UNIVERSIDAD AUTÓNOMA DE MADRID

Departamento de Matemáticas

MORSE-SARD TYPE THEOREMS IN $\mathbb{R}^n$ AND IN BANACH SPACES

Teoremas de tipo Morse-Sard en $\mathbb{R}^n$ y en espacios de Banach

Memoria para optar al grado de
Doctor en Matemáticas
presentada por

Miguel García Bravo

Bajo la dirección de

Daniel Azagra Rueda

Madrid, Septiembre de 2020

Agradecimientos

Me gustaría mencionar en pocas palabras las personas que con su ayuda han hecho posible el desarrollo del presente trabajo.

En primer lugar y más importante querría agradecer a mi director de tesis Daniel Azagra toda la ayuda que me ha dado durante estos últimos cuatro años. Es indudable que para mí Daniel ha sido un excepcional modelo a seguir en estos comienzos de mi carrera investigadora. Quiero destacar entre otras cosas sus grandes capacidades como matemático, su plena disponibilidad a atenderme cuando lo necesitaba y sus buenos consejos a la hora de resolver los problemas que se me iban planteando.

Por otro lado he tenido la suerte de poder trabajar con otras excelentes personas. Entre ellos, dar las gracias a los profesores Juan Ferrera y Javier Gómez Gil por unos meses de agradable colaboración sobre el tema de las funciones subdiferenciales y la propiedad de Lusin. Agradecer a Tadeusz Dobrowolski por su ayuda a la ahora de resolver uno de los problemas principales de esta tesis y que más tiempo nos ha llevado. Y por último quisiera agradecer también al profesor Robert Deville por recibirme durante tres fantásticos meses en Burdeos. Durante este tiempo Robert supo contagiarme su gran pasión por las matemáticas, y pude aprender incontables teoremas y problemas interesantes. Agradecer de igual forma su gran hospitalidad durante toda mi estancia allí.

En la parte burocrática dar las gracias a Pepe Moreno, mi tutor de tesis, por su tiempo y disponibilidad, así como a la gente del ICMAT, UAM y UCM que resolvieron aquellas dudas que trascendían las matemáticas. También agradezco al Programa Internacional de Becas de Doctorado La Caixa-Severo Ochoa el haberme financiado durante este tiempo.

Es indudable que estos cuatro años de trabajo hubieron sido más aburridos si no hubiese tenido la oportunidad de pasar los ratos en los que no hacía investigación junto a mis amigos. A todos ellos, dentro y fuera de la Universidad, por los grandes momentos pasados, gracias.

Finalmente me gustaría acabar dedicando esta tesis a toda mi familia.

La elaboración de esta tesis doctoral se ha producido en el Instituto de Ciencias Matemáticas de Madrid (CSIC-UAM-UC3M-UCM) y ha sido financiada por el programa de becas predoctorales La Caixa-Severo Ochoa 2016.

Index

Resumen

En 1934 Hassler Whitney prueba su famoso teorema de extensión, que da condiciones necesarias y suficientes para que una función continua definida en un conjunto cerrado C arbitrario de $\mathbb{R}^n$, junto con un conjunto de formas j-lineales $j = 1, \dots, k$, se pueda extender a una función C^k definida en todo $\mathbb{R}^n$ cuyas derivadas en C hasta el orden k son dichas formas j-lineales prefijadas (ver [151]). Con la ayuda de este resultado, en 1935 construyó una función $f : \mathbb{R}^2 \to \mathbb{R}$ de clase C^1 tal que su conjunto de valores críticos tenía medida positiva (véase [152]). Las condiciones que permiten que tal ejemplo ocurra, en el que las imágenes de conjuntos críticos puedan tener medida positiva, eran malamente entendidas en aquella época. En los próximos años se obtuvo una visión más completa del problema.

Teorema de Morse-Sard: *Si $f : \mathbb{R}^n \to \mathbb{R}^m$ es una función de clase C^k, donde $k \geq \max\{1, n-m+1\}$, entonces su conjunto de puntos críticos $C_f := \{x \in \mathbb{R}^n : \, rango \, Df(x) \text{ no es máximo}\}$ satisface que $\mathcal{L}^m(f(C_f)) = 0$, donde $\mathcal{L}^m$ denota la medida de Lebesgue en $\mathbb{R}^m$.*

Al conjunto $f(C_f)$ se le conoce por el nombre de conjunto de valores críticos. Diremos que una función diferenciable $f : \mathbb{R}^n \to \mathbb{R}^m$ satisface la propiedad de Morse-Sard si $\mathcal{L}^m(f(C_f)) = 0$. Se trata de un resultado con numerosas aplicaciones en diversas ramas de las matemáticas como topología diferencial, sistemas dinámicos, ecuaciones en derivadas parciales o teoría del grado. Este teorema tan importante es la motivación del presente trabajo.

La estructura de los capítulos será la siguiente. En el Capítulo 1 se ofrecerá una versión del teorema de Morse-Sard clásico donde se trabaja con la diferenciabilidad aproximada. La propiedad de Morse-Sard tiene relación con la propiedad de Lusin, y siguiendo esta línea en el Capítulo 2 se demuestra que las funciones subdiferenciales tienen la propiedad de Lusin de clase C^1 y C^2. Los Capítulos 4 y 5 de la tesis derivan del intento de plantearse versiones del teorema de Morse-Sard en el contexto de espacios de Banach de dimensión infinita, para los cuales una herramienta clave es la extracción difeomorfa de conjuntos cerrados en espacios de Banach, un tema tratado en el Capítulo 3. Se será más preciso a lo largo del resumen.

El teorema de Morse-Sard clásico fue probado por primera vez por Anthony P. Morse para el caso de funciones con valores reales en el 1939 y más tarde en 1942 fue probado por Arthur Sard para el de caso de valores vectoriales (véanse [125] y [140] respectivamente). Gracias al famoso contraejemplo de Whitney de 1934 se sabe que este resultado es óptimo en la escala de espacios C^j. La construcción del contraejemplo de Whitney se puede generalizar para asegurar la existencia de funciones $f : \mathbb{R}^n \to \mathbb{R}^m$, con $m < n$ y de clase C^{n-m} tales que $\mathcal{L}^m(f(C_f)) > 0$. También destacamos el artículo de Kaufman [108] que da un método para construir funciones $f : [0,1]^{n+1} \to [0,1]^n$ sobreyectivas y de clase C^1 tales que $rango(Df(x)) = 1$ para todo $x \in [0,1]^{n+1}$.

Sin embargo, ¿es posible que para ciertos espacios de funciones intermedios entre C^{k-1} y C^k (con $k = n - m + 1$) que mantengan suficientes buenas propiedades de regularidad, la propiedad de Morse-Sard se siga satisfaciendo? Una clase de espacios intermedios natural entre C^{k-1} y C^k son los espacios Hölder $C^{k-1,t}$, con $t \in (0,1]$, que se definen como el conjunto de aquellas funciones f de clase C^{k-1} tales que existe $M > 0$ de modo que $||D^{k-1}f(x) - D^{k-1}f(y)|| \leq M|x-y|^t$, para todo $x, y \in \mathbb{R}^n$. En 1986 Norton estudió las propiedades de Morse-Sard que podían tener este tipo de funciones y en particular probó que si $n > m$, $f : \mathbb{R}^n \to \mathbb{R}^m$ es de clase $C^{n-m,t}$ y C_f tiene medida Hausdorff $\mathcal{H}^{n+t-1}$−nula, entonces $\mathcal{L}^m(f(C_f)) = 0$ (ver [130, Theorem 1 y 2]). Unos años más tarde, Bates en

1993 demostró en [31] que las funciones $f : \mathbb{R}^n \to \mathbb{R}^m$ pertenecientes a la clase $C^{k-1,1}$ con $k \geq \max\{1, n-m+1\}$ satisfacen la propiedad de Morse-Sard, aunque en general esto falla para $C^{k-1,t}$ con $t < 1$ (ver [30] o [126]).

Por otro lado De Pascale daría en 2001, [58], la primera generalización del teorema de Morse-Sard para espacios de Sobolev, en la que entran en juego propiedades de integrabilidad de las funciones (como por ejemplo la desigualdad de Morrey). Probó que si $f \in W^{k,p}_{loc}(\mathbb{R}^n;\mathbb{R}^m)$ con $n > m$, $k = n-m+1$ y $p > n$ entonces $\mathcal{L}^m(f(C_f)) = 0$. Téngase en cuenta que por el teorema de inclusión de Sobolev, $W^{k,p}_{loc} \hookrightarrow C^{k-1,1-\frac{n}{p}}$ (identificando cada función f con su representante preciso[1] f^*), y como $k \geq 2$ si $n > m$, el conjunto crítico C_f puede definirse de la manera usual ya que f^* es al menos de clase C^1. El teorema de Morse-Sard de De Pascale fue revisado por Figalli en 2008, el cual da una prueba diferente, algo más sencilla y con la ventaja de ser independiente del teorema de Morse-Sard clásico (ver [82]).

Mucho más recientemente, Bourgain, Korobkov y Kristensen ([44], 2015) dan una versión del teorema para los espacios $W^{n,1}(\mathbb{R}^n;\mathbb{R})$ y $BV_n((\mathbb{R}^n;\mathbb{R})$ utilizando técnicas totalmente diferentes y más sofisticadas (ver también [43] para el caso $n = 2$). Este último espacio es el de las funciones cuyas derivadas distribucionales de orden n son medidas de Radon en $\mathbb{R}^n$. Gracias a [65] se sabe que tales funciones admiten un representante continuo que es diferenciable en $\mathcal{H}^1-$casi todo punto. Por otro lado Bourgain, Korobkov y Kristensen también establecen una propiedad N de Lusin para estas funciones afirmando que $\mathcal{L}^1(f(E)) = 0$ siempre que $\mathcal{H}^1(E) = 0$. Luego el conjunto de puntos donde f no es diferenciable y por tanto C_f no puede definirse es mandado a un conjunto de medida nula. Y siguiendo las líneas de este trabajo, las últimas generalizaciones de las que se dispone, y muy probablemente las más finas que se pueden conseguir dentro del mundo de los espacios de Sobolev, se deben a Korobkov, Kristensen y Hajlasz. Estos prueban en [95] el teorema para el caso de funciones en el espacio de Sobolev-Lorentz $W_1^{k,\frac{n}{k}}(\mathbb{R}^n;\mathbb{R}^m)$, $k = n-m+1$, $m \leq n$, formado por funciones del espacio de Sobolev $W^{k,\frac{n}{k}}$ y cuyas derivadas distribucionales de orden k además están en el espacio de Lorentz $L_{\frac{n}{k},1}$ (véanse los artículos [113, 114, 95] para más detalles). Obsérvese que el caso $m = 1$, $W_1^{n,1}(\mathbb{R}^n;\mathbb{R}) = W^{n,1}(\mathbb{R}^n;\mathbb{R})$ se corresponde con el ya tratado por Bourgain, Korobkov y Kristensen en [44]. Estos espacios poseen las mínimas condiciones de integrabilidad que garantizan la continuidad de las aplicaciones (asumiendo siempre que trabajamos con el representante preciso). Sin embargo estas aplicaciones no tienen por qué ser diferenciables en todo punto, y esto puede crear problemas a la hora de definir el conjunto de puntos críticos. Por suerte se sabe que las funciones $f \in W_1^{k,\frac{n}{k}}(\mathbb{R}^n;\mathbb{R}^m)$ son diferenciables en $\mathcal{H}^{\frac{n}{k}}-$ casi todo punto (ver [114, Theorem 2.2]) y que, gracias a una propiedad N de Lusin, la imagen de este conjunto de *puntos malos*, llámese A_f, tiene medida cero. Esta propiedad N de Lusin establece que si $f \in W_1^{k,n/k}(\mathbb{R}^n;\mathbb{R}^m)$, $1 \leq k \leq n$ y $E \subset \mathbb{R}^n$ es tal que $\mathcal{H}^s(E) = 0$ entonces $\mathcal{H}^s(f(E)) = 0$, para $s \in [\frac{n}{k}, n]$ (ver [95, Theorem 2.3]). El conjunto de puntos críticos se define entonces como $C_f := \{x \in \mathbb{R}^n \setminus A_f \ : \ rango\, Df(x) \text{ no es máximo}\}$. Este interés por saber que el conjunto de puntos malos donde no podemos definir la diferencial es enviado a un conjunto de medida nula se debe en parte a que una de las consecuencias más directas del teorema de Morse-Sard es obtener la C^1-regularidad de casi todos los conjuntos de nivel.

