Universitat de Barcelona
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Mètode de Newton en espais de Banach: Teorema de Kantorovich i aplicacions

Autor: Daniel Fernández Benítez

Director: Dr. Joan Carles Tatjer Montaña
Realitzat a: Departament Matemàtiques i Informàtica

Barcelona, 20 de juny de 2021

Resum

Donada una equació $Fx = \vec{0}$, on $F\colon X \mapsto Y$, X i Y espais de Banach de la mateixa dimensió, el Teorema de Kantorovich dona condicions suficients perquè, donat un $x^0 \in X$, l'equació tingui una solució x^* propera a x^0, localment única, i que el mètode de Newton amb condició inicial x^0 hi convergeixi; a més dona estimacions rigoroses de regions on x^* existeix i és única.

En aquest treball enunciem el Teorema de Kantorovich i el demostrem per dues vies: recurrència i successions majorants. També presentem algunes variants i resultats similars, i aplicacions i exemples de com fer servir el Teorema.

Resumen

Dada una ecuación $Fx = \vec{0}$, donde $F\colon X \mapsto Y$, X e Y espacios de Banach de la misma dimensión, el Teorema de Kantorovich da condiciones suficientes para que, dada $x^0 \in X$, la ecuación tenga una solución x^* próxima a x^0, localmente única, y que el método de Newton con condición inicial x^0 converja a ella; además da estimaciones precisas de regiones donde x^* existe y es única.

En este trabajo enunciamos el Teorema de Kantorovich i lo demostramos de dos formas distintas: por recurrencia y por sucesiones mayorantes. También presentamos algunas variantes y resultados similares, y aplicaciones y ejemplos de cómo usar el Teorema.

Abstract

Let the equation $Fx = \vec{0}$, where $F\colon X \mapsto Y$, X and Y are Banach spaces of like dimension, Kantorovich's Theorem gives sufficient conditions for the equation to have a solution x^*, locally unique, close to a previously picked $x^0 \in X$, and that the Newton's method with initial condition x^0 converges to it. Moreover, gives sharp estimations of the regions where x^* exists and is unique.

In this thesis we state Kantorovich's Theorem and prove it in two different ways: by recurrence and by majorant sequences. Besides, some variants and like results are stated, along with applications and examples on how to make use of the Theorem.

Índex

1 Introducció

Donada una aplicació $F\colon D \mapsto \mathbb{R}^n$, $D \subseteq \mathbb{R}^n$, ens plantegem el problema de trobar solucions a equacions del tipus $F(x) = y_0$ amb $y_0 \in \mathbb{R}^n$ fixat; o, com és més habitual, absorbint y_0 en la funció, resultant equacions

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

El mateix problema es pot plantejar substituint $\mathbb{R}^n$ per $\mathbb{C}^n$ o per espais de Banach arbitraris, que és el que suposarem en endavant, quedant aleshores $F\colon D \mapsto Y$, $D \subseteq X$, on X i Y són espais de Banach sobre el mateix cos K, $K = \mathbb{R}$ o $K = \mathbb{C}$, i de la mateixa dimensió, no necessàriament finita. En aquest context més general, diem que F és un operador.

Les qüestions fonamentals a discutir de l'equació (1.1) són l'*existència* de solucions i el seu *càlcul*. En general no podrem calcular explícitament solucions (mètodes directes) i ens haurem de conformar amb aproximar-les (mètodes iteratius). Un *mètode iteratiu* consisteix a generar successivament millors aproximacions a la solució d'un problema —per a nosaltres (1.1)— a partir d'una dada inicial o llavor; és a dir, generar una successió $(x^k)_{k\geq 0}$, $x^{k+1} = G(x^k)$, on x^0 és la llavor, tal que $\lim_{k\to\infty} x^k = x^*$ i $F(x^*) = \vec{0}$.

Sovint, l'existència d'una solució de (1.1) i la convergència d'un mètode iteratiu a ella es proven de manera separada; el Teorema de Kantorovich adreça ambdues qüestions alhora per al mètode iteratiu de Newton

$$x^{k+1} = x^k - F'(x^k)^{-1} F x^k \qquad \big(\forall\, k \geq 0\big) \tag{1.2}$$

provant l'existència (i unicitat local) d'una solució x^* propera a x^0, i la convergència de (1.2) a ella. Més encara, dona regions d'existència i unicitat —un conjunt a on x^* pertany, i un altre en què hi és única— i també fites d'error per al mètode iteratiu. Com veurem, el Teorema de Kantorovich també és vàlid per al mètode iteratiu de Newton simplificat

$$x^{k+1} = x^k - F'(x^0)^{-1} F x^k \qquad (k \geq 0) \tag{1.3}$$

que resulta més fàcil de fer servir en la pràctica, però a canvi d'una convergència més lenta.

El mètode de Newton es pot modificar de moltes altres formes, que reben el nom genèric de mètodes quasi-Newton

$$x^{k+1} = x^k - B(x^k)^{-1} F x^k \qquad (k \geq 0) \tag{1.4}$$

$$x^{k+1} = x^k - B_k^{-1} F x^k \qquad (k \geq 0) \tag{1.5}$$

on per a cada $x \in D$, $B(x)$ és un operador lineal relacionat amb $F'(x)$ i $(B_k)_{k\geq 0}$ és una successió d'operadors lineals. Els mètodes quasi-Newton es poden veure com alternatives entremig dels mètodes de Newton i Newton simplificat, i voldríem disposar de resultats similars al de Kantorovich per a aquests mètodes, dels quals només esmentarem breument alguns

D'un mètode iteratiu ens interessen, fonamentalment, tres aspectes: que el mètode estigui ben definit, que convergeixi a una solució, i la facilitat de realitzar els càlculs dels iterats. Que estigui ben definit vol dir que la llavor triada x^0 generi realment una successió $x^{k+1} = G(x^k)$, $k \geq 0$. En el cas de (1.2) tenim $G(x) = x - F'(x)^{-1} F(x)$, on cal $F'(x^k)$ invertible i $x^k \in D$, $\forall\, k \geq 0$. Quant a la dificultat de calcular cada iterat, en cada pas hem d'avaluar F i F' i resoldre l'equació lineal $F'(x^k)y = F(x^k)$, cosa que motiva l'ús d'alternatives (1.3) (1.4) (1.5) al mètode de Newton, per ———-per no haver de resoldre un nou sistema en cada pas, sinó només de tant en tant, o reduir o evitar el cost de calcular explícitament $F'(x^k)$ en cada pas. Sobre la convergència dels mètodes iteratius, podem distingir tres tipus de resultats

- De *convergència local*, que assumint l'existència d'una solució, assegura que hi ha un entorn d'ella en què, per a tota dada inicial dins l'entorn, el mètode iteratiu està definit i convergeix a la solució esmentada.
- De *convergència global*, que assegura la convergència del mètode per a tota dada inicial de l'espai, o almenys en un conjunt gran dins l'espai. Un exemple en són els mètodes iteratius de Jacobi i Gauss-Seidel per a l'aproximació de solucions de sistemes d'equacions lineals.
- De *convergència semilocal*, que donada una llavor, i sense assumir l'existència de cap solució, garanteix que el mètode iteratiu està definit i convergeix a una solució —presumible ment propera a la llavor—.

La distinció entre aquests grups no és del tot clara i la seva descripció pot variar segons la font consultada; també pot no ser evident en quin grup classificar un resultat. Aquí hem escollit donar els trets mínims en què coincideixen totes les descripcions; una formulació més específica es pot consultar en els prefacis de [3, 4]. El concepte de resultat semilocal com a alternativa als resultats locals o globals sembla que apareix arran del Teorema de Kantorovich, éssent ja emprat el terme en [16] i també descrit en [18], aquest últim sense fer ús explícit de la paraula «semilocal».

El Teorema de Kantorovich (també anomenat de Newton-Kantorovich) és, doncs, un teorema de convergència semilocal per al mètode iteratiu de Newton en espais de Banach. En síntesi: si x^0 gairebé resol (1.1) —$Fx^0 \approx \vec{0}$— aleshores hi ha una solució de veritat molt a prop i el mètode de Newton hi convergeix. Més encara, ens diu que la successió (1.2) resta a una bola de centre x^0, i la solució és única en una altra bola de centre x^0, i ens dona fites per al radis d'aquestes boles: regions d'existència i unicitat.

Això fa que el teorema sigui important per diversos motius. Primer, malgrat no fou el primer resultat sobre la convergència del mètode de Newton en dimensió superior a 1, sí fou el primer amb un enunciat únic i general per a tots els casos: espais de Banach arbitraris. Segon, el teorema no només prova la convergència del mètode de Newton, sinó que prova l'existència (i unicitat local) d'una solució de (1.1); és a dir, no sols és una eina de càlcul, sinó una eina per a l'estudi teòric d'equacions de tot tipus.

No obstant, el Teorema de Kantorovich presenta el problema d'haver de trobar una dada inicial x^0 prou propera a una solució x^*, que a més haurà de satisfer unes condicions bastant restrictives. A més x^* serà, en general, una solució simple. Trobar aquest x^0 és un problema que haurem de resoldre pel nostre compte. Una solució x^* de (1.1) és simple si $F'(x^*)$ existeix, és bijectiu com a operador lineal continu, i el seu invers $F'(x^*)^{-1}$ també és continu; en síntesi, $F'(x^*)^{-1} \in \mathscr{B}(Y,X)$ (veure apartat 2.2).

1.1 Notació

La següent notació serà freqüent en tot el text.

- K per denotar el cos dels reals o dels complexos.
- Fx o $F(x)$ per a denotar la imatge de x per l'operador F.
- $F'(x)$ per denotar la derivada de l'operador F en el punt x. Per evitar recarregar la notació, escriurem l'argument sense parèntesi: $F'(x)h$ és la imatge de h per l'operador lineal $F'(x)$.
- $(x_n)_{n\geq 0}$ denotarà una successió en un conjunt; també $(x^n)_{n\geq 0}$ quan no hi hagi perill de confondre x^n amb una operació algebraica.

- $[a,b]$ per a denotar el segment tancat que uneix els punts $a,b \in X$. És a dir, el conjunt $\{at+(1-t)b : t \in [0,1]\}$. Anàlogament (a,b), $[a,b)$, $(a,b]$.

1.2 Objectius

El principal objectiu del treball és enunciar i demostrar el Teorema de Kantorovich. Les demostracions d'aquest teorema es divideixen en dues classes, per recurrència i per majoració (veure secció §3); en farem una de cada tipus, comparant el resultat obtingut i els avantatges i inconvenients de cadascuna. També apuntarem algunes variants del Teorema i resultats similars, ja sigui perquè es vol la convergència semilocal de (1.2) sota altres hipòtesis o amb altres tesis, o perquè es vol un resultat similar per a alguna variant del mètode de Newton (1.3)-(1.5) o altres.

També donem exemples d'aplicació del Teorema, mostrant com encetar les dificultats d'aplicar-lo i les diverses utilitats del Teorema: com a eina de càlcul i com a eina teòrica per a provar existència i unicitat (local) de solucions.

1.3 Estructura de la memòria

El treball consta de 6 seccions i s'ha intentat que sigui el més autocontingut possible, malgrat les limitacions d'espai; tot allò que no es demostra s'indica a on consultar-ho.

Els requisits d'anàlisi funcional es troben en la secció §2, que es limiten a introduir els operadors lineals continus i el concepte de derivació d'operadors; culminant amb els teoremes de Taylor i del Valor Mitjà de Lagrange, que emprarem amb profusió. En la secció §3 es demostra el Teorema de Kantorovich de dues formes diferents: recurrència i majoració; la primera fidel a l'original de Kantorovich i l'altra seguint idees posteriors de Rheinboldt i Ortega. La secció culmina amb un resum comparant ambdues demostracions i recollint altra informació rellevant relativa al Teorema.

En la secció §4 apuntem algunes variacions del Teorema de Kantorovich i resultats similars per a variacions del mètode de Newton; un recull exhaustiu de totes les variants és impossible, en donarem només uns pocs. En la secció §5 donem exemples concrets d'aplicació del Teorema de Kantorovich mostrant les diverses possibilitats que té el Teorema: com a eina de càlcul per a garantir que el mètode de Newton convergeix, i com a eina teòrica per a provar l'existència i unicitat local de solucions.

2 Requisits

En aquesta secció donem les nocions d'anàlisi funcional que necessitem en tot el treball: espais normats i de Banach, operadors lineals continus, derivació d'operadors i teoremes de Lagrange i Taylor. La part d'anàlisi funcional lineal —els dos primers apartats— es pot consultar en qualsevol llibre d'anàlisi funcional [13, 17], mentre que per a la derivació d'operadors i els teoremes de Taylor i Lagrange es pot consultar en [12, 1] o també [16, cap.3].

Què és l'anàlisi funcional? Quan fem anàlisi real o complexa, treballem sobre els espais vectorials $\mathbb{R}^n$ o $\mathbb{C}^n$ respectivament; l'anàlisi funcional és 'el mateix' però substituint aquests espais vectorials per espais de funcions. Aquests espais de funcions acostumen a ésser —i per nosaltres ho seran— espais vectorials (reals o complexos) dotats d'una norma, respecte de la qual en són complets. És a dir, que en anàlisi funcional es treballa amb 'còpies' de $\mathbb{R}^n$ i $\mathbb{C}^n$ de dimensió arbitrària, sovint no finita, ja que en copiem les propietats algebraiques i topològiques més rellevants: estructura d'espai vectorial normat que, com a espai mètric, és complet.

Aquests espais vectorials sempre són sobre el cos dels reals o dels complexos, segons es digui, i com s'ha dit a la introducció, denotarem en tot el text per K; és a dir $K = \mathbb{R}$ o $K = \mathbb{C}$, i per $|\lambda|$ el valor absolut o mòdul de $\lambda \in K$.

En anàlisi funcional, els espais normats acostumen a ésser espais de funcions; per això, qualsevol aplicació entre espais normats, o subconjunts d'espais normats, l'anomenem *operador*. Distingim el cas en què l'espai d'arribada d'un operador és el cos K, aleshores l'anomenem *funcional*.

2.1 Espais normats i de Banach

2.1.1 Norma i espais normats

Definició 2.1. Sigui X un espai vectorial (real o complex); una *norma* sobre X és una aplicació $\|\cdot\| : X \mapsto \mathbb{R}$ que satisfà les següents propietats

1. $\forall x \in X,\ \|x\| \geq 0$
2. $\|x\| = 0 \Leftrightarrow x = \vec{0}$
3. $\forall \lambda \in K\, \forall x \in X,\ \|\lambda x\| = |\lambda|\,\|x\|$ (homogeneitat)
4. $\forall x, y \in X,\ \|x + y\| \leq \|x\| + \|y\|$ (desigualtat triangular)

Diem aleshores que X, amb la norma $\|\cdot\|$, que denotem conjuntament per $\big(X, \|\cdot\|\big)$, és un espai vectorial normat, o simplement *espai normat*.

Observació 2.2. La propietat de desigualtat triangular d'una norma garanteix que per a tot x, y es té

$$\big|\|x\| - \|y\|\big| \leq \|x - y\| \tag{2.1}$$

La norma ens permet introduir un altre concepte: conjunt fitat. Un conjunt $A \subseteq X$ diem que està *fitat* respecte de la norma, si el conjunt $\big\{\|x\| \,:\, x \in A\big\} \subseteq \mathbb{R}$ està fitat.

Un concepte algebraic que necessitarem és el de convexitat. Un conjunt $C \subseteq X$ diem que és *convex* si per a tota parella de punts de C el segment tancat que els uneix també és en C: $\forall\, a, b \in C,\ [a, b] \subseteq C$.

A partir d'una norma $\|\cdot\|$ sempre podem definir una distància en X, $d\colon X \times X \mapsto \mathbb{R}$, $d(x, y) = \|x - y\|$. L'espai vectorial normat esdevé així un espai mètric, i podem introduir conceptes d'anàlisi: entorn d'un punt, conjunts oberts, tancats i compactes; convergència i límits, continuïtat d'aplicacions, etc.

2.1.2 Successions, límits i completesa

Una successió $(x_n)_{n\geq 0} \subseteq X$ en l'espai normat $\big(X, \|\cdot\|\big)$ diem que és *convergent* si existeix $x \in X$ tal que

$$\forall \epsilon > 0\, \exists N \in \mathbb{N} : \big(n \geq N \Rightarrow \big\|x_n - x\big\| \leq \epsilon\big)$$

Diem aleshores que x és el *límit de la successió*, i ho denotem per $\lim_{n\to\infty} x_n = x$ o $x_n \xrightarrow{n\to\infty} x$. Diem que la successió és de Cauchy si compleix

$$\forall \epsilon > 0\, \exists N \in \mathbb{N} : \big(n, m \geq N \Rightarrow \big\|x_n - x_m\big\| \leq \epsilon\big)$$

Un espai normat $(X, \|\cdot\|)$ en què tota successió de Cauchy és convergent diem que és un *espai de Banach*.

Un tipus especial de successió de Cauchy és el de *successió contractiva*: una successió és contractiva si hi ha $0 < \alpha < 1$ tal que per a tot $n \geq 0$, $\|x_{n+2} - x_{n+1}\| \leq \alpha \|x_{n+1} - x_n\|$.

Alguns exemples clàssics d'espais de Banach de dimensió finita en són $(K, |\cdot|)$: $\mathbb{R}$ amb el valor absolut, $\mathbb{C}$ amb el mòdul complex. També K^n amb les p-normes ($1 \leq p < \infty$) i la del suprem

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \qquad \|x\|_\infty = \max\{\,|x_1|, \ldots, |x_n|\}$$

Altres de dimensió no finita en són els espais de successions ℓ_p, amb la seva p-norma, ($1 \leq p < \infty$)

$$\ell_p = \left\{(x_n)_{n\geq 0} \subseteq K : \sum_{n\geq 0} |x_n|^p < \infty\right\} \qquad \|x\|_p = \left(\sum_{n\geq 0} |x_n|^p\right)^{\frac{1}{p}}$$

i l'espai ℓ_∞ de les successions afitades, amb la norma del suprem

$$\ell_\infty = \left\{(x_n)_{n\geq 0} \subseteq K : \sup_n |x_n| < \infty\right\} \qquad \|x\|_\infty = \sup_n |x_n|$$

Un altre espai de Banach de dimensió no finita és $\mathscr{C}([0,1], \mathbb{R})$: l'espai de les funcions reals contínues sobre $[0,1]$, amb la norma del suprem $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$. Aquest i molts altres exemples es poden trobar discutits en [17, 13].

2.1.3 Nocions topològiques

Introduïm algunes nocions topològiques dels espais mètrics adaptades a la notació dels espais normats. Assumim un espai normat $(X, \|\cdot\|)$.

Convé aclarir que, malgrat siguin en un altre apartat, els conceptes de convergència de successions, successió de Cauchy i espai de Banach, són conceptes topològics també.

- *Bola oberta* de radi $r > 0$ i centre $a \in X$, $B(a, r) = \{x \in X : \|x - a\| < r\}$. I la seva variant tancada $\overline{B(a, r)} = \{x \in X : \|x - a\| \leq r\}$.
- Un conjunt $V \subseteq X$ és un *entorn del punt a* si hi ha $r > 0$ tal que $B(a, r) \subseteq V$. Diem aleshores que a és un *punt interior* a V.
- Un conjunt $V \subseteq X$ és *obert* si és entorn de tots els seus punts; és a dir, si tots els seus punts són interiors: $\forall x \in V\, \exists r > 0 : B(x, r) \subseteq V$.
- Un punt a és *d'acumulació* d'un conjunt $V \subseteq X$ si es pot aproximar per punts de $V \setminus \{a\}$; és a dir, $\exists (x_n)_{n\geq 0} \subseteq V \setminus \{a\} : \lim_{n\to\infty} x_n = a$. Denotem el conjunt de punts d'acumulació de V per V'.
- Un punt a és *adherent* a un conjunt $V \subseteq X$ si es pot aproximar per punts de V; és a dir, $\exists (x_n)_{n\geq 0} \subseteq V : \lim_{n\to\infty} x_n = a$. Denotem el conjunt de punts adherents a V per $\overline{V} = V \cup V'$.
- Un conjunt $V \subseteq X$ és *tancat* si conté tots els seus punts adherents; és a dir, si $(x_n)_{n\geq 0} \subseteq V$ convergent $\Rightarrow \lim_{n\to\infty} x_n \in V$. En notació: $V = \overline{V}$.

Diem que un operador $F\colon D \mapsto Y$, $D \subseteq X$ té *límit en un punt* $a \in D'$ si existeix $l \in Y$ per a què es té

$$\forall \epsilon > 0 \, \exists \delta_{a,\epsilon} > 0 : \left(x \in D,\, 0 < \|x - a\| \leq \delta_{a,\epsilon} \Rightarrow \left\| F(x) - l \right\| \leq \epsilon \right)$$

diem aleshores que l n'és el límit i ho denotem per $\lim_{x\to a} F(x) = l$.

L'operador F és *continu en el punt* $a \in D$, quan es compleix alguna de les següents condicions, que en espais normats sempre són equivalents. Observem també que tota aplicació és contínua en els seus punts aïllats.

1. $\forall \epsilon > 0 \, \exists \delta_{a,\epsilon} > 0 : \left(x \in D,\, \|x - a\| \leq \delta_{a,\epsilon} \Rightarrow \left\| F(x) - F(a) \right\| \leq \epsilon \right)$
2. Per a tota successió $(x_n)_{n\geq 0} \subseteq D$ tal que $\lim_{n\to\infty} x_n = a$, es té $\lim_{n\to\infty} F(x_n) = F(a)$

Diem que F és un *operador continu* sobre D quan és continu en tots els punts de D. Diem també que és *uniformement continu* sobre D si es compleix

$$\forall \epsilon > 0 \, \exists \delta_{\epsilon} > 0 : \left(x, y \in D,\, \|x - y\| \leq \delta_{\epsilon} \Rightarrow \left\| F(x) - F(y) \right\| \leq \epsilon \right)$$

Diem que l'operador F *és de Lipschitz* sobre D si hi ha $M > 0$ tal que $\forall x, y \in D$, $\left\| Fx - Fy \right\| \leq M \|x - y\|$.

Observació 2.3. En un espai normat $\left(X, \|\cdot\|\right)$, la norma $\|\cdot\| : X \mapsto \mathbb{R}$ sempre és (uniformement) contínua, degut a (2.1).

Observació 2.4. Les propietats algebraiques i de composició dels límits de successions i operadors, i de la continuïtat d'operadors, són les que s'esperen: les que coneixem a K^n.

2.2 Operadors lineals continus

Ens interessem pels operadors entre espais normats que tenen un bon comportament algebraic i topològic: els operadors lineals continus.

Quant a l'aspecte algebraic, un operador lineal entre espais normats X i Y (sempre sobre el mateix cos $K = \mathbb{R}$ o $K = \mathbb{C}$) és simplement una aplicació lineal entre ells com a espais vectorials. Quant a l'aspecte topològic, quan els espais normats són de dimensió finita, els operadors lineals són automàticament continus; malauradament, això no és necessàriament cert per a espais de dimensió no finita.

Necessitem, per tant, caracteritzar la continuïtat dels operadors lineals.

2.2.1 Continuïtat

Degut a la linealitat, ja sospitem certes condicions equivalents: continuïtat i continuïtat en $\vec{0}$, i continuïtat en $\vec{0}$ i ser de Lipschitz. El punt clau està en la circumferència de l'espai X, $\left\{x \in X : \|x\| = 1\right\}$.

Proposició 2.5 (Caracterització continuïtat)**.** *Sigui $L\colon X \mapsto Y$ un operador lineal entre espais normats. Són equivalents:*

1. *L és continu*
2. *L és continu en $\vec{0}$*
3. $\exists M > 0 : \left(\forall x \in X,\, \|Lx\| \leq M \|x\| \right)$
4. *L és de Lipschitz*

5. *L és uniformement continu*
6. *L està fitat en* $\{x \in X : \|x\| = 1\}$
7. *L està fitat en* $\{x \in X : \|x\| \le 1\}$
8. *L envia conjunts fitats a conjunts fitats (preserva l'acotació).*

La proposició recull un grapat de caracteritzacions, de les quals 3, 6 i 7 són les que més es fan servir. Les importants a demostrar són la 3 i la 6, la resta surten de jugar amb elles.

Demostració. La prova consisteix a veure que 1, 2, 3, 4 i 5 són equivalents, i que 3, 6, 7 i 8 també ho són; això prova que totes 8 són equivalents.

$\underline{1 \Rightarrow 2}$ Obvi.

$\underline{2 \Rightarrow 3}$ La continuïtat en $\vec{0}$ ens diu que hi ha $\delta > 0$ tal que si $x \in X$ i $\|x\| \le \delta$, aleshores $\|L(x)\| \le 1$.

Per a cada $x \ne \vec{0}$ definim $u = \frac{x}{\|x\|}$, així $x = \|x\|\, u$ amb $\|u\| = 1$. Llavors tenim $\|\delta u\| \le \delta$ i per tant

$$\|Lx\| = \|x\| \|Lu\| = \|x\| \frac{1}{\delta} \|L(\delta u)\| \le \|x\| \frac{1}{\delta}$$

d'on, prenent $M = \frac{1}{\delta}$, deduïm que $\|Lx\| \le M \|x\|$

$\underline{3 \Rightarrow 4}$ Siguin $u, v \in X$ i definim $w = u - v$. Aleshores, per hipòtesi i la linealitat de L, sabem que

$$\|Lu - Lv\| = \|Lw\| \le M \|w\| = M \|u - v\|$$

i per tant L compleix la condició de Lipschitz.

$\underline{4 \Rightarrow 5 \Rightarrow 1}$ Obvi

$\underline{3 \Rightarrow 8 \Rightarrow 7 \Rightarrow 6}$ Obvi

$\underline{6 \Rightarrow 3}$ Sigui M una fita superior del conjunt $\{\|Lx\| : x \in X, \|x\| = 1\}$. Assumim $x \ne \vec{0}$, fem $u = \frac{x}{\|x\|}$, aleshores $x = \|x\|\, u$ amb $\|u\| = 1$. Per la linealitat de L tenim

$$\|Lx\| = \|x\| \|Lu\| \le \|x\| M$$

Si $x = \vec{0}$ també es compleix la desigualtat ja que $L(\vec{0}) = \vec{0}$. □

Aquesta proposició motiva que als operadors lineals continus se'ls anomeni també *operadors lineals fitats*, i introduir el concepte de norma d'un operador lineal fitat.

Definició 2.6 (Continuïtat)**.** Un operador lineal $L\colon X \mapsto Y$ entre espais normats és continu (o fitat) si compleix qualsevol de les condicions de la proposició 2.5.

Denotem per $\mathscr{L}(X, Y)$ l'espai dels operadors lineals de X en Y, i per $\mathscr{B}(X, Y)$ el subespai dels lineals continus.

Resulta curiós que la continuïtat —concepte topològic— d'un operador lineal quedi caracteritzada per preservar l'acotació, que no és una propietat topològica (no és invariant per homeomorfismes).

Definició 2.7 (Norma)**.** Sigui $L \in \mathscr{B}(X,Y)$, definim la seva norma, $\|L\|$, de les següents formes equivalents[1]

$$\begin{aligned}\big\|L\big\| &= \inf\big\{M \geq 0 : \forall x \in X,\ \|Lx\| \leq M\|x\|\big\} = \sup\Big\{\|Lx\| \big/ \|x\| \ :\ x \in X\backslash\{\vec{0}\}\Big\}\\ &= \sup\big\{\|Lx\| \ :\ x \in X,\ \|x\| \leq 1\big\} = \sup\big\{\|Lx\| \ :\ x \in X,\ \|x\| = 1\big\}\end{aligned}$$

amb aquesta norma $\mathscr{B}(X,Y)$ esdevé un espai normat i es té $\forall x \in X, \|Lx\| \leq \|L\|\,\|x\|$.

