Métodos iterativos sobre grupos de Lie

María del Pilar Astudillo Fernández

Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Departamento de Matemáticas
Doctorado en Ciencias Matemáticas
Popayán
Diciembre de 2022

Métodos iterativos sobre grupos de Lie

María del Pilar Astudillo Fernández

Tesis presentada como requisito parcial para optar al título de:
Doctor en Ciencias Matemáticas

Director
Dr. Willy Will Sierra Arroyo
Profesor de la Universidad del Cauca

Codirector
Dr. Rodrigo Alberto Castro Marín
Profesor de la Universidad de Valparaíso

Universidad del Cauca
Facultad de Ciencias Naturales, Exactas y de la Educación
Departamento de Matemáticas
Doctorado en Ciencias Matemáticas
Popayán
Diciembre de 2022

Agradecimientos

Agradece a la llama su luz, pero no olvides el pie del candil que,
constante y paciente, la sostiene en la sombra.
Tagore.
Premio Nobel de Literatura, 1913.

Gracias a todos los que me han irradiado con energía positiva a lo largo de este camino,
no necesito enlistarlos...,ustedes saben que están en mi corazón.
Un sentido agradecimiento a mi señora madre
por ser, toda mi vida, el pie del candil.
María del Pilar.

Resumen

Realizamos un análisis de convergencia semilocal para algunos métodos iterativos sobre grupos de Lie. En particular estudiamos el método de Newton bajo las condiciones clásicas tipo Kantorovich e incluimos un método tipo secante y el método de King-Werner. Los métodos introducidos permiten aproximar ceros de funciones definidas en un grupo de Lie con valores en su respectiva álgebra. Un análisis del orden de convergencia del método de la secante también es presentado.

Palabras clave: Grupo de Lie, álgebra de Lie, métodos iterativos, método de Newton, método de la secante, diferencias divididas, método de King-Werner.

Abstract

We perform a semilocal convergence analysis for some iterative methods on Lie groups. In particular, we study Newton's method under classical Kantorovich-type conditions and include a secant-type method and King-Werner's method. The introduced methods allow us to approximate zeros of functions defined on a Lie group with values in their respective algebra. An analysis of the order of convergence of the secant method is also presented.

Keywords: Lie group, Lie algebra, iterative methods, Newton's method, secant method, divided differences, King-Werner method.

Productos de la investigación

Artículos

Semilocal Convergence Analysis For The Secant Method On Lie Groups, (Preprint. Con R. Castro y W. Sierra).

Ponencias

- *Métodos Iterativos en grupos de Lie*. XXXIII Jornada de Matemática de la Zona Sur, Universidad de la Frontera, Temuco, Chile, abril 20–24, 2021.
- *Métodos numéricos en variedades Riemannianas*. XXI Congreso Colombiano de Matemáticas, Universidad del Cauca, Popayán, Colombia, junio 10–20, 2019.

Índice general

Capítulo 1

Introducción

La solución de algunos problemas en diversas áreas de la ciencia y la ingeniería puede ser planteada como una raíz de ecuaciones del tipo

$$F(x) = 0, \tag{1.1}$$

donde F es una función definida en un espacio apropiado, el cual depende de las condiciones del problema a resolver. Por ejemplo, algunos problemas pueden ser reducidos a solucionar un sistema de ecuaciones (lineales o no lineales), lo cual resulta equivalente a hallar los ceros de una función del tipo $F : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Este caso es ampliamente estudiado en los cursos básicos de álgebra lineal. Otro tipo de problemas, como por ejemplo hallar una solución $x \in C[a, b]$ de una ecuación de la forma

$$x(s) = u(s) + \int_a^b K(s, t)H(x)(t)dt, \quad a \leq s \leq b, \tag{1.2}$$

donde $u \in C[a, b]$, $K \in C[a, b]^2$ y $H : C[a, b] \to C[a, b]$ son conocidas, es equivalente a hallar los ceros del operador $F : C[a, b] \to C[a, b]$ dado por

$$F(x)(s) = x(s) - u(s) - \int_a^b K(s, t)H(x)(t)dt, \quad a \leq s \leq b.$$

Ecuaciones integrales de la forma (1.2) son conocidas como ecuaciones del tipo Hammerstein, las cuales aparecen en problemas físicos relacionados con electromagnetismo y dinámica de fluidos. En otros contextos también surgen problemas que se reducen a determinar los ceros de campos vectoriales sobre variedades Riemannianas y otros donde el escenario natural son grupos de Lie. Es en este último donde se centra la presente tesis.

En general, a problemas como los anteriormente expuestos no se les puede determinar una solución exacta o explícita, razón por la cual se ha hecho necesario desarrollar métodos para aproximar dichas soluciones. A lo largo de la historia de las matemáticas se han desarrollado diversos procesos de aproximación de soluciones a algunos problemas, como por ejemplo la longitud o el área de una circunferencia o la solución de ecuaciones en una variable real, desde la aparición de las computadoras, tales procesos se han vuelto más relevantes. Un sin número de métodos numéricos han sido propuestos con el objetivo de abordar problemas como los descritos arriba, y en particular los conocidos como métodos iterativos son de especial interés. En términos generales, un método (o proceso) iterativo en un espacio métrico E es una sucesión $(x_n) \subset E$ de la forma

$$\begin{aligned} &x_0, \ldots, x_k \in E \quad \text{dados;} \\ &x_{n+1} = G(x_0, \ldots, x_n), \quad n \geq k. \end{aligned}$$

Como es señalado, por ejemplo en [10], una vez un proceso iterativo ha sido planteado, surgen diferentes problemas alrededor de ellos, algunos de los cuales suelen abordarse de manera independiente. Un primer problema es si el algoritmo está bien definido. Por ejemplo, algunos métodos requieren la evaluación de $F(x_n)$, lo cual implica que se debe garantizar que x_n pertenezca al dominio de definición de F. Situación similar, como ocurre con el método de Newton, hay que resolver si el algoritmo requiere la existencia de $[F'(x_n)]^{-1}$. Un segundo problema es garantizar que la sucesión (x_n) generada por el proceso iterativo converja y su límite sea en efecto una solución de (1.1). En relación con este problema, en general se estudian dos tipos de resultados de convergencia, a saber: resultados de convergencia local y semilocal. En el primero de ellos, se asume la existencia de una solución x^* del problema (1.1) y se estudia bajo qué hipótesis sobre un entorno U de x^* la sucesión (x_n) está contenida en U y converge a x^* para toda elección de datos iniciales $x_0, \ldots, x_k \in U$. Para el caso de teoremas de convergencia semilocal, no se asume la existencia de una solución de $F(x) = 0$. En este tipo de resultados se estudian hipótesis sobre datos iniciales $x_0, \ldots, x_k$ que garanticen que la sucesión generada converja a una solución de $F(x) = 0$. Existe otro tipo de resultados de convergencia, llamados teoremas de convergencia global. En estos se investigan condiciones bajo las cuales, comenzando la iteración en cualquier parte del dominio, ésta converge a un cero x^* de F. Otro problema que se estudia en relación con métodos iterativos es la velocidad de convergencia del método, esta usualmente se obtiene como consecuencia de estimativos de error previamente calculados. Este trabajo lo centramos en un análisis de convergencia semilocal para algunos métodos iterativos en grupos de Lie.

Sin duda, el método iterativo mas simple y probablemente el mas conocido, es el método de Newton, definido para una función real $f : I \subseteq \mathbb{R} \to \mathbb{R}$ por la iteración

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad n \geq 0,$$

con $x_0 \in I$ dado. La primera versión de este método fue presentada por Newton en 1669 e incluida formalmente en su libro "*Método de las Fluxiones*", el cual fue redactado en 1671 y publicado solo hasta 1736. Si bien el método fue inicialmente propuesto por Newton, Joseph Raphson hizo algunos aportes importantes, incluyendo la forma iterativa, razón por la cual este procedimiento iterativo es conocido como método de Newton-Raphson. El estudio realizado por J. Raphson fue incluido en su libro "*Aequationum Universalis*" publicado en 1690. Otros ejemplos de métodos iterativos (en una dimensión) importantes que surgieron para solucionar cierto tipo de ecuaciones en una variable real son:

a) El método de la secante, el cual es definido por la iteración

$$x_{n+1} = \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n), \qquad n \geq 1,$$

con $x_0, x_1 \in I$ dados. Detalles históricos sobre el método de la secante pueden ser consultados en [49].

b) El método de Halley, propuesto por Edmund Halley en 1708 cuando buscaba aproximar una raíz cúbica de un entero de la forma $a^3 + b$, con $a, b \in \mathbb{Z}$, ver [26] para mas detalles. El método es definido por la fórmula

$$x_{n+1} = x_n - \frac{2f'(x_n)}{2f'(x_n)^2 - f(x_n)f''(x_n)} f(x_n), \qquad n \geq 0,$$

con $x_0 \in I$ dado.

c) El método definido por la sucesión

$$x_{n+1} = x_n - \left(1 + \frac{1}{2}\frac{f(x_n)f''(x_n)}{f'(x_n)^2}\right)\frac{f(x_n)}{f'(x_n)}, \qquad n \geq 0,$$

con $x_0 \in I$ dado, se atribuye al matemático ruso Pafnuty Lvovich Chebyshev, quien lo formuló en 1838. Debido a que el método también se atribuye a Leonhard Euler, se conoce en la literatura también como método de Euler-Chebyshev, ver [30].

Es de anotar que los métodos iterativos mas importantes, dentro de los cuales se encuentran los expuestos anteriormente, tienen su versión multidimensional; la del método de Newton fue obtenida por Augustin Cauchy en 1829.

Como es natural, y debido al hecho que los métodos iterativos demostraron ser una herramienta eficaz para resolver problemas tanto en el contexto real como en el multidimensional, desde el siglo anterior algunos de tales métodos han sido extendidos a otros escenarios donde también surgen problemas que pueden ser abordados con estas herramientas. El primer paso en esta dirección se dio en 1948, cuando Leonid Kantorovich, matemático ruso y premio nobel de economía 1975, extendió el método de Newton-Raphson a espacios de Banach [40]. Este trabajo proporcionó un nuevo enfoque para resolver ecuaciones integrales, ecuaciones diferenciales, problemas variacionales, entre otros. En su primer trabajo, Kantorovich consideró una serie de condiciones sobre el punto inicial x_0 de la iteración, hoy conocidas como condiciones de Kantorovich. En especial supuso que la segunda derivada de Fréchet del operador F involucrado era acotada en un cierto entorno de x_0. Desde entonces han surgidos muchos trabajos en los cuales se analiza la convergencia de la iteración de Newton en espacios de Banach, bajo diferentes condiciones, incluyendo entre ellas condiciones tipo Lipschitz sobre la primera derivada de Fréchet del operador. Referimos al libro [27] para un estudio del método de Newton en espacios de Banach bajo estas nuevas hipótesis. Además del método de Newton, los principales métodos iterativos estudiados previamente en $\mathbb{R}$ tienen su versión en espacios de Banach, destacamos el método de la secante [55], el método de Halley [13, 22], el método de Chevyshev [12] y el método Chevyshev-Halley [6, 32]. Otros nuevos métodos también han sido propuestos, de especial interés han sido los métodos multipasos, dentro de los cuales destacamos el método de punto medio [11] y el método de King-Werner [59, 60]. De los métodos iterativos señalados anteriormente, el método de la secante, también conocido como el método de las cuerdas o Regula falsi, es un método que no requiere la diferenciabilidad del operador involucrado, de ahí su importancia. El método fue introducido en espacios de Banach en 1961 por A. Sergeev [55], ver también los trabajos de J. Schmidt [51, 52, 53]. Para su formulación en el contexto de espacios de Banach, los autores usaron una extensión del concepto de diferencias divididas que había sido introducida previamente por Schröder [54] en 1956. Uno de los objetivos centrales de la presente propuesta es realizar un análisis de convergencia semilocal del método de la secante en grupos de Lie.

Siguiendo las ideas de Kantorovich y otros autores en el contexto de espacios de Banach, con el objetivo de aproximar ceros de campos vectoriales sobre variedades Riemannianas o hallar puntos críticos de funciones reales definidas sobre dichas estructuras, distintos métodos iterativos han sido también extendidos al contexto de variedades Riemannianas. Por ejemplo, existen extensiones del método de Newton-Kantorovich, del método de Newton modificado, del método de Chebyshev, entre otros. Más detalles sobre extensiones a variedades Riemannianas se pueden consultar en [1, 2, 3, 4, 5, 7, 14, 18, 20, 21, 23, 24, 29, 42, 44].

De manera similar a lo hecho en el contexto de variedades Riemannianas, métodos iterativos han sido recientemente estudiados en grupos de Lie. La teoría de grupos de Lie, la cual es ampliamente usada en física, matemáticas y economía, fue introducida por el matemático noruego Marius Sophus Lie en una serie de trabajos publicados a partir de 1870. Su motivación inicial de estudio consistía en identificar grupos que jugaran un rol para soluciones de ecuaciones diferenciales, análogo al rol que cumplen los grupos de Galois para polinomios. Para otros aspectos acerca de la historia de grupos de Lie puede consultarse, por ejemplo, [35]. En relación con métodos numéricos sobre grupos de Lie, varias iteraciones de Newton sobre estas estructuras han sido estudiadas. Por ejemplo, en el 2000, B. Orwen y B. Welfert [48], definen la iteración de Newton para aproximar una solución de $F(x) = \mathbf{0}$, donde F es ahora una función de un grupo de Lie en su correspondiente álgebra de Lie. Los autores prueban que sus resultados son independientes de la métrica considerada en el grupo. En 2007, J. Wang y C. Li [43] presentan un análisis de convergencia para el método de Newton sobre grupos de Lie, bajo condiciones tipo Kantorovich. En este estudio, no se supone previamente la existencia de una solución de $F(x) = \mathbf{0}$. Otros resultados sobre el método de Newton en el contexto de grupos de Lie, pueden ser consultados en [8, 15, 36, 44, 45].

Con el fin de presentar algunos resultados teóricos en relación con métodos iterativos en grupos de Lie, y en especial con el método de la secante en este contexto, el presente trabajo se ha organizado como sigue. En el Capítulo 2 precisamos algunos conceptos básicos sobre métodos iterativos, tanto en espacios de Banach como en variedades Riemannianas, e incluimos también los aspectos requeridos sobre la teoría general de grupos de Lie y álgebras de Lie. En este capítulo, elementos de cálculo diferencial en grupos de Lie son presentados. En el Capítulo 3 introducimos el concepto de diferencias divididas sobre grupos de Lie con el fin de definir el método de la secante en este contexto y establecer condiciones que garanticen la convergencia del método; este es el contenido del Teorema 3.1. Adicionalmente probamos que el método de la secante en grupos de Lie, como en el caso de espacios de Banach, tiene orden de convergencia $(1+\sqrt{5})/2$. También, en este capítulo presentamos un análisis de convergencia semilocal del método de Newton-Kantorovich haciendo uso de la técnica de relaciones de recurrencia. Finalmente, en el Capítulo 4, ilustramos nuestro resultado con un ejemplo para un campo vectorial definido desde el grupo de matrices ortogonales $\mathcal{SO}_2(\mathbb{R})$ sobre su álgebra de Lie, esto es, el álgebra de matrices antisimétricas $\mathfrak{so}_2(\mathbb{R})$.

Capítulo 2

Preliminares básicos y notaciones

En esta sección hacemos una breve descripción de las definiciones, teoremas y demás elementos necesarios para desarrollar los resultados principales de la presente tesis. Como punto de partida hacemos una introducción de algunos de los métodos iterativos estudiados en el contexto de espacios de Banach para posteriormente explicar algunos de tales métodos en variedades Riemannianas. Finalmente presentamos un resumen de la teoría de grupos de Lie, el cual es el escenario en el que desarrollamos los resultados de la tesis.

Como fue mencionado en la introducción, la solución a muchos problemas en diversos escenarios suele presentarse como una raíz de cierta ecuación del tipo

$$F(x) = 0, \tag{2.1}$$

cuya formulación depende obviamente de las condiciones dadas en el problema. Ya que en general no es posible obtener una solución exacta de $F(x) = 0$, se hace necesario desarrollar métodos para hallar una solución aproximada, los procesos iterativos han demostrado ser una herramienta apropiada para abordar este problema.

2.1. Métodos iterativos en espacios de Banach

Con el objetivo de presentar algunos de los métodos iterativos mas conocidos en espacios de Banach, iniciamos esta sección con la definición de derivada de un operador entre espacios de Banach.

Definición 2.1. *Sean X, Y espacios de Banach, Ω abierto en X y $F : \Omega \subseteq X \to Y$. Dado $x_0 \in \Omega$, si existe un operador lineal continuo $L : X \to Y$ tal que*

$$\lim_{\|\Delta x\| \to 0} \frac{\|F(x_0 + \Delta x) - F(x_0) - L(\Delta x)\|}{\|\Delta x\|} = 0,$$

diremos que F es Fréchet diferenciable en x_0 y $L := F'(x_0)$ es llamada la primera derivada Fréchet de F en x_0.

De manera inductiva definimos las derivadas de orden superior de F. Por ejemplo, si F es Fréchet diferenciable en Ω y $B(X, Y)$ denota el espacio de operadores lineales acotados de X en Y, entonces $F' : \Omega \to B(X, Y)$. Por lo tanto, $F''(x_0)$ es un operador lineal de X en $B(X, Y)$. Es habitual identificar $F''(x_0) : X \to L(X, Y)$ con la forma bilineal $b : X \times X \to Y$ dada por $b(u, v) = [F''(x_0)(u)](v)$, $u, v \in X$.

Con la notación anterior, el hoy conocido como método de Newton-Kantorovich se define como una extensión natural del método de Newton en $\mathbb{R}$, a saber:

$$x_{n+1} = x_n - F'(x_n)^{-1} F(x_n), \quad n \geq 0,$$

con $x_0 \in \Omega$ dado. Este método se puede considerar como el que dio origen al estudio de métodos iterativos en espacios de Banach, fue estudiado por primera vez en 1948 por el matemático soviético Leonid Kantoróvich, ver [40]. El análisis de convergencia del método de Newton-Kantorovich se ha realizado bajo diferentes hipótesis sobre F; se han obtenido tanto resultados de convergencia semilocal como de convergencia local. En el primer trabajo desarrollado por Kantorovich, la convergencia del método se demuestra bajo las siguientes condiciones sobre F y el punto inicial x_0.

- H1. Existe $[F'(x_0)]^{-1}$ en $B(Y, X)$ tal que $\|[F'(x_0)]^{-1}\| \leq \beta$, para alguna constante positiva β;
- H2. $\|[F'(x_0)]^{-1} F(x_0)\| \leq \eta$, para alguna constante positiva η;
- H3. existe una constante positiva M tal que $\|F''(x)\| \leq M$ para todo $x \in \Omega$.

El resultado de convergencia semilocal es como sigue:

Teorema 2.1 (Kantorovich)**.** *Sean X y Y espacios de Banach, $\Omega \subset X$ abierto y convexo y $F : \Omega \subset X \to Y$ un operador dos veces continuamente Fréchet-diferenciable en Ω. Supongamos que para $x_0 \in \Omega$, las condiciones H1, H2 y H3 se satisfacen y $h := M\beta\eta$*

cumple que $h \leq 1/2$. *Suponga además que* $\overline{B(x_0, t^*)} \subset \Omega$, *donde* $t^* = \dfrac{1-\sqrt{1-2h}}{M\beta}$. *Entonces la sucesión* (x_n) *dada por* (2.1), *está bien definida,* $x_n \in B(x_0, t^*)$ *para todo* $n \geq 0$ *y* (x_n) *converge a una solución* $x^* \in \overline{B(x_0, t^*)}$ *de* $F(x) = 0$.

Mas recientemente se ha probado la convergencia de (2.1) bajo otras condiciones sobre F. Por ejemplo, H. Zhengda [62] presentó un análisis de convergencia semilocal para el método (2.1) bajo una condición Lipschitz sobre la segunda derivada del operador. Zhengda consideró las siguientes condiciones:

C1. existe $[F'(x_0)]^{-1}$ en $B(Y, X)$ tal que para constantes positivas η, γ,

$$\|[F'(x_0)]^{-1}F(x_0)\| \leq \eta \quad \text{y} \quad \|[F'(x_0)]^{-1}F''(x_0)\| \leq \gamma;$$

C2. existe una constante positiva K tal que $\|[F'(x_0)]^{-1}[F''(x) - F''(y)]\| \leq K\|x-y\|$, para todo $x, y \in \Omega$;

C3. $6\eta\gamma^3 + 9\eta^2K^2 + 18\eta\gamma K - 3\gamma^2 - 8K \leq 0$,

bajo las cuales probó que la sucesión (2.1), iniciando en x_0, converge a la única solución de (2.1) en $\overline{B(x_0, t^*)} \cap B(x_0, t^{**})$, donde $0 < t^* \leq t^{**}$ son dos raíces positivas del polinomio

$$p(t) = \frac{1}{6}Kt^3 + \frac{1}{2}\gamma t^2 - t + \eta,$$

ver [62, Teorema 2.1].