En el campo de la mecánica de fluidos, un ejemplo de las potentes aplicaciones de este tipo de refinamientos del teorema de Morse-Sard para espacios de Sobolev puede verse en el artículo de Korobkov, Pileckas y Russo del 2015, [112], que da una solución al llamado problema de Leray para el sistema de Navier-Stokes estable.

Aparte de los refinamientos para la funciones de Sobolev, podemos encontrar versiones recientes del Teorema de Morse-Sard en muchos otros contextos. Por ejemplo Barbet, Dambrine y Daniilidis demuestran en [28], y posteriormente junto a Rifford en [29], que para una subclase especial de funciones

[1]Si $f \in L^1_{loc}(\mathbb{R}^n;\mathbb{R}^m)$, entonces el representante preciso de f se define como

$$f^*(x) := \begin{cases} \lim_{r\to 0} \frac{1}{\mathcal{L}^n(B(x,r))} \int_{B(x,r)} f(y)\, dy & \text{si el límite existe} \\ 0 & \text{en otro caso.} \end{cases}$$

Lipschitz $f : \mathbb{R}^n \to \mathbb{R}^m$, definidas a través de una selección finita, o infinita indexada sobre un compacto numerable, de funciones de clase C^k con $k \geq n - m + 1$, el conjunto de valores críticos de tipo Clarke (concepto más débil que el de valor crítico usual) tiene medida nula. Se trata de una versión no regular del Teorema de Morse-Sard.

Por otro lado en 2018 Azagra, Ferrera y Gómez-Gil prueban que si $f \in C^{n-m}(\mathbb{R}^n; \mathbb{R}^m)$, $m \leq n$, satisface

$$\limsup_{y \to x} \frac{|D^{n-m}f(y) - D^{n-m}f(x)|}{|y - x|} < \infty$$

para todo $x \in \mathbb{R}^n$ (esto es que $D^{n-m}f$ es una función de Stepanov), entonces $\mathcal{L}^m(f(C_f)) = 0$. Si $n = m$ tendríamos que f es diferenciable en casi todo punto, pero a la vez f mandaría conjuntos $\mathcal{L}^n-$nulos en conjuntos $\mathcal{L}^n-$nulos, luego el conjunto de puntos de no diferenciabilidad no nos preocuparía. Los autores dedujeron este resultado de un teorema abstracto y más poderoso de tipo Morse-Sard, concretamente [18, Theorem 1.5]. Este teorema también les permitió recuperar el resultado de De Pascale. Y prácticamente al mismo tiempo, en [17], estos tres autores también encontraron versiones del teorema de Morse-Sard para funciones $f : \mathbb{R}^n \to \mathbb{R}$ subdiferenciales y con desarrollos de Taylor de orden $n - 1$ (este hecho fue uno de los gérmenes de los resultados desarrollados en el Capítulo 2 de esta tesis, donde se demuestra que las funciones subdiferenciales tienen también la propiedad de Lusin de clase C^1 y C^2).

En el Capítulo 1 de esta tesis nosotros hacemos una humilde contribución a todos estos teoremas de Morse-Sard. Para entender la motivación de nuestros resultados hay que observar que las funciones $f : \mathbb{R}^n \to \mathbb{R}^m$ pertencientes a cualquier espacio comentado anteriormente, que se han demostrado tienen la propiedad de Morse-Sard, también resultan satisfacer la propiedad de Lusin de clase C^k, con $k = n - m + 1$. Esto significa que dado $\varepsilon > 0$, existe una función $g : \mathbb{R}^n \to \mathbb{R}^m$ de clase C^k tal que $\mathcal{L}^n(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}) < \varepsilon$. Si $k = 0$, C^0 denota el espacio de las funciones continuas. En [44, 113, 114, 95] la propiedad de Lusin y la propiedad de Morse-Sard se estudian simultaneamente y no sólo se tiene que la clase de funciones consideradas en esos artículos tienen la propiedad de Lusin de la clase correspondiente, sino que esta propiedad es esencial en sus demostraciones del teorema de Morse-Sard.

Para buscar buenas caracterizaciones de la propiedad de Lusin es necesario trabajar con otro concepto de límite, diferente al usual, llamado límite aproximado. Diremos que $\operatorname{ap}\lim_{y\to x} f(y) = L$ si existe un conjunto A_x con densidad uno en x [2] tal que $\lim_{y\to x, y\in A_x} f(y) = L$ (lo mismo para $\limsup$ y $\liminf$). Si $L = f(x)$ decimos que f es aproximadamente continua en x. Una de las principales razones para introducir este concepto es que se tiene la equivalencia entre funciones que son aproximadamente continuas en casi todo punto, funciones medibles y funciones que satisfacen la propiedad de Lusin de clase C^0. Del mismo modo se puede definir la diferenciabilidad aproximada de orden $k \geq 1$ en un punto x si existe un polinomio de orden k centrado en x, $p_k(x; y)$, con $p_k(x; x) = f(x)$ tal que

$$\operatorname{ap}\lim_{y \to x} \frac{|f(y) - p_k(x; y)|}{|y - x|^k} = 0.$$

Y diremos que f tiene un $(k-1)-$polinomio de Taylor aproximado en x si existe un polinomio de orden $k - 1$ centrado en x, $p_{k-1}(x; y)$, con $p_{k-1}(x; x) = f(x)$, tal que

$$\operatorname{ap}\limsup_{y \to x} \frac{|f(y) - p_{k-1}(x; y)|}{|y - x|^k} = 0.$$

La diferenciabilidad aproximada es una propiedad más débil que la posesión de polinomios de Taylor y que la regularidad C^k usual.

[2]Se dice que un conjunto $A \subset \mathbb{R}^n$ tiene densidad uno en $x \in \mathbb{R}^n$ si

$$\lim_{r \to 0} \frac{\mathcal{L}^n(B(x, r) \cap A)}{\mathcal{L}^n(B(x, r))} = 1.$$

Resulta que gracias a los resultados independientes de Isakov [102] y Liu y Tai [121], la propiedad de Lusin de clase C^k es equivalente a ser aproximadamente diferenciable de orden k en casi todo punto $x \in \mathbb{R}^n$ (y también equivalente a tener $(k-1)$-polinomios de Taylor aproximados en casi todo punto).

En vista de estos hechos nos podemos preguntar si es posible probar versiones del teorema de Morse-Sard para funciones $f : \mathbb{R}^n \to \mathbb{R}^m$ que simplemente supongamos sean aproximadamente diferenciables de orden $k = n - m + 1$ en *suficientes* puntos.

En el Capítulo 1 mostraremos como, combinando algunas de las estrategias y herramientas que son comunes a lo largo de la literatura con la idea de la prueba de [121, Theorem 1] y un argumento de inducción, uno puede probar el siguiente resultado no regular de tipo Morse-Sard.

Teorema 1. *Sea $n \geq m$ y $f : \mathbb{R}^n \to \mathbb{R}^m$ una función Borel. Supóngase que f es aproximadamente diferenciable de orden 1 en $\mathcal{H}^m-$casi todo punto y satisface*

(a) ap $\limsup_{y\to x} \frac{|f(y)-f(x)|}{|y-x|} < +\infty$ para todo $x \in \mathbb{R}^n \setminus N_0$, donde N_0 es un conjunto contable, y

(b) (si $n > m$) ap $\lim_{y\to x} \frac{|f(y)-f(x)-\cdots \frac{F_i(x)}{i!}(y-x)^i|}{|y-x|^i} = 0$ para todo $i = 2, \dots, n-m+1$ y para todo $x \in \mathbb{R}^n \setminus N_i$, donde cada conjunto N_i es $(i+m-2)$-σ-finito y los coeficientes $F_i(x)$ (aplicaciones i-multilineales y simétricas) son funciones Borel,

entonces f tiene la propiedad de Morse-Sard (eso es, la imagen del conjunto crítico de f es nulo con respecto a la medida de Lebesgue en $\mathbb{R}^m$).

A pesar de que f es posible que no sea diferenciable en ciertos puntos, si denotamos por AppDiff(f) al conjunto de puntos donde f es aproximadamente diferenciable, estamos entendiendo como su conjunto de puntos críticos a $C_f = \{x \in \text{AppDiff}(f) : \text{rango}(F_1(x)) \text{ no es máximo}\}$. Además decimos que un conjunto es s-σ-finito si puede escribirse como unión contable de conjuntos de $\mathcal{H}^s$-medida finita. Incluimos otra versión de este teorema en la que no suponemos la medibilidad Borel de la función f.

Teorema 2. *Sea $f : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$. Supongamos que para todo $x \in \mathbb{R}^n$ y $i = 1 \dots, n-m$ existen aplicaciones $i-$multilineales y simétricas $F_i(x)$ tales que*

(a) $\text{ap}\limsup_{y\to x} \frac{|f(y)-f(x)|}{|y-x|} < +\infty$ *para todo $x \in \mathbb{R}^n \setminus N_0$, donde N_0 es un conjunto contable.*

(b) El conjunto $N_1 := \{x \in \mathbb{R}^n : \text{ap}\limsup_{y\to x} \frac{|f(y)-f(x)-F_1(x)(y-x)|}{|y-x|^2} = +\infty\}$ *es $\mathcal{H}^m-$nulo.*

(c) Para cada $i = 2, \dots, n-m$ el conjunto

$$N_i := \{x \in \mathbb{R}^n : \text{ap}\limsup_{y\to x} \frac{\left|f(y) - f(x) - \sum_{j=1}^{i} \frac{F_j(x)}{j!}(y-x)^j\right|}{|y-x|^{i+1}} = +\infty\}$$

es contablemente $(\mathcal{H}^{i+m-1}, i+m-1)$ rectificable de clase C^i.

Entonces $\mathcal{L}^m(f(C_f)) = 0$.

Un conjunto $N \subset \mathbb{R}^n$ se dice contablemente $(\mathcal{H}^s, s)$ rectificable de cierta clase C^k si existe un conjunto contable de subvariedades $s-$dimensionales A_j de clase C^k tales que $\mathcal{H}^s(N \setminus \bigcup_{j=1}^{\infty} A_j) = 0$. Ambos Teoremas 1 y 2 siguen siendo ciertos si reemplazamos $\mathbb{R}^n$ por un conjunto abierto $U \subset \mathbb{R}^n$. En el Teorema 1.23 del Capítulo 1 ofrecemos también una interesante variante de este resultado.

Los Teoremas 1 y 2 generalizan las versiones del teorema de Morse-Sard dadas por Bates y por el Apéndice de [17]. Además no son más fuertes, ni más débiles, que las versiones de [44, 113, 114] para BV_n o funciones de Sobolev con exponentes más pequeños; véase la última Sección 1.4 del Capítulo 1 para más ejemplos y comentarios. En particular, gracias a los trabajos de Norton, incluimos un ejemplo que muestra que nuestros resultados son finos en el sentido de que existen funciones $f : \mathbb{R}^n \to \mathbb{R}$ de clase $C^{n-1,t}$, $t \in (0, 1)$ que son aproximadamente diferenciables salvo en un conjunto de $\mathcal{H}^{n-1+t}-$medida nula y que no satisfacen la propiedad de Morse-Sard.

Los resultados del Capítulo 1 están publicados en [21].

Por lo comentado hasta ahora queda de manifiesto por tanto la relación entre el teorema de Morse-Sard y la propiedad de Lusin. Esto llevó al autor de la presente tesis a un estudio más profundo de esta segunda propiedad, que gracias a una fructífera colaboración con los matemáticos Azagra, Ferrera y Gómez-Gil, dio lugar a la demostración de que las funciones subdiferenciales satisfacen la propiedad de Lusin de clase C^1 y C^2. Esto aparecerá en el Capítulo 2 de la tesis. Pero antes de enunciar nuestros descubrimientos comentemos un poco más en profundidad la historia de la propiedad de Lusin, que en muchos casos es paralela a la del teorema de Morse-Sard.

El teorema clásico de Lusin, de 1912, dice que dado un conjunto medible Lebesgue $\Omega \subseteq \mathbb{R}^n$, dada una función medible Lebesgue $f : \Omega \to \mathbb{R}$ y dado $\varepsilon > 0$, existe $g : \Omega \to \mathbb{R}$ continua tal que $\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\}) < \varepsilon$. A partir de entonces diversos autores han demostrado que si se suponen más propiedades de regularidad sobre la f, la función aproximante g podrá tomarse de clase C^k para algún $k \geq 1$. Diremos que $f : \Omega \to \mathbb{R}$ satisface la propiedad de Lusin de clase C^k, $k \geq 0$, si para cada $\varepsilon > 0$ existe $g \in C^k(\Omega; \mathbb{R})$ tal que $\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\}) < \varepsilon$.