Observació 2.8. La norma en l'espai $\mathscr{B}(X,Y)$ depèn de les normes de X i Y: si es canvia alguna d'elles tenim una norma diferent per a les aplicacions lineals.

Observació 2.9. L'espai $\mathscr{B}(X,Y)$ pot, o no, ser de Banach, però si Y ho és, aleshores $\mathscr{B}(X,Y)$ també ho és (veure tema 2 de [17] o [13, sec. 2]).

Observació 2.10. Quan $X = Y = K^n$ les aplicacions lineals són matrius, i la norma de l'aplicació és exactament el que en anàlisi numèrica anomenem norma matricial subordinada.

2.2.2 Propietats de la norma

Les normes dels operadors lineals es comporten bé amb la *composició d'operadors*. Si $L_1 \in \mathscr{B}(X,Y)$ i $L_2 \in \mathscr{B}(Y,Z)$, la composició $L = L_2L_1 \colon X \mapsto Z$ té norma $\|L\| \leq \|L_2\|\,\|L_1\|$ ja que si $x \in X$

$$\big\|L(x)\big\| = \big\|L_2L_1(x)\big\| \leq \big\|L_2\big\|\big\|L_1(x)\big\| \leq \big\|L_2\big\|\big\|L_1\big\|\,\|x\|$$

òbviament, sempre que considerem en Y la mateixa norma per a L_1 i L_2.

Si $L \in \mathscr{B}(X,X)$ és injectiu i $\dim X = \infty$, L no necessàriament és exhaustiu. L'*operador invers*, $L^{-1} \in \mathscr{L}(Z,X)$, $Z = L(X)$, malgrat ésser lineal, no necessàriament és continu.

Exemple 2.11. En l'espai de les successions sobre el cos K, $K^{\mathbb{N}} = \big\{(x_n)_{n\geq 0} \subseteq K\big\}$, l'operador $(x_0, x_1, \ldots) \mapsto (0, x_0, x_1, \ldots)$ és lineal i injectiu, però no exhaustiu.

Exemple 2.12. En l'espai de les successions complexes quasi-nul·les (tots els termes són zero llevat, com a màxim, un nombre finit d'ells), φ, amb la norma $\|z\|_\infty = \max\{|z_n| : n \geq 1\}$, definim l'operador lineal $L\colon (\varphi, \|\cdot\|_\infty) \mapsto (\varphi, \|\cdot\|_\infty)$ donada per

$$L(z_1, z_2, z_3, \ldots) = (z_1, \tfrac{1}{2}z_2, \tfrac{1}{3}z_3, \ldots)$$

Clarament és bijectiu i $\|L\| = 1$, però l'operador invers $L^{-1}(y_1, y_2, y_3, \ldots) = (y_1, 2y_2, 3y_3, \ldots)$ no és un operador lineal fitat ja que, si $e_n = (e^1, e^2, e^3, \ldots) \in \varphi$ és el vector tal que $e^j = 0$ quan $j \neq n$ i $e^n = 1$, aleshores $\|L^{-1}(e_n)\| / \|e_n\| = \|L^{-1}(e_n)\| = n\,\|e_n\| = n$.

Un resultat per assegurar la continuïtat de la inversa d'un operador lineal continu, és el famós Teorema del Punt Fix de Banach[2] en la seva versió per a operadors lineals. Demostracions en [1], [16, cap. 2] i [13, sec. 3].

Proposició 2.13. *Sigui X espai de Banach i $L \in \mathscr{B}(X,X)$. Es té*

[1] Es demostra fàcilment a partir dels raonaments fets en la proposició 2.5. Es pot trobar una demostració detallada en [13, sec. 2.2] i en el tema 2 de [17].

[2] El teorema, en la versió general i la aquí enunciada, ja es troba en el seu article del 1922, *Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales*; que hagué de publicar com a requisit per a obtenir el seu doctorat en la Universitat de Lwow.

1. Si $\|L\| < 1$*, aleshores* $Id - L$ *és bijectiu, l'invers és continu i*

$$\left\|(Id - L)^{-1}\right\| \leq \frac{1}{1 - \|L\|}$$

2. Si $\left\|Id - L\right\| < 1$*, aleshores* L *és bijectiu, l'invers és continu i*

$$\left\|L^{-1}\right\| \leq \frac{1}{1 - \|Id - L\|}$$

2.3 Derivació d'operadors

Introduïm els conceptes de derivada de Gateaux i Fréchet d'un operador i en presentem les propietats bàsiques, a destacar sobretot la regla de la cadena. Una exposició més detallada es pot trobar en [16, cap. 3] per a $\mathbb{R}^n$ i [12, cap. XVII] per a espais de Banach arbitraris.

2.3.1 Notació de Landau

La notació de la «o petita» és una manera concisa i àgil d'expressar la velocitat amb què una aplicació convergeix al seu límit. Per la seva utilitat, n'estenem l'ús a espais normats en general.

Definició 2.14 (o petita de Landau)**.** Siguin X i Y espais normats, i $F, f\colon D \mapsto Y$, $D \subseteq X$. Suposem $\vec{0} \in D'$ (D' s'ha definit en 2.1.3), i que $\lim_{x\to\vec{0}} f(x) = \vec{0}$.

Diem que «F és o petita de f en $x = \vec{0}$», denotat per $F(x) = \mathrm{o}\big(f(x)\big)$, quan $\|F(x)\| = \mathrm{o}\big(\|f(x)\|\big)$. És a dir

$$\forall \epsilon > 0\, \exists \delta > 0 : \big(x \in D, 0 < \|x\| \leq \delta \Rightarrow \|F(x)\| \leq \epsilon \|f(x)\|\big)$$

En síntesi: $\|F(x)\| \big/ \|f(x)\|$ té límit 0 quan x s'atansa a $\vec{0}$.

Un ús habitual d'aquesta notació és expressar una funció desconeguda: $f(x) = g(x) + \mathrm{o}\big(h(x)\big)$, que s'ha d'entendre com que, prop de $x = \vec{0}$, l'error d'aproximar $f(x)$ per $g(x)$ és una funció desconeguda que tendeix a $\vec{0}$ més ràpid que $h(x)$. És a dir, $f(x) - g(x) = \mathrm{o}\big(h(x)\big)$.

Un abús de notació que farem serà escriure coses com $\mathrm{o}\big(f(x)^n\big)$ i $\mathrm{o}\Big(t^3\Big)$; entenent que la primera denota una funció desconeguda, $g(x)$, tal que $\|g(x)\| = \mathrm{o}\big(\|f(x)\|^n\big)$, i la segona una funció desconeguda $g(t)$, $t \in K$, tal que $\|g(t)\| = \mathrm{o}\Big(|t^3|\Big)$.

Els següents exemples mostren com fer servir aquesta notació; a més, ens seran útils en la demostració de la regla de la cadena, en l'apartat 2.3.3.

Exemple 2.15. Siguin $g\colon A \mapsto X$, $A \subseteq K$, tal que $0 \in A'$, i $L \in \mathscr{B}(X, Y)$. Suposem $g(t) = \mathrm{o}\Big(t^3\Big)$. Sabem que, fixat $\epsilon > 0$, hi ha $\delta > 0$ tal que $\big(t \in A, 0 < |t| \leq \delta \Rightarrow \|g(t)\| \leq \frac{\epsilon}{\|L\|}|t^3|\big)$, per tant

$$\left\|L(g(t))\right\| \leq \|L\|\,\|g(t)\| \leq \epsilon |t^3|$$

Hem provat doncs $L(g(t)) = \mathrm{o}\Big(t^3\Big)$, en notació $L\Big(\mathrm{o}\Big(t^3\Big)\Big) = \mathrm{o}\Big(t^3\Big)$.

Exemple 2.16. Siguin $f, g\colon A \mapsto X$, $A \subseteq K$, tal que $0 \in A'$. Suposem que $f(t) = \mathrm{o}\big(tv + g(t)\big)$ i $g(t) = \mathrm{o}(t)$, aleshores fixat $1 > \epsilon > 0$ hi ha $\delta > 0$ tal que si $0 < |t| \leq \delta$, $t \in A$, es té

$$\begin{aligned}
&\left\|g(t)\right\| \leq \epsilon |t| \leq |t| \\
&\left\|f(t)\right\| \leq \frac{\epsilon}{\|v\| + 1}\left\|tv + g(t)\right\| \leq \frac{\epsilon}{\|v\| + 1}|t|\big(\|v\| + 1\big) = \epsilon |t|
\end{aligned}$$

cosa que prova $f(t) = \mathrm{o}(t)$; en notació $\mathrm{o}\big(tv + \mathrm{o}(t)\big) = \mathrm{o}(t)$.

2.3.2 Derivades direccional, de Gateaux i de Fréchet

La derivada d'una funció $f\colon I \mapsto \mathbb{R}$, $I \subseteq \mathbb{R}$ obert, en un punt $a \in I$, denotada per $f'(a)$, és el límit del quocient incremental

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} = f'(a)$$

que geomètricament es correspon amb el pendent de la recta tangent a f en el punt a. Ho podem interpretar com associar la recta $l(h) = f(a) + f'(a)h$ al punt a, i que és la recta que millor aproxima f en el punt a: $f(a+h) - l(h) = \mathrm{o}\big(h\big)$.

Aquesta última visió és la que es generalitza a aplicacions $f\colon D \mapsto K^m$, $D \subseteq K^n$ obert; i aquí reproduïm el mateix camí: derivada direccional, derivada en totes direccions (G-derivabilitat), i derivada total (F-derivabilitat).

Definició 2.17. Sigui un operador $F\colon D \mapsto Y$, $D \subseteq X$, X i Y espais normats, i a un punt interior de D. Definim els següents conceptes

1. *Derivada direccional.* La derivada de F en a en la direcció $h \in X \setminus \{\vec{0}\}$, denotada per $\partial_h F(a)$ o $F'_h(a)$, és el límit del quocient incremental

   $$\lim_{t\to 0} \frac{1}{t}\big(F(a+th) - Fa\big) = \partial_h F(a)$$

   equivalentment

   $$F(a+th) - Fa - tF'_h(a) = \mathrm{o}(t)$$

2. *Derivada de Gateaux.* Diem que F és G-derivable en a quan totes les derivades direccionals en a existeixen i es poden donar per un operador lineal continu. És a dir, quan existeix $L \in \mathscr{B}(X,Y)$ tal que per a tot $h \in X \setminus \{\vec{0}\}$

   $$\lim_{t\to 0} \frac{1}{t}\big\|F(a+th) - Fa - tL(h)\big\| = 0$$

   equivalentment

   $$F(a+th) - Fa - tL(h) = \mathrm{o}(t)$$

3. *Derivada de Fréchet.* Diem que F és F-derivable en a quan hi ha un operador lineal que estima qualsevol increment, no només sobre rectes. És a dir, quan hi ha $L \in \mathscr{B}(X,Y)$ tal que

   $$\lim_{h\to \vec{0}} \frac{1}{\|h\|_X}\big\|F(a+h) - Fa - Lh\big\|_Y = 0$$

   equivalentment

   $$F(a+h) - Fa - Lh = \mathrm{o}\big(h\big)$$

Les derivades de Gateaux i de Fréchet es denoten de la mateixa manera $L = F'(a)$.

Observació 2.18. Per als operadors $g\colon A \mapsto X$, $A \subseteq K$, tots tres conceptes de derivabilitat són equivalents, ja que $\dim K = 1$. Direm senzillament que són derivables.

Observació 2.19 (Derivació sobre intervals). Per als teoremes de Taylor i Lagrange (apartat 2.4), voldrem derivar operadors $g\colon [0,1] \mapsto Y$, on $[0,1] \subseteq \mathbb{R}$ i Y un espai normat sobre el cos K. En aquests cas, formalment, pot haver-hi un problema; si Y és un espai complex el concepte de derivada $g'(t_0)$ no té sentit: hauria d'ésser un operador lineal entre $\mathbb{R}$ i Y, espais normats sobre cossos diferents. No obstant, els cossos $\mathbb{R}$ i $\mathbb{C}$ estan estretament relacionats i això no suposa realment un problema, ja que tot espai vectorial complex és també un espai vectorial real. Per tant, a l'hora de derivar l'operador g, si Y és un espai normat real ($K = \mathbb{R}$), no hi ha cap problema i s'aplica l'observació anterior; en canvi, si $K = \mathbb{C}$ mirem Y com a espai normat real i procedim com si $K = \mathbb{R}$.

La derivada 'bona' és la derivada en sentit de Fréchet, ja que és la que reté les propietats desitjables de la derivada que coneixem a K i K^n. La relació entre els tipus de derivada que hem donat i la continuïtat és la mateixa que en K^n

$$\text{derivada direccional} \underset{\Leftarrow}{\overset{\not\Rightarrow}{}} \text{G-derivada} \underset{\Leftarrow}{\overset{\not\Rightarrow}{}} \text{F-derivada} \underset{\not\Leftarrow}{\overset{\Rightarrow}{}} \text{continuïtat}$$

Resumim algunes propietats rellevants de les derivades en la següent proposició, que es demostren fàcilment. El punt 8 sí es demostra en l'apartat 2.4, com a corol·lari del Teorema del Valor Mitjà de Lagrange.

Proposició 2.20. *Sigui $F\colon D \mapsto Y$, $D \subseteq X$, i $a \in D$ un punt interior.*

1. *L'operació de derivar en un punt és lineal (en tots tres sentits).*
2. *Si existeix la derivada direccional $\partial_h F(a)$, aleshores $G(t) = F(a+th)$ és continu en $t = 0$.*
3. *Si existeix la derivada direccional $\partial_h F(x)$ per a tot $x \in (a, a+h)$, aleshores $G(t) = F(a+th)$ és continu en $(0,1)$.*
4. *Si existeix la G-derivada en a, aleshores, per a tot $h \in X \setminus \{\vec{0}\}$, $G(t) = F(a+th)$ és continu en $t = 0$.*
5. *Si F és G-derivable en un obert convex $C \subseteq D$, aleshores $G(t) = F(a+t(b-a))$ és continu en $[0,1]$ per a qualssevol $a, b \in C$.*
6. *Si existeix la F-derivada en a, aleshores F és continu en $x = a$.*
7. *Si existeix la F-derivada en a, aleshores també existeix la G-derivada en a i ambdues coincideixen.*
8. *Si F és G-derivable en un entorn obert de a i F' és continu en a, aleshores F és F-derivable en a.*

Els punts 7 i 8 motiven introduir el concepte d'operador de classe $\mathscr{C}^1$.

Definició 2.21. Diem que un operador $F\colon D \mapsto Y$, $D \subseteq X$ obert, és de classe $\mathscr{C}^1$, quan és derivable (en sentit de Gateaux o Fréchet) i F' és contínua.

2.3.3 Regla de la cadena i exemples

Una propietat important i que necessitarem per a obtenir els teoremes de Taylor y Lagrange és la regla de la cadena.

Proposició 2.22. *Siguin X, Y i Z espais normats, $C \subseteq X$, $D \subseteq Y$. Suposem $C \xrightarrow{G} D \xrightarrow{F} Z$ amb F F-derivable en $G(a) \in D$ i G G-derivable en $a \in C$.*

Aleshores FG és G-derivable en a i $(FG)'(a) = F'(G(a))G'(a)$.

Demostració. Volem provar que, per a tot $h \in X \setminus \{\vec{0}\}$, es té

$$FG(a+th) = FG(a) + t\big(F'(G(a))G'(a)\big)h + o(t)$$

Per hipòtesi tenim ($y = G(a)$)

$$G(a+th) = Ga + tG'(a)h + o(t) \qquad\qquad F(y+h) = Fy + F'(y)h + o\big(h\big)$$

Per tant, si fixem $h \in X$, $h \neq \vec{0}$, en un entorn de $t = 0$ tenim $(v = G'(a)h)$

$$\begin{aligned}
FG(a+th) &= F\big(G(a+th)\big) = F\big(G(a) + G'(a)(th) + \mathrm{o}(t)\big) \\
&= F\big(y + tv + \mathrm{o}(t)\big) = F(y) + F'(y)\big(tv + \mathrm{o}(t)\big) + \mathrm{o}\big(tv + \mathrm{o}(t)\big) \\
&= F(y) + tF'(y)v + F'(y)\big(\mathrm{o}(t)\big) + \mathrm{o}(t) \\
&= F(y) + tF'(y)v + \mathrm{o}(t) + \mathrm{o}(t) \\
&= FG(a) + t\big(F'(G(a))G'(a)\big)h + \mathrm{o}(t)
\end{aligned}$$

com volíem. Els punts claus són la segona igualtat (G-derivabilitat de G) i la quarta (F-derivabilitat de F); després fem servir $\mathrm{o}(t) + \mathrm{o}(t) = \mathrm{o}(t)$, $\mathrm{o}\big(tv + \mathrm{o}(t)\big) = \mathrm{o}(t)$ i $F'(y)\big(\mathrm{o}(t)\big) = \mathrm{o}(t)$; això es prova de manera similar a com s'ha fet en els exemples de la l'apartat 2.3.1. □

Corol·lari 2.23. *En les hipòtesis de la proposició anterior. Si F és F-derivable en $G(a)$ i G és F-derivable en a, aleshores FG és F-derivable en a i $(FG)'(a) = F'(G(a))G'(a)$.*

La composició d'operadors F-derivables és F-derivable.

Les hipòtesis no es poden afeblir: no és cert, en general, que la composició d'operadors G-derivables sigui G-derivable (un exemple en [16, pàg. 63]). Fem alguns casos particulars d'aquests resultats, que necessitarem més endavant, en l'apartat 2.4.

Exemple 2.24. Sigui $F\colon D \mapsto Y$, $D \subseteq X$ obert. Si F és G-derivable en $(a,b) \subseteq D$, aleshores $f(t) = F(a+th)$, $h = b - a$, $t \in (0,1)$, és derivable en $(0,1)$ amb $f'(t) = F'(a+th)h = \partial_h F(a+th)$. Fem $u = a + t_0 h$, $s \in \mathbb{R}$

$$\begin{aligned}
f(t_0 + s) &= F(a + (t_0+s)h) = F(u + sh) = F(u) + F'(u)(sh) + \mathrm{o}(s) \\
&= f(t_0) + sF'(a+t_0h)h + \mathrm{o}(s) = f(t_0) + s\partial_h F(a+t_0h) + \mathrm{o}(s)
\end{aligned}$$

d'on obtenim, com volíem, $f'(t_0) = F'(a+t_0h)h$. Això no contradiu els resultats anteriors, perquè en pertorbar la variable t continuem restant a la recta $\{a + \lambda h\}_{\lambda \in K}$; per tant, de fet estem calculant una derivada direccional.

Exemple 2.25. Sigui $\phi\colon K \mapsto X$, $\phi(t) = a + th$, $a, h \in X$. La derivada en un punt t_0 és un operador lineal $\phi'(t_0) \in \mathscr{B}(K,X) \cong X$. ϕ és derivable en tot punt i $\phi'(t) = h$, $\forall t \in K$.

$$\phi(t_0+s) = a + (t_0+s)h = a + t_0h + sh = \phi(t_0) + sh + \vec{0} = \phi(t_0) + sh + \mathrm{o}(s)$$

Exemple 2.26. Si $L \in \mathscr{B}(X,Y)$, la F-derivada en $x \in X$ és un operador lineal $T = L'(x) \in \mathscr{B}(X,Y)$ tal que

$$L(x+h) - L(x) - T(h) = \mathrm{o}\big(h\big)$$

si prenem $T = L$ es verifica la definició de F-derivabilitat, ja que $\vec{0} = \mathrm{o}\big(h\big)$. Per tant, l'operador lineal L té F-derivada constant: $\forall x \in X$, $L'(x) = L$

Exemple 2.27. Siguin $\mathbb{R} \xrightarrow{\phi} D \xrightarrow{F} Y$, $D \subseteq X$ obert, $\phi(t) = a + th$. Si prenem $L \in \mathscr{B}(Y,Z)$, i assumint que F és F-derivable; per la regla de la cadena tenim

$$\big(LF\phi\big)'(t) = \big(LF\big)'(\phi(t))\phi'(t) = \big(LF'(a+th)\big)h$$

si Y, Z són espais normats sobre $\mathbb{C}$, això continua éssent vàlid ja que $\phi(t)$ es pot estendre a $\Phi(z) = a + zh$, $z \in \mathbb{C}$, (observació 2.19).

2.3.4 Derivada segona d'un operador

Si un operador $F\colon D \mapsto Y$, $D \subseteq X$ obert, és derivable (en el sentit de Gateaux o de Fréchet); podem definir l'operador de derivació, que assigna a cada punt de D la derivada de F en el punt. És a dir, un operador $D \mapsto \mathscr{B}(X,Y)$, $x \mapsto F'(x)$; que denotarem per F'.

Aquest nou operador pot ser, també, derivable en algun punt a de l'obert D; en tal cas, la derivada de F' en a és el que anomenem la *derivada segona* de l'operador F en a, que denotarem per $(F')'(a) = F''(a)$. Si l'operador F' és derivable en l'obert D, podem replicar el mateix procés que hem fet amb F, i definir l'operador $F''\colon D \mapsto \mathscr{B}\big(X, \mathscr{B}(X,Y)\big)$, que assigna a cada punt la derivada segona de F en el punt.

Com calculem la norma de la segona derivada? Bé, $F''(a) \in \mathscr{B}\big(X, \mathscr{B}(X,Y)\big)$, per tant, $\forall h \in X$, $F''(a)h \in \mathscr{B}(X,Y)$ i així $\|F''(a)h\| = \sup_{\|k\|=1} \left\|F''(a)hk\right\|$ d'on deduïm

$$\left\|F''(a)\right\| = \sup_{\|h\|=1} \left\|F''(a)h\right\| = \sup_{\|h\|=1} \sup_{\|k\|=1} \left\|F''(a)hk\right\| \tag{2.2}$$

i per tant $\forall h, k \in X$

$$\left\|F''(a)hk\right\| \le \left\|F''(a)h\right\| \|k\| \le \left\|F''(a)\right\| \|h\| \|k\| \tag{2.3}$$

Des d'un punt de vista pràctic, el que fem és pensar la segona derivada en un punt com una aplicació bilineal $F''(a) \in \mathscr{L}(X \times X, Y)$, i li associem una norma per mitjà de fixar una variable.

2.4 Teoremes de Lagrange i Taylor

Dos resultats importants en càlcul diferencial d'una variable real són, sense dubte, el Teorema del Valor Mitjà de Lagrange i el Teorema de Taylor. Volem portar aquests resultats a espais normats en general, per tal de treure'n profit també en ells.

Aquests teoremes es poden obtenir, com a mínim, per tres vies. La primera, a partir del teorema de Hahn-Banach i la regla de la cadena; que és el que originalment va fer Kantorovich [6, 7, 8] (traduïts a l'anglès en [11]) i que també fa en [12, sec. XVII.1.1]. La segona, a partir del Teorema Fonamental del Càlcul, que requereix reproduir la integral de Riemann per a operadors $f\colon [a,b] \mapsto Y$, $[a,b] \subseteq X$, amb Y espai de Banach (per exemple [1, sec. 20]). La tercera, que és la que prendrem, s'assembla a la segona però evita fer ús explícit de la integral.

Les dues primeres tenen l'avantatge d'ésser més semblants a allò que faríem en dimensió finita, tant en raonament com en notació. La tercera, però, té l'avantatge que només requereix la norma, no cal introduir nous conceptes. La feina es concentra en el següent lema[3], que és bàsicament 'integrar' i fitar, però evitant els següents passos intermedis amb integrals

$$\left\|g(1) - g(0)\right\| = \left\|\int_0^1 g'(t)\,\mathrm{d}t\right\| \le \int_0^1 \left\|g'(t)\right\| \mathrm{d}t \le \int_0^1 \alpha'(t)\,\mathrm{d}t = \alpha(1) - \alpha(0)$$

Lema 2.28. *Siguin $g\colon [0,1] \mapsto Y$ i $\alpha\colon [0,1] \mapsto \mathbb{R}$, Y espai normat. Si g i α són contínues en $[0,1]$, derivables en $(0,1)$ i $\forall t \in (0,1)$, $\|g'(t)\| \le \alpha'(t)$; aleshores es compleix*

$$\left\|g(1) - g(0)\right\| \le \alpha(1) - \alpha(0)$$

[3] El lema és una versió lleugerament modificada d'un resultat present en els apunts de Rafael Payá Albert del curs «análisis matemático I», tema 10. Disponible al seu web personal de la Universidad de Granada, https://www.ugr.es/~rpaya/.

Recordem que Y pot ésser un espai normat no real, sinó complex; en tal cas donem sentit a la derivada $g'(t_0)$ pensant l'espai Y com a espai real, no complex (veure observació 2.19). Això no afecta les demostracions, només és una formalitat per no haver de distingir quan Y és un espai normat sobre $K = \mathbb{R}$ o $K = \mathbb{C}$, i obtenim alhora els resultats en el cas $K = \mathbb{R}$ i $\mathbb{C}$.

Demostració. Per a cada $\epsilon > 0$ definim el conjunt $A_\epsilon \subseteq [0,1]$ dels nombres reals $0 \leq t \leq 1$ tals que

$$\left\| g(t) - g(0) \right\| \leq \alpha(t) - \alpha(0) + \epsilon t \tag{2.4}$$

si aconseguim provar que $\forall \epsilon > 0$, $\sup A_\epsilon = 1$, haurem provat el lema.