Es abundante la literatura existente sobre el método de Newton-Kantorovich. Por ejemplo, un bonito texto donde se realiza un estudio profundo del método, incluyendo un análisis de convergencia bajo diferentes tipos de condiciones, es *Newton's method: an updated approach of Kantorovich's theory*, ver [27]. Buscando mejorar la velocidad de convergencia, en las últimas décadas se han propuesto otros métodos iterativos en espacios de Banach, en esta dirección los métodos multipasos han mostrado ser bastante eficientes; uno de especial interés es el conocido como método de King-Werner, el cual es un método de dos pasos estudiado por W. Werner [59, 60], pero fue propuesto inicialmente por R. King [41]. El método de King-Werner está definido por

$$\begin{cases} y_n = x_n - F'(\frac{x_{n-1}+y_{n-1}}{2})^{-1}F(x_n), \\ x_{n+1} = x_n - F'(\frac{x_n+y_n}{2})^{-1}F(x_n), \end{cases} \tag{2.2}$$

donde $x_0 \in \Omega \subset X$ es dado y $y_0 = x_0 - F'(x_0)^{-1}F(x_0)$. Como en el caso del método de Newton, la convergencia del método de King-Werner también ha sido estudiada bajo diferentes condiciones sobre el operador F, algunos de tales resultados pueden ser consultados en [17, 34, 47, 50, 33, 57].

Resulta útil también el estudio de métodos que no requieren la diferenciabilidad del operador involucrado; el más conocido de estos métodos es el método de la secante. Para su formulación en el contexto de espacios de Banach es necesario introducir la noción de diferencias divididas. Con la notación anterior, un operador $[x, y; F] \in B(X, Y)$ será una diferencia dividida de primer orden de $F : X \to Y$ en el par de puntos $x, y \in X$, con $x \neq y$, si se satisface

$$[x, y; F]\,(x - y) = F\,(x) - F\,(y)\,. \tag{2.3}$$

Notemos que en el caso $X = Y = \mathbb{R}$, la función lineal

$$\varphi(t) = \frac{F(x) - F(y)}{x - y}\,t$$

satisface (2.3). La definición de diferencia dividida es el paso central para darle sentido en espacios de dimensión superior a cocientes del tipo $\dfrac{x - y}{F(x) - F(y)}$ que aparecen en el caso real.

Con esta notación, en [55] (ver también [53]) el método de la secante en espacios de Banach es descrito por el siguiente algoritmo:

$$x_{n+1} = x_n - [x_{n-1}, x_n; F]^{-1}\,F\,(x_n)\,, \qquad x_{-1}, x_0 \text{ dados}, \qquad n \geq 0. \tag{2.4}$$

Resultados relacionados con la convergencia del método secante (2.4), incluyendo un análisis del orden de convergencia del método, pueden consultarse en [16, 31, 37, 53].

2.2. Métodos iterativos en variedades Riemannianas

Otro escenario en el cual se han extendido los métodos iterativos son las variedades Riemannianas. Recientemente se han propuesto versiones del método de Newton, método de la secante, entre otros métodos a este nuevo contexto. Para una mejor comprensión de algunos métodos iterativos que presentaremos, introducimos los elementos de geometría Riemanniana requeridos para tal fin. Referimos a [25] para un estudio mas profundo de los elementos de geometría Riemanniana aquí expuestos.

Sea (M,g) una variedad Riemanniana m-dimensional. Dado $p \in M$, T_pM denotará el espacio tangente a M en p y $\langle \cdot, \cdot \rangle_p$ el producto escalar sobre T_pM, el cual induce la norma $\| \cdot \|_p = \langle \cdot, \cdot \rangle_p^{1/2}$, donde el subíndice p puede ser omitido siempre que no exista posibilidad de confusión. El fibrado tangente de M es definido por

$$TM := \{(p,v) \mid p \in M \quad \text{y} \quad v \in T_pM\} = \bigcup_{p \in M} T_pM,$$

el cual puede ser dotado con una estructura diferenciable que lo convierte en una variedad $2m-$dimensional. Si F es una función diferenciable entre variedades, su derivada en el punto p la denotaremos en adelante por F'_p. Un campo vectorial X sobre M es una función que asigna a cada punto $p \in M$ un vector tangente

$$X_p := X(p) \in T_pM.$$

Diremos que el campo vectorial X es diferenciable si la función $X : M \to TM$ es diferenciable. En adelante, $\mathcal{X}(M)$ denotará el espacio de todos los campos vectoriales diferenciables sobre M y usaremos $\mathcal{D}(M)$ para denotar el anillo de funciones diferenciables de valor real sobre M. Dado $X \in \mathcal{X}(M)$ y $f \in \mathcal{D}(M)$, $f'(X) = X(f)$ representa la derivada direccional de f en la dirección X, donde f' denota la diferencial de f.

Si $\gamma : [a,b] \longrightarrow M$ es una curva suave por partes, la longitud de γ se define por la fórmula

$$l(\gamma) = \int_a^b ||\gamma'(t)|| \, dt = \int_a^b \langle \gamma'(t), \gamma'(t) \rangle^{1/2} \, dt.$$

La distancia Riemanniana desde p hasta q es definida por

$$d(p,q) := \inf_{\gamma} l(\gamma),$$

donde el ínfimo es tomado sobre todas las curvas suaves γ con extremos p y q. La topología inducida por la métrica d coincide con la topología de la variedad M.

Una conexión afín ∇ sobre M es una función

$$\nabla : \mathcal{X}(M) \times \mathcal{X}(M) \longrightarrow \mathcal{X}(M)$$

que asigna a la pareja de campos vectoriales (X,Y) el campo vectorial $\nabla(X,Y) := \nabla_X Y$ satisfaciendo las siguientes condiciones:

i) $\nabla_X Y$ es $\mathcal{D}(M)-$lineal en X :

$$\nabla_{fX+gY} Z = f \nabla_X Z + g \nabla_Y Z;$$

ii) $\nabla_X Y$ es $\mathbb{R}-$lineal en Y :

$$\nabla_X (aY + bZ) = a\nabla_X Y + b\nabla_X Z;$$

iii) ∇ satisface la regla del producto

$$\nabla_X (fY) = f\nabla_X Y + X(f)\, Y,$$

donde $X, Y, Z \in \mathcal{X}(M)$, $a, b \in \mathbb{R}$, y $f, g \in \mathcal{D}(M)$. El campo vectorial $\nabla_X Y$ es llamado la derivada covariante de Y con respecto a X para la conexión ∇. Expresando $\nabla_X Y$ en coordenadas puede ser probado que los valores de $\nabla_X Y(p)$, $p \in M$, dependen solo de los valores de Y en una vecindad de p y de $X(p)$, lo cual garantiza que $\nabla_v Y = \nabla_X Y\mid_p$, donde X es cualquier campo vectorial satisfaciendo $X_p = v$, está bien definido. De igual manera, dado $p \in M$, la función lineal $\mathcal{D}Y(p) : T_pM \longrightarrow T_pM$ dada por $\mathcal{D}Y(p)(v) = \nabla_X Y(p)$, donde X es cualquier campo vectorial tal que $X(p) = v$, está bien definida.

Consideremos ahora una curva $\gamma : [a, b] \longrightarrow M$ y un campo vectorial Y a lo largo de γ. Diremos que Y es paralelo a lo largo de γ si $\nabla_{\gamma'(t)} Y = 0$ para todo t. La conexión afín es compatible con la métrica $\langle \cdot, \cdot \rangle$, cuando para toda curva suave γ y cualquier par de campos vectoriales paralelos P y P' a lo largo de γ, tenemos que $\langle P, P' \rangle$ es constante. Diremos que ∇ es simétrica si

$$\nabla_X Y - \nabla_Y X = [X, Y], \quad \text{para todo} \quad X, Y \in \mathcal{X}(M),$$

donde $[X, Y]$ es el campo vectorial definido por

$$[X, Y]_p f = (X_p Y - Y_p X) f, \quad \text{para todo} \quad f \in \mathcal{D}(M).$$

El Teorema Levi–Civita establece que existe una única conexión afín ∇ sobre M que es simétrica y compatible con la métrica, la cual es conocida como la conexión de Levi-Civita. En adelante solo consideramos dicha conexión. Diremos que una curva parametrizada $\gamma : I \subseteq \mathbb{R} \longrightarrow M$ es una geodésica si el campo vectorial γ' es paralelo a lo largo de γ, caso en el cual $||\gamma'||$ es constante.

Como es conocido, una variedad Riemanniana es completa si para cualquier $p \in M$ toda geodésica partiendo de p está definida para todo $-\infty < t < \infty$. El Teorema de Hopf-Rinow establece que si M es completa entonces cualquier par de puntos en M pueden ser unidos por una geodésica minimal γ, esto es $\ell(\gamma) = d(p, q)$. Más aún, (M, d) es un espacio métrico completo. En lo que sigue, cada variedad será asumida completa.

Si $v \in T_pM$ entonces existe una única geodésica γ tal que $\gamma(0) = p$ y $\gamma'(0) = v$. El punto $\gamma(1)$ es llamado la imagen de v por el mapeo exponencial en p. En consecuencia, la función

$$\exp_p : T_pM \longrightarrow M$$

dada por $\exp_p(v) = \gamma(1)$ está bien definida. No es difícil ver que para cualquier t, $\gamma(t) = \exp_p(tv)$.

Ahora recordamos la noción de transporte paralelo.

Definición 2.2. *Sea $\gamma : \mathbb{R} \longrightarrow M$ una curva suave por partes. El transporte paralelo a lo largo de γ, denotado por $P_{\gamma,\cdot,\cdot}$ es definido por*

$$\begin{array}{rccc} P_{\gamma,a,b}: & T_{\gamma(a)}M & \longrightarrow & T_{\gamma(b)}M \\ & v & \longmapsto & V(\gamma(b)), \end{array}$$

para todo $a, b \in \mathbb{R}$, donde V es el único campo vectorial paralelo a lo largo de γ tal que $V(\gamma(a)) = v$.

Es conocido que $P_{\gamma,a,b}$ es una isometría entre los espacios tangentes $T_{\gamma(a)}M$ y $T_{\gamma(b)}M$. Su inversa es el transporte paralelo a lo largo de γ, recorrido de $\gamma(b)$ a $\gamma(a)$, a saber $P_{\gamma,b,a}$. Además tenemos que $P_{\gamma,b,d} \circ P_{\gamma,a,b} = P_{\gamma,a,d}$ y, si γ es geodésica, $P_{\gamma,a,b}(\gamma'(a)) = \gamma'(b)$.

Dado $X \in \mathcal{X}(M)$, diremos que su derivada covariante $\mathcal{D}X$ es Lipschitz con constante $L > 0$, si para cualquier geodésica γ y $a, b \in \mathbb{R}$, se tiene que

$$\|P_{\gamma,b,a}\mathcal{D}X(\gamma(b))P_{\gamma,a,b} - \mathcal{D}X(\gamma(a))\| \leq L \int_a^b \|\gamma'(t)\|\, dt.$$

En este caso escribimos $\mathcal{D}X \in Lip_L(M)$.

De manera análoga al contexto de Banach, el problema ahora es hallar una solución aproximada de la ecuación

$$X(p) = 0,$$

donde X es un campo vectorial sobre una variedad Riemanniana M. Para ilustración, presentamos a continuación el método de Newton en variedades, ver [29] para mas detalles sobre este método.

Método de Newton-Kantorovich en variedades Riemannianas

El algoritmo para el método de Newton en variedades Riemannianas, propuesto por O. P. Ferreira y B. F. Svaiter en [29], se define por

$$P_{k+1} := \exp_{p_k}(-\mathcal{D}X(p_k)^{-1}X(p_k)), \qquad k = 0, 1, 2, \ldots \tag{2.5}$$

con $p_0 \in M$ dado. El siguiente resultado proporciona una condición suficiente para garantizar la convergencia del método (2.5).

Teorema 2.2. *Sean M una variedad Riemanniana, Ω un subconjunto abierto de M y X un campo vectorial continuo sobre $\overline{\Omega}$, el cual es C^1 en Ω con $\mathcal{D}X \in Lip_L(\Omega)$. Fijemos $p_0 \in \Omega$ y supongamos que $\mathcal{D}X(p_0)$ es no singular y que para algún $a > 0$ y $b \geq 0$,*

$$\|\mathcal{D}X(p_0)^{-1}\| \leq a, \qquad \|\mathcal{D}X(p_0)^{-1}X(p_0)\| \leq b, \qquad l := abL \leq \frac{1}{2}$$

y

$$B(p_0, t_*) \subseteq \Omega,$$

donde $t_* = \frac{1}{aL}(1 - \sqrt{1-2l})$. Entonces la sucesión (p_k) generada por (2.5) con punto inicial p_0 está bien definida, contenida en $B(p_0, t_*)$ y converge a un punto p_*, el cual es el único cero de X en $B[p_0, t_*]$.

Para otros métodos iterativos en variedades Riemannianas, consultar [3, 5, 20].

2.3. Grupos de Lie

Un grupo de Lie G es un grupo dotado de una estructura de variedad diferenciable, para el cual las operaciones del grupo resultan diferenciables. Esto es, las funciones

$$m_G : G \times G \to G \qquad \text{y} \qquad i_G : G \to G$$

dadas por $m_G(x, y) := x \cdot y := xy$ y $i_G(x) := x^{-1}$, son $\mathcal{C}^\infty$. La dimensión del grupo de Lie es por definición la misma de G como variedad diferenciable. Asumiremos en adelante que G es finito dimensional y denotaremos por e el elemento neutro del grupo de Lie G.

Para cada $g \in G$, la traslación a izquierda por g, denotada por L_g, es la función de G en G definida por

$$L_g(x) := gx, \qquad \text{para todo } x \in G.$$

Resaltamos que L_g es un difeomorfismo de G sobre si mismo con inversa $(L_g)^{-1} = L_{g^{-1}}$. Por lo tanto, dado $g \in G$, la diferencial de L_g en e, denotada por $(L'_g)_e$, es un isomorfismo de T_eG sobre T_gG, cuya inversa es $(L'_{g^{-1}})_g$.

Diremos que un campo vectorial $X \in \mathcal{X}(G)$ es invariante a izquierda si para todo $g \in G$,

$$(L'_g)_h(X_h) = X_{gh}, \qquad \text{para todo } h \in G.$$

Por ejemplo, todo $u \in T_eG$ induce un campo vectorial X^u invariante a izquierda dado por

$$X^u(g) := X^u_g = (L'_g)_e(u), \qquad g \in G.$$

Es fácil ver que si X es un campo invariante a izquierda, entonces $X = X^u$, donde $u = X_e$. Esto muestra que existe una correspondencia biyectiva entre los campos invariantes a izquierda y el espacio T_eG. Ya que puede ser probado que el corchete de Lie $[X, Y]$ de dos campos invariantes a izquierda $X, Y \in \mathcal{X}(G)$, es invariante a izquierda, entonces la operación

$$[\cdot, \cdot] : T_eG \times T_eG \to T_eG$$

dada por

$$[u, v] := [X^u, X^v]_e, \qquad \text{para todo } u, v \in T_eG,$$

está bien definida. Además, $[\cdot, \cdot]$ es bilineal y satisface $[u, v] = -[v, u]$, para todo $u, v \in T_eG$, y la identidad de Jacoby

$$[u, [v, w]] + [v, [w, u]] + [w, [u, v]] = \mathbf{0}, \quad \text{para todo} \quad u, v, w \in T_eG.$$

Esto es, T_eG con la operación $[\cdot, \cdot]$ es una álgebra de Lie, la cual en adelante denotaremos por $\mathfrak{g}$. Aquí y en adelante $\mathbf{0}$ denotará el elemento neutro de $\mathfrak{g}$.

Es variada la literatura sobre grupos de Lie y sus respectivas álgebras. Puede consultarse, por ejemplo, en [28, 38, 56, 58]. A modo de ambientación presentamos algunos ejemplos clásicos de grupos de Lie y sus respectivas álgebras.

Ejemplo 2.1. *Grupo lineal general.*

Consideremos la variedad diferenciable $M(n, \mathbb{R})$ de todas las matrices reales de orden n y sea $\mathcal{GL}(n, \mathbb{R})$ el subconjunto de $M(n, \mathbb{R})$ formado por las matrices invertibles de orden n con coeficientes reales. Se puede verificar que $\mathcal{GL}(n, \mathbb{R})$ es abierto en $M(n, \mathbb{R})$ y es un grupo de Lie bajo la multiplicación matricial, el cual es denominado grupo lineal general. Cuando $n = 1$ escribimos $\mathbb{R}^*$. En este caso $\mathfrak{g}$ es isomorfo a $M(n, \mathbb{R})$ con el corchete de Lie dado por el conmutador $[A, B] = AB - BA$.

Ejemplo 2.2. *El grupo especial ortogonal $\mathcal{SO}_n(\mathbb{R})$.*

El conjunto $\mathcal{SO}_n(\mathbb{R})$ de las matrices reales de orden n que son ortogonales y tienen determinante igual a 1 es un subgrupo de Lie de $\mathcal{GL}(n, \mathbb{R})$. Este grupo es conocido como el grupo especial ortogonal y se puede identificar con el grupo de rotaciones del espacio $\mathbb{R}^n$. $\mathcal{SO}_n(\mathbb{R})$ es un grupo de Lie real, compacto y de dimensión $\frac{n(n-1)}{2}$. Su álgebra de Lie es el espacio $\mathfrak{g}$ de matrices reales de orden n antisimétricas, con el corchete de Lie dado por el conmutador.

Ejemplo 2.3. *El grupo circular.*

El círculo $\mathbb{S}^1 = \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2=1\}$ *puede dotarse con una estructura de variedad diferenciable. Al identificarlo con el círculo unidad* $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ *hereda su estructura de grupo. Se tiene en consecuencia que* $\mathbb{T}$ *es un grupo de Lie. El álgebra de Lie de* $\mathbb{T}$ *es* $\mathbb{R}$.

Observación 2.1. *Son de interés otros subgrupos de Lie del grupo lineal general* $\mathcal{GL}(n,\mathbb{R})$. *Por ejemplo, el grupo lineal especial* $\mathcal{SL}(n,\mathbb{R})$ *que consiste en aquellas matrices con determinante igual a uno y cuya álgebra* $\mathfrak{sl}(n,\mathbb{R})$ *está constituida por las matrices reales de orden* n *con traza nula.*

Fijemos ahora un producto interno $\langle\cdot,\cdot\rangle_e$ sobre el álgebra de Lie $\mathfrak{g}$ del grupo de Lie G. Entonces el isomorfismo $(L'_g)_e$ induce un producto interno en el espacio tangente T_gG, el cual es definido de la siguiente manera:

$$\langle v,w\rangle_g = \left\langle \left(L'_{g^{-1}}\right)_g(v), \left(L'_{g^{-1}}\right)_g(w)\right\rangle_e, \quad v,w \in T_gG. \tag{2.6}$$

En particular,

$$\|v\| := \|v\|_g = \left\|(L'_{g^{-1}})_g(v)\right\|_e, \qquad \text{para todo } g \in G, v \in T_gG.$$

El producto interno (2.6) define de manera natural una métrica Riemanniana sobre el grupo de Lie G, la cual convierte a G en una variedad Riemanniana. En adelante omitiremos los subíndices en la norma y en el producto interno, si no hay peligro de confusión. Como en el caso de una variedad Riemanniana, la longitud de arco de una curva suave a trozos $c : [0,1] \to G$, uniendo los puntos $x = c(0)$ y $y = c(1)$, está dada por

$$\ell(c) := \int_0^1 \|c'(t)\|\, dt.$$

La distancia entre dos puntos $x, y \in G$ es definida por $d(x,y) = \operatorname{ínf}_c \ell(c)$, donde el ínfimo involucrado es tomado sobre el conjunto de todas las curvas suaves a trozos $c : [0,1] \to G$ que unen los puntos x, y. De esta manera sobre (G,d) hay una estructura de espacio métrico, la cual asumiremos siempre completa. Como es habitual, definimos la bola abierta de centro $x \in G$ y radio $r > 0$ por $B(x,r) := \{y \in G \mid d(x,y) < r\}$. La adherencia de $A \subseteq G$ será denotada por $\overline{A}$.

2.3.1. La función exponencial de un grupo de Lie

Una de las construcciones mas importantes en la teoría de grupos de Lie, y que será esencial en este trabajo, es la función exponencial. Esta es una extensión de la exponencial de una matriz estudiada en un curso de álgebra. Con el fin de presentar este concepto, recordemos que si X es un campo vectorial sobre una variedad diferenciable M, una curva suave $\gamma : I \subseteq \mathbb{R} \to M$, donde I es un intervalo que contiene a 0, se dice que es una curva integral de X si satisface

$$\gamma'(t) = X(\gamma(t)), \qquad \text{para todo } t \in I.$$

Si no existe una extensión de γ satisfaciendo la ecuación diferencial anterior, decimos que γ es una curva integral máximal de X. El campo vectorial X se dice que es completo si todas sus curvas integrales maximales están definidas en $\mathbb{R}$. Si este es el caso, un resultado de existencia y unidad de ecuaciones diferenciales permite demostrar que para todo $p \in M$, existe una única curva integral $\gamma : \mathbb{R} \to M$ tal que $\gamma(0) = p$.

