(\mu w) = \overline{\mu}\sigma(w)$. En particular, σ no es lineal. Cuando apliquemos la conjugación sobre un vector $w \in E_{\mathbb{C}}$ emplearemos la notación reducida $\sigma(w) = \overline{w}$. Una vez introducida la aplicación conjugación se puede comprobar que si $w \in \mathbb{C}^n$ es vector propio de $T_{\mathbb{C}}$ para el valor propio μ, entonces $\overline{w}$ es vector propio de $T_{\mathbb{C}}$ para el valor propio $\overline{\mu}$. En efecto, si $w = u + iv$ se tiene

$$T_{\mathbb{C}}\overline{w} = T_{\mathbb{C}}(u - iv) = Tu - iTv = \overline{Tu + iTv} = \overline{T_{\mathbb{C}}w} = \overline{\mu\, w} = \overline{\mu}\,\overline{w}.$$

Además, la multiplicidad de $\overline{w}$ coincide con la de w. El siguiente resultado responde a la cuestión de cuándo subespacios y operadores complejos son los complexificados de análogos reales.

Proposición 3.5.11. *Sean W un subespacio vectorial de $E_{\mathbb{C}}$ y $L : E_{\mathbb{C}} \to E_{\mathbb{C}}$ lineal.*

1. *W es el complexificado de un subespacio V de E si y sólo si $\sigma(W) \subset W$.*
2. *L es es complexificado de un operador lineal T de E si y sólo si $L\sigma = \sigma L$, es decir L y σ conmutan.*

Demostración. 1. Supongamos $W = V_{\mathbb{C}}$, siendo V subespacio vectorial de W, y tomemos $w \in W$. El objetivo es verificar que $\overline{w} \in W$. Por ser $W = V_{\mathbb{C}}$, existirán $u, v \in V$ tales que $w = u + iv$. Entonces, $\overline{w} = u - iv = u + i(-v)$ y como V es subespacio vectorial en particular $-v \in V$ y se tiene $\overline{w} \in W$.

Recíprocamente, supongamos que $\sigma(W) \subset W$. Debemos encontrar algún subespacio vectorial V de E de manera que $V_{\mathbb{C}} = W$, y para ello definimos

$$V = \{\Re w \colon w \in W\}.$$

Que V es subespacio vectorial real de E es inmediato. En efecto, si $u, v \in V$ entonces $u = \Re w,\ v = \Re z$ y por tanto $u + v = \Re(w + z)$. Como $w + z \in W$ se tiene que $u + v \in V$. Igualmente, si $\lambda \in \mathbb{R}$ entonces $\lambda v = \lambda \Re w = \Re(\lambda w)$ y como $\lambda w \in W$ se tiene $\lambda v \in V$. Falta probar que $V_{\mathbb{C}} = W$. Primero, dado $u + iv \in V_{\mathbb{C}}$ se tiene $u = \Re w,\ v = \Re z$ con $w, z \in W$. Como quiera que se tienen las igualdades

$$\Re w = \frac{w + \overline{w}}{2}, \qquad \Re z = \frac{z + \overline{z}}{2}$$

y $\overline{w},\ \overline{z} \in W$ por hipótesis, entonces $u, v \in W$ y en consecuencia $u + iv \in W$. Por otra parte, dado $w \in W$ se tiene $w = u + iv$. Pero $u = \Re w \in V$ y $v = -\Re(iw) \in V$ y concluimos $w = u + iv \in V_{\mathbb{C}}$.

2. Si L es el complexificado de T entonces dado $w = u + iv \in E_{\mathbb{C}}$ se tiene

$$L\sigma(w) = T_{\mathbb{C}}(u - iv) = Tu - iTv = \sigma(Tu + iTv) = \sigma Lw.$$

Como w es arbitrario se tiene que L y σ conmutan.

Recíprocamente, si L y σ conmutan entonces dado $w = u + iv \in E_{\mathbb{C}}$ la igualdad $L\sigma(w) = \sigma(Lw)$ implica $\overline{Lu} = Lu$ y $\overline{Lv} = Lv$ de donde concluimos $Lu, Lv \in E$. En particular, la restricción $L : E \to E$ está bien definida. Ahora es claro que definiendo $T : E \to E$ dado por $Tu = Lu$ la igualdad $L = T_{\mathbb{C}}$ es trivial.

$\square$

Una vez introducidos estos conceptos de álgebra lineal, seguimos con la búsqueda de la forma de Jordan real de A correspondiente a un valor propio $\mu \in \mathbb{C}$. De igual forma al caso real, el objetivo es encontrar un cierto subespacio de $\mathbb{R}^n$ y una base $\mathcal{B}$ suya de manera que al restringir el operador A a tal subespacio, su matriz en la base $\mathcal{B}$ sea

$$J(\mu) = \begin{pmatrix} E_{r_1}(\mu) & & \\ & \ddots & \\ & & E_{r_{m_g(\mu)}}(\mu) \end{pmatrix}.$$

Mimetizando el caso real, se tiene $\dim \ker(A_{\mathbb{C}} - \mu I_n) = m_g(\mu)$ y la cadena de subespacios

$$\{0\} \subset \ker(A_{\mathbb{C}} - \mu I_n) \subset \ker(A_{\mathbb{C}} - \mu I_n)^2 \subset \cdots \subset \ker(A_{\mathbb{C}} - \mu I_n)^k \subset \cdots$$

se estabiliza para cierto p para el cual $\dim \ker(A_{\mathbb{C}} - \mu I_n)^p = m_a(\mu)$. Definimos de forma análoga al caso real el subespacio propio generalizado de $A_{\mathbb{C}}$ correspondiente a μ como

$$E_{\mathbb{C}}(A_{\mathbb{C}}, \mu) = \ker(A_{\mathbb{C}} - \mu I_n)^p.$$

La peculiaridad es que si μ es complejo con $b \neq 0$, $E_{\mathbb{C}}(A_{\mathbb{C}}, \mu)$ no es el complexificado de ningún subespacio vectorial real, básicamente por el hecho de que si μ es valor propio de A entonces $\overline{\mu}$ es también valor propio. Sin embargo, la suma directa $E_{\mathbb{C}}(A_{\mathbb{C}}, \mu) \oplus E(A_{\mathbb{C}}, \overline{\mu})$ sí que es el complexificado de algún subespacio vectorial real, ya que en virtud de la Proposición 3.5.11 basta probar que esta suma directa es invariante por la operación conjugación. Como quiera que $\sigma(E_{\mathbb{C}}(A_{\mathbb{C}}, \mu)) = E_{\mathbb{C}}(A_{\mathbb{C}}, \overline{\mu}))$, se tiene que la suma directa anterior es el complexificado de algún subespacio vectorial real V_μ. Más aún, la dimensión de $E_{\mathbb{C}}(A_{\mathbb{C}}, \mu) \oplus E_{\mathbb{C}}(A_{\mathbb{C}}, \overline{\mu})$ como espacio vectorial complejo es $2m_a(\mu)$ y por tanto la dimensión de V_μ, como espacio vectorial real, es igualmente $2m_a(\mu)$. El problema se reduce pues a calcular una base del subespacio real V_μ. Detallamos a continuación los pasos.

1. La primera parte es totalmente análoga al caso real, con la diferencia fundamental que ahora estamos considerando $A_{\mathbb{C}}$ como endomorfismo de $\mathbb{C}^n$ y no de $\mathbb{R}^n$. Entonces, $E_{\mathbb{C}}(A_{\mathbb{C}}, \mu)$ se descompone como una suma directa de subespacios cíclicos,

   $$E_{\mathbb{C}}(A_{\mathbb{C}}, \mu) = Z(w_1) \oplus \cdots \oplus Z(w_{m_g(\mu)}),$$

   con $w_1, ..., w_{m_g(\mu)} \in \mathbb{C}^n$. De hecho, cada $Z(w_k)$ tiene como base cíclica

   $$\mathcal{B}_k = \{w_k, (A_{\mathbb{C}} - \mu I_n)w_k, ..., (A_{\mathbb{C}} - \mu I_n)^{\mathrm{nil}(w_k)-1} w_k\},$$

   y $(A_{\mathbb{C}} - \mu I_n)^{\mathrm{nil}(w_k)-1} w_k \in \ker(A_{\mathbb{C}} - \mu I_n)$. Concatenando estas bases obtendríamos una base $\mathcal{B}$ respecto de la cual la matriz del operador $A_{\mathbb{C}}$ es de la forma

   $$J(\mu) = \begin{pmatrix} D_{r_1}(\mu) & & \\ & \ddots & \\ & & D_{r_{m_g(\mu)}}(\mu) \end{pmatrix}, \qquad D_{r_j}(\mu) = \begin{pmatrix} \mu & & & \\ 1 & \mu & & \\ & \ddots & \ddots & \\ & & 1 & \mu \end{pmatrix}.$$

   Esto no puede ser la representación de Jordan real, puesto que μ es complejo. De hecho, $E_{\mathbb{C}}(A_{\mathbb{C}}, \mu)$ no es el complexificado de ningún subespacio real puesto que ni siquiera $Z(w_k)$ lo es. Ahora bien, $Z(w_k) \oplus Z(\overline{w_k})$ sí que es complexificado de algún subespacio real puesto que la aplicación conjugación cumple

$$\sigma(Z(w_k) \oplus Z(\overline{w_k})) = Z(\overline{w_k}) \oplus Z(w_k) = Z(w_k) \oplus Z(\overline{w_k}).$$

En consecuencia, $E_{\mathbb{C}}(A_{\mathbb{C}}, \mu) \oplus E_{\mathbb{C}}(A_{\mathbb{C}}, \overline{\mu})$ es complexificado de algún subespacio real.

2. Consideremos una base cíclica $\{w_1, ..., w_r\}$ de cierto bloque $D_r(\mu)$, donde $w_{k+1} = (A_{\mathbb{C}} - \mu I_n)w_k$ para $k = 1, ..., r-1$, o expresado de otra forma, $A_{\mathbb{C}}w_k = \mu w_k + w_{k+1}$ para $k = 1, ..., r-1$ y $A_{\mathbb{C}}w_r = \mu w_r$. Si descomponemos $w_k = u_k + iv_k$, con $u_k, v_k \in \mathbb{R}^n$, se tiene

$$\begin{aligned} A_{\mathbb{C}}w_k = Au_k + iAv_k &= (a+bi)(u_k + iv_k) + (u_{k+1} + iv_{k+1}) \\ &= (au_k - bv_k - u_{k+1}) + i(av_k + bu_k + v_{k+1}), \end{aligned}$$

donde ahora enfatizamos que A es endomorfismo real actuando sobre $u_k, v_k \in \mathbb{R}^n$, y en consecuencia

$$Au_k = au_k - bv_k - u_{k+1}, \qquad Av_k = av_k + bu_k + v_{k+1}, \quad k = 1, ..., r-1.$$

Además, $A_{\mathbb{C}}w_r = \mu w_r$ implica

$$Au_r + iAv_r = (a+bi)(u_r + iv_r) = (au_r - bv_r) + i(av_r + bu_r).$$

Por tanto, la matriz de A en $\{v_1, u_1, v_2, u_2, ..., v_r, u_r\}$ es

$$E_r(\mu) = \left(\begin{array}{cc|cc|ccc|cc} a & -b & & & & & & & \\ b & a & & & & & & & \\ \hline 1 & 0 & a & -b & & & & & \\ 0 & 1 & b & a & & & & & \\ & & & & \ddots & & & & \\ & & & & & a & -b & & \\ & & & & & b & a & & \\ \hline & & & & & 1 & 0 & a & -b \\ & & & & & 0 & 1 & b & a \end{array} \right).$$

Más aún, el subespacio real $V = \operatorname{span}(\{v_1, u_1, v_2, u_2, ..., v_r, u_r\})$ cumple $V_{\mathbb{C}} = Z(w_1) \oplus Z(\overline{w_1})$.

3. Con esto hemos representado un bloque arbitrario $D_{r_j}(\mu)$, obteniendo un bloque del doble de tamaño $E_{r_j}(\mu)$ que representa tanto a μ como $\overline{\mu}$. Repitiendo este proceso para cada bloque y concatenando las bases, se tendrá una base $\mathcal{B}$ de $2n$ vectores respecto de la cual la matriz de A en $\mathcal{B}$ es precisamente

$$J(\mu) = \begin{pmatrix} E_{r_1}(\mu) & & \\ & \ddots & \\ & & E_{r_{m_g(\mu)}}(\mu) \end{pmatrix}.$$

Ejemplo 3.9. Consideremos la matriz

$$A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 2 & 3 & -2 & -1 \\ 2 & 2 & 1 & 0 \\ -2 & 0 & 2 & 3 \end{pmatrix}.$$

El polinomio característico de A es

$$p(\lambda) = \det(A - \lambda I_4) = 25 - 40\lambda + 26\lambda^2 - 8\lambda^3 + \lambda^4 = (\lambda(\lambda - 4) + 5)^2,$$

cuyas raíces son $\lambda = 2 \pm i$ con $m_a(2 \pm i) = 2$. Como A tiene valores propios complejos, necesitamos trabajar con el operador complexificado $A_{\mathbb{C}}$ en $\mathbb{C}^4$, en el cual consideramos coordenadas $w = (x, y, z, t) \in \mathbb{C}^4$ y $w = u + iv$ con $u, v \in \mathbb{R}^4$. Calculamos $\ker(A_{\mathbb{C}} - (2 + i)I_4)$, obteniendo el sistema de ecuaciones complejas

$$\begin{pmatrix} -1-i & 0 & 1 & 1 \\ 2 & 1-i & -2 & -1 \\ 2 & 2 & -1-i & 0 \\ -2 & 0 & 2 & 1-i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad x, ..., t \in \mathbb{C}.$$

Este sistema tiene como solución el espacio de dimensión 1 generado por $(1 - i, -1 + i, 0, 2)$. En consecuencia, $m_g(2+i) = 1$ y la matriz no es diagonalizable en $\mathbb{C}$. Calculamos pues $\ker(A_{\mathbb{C}} - (2+i)I_4)^2$,

$$(A_{\mathbb{C}} - (2+i)I_4)^2 \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 2i & 2 & -2i & -2i \\ -2-4i & -4-2i & 4i & 2i \\ -4i & -4i & -2+2i & 0 \\ 4+4i & 4 & -2-4i & -2-2i \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Resolviendo este sistema nos queda el subespacio de dimensión 2 definido por las ecuaciones

$$x = -i(y + t), \qquad z = -y - (1 + i)t$$

Una base de este subespacio es $\ker(A_{\mathbb{C}} - (2 + i)I_4)^2 = \operatorname{span}((1, 0, 1 - i, i), (1, i, 2, 0))$. Además, notemos que $w_1 = (1, i, 2, 0) \in \ker(A_{\mathbb{C}} - (2 + i)I_4)^2 \setminus \ker(A_{\mathbb{C}} - (2 + i)I_4)$ y podemos definir $w_2 = (A_{\mathbb{C}} - (2 + i)I_4)w_1 = (1 - i, -1 + i, 0, 2)$. En consecuencia,

$$E_{\mathbb{C}}(A_{\mathbb{C}} - (2 + i)I_4) = \ker(A_{\mathbb{C}} - (2 + i)I_4)^2 = Z(w_1).$$

Si tomamos como base de $\mathbb{C}^4$ el conjunto $\mathcal{B}_{\mathbb{C}} = \{w_1, w_2, \overline{w_1}, \overline{w_2}\}$ se tiene que la matriz de $A_{\mathbb{C}}$ (visto como endomorfismo de $\mathbb{C}^4$) en $\mathcal{B}_{\mathbb{C}}$ es

$$\begin{pmatrix} 2+i & 0 & 0 & 0 \\ 1 & 2+i & 0 & 0 \\ 0 & 0 & 2-i & 0 \\ 0 & 0 & 1 & 2-i \end{pmatrix} = \operatorname{diag}(D_2(2 + i), D_2(2 - i)).$$

En particular,

$$\mathbb{C}^4 = E_{\mathbb{C}}(A_{\mathbb{C}} - (2 + i)I_4) \oplus E_{\mathbb{C}}(A_{\mathbb{C}} - (2 - i)I_4).$$

Sabemos que esta no es la forma canónica de Jordan real de A, y para calcularla necesitamos definir la base de $\mathbb{R}^4$ dada por

$$\mathcal{B} = \{\Im w_1, \Re w_1, \Im w_2, \Re w_2\}.$$

Ahora definiendo P como la matriz cuyas columnas son los vectores de $\mathcal{B}$ podemos concluir que $A = PJP^{-1}$, donde

$$J = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & -1 \\ 0 & 1 & 1 & 2 \end{pmatrix} = E_2(2+i), \qquad P = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

3.5.3. La forma canónica de Jordan de matrices 3×3

A continuación vamos a realizar un estudio en profundidad de la forma canónica de Jordan de matrices 3×3, dependiendo de los valores propios y sus multiplicidades.

1. A tiene tres valores propios distintos. Se tiene uno de los siguientes casos.

   - Los tres valores propios son reales, $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$. Necesariamente, $\dim \ker(A - \lambda_k I_3) = 1$ y A es diagonalizable. Tomando $v_k \in \ker(A - \lambda_k I_3)$, como base $\mathcal{B} = \{v_1, v_2, v_3\}$ y $P = (v_1, v_2, v_3)$ se tiene

     $$A = PJP^{-1}, \qquad J = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix}$$

     La exponencial e^{tA} viene dada por

     $$e^{tA} = P \begin{pmatrix} e^{t\lambda_1} & & \\ & e^{t\lambda_2} & \\ & & e^{t\lambda_3} \end{pmatrix} P^{-1}.$$

   - Los valores propios son $\lambda \in \mathbb{R},\ \mu, \overline{\mu} \in \mathbb{C}$. Igualmente A es diagonalizable en $\mathbb{C}^3$. Tomamos $v \in \ker(A - \lambda I_3)$, que necesariamente vuelve a tener dimensión 1. Ahora bien, si $\mu = a + bi$ y w es vector propio para μ, entonces descomponemos $w = w_1 + iw_2$ y se tiene

     $$Aw_1 = aw_1 - bw_2, \qquad Aw_2 = bw_1 + aw_2.$$

     Tomando como base $\mathcal{B} = \{v, w_2, w_1\}$ y $P = (v, w_2, w_1)$ se tiene

     $$A = PJP^{-1}, \qquad J = \begin{pmatrix} \lambda & & \\ & a & -b \\ & b & a \end{pmatrix}.$$

     La exponencial e^{tA} viene dada por

     $$e^{tA} = P \begin{pmatrix} e^{t\lambda} & & \\ & e^{at}\cos(bt) & -e^{at}\sin(bt) \\ & e^{at}\sin(bt) & e^{at}\cos(bt) \end{pmatrix} P^{-1}.$$

     En el caso 3×3 no pueden salir valores propios complejos fuera de este caso.

2. Dos valores propios reales, λ_1, λ_2, con multiplicidades algebraicas $m_a(\lambda_1) = 1, m_a(\lambda_2) = 2$. El valor propio λ_1 no presenta problemas. Tomamos $v_1 \in \ker(A - \lambda_1 I_3)$, el cual tiene dimensión 1. Ahora estudiamos $\ker(A - \lambda_2 I_3)$, pudiendo tener dos posibilidades.

   - $\dim \ker(A - \lambda_2 I_3) = 2$. En este caso A es diagonalizable y tenemos dos bloques elementales de Jordan. Se tomarían $v_2, v_3 \in \ker(A - \lambda_2 I_3)$ linealmente independientes, la base $\mathcal{B} = \{v_1, v_2, v_3\}$ y la matriz $P = (v_1, v_2, v_3)$, obteniendo

$$A = PJP^{-1}, \qquad J = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_2 \end{pmatrix}.$$

     La exponencial e^{tA} viene dada por

$$e^{tA} = P \begin{pmatrix} e^{t\lambda_1} & & \\ & e^{t\lambda_2} & \\ & & e^{t\lambda_2} \end{pmatrix} P^{-1}.$$

   - $\dim \ker(A - \lambda_2 I_3) = 1$. En este caso necesitamos considerar el subespacio generalizado $\ker(A - \lambda_2 I_3)^2$, el cual necesariamente tiene dimensión 2. Sea $v_2 \in \ker(A - \lambda_2 I_3)^2 \setminus \ker(A - \lambda_2 I_3)$ y definimos $v_3 = (A - \lambda_2 I_3)v_2$. Entonces, $v_2, v_3 \in \ker(A - \lambda_2 I_3)^2$, son linealmente independientes y en particular $v_3 \in \ker(A - \lambda_2 I_3)$. Por último, definiendo de forma usual $P = (v_1, v_2, v_3)$, se tiene

$$A = PJP^{-1}, \qquad J = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & 1 & \lambda_2 \end{pmatrix}.$$

     La exponencial e^{tA} viene dada por

$$e^{tA} = P \begin{pmatrix} e^{t\lambda_1} & & \\ & e^{t\lambda_2} & \\ & te^{\lambda_2} & e^{t\lambda_2} \end{pmatrix} P^{-1}.$$

   - Un único valor propio λ. Si $\dim \ker(A - \lambda I_3) = 3$, la matriz es un múltiplo de la identidad y hemos acabado. En efecto, calculando P se tendría

$$A = PJP^{-1} = P(\lambda I_3)P^{-1} = \lambda I_3, \qquad e^{tA} = e^{t\lambda} I_3.$$

     Por tanto suponemos $\dim \ker(A - \lambda I_3) \leq 2$, teniendo que estudiar los subcasos siguientes.

     - $\dim \ker(A - \lambda I_3) = 1$. Existe un único bloque de Jordan y por tanto la torre de subespacios $(A - \lambda I_3)^k$ es de la forma

$$\{0\} \subset \ker(A - \lambda I_3) \subset \ker(A - \lambda I_3)^2 \subset \ker(A - \lambda I_3)^3 = \mathbb{R}^3.$$

       Sea pues $v_1 \in \ker(A - \lambda I_3)^3 \setminus \ker(A - \lambda I_3)^2$ y definimos $v_2 = (A - \lambda I_3)v_1,\ v_3 = (A - \lambda I_3)v_2$. En particular $v_3 \in \ker(A - \lambda I_3)$ y es vector propio de λ. Si definimos $P = (v_1, v_2, v_3)$ se tiene

$$A = PJP^{-1}, \qquad J = \begin{pmatrix} \lambda & & \\ 1 & \lambda & \\ & 1 & \lambda \end{pmatrix}.$$

La exponencial e^{tA} viene dada por

$$e^{tA} = P \begin{pmatrix} e^{t\lambda} & & \\ te^{t\lambda} & e^{t\lambda} & \\ \frac{t^2}{2}e^{t\lambda} & te^{\lambda} & e^{t\lambda} \end{pmatrix} P^{-1}.$$

- $\dim\ker(A-\lambda I_3) = 2$. En este último caso se tienen dos bloques de Jordan. Sea pues $v_1 \in \ker(A-\lambda I_3)^2 \setminus \ker(A-\lambda I_3)$ y definimos $v_2 = (A-\lambda I_3)v_1$. Estos dos vectores nos darán el primer bloque elemental de Jordan. Para el segundo bloque, basta completar a una base de $\ker(A-\lambda I_3)$ mediante algún v_3 linealmente independiente con v_2. Definiendo $P = (v_1, v_2, v_3)$ se tiene

$$A = PJP^{-1}, \qquad J = \begin{pmatrix} \lambda & & \\ 1 & \lambda & \\ & & \lambda \end{pmatrix}.$$

La exponencial e^{tA} viene dada por

$$e^{tA} = P \begin{pmatrix} e^{t\lambda} & & \\ te^{\lambda} & e^{t\lambda} & \\ & & e^{t\lambda} \end{pmatrix} P^{-1}.$$

3.6. Solución de la ecuación completa $x' = Ax + b(t)$

Consideremos ahora el problema no homogéneo,

$$x' = Ax + b(t), \tag{3.6.1}$$

con $A \in \mathcal{M}_n(\mathbb{R})$ y $b : I \subset \mathbb{R} \to \mathbb{R}$ continua. Este problema ya fue estudiado en la sección 3.1.3 cuando A era una aplicación $A(t)$. La diferencia ahora es que conocemos un método para obtener una matriz fundamental, a saber $\Phi(t) = e^{tA}$. En consecuencia, la solución general del sistema no homogéneo con A constante es

$$x = e^{tA}\lambda + e^{tA}\int e^{-tA}b(t)\,dt, \qquad \lambda \in \mathbb{R}^n.$$

Más aún, para la condición inicial $x(t_0) = x_0 \in \mathbb{R}^n$ sabemos que una matriz fundamental que cumple $\Phi(t_0) = I_n$ es $\Phi(t) = e^{(t-t_0)A}$, con lo que concluimos que la única solución del problema de Cauchy es

$$x = e^{(t-t_0)A}x_0 + e^{(t-t_0)A}\int_{t_0}^{t} e^{(t-s)A}b(s)\,ds.$$

Existe una alternativa para encontrar soluciones particulares de (3.6.1) si el término $b(t)$ tiene una expresión concreta: el *método de los coeficientes indeterminados*, cuya prueba se omite.

Teorema 3.6.1. *Supongamos que $b(t)$ es de la forma*

$$b(t) = e^{at}\left(p(t)\cos(bt) + q(t)\sin(bt)\right), \tag{3.6.2}$$

donde $p, q : I \to \mathbb{R}^n$ son polinomios en cada una de sus componentes de grado a lo sumo $k \geq 0$. Sea $\mu = a + bi$.

1. *Si μ no es valor propio de A, entonces (3.6.1) tiene una solución particular de la forma*
$$x_p(t) = e^{at}(r(t)\cos(bt) + s(t)\sin(bt)),$$
donde $r, s : I \to \mathbb{R}^n$ son polinomio en cada una de sus componentes de grado a lo sumo k.

2. *Si μ es valor propio de A, sea $D_l(a)$ o $E_l(a+bi)$ el bloque elemental de Jordan de mayor tamaño en la forma canónica de Jordan de A. Entonces, (3.6.1) tiene una solución particular de la forma*
$$x_p(t) = e^{at}(r(t)\cos(bt) + s(t)\sin(bt)),$$
donde $r, s : I \to \mathbb{R}^n$ son polinomio en cada una de sus componentes de grado a lo sumo $k+l$.

Ejemplo 3.10. Resolvamos el problema de Cauchy
$$\begin{cases} x' = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} x + \begin{pmatrix} \cos(2t) \\ -\sin(2t) \end{pmatrix}, \\ x(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \end{cases}$$
mediante el método de los coeficientes indeterminados y el del cálculo de la matriz exponencial.

1. Si lo hacemos mediante el cálculo de la exponencial, lo primero es calcular la exponencial de At. Los valores propios de A son $\mu = 1 \pm 2i$ y por tanto buscamos $w = (w_1, w_2) \in \mathbb{C}^2$ tal que $Aw = \mu w$. Esto nos da un espacio de dimensión compleja 1, generado por el vector $(1, i)$. Ahora bien, tomando la base real
$$\{(0,1), (1,0)\},$$
definiríamos $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Esta matriz no es más que la operación de permutación de filas para matrices 2×2 y por tanto se tiene que $A = PJP^{-1} = J$. Esto nos dice que A coincide con su forma de Jordan respecto de esta base. Por tanto, una matriz fundamental del sistema homogéneo es
$$\Phi(t) = e^{tA} = e^t \begin{pmatrix} \cos(2t) & -\sin(2t) \\ \sin(2t) & \cos(2t) \end{pmatrix}. \tag{3.6.3}$$
Por una parte,
$$e^{tA}x_0 = e^t \begin{pmatrix} \cos(2t) & -\sin(2t) \\ \sin(2t) & \cos(2t) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} e^t\cos(2t) \\ -e^t\sin(2t) \end{pmatrix}.$$
Por otra parte,
$$e^{(t-s)A}b(s) = e^{t-s} \begin{pmatrix} \cos(2(t-s)) & -\sin(2(t-s)) \\ \sin(2(t-s)) & \cos(2(t-s)) \end{pmatrix} \begin{pmatrix} \cos(2s) \\ -\sin(2s) \end{pmatrix} = \begin{pmatrix} e^{t-s}\cos(2s) \\ -e^{t-s}\sin(2s) \end{pmatrix}.$$
En consecuencia,
$$\int_0^t e^{(t-s)A}b(s)\,ds = \int_0^t \begin{pmatrix} e^{t-s}\cos(2t) \\ -e^{t-s}\sin(2t) \end{pmatrix} ds = (e^t - 1)\begin{pmatrix} \cos(2t) \\ -\sin(2t) \end{pmatrix}.$$
Concluimos que la solución del problema es
$$x(t) = (2e^t - 1)\begin{pmatrix} \cos(2t) \\ -\sin(2t) \end{pmatrix}.$$

2. Si lo hacemos mediante el método de los coeficientes indeterminados, observamos el término independiente $b(t)$ y comprobamos que es de la forma (3.6.2) para los valores $a = 0,\ b = 2,\ p(t) = (1,0),\ q(t) = (0,1)$. Como $2i$ no es valor propio de la matriz de coeficientes, concluimos que una solución particular es de la forma

$$x_p = \begin{pmatrix} A \\ B \end{pmatrix} \cos(2t) + \begin{pmatrix} C \\ D \end{pmatrix} \sin(2t),$$

con $A, B, C, D \in \mathbb{R}$. Si forzamos a que x_p sea solución particular obtenemos $A = -1,\ B = C = 0,\ D = 1$, concluyendo

$$x_p = \begin{pmatrix} -\cos(2t) \\ \sin(2t) \end{pmatrix}.$$

La solución general del sistema es

$$x(t) = \Phi(t)\lambda + x_p(t), \qquad \lambda \in \mathbb{R}^n,$$

siendo $\Phi(t)$ la matriz fundamental (3.6.3). Por tanto,

$$x(t) = e^t \begin{pmatrix} \cos(2t) & -\sin(2t) \\ \sin(2t) & \cos(2t) \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} + \begin{pmatrix} -\cos(2t) \\ \sin(2t) \end{pmatrix}, \qquad \lambda_1, \lambda_2 \in \mathbb{R}.$$

Si imponemos la condición de Cauchy $x(0) = (1,0)$, obtenemos el vector $\lambda = (2,0)$ y la única solución es

$$x(t) = (2e^t - 1) \begin{pmatrix} \cos(2t) \\ -\sin(2t) \end{pmatrix}.$$

3.7. Ecuación lineal de orden n uno-dimensional con coeficientes constantes

Un caso particular surge al tomar funciones constantes $a_0, ..., a_{n-1}$ en la EDO (3.2.1), esto es

$$x^{n)} + a_{n-1}x^{n-1)} + \cdots + a_1 x' + a_0 x = b(t), \tag{3.7.1}$$

con $a_{n-1}, ..., a_0 \in \mathbb{R}$ y $b : I \to \mathbb{R}$ continua.

Supongamos inicialmente $b(t) \equiv 0$, es decir estudiamos la EDO homogénea. Como otras tantas cuestiones matemáticas, Leonhard Euler abordó el estudio de esta EDO, y para ello buscó soluciones de la forma $x = e^{\lambda t}$ para cierto $\lambda \in \mathbb{R}$. Teniendo en cuenta que $x^{k)} = \lambda^k e^{\lambda t}$, obtenemos

$$a_0 e^{\lambda t} + a_1 \lambda e^{\lambda t} + \cdots + a_{n-1}\lambda^{n-1} e^{\lambda t} + \lambda^n e^{\lambda t} = 0.$$

Como $e^{\lambda t}$ no se anula, se tiene que la función $x = e^{\lambda t}$ es solución si y sólo si λ es solución del polinomio

$$a_0 + \lambda a_1 + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n = 0.$$

Toda raíz real de esta ecuación proporcionará una solución exponencial de la EDO, pero es posible que encontremos raíces complejas $\lambda = a \pm bi$. En tal caso, Euler observó a raíz de la identidad

$$e^{it} = \cos t + i \sin t$$

que las funciones $e^{at}\cos(bt)$ y $e^{at}\sin(bt)$ son igualmente soluciones de la EDO. Vamos a formalizar esta brillante idea de Euler para obtener todas las soluciones de (3.7.1) y comprobar que Euler estaba bastante bien encaminado.