Raonarem per contradicció: suposem que hi ha $\epsilon > 0$ tal que $s_\epsilon = \sup A_\epsilon < 1$. El conjunt A_ϵ no és buit, ja que $0 \in A_\epsilon$, i està fitat; per tant $s_\epsilon = \sup A_\epsilon$ té sentit. El que hi ha a banda i banda de (2.4) és continu, així doncs, per continuïtat, la desigualtat se satisfà en un entorn de $t = 0$ —per tant $s_\epsilon > 0$— i també en s_ϵ. Tenim aleshores

$$\left\| g(s_\epsilon) - g(0) \right\| \leq \alpha(s_\epsilon) - \alpha(0) + \epsilon s_\epsilon$$

Arribarem a una contradicció veient que hi ha $r \in A_\epsilon$, $r > s_\epsilon$. Sigui $r \in (s_\epsilon, 1]$

$$\begin{aligned} \left\| g(r) - g(0) \right\| &\leq \left\| g(r) - g(s_\epsilon) \right\| + \left\| g(s_\epsilon) - g(0) \right\| \\ &\leq \left\| g(r) - g(s_\epsilon) \right\| + \alpha(s_\epsilon) - \alpha(0) + \epsilon s_\epsilon \end{aligned} \tag{2.5}$$

prenem $h = r - s_\epsilon$; la derivabilitat de g i α en $t = s_\epsilon \in (0,1)$ ens permet assegurar que hi ha $\delta > 0$ tal que si $1 \geq s_\epsilon + \delta \geq r > s_\epsilon$, aleshores

$$\left\| g(s_\epsilon + h) - g(s_\epsilon) - g'(s_\epsilon)h \right\| \leq \frac{\epsilon h}{2} \qquad \left| \alpha(s_\epsilon + h) - \alpha(s_\epsilon) - \alpha'(s_\epsilon)h \right| \leq \frac{\epsilon h}{2}$$

d'on deduïm

$$\left\| g(r) - g(s_\epsilon) \right\| = \left\| g(s_\epsilon + h) - g(s_\epsilon) \right\| \leq \frac{\epsilon h}{2} + \left\| g'(s_\epsilon)h \right\| \leq \frac{\epsilon h}{2} + \alpha'(s_\epsilon)h$$

posant això en (2.5) tenim

$$\begin{aligned} \left\| g(r) - g(0) \right\| &\leq \frac{\epsilon h}{2} + \alpha'(s_\epsilon)h + \alpha(s_\epsilon) - \alpha(0) + \epsilon s_\epsilon \\ &= \alpha'(s_\epsilon)h + \alpha(s_\epsilon) - \alpha(0) + \frac{\epsilon h}{2} + \epsilon s_\epsilon \\ &= -\big(\alpha(s_\epsilon + h) - \alpha(s_\epsilon) - \alpha'(s_\epsilon)h\big) + \alpha(r) - \alpha(0) + \frac{\epsilon h}{2} + \epsilon s_\epsilon \\ &\leq \frac{\epsilon h}{2} + \alpha(r) - \alpha(0) + \frac{\epsilon h}{2} + \epsilon s_\epsilon \\ &= \alpha(r) - \alpha(0) + \epsilon r \end{aligned}$$

Concloem així que $(s_\epsilon, s_\epsilon + \delta] \subseteq A_\epsilon$, contradient que s_ϵ era el suprem de A_ϵ. Per tant $\sup A_\epsilon = 1$ per a tot $\epsilon > 0$ i el lema queda provat. $\square$

2.4.1 Teorema del Valor Mitjà de Lagrange

El teorema del valor mig de Lagrange per a una funció $f\colon [a,b] \mapsto \mathbb{R}$, $[a,b] \subset \mathbb{R}$, és (2.6), on $c \in (a,b)$; sovint, però, es fa servir (2.7), on $M = \sup\{|f'(x)| : x \in (a,b)\}$, que n'és una versió més feble.

$$f(b) - f(a) = f'(c)(b-a) \tag{2.6}$$
$$\left|f(b) - f(a)\right| \leq M|b-a| \tag{2.7}$$

El teorema es pot generalitzar fàcilment a funcions escalars $f\colon C \mapsto \mathbb{R}$, $C \subseteq \mathbb{R}^n$ obert convex. En canvi és fals per a funcions $f\colon C \mapsto \mathbb{R}^m$, $C \subseteq \mathbb{R}^n$ obert convex. Un exemple senzill és la funció $f\colon [0,2\pi] \mapsto \mathbb{R}^2$, $f(t) = (\cos(t), \sin(t))$; no hi ha cap $c \in (0,2\pi)$ tal que $f(2\pi) - f(0) = f'(c)(2\pi - 0)$.

Afortunadament, però, la desigualtat (2.7) sí es manté en espais normats en general.

Teorema 2.29. *Siguin $F\colon C \mapsto Y$, $C \subseteq X$ obert convex, X i Y dos espais normats. Si F és G-derivable en C, aleshores per a qualssevol $a, b \in C$ es compleix*

$$\left\|F(b) - F(a)\right\| \leq \sup_{x \in (a,b)} \left\|F'(x)\right\| \|b-a\| \tag{2.8}$$

Demostració. Assumim $M = \sup\{\|F'(x)\| : x \in (a,b)\} < \infty$; altrament (2.8) es compleix trivialment.

Definim $h = b - a$ i $g\colon [0,1] \mapsto Y$, $g(t) = F(a + th)$. L'operador g és continu en $[0,1]$ per ésser F G-derivable en un entorn de $[a,b]$ —proposició 2.20—; per la mateixa raó, també és derivable en $(0,1)$ —veure exemples apartat 2.3.3—. Derivant tenim

$$g'(t) = F'(a+th)h \qquad\qquad \left\|g'(t)\right\| \leq \left\|F'(a+th)\right\| \|h\| \leq M\|h\|$$

Definint $\alpha(t) = tM\left\|h\right\|$, $t \in [0,1]$ i aplicant el lema 2.28 obtenim (2.8)

$$\left\|F(b) - F(a)\right\| = \left\|g(1) - g(0)\right\| \leq \alpha(1) - \alpha(0) = M\|h\|$$

□

Aquest teorema té dues aplicacions interessants i immediates. Una ens servirà per a analitzar el mètode de Newton simplificat, i l'altra és una condició de suficiència per a la F-derivabilitat: el punt 8 de la proposició 2.20.

Proposició 2.30. *Siguin $F\colon C \mapsto Y$, $C \subseteq X$ obert convex, X i Y dos espais normats. Si F és G-derivable, aleshores per a qualssevol $x, y, z \in C$ es compleix*

$$\left\|F(z) - F(y) - F'(x)(z-y)\right\| \leq \sup_{\overline{x} \in (z,y)} \left\|F'(\overline{x}) - F'(x)\right\| \|z-y\| \tag{2.9}$$

Es tracta d'aplicar el teorema 2.29 a l'operador $G(w) = F(w) - F'(x)w$. Clarament G satisfà les condicions del teorema 2.29 i $G'(w) = F'(w) - F'(x)$. Per tant (2.9) és precisament

$$\left\|G(z) - G(y)\right\| \leq \sup_{\overline{x} \in (z,y)} \left\|G'(\bar{x})\right\| \|z-y\|$$

Proposició 2.31 (Punt 8 proposició 2.20)**.** *Siguin $F\colon C \mapsto Y$, $C \subseteq X$ obert, X i Y dos espais normats. Si F és G-derivable i F' és contínua en $x \in C$, aleshores F és F-derivable en x.*

Demostració. Ens restringim a un entorn obert i convex del punt x. La continuïtat de F' en x vol dir que, fixat $\epsilon > 0$, hi ha $\delta > 0$ tal que $\|h\| \le \delta \Rightarrow \left\|F'(x+h) - F'(x)\right\| \le \epsilon$.

Fent servir (2.9) tenim el següent per a tot $h \in B(\vec{0}, \delta)$

$$\left\|F(x+h) - F(x) - F'(x)h\right\| \le \sup_{\overline{x} \in (x, x+h)} \left\|F'(\bar{x}) - F'(x)\right\| \|h\| \le \epsilon \|h\|$$

que prova que F és F-derivable en x. □

2.4.2 Teorema de Taylor

El teorema de Taylor per a una funció $f\colon I \mapsto \mathbb{R}$, $I \subseteq \mathbb{R}$ obert, en un punt $a \in I$, és

$$f(a+h) = T_n(h) + R_n(h) = f(a) + f'(a)h + \cdots + \frac{1}{n!} f^{n)}(a)h^n + R_n(h) \tag{2.10}$$

on $R_n(h) = \mathrm{o}\big(h^n\big)$ és l'error comés en aproximar $f(a+h)$ per $T_n(h)$. Aquest reste es pot expressar de diferents formes, entre elles la forma de Lagrange (2.11) per a algun c entre a i $a+h$. Habitualment aquest punt intermedi c és desconegut i és comú fer servir (2.12), on $M = \sup\big\{|f^{n+1)}(x)| : x \text{ entre } a \text{ i } a+h\big\}$.

$$f(a+h) - T_n(h) = R_n(h) = \frac{1}{(n+1)!} f^{n+1)}(c) h^{n+1} \tag{2.11}$$

$$\big|f(a+h) - T_n(h)\big| = \big|R_n(h)\big| \le \frac{1}{(n+1)!} M h^{n+1} \tag{2.12}$$

La versió (2.11) es pot generalitzar fàcilment a funcions escalars $f\colon C \mapsto \mathbb{R}$, $C \subseteq \mathbb{R}^n$ obert convex. En canvi no és certa per a funcions $f\colon C \mapsto \mathbb{C}^m$, $C \subseteq \mathbb{C}^n$ obert convex. Un exemple senzill és l'aplicació $f\colon \big\{|z| < 7\big\} \mapsto \mathbb{C}$, $f(z) = \mathrm{e}^{iz}$; no hi ha cap $c \in (0, 2\pi)$ tal que $f(2\pi) = f(0) + f'(0)(2\pi) + \frac{1}{2} f''(c)(2\pi)^2$.

La versió (2.12), en canvi, sí es manté en espais normats en general. Nosaltres només farem el cas $n = 1$, però el resultat és vàlid per a tot $n \ge 1$ (veure [1]).

Teorema 2.32 (Taylor)**.** *Siguin $F\colon C \mapsto Y$, $C \subseteq X$ obert convex, X i Y espais normats. Si F té G-derivada segona en C, aleshores per a qualssevol $a, b \in C$ es té*

$$\left\|F(b) - F(a) - F'(a)(b-a)\right\| \le \sup_{x \in (a,b)} \frac{1}{2} \left\|F''(x)\right\| \|b-a\|^2$$

Demostració. Fem $h = b - a$ i assumim $M = \sup\big\{\|F''(x)\| : x \in (a,b)\big\} < \infty$, altrament la desigualtat se satisfà trivialment. Definim $f\colon [0,1] \mapsto Y$, $f(t) = F(a+th) - F(a) - tF'(a)h$, i derivant-la dos cops (veure exemple 2.24) tenim

$$f'(t) = F'(a+th)h - F'(a)h \qquad\qquad f''(t) = F''(a+th)hh$$

Definim ara, per a $t \in (0,1)$ fixat, $g(s) = f'(st)$, $s \in [0,1]$. Es té, per a $0 < s < 1$

$$g'(s) = f''(st)t = tF''(a+sth)hh \qquad\qquad \left\|g'(s)\right\| \le tM\|h\|^2$$

Pel lema 2.28 amb $\beta(s) = stM\|h\|^2$, $s \in [0,1]$

$$\left\|f'(t)\right\| = \left\|g(1) - g(0)\right\| \le \beta(1) - \beta(0) = tM\|h\|^2$$

i això es té per a tot $t \in (0,1)$.

Prenem $\alpha(t) = \frac{t^2}{2} M \|h\|^2$, $t \in [0,1]$, i reapliquem el lema 2.28

$$\|f(1)\| = \|f(1) - f(0)\| \leq \alpha(1) - \alpha(0) = \frac{1}{2} M \|h\|^2$$

□

Una versió lleugerament més general del Teorema de Taylor és la següent.

Proposició 2.33. *Sigui $F\colon C \mapsto Y$, $C \subseteq X$ obert convex, X i Y espais normats. Si F és G-derivable en C i per a tot $x, y \in C$ es té*

$$\|F'(x) - F'(y)\| \leq \gamma \|x - y\|^p$$

per a algun $p \in [0, \infty)$ i $\gamma \in (0, \infty)$, aleshores per a qualssevol $a, b \in C$

$$\|F(b) - F(a) - F'(a)(b-a)\| \leq \frac{\gamma}{p+1} \|b - a\|^{p+1}$$

Demostració. Fem $h = b - a$ i definim $g\colon [0,1] \mapsto Y$, $g(t) = F(a + th) - tF'(a)h$. Derivant (veure exemple 2.24) tenim

$$g'(t) = F'(a + th)h - F'(a)h = \big(F'(a + th) - F'(a)\big)h$$
$$\|g'(t)\| \leq \|F'(a + th) - F'(a)\| \, \|h\| \leq \gamma \|th\|^p \, \|h\|$$

per a tot $t \in (0,1)$.

Prenent $\alpha(t) = \frac{\gamma}{p+1} t^{p+1} \|h\|^{p+1}$, $t \in [0,1]$ i aplicant el lema 2.28 tenim

$$\|F(a + h) - F(a) - F'(a)h\| = \|g(1) - g(0)\| \leq \|\alpha(1) - \alpha(0)\| \leq \frac{\gamma}{p+1} \|h\|^{p+1}$$

□

3 Teorema de Kantorovich

Entre el 1948 i el 1951 Leonid Vitaliyevich Kantorovich publicà dues demostracions diferents del que avui es coneix com Teorema de Kantorovich: un teorema de convergència semilocal per al mètode de Newton en espais de Banach (veure Introducció). El Teorema també es coneix com Teorema de Newton-Kantorovich, però en aquest treball prenem la primera denominació.

La primera demostració estava basada en relacions de recurrència i la segona en el principi de majoració. Donarem la primera prova i, quant a la segona, explicarem el principi de majoració, però en donarem una demostració diferent (Ortega [15]) basada en successions majorants. Prèviament introduïm, superficialment, les successions majorants i les idees i resultats de Rheinboldt [18] amb què complementem la prova d'Ortega, i que serveixen per a obtenir resultats de convergència semilocal similars al de Kantorovich per a variacions del mètode de Newton (1.2)-(1.5) i altres, que es poden trobar en [18] i [16, cap. 12].

La primera demostració (1948) és un breu article [6], que durant els mesos vinents complementà amb altres dos més detallats [7, 8] (tots tres traduïts a l'anglès en [11]), però que essencialment contenien la mateixa demostració. Aquest tipus de demostració s'anomena *per recurrència*, perquè consisteix a estudiar la relació de recurrència de (1.2) i obtenir-ne successions reals, també definides recurrentment, a partir de les quals deduir la convergència del mètode iteratiu. Al 1951 publicà una demostració diferent [9, 10], basada en el *principi de majoració* que, essencialment, consisteix a comparar el mètode iteratiu de Newton (1.2) en un espai de Banach, amb un mètode iteratiu en $\mathbb{R}$, a partir del qual deduïm les propietats de convergència del primer (més detalls en l'apartat 3.3.1).

En aquesta secció i en totes les vinents, tots els **espais normats seran de Banach**.

3.1 Condicions de Kantorovich

Kantorovich formulà inicialment el seu teorema per a operadors $F\colon X \mapsto Y$ sobre espais de Banach X i Y amb segona derivada de Fréchet [6, 7, 8] (traduïts a l'anglès en [11]). El Teorema, però, també és vàlid si substituïm X per $D \subseteq X$ obert; que és la formulació habitual avui dia i la que prenem. El teorema demana certes condicions sobre l'operador F i un punt x^0, que podem interpretar com una solució aproximada de l'equació $Fx = \vec{0}$. Aquestes condicions de Kantorovich són

K1) $\Gamma_0 = F'(x^0)^{-1} \in \mathscr{B}(Y,X)$ amb $\|\Gamma_0\| \le \beta$.

K2) $\|\Gamma_0 F x^0\| \le \eta$.

K3) Hi ha $R, M > 0$ tals que $B(x^0,R) \subseteq D$ i $\|F''(x)\| \le M$, per a tot $x \in B(x^0,R)$.

K4) $h = \beta M \eta \le \frac{1}{2}$.

El valor de R haurà de satisfer una condició extra que demana el Teorema, però una tria segura és, com veurem, qualsevol $R > 2\eta$ tal que $B(x^0,R) \subseteq D$. El valor de M que puguem trobar dependrà també del valor de R. Una altra formulació habitual de K3) és substituir $B(x^0,R)$ per un obert convex $D_0 \subseteq D$, que ens serà útil en §5. Això no altera en absolut les demostracions del Teorema, i per tant és igualment vàlid.

Les condicions K1) i K2) s'imposen sobre el punt x^0, mentre que la condició K3) sobre l'operador F, i obliga a que sigui $\mathscr{C}^1$ sobre $B(x^0,R)$ (veure definició 2.21). La condició K4) és la condició clau: lliga les altres i és una mesura de com de petit ha d'ésser $\|Fx^0\|$, és a dir, com de prop està x^0 de resoldre $Fx = \vec{0}$.

A partir d'aquestes condicions Kantorovich garantia l'existència d'una solució x^*, i donava dos entorns de x^0, un en què la solució existeix i un altre en què hi és única. És a dir, donava una regió d'existència i una altra d'unicitat.

Les condicions de Kantorovich es poden modificar de diverses maneres i amb diferents propòsits: per a obtenir condicions més fàcils de satisfer o més ajustades a un tipus de problemes concrets, regions d'existència i unicitat més àmplies, o ampliar la classe d'operadors a què aplicar el teorema, etc. La condició K3) es pot substituir per una condició de Lipschitz sobre F' en $B(x^0,R)$; que igual que K3), obliga a que F sigui $\mathscr{C}^1$ sobre $B(x^0,R)$.

K3') $\exists\, R, M > 0$ tals que $B(x^0,R) \subseteq D$ i $\|F'(x) - F'(y)\| \le M\,\|x-y\|$, per a tot $x, y \in B(x^0,R)$.

Igual que hem dit dalt amb la condició K3), la condició K3') també es pot formular substituint $B(x^0,R)$ per un obert convex $D_0 \subseteq D$, i això no altera en absolut les demostracions del Teorema.

Les condicions K1) i K3) (o K3')) es poden substituir per altres similars, però sobre una regió $D_0 \subseteq D$ obert convex. Un exemple en són les condicions de Mysovskikh, que permeten ampliar K4) i obtenir una demostració més senzilla, però també una mica més difícil d'aplicar (veure secció §4).

M1) $\forall x \in D_0$, $F'(x)^{-1} \in \mathscr{B}(Y,X)$ i $\|F'(x)^{-1}\| \leq \beta$.

M2) Hi ha $x^0 \in D_0$ tal que $\|\Gamma_0 F x^0\| \leq \eta$, $\Gamma_0 = F'(x^0)^{-1}$.

M3) $\exists M > 0$ tal que $\left\|F'(x) - F'(y)\right\| \leq M\|x-y\|$, per a tot $x, y \in D_0$.

M4) $h = \beta M \eta < 2$.

Les condicions de Kantorovich és poden modificar de moltes formes, donant lloc a una varietat de resultats similars al de Kantorovich. En la secció §4 n'exposem alguns.

3.2 Demostració per recurrència

La demostració per recurrència que donem és essencialment la que dona Kantorovich en [7] (traduït a l'anglès en [11]), però evitant petites parts en què fa servir integrals o el Teorema de Hahn-Banach. Una altra demostració per recurrència, ben exposada, comentada i fidel a l'original de Kantorovich, es troba en [1]. Una altra que fa ús d'integrals es troba en [3]. No obstant totes són gairebé iguals, éssent la distinció de fer servir integrals o no més aviat una qüestió de notació i comoditat.

Recordem que el mètode de Newton per a un operador F amb condició inicial x^0 és

$$x^{k+1} = x^k - \Gamma_k F x^k \qquad \left(\forall k \geq 0\right) \tag{3.1}$$

on $\Gamma_k = F'(x^k)^{-1} \in \mathscr{B}(Y,X)$.

Teorema 3.1 (Kantorovich, 1948)**.** *Sigui l'operador $F\colon D \mapsto Y$, $D \subseteq X$ obert, amb G-derivada segona sobre D. Suposem que es compleixen les condicions de Kantorovich K1)-K4).*

A més definim $r^* = \frac{1-\sqrt{1-2h}}{h}\eta$, $r^{**} = \frac{1+\sqrt{1-2h}}{h}\eta$; i suposem que $\overline{B(x^0, r^*)} \subseteq B(x^0, R)$.

Aleshores, la successió (3.1) resta a $\overline{B(x^0, r^*)}$, convergeix a una solució x^* de $Fx = \vec{0}$, única en $\overline{B(x^0, r^*)}$, i satisfà

$$\|x^* - x^k\| \leq \frac{1}{2^{k-1}}\big(2h\big)^{2^k-1}\eta \qquad \left(\forall k \geq 0\right) \tag{3.2}$$

si a més $h < \frac{1}{2}$, la solució x^* és única en $B(x^0, r^{**}) \cap B(x^0, R)$.

La demostració és conceptualment senzilla: cada nou iterat de (3.1) és a $\overline{B(x^0, r^*)}$ i compleix les condicions K1), K2) i K4) amb altres valors de β, η, h. Malgrat aquesta senzillesa, és una mica llarga si es vol fer amb claredat; per tant la dividirem en lemes. En tots els lemes s'assumeixen les hipòtesis del teorema 3.1 i denotem per $(x^k)_{k\geq 0}$ la successió del iterats de Newton (3.1).

Lema 3.2. *Sigui $\varphi\colon [0, 1/2] \mapsto \mathbb{R}$, $\varphi(t) = \frac{1}{t}(1 - \sqrt{1-2t})$. La funció és creixent i $\forall t \in [0, 1/2]$, $1 \leq \varphi(t) \leq 2$.*

Demostració. Només cal observar que, si $t \neq 0$, $\varphi(t) = \frac{1}{t}(1-\sqrt{1-2t}) = \frac{2}{1+\sqrt{1-2t}}$, que és contínua i creixent en $[0, 1/2]$. □

Lema 3.3. *Siguin $x, y \in B(x^0, R)$. Si $\Gamma = F'(x)^{-1} \in \mathscr{B}(Y,X)$ i $\|\Gamma\|\, M\, \|x-y\| \leq \alpha < 1$, aleshores $F'(y)$ és bijectiu amb invers continu: $F'(y)^{-1} \in \mathscr{B}(Y,X)$. A més,*

$$\left\|F'(y)^{-1}\right\| \leq \frac{\|\Gamma\|}{1-\alpha}$$

Demostració. Veiem que $\Gamma F'(y) \in \mathscr{B}(X,X)$ és bijectiu amb invers continu, això provarà que $F'(y) \in \mathscr{B}(X,Y)$ és bijectiu i amb invers continu:

$$[\Gamma F'(y)]^{-1} F'(x)^{-1} = \left(F'(x)[\Gamma F'(y)]\right)^{-1} = \left([\Gamma^{-1}\Gamma]F'(y)\right)^{-1} = F'(y)^{-1}$$

Per a veure que $\Gamma F'(y)$ és bijectiu amb invers continu fem servir la proposició 2.13

$$\begin{aligned}\left\|Id_X - \Gamma F'(y)\right\| &= \left\|\Gamma\left(F'(x) - F'(y)\right)\right\| \le \|\Gamma\| \left\|F'(x) - F'(y)\right\| \\ &\le \|\Gamma\| \sup_{z\in(y,x)} \left\|F''(z)\right\| \|x-y\| \le \alpha < 1\end{aligned}$$

on la introducció del suprem és pel Teorema de Lagrange 2.29 aplicat a F' sobre $B(x^0,R)$. De la proposició 2.13 tenim que $\Gamma F'(y)$ és bijectiu amb $\left\|[\Gamma F'(y)]^{-1}\right\| \le \left(1 - \|Id_X - \Gamma F'(y)\|\right)^{-1}$. Fent servir que $t \mapsto (1-t)^{-1}$, $t \in [0,1)$, és creixent, obtenim $\left\|\left[\Gamma F'(y)\right]^{-1}\right\| \le (1-\alpha)^{-1}$; per tant

$$\left\|F'(y)^{-1}\right\| = \left\|F'(y)^{-1}\Gamma^{-1}\Gamma\right\| \le \left\|F'(y)^{-1}\Gamma^{-1}\right\| \|\Gamma\| = \left\|\left[\Gamma F'(y)\right]^{-1}\right\| \|\Gamma\| \le \frac{\|\Gamma\|}{1-\alpha}$$

□

Lema 3.4 (Pas inicial d'inducció del teorema 3.1)**.** *Volem veure que existeix* $x^1 \in \overline{B(x^0,r^*)}$ *i satisfà també les condicions K1), K2) i K4).*

Demostració. Fem $\beta_0 = \beta$, $\eta_0 = \eta$, $h_0 = h$; i definim $G_0\colon D \mapsto X$, $G_0(x) = x - \Gamma_0 Fx$, que té sentit ja que $\Gamma_0 = F'(x^0)^{-1} \in \mathscr{B}(Y,X)$. Clarament es té $x^1 = G_0 x^0$.

x^1 és a $\overline{B(x^0,r^*)}$. Això és una conseqüència directa del lema 3.2

$$\|x^1 - x^0\| = \left\|G_0 x^0 - x^0\right\| = \left\|\Gamma_0 F x^0\right\| \le \eta_0 \le \frac{1-\sqrt{1-2h_0}}{h_0}\eta_0 = r^*$$

Condició K1). L'operador $\Gamma_1 = F'(x^1)^{-1} \in \mathscr{B}(Y,X)$ té sentit i $\|\Gamma_1\| \le \frac{\beta_0}{1-h_0} = \beta_1$. Això és una conseqüència directa del lema 3.3 amb $x = x^0$, $y = x^1$ i $\alpha = h_0$

$$\left\|\Gamma_1\right\| \le \frac{\left\|\Gamma_0\right\|}{1-h_0} \le \frac{\beta_0}{1-h_0} = \beta_1$$

Condició K2). $\left\|\Gamma_1 F x^1\right\| \le \frac{h_0\eta_0}{2(1-h_0)} = \eta_1$. Per definició de $x^1 = G_0 x^0$ es té $Fx^0 + F'(x^0)(x^1 - x^0) = \vec{0}$, per tant

$$\begin{aligned}\left\|\Gamma_1 F x^1\right\| &\le \beta_1 \left\|Fx^1 - Fx^0 - F'(x^0)(x^1-x^0)\right\| \le \frac{\beta_1}{2} \sup_{y\in(x^0,x^1)} \left\|F''(y)\right\| \|x^1 - x^0\|^2 \\ &\le \frac{\beta_1}{2} M\eta_0^2 = \frac{\beta_0}{2(1-h_0)} M\eta_0^2 = \frac{h_0\eta_0}{2(1-h_0)} = \eta_1\end{aligned}$$

on la introducció del suprem és pel Teorema de Taylor 2.32 aplicat a F sobre $B(x^0,R)$.

Condició K4). Prenem $h_1 = \beta_1 M \eta_1$ que verifica $h_1 \le \frac{1}{2}$

$$h_1 = \frac{\beta_0}{1-h_0} M \frac{h_0\eta_0}{2(1-h_0)} = \frac{h_0^2}{2(1-h_0)^2} \le 2h_0^2 \le \frac{1}{2}$$

□

En resum: $x^1 \in \overline{B(x^0,r^*)}$ i les condicions K1), K2) i K4) se satisfan amb x^1 substituint β_0, η_0, h_0 per β_1, η_1, h_1. Ara podem definir $G_1\colon D \mapsto X$, $G_1(x) = x - \Gamma_1 F(x)$, que té sentit ja que $\Gamma_1 \in \mathscr{B}(Y,X)$, i determinar el següent iterat $x^2 = G_1 x^1$.

Reiterant el procés, assumim que hem obtingut punts $(0 \le k \le n)$, $x^k \in \overline{B(x^0,r^*)}$ i nombres β_k, η_k, h_k, donats per les relacions de recurrència $(0 \le k < n)$

$$\beta_{k+1} = \frac{\beta_k}{1-h_k} \qquad \eta_{k+1} = \frac{h_k\eta_k}{2(1-h_k)} \qquad h_k = M\beta_k\eta_k \tag{3.3}$$

$$x^{k+1} = x^k - \Gamma_k F x^k \tag{3.4}$$

on $\Gamma_k = F'(x^k)^{-1} \in \mathscr{B}(Y,X)$. I a més satisfan les relacions següents per a $1 \le k \le n$

$$\|\Gamma_k\| \le \beta_k \tag{3.5}$$

$$h_k = \frac{h_{k-1}^2}{2(1-h_{k-1})^2} \le 2h_{k-1}^2 \le \frac{1}{2} \tag{3.6}$$

$$\eta_k \le h_{k-1}\eta_{k-1} \tag{3.7}$$

$$\|x^k - x^{k-1}\| = \|\Gamma_{k-1}Fx^{k-1}\| \le \eta_{k-1} \tag{3.8}$$

El pas inductiu consisteix a veure que tot això també és cert amb $n+1$: existeixen $x^{n+1} \in \overline{B(x^0,r^*)}$, β_{n+1}, η_{n+1}, h_{n+1}, compleixen (3.3) (3.4) amb $k=n$, i se satisfan (3.5) - (3.8) amb $k = n+1$.

Fem notar que les successions reals $(\beta_k)_{k\ge 0}$, $(\eta_k)_{k\ge 0}$, $(h_k)_{k\ge 0}$ queden determinades per β_0, η_0, h_0 i les relacions de recurrència de (3.3). Per tant, existeixen amb independència del mètode de Newton; és a dir, els iterats x^k. Té sentit, doncs, el següent lema, que es demostra fàcilment per inducció.