En el caso especial de un grupo de Lie G, se tiene que todo campo invariante a izquierda $X \in \mathcal{X}(G)$ es completo [38, Proposición 9.2.1]. Así, dado $u \in \mathfrak{g}$, existe una única curva $\gamma := \gamma_u : \mathbb{R} \to G$ tal que

$$\gamma(0) = e \qquad \text{y} \qquad \gamma'(t) = X^u(\gamma(t)) = (L'_{\gamma(t)})(u), \quad t \in \mathbb{R}, \tag{2.7}$$

de donde

$$\gamma'(0) = X^u(e) = u \qquad \text{y} \qquad \|\gamma'(t)\| = \|u\|\,, \quad t \in \mathbb{R}. \tag{2.8}$$

Definimos la función exponencial $\exp : \mathfrak{g} \to G$ por $\exp(u) = \gamma_u(1)$. Una de las propiedades fundamentales de la función exponencial es que describe completamente a la curva integral γ_u, en el sentido que

$$\gamma_u(t) = \exp(tu), \qquad \text{para todo } t \in \mathbb{R}.$$

Otra propiedad a destacar es que exp es un homomorfismo de grupos, esto es,

$$\gamma_u(t+s) = \gamma_u(t)\gamma_u(s), \qquad t, s \in \mathbb{R}.$$

Además, como en el caso de las variedades diferenciables, se puede probar que existe un conjunto abierto $U \subseteq \mathfrak{g}$, con $\mathbf{0} \in U$, tal que la restricción $\exp|_U$ es un difeomorfismo sobre $\exp(U)$. Esto es consecuencia del hecho que $\exp'_{\mathbf{0}}(u) = u$, para todo $u \in \mathfrak{g}$.

Si bien la igualdad $\exp(u+v) = \exp(u)\exp(v)$ no siempre es válida, si lo es en el caso cuando $[u, v] = \mathbf{0}$, lo cual se tiene, por ejemplo, cuando G es un grupo de Lie abeliano.

Observación 2.2. *En el caso general, cuando u, v tienen norma suficientemente pequeña, entonces existe $w \in \mathfrak{g}$ tal que*

$$\exp(u)\exp(v) = \exp(w).$$

Existen diversas expresiones para w, una de ellas, la cual es conocida como fórmula de Baker-Campbell-Hausdorff, está dada por

$$w = u + v + \frac{1}{2}[u, v] + \frac{1}{12}\left([u, [u, v]] + [v, [v, u]]\right) - \frac{1}{24}[v, [u.[u, v]]] + \cdots, \tag{2.9}$$

donde los términos que siguen contienen conmutadores superiores de u, v. Dado que el corchete $[\cdot,\cdot]$ es una forma bilineal, existe una constante M tal que

$$\|[u, v]\| \leq M \|u\| \|v\|, \qquad \text{para todo } u, v \in \mathfrak{g}.$$

Luego, la serie en el segundo miembro de (2.9) converge absolutamente si, por ejemplo, $\|u\|, \|v\| \leq 1/M^2$.

Ejemplo 2.4. *La exponencial matricial*

Consideremos el grupo lineal general $\mathcal{GL}(n, \mathbb{R})$ y su respectiva álgebra de Lie $M(n, \mathbb{R})$. El campo invariante a izquierda inducido por una matriz $A \in M(n, \mathbb{R})$ está dado por

$$X^A(g) = (L'_g)_I(A) = \frac{d}{dt}\Big|_{t=0} L_g(I + tA) = gA,$$

donde I denota la matriz identidad de tamaño n. Es conocido que la única solución del problema de valor inicial

$$\gamma(0) = I \qquad y \qquad \gamma'(t) = X^A(\gamma(t)) = \gamma(t)A,$$

es la curva $\gamma_A : \mathbb{R} \to \mathcal{GL}(n, \mathbb{R})$ definida por

$$\gamma_A(t) = e^{tA} = \sum_{k=0}^{\infty} \frac{1}{k!} t^k A^k.$$

En consecuencia, en este caso la "exponencial de Lie"de A coincide con la función exponencial matricial en A, esto es, $\exp(A) = e^A$.

2.3.2. Diferenciabilidad

Consideramos una función diferenciable $F : G \to \mathfrak{g}$. Entonces la diferencial de F en un punto $g \in G$ es la transformación lineal $F'_g : T_gG \to \mathfrak{g}$ dada por

$$F'_g(v_g) = \frac{d}{dt}\Big|_{t=0} (F \circ \alpha)(t),$$

donde $\alpha : (-\delta, \delta) \to G$ satisface $\alpha(0) = g$ y $\alpha'(0) = v_g$, para algún $\delta > 0$. En particular, considerando la curva $\alpha(t) = g \exp\left[t(L'_{g^{-1}})_g(v_g)\right]$, obtenemos

$$F'_g(v_g) = \frac{d}{dt}\bigg|_{t=0} F\left(g \exp\left[t(L'_{g^{-1}})_g(v_g)\right]\right).$$

Definimos en este contexto la primera derivada de F en g, la cual denotaremos por dF_g, como el operador lineal

$$dF_g := (F \circ L_g)'_e : \mathfrak{g} \to \mathfrak{g}. \tag{2.10}$$

De la discusión anterior se sigue que para todo $v \in \mathfrak{g}$,

$$dF_g(v) = \frac{d}{dt}\bigg|_{t=0} F\left(g \exp(tv)\right) = \frac{d}{dt}\bigg|_{t=0} F\left(c(t)\right), \tag{2.11}$$

donde $c(t) = g \exp(tv)$, $t \in \mathbb{R}$. Para más detalles sobre esta construcción, ver [43].

Observación 2.3. *De la igualdad $c(s+t) = c(s)\exp(tv)$ se sigue que*

$$dF_{c(s)}(v) = \frac{d}{dt}\bigg|_{t=0} F\left(c(t+s)\right) = \frac{d}{dt}\bigg|_{t=s} F\left(c(t)\right),$$

para todo $v \in \mathfrak{g}$. Así, si $\alpha(t) = F(c(t))$, entonces

$$F(c(1)) - F(c(0)) = \int_0^1 \alpha'(t)dt = \int_0^1 dF_{c(t)}(v)dt. \tag{2.12}$$

Este es el contenido de Lemma 1 en [45].

Si asumimos F de clase C^2, tenemos que la función $dF : G \to \mathrm{End}(\mathfrak{g})$, donde $\mathrm{End}(\mathfrak{g})$ denota el espacio de transformaciones lineales de $\mathfrak{g}$ en si mismo, es diferenciable. De manera análoga a la definición de dF, definimos la segunda derivada de F en $g \in G$ como el operador lineal

$$d^2F_g := (dF \circ L_g)'_e : \mathfrak{g} \to \mathrm{End}(\mathfrak{g}).$$

Así, como en Observación 2.3 obtenemos

$$d^2F_{c(s)}(v) = \frac{d}{dt}\bigg|_{t=s} dF_{c(t)}, \tag{2.13}$$

para todo $v \in \mathfrak{g}$ y $c(t) = g\exp(tv)$, $t \in \mathbb{R}$. Si como en el caso de espacios de Banach, adoptamos la notación $d^2F_g(u,w) := [d^2F_g(u)](w)$, para $u, w \in \mathfrak{g}$, se sigue de la Observación 2.3 y (2.13) que

$$\begin{aligned} d^2F_{c(s)}(v,v) &= \frac{d}{dt}\bigg|_{t=s} dF_{c(t)}(v) \\ &= \frac{d}{dt}\bigg|_{t=s} \frac{d}{d\tau}\bigg|_{\tau=t} F(c(\tau)) \\ &= \frac{d^2}{dt^2}\bigg|_{t=s} F(c(t)). \end{aligned}$$

De manera inductiva definimos d^nF_g; la $n-$ésima derivada de $F : G \to \mathfrak{g}$ en un punto $g \in G$. Procediendo como antes, podemos obtener

$$d^nF_{c(s)}(v,\ldots,v) = \frac{d^n}{dt^n}\bigg|_{t=s} F(c(t)), \tag{2.14}$$

para todo $v \in \mathfrak{g}$ y $c(t) = g\exp(tv)$, $t \in \mathbb{R}$.

Capítulo 3

Algunos métodos iterativos en grupos de Lie

En este capítulo abordaremos el problema (2.1) en el caso de una función $F : G \to \mathfrak{g}$, definida desde un grupo de Lie G a su respectiva álgebra de Lie $\mathfrak{g}$. El propósito es presentar métodos iterativos que aproximen a una solución de $F(x) = 0$. Comenzamos con un método tipo secante.

3.1. Un método secante en grupos de Lie

3.1.1. Diferencias divididas sobre grupos de Lie

El método secante es una de las variantes más conocidas del método de Newton, el cual puede aplicarse a una función no necesariamente diferenciable, ya que el método no requiere calcular la derivada en cada iteración. En su lugar se hace uso de una aproximación adecuada para la primera derivada de la función. En el caso unidimensional, dada $f : I \subseteq \mathbb{R} \to \mathbb{R}$, $f'(x)$ es aproximada por una diferencia dividida del tipo

$$\frac{f(x+h) - f(x)}{h},$$

donde h es elegida adecuadamente. Más generalmente, en espacios de Banach, A. Sergeev [55], ver también [37, 53], usó la definición de diferencia dividida de primer orden (2.3) introducida por J. Schröder [54] para extender el método secante a este contexto. Siguiendo estas ideas, en [20] se presenta una definición de diferencias divididas en el contexto de variedades Riemannianas. Usaremos esto como punto de partida con el fin de proveer una definición de diferencias divididas de primer orden para una función F definida desde un grupo de Lie G en su correspondiente álgebra $\mathfrak{g}$.

Con la notación previa, consideremos una curva regular $c : I \subseteq \mathbb{R} \to G$, $s, s+h \in I$ y $F : G \to \mathfrak{g}$ continua. Entonces, existe un único $u \in \mathfrak{g}$ para el cual

$$\left(L'_{c(s)}\right)_e(u) = c'(s). \tag{3.1}$$

Diremos que un operador lineal $\theta : \mathfrak{g} \to \mathfrak{g}$ es una diferencia dividida de primer orden para F en los puntos $c(s)$ y $c(s+h)$, con dirección $c'(s)$, si se satisface la igualdad

$$\theta(u) = \frac{1}{h}(F(c(s+h)) - F(c(s))). \tag{3.2}$$

En lo que sigue, si no hay lugar a confusión, escribimos $\theta := [c(s), c(s+h); F]$.

Observación 3.1. *Note que si consideramos la curva sobre $\mathfrak{g}$ dada por $\alpha(t) = F(c(t))$ y F diferenciable, se tiene que*

$$\begin{aligned}\alpha'(s) &= F'_{c(s)}(c'(s)) \\ &= F'_{c(s)}((L'_{c(s)})_e(u)) \\ &= dF_{c(s)}(u),\end{aligned}$$

lo cual muestra que θ resulta ser una buena aproximación del operador dF definido en (2.10).

Ejemplo 3.1. *Sea $F : G \to \mathfrak{g}$ una función de clase C^1, $g \in G$, y $u \in \mathfrak{g}$. Si $\gamma := \gamma_u$ satisface (2.7), entonces la curva $c(t) = g\exp(tu) = g\gamma(t)$, $t \in \mathbb{R}$, cumple que*

$$c'(t) = (dL_g)_{\gamma(t)}(\gamma'(t)), \qquad t \in \mathbb{R},$$

de donde $c'(0) = (dL_g)_e(\gamma'(0)) = (dL_g)_e(u)$. Así, por (2.12), la transformación lineal $\theta : \mathfrak{g} \to \mathfrak{g}$, definida por

$$\theta(v) = \int_0^1 dF_{c(t)}(v)dt, \tag{3.3}$$

satisface

$$\theta(u) = \int_0^1 dF_{c(t)}(u)dt = F(c(1)) - F(c(0)).$$

Por lo tanto θ es una diferencia dividida de primer orden de F en los puntos $c(0) = g$ y $c(1) = g\exp(u)$, con dirección $c'(0) = (dL_g)_e(u)$.

3.1.2. Un algoritmo tipo secante en grupos de Lie

Estamos listos para presentar el algoritmo tipo secante que permitirá aproximar soluciones a problemas del tipo $F(x) = 0$, en el caso de una función continua del tipo $F : \Omega \subseteq G \to \mathfrak{g}$, donde Ω es un subconjunto abierto del grupo de Lie G.

Dados $p_{-1}, p_0 \in \Omega$, definimos la sucesión secante (p_n), por

$$\begin{aligned} v_n &= -[p_{n-1}, p_n; F]^{-1} F(p_n) \\ p_{n+1} &= p_n \exp(v_n), \end{aligned} \tag{3.4}$$

$n = 0, 1, 2, \ldots$.

3.1.3. Análisis de convergencia semilocal del algoritmo (3.4)

Para el análisis de convergencia del algoritmo secante (3.4), asumiremos que F satisface cierto tipo de condiciones de continuidad tipo Lipschitz; éstas han sido estudiadas en otros escenarios como espacios de Banach y variedades Riemannianas, ver [37] para el caso de espacios de Banach y [20] para el caso de variedades Riemannianas.

Sea $\omega : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ una función no decreciente y continua en cada uno de sus argumentos. Diremos que la aplicación continua $F : \Omega \subseteq G \to \mathfrak{g}$ satisface la $\omega-$condición sobre Ω, si la desigualdad

$$\|[p, q; F] - [x, y; F]\| \leq \omega(d(p, x); d(q, y)) \tag{3.5}$$

se cumple para todo $p, q, x, y \in \Omega$, con $[\,\cdot\,, \cdot\,; F]$ satisfaciendo (3.2).

Asumiremos que F satisface las siguientes condiciones:

1. $L_0 = [p_{-1}, p_0; F]$ es invertible y $\left\|L_0^{-1}\right\| \leq \beta$, para algún $\beta \geq 0$;
2. $\left\|L_0^{-1} F(p_0)\right\| \leq \eta, \quad$ para algún $\quad \eta \geq 0$;
3. la ecuación
$$u = \left(\frac{a(u)b(u)}{1 - c(u)} + a(u) + 1\right)\eta,$$
donde $\alpha = d(p_{-1}, p_0)$ y
$$a(t) = \frac{\beta\omega(\alpha, t)}{1 - \beta\omega(\alpha, t)}, \quad b(t) = \frac{\beta\omega(t, 2t)}{1 - \beta\omega(\alpha + t, t)}, \quad c(t) = \frac{\beta\omega(2t, 2t)}{1 - \beta\omega(\alpha + t, t)},$$

para $t \in \mathbb{R}$, tiene al menos un cero positivo. Al menor de estos ceros lo denotaremos por R.

Teorema 3.1. *Además de las condiciones anteriores, supongamos que la $\omega-$condición (3.5) es satisfecha, y que también*

$$\beta\omega(R+\alpha, R) < 1, \quad c(R) < 1 \quad y \quad \overline{B(p_0, R)} \subset \Omega.$$

Entonces la sucesión secante (3.4) está bien definida, está contenida en $B(p_0, R)$ y converge a la única solución p^* de la ecuación $F(p) = \mathbf{0}$ en $\overline{B(p_0, R)}$.

Demostración. Para probar que la sucesión (p_n) está bien definida y contenida en $B(p_0, R)$, procederemos por inducción. Definamos $L_n := [p_{n-1}, p_n; F]$, con $n \geq 0$, y consideremos la familia de curvas $\{\gamma_n : [0,1] \to G \mid n \geq 0\}$ definidas por $\gamma_n(t) := p_n \exp(tv_n)$. Notemos que, de una manera similar a (2.8),

$$\|\gamma_n'(t)\| = \|v_n\|, \quad \text{para todo} \quad 0 \leq t \leq 1, \tag{3.6}$$

de donde, por definición de (p_n),

$$d(p_n, p_{n+1}) \leq \int_0^1 \|\gamma_n'(t)\| \, dt = \|v_n\|, \quad n = 0, 1, \ldots. \tag{3.7}$$

Por otra parte, la diferencia dividida $[p_n, p_{n+1}; F]$ en los puntos p_n, p_{n+1} con dirección $\gamma_n'(0) = (dL_{p_n})_e(v_n)$, satisface

$$[p_n, p_{n+1}; F](v_n) = F(p_{n+1}) - F(p_n), \quad n = 0, 1, \ldots. \tag{3.8}$$

Se sigue de aquí y (3.4) que

$$\begin{aligned} F(p_{n+1}) &= [p_n, p_{n+1}; F](v_n) - [p_{n-1}, p_n; F](v_n) \\ &= (L_{n+1} - L_n)(v_n), \quad n = 0, 1, \ldots. \end{aligned} \tag{3.9}$$

lo que en virtud de (3.5) implica

$$\|F(p_{n+1})\| \leq \omega(d(p_{n-1}, p_n), d(p_n, p_{n+1})) \|v_n\|, \quad n = 0, 1, \ldots. \tag{3.10}$$

De otro lado, por hipótesis L_0^{-1} existe y por lo tanto, sustituyendo $n = 0$ en (3.4), tenemos que

$$v_0 = -L_0^{-1} F(p_0) \qquad \text{y} \qquad p_1 = p_0 \exp(v_0)$$

están bien definidos. Probemos que $p_1 \in B(p_0, R) \subset \Omega$. Se sigue de (3.7) que

$$d(p_0, p_1) \leq \|v_0\| \leq \eta,$$

siendo la última desigualdad consecuencia de (3.4) y la condición (2). De aquí y la igualdad

$$R = \left(\frac{ba}{1-c} + a + 1 \right) \eta = K\eta, \tag{3.11}$$

donde $a = a(R)$, $b = b(R)$ y $c = c(R)$, podemos concluir que $d(p_0, p_1) \leq \eta < R$. Así $p_1 \in B(p_0, R)$. Consideremos ahora el caso $n = 1$. Por definición de (p_n) tenemos que

$$v_1 = -L_1^{-1} F(p_1) \qquad \text{y} \qquad p_2 = p_1 \exp(v_1).$$

Entonces, para probar que p_2 está bien definido, es suficiente verificar que el operador L_1 es invertible. De hecho, por la $\omega-$condición (3.5) y las propiedades de ω, obtenemos

$$\begin{aligned} \left\| I - L_0^{-1} L_1 \right\| &\leq \left\| L_0^{-1} \right\| \left\| L_0 - L_1 \right\| \\ &\leq \beta\omega(d(p_{-1}, p_0), d(p_0, p_1)) \\ &\leq \beta\omega(\alpha, R) \\ &\leq \beta\omega(R + \alpha, R) \\ &< 1, \end{aligned} \tag{3.12}$$

de lo cual se sigue, por el lema de Banach, que L_1^{-1} existe y además,

$$\left\| L_1^{-1} \right\| \leq \frac{\left\| L_0^{-1} \right\|}{1 - \left\| L_0^{-1} \right\| \left\| L_0 - L_1 \right\|} \leq \frac{\beta}{1 - \beta\omega(\alpha, R)}.$$

Probemos ahora que $d(p_2, p_0) < R$. Por un lado, de (3.4) y (3.7) se tiene que

$$d(p_1, p_2) \leq \|v_1\| \leq \left\| L_1^{-1} \right\| \|F(p_1)\|.$$

Por otro lado, por (3.10),

$$\|F(p_1)\| \leq \omega(d(p_{-1}, p_0), d(p_0, p_1)) \|v_0\| \leq \omega(\alpha, R)\eta$$

y, en consecuencia, $\|v_1\| \leq a \|v_0\|$. Adicionalmente,

$$\begin{aligned} d(p_2, p_0) &\leq d(p_2, p_1) + d(p_1, p_0) \\ &\leq \|v_1\| + \|v_0\| \\ &\leq (a + 1)\eta. \end{aligned}$$

De aquí y (3.11) se sigue $d(p_2, p_0) < R$, y por lo tanto $p_2 \in B(p_0, R)$. Para el caso $n = 2$ primero notamos que de (3.5) y del hecho que p_1 y $p_2 \in B(p_0, R)$, se sigue que

$$\begin{aligned} \left\| I - L_0^{-1} L_2 \right\| &\leq \left\| L_0^{-1} \right\| \left\| L_0 - L_2 \right\| \\ &\leq \beta\omega(d(p_{-1}, p_1), d(p_0, p_2)) \\ &\leq \beta\omega(d(p_{-1}, p_0) + d(p_0, p_1), d(p_0, p_2)) \\ &\leq \beta\omega(\alpha + R, R) < 1, \end{aligned}$$

de donde L_2 es invertible y por consiguiente $p_3 = p_2 \exp(v_2)$, con $v_2 = -L_2^{-1}F(p_2)$, está bien definido. También se concluye que

$$\left\|L_2^{-1}\right\| \leq \frac{\beta}{1 - \beta\omega(\alpha + R, R)}.$$

Es claro que siguiendo el procedimiento anterior, también se puede concluir que para todo $n \geq 2$,

$$\left\|I - L_0^{-1}L_n\right\| \leq \beta\omega(\alpha + R, R) < 1, \qquad \text{si} \quad p_{n-1}, p_n \in B(p_0, R). \tag{3.13}$$

Si lo anterior se tiene, podríamos obtener por el lema de Banach que el operador L_n es invertible, así como también una cota uniforme para la norma $\|L_n^{-1}\|$; la cota solo dependería de β, α y R.