Comenzamos destacando que la EDO (3.7.1) sabemos resolverla perfectamente sin más que transformarla en el sistema $x' = Ax$, con A dada por (3.2.2), obtener la forma canónica de Jordan de A y calcularle la exponencial. No obstante, en este caso concreto de coeficientes constantes, no va a ser necesario el pasar al sistema matricial equivalente y realizar el cálculo de la exponencial. Sea pues

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \ddots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}, \tag{3.7.2}$$

y calculemos su polinomio característico, $\det(A - \lambda I_n)$. Para ello, desarrollemos el determinante de $A - \lambda I_n$,

$$\det(A - \lambda I_n) = \begin{vmatrix} -\lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & -\lambda & 1 & 0 & \ddots & \vdots \\ 0 & 0 & -\lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 0 & -\lambda & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-2} & -a_{n-1} - \lambda \end{vmatrix},$$

por la última fila, obteniendo

$$(-1)^{n+1}(-a_0)\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\lambda & 1 & 0 & \ddots & \vdots \\ 0 & -\lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -\lambda & 1 \end{vmatrix} + (-1)^{n+1}a_1\begin{vmatrix} -\lambda & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \ddots & \vdots \\ 0 & -\lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -\lambda & 1 \end{vmatrix} + (-1)^{n+1}(-a_2)\begin{vmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ 0 & -\lambda & 0 & \ddots & \vdots \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -\lambda & 1 \end{vmatrix}$$

$$+ \cdots + a_{n-2}\begin{vmatrix} -\lambda & 1 & 0 & 0 & 0 \\ 0 & -\lambda & 1 & 0 & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -\lambda & 1 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} - (a_{n-1} + \lambda)\begin{vmatrix} -\lambda & 1 & 0 & 0 & 0 \\ 0 & -\lambda & 1 & 0 & \ddots \\ 0 & 0 & -\lambda & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & 0 & 0 & 0 & -\lambda \end{vmatrix}.$$

Ahora bien, observemos que el primer determinante vale 1, el segundo vale $-\lambda$, el tercero vale λ^2 y así sucesivamente. En general, el determinante k-ésimo vale $(-1)^k \lambda^k$. Concluimos pues

$$\det(A - \lambda I_n) = (-1)^{n+2}\left(a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_{n-2}\lambda^{n-2} + a_{n-1}\lambda^{n-1} + \lambda^n\right).$$

El desarrollo anterior motiva la siguiente definición

Definición 3.7.1. Dada la EDO lineal homogénea de orden n, (3.7.1), se define su polinomio característico como el polinomio

$$p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_{n-2}\lambda^{n-2} + a_{n-1}\lambda^{n-1} + \lambda^n. \tag{3.7.3}$$

Referente a la forma canónica de Jordan de A, podemos calcularla sin necesidad de realizar cálculos de matrices de paso ni de subespacios de operadores nilpotentes. En efecto, como las raíces de $p(\lambda)$ coinciden con los valores propios de A dada por (3.7.2), observamos que cada raíz $\lambda = a+bi$ de $p(\lambda)$ define una única caja $D_r(\lambda)$ ó $E_r(\lambda)$, dependiendo de si λ es real o complejo, en la forma canónica de Jordan de A. Para ello notemos que el operador $A - \lambda I_n$ tiene rango al menos $n-1$, ya que la submatriz obtenida al suprimir la última fila y la primera columna tiene determinante

$$\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\lambda & 1 & 0 & \ddots & \vdots \\ 0 & -\lambda & 1 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -\lambda & 1 \end{vmatrix} = 1 \neq 0,$$

y por tanto el rango de $A - \lambda I_n$ es al menos $n-1$. Ahora bien, no puede ser n por ser λ valor propio y en consecuencia $\dim\ker(A - \lambda I_n) \geq 1$. Esto nos dice que la multiplicidad geométrica de λ es 1 y sólo hay una caja de Jordan, precisamente de tamaño $m_a(\lambda)$. Como quiera que el subespacio generalizado $E(A, \lambda)$ tiene dimensión $m_a(\lambda)$, concluimos que la cadena de subespacios

$$\ker(A - \lambda I_n) \subset \ker(A - \lambda I_n)^2 \subset \cdots \subset \ker(A - \lambda I_n)^p = \ker(A - \lambda I_n)^{p+1} = \cdots$$

se estabiliza exactamente para $p = m_a(\lambda)$ y en cada salto $\ker(A - \lambda I_n)^k \subset \ker(A - \lambda I_n)^{k+1}$ la dimensión aumenta exactamente en 1. Para obtener una base de $E(A, \lambda)$ que represente a la caja $D_r(\lambda)$ ó $E_r(\lambda)$, basta tomar $v_1 \in \ker(A - \lambda I_n)^{m_a(\lambda)} \setminus \ker(A - \lambda I_n)^{m_a(\lambda)-1}$ y definir por recurrencia

$$v_k = (A - \lambda I_n)v_{k-1}, \qquad k = 2, ..., m_a(\lambda).$$

Por último, si λ es complejo entonces tendríamos que tomar $\{\Im v_k, \Re v_k\}$ como base para obtener la forma real del bloque $E_r(\lambda)$.

El siguiente resultado es consecuencia del Corolario 3.4.7 y de la existencia de un único bloque de Jordan para cada valor propio λ.

Teorema 3.7.2. *Un sistema fundamental de soluciones de la EDO homogénea* (3.7.1) *está formado por*

$$\{t^k e^{at}\cos(bt), t^k e^{at}\sin(bt)\}, \tag{3.7.4}$$

con $a + bi,\ b \geq 0$, *recorriendo el conjunto de los valores propios y siendo* $k = 0, ..., m_a(\lambda) - 1$.

Demostración. Sean $\lambda_1, ..., \lambda_r$ los valores propios de A. Si λ_j es real, entonces $b = 0$ y tenemos un total de $m_a(\lambda)$ funciones independientes entre ellas. Si λ_j es complejo entonces necesariamente $\overline{\lambda_j}$ es también valor propio de A, las multiplicidades de λ_j y $\overline{\lambda_j}$ coinciden y su suma es el doble de la multiplicidad de λ_j. En particular, se puede imponer $b \geq 0$, ya que las funciones para $b < 0$ son dependientes con las de $b > 0$. Esto nos da un total de $2m_a(\lambda_j)$ funciones, sin contar con las que definen el valor propio $\overline{\lambda_j}$. Como la suma de las multiplicidades es precisamente n, en total tenemos n funciones linealmente independientes.

Falta comprobar que, en efecto, son soluciones de la EDO (3.7.1). Fijemos un valor propio $\lambda = a + bi$ de A dada por (3.2.2), o de forma equivalente una raíz del polinomio característico (3.7.3), y sea $m_a(\lambda)$ su multiplicidad algebraica. Para cada $k = 0, ..., m_a(\lambda) - 1$ definimos $x_k(t) = t^k e^{\lambda t}$. Para $k = 0$ es inmediato que se tiene

$$x_0^{j)}(t) = \lambda^j x_0(t), \qquad j = 0, .., n,$$

y por tanto al sustituir en la EDO homogénea (3.7.1) concluimos

$$x_0(t)(\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + \lambda^n a_1 + a_0) = x_0(t)p(\lambda) = 0.$$

Estudiemos ahora $x_1(t) = te^{t\lambda}$. Es inmediato comprobar que se cumple $x_1^{j)}(t) = j\lambda^{j-1}e^{t\lambda} + \lambda^j x_1(t)$. Sustituyendo $x_1(t)$ en la EDO homogénea y agrupando convenientemente se tiene

$$te^{\lambda t}(a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n) + e^{\lambda t}(a_1 + 2a_2\lambda + \cdots + (n-1)a_{n-1}\lambda^{n-2} + n\lambda^{n-1}) = 0,$$

donde el segundo paréntesis vale cero por ser λ raíz de multiplicidad $m_a(\lambda) - 1$ y por tanto $p^{j)}(\lambda) = 0$ para todo $j = 0, ..., m_a(\lambda) - 2$. Por inducción es ahora fácil ver que $x_k(t)$ es solución para todo $k = 0, ..., m_a(\lambda) - 1$. $\square$

Ejemplo 3.11. Consideremos la EDO

$$x^{iv)} + 4x''' + 5x'' + 4x' + 4x = 0.$$

Su polinomio característico es

$$p(\lambda) = \lambda^4 + 4\lambda^3 + 5\lambda^2 + 4\lambda + 4 = 0,$$

el cual tiene como raíces -2 (doble) y $\pm i$. Su solución general es de la forma

$$x = c_1 e^{-2t} + c_2 te^{-2t} + c_3 \cos t + c_4 \sin t$$

3.7.1. Ecuación completa. Método de los coeficientes indeterminados

Queda sólo pendiente considerar de forma particular la EDO completa

$$x^{n)} + a_{n-1}x^{n-1)} + \cdots + a_1 x' + a_0 x = b(t). \tag{3.7.5}$$

Poco hay que contar. De la ecuación homogénea podemos calcular el polinomio característico y obtener n funciones independientes, $x_1, ..., x_n$. Para obtener una solución particular x_p bastaría aplicar el método de variación de las constantes explicado en 3.2.1, ya que conocemos una matriz fundamental del sistema correspondiente. No obstante, se puede aplicar igualmente una versión del método de los coeficientes indeterminados introducido en 3.6.

Teorema 3.7.3. *Supongamos $b(t) = e^{at}(p(t)\cos(bt) + q(t)\sin(bt))$, donde p, q son polinomios de grado a lo sumo $k \geq 0$. Sea $\lambda = a + bi$.*

1. *Si λ no es raíz del polinomio característico* (3.7.3)*, entonces* (3.7.5) *tiene una solución particular de la forma*
$$e^{at}(r(t)cos(bt) + s(t)\sin(bt)),$$
con r, s polinomios de grado a lo sumo k.

2. *Si λ es raíz del polinomio característico* (3.7.3) *de multiplicidad l, entonces* (3.7.5) *tiene una solución particular de la forma*
$$t^l e^{at}(r(t)\cos(bt) + s(t)\sin(bt)),$$
con r, s polinomios de grado a lo sumo k.

Para finalizar, enunciamos un principio de superposición de soluciones que será de utilidad para aplicar el método de los coeficientes indeterminados. La prueba se deja como ejercicio por su simpleza.

Teorema 3.7.4. *Consideremos la EDO*
$$x^{n)} + a_{n-1}(t)x^{n-1)} + \cdots + a_1(t)x' + a_0(t)x = b_1(t) + b_2(t) + \cdots + b_r(t), \tag{3.7.6}$$
con $b_1, ..., b_r : I \subset \mathbb{R} \to \mathbb{R}$ continuas. Para cada $k = 1, .., r$, sea x_k solución particular de
$$x^{n)} + a_{n-1}(t)x^{n-1)} + \cdots + a_1(t)x' + a_0(t)x = b_k(t).$$
Entonces, $x_p = x_1 + \cdots + x_r$ es solución particular de (3.7.6)*.*

Ejemplo 3.12. Consideremos la EDO
$$x''' - 4x' = t + 3\cos t + e^{-2t}.$$
Su polinomio característico es
$$p(\lambda) = \lambda^3 - 4\lambda,$$
que tiene como raíces $0, 2, -2$. Su solución general es de la forma
$$x = c_1 + C_2e^{2t} + c_3e^{-2t}.$$
Para hallar una solución particular, aplicamos el principio de superposición dado por el Teorema 3.7.4. Por tanto, debemos encontrar soluciones particulares de las EDOs
$$x''' - 4x' = t, \qquad x''' - 4x' = 3\cos t, \qquad x''' - 4x' = e^{-2t}.$$

- En el primer caso $b(t) = t$ y por tanto $a = b = 0$, $p(t) = t$ y $q(t) = 0$. Como 0 es raíz de $p(\lambda)$ de multiplicidad 1, debemos buscar la solución particular de la forma
$$x_{p_1} = t(At + B).$$
Sustituyendo en la EDO se tiene $A = -1/8$ y $B = 0$, por lo que $x_{p_1} = -t^2/8$.

- Aquí $b(t) = 3\cos t$ y por tanto $a = 0,\ b = 3$, $p(t) = 3$ y $q(t) = 0$. Como $3i$ no es raíz de $p(\lambda)$, la solución particular será de la forma
$$x_{p_2} = A\cos t + B\sin t,$$
y sustituyendo en la EDO concluimos $A = 0, B = -3/5$.

- Por último, $b(t) = e^{-2t}$ y se tiene $a = -2,\ b = 0$. Como -2 también es raíz de $p(\lambda)$, la solución particular será de la forma

  $$x_{p_3} = Ate^{-2t}.$$

  Sustituyendo en la EDO concluimos $A = 1/8$.

Tras estas discusiones podemos afirmar que la función

$$x_p = -\frac{t^2}{8} - \frac{3}{5}\sin t + \frac{t}{8}e^{-2t}$$

es solución particular.

Ejercicios tema 3

1. Encontrar funciones $a, b \in C(I)$ de modo que t, t^2 sean soluciones de la EDO lineal
$$x'' + a(t)x' + b(t)x = 0.$$
¿Puede el intervalo ser toda la recta real?

2. Construir un sistema de ecuaciones diferenciales que tenga como soluciones a las funciones
$$x_1(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad x_2(t) = \begin{pmatrix} \arctan t \\ \frac{1}{1+t^2} \end{pmatrix}$$
¿Es este sistema el resultado de reducir el orden de una EDO lineal de orden superior?

3. Dada $\Phi : I \subset \mathbb{R} \to \mathcal{M}_n(\mathbb{R})$ de clase C^1, probar que si $\det \Phi(t) \neq 0$ para todo $t \in I$ entonces $\Phi(t)$ es matriz fundamental de algún sistema homogéneo de la forma $x' = A(t)x$.

4. Se consideran las funciones $\sin t, e^t, e^{-t}$. Determinar el menor entero positivo n para el cual tales funciones sean solución de una EDO lineal de orden n. ¿Es única la EDO? ¿Y si se impone que el dominio de definición sea $\mathbb{R}$?

5. Sean $x_1, ..., x_n \in C^n(I)$ tales que
$$W(t; x_1, ..., x_n) \neq 0, \qquad \forall t \in I.$$
Demostrar que existe una única ecuación lineal homogénea de orden n y con coeficiente 1 para la derivada n-ésima,
$$x^{n)} + a_{n-1}(t)x^{n-1)} + \cdots + a_0(t)x = 0,$$
que tiene a $x_1, ..., x_n$ como soluciones. Dar una expresión de los coeficientes $a_k(t)$ en términos de x_k.

6. Consideremos la EDO lineal de orden n homogénea,
$$x^{n)} + a_{n-1}(t)x^{n-1)} + \cdots + a_1(t)x' + a_0(t)x = 0,$$
y sea x una solución no trivial. Probar que los ceros de x son aislados.

7. Sea $A : \mathbb{R} \to \mathcal{M}_n(\mathbb{R})$ continua y supongamos que existe $M > 0$ tal que
$$||A(t)|| \leq M, \qquad \forall t \in \mathbb{R}.$$
Sea x solución de $x' = A(t)x$.
   a) Dado $\lambda \in \mathbb{R}$, obtener la ecuación diferencial satisfecha por $y_\lambda(t) = e^{-\lambda t}x(t)$.
   b) Probar que la función $\varphi(t) = ||y_\lambda(t)||^2$ es derivable y que
$$-(M + \lambda)\varphi(t) \leq \frac{1}{2}\varphi'(t) \leq (M - \lambda)\varphi(t), \qquad \forall t \in \mathbb{R}.$$
   c) Deducir del apartado anterior la existencia de intervalos de valores de λ para los que $\lim_{t\to\infty} ||y_\lambda(t)||$ es o bien 0 o bien ∞.

8. Consideremos la EDO lineal de orden n homogénea,

$$x^{n)} + a_{n-1}(t)x^{n-1)} + \cdots + a_1(t)x' + a_0(t)x = 0,$$

definida en cierto $I \subset \mathbb{R}$, y sea x_1 una solución que no se anula en $J \subset I$.

a) Probar que el cambio de variable $x(t) = z(t)x_1(t)$ reduce la ecuación en J a una EDO lineal homogénea de orden $n-1$ en la variable $w(t) = z'(t)$.

b) Si tal EDO homogénea en w tiene a $w_2, ..., w_n$ como sistema fundamental y definimos $z_k(t) = \int w_k(t)$, probar que $x_1, z_2x_1, ..., z_nx_1$ es sistema fundamental en J de la EDO de partida.

c) Consideremos el caso $n = 2$,

$$x'' + a_1(t)x' + a_0(t)x = 0.$$

Si x_1 es solución de esta EDO que no se anula en J, probar que x_1 y

$$x_2(t) = x_1(t) \int \frac{1}{x_1(t)^2} e^{-\int a_1(t)\,dt}\,dt$$

son un sistema fundamental de la EDO en x_2.

d) Como aplicación, encontrar la solución general de

$a)$ $(1+t^2)x'' - 2tx' + 2x = 0,$

$b)$ $(1+t^2)x'' + (t - 2\sqrt{1+t^2})x' + x = 0,$ $\qquad x(t) = e^{\text{arcsinh}(t)}$ es solución,

$c)$ $x'' - 2\tanh t x' + (1 - 2/\cosh^2 t)x = 0.$

9. Se considera el sistema $x' = A(t)x + b(t)$, donde $A(t), b(t)$ son periódicas de periodo $T > 0$.

a) Probar que si x es solución, entonces $z(t) = x(t+T)$ es también solución.

b) Probar que una solución x es T-periódica si y sólo si $x(t_0) = x(t_0+T)$ para algún $t_0 \in \mathbb{R}$.

10. Consideremos el sistema $x' = A(t)x$.

a) Si $A(t)$ se descompone como

$$A(t) = f(t)I_n + g(t)B,$$

con $f, g : I \subset \mathbb{R} \to \mathbb{R}$ y B diagonalizable en $\mathbb{R}$, encontrar una expresión para la solución de $x' = A(t)x$ en términos de $f(t), g(t)$ y la exponencial de B.

b) Usar lo anterior para resolver el sistema

$$\begin{cases} x' = (t - e^t)x + e^t y, \\ y' = e^t x + ty. \end{cases}$$

11. Resolver el sistema $x' = Ax$, donde la matriz A viene dada en cada caso por:

a) $\begin{pmatrix} 7 & -1 & 6 \\ -10 & 4 & -12 \\ -2 & 1 & -1 \end{pmatrix}$ b) $\begin{pmatrix} 0 & 2 & -1 \\ -1 & 3 & 0 \\ 1 & -1 & 3 \end{pmatrix}$ c) $\begin{pmatrix} 0 & -1 & 2 \\ 1 & -2 & 2 \\ 2 & -2 & 1 \end{pmatrix}$

d) $\begin{pmatrix} 3 & 6 & 2 \\ -1 & -2 & -1 \\ 1 & 3 & 2 \end{pmatrix}$ e) $\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & 0 & 0 \\ 1 & 1 & 2 & 1 \\ -1 & -1 & 1 & 2 \end{pmatrix}$ f) $\begin{pmatrix} 0 & 0 & 0 & -8 \\ 1 & 0 & 0 & 16 \\ 0 & 1 & 0 & -14 \\ 0 & 0 & 1 & 6 \end{pmatrix}$

g) $\begin{pmatrix} 3 & 2 & 1 & 1 \\ -1 & 0 & -1 & -1 \\ 0 & 1 & 3 & 1 \\ 1 & 1 & 0 & 2 \end{pmatrix}$ h) $\begin{pmatrix} 1 & -2 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -2 \\ 0 & 0 & 1 & -1 \end{pmatrix}$ i) $\begin{pmatrix} 3 & 1 & -1 & -1 \\ 0 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$

j) $\begin{pmatrix} 3 & -4 & 2 & 0 \\ 1 & 2 & 1 & 1 \\ 0 & 1 & 2 & 1 \\ -1 & 2 & -2 & 1 \end{pmatrix}$

12. Resolver las siguientes EDOs.

a) $x'' - 3x' + 7x = 5te^{2t}$. b) $x'' + 4x = 5\sin(3t) + \cos(3t) + \sin(2t)$.

c) $x'' - 2x' + 3x = t^3 + \sin t$. d) $x''' - 6x'' - 6x + 11 = 2te^{-t}$.

e) $x'' + 4x' + 4x = t^{-2}e^{-2t}$. f) $x'' - 3x' + 2x = 1/(1 + e^{-t})$.

13. Supongamos que $A \in \mathcal{M}_4(\mathbb{R})$ satisface

$$\begin{aligned} A(1,2,0,0)^T &= (2,2,0,0)^T, \\ A(1,0,0,0)^T &= (2,1,1,3)^T, \\ A(1,1,1,3)^T &= (1,1,1,3)^T, \\ A(1,1,2,0)^T &= (1,1,2,0)^T. \end{aligned}$$

Resolver el sistema $x' = Ax$ sin calcular explícitamente la matriz A.

14. Sea $A \in \mathcal{M}_n(\mathbb{R})$. Si A tiene algún valor propio imaginario puro, deducir que existen soluciones periódicas no triviales del sistema $x' = Ax$. Calcular dicho periodo en términos del valor propio.

15. Dar una condición sobre los valores propios de $A \in \mathcal{M}_n(\mathbb{R})$ para que el sistema $x' = Ax$ tenga alguna solución polinómica. Dar una cota óptima del grado de tal polinomio.

16. Sea $A \in \mathcal{M}_7(\mathbb{R})$ con valores propios $\{1, \pm i\}$. Sabemos además se tiene la cadena de subespacios

$$\{0\} \subset \ker(A - iI_7) \subset \ker(A - iI_7)^2 \subset \ker(A - iI_7)^3,$$

donde todas las inclusiones se entienden estrictas. Hallar e^J, donde J es la forma canónica de Jordan de A. Deducir que el espacio de soluciones periódicas de $x' = Ax$ tiene dimensión 2.

17. Sea $A \in \mathcal{M}_7(\mathbb{R})$, con valores propios $\lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 3$ con multiplicidades algebraicas y geométricas satisfaciendo

$$\begin{aligned} m_a(1) &= 2, & m_g(1) &= 1, \\ m_a(2) &= 4, & m_g(2) &= 2, \\ m_a(3) &= 1, & m_g(3) &= 1. \end{aligned}$$

¿Se puede deducir de manera unívoca la forma canónica de Jordan de A? Indicar la información necesaria para determinar J en términos de sus subespacios propios generalizados.

18. Dada S una superficie en $\mathbb{R}^3$, se definen la *curvatura de Gauss* y la *curvatura media* de S como

$$K = \kappa_1 \kappa_2, \qquad H = \frac{1}{2}(\kappa_1 + \kappa_2),$$

donde κ_1, κ_2 son las *curvaturas principales* de S. Sea

$$\alpha : I \subset \mathbb{R} \to \mathbb{R}^3, \qquad \alpha(t) = (x(t), 0, z(t)), \quad x(t) > 0,\ \forall t \in I,$$

una curva plana en $\mathbb{R}^3$ y sea S la superficie obtenida al rotar α alrededor del eje z. Si $\alpha(t)$ está *parametrizada por el arco*, esto es $||\alpha'(t)|| = 1$, las *curvaturas principales* de S vienen dadas por las funciones

$$\kappa_1(t) = x'z'' - x''z', \qquad \kappa_2 = \frac{z'}{x}.$$

a) Usando la condición de estar parametrizada por el arco, escribir κ_1, κ_2 únicamente en términos de x y sus derivadas.

b) Clasificar las superficies rotacionales con curvatura de Gauss constante $K = K_0 \in \mathbb{R}$, en términos de K_0.

c) Clasificar las superficies rotacionales con curvatura media $H = 0$.

Capítulo 4

Teoría fundamental

En los capítulos anteriores hemos realizado un estudio acerca de la integración de ecuaciones diferenciales ordinarias. No obstante, y tal y como se ha puesto de manifiesto, el número de EDOs cuya solución se puede obtener, incluso únicamente de forma implícita, es más bien reducido. Sin ir más lejos, sabemos por el Teorema Fundamental del Cálculo que la EDO

$$x' = f(t),$$

siendo $f : I \subset \mathbb{R} \to \mathbb{R}$ continua, tiene existencia, aunque en la mayoría de las veces el cálculo de una primitiva explícita de f se antoja imposible. En general, dada una EDO

$$x' = f(t, x),$$

nos gustaría poder garantizar existencia y unicidad para el problema de Cauchy asociado a una condición inicial $x(t_0) = x_0$. Salvo en casos muy concretos estudiados en el tema 2, como las EDOs lineales (o reducibles a éstas), o en EDOs que se pueden tratar invocando el Teorema de la función implícita, no disponemos de herramientas para garantizar la existencia y unicidad de un problema de Cauchy. En el tema 3 dimos el primer resultado de existencia y unicidad para una EDO lineal matricial de la forma

$$x' = A(t)x + b(t),$$

pero poco más sabemos si de nuevo abordamos una EDO arbitraria $x' = f(t, x)$. El último tema de la asignatura se centra en la teoría clásica fundamental de existencia y unicidad de ecuaciones diferenciales ordinarias y pretende dar respuesta a estas cuestiones.

4.1. Existencia y unicidad local

A lo largo de este tema, y salvo que se especifique lo contrario en algún caso particular, denotaremos por $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ a un abierto, por $(t, x) \in \Omega$ a sus coordenadas y por $f : \Omega \to \mathbb{R}^n$ a una función continua. Dado $(t_0, x_0) \in \Omega$, el objeto de estudio de este tema es el problema de Cauchy

$$\begin{cases} x' = f(t, x) \\ x(t_0) = x_0. \end{cases} \tag{4.1.1}$$

Si denotamos por $f = (f_1, ..., f_n)$, con cada $f_k : \Omega \to \mathbb{R}$ continua, entonces (4.1.1) es realmente un sistema de n ecuaciones con vector incógnita de n componentes, $x = (x_1, ..., x_n)$. En el caso $n = 1$

obtenemos simplemente una EDO en forma normal. Ya sabemos del tema 1 que toda EDO escalar de orden n,

$$x^{n)} = F(t, x, x', ..., x^{n-1)}),$$

se puede reducir a un sistema equivalente de primer orden. Así pues, el problema de Cauchy (4.1.1) pretende considerar la situación más general posible dentro de toda la casuística ya estudiada.

En el tema 1 probamos que el problema (4.1.1) admite una versión integral que será de gran utilidad, la cual recordamos.

Lema 4.1.1 (Condición integral de Volterra)**.** *Una función continua $x : I \subset \mathbb{R} \to \mathbb{R}^n$ es solución de* (4.1.1) *si y sólo si para todo $t \in I$ se tiene $(t, x(t)) \in \Omega$ y*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds,$$

donde la integral anterior se entiende componente a componente de forma vectorial,

$$\int_a^b f(s, x(s))\, ds = \left(\int_a^b f_1(s, x(s))\, ds, \dots, \int_a^b f_n(s, x(s))\, ds \right).$$

El siguiente resultado es más propio de un curso de Análisis y se incluye para hacer las notas lo más autocontenidas posibles.

Proposición 4.1.2. *Si denotamos por $||\cdot||$ a la norma euclídea usual, y $f : [a, b] \to \mathbb{R}^n$ es continua, entonces se satisfacen las siguientes propiedades.*

1. $\left|\left| \int_a^b f(t)\, dt \right|\right| \leq \int_a^b ||f(t)||\, dt.$
2. $\left|\left| \int_a^b f(t)\, dt \right|\right| \leq ||f||_\infty |b - a|$, *siendo* $||f||_\infty = \max\{||f(t)|| : t \in [a, b]\}$.

Demostración. 1. La demostración de la primera propiedad hace uso de la desigualdad de Cauchy-Schwarz, escrita en forma de producto escalar como

$$\langle x, y\rangle^2 \leq ||x||^2 ||y||^2,$$

o de forma equivalente en coordenadas, si $(x_1, ...x_n), (y_1, ..., y_n) \in \mathbb{R}^n$,

$$\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leq \left(\sum_{k=1}^{n} x_k^2\right)\left(\sum_{k=1}^{n} y_k^2\right).$$

La norma de la integral de f es

$$\left|\left| \int_a^b f(t)\, dt \right|\right|^2 = \sum_{k=1}^{n} \left(\int_a^b f_k(t)\, dt \right)^2 = \sum_{k=1}^{n} \int_a^b f_k(t)\, dt \int_a^b f_k(t)\, dt = \int_a^b \sum_{k=1}^{n} z_k f_k(t)\, dt,$$

donde en la última igualdad hemos llamado $z_k = \int_a^b f_k(t)\, dt$ y usado la aditividad de la integral. Por otra parte, la desigualdad de Cauchy-Schwarz asegura

$$\sum_{k=1}^{n} z_k f_k(t) \leq \sqrt{\sum_{k=1}^{n} z_k^2} \sqrt{\sum_{k=1}^{n} f_k(t)^2} = \left|\left| \int_a^b f(t)\, dt \right|\right| ||f(t)||,$$

y en consecuencia

$$\left|\left|\int_a^b f(t)\,dt\right|\right| \leq \int_a^b ||f(t)||\ dt.$$

2. La segunda propiedad es inmediata de la primera y del hecho de que por definición,

$$||f(t)|| \leq ||f||_\infty.$$

Ahora basta tomar integrales en ambos lados y observar que el de la derecha es constante para concluir con lo deseado.

□

Respecto de la primera desigualdad, un comentario. Podemos sustituir la norma euclídea por cualquier norma de $\mathbb{R}^n$, puesto que son todas equivalentes. El motivo de hacer la prueba para la norma euclídea es más bien para recordar al alumno la desigualdad de Cauchy-Schwarz y mostrar este tipo de argumentos que se repiten en multitud de argumentos lógicos. En particular, si la norma $||\cdot||$ usada en $\mathbb{R}^n$ es la del máximo,

$$||x||_\infty = \text{máx}\{|x_1|, ..., |x_n|\}, \qquad x = (x_1, ..., x_n) \in \mathbb{R}^n,$$

entonces la primera desigualdad es inmediata de demostrar. En efecto,

$$\begin{aligned}\left|\left|\int_a^b f(t)\,dt\right|\right|_\infty &= \underset{k=1,\dots,n}{\text{máx}} \left\{\left|\int_a^b f_k(t)\,dt\right|\right\} \leq \underset{k=1,\dots,n}{\text{máx}} \left\{\int_a^b |f_k(t)|\ dt\right\} \\ &= \int_a^b \underset{k=1,\dots,n}{\text{máx}} \{|f_k(t)|\} = \int_a^b ||f(t)||_\infty\,dt.\end{aligned}$$

Obviamente la segunda siempre es consecuencia trivial de la primera.

Por último, necesitamos la siguiente versión del Teorema del Valor Medio para aplicaciones vectoriales.