Lema 3.5. *Les successions $(\beta_k)_{k\ge 0}$, $(\eta_k)_{k\ge 0}$ i $(h_k)_{k\ge 0}$ satisfan (3.3) per a tot $k \ge 0$, i (3.6) (3.7) per a tot $k \ge 1$.*

Per a completar el pas inductiu de la demostració del teorema 3.1 només resta provar que existeix $x^{n+1} \in \overline{B(x^0,r^*)}$ i (3.5) (3.8) són certes amb $k = n+1$. Per a això i per a la unicitat de la solució, necessitarem el següent lema, que és un punt clau de la demostració del teorema.

Lema 3.6. *Siguin $\varphi, \psi\colon [0, 1/2] \mapsto \mathbb{R}$, $\varphi(t) = \frac{1}{t}(1-\sqrt{1-2t})$, $\psi(t) = \frac{1}{t}(1+\sqrt{1-2t})$. Se satisfan les identitats*

$$\eta_k\varphi(h_k) - \eta_{k+1}\varphi(h_{k+1}) = \eta_k \qquad (\forall k \ge 0) \tag{3.9}$$

$$\eta_k\psi(h_k) - \eta_{k+1}\psi(h_{k+1}) = \eta_k \qquad (\forall k \ge 0) \tag{3.10}$$

i a més $\sum_{k\ge 0}\eta_k = \eta_0\varphi(h_0) = r^*$.

Demostració. La demostració es basa en què (3.6) (3.7) són certes $\forall k \ge 1$ (lema 3.5).

De (3.6) es té la identitat (3.9); només farem aquesta: la identitat (3.10) es prova igual.

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

De (3.7) es té $\eta_k \le \frac{1}{2}\eta_{k-1} \le 2^{-k}\eta_0$, d'on inferim $\lim_{k\to\infty}\eta_k = 0$. La successió $(\varphi(h_k))_{k\ge 0}$ està afitada, pel lema 3.2; així doncs $\lim_{k\to\infty}\eta_k\varphi(h_k) = 0$, d'on concloem

$$\sum_{k\ge 0}\eta_k = \lim_{k\to\infty}\sum_{j=0}^{k}\big(\eta_j\varphi(h_j) - \eta_{j+1}\varphi(h_{j+1})\big) = \lim_{k\to\infty}\big(\eta_0\varphi(h_0) - \eta_{k+1}\varphi(h_{k+1})\big) = r^*$$

□

Finalment, completem el pas inductiu de la demostració.

Lema 3.7 (Completar el pas inductiu del teorema 3.1)**.** *Existeix $x^{n+1} \in \overline{B(x^0, r^*)}$ i (3.5) (3.8) es compleixen amb $k = n+1$.*

Demostració. Recordem que, per hipòtesis d'inducció, teníem $x^0, \ldots, x^n \in \overline{B(x^0, r^*)}$ donats per (3.4) $(0 \le k < n)$ i satisfent (3.5) (3.8) per a tot $1 \le k \le n$.

Condició K2). Atès que $x^n \in \overline{B(x^0, r^*)}$ podem definir $x^{n+1} = x^n - \Gamma_n F x^n$, $\Gamma_n = F'(x^n)^{-1} \in \mathscr{B}(Y, X)$. En particular, per a cada $0 \le k \le n$, es compleix

$$F(x^k) + F'(x^k)(x^{k+1} - x^k) = \vec{0} \tag{3.11}$$

Comprovem que $\|x^{n+1} - x^n\| \le \eta_n$ fent servir (3.5) i (3.11), el Teorema de Taylor 2.32 sobre $B(x^0, R)$, i (3.3)

$$\begin{aligned}
\|x^{n+1} - x^n\| &= \big\|\Gamma_n F x^n\big\| \le \beta_n \big\|F x^n - F x^{n-1} - F'(x^{n-1})(x^n - x^{n-1})\big\| \\
&\le \frac{\beta_n}{2}\sup_{y\in(x^{n-1},x^n)}\big\|F''(y)\big\|\,\|x^n - x^{n-1}\|^2 \le \frac{\beta_n M\eta_{n-1}^2}{2} \\
&= \frac{\beta_{n-1}M\eta_{n-1}^2}{2(1-h_{n-1})} = \frac{h_{n-1}\eta_{n-1}}{2(1-h_{n-1})} = \eta_n
\end{aligned}$$

cosa que prova (3.8) per a $n+1$, com volíem.

x^{n+1} és a $\overline{B(x^0, r^*)}$. Hem de provar $\|x^{n+1} - x^0\| \le r^*$; acabem de veure que (3.8) es compleix per a $1 \le k \le n+1$, per tant

$$\|x^{n+1} - x^0\| \le \sum_{j=0}^{n}\|x^{j+1} - x^j\| \le \sum_{j=0}^{n}\eta_j \tag{3.12}$$

i pel lema 3.6 la suma és telescòpica i afitada per $\sum_{k\ge 0}\eta_k = r^*$.

Condició K1). Ara que sabem que $x^{n+1} \in \overline{B(x^0, r^*)} \subseteq B(x^0, R)$, sabem que existeix $F'(x^{n+1})$ i volem veure que $\Gamma_{n+1} = F'(x^{n+1})^{-1} \in \mathscr{B}(Y, X)$. Novament ho fem via el lema 3.3 amb $x = x^n$, $y = x^{n+1}$ i $\alpha = h_n$, també recordant (3.3)

$$\big\|\Gamma_{n+1}\big\| \le \frac{\big\|\Gamma_n\big\|}{1-h_n} \le \frac{\beta_n}{1-h_n} = \beta_{n+1}$$

cosa que prova (3.5) per a $n+1$, com volíem.

Condició K4). Del lema 3.5 tenim, en particular, $h_{n+1} = M\beta_{n+1}\eta_{n+1} \le \frac{1}{2}$. □

En resum: amb el raonament d'inducció complert, tenim que $(x^k)_{k\geq 0} \subseteq \overline{B(x^0, r^*)}$, les relacions (3.5)-(3.8) són vàlides per a tot $k \geq 1$, i (3.11) és certa per a tot $k \geq 0$. En particular, $\lim_{k\to\infty} h_k = \lim_{k\to\infty} \eta_k = 0$.

Ens centrem en la convergència de la successió $(x^k)_{k\geq 0}$ a una solució $x^* \in \overline{B(x^0, r^*)}$ de $Fx = \vec{0}$, la seva unicitat i l'ordre de convergència.

Lema 3.8 (Convergència a una solució i ordre de convergència)**.** *La successió dels iterats de Newton $(x^k)_{k\geq 0}$ convergeix $\lim_{k\to\infty} x^k = x^* \in \overline{B(x^0, r^*)}$ i $Fx^* = \vec{0}$. A més, satisfà (3.2).*

Demostració. Veiem que la successió $(x^k)_{k\geq 0}$ és de Cauchy utilitzant la identitat (3.9) ($n \geq 0$, $m \geq 1$)

$$\begin{aligned}\|x^{n+m} - x^n\| &\leq \sum_{j=n}^{n+m-1} \|x^{j+1} - x^j\| \leq \sum_{j=n}^{n+m-1} \eta_j = \sum_{j=n}^{n+m-1} \big(\eta_j \varphi(h_j) - \eta_{j+1}\varphi(h_{j+1})\big) \\ &= \eta_n \varphi(h_n) - \eta_{n+m}\varphi(h_{n+m}) \leq \eta_n \varphi(h_n) \leq 2\eta_n\end{aligned}$$

atès que $\eta_k \xrightarrow{k\to\infty} 0$, $(x^k)_{k\geq 0}$ és de Cauchy i té límit $\lim_{k\to\infty} x^k = x^* \in \overline{B(x^0, r^*)}$. En concret, fent $m \to \infty$, tenim

$$\|x^* - x^k\| \leq 2\eta_k \qquad \big(\forall k \geq 0\big) \tag{3.13}$$

Per (3.7) i (3.6) tenim

$$\begin{aligned}\eta_k &\leq h_{k-1}\eta_{k-1} \leq h_{k-1}h_{k-2}\eta_{k-2} \leq \cdots \leq h_{k-1}\cdots h_0\eta_0 \\ 2h_k &\leq (2h_{k-1})^2 \leq (2h_{k-2})^{2^2} \leq \cdots \leq (2h_0)^{2^k} \leq 1\end{aligned}$$

d'on deduïm el següent, que multiplicat per 2 i junt amb (3.13), dona (3.2)

$$\eta_k \leq h_{k-1}\cdots h_0\eta_0 \leq 2^{-k}(2h_0)^{2^{k-1}+2^{k-2}+\cdots+1}\eta_0 = \frac{\eta_0}{2^k}(2h_0)^{2^k-1}$$

Per a provar que $Fx^* = \vec{0}$ recordem que $\forall k \geq 0$, $Fx^k + F'(x^k)(x^{k+1} - x^k) = \vec{0}$; per tant

$$\begin{aligned}\big\|Fx^k\big\| &= \big\|F'(x^k)(x^{k+1} - x^k)\big\| \leq \Big(\big\|F'(x^0)\big\| + \big\|F'(x^k) - F'(x^0)\big\|\Big)\|x^{k+1} - x^k\| \\ &\leq \Big(\big\|F'(x^0)\big\| + Mr^*\Big)\|x^{k+1} - x^k\| \xrightarrow{k\to\infty} 0\end{aligned}$$

per la condició K3) F és $\mathscr{C}^1$ en $B(x^0, R)$, i per tant contínua; de la continuïtat de F en x^* es té $\vec{0} = \lim_{k\to\infty} Fx^k = Fx^*$. □

Lema 3.9 (Unicitat de la solució)**.** *La solució x^* és única en $\overline{B(x^0, r^*)}$. Si a més, $h < \frac{1}{2}$, és única en $B(x^0, R) \cap B(x^0, r^{**})$.*

Demostració. Suposem que hi ha una altra solució $y^* \in B(x^0, R)$ de $Fx = \vec{0}$ tal que $\|y^* - x^0\| = \theta r^{**}$, per a algun $\theta \in (0, 1]$. Prenem $\psi(t) = \frac{1+\sqrt{1-2t}}{t}$, aleshores $\|y^* - x^0\| = \theta\eta_0\psi(h_0)$.

Definim la successió d'operadors $G_k \colon B(x^0, R) \mapsto X$, $G_k x = x - \Gamma_k Fx$, on $\Gamma_k = F'(x^k)^{-1} \in \mathscr{B}(Y, X)$. Observem que $\forall k \in \mathbb{N}$, $G_k y^* = y^*$, $G_k x^* = x^*$, $x^{k+1} = G_k x^k$, $G_k'(x^k) = 0$, $G_k''(x) =$

$-\Gamma_k F''(x)$

$$
\begin{aligned}
\|y^* - x^1\| &= \left\|G_0 y^* - G_0 x^0\right\| = \left\|G_0 y^* - G_0 x^0 - G_0'(x^0)(y^* - x^0)\right\| \\
&\le \frac{1}{2} \sup_{y \in (x^0, y^*)} \left\|G_0''(y)\right\| \|y^* - x^0\|^2 \le \frac{\beta_0 M}{2} (\theta \eta_0 \psi(h_0))^2 \\
&\le \frac{\theta^2}{2} h_0 \eta_0 (\psi(h_0))^2 = \theta^2 \frac{\eta_0}{h_0} \left(1 - h_0 + \sqrt{1 - 2h_0}\right) = \theta^2 (\eta_0 \psi(h_0) - \eta_0)
\end{aligned}
$$

per la identitat (3.10) concloem que

$$\|y^* - x^1\| \le \theta^2 (\eta_0 \psi(h_0) - \eta_0) = \theta^2 \eta_1 \psi(h_1)$$

Assumim que $\|y^* - x^k\| \le \theta^{2^k} \eta_k \psi(h_k)$, i veiem que també és cert per a $k+1$

$$
\begin{aligned}
\|y^* - x^{k+1}\| &= \|G_k y^* - G_k x^k\| = \|G_k y^* - G_k x^k - G_k'(x^k)(y^* - x^k)\| \\
&\le \frac{1}{2} \sup_{y \in (x^k, y^*)} \left\|G_k''(y)\right\| \|y^* - x^k\|^2 \le \frac{\beta_k M}{2} \left(\theta^{2^k} \eta_k \psi(h_k)\right)^2 \\
&= \frac{\theta^{2^{k+1}}}{2} h_k \eta_k (\psi(h_k))^2 = \theta^{2^{k+1}} \frac{\eta_k}{h_k} \left(1 - h_k + \sqrt{1 - 2h_k}\right) \\
&\le \theta^{2^{k+1}} (\eta_k \psi(h_k) - \eta_k) = \theta^{2^{k+1}} \eta_{k+1} \psi(h_{k+1})
\end{aligned}
$$

Per a acabar, notem que $\eta_k \psi(h_k) \le 2\frac{\eta_k}{h_k} = \frac{2}{\beta_k M}$. Si $h_0 < \frac{1}{2}$ es té $r^* < r^{**}$ i prenent $\theta < 1$, tenim $x^* = y^*$; ja que $\|y^* - x^k\| \le \theta^{2^k} \frac{2}{\beta_k M} \le \theta^{2^k} \frac{2}{\beta_0 M} \xrightarrow{k \to \infty} 0$. Així x^* és l'única solució en $B(x^0, r^{**}) \cap B(x^0, R)$.

Si $h_0 = \frac{1}{2}$, $r^* = r^{**} = 2\eta_0$, $\theta \le 1$ i $\eta_k \psi(h_k) = 2\eta_k$; aquí també $x^* = y^*$, ja que $\|y^* - x^k\| \le 2\eta_k \xrightarrow{k \to \infty} 0$. Així x^* és l'única solució en $\overline{B(x^0, r^*)}$. $\square$

Observació 3.10 (Tria del radi R). A l'hora d'aplicar el teorema, hem de calcular η, β, M i h; els dos primers són independents del valor de R, però no els dos últims. Una elecció habitual és $R > 2\eta$, ja que $0 \le h \le 1/2 \Rightarrow r^* = \eta \frac{1 - \sqrt{1 - 2h}}{h} \le 2\eta$ pel lema 3.2.

Observació 3.11. Un punt clau de la demostració, sinó el més important, és garantir $\sum_{k \ge 0} \eta_k = l < \infty$. Assumint $(x^k)_{k \ge 0} \subseteq B(x^0, R) \subseteq D$, per a algun $R > 0$, atesa la relació $\|x^{k+1} - x^0\| \le \sum_{j=0}^{k} \eta_j \le l$, la convergència d'aquesta sèrie permet assegurar que $(x^k)_{k \ge 0} \subseteq \overline{B(x^0, l)}$, que és de Cauchy i obtenir una fita de l'error $\|x^* - x^k\|$. A més, podem fixar $r^* = l$. És a dir, la convergència de $(x^k)_{k \ge 0}$, la localització de $x^* \in \overline{B(x^0, r^*)}$ i les fites d'error depenen íntimament de la sèrie $\sum_{k \ge 0} \eta_k$.

Observació 3.12 (Simplicitat de la solució). Una conseqüència del Teorema de Kantorovich és que la solució x^* és necessàriament simple quan $h < 1/2$: $F'(x^*)$ és bijectiu i $F'(x^*)^{-1} \in \mathscr{B}(Y, X)$. Això és degut al lema 3.3, que garanteix que $F'(x^*)^{-1} \in \mathscr{B}(Y, X)$ quan $x^* \in B(x^0, 1/\beta M)$, cosa que es dona si $h < 1/2$. Quan $h = 1/2$ es té $r^* = \frac{1}{\beta M}$, aleshores la solució x^* pot no ser simple (veure exemple 3.41).

3.2.1 Mètode de Newton simplificat per recurrència

Repassant amb cura la demostració del Teorema de Kantorovich 3.1, veiem que podem adaptar-la al mètode de Newton simplificat

$$x^{k+1} = x^k - \Gamma_0 F x^k \qquad (\forall k \ge 0) \tag{3.14}$$

on $\Gamma_0 = F'(x^0)^{-1} \in \mathscr{B}(Y,X)$. La demostració que donem és gairebé idèntica a la que dona Kantorovich en [8]: només hem introduit algun canvi menor per afavorir-ne la claredat. Assumint $(x^k)_{k\geq 0} \subseteq \overline{B(x^0,r^*)}$, les relacions de recurrència (3.3) esdevenen, per a aquest mètode, $\big(\forall k \geq 0\big)$

$$\beta_k = \beta_0 \qquad h_k = \beta_0 M \eta_k \qquad \eta_{k+1} = \beta_0 M \eta_k r^* \tag{3.15}$$

amb $\|x^{k+1} - x^k\| = \|\Gamma_0 F x^k\| \leq \eta_k$, $\forall k \geq 0$.

Teorema 3.13. *Mantenim les hipòtesis del teorema 3.1, però ara $h < \frac{1}{2}$ i $r^{**} = \frac{\eta}{h} = \frac{1}{\beta M}$.*

Aleshores la successió (3.14) resta a $\overline{B(x^0,r^*)}$, convergeix a una solució x^* de $Fx = \vec{0}$, única en $B(x^0,r^{**}) \cap B(x^0,R)$, i satisfà

$$\|x^* - x^k\| \leq \frac{\alpha^k}{1-\alpha}\eta \qquad \big(\forall k \geq 0\big) \tag{3.16}$$

on $\alpha = 1 - \sqrt{1-2h} = \beta M r^*$.

Es tracta d'adaptar la prova del teorema 3.1 per al mètode de Newton simplificat (3.14); també dividim aquesta demostració en lemes. Saltem directament al pas d'inducció.

Lema 3.14 (Pas inductiu teorema 3.13)**.** *Fem $\eta_0 = \eta$, $\beta_0 = \beta$ i $h_0 = h$ i considerem les successions $(\beta_k)_{k\geq 0}$, $(h_k)_{k\geq 0}$, $(\eta_k)_{k\geq 0}$ definides per (3.15), en particular*

$$\eta_{k+1} = \beta_0 M \eta_k r^* = \eta_k(1 - \sqrt{1-2h_0}) \qquad \big(\forall k \geq 0\big) \tag{3.17}$$

Suposem que, per a $1 \leq k \leq n$, $x^k \in \overline{B(x^0,r^*)}$ i $\|x^k - x^{k-1}\| \leq \eta_{k-1}$; aleshores existeix $x^{n+1} \in \overline{B(x^0,r^*)}$ i $\|x^{n+1} - x^n\| \leq \eta_n$.

Per tant, $(x^k)_{k\geq 0} \subseteq \overline{B(x^0,r^*)}$ està ben definida i $\forall k \geq 0$, $\|x^{k+1} - x^k\| \leq \eta_k$.

Demostració. $\underline{x^{n+1} \in \overline{B(x^0,r^*)}}$. Hem de provar que x^{n+1} existeix i $\|x^{n+1} - x^0\| \leq r^*$.

Definim l'operador $G\colon B(x^0,R) \mapsto X$, $Gx = x - \Gamma_0 F x$, que té sentit ja que $\Gamma_0 \in \mathscr{B}(Y,X)$. Així doncs, $x^{n+1} = Gx^n$ existeix; observem que $G'(x^0) = 0$ i $G''(x) = -\Gamma_0 F''(x)$

$$\begin{aligned}
\|x^{n+1} - x^0\| &\leq \|x^{n+1} - x^1\| + \|x^1 - x^0\| \\
&= \|Gx^n - Gx^0 - G'(x^0)(x^n - x^0)\| + \eta_0 \\
&\leq \frac{1}{2} \sup_{y\in(x^0,x^n)} \big\|G''(y)\big\| \, \|x^n - x^0\|^2 + \eta_0 \\
&\leq \frac{\beta_0 M}{2}(r^*)^2 + \eta_0 = \frac{\beta_0 M}{2}\big(\frac{\eta_0}{h_0}(1 - \sqrt{1-2h_0})\big)^2 + \eta_0 \\
&= \frac{\eta_0}{h_0}(1 - h_0 - \sqrt{1-2h_0}) + \eta_0 = r^*
\end{aligned}$$

com volíem. La introducció del suprem és pel Teorema de Taylor 2.32 aplicat a G sobre $B(x^0,R)$.

$\underline{\|x^{n+1} - x^n\| \leq \eta_n}$. Sabent que $x^{n+1} \in \overline{B(x^0,r^*)} \subseteq B(x^0,R)$, podem aplicar la proposició 2.30 a l'operador G i el teorema de Lagrange 2.29 a G'

$$\begin{aligned}
\|x^{n+1} - x^n\| &= \big\|Gx^n - Gx^{n-1} - G'(x^0)(x^n - x^{n-1})\big\| \\
&\leq \sup_{y\in(x^{n-1},x^n)} \big\|G'(y) - G'(x^0)\big\| \, \|x^n - x^{n-1}\| \\
&\leq \sup_{y\in(x^{n-1},x^n)} \sup_{z\in(x^0,y)} \big\|G''(z)\big\| \, \|y - x^0\| \, \|x^n - x^{n-1}\| \leq \beta_0 M r^* \eta_{n-1} = \eta_n
\end{aligned}$$

això completa la inducció i concloem $(x^k)_{k\geq 0} \subseteq \overline{B(x^0,r^*)}$ amb $\|x^{k+1} - x^k\| \leq \eta_k$, $\forall k \geq 0$. □

Lema 3.15 (Convergència a una solució i ordre de convergència)**.** *La successió dels iterats de Newton simplificat* $(x^k)_{k\geq 0}$ *convergeix* $\lim_{k\to\infty} x^k = x^* \in \overline{B(x^0, r^*)}$ *i* $Fx^* = \vec{0}$*. A més, satisfà (3.16).*

Demostració. Veiem que la successió és de Cauchy ($n \geq 0$, $m \geq 1$)

$$\|x^{n+m} - x^n\| \leq \sum_{j=n}^{n+m-1} \|x^{j+1} - x^j\| \leq \sum_{j=n}^{n+m-1} \eta_j$$

Per (3.17) tenim que $\forall k \geq 0, \eta_k = \alpha^k \eta_0$, per tant

$$\|x^{n+m} - x^n\| \leq \sum_{j=n}^{n+m-1} \eta_j = \sum_{j=n}^{n+m-1} \eta_0 \alpha^j \leq \eta_0 \alpha^n \sum_{j\geq 0} \alpha^j = \eta_0 \frac{\alpha^n}{1-\alpha}$$

atès que $\alpha^k \xrightarrow{k\to\infty} 0$, ja que $h_0 < \frac{1}{2}$, $(x^k)_{k\geq 0}$ és de Cauchy i té límit $\lim_{k\to\infty} x^k = x^* \in \overline{B(x^0, r^*)}$. A més, fent $m \to \infty$ tenim (3.16).

Per a provar que $Fx^* = \vec{0}$ recordem que $\forall k \geq 0$, $Fx^k + F'(x^0)(x^{k+1} - x^k) = \vec{0}$; per tant

$$\left\|Fx^k\right\| = \left\|F'(x^0)(x^{k+1} - x^k)\right\| \leq \left\|F'(x^0)\right\| \|x^{k+1} - x^k\| \xrightarrow{k\to\infty} 0$$

per la condició K3) F és $\mathscr{C}^1$ en $B(x^0, R)$, i per tant contínua; de la continuïtat de F en x^* es té $\vec{0} = \lim_{k\to\infty} Fx^k = Fx^*$. □

Lema 3.16 (Unicitat de la solució)**.** *La solució* x^* *és única en* $B(x^0, r^{**}) \cap B(x^0, R)$.

Demostració. Sigui $y^* \in B(x^0, R)$ una altra solució de $Fx = \vec{0}$ tal que $\|y^* - x^0\| < r^{**}$. Notem que $r^{**} = \frac{\eta_0}{h_0} > 2\eta_0 > \frac{\eta_0}{h_0}(1 - \sqrt{1-2h_0}) = r^*$, per tant $\sup\left\{\|y - x^0\| \,:\, y \in (x^*, y^*)\right\} < r^{**}$

$$\begin{aligned}\|y^* - x^*\| &= \left\|Gy^* - Gx^* - G'(x^0)(y^* - x^*)\right\| \leq \sup_{y\in(x^*,y^*)} \left\|G'(y) - G'(x^0)\right\| \|y^* - x^*\| \\ &\leq \sup_{y\in(x^*,y^*)} \sup_{z\in(x^0,y)} \left\|G''(z)\right\| \|y - x^0\| \|y^* - x^*\| \\ &< \beta_0 M r^{**} \|y^* - x^*\| = \|y^* - x^*\|\end{aligned}$$

arribant així a la contradicció que $\|y^* - x^*\| < \|y^* - x^*\|$; per tant x^* és l'única solució de $Fx = \vec{0}$ en $B(x^0, R) \cap B(x^0, r^{**})$. □

Observació 3.17 (Tria del radi R). Per a aplicar el teorema 3.13, hem de calcular η, β, M i h; els dos primers són independents del valor de R, però no els dos últims. Una elecció habitual és $R \geq 2\eta$, ja que, pel lema 3.2, $h < \frac{1}{2} \Rightarrow r^* = \frac{\eta}{h}(1 - \sqrt{1-2h}) < 2\eta$.

Observació 3.18 (Radi d'unicitat r^{**} i equivalència Newton i Newton simplificat). Les hipòtesis i les tesis dels teoremes 3.13 i 3.1 són molt semblants; obtenim així una certa equivalència entre el mètode de Newton (3.1) i el mètode de Newton simplificat (3.14). En concret, si es pot aplicar el teorema 3.13, també es pot aplicar el teorema 3.1; obtenim així un radi d'existència major per al mètode de Newton simplificat $r^{**} = \frac{\eta}{h}\left(1 + \sqrt{1-2h}\right)$.

Observació 3.19 (Reiniciar el mètode). En el pas inductiu de la demostració del teorema 3.13 hem vist, de fet, que $(x^k)_{k\geq 1} \subseteq \overline{B(x^1, \frac{\alpha}{2} r^*)} \subset \overline{B(x^0, r^*)}$. Per tant, pot resultar convenient, un cop trobat x^0, reaplicar el teorema amb $\overline{x}^0 = x^1$ i els valors $\overline{\beta}_0 = \frac{\beta_0}{1-h_0}$, $\overline{\eta}_0 = \frac{h_0\eta_0}{2(1-h_0)}$, $\overline{h}_0 = M\overline{\beta}_0\overline{\eta}_0$ que donen un menor radi d'existència i fita d'error dels iterats $\overline{r}^* = \frac{\overline{\eta}_0}{\overline{h}_0}\left(1 - \sqrt{1-2\overline{h}_0}\right)$ i $\overline{\alpha} = 1 - \sqrt{1-2\overline{h}_0}$.

Observació 3.20 (Mètode quasi-Newton). La fita (3.16) només garanteix ordre de convergència lineal. L'observació anterior suggereix una primera millora, un primer mètode quasi-Newton: reaplicar el teorema periòdicament, per exemple, cada 3 iterats, o cada 5 iterats. Això permet actualitzar el valor d'α cada cert temps i obtenir així una millor fita (3.16).

3.3 Demostració per successions majorants

Donem una demostració del Teorema de Kantorovich 3.36 basada en successions majorants: prenem com a referència la que donà Ortega [15], que conjuntament amb Rheinboldt milloraren en [16, sec. 12.6] amb les seves idees d'aquest últim [18]; també incorporem l'aportació de Tapia [19] sobre la fita d'error. Prèviament, fem una breu introducció a les successions majorants i el seu ús per a provar la convergència de successions en espais de Banach.

Aquesta introducció prèvia és un resum concís de les seccions 12.1, 12.3, 12.4 i 12.5 de [16], prenent com a punt de partida estendre el Teorema del Punt Fix de Banach 3.21. L'objectiu és mostrar la utilitat de les idees de Rheinboldt per a provar la convergència del mètode de Newton (1.2) i les seves variants (1.3)-(1.5); en síntesi, Rheinboldt introdueix un problema auxiliar d'equacions en diferències a partir del qual pot deduir propietats de la convergència del mètode iteratiu.