Verifiquemos que $d(p_3, p_0) < R$. En virtud de (3.10) y el hecho que $p_1, p_2 \in B(p_0, R)$, tenemos que

$$\|F(p_2)\| \leq \omega(d(p_0, p_1), d(p_1, p_2)) \|v_1\| \leq \omega(R, 2R) \|v_1\|,$$

lo cual implica que

$$\|v_2\| = \left\|-L_2^{-1}F(p_2)\right\| \leq \frac{\beta\omega(R, 2R)}{1 - \beta\omega(\alpha + R, R)} \|v_1\| = b \|v_1\|.$$

Así, ya que $\|v_1\| \leq a \|v_0\|$, obtenemos

$$\begin{aligned} d(p_3, p_0) &\leq d(p_3, p_2) + d(p_2, p_1) + d(p_1, p_0) \\ &\leq \|v_2\| + \|v_1\| + \|v_0\| \\ &\leq (ab + a + 1) \|v_0\| \\ &\leq (ab + a + 1)\eta < R, \end{aligned}$$

y por lo tanto $p_3 \in B(p_0, R)$. El siguiente paso es probar que las siguientes afirmaciones son válidas para todo $n \geq 3$.

(I) Existe L_n^{-1} y
$$\left\|L_n^{-1}\right\| \leq \frac{\beta}{1 - \beta\omega(R + \alpha, R)};$$

(II) $p_{n+1} \in B(p_0, R)$.

Para el caso $n = 3$, vemos de (3.13) y el comentario que le sigue, que L_3 es invertible y

$$\left\|L_3^{-1}\right\| \leq \frac{\beta}{1 - \beta\omega(R + \alpha, R)}.$$

Ahora, del hecho que $p_1, p_2, p_3 \in B(p_0, R)$ y (3.10), inferimos que

$$\|F(p_3)\| \leq \omega(d(p_1, p_2), d(p_2, p_3)) \|v_2\| \leq \omega(2R, 2R) \|v_2\|,$$

de lo cual se deduce

$$\begin{aligned} \|v_3\| &\leq \left\|L_3^{-1}\right\| \|F(p_3)\| \\ &\leq \frac{\beta\omega(2R, 2R)}{1 - \beta\omega(\alpha + R, R)} ab \|v_0\| \\ &= cab \|v_0\|. \end{aligned}$$

Así,

$$\begin{aligned} d(p_4, p_0) &\leq d(p_4, p_3) + d(p_3, p_2) + d(p_2, p_1) + d(p_1, p_0) \\ &\leq \|v_3\| + \|v_2\| + \|v_1\| + \|v_0\| \\ &\leq ((c + 1)ab + a + 1)\eta \\ &= \left(\frac{1 - c^2}{1 - c} ab + a + 1\right)\eta \\ &< R \end{aligned}$$

y, en consecuencia, $p_4 \in B(p_0, R)$.

Supongamos ahora que (I) y (II) se satisfacen para $k = 1, \ldots, n$ y comprobemos el caso $k = n + 1$, $n > 3$. Procediendo como en el caso $n = 3$, la hipótesis $p_k \in B(p_0, R)$ para $k = 1, \ldots, n + 1$, permite concluir de (3.13) que L_{n+1}^{-1} existe y

$$\left\|L_{n+1}^{-1}\right\| < \frac{\beta}{1 - \beta\omega(\alpha + R, R)}.$$

De otro lado, por hipótesis de inducción y (3.10), tenemos

$$\|F(p_{n+1})\| \leq \omega(2R, 2R) \|v_n\|,$$

de donde

$$\|v_{n+1}\| \leq \left\|L_{n+1}^{-1}\right\| \|F(p_{n+1})\| \leq c \|v_n\| \leq \cdots \leq c^{n-1} \|v_2\| \leq c^{n-1} ab \|v_0\|$$

y por lo tanto

$$d(p_{n+1}, p_{n+2}) \leq \|v_{n+1}\| \leq c^{n-1} ab \|v_0\|. \tag{3.14}$$

Se concluye de aquí que

$$\begin{aligned} d(p_{n+2}, p_0) &\leq d(p_{n+2}, p_{n+1}) + d(p_{n+1}, p_n) + \cdots + d(p_2, p_1) + d(p_1, p_0) \\ &\leq \quad \|v_{n+1}\| + \|v_n\| + \cdots + \|v_1\| + \|v_0\| \\ &\leq ((c^{n-1} + c^{n-2} + \cdots + 1)ab + a + 1)\eta \\ &= \left(\frac{1 - c^n}{1 - c} ab + a + 1 \right) \eta \\ &< R, \end{aligned}$$

o equivalentemente $p_{n+2} \in B(p_0, R)$. Esto completa la prueba de (I) y (II).

El siguiente paso es probar que (p_n) converge. Procediendo como en la prueba de la desigualdad anterior se obtiene,

$$\begin{array}{rcl} d(p_n, p_{n+m}) & \leq & d(p_n, p_{n+1}) + d(p_{n+1}, p_{n+2}) + \cdots + d(p_{n+m-1}, p_{n+m}) \\ & \leq & (1 + c + \cdots + c^{m-1}) \|v_n\| \\ & < & \left(\dfrac{1}{1 - c} \right) c^{n-2} ab \|v_0\|, \end{array}$$

para todo $n, m \geq 3$. En consecuencia, (p_n) es una sucesión de Cauchy en $B(p_0, R)$ y por lo tanto existe $p^* \in \overline{B(p_0, R)}$ tal que (p_n) converge a p^*. Se sigue también de la desigualdad $\|F(p_{n+1})\| \leq \omega(2R, 2R) \|v_n\|$, la continuidad de F y la condición $v_n \to \mathbf{0}$, que $F(p^*) = \mathbf{0}$.

Finalmente, supongamos que existe otra solución $q^* \in \overline{B(p_0, R)}$ de $F(p) = \mathbf{0}$ y consideremos una diferencia dividida $\theta = [q^*, p^*; F]$. Entonces

$$\theta(v^*) = F(p^*) - F(q^*) = \mathbf{0}, \quad \text{para algún} \quad v^* \in \mathfrak{g}.$$

Dado que

$$\begin{array}{rcl} \left\| I - L_0^{-1}\theta \right\| & \leq & \left\| L_0^{-1} \right\| \|L_0 - \theta\| \\ & \leq & \beta\omega(d(p_{-1}, q^*), d(p_0, p^*)) \\ & \leq & \beta\omega(d(p_{-1}, p_0) + d(p_0, q^*), R) \\ & \leq & \beta\omega(\alpha + R, R) < 1, \end{array}$$

existe θ^{-1} y así $v^* = \theta^{-1}(\mathbf{0}) = \mathbf{0}$. Por consiguiente, $p^* = q^*$. □

3.1.4. Orden de convergencia del método de la secante

Probaremos en esta sección que, como en el caso de espacios de Banach, el orden de convergencia del método de la secante estudiado en la sección anterior es $(1 + \sqrt{5})/2$. Comenzamos deduciendo un teorema de Taylor en grupos de Lie.

Observación 3.2. *Como antes, sea G un grupo de Lie con álgebra de Lie $\mathfrak{g}$. Si $F : G \to \mathfrak{g}$ es suficientemente diferenciable, entonces la curva en $\mathfrak{g}$ definida por $\alpha(t) = F(c(t))$, satisface*

$$\alpha(t) = \alpha(0) + t\alpha'(0) + \cdots + \frac{t^n}{n!}\alpha^{(n)}(0) + \int_0^t \frac{(t-s)^n}{n!}\alpha^{n+1}(s)ds.$$

Se concluye en virtud de (2.14), que

$$F(g\exp(tv)) = F(g) + tdF_g(v) + \cdots + \frac{t^n}{n!}d^nF_g(v^n) + \int_0^t \frac{(t-s)^n}{n!}d^{n+1}F_{c(s)}(v^{n+1})ds, \quad (3.15)$$

donde v^k denota la $k-$tupla $(v, \ldots, v)$. Si adoptamos la notación $O_n(w)$ (notación O grande) para representar $O(\|w\|^n)$, con $w \in \mathfrak{g}$, se tiene

$$F(g\exp(tv)) = F(g) + tdF_g(v) + \cdots + \frac{t^n}{n!}d^nF_g(v^n) + O_{n+1}(v), \quad (3.16)$$

para $g \in G$ y $v \in \mathfrak{g}$.

En el resto de la sección, asumiremos F de clase C^3.

Dados $x \in G$ y $v \in \mathfrak{g}$, consideremos la curva $c(t) = x\exp(tv)$, $0 \leq t \leq 1$, y la diferencia dividida $\theta : \mathfrak{g} \to \mathfrak{g}$ del Ejemplo 3.1, la cual está definida por

$$\theta(u) = \int_0^1 dF_{c(t)}(u)dt, \qquad u \in \mathfrak{g}.$$

Procediendo como en la Observación 3.2, pero con las curvas $\alpha(t) = dF_{c(t)}$ y $\beta(t) = d^2F_{c(t)}$, se obtiene

$$dF_{c(t)} = dF_x + td^2F_x(v, \cdot) + \int_0^t (t-s)d^3F_{c(s)}(v, v, \cdot)ds \quad (3.17)$$

y también

$$d^2F_{c(t)}(\cdot, \cdot) = d^2F_x(\cdot, \cdot) + \int_0^t d^3F_{c(s)}(v, \cdot, \cdot)ds. \quad (3.18)$$

Por lo tanto, por definición de θ, se infiere que

$$\theta(\cdot) = dF_x(\cdot) + \frac{1}{2}d^2F_x(v, \cdot) + O_2(v). \quad (3.19)$$

Definición 3.1. *Sea (x_n) una sucesión en un grupo de Lie G que converge a $x^* \in G$. Para n suficientemente grande, escribamos $x_n = x^* \exp(u_n)$, con $u_n \in \mathfrak{g}$. Diremos que (x_n) converge a x^* con orden al menos $\alpha > 0$, si*

$$\|u_{n+1}\| \leq \lambda \|u_n\|^{\alpha},$$

para alguna constante positiva λ y n suficientemente grande.

Teorema 3.2. *Suponga que la sucesión (p_n) definida en (3.4) converge a p^* y dF_{p^*} es invertible. Entonces el orden de convergencia del método es $(1+\sqrt{5})/2$.*

Demostración. Para n suficientemente grande, existe $u_n \in \mathfrak{g}$ tal que

$$p_n = p^* \exp(u_n) \qquad \text{y} \qquad d(p^*, p_n) \leq \|u_n\| = \ell(\sigma_n),$$

donde $\sigma_n(t) = p^* \exp(tu_n)$, $0 \leq t \leq 1$. Consideremos también las curvas $\gamma_n(t) = p_n \exp(tv_n)$, $0 \leq t \leq 1$, y la diferencia dividida

$$L_n(w) := [p_{n-1}, p_n; F](w) = \int_0^1 dF_{\gamma_{n-1}(t)}(w)dt, \quad w \in \mathfrak{g}.$$

Así, de (3.19) se sigue que

$$L_n = dF_{p_{n-1}} + \frac{1}{2} d^2 F_{p_{n-1}}(v_{n-1}, \cdot) + O_2(v_{n-1}) \tag{3.20}$$

Ahora, por (3.17),

$$\begin{aligned} dF_{p_{n-1}} &= dF_{p^*} + d^2 F_{p^*}(u_{n-1}, \cdot) + O_2(u_{n-1}) \\ &= dF_{p^*}[I + L_*(u_{n-1}, \cdot) + O_2(u_{n-1})], \end{aligned} \tag{3.21}$$

donde $L_* = dF_{p^*}^{-1} d^2 F_{p^*}$. Similarmente, (3.18) implica

$$d^2 F_{p_{n-1}}(v_{n-1}, \cdot) = d^2 F_{p^*}(v_{n-1}, \cdot) + O_2(u_{n-1}, v_{n-1}),$$

con $O_2(u_{n-1}, v_{n-1}) = O(\|u_{n-1}\| \|v_{n-1}\|)$. De aquí, (3.20) y (3.21), se concluye que

$$\begin{aligned} L_n = dF_{p^*}[I + L_*(u_{n-1}, \cdot) + O_2(u_{n-1})] + \\ + \frac{1}{2}[d^2 F_{p^*}(v_{n-1}, \cdot) + O_2(u_{n-1}, v_{n-1})] + O_2(v_{n-1}), \end{aligned}$$

de donde

$$L_n = dF_{p^*}\left[I + \frac{1}{2} L_*(2u_{n-1} + v_{n-1}, \cdot) + \tilde{O}_2(u_{n-1}, v_{n-1})\right]$$

y, en consecuencia,

$$L_n^{-1} = \left[I - \frac{1}{2} L_*(2u_{n-1} + v_{n-1}, \cdot) + \tilde{O}_2(u_{n-1}, v_{n-1}) \right] dF_{p^*}^{-1}. \tag{3.22}$$

Aquí, $\tilde{O}_2(u_{n-1}, v_{n-1}) = O(\|u_{n-1}\|^2) + O(\|u_{n-1}\| \, \|v_{n-1}\|) + O(\|v_{n-1}\|^2)$. De otro lado, como

$$p_{n+1} = p_n \exp(v_n) = p^* \exp(u_n) \exp(v_n) \quad \text{y} \quad p_{n+1} = p^* \exp(u_{n+1}),$$

entonces,

$$\exp(u_n) \exp(v_n) = \exp(u_{n+1}) \qquad \text{para todo } n.$$

Se sigue de la Observación 2.2 que, para valores de n suficientemente grande,

$$u_{n+1} = u_n + v_n + w_n,$$

donde

$$\begin{aligned} w_n &= \frac{1}{2}[u_n, v_n] + \frac{1}{12}\left([u_n, [u_n, v_n]] + [v_n, [v_n, u_n]]\right) - \frac{1}{24}[v_n, [u_n.[u_n, v_n]]] + \cdots \\ &= O_2(u_n, v_n) + o(\|u_n\| \, \|v_n\|). \end{aligned} \tag{3.23}$$

Así, para tales valores de n, se sigue de (3.4) que

$$u_{n+1} = u_n - L_n^{-1} F(p_n) + w_n,$$

y por (3.22) llegamos a la igualdad

$$u_{n+1} = u_n - \left\{ \left[I - \frac{1}{2} L_*(2u_{n-1} + v_{n-1}, \cdot) + \tilde{O}_2(u_{n-1}, v_{n-1}) \right] dF_{p^*}^{-1} \right\} F(p_n) + w_n.$$

Ahora, aplicando (3.16) a la curva σ_n, se concluye que

$$\begin{aligned} F(p_n) &= F(p^*) + dF_{p^*}(u_n) + \frac{1}{2} d^2 F_{p^*}(u_n, u_n) + O_3(u_n) \\ &= dF_{p^*} \left[u_n + \frac{1}{2} dF_{p^*}^{-1} d^2 F_{p^*}(u_n, u_n) + O_3(u_n) \right] \\ &= dF_{p^*} \left[u_n + \frac{1}{2} L_*(u_n, u_n) + O_3(u_n) \right]. \end{aligned}$$

En consecuencia, y teniendo en cuenta que $u_n = u_{n-1} + v_{n-1} + w_{n-1}$, obtenemos por bilinealidad de L_* que

$$\begin{aligned} u_{n+1} = u_n - &\left[I - \frac{1}{2} L_*(u_{n-1} + u_n, \cdot) + \frac{1}{2} L_*(w_{n-1}, \cdot) + \tilde{O}_2(u_{n-1}, v_{n-1}) \right] \\ &\qquad \left[u_n + \frac{1}{2} L_*(u_n, u_n) + O_3(u_n) \right] + w_n. \end{aligned}$$

Se obtiene de aquí y (3.23) que

$$u_{n+1} = \frac{1}{2} L_*(u_{n-1}, u_n) + o(\|u_n\| \, \|u_{n-1}\|) + O_3(u_n), \tag{3.24}$$

para n suficientemente grande.

Para finalizar la prueba, usamos un procedimiento estándar. Supongamos a priori que

$$\lim_{n\to\infty} \frac{\|u_{n+1}\|}{\|u_n\|^{\alpha}} = c, \quad \text{para algunas constantes } c > 0 \quad \text{y} \quad 1 \leq \alpha.$$

Entonces también se tendría

$$\lim_{n\to\infty} \frac{\|u_n\|^{\alpha}}{\|u_{n+1}\|} = \frac{1}{c} \qquad \text{y} \qquad \lim_{n\to\infty} \frac{\|u_{n-1}\|}{\|u_n\|^{1/\alpha}} = \frac{1}{c^{1/\alpha}}.$$

De otro lado, por (3.24),

$$\|u_{n+1}\| \leq C \|u_n\| \, \|u_{n-1}\| + o(\|u_n\| \, \|u_{n-1}\|) + O_3(u_n),$$

donde $C = \|L_*\|/2$. Así,

$$\begin{aligned} \|u_n\|^{\alpha-1-1/\alpha} \leq C \frac{\|u_n\|^{\alpha}}{\|u_{n+1}\|} \frac{\|u_{n-1}\|}{\|u_n\|^{1/\alpha}} + \frac{\|u_n\|^{\alpha}}{\|u_{n+1}\|} \frac{\|u_{n-1}\|}{\|u_n\|^{1/\alpha}} \frac{o(\|u_n\| \, \|u_{n-1}\|)}{\|u_n\| \, \|u_{n-1}\|} + \\ + \frac{\|u_n\|^{\alpha}}{\|u_{n+1}\|} \frac{O_3(u_n)}{\|u_n\|^{1+1/\alpha}}. \end{aligned}$$

Como el segundo miembro de la desigualdad anterior tiene límite positivo, necesariamente debemos tener que $\alpha - 1 - 1/\alpha = 0$, lo cual implica $\alpha = (1+\sqrt{5})/2$. Esto completa la prueba. □

3.2. El método Newton Kantorovich

Como hemos descrito en la introducción, uno de los métodos mas estudiados en diferentes contextos es el método de Newton. En el caso de grupos de Lie, se ha realizado un análisis de convergencia para el método de Newton usando diferentes hipótesis, como puede ser visto en [8, 9, 43, 46, 48, 39].

En esta sección presentamos un análisis de convergencia semilocal del método de Newton usando la técnica de relaciones de recurrencia de Kantorovich.

Teorema 3.3. *Sea Ω un subconjunto abierto y convexo de G y $F : \Omega \subset G \to \mathfrak{g}$ de clase C^2. Supongamos que se cumplen las siguientes condiciones:*

1. *para $x_0 \in \Omega$, existe $dF_{x_0}^{-1} \in End(\mathfrak{g})$ y $\beta > 0$ tal que $\left\|dF_{x_0}^{-1}\right\| \leq \beta$;*
2. *existe $\eta > 0$ tal que $\left\|dF_{x_0}^{-1}F(x_0)\right\| \leq \eta$;*
3. *existe $R > 0$ y $M > 0$ que satisfacen $\|dF_x^2\| \leq M$, para todo $x \in B(x_0, R)$;*
4. *si $h := \beta\eta M \leq \frac{1}{2}$, tenemos que $B(x_0, \rho^*) \subseteq B(x_0, R)$, donde*
$$\rho^* = \frac{1 - \sqrt{1 - 2h}}{h}\eta.$$

Entonces la sucesión (x_n), con punto inicial $x_0 \in \Omega$, definida por

$$x_{n+1} := x_n \exp[-dF_{x_n}^{-1}(F(x_n))], \tag{3.25}$$

converge a una solución x^* de la ecuación $F(x) = \mathbf{0}$ y $x^*, x_n \in B(x_0, \rho^*)$ para todo n. Además, si $h < \frac{1}{2}$ la solución es única en $B(x_0, \rho^{**}) \cap B(x_0, R)$, donde $\rho^{**} = \frac{1+\sqrt{1-2h}}{h}\eta$. En el caso $h = 1/2$, x^* es la única solución en $B(x_0, \rho^*)$. También,

$$d(x_n, x^*) \leq \frac{1}{2^{n-1}}(2h)^{2^n-1}\eta, \tag{3.26}$$

para todo $n = 0, 1, 2, \ldots$.

Demostración. Notemos primero que por ítem 1, $x_1 := x_0 \exp[-dF_{x_0}^{-1}(F(x_0))]$ está bien definido. Además, definiendo

$$\beta_0 := \beta, \qquad \eta_0 := \eta \quad \text{y} \quad h_0 := h,$$

se sigue de ítem 2 que $\left\|dF_{x_0}^{-1}F(x_0)\right\| \leq \eta_0$. Ahora, considerando la curva

$$c_0(t) = x_0 \exp(tu_0), \quad 0 \leq t \leq 1, \quad \text{con } u_0 = -dF_{x_0}^{-1}F(x_0),$$

la cual tiene extremos x_0 y x_1, obtenemos que

$$d(x_1, x_0) \leq \left\|dF_{x_0}^{-1}F(x_0)\right\| \leq \eta_0 = \eta < \rho^*.$$

La última desigualdad es una consecuencia de la definición de ρ^* y la desigualdad

$$\frac{1 - \sqrt{1 - 2h_0}}{h_0} > 1.$$

De esto concluimos que $x_1 \in B(x_0, \rho^*)$.

El siguiente paso es probar que x_2 está bien definido, para lo cual es suficiente garantizar que dF_{x_1} es invertible. Tenemos por (3.17) que

$$dF_{x_1}(\cdot) - dF_{x_0}(\cdot) = \int_0^1 d^2F_{c(t)}(u_0, \cdot)dt,$$

de donde se sigue que

$$\|dF_{x_1} - dF_{x_0}\| \leq \int_0^1 \left\|d^2F_{c(t)}\right\| \|u_0\| \leq M\eta_0.$$

Así, por hipótesis tenemos que

$$\left\|I - dF_{x_0}^{-1}dF_{x_1}\right\| \leq \left\|dF_{x_0}^{-1}\right\| \|dF_{x_1} - dF_{x_0}\| \leq \beta_0 M\eta_0 \leq 1/2 < 1.$$

El lema de Banach implica que dF_{x_1} es invertible y por lo tanto x_2 está bien definido.