Lema 4.1.3 (Teorema del Valor Medio para aplicaciones vectoriales)**.** *Sea* $f : [a, b] \to \mathbb{R}^n$ *de clase* C^1 *en* $[a, b]$*. Entonces, existe* $c \in (a, b)$ *tal que*

$$||f(b) - f(a)|| \leq ||f'(c)||(b - a).$$

Demostración. Definimos $\varphi(t) = \langle f(b)-f(a), f(t)\rangle$, la cual es una función real de variable real, definida en $t \in [a, b]$, diferenciable por serlo f. El Teorema del Valor Medio usual asegura la existencia de $c \in (a, b)$ tal que

$$\varphi(b) - \varphi(a) = \varphi'(c)(b - a).$$

Ahora bien, $\varphi(b) - \varphi(a) = ||f(b) - f(a)||^2$ y $\varphi'(c) = \langle f(b) - f(a), f'(c)\rangle$, lo que implica

$$||f(b) - f(a)||^2 = \langle f(b) - f(a), f'(c)\rangle(b - a) \leq ||f(b) - f(a)||\,||f'(c)||(b - a).$$

donde en la última desigualdad hemos usado Cauchy-Schwarz. Si $f(b) = f(a)$, la desigualdad se da de forma trivial. En caso contario, dividimos por $||f(b) - f(a)||$ y concluimos

$$||f(b) - f(a)|| \leq ||f'(c)||(b - a).$$

□

4.1.1. El Teorema del punto fijo de Banach

Existen diferentes formas de probar el resultado principal de existencia y unicidad de (4.1.1), y la que nosotros vamos a dar en este curso está basada en el Teorema del punto fijo de Banach. Antes de recordar este resultado, y dar una demostración para hacer estas notas lo más autocontenidas posibles, necesitamos unas definiciones previas.

Definición 4.1.4. Se dice que $f : \Omega \to \mathbb{R}^n$ es *lipschitziana* respecto a la segunda variable en Ω si existe una constante positiva llamada *constante de Lipschitz* y denotada por L_f tal que para todo $(t, x), (t, y) \in \Omega$ se tiene

$$||f(t, x) - f(t, y)|| \leq L_f ||x - y||.$$

Si $L_f < 1$, se dice que f es *contractiva*.

Una condición menos restrictiva para una función y que nos seguirá siendo de utilidad en las demostraciones de los resultados sucesivos es la lipschitzianidad local.

Definición 4.1.5. Se dice que $f : \Omega \to \mathbb{R}^n$ es *localmente lipschitziana* respecto a la segunda variable en Ω si para cada $(t, x) \in \Omega$ existe un entorno $U_{(t,x)} \subset \Omega$ tal que la restricción de f a $U_{(t,x)}$ es lipschitziana respecto de la segunda variable.

El siguiente resultado da una condición suficiente para que una función sea localmente lipschitziana que es inmediata de comprobar.

Proposición 4.1.6. *Sea $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ continua y de clase C^1 respecto de la segunda variable. Entonces, f es localmente lipschitziana.*

Demostración. Denotemos $D_x f(t, x)$ a la matriz Jacobiana correspondiente, es decir

$$D_x f(t, x) = \frac{\partial f}{\partial x}(t, x) = \frac{\partial(f_1, ..., f_n)}{\partial(x_1, ..., x_n)}(t, x).$$

Para cada $(t_0, x_0) \in \Omega$ podemos tomar como entorno contenido en Ω un compacto de la forma $K = [t_0 - \varepsilon, t_0 + \varepsilon] \times \overline{B}(x_0, r)$, con ε, r suficientemente pequeños. En particular, $f_{|_K}$ es C^1 y $||D_x f||$ alcanza su máximo en K; llamemos L_f a tal máximo.

La idea es aplicar el Teorema del Valor Medio para funciones vectoriales a una función de variable real adecuada. Para ello, observemos que para cada $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ se tiene que $\overline{B}(x_0, r)$ es convexo y en consecuencia dados $(t, x), (t, y) \in K$, el segmento $\gamma(s) = (t, (1 - s)x + sy),\ s \in [0, 1]$ cumple $\gamma(s) \in K$ para todo s y además $\gamma(0) = x, \gamma(1) = y$. En particular, la función vectorial $g : [0, 1] \to \mathbb{R}^n$ dada por $g(s) = f(t, \gamma(s))$ está bien definida y cumple $g(0) = f(t, x)$ y $g(1) = f(t, y)$. El Teorema 4.1.3 asegura $||g(1) - g(0)|| \leq ||g'(c)||$ para cierto $c \in (0, 1)$. Sustituyendo se tiene

$$||f(t, y) - f(t, x)|| \leq ||D_x f(t, \gamma(c))\gamma'(c)|| \leq ||D_x f(t, \gamma(c))||\, ||y - x|| \leq L_f ||x - y||,$$

donde en la penúltima desigualdad hemos usado que $\gamma'(c) = y - x$ y que $||Av|| \leq ||A||\, ||v||$ para toda matriz A y todo vector v. De aquí se concluye que f es lipschitziana en su segunda variable. □

Observación 4.1. La proposición anterior se puede generalizar para obtener lipschitzianidad global de la siguiente forma. Denotemos por $\pi_1 : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ a la proyección en el primer factor dada por $\pi_1(t, x) = t$, y para cada $t \in \pi_1(\Omega)$ definimos $\Omega_t = \{x \in \mathbb{R}^n : (t, x) \in \Omega\}$. Por ejemplo, si $\Omega = I \times \Omega'$ entonces $\pi_1(\Omega) = I$ y $\Omega_t = \{x \in \mathbb{R}^n\colon (t,x) \in \Omega\} = \Omega'$. Supongamos que existe $D_x f(t,x)$, que está acotada en Ω y Ω_t es convexo para cada $t \in \pi_1(\Omega)$. Entonces $f(t,x)$ es lipschitziana en Ω respecto a su segunda variable. En efecto, para cada $t \in \pi_1(\Omega)$ y dados $x, y \in \Omega_t$, consideramos $\gamma(s) = (1-s)x + sy,\ s \in [0,1]$ el segmento que une x con y, el cual está incluido en Ω_t para todo $s \in [0,1]$ por ser convexo. Defininendo $g(s) = f(t, \gamma(s))$, la cual está bien definida, terminamos como en la demostración de la Proposición 4.1.6.

El siguiente es uno de los resultados fundamentales en la teoría del Análisis Funcional y que tiene incidencia en prácticamente todas las disciplinas de las Matemáticas. Incluimos una prueba del mismo.

Teorema 4.1.7 (Teorema del punto fijo de Banach)**.** *Sea (X,d) un espacio de Banach y $T : X \to X$ una aplicación contractiva. Entonces, existe un único $x \in X$ tal que $Tx = x$. Además, tal punto fijo es* atractor *de T en el siguiente sentido: si para cualquier $p \in X$ definimos por recurrencia $p_0 = p,\ p_n = Tp_{n-1} = T^n p$, entonces*

$$\lim_{n\to\infty} p_n = \lim_{n\to\infty} T^n p = x.$$

Demostración. Por ser T contractiva, existe $k \in (0,1)$ tal que $d(Tx,Ty) \leq kd(x,y)$. Sea $p \in X$ arbitrario y definamos la sucesión $p_0 = p,\ p_n = Tp_{n-1} = T^n p$. Si probamos que $\{p_n\}$ es una sucesión de Cauchy, la completitud de (X,d) implicará que $\{p_n\}$ converge a un cierto $x \in X$. Dados $n, m \in \mathbb{N}$ arbitrarios, se tiene

$$d(p_{n+m}, p_n) = d(Tp_{n+m-1}, Tp_{n-1}) \leq kd(p_{n+m-1}, p_{n-1}) \leq k^2 d(p_{n+m-2}, p_{n-2}) \leq \cdots \leq k^n d(p_m, p_0).$$

En particular, para cada $j = 0, ..., m-1$ se tiene $d(p_{j+1}, p_j) = k^j d(p_1, p_0)$ y en consecuencia por la desigualdad triangular deducimos

$$d(p_m, p_0) \leq \sum_{j=0}^{m-1} d(p_{j+1}, p_j) = \sum_{j=0}^{m-1} k^j d(p_1, p_0) = (1+k+k^2+\cdots+k^{j-1}) d(p_1, p_0) = \frac{1}{1-k} d(p_1, p_0).$$

Juntando ambas desigualdades concluimos

$$d(p_{n+m}, p_n) \leq \frac{k^n}{1-k} d(p_1, p_0).$$

Como $\lim_{n\to\infty} k^n = 0$ y n, m se escogieron arbitrarios concluimos que $\{p_n\}$ es de Cauchy y por completitud de (X,d) es convergente a un cierto $x \in X$. Como T es contractiva en particular es continua y por tanto

$$Tx = T\left(\lim_{n\to\infty} p_n\right) = \lim_{n\to\infty} Tp_n = \lim_{n\to\infty} p_{n+1} = x,$$

es decir x es punto fijo de T. Finalmente, probemos que x es el único punto fijo de T. Por reducción al absurdo, si existiese $y \in X$ tal que $Ty = y$ con $y \neq x$, entonces

$$d(x,y) = d(Tx,Ty) \leq kd(x,y) < d(x,y),$$

lo cual es absurdo. $\square$

El siguiente corolario tiene su propio interés y nos será de utilidad en la demostración del resultado principal de esta sección.

Corolario 4.1.8. *Sea (X,d) un espacio de Banach y $T : X \to X$ una aplicación continua y tal que para algún $n \in \mathbb{N}$ el operador T^n es contractivo. Entonces, T posee un único punto fijo que coincide con el único punto fijo de T^n.*

Demostración. Por ser T^n contractiva, el Teorema del punto fijo de Banach asegura la existencia de un único $x \in X$ tal que $T^n x = x$, el cual además es atractor de T^n. Ahora bien, de la ecuación $T^n x = x$ se tiene

$$Tx = T^{n+1}x = T^n(Tx),$$

de donde deducimos que Tx es igualmente punto fijo de T^n. Por la unicidad del punto fijo, necesariamente $Tx = x$. □

4.1.2. El Teorema de Picard-Lindelöf

El objetivo de esta sección es demostrar uno de los principales resultados en la teoría de Ecuaciones Diferenciales Ordinarias, el cual enunciamos a continuación.

Teorema 4.1.9 (de Picard-Lindelöf)**.** *Sean $\varepsilon, r > 0$ y definamos $K = I_\varepsilon(t_0) \times \overline{B}(x_0, r) \subset \mathbb{R} \times \mathbb{R}^n$, donde $I_\varepsilon(t_0) = [t_0 - \varepsilon, t_0 + \varepsilon]$. Sea $f : K \to \mathbb{R}$ continua y Lipschitziana respecto a la segunda variable en K y tomemos $M = \text{máx}\{||f(t,x)||\colon (t,x) \in K\}$. Entonces, existe una única solución del problema de Cauchy*

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

definida en el intervalo $I_\alpha(t_0)$, siendo $\alpha = \text{mín}\{\varepsilon, r/M\}$.

Antes de probar el teorema, debemos hacer unos comentarios.

1. La norma respecto de la cual se calcula la bola $\overline{B}(x_0, r)$ en $\mathbb{R}^n$ es arbitraria; realmente da igual puesto que todas las normas en $\mathbb{R}^n$ son equivalentes. Fijada una norma $||\cdot||$ en $\mathbb{R}^n$, en $\mathbb{R} \times \mathbb{R}^n$ usaremos la norma $||(t,x)|| = \text{máx}\{|t|, ||x||\}$.

2. La elección del radio α del intervalo de definición, aunque pudiera parecer extraña, es totalmente normal. Por una parte necesitamos $\alpha \leq \varepsilon$ para que la primera variable de $f(t,x)$ pueda ser correctamente evaluada. De manera similar, la condición $\alpha \leq r/M$ surge de la siguiente discusión. Dada $x : I_\alpha(t_0) \to \mathbb{R}^n$ solución del problema de Cauchy, se tiene $||x'(t)|| = ||f(t, x(t))|| \leq M$. En consecuencia, el Teorema 4.1.3 asegura la existencia de $\xi \in (t_0, t)$ si $t > t_0$, ó $\xi \in (t, t_0)$ si $t < t_0$, tal que

$$||x(t) - x(t_0)|| \leq ||x'(\xi)||\,|t - t_0| = ||f(\xi, x(\xi))||\,|t - t_0| \leq M|t - t_0| \leq M\alpha \leq r.$$

En particular, $x(t) \in \overline{B}(x_0, r)$ para $t \in I_\alpha(t_0)$ y $f(t, x(t))$ está bien definida en su segunda variable.

3. Siguiendo con la elección de α, se ha tomado como el mínimo entre ε y r/M pero nada impide tomar valores menores que este mínimo. Esto nos dice que el problema de Cauchy (4.1.1) tiene existencia y unicidad en cada intervalo $I_\alpha(t_0)$ con $\alpha \leq \text{mín}\{\varepsilon, r/M\}$. El tomar α de esta forma nos garantiza el mayor intervalo posible en estas circunstancias.

4. Si por algún casual se toma $K = [t_0, t_0 + \varepsilon]$ ó $K = [t_0 - \varepsilon, t_0]$, el resultado sigue siendo cierto; es decir, el Teorema 4.1.9 también da existencia y unicidad local en los extremos de un intervalo compacto. No obstante, la prueba se realizará para t_0 interior y será el lector interesado el que tendrá que ir cambiando en cada etapa de la demostración el comportamiento de frontera que podría tener t_0.

Demostración. Con la notación $I_\alpha(t_0) = [t_0 - \alpha, t_0 + \alpha]$, sea $X = C^0\left(I_\alpha(t_0), \overline{B}(x_0, r)\right)$ el espacio de las funciones continuas del intervalo $I_\alpha(t_0)$ con imagen en la bola cerrada $\overline{B}(x_0, r)$. Si consideramos en X la distancia asociada a la norma uniforme,

$$d_\infty(x, y) = ||x - y||_\infty = \sup\{||x(t) - y(t)||\colon t \in I_\alpha(t_0)\},$$

se tiene que (X, d_∞) es un espacio de Banach. Dada $\phi \in X$, definimos la función $T\phi$ mediante

$$(T\phi)(t) = x_0 + \int_{t_0}^{t} f(s, \phi(s))\, ds.$$

Así definido, T es un operador $T : X \to C^0(I_\alpha(t_0))$, puesto que cada $T\phi$ es continua de forma trivial. No obstante, podría pasar que la función $T\phi$ no tuviese su imagen contenida en $\overline{B}(x_0, r)$. Veamos que no es este el caso y de hecho T es un operador de X en sí mismo. Para ello basta comprobar que $T(X) \subset X$. Dada $\phi \in X$ hay que probar que $(T\phi)(t) \in \overline{B}(x_0, r)$ para todo $t \in I_\alpha(t_0)$. Ahora bien,

$$||(T\phi)(t) - x_0|| \leq \int_{t_0}^{t} ||f(s, \phi(s))||\, ds \leq M|t - t_0| \leq M\alpha \leq r,$$

y concluimos $(T\phi)(t) \in \overline{B}(x_0, r)$ para todo $t \in I_\alpha(t_0)$, probando $T(X) \subset X$.

A continuación probamos que T es un operador continuo y que alguna potencia T^n es contractiva. Esto nos dará la existencia de un único punto fijo y en consecuencia de una única solución del problema de Cauchy. Sean $x, y \in X$ y $t \in I_\alpha(t_0)$ arbitrario.

$$\begin{aligned}
||(Tx)(t) - (Ty)(t)|| &= \left|\left|\int_{t_0}^{t} f(s, x(s)) - f(s, y(s))\, ds\right|\right| \\
&\leq \int_{t_0}^{t} ||f(s, x(s)) - f(s, y(s))||\; ds \\
&\leq \int_{t_0}^{t} L_f\, ||x(s) - y(s)||\; ds \\
&\leq L_f|t - t_0|\, ||x - y||_\infty \leq L_f\alpha||x - y||_\infty.
\end{aligned}$$

Como $t \in I_\alpha(t_0)$ se escogió de manera arbitraria se tiene $||Tx - Ty||_\infty \leq L_f\alpha||x - y||_\infty$ y en particular T es un operador continuo. Ahora probaremos por inducción sobre $p \in \mathbb{N}$ la siguiente desigualdad para $t \in I_\alpha(t_0)$ arbitrario:

$$||(T^p x)(t) - (T^p y)(t)|| \leq \frac{L_f^p|t - t_0|^p}{p!}||x - y||_\infty. \tag{4.1.2}$$

El caso base para $p = 1$ ha sido ya probado. Supuesta cierta para p y dado de nuevo t arbitrario,

$$\begin{aligned}
||(T^{p+1}x)(t) - (T^{p+1}y)(t)|| &= \left|\left|\int_{t_0}^{t} f(s,(T^p x)(s)) - f(s,(T^p y)(s))\,ds\right|\right| \\
&\leq \int_{t_0}^{t} ||f(s,(T^p x)(s)) - f(s,(T^p y)(s))||\;ds \\
&\leq \int_{t_0}^{t} L_f\,||(T^p x)(s) - (T^p y)(s)||\;ds \\
&\leq L_f \frac{L_f^p}{p!}||x-y||_\infty \int_{t_0}^{t} |s-t_0|^p\,ds = \frac{L_f^{p+1}|t-t_0|^{p+1}}{(p+1)!}||x-y||_\infty.
\end{aligned}$$

Así se tiene (4.1.2) y como t era arbitrario en $I_\alpha(t_0)$ concluimos $||T^p x - T^p y||_\infty \leq \frac{L_f^p \alpha^p}{p!}||x-y||_\infty$. Como $\lim_{p\to\infty} \frac{L_f^p\alpha^p}{p!} = 0$, existe p_0 suficientemente grande tal que

$$||T^{p_0}x - T^{p_0}y||_\infty < ||x-y||_\infty$$

y en virtud del Corolario 4.1.8 al ser T^{p_0} contractivo concluimos que T tiene un único punto fijo en X, es decir $x = Tx$. Esto implica que para todo $t \in I_\alpha(t_0)$ se tiene

$$x(t) = (Tx)(t) = x_0 + \int_{t_0}^{t} f(s,x(s))\,ds.$$

En particular $x \in C^1(I_\alpha(t_0))$ y es solución del problema de Cauchy, lo que concluye la demostración. □

Observación 4.2. Un par de observaciones.

1. En la demostración de la continuidad de T, se concluye

   $$||Tx - Ty||_\infty \leq L_f\alpha||x-y||_\infty.$$

   En este punto ya podríamos finalizar la demostración sin más que tomar α con la condición adicional $\alpha < 1/L_f$. Aunque este argumento nos da la existencia en un intervalo $I_\alpha(t_0)$ posiblemente menor que el óptimo construido en la demostración del teorema, esto no es restrictivo como veremos en el apartado de prolongación y existencia de soluciones maximales.

2. En la prueba del Teorema 4.1.9 se encuentra implícita la misma idea de definir iterantes de Picard, tal y como se hizo en la demostración del problema de Cauchy del sistema lineal $x' = A(t)x + b(t)$ en el tema 3. En efecto, por ser T un operador contractivo en el espacio completo $X = C^0\left(I_\alpha(t_0), \overline{B}(x_0,r)\right)$, el único punto fijo de T es atractor de cualquier función de X. Si comenzamos con la función $x_0(t) = x_0 \in \mathbb{R}^n$, entonces la sucesión $\{x_n\}$ en X definida de forma recursiva por

   $$x_{n+1}(t) = (Tx_n)(t) = x_0 + \int_{t_0}^{t} f(s,x_n(s))\,ds$$

   cumple que converge uniformemente en el compacto $I_\alpha(t_0)$ a una función límite x que resulta ser la solución de (4.1.1). Esta es una demostración alternativa del Teorema 4.1.9 y las ideas principales ya se desarrollaron en el mencionado teorema de existencia y unicidad del problema de Cauchy para el sistema diferencial $x' = A(t)x + b(t)$.

Sabemos que la hipótesis de lipschitzianidad no se puede rebajar a que $f(t,x)$ sea únicamente continua. En efecto, consideremos el problema de Cauchy

$$\begin{cases} x' = 3x^{2/3}, \\ x(t_0) = x_0. \end{cases}$$

Si $x_0 \neq 0$ la función $f(t,x) = 3x^{2/3}$ es lipschitziana en un entorno de $t = t_0$ y por tanto el problema tiene existencia y unicidad local dado por la función $x(t) = (t + x_0^{1/3} - t_0)^3$. Ahora bien, justo para el instante $t_* = t_0 - x_0^{1/3}$ la función se anula y observamos que igualmente la función constantemente cero es solución del problema de Cauchy, perdiendo la unicidad. El motivo es que $f(t,x) = 3x^{2/3}$ no es localmente lipschitziana en ningún entorno que contiene a 0. De hecho, el problema de Cauchy

$$\begin{cases} x' = 3x^{2/3}, \\ x(t_0) = 0, \end{cases}$$

tiene infinitas soluciones dadas por la familia 2-paramétrica en términos de $c_1 \leq 0 \leq c_2$ por

$$x(t) = \begin{cases} (t-c_1)^3 & \text{si} \quad t \leq c_1, \\ 0 & \text{si} \quad c_1 \leq t \leq c_2, \\ (t-c_2)^3 & \text{si} \quad t \geq c_2. \end{cases}$$

En la mayoría de ocasiones la función $f(t,x)$ estará definida en un abierto arbitrario Ω y en particular no tenemos directamente una estructura de producto $\Omega = I_\varepsilon(t_0) \times \overline{B}(x_0, r)$ con ambos factores compactos.

No obstante, podemos suponer que la función $f(t,x)$ es localmente lipschitziana (Definición 4.1.5) y obtener el siguiente resultado de existencia local en cualquier abierto.

Corolario 4.1.10 (Teorema de Picard-Lindelöf en un abierto)**.** *Sea $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ un abierto y $f : \Omega \to \mathbb{R}^n$ una función continua y localmente lipschitziana respecto de la segunda variable. Entonces, para todo $(t_0, x_0) \in \Omega$ existen $\varepsilon, r > 0$ suficientemente pequeños de forma que $V_{(t_0,x_0)} = I_\varepsilon(t_0) \times \overline{B}(x_0, r)$ está contenido en Ω y tal que el problema de Cauchy*

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

tiene una única solución $x : I_\varepsilon(t_0) \to \overline{B}(x_0, r)$.

Demostración. Sea K entorno compacto de (t_0, x_0) en Ω. Entonces, $f_{|K}$ es lipschitziana. En efecto, sea $\cup_{i \in I} U_i$ recubrimiento por abiertos de K y sea L_{f_i} la constante de Lipschitz de $f_{|U_i}$. Por ser K compacto, existe un subrecubrimiento finito $\cup_{i=1}^n U_i$. Sea $L_{f_\infty} = \text{máx}_{i=1,..,n} L_{f_i}$. Entonces, para todo $(t,x), (t,y) \in K$ se tiene

$$||f(t,x) - f(t,y)|| \leq L_{f_\infty} ||x - y||,$$

lo que prueba la lipschitzianidad de $f_{|K}$.

Por otra parte, sea $M = \text{máx}_{(t,x) \in K} ||f(t,x)||$ y tomemos ε suficientemente pequeño para que el entorno $V_{(t_0,x_0)} = I_\varepsilon(t_0) \times \overline{B}(x_0, \varepsilon M)$ esté contenido en K. Ahora bastaría aplicar el Teorema de Picard-Lindelöf (Teorema 4.1.9) en $V_{(t_0,x_0)}$. □

Notemos que en los resultados anteriores hemos probado la existencia de una solución del problema de Cauchy (4.1.1) siempre y cuando la función f estuviese definida en un conjunto de la forma $I_\varepsilon(t_0) \times \overline{B}(x_0, r)$. En tal caso, la solución $x(t)$ está definida en el entorno compacto $I_\alpha(t_0)$ de la condición inicial t_0, siendo $\alpha = \min\{\varepsilon, r/M\}$. En particular, la condición $\alpha \leq r/M$ es fundamental para garantizar que $x(t) \in \overline{B}(x_0, r)$ para todo $t \in I_\alpha(t_0)$. En el caso particular en que $\Omega = [a, b] \times \mathbb{R}^n$, la segunda condición se da siempre de manera trivial y podemos probar el siguiente resultado global; a cambio, necesitamos imponer una condición más fuerte sobre la función f.

Teorema 4.1.11 (de Picard-Lindelöf en una banda)**.** *Sean $I \subset \mathbb{R}$ un intervalo y $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ una función continua y (globalmente) lipschitziana respecto a la segunda variable. Entonces, el problema de Cauchy*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0, \end{cases}$$

posee una única solución globalmente definida en I.

Demostración. Dado $t_0 \in I$, el primer paso es probar la existencia y unicidad del problema de Cauchy en cualquier intervalo compacto $[a, b] \subset I$ tal que $t_0 \in [a, b]$. Para ello, podemos rehacer la demostración del Teorema 4.1.9 considerando ahora el operador T definido en el espacio normado $X = C([a, b], \mathbb{R}^n)$ de funciones continuas de $[a, b]$ a $\mathbb{R}^n$ dotado de la norma uniforme $|| \cdot ||_\infty$, el cual es también un espacio completo. En primer lugar, observemos que no es necesario que f sea acotada, puesto que la imagen Tx vuelve a ser un elemento de X al tener como codominio todo $\mathbb{R}^n$ y no tener que restringirnos a una bola. Ahora bien, por ser f globalmente lipschitziana en la segunda variable, existe $L_f > 0$ tal que

$$||f(t, x) - f(t, y)|| \leq L_f ||x - y||,$$

para todo $x, y \in \mathbb{R}^n$. Aquí la clave es que la constante L_f es uniforme en todo $\mathbb{R}^n$. En particular, dadas $x, y \in X$ y $t \in [a, b]$, se tiene

$$\begin{aligned} ||(Tx)(t) - (Ty)(t)|| &= \left|\left| \int_{t_0}^t f(s, x(s)) - f(s, y(s))\, ds \right|\right| \\ &\leq \int_{t_0}^t ||f(s, x(s)) - f(s, y(s))||\, ds \\ &\leq \int_{t_0}^t L_f \, ||x(s) - y(s)||\, ds \\ &\leq L_f |t - t_0| \, ||x - y||_\infty \leq L_f (b - a) ||x - y||_\infty. \end{aligned} \tag{4.1.3}$$

Igual que en la prueba del Teorema 4.1.9 se prueba por inducción

$$||T^p x - T^p y||_\infty \leq \frac{L_f^p (b - a)^p}{p!} ||x - y||_\infty.$$

Existe por tanto p_0 suficientemente grande tal que el iterado T^{p_0} es contractivo, lo que en virtud del Corolario 4.1.8 implica la existencia y unicidad de una solución del problema de Cauchy en $[a, b]$. Sea ahora $K_n = [a_n, b_n]$ una exhaución de I por intervalos compactos tales que $t_0 \in K_n$, esto es $K_n \subset K_{n+1}$ y $\cup_{n \geq 1} K_n = I$. Sabemos que existe una única solución $x_n : K_n \to \mathbb{R}^n$ de (4.1.1). Basta ahora definir $x : I \to \mathbb{R}^n$ mediante $x(t) = x_n(t)$ si $t \in K_n$, la cual está bien definida por unicidad en cada K_n y cumple que de clase C^1 en I, por serlo cada x_n en K_n y ser la derivabilidad un concepto local. Finalmente, dado $t \in I$ se tiene $t \in K_n$ para cierto $n \in \mathbb{N}$ y

$$x'(t) = x'_n(t) = f(t, x_n(t)) = f(t, x(t)),$$

concluyendo que x es solución de (4.1.1) en I.

Llegados a este punto, creemos que es constructivo dar otra demostración que será útil para comenzar a entender conceptos de prolongación que se verán posteriormente. De la última desigualdad en (4.1.3) concluimos

$$||Tx - Ty||_\infty \leq L_f(b-a)|x-y||_\infty.$$

Tomando $[a, b]$ tal que $L_f(b-a) < 1$ se tiene que T es contractivo. En particular, para cada intervalo con esta longitud tenemos existencia y unicidad. Sea $t_0 \in I$ y definimos $h = b - a$. Entonces, existe una única solución $x : [t_0, t_0+h \to \mathbb{R}^n$. Sea $t_1 = t_0+h/2$ y $x_1 = x(t_1)$ y consideremos el problema de Cauchy

$$\begin{cases} x' = f(t, x), \\ x(t_1) = x_1. \end{cases}$$

Este problema de Cauchy tiene existencia y unicidad en el intervalo compacto $[t_1, t_1+h]$, llamémosla $y : [t_1, t_1 + h] \to \mathbb{R}^n$. Ahora bien, $[t_0, t_0 + h] \cap [t_1, t_1 + h] = [t_1, t_0 + h] = [t_1, t_1 + h/2]$ y por tanto $x = y$ en esta intersección por unicidad. En consecuencia, hemos *extendido* la solución x al intervalo $[t_0, t_0 + 3h/2]$. Repitiendo este proceso acabaríamos de extender x a todo $[t_0, \infty) \cap I$. El argumento a la izquierda de t_0 es similar.

$\square$

Observación 4.3. Un par de comentarios referentes a este resultado.

1. Aunque esta prueba nos debe recordar a la dada en el Teorema 3.1.4 del tema 3 para el sistema diferencial lineal $x' = A(t)x + b(t)$, hay que enfatizar que tal resultado no es un corolario del Teorema 4.1.11. En efecto, aunque la aplicación $f(t, x) = A(t)x + b(t)$ está definida en $f : I \times \mathbb{R}^n \to \mathbb{R}^n$, no tiene por qué ser globalmente Lipschitziana, puesto que la norma matricial $||A(t)||$ podría hacerse arbitrariamente grande. No obstante, esta función cumple una acotación *sublineal* de la forma

$$||f(t, x)|| \leq \alpha(t) + \beta(t)||x||,$$

lo cual es suficiente para probar la definición global en I de las soluciones de la EDO correspondiente; véase el ejercicio 16.

2. La hipótesis de lipschitzianidad global es, como no puede ser de otra forma, clave. Por ejemplo, es sabido que el problema $x' = x^2$ con condición inicial $x(0) = 1$ tiene como única solución

$$x(t) = \frac{1}{1-t}, \qquad t \in (-\infty, 1),$$

y sin embargo $f(t, x) = x^2$ está definida para todo $t \in \mathbb{R}$. El problema radica en la cadena de desigualdades (4.1.3) y en concreto en no poder tomar de manera uniforme una constante L_f para todas las funciones $x, y \in X$. De hecho, en este ejemplo concreto $f(t, x) = x^2$, lo que sucede es que cuando la norma de x crece así lo hace la constante de Lipschitz, debido a que se cumple

$$|x^2 - y^2| = |x + y||x - y| \leq L_f|x - y|,$$

si y sólo si $|x + y| \leq L_f$, lo cual no es cierto globalmente en $\mathbb{R}$.

4.2. Existencia local

Cuando la función $f(t,x)$ es sólo continua sin hipótesis adicionales sobre la lipschitzianidad, las conclusiones del Teorema de Picard-Lindelöf 4.1.9 y sus consecuencias local y global, Corolario 4.1.10 y Teorema 4.1.11 respectivamente, no son ciertas. Sin centrarnos en la unicidad, la mera existencia de una solución depende fuertemente de la lipschitzianidad de $f(t,x)$. Por ejemplo a la hora de definir los iterantes de Picard

$$x_{n+1} = (Tx_n)(t) = x_0 + \int_{t_0}^{t} f(s, x_n(s))\, ds,$$

y probar que éstos convergen uniformemente a una función límite que será eventualmente la solución de la EDO; o equivalentemente, para probar que un cierto iterado T^n es contractivo.