3.3.1 Principi de majoració i successions majorants

Ambdós conceptes estan relacionats, però són subtilment diferents; la diferència es fa palesa en el desenvolupament teòric que permeten fer un i altre per a arribar al Teorema de Kantorovich 3.36. Introduim el principi de majoració seguint [12, cap. XVIII], que també reprodueixen meticulosament Ezquerro i Hernández [3, cap. 1].

L. V. Kantorovich introduí el *principi de majoració* per a donar una nova prova [9, 10] del seu teorema. El principi, en síntesi, relaciona un mètode iteratiu en un espai de Banach amb un mètode iteratiu en $\mathbb{R}$ a partir del qual inferir les propietats de convergència del primer. Amb més precisió, Kantorovich considera l'equació

$$x = S(x) \tag{3.18}$$

on $S\colon B(x^0,R) \mapsto X$ és un operador definit en una bola $B(x^0,R) \subseteq X$, X espai de Banach. També considera l'equació real

$$t = \phi(t) \tag{3.19}$$

on $\phi\colon [t_0,t'] \mapsto \mathbb{R}$ és una funció real definida en l'interval $[t_0,t'] \subseteq [t_0,t_0+R)$ per a un $t_0 \in \mathbb{R}$. Kantorovich diu que l'equació (3.19) *majora* l'equació (3.18) —o que la funció ϕ majora l'operador S— si per a tot $x \in B(x^0,R)$ i tot $t \in [t_0,t']$ es compleixen

$$\left\| Sx^0 - x^0 \right\| \leq \phi(t_0) - t_0 \tag{3.20}$$

$$\left\| S'(x) \right\| \leq \phi'(t) \quad \text{quan} \quad \|x - x^0\| \leq t - t_0 \tag{3.21}$$

Amb això es pot veure que la convergència de la successió $t_{k+1} = \phi(t_k)$, $k \geq 0$, implica la convergència de la successió $x^{k+1} = Sx^k$, $k \geq 0$, en X. Concretament guarden la següent relació

$$\|x^{k+1} - x^k\| \leq t_{k+1} - t_k \qquad \left(\forall\, k \geq 0\right) \tag{3.22}$$

a partir d'aquí Kantorovich prova una sèrie de resultats que permeten, a partir d'imposar condicions al mètode iteratiu $t_{k+1} = \phi(t_k)$, $k \geq 0$, inferir l'existència i unicitat local d'una solució $x^* \in B(x^0,R)$ de (3.18), donar regions d'existència i unicitat de x^*, i fites per a l'error $\|x^* - x^k\|$. Aquest desenvolupament teòric es pot trobar en les referències donades dalt.

El concepte de *successió majorant* radica en ignorar les condicions (3.20) (3.21) (i per tant la funció $\phi(t)$), però quedar-se amb (3.22), que encara permet inferir les propietats de convergència de $x^{k+1} = Sx^k$, x^0, a partir de la convergència de $(t_k)_{k\geq 0}$. Sembla que aquest concepte fou definit explícitament per primera vegada per Ortega [15] i Rheinboldt [18], segons diuen ells mateixos [16, p. 414].

Fem notar que ambdues vies tenen problemes similars: trobar una $\phi(t)$ que compleixi (3.20) (3.21), o trobar una successió $(t_k)_{k\geq 0}$ que verifiqui (3.22). Dediquem els següents apartats al problema de trobar una tal successió $(t_k)_{k\geq 0}$.

3.3.2 Contraccions i contraccions iterades

El concepte de contracció iterada busca ampliar el Teorema del Punt Fix de Banach a una classe més gran d'aplicacions; recordem primer aquest teorema.

En un espai de Banach $\left(X, \|\cdot\|\right)$, una *contracció o operador contractiu* sobre $D \subseteq X$, és un operador $G\colon D \mapsto X$ per a què $\exists\, \alpha \in (0,1) : \forall x, y \in D, \left\|Gx - Gy\right\| \leq \alpha \|x - y\|$.

Lema 3.21 (Tª Punt Fix de Banach)**.** *Sigui $G\colon D \mapsto X$, $D \subseteq X$, una contracció. Si $G(D) \subseteq D$ i D és tancat, aleshores per a tot $x^0 \in D$ la successió $x^{k+1} = Gx^k$, $k \geq 0$, és convergent a un punt fix (únic en D) de G, $p \in D$; a més es té*

$$\|x^k - p\| \leq \frac{\alpha^k}{1-\alpha} \|x^1 - x^0\| \qquad \left(\forall k \geq 0\right)$$

La demostració és redueix a tres punts: primer, l'aplicació G és Lipschitz; segon, les successions $(x^k)_{k\geq 0}$, $x^{k+1} = G(x^k)$ són contractives (veure apartat 2.1.2), i per tant de Cauchy; i tercer, dels dos punts anteriors i que D és tancat en X, es té que totes aquestes successions convergeixen a un punt fix (únic en D) de G. La demostració es pot consultar en qualsevol llibre d'anàlisi funcional, per exemple [12, sec. XVI.1] o [13, 17]; també es pot trobar en [1] una versió semilocal.

La dificultat a l'hora d'aplicar aquest resultat radica a trobar un conjunt $D \subseteq X$ tancat tal que $G(D) \subseteq D$ i $G_{\restriction D}$ és una contracció. Això és el que volem ampliar, però mantenint la propietat essencial que G genera successions contractives, almenys per a alguns punts del domini.

Definició 3.22. Un operador $G\colon D \mapsto X$, $D \subseteq X$, és una *contracció iterada* sobre D si hi ha $0 < \alpha < 1$ tal que

$$\left\|GGx - Gx\right\| \leq \alpha \left\|Gx - x\right\|$$

sempre que $x, Gx \in D$.

La diferència subtil amb les contraccions és que una contracció iterada G només contrau les distàncies entre un punt x i la seva imatge Gx, si també és al domini.

Tota contracció és un contracció iterada, però no al contrari. Igual que les contraccions, una contracció iterada genera successions contractives, i per tant de Cauchy, si podem garantir que la successió resta a D. A diferència de les contraccions, però, una contracció iterada pot no ésser contínua, no tenir punts fixos, o tenir-ne més d'un.

Establim una 'còpia' feble del lema 3.21 per a les contraccions iterades, que es demostra de forma gairebé idèntica. Bàsicament afeblim les condicions sobre l'operador G a canvi d'introduir una altra sobre un punt x^0.

Lema 3.23. *Sigui $G\colon D \mapsto X$, $D \subseteq X$ tancat, una contracció iterada. Si per a algun $x^0 \in D$ la successió $(x^k)_{k\geq 0}$, $x^{k+1} = G(x^k)$, resta a D, aleshores $(x^k)_{k\geq 0}$ és contractiva amb $\lim_{k\to\infty} x^k = p \in$*

D, es compleix

$$\|p - x^k\| \le \frac{\alpha^k}{1-\alpha}\|x^1 - x^0\| \qquad \big(\forall k \ge 0\big)$$

i, si G és continu en p, aleshores $Gp = p$.

A l'hora d'aplicar aquest lema, tenim dos problemes similars al que teníem amb el Punt Fix de Banach: trobar $D \subseteq X$ tancat tal que $G_{\restriction D}$ és una contracció iterada, i trobar $x^0 \in D$ tal que $(x^k)_{k\ge 0} \subseteq D$.

3.3.3 Successions majorants

Continuem ampliant el Teorema del Punt Fix de Banach, ara, però, volem introduir fites no lineals de $\|Gy - Gx\|$ i $\|GGx - Gx\|$. La idea és poder deduir la convergència d'una successió $x^{k+1} = Gx^k$, $k \ge 0$, en un espai de Banach, a partir de la convergència d'una successió real.

Definició 3.24. Sigui $(x^k)_{k\ge 0} \subseteq X$. Una successió creixent $(t_k)_{k\ge 0} \subseteq [0,\infty)$ que compleix

$$\|x^{k+1} - x^k\| \le t_{k+1} - t_k \qquad \big(\forall k \ge 0\big)$$

diem que és una *successió majorant* de, o que majora la successió, $(x^k)_{k\ge 0}$.

Efectivament la convergència d'una successió majorant implica la convergència de la successió majorada.

Lema 3.25. *Sigui $(x^k)_{k\ge 0} \subseteq X$ una successió majorada per $(t_k)_{k\ge 0}$. Si $\lim_{k\to\infty} t_k = t^* < \infty$, aleshores $\lim_{k\to\infty} x^k = x^* \in X$ i es té*

$$\|x^* - x^k\| \le t^* - t_k \tag{3.23}$$

Demostració. La convergència de la successió majorant $(t_k)_{k\ge 0}$ ens diu que, fixat $\epsilon > 0$, hi ha $H \in \mathbb{N}$ tal que $k \ge H \Rightarrow 0 \le t^* - t_k \le \epsilon$.

Això obliga que $(x^k)_{k\ge 0}$ sigui de Cauchy, i per tant convergent per ésser X espai de Banach. Prenent $n \ge H$, $m \ge 1$

$$\|x^{n+m} - x^n\| \le \sum_{j=n}^{m+n-1} \|x^{j+1} - x^j\| \le \sum_{j=n}^{m+n-1} t_{j+1} - t_j = t_{m+n} - t_n \le t^* - t_n \le \epsilon$$

De la continuïtat de la norma (observació 2.3), tenim (3.23) prenent el límit $m \to \infty$. □

Proposició 3.26. *Siguin un operador $G\colon D \mapsto X$, $D \subseteq X$ tancat, i una funció creixent $\varphi\colon [0,\infty) \mapsto [0,\infty)$. Suposem*

1. *$\|GGx - Gx\| \le \varphi\big(\|Gx - x\|\big)$, sempre que $x, Gx \in D$.*
2. *Hi ha $x^0 \in D$ tal que $(x^k)_{k\ge 0} \subseteq D$, $x^{k+1} = Gx^k$.*
3. *La successió $t_{k+1} - t_k = \varphi(t_k - t_{k-1})$, $t_0 = 0$, $t_1 \ge \|Gx^0 - x^0\|$ convergeix a $t^* < \infty$.*

Aleshores $\lim_{k\to\infty} x^k = x^* \in D$, es compleix (3.23) i, si G és continu en x^*, $Gx^* = x^*$.

Demostració. Bàsicament hem de comprovar que podem aplicar el lema precedent: $(t_k)_{k\geq 0}$ majora $(x^k)_{k\geq 0}$.

Procedim per inducció; per hipòtesi $\|x^1 - x^0\| \leq t_1 - t_0$, suposem que $\|x^j - x^{j-1}\| \leq t_j - t_{j-1}$, $(j \geq 1)$, volem veure que $\|x^{j+1} - x^j\| \leq t_{j+1} - t_j$

$$\|x^{j+1} - x^j\| = \left\|G(Gx^{j-1}) - Gx^{j-1}\right\| \leq \varphi\big(\|x^j - x^{j-1}\|\big) \leq \varphi(t_j - t_{j-1}) = t_{j+1} - t_j$$

on hem aplicat la definició de la successió $(x^k)_{k\geq 0}$, el punt 1, la monotonia de φ, i el punt 3.

Òbviament, si G és continu en x^*, aleshores $x^* = \lim_{k\to\infty} x^{k+1} = \lim_{k\to\infty} Gx^k = Gx^*$. □

El lema 3.23 és un cas particular d'aquest resultat, amb $\varphi(t) = \alpha t$. Si substituïm el punt 1 per $\|Gx - Gy\| \leq \varphi\big(\|x - y\|\big)$ i el punt 2 per $G(D) \subseteq D$, obtenim el Teorema del Punt Fix de Banach 3.21 prenent $\varphi(t) = \alpha t$.

Aquest resultat, malgrat més general que el lema 3.23, continua tenint els mateixos problemes —trobar x^0 i D tancat tals que $\|GGx - Gx\| \leq \varphi\big(\|Gx - x\|\big)$ i $(x^k)_{k\geq 0} \subseteq D$— i n'afegeix un de nou: resoldre l'equació en diferències (3.24), tot garantint-ne la convergència de la solució

$$\begin{cases} t_k = f(t_{k-1}, t_{k-2}) = t_{k-1} + \varphi(t_{k-1} - t_{k-2}) \\ t_1 \geq \|x^1 - x^0\| \\ t_0 = 0 \end{cases} \tag{3.24}$$

Això no és un problema fàcil en general, i no és menor: conèixer la solució $(t_k)_{k\geq 0}$ i el seu límit t^* ens permeten localitzar $(x^k)_{k\geq 0}$ i x^* —garantir que són a $\overline{B(x^0, t^*)}$— i tenir bones fites per a l'error $\|x^* - x^k\| \leq t^* - t_k$. Una opció és conformar-se amb una successió $(t'_k)_{k\geq 0}$ que afiti superiorment la solució i de la qual coneixem el límit t'^* i sabem calcular bones fites per a $t'^* - t'_k$; així $t^* \leq t'^*$ i $t^* - t_k \leq t'^* - t'_k$.

Observació 3.27. Fer servir la proposició 3.26 directament resulta complicat, ja que el punt 2 pot ésser difícil d'assegurar. La seva utilitat és més aviat introduir el problema auxiliar de l'equació en diferències (3.24) per, precisament, garantir que es compleix el punt 2 via el lema 3.25. La proposició, doncs, més que un resultat aplicable és un algorisme. Els passos a seguir són: primer, trobar una $\varphi(t)$ que verifiqui el punt 1; segon, plantejar l'equació en diferències (3.24) del punt 3 per a algun x^0, i intentar veure que $\overline{B(x^0, t^*)} \subseteq D$. Aleshores, pel lema 3.25, $(x^k)_{k\geq 0} \subseteq \overline{B(x^0, t^*)}$, i tenim el punt 2: ja podem 'aplicar' la proposició.

El teorema de Mysovskikh ens servirà d'exemple d'aplicació per al mètode de Newton (1.2)

Teorema 3.28 (Mysovskikh, 1950)**.** *Sigui l'operador G-derivable $F\colon D_0 \mapsto Y$, $D_0 \subseteq X$ obert convex. Suposem que es compleixen les condicions de Mysovskikh M1)-M4) (veure apartat 3.1).*

Definim a més $\alpha = \frac{1}{2}h < 1$, i assumim $\overline{B(x^0, r^*)} \subseteq D_0$ amb $r^* = \eta \sum_{j\geq 0} \alpha^{2^j - 1}$. Aleshores els iterats de Newton (1.2) estan ben definits, resten a $\overline{B(x^0, r^*)}$, convergeixen a una solució x^* de $Fx = \vec{0}$ i satisfan

$$\|x^* - x^k\| \leq \frac{\alpha^{2^k - 1}}{1 - \alpha^{2^k}}\eta \tag{3.25}$$

Demostració. Definim $Gx = x - F'(x)^{-1}Fx$, $x \in D_0$, que té sentit per la condició M1), i suposem $x, Gx \in D_0$

$$\left\|GGx - Gx\right\| = \left\|F'(Gx)^{-1}F(Gx)\right\| \leq \beta\left\|F(Gx) - Fx - F'(x)(Gx - x)\right\| \leq \frac{\beta M}{2}\left\|Gx - x\right\|^2$$

on hem aplicat la proposició 2.33 a F sobre D_0. Per tant $\|GGx - Gx\| \leq \varphi\big(\|Gx - x\|\big)$, amb $\varphi(t) = \frac{\beta M}{2} t^2$. L'equació en diferències (3.24) esdevé en aquest cas

$$t_{k+1} - t_k = \frac{\beta M}{2}(t_k - t_{k-1})^2, \qquad t_1 = \eta, \qquad t_0 = 0 \qquad \big(\forall k \geq 1\big)$$

clarament $t_2 - t_1 = \frac{1}{2}\beta M \eta^2 = \alpha\eta$; per inducció es veu que $t_{k+1} - t_k \leq \eta\alpha^{2^k - 1}$, $\forall k \geq 0$. Així

$$t^* = \lim_{k\to\infty} t_k = \lim_{k\to\infty} \sum_{j=0}^{k-1} t_{j+1} - t_j \leq \eta \sum_{k \geq 0} \alpha^{2^k - 1} = r^*$$

Pel lema 3.25, $(x^k)_{k\geq 0}$ resta a $\overline{B(x^0, r^*)}$, convergeix a $x^* \in \overline{B(x^0, r^*)}$ i es té (3.25) ja que $(x = \alpha^{2^k})$

$$t^* - t_k \leq \eta \sum_{j \geq k} \alpha^{2^j - 1} = \frac{\eta}{\alpha} \sum_{j \geq 0} \alpha^{2^{j+k}} = \frac{\eta}{\alpha} \sum_{j \geq 0} x^{2^j} \leq \frac{\eta}{\alpha} \sum_{j \geq 1} x^j \leq \eta \frac{\alpha^{2^k - 1}}{1 - \alpha^{2^k}}$$

Per a provar que $Fx^* = \vec{0}$ recordem que $\forall k \geq 0$, $Fx^k + F'(x^k)(x^{k+1} - x^k) = \vec{0}$; per tant

$$\begin{aligned} \left\|Fx^k\right\| = \left\|F'(x^k)(x^{k+1} - x^k)\right\| &\leq \left(\left\|F'(x^0)\right\| + \left\|F'(x^k) - F'(x^0)\right\|\right) \|x^{k+1} - x^k\| \\ &\leq \left(\left\|F'(x^0)\right\| + Mr^*\right) \|x^{k+1} - x^k\| \xrightarrow{k\to\infty} 0 \end{aligned}$$

de la continuïtat de F en x^* es té $\vec{0} = \lim_{k\to\infty} Fx^k = Fx^*$. □

El teorema original de Mysovskikh és diferent del que hem enunciat; en ell també es dona una regió d'unicitat de la solució (veure secció §4).

Observació 3.29. En la demostració no hem resolt l'equació en diferències, sinó que hem fitat la solució per una altra successió $t_{k+1} \leq \eta \sum_{j=0}^{k} \alpha^{2^j - 1}$. No obstant, en aquest cas sí es pot resoldre fàcilment i la solució és, precisament, $t_{k+1} = \eta \sum_{j=0}^{k} \alpha^{2^j - 1}$.

Observació 3.30. La condició M1) de Mysovskikh obliga que el teorema de Mysovskikh 3.28 només serveixi per a trobar zeros simples de l'operador F.

Majoració més general El que hem vist fins ara encara no és suficient per a provar (completament) el Teorema de Kantorovich; en particular, per a obtenir la regió d'unicitat de la solució que dona la demostració per recurrència (teorema 3.1). Per a això introduïm, sense demostració i de forma simplificada, els conceptes i resultats de Rheinboldt que adrecen aquesta part (veure [16, sec. 12.5] o, per a ampliar, [18]).

El problema radica en què, per al Teorema de Kantorovich, necessitem fites de $\|GGx - Gx\|$ que depenguin també del punt inicial x^0: $\|GGx - Gx\| \leq \varphi\big(\|Gx - x\|, \|Gx - x^0\|\big)$. Això ens porta a un resultat idèntic a la proposició 3.26, però amb l'equació en diferències

$$\begin{cases} t_k = f(t_{k-1}, t_{k-2}) = t_{k-1} + \varphi(t_{k-1} - t_{k-2}, t_{k-1}) \\ t_1 \geq \|x^1 - x^0\| \\ t_0 = 0 \end{cases} \tag{3.26}$$

on $\varphi(s,t)$, $\varphi\colon [0,\infty)^2 \mapsto [0,\infty)$ és creixent en cada variable. No obstant, aquest resultat garanteix l'existència d'un punt fix de G, però no la seva unicitat; per a això caldrà una formulació lleugerament modificada (teorema 3.32).

L'estudi de l'equació en diferències (3.26) pot simplificar-se notablement amb el concepte d'integral primera que introdueix Rheinboldt en [18]. Una *integral primera* de l'equació en diferències (3.26) és una funció contínua $\psi\colon [0,\infty) \mapsto [0,\infty)$ tal que, si $(t'_k)_{k\geq 0}$ satisfà

$$t'_k = \psi(t'_{k-1}), \qquad t'_0 = 0 \qquad \bigl(\forall k \geq 1\bigr)$$

aleshores també satisfà $t'_{k+1} - t'_k = \varphi(t'_k - t'_{k-1}, t'_k)$. Així, una integral primera $\psi(t)$ dona una solució de (3.26) amb condicions inicials $t_0 = 0$, $t_1 = \psi(0)$ i si convergeix ho fa a un punt fix de $\psi(t)$; una integral primera, per tant, redueix el problema de l'equació en diferències a estudiar un mètode iteratiu d'un pas (cada iterat depèn només de l'anterior). Rheinboldt dona una condició suficient perquè una funció sigui una integral primera de (3.26)

$$\psi(s) - \psi(t) = \varphi(s-t, t) \qquad \bigl(\forall s \geq t \geq 0\bigr) \tag{3.27}$$

en tal cas, $\psi(t)$ és creixent.

Amb el concepte d'integral primera podem encetar ja la unicitat de la solució del mètode iteratiu majorat per la solució de (3.26); per a això ens ajudarà el lema de Kantorovich.

Lema 3.31 (Kantorovich)**.** *Sigui $\psi\colon [t_0, s_0] \mapsto \mathbb{R}$, $[t_0, s_0] \subseteq [0,\infty)$, creixent. Si $t_0 \leq \psi(t_0)$ i $s_0 \geq \psi(s_0)$, les successions $t_{k+1} = \psi(t_k)$ i $s_{k+1} = \psi(s_k)$ estan ben definides, són creixent i decreixent, respectivament, i $\lim_{k\to\infty} t_k = t^* \leq s^* = \lim_{k\to\infty} s_k$.*

Si $\psi(t)$ és contínua, t^* i s^* són el menor i el major punts fixos de $\psi(t)$ en $[t_0, s_0]$.

Si podem assegurar que en $[t_0, s_0]$ només hi ha un punt fix, aleshores $t^* = s^*$ i $\lim_{k\to\infty}(s_k - t_k) = 0$; aquesta és la base de l'argument de Rheinboldt per a obtenir una regió d'unicitat per a un punt fix dels mètodes iteratius de Newton i similars.

Teorema 3.32 (Regió d'unicitat)**.** *Siguin un operador $G\colon D \mapsto X$, $D \subseteq X$ tancat, i una funció $\varphi\colon [0,\infty)^2 \mapsto [0,\infty)$ creixent en cada variable. Suposem que hi ha $x^0 \in D$ tal que*

1. *$\|Gy - Gx\| \leq \varphi\bigl(\|y - x\|, \|x - x^0\|\bigr)$, sempre que $x, y \in D$.*
2. *$(x^k)_{k\geq 0} \subseteq D$, $x^{k+1} = Gx^k$.*
3. *La solució $(t_k)_{k\geq 0}$ de l'equació en diferències (3.26) convergeix a $t^* < \infty$.*

Aleshores $\lim_{k\to\infty} x^k = x^* \in D$, es compleix (3.23) i, si G és continu en x^*, $Gx^* = x^*$.

Si a més, $\psi(t)$ satisfà (3.27) amb $\psi(0) = \|Gx^0 - x^0\|$, x^* és l'únic possible punt fix de G en $\overline{B(x^0, t^*)} \cap D$. Si per a algun $t' > t^*$, $\psi(t) < t$ quan $t \in (t^*, t')$; aleshores x^* és l'únic possible punt fix de G en $B(x^0, t') \cap D$.

Aquest resultat presenta el mateix problema que la proposició (3.26) a l'hora d'aplicar-lo, per tant la mateixa estratègia també ens servirà aquí (veure observació 3.27).

Rheinboldt aplica aquest resultat a provar el següent teorema, que constitueix la base d'una sèrie de resultats d'unicitat per a mètodes de Newton 'simples' [18, sec. 3] —$x^{k+1} = x^k - \Gamma F x^k$, Γ proper a $F'(x^0)^{-1}$— i Newton aproximats [18, sec. 4] —$x^{k+1} = x^k - A(x^k)^{-1} F x^k$, on, per a cada x, $A(x)$ és un operador lineal relacionat amb $F'(x)$—. Nosaltres ens limitarem només al mètode de Newton simplificat (1.3).

Teorema 3.33. *Sigui l'operador G-derivable $G\colon D_0 \mapsto X$, $D_0 \subseteq X$ obert convex. Prenem $x^0 \in D_0$ i suposem*

1. *$\|G'(x^0)\| \leq \delta < 1$.*

2. $\|Gx^0 - x^0\| \leq \eta$.
3. *Per a tot* $x, y \in D_0$, *es té* $\|G'(y) - G'(x)\| \leq \gamma \|y - x\|$.
4. $h = \gamma\eta/(1-\delta)^2 \leq \frac{1}{2}$.

A més definim $r^* = \frac{\eta}{(1-\delta)h}(1 - \sqrt{1-2h})$ *i* $r^{**} = \frac{\eta}{(1-\delta)h}(1 + \sqrt{1-2h})$ *i assumim* $\overline{B(x^0, r^*)} \subseteq D_0$.

Aleshores, la successió $x^{k+1} = Gx^k$, $k \geq 0$, *està ben definida, resta a* $\overline{B(x^0, r^*)}$ *i convergeix a un punt fix* x^*, *únic en* $\overline{B(x^0, r^*)}$, *de* G. *Quan* $h < 1/2$, x^* *és únic en* $B(x^0, r^{**}) \cap D_0$.

En particular, $(x^k)_{k\geq 0}$ *està majorada per la successió* $t_{k+1} = \frac{1}{2}\gamma t_k^2 + \delta t_k + \eta$, $t_0 = 0$.

Demostració. Si $x, y \in D_0$, aleshores per la proposició 2.33 i els punts 1 i 3

$$\begin{aligned}\left\|Gy - Gx\right\| &\leq \left\|Gy - Gx - G'(x)(y-x)\right\| + \left\|G'(x) - G'(x^0)\right\| \|y - x\| + \left\|G'(x^0)\right\| \|y - x\| \\ &\leq \frac{\gamma}{2} \|y - x\|^2 + \gamma \|x - x^0\| \|y - x\| + \delta \|y - x\|\end{aligned}$$

per tant $\|Gy - Gx\| \leq \varphi\big(\|y - x\|, \|x - x^0\|\big)$, on $\varphi(s, t) = \frac{\gamma}{2}s^2 + \gamma ts + \delta s$. L'equació en diferències (3.26) esdevé

$$t_{k+1} = t_k + \varphi(t_k - t_{k-1}, t_k), \qquad t_1 = \eta, \qquad t_0 = 0$$

que té com a una integral primera $\psi(t) = \frac{\gamma}{2}t^2 + \delta t + \eta$, ja que fàcilment es comprova (3.27); a més $\psi(0) = \eta$, per tant la solució de l'equació en diferències és $t_{k+1} = \psi(t_k)$, $t_0 = 0$.

La funció $\psi(t)$ és contínua, creixent i té dos punts fixos $r^* \leq r^{**}$; com que $t_0 \leq \psi(t_0) \leq r^*$, ha d'ésser $t_k \xrightarrow{k\to\infty} r^*$. Així, prenent $y = Gx$ tenim, inductivament, $(x^k)_{k\geq 0} \subseteq \overline{B(x^0, r^*)}$ està ben definida i majorada per $(t_k)_{k\geq 0}$; pel lema 3.25, és convergent a $x^* \in \overline{B(x^0, r^*)}$. Per tant, pel teorema 3.32, x^* és l'únic possible punt fix de G en $\overline{B(x^0, r^*)}$.