Después del resultado de Lusin, la primera extensión relevante se debió a H. Federer en 1944 (probado implícitamente en [73, p. 442]). Mostró que si f era diferenciable en casi todo punto, entonces f tenía la propiedad de Lusin de clase C^1. En particular las funciones Lipschitzianas y las funciones del espacio de Sobolev $W^{1,p}(\mathbb{R}^n; \mathbb{R})$ con $n < p \leq \infty$ tienen la propiedad de Lusin de clase C^1. Este resultado era una simple consecuencia del teorema de extensión de Whitney, el cual seguirá siendo pieza clave en todas las generalizaciones y versiones posteriores del teorema de Lusin. Obsérvese también que la condición de ser diferenciable en casi todo punto es equivalente a que $\limsup_{y \to x} \frac{|f(y)-f(x)|}{|y-x|} < \infty$ para casi todo punto.

Avanzamos ahora hasta el 1951, cuando Whitney en [153] demuestra que una función $f : \Omega \to \mathbb{R}$ satisface la propiedad de Lusin de clase C^1 si y solo si tiene derivadas parciales aproximadas en casi todo punto. Y habría que esperar hasta el 1994, momento en el que Liu y Tai consiguen probar que una función tiene la propiedad de Lusin de clase C^k, para cualquier $k \geq 1$, si y sólo si es aproximadamente diferenciable de orden k en casi todo punto (y si y sólo si tiene$(k-1)$-polinomios de Taylor aproximados en casi todo punto).

Aparte de esta bonita caracterización de la propiedad de Lusin está la pregunta de qué espacios de funciones la satifacen para cierta clase C^k.

Sobre las generalizaciones del teorema de Lusin para funciones de Sobolev, desde 1933 se sabía que las funciones de $W^{1,1}(\mathbb{R}^n)$ tienen derivadas parciales en casi todo punto (Nikodym [129], o ver también [71, Lemma 4.9.2]), así que usando el resultado de Whitney anterior tenemos automáticamente la propiedad de Lusin de clase C^1. La generalización para el caso de $W^{k,p}(\mathbb{R}^n)$, $1 < p < \infty$, es debida a Calderón y Zygmund [49], quienes en 1961 probaron que este espacio tenía la propiedad de Lusin de clase C^k. Más tarde Liu [120] añade al teorema de Calderón y Zygmund la propiedad de que si $f \in W^{k,p}(\mathbb{R}^n)$ y $g \in C^k(\mathbb{R}^n)$ es la función aproximante, entonces g también aproxima a f en la norma de Sobolev. Una versión más fuerte apareció con el trabajo de Michael y Ziemer de 1984 publicado en [124], donde los conjuntos excepcionales se toman en términos de capacidades de Bessel (ver [155, Section 2.6] para las definiciones): Si $1 < p < \infty$, $l = 0, 1, \ldots, k$, $\varepsilon > 0$ y $f \in W^{k,p}_{loc}(\mathbb{R}^n)$, entonces existe $g \in C^l(\mathbb{R}^n)$ tal que $B_{k-l,p}(\{x \in \mathbb{R}^n : f(x) \neq g(x)$ ó $D^j f(x) \neq D^j g(x)$ para algún $j = 1, \ldots, l\}) < \varepsilon$ y $||f-g||_{l,p} \leq \varepsilon$. Para $k = l$, la capacidad de Bessel $B_{0,p}$ coincide con la medida de Lebesgue y por tanto se recupera el resultado de Liu anterior. Obsérvese que a medida que $l = 0, 1 \ldots, k$ va decreciendo los conjuntos *excepcionales* son más cada vez pequeños, pues en general si $0 \leq a < b$ entonces $B_{a,p}(C) \geq B_{b,p}(C)$ para todo $C \subset \mathbb{R}^n$. Bojarski, Hajlasz y Strzelecki en 2002 extienden el resultado de Michael y Ziemer al poder tomar la aproximación en el espacio de Sobolev de un orden mayor, es decir $g \in W^{l+1,p}$ y la cual aproxima a f también en la norma de este espacio, para $l = 0, \ldots, k-1$. Luego aparecieron los resultados de Bourgain, Korobkov y Kristensen que consideran los espacios $W^{k,1}$ y BV_k y donde los conjuntos excepcionales ahora tienen contenido Hausdorff $\mathcal{H}^l_\infty$ pequeño (para más detalles léase [44, Theorem 3.1 y 6.2]). Finalmente Korobkov y Krsitensen en [114, Theorem 2.1] demuestran el siguiente resultado de aproximación tipo Lusin para espacios de Sobolev-Lorentz: sean $k, l \in \{1, \ldots, n\}$, $l \leq k$, $p \in (1, \infty)$ y $f \in W_1^{k,p}(\mathbb{R}^n; \mathbb{R}^m)$, entonces para todo $\varepsilon > 0$ existe un conjunto abierto $U \subset \mathbb{R}^n$ y una función $g \in C^l$ tal que $\mathcal{H}_\infty^{n-(k-l)p}(U) < \varepsilon$ y $f = g$, $D^j f = D^j g$ en $\mathbb{R}^n \setminus U$ para $j = 1, \ldots, l$, donde $\mathcal{H}_\infty^t$ denota el contenido Hausdorff. Esta propiedad resulta clave para demostrar los resultados que aparecen en [95].

Para la clase especial de funciones convexas $f : \mathbb{R}^n \to \mathbb{R}$, Alberti y Imonkulov [2, 101] mostraron que toda función convexa tiene la propiedad de Lusin de clase C^2 (siendo la función aproximante no necesariamente convexa); ver también [1] para un problema relacionado. Más recientemente Azagra y Hajlasz [23] han probado que la aproximación en el sentido de Lusin puede tomarse de clase $C^{1,1}_{loc}$ y convexa si y sólo si f es esencialmente coerciva (en el sentido de que $\lim_{|x|\to\infty} f(x) - l(x) = \infty$ para cierta función lineal l) o bien f ya es de clase $C^{1,1}_{loc}$ (en cuyo caso tomar la propia f como la aproximación es la única opción).

En el Capítulo 2 de este tesis nosotros estamos interesados en las posibles versiones de este tipo de resultados en el mundo de la subdiferenciabilidad. Se probará que las funciones subdiferenciables satisfacen las propiedades de Lusin de clase C^1 or C^2. Por subdiferenciales entendemos el subdiferencial Fréchet, el subdiferencial proximal, o el subdiferencial de viscosidad de segundo orden. Recordamos al lector que el subdiferencial Fréchet de una función $f : \mathbb{R}^n \to \mathbb{R}$ en $x \in \mathbb{R}^n$ se define como el conjunto de $\xi \in \mathbb{R}^n$ tales que $\liminf_{y\to x} \frac{f(y)-f(x)-\langle\xi,y\rangle}{|y-x|} \geq 0$, y que el subdiferencial proximal de f en x se define como el conjunto de $\xi \in \mathbb{R}^n$ tales que $\liminf_{y\to x} \frac{f(y)-f(x)-\langle\xi,y\rangle}{|y-x|^2} > -\infty$. La pregunta en su forma más general que intentamos resolver es: dado $k \in \mathbb{N}$ y una función $f : \mathbb{R}^n \to \mathbb{R}$, supongamos que para casi todo $x \in \mathbb{R}^n$ existe un polinomio $p_{k-1}(x; y)$ centrado en x, con $p_{k-1}(x; x) = f(x)$ y de grado menor o igual que $k - 1$ tal que

$$\liminf_{y\to x} \frac{f(y) - p_{k-1}(x; y)}{|y - x|^k} > -\infty.$$

¿Se sigue que f tiene la propiedad de Lusin de clase C^k? Como ya se ha dicho, en el Capítulo 2 daremos una respuesta positiva para el caso $k = 1, 2$, pero negativa para $k \geq 3$. Todo esto podría tener aplicaciones en análisis no diferenciable o en la teoría de soluciones de viscosidad de PDE como las ecuaciones de Hamilton-Jacobi.

El enunciado preciso para $k = 1$ es:

Teorema 3. *Sea $\Omega \subset \mathbb{R}^n$ un conjunto medible Lebesgue, y $f : \Omega \to \mathbb{R}$ una función. Supongamos que para casi todo $x \in \Omega$ tenemos*

$$\liminf_{y\to x} \frac{f(y) - f(x)}{|y - x|} > -\infty.$$

Entonces, para todo $\varepsilon > 0$ existe una función $g \in C^1(\mathbb{R}^n)$ tal que

$$\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\}) \leq \varepsilon.$$

Si $f : \mathbb{R}^n \to \mathbb{R}$ es una función medible y $\Omega = \{x \in \mathbb{R}^n : \partial^- f(x) \neq \emptyset\}$ es el conjunto de puntos donde el subdiferencial Fréchet de f es no vacía, entonces se sigue del Teorema 3 que para cada $\varepsilon > 0$ existe una función $g \in C^1(\mathbb{R}^n)$ tal que

$$\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\}) \leq \varepsilon.$$

Para el caso de regularidad C^2:

Teorema 4. *Sea $\Omega \subset \mathbb{R}^n$ un conjunto medible Lebesgue, y sea $f : \Omega \to \mathbb{R}$ sea una función tal que para casi todo $x \in \Omega$ existe un vector $\xi_x \in \mathbb{R}^n$ tal que*

$$\liminf_{y\to x} \frac{f(y) - f(x) - \langle\xi_x, y - x\rangle}{|y - x|^2} > -\infty.$$

Entonces para todo $\varepsilon > 0$ *existe una función* $g \in C^2(\mathbb{R}^n)$ *tal que*

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

Si $f : \mathbb{R}^n \to \mathbb{R}$ es una función medible y $\Omega = \{x \in \mathbb{R}^n : \partial_P f(x) \neq \emptyset\}$ es el conjunto de puntos donde el subdiferencial proximal de la f es no vacía, entonces se sigue del Teorema 4 que para cada $\varepsilon > 0$ existe una función $g \in C^2(\mathbb{R}^n)$ tal que

$$\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\}) \leq \varepsilon.$$

En la Sección 2.4 ofrecemos algunos ejemplos que muestran que esta clase de resultados ya no son ciertos para *subexpansiones de Taylor* de órdenes mayores. Una de las principales razones de por qué esto no funciona para $k \geq 3$ reside en el hecho de que poseer subdiferenciales no vacías de orden 2 implica que las subdiferenciales de orden $k \geq 3$ tampoco son vacías (ver por ejemplo [17, Proposition 1.1]).

Todos los resultados del Capítulo 2 se pueden encontrar publicados en [19].

Cambiemos ahora drásticamente de contexto. Todo lo desarrollado en los Capítulos 3, 4 y 5 es en espacios de dimensión infinita. En el Capítulo 3 establecemos nuevos resultados acerca de la extracción difeomorfa de ciertos conjuntos cerrados dentro de espacios de Banach de dimensión infinita. Estos resultados son pieza clave para nuestras demostraciones de los teoremas más importantes de los Capítulos 4 y 5 de la tesis. Es preferible exponer primero las motivaciones y contexto histórico que nos llevan a desarrollar estos útlimos dos capítulos, y dejar para el final de este resumen todo lo concerniente a la extracción difeomorfa de cerrados en espacios de Banach.

Tras ver la importancia que tiene el teorema de Morse-Sard clásico en la literatura es natural preguntarse, ¿qué ocurre si buscamos análogos del teorema para funciones $f : M \to N$ que actúan entre variedades de dimensión infinita?. ¿Qué condiciones de regularidad debemos imponer a las funciones para que su conjunto de valores críticos $f(C_f)$ sea pequeño en algún sentido? Aquí definimos el conjunto de puntos críticos por $C_f := \{x \in M : Df(x) \text{ no es un operador sobreyectivo}\}$. En general M y N denotarán espacios de Banach o bien variedades modeladas en espacios de Banach.

Esta pregunta fue estudiada por primera vez por Smale en 1965, quién en [142] probó que si M y N son variedades separables, conexas y C^∞, modeladas en espacios de Banach, y $f : M \to N$ es una aplicación C^r de Fredholm (esto es, toda diferencial $Df(x)$ es un operador Fredholm entre los correspondientes espacios tangentes) entonces $f(C_f)$ es de primera categoría, y en particular $f(C_f)$ no tiene puntos interiores, siempre que $r > \text{máx}\{\text{índice}(Df(x)), 0\}$ para todo $x \in M$; aquí índice$(Df(x))$ significa el índice del operador Fredholm $Df(x)$, que es la diferencia entre las dimensiones del núcleo de $Df(x)$ y la codimensión de la imagen de $Df(x)$, las cuales son finitas. Por supuesto, estas suposiciones son muy restrictivas ya que, por ejemplo, si M es infinito-dimensional entonces ninguna función $f : M \to \mathbb{R}$ es de Fredholm.