El objetivo de esta sección es presentar el Teorema de Peano, que garantiza la existencia de soluciones del problema de Cauchy (4.1.1) bajo la hipótesis de continuidad de la función $f(t,x)$. Este resultado se obtiene sin dificultad como una sencilla consecuencia del Teorema de Picard-Lindelöf 4.1.9 y dos importantes resultados del Análisis Matemático: los Teoremas de aproximación de Stone-Weierstrass y de Ascoli-Arzela. La dificultad pues de probar el Teorema de Peano radica en la demostración de estos dos úlitmos resultados.

4.2.1. Los Teoremas de Stone-Weierstrass y Ascoli-Arzela

El primer resultado de esta sección nos permite aproximar uniformemente por polinomios cualquier función continua definida en un compacto. De éste, omitimos su prueba.

Teorema 4.2.1 (de aproximación de Stone-Weierstrass)**.** *Sean $K \subset \mathbb{R}^n$ compacto y $f : K \to \mathbb{R}^k$ una aplicación continua. Entonces, existe una sucesión $\{p_n\}$ de polinomios que la aproxima uniformemente. Esto es, para todo $\varepsilon > 0$ existe $n_0(\varepsilon) \in \mathbb{N}$ tal que para todo $n > n_0$ se tiene*

$$||p_n(x) - f(x)|| < \varepsilon, \qquad \forall x \in K.$$

Para poder enunciar el Teorema de Ascoli-Arzela se necesitan una serie de conceptos previos. Enunciaremos estos resultados para un intervalo $I \subset \mathbb{R}$ cualquiera, aunque en general se pueden sustituir el dominio y codominio de las funciones involucradas por subconjuntos adecuados de espacios métricos arbitrarios.

Definición 4.2.2. Una sucesión de funciones $\varphi_n : I \to \mathbb{R}^n$ se dice equicontinua en t_0 si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para todo $m \in \mathbb{N}$ y para todo $t \in I$ con $|t - t_0| < \delta$ se tiene

$$||\varphi_m(t) - \varphi_m(t_0)|| < \varepsilon.$$

La sucesión se dice equicontinua en I si lo es en cada uno de sus puntos. Más aún, la sucesión se dice uniformemente equicontinua en I si para todo $\varepsilon > 0$ existe un $\delta > 0$ uniforme para todo $t, t' \in I$.

Obsérvese que la equicontinuidad es una condición de uniformidad: el valor de δ únicamente depende de t_0 y de ε, pero no de la función de la sucesión. En particular, toda sucesión equicontinua en t_0 es continua en t_0. Más aún, la equicontinuidad va bien con la convergencia puntual, como se exhibe en el siguiente resultado.

Proposición 4.2.3. *Sea $\varphi_n : I \to \mathbb{R}^n$ una sucesión.*

1. *Si $\{\varphi_n\}$ es equicontinua en $t_0 \in I$ y converge puntualmente a φ_∞, entonces φ_∞ es continua en t_0.*

2. *Si $\{\varphi_n\}$ es equicontinua y converge puntualmente a φ_∞, entonces $\{\varphi_n\} \cup \{\varphi_\infty\}$ es equicontinua.*

Demostración. 1. Fijemos $t_0 \in I$ y sea $\varepsilon > 0$. Sabemos que existe $\delta > 0$, dependiendo de t_0 y ε tal que para cualquier $t \in I$ con $|t - t_0| < \delta$ se tiene $||\varphi_n(t) - \varphi_n(t_0)|| < \varepsilon$. Como $\{\varphi_n\}$ converge puntualmente a φ_∞ en I, basta tomar límite en la desigualdad anterior para concluir $||\varphi_\infty(t) - \varphi_\infty(t_0)|| \leq \varepsilon$, lo cual implica la continuidad de φ_∞ en t_0.

2. Es similar a la anterior. Sea $t_0 \in I$ fijo. Por equicontinuidad en t_0, dado $\varepsilon > 0$ existe $\delta > 0$ tal que para todo $t \in I$ con $|t - t_0| < \delta$ se tiene $||\varphi_n(t) - \varphi_n(t_0)|| < \varepsilon$. Si se toma límite otra vez se tiene $||\varphi_\infty(t) - \varphi_\infty(t_0)|| \leq \varepsilon$. Leyendo entre líneas observamos pues que la función φ_∞ cumple el criterio de equicontinuidad en t_0, el cual era arbitrario, y por tanto $\{\varphi_n\} \cup \{\varphi_\infty\}$ es equicontinua.

□

Sabemos que en general, si una sucesión de funciones definidas en un compacto converge puntualmente, la convergencia no es necesariamente uniforme. El siguiente resultado prueba que añadiendo la hipótesis de equicontinuidad entonces sí se garantiza la convergencia uniforme, y además la equicontinuidad también adquiere la uniformidad.

Proposición 4.2.4. *Toda sucesión de funciones equicontinuas definidas en un compacto es uniformemente equicontinua. Si además la sucesión converge puntualmente, entonces la convergencia es uniforme.*

Demostración. Sea $\{x_n : K \to \mathbb{R}^n\}$ sucesión equicontinua definidas en K compacto. Fijemos $\varepsilon > 0$ y consideremos el recubrimiento por abiertos de K

$$\mathcal{U} = \{B(t, \delta(t, \varepsilon/2)/2) \colon t \in K\},$$

donde $\delta(t, \varepsilon/2) > 0$ es el valor de δ dependiendo de t y $\varepsilon/2$ dado por la condición de equicontinuidad en t, pero que no depende de ninguna de las funciones φ_n. Por compacidad de K, podemos extraer un subrecubrimiento finito, obteniendo $t_1, ..., t_p$ y valores $\delta(t_1, \varepsilon/2), ..., \delta(t_p, \varepsilon/2)$. Definimos $\delta = \text{mín}_{k=1,...,p}\{\delta(t_k, \varepsilon/2)/2\}$.

Sean ahora $t, t' \in K$ tales que $|t - t'| < \delta$ y fijemos $n \in \mathbb{N}$ arbitrario. Existe un índice $k \in \{1, ..., p\}$ tal que $t \in B(t_k, \delta(t_k, \varepsilon/2)/2)$, es decir $|t - t_k| < \delta(t_k, \varepsilon/2)/2$. Ahora la desigualdad triangular implica

$$|t' - t_k| = |t' - t + t - t_k| \leq |t' - t| + |t - t_k| < \delta + \delta(t_k, \varepsilon/2)/2 \leq \delta(t_k, \varepsilon/2)/2 + \delta(t_k, \varepsilon/2)/2 = \delta(t_k, \varepsilon/2),$$

y por tanto $t' \in B(t_k, \delta(t_k, \varepsilon/2))$. Es decir, podemos aplicar la condición de equicontinuidad de la sucesión $\{\varphi_n\}$ en t_k para los valores t, t', obteniendo

$$||\varphi_n(t) - \varphi_n(t')|| = ||\varphi_n(t) - \varphi_n(t_k)|| + ||\varphi_n(t_k) - \varphi_n(t')|| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

lo que implica la equicontinuidad uniforme.

Para probar que convergencia puntual implica convergencia uniforme, llamemos $x_\infty(t)$ a la función límite puntual de $x_n(t)$ para cada $t \in K$. Sabemos por 4.2.3 que $\{x_n\} \cup \{x_\infty\}$ sigue siendo equicontinua y de hecho es uniformemente equicontinua. Fijemos $\varepsilon > 0$ y sea $\delta > 0$ el valor dependiendo únicamente de $\varepsilon/3$ dado por la equicontinuidad uniforme. Consideremos el recubrimiento por abiertos de K,

$$\mathcal{U} = \{B(t,\delta) \colon t \in K\}.$$

Extraemos un recubrimiento finito, obteniendo $t_1, ..., t_p \in K$ tales que $K = \cup_{i=1}^p B(t_i, \delta)$. Ahora bien, como $x_n(t) \to x_\infty(t)$ puntualmente en particular lo hará en cada t_k. Por ser una cantidad finita, existe $n_0 \in \mathbb{N}$ tal que para todo $n > n_0$ se tiene

$$||x_n(t_k) - x_\infty(t_k)|| < \varepsilon/3.$$

Por otra parte, por ser $\mathcal{U}$ un recubrimiento de K, para todo $t \in K$ existe $k \in \{1, ..., p\}$ tal que $t \in B(t_k, \delta)$, lo que implica por equicontinuidad uniforme

$$||x_n(t) - x_n(t_k)|| < \varepsilon/3,$$

para todo $n \in \mathbb{N} \cup \{\infty\}$. Finalmente, para cualquier $t \in K$ y para todo $n > n_0$ se tiene

$$||x_n(t) - x_\infty(t)|| \leq ||x_n(t) - x_n(t_k)|| + ||x_n(t_k) - x_\infty(t_k)|| + ||x(t_k) - x_\infty(t)|| < \varepsilon.$$

□

Definición 4.2.5. Una sucesión de funciones $\varphi_n : I \to \mathbb{R}^n$ se dice puntualmente acotada si para cada $t \in I$ la sucesión $\{\varphi_n(t)\}$ está acotada, esto es existe M_t sólo dependiendo de t tal que $||\varphi_n(t)|| < M_t$ para todo $n \in \mathbb{N}$. Más aún, si existe $M > 0$ tal que $||\varphi_n(t)|| < M$ para todo $t \in I$ y para todo $n \in \mathbb{N}$, la sucesión se dice uniformemente acotada.

Notemos que incluso si I se escoge compacto y $\varphi_n : [a,b] \to \mathbb{R}^n$ son continuas y puntualmente acotadas, no necesariamente $\{\varphi_n\}$ ha de ser uniformemente acotada. En efecto, para cada $n \in \mathbb{N}$ con $n > 2$ definimos la familia $\varphi_n : [0,1] \to \mathbb{R}$ por

$$\varphi_n(t) = \begin{cases} n & \text{si } t < 1/n, \\ 0 & \text{si } t \in [2/n, 1], \end{cases}$$

y extendemos de forma continua tal que $\varphi_n(t) \in (0, n)$ para todo $t \in (\frac{1}{n}, \frac{2}{n})$. Así definida, $\varphi_n(t)$ está puntualmente acotada. En efecto, para cada $t \in [0,1]$ se tiene que existe $n_0 \in \mathbb{N}$ de forma que $\frac{1}{n_0} < t$ y por tanto para todo $n > n_0$ obtenemos $\varphi_n(t) = 0$. Para una cantidad finita podemos tomar $M_t = \text{máx}_{2 \leq n \leq n_0} |\varphi_n(t)| = n_0$ y se tiene la acotación puntual. No obstante, φ_n no puede estar unifórmemente acotada puesto que toma valores arbitrariamente grandes.

Estamos ahora preparados para enunciar el Teorema de Ascoli-Arzela. Este resultado se puede introducir de distintas formas equivalentes y según el contexto cada vez más general. A nosotros nos basta la siguiente versión.

Teorema 4.2.6 (de Ascoli-Arzela)**.** *Sea $\varphi_n : [a,b] \to \mathbb{R}^n$ una sucesión equicontinua y puntualmente acotada. Entonces, existe una parcial $\{\varphi_{\sigma(n)}\}$ que converge uniformemente a una función continua $\varphi_\infty : [a,b] \to \mathbb{R}^n$.*

Por intentar que estas notas cubran lo máximo posible dentro de las limitaciones del curso, enunciamos una versión del Teorema de Ascoli-Arzela para espacios métricos generales. La notación $\mathcal{F} \subset C(E, F)$ denota una familia arbitraria de funciones continuas entre espacios métricos E, F. Además, tal familia se dice *relativamente compacta* si para toda sucesión $\{\varphi_n\} \subset \mathcal{F}$, existe una parcial convergente a un elemento de $\mathcal{F}$.

Teorema 4.2.7 (de Ascoli-Arzela)**.** *Sean (E, d) compacto y (F, d') completo. Una familia $\mathcal{F} \subset C(E, F)$ es relativamente compacta respecto a la distancia d_∞ de E si y sólo si se verifican*

1. *$\mathcal{F}$ es equicontinua,*
2. *para cada $x \in E$, la familia $\mathcal{F}(x) = \{\varphi(x) \colon \varphi \in \mathcal{F}\}$ es relativamente compacta en F.*

4.2.2. El Teorema de Peano

En esta sección vamos a probar que podemos garantizar existencia del problema de Cauchy (4.1.1) bajo la única hipótesis de que $f : \Omega \to \mathbb{R}^n$ sea continua. La relajación de esta hipótesis en detrimento de dejar de exigir lipschitzianidad local nos impide seguir garantizando la unicidad de soluciones.

Teorema 4.2.8 (de Peano)**.** *Sea $K = I_\varepsilon(t_0) \times \overline{B}(x_0, r)$ y supongamos que $f : K \to \mathbb{R}^n$ es continua. Sea $M = \max\{||f(t, x)|| \colon (t, x) \in K\}$. Entonces, existe una solución del problema de Cauchy*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0, \end{cases}$$

definida en $I_\alpha(t_0)$, donde $\alpha = \min\{\varepsilon, r/M\}$.

Sea ahora Ω abierto y $f : \Omega \to \mathbb{R}^n$ continua. Entonces, el problema de Cauchy tiene una solución en $I_\varepsilon(t_0)$, para $\varepsilon > 0$ suficientemente pequeño.

Antes de probar el resultado, repetimos uno de los comentarios que antecedieron a la prueba del Teorema 4.1.9 de Picard-Lindelöf: podemos tomar igualmente $K = [t_0, t_0 + \varepsilon]$ ó $K = [t_0 - \varepsilon, t_0]$, es decir la condición inicial se puede encontrar en uno de los extremos del intervalo.

Demostración. De acuerdo con el Teorema de aproximación de Stone-Weierstrass (Teorema 4.2.1), tomamos $\{p_m\}$ una sucesión de polinomios que aproxima uniformemente a f en K. Además, salvo una parcial no es restrictivo suponer $\sup\{||p_m(t, x)|| \colon (t, x) \in K\} < M$ para cada m. Por el Teorema de Picard-Lindelöf 4.1.9, el problema de Cauchy

$$\begin{cases} x' = p_m(t, x), \\ x(t_0) = x_0, \end{cases}$$

tiene existencia y unicidad $x_m : I_\alpha(t_0) \to \mathbb{R}^n$, con $\alpha = \min\{\varepsilon, r/M\}$ uniformemente en m. En particular, la imagen de cada $x_m(t)$ está contenida en $\overline{B}(x_0, r)$ para todo $t \in I_\alpha(t_0)$ y para todo m, es decir

$$||x_m(t)|| \leq ||x_0|| + r,$$

y deducimos que $\{x_m(t)\}$ está uniformemente acotada; en particular, puntualmente. Por otro lado, la cota uniforme $||p_m(t, x)|| < M$ nos permite afirmar que dados $t_1, t_2 \in I_\alpha(t_0)$ se tiene

$$||x_m(t_1) - x_m(t_2)|| = \left|\left| \int_{t_1}^{t_2} p_m(s, x_m(s))\, ds \right|\right| \leq M|t_1 - t_2|.$$

Fijado ahora $\varepsilon > 0$ basta definir $\delta = \varepsilon/M$ y en consecuencia $||x_m(t_1) - x_m(t_2)|| < \varepsilon$ siempre que $|t_1 - t_2| < \delta$, lo que garantiza la equicontinuidad de $\{x_m\}$ (de hecho uniforme). El Teorema de Ascoli-Arzela (Teorema 4.2.6) implica la existencia de una parcial de $\{x_m\}$, que seguimos denotando por comodidad $\{x_m\}$, que converge uniformemente a una función $x : I_\alpha(t_0) \to \mathbb{R}^n$ continua y con valores en $\overline{B}(x_0, r)$.

Ahora bien, las funciones $\{p_m(t, x_m(t))\}$ convergen uniformemente a la función $f(t, x(t))$. En efecto, fijemos $t \in I_\alpha(t_0)$ y sea $\varepsilon > 0$. Consideremos la estimación

$$||p_m(t, x_m(t)) - f(t, x(t))|| \leq ||p_m(t, x_m(t)) - f(t, x_m(t))|| + ||f(t, x_m(t)) - f(t, x(t))||.$$

Para el primer miembro de la derecha existe $m_0 \in \mathbb{N}$ tal que $||p_m(t, x_m(t) - f(t, x_m(t))|| < \varepsilon$ para todo $m > m_0$, puesto que $\{p_m\}$ converge uniformemente a f. De igual forma, existe $m_1 \in \mathbb{N}$ tal que $||f(t, x_m(t)) - f(t, x(t))|| < \varepsilon$ para todo $m > m_1$, esta vez por la convergencia uniforme de x_m a x y la continuidad uniforme de f, ésta última garantizada por ser f continua en un compacto.

Teniendo en cuenta que cada x_m cumple la ecuación integral

$$x_m(t) = x_0 + \int_{t_0}^{t} p_m(s, x_m(s))\, ds,$$

la convergencia uniforme de $\{x_m\}$ a x nos permite tomar límites para obtener

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds,$$

y finalmente concluir que x es solución del problema de Cauchy. □

De igual forma que en el Teorema de Picard-Lindelöf, existe una versión global del Teorema de Peano para $f(t, x)$ definida en una banda, pero en este caso hay que exigir la acotación de f como hipótesis adicional. A continuación, presentamos una demostración basada en el conocido como método de Tonnelli, la cual puede resultar algo complicada. No obstante, cuando veamos los apartados de prolongación y de comportamiento de las soluciones maximales, daremos pruebas más sencillas.

Teorema 4.2.9 (de Peano en una banda)**.** *Sea $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ continua y acotada, con $I \subset \mathbb{R}$ un intervalo arbitrario. Entonces, para cada $(t_0, x_0) \in I \times \mathbb{R}^n$ el problema de Cauchy*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0, \end{cases}$$

tiene una solución definida en I.

Demostración. Probaremos el resultado en el caso $I = [a, b]$ usando un argumento conocido como método de Tonnelli. Para el caso en que I es arbitrario, probaremos este teorema como una consecuencia de los resultados de prolongación de soluciones. Para cada $n \in \mathbb{N}$ consideremos la partición del intervalo $[a, b]$ dada por $t_0 = a,\ t_k = a + k\frac{b-a}{n}$, para $k = 0, ..., n$. Denotando $h = \frac{b-a}{n}$ se define la sucesión

$$x_n(t) = \begin{cases} x_0 & \text{si} \quad a \leq t \leq t_1, \\ x_0 + \int_a^{t-h} f(s, x_k(s))\, ds & \text{si} \quad t_k \leq t \leq t_{k+1}, \end{cases}, \qquad k = 1, 2, ..., k-1.$$

Antes de continuar la prueba, veamos el comportamiento de estas funciones. Si se tiene $n = 1$ entonces la partición es el propio intervalo $[a, b]$ y $x_1(t) = x_0$ para todo $t \in [a, b]$. Si $n = 2$ y $h = (b-a)/2$ entonces $t_0 = a$, $t_1 = a + h$ es el punto medio y $t_2 = b$. Si $t \in [a, t_1]$ entonces $x_2(t) = x_0$, pero si $t \in [t_1, b]$ entonces debemos estudiar el comportamiento del parámetro $s \in [a, t-h]$. En efecto, para $t = t_1$ se tiene $s = a$, mientras que si $t = t_2$ entonces $s \in [a, t_1]$. En cualquier caso, $s \in [a, t-h] \subset [a, t_1]$ y se tiene la función

$$x_2(t) = \begin{cases} x_0 & \text{si} \quad a \le t \le t_1, \\ x_0 + \int_a^{t-h} f(s, x_0)\, ds & \text{si} \quad t_1 \le t \le b. \end{cases}$$

Si $n = 3$ y $h = (b-a)/3$ entonces se tiene la partición $t_0 = a < t_1 < t_2 < t_3 = b$ y de nuevo $x_3(t) = x_0$ para $t \in [a, t_1]$. Si $t \in [t_1, t_2]$ entonces $s \in [a, t-h] \subset [a, t_1]$ vuelve a implicar que $x_3(s) = x_0$ y deducimos $x_3(t) = x_0 + \int_a^{t-h} f(s, x_0)\, ds$. Ahora, si $t \in [t_2, b]$ entonces $s \in [a, t-h] \subset [a, t_2]$, siendo $s = t_2$ precisamente para $t = b$. En este punto hay que distinguir casos. Fijado $t \in [t_2, b]$ se tiene $s \in [a, t-h] \subset [a, t_2]$ y por tanto

$$x_0 + \int_a^{t-h} f(s, x_3(s))\, ds = x_0 + \int_a^{t_1} f(s, x_3(s))\, ds + \int_{t_1}^{t-h} f(s, x_3(s))\, ds.$$

En la segunda integral se tiene $s \in [t_1, t_2]$ y aquí ya sabemos que $x_3(s) = x_0 + \int_a^{t-h} f(s, x_0)\, ds$. Concluyendo, $x_3(s) = x_0$ si $s \in [a, t_1]$, pero cuando $s \in [t_1, t_2]$ se tiene $x_3(s) = x_0 + \int_a^{s-h} f(u, x_3(u))$, donde ahora hemos vuelto a retardar el tiempo y $u \in [a, t_1]$, por lo que $x_3(u) = x_0$. Se tiene pues

$$\begin{aligned} x_3(t) &= \begin{cases} x_0 & \text{si} \quad a \le t \le t_1, \\ x_0 + \int_a^{t-h} f(s, x_0)\, ds & \text{si} \quad t_1 \le t \le t_2, \\ x_0 + \int_a^{t_1} f(s, x_0)\, ds + \int_{t_1}^{t-h} f(s, x_3(s))\, ds & \text{si} \quad t_2 \le t \le b, \end{cases} \\ &= \begin{cases} x_0 & \text{si} \quad a \le t \le t_1, \\ x_0 + \int_a^{t-h} f(s, x_0)\, ds & \text{si} \quad t_1 \le t \le t_2, \\ x_0 + \int_a^{t_1} f(s, x_0)\, ds + \int_{t_1}^{t-h} f\left(s, x_0 + \int_a^{s-h} f(u, x_0)\, du\right) ds & \text{si} \quad t_2 \le t \le b. \end{cases} \end{aligned}$$

Sean ahora $t, t' \in [a, b]$. Para cada $n \in \mathbb{N}$ se tiene $||x_n(t) - x_n(t')|| \le M|t - t'|$ y por tanto es uniformemente equicontinua sin más que tomar $\delta = \varepsilon/M$ para cada $\varepsilon > 0$ dado. Por otra parte, dado $t \in [a, b]$ arbitrario se tiene

$$||x_n(t)|| \le ||x_0|| + \int_a^{t-h} ||f(s, x_n(s))||\, ds \le ||x_0|| + M(b-a),$$

y por tanto $\{x_n\}$ es uniformemente acotada. El Teorema de Ascoli-Arzela asegura la existencia de una parcial unifórmemente convergente de $\{x_n\}$. Sea $x : [a, b] \to \mathbb{R}^n$ su límite uniforme, el cual es una función continua por ser límite uniforme de funciones continuas. Si denotamos por $h_n = (b-a)/n$ se cumple

$$x_n(t) = x_0 + \int_a^{t-h_n} f(s, x_n(s))\, ds = x_0 + \int_a^t f(s, x_n(s))\, ds - \int_{t-h_n}^t f(s, x_n(s))\, ds.$$

Por un lado, la convergencia uniforme de x_n a x permite permutar integral y límite, obteniendo $\int_a^t f(s, x_n(s))\, ds \to \int_a^t f(s, x(s))\, ds$. Por otro lado estimamos la integral sobrante, concluyendo

$$\left|\left|\int_{t-h_n}^t f(s, x_n(s))\, ds\right|\right| \le M\frac{b-a}{n} \to 0$$

cuando $n \to \infty$. Esto nos dice que la función límite cumple la integral de Volterra

$$x(t) = x_0 + \int_a^t f(s, x(s))\, ds$$

y por tanto $x \in C^1([a, b])$ es solución. $\square$

4.3. Prolongación de soluciones

En la sección anterior hemos probado existencia de soluciones del problema de Cauchy (4.1.1) cuando la función f está definida en un dominio compacto de la forma $I_\alpha(t_0) \times \overline{B}(x_0, r)$. Salvo en los Teoremas 4.1.11 y 4.2.9 en los que f está definida en una banda $I \times \mathbb{R}^n \to \mathbb{R}^n$ y sus soluciones están definidas en todo I, la existencia de soluciones está limitada a intervalos de definición compactos. El objetivo de esta sección es estudiar cuándo podemos garantizar que una solución de un problema de Cauchy está definida en el mayor intervalo posible, y dar condiciones a priori que nos permitan conocer la topología de dicho intervalo.

Definición 4.3.1. Una solución $x : I \subset \mathbb{R} \to \mathbb{R}^n$ del problema de Cauchy (4.1.1) se dice *prolongable* si existe otra solución $y : J \subset \mathbb{R} \to \mathbb{R}^n$ de (4.1.1) tal que $I \subset J$ e $y_{|I} = x$. En caso contrario, la solución x se dice *maximal*.

Otro problema de interés que surge al prolongar soluciones es la posible pérdida de unicidad de la solución al extenderla. El Teorema de Picard-Lindelöf asegura la unicidad de soluciones en el entorno donde garantiza la existencia, pero podría pasar que al prolongar una solución llegásemos a un intervalo donde el Teorema de Picard-Lindelöf deja de ser aplicable y entonces nos preguntamos si en caso de poder prolongar la solución hasta una no prolongable, ésta será también única. Aquí podemos volver a mencionar el problema

$$\begin{cases} x' = 3x^{2/3}, \\ x(0) = x_0. \end{cases}$$

Si $x_0 \neq 0$ entonces la única solución alrededor de $t = 0$ es $x(t) = (t + x_0^{1/3})^3$, la cual está definida en todo $\mathbb{R}$. El problema es que en el momento en que x se anula, en el instante $t_* = -x_0^{1/3}$, perdemos unicidad de la solución. Este es un ejemplo de un problema de Cauchy con las siguientes propiedades.

1. Alrededor de cada punto $(t_0, x_0) \in \mathbb{R}^2$, con $x_0 \neq 0$, pasa una única solución de manera local.

2. Si $x_0 \neq 0$, la solución única local se puede extender a infinitas soluciones maximales, todas definidas en $\mathbb{R}$.

El contexto en el que trabajaremos será el de un problema de Cauchy para una EDO $x' = f(t, x)$, con $f : \Omega \to \mathbb{R}^n$ y Ω abierto. Si Ω no es abierto entonces los resultados que vamos a enunciar y demostrar pueden no ser verdad. Resumiendo, los tres problemas que vamos a tratar son los siguientes:

1. Si existe una única solución de cada problema de Cauchy en un pequeño entorno de la condición inicial, ¿existe una única solución global para cada problema de Cauchy? En otras palabras, ¿unicidad local implica unicidad global? Veremos que sí, y el procedimiento para probar este resultado será un argumento de pegado de soluciones.

2. Dada una solución $x : I \subset \mathbb{R} \to \mathbb{R}^n$, ¿existe un intervalo J con $I \subset J$ y una solución $y : J \to \mathbb{R}^n$ tal que $y_{|_I} = x$? Es decir, nos preguntamos si la solución x es maximal o si por el contrario es prolongable. Veremos que si Ω es abierto entonces una solución que no sea maximal siempre puede ser prolongada, obteniendo eventualmente una solución maximal definida sobre un intervalo abierto.

3. ¿Cómo saber si una solución es maximal? Básicamente vamos a probar que la solución maximal se aproxima a la frontera del abierto donde está definida la ecuación tanto como se desee.

4.3.1. Primer problema. Unicidad local y global

Teorema 4.3.2. *Sea Ω un abierto y supongamos que que cumple la siguiente condición: para cada $(t_0, x_0) \in \Omega$ existe un intervalo $I(t_0, x_0)$, entorno de t_0, tal que el problema de Cauchy*

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

tiene existencia y unicidad sobre $I(t_0, x_0)$. Sean, $x : I \to \mathbb{R}^n, y : J \to \mathbb{R}^n$ son dos soluciones de la EDO $x' = f(t,x)$ y supongamos que existe $t_0 \in I \cap J$ tal que $x(t_0) = y(t_0)$. Entonces, ambas soluciones coinciden en $I \cap J$.

En el caso particular en que $f(t,x)$ sea localmente lipschitziana en el abierto Ω, las hipótesis de este resultado se tienen inmediatamente. En el caso del Teorema de Peano donde sólo tenemos garantizada existencia del problema de Cauchy, la unicidad local debe ser impuesta como hipótesis adicional.

Demostración. Sean x, y dos soluciones de la EDO $x' = f(t,x)$, teniendo cada una como dominio de definición I, J, respectivamente. Notemos que no sabemos nada más de las funciones x, y, como por ejemplo qué condición inicial cumplen. Sólo sabemos que sus dominios tienen intersección no vacía por la existencia de $t_0 \in I \cap J$ tal que $x(t_0) = y(t_0)$. El objetivo es demostrar que $x(t) = y(t)$ para todo $t \in I \cap J$ y para tal fin definimos

$$\mathcal{I} = \{t \in I \cap J \colon x(t) = y(t)\}.$$

Claramente $\mathcal{I} \neq \varnothing$ por ser $t_0 \in \mathcal{I}$. Además, $\mathcal{I}$ es un intervalo por ser intersección de dos intervalos y por tanto es conexo. Para probar que $\mathcal{I} = I \cap J$ basta probar que $\mathcal{I}$ es abierto y cerrado en la topología de $I \cap J$. Que $\mathcal{I}$ es cerrado se tiene inmediatamente por ser x, y funciones continuas. En efecto, si $\{t_n\} \subset \mathcal{I}$ cumple $t_n \to t_\infty \in I \cap J$, entonces $x(t_n) = x(y_n)$ y tomando límite y por unicidad del mismo se tiene $x(t_\infty) = y(t_\infty)$. Dado ahora $t_0 \in \mathcal{I}$, el objetivo es probar que existe un entorno de t_0 en la topología relativa de $I \cap J$ y contenido en $I \cap J$; es decir que existe un entorno de t_0 donde x e y coinciden. Por ser $t_0 \in \mathcal{I}$ de la propia definición se tiene $x_0 := x(t_0) = y(t_0)$. Si por algún casual $I \cap J = \{t_0\}$, hemos acabado de forma trivial. De lo contrario, t_0 es interior a $I \cap J$ en la topología relativa. Por hipótesis, existe un intervalo $I(t_0, x_0)$, entorno de t_0, tal que el problema de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

tiene una única solución sobre $I(t_0, x_0)$. Ahora bien, $I(t_0, x_0) \cap I \cap J$ es un entorno de t_0 en la topología relativa de $I \cap J$, y como x, y son soluciones de $x' = f(t,x)$ que coinciden en t_0, deben coincidir sobre todo $I(t_0, x_0) \cap I \cap J$. Así pues,

$$t_0 \in \overset{\circ}{\widehat{I(t_0, x_0)}} \cap I \cap J \subset \mathcal{I},$$

y t_0 es interior a $\mathcal{I}$ en la topología relativa de $I \cap J$. $\square$

Corolario 4.3.3. *En las condiciones del Teorema 4.3.2, sean $x : I \to \mathbb{R}^n$ e $y : J \to \mathbb{R}^n$ soluciones de $x' = f(x, y)$ tales que $I \cap J \neq \varnothing$ y $x(t_0) = y(t_0)$ para algún $t_0 \in I \cap J$. Entonces,*

$$z(t) = \begin{cases} x(t) & \text{si} \quad t \in I, \\ y(t) & \text{si} \quad t \in J, \end{cases}$$

está bien definida y es solución de $x' = f(t, x)$.