Ahora probaremos que se cumplen las mismas hipótesis del teorema, pero reemplazando x_1 por x_0 y constantes por definir β_1, η_1 y h_1, en lugar de β_0, η_0 y h_0, respectivamente.

Aplicando el lema de Banach en la desigualdad anterior también se sigue que si $\Gamma_0 = dF_{x_0}^{-1}$ y $\Gamma_1 = dF_{x_1}^{-1}$, entonces

$$\|\Gamma_1\| \leq \frac{\|\Gamma_0\|}{1 - \|I - \Gamma_0 dF_{x_1}\|} = \frac{\beta_0}{1 - h_0} := \beta_1,$$

ya que $h_0 = \beta_0 M\eta_0$. De otro lado, por definición de u_0 y la fórmula de Taylor (3.15), tenemos que

$$F(x_1) = F(x_0) + dF_{x_0}(u_0) + \int_0^1 (1-s)d^2F_{c(s)}(u_0, u_0)ds = \int_0^1 (1-s)d^2F_{c(s)}(u_0, u_0)ds,$$

de donde se concluye

$$\|F(x_1)\| = \left\|\int_0^1 (1-s)d^2F_{c(s)}(u_0, u_0)ds\right\| \leq \frac{1}{2}M\|u_0\|^2 \leq \frac{1}{2}M\eta_0^2.$$

Se sigue de aquí, la cota obtenida para $\|\Gamma_1\|$ y la igualdad $h_0 = h = \beta_0\eta_0 M$, que

$$\|\Gamma_1 F(x_1)\| \leq \|\Gamma_1\| \|F(x_1)\| \leq \frac{\beta_0 M\eta_0^2}{2(1-h_0)} = \frac{h_0\eta_0}{2(1-h_0)} := \eta_1.$$

Notar que si consideramos la curva

$$c_1(t) = x_1 \exp(tu_1), \quad 0 \leq t \leq 1, \quad \text{con } u_1 = -\Gamma_1 F(x_1),$$

se concluye del estimativo anterior que

$$d(x_2, x_1) \leq \ell(c_1) = \|u_1\| \leq \eta_1 \leq h_0\eta_0 \leq \eta_0/2,$$

ya que la condición $h_0 = \beta_0\eta_0 M \leq 1/2$ implica $2(1-h_0) \geq 1$. Concluimos también que

$$h_1 := \beta_1\eta_1 M = \frac{\beta_0}{1-h_0}\frac{h_0\eta_0}{2(1-h_0)}M = \frac{h_0^2}{2(1-h_0)^2} \leq \frac{1}{2}.$$

Esto muestra que se satisfacen las condiciones del teorema en x_1, reemplazando β_1, η_1 y h_1 por β_0, η_0 y h_0, respectivamente.

Ahora definamos

$$\rho_1 = \frac{1-\sqrt{1-2h_1}}{h_1}\eta_1.$$

De la igualdad

$$\sqrt{1-2h_1} = \left(1 - \frac{h_0^2}{(1-h_0)^2}\right)^{1/2} = \frac{\sqrt{1-2h_0}}{1-h_0}$$

y la definición de h_1, η_1, es fácil obtener

$$\rho_1 = \left(\frac{(1-h_0)-\sqrt{1-2h_0}}{h_0}\right)\eta_0 = \left(\frac{1-\sqrt{1-2h_0}}{h_0} - 1\right)\eta_0 = \rho^* - \eta_0.$$

Así,

$$\begin{aligned} x \in B(x_1, \rho_1) &\Rightarrow d(x, x_0) \leq d(x, x_1) + d(x_1, x_0) < \rho_1 + \eta_0 \\ &\Rightarrow d(x, x_0) < \rho^* \\ &\Rightarrow x \in B(x_0, \rho^*), \end{aligned}$$

de donde

$$B(x_1, \rho_1) \subseteq B(x_0, \rho^*) \subseteq B(x_0, R).$$

Notar que como $d(x_1, x_2) \leq \eta_1 \leq \rho_1$, podemos concluir que $x_2 \in \overline{B(x_0, \rho^*)} \subseteq \overline{B(x_0, R)} \subseteq \Omega$.

Supongamos que se ha probado la buena definición de x_{k+1}, $x_{k+1} \in \overline{B(x_0, \rho^*)}$, con $0 \leq k \leq n$, y se han definido las constantes correspondientes

$$\beta_k := \frac{\beta_{k-1}}{1-h_{k-1}}, \qquad \eta_k := \frac{h_{k-1}\eta_{k-1}}{2(1-h_{k-1})}, \qquad h_k := \beta_k\eta_k M$$

y el radio

$$\rho_k = \frac{1-\sqrt{1-2h_k}}{h_k}\eta_k, \qquad 1 \le k \le n.$$

Asumiremos que $\Gamma_k = dF_{x_k}^{-1}$ existe y

$$\|\Gamma_k\| \le \beta_k = \frac{\beta_{k-1}}{1-h_{k-1}}, \qquad 1 \le k \le n, \tag{3.27}$$

y además

$$\|F(x_k)\| \le \frac{1}{2}M\eta_{k-1}^2 \quad \text{y} \quad B(x_k,\rho_k) \subseteq B(x_0,\rho^*), \quad 1 \le k \le n. \tag{3.28}$$

De manera similar a lo deducido para h_1, tenemos

$$h_k = \beta_k\eta_k M = \frac{\beta_{k-1}}{1-h_{k-1}}\frac{h_{k-1}\eta_{k-1}}{2(1-h_{k-1})}M = \frac{h_{k-1}^2}{2(1-h_{k-1})^2}, \qquad 1 \le k \le n.$$

En particular, se sigue de aquí y del estimativo $h_1 \le 1/2$, que

$$h_2 \le \frac{h_1^2}{2(1-h_1)^2} \le 2h_1^2 \le \frac{1}{2}.$$

Siguiendo este proceso es posible concluir que

$$h_k \le 2h_{k-1}^2 \le \frac{1}{2}, \qquad 1 \le k \le n. \tag{3.29}$$

Ahora consideremos las curvas

$$c_k(t) = x_k\exp(tu_k), \quad 0 \le t \le 1, \quad \text{con } u_k = -\Gamma_k F(x_k), \qquad 1 \le k \le n,$$

las cuales tienen extremos en x_k y x_{k+1}, respectivamente. Entonces

$$d(x_{k+1},x_k) \le \ell(c_k) = \|u_k\| \le \|\Gamma_k\|\,\|F(x_k)\|, \qquad 1 \le k \le n.$$

Si sigue de las hipótesis inductivas que

$$d(x_{k+1},x_k) \le \|u_k\| \le \frac{\beta_{k-1}}{1-h_{k-1}}\frac{1}{2}M\eta_{k-1}^2 = \frac{h_{k-1}\eta_{k-1}}{2(1-h_{k-1})} = \eta_k, \qquad 1 \le k \le n. \tag{3.30}$$

Probemos que x_{n+2} está bien definido, lo cual sigue de la existencia de $\Gamma_{n+1} := dF_{x_{n+1}}^{-1}$. Para ver esto, primero observemos que de (3.17) se sigue que

$$dF_{x_{n+1}} - dF_{x_n} = \int_0^1 d^2F_{c_n(t)}(u_n,\cdot)dt,$$

lo cual implica

$$\begin{aligned}\left\|I-\Gamma_n dF_{x_{n+1}}\right\| &\leq \|\Gamma_n\|\left\|dF_{x_n}-dF_{x_{n+1}}\right\| \\ &\leq \|\Gamma_n\|\left\|d^2F_{c_n(t)}\right\|\|u_n\| \\ &\leq \beta_n M\eta_n \\ &= h_n \leq 1/2.\end{aligned}$$

Del lema de Banach se concluye que Γ_{n+1} existe y x_{n+2} está bien definido. Además tenemos el estimativo

$$\|\Gamma_{n+1}\| \leq \frac{\|\Gamma_n\|}{1-\left\|I-\Gamma_n dF_{x_{n+1}}\right\|} \leq \frac{\beta_n}{1-h_n} := \beta_{n+1},$$

lo cual es (3.27) para $k=n+1$. De otro lado, de la fórmula de Taylor (3.15) uno tiene que

$$\begin{aligned}F(x_{n+1}) &= F(x_n)+dF_{x_n}(u_n)+\int_0^1(1-s)d^2F_{c_n(s)}(u_n,u_n)ds \\ &= \int_0^1(1-s)d^2F_{c_n(s)}(u_n,u_n)ds,\end{aligned}$$

siendo la última igualdad una consecuencia de $u_n=\Gamma_n F(x_n)$. De aquí obtenemos

$$\|F(x_{n+1})\| \leq \frac{1}{2}M\|u_n\|^2 \leq \frac{1}{2}M\eta_n^2.$$

Esto es (3.28) para $k=n+1$. De lo anterior, considerando la curva

$$c_{n+1}(t)=x_{n+1}\exp(tu_{n+1}), \quad 0\leq t\leq 1, \quad \text{con } u_{n+1}=-\Gamma_{n+1}F(x_{n+1}),$$

se concluye que

$$d(x_{n+1},x_{n+2}) \leq \|u_{n+1}\| \leq \|\Gamma_{n+1}\|\|F(x_{n+1})\| \leq \frac{M\beta_{n+1}\eta_n^2}{2} = \frac{h_n\eta_n}{2(1-h_n)} := \eta_{n+1},$$

y por lo tanto

$$h_{n+1} := \beta_{n+1}\eta_{n+1}M = \frac{h_n^2}{2(1-h_n)^2} \leq 2h_n^2 \leq \frac{1}{2}.$$

Usando reiteradamente (3.29), se obtiene

$$h_k \leq 2h_{k-1}^2 \leq 2\cdot 2^2 h_{k-2}^4 \leq \cdots \leq \frac{1}{2}(2h_0)^{2^k}, \quad 1\leq k\leq n+1,$$

y consecuentemente

$$\eta_{k+1} \leq h_k \eta_k \leq h_k h_{k-1} \eta_{k-1} \leq \cdots \leq h_k h_{k-1} \cdots h_1 h_0 \eta_0.$$

Así,

$$\begin{aligned} \eta_{k+1} &\leq \left[\frac{1}{2}(2h_0)^{2^k}\right]\left[\frac{1}{2}(2h_0)^{2^{k-1}}\right]\cdots\left[\frac{1}{2}(2h_0)^{2}\right] h_0\eta_0 \\ &\leq \frac{1}{2^{k+1}}(2h_0)^{2^{k+1}-1}\eta_0, \end{aligned}$$

para todo $1 \leq k \leq n$. Resta probar que $x_{n+2} \in B(x_0, \rho^*)$. De la fórmula

$$h_{k+1} = \frac{h_k^2}{2(1-h_k)^2}, \qquad 0 \leq k \leq n,$$

se sigue que para estos valores de k,

$$\sqrt{1-2h_{k+1}} = \sqrt{1 - \frac{h_k^2}{(1-h_k)^2}} = \frac{\sqrt{1-2h_k}}{1-h_k},$$

de donde

$$\begin{aligned} \frac{1-\sqrt{1-2h_{k+1}}}{h_{k+1}} &= \frac{1-h_k-\sqrt{1-2h_k}}{h_k}\frac{h_k}{h_{k+1}(1-h_k)} \\ &= \frac{1-h_k-\sqrt{1-2h_k}}{h_k}\frac{2(1-h_k)}{h_k} \\ &= \left(\frac{1-\sqrt{1-2h_k}}{h_k}-1\right)\frac{2(1-h_k)}{h_k}. \end{aligned}$$

Por definición de η_{k+1} se concluye que

$$\rho_{k+1} = \frac{1-\sqrt{1-2h_{k+1}}}{h_{k+1}}\eta_{k+1} = \left(\frac{1-\sqrt{1-2h_k}}{h_k}-1\right)\eta_k = \rho_k - \eta_k,$$

para todo $0 \leq k \leq n$. Luego,

$$\begin{aligned} x \in B(x_{n+1}, \rho_{n+1}) &\Rightarrow d(x, x_n) \leq d(x, x_{n+1}) + d(x_{n+1}, x_n) \\ &\Rightarrow d(x, x_n) \leq \rho_{n+1} + \eta_n = \rho_n \\ &\Rightarrow x \in B(x_n, \rho_n), \end{aligned}$$

lo cual implica, en virtud de la hipótesis inductiva (3.28), que

$$B(x_{n+1}, \rho_{n+1}) \subseteq B(x_n, \rho_n) \subseteq B(x_0, \rho^*).$$

Como consecuencia, $x_{n+2} \in \overline{B(x_0, \rho^*)}$, ya que $d(x_{n+2}, x_{n+1}) \leq \eta_{n+1} \leq \rho_{n+1}$. Por inducción matemática se concluye que (x_n) está bien definida, $(x_n) \subseteq \overline{B(x_0, \rho^*)}$ y los estimativos (3.27), (3.28), (3.29) y (3.30) se cumplen para todo entero positivo n. También se concluye que

$$\rho_{n+1} = \rho_n - \eta_n \qquad \text{y} \qquad \eta_n \leq \frac{1}{2^n}(2h_0)^{2^n - 1}\eta_0, \qquad \text{para todo } n. \tag{3.31}$$

Finalmente probemos que la sucesión (x_n) es de Cauchy. Primero notemos que por definición de ρ_n y la condición $0 \leq h_n \leq 1/2$, se tiene que $\eta_n \leq \rho_n \leq 2\eta_n$, para todo entero positivo n.

Dados enteros positivo n, m, por (3.31) tenemos

$$\begin{aligned}
d(x_n, x_{n+m}) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{n+m-1}, x_{n+m}) \\
&\leq \eta_n + \eta_{n+1} + \cdots + \eta_{n+m-1} \\
&\leq (\rho_n - \rho_{n+1}) + (\rho_{n+1} - \rho_{n+2}) + \cdots + (\rho_{n+m-1} - \rho_{n+m}) \\
&\leq \rho_n - \rho_{n+m}.
\end{aligned}$$

De aquí y (3.31) obtenemos

$$d(x_n, x_{n+m}) \leq \rho_n \leq 2\eta_n \leq 2\frac{1}{2^n}(2h_0)^{2^n-1}\eta_0 = \frac{1}{2^{n-1}}(2h_0)^{2^n-1}\eta_0,$$

de donde se concluye que $(x_n) \subseteq \overline{B(x_0, \rho^*)}$ es de Cauchy y por lo tanto existe $x^* \in \overline{B(x_0, \rho^*)}$ tal que $x_n \to x^*$ cuando $n \to \infty$. En particular, tomando límite en la desigualdad anterior cuando $m \to \infty$, se concluye (3.26).

Ahora es claro que del estimativo (3.28) para $\|F(x_n)\|$ y la continuidad de F, se concluye que $F(x^*) = \mathbf{0}$. $\square$

3.3. Método de King-Werner en Grupos de Lie

Esta sección está dedicada a realizar un análisis de convergencia semilocal para una versión en grupos de Lie del método de King-Werner, ver (2.2) para la versión en espacios de Banach del método de King-Werner. Asumiremos que G es un grupo de Lie abeliano con álgebra de Lie $\mathfrak{g}$ y $F : G \to \mathfrak{g}$ es de clase C^2. El método es definido como sigue:

Sea $x_0 \in G$ tal que $dF_{x_0} : \mathfrak{g} \to \mathfrak{g}$ es invertible y defina

$$\begin{aligned} \tilde{v}_0 &= -dF_{x_0}^{-1}(F(x_0)); \\ z_0 &= x_0 \exp(\tilde{v}_0); \\ g_0 &= x_0 \exp(\tilde{v}_0/2); \\ v_0 &= -dF_{g_0}^{-1}(F(x_0)); \\ x_1 &= x_0 \exp(v_0). \end{aligned} \tag{3.32}$$

Para $n \geq 1$, defina

$$\begin{aligned} \tilde{v}_n &= -dF_{g_{n-1}}^{-1}(F(x_n)); \\ z_n &= x_n \exp(\tilde{v}_n); \\ g_n &= x_n \exp(\tilde{v}_n/2); \\ v_n &= -dF_{g_n}^{-1}(F(x_n)); \\ x_{n+1} &= x_n \exp(v_n). \end{aligned} \tag{3.33}$$

Observación 3.3. *De las igualdades anteriores, la fórmula* $[\exp(v)]^{-1} = \exp(-v)$, $v \in \mathfrak{g}$, *y teniendo en cuenta que* G *es abeliano, se obtiene que*

$$\begin{aligned} x_{n+1} &= x_n \exp(v_n) \\ &= g_n \exp(-\tilde{v}_n/2) \exp(v_n) \\ &= g_n \exp(v_n - \tilde{v}_n/2) \\ &= g_n \exp(w_n), \end{aligned}$$

donde $w_n = v_n - \tilde{v}_n/2$, $n \geq 0$. *Así, la curva* $c_n(t) = g_n \exp(tw_n)$, $0 \leq t \leq 1$, *satisface* $c_n(0) = g_n$ *y* $c_n(1) = x_{n+1}$.

Similarmente, la curva $\beta_n(t) = x_n \exp(t\tilde{v}_n/2)$, $0 \leq t \leq 1$, *satisface* $\beta_n(0) = x_n$ *y* $\beta_n(1) = g_n$. *Además,*

$$x_{n+1} = x_n \exp(v_n) = z_n \exp(-\tilde{v}_n) \exp(v_n) = z_n \exp(u_n),$$

donde $u_n = v_n - \tilde{v}_n$. *y la curva* $\alpha_n(t) = g_n \exp(-t\tilde{v}_n/2)$, $0 \leq t \leq 1$, *satisface* $\alpha_n(0) = g_n$ *y* $\alpha_n(1) = x_n$.

3.3.1. Análisis de convergencia semilocal del método de King-Werner

Asumiremos que se satisfacen las siguientes condiciones, las cuales son similares a las consideradas en el contexto de espacios de Banach. Las constantes η, M y K son asumidas positivas.

C1 $\left\| dF_{x_0}^{-1}(F(x_0)) \right\| = \|\tilde{v}_0\| \leq \eta$;

C2 $\left\| dF_{x_0}^{-1}\, d^2F_{x_0} \right\| \leq M$;

C3 $\left\| dF_{x_0}^{-1}(d^2F_x - d^2F_y) \right\| \leq Kd(x,y)$, para todo $x, y \in G$.

Además supongamos que

$$6\eta M^3 + 9\eta^2 K^2 + 18\eta MK - 3M^2 - 8K \leq 0. \tag{3.34}$$

Esta condición es estándar cuando se usa el conocido principio del mayorante de Kantorovich, ver [27, 62] para mas detalles. La condición (3.34) garantiza que el polinomio

$$g(t) = \frac{K}{6}t^3 + \frac{M}{2}t^2 - t + \eta$$

tiene una raíz negativa y dos raíces positivas $0 < t^* \leq t^{**}$. Este resultado puede ser hallado en [62], donde se usa una técnica similar para probar la convergencia del método de Newton en espacios de Banach bajo condiciones tipo Lipschitz sobre la derivada de Fréchet del operador involucrado. En ese mismo artículo está probado que el polinomio g satisface las siguientes propiedades:

1. g es decreciente en el intervalo $[0, t^*]$;
2. $g > 0$ en $[0, t^*)$;
3. $g'' \geq 0$ en $[0, t^*]$;
4. g'' es creciente en $[0, t^*]$.

Lema 3.1. *Sean $(s_n), (t_n)$ las sucesiones reales definidas por*

$$s_n = t_n - \frac{g(t_n)}{g'\left(\frac{t_{n-1}+s_{n-1}}{2}\right)} \qquad y \qquad t_{n+1} = t_n - \frac{g(t_n)}{g'\left(\frac{t_n+s_n}{2}\right)},$$

donde $t_0 \in [0, t^*)$ y $s_0 = t_0 - \frac{g(t_0)}{g'(t_0)}$. Entonces

$$0 \leq t_n \leq s_n \leq t_{n+1} < t^*, \qquad n = 0, 1, \ldots,$$

y (t_n) converge a t^*.

Una prueba del lema anterior puede ser hallada en [19].

El objetivo del principio del mayorante es construir una sucesión real que "controle" a la sucesión (x_n), de tal forma que la convergencia de la sucesión real garantice la convergencia de la sucesión (x_n). En nuestro caso probaremos que

$$d(x_{n+1}, x_n) \leq t_{n+1} - t_n, \qquad \text{para todo } n.$$

El lema anterior y el hecho que estamos asumiendo G completo, garantiza la convergencia de (x_n).

Comencemos probando los siguientes estimativos, los cuales permiten "controlar" el comportamiento de F en términos de g. Por comodidad asumiremos en adelante $t_0 = 0$.

Proposición 3.1. *Bajo las mismas hipótesis anteriores, tenemos:*

H1. $\left\| dF_{x_0}^{-1}(F(x_0)) \right\| \leq -\frac{g(0)}{g'(0)}$;

H2. $\left\| dF_{x_0}^{-1}\, d^2F_x \right\| \leq g''(t)$, *para todo* $x \in G$ *tal que* $d(x, x_0) \leq t \leq t^*$;

H3. $\left\| dF_{x_0}^{-1}[d^2F_x - d^2F_y] \right\| \leq |g''(t) - g''(s)|$, *para todo* $x, y \in G$ *y* $t, s \in [0, t^*]$ *con* $d(x, y) \leq |t - s|$.