Si $h < 1/2$, $r^* < r^{**}$, i es té $\psi(t) < t$ per a tot $t \in (r^*, r^{**})$. Per tant, pel teorema 3.32, quan $h < 1/2$ l'únic possible punt fix de G en $B(x^0, r^{**}) \cap D_0$ és x^*.

Per a veure que $Gx^* = x^*$ observem que

$$\begin{aligned}\left\|Gx^* - x^*\right\| &\leq \left\|Gx^* - x^{k+1}\right\| + \|x^{k+1} - x^*\| = \left\|Gx^* - Gx^k\right\| + \|x^{k+1} - x^*\| \\ &\leq \left\|Gx^* - Gx^k - G'(x^0)(x^* - x^k)\right\| + \left\|G'(x^0)\right\| \|x^* - x^k\| + \|x^{k+1} - x^*\|\end{aligned}$$

i fent servir la proposició 2.30 amb G sobre D_0 tenim

$$\left\|Gx^* - x^*\right\| \leq \gamma r^* \|x^* - x^k\| + \left\|G'(x^0)\right\| \|x^* - x^k\| + \|x^{k+1} - x^*\| \xrightarrow{k\to\infty} 0$$

□

La redacció del teorema s'ha fet deliberadament semblant a la del Teorema de Kantorovich, per tal de ressaltar-ne la similitud. Com hem dit, l'apliquem al mètode de Newton simplificat.

Teorema 3.34. *Sigui l'operador G-derivable* $F\colon D \mapsto Y$, $D \subseteq X$ *obert. Suposem que es compleixen les condicions K1), K2), K3') i K4) (veure 3.1).*

Definim a més $r^* = \frac{1-\sqrt{1-2h}}{h}\eta$ *i* $r^{**} = \frac{1+\sqrt{1-2h}}{h}\eta$; *i suposem* $\overline{B(x^0, r^*)} \subseteq B(x^0, R)$.

Aleshores la successió $x^{k+1} = x^k - F'(x^0)^{-1}Fx^k$ *està ben definida, resta a* $\overline{B(x^0, r^*)}$, *convergeix a una solució* x^* *de* $Fx = \vec{0}$, *única en* $\overline{B(x^0, r^*)}$ *i satisfà*

$$\|x^* - x^k\| \leq \frac{\eta}{h}\big(1 - \sqrt{1-2h}\big)^k \qquad \big(\forall k \geq 0\big) \tag{3.28}$$

si a més, $h < 1/2$, *la solució* x^* *és única en* $B(x^0, r^{**}) \cap B(x^0, R)$.

Demostració. Fem $\Gamma_0 = F'(x^0)^{-1}$, que existeix per K1); i definim $G\colon D \mapsto X$, $Gx = x - \Gamma_0 Fx$.

Fàcilment es comprova que es verifiquen les hipòtesis del teorema anterior amb $D_0 = B(x^0, R)$, $\delta = 0$, $\gamma = \beta M$. Quant a (3.28), del teorema anterior tenim que $(x^k)_{k\geq 0}$ està majorada per $t_{k+1} = \psi(t_k) = \frac{1}{2}\beta M t_k^2 + \eta$; per tant es compleix (3.23) i es comprova per inducció que $r^* - t_{k+1} = \psi(r^*) - \psi(t_k) \leq \beta M r^*(r^* - t_k) \leq (\beta M r^*)^k (r^* - t_0) = \frac{\eta}{h}\left(1 - \sqrt{1-2h}\right)^{k+1}$ □

Observació 3.35. Quan $h = \frac{1}{2}$ la fita (3.28) resulta inútil, però com que sabem que el mètode convergeix, podem intentar reaplicar el teorema amb x^1.

3.3.4 Demostració per successions majorants

Finalment, demostrem el Teorema de Kantorovich 3.36, prenent com a guia la demostració d'Ortega [15] i afegint les aportacions de Tapia [19]. Recordem que el mètode de Newton és

$$x^{k+1} = x^k - F'(x^k)^{-1} F x^k \qquad \left(k \geq 0\right) \tag{3.29}$$

Teorema 3.36. *Sigui l'operador G-derivable $F\colon D \mapsto Y$, $D \subseteq X$ obert. Suposem que es compleixen les condicions K1), K2), K3') i K4).*

A més definim $r^* = \frac{1-\sqrt{1-2h}}{h}\eta$, $r^{**} = \frac{1+\sqrt{1-2h}}{h}\eta$; i suposem que $\overline{B(x^0, r^*)} \subseteq B(x^0, R)$.

Aleshores, la successió $x^{k+1} = x^k - F'(x^k)Fx^k$ està ben definida, resta a $\overline{B(x^0, r^*)}$, convergeix a una solució x^* de $Fx = \vec{0}$, única en $\overline{B(x^0, r^*)}$, i satisfà

$$\|x^* - x^k\| \leq \frac{1}{2^{k-1}}\left(2h\right)^{2^k - 1}\eta \qquad \left(\forall k \geq 0\right) \tag{3.30}$$

si a més $h < \frac{1}{2}$, la solució x^* és única en $B(x^0, r^{**}) \cap B(x^0, R)$.

Dividirem la prova en lemes, que n'aïllaran els punts essencials, tal com fa Ortega. En els lemes que segueixen s'assumeixen les hipòtesis del teorema 3.36. La idea de la demostració és veure que la successió dels iterats de Newton (3.29) està majorada per $(t_k)_{k\geq 0}$, la successió dels iterats de Newton del polinomi $p(t) = \frac{\beta M}{2}t^2 - t + \eta$, amb condició inicial $t_0 = 0$.

El primer lema ens dona informació sobre l'operador $Gx = x - F'(x)^{-1}Fx$. Observem que $r^* = \frac{1}{\beta M}(1 - \sqrt{1-2h}) \leq \frac{1}{\beta M}$.

Lema 3.37. *Sigui $Q = \left\{x \in B(x^0, R) : \|x - x^0\| < \frac{1}{\beta M}\right\}$.*

1. *$\forall x \in Q$, $F'(x)^{-1} \in \mathscr{B}(Y, X)$ i*

$$\|F'(x)^{-1}\| \leq \frac{\beta}{1 - \beta M \|x - x^0\|}$$

2. *Si x i $Gx = x - F'(x)^{-1}Fx$ són a Q, aleshores*

$$\|GGx - Gx\| \leq \frac{1}{2}\,\frac{\beta M \|Gx - x\|^2}{1 - \beta M \|Gx - x^0\|}$$

Demostració. Punt 1. Això és un cas particular del lema 3.3. El lema 3.3 es recolzava en què F tenia G-derivada segona i complia la condició K3), però es demostra igual en el nostre cas: F G-derivable + condició K3').

Punt 2. Si $x, Gx \in Q$, aleshores $Fx + F'(x)(Gx - x) = \vec{0}$. Així doncs

$$\begin{aligned}\|GGx - Gx\| = \|F'(Gx)^{-1}F(Gx)\| &\leq \frac{\beta}{1-\beta M\|x-x^0\|}\|F(Gx)\| \\ &= \frac{\beta}{1-\beta M\|x-x^0\|}\|F(Gx) - Fx - F'(x)(Gx-x)\|\end{aligned}$$

aplicant la proposició 2.33 a F sobre Q obtenim la desigualtat. □

Afegim a l'article d'Ortega el següent lema sobre el mètode de Newton per a paràboles, que crec que farà més clar el lema que ve després. Geomètricament és obvi: si un dibuixa la construcció dels iterats de Newton, les tesis del lema són evidents.

Lema 3.38. *Sigui $p(t) = c_2 t^2 + c_1 t + c_0$ una funció polinòmica real de grau 2, amb dues arrels reals (potser iguals) $a \leq b$.*

Per a qualsevol $t_0 < a$, la successió dels iterats del mètode de Newton, $t_{k+1} = t_k - \frac{p(t_k)}{p'(t_k)}$, està ben definida, és estrictament creixent i convergeix $\lim_{k\to\infty} t_k = a$.

Demostració. Els iterats de Newton són exactament els mateixos tant per a la funció $p(t)$ com per a $-p(t)$. Per tant, suposem que la paràbola està orientada cap a dalt ($c_2 > 0$).

Així per a $t_0 < a$ tenim $p'(t_0) < 0 < p(t_0)$; per tant l'equació $p(t_0) + p'(t_0)(t - t_0) = 0$ té una única solució, $t_1 > t_0$. Com que la gràfica de $p(t)$ queda per sobre de la recta tangent en qualsevol punt, podem assegurar que $a > t_1$. Repetint el procés trobem $a > t_2 > t_1 > t_0$.

Raonant per inducció es conclou que la successió està ben definida, és estrictament creixent i està afitada superiorment per a; per tant ha de convergir, i només ho pot fer a la arrel a. □

El lema té una conseqüència que és crucial per a estudiar l'error $\|x^* - x^k\|$.

Corol·lari 3.39. *La successió $(t_k)_{k\geq 0}$ del lema anterior satisfà la identitat*

$$t_{k+1} - t_k = -c_2\frac{(t_k - t_{k-1})^2}{c_1 + 2c_2 t_k} \qquad (\forall k \geq 1)$$

Demostració. Atès que $p(t)$ és un polinomi de grau 2, i recordant que $p(t_k) + p'(t_k)(t_{k+1} - t_k) = 0$, $\forall k \geq 0$, es té, per a tot $k \geq 1$

$$p(t_k) = p(t_{k-1}) + p'(t_{k-1})(t_k - t_{k-1}) + \tfrac{1}{2}p''(t_{k-1})(t_k - t_{k-1})^2 = c_2(t_k - t_{k-1})^2$$

per tant, els iterats de Newton són $t_{k+1} = t_k - \frac{p(t_k)}{p'(t_k)} = t_k - \frac{c_2(t_k - t_{k-1})^2}{c_1 + 2c_2 t_k}$, com volíem veure. □

Lema 3.40. *La successió (3.29) està ben definida i està majorada per $(t_k)_{k\geq 0}$*

$$t_{k+1} = t_k - \frac{\frac{\beta M}{2}t_k^2 - t_k + \eta}{\beta M t_k - 1} \qquad t_0 = 0 \qquad (\forall k \geq 0)$$

a més $\lim_{k\to\infty} t_k = r^*$, on r^* és el del teorema 3.36.

Demostració. Fem $p(t) = \frac{\beta M}{2} t^2 - t + \eta$, les seves arrels són $r^* \le r^{**}$, i la successió $(t_k)_{k\ge 0}$ són els iterats de Newton per a la funció real $p(t)$ amb $t_0 = 0 < r^*$. Per tant, pel lema 3.38, $(t_k)_{k\ge 0}$ és estrictament creixent i convergeix a r^*. Així doncs, només hem de veure que $(x^k)_{k\ge 0}$ està ben definida i $\|x^{k+1} - x^k\| \le t_{k+1} - t_k$ per a tot $k \ge 0$.

Fem la prova per inducció. Definim $Q = B(x^0, R) \cap B(x^0, \frac{1}{\beta M})$ i l'operador $G\colon Q \mapsto X$, $Gx = x - F'(x)^{-1}Fx$, que té sentit perquè $F'(x)$ és bijectiu $\forall x \in Q$, pel lema 3.37.

Clarament $x^1 = Gx^0$ existeix i $\|x^1 - x^0\| \le \eta = t_1 - t_0 < r^*$; per tant $x^1 \in B(x^0, r^*) \subseteq Q$.

Suposem que existeixen $x^1, \dots, x^n \in B(x^0, r^*)$ i es compleix $\|x^{k+1} - x^k\| \le t_{k+1} - t_k$ per a tot $0 \le k < n$; volem veure que també existeix $x^{n+1} \in B(x^0, r^*)$ i $\|x^{n+1} - x^n\| \le t_{n+1} - t_n$. Com que $x^n \in Q$, $x^{n+1} = Gx^n$ existeix i, aplicant el lema 3.37 i la hipòtesi d'inducció

$$\begin{aligned}\|x^{n+1} - x^n\| &= \left\|GGx^{n-1} - Gx^{n-1}\right\| \le \frac{1}{2} \frac{\beta M \left\|Gx^{n-1} - x^{n-1}\right\|^2}{1 - \beta M \left\|Gx^{n-1} - x^0\right\|} \\ &= \frac{1}{2} \frac{\beta M \|x^n - x^{n-1}\|^2}{1 - \beta M \|x^n - x^0\|} \le \frac{1}{2} \frac{\beta M (t_n - t_{n-1})^2}{1 - \beta M \|x^n - x^0\|}\end{aligned}$$

atès que $t \mapsto \frac{1}{1-t}$, $t \in [0,1)$, és una funció creixent i $\|x^n - x^0\| \le \sum_{j=0}^{n-1} t_{j+1} - t_j = t_n$

$$\|x^{n+1} - x^n\| \le \frac{1}{2} \frac{\beta M (t_n - t_{n-1})^2}{1 - \beta M t_n} = t_{n+1} - t_n \tag{3.31}$$

on l'última igualtat és pel corol·lari 3.39 aplicat al polinomi $p(t)$. Per a completar la inducció només resta veure que $x^{n+1} \in B(x^0, r^*)$, però això és obvi ja que $(t_k)_{k\ge 0}$ és estrictament creixent: $\|x^{n+1} - x^0\| \le t_{n+1} - t_0 = t_{n+1} < r^*$ □

Demo teorema 3.36. La prova és ara immediata: els lemes 3.40 i 3.25 garanteixen que la successió (3.29) $(x^k)_{k\ge 0}$ està ben definida, resta a $B(x^0, r^*) \subseteq B(x^0, R) \cap B(x^0, \frac{1}{\beta M})$, i convergeix a $x^* \in \overline{B(x^0, r^*)}$.

Per a provar que $Fx^* = \vec{0}$ recordem que $\forall k \ge 0$, $Fx^k + F'(x^k)(x^{k+1} - x^k) = \vec{0}$; per tant

$$\begin{aligned}\left\|Fx^k\right\| &= \left\|F'(x^k)(x^{k+1} - x^k)\right\| \le \left(\left\|F'(x^0)\right\| + \left\|F'(x^k) - F'(x^0)\right\|\right) \|x^{k+1} - x^k\| \\ &\le \left(\left\|F'(x^0)\right\| + Mr^*\right) \|x^{k+1} - x^k\| \xrightarrow{k\to\infty} 0\end{aligned}$$

per la condició K3') F és $\mathscr{C}^1$ en $B(x^0, R)$, i per tant contínua; de la continuïtat de F en x^* es té $\vec{0} = \lim_{k\to\infty} Fx^k = Fx^*$.

La regió d'unicitat s'obté via el mètode de Newton simplificat: teorema 3.34. Quant a la fita d'error (3.30), Ortega no la dona ni la demostra en [15]; en una demostració posterior [16, teo. 12.6.2] Ortega i Rheinboldt proven per inducció que $\forall k \ge 0$, $t_{k+1} - t_k \le \eta 2^{-k}$, d'on dedueixen, també per inducció, $r^* - t_k \le \left(\beta M 2^k\right)^{-1} (2h)^{2^k}$. Tapia fa un raonament similar, però obté una fita lleugerament millor i a més prova que és òptima en un cert sentit. □

Fites òptimes de Tapia Tapia obté en [19] obté la següent fita alternativa a (3.30), que és millor ja que $1 - \sqrt{1-2h} \le 2h$ quan $0 \le h \le 1/2$ (lema 3.2)

$$\|x^* - x^k\| \le \frac{\eta}{h} \frac{\left(1 - \sqrt{1-2h}\right)^{2^k}}{2^k} \tag{3.32}$$

i prova que és la millor en el sentit formulat en el punt 4. Desglossem el seu raonament en els següents punts.

1. Els iterats de Newton $(t_k)_{k\geq 0}$, $t_0 = 0$, del polinomi $p(t) = \frac{\beta M}{2} t^2 - t + \eta$, $(0 \leq h \leq 1/2)$, estan fitats pels iterats de Newton $(t'_k)_{k\geq 0}$, $t'_0 = 0$, del polinomi $p(t)$ quan $h = 1/2$. És a dir, $\forall h \in [0, 1/2], \forall k \geq 0,\ t_k \leq t'_k$.
2. La successió $(t'_k)_{k\geq 0}$ satisfà $\eta - \frac{1}{2} t'_k = \eta 2^{-k}$; per tant $\forall k \geq 0$, $(\eta - \frac{1}{2} t_k)^{-1} \leq 2^k/\eta$.
3. La successió $(t_k)_{k\geq 0}$ satisfà $r^* - t_{k+1} = \frac{h(r^*-t_k)^2}{2(\eta - h t_k)} \leq \frac{h}{\eta} 2^{k-1}(r^* - t_k)^2$.
4. Cerquem el millor $\alpha(h)$ tal que $\forall k \geq 0$, $r^* - t_k \leq \frac{\eta}{h} \frac{(\alpha(h))^{2^k}}{2^k} \Rightarrow r^* - t_{k+1} \leq \frac{\eta}{h} \frac{(\alpha(h))^{2^{k+1}}}{2^{k+1}}$.

El primer punt es té revisant (3.31) i reescrivint-la així (fixem $0 \leq h = \beta M \eta \leq 1/2$)

$$t_{k+1} - t_k = \frac{1}{2} \frac{h(t_k - t_{k-1})^2}{\eta - h t_k} \leq \frac{1}{2} \frac{\frac{1}{2}(t_k - t_{k-1})^2}{\eta - \frac{1}{2} t_k}, \qquad t_0 = 0, \qquad t_1 = \eta \qquad (\forall k \geq 1)$$

És clar que $(t'_k)_{k\geq 0}$, $t'_0 = 0$, $t'_1 = \eta$, està generada per la relació de recurrència $t'_{k+1} - t'_k = \frac{1}{2} \frac{(1/2)(t'_k - t'_{k-1})^2}{\eta - t'_k/2}$; de la mateixa manera que $(t_k)_{k\geq 0}$, $t_0 = 0$, $t_1 = \eta$, ho està per la recurrència $t_{k+1} - t_k = \frac{1}{2} \frac{h(t_k - t_{k-1})^2}{\eta - h t_k}$. De la desigualtat anterior, per tant, $\forall k \geq 0$, $t_{k+1} - t_k \leq t'_{k+1} - t'_k$; del fet que $t'_0 = t_0 = 0$, $t'_1 = t_1 = \eta$, se segueix $t_k \leq t'_k$ per a tot $k \geq 0$.

El segon punt es verifica fàcilment per inducció. En el tercer punt, la desigualtat és evident a partir del punt anterior; la igualtat, en canvi, requereix una mica de feina. Els iterats de Newton del polinomi $p(t)$, $(t_k)_{k\geq 0}$, $t_0 = 0$, són $t_{k+1} = t_k - \frac{p(t_k)}{p'(t_k)} = t_k - \frac{(\beta M/2) t_k^2 - t_k + \eta}{\beta M t_k - 1} = \frac{(\beta M/2) t_k^2 - \eta}{\beta M t_k - 1}$; amb això i $(r^*)^2 = \frac{2}{\beta M}(r^* - \eta)$ es comprova la igualtat desitjada

$$r^* - t_{k+1} = \frac{r^* - \beta M r^* t_k + (\beta M/2) t_k^2 - \eta}{1 - \beta M t_k} = \frac{\beta M}{2} \frac{(r^* - t_k)^2}{1 - \beta M t_k} = \frac{h}{2} \frac{(r^* - t_k)^2}{\eta - h t_k}$$

El quart punt és evident a partir del tercer; per a trobar el millor $\alpha(h)$ observem que $r^* - t_0 = \frac{\eta}{h}(1 - \sqrt{1 - 2h})$, per tant, fixat $h \in [0, 1/2]$, $\alpha(h) = 1 - \sqrt{1 - 2h}$ és el menor valor possible que satisfà el punt quart.

3.4 Comentaris finals

Fem un resum del més rellevant d'aquesta secció, comparant ambdues demostracions i la informació obtinguda de cadascuna; també afegim altra informació relativa al Teorema de Kantorovich i la seva aplicació que no s'ha vist en el text.

Hem formulat i demostrat el Teorema de Kantorovich de dues formes subtilment diferents: una demana que la segona derivada estigui afitada i l'altre ho canvia per una condició de Lipschitz sobre la primera derivada. La primera prova és per recurrència (teorema 3.1) i la segona per successions majorants (teorema 3.36). Aquesta distinció —condicions K3) i K3')— és innòcua, ja que realment no afecta les demostracions dels teoremes i, en la pràctica, la dificultat de comprovar que es compleixen és similar. No obstant, hi ha altres variacions d'aquesta i altres condicions que sí porten a modificacions substancials del Teorema (veure secció §4). Una modificació habitual i que no altera en absolut les demostracions donades aquí, és substituir la bola $B(x^0, R)$ per un obert convex $D_0 \subseteq D$; això és innocu a nivell teòric, però sí té una certa importància des d'un punt de vista pràctic, com expliquem i veiem en §5.

També hem vist que el Teorema és vàlid no sols per al mètode de Newton, sinó també per al mètode de Newton simplificat (teoremes 3.13 i 3.34). És interessant, però, que aquesta equivalència del Teorema de Kantorovich entre ambdós mètodes s'obté de manera diferent en cada cas. En la demostració per recurrència, la prova important és la del mètode de Newton, quedant la de Newton simplificat com un corol·lari; en canvi, en la demostració per successions majorants, per a completar la prova per al mètode de Newton necessitem la prova per al de Newton simplificat, altrament la regió d'unicitat seria menor. En la demostració del propi Kantorovich amb el principi de majoració, això és més accentuat (secció XVIII teorema 5 de [12]): el Teorema s'estableix primer per al mètode de Newton simplificat, quedant el de Newton com una conseqüència directa. Val a dir que amb la demostració per recurrència s'estableix una equivalència parcial ($h < 1/2$), mentre que amb la de successions majorants (i també amb el principi de majoració) s'aconsegueix una equivalència total ($h \leq 1/2$).

La prova per recurrència és conceptualment senzilla, però feixuga, i difícilment es pot reproduir per als mètodes quasi-Newton (1.4)(1.5); la prova per successions majorants, en canvi, és més simple. Les successions majorants ofereixen més flexibilitat i permeten reproduir la prova per als mètodes quasi-Newton [18], però a canvi té el cost d'haver d'introduir un problema auxiliar d'una equació en diferències per poder trobar una successió majorant. Malgrat tot això, la demostració per recurrència té una idea de fons similar a les successions majorants (observació 3.11).

També hem vist que el Teorema garanteix que la solució x^* és simple ($F'(x^*)$ és bijectiu i $F'(x^*)^{-1}$ és continu) quan $h < 1/2$ (observació 3.12). Quan $h = 1/2$, en canvi, pot ocórrer que no sigui simple, com mostra el següent exemple.

Exemple 3.41. Prenem $f\colon \mathbb{R} \mapsto \mathbb{R}$, $f(x) = (x-2)^2$; clarament té un zero $x = 2$ que no és simple, però el mètode de Newton es pot aplicar a qualsevol $x \neq 2$. El Teorema de Kantorovich 3.1 es pot aplicar prenent $x_0 = 1.9$, $\beta = 1/f'(x_0) = 5$, $\eta = f(x_0)/f'(x_0) = 1/20$, $R = \infty$, $M = 2$; això fa que $h = \beta M \eta = \frac{1}{2}$. Per tant, el mètode de Newton $x_{k+1} = x_k - f(x_k)/f'(x_k)$, $x_0 = 1.9$ convergeix al zero no simple $x = 2$.

En sentit contrari, és natural esperar que el Teorema de Kantorovich pugui aplicar-se prop d'una solució simple, o inclús amb ella mateixa com a condició inicial; en aquest sentit tenim els següents resultats.

Proposició 3.42 (Radi d'unicitat local)**.** *Siguin un operador amb G-derivada segona $F\colon D \mapsto Y$, $D \subseteq X$ obert, i $x^* \in D$ una solució simple de l'equació $Fx = \vec{0}$.*

Suposem que $\left\| F'(x^*)^{-1} \right\| \leq \beta^*$, hi ha $M, R > 0$ tal que $B(x^*, R) \subseteq D$ i $\forall x \in B(x^*, R)$, $\left\| F''(x) \right\| \leq M$. Aleshores x^* és l'única solució en $B(x^*, \frac{2}{\beta^* M}) \cap B(x^*, R)$.

Proposició 3.43. *Mantenim les hipòtesis de la proposició anterior. Les hipòtesis del Teorema de Kantorovich 3.1 se satisfan per a qualsevol $x^0 \in \overline{B(x^*, \frac{1}{4\beta^* M})} \cap B(x^*, R)$. Se satisfan amb $h = 1/2$ si, i només si, $\|x^* - x^0\| = (4\beta^* M)^{-1}$.*

La demostració de la primera és simplement aplicar el Teorema de Kantorovich 3.1 prenent $x^0 = x^*$, $\eta = 0$ (aleshores $r^* = 0$ i $r^{**} = (\beta^* M)^{-1}$). La segona la comenta, sense demostració, Kantorovich en [8] (traduït a l'anglès en[11]); una demostració detallada es pot trobar en [1, teo. 26.1], que el mateix autor millorà després augmentant $(4\beta^* M)^{-1}$ a $(2-\sqrt{2})(2\beta^* M)^{-1}$ i mostrant que aquest últim valor és òptim [2].

Malgrat l'exemple 3.41, quan la solució x^* de $Fx = \vec{0}$ no és simple, les hipòtesis del Teorema de Kantorovich poden no complir-se per a punts arbitràriament propers a x^* (veure l'exemple de [1, pp. 190-191]).

Finalment, quant a l'obtenció de fites per a l'error $\|x^* - x^k\|$ i la seva optimalitat, tenim un article de Gragg i Tapia [5] en què donen fites, tant inferiors com superiors, de l'error i discuteixen l'optimalitat.

4 Variants del Teorema i resultats similars

A la literatura matemàtica hi ha una infinitat de variacions del Teorema de Kantorovich o resultats que s'hi assemblen. Llistar-los tots seria impossible, n'apuntem només alguns. Bons punts de partida per trobar versions o resultats similars al Teorema, són els articles d'en I. K. Argyros, els llibres d'Ezquerro i Hernández [3, 4] i el d'Ortega i Rheinboldt [16], i les referències que ells es troben.

Recordem que el mètode de Newton per a un operador F, amb condició inicial x^0, és

$$x^{k+1} = x^k - \Gamma_k F x^k \qquad \big(\forall k \geq 0\big) \tag{4.1}$$

on $\Gamma_k = F'(x^k)^{-1} \in \mathscr{B}(Y, X)$.

Mètode quasi-Newton i pertorbacions Les primeres modificacions del Teorema de Kantorovich en què podem pensar són en els mètodes quasi-Newton (veure Introducció) i també el mètode de Newton assumint que no podem calcular F exactament. Molts resultats en aquest sentit es poden trobar en els teoremes, notes i exercicis de [16] al llarg del capítol 12.

Mysovskikh Al 1950 I. P. Mysovskikh publicà un article [14] en què complementava els recents resultats d'en L. V. Kantorovich sobre equacions funcionals $Fx = \vec{0}$ i els mètodes de Newton i Newton simplificat [6, 7, 8]. Presentava dos teoremes teòrics i dos sobre aplicacions a sistemes d'equacions algebraiques i equacions integrals.

Aquest teorema l'hem enunciat i provat en aquest text (teorema 3.28). No obstant la formulació original d'en Mysovskikh (i per descomptat demostració) era diferent i, sobretot, ens interessa destacar que incloïa una menció sobre la regió d'unicitat de la solució: la solució x^* que dona el teorema 3.28 és única en $B(x^0, \frac{2}{\beta M})$; sempre que sobre ella es mantingui la condició M3) (veure 3.1), i $h < \lambda \approx 1.13$, on λ és la solució de l'equació $h \sum_{j \geq 0} ({}^h\!/_2)^{2^j - 1} = 2$.