Demostración. Que $z(t)$ está bien definida y es continua se tiene por el Teorema 4.3.2. Basta ver que $z \in C^1$ y que es solución de la EDO $x' = f(t, x)$. Si $I \cap J$ no se reduce a un punto, entonces dado $t_0 \in I \cap J$ o bien t_0 es interior a I o bien es interior a J; supongamos que $t_0 \in \overset{\circ}{I}$. Entonces, z restringida a un entorno de t_0 coincide con x, la cual es solución de la EDO y en particular C^1.

El caso no trivial a tratar es si $I \cap J = \{t_0\}$. Como el problema de Cauchy tiene existencia y unicidad local, llamamos $x_0 = x(t_0) = y(t_0)$ y consideramos $\hat{z} : [t_0 - \varepsilon, t_0 + \varepsilon] \to \mathbb{R}^n$ la solución de la EDO con condición inicial $\hat{z}(t_0) = x_0$. De hecho, no es restrictivo tomar $\varepsilon > 0$ suficientemente pequeño para que el intervalo $[t_0 - \varepsilon, t_0 + \varepsilon]$ sea el que satisface la hipótesis del Teorema 4.3.2 de unicidad local alrededor de t_0. Por unicidad, en $[t_0 - \varepsilon, t_0] \cap I$ se tiene $\hat{z} = x$, e igualmente en $[t_0, t_0 + \varepsilon] \cap J$ se tiene $\hat{z} = y$. Esto nos dice que para todo $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ las funciones $\hat{z}$ y z coinciden y de nuevo por unicidad necesariamente $\hat{z} = z$. $\square$

La respuesta al problema de unicidad local implica unicidad global se responde mediante el siguiente resultado.

Teorema 4.3.4. *Supongamos que el problema de Cauchy*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0, \end{cases}$$

tiene existencia y unicidad local para cada condición inicial y sea $x : I \to \mathbb{R}^n$ una solución maximal. Si $y : J \to \mathbb{R}^n$ es solución de $x' = f(t, x)$ y existe $t_0 \in J$ tal que $x(t_0) = y(t_0)$ entonces necesariamente $J \subset I$ y $x_{|_J} = y$.

Demostración. Como $t_0 \in I \cap J$, el Corolario 4.3.3 asegura que la función

$$z(t) = \begin{cases} x(t) & \text{si} \quad t \in I, \\ y(t) & \text{si} \quad t \in J, \end{cases}$$

está bien definida y es solución del problema de Cauchy. Como x es maximal necesariamente $I \cup J = I$ y en particular $J \subset I$. En tal caso, el Teorema 4.3.2 asegura que x, y coinciden en $I \cap J = J$ y en consecuencia $x_{|_J} = y$. $\square$

4.3.2. Segundo problema. Existencia de soluciones maximales

Como hemos puesto de manifiesto en el comienzo de esta sección, las soluciones cuya existencia viene dada por los Teoremas de Picard-Lindelöf o de Peano en sus versiones locales están definidas en un intervalo cerrado $I_\alpha(t_0) = [t_0 - \alpha, t_0 + \alpha]$. En particular, en el caso de la versión local en un abierto Ω, la gráfica de la solución cumple $(t, x(t)) \in \Omega$ para todo $t \in I_\alpha(t_0)$ y en particular los puntos extremos cumplen $(t_0 \pm \alpha, x(t_0 \pm \alpha)) \in \Omega$. Esto nos permite volver a considerar el problema de Cauchy $x' = f(t, x)$ con estas condiciones iniciales, para las cuales existirán nuevas soluciones. Lo interesante de esta cuestión es que estas soluciones necesariamente coinciden con x en un entorno de los instantes $t_0 \pm \alpha$, obteniendo una prolongación de nuestra solución original.

En general, a partir de ahora enunciaremos los resultados de prolongación a la derecha del intervalo, pero esta situación debe entenderse de igual manera siempre por la izquierda (en caso que sea posible). Recordamos, e incidimos, que el conjunto Ω donde está definida la función $f(t, x)$ es abierto de $\mathbb{R} \times \mathbb{R}^n$. Además, salvo que se especifique lo contrario, la única hipótesis sobre $f : \Omega \to \mathbb{R}^n$ es su continuidad.

El siguiente resultado nos dice que si una solución está definida en un intervalo abierto y acotado y tiene límite en sus extremos, ésta se puede extender hasta los extremos.

Lema 4.3.5. *Sean $f : \Omega \to \mathbb{R}^n$ continua y $x : (a, b) \to \mathbb{R}^n$ una solución del problema de Cauchy*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0. \end{cases}$$

Supongamos que $\lim_{t \to b} x(t) = x_1$ y además $(b, x_1) \in \Omega$. Entonces, la extensión continua $\tilde{x} : (a, b] \to \mathbb{R}^n$ dada por

$$\tilde{x}(t) = \begin{cases} x(t) & \text{si} \quad t \in (a, b), \\ x_1 & \text{si} \quad t = b, \end{cases}$$

es también solución del problema de Cauchy.

Demostración. Sabemos que $\tilde{x}$ es solución si y sólo si para cada $t \in (a, b]$ se cumple la ecuación integral

$$\tilde{x}(t) = x_0 + \int_{t_0}^{t} f(s, \tilde{x}(s))\, ds.$$

Observemos que para cada $t \in (a, b]$ la función $t \mapsto f(t, \tilde{x}(t))$ es continua por ser $f(t, x)$ y $\tilde{x}$ funciones continuas. Además, si $t \neq b$ la ecuación integral es cierta por ser $\tilde{x}(t) = x(t)$ solución. Ahora bien,

$$\tilde{x}(b) = \lim_{t \to b} \tilde{x}(t) = x_0 + \lim_{t \to b} \int_{t_0}^{t} f(s, \tilde{x}(s))\, ds = x_0 + \int_{t_0}^{b} f(s, \tilde{x}(s))\, ds,$$

por ser $t \mapsto \int_{t_0}^{t} f(s, \tilde{x}(s))\, ds$ una función continua (es de hecho C^1 por el Teorema Fundamental del Cálculo). Como $\tilde{x}$ satisface la ecuación integral en todo $t \in (a, b]$, es solución del problema de Cauchy. □

El siguiente resultado nos proporciona una herramienta fundamental para poder extender soluciones. Lo interesante del mismo es su dualidad: bien podemos *pegar* dos soluciones que coinciden en un punto, o extender una solución dada sin necesidad de tener otra solución candidata a pegar.

Lema 4.3.6. *Sean $f : \Omega \to \mathbb{R}^n$ continua y $x : [a,b] \to \mathbb{R}^n$ una solución de la EDO $x' = f(t,x)$. Entonces, x es prolongable a un intervalo $[a, b+\varepsilon]$ para cierto $\varepsilon > 0$.*

Demostración. Por ser Ω abierto, se tiene $(b, x(b)) \in \Omega$. Podemos considerar pues el problema de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(b) = x_0. \end{cases}$$

Sabemos por el Teorema de Peano que existe una solución de dicho problema $y : [b-\varepsilon, b+\varepsilon] \to \mathbb{R}^n$. Consideremos la función

$$z(t) = \begin{cases} x(t) & \text{si} \quad t \in [a,b], \\ y(t) & \text{si} \quad t \in [b, b+\varepsilon], \end{cases}$$

la cual está bien definida y es continua. Es evidente que z extiende a x al intervalo $[a, b+\varepsilon]$ y que z es de clase C^1 en todos los puntos salvo quizás en $t = b$. Si calculamos las derivadas laterales de $z(t)$ en $t = b$ y usamos que $f(t,x)$ es continua y que x, y son soluciones del problema de Cauchy, obtenemos

$$\lim_{t \to b^-} z'(t) = \lim_{t \to b^-} x'(t) = f(b, x_0) = \lim_{t \to b^+} y'(t) = \lim_{t \to b^+} z'(t).$$

Por tanto, $z'(t)$ tiene ambos límites laterales iguales y se tiene $z'(b) = f(b, z(b))$. Concluimos que la función z' es continua en $t = b$ y por tanto z es de clase C^1 en $[a, b+\varepsilon]$. □

En particular, hemos probado que si $x : [a,b] \to \mathbb{R}^n$ e $y : [b,c] \to \mathbb{R}^n$ son soluciones de $x' = f(t,x)$ tales que $x(b) = y(b)$, entonces ambas soluciones pueden ser pegadas en $t = b$ para obtener una extensión $z : [a,c] \to \mathbb{R}^n$. Con estos resultados podemos dar una demostración alternativa del Teorema 4.2.9 de Peano en una banda.

Demostración del Teorema 4.2.9 de Peano en una banda. Supongamos primero $I = [a,b]$. Sea pues $f : [a,b] \times \mathbb{R}^n \to \mathbb{R}^n$ una función continua y acotada y denotemos $M = \sup\{||f(t,x)|| \colon (t,x) \in [a,b] \times \mathbb{R}^n\}$. Dado $t_0 \in [a,b]$, primero construiremos una solución en el intervalo $[t_0, b]$ y posteriormente otra solución en el intervalo $[a, t_0]$ para posteriormente pegarlas en virtud del Lema 4.3.6. Obviamente, si $t_0 = a, b$ entonces nos ahorramos una de las etapas. Supongamos pues $t_0 < b$ y construyamos una solución en $[t_0, b]$. Sean $\varepsilon = b - t_0$ y $r = M\varepsilon$ y consideramos $f : [t_0, t_0+\varepsilon] \times \overline{B}(x_0, r) \to \mathbb{R}^n$, la cual está bien definida, especialmente en su segunda variable, ya que en ésta su dominio es todo $\mathbb{R}^n$. El Teorema de Peano 4.2.8 asegura la existencia de una solución $x_1 : [t_0, t_0+\alpha] \to \overline{B}(x_0, r)$, siendo $\alpha = \min\{\varepsilon, r/M\} = \varepsilon = b - t_0$ y en consecuencia $x_1 : [t_0, b] \to \mathbb{R}^n$. Ahora construimos $x_2 : [a, t_0] \to \overline{B}(x_0, r)$ otra solución y concluimos por el Lema 4.3.6 que el pegado de estas dos funciones en t_0 define una solución $x : [a,b] \to \overline{B}(x_0, r)$.

Consideremos ahora $I \subset \mathbb{R}$ arbitrario y $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ continua y acotada. De nuevo, dado $t_0 \in I$ el objetivo es construir dos soluciones; una a la derecha de t_0 y la otra a la izquierda de t_0 para posteriormente pegarlas. Haremos el caso a la derecha. Supongamos $[t_0, \infty) \cap I = [t_0, b)$ con $t_0 < b \leq \infty$. Vamos a tomarnos una sucesión $\{t_n\}_{n \geq 0} \subset [t_0, b)$ tal que $t_n \nearrow b$. La idea es ir construyendo soluciones en cada intervalo $[t_n, t_{n+1}]$ e ir pegándolas en sus extremos gracias al Lema 4.3.6.

Si llamamos $\varepsilon_0 = t_1 - t_0$ y $r_0 = \varepsilon_0 M$, entonces $\alpha_0 = \min\{\varepsilon_0, r_0/M\} = \varepsilon_0$ y podemos aplicar el Teorema de Peano 4.2.8 para construir una solución del problema de Cauchy $x_0 : [t_0, t_1] \to \mathbb{R}^n$ globalmente definida en $[t_0, t_1]$ y con imagen en $\overline{B}(x_0, r_0)$. En este punto, la hipótesis de acotación global de f es clave, ya que estamos definiendo el radio de la bola en términos del supremo de

$||f(t,x)||$, sin preocuparnos si f en esta bola va a superar tal supremo. Es decir, si f no estuviese acotada y partiésemos de un compacto $K_0 = [t_0, t_1] \times \overline{B}(x_0, r_0)$ con un radio arbitrario, llamaríamos $M_0 = \text{máx}\{||f(t,x)||\colon (t,x) \in K_0\}$ y consideraríamos $f : K_0 \to \mathbb{R}^n$. Ahora bien, si $\varepsilon_0 = t_1 - t_0$ entonces $\alpha_0 = \text{mín}\{\varepsilon_0, r/M_0\}$ y es posible que r/M_0 fuese menor que ε_0. Si considerásemos el nuevo radio $r_* = \varepsilon_0 M_0$ y el nuevo compacto $K_1 = [t_0, t_1] \times \overline{B}(x_0, r_*)$, podría ocurrir que $M_1 = \text{máx}\{||f(t,x)||\colon (t,x) \in K_1\}$ fuese mayor que M_0 y en tal caso $\alpha_* = \text{mín}\{\varepsilon_0, r_*/M_1\} = \text{mín}\{\varepsilon_0, \varepsilon_0 M_0/M_1\} = \varepsilon_0 M_0/M_1$, haciendo disminuir el valor de α_* e imposibilitándonos llegar a t_1.

En cualquier caso, como f está cotada podemos construir una solución x_1 que satisface la condición inicial $x(t_1) = x_0(t_1) = x_1$ y definida en todo $[t_1, t_2]$. Para ello, definimos $\varepsilon_1 = t_2 - t_1$ y $r_1 = \varepsilon_1 M$ y por tanto $\alpha_1 = \text{mín}\{\varepsilon_1, r_1/M\} = \varepsilon_1$. En general, dada x_{n-1}, definiríamos $x_n : [t_n, t_{n+1}] \to \mathbb{R}^n$ como una solución del problema de Cauchy con condición inicial $x(t_n) = x_{n-1}(t_n) = x_n$ y definida en todo $[t_n, t_{n+1}]$. Por una parte, es claro que se tiene $[t_0, b) = \cup_{n=1}^{\infty}[t_n, t_{n+1}]$ y por otra parte, podemos definir la función $x : [t_0, b) \to \mathbb{R}^n$ tal que $x(t) = x_n(t)$ si $t \in [t_n, t_{n+1}]$. Esta función está bien definida, puesto que para cada $t \in [t_0, b)$ existe un único $n \in \mathbb{N}$ tal que $t \in [t_n, t_{n+1}]$, salvo que t sea exactamente uno de los extremos en los que precisamente cada par de funciones consecutivas x_k, x_{k+1} coinciden.

De igual manera se considera $(a, t_0] = (-\infty, t_0] \cap I$ con $-\infty \leq a < t_0$ y se construye una solución $y : (a, t_0] \to \mathbb{R}^n$. Finalmente, las soluciones x e y se pegan en t_0 en virtud de nuevo del Lema 4.3.6 para obtener una solución definida en todo I. □

Observación 4.4. Como hemos puesto de manifiesto, la acotación global de f es fundamental. Consideremos el problema de Cauchy

$$\begin{cases} x' = x^2, \\ x(0) = 1. \end{cases}$$

Como $f(t,x) = x^2$ está definida en $\mathbb{R}^2$ y es C^1, en particular es localmente Lipschitziana y cada problema de Cauchy tiene existencia y unicidad local. La única solución de tal problema de Cauchy viene dada por la función $x(t) = \frac{1}{1-t}$, $t \in (-\infty, 1)$, la cual es maximal en el sentido de la Definición (4.3.1). Que el intervalo de definición de x no sea todo $\mathbb{R}$ es consecuente con la falta de acotación de x^2 en $\mathbb{R}$. Es decir, incluso restringiendo $f : [0,2] \times \mathbb{R} \to \mathbb{R}$, la función $f(t,x)$ no está acotada y esto hace que las funciones sucesivas $\{x_n\}$ que se van pegando cumplan que sus intervalos de definición $[t_n, t_{n+1}]$ convergen a $t = 1$ antes incluso de poder converger a $t = 2$.

El siguiente resultado resuelve el problema de la prolongación en el contexto más general posible.

Teorema 4.3.7. *Toda solución $x : I \to \mathbb{R}^n$ del problema de Cauchy*

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

puede prolongarse a una solución maximal $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$, con $-\infty \leq \omega_- < \omega_+ \leq \infty$. Si en particular $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ es abierto y $f : \Omega \to \mathbb{R}^n$ es continua, φ está definida en un intervalo abierto.

Demostración. Dada $x : I \to \mathbb{R}^n$ solución, escribimos el par (x, I) para denotar la solución x definida sobre I. Con esta notación consideramos el conjunto

$$\mathcal{A} = \{(z, K)\colon I \subset K,\ \text{y}\ z_{|_I} = x\},$$

donde $z : K \to \mathbb{R}^n$ es solución del problema de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0. \end{cases}$$

Definimos en $\mathcal{A}$ la relación binaria $' \leq'$ mediante $(K_1, z_1) \leq (K_2, z_2)$ si y sólo si $K_1 \subset K_2$ y $z_{2|_{K_1}} = z_1$. Así definida, $' \leq'$ es un orden parcial sobre $\mathcal{A}$. Puesto que $\mathcal{A} \neq \varnothing$ ya que $(x, I) \in \mathcal{A}$, si demostramos que todo subconjunto $\mathcal{B} \subset \mathcal{A}$ totalmente ordenado tiene una cota superior en $\mathcal{A}$, el Lema de Zorn garantizará la existencia de un elemento maximal en $\mathcal{A}$ para este orden.

Sea pues $\mathcal{B} = \{(K_i, z_i) \colon i \in C\}$ totalmente ordenado, donde C es un conjunto arbitrario de índices y cada K_i es un intervalo. Definimos $K = \cup_{i \in C} K_i$ y $z : K \to \mathbb{R}^n$ de la forma $z(t) = z_i(t)$ si $i \in K_i$. Primero notemos que K es en sí mismo un intervalo, puesto que es unión arbitraria de intervalos con intersección no vacía ya que $I \subset K_i$ para todo $i \in C$. Por otra parte, $z(t)$ está bien definida puesto que si $t \in K_i \cap K_j$ entonces necesariamente $K_i \subset K_j$, o viceversa. En cualquier caso, $z_i(t) = z_j(t)$ y en particular $z_{|_{K_i}} = z_i$. Para concluir que (z, K) es cota superior de $\mathcal{B}$ basta ver que de hecho $z : K \to \mathbb{R}^n$ es solución del problema de Cauchy, es decir que $z \in C^1$ y $z'(t) = f(t, z(t))$ para todo $t \in K$. Sea pues $t \in K$ arbitrario. Existe K_i tal que $t \in K_i$; de hecho puede suponerse que t es interior a K_i. En cualquier caso, $z_{|_{K_i}} = z_i$ y en particular z es de clase C^1 en t por serlo z_i. Además, se tiene $z'(t) = z_i'(t) = f(t, z_i(t)) = f(t, z(t))$. Como $t \in K$ era arbitrario concluimos que $z : K \to \mathbb{R}^n$ es solución de la EDO y por tanto $(z, K) \in \mathcal{A}$, lo que prueba que es cota superior de $\mathcal{B}$.

En esta situación el Lema de Zorn asegura la existencia de un elemento maximal en $\mathcal{A}$, esto es existe $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$ solución del problema de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

la cual no puede prolongarse más en su dominio.

Para el caso particular en que $f : \Omega \to \mathbb{R}^n$ sea continua y esté definida en un abierto, probemos que el intervalo de definición de φ es de la forma (ω_-, ω_+) y en particular abierto. Por reducción al absurdo supongamos que no es abierto y no es todo $\mathbb{R}$. Sin perder generalidad, supongamos que a la derecha de t_0 el intervalo es de la forma $[t_0, \omega_+]$. Ahora bien, como la gráfica $(t, \varphi(t)) \subset \Omega$ para cada $t \in (\omega_-, \omega_+]$ y Ω es abierto, necesariamente $(\omega_+, \varphi(\omega_+))$ es un punto del interior de Ω. En este punto podemos aplicar el Lema 4.3.6 para extender φ a una función $\tilde{\varphi} : [t_0, \omega_+ + \varepsilon] \to \mathbb{R}^n$ para cierto $\varepsilon > 0$, y tal extensión sigue siendo solución de la EDO. En consecuencia $(\tilde{\varphi}, (\omega_-, \omega_+ + \varepsilon]) \in \mathcal{A}$ y esto contradice la maximalidad de $(\varphi, (\omega_-, \omega_+))$ en $\mathcal{A}$. □

La notación que usaremos para denotar la solución maximal será $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$. Como consecuencia de los resultados anteriormente probados, podemos resumir el segundo problema de prolongación de la siguiente forma.

Teorema 4.3.8. *Sea $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ abierto y $f : \Omega \to \mathbb{R}^n$ continua. Para cada problema de Cauchy*

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

existe una solución maximal $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$. Si además f es localmente lipschitziana en su segunda variable, la solución maximal es única.

Demostración. El Teorema de Peano 4.2.8 garantiza la existencia de una solución $x : I_\alpha(t_0) \to \mathbb{R}^n$ del problema de Cauchy para cada $(t_0, x_0) \in \Omega$. Por el Teorema 4.3.7, la solución x se puede extender a una solución maximal $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$, la cual necesariamente debe estar definida en un intervalo abierto por ser f continua y Ω abierto.

Referente a la segunda afirmación, si f es localmente lipschitziana en su segunda variable en Ω, tenemos existencia y unicidad local para cada problema de Cauchy. Ahora terminaríamos como en la demostración del Teorema 4.3.4. □

Notemos que el Teorema 4.3.7 sólo nos dice que una solución del problema de Cauchy se puede prolongar a una maximal, pero si para una misma condición inicial $x(t_0) = x_0$ existen varias soluciones, cada una de ellas puede ser prolongada a una maximal. En general, cada solución maximal puede estar definida en un intervalo distinto.

4.3.3. Tercer problema. Comportamiento de las soluciones maximales

El objetivo de esta sección es abordar el último problema de la prolongación: estudiar el comportamiento de las soluciones maximales cuando el parámetro tiende a los extremos del intervalo maximal de definición. La siguiente definición se antoja totalmente necesaria para tal fin.

Definición 4.3.9. Sea $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ y $\gamma : I = (\omega_-, \omega_+) \to \mathbb{R}^n$ una curva tal que $\mathrm{graf}(\gamma) = \{(t, \gamma(t)) \colon t \in I\} \subset \Omega$. Se dice que γ tiende a la frontera de Ω cuando $t \to \omega_+$ si para todo compacto $K \subset \Omega$ existe $\varepsilon > 0$ tal que si $t \in (\omega_+ - \varepsilon, \omega_+)$ se tiene $(t, \gamma(t)) \notin K$.

De forma análoga se define cuando $t \to \omega_-$ si para todo compacto K existe $\varepsilon > 0$ tal que si $t \in (\omega_-, \omega_- + \varepsilon)$ se tiene $(t, \gamma(t)) \notin K$.

Intuitivamente, una curva tiende a la frontera de Ω si se *sale* de cualquier compacto $K \subset \Omega$ en tiempo finito y no vuelve a él.

Notemos que esta definición tiene sentido incluso si $\Omega = \mathbb{R} \times \mathbb{R}^n$; lo que estaríamos diciendo es que la gráfica de γ es no acotada. Además, esta observación se refiere a la gráfica $(t, \gamma(t))$ y no únicamente a la propia imagen $\gamma(t)$. Podría darse el caso $t \in (-\infty, \infty)$, que $\gamma(t)$ estuviese contenida en una región acotada y $t \to \infty$ hiciese que γ tendiese a la frontera de Ω. Por ejemplo, $\gamma(t) = (\cos t, \sin t)$ es una circunferencia de radio 1 (en particular compacta), pero su gráfica $\mathrm{graf}(\gamma) = \{(t, \cos t, \sin t) \colon t \in \mathbb{R}\}$ es una hélice que se sale de cualquier compacto de $\mathbb{R}^3$.

Recordemos que si $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ se suponía compacto entonces no podemos asegurar la conclusión final del Teorema 4.3.7 que probaba que la solución maximal tenía que estar definida en un abierto. El siguiente resultado prueba que de hecho este hecho es imposible.

Teorema 4.3.10. *Sean $K \subset \mathbb{R} \times \mathbb{R}^n$ compacto, $f : K \to \mathbb{R}^n$ continua y $\varphi : I \to \mathbb{R}^n$ una solución maximal del problema de Cauchy*

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0, \end{cases}$$

Entonces, $I = [\omega_-, \omega_+]$ y $(\omega_\pm, \varphi(\omega_\pm))$ pertenecen a ∂K.

Demostración. Sabemos que la solución maximal cumple $(t, \varphi(t)) \in K$ para todo $t \in I$ por ser solución y en particular $I \neq \mathbb{R}$. Sea pues $\omega_+ < \infty$ el extremo superior de I. Primero probaremos que necesariamente $\omega_+ \in I$ y posteriormente que $(\omega_+, \varphi(\omega_+)) \in \partial K$. Como de costumbre, llamamos $M = \max_{(t,x) \in K} ||f(t, x)||$, que existe por ser K compacto y f continua.

Por reducción al absurdo, supongamos $\omega_+ \notin I$ y consideremos el intervalo $[t_0, \omega_+)$. Sea $\{t_n\} \subset I$ una sucesión tal que $t_n \to \omega_+$. Dados $n, m \in \mathbb{N}$ la ecuación integral de Volterra (3.1.2) implica

$$||\varphi(t_n) - \varphi(t_m)|| \leq \int_{t_n}^{t_m} ||f(s, \varphi(s))||\, ds \leq M|t_n - t_m| \to 0, \qquad n, m \to \infty.$$

En particular, la sucesión $\{\varphi(t_n)\}$ es de Cauchy y existe una parcial $\varphi(t_{\sigma(n)}) \to L_+$. Además, L_+ tiene que ser un valor finito puesto que $(t_{\sigma(n)}, \varphi(t_{\sigma(n)})) \in K$, el cual es compacto, luego $(\omega_+, L_+) \in K$. La acotación dada por la integral de Volterra nos dice que de hecho *cualquier* otra sucesión $t'_n \to \omega_+$ debe cumplir $\varphi(t'_n) \to L_+$. En efecto, por una parte

$$||\varphi(t'_n) - \varphi(t_{\sigma(n)})|| \leq M|t'_n - t_{\sigma(n)}| \to 0, \qquad n \to \infty,$$

y por otra parte

$$||\varphi(t'_n) - L_+|| \leq ||\varphi(t'_n) - \varphi(t_{\sigma(n)})|| + ||\varphi(t_{\sigma(n)}) - L_+|| \to 0, \qquad n \to \infty.$$

Concluimos que $\lim_{t \to \omega_+} \varphi(t) = L_+$ existe y es finito, y además $(\omega_+, L_+) \in K$. Pero ahora podemos aplicar el Lema 4.3.5 para extender φ al extremo superior ω_+ obteniendo una solución prolongada, en contradicción con la maximalidad de φ. Con esto probamos que $\omega_+ \in I$.

Ya sabemos que $(\omega_+, \varphi(\omega_+)) \in K$. Si $(\omega_+, \varphi(\omega_+)) \notin \partial K$ necesariamente pertenece al interior de K. En estas condiciones el Lema 4.3.6 nos permitiría prolongar la solución φ a la derecha hasta $[t_0, \omega_+ + \alpha]$, para $\alpha > 0$ suficientemente pequeño. Esto es contradictorio con la maximalidad de φ.

Ahora repetiríamos el argumento a la izquierda de t_0, concluyendo que $\omega_- \in I$ y $(\omega_-, \varphi(\omega_-)) \in \partial K$. Necesariamente se tiene $I = [\omega_-, \omega_+]$ y $(\omega_\pm, \varphi(\omega_\pm)) \in \partial K$. □

El siguiente resultado es un lema técnico que usaremos en la prueba del último resultado del curso y es conocido en la literatura como Lema de Wintner.

Lema 4.3.11 (Lema de Wintner)**.** *Sean $f : \Omega \to \mathbb{R}^n$ y $x : (\omega_-, \omega_+) \to \mathbb{R}^n$, con $\omega_+ < \infty$, una solución de $x' = f(t, x)$. Supongamos que existe una sucesión $\{t_n\} \subset (\omega_-, \omega_+)$ tal que $t_n \to \omega_+$ y $x(t_n) \to L_+ \in \mathbb{R}^n$. Si existe un entorno V de (ω_+, L_+) tal que f es acotada sobre $V \cap \Omega$, entonces $L_+ = \lim_{t \to \omega_+} x(t)$.*

En particular, si el valor $f(t_0, L_+)$ está definido o puede ser definido de forma que f sea continua en (ω_+, L_+), entonces x se puede prolongar al intervalo $(\omega_-, \omega_+]$ de forma C^1 y es solución de $x' = f(t, x)$.

Demostración. Sea $t_n \to \omega_+$ tal que $(t_n, x(t_n)) \to (\omega_+, L_+)$. Aquí incidimos en que $(\omega_+, L_+) \in \overline{\Omega}$, pudiendo pertenecer a su frontera si Ω fuese abierto. Denotamos igualmente por f a la restricción $f : V \cap \Omega \to \mathbb{R}^n$, la cual está acotada por hipótesis y sea $M = \sup\{||f(t, x)\colon (t, x) \in V \cap \Omega\}$. Más aún, dado $\varepsilon > 0$ pequeño para que $\omega_+ - \varepsilon > \omega_-$, no es restrictivo suponer $V = [\omega_+ - \varepsilon, \omega_+] \times \overline{B}(L_+, \varepsilon)$.

Supongamos que el enunciado es falso. Existe $\varepsilon > 0$ y $t'_n \to \omega_+$ tal que $||x(t'_n) - L_+|| \geq \varepsilon$ para todo n. Notemos que podemos suponer sin perder generalidad que $t'_n \in [\omega_+ - \varepsilon, \omega_+)$ para todo n. Además, por ser (ω_+, L_+) un punto límite, para cada n podemos encontrar $t''_n > t'_n$ tal que $t''_n \to \omega_+$ y $||x(t''_n) - L_+|| < \varepsilon/2$ (bastaría tomar una subsucesión adecuada de t_n). Si nos centramos en el intervalo $[t'_n, t''_n]$ y definimos la función real

$$g : [t'_n, t''_n] \to \mathbb{R}, \qquad g(t) = ||x(t) - L_+||,$$

se tiene $g(t_n') \geq \varepsilon$ y $g(t_n'') < \varepsilon/2$ y por tanto existe $s_n \in (t_n', t_n'')$ (maximal) tal que $g(s_n) = \varepsilon$. En particular, $g(t) \leq \varepsilon$ para todo $t \in [s_n, t_n'']$ lo cual implica $||x(t)-L_+|| \leq \varepsilon$, es decir $x(t) \in \overline{B}(L_+, \varepsilon)$. Además, por ser $t \in [s_n, t_n''] \subset [t_n', t_n'']$ se tiene en particular $t \in [\omega_+ - \varepsilon, \omega_+]$ y en consecuencia $(t, x(t)) \in V \cap \Omega$ para todo $t \in [s_n, t_n'']$. Por una parte, la desigualdad triangular inversa implica

$$||x(t_n'') - x(s_n)|| = ||x(t_n'') - L_+ + L_+ - x(s_n)|| \geq ||x(s_n) - L_+|| - ||x(t_n'') - L_+|| = \varepsilon/2.$$

Pero por otra parte, la ecuación integral de Volterra implica

$$||x(t_n'') - x(s_n)|| \leq \int_{s_n}^{t_n''} ||f(t, x(t))||\, dt \leq M|t_n'' - s_n|,$$

que tiende a cero cuando $n \to \infty$, llegando a contradicción. □

Podría pensarse, erróneamente, que se podría haber evitado la enrevesada construcción de la sucesión s_n y proceder directamente como en la demostración del Teorema 4.3.10, probando que *cualquier* sucesión $t_n' \to \omega_+$ cumple $x(t_n') \to L_+$. No obstante, a poco que se analiza la situación debemos darnos cuentas que esta forma de proceder es errónea. En efecto, si $t_n' \to \omega_+$ entonces no podemos garantizar que $(t_n', x(t_n')) \in V \cap \Omega$, que es donde podemos acotar f por su supremo en la desigualdad integral de Volterra. Es decir, en el paso

$$||\varphi(t_{\sigma(n)}) - \varphi(t_n')|| \leq \int_{t_{\sigma(n)}}^{t_n'} ||f(s, \varphi(s))||\, ds \leq M|t_{\sigma(n)} - t_n'| \to 0, \qquad n \to \infty,$$

no tenemos garantizado que dado $s \in [t_{\sigma(n)}, t_n']$ se tenga $(s, \varphi(s)) \in V \cap \Omega$ para acotar $||f(s, \varphi(s))|| \leq M$.