Demostración. $H1$ sigue inmediatamente de $C1$ y la igualdad $-g(0)/g'(0) = \eta$. Para probar $H2$ notemos que

$$dF_{x_0}^{-1}\, d^2F_x = dF_{x_0}^{-1}\, d^2F_{x_0} + dF_{x_0}^{-1}\, [d^2F_x - d^2F_{x_0}].$$

De aquí y $C2$ concluimos que

$$\left\| dF_{x_0}^{-1}\, d^2F_x \right\| \leq M + \left\| dF_{x_0}^{-1}\, [d^2F_x - d^2F_{x_0}] \right\|,$$

de donde, en virtud de $C3$,

$$\begin{aligned} \left\| dF_{x_0}^{-1}\, d^2F_x \right\| &\leq M + Kd(x, x_0) \\ &\leq M + Kt \\ &= g''(t), \end{aligned}$$

si $0 \leq t \leq t^*$ y $d(x, x_0) \leq t$. Finalmente, $H3$ sigue directamente de $C3$ y $g''(t) = Kt + M$. □

Continuamos con el siguiente lema, el cual será útil en la prueba del teorema principal de esta sección.

Lema 3.2. *Con la notación definida en Observación 3.3,*

$$F(x_{n+1}) = \frac{1}{4}\int_0^1 (1-t)[d^2F_{c_n(t)} - d^2F_{\alpha_n(t)}](\tilde{v}_n, \tilde{v}_n)dt + \\ + \int_0^1 (1-t)d^2F_{c_n(t)}(u_n, u_n)dt + \int_0^1 (1-t)d^2F_{c_n(t)}(u_n, \tilde{v}_n)dt.$$

Demostración. La fórmula de Taylor (3.15) aplicada a la curva c_n garantiza que

$$F(x_{n+1}) = F(g_n) + dF_{g_n}(w_n) + \int_0^1 (1-t)d^2F_{c_n(t)}(w_n, w_n)dt,$$

lo cual implica

$$F(x_{n+1}) = F(g_n) + dF_{g_n}(v_n) + dF_{g_n}(-\tilde{v}_n/2) + \int_0^1 (1-t)d^2F_{c_n(t)}(w_n, w_n)dt.$$

Así, por (3.33), se tiene

$$F(x_{n+1}) = F(g_n) - F(x_n) + dF_{g_n}(-\tilde{v}_n/2) + \int_0^1 (1-t)d^2F_{c_n(t)}(w_n, w_n)dt. \tag{3.35}$$

De otro lado, de la fórmula de Taylor (3.15) aplicada a la curva α_n se sigue que

$$F(x_n) = F(g_n) + dF_{g_n}(-\tilde{v}_n/2) + \int_0^1 (1-t)d^2F_{\alpha_n(t)}(-\tilde{v}_n/2, -\tilde{v}_n/2)dt,$$

de donde, remplazando en (3.35) y usando la definición de u_n, w_n, se concluye que

$$F(x_{n+1}) = -\frac{1}{4}\int_0^1 (1-t)d^2F_{\alpha_n(t)}(\tilde{v}_n, \tilde{v}_n)dt + \int_0^1 (1-t)d^2F_{c_n(t)}(u_n + \tilde{v}_n/2, u_n + \tilde{v}_n/2)dt.$$

De aquí se sigue el lema, usando la bilinealidad y la simetría de d^2F_x, $x \in G$. $\square$

Observación 3.4. *En primer lugar, ya que $d(T \circ F) = T \circ dF$ para todo endomorfismo $T : \mathfrak{g} \to \mathfrak{g}$, la conclusión del lema resulta válida si cambiamos F por $T \circ F$.*

También, el mismo procedimiento realizado en la prueba del lema puede ser aplicado al polinomio g para obtener la fórmula

$$\begin{aligned} g(t_{n+1}) = \frac{1}{4}\int_0^1 (1-t)[g''(\tau t_{n+1} + (1-\tau)m_n) - g''(\tau t_n + (1-\tau)m_n)](s_n - t_n)^2 d\tau \\ + \int_0^1 (1-t)g''(\tau t_{n+1} + (1-\tau)m_n)(t_{n+1} - s_n)^2 d\tau \\ + \int_0^1 (1-t)g''(\tau t_{n+1} + (1-\tau)m_n)(t_{n+1} - s_n)(s_n - t_n) d\tau, \end{aligned}$$

donde $m_n = (s_n + t_n)/2$, *para todo* $n \geq 0$.

Ahora probaremos el teorema principal de la sección.

Teorema 3.4. *Con la notación anterior y asumiendo* $t_0 = 0$, *tenemos los siguientes estimativos:*

a) $\left\| dF_{x_0}^{-1}(F(x_k)) \right\| \leq g(t_k)$;

b) $\left\| dF_{x_k}^{-1}\, dF_{x_0} \right\| \leq -\frac{1}{g'(t_k)}$;

c) $\|\tilde{v}_k\| \leq s_k - t_k$;

d) $\left\| dF_{g_k}^{-1}\, dF_{x_0} \right\| \leq -\frac{1}{g'(m_k)}$;

e) $\|v_k\| \leq t_{k+1} - t_k$;

f) $\|u_k\| \leq t_{k+1} - s_k$,

para todo $k = 0, 1, \ldots$.

Demostración. La prueba la haremos por inducción sobre k.

El ítem a) para $k = 0$ sigue por $H1$ y el hecho que $g'(0) = -1$. El ítem b) para $k = 0$ se sigue como consecuencia de las igualdades $g'(0) = -1$ y $dF_{x_0}^{-1}\, dF_{x_0} = I$, donde I denota el operador identidad sobre el álgebra $\mathfrak{g}$. Para el ítem c), note que

$$\tilde{v}_0 = -dF_{x_0}^{-1}(F(x_k)) \qquad \text{y} \qquad s_0 - t_0 = s_0 = -\frac{g(0)}{g'(0)}.$$

Luego $c)$ para $k=0$ sigue por $H1$. Para $d)$ probaremos primero que dF_{g_0} es invertible, para lo cual usaremos el lema de Banach. Usando (3.17) obtenemos

$$\begin{aligned} dF_{x_0}^{-1}\, dF_{g_0} - I &= dF_{x_0}^{-1}\, [dF_{g_0} - dF_{x_0}] \\ &= \int_0^1 dF_{x_0}^{-1}\, d^2_{\beta_0(t)}(\tilde{v}_0/2, \cdot) dt, \end{aligned}$$

donde recordemos que $\beta_0(t) = x_0 \exp(t\tilde{v}_0/2)$, $0 \leq t \leq 1$. De aquí y $C1$ se sigue que

$$d(\beta(t), x_0) \leq \frac{t}{2}\, \|\tilde{v}_0\| \leq \frac{\eta}{2} \leq t^*,$$

ya que $\eta = g(0) = 0 - g(0)/g'(0) = s_0 \leq t^*$. Se obtiene por $H2$ que

$$\begin{aligned} \left\| dF_{x_0}^{-1}\, dF_{g_0} - I \right\| &\leq \int_0^1 \left\| dF_{x_0}^{-1}\, d^2_{\beta_0(t)}(\tilde{v}_0/2, \cdot) \right\| dt \\ &\leq \int_0^1 g''\left(\frac{t}{2}\, \|\tilde{v}_0\|\right) \frac{\|\tilde{v}_0\|}{2} dt \\ &= g'\left(\frac{\|\tilde{v}_0\|}{2}\right) - g'(0) \\ &= 1 + g'\left(\frac{\|\tilde{v}_0\|}{2}\right) < 1. \end{aligned} \tag{3.36}$$

Concluimos del lema de Banach que $dF_{x_0}^{-1}\, dF_{g_0}$ es invertible y en consecuencia también lo es dF_{g_0}. Además tenemos el estimativo

$$\left\| dF_{g_0}^{-1}\, dF_{x_0} \right\| \leq \frac{1}{1 - \left\| dF_{x_0}^{-1}\, dF_{g_0} - I \right\|} \leq -\frac{1}{g'\left(\frac{\|\tilde{v}_0\|}{2}\right)} \leq -\frac{1}{g'(m_0)},$$

donde la última desigualdad es consecuencia de

$$\frac{\|\tilde{v}_0\|}{2} \leq \frac{\eta}{2} = \frac{s_0}{2} = \frac{t_0 + s_0}{2} = m_0$$

y del hecho que g' es creciente en $[0, t^*]$. Esto prueba $d)$ en el caso $k=0$. Para probar $e)$ en este caso, notemos que

$$\|v_0\| = \left\| dF_{g_0}^{-1}(F(x_0)) \right\| = \left\| dF_{g_0}^{-1}\, dF_{x_0}\, dF_{x_0}^{-1}(F(x_0)) \right\| \leq \left\| dF_{g_0}^{-1}\, dF_{x_0} \right\| \left\| dF_{x_0}^{-1}(F(x_0)) \right\|.$$

El resultado se sigue por ítem $a)$, ítem $d)$ y definición de t_1. Ahora probemos $f)$ para $k=0$. Tenemos por definición de u_n que

$$u_0 = v_0 - \tilde{v}_0 = dF_{x_0}^{-1}(F(x_0)) - dF_{g_0}^{-1}(F(x_0)) = \left(dF_{x_0}^{-1}\, dF_{g_0} - I\right) dF_{g_0}^{-1}(F(x_0)).$$

De lo probado en (3.36) y el estimativo anterior para $v_0 = -dF_{g_0}^{-1}(F(x_0))$, se concluye que

$$\begin{aligned}\|u_0\| &\leq \left\|dF_{x_0}^{-1}\, dF_{g_0} - I\right\| \left\|dF_{g_0}^{-1}(F(x_0))\right\| \\ &\leq (g'(\|\tilde{v}_0\|/2) - g'(0))\, \|v_0\| \\ &\leq (g'(m_0) - g'(0))\,(t_1 - t_0) \\ &= t_1 - s_0,\end{aligned}$$

donde la última igualdad se sigue por calculo directo, teniendo en cuenta la definición de t_1, s_0 y el hecho que $g'(0) = 1$. Esto prueba $e)$ para $k = 0$.

Ahora supongamos que $a)-e)$ se satisfacen para $k = 0, \ldots, n$ y probemos que también se cumplen para $k = n+1$. Observar que $e)$ implica que para $0 \leq i \leq j \leq n+1$,

$$d(x_i, x_j) \leq d(x_i, x_{i+1}) + \cdots + d(x_{j-1}, x_j) \leq \|v_i\| + \cdots + \|v_{j-1}\| \leq t_j - t_i. \tag{3.37}$$

Probemos $a)$ para $k = n+1$. Por un lado, aplicando el Lema 3.2 a $dF_{x_0}^{-1}(F(x_{n+1}))$ en lugar de F, obtenemos

$$\begin{aligned}dF_{x_0}^{-1}(F(x_{n+1})) = \frac{1}{4}\int_0^1 &(1-t)dF_{x_0}^{-1}[d^2F_{c_n(t)} - d^2F_{\alpha_n(t)}](\tilde{v}_n, \tilde{v}_n)dt \\ &+ \int_0^1 (1-t)dF_{x_0}^{-1}d^2F_{c_n(t)}(u_n, u_n)dt \\ &\qquad + \int_0^1 (1-t)dF_{x_0}^{-1}d^2F_{c_n(t)}(u_n, \tilde{v}_n)dt.\end{aligned} \tag{3.38}$$

Por otro lado, la definición de c_n y α_n dada en la Observación 3.3, implica que $c_n(t) = \alpha_n(t)\exp(tv_n)$. Por lo tanto, la hipótesis inductiva para $e)$ garantiza que

$$d(c_n(t), \alpha_n(t)) \leq t\,\|v_n\| \leq t(t_{n+1} - t_n) = (tt_{n+1} + (1-t)m_n) - (tt_n + (1-t)m_n).$$

De aquí, $H3$ y la hipótesis inductiva para $c)$, se sigue el siguiente estimativo para el primer término de (3.38)

$$\begin{aligned}&\frac{1}{4}\int_0^1 (1-t)dF_{x_0}^{-1}[d^2F_{c_n(t)} - d^2F_{\alpha_n(t)}](\tilde{v}_n, \tilde{v}_n)dt \leq \\ &\quad \frac{1}{4}\int_0^1 (1-\tau)\left(g''(\tau t_{n+1} + (1-\tau)m_n) - g''(\tau t_n + (1-\tau)m_n)\right)(s_n - t_n)^2 d\tau.\end{aligned} \tag{3.39}$$

Similarmente, como

$$c_n(t) = g_n \exp(tw_n) = x_n \exp(\tilde{v}_n/2)\exp(tv_n - t\tilde{v}_n/2) = x_n \exp(tv_n + (1-t)\tilde{v}_n/2),$$

deducimos de la hipótesis inductiva y (3.37) que

$$\begin{aligned} d(c_n(t), x_0) &\leq d(c_n(t), x_n) + d(x_n, x_0) \\ &\leq \|tv_n + (1-t)\tilde{v}_n/2\| + t_n \\ &\leq t\,\|v_n\| + \frac{1-t}{2}\,\|\tilde{v}_n\| + t_n \\ &\leq t(t_{n+1} - t_n) + \frac{1-t}{2}(s_n - t_n) + t_n \\ &= tt_{n+1} + (1-t)m_n. \end{aligned}$$

Así, de la hipótesis inductiva y $H2$ se sigue que

$$\int_0^1 (1-t)dF_{x_0}^{-1}d^2F_{c_n(t)}(u_n, u_n)dt \leq \int_0^1 (1-\tau)g''(\tau t_{n+1} + (1-\tau)m_n)(t_{n+1} - s_n)^2 d\tau$$

y también

$$\int_0^1 (1-t)dF_{x_0}^{-1}d^2F_{c_n(t)}(u_n, u_n)dt \leq \int_0^1 (1-\tau)g''(\tau t_{n+1} + (1-\tau)m_n)(t_{n+1} - s_n, s_n - t_n)d\tau.$$

De estos dos últimos estimativos, (3.39) y (3.38) concluimos que

$$\begin{aligned} \left\|dF_{x_0}^{-1}(F(x_{n+1}))\right\| \leq \frac{1}{4}\int_0^1 (1-\tau)\left(g''(\tau t_{n+1} + (1-\tau)m_n) - g''(\tau t_n + (1-\tau)m_n)\right)(s_n - t_n)^2 d\tau \\ + \int_0^1 (1-\tau)g''(\tau t_{n+1} + (1-\tau)m_n)(t_{n+1} - s_n)^2 d\tau \\ + \int_0^1 (1-\tau)g''(\tau t_{n+1} + (1-\tau)m_n)(t_{n+1} - s_n, s_n - t_n)d\tau. \end{aligned}$$

De la expresión para $g(t_{n+1})$ dada en la Observación 3.4 deducimos $a)$ para $k = n+1$.

Para $b)$, consideremos la curva $\lambda_n(t) = x_n \exp(tv_n)$, $0 \leq t \leq 1$, la cual tiene extremos en x_n y x_{n+1}. Entonces,

$$dF_{x_0}^{-1}dF_{x_{n+1}} - dF_{x_0}^{-1}dF_{x_n} = \int_0^1 dF_{x_0}^{-1}d^2F_{\lambda_n(t)}(v_n, \cdot)dt,$$

De aquí, $H2$, las hipótesis inductiva $b)$ y $e)$ y el estimativo

$$d(\lambda_n(t), x_0) \leq d(\lambda_n(t), x_n) + d(x_n, x_0) \leq t\,\|v_n\| + t_n \leq tt_{n+1} + (1-t)t_n,$$

obtenemos

$$\begin{aligned}\left\|[dF_{x_0}^{-1}dF_{x_n}]^{-1}dF_{x_0}^{-1}dF_{x_{n+1}} - I\right\| &\leq \left\|dF_{x_n}^{-1}dF_{x_0}\right\| \left\|dF_{x_0}^{-1}dF_{x_{n+1}} - dF_{x_0}^{-1}dF_{x_n}\right\| \\ &\leq -\frac{1}{g'(t_n)}\int_0^1 g''(\tau t_{n+1} + (1-\tau)t_n)(t_{n+1}-t_n)d\tau \\ &= -\frac{1}{g'(t_n)}(g'(t_{n+1}) - g'(t_n)) \\ &= 1 - \frac{g'(t_{n+1})}{g'(t_n)} < 1.\end{aligned}$$

Del lema de Banach se sigue que $[dF_{x_0}^{-1}dF_{x_n}]^{-1}dF_{x_0}^{-1}dF_{x_{n+1}}$ es invertible y por lo tanto también lo es $dF_{x_{n+1}}$. Además,

$$\begin{aligned}\left\|dF_{x_{n+1}}^{-1}dF_{x_0}\right\| &\leq \frac{\left\|dF_{x_n}^{-1}dF_{x_0}\right\|}{1 - \left\|dF_{x_n}^{-1}dF_{x_0}\right\| \left\|dF_{x_0}^{-1}dF_{x_{n+1}} - dF_{x_0}^{-1}dF_{x_n}\right\|} \\ &\leq -\frac{1}{g'(t_{n+1})},\end{aligned}$$

lo cual es $b)$ para $k = n+1$.

El ítem $c)$ para $k = n+1$ sigue inmediatamente de hipótesis inductiva $d)$ y lo probado para ítem $a)$ en el caso $k = n+1$. En efecto,

$$\|\tilde{v}_{n+1}\| = \left\|dF_{g_n}^{-1}(F(x_{n+1}))\right\| \leq \left\|dF_{g_n}^{-1}dF_{x_0}\right\| \left\|dF_{x_0}^{-1}F(x_{n+1})\right\| \leq -\frac{g(t_{n+1})}{g'(m_n)} = s_{n+1} - t_{n+1}.$$

Para probar $d)$, primero notamos que

$$\begin{aligned}dF_{x_{n+1}}^{-1}dF_{g_{n+1}} - I &= dF_{x_{n+1}}^{-1}[dF_{g_{n+1}} - dF_{x_{n+1}}] \\ &= dF_{x_{n+1}}^{-1}\, dF_{x_0}\, dF_{x_0}^{-1}[dF_{g_{n+1}} - dF_{x_{n+1}}].\end{aligned}$$

Así, por lo probado para $b)$ para $k = n+1$ y considerando la curva β_{n+1} definida en Observación 3.3, obtenemos

$$\left\|dF_{x_{n+1}}^{-1}dF_{g_{n+1}} - I\right\| \leq -\frac{1}{g'(t_{n+1})}\int_0^1 \left\|dF_{x_0}^{-1}d^2F_{\beta_{n+1}(t)}(\tilde{v}_{n+1}/2, \cdot)\right\| dt.$$

Ahora, por (3.37), definición de β_{n+1} y lo probado para $c)$ en $k = n+1$,

$$d(\beta_{n+1}(t), x_0) \leq d(\beta_{n+1}(t), x_{n+1}) + d(x_{n+1}, x_0) \leq t\frac{\|\tilde{v}_{n+1}\|}{2} + t_{n+1} \leq t\frac{s_{n+1} - t_{n+1}}{2} + t_{n+1},$$

lo cual da

$$d(\beta_{n+1}(t), x_0) \leq t_{n+1} + t(m_{n+1} - t_{n+1}) \leq t^*.$$

De aquí y $H2$ y el hecho $\tilde{v}_{n+1}/2 \leq (s_{n+1} - t_{n+1})/2 = m_{n+1} - t_{n+1}$,

$$\begin{aligned}
\left\| dF_{x_{n+1}}^{-1} dF_{g_{n+1}} - I \right\| &\leq -\frac{1}{g'(t_{n+1})} \int_0^1 \left\| dF_{x_0}^{-1} d^2 F_{\beta_{n+1}(t)} \right\| \|\tilde{v}_{n+1}/2\| \, dt. \\
&\leq -\frac{1}{g'(t_{n+1})} \int_0^1 g''(t_{n+1} + \tau(m_{n+1} - t_{n+1}))(m_{n+1} - t_{n+1}) d\tau \\
&= -\frac{1}{g'(t_{n+1})} \left(g'(m_{n+1}) - g'(t_{n+1}) \right) \\
&= 1 - \frac{g'(m_{n+1})}{g'(t_{n+1})} < 1.
\end{aligned}$$

Esto muestra que $dF_{x_{n+1}}^{-1} dF_{g_{n+1}}$ es invertible y además se concluye

$$\left\| dF_{g_{n+1}}^{-1} dF_{x_{n+1}} \right\| \leq \frac{1}{1 - \left\| dF_{x_{n+1}}^{-1} dF_{g_{n+1}} - I \right\|} \leq \frac{g'(t_{n+1})}{g'(m_{n+1})}.$$

Luego, por lo probado para $b)$ en $k = n + 1$,

$$\left\| dF_{g_{n+1}}^{-1} dF_{x_0} \right\| \leq \left\| dF_{g_{n+1}}^{-1} dF_{x_{n+1}} \right\| \left\| dF_{x_{n+1}}^{-1} dF_{x_0} \right\| \leq -\frac{1}{g'(m_{n+1})},$$

lo que demuestra $d)$ para $k = n + 1$.

El ítem $e)$ para $k = n + 1$ se deduce de

$$\|v_{n+1}\| = \left\| dF_{g_{n+1}}^{-1}(F(x_{n+1})) \right\| \leq \left\| dF_{g_{n+1}}^{-1} dF_{x_0} \right\| \left\| dF_{x_0}^{-1}(F(x_{n+1})) \right\|,$$

de lo probado para $a)$ y $d)$ en $k = n + 1$ y de la definición de la sucesión (t_n).