Diferents variacions del teorema de Mysovskikh, en el sentit del paràgraf inicial d'aquesta secció (mètodes quasi-Newton o pertorbats), es poden trobar en [18] i [16, cap. 12].

Mètode de Newton amortit Una modificació interessant del mètode de Newton $x^{k+1} = x^k - F'(x^k)^{-1} F x^k$ és exigir que la norma decreixi en cada iteració: $\|Fx^{k+1}\| \leq \|Fx^k\|$. Una opció és

$$x^{k+1} = x^k - \lambda_k F'(x^k)^{-1} F x^k$$

on els λ_k s'escullen perquè es complexi la condició de reducció de la norma. Aquests λ_k es poden determinar de moltes maneres, una clàssica n'és la regla d'Armijo. No obstant, en el nostre cas, amb les condicions de Kantorovich o similars podem fer servir el següent resultat extret de [16], Exercici 7.1-1.

Lema 4.1. *Sigui l'operador $F\colon D \mapsto Y$, $D \subseteq X$ obert convex, G-derivable en D. Suposem que existeix $y = F'(x)^{-1} F x \in D$, aleshores hi ha $\lambda > 0$ tal que $\big\|F(x - \lambda y)\big\| \leq \big\|Fx\big\|$.*

Si a més, F' compleix

$$\left\|F'(x-\lambda y)-F'(x)\right\| \le M|\lambda|\,\|y\| \qquad \big(\forall \lambda\in(0,1)\big)$$

aleshores qualsevol $\lambda \le \|Fx\|\,\frac{1}{M\|y\|^2}$ *serveix.*

5 Exemples i aplicacions del Teorema de Kantorovich

Exposem un seguit d'exemples de com fer servir el Teorema de Kantorovich, tant en dimensió finita com infinita. Centrarem la nostra atenció, especialment, en emprar-lo per a provar l'existència de solucions d'equacions; malgrat també hi ha algun exemple sobre altres aspectes del Teorema, com les regions d'existència i unicitat local d'una solució.

Per aplicar el Teorema de Kantorovich a una equació $Fx=\vec{0}$, cal trobar una condició inicial x^0 i un obert convex $x^0\in D_0$ i comprovar que se satisfan les hipòtesis del Teorema (teoremes 3.1 o 3.36). La validesa de l'ús d'un obert convex D_0 en comptes d'una bola $B(x^0,R)$ en el Teorema de Kantorovich s'explicà en l'apartat 3.1.

El problema és, doncs, aquesta primera part de trobar x^0 i D_0; en els diferents exemples es veurà com adreçar aquesta qüestió, que dependrà molt del problema concret. També cal tenir present un altre problema relacionat: el Teorema troba zeros simples de F, però pot no detectar zeros no simples (veure apartat 3.4). Per tant, estudiar F' —saber on és invertible amb invers continu— serà un assumpte a tenir en compte a l'hora de cercar zeros de F, per saber si n'hi ha més dels que haguem pogut trobar amb el Teorema de Kantorovich.

Així doncs, de cara a aplicar el Teorema per a *provar l'existència d'un zero* d'un operador, ens interessarà, primer, localitzar possibles zeros i saber si podrien ésser simples o no, i aleshores comprovar les hipòtesis del Teorema per a demostrar-ne la seva existència. És a dir, farem el següent

1. Triar un D_0 obert convex que creiem que conté algun zero i tal que hi puguem fitar la segona derivada de l'operador F.
2. Estudiar si hi ha $x\in D_0$ tals que l'operador lineal $F'(x)$ no és bijectiu o $F'(x)^{-1}$ no és continu: saber si D_0 pot contenir un zero no simple.
3. Cercar $x^0\in D_0$ amb què aplicar el Teorema de Kantorovich per a provar que F té un zero en D_0.

Càlcul i accessibilitat de les arrels d'un nombre real Aquest exemple se centra en mostrar la gran discrepància que hi ha entre dos conjunts de condicions inicials x^0; el de les que fan convergir el mètode de Newton, i el de les que satisfan les hipòtesis del Teorema de Kantorovich.

Aquesta idea de quines condicions inicials fan que un mètode iteratiu convergeixi a alguna solució d'una equació concreta, és el que anomenem *accessibilitat d'un mètode iteratiu.* Ens basem en dues formulacions d'aquesta idea que fan Ezquerro i Hernández [4, sec. 1.3].

Definició 5.1 (Conca d'atracció)**.** Donat un mètode iteratiu (veure Introducció) $G(x)$ i una solució x^* d'una equació, anomenem *conca d'atracció* de x^* el conjunt de condicions inicials x^0 per a què el mètode iteratiu convergeix a x^*

$$C(x^*)=\left\{x^0:\ \lim_{k\to\infty}x^k=x^*,\ x^{k+1}=Gx^k\ \forall k\ge 0\right\}$$

Definició 5.2 (Regió d'accessibilitat)**.** Assumim que tenim un resultat de convergència semilocal (veure Introducció) per a un mètode iteratiu $G(x)$, com ho és el de Kantorovich per al mètode de Newton. Anomenem *regió d'accessibilitat* el conjunt de condicions inicials x^0 per a què se satisfan les hipòtesis d'aquest resultat de convergència semilocal.

Considerem el problema d'aproximar l'arrel n-éssima real de $a > 0$, $\sqrt[n]{a}$. Això és aproximar el zero real i positiu de l'equació

$$x^n - a = 0 \tag{5.1}$$

Per a fer servir el mètode de Newton necessitem una condició inicial x_0 per a què el mètode estigui definit i convergeixi a $\sqrt[n]{a}$. En aquest cas, $\sqrt[n]{a}$ és una solució simple (de fet totes les solucions ho són) de (5.1); per tant, podem estar segurs que, en un entorn d'ella, el mètode de Newton està definit i hi convergeix. De fet podem assegurar que es compleixen les hipòtesis del Teorema de Kantorovich en un entorn de la solució $\sqrt[n]{a}$ (proposició 3.43).

Una opció segura seria aplicar el mètode de la bisecció fins a obtenir una aproximació prou propera de $\sqrt[n]{a}$ amb què iniciar el mètode de Newton. No obstant, en aquest cas, la conca d'atracció $C(\sqrt[n]{a})$ és molt gran i fàcil de calcular, podent escollir qualsevol $x_0 > 0$

$$C(\sqrt[n]{a}) = \begin{cases} (0,\infty) & n > 0 \text{ parell} \\ \mathbb{R} \setminus E & n > 0 \text{ senar} \end{cases}$$

on $E \subseteq (-\infty, 0]$ és un conjunt numerable. No ens hi entretindrem perquè no és l'objectiu, però això és molt fàcil de veure amb la gràfica de la funció $f(x) = x^n - a$. Per exemple, fàcilment $(0,\infty) \subseteq C(\sqrt[n]{a})$, i l'argument és similar al del lema 3.38: la gràfica de la funció $f_{\restriction(0,\infty)}$ queda per sobre de la recta tangent en qualsevol punt de $(0,\infty)$; és a dir, $f(x)$ és convexa en $(0,\infty)$.

Veiem, per tant, que l'equació (5.1) és un cas extrem en què *no importa quina condició inicial* $x_0 > 0$ *agafem*, el mètode de Newton convergirà a la solució $\sqrt[n]{a}$. Les regions d'accessibilitat, en canvi, i com veurem ara, són molt més petites. Per veure-ho bé considerem l'equació (5.1) en els complexos i amb valors concrets $n = 3$ i $a = 1$: cerquem solucions de l'equació

$$f(z) = z^3 - 1 = 0 \tag{5.2}$$

La figura A.1 a la pàgina 46 mostra les conques d'atracció de les solucions 1, $\exp(\frac{2}{3}\pi \mathrm{i})$, $\exp(\frac{4}{3}\pi \mathrm{i})$ de (5.2) en el rectangle $R = \{-2 \le \mathrm{Re}(z), \mathrm{Im}(z) \le 2\}$ i la regió d'accessibilitat en el disc $D_0 = \{|z| < 2\}$. Les imatges fan palesa la gran diferència que pot haver entre aquests conjunts.

Com es dibuixen? Prenem una xarxa de punts en R. Les conques d'atracció es dibuixen assignant a cada solució un color i prenent cada punt com a condició inicial del mètode de Newton; el punt es dibuixa només si s'atansa a alguna de les solucions, i amb el color de tal solució. La regió d'accessibilitat per al Teorema de Kantorovich, amb la definició donada dalt, és complicada de calcular; però podem fer-nos una idea simplificant la situació: fixem l'obert convex D_0, que conté les solucions, i a on sabem fitar la segona derivada $|f''(z)| \le 12 = M$. Aleshores en cada punt z_0 de la xarxa en R tal que $z_0 \in D_0$, comprovem les hipòtesis del Teorema de Kantorovich 3.1 i el dibuixem només si se satisfan; les comprovacions es redueixen a calcular

$$\beta = \frac{1}{|f'(z_0)|} \qquad \eta = \beta |f(z_0)| \qquad h = \beta M \eta \qquad r^* = \frac{1 - \sqrt{1 - 2h}}{\beta M}$$

i verificar $h \le \frac{1}{2}$ i $\overline{B(x^0, r^*)} \subseteq D_0$. S'inclouen a l'annex B els programes en llenguatge C i les ordres del Gnuplot per a crear imatges de les conques d'atracció i de la regió d'accessibilitat.

En resum, és esperable que la regió d'accessibilitat sigui bastant petita en comparació a les conques d'atracció; per tant, una estratègia per a provar l'existència de solucions és fer servir el Teorema de Kantorovich *a posteriori*: iterar primer fins a atansar-nos a una solució, i prendre un d'aquests iterats com a condició inicial per al Teorema; provant així, formalment, l'existència de la solució que s'ha aproximat.

Càlcul dels zeros d'un polinomi complex Aquesta aplicació se centra en emprar el Teorema de Kantorovich per a trobar i calcular els zeros d'un polinomi complex. També mostrarem l'accessibilitat d'aquests zeros (veure l'exemple anterior).

Volem calcular els zeros del polinomi complex $p(z) = z^4 - 2z^3 + 7z - 3$. Ja sabem que existeixen (Teorema Fonamental d'Àlgebra). Hi ha moltes formes de fer això, i més fàcils que emprar el Teorema de Kantorovich, però no ens complicarem, ja que només volem mostrar com fer servir el Teorema, ja aquesta estratègia també ens servirà quan no sapiguem de l'existència de zeros.

Seguint l'esquema plantejat a l'inici d'aquesta secció, hem de localitzar els zeros (trobar un obert convex D_0 que els contingui) i intentar esbrinar si n'hi ha algun de no simple.

Quant a la simplicitat dels zeros, en aquest cas és molt senzill provar que tots els zeros són simples, ja que $p(z)$ és un polinomi, i un polinomi només té zeros simples si, i només si, $\operatorname{mcd}(p(z), p'(z)) = 1$. Cosa que es comprova fàcilment.

Quant a la localització dels zeros, ens saltem els detalls, però es pot veure que tots són dins el disc $D_0 = \{|z| < 2.5\}$; això es pot veure amb el Teorema de Rouché[4]. De fet, amb una anàlisi més a fons es pot veure que té dos zeros reals, un en $(-2, 0)$, i l'altre prop de $z_0 = \frac{1}{2}$.

Per a calcular els zeros de $p(z)$ podem prendre una xarxa de punts en el rectangle $Q = [-2.5, 2.5] \times [-2.5, 2.5]$ —que conté D_0— i, en cada punt de la xarxa, comprovar si se satisfan les hipòtesis del Teorema de Kantorovich 3.1, i en cas afirmatiu, engeguem el mètode de Newton per a aproximar el zero trobat. Això es pot fer fàcilment modificant lleugerament el programa 2 de l'annex B; aquest procés ens donarà la llista dels quatre zeros de $p(z)$

$$z_0 \approx 1.564 - 1.339\mathrm{i} \qquad z_1 = \overline{z_0} \qquad z_2 \approx -1.577 \qquad z_3 \approx 0.449$$

Aquesta mateixa estratègia es pot fer servir per a trobar i calcular solucions de sistemes d'equacions reals o complexes $Fx = \vec{0}$, quan no sabem, a priori, si existeixen zeros o no en un obert convex D_0, com veurem tot seguit en el següent exemple.

Reprenent les idees d'accessibilitat introduides en l'exemple anterior (definicions 5.1 5.2), la figura A.2 a la pàgina 47 mostra l'accessibilitat dels zeros de $p(z)$. El mètode per a dibuixar la imatge ha sigut el mateix que s'ha descrit en l'exemple anterior.

Sistema d'equacions en $\mathbb{R}^2$ Aquest exemple del Teorema de Kantorovich se centra fer-lo servir per a provar l'existència i unicitat local de zeros d'un operador.

Plantegem el sistema de dues equacions en dues incògnites extret de [20, pàg. 51]

$$\begin{cases} f_1(x, y) = x^2 + y^2 = 1 \\ f_2(x, y) = \sin(\pi x/2) + y^3 = 0 \end{cases}$$

[4]Wikipedia. *Teorema de Rouché* [en línia]. Disponible a `https://es.wikipedia.org/wiki/Teorema_de_Rouch%C3%A9` [consulta: 19 de juny de 2021].

Novament, seguint el plantejament descrit a l'inici de la secció, primer intentem obtenir una regió a on poden ser les solucions. Observem que $-f_2(x,y) = f_2(-x,-y)$, per tant és suficient cercar solucions en $\mathbb{R}\times[0,\infty)$. La primera equació ens diu $0 < |x|,|y| < 1$, i amb això, de la segona tenim que x i y han de tenir diferent signe. En suma, podem restringir la cerca de solucions a la regió $Q = (-1,0)\times(0,1)$.

Ara toca preocupar-se de la possible existència de solucions no simples; examinem la derivada de $F(x,y) = (f_1(x,y)-1, f_2(x,y))$ en Q

$$F'(x_0,y_0) = \begin{pmatrix} 2x_0 & 2y_0 \\ \frac{\pi}{2}\cos(\pi x_0/2) & 3y_0^2 \end{pmatrix} \qquad \det F'(x_0,y_0) = 6x_0y_0^2 - \pi y_0\cos(\pi x_0/2)$$

on $\det F'(x_0,y_0) = 0 \Leftrightarrow 6x_0y_0 - \pi\cos(\pi x_0/2) = 0$. Per a cada $y_0 \in (0,1)$, $\det F'(x_0,y_0)$ és una funció estrictament creixent respecte de $x_0 \in (-1,0)$; atès que $\det F'(0,y_0) < 0$, concloem que $F'(x_0,y_0)$ és regular per a tot $(x_0,y_0)\in Q$. Per tant si F té zeros en Q, són simples.

Per a provar l'existència d'un zero, prenem una xarxa de punts sobre Q i comprovem si es verifiquen les hipòtesis del Teorema de Kantorovich en cada punt de la xarxa. Això ens permetrà obtenir solucions, però no podem saber si les hem trobat totes o no. Sabent que hi ha solucions simples, la proposició 3.42 ens pot ajudar a escollir la mida de la xarxa de punts.

En aquest cas trobem $p^* \approx (-0.4761, 0.8794) \in Q$ i es pot veure amb el gràfic de F, que és l'única solució en Q, però en altres casos en què no disposem d'aquesta ajuda, no sabríem si hem d'escollir una xarxa més fina o els errors de càlcul ens impedeixen trobar-la.

Òbviament aquesta estratègia de localitzar zeros, i iterar sobre una xarxa de punts, presenta problemes evidents a mida que augmenta la dimensió de l'espai n. I per descomptat

Teorema de la Funció Inversa Aquesta aplicació del Teorema de Kantorovich se centra en les regions d'existència i unicitat que dona el Teorema.

El Teorema de la Funció Inversa és un resultat important de càlcul diferencial en diverses variables: dona condicions suficients perquè una funció $F\colon D\mapsto\mathbb{R}^n$, $D\subseteq\mathbb{R}^n$ obert, sigui localment invertible en un punt $x^0\in D$ i amb la mateixa regularitat que F. És a dir, garanteix que hi ha $U\subseteq D$, entorn obert de x^0, i V, entorn obert de Fx^0, tals que $F_U\colon U\mapsto V$, $F_U(x) = Fx$, és bijectiva i F_U^{-1} és $\mathscr{C}^r$ si F_U també ho és ($r\geq 1$).

Amb el Teorema de Kantorovich podrem donar estimacions concretes d'aquests conjunts oberts que el Teorema de la Funció Inversa diu que existeixen, però no precisa.

Una interpretació del Teorema de la Funció Inversa és com a resposta afirmativa a les següents preguntes; fixem $y^*\in\mathbb{R}^n$ i considerem l'equació

$$Fx = y^* \tag{5.3}$$

de la qual en coneixem una solució x^*. És x^* una solució localment única (té un entorn en què no n'hi ha cap altra)? Si variem lleugerament y^*, l'equació encara té solució (localment única)? La solució x^* varia contínuament respecte de y^*?

Teorema 5.3 (Teorema Funció Inversa)**.** *Siguin $F\colon D\mapsto\mathbb{R}^n$, $D\subseteq\mathbb{R}^n$ obert, i $x^0\in D$. Suposem que F és $\mathscr{C}^1$ en un entorn de x^0 i $F'(x^0)$ és invertible.*

Aleshores hi ha $U\subseteq D$ entorn obert de x^0 i V entorn obert de Fx^0 tals que $F_U\colon U\mapsto V$ és bijectiva, F_U^{-1} és $\mathscr{C}^1$ i $\forall x\in U$, $(F_U^{-1})'(Fx) = F'(x)^{-1}$.

Les demostracions poden ésser més o menys complexes, però sempre tenen dues parts; una en què es prova que existeixen U, V i que F_U és bijectiva, i una altra en què es prova que la regularitat de F_U^{-1} és la mateixa que la de F_U. Una demostració basada en el Teorema del Punt Fix de Banach es pot veure en [16, teo. 5.2.1].

El teorema, però, no dona pistes de com són aquests oberts U i V. Aquí ens pot ajudar el Teorema de Kantorovich, ja que el que volem és que $\forall y^* \in V$ (5.3) tingui una única solució $x^* \in U$; però té el cost d'afegir noves hipòtesis. Enunciem el teorema amb tota generalitat: per a operadors entre espais de Banach.

Teorema 5.4. *Siguin $F\colon D \mapsto Y$, $D \subseteq X$ obert, F-derivable i $x^0 \in D$. Suposem que*

1. $\Gamma_0 = F'(x^0)^{-1} \in \mathscr{B}(Y, X)$ *amb* $\left\|\Gamma_0\right\| \leq \beta$.
2. $\exists R,\, M > 0$ *tals que* $B(x^0, R) \subseteq D$ *i* $\left\|F'(x) - F'(z)\right\| \leq M\,\|x - z\|$*, per a tot* $x, z \in B(x^0, R)$.
3. *Es compleix* $\frac{1}{\beta M} \leq R$.

Aleshores hi ha $U \subseteq D$ entorn obert de x^0 i V entorn obert de $y^0 = Fx^0$ tals que $F_U\colon U \mapsto V$ és bijectiva, F_U^{-1} és $\mathscr{C}^1$ i $\forall x \in U$, $(F_U^{-1})'(Fx) = F'(x)^{-1}$.

En concret podem escollir $V = B(y^0, 1/2\beta^2 M)$, $U = F^{-1}(V) \cap B(x^0, \frac{1}{\beta M})$; aleshores ($\|F'(x^0)\| \leq \alpha$)

$$\overline{B\left(x^0, \frac{1}{\beta M}\frac{1}{\beta\alpha + \sqrt{(\beta\alpha)^2 + 1}}\right)} \subseteq U \subseteq B(x^0, \frac{1}{\beta M})$$

Demostració. La regularitat de F_U^{-1} es prova fàcilment i no la farem.

Fixat $y^* \in B(y^0, 1/2\beta^2 M)$, volem veure que podem aplicar el Teorema de Kantorovich 3.36 a l'equació

$$Gx = Fx - y^* = \vec{0} \tag{5.4}$$

per a això comprovem que es compleixen les condicions de Kantorovich K1), K2), K3') i K4). Les condicions K1) i K3') són els punts 1 i 2 de les hipòtesis; quant a la condició K2)

$$\left\|G'(x^0)^{-1}Gx^0\right\| = \left\|F'(x^0)^{-1}(Fx^0 - y^*)\right\| \leq \beta\,\|y^* - y^0\| = \overline{\eta}$$

aquest $\overline{\eta}$ l'hem d'escollir de tal manera que es compleixi la condició K4): $h = \beta M\overline{\eta} = \beta^2 M\,\|y^* - y^0\| \leq \frac{1}{2}$. Així doncs, prenem $\overline{\eta} = 1/2\beta M$ i per a tot $y^* \in V = B(y^0, 1/2\beta^2 M)$, tenim $h < 1/2$.

Per a poder aplicar el Teorema de Kantorovich 3.36 ens falta comprovar que $r^* = \frac{1-\sqrt{1-2h}}{h}\overline{\eta} < R$, però això es té perquè $h < 1/2 \Rightarrow r^* = \frac{1-\sqrt{1-2h}}{\beta M} < \frac{1}{\beta M}$, i pel punt 3 de les hipòtesis. Per tant, aplicant el Teorema de Kantorovich, (5.4) té solució $x^* \in \overline{B(x^0, r^*)} \subsetneq B(x^0, 1/\beta M) \subseteq D$ i és única en $B(x^0, r^{**}) \cap B(x^0, R)$, $r^{**} = \frac{1+\sqrt{1-2h}}{h}\overline{\eta}$, per tant única en $B(x^0, 1/\beta M)$. Així, podem prendre $U = F^{-1}(V) \cap B(x^0, 1/\beta M)$ i $F_U\colon U \mapsto V$ és bijectiva.

Per trobar $r > 0$ tal que $B(x^0, r) \subseteq U$, fem servir la proposició 2.33 amb F sobre $B(x^0, R)$

$$\left\|Fx - Fx^0\right\| \leq \left\|Fx - Fx^0 - F'(x^0)(x - x^0)\right\| + \left\|F'(x^0)(x - x^0)\right\| \leq \frac{1}{2}M\,\|x - x^0\|^2 + \alpha\,\|x - x^0\|$$

volem que $x \in U$ i $Fx \in V$; és a dir, $r \leq \frac{1}{\beta M}$ i $\frac{1}{2}Mr^2 + \alpha r \leq \frac{1}{2\beta^2 M}$. Prenent r com la solució positiva de l'equació $\frac{1}{2}Mr^2 + \alpha r = \frac{1}{2\beta^2 M}$, tenim $r = \frac{1}{\beta M}\frac{1}{\beta\alpha + \sqrt{(\beta\alpha)^2+1}} \leq \frac{1}{\beta M}\frac{1}{1+\sqrt{1+1}} = \frac{\sqrt{2}-1}{\beta M} < \frac{1}{\beta M}$. Per tant $\overline{B(x^0, r)} \subseteq U$. $\square$

Un dels usos habituals del Teorema de la Funció Inversa és explotar la regularitat de la inversa (local) per aproximar la solució de l'equació $Fx = y$ en funció de y, $x(y) = F_U^{-1}(y)$. Assumint $Fx^* = y^*$ coneguts, tenim el desenvolupament de Taylor:

$$x(y^* + h) = x(y^*) + x'(y^*)h + \mathrm{o}\big(h\big) = x^* + F'(x^*)^{-1}h + \mathrm{o}\big(h\big) \approx x^* + F'(x^*)^{-1}h$$

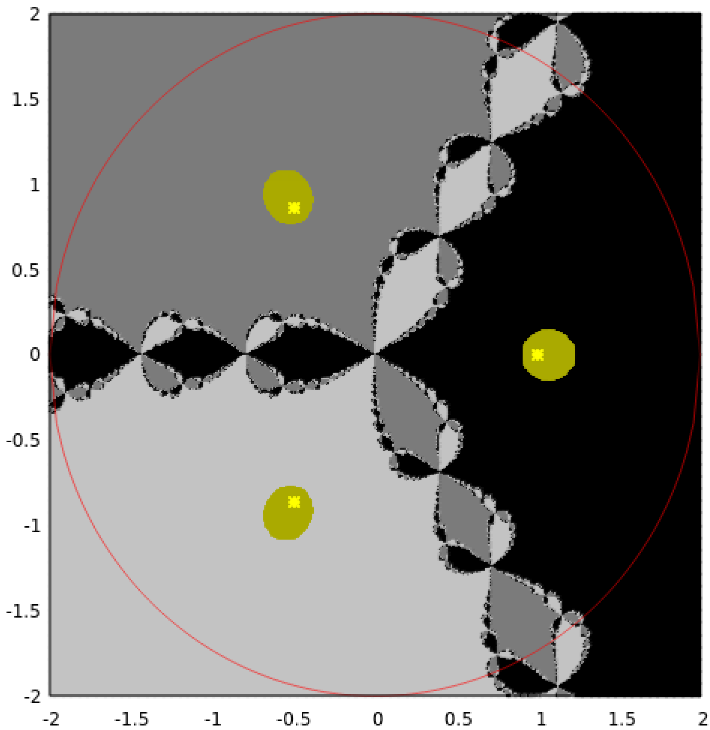
Un altre ús, habitual en geometria diferencial i càlcul integral, és garantir que una aplicació $f\colon \mathbb{R}^n \mapsto \mathbb{R}^n$ defineix un canvi de variables, almenys en un entorn d'un punt x^0 del domini.

6 Conclusions

Aquest treball de fi de grau és sobre un tema que, inicialment, no esparava ni havia pensat: el Teorema de Kantorovich; però que en parlar-ne amb el tutor, despertà el meu interés. El treball ha sigut una feina de cerca, lectura i síntesi de molta informació, més de la que surt en aquesta memòria, i que de vegades m'ha aclaparat una mica. A això cal afegir que no tenia coneixements d'anàlisi funcional i els he hagut d'adquirir en el decurs de l'assignatura del treball de fi de grau. En aquest sentit crec que m'ha ajudat a crèixer com a futur matemàtic.

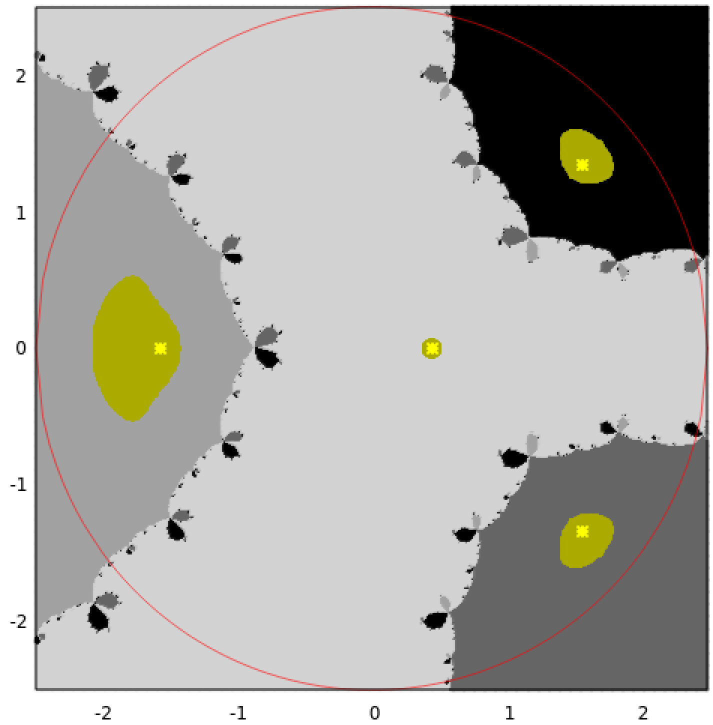
A Imatges de la secció 5

Figura A.1: Accessibilitat dels zeros $z^3 - 1 = 0$



Les conques d'atracció dels zeros 1, $\exp(\frac{2}{3}\pi i)$, $\exp(\frac{4}{3}\pi i)$ dins el rectangle $[-2,2] \times [-2,2]$ es mostren en escala de grisos. Les zones en groc verdós tènue són a on es compleixen les hipòtesis del Teorema de Kantorovich, prenent $D_0 = \{|z| < 2\}$. El contorn disc D_0 es mostra en vermell. Els zeros es mostren en groc intens.