Estamos ahora preparados para enunciar el resultado principal de esta sección, el cual caracteriza a las soluciones maximales por salirse de cualquier compacto.

Teorema 4.3.12. *Sean Ω abierto, $f : \Omega \to \mathbb{R}^n$ continua y $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$ una solución maximal de $x' = f(t, x)$. Entonces, φ tiende a la frontera de Ω cuando $t \to \omega_\pm$.*

Demostración. Demostraremos sólo el caso a la derecha, es decir para ω_+. Por reducción al absurdo, supongamos que φ no tiende a la frontera de Ω cuando $t \to \omega_+$, en sentido de la Definición 4.3.9. Esto significa que existe un compacto $K \subset \Omega$ tal que para cada $\varepsilon > 0$ existe $t_\varepsilon \in (\omega_+ - \varepsilon, \omega_+)$ tal que $(t_\varepsilon, \varphi(t_\varepsilon)) \in K$. En particular, podemos tomar $\varepsilon = 1/n$ y construir una sucesión $\{t_n\} \subset (\omega_-, \omega_+)$ tal que $t_n \to \omega_+$ y $(t_n, \varphi(t_n)) \in K$. De aquí ya tenemos una implicación: $\omega_+ < \infty$ ya que la proyección $\pi_1(t_n, \varphi(t_n)) = t_n$ restringida a K es continua, por tanto $\pi_1(K)$ es compacto y $t_n \to \omega_+ \in \pi_1(K)$. Más aún, la compacidad de K garantiza la existencia de una parcial $\{t_{\sigma(n)}\}$ tal que $(t_{\sigma(n)}, \varphi(t_{\sigma(n)}))$ es convergente a un punto de K. Por una parte $t_{\sigma(n)} \to \omega_+$ y denotando $L_+ = \lim \varphi(t_{\sigma(n)})$ se tiene $(\omega_+, L_+) \in K$ y en particular L_+ es finito. Ahora podemos tomar V un entorno compacto de (ω_+, L_+) y entonces $f : V \cap \Omega \to \mathbb{R}^n$ estará trivialmente acotada, por lo que el Lema de Wintner 4.3.11 asegura $\lim_{t \to \omega_+} \varphi(t) = L_+$. De hecho, $(\omega_+, L_+) \in K \subset \Omega$ y por tanto (ω_+, L_+) es punto interior de Ω.

En estas condiciones, el Lema 4.3.5 nos permitiría extender φ por continuidad al intervalo $(\omega_-, \omega_+]$; de hecho el Lema 4.3.6 nos permitiría extender φ un poco más a la derecha a un intervalo de la forma $[t_0, \omega_+ + \varepsilon]$. En cualquier caso, obtendríamos una solución del problema de Cauchy que extiende a φ, lo cual contradice su maximalidad. □

Finalizamos el tema resumiendo el posible comportamiento de una solución maximal, y que será lo que eventualmente se usará en la mayoría de argumentos para probar la maximalidad o las propiedades de ciertas soluciones. Como de costumbre, lo enunciamos a la derecha de su intervalo maximal de definición, siendo totalmente análogo a la izquierda.

Teorema 4.3.13. *Sea $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$ una solución maximal de $x' = f(t,x)$, con $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ continua. Si $\omega_+ < \infty$, se cumple una de las siguientes alternativas:*

1. $\lim_{t\to\omega_+} ||\varphi(t)|| = \infty$;
2. *existe $\{t_n\} \to \omega_+$ tal que $\varphi(t_n) \to L_+ \in \mathbb{R}^n$ y además $(\omega_+, L_+) \in \partial\Omega$.*

Demostración. Supongamos que no se cumple el primer punto. Existe por tanto $t_n \to \omega_+$ tal que $||\varphi(t_n)||$ está acotada y por Bolzano-Weierstrass podemos tomar $t_{\sigma(n)}$ tal que $\varphi(t_{\sigma(n)}) \to L_+ \in \mathbb{R}^n$. En particular, $(\omega_+, L_+) \in \overline{\Omega}$. Si $(\omega_+, L_+) \in \text{int}(\Omega)$ y tras restringir f a un entorno compacto de (ω_+, L_+), el Lema de Wintner aseguraría $\lim_{t\to\omega_+} \varphi(t) = L_+$. En esta situación los Lemas 4.3.5 y 4.3.6 implican que podríamos tomar la extensión $\varphi : (\omega_-, \omega_+ + \varepsilon] \to \mathbb{R}^n$, para cierto $\varepsilon > 0$ suficientemente pequeño, lo que contradice el carácter maximal de φ. Queda por tanto como única opción $(\omega_+, L_+) \in \partial\Omega$. □

Corolario 4.3.14. *Sea $f : (a,b) \times \mathbb{R}^n \to \mathbb{R}^n$ continua y $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$ una solución maximal de $x' = f(t,x)$. Si $\omega_+ < b$ entonces $\lim_{t\to\omega_+} ||\varphi(t)|| = \infty$.*

Demostración. Damos dos demostraciones de este hecho para que el alumno tenga herramientas suficientes para aplicar la teoría de prolongación, una basada en el Teorema 4.3.13 y otra basada en el Teorema 4.3.12.

1. La demostración es inmediata a partir del Teorema anterior. Si $||\varphi(t)||$ no tendiese a infinito cuando $t \to \omega_+$ entonces $(t, \varphi(t))$ tendría alguna parcial convergente a un punto finito (ω_+, L_+). Como $\omega_+ < b$ necesariamente $(\omega_+, L_+) \in (a,b) \times \mathbb{R}^n$ y podemos encontrar V entorno compacto de (ω_+, L_+) contenido en $(a,b) \times \mathbb{R}^n$ donde f está acotada. El Lema de Wintner 4.3.11 asegura que $\lim_{t\to\omega_+} \varphi(t) = L_+$ y podríamos extender φ en virtud del Lema 4.3.5 ó 4.3.6; en cualquier caso, llegamos a contradicción.

2. Fijemos $t_0 \in (\omega_-, \omega_+)$ y para cada $n \in \mathbb{N}$ consideremos la familia de compactos $K_n = [t_0, \omega_+] \times \overline{B}(0,n) \subset (a,b) \times \mathbb{R}^n$. Por ser φ maximal existe $\varepsilon_n > 0$ tal que para todo $t \in (\omega_+ - \varepsilon_n, \omega_+)$ se tiene $(t, \varphi(t)) \notin K_n$. Ahora bien, $t \in (\omega_+ - \varepsilon_n, \omega_+) \subset [t_0, \omega_+]$ y por tanto $\varphi(t) \notin \overline{B}(0,n)$. Esto nos dice que $||\varphi(t)|| > n$ para todo $t \in (\omega_+ - \varepsilon_n, \omega_+)$, siendo precisamente la definición de que φ no está acotada.

□

En el corolario anterior, si $\omega_+ = b$ entonces existen dos posibilidades: o bien $||\varphi(t)|| \to \infty$ cuando $t \to \omega_+$, o bien $(t_n, \varphi(t_n)) \to (\omega_+, L_+) \in \partial\Omega$ para *alguna* sucesión $t_n \to \omega_+$. En este último caso, podría ser que $\lim_{t\to\omega_+} \varphi(t) = L_+$, o que $\lim_{t\to\omega_+} \varphi(t)$ no exista. Damos a continuación un ejemplo de cada posibilidad.

Ejemplo 4.1. 1. Consideremos el problema de Cauchy

$$\begin{cases} x' = x^2, \\ x(0) = 1. \end{cases}$$

La función $f(t,x) = x^2$ tiene como dominio de definición $\Omega = \mathbb{R}^2$ y es de clase C^1, en particular localmente lipschiztiana lo que implica existencia y unicidad para cada problema de Cauchy. La única solución maximal de este problema viene dada por la función

$$\varphi(t) = \frac{1}{1-t}, \qquad t \in (-\infty, 1).$$

Ahora bien, $\lim_{t\to 1^-} \varphi(t) = \infty$.

2. Consideremos el problema de Cauchy

$$\begin{cases} x' = \frac{1}{2x}, \\ x(1) = 1. \end{cases}$$

Esta vez $f(t,x) = 1/(2x)$, definida en $\Omega = \mathbb{R} \times (\mathbb{R} - \{0\})$, y en este dominio es C^1 lo que implica de nuevo existencia y unicidad local. En concreto, este problema de Cauchy tiene como única solución $\varphi(t) = \sqrt{t}$ definida en $(0,\infty)$. En este caso, $\lim_{t\to 0} \varphi(t) = 0$ y además $(0,0) \in \partial\Omega$. Esta solución puede extenderse por continuidad a $t = 0$, pero no de forma C^1 ya que su derivada tiende a infinito.

3. Consideremos ahora la función $x(t) = \sin\frac{1}{t}$ definida para $t \in (0,\infty)$, la cual es trivialmente solución de la EDO
$$x' = -\frac{1}{t^2}\cos\frac{1}{t},$$
pero $x(t)$ no tiene límite cuando $t \to 0^+$. En efecto, las sucesiones
$$t_n = \frac{1}{\frac{\pi}{2} + 2n\pi}, \qquad t_n' = \frac{1}{\frac{3\pi}{2} + 2n\pi},$$
cumplen $t_n, t_n' \to 0$ y $x(t_n) = 1,\ x(t_n') = -1$. En particular, el límite $\lim_{t\to 0} x(t)$ no existe. Observamos que la función $f(t,x) = -\frac{1}{t^2}\cos\frac{1}{t}$ que define la EDO no está acotada en ningún entorno de $t = 0$, pero esto no es condición suficiente para garantizar la no existencia del límite $\lim_{t\to\omega_+} \varphi(t)$ precisamente por el ejemplo anterior.

4. Ahora bien, una condición suficiente para garantizar que existe el límite $\lim_{t\to\omega_+} \varphi(t)$ es que $f(t,x)$ esté acotada en Ω y que ω_+ sea finito. En efecto, si M es una cota superior de $||f(t,x)||$ en Ω y $\{t_n\} \to \omega_+$, dados $m,n \in \mathbb{N}$ podemos hacer uso de la ecuación integral de Volterra, obteniendo
$$||\varphi(t_n) - \varphi(t_m)|| \le \int_{t_m}^{t_n} ||f(s,\varphi(s))||\, ds \le M|t_n - t_m| \to 0,$$
lo que significa que $\{x(t_n)\}$ es una sucesión de Cauchy y por tanto convergente a un cierto L_+. Si ahora t_n' es otra sucesión convergiendo a ω_+, acotamos igualmente usando la integral de Volterra y se tiene
$$||\varphi(t_n') - L_+|| \le ||\varphi(t_n') - \varphi(t_n)|| + ||\varphi(t_n) - L_+||.$$

El segundo término tiende a cero, mientras que el primero se acotaría de nuevo usando Volterra y concluimos $\varphi(t'_n) \to L_+$, es decir, $\lim_{t\to\omega_+} \varphi(t) = L_+$. Naturalmente, $(\omega_+, L_+) \in \partial\Omega$.

Nótese que no podemos pedir únicamente acotación de f en un entorno V de (ω_+, L_+) puesto que a la hora de acotar la desigualdad de Volterra necesitamos garantizar que $(s, \varphi(s)) \in V$ para todo $s \in [t_n, t_m]$ (véase el comentario tras la demostración del Lema de Wintner 4.3.11).

Finalmente, tal y como hicimos tras la demostración del Lema 4.3.6, damos una demostración alternativa del Teorema de Peano en una banda 4.2.9 usando las propiedades de las soluciones maximales.

Demostración del Teorema 4.2.9 de Peano en una banda. Sea $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ continua y acotada y $\varphi : (\omega_-, \omega_+) \to \mathbb{R}^n$ una solución maximal del problema de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0. \end{cases}$$

Supongamos por reducción al absurdo, y sin perder generalidad, $[t_0, \omega_+) \subset I$ de forma estricta, es decir $\omega_+ < \sup[t_0, \infty) \cap I$. En particular, $\omega_+ < \infty$ y sabemos que cuando $t \to \omega_+$, el comportamiento de φ está dado por el Teorema 4.3.13. Más aún, si $I \cap [t_0, \infty) = [t_0, b\rangle$, donde por $\rangle$ entendemos que el extremo b puede ser abierto ó cerrado, en particular $\omega_+ < b$ y el Corolario 4.3.14 implica que $\lim_{t\to\omega_+} ||\varphi(t)|| = \infty$; digamos $\varphi(t) \to \infty$. Por ser $\omega_+ < \infty$, el Teorema del valor medio asegura la existencia de una parcial $t_n \to \omega_+$ tal que $||\varphi'(t_n)|| \to \infty$. Ahora bien,

$$||\varphi'(t_n)|| = ||f(t_n, \varphi(t_n))|| \leq M,$$

donde M es una cota para $||f(t,x)||$ en todo $I \times \mathbb{R}^n$. Esto contradice que $||\varphi'(t_n)|| \to \infty$ y por tanto $||\varphi(t)||$ no puede *explotar* en tiempo finito. La única posibilidad es que $\omega_+ = \sup[t_0, \infty) \cap I$. $\square$

4.4. Continuidad y diferenciabilidad respecto de condiciones iniciales y parámetros

En Física de bachillerato e incluso en cursos avanzados, se aproxima el movimiento de un péndulo por el de un oscilador armónico. Más concretamente, si denotamos por $\theta(t)$ a la amplitud angular de un péndulo de longitud $L > 0$ sobre el que únicamente actúa la fuerza de la gravedad g, se tiene que θ es solución de la EDO de segundo orden

$$\theta'' + \frac{g}{L} \sin\theta = 0.$$

Por otra parte, la ecuación del oscilador armónico viene dada por $\theta'' + k\theta = 0$, donde $k > 0$. En particular, tomando $k = g/L$ y debido a que $\lim_{\theta\to 0} \frac{\sin\theta}{\theta} = 1$, las dos ecuaciones *se parecen* cerca del origen y es esperable que para pequeñas oscilaciones, esto es $\theta \approx 0$, las soluciones de ambas ecuaciones estén cerca.

El objetivo de esta sección es desarrollar una teoría alrededor de la siguiente cuestión: supongamos que $x(t), y(t)$ son soluciones de los problemas de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases} \qquad \begin{cases} y' = g(t,y), \\ y(\tilde{t_0}) = y_0. \end{cases}$$

Supongamos que las aplicaciones f, g están *próximas*, así como las condiciones iniciales (t_0, x_0), $(\tilde{t_0}, y_0)$. ¿Podemos asegurar que x, y están *próximas*? Y más importante aún, ¿qué significa *estar próximas*?

4.4.1. Dependencia continua respecto de condiciones iniciales

El resultado principal de esta sección es el siguiente.

Teorema 4.4.1. *Sea $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ continua y tal que cada problema de Cauchy asociado a $x' = f(t,x)$ tiene una única solución. Sea $(t_0, x_0) \in \Omega$ y supongamos que la solución maximal φ_0 del problema de Cauchy*

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

está definida en $[a,b]$, con $t_0 \in (a,b)$. Sea $(t_n, x_n) \in \Omega$ tal que $(t_n, x_n) \to (t_0, x_0)$. Entonces, existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$ el problema de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(t_n) = x_n, \end{cases}$$

tiene solución maximal φ_n definida en $[a,b]$ y se cumple

$$\lim_{n\to\infty} \varphi_n(t) = \varphi_0(t),$$

uniformemente en $[a,b]$.

Observación 4.5. Un par de observaciones.

1. La hipótesis de unicidad de soluciones para cada problema de Cauchy es clave. En efecto, para cada $n \in \mathbb{N}$ consideremos el problema de Cauchy

$$\begin{cases} x' = 3x^{2/3}, \\ x(0) = \frac{1}{n^3}, \end{cases}$$

   el cual tiene existencia y unicidad dada por la función $x_n(t) = (t+\frac{1}{n})^3$. Ahora bien, $(0, 1/n^3) \to (0,0)$ y en esta condición inicial tenemos como soluciones $x(t) = t^3$ y $x(t) = 0$ para todo t. Por una parte, $x_n(t)$ converge uniformemente a $x(t) = t^3$, pero no lo hace a $x(t) = 0$.

2. Es esencial que el intervalo de definición de x sea compacto. Lo comprobaremos con dos ejemplos: uno con un intervalo no acotado y otro con el intervalo acotado (y necesariamente abierto).

   - En el primero, consideremos el problema de Cauchy

$$x' = x, \qquad x(0) = \frac{1}{n},$$

     con solución única dada por $x_n(t) = \frac{1}{n}e^t$, y el problema límite con $x(0) = 0$, con solución única $x(t) = 0$ para todo t. Es sabido que la sucesión $x_n(t)$ converge uniformemente a $x(t) = 0$ sobre cada intervalo compacto $[a,b]$ conteniendo a 0. Sin embargo, esta convergencia deja de ser uniforme en $[-1, \infty)$.

   - En el segundo, consideremos el problema de Cauchy

$$x' = x^2, \qquad x(0) = 1 + \frac{1}{n},$$

con solución única dada por $x_n(t) = \frac{1+n}{n-t(1+n)}$, y el problema límite con $x(0) = 1$, con solución única $x(t) = \frac{1}{1-t}$. Por una parte, x_n está definida para $t \in (-\infty, \frac{n}{1+n})$. Para cada intervalo $[a, b] \subset (-\infty, 1)$ conteniendo a 0, existe n_0 suficientemente grande tal que $[a, b] \subset (-\infty, \frac{n}{1+n})$ para cada $n \geq n_0$ y aquí sería aplicable el teorema. No obstante, este resultado no se puede dar en cualquier intervalo con extremo superior 1 abierto, puesto que todas las soluciones x_n dejan de estar definidas y *explotan* antes de $t = 1$.

Antes de probar este resultado vamos a darle una interpretación *visual* y para ello consideremos el caso en que $x \in \mathbb{R}$. Sea φ_n la solución maximal del problema $x' = f(t, x)$ con $x(t_n) = x_n$, para $n \in \mathbb{N} \cup \{0\}$. La convergencia uniforme de φ_n a φ_0 en $[a, b]$ asegura que para todo $\epsilon > 0$ existe un n_0 a partir del cual $|\varphi_n(t) - \varphi_0(t)| < \epsilon$ para todo $t \in [a, b]$. Definimos el *entorno tubular de radio* ϵ *de la gráfica de* φ_0 como

$$U_\epsilon = \{(t, x) \in [a, b] \times \mathbb{R} \colon |x - \varphi_0(t)| < \epsilon\}.$$

La conclusión de este resultado es que no importa cómo de pequeño tomemos el radio de U_ϵ que siempre podremos encontrar un instante en nuestra sucesión de forma que todas las gráficas $(t, \varphi_n(t))$ distan de la gráfica de φ_0 una cantidad menor que ϵ.

Previo a la demostración, necesitaremos el siguiente lema sobre convergencia uniforme de una familia de funciones equicontinuas.

Lema 4.4.2. *Sea* $\{\varphi_n\}$, $\varphi_n : [a, b] \to \mathbb{R}^n$ *equicontinua y puntualmente acotada. Supongamos que toda subsucesión* $\{\varphi_{n_k}\}$ *que converge uniformemente lo hace a una misma función* $\varphi_\infty : [a, b] \to \mathbb{R}^n$. *Entonces,* $\{\varphi_n\}$ *converge uniformemente a* φ_∞ *en* $[a, b]$.

Demostración. Notemos que $\{\varphi_n\}$ se encuentra en las condiciones del Teorema de Ascoli-Arzela 4.2.6 y por tanto existe una subsucesión que converge uniformemente a cierta $\varphi_\infty : [a, b] \to \mathbb{R}^n$ continua. En particular, φ_∞ existe y está definida sin ambigüedad.

Supongamos que $\{\varphi_n\}$ no converge uniformemente a φ_∞. Entonces existe $\epsilon > 0$, una subsucesión $\{\varphi_{n_k}\}$ e instantes t_{n_k} tales que

$$||\varphi_{n_k}(t_k) - \varphi_\infty(t_k)|| > \epsilon.$$

Ahora bien, la propia subsucesión $\{\varphi_{n_k}\}$ es equicontinua y está puntualmente acotada por serlo la propia $\{\varphi_n\}$ y el Teorema de Ascoli-Arzela asegura la existencia de otra subsucesión $\{\varphi_{n_{kl}}\}$ que converge unifórmemente y necesariamente a φ_∞ por hipótesis. En particular, si n_{kl} es suficientemente grande se tendrá

$$||\varphi_{n_{kl}}(t) - \varphi_\infty(t)|| < \epsilon, \qquad \forall t \in [a, b],$$

en contradicción con $||\varphi_{n_k}(t_k) - \varphi_\infty(t_k)|| > \epsilon$. □

Estamos en disposición de probar el Teorema 4.4.1.

Demostración. Sea $\Gamma = \{(t, \varphi_0(t)) \colon t \in [a, b]\}$ la gráfica de φ_0. Como Γ es compacto y Ω abierto, existen compactos $K_0, K \subset \Omega$ con $\Gamma \subset K_0$ y $K_0 \subset \text{int}(K)$. Además, ya que $\text{dist}(K_0, \Omega \backslash K) > 0$, existen $\epsilon, r > 0$ tales que para cualquier $(t, x) \in K_0$ se tiene $[t - \epsilon, t + \epsilon] \times \overline{B}(x, r) \subset K$. Finalmente, fijamos $n_1 \in \mathbb{N}$ tal que si $n \geq n_1$ entonces $(t_n, x_n) \in K_0$ y dentamos por $M = \text{máx}_{(t,x) \in K} ||f(t, x)||$.

Definamos $\alpha = \text{mín}\{\epsilon, r/M\}$; esto nos debería resultar familiar con la elección de la amplitud del intervalo en el Teorema de Picard-Lindelöf 4.1.9 y en el Teorema de Peano 4.2.8. Pues bien, para cada $n \geq n_1$ y $n = \infty$ se tiene que φ_n está definida en $[t_n - \alpha, t_n + \alpha]$ y toma valores en $\overline{B}(x_n, r)$. Esto es precisamente consecuencia de aplicar el Teorema de Peano 4.2.8 (no podemos aplicar el Teorema de Picard-Lindelöf 4.1.9 por no tener la condición de lipschitzianidad local). Sea ahora $\epsilon = \alpha/2$. Redefiniendo n_1 si es necesario, podemos suponer que $|t_n - t_0| < \epsilon$ para cada $n \geq n_1$ y por tanto $\varphi_n(t)$ está definida en $[t_0 - \epsilon, t_0 + \epsilon] \subset [t_n - \alpha, t_n + \alpha]$ si $n \geq n_1$ o $n = 0$.

El siguiente paso es probar que φ_n converge unifórmemente a φ_0 en $[t_0 - \epsilon, t_0 + \epsilon]$ y para ello usaremos el Lema 4.4.2. Primero debemos probar que la familia $\{\varphi_n\}$ es una familia equicontinua y puntualmente acotada. Como cada φ_n está definida en $[t_0 - \epsilon, t_0 + \epsilon]$ y toma valores en $\overline{B}(x_n, r)$, para este ϵ que tenemos fijo podemos redefinir n_1 tal que para todo $n \geq n_1$ se tenga $\overline{B}(x_n, r) \subset \overline{B}(x_0, r + \epsilon)$. En consecuencia, para todo $n \geq n_1$ se tiene que φ_n toma valores en la bola fija $\overline{B}(x_0, r + \epsilon)$, lo cual asegura la acotación puntual de φ_n. Por otra parte, dados $t, t' \in [t_0 - \epsilon, t_0 + \epsilon]$ basta tomar $\delta < \epsilon/M$ para concluir

$$||\varphi_n(t) - \varphi_n(t')|| = || \int_{t'}^{t} f(s, \varphi_n(s))\, ds|| \leq M|t - t'| < \epsilon,$$

lo cual implica equicontinuidad y además uniforme.

Sea ahora $\{\varphi_{n_k}\}$ una subsucesión de $\{\varphi_n\}$ y supongamos que converge a uniformemente a cierta φ_∞ en $[t_0 - \epsilon, t_0 + \epsilon]$; existe al menos una subsucesión convergiendo a cierta φ_∞ por el Teorema de Ascoli-Arzela. Demostremos que φ_∞ coincide con φ_0. La integral de Volterra de cada φ_{n_k} implica

$$\varphi_{n_k}(t) = x_{n_k} + \int_{t_{n_k}}^{t} f(s, \varphi_{n_k}(s))\, ds = x_{n_k} + \int_{t_0}^{t} f(s, \varphi_{n_k}(s))\, ds + \int_{t_{n_k}}^{t_0} f(s, \varphi_{n_k}(s))\, ds.$$

Como $||f||$ está acotada en K por M se tiene

$$\left|\left|\int_{t_{n_k}}^{t_0} f(s, \varphi_{n_k}(s))\, ds\right|\right| \leq M|t_0 - t_{n_k}| \to 0, \quad k \to \infty,$$

y en consecuencia

$$\varphi_\infty(t) = x_0 + \int_{t_0}^{t} f(s, \varphi_\infty(s))\, ds,$$

donde se ha usado la convergencia uniforme de $\{\varphi_{n_k}\}$ a φ_∞ para permutar límite e integral. Esto nos dice que φ_∞ es solución del problema de Cauchy para $x' = f(t, x)$ con condición inicial $x(t_0) = x_0$, y la unicidad de cada de problema de Cauchy concluye $\varphi_\infty = \varphi_0$ en $[t_0 - \epsilon, t_0 + \epsilon]$. Por el Lema 4.4.2 se tiene de hecho que $\{\varphi_n\}$ converge uniformemente a φ_0.

Si $[a, b] \subset [t_0 - \epsilon, t_0 + \epsilon]$, ya hemos terminado, puesto que a su vez $[t_0 - \epsilon, t_0 + \epsilon] \subset [t_n - \alpha, t_n + \alpha]$ siendo éste último el intervalo de existencia de cada φ_n garantizado por el Teorema de Peano, el cual estará trivialmente contenido en el maximal; llamémoslo I_n. Supongamos pues sin perder generalidad que $t_0 + \epsilon > b$ y definamos la sucesión $\tilde{x_n} = \varphi_n(t_0 + \epsilon)$ que converge a $(t_0 + \epsilon, \tilde{x_0})$, siendo $\tilde{x_0} = \varphi_0(t_0 + \epsilon)$. Consideremos la solución maximal $\tilde{\varphi_n}$ del problema de Cauchy

$$\begin{cases} x' = f(t, x), \\ x(t_0 + \epsilon) = \tilde{x_n}. \end{cases}$$

Repitiendo el proceso anterior, podemos encontrar $n_2 > n_1$ tal que $\tilde{\varphi_n}$ está ahora definida en $[t_0, t_0 + 2\epsilon]$ y converge uniformemente a $\tilde{\varphi_0}$ en tal intervalo. Ahora bien, φ_n es igualmente solución de este problema de Cauchy y por el Lema 4.3.6 las soluciones φ_n y $\tilde{\varphi_n}$ pueden ser pegadas en $(t_0 + \epsilon, \tilde{x_n})$. Por unicida de cada problema de Cauchy, $\tilde{\varphi_n}$ es la restricción a $[t_0, t_0 + 2\epsilon]$ de la función φ_n definida en $[t_0 - \epsilon, t_0 + 2\epsilon]$. Además, por unicidad del límite uniforme, $\tilde{\varphi_\infty}$ es la restricción de φ_0 a $[t_0, t_0 + 2\epsilon]$. Repitiendo este proceso un número finito de pasos, tanto a la derecha como a la izquierda, existirá un cierto $n_0 \in \mathbb{N}$ tal que $[a, b] \subset [t_0 - \epsilon, t_0 + \epsilon]$ y para cada $n \geq n_0$ las soluciones maximales de φ_n cumplen $[a, b] \subset I_n$ y $\{\varphi_n\}$ converge uniformemente a φ_0 en $[a, b]$. $\square$

Para la siguiente definición necesitamos introducir algo de notación. Supongamos que cada problema de Cauchy asociado a $x' = f(t, x)$ tiene existencia y unicidad. Para cada $(t_0, x_0) \in \Omega$, sea $I(t_0, x_0)$ el intervalo maximal de definición de la solución maximal φ que cumple $\varphi(t_0) = x_0$. Denotaremos por $\varphi = \varphi_{(t_0,x_0)}$ para explicitar la dependencia y unicidad de cada solución respecto de las condiciones inciales.

Definición 4.4.3. Sea $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ continua tal que cada problema de Cauchy asociado a $x' = f(t, x)$ tiene existencia y unicidad. Sea

$$\mathcal{A} = \{(t, t_0, x_0) \colon (t_0, x_0) \in \Omega,\ t \in I(t_0, x_0)\}.$$

Se define el *flujo* de la EDO $x' = f(t, x)$ como la aplicación

$$\Phi : \mathcal{A} \to \mathbb{R}^n, \qquad \Phi(t, t_0, x_0) = \varphi_{(t_0,x_0)}(t). \tag{4.4.1}$$

La idea del flujo es bastante intuitiva. Para cada condición inicial $(t_0, x_0) \in \Omega$, el flujo nos dice cómo evoluciona la única solución que pasa por tal punto en cada instante de tiempo. El Teorema 4.4.1 nos permite deducir la siguiente consecuencia sobre el flujo.

Corolario 4.4.4. *El conjunto $\mathcal{A}$ es abierto y el flujo $\Phi : \mathcal{A} \to \mathbb{R}^n$ es una aplicación continua.*

Demostración. Sea $(\tau_n, t_n, x_n) \in \mathcal{A}$ una sucesión convergente a cierto $(\tau_0, t_0, x_0) \in \mathcal{A}$. Sea φ_n la solución maximal del problema de Cauchy asociado a $x' = f(t, x)$ con condición inicial $x(t_n) = x_n$, para $n \in \mathbb{N} \cup \{0\}$. Sea $[a, b]$ conteniendo a τ_0. Dado $\delta > 0$, el Teorema 4.4.1 asegura la existencia de $n_0 \in \mathbb{N}$ tal que φ_n está definida en $[\tau_0 - \delta, \tau_0 + \delta]$ y φ_n converge uniformemente a φ_0. En particular, $\tau_n \in [\tau_0 - \delta, \tau_0 + \delta]$. Como la convergencia es uniforme también lo es puntual y por tanto $\varphi_n(\tau_n) \to \varphi_0(\tau_0)$, pero esto no es más que escribir $\Phi(\tau_n, t_n, x_n) \to \Phi(\tau_0, t_0, x_0)$.