En general, cualquier intento de adaptar el teorema de Morse-Sard a dimensión infinita tendrá que imponer fuertes restricciones porque, como muestra un contraejemplo de Kupka [116], hay funciones $f : \ell_2 \to \mathbb{R}$ de clase C^∞ tales que su conjunto de valores críticos $f(C_f)$ contiene intervalos. Como muestran Bates y Moreira en [32], uno puede incluso hacer que f sea un polinomio de grado 3. En concreto uno puede tomar la función

$$f(\sum_{n=1}^{\infty} x_n e_n) = \sum_{n=1}^{\infty}(3 \cdot 2^{-\frac{n}{3}} x_n^2 - 2x_n^3),$$

cuyo conjunto de puntos críticos es $C_f = \{\sum_{n=1}^{\infty} x_n e_n : x_n \in \{0, 2^{-\frac{n}{3}}\}\}$ y tal que $\mathcal{H}^3(C_f) < +\infty$ y $f(C_f) = [0, 1]$.

Por suerte, para muchas aplicaciones del teorema de Morse-Sard, es a menudo suficiente que cualquier aplicación continua dada pueda ser uniformemente aproximada por una aplicación cuyo conjunto de valores críticos sea pequeño en algún sentido; es por tanto lógico preguntarse qué aplicaciones entre qué espacios de Banach satisfacen tal propiedad de aproximación. Yendo en esta dirección, Eells y McAlpin establecieron el siguiente teorema [69]: Si E es un espacio de Hilbert separable, entonces toda función continua de E en $\mathbb{R}$ puede aproximarse uniformemente por una función f de clase C^∞ cuyo conjunto de valores críticos $f(C_f)$ es de medida cero. Esto les permitió deducir una versión de este teorema para aplicaciones entre variedades C^∞, M y N, modeladas en un espacio de Hilbert E y un espacio de Banach F respectivamente, lo que ellos llamaron un *teorema de Morse-Sard aproximado*: Toda función continua de M en N puede ser uniformemente aproximada por una función $f : M \to N$ de clase C^∞ tal que $f(C_f)$ tiene interior vacio. No osbtante, como se observó en [69, Remark 3A], tenemos $C_f = M$ en el caso de que F sea infinito-dimensional (así que, aunque el conjunto de valores críticos es relativamente pequeño, el conjunto de puntos críticos de f es enorme, lo que es algo decepcionante). Un resultado parecido de la misma época consigue además la aproximación en las derivadas, siempre que la función inicial sea C^1. Se trata de un resultado de Moulis [128, p. 331] que dice lo siguiente: para toda función C^1 $f : E \to F$, donde E es un espacio de Hilbert separable infinito-dimensional y F es un espacio de Hilbert separable, y para toda función continua $\varepsilon : E \to (0, \infty)$ existe una función C^∞ $g : E \to F$ tal que $||f(x) - g(x)|| \leq \varepsilon(x)$, $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ para todo $x \in E$ y tal que $g(C_g)$ tiene interior vacío en F.

En [12], un resultado mucho más fuerte fue obtenido por Azagra y Cepedello-Boiso: Si M es una variedad C^∞ modelada en espacios infinito-dimensionales separables de Hilbert E, entonces toda aplicación continua de M en $\mathbb{R}^m$ puede ser uniformemente aproximada por funciones C^∞ *sin puntos críticos.* Este tipo de enunciado es el más fuerte que se puede conseguir en el contexto de teoremas aproximados de Morse-Sard. Desafortunadamente, puesto que parte de la prueba requiere el uso de las buenas propiedades de la norma Hilbertiana, esto no puede extenderse de manera directa a otros espacios de Banach. Por otro lado una parte importante de la demostración requiere la posibilidad de encontrar difeomorfismos $h : E \to E \setminus X$, donde X es un conjunto cerrado localmente compacto, y que estén arbitariamente cerca de la identidad, concretamente se usa [150, Theorem 1]. Conviene recordar esta idea, pues volverá a ser clave para todo el desarrollo de los Capítulos 3, 4 y 5 de esta tesis.

P. Hájek y M. Johannis [91] establecieron un resultado similar para $m = 1$ en el caso de que E sea un espacio separable de Banach que contiene c_0 y admita funciones meseta de clase C^k (entendemos por función meseta a toda función continua $\lambda : E \to [0, \infty)$ cuyo soporte $\text{sop}(f) = \overline{\{x \in E : \lambda(x) \neq 0\}}$ sea acotado). En este caso las funciones aproximantes son de clase C^k. El método que emplean Hájek y Johannis está basado en el resultado de que las funciones con valores reales de clase C^∞ definidas en c_0 y que localmente dependen en un número finito de coordenadas[3] son densas en el espacio de las funciones continuas definidas en c_0 y con valores reales (ver [60]). Sin embargo, los autores ya indicaron que su método no se aplica cuando el espacio E es reflexivo, dejando fuera por tanto los espacios de Banach clásicos ℓ_p y L^p, $1 < p < \infty$. Un poco más tarde, Azagra y Mar Jiménez en 2007 caracterizaron la clase de espacios de Banach separables E tales que para toda función continua $f : E \to \mathbb{R}$ y $\varepsilon : E \to (0, \infty)$ existe una función $g : E \to \mathbb{R}$ de clase C^1 para la que $||f(x) - g(x)|| \leq \varepsilon(x)$ y $g'(x) \neq 0$ para todo $x \in E$, como aquellos espacios de Banach infinito-dimensionales con dual separable (véase [25]).

Hasta ahora, hay algunas consecuencias *buenas* de estos resultados, todos ellas relacionadas de alguna forma con teoremas de tipo Morse-Sard. Pero hay también algunas *malas* consecuencias. Una de ellas, puesto que el conjunto de funciones diferenciables sin puntos críticos es denso en el conjunto de funciones continuas, es que hay conjuntos bastantes grandes de funciones para los que ninguna teoría de Morse concebible es válida.

Entre las *buenas* consecuencias hay algunas interesantes como la existencia de un teorema de Hahn Banach no lineal o la construcción de ejemplos de funciones $f : E \to \mathbb{R}$ que no cumplen el teorema de Rolle y que tienen un soporte prefijado (ver [25]). Es bien conocido que el teorema de Rolle en general

[3]Decimos que una función f definida en un espacio de Banach E localmente depende de un número finito de coordenadas si para todo $x \in E$ existe un número natural l_x, un entorno abierto U_x de x, ciertos funcionales $L_1, \ldots, L_{l_x} \in E^*$ y una función $\gamma : \mathbb{R}^{l_x} \to \mathbb{R}$ tal que

$$f(y) = \gamma(L_1(y), \ldots, L_{l_x}(y))$$

para todo $y \in U_x$.

falla en dimensión infinita; el primer ejemplo de este tipo se debe a Shkarin en [141]. El lector interesado también puede echar un vistazo a los artículos [77, 22, 24], aunque se comentará más en detalle este tema en la Sección 5.6.

Aunque es obvio, cabe mencionar que en dimensión finita este tipo de resultados de aproximación de funciones continuas por funciones diferenciables sin puntos críticos no se puede dar. Basta pensar por ejemplo en $f:\mathbb{R}^n\to\mathbb{R}^m$ definida como $f(x)=(f^1(x),\ldots,f^m(x))=(|x|,0,\ldots,0)$, donde $|\cdot|$ denota la norma euclidea usual. Si queremos aproximar uniformemente f por otra g de clase C^1, debemos tener que $g^1(x)$ aproxima $|x|$ y por lo tanto g^1 debe alcanzar un mínimo global en cierto punto x_0 para el cual se debe tener $\nabla g^1(x_0)=(0,\ldots,0)$.

Nuestra preocupación en el desarrollo del Capítulo 4, fue intentar reemplazar los pares $(\ell_2,\mathbb{R}^m)$ o $(E,\mathbb{R})$, de los resultados de Azagra y Cepedello y de Azagra y Mar Jiménez, por otros pares de la forma (E,F) donde E fuese un espacio de Banach lo más general posible y F sea un cociente de E, pudiendo ser de dimensión infinita. La hipótesis de considerar F como cociente de E es obligatoria ya que si no hay operadores sobreyectivos de E sobre F entonces todos los puntos son críticos para todas las funciones diferenciables.

Una de las claves para los resultados que presentamos es que el espacio de Banach de salida E tenga cierta estructura "compuesta". En un caso con esto nos referiremos a que E sea isomorfo a su cuadrado y en otro caso a que E posea base incondicional.

Para espacios E que son reflexivos y son isomorfos a su cuadrado tenemos el siguiente resultado.

Teorema 5. *Sea E un espacio de Banach separable reflexivo de dimensión infinita, y F un espacio de Banach. En el caso de que F sea infinito-dimensional, asumamos además que:*

1. *E es isomorfo a $E\oplus E$.*
2. *Existe un operador lineal acotado de E sobre F (equivalentemente, F es un espacio cociente de E).*

Entonces, para toda aplicación continua $f:E\to F$ y toda función continua $\varepsilon:E\to(0,\infty)$ existe una aplicación $g:E\to F$ de clase C^1 tal que $||f(x)-g(x)||\le\varepsilon(x)$ y $Dg(x):E\to F$ es un operador lineal sobreyectivo para cada $x\in E$.

Téngase en cuenta que existen espacios de Banach separables y reflexivos E tales que no son isomorfos a $E\oplus E$. El primer ejemplo de tal espacio fue dado por Fiegel en 1972 [81].

Para espacios que no son necesariamente reflexivos pero tienen una base de Schauder apropiada tenemos lo siguiente.

Teorema 6. *Sea E un espacio de Banach infinito-dimensional, y F un espacio de Banach tales que:*

(i) E tiene una norma equivalente localmente uniformemente convexa $\|\cdot\|$ que es C^1.

(ii) $E=(E,\|\cdot\|)$ tiene una base (normalizada) de Schauder $\{e_n\}_{n\in\mathbb{N}}$ tal que para todo $x=\sum_{j=1}^{\infty}x_je_j$ y todo $j_0\in\mathbb{N}$ tenemos que

$$\left\|\sum_{j\in\mathbb{N},\,j\neq j_0}x_je_j\right\|\le\left\|\sum_{j\in\mathbb{N}}x_je_j\right\|.$$

(iii) En el caso de que F es infinito-dimensional, existe un subconjunto $\mathbb{P}$ de $\mathbb{N}$ tal que ambos $\mathbb{P}$ y $\mathbb{N}\setminus\mathbb{P}$ son infinitos y, para cada subconjunto infinito J de $\mathbb{P}$, existe un operador lineal acotado de $\overline{\mathrm{span}}\{e_j:j\in J\}$ sobre F (equivalentemente, F es un espacio cociente de $\overline{\mathrm{span}}\{e_j:j\in J\}$).

Entonces, para toda aplicación continua $f:E\to F$ y toda función continua $\varepsilon:E\to(0,\infty)$ existe una aplicación $g:E\to F$ de clase C^1 tal que $||f(x)-g(x)||\le\varepsilon(x)$ y $Dg(x):E\to F$ es un operador lineal sobreyectivo para cada $x\in E$.

Las demostraciones de estos teoremas, que fueron un trabajo conjunto con Azagra y Dobrowolski, se encuentran en el Capítulo 4 y están publicados en [16].

La parte (ii) del Teorema 6 es equivalente al hecho de que, para todo conjunto $A \subset \mathbb{N}$ (equivalentemente, todo conjunto finito $A \subset \mathbb{N}$) se tiene $\|P_A\| \leq 1$, donde P_A representa la proyección $P_A(x) = \sum_{j\in A} x_j e_j$. Esto, en particular, implica que $\{e_n\}_{n\in N}$ es una base incondicional, con constante supresora incondicional K_s igual a 1; para más detalles ver [4, p. 53] o [3].

Estos resultados engloban los casos clásicos de espacios de Banach E como c_0, ℓ_p, L^p, $1 < p < \infty$, siendo F un espacio de Banach tal que existen operadores lineales y acotados de E sobre F. En particular el resultado también se aplica a los espacios de Sobolev $W^{k,p}(\mathbb{R}^n)$ con $1 < p < \infty$ ya que son isomorfos a $L^p(\mathbb{R}^n)$ (véase [132, Teorema 11]). Además será cierto para espacios menos clásicos como por ejemplo $C(K)$, siendo K un compacto metrizable contable (en general cualquier predual isométrico de ℓ_1) o como el espacio de James J. En este último caso podemos tomar como espacio de llegada F cualquier cociente de cualquier subespacio complementado infinito-dimensional y reflexivo de J.