Figura A.2: Accessibilitat dels zeros de $z^4 - 2z^3 + 7z - 3 = 0$



Les conques d'atracció dels zeros de l'equació dins el rectangle $[-2.5, 2.5] \times [-2.5, 2.5]$ es mostren en escala de grisos. Les zones en groc verdós tènue són a on es compleixen les hipòtesis del Teorema de Kantorovich, prenent $D_0 = \{|z| < 2.5\}$. El contorn disc D_0 es mostra en vermell. Els zeros es mostren en groc intens:
$z_0 \approx 1.564 - 1.339\mathrm{i}$, $z_1 = \overline{z_0}$, $z_2 \approx -1.577$, $z_3 \approx 0.449$.

B Programes en C

Programa 1: conques-atraccio-complexos.c

```
/*
  Programa per a calcular les conques d'atracció del mètode iteratiu de Newton
per a una funció complexa.
  El programa escriu els punts de les conques d'atracció a la sortida estàndar;
per tant cal redirigir la sortida per a desar-ho a un arxiu i poder fer-ne
imatges amb el Gnuplot.

      ./conques-atraccio-complexos.out > arxiu_escriure_punts.dad


  Autor: Daniel Fernández Benítez.
  Última modificació: 18/06/2021.


  L'usuari proporciona una funció complexa, tots els seus zeros i un rectangle
en el pla complex.
  Aleshores es crea una xarxa de punts en el rectangle i cada punt es pren
com condició inicial del Mètode de Newton. Si els iterats de Newton s'atansen
a algun zero de la funció, aleshores la condició inicial pertany a la conca d'atracció d
seguint el següent format

    <part_real> <part_imaginària> <índex_zero>

on «índex_zero» és l'índex del zero en el vector de zeros que l'usuari
ha de donar.


________________________________
  Les dades que l'usuari pot tenir interès a modificar (funcio, constants, etc.)
es troben en XARXA_PUNTS i MÈTODE DE NEWTON.
________________________________


  El programa s'ha escrit definint com a macros del preprocessador aquells
càlculs que s'han de repetir molts cops i que, altrament, caldria escriure
en una funció. D'aquesta manera s'evita que el programa estigui constantment
fent crides a funcions, i així accelerem l'execució del programa.



POSSIBLES MILLORES
 - Potser fer els gràfics amb Gnuplot directament
  des del programa el mòdul http://ndevilla.free.fr/gnuplot/
*/
```

```c
#include<stdio.h>
#include<stdlib.h>
#include<complex.h>
#include<math.h>
#include<stdbool.h>

/*----------------------------------------------------------------
                    MÈTODE DE NEWTON
----------------------------------------------------------------*/

/*Informació relativa a l'execució del mètode de Newton:
condicions d'aturada, la funció i els seus zeros.*/


/*Nombre màxim d'iteracions*/
#define MAX_ITER 25

/*tolerància, al quadrat, en la proximitat a un zero de f(z) (TOL = 1E-3)*/
#define TOL2_CONCA 1E-6
/*tolerància, al quadrat, per considerar un nombre com 0 (TOL = 1E-12)*/
#define TOL2_NULLITAT 1E-24


/*funció a què aplicar el mètode de Newton*/
#define funcio(z) (z * z * z - 1.)
//#define funcio(z) (-3 + z * (7 + z * z * (-2 + z)))

/*derivada de la funció a què aplicar el mètode de Newton*/
#define funcio_1(z) (3 * z * z)
//#define funcio_1(z) (7 + z * z * (-6 + 4 * z))

/*quantitat de zeros que té funcio(z)*/
#define Q_ZEROS 3
//#define Q_ZEROS 4


/*Zeros de la funció definida dalt*/
//arrel quadrada de 3
#define AQ_3 1.73205080756887729352
double complex const g_zeros[Q_ZEROS] = {
    1.,
    -1. / 2. + AQ_3 / 2. * I,
    -1. / 2. - AQ_3 / 2. * I
};

/*Zeros de la funcio definida dalt*/
```

```
/*
double complex const g_zeros[Q_ZEROS] = {
  1.564432070603465 + 1.338737397039180 * I,
  1.564432070603465 - 1.338737397039180 * I,
  -1.577440734126577,
  .4485765929196463
};
*/
/*------------------------------------------------------------
                    FI MÈTODE DE NEWTON
------------------------------------------------------------*/


/*------------------------------------------------------------
                        XARXA PUNTS
------------------------------------------------------------*/

/*Domini rectangular en què crear la xarxa de punts*/
#define X_MIN -2.
#define Y_MAX 2.
#define X_MAX 2.
#define Y_MIN -2.

/*quantitat de punts en què dividir cada costat del rectangle*/
#define Q_PUNTS_XARXA_X 512
#define Q_PUNTS_XARXA_Y Q_PUNTS_XARXA_X

/*distància (horitzontal i vertical) entre punts de la xarxa*/
#define XARXA_PAS_X  ((X_MAX - X_MIN) / (Q_PUNTS_XARXA_X - 1))
#define XARXA_PAS_Y  ((Y_MAX - Y_MIN) / (Q_PUNTS_XARXA_Y - 1))
/*------------------------------------------------------------
                      FI XARXA PUNTS
------------------------------------------------------------*/


/*mòdul al quadrat d'un complex (cabs(z) ^ 2)*/
#define cmodul2(z) (creal(z) * creal(z) + cimag(z) * cimag(z))



char * g_nom_prog;


void conques_atraccio();




int main(int argc, char *argv[]){
```

```
  unsigned int i;


  /*processar arguments del programa*/
  g_nom_prog = argv[0];


  /*calcular i escriure les conques d'atracció*/
  conques_atraccio();


  /*escrivim també els zeros de la funció*/
  for(i = 0; i < Q_ZEROS; i++){
    printf("%13.7le %13.7le %u\n", creal(g_zeros[i]), cimag(g_zeros[i]), Q_ZEROS);
  }


  return EXIT_SUCCESS;
}


/*
DESC: Calcula a quina conca d'atracció pertany cada punt de la xarxa,
i ho escriu a la sortida estàndar.
  Amb cada punt de la xarxa, s'engega el mètode Newton i es comprova si s'atansa
a algun dels zeros de la funció. En cas afirmatiu escrivim a la sortida estàndar
el punt i l'índex del zero (índex en el vector «g_zeros»).

PARAMS:

RETORN:
*/
void conques_atraccio(){
  unsigned int i, j;
  double complex z, z_0;
  unsigned int conca;
  const double dx = XARXA_PAS_X, dy = XARXA_PAS_Y;


  /*Engeguem el mètode de Newton per a cada punt de la xarxa
  (d'esquerra a dreta i de dalt a baix)
  i mirem cap a quin zero s'atansa*/
  for(i = 0; i < Q_PUNTS_XARXA_Y; i++){
    for(j = 0; j < Q_PUNTS_XARXA_X; j++){
      unsigned int q_iter;
      bool es_conca_trobada;
      double complex f_1;
```

```
    /*mètode de Newton amb condició inicial*/
    z_0 = z = (X_MIN + j * dx) + (Y_MAX - i * dy) * I;
    es_conca_trobada = false;
    q_iter = 0;
    while(!es_conca_trobada && q_iter < MAX_ITER){
      /*Saltem el punt si la derivada en «z» s'anul·la*/
      f_1 = funcio_1(z);
      if(cmodul2(f_1) < TOL2_NULLITAT){
        break;
      }

      /*iteració Newton i comprovar si és prop d'un zero*/
      z = z - funcio(z) / f_1;

      for(conca = 0; !es_conca_trobada && conca < Q_ZEROS; conca++){
        if(cmodul2(g_zeros[conca] - z) < TOL2_CONCA){
          es_conca_trobada = true;
        }
      }
      if(es_conca_trobada){
        conca--;
      }


      q_iter++;
    }


    /*escrivim la dada inicial i la conca a què pertany*/
    if(es_conca_trobada){
      printf("%13.7le %13.7le %u\n", creal(z_0), cimag(z_0), conca);
    }
  }
 }
}
```

Programa 2: conques-Kantorovich-complexos.c

```
/*
  Programa per a calcular els punts en què se satisfan
les hipòtesis del Teorema de Kantorovich per a una funció complexa donada.
  El programa escriu els punts a la sortida estàndar; per tant cal redirigir
la sortida per a desar-ho a un arxiu i poder fer-ne imatges amb el Gnuplot.

    ./conques-atraccio-complexos.out > arxiu_escriure_punts.dad


  Autor: Daniel Fernández Benítez.
```

```
Última modificació: 18/06/2021.


L'usuari proporciona una funció complexa i una bola oberta en el pla complex
a on la segona derivada de la funció està afitada.
Aleshores es crea una xarxa de punts en un rectangle que conté la bola,
i en cada punt de la xarxa que sigui dins la bola es comproven les hipòtesis
del Teorema de Kantorovich. Tots els punts en què se satisfan les hipòtesis
de Kantorovich s'escriuen a la sortida estàndar.


---------------------------------------
Les dades que l'usuari pot tenir interès a modificar (funcio, constants, etc.)
es troben en XARXA_PUNTS i TEOREMA KANTOROVICH.
---------------------------------------


El programa s'ha escrit definint com a macros del preprocessador aquells
càlculs que s'han de repetir molts cops i que, altrament, caldria escriure
en una funció. D'aquesta manera s'evita que el programa estigui constantment
fent crides a funcions, i així accelerem l'execució del programa.



POSSIBLES MILLORES
 - Potser fer els gràfics amb Gnuplot directament des del programa
  amb el mòdul http://ndevilla.free.fr/gnuplot/
*/



#include<stdio.h>
#include<stdlib.h>
#include<complex.h>
#include<math.h>
#include<stdbool.h>


/*-----------------------------------------------------------------------
                          TEOREMA KANTOROVICH
-----------------------------------------------------------------------*/
/*tolerància per considerar un nombre com 0*/
#define TOL_NULLITAT 1E-12


/*funció a què aplicar el mètode de Newton*/
#define funcio(z) (z * z * z - 1.)
//#define funcio(z) (-3 + z * (7 + z * z * (-2 + z)))
```

```
/*derivada de la funció a què aplicar el mètode de Newton*/
#define funcio_1(z) (3 * z * z)
//#define funcio_1(z) (7 + z * z * (-6 + 4 * z))


/*Bola que conté els zeros de la funció definida dalt (f(z))*/
#define BOLA_CENTRE (0. + 0. * I)
#define BOLA_RADI 2.
//#define BOLA_RADI 2.5


/*M = fita de |f''(z)| en la bola*/
#define KANTOR_M 12.
//#define KANTOR_M 45.

/*Condició Kantorovich h = (beta)*M*(eta) <= 1/2*/
#define MAX_H_KANTOR 1. / 2
/*---------------------------------------------------------------------------
                         FI TEOREMA KANTOROVICH
---------------------------------------------------------------------------*/


/*---------------------------------------------------------------------------
                              XARXA PUNTS
---------------------------------------------------------------------------*/
/*Domini rectangular en què crear la xarxa de punts.
Ha de contenir la bola definida dalt*/
#define X_MIN -2.
#define Y_MAX 2.
#define X_MAX 2.
#define Y_MIN -2.

/*quantitat de punts en què dividir cada costat del rectangle*/
#define Q_PUNTS_XARXA_X 512
#define Q_PUNTS_XARXA_Y Q_PUNTS_XARXA_X

/*distància (horitzontal i vertical) entre punts de la xarxa*/
#define XARXA_PAS_X  ((X_MAX - X_MIN) / (Q_PUNTS_XARXA_X - 1))
#define XARXA_PAS_Y  ((Y_MAX - Y_MIN) / (Q_PUNTS_XARXA_Y - 1))
/*---------------------------------------------------------------------------
                            FI XARXA PUNTS
---------------------------------------------------------------------------*/


/*mòdul al quadrat d'un complex (cabs(z) ^ 2)*/
#define cmodul2(z) (creal(z) * creal(z) + cimag(z) * cimag(z))

/*Comprova que un complex pertany a la bola (tancada) definida dalt*/
#define pertany_bola(z) \
```

```
  (cmodul2(z - BOLA_CENTRE) + TOL_NULLITAT <= BOLA_RADI * BOLA_RADI)

/*comprova que una bola (tancada) B(z, r)
està continguda en la bola (oberta) definida dalt*/
#define es_bola_dins(z, r) \
  (cmodul2(z - BOLA_CENTRE) + TOL_NULLITAT < (BOLA_RADI - r) * (BOLA_RADI - r))



char * g_nom_prog;


void conques_Kantorovich();




int main(int argc, char *argv[]){
  /*processar arguments del programa*/
  g_nom_prog = argv[0];


  /*calcular i escriure els punts en què es compleix la condició de Kantorovich*/
  conques_Kantorovich();


  return EXIT_SUCCESS;
}


/*
DESC: Comprova en quins punts dins la bola es compleixen les hipòtesis
del Teorema de Kantorovich i els escriu a la sortida estàndar.
  Per a cada punt de la xarxa es comprova el següent
    - si el punt és dins la bola.
    - si el valor h = (beta) * M * (eta) és menor que 1/2.
    - si la regió d'existència, que dona el Teorema per al punt, és dins la bola.

si compleix tot això, compleix les hipòtesis del Teorema de Kantorovich
i escrivim el punt a la sortida estàndar.

PARAMS:

RETORN:
*/
void conques_Kantorovich(){
  unsigned int i, j;
  double complex z;
  double beta, eta, h, r_exist;
```

```
  const double dx = XARXA_PAS_X, dy = XARXA_PAS_Y;


  /*comprovem les hipòtesis del Tª Kantorovich
  en els punts de la xarxa que pertanyen a la bola*/
  for(i = 0; i < Q_PUNTS_XARXA_Y; i++){
    for(j = 0; j < Q_PUNTS_XARXA_X; j++){
      z = (X_MIN + j * dx) + (Y_MAX - i * dy) * I;

      /*comprovem si pertany a la bola*/
      if(! pertany_bola(z)){
        continue;
      }



      /*càlcul de h = (beta) * M * (eta)   */
      /*beta = |1 / f'(z)| */
      beta = cabs(funcio_1(z));
      if(beta < TOL_NULLITAT){
        continue;
      }
      beta = 1. / beta;

      /*eta = |f(z) / f'(z)|*/
      eta = cabs(funcio(z)) * beta;
      h = beta * KANTOR_M * eta;


      /*si es compleixen les hipòtesis del Tª Kantorovich, escrivim el punt*/
      /*r = (eta) * (1 - sqrt(1 - 2h)) / h) =
          = 2h / (beta * M * (1 + sqrt(1 - 2h)))*/
      r_exist = 2. * h / (beta * KANTOR_M * (1 + sqrt(1 - 2 * h)));
      if(h + TOL_NULLITAT <= MAX_H_KANTOR && es_bola_dins(z, r_exist)){
        printf("%13.7le %13.7le\n", creal(z), cimag(z));
      }
    }
  }
}
```

Programa 3: Newton-Raphson-complex.c

```
/*
  Mètode de Newton–Raphson en el pla complex.

  Autor: Daniel Fernández Benítez.
  Última modificació: 19/06/2021.


  L'usuari proporciona una funció complexa a què aplicar el mètode de Newton–
```

```
Raphson, junt amb una condició inicial i un valor per al criteri d'aturada
del mètode iteratiu.
  Com a criteri d'aturada del mètode iteratiu escollim que la distància
entre dos iterats consecutius sigui inferior a un valor prefixat proporcionat
per l'usuari. També s'imposa un nombre màxim d'iteracions.


-----------------------------------
  Les dades que l'usuari pot tenir interès a modificar (funcio, condició inicial,
constants, etc.) es troben en MÈTODE DE NEWTON.
-----------------------------------


POSSIBLES MILLORES:
  - Rebre la condició inicial i la distància entre iterats com arguments
    del programa.
  - Afegir el mètode de Newton amortiguat i permetre que l'usuari en triï
    quin fer servir amb un argument del programa.
*/



#include<stdio.h>
#include<stdlib.h>
#include<complex.h>
#include<math.h>
#include<stdbool.h>


/*---------------------------------------------------------------------
                          MÈTODE DE NEWTON
---------------------------------------------------------------------*/

/*Informació relativa a l'execució del mètode de Newton:
condició inicial, condicions d'aturada, la funció.*/


/*Nombre màxim d'iteracions*/
#define MAX_ITER 15
/*criteri d'aturada: distància entre iterats consecutius*/
#define DIST_ITERS_CONSEC 1E-10

/*tolerància, al quadrat, per considerar un nombre com 0 (TOL = 1E-12)*/
#define TOL2_NULLITAT 1E-24


/*condició inicial per al mètode de Newton*/
#define NEWTON_COND_INI (-1. + 1. * I)
```

```
/*funció a què aplicar el mètode de Newton*/
#define funcio(z) (z * z * z - 1.)
//#define funcio(z) (-3 + z * (7 + z * z * (-2 + z)))

/*derivada de la funció a què aplicar el mètode de Newton*/
#define funcio_1(z) (3. * z * z)
//#define funcio_1(z) (7 + z * z * (-6 + 4 * z))
/*--------------------------------------------------------------------
                    FI MÈTODE DE NEWTON
--------------------------------------------------------------------*/


/*mòdul al quadrat d'un complex (cabs(z) ^ 2)*/
#define cmodul2(z) (creal(z) * creal(z) + cimag(z) * cimag(z))



char * g_nom_prog;


int newton_raphson(double complex *, const double);




int main(int argc, char *argv[]){
  int q_iter;
  double complex z_0, z_sol;


  /*processar arguments del programa*/
  g_nom_prog = argv[0];


  /*iniciar el mètode iteratiu de Newton*/
  z_sol = z_0 = NEWTON_COND_INI;
  q_iter = newton_raphson(&z_sol, DIST_ITERS_CONSEC);


  /*informem sobre el resultat*/
  if(z_sol == z_0){
    printf("Hi ha hagut algun problema amb el procés d'iteració," \
        " en la iteració número %d\n", q_iter);
  }else{
    printf("iteracions: %d\n" \
        "z_0: %16.9le %+16.9le i\n" \
        "z_sol: %16.9le %+16.9le i\n", q_iter, creal(z_0), cimag(z_0),
        creal(z_sol), cimag(z_sol));
  }
```

```
  return EXIT_SUCCESS;
}


/*
DESC: Mètode de Newton en el pla complex.
  Executa el mètode iteratiu de Newton amb la condició inicial donada, El mètode
s'atura amb èxit si s'ha satisfet el criteri d'aturada: la distància
entre dos iterats consecutius és inferior a «dist_iters_atur»

              |x_{k+1} - x_k| < «dist_iters_atur».

  En tal cas es retorna l'últim iterat calculat, desant-lo al paràmetre «z».
En qualsevol altre cas d'aturada, «z» no es modifica: la derivada s'anula
en un iterat o s'arriba al màxim d'iteracions.

PARAMS:
  - *z, apuntador a la condició inicial del mètode de Newton. Es desarà
    l'últim iterat calculat en cas d'aturar-se amb èxit.
  - dist_iters_atur, la distància entre dos iterats consecutius per a aturar
    el

RETORN: El nombre d'iteracions fetes fins que el mètode s'ha aturat.
*/
int newton_raphson(double complex *z, const double dist_iters_atur){
  const double dist2_iters_atur = dist_iters_atur * dist_iters_atur;
  int q_iter;
  double complex z0, z1, f_1;
  bool es_divisio_zero;


  if(z == NULL){
    return 0;
  }


  /*mètode de Newton-Raphson*/
  q_iter = 0;
  z0 = z1 = *z;
  es_divisio_zero = false;
  do{
    z0 = z1;

    /*abans d'iterar, comprovem si la derivada s'hi anul·la*/
    f_1 = funcio_1(z0);
    if(cmodul2(f_1) < TOL2_NULLITAT){
```

```
      es_divisio_zero = true;
      continue;
    }

    /*iteració Newton*/
    z1 = z0 - funcio(z0) / f_1;


    q_iter++;
  }while(! es_divisio_zero && q_iter < MAX_ITER &&
      cmodul2(z1 - z0) > dist2_iters_atur);

  /*desem la solució a «z»*/
  if(! es_divisio_zero && cmodul2(z1 - z0) <= dist2_iters_atur){
    *z = z1;
  }


  return q_iter;
}
```

Imatges amb Gnuplot: dibuixar.gpl

```
# Ordres Gnuplot per a crear imatges amb la sortida dels programes en C


#Fixem un format de sortida que permeti modificar la mida dels punts (linewidth)
#per exemple «qt». També serveixen «postscript» o «pdfcairo».
set term qt size 500, 500

#Descomentar per a sortida en postscript
#set term postscript
#set output 'nom-arxiu-sortida-postscript'

#Descomentar per a sortida en pdf
#set term pdfcairo
#set output 'nom-arxiu-sortida-pdf'



#Treiem la caixa que mostra el gradient de la paleta
unset colorbox

#No mostrar els noms dels arxius per defecte
set key noautotitle

#Paleta de colors en escala de grisos o color
set palette grey
#set palette color
```

```
#Fixar les dimensions del llenç, segons convingui
#set xrange[-2:2]
#set yrange[-2:2]


#----------------
#Crear imatge de les conques d'atracció.
#Assumim que els punts són en un arxiu <conques-atraccio.dad>
#----------------

#La tercera columna serveix per a donar color a cada punt: es pren l'índex
#del color en la paleta.
plot 'conques-atraccio.dad' u 1:2:3 w dots lw 2.5 lc palette z


#----------------
#Crear imatge de l'accessibilitat del Teorema de Kantorovich (conques Kantorovich)
#Assumim que els punts són en un arxiu <conques-Kantorovich.dad>
#----------------

#Dibuixar-les en groc
plot 'conques-Kantorovich.dad' u 1:2 w dots lw 2 lc "#aaaa00"




#----------------
#Crear una imatge conjunta amb les conques d'atracció les de Kantorovich
#i els zeros de la funció.
#Assumim que els zeros de la funció de què hem calculat les conques
#són en un arxiu <zeros.dad>
#----------------

#Les conques de Kantorovich en groc verdós, i els zeros en groc fort.
plot 'conques-atraccio.dad' u 1:2:3 w dots lw 2.5 lc palette z, \
  'conques-Kantorovich.dad' u 1:2 w dots lw 2 lc "#aaaa00", \
  'zeros.dad' u 1:2 w points lw 2 lc "#ffff00"
```

Referències

[1] B. Rall, Louis. *Computational solution of nonlinear operator equacions*. New York, Robert E. Krieger Publising Co, 1969. Reimpressió New York: John Wiley & Sons, 1979. ISBN 0-88275-667-2.

[2] B. Rall, Louis. A note on the convergence on Newton's method. *SIAM J. Numer. Anal.* 1974. Vol. 11, núm. 1, p. 34-36.

[3] Ezquerro Fernández, J. A.; Hernández Verón, M. A. *Newton's method: an updated approach of Kantorovich's theory*. Birkhäuser, 2017. ISBN 978-3-319-55975-9.

[4] Ezquerro Fernández, J. A.; Hernández Verón, M. A. *Mild differentiability conditions for Newton's method in Banach spaces*. Birkhäuser, 2020. ISBN 978-3-030-48702-7.

[5] Gragg, W. B.; Tapia, R. A. Error bounds for the Newton-Kantorovich theorem. *SIAM J. Num. Anal.* 1974. Vol. 11, núm. 1, p. 10-13.

[6] Kantorovich, L. V. On Newton's method for functional equacions (Rus). *Dokl. Akad. Nauk. SSSR*. 1948. Vol. 59, núm. 7, p. 1237-1240.

[7] Kantorovich, L. V. Functional analysis and applied mathematics (Rus). *Uspekhi Mat. Nauk*. 1948. Vol. 3, núm. 6, p. 89-185.

[8] Kantorovich, L. V. On the Newton method (Rus). *Trudy Math. Inst. Steklov.* 1949. Vol. 28, p. 104-144.

[9] Kantorovich, L. V. The majorant principle and Newton's method (Rus). *Dokl. Akad. Nauk. SSSR*. 1951. Vol. 76, p. 17-20.

[10] Kantorovich, L. V. Some further applications of the principle of majorants (Rus). *Dokl. Akad. Nauk. SSSR*. 1951. Vol. 80, p. 849-852.

[11] Kantorovich, L. V. *Selected works part II*. Editat per S. S. Kutateladze, J. V. Romanovsky. Traducció del rus a l'anglès per A. B. Sossinkii. Amsterdam: Gordon and Breach Science Publishers SA. 1996.

[12] Kantorovich, L. V.; Akilov, G. P. *Functional analysis*. Segona edició. Oxford: Pergamon Press, 1982.

[13] Llorens Fuster, Enrique. Apunts de l'assignatura optativa d'anàlisi funcional de la Universitat de València [en línia]. 2016, disponible a `https://www.uv.es/~llorens/` [consulta: 16 de juny de 2021].

[14] Mysovskikh, I. P. On the convergence of L. V. Kantorovich's method for solution of functional equacions and its applications (Rus). *Dokl. Akad. Nauk. SSSR*. 1950. Vol. 70, pp. 565-568.

[15] Ortega, J. M. The Newton-Kantorovich theorem. *The Amer. Math. Monthly*. 1968. Vol. 75, núm. 6, p. 658-660.

[16] Ortega, J. M.; Rheinboldt, W. C. *Iterative solutions of nonlinear equacions in several variables*. Orlando: Academic Press Inc, 1970.

[17] Payá Albert, Rafael. Apunts de l'assignatura d'anàlisi funcional de la Universidad de Granada [en línia]. 2020, disponible a `https://www.ugr.es/~rpaya/docencia.htm` [consulta: 16 de juny de 2021].

[18] Rheinboldt, W. C. A unified convergence theory for a class of iterative processes. *SIAM J. Num. Anal.* 1968. Vol. 5, núm. 1, p. 42-63.

[19] Tapia, R. A. The Kantorovich theorem for Newton's method. The *Amer. Math. Monthly*. 1971. Vol. 78, núm. 4, p. 389-392.

[20] Quarteroni, A.; Saleri, F. *Cálculo científico con MATLAB i Octave*. Milan: Springer-Verlag Itàlia, 2006. Traducció d'Alfredo Bermúdez.

[21] Wikipedia. *Teorema de Rouché*. Disponible a `https://es.wikipedia.org/wiki/Teorema_de_Rouch%C3%A9` [consulta: 19 de juny de 2021].

[22] Institut d'Estudis Catalans. *Diccionari de la Llengua Catalana* [en línia]. Disponible a `https://dlc.iec.cat/` [consulta: 20 de juny de 2021].

[23] Universitat Politècnica de Catalunya; Enciclopèdia Catalana. *Diccionari de matemàtiques i estadística* [en línia]. 2a ed. Barcelona: Institut d'Estudis Catalans, 2021. Disponible a `https://cit.iec.cat/obresx.asp?obra=DME` [consulta: 20 de juny de 2021].

[24] Free Software Foundation. Documentació de la implementació de GNU de la biblioteca estàndar de C (GNU C Library) [en línia]. Versió de la biblioteca 2.32. Disponible a `https://www.gnu.org/software/libc/documentation.html` [consulta: 18 de juny de 2021].

[25] Joyanes Aguilar, Luis; Zahonero Martínez, Ignacio. *Programación en C: metodologia, algoritmos y estructura de datos*. Madrid: McGraw-hill, 2001.