Que $\mathcal{A}$ es abierto se prueba de forma similar. Sea $(t, t_0, x_0) \in \mathcal{A}$ y tomemos $\delta > 0$ tal que $[t - \delta, t + \delta] \subset I(x_0, t_0)$. Necesitamos encontrar un entorno U de (t_0, x_0) en Ω tal que $[t - \delta, t + \delta] \subset I(t_1, x_1)$ para cualquier otro $(t_1, x_1) \in \Omega$. Pero esto precisamente es el contrarecíproco del Teorema 4.4.1. En efecto, supongamos que tal entorno no existe. Podemos encontrar pues una sucesión $(t_n, x_n) \to (t_0, x_0)$ de manera que $[t_0 - \delta, t_0 + \delta] \not\subset I(t_n, x_n)$. Pero es que a partir de cierto $n_0 \in \mathbb{N}$, cada solución maximal $\varphi_{(t_n,x_n)}$ debe cumplir $[t_0 - \delta, t_0 + \delta] \subset I(t_n, x_n)$, llegando a una contradicción. Por tanto tal entorno U de (t_0, x_0) debe existir y $[t - \delta, t + \delta] \times U \subset \mathcal{A}$ es el entorno de (t, t_0, x_0) buscado, probando que $\mathcal{A}$ es abierto. $\square$

4.4.2. Dependencia continua respecto de parámetros

Otro problema de parecida índole al de dependencia continua respecto a condiciones iniciales es considerar dependencia respecto de parámetros. Por ejemplo, en la ecuación del péndulo

$$x'' + \frac{g}{L} \sin x = 0,$$

tenemos dos parámetros, g, L. Podemos pensar que si dos parámetros L_1, L_2 son muy parecidos, también serán parecidas las oscilaciones de las soluciones asociadas. De forma más concreta, dadas condiciones iniciales $x(t_0) = x_0, x'(t_0) = v_0$, denotemos por $\varphi_{L_1}, \varphi_{L_2}$ a las soluciones maximales para los parámetros L_1, L_2, respectivamente. Lo deseable y esperable es que las soluciones sean continuas con respecto al parámetro longitud L; tanto ellas como sus derivadas. Para tal fin, es necesario expresar esta EDO de segundo orden como el sistema de primer orden equivalente

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} y \\ -\frac{g}{L}\sin x \end{pmatrix} = f(t, (x, y)).$$

Nos disponemos a dar una formulación general. Sea $\lambda = (\lambda_1, ..., \lambda_s) \in \Lambda \subset \mathbb{R}^s$ un vector de parámetros y $f(t, x, \lambda)$ una aplicación continua definida en un abierto conexo $\Omega \times \Lambda \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^s$ y con valores en $\mathbb{R}^n$. Supongamos además que cada problema de Cauchy asociado a la EDO $x' = f(t, x, \lambda)$ tiene existencia y unicidad para cualquier condición inicial en Ω. Fijado $(t_0, x_0) \in \Omega$, para cada $\lambda \in \Lambda$ denotemos por $\varphi_\lambda(t)$ a la solución maximal del problema de Cauchy

$$\begin{cases} x' = f(t, x, \lambda), \\ x(t_0) = x_0. \end{cases} \tag{4.4.2}$$

Teorema 4.4.5. *En las condiciones anteriores, sea $\{\lambda_n\} \subset \Lambda$ una sucesión convergente a $\lambda_0 \in \Lambda$ y supongamos que φ_{λ_0} está definida en $[a, b]$. Entonces, existe $n_0 \in \mathbb{N}$ tal que si $n \geq n_0$ la solución φ_{λ_n} está definida en $[a, b]$ y se cumple*

$$\lim_{n\to\infty} \varphi_{\lambda_n}(t) = \varphi_{\lambda_0}(t),$$

uniformemente en $[a, b]$.

Observación 4.6. Otras tantas observaciones antes de la demostración.

1. El enunciado del Teorema 4.4.5 es totalmente similar en cuanto a hipótesis y tesis al del Teorema 4.4.1. Probaremos que de hecho es una consecuencia.

2. De nuevo es esencial que el intervalo de definición se tome compacto. Consideremos el problema de Cauchy
$$x' = \lambda x, \qquad x(0) = 1,$$
cuya única solución es $x_\lambda(t) = e^{\lambda t}$. Tomemos $\lambda_n = 1/n$ y $\lambda_0 = 0$. Entonces, $x_{\lambda_n}(t) = e^{t/n}$ y $x_0 = 1$. Es cierto que sobre cada $[a, b]$ se tiene $x_{\lambda_n}(t) \to 0$ uniformemente, pero no lo es en todo $\mathbb{R}$. En efecto, si tomamos la sucesión $t_n = n$ se tiene $x_{\lambda_n}(n) = e \not\to 1$.

3. La teoría de integrales dependientes de parámetros se engloba de cierta forma en el Teorema 4.4.5. Supongamos que $h : \mathbb{R} \times \Lambda \to \mathbb{R}$ es continua y definamos
$$F : \Lambda \to \mathbb{R}, \qquad F(\lambda) = \int_a^b h(t, \lambda)\, dt.$$
El Teorema 4.4.5 garantiza que F es continua. Para probar este hecho, definamos la aplicación
$$f : \mathbb{R} \times \mathbb{R} \times \Lambda \to \mathbb{R}, \qquad f(t, x, \lambda) = h(t, \lambda).$$

El problema de Cauchy

$$\begin{cases} x' = f(t, x, \lambda), \\ x(a) = 0, \end{cases}$$

tiene como única solución

$$x_\lambda(t) = \int_a^t h(s, \lambda)\, ds.$$

Entonces, $F(\lambda) = x_\lambda(b)$ y por tanto F es continua en virtud del Teorema 4.4.5.

Ahora sí, demostremos el Teorema 4.4.5.

Demostración. La idea consiste en transformar el problema de dependencia de parámetros en otro de dependencia de condiciones iniciales, haciendo que el parámetro λ pase a ser una condición inicial.

Dado el espacio de parámetros $\Lambda \subset \mathbb{R}^s$ y $f(t, x, \lambda)$ definida en $\Omega \times \Lambda \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^s$, definimos $\mathcal{O} = \Omega \times \Lambda$ e identificamos $(t, x, \lambda) \in \mathcal{O}$ como (t, y), donde $y = (x, \lambda)$. Definimos $F : \mathcal{O} \to \mathbb{R}^{n+s}$ dada por $F(t, y) = (f(t, x, \lambda), \mathbf{0}_s)$, siendo $\mathbf{0}_s$ el vector nulo en $\mathbb{R}^s$. Fijado λ, que $y(t) = (x(t), \lambda(t))$ sea solución de

$$\begin{cases} y' = F(t, y), \\ y(t_0) = (x_0, \lambda), \end{cases} \tag{4.4.3}$$

es equivalente a que se cumpla

$$\begin{cases} (x, \lambda)' = (f(t, x, \lambda), \mathbf{0}_s) \\ (x, \lambda)(t_0) = (x_0, \lambda). \end{cases} \tag{4.4.4}$$

Por una parte, $\lambda'(t) = \mathbf{0}_s$ asegura que $\lambda(t)$ es constante y como $\lambda(t_0) = \lambda$ se tiene $\lambda(t) = \lambda$ para todo t. Por otra parte, estamos pidiendo que x sea solución del problema de Cauchy (4.4.2). Por tanto, $y(t) = (x(t), \lambda)$ es solución de (4.4.3) si y sólo si x es solución de (4.4.2). En tal caso y en términos del flujo para la EDO (4.4.3), si φ_λ es la solución maximal de (4.4.2) y $\Phi(t, t_0, (x_0, \lambda))$ es la solución maximal de (4.4.3), ambas están relacionadas por

$$\Phi(t, t_0, (x_0, \lambda)) = (\varphi_\lambda(t), \lambda).$$

En otras palabras, $(\varphi_\lambda(t), \lambda)$ es la solución maximal de (4.4.4). En particular, los intervalos maximales de φ y $\Phi(t, t_0, (x_0, \lambda))$ necesariamente coinciden. Finalmente, dada $\lambda_n \to \lambda_0$, consideremos la sucesión $\{(t_0, x_0, \lambda_n)\} \subset \mathcal{O}$ y sea $\varphi_{\lambda_n}(t)$ la solución maximal para el problema de Cauchy $x' = f(t, x, \lambda_n)$ con $x(t_0) = x_0$. Por otra parte, sea $\Phi(t, t_0, (x_0, \lambda_n))$ la solución del problema de Cauchy (4.4.3) con condición inicial $x(t_0) = (x_0, \lambda_n)$. Para ésta última, el Teorema 4.4.1 asegura

$$\lim_{n \to \infty} (\varphi_{\lambda_n}(t), \lambda_n) = \lim_{n \to \infty} \Phi(t, t_0, (x_0, \lambda_n)) = \Phi(t, t_0, (x_0, \lambda_0)) = (\varphi_{\lambda_0}(t), \lambda),$$

uniformemente en cada intervalo compacto $[a, b]$ conteniendo a t_0, lo que implica la convergencia uniforme en $[a, b]$ de $\varphi_{\lambda_n}(t)$ a $\varphi_{\lambda_0}(t)$. □

Observemos que en la demostración del Teorema 4.4.5 podríamos haber tomado una sucesión $(t_n, x_n, \lambda_n) \to (t_0, x_0, \lambda_0)$ y concluir igualmente la convergencia uniforme en $[a, b]$ de las soluciones $\varphi_{(t_n, x_n, \lambda_n)}(t)$ a $\varphi_{(t_0, x_0, \lambda_0)}(t)$, sin más que considerar los problemas de Cauchy equivalentes

$$\begin{cases} x' = f(t, x, \lambda_n), \\ x(t_n) = x_n, \end{cases} \qquad \begin{cases} (x, \lambda)' = (f(t, x, \lambda_n), \mathbf{0}_s), \\ (x, \lambda)(t_n) = (x_n, \lambda_n). \end{cases}$$

Esta misma idea de permutar condiciones iniciales y parámetros nos permite concluir el siguiente resultado.

Corolario 4.4.6. *En las condiciones anteriores, para cada* $(t_0, x_0, \lambda) \in \Omega \times \Lambda$*, sea* $\varphi_{(t_0,x_0,\lambda)}$ *la solución maximal del problema de Cauchy*

$$\begin{cases} x' = f(t, x, \lambda), \\ x(t_0) = x_0, \end{cases}$$

definida en su intervalo maximal denotado por $I(t_0, x_0, \lambda)$*. Sea*

$$\mathcal{B} = \{(t, t_0, x_0, \lambda) \colon (t_0, x_0, \lambda) \in \Omega \times \Lambda,\ t \in I(t_0, x_0, \lambda)\},$$

y el flujo

$$\Psi : \mathcal{B} \to \mathbb{R}^n, \qquad \Psi(t, t_0, x_0, \lambda) = \varphi_{(t_0,x_0,\lambda)}(t). \tag{4.4.5}$$

Entonces, $\mathcal{B}$ *es abierto y* Ψ *es continua. Más aún,* Ψ *está relacionado con el flujo* Φ *definido en* (4.4.1) *por*

$$\Phi(t, t_0, (x_0, \lambda)) = (\Psi(t, t_0, x_0, \lambda), \lambda). \tag{4.4.6}$$

Demostración. Que Ψ es continua se deduce de forma secuencial usando los Teoremas 4.4.1 y 4.4.5. Que $\mathcal{B}$ es abierto se deduce igual que en el Corolario 4.4.4, encontrando un entorno de (t_0, x_0, λ_0) que contenga a $[t - \delta, t + \delta]$ para $\delta > 0$, usando la convergencia uniforme de $\varphi_{(t_n,x_n,\lambda_n)}$ a $\varphi_{(t_0,x_0,\lambda_0)}$ en $[t - \delta, t + \delta]$. □

Observemos que el Teorema 4.4.5 responde parcialmente a una de las cuestiones planteadas en la introducción de este capítulo: si $f(t, x)$ y $g(t, x)$ son aplicaciones que están próximas, ¿cómo de próximas están sus soluciones? En efecto, fijamos $(t_0, x_0) \in \Omega$ y tomamos $\{\lambda_n\} \subset \Lambda$ una sucesión tal que $\lambda_n \to \lambda_0 \in \Lambda$. Dada $f(t, x, \lambda)$ continua tal que cada problema de Cauchy tiene existencia y unicidad, para cada $n \in \mathbb{N} \cup \{0\}$ definimos $f_n(t, x) = f(t, x, \lambda_n)$ y denotamos por φ_n a la solución maximal del problema de Cauchy asociado a $x' = f_n(t, x)$ con $x(t_0) = x_0$. Entonces, $\varphi_n(t) \to \varphi_0(t)$ uniformemente en cada compacto $[a, b]$ conteniendo a t_0. El siguiente resultado generaliza esta idea para una sucesión de aplicaciones $\{f_n\}$ que convergen a una cierta aplicación límite f_0. Con el mismo esfuerzo, haremos variar las condiciones iniciales.

Corolario 4.4.7. *Dado* $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ *abierto. Para cada* $n \in \mathbb{N} \cup \{0\}$ *consideramos* $f_n : \Omega \to \mathbb{R}^n$ *continua y supongamos que* $\{f_n\}$ *converge uniformemente a* f_0 *sobre compactos y que cada problema de Cauchy asociado a* $x' = f_n(t, x)$ *tiene existencia y unicidad. Sean* $(t_n, x_n) \in \Omega$ *tal que para* $n \geq 1$ *se tiene* $(t_n, x_n) \to (t_0, x_0)$*. Supongamos que la solución maximal* φ_0 *del problema de Cauchy*

$$\begin{cases} x' = f_0(t, x), \\ x(t_0) = x_0, \end{cases}$$

está definida en $[a, b]$ *con* $t_0 \in (a, b)$*. Entonces, existe* $n_0 \in \mathbb{N}$ *tal que si* $n \geq n_0$ *la solución maximal* φ_n *del problema de Cauchy*

$$\begin{cases} x' = f_n(t, x), \\ x(t_n) = x_n, \end{cases}$$

está definida en $[a, b]$ *y se cumple*

$$\lim_{n \to \infty} \varphi_n(t) = \varphi_0(t).$$

Obviamente, si $f_n(t, x) = f(t, x)$ para todo $n \in \mathbb{N} \cup \{0\}$, entonces el resultado es simplemente el Teorema 4.4.1.

Demostración. La demostración es totalmente análoga a la del Teorema 4.4.1 y sólo detallamos las diferencias más notables.

Primero, como f_n converge uniformemente a f_0 sobre compactos, se puede tomar $M > 0$ una cota común para todas las normas $||f(t,x)||$ en K. Esta cota común nos permite probar de igual forma la equicontinuidad uniforme de la sucesión φ_n en $[t_0 - \epsilon, t_0 + \epsilon]$.

Ahora debemos probar que si una subsucesión φ_{n_k} converge uniformemente a cierta φ_∞, entonces necesariamente $\varphi_\infty = \varphi_0$. Para ello, expresamos

$$\varphi_{n_k}(t) = x_{n_k} + \int_{t_{n_k}}^{t} f_{n_k}(s, \varphi_{n_k})\, ds = x_{n_k} + \int_{t_0}^{t} f_{n_k}(s, \varphi_{n_k}(s))\, ds + \int_{t_{n_k}}^{t_0} f_{n_k}(s, \varphi_{n_k}(s))\, ds,$$

y usamos la acotación uniforme de $||f_{n_k}|| \leq M$ y la convergencia uniforme de f_{n_k} a f_0 para permutar integral y límite para concluir

$$\varphi_\infty(t) = x_0 + \int_{t_0}^{t} f_0(s, \varphi_\infty(s))\, ds,$$

y por unicidad $\varphi_\infty = \varphi_0$. Finalmente, repetimos el argumento de prolongación del intervalo $[t_0 - \epsilon, t_0 + \epsilon]$ en caso de que $[a, b]$ no esté contenido en tal intervalo para concluir que a partir de cierto n_0 todas las soluciones maximales φ_n contienen a $[a, b]$ en su intervalo maximal de definición y φ_n converge uniformemente a φ_0 en $[a, b]$. $\square$

Aproximación de la solución del péndulo simple

Consideremos la ecuación del péndulo simple

$$\theta'' + \frac{g}{L} \sin\theta = 0, \qquad \theta(0) = \epsilon > 0, \quad \theta'(0) = 0.$$

Es lógico pensar que para valores de ϵ suficientemente pequeños podamos usar la aproximación $\sin\theta \approx \theta$ y exista una similitud con las soluciones del oscilador armónico

$$\phi'' + \frac{g}{L}\phi = 0, \qquad \phi(0) = \epsilon, \quad \phi'(0) = 0.$$

Para comenzar, consideramos el cambio de variable $\theta = \epsilon x$ y pasamos a un sistema de primer orden con coordenadas (x, y) y el cambio usual $x' = y$, obteniendo

$$\begin{pmatrix} x \\ y \end{pmatrix}' = \begin{pmatrix} y \\ -\frac{g}{L}\frac{\sin(\epsilon x)}{\epsilon x} x \end{pmatrix} := f_\epsilon(x, y), \qquad x(0) = 1, \quad y(0) = 0.$$

Observemos que f_ϵ es una función continua y de hecho es localmente lipschitziana alrededor de cada punto $(0, y)$; fuera de estos puntos es trivial por ser C^1. Por tanto, para cada $\epsilon > 0$ existe una única solución $\gamma_\epsilon = (x_\epsilon, y_\epsilon)$ del sistema $\gamma_\epsilon' = f_\epsilon(\gamma_\epsilon)$ con condición incial $\gamma_\epsilon(0) = (1, 0)$. Ahora bien,

$$\lim_{\epsilon \to 0} f_\epsilon(x, y) = (y, -\frac{g}{L}x),$$

que se corresponde con el sistema lineal obtenido al reducir el orden de la ecuación del oscilador armónico. Si denotamos por γ_0 a la solución del sistema $\gamma' = f_0(\gamma)$ con condición inicial $\gamma(0) = (1, 0)$, se tiene

$$\gamma_0(t) = \left(\cos\sqrt{\frac{g}{L}}t, -\sqrt{\frac{g}{L}} \sin\sqrt{\frac{g}{L}}t \right).$$

En estas condiciones, el Teorema 4.4.5 de dependencia continua respecto de parámetros asegura que $\lim_{\epsilon\to 0}\gamma_\epsilon = \gamma_0$ uniformemente sobre cada compacto conteniendo a $t=0$. En otras palabras,

$$\lim_{\epsilon\to 0} x_\epsilon(t) = \cos\sqrt{\frac{g}{L}}t, \qquad \lim_{\epsilon\to 0} x'_\epsilon(t) = -\sqrt{\frac{g}{L}}\sin\sqrt{\frac{g}{L}}t.$$

Deshaciendo el cambio $\theta = \epsilon x$ y $\theta' = \epsilon y$ concluimos finalmente

$$\theta_\epsilon(t) = \epsilon\cos\sqrt{\frac{g}{L}}t + o(\epsilon), \qquad \theta'_\epsilon(t) = -\epsilon\sqrt{\frac{g}{L}}\sin\sqrt{\frac{g}{L}}t + o(\epsilon),$$

los cuales son uniformes en $t\in[a,b]$ cuando $\epsilon\to 0$.

Observación 4.7. Hemos usado la notación *o pequeña*, que se define como sigue. Dadas dos funciones $f(t), g(t)$, se dice que $f(t) = o(g(t))$ cuando $t\to 0$ si

$$\lim_{t\to 0}\frac{f(t)}{g(t)} = 0.$$

Escribir pues $o(\epsilon)$ denota una cierta función que depende de ϵ y tal que $\lim_{\epsilon\to 0}\frac{o(\epsilon)}{\epsilon} = 0$.

4.4.3. Dependencia diferenciable respecto a condiciones iniciales

El objetivo de esta sección es estudiar la siguiente cuestión. Sea $\Omega\subset\mathbb{R}\times\mathbb{R}^n$ abierto y $f:\Omega\to\mathbb{R}^n$ continua y tal que cada problema de Cauchy asociado a $x' = f(t,x)$ tiene existencia y unicidad. Supongamos además que la aplicación f tiene *buenas* propiedades de diferenciabilidad. Si x, y son soluciones de esta EDO y satisfacen condiciones iniciales *cercanas*, sabemos por el Teorema 4.4.1 que ambas soluciones deben estar inicialmente cercanas. ¿Podemos afirmar algo similar sobre las derivadas de x e y? ¿Cuáles son esas buenas propiedades de diferenciabilidad necesarias como hipótesis sobre f?

La mayor dificultad que nos encontraremos no es tanto la complejidad de los argumentos que involucran las demostraciones, sino la notación usada que puede ser algo enreversada, sobre todo si se quiere escribir con suficiente formalidad. Pedimos paciencia al alumno, sobre todo a las alturas del curso en las que se ven estos resultados, y lo animamos a no caer en la desesperación; no es para tanto.

Comenzaremos estudiando el problema de Cauchy

$$\begin{cases} x' = f(t,x), \\ x(t_0) = x_0, \end{cases}$$

donde las condiciones iniciales (t_0, x_0) son susceptibles de variar. El resultado principal de esta sección es el siguiente.

Teorema 4.4.8. *Sea $f:\Omega\subset\mathbb{R}\times\mathbb{R}^n\to\mathbb{R}^n$ una aplicación continua tal que la matriz jacobiana $\frac{\partial f}{\partial x}(t,x)$ existe y la aplicación $(t,x)\in\Omega\mapsto\frac{\partial f}{\partial x}(t,x)$ es continua. Entonces, el flujo $\Phi(t,t_0,x_0)$ definido en (4.4.1) es una aplicación diferenciable. Más aún, las derivadas parciales vienen dadas* *por*

$$\begin{aligned}\frac{\partial \Phi}{\partial t}(t,t_0,x_0) &= \frac{d}{dt}\Phi(t,t_0,x_0),\\ \frac{\partial \Phi}{\partial t_0}(t,t_0,x_0) &= y(t),\\ \frac{\partial \Phi}{\partial x_0}(t,t_0,x_0) &= Y(t),\end{aligned}$$

donde $y(t)$ *es la única solución de*

$$y' = \frac{\partial f}{\partial x}(t,\Phi(t,t_0,x_0))y, \qquad y(t_0) = -f(t_0,x_0),$$

e $Y(t)$ *es la única solución (matriz) de*

$$Y' = \frac{\partial f}{\partial x}(t,\Phi(t,t_0,x_0))Y, \qquad Y(t_0) = I_n.$$

Observación 4.8. Antes de probar este resultado, unos comentarios.

1. Las derivadas parciales $\frac{\partial \Phi}{\partial t}$ y $\frac{\partial \Phi}{\partial t_0}$ son vectores en $\mathbb{R}^n$, mientras que $\frac{\partial \Phi}{\partial x}$ es una matriz cuadrada de orden n. Observemos que y e Y son soluciones de la misma EDO, pero y es vectorial e Y es matricial: de hecho es la matriz fundamental principal en t_0. En ambos casos, $\frac{\partial f}{\partial x}$ es la matriz jacobiana de la aplicación $f(t,x)$ respecto de las variables $x = (x_1, ..., x_n)$, calculada en el punto (t,x).

2. Para entender de dónde provienen estas expresiones para las derivadas parciales de Φ, vamos a suponer que todo es suficientemente regular y derivar en la igualdad

$$\frac{d}{dt}\Phi(t,t_0,x_0) = f(t,\Phi(t,t_0,x_0)). \tag{4.4.7}$$

Derivando respecto de t_0 y cruzando las derivadas se tiene

$$\frac{d}{dt}\left(\frac{\partial \Phi}{\partial t_0}(t,t_0,x_0)\right) = \frac{\partial f}{\partial x}(t,\Phi(t,t_0,x_0))\frac{\partial \Phi}{\partial t_0}(t,t_0,x_0).$$

Denotando por $y(t) = \frac{\partial \Phi}{\partial t_0}(t,t_0,x_0)$, se tiene que y es solución de la EDO lineal

$$y' = \frac{\partial f}{\partial x}(t,\Phi(t,t_0,x_0))y.$$

Para determinar la condición inicial, consideramos el flujo $x_0 = \Phi(t,t,x_0)$, derivamos respecto de t y sustituimos en (t_0,t_0,x_0), obteniendo

$$\frac{d}{dt}\Phi(t_0,t_0,x_0) + \frac{\partial \Phi}{\partial t_0}(t_0,t_0,x_0) = 0,$$

o de forma equivalente

$$y(t_0) = -f(t_0,x_0).$$

Para la última identidad, derivamos (4.4.7) respecto de x_0 y permutamos derivadas cruzadas, obteniendo
$$\frac{d}{dt}\left(\frac{\partial \Phi}{\partial x_0}(t,t_0,x_0)\right) = \frac{\partial f}{\partial x}(t,\Phi(t,t_0,x_0))\frac{\partial \Phi}{\partial x_0}(t,t_0,x_0).$$
Ahora la diferencia es que $Y(t) = \frac{\partial \Phi}{\partial x_0}(t,t_0,x_0)$ es una matriz cuadrada de orden n y no un vector. Para deducir la condición inicial derivamos $\Phi(t_0,t_0,x_0) = x_0$ respecto de x_0, concluyendo
$$Y(t_0) = \frac{\partial \Phi}{\partial x_0}(t_0,t_0,x_0) = \frac{\partial x_0}{\partial x_0} = I_n.$$

3. La ecuación lineal
$$y' = \frac{\partial f}{\partial x}(t,\Phi(t,t_0,x_0))y, \qquad y(t_0) = -f(t_0,x_0),$$
se suele llamar ecuación variacional alrededor de la solución $\Phi(t,t_0,x_0)$, la cual es necesaria conocer de forma explícita para resolver la ecuación variacional.

Para demostrar el Teorema 4.4.8 necesitamos un lema previo, el cual enunciamos y probamos a continuación.

Lema 4.4.9. *Consideremos la EDO $x' = f(t,x)$ con condición inicial $x(t_0) = x_0$ y sea Φ el flujo definido en* (4.4.1). *Dado $v \in \mathbb{R}^n$ definimos el cociente incremental*
$$\delta_h(t) = \frac{\Phi(t,t_0,x_0+hv) - \Phi(t,t_0,x_0)}{h}, \tag{4.4.8}$$
el cual está bien definido si $t \in [a,b]$ y $h \in [-\epsilon,\epsilon] - \{0\}$ para $\epsilon > 0$ suficientemente pequeño. Entonces, existe una aplicación continua
$$A : [a,b] \times [-\epsilon,\epsilon] \to \mathcal{M}_n(\mathbb{R}),$$
tal que para todo $t \in [a,b]$ y $h \neq 0$ se tiene
$$\delta_h'(t) = A(t,h)\delta_h(t), \qquad A(t,0) = \frac{\partial f}{\partial x}(t,\Phi(t,t_0,x_0)).$$

Demostración. La demostración del lema utiliza un *sucedáneo* del Teorema del Valor Medio para funciones vectoriales, el cual enunciamos a continuación: *dada $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ de clase C^1 y dos puntos $x,y \in \Omega$ tales que el segmento $[x,y] \subset \Omega$, se cumple*
$$f(x) - f(y) = \left(\int_0^1 Jf(\lambda x + (1-\lambda)y)\,d\lambda\right)(x-y),$$
donde $Jf(\lambda x + (1-\lambda)y)$ es la matriz jacobiana de f evaluada en tal punto y la integral se entiende coordenada a coordenada. La demostración de esta propiedad es inmediata.
$$\begin{aligned} f(x) - f(y) &= \int_0^1 \frac{d}{d\lambda} f(\lambda x + (1-\lambda)y)\,d\lambda = \int_0^1 Jf(\lambda x + (1-\lambda)y)(x-y)\,d\lambda \\ &= \left(\int_0^1 Jf(\lambda x + (1-\lambda)y)\,d\lambda\right)(x-y), \end{aligned}$$

donde en la primera igualdad se ha usado la regla de Barrow para funciones de una variable, en la segunda la regla de la cadena y en la tercera la linealidad de la integral y del producto de matrices con vectores.

Ahora podemos demostrar el lema aplicando el resultado anterior a $f(t, x)$ en su segunda variable.

$$\begin{aligned}\delta_h'(t) &= \frac{1}{h}\left(\frac{d}{dt}\Phi(t,t_0,x_0+hv) - \frac{d}{dt}\Phi(t,t_0,x_0)\right) = \frac{1}{h}\left(f(t,\Phi(t,t_0,x_0+hv)) - f(t,\Phi(t,t_0,x_0))\right)\\ &= \underbrace{\frac{1}{h}\left(\int_0^1 \frac{\partial f}{\partial x}(t,\lambda\Phi(t,t_0,x_0+hv)+(1-\lambda)\Phi(t,t_0,x_0))\,d\lambda\right)}_{A(t,h)}(\Phi(t,t_0,x_0+hv)-\Phi(t,t_0,x_0))\\ &= A(t,h)\delta_h(t).\end{aligned}$$

Por último, que $A(t, 0) = \frac{\partial f}{\partial x}(t, \Phi(t, t_0, x_0))$ se prueba, por ejemplo, usando la dependencia continua respecto a condiciones iniciales y el Teorema de la convergencia dominada. Para probar que dado la aplicación $(t, h) \mapsto A(t, h)$, con $h \neq 0$, es continua, se usa de nuevo la dependencia continua respecto a condiciones iniciales y la teoría de integrales dependientes de parámetros. □

Ahora sí, estamos en condiciones de probar el Teorema 4.4.8.

Demostración. Observemos que se tiene

$$\frac{\partial \Phi}{\partial t}(t, t_0, x_0) = f(t, \Phi(t, t_0, x_0)),$$

la cual es continua por la continuidad de f y Φ.

Dado $(t_0, x_0) \in \Omega$, sea $[a, b]$ tal que $[a, b] \subset I(t_0, x_0)$, o de forma equivalente $(t, t_0, x_0) \in \mathcal{A}$ para todo $t \in [a, b]$. Vamos a probar que las derivadas direccionales de Φ con respecto a x_0 existen y son continuas. La existencia y continuidad de la derivada parcial respecto a t_0 es similar.