Más aún, podemos conseguir que las funciones aproximantes tenga más regularidad que simplemente ser C^1. Para ello es necesario usar los resultados de Nicole Moulis en aproximación C^1-fina en espacios de Banach [128], o bien los resultados más generales de [93, Corollary 7.96]. Se logra entonces aproximar uniformemente funciones continuas $f : E \to F$, donde $E = c_0, \ell_p, L^p$, $1 < p < \infty$, por funciones C^k sin puntos críticos, donde k denota el orden de regularidad del espacio E en cada caso. Por ejemplo para c_0 o ℓ_p, L^p con p par se consigue regularidad C^∞.

Una aplicación directa de nuestros teoremas es que para todas las parejas de espacios de Banach (E, F) a las que se puede aplicar o bien Teorema 5 ó 6, toda función continua $f : E \to F$ puede aproximarse uniformemente por aplicaciones abiertas de clase C^k (siendo k el orden de regularidad del espacio E).

La demostración de los Teoremas 5 y 6 consta esencialmente de dos partes. En primer lugar, puesto que el resultado es invariante por difeomorfismos, bastará probarlo para funciones continuas $f : S^+ = \{(u,t) : u \in E, t > 0, \|u\|^2 + t^2 = 1\} \to F$ definidas en la esfera unidad superior del espacio $Y = E \times \mathbb{R}$, el cual dotamos de la norma $|(u,t)| = \left(\|u\|^2 + t^2\right)^{1/2}$. De este modo definiremos una primera aproximación uniforme $\varphi : S^+ \to F$ de clase C^1,

$$\varphi(y) = \sum_{n\in\mathbb{N}} \psi_n(y)(f(y_n) + T_n(y)),$$

donde $\{\psi_n\}_{n\in\mathbb{N}}$ son particiones de la unidad diferenciables especialmente construidas en S^+, $y_n \in \operatorname{supp}(\psi_n)$ y donde estamos perturbando localmente por cieros operadores lineales sobreyectivos $T_n : Y \to F$. Esta construcción es realizada con sumo cuidado de tal modo que $||f(x) - \varphi(x)|| \leq \varepsilon(x)/2$ para $x \in S^+$ y tal que el conjunto de puntos críticos C_φ esté dentro de un conjunto difeomorficamente extractible (éste será un conjunto cerrado que localmente está contenido en el grafo de una función continua definida en un subespacio complementado infinito-codimensional de E y tomando valores en su complementario lineal de Y). En este primer paso son importantes la separabilidad del espacio E, el poder trabajar con una norma equivalente de clase C^1, y la existencia de *muchos* operadores lineales acotados y sobreyectivos $S_n : E \to F$. Estos operadores se definirán desde ciertos subespacios complementados E_n de codimensión infinita en E y tales que $E_n \cap E_m = \{0\}$ para todo $n \neq m$. Aquí es donde adquiere su importancia el poder descomponer el espacio E en suficientes subespacios complementados de codimensiones infinitas, esto es, poder escribir $E = E_1 \oplus E_2$ con E_1 y E_2 isomorfos a E, o bien disponer de una base incondicional en el espacio. Una vez que se ha realizado este primer paso ya simplemente falta extraer difeomorficamente el conjunto de puntos críticos de esta función aproximante φ. Debemos encontrar un difeomorfismo $h : S^+ \to S^+ \setminus C_\varphi$ de clase C^1 tal que $\{\{x, h(x)\} : x \in S^+\}$ refina $\mathcal{G}$ (en otras palabras, h está *limitada por* $\mathcal{G}$), donde $\mathcal{G}$ es un recubrimiento abierto de S^+ por bolas abiertas $B(z, \delta_z)$ elegidas de tal modo que si $x, y \in B(z, \delta_z)$ entonces

$$\|\varphi(y) - \varphi(x)\| \leq \frac{\varepsilon(z)}{4} \leq \frac{\varepsilon(x)}{2}.$$

La existencia de tales difeomorfismos, viene dada por los resultados incluidos en el Capítulo 3. Tomando $g := \varphi \circ h$ habríamos acabado.

El Capítulo 5 tiene un aroma algo diferente. Supongamos que $f : E \to F$ es una función de clase C^1 y que sabemos que su conjunto de puntos críticos C_f está incluido en cierto abierto U. ¿Dada una función continua $\varepsilon : E \to (0, \infty)$, es posible encontrar $\varphi : E \to F$ de clase C^1 sin puntos críticos, con $||f(x) - \varphi(x)|| \leq \varepsilon(x)$, y tal que $f = \varphi$ fuera de U?

Responderemos a esta pregunta afirmativamente para el caso en el que E sea c_0 o ℓ_p, $1 < p < \infty$, y $F = \mathbb{R}^d$. Además, en el caso de c_0 se puede conseguir que $||Df(x) - D\varphi(x)|| \leq \varepsilon(x)$ para todo $x \in c_0$.

El camino tomado para atacar este problema es el siguiente. Primero construimos una función $g : U \to \mathbb{R}^d$ de clase C^1 tal que $|f(x) - g(x)| \leq \varepsilon(x)/2$ y $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ y tal que C_g es o bien el conjunto vacio para el caso de c_0, o bien está localmente contenido en una unión finita de subespacios complementados de codimensión infinita en E para el caso de ℓ_p. Las técnicas empleadas por Nicole Moulis acerca de aproximación C^1-fina de [128], y que también fueron utilizadas en [20], serán de gran ayuda. En segundo lugar extendemos la función g al espacio entero E haciéndole igual a f fuera de U. Debido a la aproximación C^1-fina del primer paso esta extensión es todavía de clase C^1 en E. Para el caso de c_0 habríamos acabado. Para el caso de ℓ_p debemos encontrar un difeomorfismo C^1 $h : E \to E \setminus C_g$ que será la identidad fuera de U y tal que $\{\{x, h(x)\} : x \in E\}$ refina $\mathcal{G} = \bigcup_{z \in E} B(z, \delta_z)$, donde $\delta_z > 0$ es elegido de tal modo que si $x, y \in B(z, \delta_z)$ entonces $|g(y) - g(x)| \leq \frac{\varepsilon(z)}{4} \leq \frac{\varepsilon(x)}{2}$. La existencia de tal difeomorfismo h se sigue de nuevo por los resultados del Capítulo 3. Entonces la aplicación $\varphi(x) := g(h(x))$ no tiene puntos críticos, es igual a f fuera de U y satisface $|f(x) - \varphi(x)| \leq \varepsilon(x)$ para todo $x \in E$.

Recordemos por un momento el ya comentado trabajo de Moulis, [128]. En su artículo Moulis demuestra que para toda función $f : E \to F$ de clase C^1, donde E es un espacio de Hilbert separable infinito-dimensional y F es un espacio de Hilbert separable, y para toda función continua $\varepsilon : E \to (0, \infty)$ existe una función $g : E \to F$ de clase C^∞ tal que $||f(x) - g(x)|| \leq \varepsilon(x)$, $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ para todo $x \in E$ y tal que $g(C_g)$ tiene interior vacío en F. Al compararlo con nuestro resultado es claro que nosotros mejoramos la conclusión al ser capaces de obtener $C_g = \emptyset$ y al considerar otros espacios de Banach, no necesariamente Hilbertianos. Sin embargo para el caso Hilbert no podemos escribir como espacio de llegada un espacio de Banach infinito-dimensional como hace Moulis y tampoco podemos conseguir la aproximación en las derivadas.

A continuación se presentan enunciados precisos de los resultados obtenidos en el Capítulo 5. En el primero se trata el caso en el que E sea un espacio de Banach infinito-dimensional con base incondicional y norma equiavelente C^1 que localmente depende de un número finito de coordenadas (en particular c_0). En el segundo consideramos E como espacio de Banach infinito-dimensional con norma equivalente C^1 estrictamente convexa y con base $1-$supresora incondicional (en particular ℓ_p, $1 < p < \infty$).

Teorema 7. *Sea E un espacio de Banach infinito-dimensional con base incondicional y norma equivalente C^1 que localmente depende de un número finito de coordenadas. Sea $f : E \to \mathbb{R}^d$ una función C^1 y $\varepsilon : E \to (0, \infty)$ una función continua. Tomar cualquier conjunto abierto U tal que $C_f \subset U$. Entonces existe una función C^1 $\varphi : E \to \mathbb{R}^d$ tal que,*

1. *$|f(x) - \varphi(x)| \leq \varepsilon(x)$ para todo $x \in E$.*

2. *$f(x) = \varphi(x)$ para todo $x \in E \setminus U$.*

3. *$||Df(x) - D\varphi(x)|| \leq \varepsilon(x)$ para todo $x \in E$; y*

4. *$D\varphi(x)$ es sobreyectivo para todo $x \in E$.*

Teorema 8. *Sea E un espacio de Banach infinito-dimensional con norma equivalente C^1 estrictamente convexa y con base $1-$supresora incondicional $\{e_n\}_{n \in \mathbb{N}}$, eso es una base de Schauder tal que para todo*

$x = \sum_{j=1}^{\infty} x_j e_j$ y cada $j_0 \in \mathbb{N}$ tenemos que

$$||\sum_{j\in\mathbb{N}, j\neq j_0} x_j e_j|| \leq ||\sum_{j\in\mathbb{N}} x_j e_j||.$$

Sea $f : E \to \mathbb{R}^d$ una función C^1 y $\varepsilon : E \to (0,\infty)$ una función continua. Sea U cualquier conjunto abierto tal que $C_f \subset U$. Entonces existe una función C^1 $\varphi : E \to \mathbb{R}^d$ tal que,

1. *$|f(x) - \varphi(x)| \leq \varepsilon(x)$ para todo $x \in E$.*

2. *$f(x) = \varphi(x)$ para todo $x \in E \setminus U$.*

3. *$D\varphi(x)$ es sobreyectivo para todo $x \in E$.*

Los resultados del Capítulo 5 se encuentran publicados en [87].

Es momento de hablar sobre los resultados obtenidos en el Capítulo 3, que como se ha dicho son imprescindibles para el desarrollo de los teoremas de Morse-Sard aproximados en espacios de Banach de dimensión infinita de los Capítulos 4 y 5.

Comenzaremos con una introducción histórica sobre la eliminación difeomorfa de cerrados en espacios de Banach.

Lo que uno podría llamar *teoría de eliminación* (en ocasiones también nos referiremos a esta como teoría de extracción) en espacios de Banach comenzó en 1953 cuando Victor L. Klee en [110] probó que si E es un espacio de Banach no reflexivo o un espacio clásico L^p y K es un subconjunto compacto de E, existe un homeomorfismo entre E y $E \setminus K$ que es la identidad fuera de un entorno abierto de K. Diremos en este caso que K es (topológicamente) elinimable o extractible. Klee también probó que para tales espacios de Banach infinito dimensionales E la esfera unidad es homeomorfa a cualquiera de sus hiperplanos cerrados, y dio una clasificación topológica de cuerpos convexos en espacios de Hilbert. Estos resultados fueron más adelante extendidos a la clase de todos los espacios normados infinito-dimensionales por C. Bessaga y el mismo Klee en [36, 37]. No hace falta indicar que estos resultados no se dan en dimensión finita: uno no puede obtener homeomorfismos entre $\mathbb{R}^n$ y $\mathbb{R}^n \setminus \{0\}$, pues el primero es contractible (homotópico a un punto) y el segundo no lo es. Luego siempre que se hable sobre teoría de eliminación debemos pensar en espacios o variedades de dimensión infinita.

El trabajo de Klee estaba motivado por los de Tychonoff [146] y Kakutani [107]. Del teorema del punto fijo de Tychonoff se sigue que, en la topología débil, la bola unidad B_E del espacio de Hilbert E tiene la propiedad del punto fijo. En la topología de la norma, sin embargo, esto es falso. S. Kakutani contruyó un homeomorfismo sin puntos fijos de B_E en sí misma. Usando este hecho mostró que la esfera unidad S_E de E es contractible y es un retracto de deformación de B_E. Kakutani preguntó: ¿Son E, B_E y S_E homemorfos?, a lo que el trabajo de Klee respondió afirmativamente.

Las pruebas originales de Klee eran de un fuerte contenido geométrico: muy bonitas, pero bastante dificiles de manejar de manera analítica. Bessaga encontró elegantes fórmulas explícitas para construir homeomorfismos eliminadores, basadas en la existencia de normas continuas no completas (no equivalentes) en todo espacio de Banach infinito-dimensional. Usando lo que él autodenominó *técnica de la norma no completa* y gracias a la existencia de una norma C^∞ no completa en todo espacio de Hilbert E, en 1966, Bessaga prueba que existe un difeomorfismo entre E y $E \setminus \{0\}$ siendo la identidad fuera de una bola (ver [34]). Consecuentemente concluyó que el espacio de Hilbert infinito-dimensional es difeomorfo a su esfera unidad.