Solo resta probar que se cumple $f)$ en $k = n + 1$. Por definición, podemos expresar u_{n+1} en la forma

$$\begin{aligned}
u_{n+1} &= v_{n+1} - \tilde{v}_{n+1} \\
&= \left(dF_{g_n}^{-1} - dF_{g_{n+1}}^{-1} \right) (F(x_{n+1})) \\
&= dF_{g_n}^{-1} \left(dF_{g_{n+1}} - dF_{g_n} \right) dF_{g_{n+1}}^{-1} \left(F(x_n + 1) \right) \\
&= \left(dF_{g_n}^{-1} dF_{x_0} \right) \left[dF_{x_0}^{-1} \left(dF_{g_{n+1}} - dF_{g_n} \right) \right] \left(dF_{g_{n+1}}^{-1} dF_{x_0} \right) \left(dF_{x_0}^{-1} (F(x_n + 1)) \right).
\end{aligned} \tag{3.40}$$

Ahora, por un lado, por lo probado para $a)$ y $d)$,

$$\begin{aligned}
\left\| dF_{g_n}^{-1} dF_{x_0} \right\| &\leq -\frac{1}{g'(m_n)}; \\
\left\| dF_{g_{n+1}}^{-1} dF_{x_0} \right\| &\leq -\frac{1}{g'(m_{n+1})}; \\
\left\| dF_{x_0}^{-1} (F(x_n+1)) \right\| &\leq g(t_{n+1}).
\end{aligned} \tag{3.41}$$

De otro lado, definiendo el vector $V = u_n + \tilde{v}_n/2 + \tilde{v}_{n+1}/2$, puede ser probado que la curva

$$r(t) = g_n \exp(tV), \qquad 0 \leq t \leq 1,$$

satisface $r(0) = g_n$ y

$$r(1) = g_n \exp(v_n - \tilde{v}_n/2 + \tilde{v}_{n+1}/2) = x_n \exp(v_n + \tilde{v}_{n+1}/2) = x_{n+1} \exp(\tilde{v}_{n+1}/2) = g_{n+1},$$

de donde

$$dF_{x_0}^{-1} \left(dF_{g_{n+1}} - dF_{g_n} \right) = \int_0^1 dF_{x_0}^{-1} d^2 F_{r(t)}(V, \cdot). \tag{3.42}$$

Ahora, por definición de g_n, el hecho que $c)$ se cumple para $k = n, n+1$, $f)$ se cumple para $k = n$ y (3.37), tenemos

$$\begin{aligned}
d(r(t), x_0) &\leq d(r(t), g_n) + d(g_n, x_n) + d(x_n, x_0) \\
&\leq t \|V\| + \|\tilde{v}_n/2\| + t_n \\
&\leq t(\|u_n\| + \|\tilde{v}_n/2\| + \|\tilde{v}_{n+1}/2\|) + \frac{s_n - t_n}{2} + t_n \\
&\leq t \left(t_{n+1} - s_n + \frac{s_n - t_n}{2} + \frac{s_{n+1} - t_{n+1}}{2} \right) + \frac{s_n + t_n}{2} \\
&= t(m_{n+1} - m_n) + m_n \leq t^*.
\end{aligned}$$

Notar que también hemos probado que $\|V\| \leq m_{n+1} - m_n$. Luego, por $H2$ y (3.42),

$$\begin{aligned}
\left\| dF_{x_0}^{-1} \left(dF_{g_{n+1}} - dF_{g_n} \right) \right\| &\leq \int_0^1 g''(\tau(m_{n+1} - m_n) + m_n) \|V\| \, d\tau \\
&\leq \int_0^1 g''(\tau(m_{n+1} - m_n) + m_n)(m_{n+1} - m_n) d\tau \\
&= g'(m_{n+1}) - g'(m_n).
\end{aligned}$$

Finalmente, (3.40), (3.41) y el estimativo anterior implican

$$\begin{aligned}
\|u_{n+1}\| &\leq \frac{g(t_{n+1})}{g'(m_n) g'(m_{n+1})} \left(g'(m_{n+1}) - g'(m_n) \right) \\
&= \left(\frac{1}{g'(m_n)} - \frac{1}{g'(m_{n+1})} \right) g(t_{n+1}) \\
&= t_{n+2} - s_{n+1}.
\end{aligned}$$

Esto completa la prueba del teorema. $\square$

Finalizamos este capítulo con el resultado que garantiza la convergencia del método de King-Werner. Note que los resultados obtenidos siguen siendo válidos si las hipótesis son restringidas a cumplirse en $\Omega = x_0 \exp(B(\mathbf{0}, t^*))$, donde $B(\mathbf{0}, t^*)$ es la bola abierta en el álgebra $\mathfrak{g}$ de centro $\mathbf{0}$ y radio t^*.

Teorema 3.5. *Asuma las notaciones e hipótesis anteriores. Entonces la sucesión (x_n) está bien definida y converge a una raíz x^* de $F(x) = \mathbf{0}$ en la clausura de Ω.*

Demostración. En el teorema anterior se probó que la sucesión está bien definida y además satisface

$$d(x_n, x_m) \leq t_n - t_m, \qquad \text{para todo } m < n.$$

En particular,

$$d(x_0, x_n) \leq t_n \leq t^*, \qquad \text{para todo } n$$

y la sucesión (x_n) es de Cauchy en G. Como G es completo, (x_n) converge a algún x^* satisfaciendo $d(x^*, x_0) \leq t^*$.

Finalmente, por continuidad de F y el ítem a) del teorema anterior, se concluye que $F(x^*) = \mathbf{0}$. $\square$

Capítulo 4

Ejemplos numéricos

4.1. Una aplicación del método de la secante

Iniciamos con un ejemplo en el cual implementamos el método secante sobre grupos de Lie definido en (3.4), para aproximar una solución de una ecuación matricial. Consideremos el conjunto de matrices

$$\mathcal{SO}_2(\mathbb{R}) = \left\{ X \in \mathbb{R}^{2\times 2} : X^T X = I,\ \det(X) = 1 \right\}$$

del Ejemplo 2.2, el cual recordemos que es un grupo de Lie cuya álgebra de Lie es

$$\mathfrak{so}_2(\mathbb{R}) = \left\{ V \in \mathbb{R}^{2\times 2} : V^T + V = 0 \right\}.$$

El problema a considerar es hallar una raíz de la función

$$F : \mathcal{SO}_2(\mathbb{R}) \to \mathfrak{so}_2(\mathbb{R}),$$

definida por

$$F(X) = X - X^T + A, \qquad \text{con} \qquad A = \begin{pmatrix} 0 & \frac{1}{4} \\ -\frac{1}{4} & 0 \end{pmatrix} \in \mathfrak{so}_2(\mathbb{R}). \tag{4.1}$$

Note que $\|A\|_F = \frac{1}{4}\sqrt{2} \leq 1$, donde $\|\cdot\|_F$ denota la norma de Frobenius. Por comodidad, en adelante omitiremos el subíndice F en $\|\cdot\|_F$. Para más detalles sobre la norma Frobenius, ver [61]. Podemos observar también de la definición de F y las propiedades del operador transpuesto que

$$F^T(X) = X^T - X - A = -F(X),$$

de donde concluimos que en efecto F está bien definida.

Como primer paso hallemos expresiones para dF y d^2F. Dados $X, Y \in \mathcal{SO}_2(\mathbb{R})$ y $U, V \in \mathfrak{so}_2(\mathbb{R})$, definimos la curva $\gamma(t) = X\exp(tU)$, $t \in \mathbb{R}$, la cual satisface $\gamma(0) = X$ y $\gamma'(0) = (L'_X)_I(U)$. Por lo tanto, de la definición de dF dada en (2.11) y la igualdad

$$\begin{aligned} F(\gamma(t)) &= \gamma(t) - [\gamma(t)]^T + A \\ &= X\exp(tU) - [X\exp(tU)]^T + A \\ &= X\exp(tU) - \exp(tU^T)X^T + A, \end{aligned}$$

deducimos que

$$\begin{aligned} dF_X(U) &= \frac{d}{dt}\Big|_{t=0} F(\gamma(t)) \\ &= \frac{d}{dt}\Big|_{t=0}[X\exp(tU)] - \frac{d}{dt}\Big|_{t=0}[\exp(tU^T)X^T] \\ &= XU - U^TX^T \\ &= XU + UX^{-1}, \end{aligned} \tag{4.2}$$

donde en la penúltima igualdad hemos utilizado los cálculos realizados en Ejemplo 2.4. Por otro lado, por (2.14) obtenemos

$$\begin{aligned} d^2F_X(W,W) &= \frac{d^2}{dt^2}\Big|_{t=0} F(X\exp(tW)) \\ &= \frac{d}{dt}\Big|_{t=0}\left(XW\exp(tW) - W^T\exp(tW^T)X^T\right) \\ &= XW^2 - (XW^2)^T, \end{aligned}$$

para todo $W \in \mathfrak{so}_2(\mathbb{R})$. De aquí y la igualdad

$$d^2F_X(U,V) = \frac{1}{4}\left\{d^2F_X(U+V,U+V) - d^2F_X(U-V,U-V)\right\}$$

podemos obtener para $(U,V) \in \mathfrak{g} \times \mathfrak{g}$,

$$\begin{aligned} d^2F_X(U,V) &= \frac{X(U+V)^2 - (X(U+V)^2)^T}{4} - \frac{X(U-V)^2 - (X(U-V)^2)^T}{4} \\ &= \frac{XUV + XVU}{2} - \frac{U^TV^TX^T + V^TU^TX^T}{2}. \end{aligned}$$

Ahora, dado que el álgebra $\mathfrak{so}_2(\mathbb{R})$ es abeliana se tiene que

$$d^2F_X(U,V) = XUV - (XUV)^T. \tag{4.3}$$

Adicionalmente, consideramos la diferencia dividida descrita en el Ejemplo 3.1, esto es,

$$[X, X\exp(U); F](W) := \int_0^1 dF_{\gamma(t)}(W)dt, \qquad W \in \mathfrak{so}_2(\mathbb{R}).$$

Ahora usaremos la fórmula para dF obtenida en (4.2) para hallar una expresión alternativa para la diferencia dividida anterior. De (4.2), la definición de $\mathcal{SO}_2(\mathbb{R})$ y de $\mathfrak{so}_2(\mathbb{R})$, se sigue que

$$\begin{aligned}[X, X\exp(U); F](W) &= \int_0^1 [\gamma(t)W + W(\gamma(t))^{-1}]dt \\ &= \int_0^1 [X\exp(tU)W + W\exp(-tU)X^{-1}]dt \\ &= [X\exp(U)U^{-1} - XU^{-1}](W) - [(X\exp(U)U^{-1} - XU^{-1})(W)]^T.\end{aligned}$$

Concluimos que para $X \in \mathcal{SO}_2(\mathbb{R})$ y $U \in \mathfrak{so}_2(\mathbb{R})$,

$$[X, X\exp(U); F](W) = G(X, U)(W) - [G(X, U)(W)]^T, \qquad W \in \mathfrak{so}_2(\mathbb{R}), \tag{4.4}$$

donde $G : M(2, \mathbb{R}) \times \mathcal{GL}(2, \mathbb{R}) \to M(2, \mathbb{R})$ está definida por

$$G(A, B) = A[\exp(B) - I]B^{-1}. \tag{4.5}$$

Observemos que si

$$G(X, U) = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \text{y} \qquad W = \begin{pmatrix} 0 & w \\ -w & 0 \end{pmatrix},$$

entonces

$$G(X, U)W = w\begin{pmatrix} -b & a \\ -d & c \end{pmatrix}.$$

Se sigue de (4.4) que

$$[X, X\exp(U); F](W) = \begin{pmatrix} 0 & w(a+d) \\ -w(a+d) & 0 \end{pmatrix} \tag{4.6}$$

y por lo tanto, si $a + d \neq 0$,

$$[X, X\exp(U); F]^{-1}(W) = \frac{1}{a+d}\begin{pmatrix} 0 & w \\ -w & 0 \end{pmatrix}. \tag{4.7}$$

Así,

$$\left\|[X, X\exp(U); F]^{-1}\right\| = \frac{1}{|a+d|}. \tag{4.8}$$

Verifiquemos que se satisfacen las condiciones de convergencia presentadas en el Teorema 3.1 del Capítulo 3. Para ello es esencial definir una función ω que cumpla la condición (3.5).

Sean $U, V \in \mathfrak{so}_2(\mathbb{R})$ y $X \in \mathcal{SO}_2(\mathbb{R})$. Estas tres matrices son de la forma

$$U = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 0 & v \\ -v & 0 \end{pmatrix}, \qquad X = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$$

con $u, v, \theta \in \mathbb{R}$. De (4.3) se deduce fácilmente que

$$d^2F_X(U,V) = \begin{pmatrix} 0 & -2uv\sin\theta \\ 2uv\sin\theta & 0 \end{pmatrix}$$

y consecuentemente,

$$\begin{aligned} \frac{\|d^2F_X(U,V)\|_F}{\|U\|_F\,\|V\|_F} &= \frac{\sqrt{(-2uv\sin\theta)^2 + (-2uv\sin\theta)^2}}{\sqrt{2u^2}\sqrt{2v^2}} \\ &= \frac{\sqrt{8}\,|\sin\theta|}{\sqrt{2}\sqrt{2}} \\ &= \sqrt{2}\,|\sin\theta|\,. \end{aligned}$$

Así, $\|d^2F_X\|_F \leq \sqrt{2}$. Si ahora $Y = X\exp(U)$, con $\|U\| = \|\exp^{-1}(X^{-1}Y)\| = d(X,Y)$, se sigue de (3.17) que

$$\begin{aligned} \left\|dF_{X\exp(U)} - dF_X\right\| &\leq \int_0^1 \left\|d^2F_{X\exp(tU)}(U,\cdot)\right\|_F dt \\ &\leq \int_0^1 \sqrt{2}\,\|U\|\,dt \\ &= \sqrt{2}\,d\,(X,Y)\,. \end{aligned}$$

Usaremos este estimativo para definir la función ω. Dados

$$P_1, P_2, Q_1, Q_2 \in \mathcal{SO}_2(\mathbb{R}) \qquad \text{y} \qquad U_1, V_1 \in \mathfrak{so}_2(\mathbb{R})$$

satisfaciendo las igualdades

$$P_2 = P_1\exp(U_1) \qquad \text{y} \qquad Q_2 = Q_1\exp(V_1), \tag{4.9}$$

se deduce que

$$
\begin{aligned}
\|[P_1,P_2;F]-[Q_1,Q_2;F]\| &= \left\| \int_0^1 dF_{P_1\exp(tU_1)}dt - \int_0^1 dF_{Q_1\exp(tV_1)}dt \right\| \\
&\leq \int_0^1 \left\| dF_{P_1\exp(tU_1)} - dF_{Q_1\exp(tV_1)} \right\| dt \\
&\leq \sqrt{2} \int_0^1 d\left(P_1\exp(tU_1), Q_1\exp(tV_1)\right) dt \\
&= \sqrt{2} \int_0^1 \left\| \exp^{-1}\left((P_1\exp(tU_1))^{-1} Q_1\exp(tV_1) \right) \right\| dt.
\end{aligned}
\tag{4.10}
$$

De otro lado, de la propiedad $[\exp(u)]^{-1}=\exp(-u)$, la cual es válida para todo $u\in\mathfrak{g}$, se sigue que

$$
\begin{aligned}
\exp^{-1}\left((P_1\exp(tU_1))^{-1}(Q_1\exp(tV_1))\right) &= \exp^{-1}\left(\left(\exp(-tU_1)P_1^{-1}\right)(Q_1\exp(tV_1))\right) \\
&= \exp^{-1}(\exp(-tU_1)) + \exp^{-1}\left(P_1^{-1}Q_1\right) + \exp^{-1}(\exp(tV_1)) \\
&= \exp^{-1}\left(P_1^{-1}Q_1\right) - tU_1 + tV_1,
\end{aligned}
$$

de donde, por (4.10),

$$
\|[P_1,P_2;F]-[Q_1,Q_2;F]\| \leq \sqrt{2}\int_0^1 \left\| \exp^{-1}\left(P_1^{-1}Q_1\right) - tU_1 + tV_1 \right\| dt. \tag{4.11}
$$

Para estimar el integrando, usamos la definición de U_1, V_1 dada en (4.9), para obtener

$$
\begin{aligned}
&\exp^{-1}\left(P_1^{-1}Q_1\right) - tU_1 + tV_1 = \exp^{-1}\left(P_1^{-1}Q_1\right) - t\exp^{-1}(P_1^{-1}P_2) + t\exp^{-1}(Q_1^{-1}Q_2) \\
&= \exp^{-1}\left(P_1^{-1}Q_1\right) - t\left(\exp^{-1}(P_1^{-1}) + \exp^{-1}(P_2)\right) + t\left(\exp^{-1}(Q_1^{-1}) + \exp^{-1}(Q_2)\right) \\
&= \exp^{-1}\left(P_1^{-1}Q_1\right) + t\left(\exp^{-1}(Q_1^{-1}) - \exp^{-1}(P_1^{-1}) + \exp^{-1}(Q_2) - \exp^{-1}(P_2)\right).
\end{aligned}
\tag{4.12}
$$

Ahora, la fórmula $\exp^{-1}(g\cdot h)=\exp^{-1}(g)+\exp^{-1}(h)$, $g\cdot h\in G$, la cual es válida en grupos de Lie abelianos, como es el caso de $\mathcal{SO}_2(\mathbb{R})$, implica que $\exp^{-1}(g^{-1})=-\exp^{-1}(g)$, para todo $g\in G$. Luego, (4.12) implica

$$
\begin{aligned}
\exp^{-1}\left(P_1^{-1}Q_1\right) - tU_1 + tV_1 &= \exp^{-1}\left(P_1^{-1}Q_1\right) + t\left(\exp^{-1}(Q_1^{-1}P_1) + \exp^{-1}(P_2^{-1}Q_2)\right) \\
&= (1-t)\exp^{-1}\left(P_1^{-1}Q_1\right) + t\exp^{-1}(P_2^{-1}Q_2),
\end{aligned}
$$

de donde, por (4.11) deducimos que

$$\begin{aligned}\|[P_1,P_2;F]-[Q_1,Q_2;F]\| &\leq \sqrt{2}\int_0^1 \left\|(1-t)\exp^{-1}\left(P_1^{-1}Q_1\right)+t\exp^{-1}(P_2^{-1}Q_2)\right\| dt\\ &\leq \sqrt{2}\int_0^1 \left((1-t)\left\|\exp^{-1}\left(P_1^{-1}Q_1\right)\right\|+t\left\|\exp^{-1}(P_2^{-1}Q_2)\right\|\right) dt\\ &= \frac{\sqrt{2}}{2}\left(\left\|\exp^{-1}\left(P_1^{-1}Q_1\right)\right\|+\left\|\exp^{-1}(P_2^{-1}Q_2)\right\|\right)\\ &= \frac{\sqrt{2}}{2}\left(d\left(P_1,Q_1\right)+d(P_2,Q_2)\right).\end{aligned}$$

Por lo tanto, definiendo

$$\omega\left(x,y\right) := \frac{\sqrt{2}}{2}\left(x+y\right), \qquad x,y\geq 0,$$

se concluye que la diferencia dividida considerada satisface la $\omega-$condición (3.5). Esto es,

$$\|[P_1,P_2;F]-[Q_1,Q_2;F]\| \leq \omega(d(P_1,Q_1),d(P_2,Q_2)).$$

Tomemos

$$P_0=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\quad P_1=\begin{pmatrix}0{,}98981 & 0{,}14237\\ -0{,}14237 & 0{,}98981\end{pmatrix}\ \text{ y }\ V_0=\begin{pmatrix}0 & 1/7\\ -1/7 & 0\end{pmatrix},$$

los cuales satisfacen

$$P_1=P_0\exp(V_0) \quad \text{y} \quad d\left(P_0,P_1\right)=\|V_0\|=0{,}20203.$$

Así, en el Teorema 3.1 podemos tomar $\alpha=0{,}20203$.