Sabemos por el Lema 4.4.9 que la aplicación δ_h dada por (4.4.8) está bien definida para t, h adecuados y cumple la ecuación lineal

$$\delta_h'(t) = A(t, h)\delta_h(t), \qquad \delta_h(t_0) = v,$$

la cual tiene existencia y unicidad. Sea $\delta_h(t, v, h)$ la solución de este problema dependiendo de las condiciones iniciales (t_0, v) y del parámetro h. Podemos aplicar los teoremas de dependencia continua respecto a condiciones iniciales y parámetros en el dominio $(a, b) \times \mathbb{R}^n \times (-\epsilon, \epsilon)$. En particular, $\delta_h(t)$ converge para $h \to 0$ a la solución del problema límite

$$y' = A(t, 0)y, \qquad y(t_0) = v,$$

que por el Lema 4.4.9 es exactamente

$$y' = \frac{\partial f}{\partial x}(t, \Phi(t, t_0, x_0))y, \qquad y(t_0) = v. \tag{4.4.9}$$

Como por otra parte para cada $v \in \mathbb{R}^n$ se define la derivada direccional de una aplicación $F : U \subset \mathbb{R}^n \to \mathbb{R}^m$ mediante

$$D_v F(x) = \lim_{h \to 0} \frac{F(x + hv) - F(x)}{h},$$

garantizamos la existencia de la derivada direccional de $\Phi(t, t_0, x_0)$ respecto de la variable x_0 en cualquier dirección $v \in \mathbb{R}^n$. Es más, si $Y(t)$ es la matriz fundamental principal del sistema (4.4.9) en t_0, entonces ya que la solución de (4.4.9) viene dada por $y(t) = Y(t)v$ concluimos

$$D_v\Phi(t, t_0, x_0) = Y(t)v.$$

Realmente, $Y(t)$ depende también de las condiciones iniciales (t_0, x_0), así que denotaremos por $Y = Y(t, t_0, x_0)$ a la solución (matricial) del sistema lineal

$$Y' = \frac{\partial f}{\partial x}(t, \Phi(t, t_0, x_0))Y, \qquad Y(t_0) = I_n, \tag{4.4.10}$$

donde x_0 es parámetro y t_0 es a la vez condición inicial y parámetro. Ahora sí podemos escribir formalmente

$$D_v\Phi(t, t_0, x_0) = Y(t, t_0, x_0)v. \tag{4.4.11}$$

A continuación, probaremos que estas derivadas direccionales son también funciones continuas de (t, t_0, x_0); o en otras palabras, que la aplicación $(t, t_0, x_0) \mapsto Y(t, t_0, x_0)$ es continua. Por hipótesis, la aplicación $\frac{\partial f}{\partial x}$ es continua y el Corolario 4.4.7 que aúna dependencia continua de condiciones iniciales y parámetros nos permite concluir que la aplicación $(t, t_0, x_0) \mapsto \frac{\partial f}{\partial x}(t, \Phi(t, t_0, x_0))$ es continua. Ya que en la ecuación matricial

$$Y' = \frac{\partial f}{\partial x}(t, \Phi(t, t_0, x_0))Y, \qquad Y(t_0) = I_n,$$

tenemos como condición inicial t_0 y como parámetros (t_0, x_0), de nuevo el Corolario 4.4.7 nos permite concluir que la aplicación $(t, t_0, x_0) \mapsto Y(t, t_0, x_0)$ es continua. En conclusión, la aplicación Φ admite derivadas direccionales continuas en cada dirección $v \in \mathbb{R}^n$, y tal derivada direccional viene dada por (4.4.11). En consecuencia,

$$\frac{\partial \Phi}{\partial x_0}(t, t_0, x_0) = Y(t, t_0, x_0).$$

En cuanto a la derivada parcial respecto de t_0, se define de forma análoga

$$\delta_h(t) = \frac{1}{h}\left(\Phi(t, t_0 + h, x_0) - \Phi(t, t_0, x_0)\right),$$

la cual es solución del problema lineal

$$\delta_h'(t) = A(t, h)\delta_h(t), \qquad \delta_h(t_0 + h) = \frac{x_0 - \Phi(t_0 + h, t_0, x_0)}{h}.$$

Además, $A(t, 0) = \frac{\partial f}{\partial x}(t, \Phi(t, t_0, x_0))$. Por una parte, $\lim_{h\to 0} \delta_h(t) = \frac{\partial \Phi}{\partial t_0}(t, t_0, x_0)$, y por otra parte $\delta_h(t)$ converge para $h \to 0$ a la solución del problema límite

$$y' = \frac{\partial f}{\partial x}(t, \Phi(t, t_0, x_0))y.$$

Finalmente, para determinar la condición inicial $y(t_0)$, bastará hacer $h \to 0$ en la condición inicial $\delta_h(t_0 + h)$ y aplicar la regla de L'Hôpital, obteniendo

$$\begin{aligned}\lim_{h\to 0} \delta_h(t_0) &= \lim_{h\to 0} \frac{x_0 - \Phi(t_0 + h, t_0, x_0)}{h} \\ &= -\frac{d}{dt}\Phi(t_0, t_0, x_0) = -f(t_0, x_0).\end{aligned}$$

De forma matricial, si $Y(t,t_0,x_0)$ es la solución de (4.4.10), entonces

$$\frac{\partial \Phi}{\partial t_0}(t,t_0,x_0) = -Y(t,t_0,x_0)f(t_0,x_0).$$

$\square$

Ejemplo 4.2. Veamos un ejemplo concreto analizando la ecuación del péndulo

$$\theta'' + \frac{g}{L}\sin\theta = 0, \qquad \theta(t_0) = \theta_0, \quad \theta'(t_0) = \omega_0.$$

Sea $\theta(t,t_0,(\theta_0,\omega_0))$ la única solución de este problema de Cauchy dependiendo de los datos iniciales $t_0,(x_0,\theta_0)$. Por ejemlpo, si $\theta_0 = \omega_0 = 0$ entonces $\theta(t,t_0,(0,0)) = 0$. Haciendo $\theta_1 = \theta$ y definiendo $\theta_2 = \theta_1'$, llegamos al sistema equivalente de primer orden

$$\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}' = \begin{pmatrix}\theta_2\\ -\frac{g}{L}\sin\theta_1\end{pmatrix} = f(t,(\theta_1,\theta_2)), \qquad \begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}(t_0) = \begin{pmatrix}\theta_0\\ \omega_0\end{pmatrix}.$$

Consideramos el flujo $\Phi(t,t_0,(\theta_0,\omega_0))$ dado por (4.4.1) y usando la definición de θ_1 y θ_2 se tiene

$$\Phi(t,t_0,(\theta_0,\omega_0)) = (\theta(t,t_0,(\theta_0,\omega_0),\theta'(t,t_0,(\theta_0,\omega_0)).$$

Vamos a trabajar alrededor de la condición inicial $(\theta_0,\omega_0) = (0,0)$, para la cual conocemos su solución explícita $\Phi(t,t,(0,0)) = (0,0)$. Ahora debemos evaluar en $(t,(0,0))$ la matriz jacobiana de f respecto de las variables (θ_1,θ_2).

$$\frac{\partial f}{\partial(\theta_1,\theta_2)}(t,(\theta_1,\theta_2)) = \begin{pmatrix}0 & 1\\ -\frac{g}{L}\cos\theta_1 & 0\end{pmatrix}, \qquad \frac{\partial f}{\partial(\theta_1,\theta_2)}(t,(0,0)) = \begin{pmatrix}0 & 1\\ -\frac{g}{L} & 0\end{pmatrix}.$$

Para calcular $\frac{\partial\Phi}{\partial t_0}(t,t_0,(0,0))$ calculamos la solución del problema lineal

$$y' = \frac{\partial f}{\partial(\theta_1,\theta_2)}(t,(0,0))y = \begin{pmatrix}0 & 1\\ -\frac{g}{L} & 0\end{pmatrix}y, \qquad y(t_0) = -f(t_0,(0,0)) = (0,0).$$

Por unicidad, $y(t) = (0,0)$ para todo t y por tanto $\frac{\partial\Phi}{\partial t_0}(t,t_0,0) = (0,0)$ para todo t.

Por otra parte, para calcular $\frac{\partial\Phi}{\partial(\theta_0,\omega_0)}(t,t_0,(0,0))$, consideramos el sistema matricial

$$Y' = \frac{\partial f}{\partial(\theta_1,\theta_2)}(t,(0,0))Y = \begin{pmatrix}0 & 1\\ -\frac{g}{L} & 0\end{pmatrix}Y, \qquad Y(t_0) = I_2.$$

Este sistema tiene como solución

$$Y(t) = \begin{pmatrix}\cos\sqrt{\frac{g}{L}}(t-t_0) & \sqrt{\frac{L}{g}}\sin\sqrt{\frac{g}{L}}(t-t_0)\\ -\sqrt{\frac{g}{L}}\sin\sqrt{\frac{g}{L}}(t-t_0) & \cos\sqrt{\frac{g}{L}}(t-t_0)\end{pmatrix}.$$

Finalmente, por ser

$$\Phi(t,0,(\theta_0,\omega_0)) = (\theta_1(t),\theta_2(t)) = (\theta(t,0,(\theta_0,\omega_0)),\theta'(t,0,(\theta_0,\omega_0))),$$

se tiene

$$\frac{\partial\Phi}{\partial(\theta_0,\omega_0)}(t,0,(0,0)) = \begin{pmatrix} \frac{\partial\theta}{\partial\theta_0}(t,t_0,(0,0)) & \frac{\partial\theta}{\partial\omega_0}(t,t_0,(0,0)) \\ \frac{\partial\theta'}{\partial\theta_0}(t,t_0,(0,0)) & \frac{\partial\theta'}{\partial\omega_0}(t,t_0,(0,0)) \end{pmatrix}$$

$$= \begin{pmatrix} \cos\sqrt{\frac{g}{L}}(t-t_0) & -\sqrt{\frac{L}{g}}\sin\sqrt{\frac{g}{L}}(t-t_0) \\ \sqrt{\frac{g}{L}}\sin\sqrt{\frac{g}{L}}(t-t_0) & \cos\sqrt{\frac{g}{L}}(t-t_0) \end{pmatrix}.$$

de donde concluimos

$$\frac{\partial\theta}{\partial\theta_0}(t,t_0,(0,0)) = \cos\sqrt{\frac{g}{L}}(t-t_0), \qquad \frac{\partial\theta}{\partial\omega_0}(t,t_0,(0,0)) = \sqrt{\frac{L}{g}}\sin\sqrt{\frac{g}{L}}(t-t_0).$$

Con esta expresión de las derivadas parciales podemos desarrollar por Taylor la función $\theta(t,t_0,(\theta_0,\omega_0))$ alrededor de $(\theta_0,\omega_0)=(0,0)$. En efecto, para (θ_0,ω_0) cercano a $(0,0)$ se tiene

$$\theta(t,t_0,(\theta_0,\omega_0)) = \theta(t,t_0,(0,0)) + \frac{\partial\theta}{\partial\theta_0}(t,t_0,(0,0))\theta_0 + \frac{\partial\theta}{\partial\theta_0}(t,t_0,(0,0))\omega_0 + R(t,t_0,\theta_0,\omega_0).$$

con

$$\lim_{(\theta_0,\omega_0)\to(0,0)} \frac{R(t,t_0,\theta_0,\omega_0)}{||(\theta_0,\omega_0)||} = 0.$$

Sustituyendo las expresiones de las derivadas parciales, tenemos la aproximación

$$\theta(t,t_0,(\theta_0,\omega_0)) = \theta_0\cos\sqrt{\frac{g}{L}}(t-t_0) + \omega_0\sqrt{\frac{L}{g}}\sin\sqrt{\frac{g}{L}}(t-t_0) + R(t,t_0,\theta_0,\omega_0).$$

Los términos de primer orden corresponden exactamente a la solución de la ecuación del oscilador armónico simple

$$\phi'' + \frac{g}{L}\phi = 0, \qquad \phi(t_0)=\theta_0, \quad \phi(t_0)=\omega_0.$$

En el Teorema 4.4.8 hemos derivado el flujo en el abierto $\mathcal{A}$ dado por (4.4.1). No obstante, el problema podría cambiar en el caso $f:\Omega\subset\mathbb{R}^2\to\mathbb{R}$, al considerar t_0 y x_0 iguales. En efecto, en tal caso el subconjunto de $\mathcal{A}$,

$$\mathcal{G} = \{(t,t_0,t_0)\colon (t_0,t_0)\in\Omega,\ t\in I(t_0,t_0)\},$$

es abierto y la ecuación variacional satisfecha por las derivadas parciales es la misma y lo que cambia es la condición inicial. Para deducirla, consideramos el flujo restringido

$$\xi:\mathcal{G}\to\mathbb{R}, \qquad \xi(t,t_0) = \Phi(t,t_0,t_0).$$

Ahora bien, $\xi(t_0,t_0)=\Phi(t_0,t_0,t_0)=t_0$. Derivando respecto de t_0 se tiene

$$\frac{d}{dt}\xi(t_0,t_0) + \frac{\partial\xi}{\partial t_0}(t_0,t_0) = \frac{d}{dt}\Phi(t_0,t_0,t_0) + \frac{\partial\Phi}{\partial t_0}(t_0,t_0,t_0) + \frac{\partial\Phi}{\partial x_0}(t_0,t_0,t_0).$$

Sustituyendo, concluimos

$$\frac{\partial\xi}{\partial t_0}(t_0,t_0) = 1 - f(t_0,t_0).$$

Si denotamos $v(t)=\frac{\partial\xi}{\partial t_0}(t,t_0)$, la ecuación variacional satisfecha es

$$v' = \frac{\partial f}{\partial x}(t,\xi(t,t_0))v, \qquad v(t_0) = 1 - f(t_0,t_0).$$

4.4.4. Dependencia diferenciable respecto a parámetros

A continuación nos centramos en el caso de diferenciabilidad respecto a parámetros, el cual no presenta dificultad añadida puesto que podemos tratar las condiciones iniciales como parámetros y viceversa a conveniencia. El resultado que probamos es el siguiente.

Teorema 4.4.10. *Sean $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ y $\Lambda \subset \mathbb{R}^s$ conjuntos abiertos y conexos, y sea $f : \Omega \times \Lambda \to \mathbb{R}^n$ continua. Supongamos que las derivadas parciales $\frac{\partial f}{\partial x}$ y $\frac{\partial f}{\partial \lambda}$ existen y las funciones*

$$(t, x, \lambda) \mapsto \frac{\partial f}{\partial x}(t, x, \lambda), \qquad (t, x, \lambda) \mapsto \frac{\partial f}{\partial \lambda}(t, x, \lambda)$$

son continuas. Sea $\Psi : \mathcal{B} \to \mathbb{R}^n$ el flujo de la EDO

$$x' = f(t, x, \lambda),$$

definido en (4.4.5). Entonces, Ψ es de clase C^1 en $\mathcal{B}$. Más, aún, fijados (t_0, x_0), el conjunto

$$\mathcal{G} = \{(t, \lambda) \colon (t_0, x_0, \lambda) \in \Omega \times \Lambda,\ t \in I(t_0, x_0, \lambda)\}$$

es abierto y se tiene

$$\frac{\partial \Psi}{\partial \lambda}(t, t_0, x_0, \lambda) = Y(t, t_0, x_0, \lambda), \qquad \forall (t, \lambda) \in \mathcal{G},$$

donde Y es la solución (matricial) del problema

$$Y' = \frac{\partial f}{\partial x}(t, \Psi(t, t_0, x_0, \lambda_0), \lambda)Y + \frac{\partial f}{\partial \lambda_0}(t, \Psi(t, t_0, x_0, \lambda_0), \lambda_0), \qquad Y(t_0) = 0. \tag{4.4.12}$$

Demostración. Que así definido $\mathcal{G}$ es abierto se deduce de que $\mathcal{B}$ es abierto y G es la proyección de $\mathcal{B}$ sobre la primera y última coordenadas, siendo las proyecciones aplicaciones continuas.

Sea $y(t) = (x(t), \lambda(t))$ y consideremos el sistema ampliado

$$\begin{cases} y' = (f(t, x, \lambda), \mathbf{0}_s) = F(t, y), \\ y(t_0) = (x_0, \lambda) = y_0. \end{cases}$$

La aplicación F está en las condiciones del Teorema 4.4.8 y por tanto el flujo

$$\Phi : \mathcal{A} \to \mathbb{R}^n, \qquad \mathcal{A} = \{(t, t_0, y_0) \colon (t_0, y_0) \in \Omega,\ t \in I(t_0, y_0)\}$$

es una aplicación C^1. Que el flujo Ψ es C^1 se tiene de la diferenciabilidad del flujo Φ y de la equivalencia dada en (4.4.6).

Queda por deducir la expresión de la ecuación variacional asociada a la parcial respecto de λ. Sabemos que la ecuación variacional asociada al flujo Φ respecto de la variable y es

$$\frac{\partial \Phi}{\partial y_0}(t, t_0, y_0) = Z(t),$$

donde Z es la solución matricial del sistema

$$Z' = \frac{\partial F}{\partial y}(t, \Phi(t, t_0, y_0))Z, \qquad Z(t_0) = I_{n+s}.$$

Ahora bien,

$$\frac{\partial F}{\partial y}(t,\Phi(t,t_0,y_0)) = \begin{pmatrix} \frac{\partial f}{\partial x}(t,\Phi(t,t_0,y_0)) & \frac{\partial f}{\partial \lambda}(t,\Phi(t,t_0,y_0)) \\ \mathbf{0}_{s\times n} & \mathbf{0}_s \end{pmatrix}, \qquad \begin{matrix} \frac{\partial f}{\partial x}(t,\Phi(t,t_0,y_0)) \in \mathcal{M}_n \\ \frac{\partial f}{\partial \lambda}(t,\Phi(t,t_0,y_0)) \in \mathcal{M}_{n\times s}. \end{matrix}$$

Por otra parte, de la relación (4.4.6) y la condición inicial $Z(t_0) = I_{n+s}$ se tiene

$$\frac{\partial \Phi}{\partial y_0}(t,t_0,y_0) = \begin{pmatrix} \frac{\partial \Psi}{\partial x_0}(t,t_0,x_0,\lambda) & \frac{\partial \Psi}{\partial \lambda}(t,t_0,x_0,\lambda) \\ \mathbf{0}_{s\times n} & I_s \end{pmatrix}.$$

Multiplicando las expresiones matriciales se tiene que se satisface la EDO

$$\left(\frac{\partial \Psi}{\partial \lambda}(t,t_0,x_0,\lambda)\right)' = \frac{\partial f}{\partial x}(t,\Psi(t,t_0,x_0,\lambda),\lambda)\frac{\partial \Psi}{\partial \lambda}(t,t_0,x_0,\lambda) + \frac{\partial f}{\partial \lambda}(t,\Psi(t,t_0,x_0,\lambda),\lambda).$$

En otras palabras, $\frac{\partial \Psi}{\partial \lambda}(t,t_0,x_0,\lambda)$ es solución de la EDO matricial (4.4.12).

Para deducir la condición inicial, consideremos el flujo $\Psi(t,t_0,x_0,\lambda)$. Como (t_0,x_0) están fijos, se tiene $\Psi(t_0,t_0,x_0,\lambda) = x_0$ para todo λ. Derivando respecto de λ se tiene

$$\frac{\partial \Psi}{\partial \lambda}(t_0,t_0,x_0,\lambda) = 0;$$

en otras palabras, $Y(t_0) = 0$. □

Ejemplo 4.3. Consideremos la EDO

$$x' = \epsilon x^2,$$

y sea $\Psi(t,t_0,x_0,\epsilon)$ su flujo; esto es, la única solución con condición inicial $x(t_0) = x_0$. Algunos casos triviales son

- Si $\epsilon = 0$ entonces $\Psi(t,t_0,x_0,0) = x_0$.
- Si $x_0 = 0$ entonces $\Psi(t,t_0,0,\epsilon) = 0$.

En este caso podemos integrar explícitamente el flujo,

$$\Psi(t,t_0,x_0,\epsilon) = \frac{x_0}{1-\epsilon(t-t_0)x_0},$$

donde Ψ está definido en todo el abierto $\mathcal{B}$ dado en (4.4.5). En este caso particular podemos derivar Ψ respecto de sus variables, obteniendo

$$\begin{aligned} \frac{\partial \Psi}{\partial t_0}(t,t_0,x_0,\epsilon) &= -\frac{\epsilon x_0^2}{(1-\epsilon(t-t_0)x_0)^2}, \\ \frac{\partial \Psi}{\partial x_0}(t,t_0,x_0,\epsilon) &= \frac{1}{(1-\epsilon(t-t_0)x_0)^2}, \\ \frac{\partial \Psi}{\partial \epsilon}(t,t_0,x_0,\epsilon) &= \frac{(t-t_0)x_0^2}{(1-\epsilon(t-t_0)x_0)^2}. \end{aligned}$$

Podemos comprobar que estas expresiones explícitas coinciden, como no puede ser de otra forma, con las soluciones de los problemas lineales correspondientes.

Ahora bien, supongamos que restringimos el dominio de parámetros $\mathcal{B}$ a un abierto concreto, por ejemplo

$$\mathcal{G} = \{(t, \epsilon, x_0, \epsilon)\} \subset \mathcal{B}.$$

Estamos ahora diciendo que el instante inicial t_0 y el parámetro ϵ coinciden. El flujo Ψ restringido a $\mathcal{G}$ y denotado por ξ es

$$\xi(t, \epsilon, x_0, \epsilon) = \frac{x_0}{1 - \epsilon(t - \epsilon)x_0}.$$

Observemos que

$$\frac{\partial \Psi}{\partial \epsilon}(t, \epsilon, x_0, \epsilon) = \frac{(t - \epsilon)x_0^2}{(1 - \epsilon(t - \epsilon)x_0)^2}, \qquad \frac{\partial \xi}{\partial \epsilon}(t, \epsilon, x_0, \epsilon) = \frac{(t - 2\epsilon)x_0^2}{(1 - \epsilon(t - \epsilon)x_0)^2}.$$

las cuales no coinciden. Esto es porque estamos derivando Ψ en dos abiertos distintos y por tanto las derivadas direccionales cambian. En la primera derivamos Ψ en todo $\mathcal{B}$, pero en la segunda derivamos Ψ *estando ya restringido a* $\mathcal{G}$.

Ejercicios tema 4

1. Probar que existe una única solución maximal del problema de Cauchy

$$\begin{cases} x' = x^2 + t, \\ x(0) = 0. \end{cases}$$

¿Puede estar definida en $\mathbb{R}$? (*Indicación: estudiar el extremo superior del intervalo maximal de definición y ayudarse de la función* $z(t) = \arctan x(t)$).

2. Sea $f : I \subset \mathbb{R} \to \mathbb{R}$ una función localmente lipschitziana. Probar que toda solución de $x' = f(x)$ es o bien creciente o bien decreciente. ¿Es cierto si únicamente se impone que f sea continua?

3. Probar el Teorema de Peano en una banda usando resultados de maximalidad y prolongación: si $f : I \times \mathbb{R}^n \to \mathbb{R}^n$ es continua y acotada y φ es una solución maximal de $x' = f(t, x)$, entonces φ está definida en todo I. (*Indicación: si b es el extremo superior de I, suponer $\omega_+ < b$, tomar $t_n \to \omega_+$ y distinguir casos entre $\varphi(t_n)$ acotada o no*).

4. Se considera el problema

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0, \end{cases}$$

donde $f : [a, b] \times \mathbb{R}^n \to \mathbb{R}^n$ es continua.

   a) Probar que dicho problema tiene al menos una solución maximal $\varphi : I \to \mathbb{R}^n$ y que o bien $I \cap [t_0, \infty) = [t_0, b]$ o bien $I \cap [t_0, \infty) = [t_0, \omega_+)$ con $\omega_+ \leq b$ y $||x(t)|| \to \infty$ cuando $t \to \omega_+$.

   b) Enunciar y deducir el resultado análogo a la izquierda de I.

5. Sean $g : \mathbb{R} \to \mathbb{R}$ lipschitziana y $f : \mathbb{R} \to \mathbb{R}$ continua. Probar que el sistema

$$\begin{cases} x' = g(x), \\ y' = f(x)y, \end{cases}$$

tiene una única solución maximal para cada condición inicial. ¿Se tiene el mismo resultado y sólo se supone g continua?

6. Probar que el problema de Cauchy

$$\begin{cases} x' = \sin(tx), \\ x(t_0) = x_0, \end{cases}$$

tiene una única solución maximal definida en todo $\mathbb{R}$. Probar que si $x_0 > 0$, entonces $x(t) > 0$ para todo t.

7. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^1. Supongamos que la ecuación $x' = f(t, x)$ tiene entre sus soluciones a las funciones constantes $x_1(t) = 1$ y $x_2(t) = -1$.

   a) Probar que la solución maximal del problema de Cauchy con condición inicial $\varphi(t_0) = x_0 \in [-1, 1]$ está definida en todo $\mathbb{R}$.

b) Probar que si $x_0 > 1$ (resp. $x_0 < -1$), entonces toda solución maximal con condición inicial $\varphi(t_0) = x_0$ está acotada inferiormente (resp. superiormente).

8. Sea $f : \mathbb{R} \to \mathbb{R}$ una función C^1 y periódica. Supongamos que existe $t_* \in \mathbb{R}$ tal que $f(t_*) = 0$. Probar que las soluciones maximales de $x' = f(x)$ están acotadas y definidas en $\mathbb{R}$.

9. Consideremos el problema de Cauchy

$$\begin{cases} x' = 1 - x^2, \\ x(0) = x_0. \end{cases}$$

Obtener todas las soluciones maximales en función de x_0. Deducir que para una misma EDO se pueden plantear distintos problemas de Cauchy (distintas condiciones iniciales) cuyas soluciones maximales pueden estar definidas en intervalos de distinta longitud.

10. Probar que la solución maximal del problema

$$\begin{cases} x' = xe^{t-x^2}, \\ x(t_0) = x_0, \end{cases}$$

está definida en todo $\mathbb{R}$.

11. Sea $f : [t_0 - \varepsilon, t_0 + \varepsilon] \times \overline{B}(x_0, r) \to \mathbb{R}^n$ continua. Definimos los iterantes de Picard de forma recursiva mediante

$$x_0(t) = x_0, \qquad x_{n+1}(t) = x_0 + \int_{t_0}^{t} f(s, x_n(s))\, ds.$$

a) Probar que $\{x_n\}$ es una sucesión unifórmemente equicontinua y unifórmemente acotada.

b) Deducir la existencia de una subsucesión $\{x_{\sigma(n)}\}$ que converge unifórmemente en $[t_0 - \varepsilon, t_0 + \varepsilon]$ a una función x_∞ que cumple $x_\infty(t_0) = x_0$.

c) ¿Es necesariamente x_∞ solución de la EDO $x' = f(t, x)$? (*Indicación: estudiar el paso al límite de los iterantes de Picard de la subsucesión* $\{x_{\sigma(n)}\}$).

12. Sea $f : \Omega \to \mathbb{R}^n$ una función continua y localmente lipschitziana en su segunda variable. Definimos en Ω la siguiente relación binaria: $(t_1, x_1) \sim (t_2, x_2)$ si y sólo si existe $x : [t_1, t_2] \to \mathbb{R}^n$ solución de $x' = f(t, x)$ tal que $x(t_1) = x_1,\ x(t_2) = x_2$. Probar que $\sim$ es una relación de equivalencia. Describir el conjunto cociente $\Omega/\sim$. Dado $(t_0, x_0) \in \Omega$, ¿cuál es la clase de equivalencia $[(t_0, x_0)]$?

13. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ una función de clase C^1. Demostrar que si $x' = f(t, x)$ tiene una solución maximal cuyo intervalo de definición no es todo $\mathbb{R}$, entonces tiene infinitas soluciones maximales con la misma propiedad (y con gráficas disjuntas dos a dos).

14. Sean $I \subset \mathbb{R}$ un intervalo y $f : I \times \mathbb{R}^n \to \mathbb{R}^n$, $\alpha : I \to [0, \infty)$ funciones continuas tales que $||f(t, x)|| \leq \alpha(t)$ para todo $(t, x) \in I \times \mathbb{R}^n$. Demostrar que las soluciones maximales de $x' = f(t, x)$ están definidas en I.

15. Demostrar la siguiente versión del Lema de Gronwall: Sean $f : I \subset \mathbb{R} \to \mathbb{R}$ y $t_0 \in I$ tales que
$$0 \leq f(t) \leq A + B \int_{t_0}^{t} f(s)\, ds,$$
con $A, B \geq 0$ constantes. Entonces, $f(t) \leq Ae^{B(t-t_0)}$.
Indicación: definir $F(t) = \int_{t_0}^{t} f(s)\, ds$, escribir la desigualdad dada por la hipótesis en términos de F y F' y multiplicar por el factor integrante $e^{-B(t-t_0)}$.

16. Sean $I \subset \mathbb{R}$ un intervalo y $f : I \times \mathbb{R}^n \to \mathbb{R}^n$, $\alpha, \beta : I \to [0, \infty)$ funciones continuas tales que $||f(t,x)|| \leq \alpha(t) + \beta(t)||x||$ para todo $(t,x) \in I \times \mathbb{R}^n$. Demostrar que las soluciones maximales de $x' = f(t,x)$ están definidas en I. *Indicación: usar el Lema de Grönwall y las ideas del ejercicio 14.*

17. Sea $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ abierto y $f : \Omega \to \mathbb{R}^n$ continua y lipschitziana respecto a la segunda variable, con constante de Lipschitz K. Probar que para cada $(t_0, x_0), (t_0, y_0) \in \Omega$ y $t \in I(t_0, x_0) \cap I(t_0, y_0)$ se verifica
$$||\varphi(t, t_0, x_0) - \varphi(t, t_0, y_0)|| \leq ||x_0 - y_0|| e^{K|t-t_0|}.$$
Indicación: usar el Lema de Grönwall.

18. Sea $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ continua y localmente lipschitziana en la segunda variable. Consideremos una sucesión $(t_k, x_0^k, ..., x_{n-1}^k) \in \Omega$ convergiendo a $(t_0, x_0, ..., x_{n-1}) \in \Omega$. Finalmente, denotemos por x_k, con $k \in \mathbb{N} \cup \{0\}$, a la solución del problema de Cauchy
$$\begin{cases} x^{n)} = f(t, x, x', ..., x^{n-1)}), \\ x(t_k) = x_0^k, \\ x'(t_k) = x_1^k, \\ \vdots \\ x^{n-1)}(t_k) = x_{n-1}^k. \end{cases}$$
Probar que para cada $j = 0, ..., n-1$ se tiene
$$\lim_{k \to \infty} x_k^{j)}(t) = x_0^{j)}(t),$$
uniformemente sobre cada $[a, b]$ conteniendo a t_0.

19. Sea $f : \Omega \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}^n$ diferenciable. Consideremos la EDO $x' = f(t, x, \lambda)$ y sea $\Phi(t, t_0, x_0, \lambda)$ su flujo. Calcular las condiciones iniciales satisfechas por $\frac{\partial \Phi}{\partial \lambda}(t, t_0, x_0, \lambda)$ en los siguientes problemas de Cauchy.

    - El instante inicial es el parámetro $\lambda \in \mathbb{R}$,
    $$\begin{cases} x' = f(t, x, \lambda), \\ x(\lambda) = x_0, \end{cases}$$
    - La condición incial es una función,
    $$\begin{cases} x' = f(t, x, \lambda), \\ x(t_0) = g(\lambda), \end{cases} \qquad g : \Lambda \subset \mathbb{R}^s \to \mathbb{R}^n.$$

- El instante y la condición iniciales dependen del parámetro $\lambda \in \mathbb{R}$,

$$\begin{cases} x' = f(t, x, \lambda), \\ x(\lambda) = g(\lambda), \end{cases} \qquad g : \Lambda \subset \mathbb{R} \to \mathbb{R}.$$

20. Consideremos el problema de Cauchy

$$\begin{cases} x' = \epsilon x^2 + x, \\ x(\epsilon) = 1, \end{cases}$$

y denotemos por $x(t, \epsilon)$ a la solución.

a) Justificar de forma rigurosa la existencia de $\frac{\partial x}{\partial \epsilon}(t, 0)$. En caso de que exista, calcularla resolviendo la ecuación variacional asociada.

b) Comparar el resultado integrando explícitamente la ecuación.

21. Consideremos el problema de Cauchy

$$\begin{cases} x'' + x + \epsilon x^3 = 0, \\ x(0) = 1, \\ x'(0) = 0. \end{cases}$$

y denotemos por $x(t, \epsilon)$ a la solución.

- Tras el cambio usual $x' = y,\ y' = -x - \epsilon x^3$, probar que la función

$$E(x, y) = \frac{x^2}{2} + \frac{y^2}{2} + \frac{\epsilon x^4}{4}$$

es constante a lo largo de las soluciones del sistema de primer orden equivalente.

- Deducir del punto anterior que si $\epsilon > 0$, las soluciones están definidas en $\mathbb{R}$.
- Calcular $\frac{\partial x}{\partial \epsilon}(t, 0)$ y deducir un desarrollo asintótico de $x(t, \epsilon)$ para ϵ suficientemente pequeño.