Después del gran resultado de Bessaga, a finales de los 60 y principio de los 70 apareció una importante cantidad de artículos tratando la eliminación homeomorfa y difeomorfa de conjuntos cerrados, bien en variedades, bien en espacios de Fréchet (espacio topológicos localmente convexos que son completos respecto de una métrica invariante por traslaciones) o bien en espacios de Banach. Las preguntas que se planteaban eran: ¿Qué espacios admiten la eliminación homeomorfa o difeomorfa de sus puntos o de sus compactos (o de otros cerrados en general)?; Si queremos eliminar homeomórficamente o difeomórficamente compactos de un espacio, ¿qué condiciones de diferenciabilidad debemos imponer al espacio en cuestión en su estructura geométrica para que esto sea posible? (observérvese que la existencia de funciones meseta con el grado de diferenciabilidad del difeomorfismo buscado es una condición necesaria); y ¿en qué casos se puede además saber que el homeomorfismo o difeomorfismo extractor está tan cerca de la identidad como queramos, es decir, que esté limitado por cualquier recubrimiento abierto?

En el mundo homeomorfo sobre teoría de eliminación destaca un resultado de R. D. Anderson, D. H. Henderson y J. West ([8], 1969). Ellos mostraron que si M es una variedad metrizable modelada en un espacio de Fréchet infinito-dimensional separable, entonces cada subvariedad N abierta, densa y con la propiedad de que todo conjunto abierto U, U y $U \cap N$ tienen el mismo tipo de homotopía, es homeomorfo a M por un homeomorfismo que puede requerirse sea la identidad en cualquier subconjunto cerrado dentro de N y esté limitado por cualquier recubrimiento. Tales subvariedades incluyen los complementarios de todos los conjuntos cerrados y localmente compactos de M.

Pero centrándonos en la teoría de extracción difeomorfa, quizás el resultado más celebrado y hoy en día clásico y fundamental es el hecho de que dos variedades de Hilbert infinito-dimensionales, separables y homotópicas son difeomorfas (véanse los trabajos de Burghelea, Kuiper, Eells, Elworthy y Moulis [47, 127, 67, 70]), para los cuales fue clave el ya citado trabajo de Bessaga.

Destacamos ahora dos trabajos que resultan claves para el desarrollo del Capítulo 3 de esta tesis. En el primero se consigue que los difeormorfismos extractores h estén tan cerca de la identidad como queremos, esto será que refinen un recubrimiento abierto dado $\mathcal{G}$ (que para todo $x \in E$ exista $U \in \mathcal{G}$ tal que $x, h(x) \in U$).

En 1969 James West probó que en el caso cuando E es un espacio de Hilbert separable o incluso una variedad de Hilbert separable, para cada conjunto K localmente compacto, cada abierto U incluyendo a K, y cada recubrimiento abierto $\mathcal{G}$ de E, existe un difeomorfismo $h : E \to E \setminus K$ de clase C^∞ tal que h es la identidad fuera de U y está limitado por $\mathcal{G}$.

Peter Renz en su tesis de 1969 probó lo siguiente: Sea E un espacio de Banach infinito-dimensional con base incondicional y con norma equivalente C^k. Entonces para cada conjunto cerrado y localmente compacto K y cada conjunto abierto U, existe un difeomorfismo C^k de $E \setminus K$ sobre $E \setminus (K \setminus U)$ que es la identidad en $E \setminus (U \setminus K)$. (Lo mismo es cierto para variedades C^k paracompactas modeladas en tales espacios de Banach).

Tras estos años de gran producción matemática en este campo encontramos esencialmente tres autores que han seguido refinando la teoría de eliminación difeomorfa en espacios de Banach desde entonces hasta nuestros días; estos son T. Dobrowoslki, D. Azagra y A. Montesinos. En todos los casos, sus trabajos no prestan atención a intentar que los difeomorfismos extractores estén arbitrariamente cerca de la identidad, en el sentido de refinar un recubrimiento dado. Esta sutil propiedad ha sido retomada tras casi 50 años en esta tesis, pues resulta ser imprescindible para nuestros teoremas sobre aproximación uniforme sin puntos críticos.

La eliminación difeomorfa de conjuntos compactos en espacios de Banach fue recuperada por Tadeusz Dobrowolski, quien desarrolló la técnica de la norma no completa de Bessaga. En su artículo [63] de 1979 Dobrowolski primeramente probó que si E es un espacio de Banach separable infinito-dimensional y si K es un conjunto débilmente compacto o un subespacio cerrado infinito-codimensional, entonces E y $E \setminus K$ son real-analíticos difeomorfos[4]. Anteriormente lo único que se sabía sobre extracción real-analítica era que un espacio de Hilbert E infinito-dimensional es real-analítico difeomorfo a $E \setminus \{0\}$ y la eliminación del punto puede ocurrir al final de una C^∞ isotopía. Esto fue probado por Burghelea y Kuiper [47]. En segundo lugar Dobrowolski mostró que para todo espacio de Banach infinito-dimensional E teniendo una norma C^k no completa y para cada conjunto compacto K en E, el espacio E es C^k difeomorfo a $E \setminus K$. En particular los espacios de Banach para los que existen inyecciones lineales y continuas en algún $c_0(\Gamma)$ (los espacios débilmente compactamente generados WCG[5] satisfacen

[4]Tales resultados no son válidos en el caso no separable: si Γ es un conjunto incontable, entonces $c_0(\Gamma)$ no es real-analítico difeomorfo a $c_0(\Gamma) \setminus \{0\}$ (ver [63, Proposition 4.7]).

[5]Un espacio de Banach E es débilmente compactamente generado si $\overline{\text{span}}(K) = E$ para algún K débilmente compacto.

esto, ver [60, p. 246], por lo que también todos los espacios separables y todos los reflexivos) tienen normas C^∞ no completas. Si, además, E tiene una norma equivalente $||\cdot||$ de clase C^k entonces uno puede deducir que $S = \{x \in E : ||x|| = 1\}$ es C^k difeomorfo a cualquier hiperplano de E. Dobrowolski también empleó sus resultados de eliminación difeomorfa para dar una clasificacion de cuerpos C^k en espacios de Banach WCG (ver [62]).

Remarcamos que hay ejemplos de espacios con normas equivalentes C^∞ que no se incluyen linealmente en ningún $c_0(\Gamma)$. Un ejemplo de tal espacio de Banach (no separable) está dado en [60, Ex. VI.8.8] y puede elegirse que sea $C(K)$ para un cierto compacto K. Así, cuando uno quiere generalizar esos resultados a todo espacio de Banach infinito-dimensional que tenga una norma C^k suave, uno encara el siguiente problema: ¿Todo espacio de Banach infinito-dimensional con una norma equivalente C^k admite una norma no completa C^k también? Para $k \geq 2$, esta intrigante pregunta fue resuelta muy recientemente por D'Alessandro y Hájek [57], pero el caso C^1 sigue sin resolverse.

Sin probar la existencia de normas no-completas diferenciables, introduciendo un tipo de función no-completa asimétrica y convexa (llamada norma asimétrica), D. Azagra mostró en 1997 [10] que todo espacio de Banach E con una (no necesariamente equivalente) norma C^k es C^k difeomorfo a $E \setminus \{0\}$ y, además, que todo hiperplano cerrado es C^k difeomorfo a la esfera $\{x \in E : ||x|| = 1\}$. Hoy en día, para el caso $k \geq 2$, usando el resultado de D'Alessandro y Hájek de [57] esto sería sencillo por una aplicación directa de [63].

En 1998 Azagra y Dobrowolski unieron esfuerzos para reforzar la técnica de la norma asimétrica introducida en [10] y generalizaron algunos resultados en eliminación diferenciable de compactos y subespacios a la clase de todos los espacios de Banach que tienen una (no necesariamente equivalante) norma C^k diferenciable. Ellos también dieron una completa clasificación diferenciable de cuerpos convexos de cada espacio de Banach. En particular, mostraron que todo cuerpo convexo diferenciable que no contine subespacios lineales en un espacio de Banach infinito-dimensional es difeomorfo a un semi-espacio.

Estos resultados les habilitaron para ampliar la clase de espacios para los que otras aplicaciones de la eliminación son válidas. Una muestra de tales aplicaciones incluyen teoremas de Garay [85, 86] concernientes a la existencia de soluciones de ecuaciones diferenciales ordinarias y secciones transversales de soluciones embudo en espacios de Banach, además de enunciados más finos de los resultados de Klee [110] en homeomorfismos periódicos sin puntos fijos.

En la geometría intrínsica y estructura de los espacios de Banach juega un papel importante la existencia de normas diferenciables, de funciones meseta diferenciables y de particiones de la unidad diferenciables. Tras los trabajos de Azagra y Dobrowolski en el que la hipótesis principal era la existencia de una norma (no necesariamente equivalente) de clase C^k surge la pregunta de qué resultados de eliminación difeormorfa son posibles suponiendo la existencia de funciones meseta o particiones de la unidad de cierta clase C^k. Estas suposiciones iniciales son naturales ya que son más débiles que la mera existencia de una norma equivalente C^k en el espacio de Banach, como mostró el ejemplo de Haydon en [99]. Este fue el problema estudiado en la tesis de Montesinos, alumno de Azagra y Jaramillo.

Como fruto a este trabajo Azagra y Montesinos añadieron dos nuevos resultados a la teoría de extracción difeomorfa en 2003. Por un lado probaron que si E es un espacio de Banach infinito-dimensional con particiones de la unidad C^k entonces, para cada conjunto débilmente compacto K y cada cuerpo estrellado A tal que $\operatorname{dist}(K, E \setminus A) > 0$, existe un difeomorfismo $h : E \to E \setminus K$ de clase C^k tal que h es la identidad fuera de A. Por otro lado si uno asume que E es un espacio de Banach infinito-dimensional con una base de Schauder entonces E tiene funciones meseta C^k si y sólo si para cada subconjunto compacto K y cada subconjunto abierto U de E que incluye a K, existe un difeomorfismo C^k $h : E \to E \setminus K$ tal que h es la identidad en $E \setminus U$.

Esta clase de resultados sobre eliminación topológica ha encontrado muchas aplicaciones interesantes en varias ramas de las matemáticas, que incluyen la teoría de puntos fijo, la clasificación topológica diferenciable de cuerpos convexos, fenómenos extraños relacionados con las ecuaciones diferenciales ordinarias y sistemas dinámicos en dimensiones infinitas, el fallo del teorema de Rolle en dimensión infinita y muchas más cosas. Ver [13, 14, 15, 25, 35, 12, 16, 87] y las referencias en ellas.

Aunque quizás la aplicación que más nos interesa, y cabe mencionar para nuestro propósitos, fue la aplicación del teorema de West [150] para probar el resultado de Azagra y Cepedello de que las funciones continuas definidas en un espacio de Hilbert separable tomando valores en $\mathbb{R}^m$ se pueden aproximar uniformemente por funciones C^∞ sin ningún punto crítico (ver [12]). Realmente aquí reside la idea de utilizar resultados de extracción difeomorfa de conjuntos cerrados en espacios de Banach para probar nuestros teoremas de Morse-Sard aproximados del Capítulo 4 y 5.

Volvemos ahora al trabajo pionero de esta tesis. Nosotros, como ya se dijo, queremos retomar la propiedad en los difeomorfismos extractores de limitar cualquier recubrimiento dado, que había sido dejada de lado por Azagra, Dobrowolski y Montesinos. En los Capítulos 4 y 5 nosotros probaremos lo siguiente.

Teorema 9. *Sea E un espacio de Banach, $p \in \mathbb{N} \cup \{\infty\}$, y $X \subset E$ un conjunto cerrado con la propiedad de que, para cada $x \in X$, existe un entorno U_x de x en E, espacios de Banach $E_{(1,x)}$ y $E_{(2,x)}$, y una aplicación continua $f_x : C_x \to E_{(2,x)}$, donde C_x es un subconjunto cerrado de $E_{(1,x)}$, tal que:*

1. $E = E_{(1,x)} \oplus E_{(2,x)}$;
2. $E_{(1,x)}$ *tiene particiones de la unidad C^k;*
3. $E_{(2,x)}$ *es infinito-dimensional y tiene (no necesariamente equivalente) una norma de clase C^k;*
4. $X \cap U_x \subset G(f_x)$, *donde*

$$G(f_x) = \{y = (y_1, y_2) \in E_{(1,x)} \oplus E_{(2,x)} : y_2 = f_x(y_1), y_1 \in C_x\}.$$

Entonces, para todo recubrimiento abierto $\mathcal{G}$ de E y todo subconjunto abierto U de E, existe un difeomorfismo C^k h de $E \setminus X$ sobre $E \setminus (X \setminus U)$ que es la identidad en $(E \setminus U) \setminus X$ y está limitado por $\mathcal{G}$. Además, la misma conclusión es cierta si reemplazamos E por un subconjunto abierto de E.