Para obtener la norma $\|[P_0,P_1;F]^{-1}\|$, calculemos $G\left(P_0,V_0\right)$. Por (4.5),

$$\begin{aligned}G\left(P_0,V_0\right) &= P_0[\exp(V_0)-I_2]V_0^{-1}\\ &= [P_1-P_0]V_0^{-1}\\ &= \begin{pmatrix}0{,}99660 & 0{,}071303\\ -0{,}071303 & 0{,}99660\end{pmatrix}.\end{aligned}$$

Concluimos de (4.8) que

$$\beta=\left\|[P_0,P_1;F]^{-1}\right\|=\frac{1}{2*0{,}99660}=0{,}50171.$$

Para calcular la constante η en el Teorema 3.1, notemos que

$$F(P_1) = \begin{pmatrix} 0 & 0{,}53474 \\ -0{,}53474 & 0 \end{pmatrix},$$

de donde, en virtud de (4.7), obtenemos

$$[P_0, P_1; F]^{-1}(F(P_1)) = 0{,}50171 \begin{pmatrix} 0 & 0{,}53474 \\ -0{,}53474 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0{,}2683 \\ -0{,}2683 & 0 \end{pmatrix}$$

Así, podemos tomar

$$\eta = \left\| [P_0, P_1; F]^{-1}(F(P_1)) \right\| = 0{,}37941.$$

Ahora, recordando que

$$a\,(u) = \frac{\beta\omega\,(\alpha, u)}{1 - \beta\omega\,(\alpha, u)}\,, \quad b\,(u) = \frac{\beta\omega\,(u, 2u)}{1 - \beta\omega\,(\alpha + u, u)}\,, \quad c\,(u) = \frac{\beta\omega\,(2u, 2u)}{1 - \beta\omega\,(\alpha + u, u)}\,,$$

podemos calcular las soluciones de

$$u = \left(\frac{b\,(u)\,a\,(u)}{1 - c\,(u)} + a\,(u) + 1 \right) \eta$$

y obtener

$$u_1 = 0{,}44166, \qquad u_2 = 0{,}44166 \quad \text{y} \quad u_3 = 2{,}2217.$$

Así, tomamos $R = u_1$. Solo queda por verificar que se satisfacen las condiciones

$$\beta\omega\,(R + \alpha, R) < 1 \qquad \text{y} \qquad c\,(R) < 1.$$

En efecto, por definición de ω,

$$\beta\omega\,(R + \alpha, R) = \beta(2R + \alpha) = 0{,}38504 < 1$$

y

$$c(R) = \frac{4R\beta}{1 - \beta(\alpha + 2R)} = 0{,}96051 < 1.$$

Así, se satisfacen todas las hipótesis del Teorema 3.1, por lo cual la sucesión secante (P_n), con punto inicial P_0, converge a la única solución de $F(X) = X - X^T + A = \mathbf{0}$ en la bola $\overline{B(P_1, R)}$.

Finalmente, usaremos la fórmula (4.7) para describir un algoritmo que permita aproximar una solución de $F(X) = \mathbf{0}$. Recordemos que por (4.4) y (4.5),

$$[P_{k-1}, P_k; F](F(P_k)) = G(P_{k-1}, V_{k-1})(F(P_k)) - [G(P_{k-1}, V_{k-1})(F(P_k))]^T$$

y además,

$$\begin{aligned} G(P_{k-1}, V_{k-1}) &= P_{k-1}[\exp(V_{k-1}) - I_2]V_{k-1}^{-1} \\ &= [P_k - P_{k-1}]V_{k-1}^{-1}, \end{aligned}$$

El algoritmo queda como sigue:

i) Fijamos A, P_0, P_1 y V_0;

ii) Para $k = 1, 2, \ldots$ calculamos

1. $G(P_{k-1}, V_{k-1})$;
2. $F(P_k) = P_k - P_k^T + A$;
3. $V_k = -\dfrac{1}{\text{traza}(G(P_{k-1}, V_{k-1}))} F(P_k)$;
4. $P_k = P_{k-1}\exp(V_{k-1})$.

La siguiente tabla presenta las primeras iteraciones con el algoritmo planteado.

i	P_i	$F(P_i)$	$\|F(P_i)\|$
0	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1/4 \\ -1/4 & 0 \end{pmatrix}$	0,35355
1	$\begin{pmatrix} 0{,}98981 & 0{,}14337 \\ -0{,}14237 & 0{,}98981 \end{pmatrix}$	$\begin{pmatrix} 0 & 0{,}53474 \\ -0{,}53474 & 0 \end{pmatrix}$	0,75624
2	$\begin{pmatrix} 0{,}99214 & -0{,}12510 \\ 0{,}12510 & 0{,}99214 \end{pmatrix}$	$\begin{pmatrix} 0 & -0{,}00019646 \\ 0{,}00019646 & 0 \end{pmatrix}$	0,00027783
3	$\begin{pmatrix} 0{,}99216 & -0{,}12500 \\ 0{,}12500 & 0{,}99216 \end{pmatrix}$	$\begin{pmatrix} 0 & -9{,}5367 \times 10^{-7} \\ 9{,}5367 \times 10^{-7} & 0 \end{pmatrix}$	$1{,}3487 \times 10^{-6}$

Tabla 4.1: Resultados para el problema $F(X) = \mathbf{0}$ usando el algoritmo (3.4).

Observación 4.1. *Con el objetivo de comparar el método de Newton-Kantorovich con el método de la secante, implementamos también el método de Newton-Kantorovich* (3.25) *a la función del ejemplo anterior. Procediendo como en el ejemplo anterior uno puede escoger* $\beta = \frac{1}{2}$, $\eta = \frac{1}{2}$ *y* $M = 2$. *Así* $h := \beta\eta M = \frac{1}{2}$, *por lo que se satisfacen las hipótesis del Teorema* 3.3. *Los datos obtenidos, partiendo del punto inicial* $P_0 = I_2$, *se presentan en la siguiente tabla.*

i	X_i	$F(X_i)$	$\|F(X_i)\|$
0	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1/4 \\ -1/4 & 0 \end{pmatrix}$	0,35355
1	$\begin{pmatrix} 0{,}99220 & -0{,}12467 \\ 0{,}12467 & 0{,}99220 \end{pmatrix}$	$\begin{pmatrix} 0 & 0{,}00065 \\ -0{,}00065 & 0 \end{pmatrix}$	0,00091981
2	$\begin{pmatrix} 0{,}99216 & -0{,}05461 \\ 0{,}12500 & 0{,}99216 \end{pmatrix}$	$\begin{pmatrix} 0 & -4{,}44739 \times 10^{-10} \\ 4{,}44739 \times 10^{-10} & 0 \end{pmatrix}$	$6{,}2896 \times 10^{-10}$
3	$\begin{pmatrix} 0{,}99216 & -0{,}12500 \\ 0{,}12500 & 0{,}99216 \end{pmatrix}$	$\begin{pmatrix} 0 & -2{,}22044 \times 10^{-16} \\ 2{,}22044 \times 10^{-16} & 0 \end{pmatrix}$	$3{,}1402 \times 10^{-16}$

Tabla 4.2: Resultados para $F(X) = \mathbf{0}$ usando el algoritmo (3.25).

Notemos que como ocurre en el caso finito dimensional, y mas generalmente en espacios de Banach, el método de Newton-Kantorovich converge más rápido a la solución de $F(X) = \mathbf{0}$ que el método de la secante.

Finalizamos con un ejemplo donde consideramos la función F de los ejemplos anteriores, pero en el caso $n = 3$.

4.2. Una aplicación del método de Newton Kantorovich (N-K)

Consideremos ahora el problema $F(X) = \mathbf{0}$, donde $F : \mathcal{SO}_3(\mathbb{R}) \to \mathfrak{so}_3(\mathbb{R})$ está dada por

$$F(X) = X - X^T + A, \qquad X \in \mathcal{SO}_3(\mathbb{R}),$$

con

$$A = \begin{pmatrix} 0 & \frac{1}{8} & -\frac{1}{4} \\ -\frac{1}{8} & 0 & \frac{1}{4} \\ \frac{1}{4} & -\frac{1}{4} & 0 \end{pmatrix} \in \mathfrak{so}_3(\mathbb{R}).$$

Recordemos que $\mathcal{SO}_3(\mathbb{R})$ es el grupo de Lie de matrices reales de orden 3, ortogonales con determinante 1, esto es,

$$\mathcal{SO}_3(\mathbb{R}) = \left\{ X \in \mathbb{R}^{3\times 3} : X^T X = I, \quad \det(X) = 1 \right\}.$$

Su álgebra de Lie $\mathfrak{so}_3(\mathbb{R})$, es el grupo de matrices reales de orden 3 antisimétricas.

Para aplicar el método de Newton Kantorovich definido en (3.25), primero verificamos que se satisfacen las hipótesis del Teorema 3.3. Los cálculos realizados en el ejemplo anterior, pero tomando $X_0 = I_3$, nos llevan a que podemos tomar $\beta = \frac{1}{2}$, $\eta = \frac{1}{2}$ y $M = 2$, por lo que $h := \beta\eta M = \frac{1}{2}$. La siguiente tabla muestra las primeras iteraciones con el algoritmo Newton-Kantorovich planteado.

i	X_i	$\|F(X_i)\|$
0	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	0,53033
1	$\begin{pmatrix} 0{,}990264 & -0{,}054345 & 0{,}128164 \\ 0{,}069924 & 0{,}990264 & -0{,}120374 \\ -0{,}120374 & 0{,}128164 & 0{,}984420 \end{pmatrix}$	0,003101
2	$\begin{pmatrix} 0{,}990147 & -0{,}054618 & 0{,}128940 \\ 0{,}070382 & 0{,}990147 & -0{,}121059 \\ -0{,}121059 & 0{,}128941 & 0{,}984233 \end{pmatrix}$	$1{,}0474 \times 10^{-6}$
3	$\begin{pmatrix} 0{,}990147 & -0{,}054618 & 0{,}128940 \\ 0{,}070382 & 0{,}990147 & -0{,}121059 \\ -0{,}121060 & 0{,}128941 & 0{,}984233 \end{pmatrix}$	$6{,}0920 \times 10^{-7}$

Tabla 4.3: Resultados para $F(X) = \mathbf{0}$ usando el algoritmo N-K.

Conclusiones

El presente trabajo de investigación inicia a partir de preguntarnos qué métodos iterativos, de la amplia gama de los que han sido estudiados en espacios de Banach, pueden ser desarrollados en el contexto de grupos de Lie. En este sentido presentamos algunos avances importantes en tres de los métodos iterativos más conocidos: el método de la secante, el método de King-Werner y el renombrado método de Newton-Kantorovich. Resaltamos que a la fecha, en el contexto de grupos de Lie, solo se había estudiado el método de Newton, sobre éste ya otros autores habían obtenido resultados de convergencia, tanto local como semilocal.

Para los tres métodos estudiados, nuestro principal interés fue realizar un análisis de convergencia, logrando obtener teoremas de convergencia semilocal para cada uno de ellos. En el caso del método de Newton-Kantorovich, estudiado en trabajos previos bajo diferentes hipótesis, presentamos un nuevo enfoque estudiando un resultado de convergencia semilocal bajo condiciones tipo Kantorovich y utilizando la técnica de relaciones de recurrencia. En el caso del método de la secante, además de un resultado de convergencia semilocal, logramos probar que su orden de convergencia coincide con el orden de convergencia del método en el contexto de espacios de Banach, para este resultado fue necesario desarrollar un teorema tipo Taylor en grupos de Lie. Es de destacar que el orden de convergencia ya es conocido para el método de Newton-Kantorovich en grupos de Lie. Finalmente, el análisis de convergencia semilocal para el método de King-Werner es presentado para el caso en que el grupo de Lie es abeliano.

A partir de los resultados obtenidos quedan abiertas varias preguntas de interés. Por ejemplo, realizar un análisis de convergencia local para los tres métodos estudiados. ¿Qué otros métodos numéricos se pueden desarrollar en grupos de Lie? Estudiar el método de King-Werner en el caso no abeliano, entre otras.

Bibliografía

[1] R.L. Adler, J.P. Dedieu, J.Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA Journal of Numerical Analysis **22** (2002), no. 3, 359–390.

[2] F. Álvarez, J. Bolte, and J. Munier, *A Unifying Local Convergence Result for Newton's Method in Riemannian Manifolds*, Found. Comput. Math. **8** (2008), no. 2, 197–226.

[3] S. Amat, I.K. Argyros, S. Busquier, R. Castro, S. Hilout, and S. Plaza, *On a bilinear operator free third order method on Riemannian manifolds*, Applied Mathematics and Computation **219** (2013), no. 14, 7429–7444.

[4] ______, *Traub-type high order iterative procedures on Riemannian manifolds*, SeMA J. **63** (2014), 27–52.

[5] S. Amat, S. Busquier, R. Castro, and S. Plaza, *Third-order methods on Riemannian manifolds under Kantorovich conditions*, Journal of Computational and Applied Mathematics **255** (2014), 106–121.

[6] I.K. Argyros, *Chebyshev-Halley-like methods in Banach spaces*, Korean Journal of Computational & Applied Mathematics **4** (1997), no. 1, 83–107.

[7] ______, *An improved unifying convergence analysis of Newton's method in Riemannian manifolds*, Journal of Applied Mathematics and Computing **25** (2007), 345–351.

[8] ______, *Newton's method on Lie groups*, Journal of Applied Mathematics and Computing **31** (2008), 217–228.

[9] ______, *Newton's method on lie groups*, Journal of Applied Mathematics and Computing **31** (2009), no. 1, 217–228.

[10] ______, *Computational Theory of Iterative Methods*, Studies in Computational Mathematics, vol. 15, Elsevier, 2014.

[11] I.K. Argyros and D. Chen, *The midpoint method for solving nonlinear operator equations in Banach space*, Applied mathematics letters **5** (1992), no. 4, 7–9.

[12] ______, *Results on the Chebyshev method in Banach spaces*, Proyecciones (Antofagasta, On line) **12** (1993), no. 2, 119–128.

[13] I.K. Argyros, D. Chen, and Q.S. Qian, *A note on the Halley method in Banach spaces*, Applied Mathematics and Computation **58** (1993), no. 2-3, 215–224.

[14] I.K. Argyros and S. Hilout, *Newton's method for approximating zeros of vector fields on Riemannian manifolds*, Journal of Applied Mathematics and Computing **29** (2009), 417–427.

[15] ______, *Extending the applicability of newton's method on lie groups*, Applied Mathematics and Computation **219** (2013), no. 20, 10355–10365.

[16] I.K. Argyros and S.K. Khattri, *On the Secant method*, Journal of Complexity **29** (2013), no. 6, 454–471.

[17] I.K. Argyros and H. Ren, *On the convergence of efficient king–werner-type methods of order* $1+\sqrt{2}$, Journal of Computational and Applied Mathematics **285** (2015), 169–180.

[18] T. Bittencourt and O. P. Ferreira, *Kantorovich's Theorem on Newton's Method under Majorant Condition in Riemannian Manifolds*, J. of Global Optimization **68** (2017), no. 2, 387–411.

[19] E. Cárdenas, R. Castro, and W. Sierra, *Convergence analysis for the king–werner method under* γ*–conditions*, Journal of Applied Mathematics and Computing (2022).

[20] R. Castro, *Higher order iterative methods on Riemannian manifolds*, Ph.D. thesis, Universidad de Santiago de Chile, Las Sophoras n 173. Estación Central, Santiago, Chile., 2011.

[21] R. Castro, J.C. Rodríguez, W. Sierra, G. Di Giorgi, and S. Gómez, *Chebyshev–Halley's method on Riemannian manifolds*, Journal of Computational and Applied Mathematics **336** (2018), 30–53.

[22] D. Chen, *Ostrowski-Kantorovich theorem and S-order of convergence of Halley method in Banach spaces*, Commentationes Mathematicae Universitatis Carolinae **34** (1993), no. 1, 153–163.

[23] J.P. Dedieu and D. Nowicki, *Symplectic methods for the approximation of the exponential map and the Newton iteration on Riemannian submanifolds*, Journal of Complexity **21** (2005), no. 4, 487–501, Festschrift for the 70th Birthday of Arnold Schonhage.

[24] J.P. Dedieu, Priouret P., and G. Malajovich, *Newton's method on Riemannian manifolds: covariant alpha theory*, IMA Journal of Numerical Analysis **23** (2003), no. 3, 395–419.

[25] M.P. do Carmo, *Riemannian geometry*, Birkhäuser, 1992.

[26] J.A. Ezquerro, J.M. Gutiérrez, M.A. Hernández, and M. A. Salanova, *El método de halley: posiblemente, el método más redescubierto del mundo*, Margarita mathematica en memoria de José Javier (Chicho) Guadalupe Hernández, Universidad de La Rioja, 2001, pp. 205–220.

[27] J.A. Ezquerro and M.A Hernández, *Newton's method: an updated approach of kantorovich's theory*.

[28] Y. Félix, Oprea J., and D. Tanré, *Algebraic models in geometry*, OUP Oxford, 2008.

[29] O.P. Ferreira and B.F. Svaiter, *Kantorovich's Theorem on Newton's Method in Riemannian Manifolds*, Journal of Complexity **18** (2002), no. 1, 304–329.

[30] M. Garcıa and J. Gutiérrez, *Notas históricas sobre el método de chebyshev para resolver ecuaciones no lineales*, Miscelánea Matemática **57** (2013), 63–83.

[31] M. Grau-Sánchez, A. Grau, and M. Noguera, *Frozen divided difference scheme for solving systems of nonlinear equations*, Journal of Computational and Applied Mathematics **235** (2011), no. 6, 1739–1743.

[32] J.M. Gutiérrez and M.A. Hernández, *A family of Chebyshev-Halley type methods in Banach spaces*, Bulletin of the Australian Mathematical Society **55** (1997), no. 1, 113–130.

[33] D. Han and X. Wang, *Point estimates on deformation newton's iterations*, Mathematica Numerica Sinica **1** (1990), no. 2, 145–156.

[34] ______, *Convergence on a deformed newton method*, Applied Mathematics and Computation **94** (1998), no. 1, 65–72.

[35] T. Hawkins, *Emergence of the theory of Lie groups: An essay in the history of mathematics 1869–1926*, Springer Science & Business Media, 2012.

[36] J. He, J. Wang, and J. Yao, *Convergence criteria of Newton's method on Lie groups*, Fixed Point Theory and Applications **2013** (2013), 1–15.

[37] M.A. Hernández and M.J. Rubio, *Semilocal convergence of the secant method under mild convergence conditions of differentiability*, Computers & Mathematics with Applications **44** (2002), no. 3-4, 277–285.

[38] J. Hilgert and K.H. Neeb, *Structure and geometry of Lie groups*, Springer Science & Business Media, 2011.

[39] J.Wang and C. Li, *Kantorovich's theorems for newton's method for mappings and optimization problems on lie groups*, IMA journal of numerical analysis **31** (2011), no. 1, 322–347.

[40] L.V. Kantorovich, *On Newton's method for functional equations (Russian)*, Dokl. Akad. Nauk. SSSR **59** (1948), 1237–1240.

[41] R.F. King, *Tangent methods for nonlinear equations*, Numerische Mathematik **18** (1971), no. 4, 298–304.

[42] C. Li and J. Wang, *Newton's method on Riemannian manifolds: Smale's point estimate theory under the γ-condition*, IMA Journal of Numerical Analysis **26** (2006), no. 2, 228–251.

[43] ______, *Kantorovich's theorem for Newton's method on Lie groups*, Journal of Zhejiang University-SCIENCE A **8** (2007), no. 6, 978–986.

[44] ______, *Newton's method for sections on Riemannian manifolds: Generalized covariant α theory*, Journal of Complexity **24** (2008), no. 3, 423–451.

[45] C. Li and J. Wang, *Kantorovich's theorems for Newton's method for mappings and optimization problems on Lie groups*, IMA Journal of Numerical Analysis **31** (2009), no. 1, 322–347.

[46] C. Li, J. Wang, and J. Dedieu, *Smale's point estimate theory for newton's method on lie groups*, Journal of Complexity **25** (2009), no. 2, 128–151.

[47] T.J. McDougall and S.J. Wotherspoon, *A simple modification of newton's method to achieve convergence of order* $1+\sqrt{2}$, Applied Mathematics Letters **29** (2014), 20–25.

[48] B. Owren and B. Welfert, *The Newton iteration on Lie groups*, BIT Numerical Mathematics **40** (2000), no. 1, 121–145.

[49] J.M. Papakonstantinou and R.A. Tapia, *Origin and evolution of the Secant method in one dimension*, The American Mathematical Monthly **120** (2013), no. 6, 500–517.

[50] H. Ren, *On the local convergence of a deformed newton's method under argyros-type condition*, Journal of mathematical analysis and applications **321** (2006), no. 1, 396–404.

[51] J.W. Schmidt, *Eine Übertragung der Regula Falsi auf Gleichungen in Banachräumen I*, ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik **43** (1963), no. 1-2, 1–8.

[52] ______, *Eine Übertragung der Regula Falsi auf Gleichungen in Banachräumen II Nichtlineare Gleichungssysteme*, ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik **43** (1963), no. 3, 97–110.

[53] ______, *Regula-falsi-Verfahren mit konsistenter Steigung und Majorantenprinzip*, Periodica Mathematica Hungarica **5** (1974), no. 3, 187–193.

[54] J. Schröder, *Nichtlineare majoranten beim verfahren der schrittweisen näherung*, Archiv der Mathematik **7** (1957), no. 6, 471–484.

[55] A.S. Sergeev, *On the method of chords*, Sibirsk. Matem. Zhournal **2** (1961), 282–289.

[56] V.S. Varadarajan, *Lie groups, Lie algebras, and their representations*, vol. 102, Springer Science & Business Media, 2013.

[57] X. Wang and S. Zheng, *On the convergence of king–werner's iteration procedure for solving nonlinear equations*, Math. Numer. Sinica **4** (1982), 70–79.

[58] F.W. Warner, *Foundations of differentiable manifolds and Lie groups*, vol. 94, Springer Science & Business Media, 2013.

[59] W. Werner, *Über ein verfahren der ordnung* $1+\sqrt{2}$ *zur nullstellenbestimmung*, Numerische Mathematik **32** (1979), no. 3, 333–342.

[60] ______, *Some supplementary results on the* $1+\sqrt{2}$ *order method for the solution of nonlinear equations*, Numerische Mathematik **38** (1982), no. 3, 383–392.

[61] X. Zhang, *Matrix analysis and applications*, Cambridge University Press, 2017.

[62] H. Zhengda, *A note on the kantorovich theorem for newton iteration*, Journal of computational and applied mathematics **47** (1993), no. 2, 211–217.