En su demostración combinaremos técnicas y métodos de los trabajos de West [150], Renz [134, 135] y Azagra y Dobrowolski [10, 14]. Observar que los conjuntos compactos en espacios de Banach que poseen base incondicional se pueden ver localmente como gráficas de conjuntos cerrados definidos en un subespacio infinito-codimensional y tomando valores en su complentario lineal. La demostración de este hecho, en el que es clave un teorema de Corson de [55], se incluye en la Sección 3.6 y hace que nuestro Teorema 9 generalice los teoremas de Renz y West.

Por último, con vistas a su utilización en el Teorema 8 del Capítulo 5 necesitaremos también la siguiente variante del teorema anterior, en la que pedimos que el conjunto extractible esté localmente contenido en la unión finita de subespacios cerrados de codimensión infinita.

Teorema 10. *Sea E un espacio de Banach con una norma de clase C^k. Tomemos un recubrimiento abierto $\mathcal{G}$ de un conjunto abierto U y un conjunto cerrado $X \subset U$ tal que para cada $x \in X$, existen un entorno abierto U_x de x y subespacios cerrados $E_1, \ldots, E_{n_x} \subset E$ complementados en E y de codimensión infinita tales que*

$$X \cap U_x \subset \bigcup_{j=1}^{n_x} E_j.$$

Entonces existe un difemorfismo C^k $h : E \to E \setminus X$ que es la identidad fuera de U y está limitado por $\mathcal{G}$.

El Teorema 9, que fue un trabajo conjunto con Azagra y Dobrowolski, se encuentra publicado en [16, Section 2] y el Teorema 10 puede encontrarse en [87, Section 2].

Introduction

In 1934 Hassler Whitney proves his famous extension theorem, which gives necessary and sufficient conditions so that a continuous function defined on an arbitrary closed set C of $\mathbb{R}^n$, together with a set of $j-$linear forms $j = 1, \ldots, k$, can be extended to a C^k function defined in all $\mathbb{R}^n$ whose derivatives in C up to order k are those prefixed $j-$linear forms (see [151]). With the help of this result, in 1935 he built a function $f : \mathbb{R}^2 \to \mathbb{R}$ of class C^1 such that its set of critical points had positive measure (see [152]). The conditions that allow such an example to occur, in which the images of critical sets could have positive measure, were poorly understood at that time. In the next years a more complete vision of the problem was obtained.

Morse-Sard Theorem: *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is a function of class C^k, where $k \geq \max\{1, n - m + 1\}$, then its set of critical points $C_f := \{x \in \mathbb{R}^n : rank\, Df(x) \text{ is not maximum}\}$ satisfies that $\mathcal{L}^m(f(C_f)) = 0$, where $\mathcal{L}^m$ denotes the Lebesgue measure on $\mathbb{R}^m$.*

The set $f(C_f)$ is called set of critical values. We say that a differentiable function $f : \mathbb{R}^n \to \mathbb{R}^m$ satisfies the Morse-Sard property if $\mathcal{L}^m(f(C_f)) = 0$. This is a result with numerous applications in various branches of mathematics as differential topology, dynamical systems, partial differential equations or degree theory. This important theorem is the motivation of the present work.

The structure of the chapters will be the following one. In Chapter 1 we provide a version of the classical Morse-Sard theorem obtained via approximate differentiability. The Morse-Sard property has relations with the Lusin property, and following this line in Chapter 2 it is proved that subdifferentiable functions have the Lusin property of class C^1 and C^2. Chapters 4 and 5 of the thesis derive from the attempt to consider versions of the Morse-Sard theorem in the context of Banach spaces of infinite dimensions, for which a key tool is the diffeomorphic extraction of closed sets in Banach spaces, a topic that is treated in Chapter 3. We will give more precisions later.

The classical Morse-Sard theorem was first proven by Anthony P. Morse for the case of functions with real values in 1939 and later in 1942 was proven by Arthur Sard for the case of vector values (see [125] and [140] respectively). Thanks to the famous Whitney's counterexample of 1934 it is known that this result is optimum within the scale of spaces C^j. The construction of Whitney's counterexample can be generalized to ensure the existence of functions $f : \mathbb{R}^n \to \mathbb{R}^m$, with $m < n$ and of class C^{n-m} such that $\mathcal{L}^m(f(C_f)) > 0$. We also highlight Kaufman's paper [108] which gives a method to build surjective functions $f : [0,1]^{n+1} \to [0,1]^n$ of class C^1 such that $rank(Df(x)) = 1$ for all $x \in [0,1]^{n+1}$.

It is natural to ask whether or not the Morse-Sard property holds for functions lying between C^{k-1} and C^k (with $k = n - m + 1$). A class of intermediate spaces between C^{k-1} and C^k are Hölder spaces $C^{k-1,t}$, with $t \in (0,1]$, that are defined as the set of all functions f of class C^{k-1} such that there exists $M > 0$ so that $||D^{k-1}f(x) - D^{k-1}f(y)|| \leq M|x-y|^t$, for all $x, y \in \mathbb{R}^n$. In 1986 Norton studied the Morse-Sard properties that could have this kind of functions and in particular he proved that if $f : \mathbb{R}^n \to \mathbb{R}^m$ is of class $C^{k,t}$ and C_f has $\mathcal{H}^{k+t+m-1}-$null Hausdorff measure, then $\mathcal{L}^m(f(C_f)) = 0$ (see [130, Theorem 1 and 2]). A few years later, Bates in 1993 proved in [31] that functions $f : \mathbb{R}^n \to \mathbb{R}^m$ belonging to the class $C^{k-1,1}$, with $k \geq \max\{1, n - m + 1\}$ satisfy the Morse-Sard property, although in general this fails for $C^{k-1,t}$ with $t < 1$ (see [30] or [126]).

On the other hand De Pascale gave in 2001, [58], the first generalization of the Morse-Sard theorem for Sobolev spaces, in which integrability properties of the functions (such as for instance Morrey's inequality) come to play. He proved that if $f \in W^{k,p}_{loc}(\mathbb{R}^n;\mathbb{R}^m)$ with $n > m$, $k = n - m + 1$ and $p > n$ then $\mathcal{L}^m(f(C_f)) = 0$. Take into account that by the Sobolev embedding theorem, $W^{k,p}_{loc} \hookrightarrow C^{k-1,1-\frac{n}{p}}$ (identifying each function f with its precise representative[6]), and since $k \geq 2$ if $n > m$, the critical set C_f can be defined in the usual way because f is at least of class C^1. De Pascale's Morse-Sard theorem was revised by Figalli in 2008, who gives a different proof, a bit simpler and with the advantage of being independent of the classical Morse-Sard theorem (see [82]).

Much more recently, Bourgain, Korobkov and Kristensen ([44], 2015) give a version of the theorem for the spaces $W^{n,1}(\mathbb{R}^n;\mathbb{R})$ and $BV_n(\mathbb{R}^n;\mathbb{R})$ using totally different and more sophisticated techniques (see also [43] for the case $n = 2$). The latter space is the one formed by functions whose distributional derivatives of order n are Radon measures in $\mathbb{R}^n$. Thanks to [65] it is known that such functions admit a continuous representative which is differentiable in $\mathcal{H}^1-$almost every point. On the other hand Bourgain, Korobkov and Kristensen also established a Lusin $N-$property for these functions stating that $\mathcal{L}^1(f(E)) = 0$ provided that $\mathcal{H}^1(E) = 0$. Then the set of points where f is not differentiable and hence C_f cannot be defined is sent to a set of null measure. In this line of research, the most recent (and possibly optimal) results are due to Korobkov, Kristensen and Hajlasz. These authors prove in [95] the theorem for the case of functions in the Sobolev-Lorentz space $W_1^{k,\frac{n}{k}}(\mathbb{R}^n;\mathbb{R}^m)$, $k = n-m+1$, $m \leq n$, formed by functions from the Sobolev space $W^{k,\frac{n}{k}}$ and whose distributional derivatives of order k in addition are in the Lorentz space $L_{\frac{n}{k},1}$ (see the papers [113, 114, 95] for more details). Observe that the case $m = 1$, $W_1^{n,1}(\mathbb{R}^n;\mathbb{R}) = W^{n,1}(\mathbb{R}^n;\mathbb{R})$ corresponds with the one treated by Bourgain, Korobkov and Kristensen in [44]. Theses spaces possess the minimal integrability conditions that assure the continuity of the mappings (always assuming that we are working with the precise representative). However these mappings need not be differentiable everywhere, and this can create problems when defining the set of critical points. Luckily it is known that the functions $f \in W_1^{k,\frac{n}{k}}(\mathbb{R}^n;\mathbb{R}^m)$ are differentiable $\mathcal{H}^{\frac{n}{k}}-$almost everywhere (see [114, Theorem 2.2]) and, thanks to a Lusin $N-$property, the image of this set of *bad points*, call it A_f, has measure zero. This Lusin $N-$property establishes that if $f \in W_1^{k,\frac{n}{k}}(\mathbb{R}^n;\mathbb{R}^m)$, $1 \leq k \leq n$ and $E \subset \mathbb{R}^n$ is such that $\mathcal{H}^s(E) = 0$ then $\mathcal{H}^s(f(E)) = 0$, for $s \in [\frac{n}{k}, n]$ (see [95, Theorem 2.3]). The set of critical points is defined then as $C_f := \{x \in \mathbb{R}^n \setminus A_f : rankDf(x) \text{ is not maximum}\}$. This interest to know that the set of bad points where we cannot define the differential is sent to a set of null measure is partially due to the fact that one of the more direct consequences of the Morse-Sard theorem is to obtain the C^1-regularity of almost all level sets.

In the field of fluid mechanics, an example of the powerful applications of this kind of refinements of the Morse-Sard theorem for Sobolev spaces can be seen in the paper of Korobkov, Pileckas and Russo from 2015, [112], that gives a solution to the so-called Leray's problem for the steady Navier-Stokes system.

Apart from these refinements for Sobolev functions, we can find recent versions of the Morse-Sard theorem in many other contexts. For example Barbet, Dambrine and Daniilidis prove in [28], and later with Rifford in [29], that for a special subclass of Lipschitz functions $f : \mathbb{R}^n \to \mathbb{R}^m$, defined through a finite selection, or infinite indexed on a countable compactum, of C^k smooth functions with $k \geq n - m + 1$, the set of Clarke critical values (a weaker concept than the usual one of critical value) has null measure. It is a non-smooth version of the Morse-Sard theorem.

On the other hand in 2018 Azagra, Ferrera and Gómez-Gil prove that in the case that $f \in C^{n-m}(\mathbb{R}^n;\mathbb{R}^m)$, $m \leq n$, satisfies

$$\limsup_{y \to x} \frac{|D^{n-m}f(y) - D^{n-m}f(x)|}{|y-x|} < \infty$$

for all $x \in \mathbb{R}^n$ (this is that $D^{n-m}f$ is a Stepanov function), then $\mathcal{L}^m(f(C_f)) = 0$. If $n = m$ we would

[6] If $f \in L^1_{loc}(\mathbb{R}^n;\mathbb{R}^m)$, then the precise representative of f is defined as

$$f^*(x) := \begin{cases} \lim_{r\to 0} \frac{1}{\mathcal{L}^n(B(x,r))} \int_{B(x,r)} f(y)\,dy & \text{if the limits exists} \\ 0 & \text{elsewhere.} \end{cases}$$

have that f is differentiable almost everywhere, but at the same time f would send $\mathcal{L}^n-$null sets to $\mathcal{L}^n-$null sets, then the set of non-differentiability points would not worry us. The authors deduced this result from an abstract and more powerful theorem of Morse-Sard type, namely [18, Theorem 1.5]. This theorem allows them to recover De Pascale's result too. And practically at the same time, in [17], these three authors also found versions of the Morse-Sard theorem for subdifferentiable functions $f : \mathbb{R}^n \to \mathbb{R}$ and Taylor expansions of order $n-1$ (this fact was one of the germs of the results developed in Chapter 2 of this dissertation, where it is shown that subdifferentiable functions also have the Lusin property of class C^1 and C^2).

In Chapter 1 of this thesis we make a humble contribution to all these Morse-Sard theorems. To understand the motivation of our results one has to observe that the functions $f : \mathbb{R}^n \to \mathbb{R}^m$ belonging to any space previously commented, which have been proved to have the Morse-Sard property, also turn out to satisfy the Lusin property of class C^k, with $k = n - m + 1$. This means that given $\varepsilon > 0$, there exists a function $g : \mathbb{R}^n \to \mathbb{R}^m$ of class C^k such that $\mathcal{L}^n(\{x \in \mathbb{R}^n : \ f(x) \neq g(x)\}) < \varepsilon$. If $k = 0$, C^0 denotes the space of continuous functions. In [44, 113, 114, 95] the Lusin property and the Morse-Sard property are studied simultaneously and not only do you get that the class of functions considered in those papers have the Lusin property of the correspondent class, but also that this property is essential in their proofs of the Morse-Sard theorem.

In order to find good characterizations of the Lusin property it is necessary to work with other concept of limit, different to the usual one, called approximate limit. We will say that $\operatorname{ap}\lim_{y\to x} f(y) = L$ if there exists a set A_x with density one at x [7] such that $\lim_{y\to x, y\in A_x} f(y) = L$ (the same for $\limsup$ and $\liminf$). If $L = f(x)$ we say that f is approximately continuous at x. One of the main reasons for introducing this concept is that one has the equivalence between approximately continuous functions at almost every point, measurable functions and functions that satisfy the Lusin property of class C^0. In the same way one can define the approximate differentiability of order $k \geq 1$ at a point x if there exists a polynomial of order k centered at x, $p_k(x; y)$, with $p_k(x; x) = f(x)$, such that

$$\operatorname{ap}\lim_{y\to x} \frac{|f(y) - p_k(x;y)|}{|y-x|^k} = 0.$$

And we will say that f has a $(k-1)-$approximate Taylor polynomial at x if there exists a polynomial of order $k-1$ centered at x, $p_{k-1}(x; y)$, with $p_{k-1}(x; x) = f(x)$, such that

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y) - p_{k-1}(x;y)|}{|y-x|^k} = 0.$$

Approximate differentiability is a weaker property than possession of Taylor polynomials and than usual regularity of class C^k. It turns out that thanks to the independent results of Isakov [102] and Liu and Tai [121], the Lusin property of class C^k is equivalent to being approximate differentiable of order k at almost every point $x \in \mathbb{R}^n$ (and also equivalent to having $(k-1)-$approximate Taylor polynomials at almost every point).

In view of these facts we may ask ourselves whether it is possible to prove versions of the Morse-Sard theorem for functions $f : \mathbb{R}^n \to \mathbb{R}^m$ that we simply assume are approximately differentiable of order $k = n - m + 1$ at *enough* points.

In Chapter 1 we show how, by combining some of the techniques and strategies that are common along the literature with the idea of the proof of [121, Theorem 1] and an induction argument, one can prove the following non-smooth Morse-Sard type result.

Theorem 1. *Let $n \geq m$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Borel function. Suppose f is approximately differentiable of order* 1 *at $\mathcal{H}^m$-almost every point and satisfies*

[7] A set $A \subset \mathbb{R}^n$ is said to have density one at $x \in \mathbb{R}^n$ if

$$\lim_{r\to 0} \frac{\mathcal{L}^n(B(x,r)\cap A)}{\mathcal{L}^n(B(x,r))} = 1.$$

(a) *ap* $\limsup_{y\to x} \frac{|f(y)-f(x)|}{|y-x|} < +\infty$ *for all* $x \in \mathbb{R}^n \setminus N_0$*, where* N_0 *is a countable set, and*

(b) *(if* $n > m$*)* *ap* $\lim_{y\to x} \frac{|f(y)-f(x)-\cdots \frac{F_i(x)}{i!}(y-x)^i|}{|y-x|^i} = 0$ *for all* $i = 2, \dots, n-m+1$ *and for all* $x \in \mathbb{R}^n \setminus N_i$*, where each set* N_i *is* $(i+m-2)$*-*σ*-finite and the coefficients* $F_i(x)$ *(*i*-multilinear and symmetric mappings) are Borel functions,*

then f *has the Morse-Sard property (that is, the image of the critical set of points of* f *is null with respect to the Lebesgue measure in* $\mathbb{R}^m$*).*

Though it is possible that f is not differentiable at some points, if we denote by AppDiff(f) the set of points where f is approximately differentiable, we understand as its critical set of points $C_f = \{x \in \text{AppDiff}(f) : \text{rank}(F_1(x)) \text{ is not maximum}\}$. Moreover we say that a set is s-σ-finite if it can be written as a countable union of sets of finite $\mathcal{H}^s-$measure. We include another version of this theorem in which we do not suppose the Borel measurability of the function f.

Theorem 2. *Let* $f : \mathbb{R}^n \to \mathbb{R}^m$*,* $m \leq n$ *so that for all* $x \in \mathbb{R}^n$ *and* $i = 1 \dots, n-m$ *there exist* $i-$*multilinear and symmetric mappings* $F_i(x)$ *such that*

(a) $\text{ap}\limsup_{y\to x} \frac{|f(y)-f(x)|}{|y-x|} < +\infty$ *for all* $x \in \mathbb{R}^n \setminus N_0$*, where* N_0 *is a countable set.*

(b) *The set* $N_1 := \{x \in \mathbb{R}^n : \text{ap}\limsup_{y\to x} \frac{|f(y)-f(x)-F_1(x)(y-x)|}{|y-x|^2} = +\infty\}$ *is* $\mathcal{H}^m-$*null.*

(c) *For each* $i = 2, \dots, n-m$ *the set*

$$N_i := \{x \in \mathbb{R}^n : \text{ap}\limsup_{y\to x} \frac{\left|f(y) - f(x) - \sum_{j=1}^{i} \frac{F_j(x)}{j!}(y-x)^j\right|}{|y-x|^{i+1}} = +\infty\}$$

is countably $(\mathcal{H}^{i+m-1}, i+m-1)$ *rectifiable of class* C^i*.*

Then $\mathcal{L}^m(f(C_f)) = 0$*.*

A set $N \subset \mathbb{R}^n$ is said to be countably $(\mathcal{H}^s, s)$ rectifiable of certain class C^k if there exists a countable set of $s-$dimensional submanifolds A_j of class C^k such that $\mathcal{H}^s(N \setminus \bigcup_{j=1}^\infty A_j) = 0$. Both Theorems 1 and 2 remain valid if we replace $\mathbb{R}^n$ with an open set U of $\mathbb{R}^n$. In Theorem 1.23 of Chapter 1 we also offer an interesting variant of this result.

Theorems 1 and 2 generalize the versions of the Morse-Sard theorem given by Bates and the Appendix of [17]. Moreover they are not stronger, nor weaker, than the versions of [44, 113, 114] for BV_n or Sobolev functions with smaller exponents; see the last Section 1.4 of Chapter 1 for more examples and comments. In particular, thanks to Norton's works, we include an example showing that our results are sharp in the sense that there exist functions $f : \mathbb{R}^n \to \mathbb{R}$ of class $C^{n-1,t}$, $t \in (0,1)$ that have an $(n-1)$-approximate Taylor polynomial everywhere except on a set of $\mathcal{H}^{n-1+t}-$null measure and do not satisfy the Morse-Sard property.

The results of Chapter 1 are published in [21].

The comments made so far show therefore the relation between the Morse-Sard theorem and the Lusin property. This led the author of this dissertation to a deeper study of this second property, which thanks to a fruitful collaboration with the mathematicians Azagra, Ferrera and Gomez-Gil, resulted in the proof that subdifferentiable functions satisfy the Lusin property of class C^1 and C^2. This will appear in Chapter 2 of the thesis. But before stating our discoveries let us comment a little more in depth the history of the Lusin property, which in many cases goes parallel to that of the Morse-Sard theorem.

The classical Lusin theorem, from 1912, says that given a Lebesgue measurable set $\Omega \subseteq \mathbb{R}^n$, given a Lebesgue measurable function $f : \Omega \to \mathbb{R}$ and given $\varepsilon > 0$, there exists $g : \Omega \to \mathbb{R}$ continuous such that $\mathcal{L}^1(\{x \in \Omega : f(x) \neq g(x)\}) < \varepsilon$. Thereafter a number of authors have proved that if more regularity properties on f are assumed, the approximating function g can be taken of class C^k for some $k \geq 1$. We will say that $f : \Omega \to \mathbb{R}$ satisfies the Lusin property of class C^k, $k \geq 0$, if for each $\varepsilon > 0$ there exists $g \in C^k(\Omega; \mathbb{R})$ such that $\mathcal{L}^n(\{x \in \Omega : \ f(x) \neq g(x)\}) < \varepsilon$.

After Lusin's result, the first relevant extension was due to H. Federer in 1944 (proved implicitly in [73, p. 442]). He showed that if f was differentiable almost everywhere, then f had the Lusin property of class C^1. In particular Lipschitz functions and functions from the Sobolev space $W^{1,p}(\mathbb{R}^n; \mathbb{R})$ with $n < p \leq \infty$ have the Lusin property of class C^1. This result was a simple consequence of the Whitney's extension theorem, which will remain a key piece in all subsequent generalizations and versions of the Lusin property. Observe that the condition about being differentiable almost everywhere is equivalent to $\limsup_{y \to x} \frac{|f(y)-f(x)|}{|y-x|} < \infty$ for almost every point.

We move now until 1951, where Whitney in [153] proves that a function $f : \Omega \to \mathbb{R}$ satisfies the Lusin property of class C^1 if and only if it has approximate partial derivatives almost everywhere. And we would have to wait until 1994, moment at which Liu and Tai get to prove that a function has the Lusin property of class C^k, for any $k \geq 1$, if and only if it is approximately differentiable of order k almost everywhere (and if and only if it has $(k-1)-$approximate Taylor polynomials almost everywhere).

Apart from this beautiful characterization of the Lusin property there is the question of which function spaces satisfy it for certain class C^k.

About the possible generalizations of the Lusin theorem for Sobolev functions, since 1933 it is known that all functions from $W^{1,1}(\mathbb{R}^n)$ have partial derivatives almost everywhere (Nikodym [129], or also see [71, Lemma 4.9.2]), hence using the previous Whitney's result we automatically have the Lusin property of class C^1. The generalization for the case $W^{k,p}(\mathbb{R}^n)$, $1 < p < \infty$, is due to Calderón and Zygmund [49], who in 1961 proved that this space had the Lusin property of class C^k. Later on Liu [120] adds Calderón and Zygmund's theorem the property that if $f \in W^{k,p}(\mathbb{R}^n)$ and $g \in C^k(\mathbb{R}^n)$ is the approximating function, then g also approximates f in the Sobolev norm. A stronger version appeared with the work of Michael and Ziemer of 1984 published in [124], where the exceptional sets are taken in terms of Bessel capacities (see [155, Section 2.6]) for definitions): If $1 < p < \infty$, $l = 0, 1 \ldots, k$, $\varepsilon > 0$ and $f \in W^{k,p}_{loc}(\mathbb{R}^n)$, then there exists $g \in C^l(\mathbb{R}^n)$ such that $B_{k-l,p}(\{x \in \mathbb{R}^n : \ f(x) \neq g(x) \text{ or } D^j f(x) \neq D^j g(x) \text{ for some}; j = 1, \ldots, l\}) < \varepsilon$ and $||f - g||_{l,p} \leq \varepsilon$. For $k = l$, the Bessel capacity $B_{0,l}$ coincides with the Lebesgue measure and hence the previous Liu's result is recovered. The reader must observe that when $l = 0, 1, \ldots, k$ decreases, the *exceptional* set gets smaller because in general if $0 \leq a < b$ then $B_{a,p}(C) \geq B_{b,p}(C)$ for all $C \subset \mathbb{R}^n$. Bojarski, Hajlasz and Strzelecki in 2002 extend the result of Michael and Ziemer by being able to take the approximation in the Sobolev space of a higher order, that is $g \in W^{l+1,p}$ which also approximates f in the norm of such space, for $l = 0, 1, \ldots, k-1$. Later on appeared the results of Bourgain, Korobkov and Kristensen, who consider the spaces $W^{k,1}$ and BV_k and where the exceptional sets now have small Hausdorff content $\mathcal{H}^l_\infty$ (for more details see [44, Theorem 2.1 and 6.2]). Finally Korobkov and Kristensen in [114, Theorem 2.1] prove the following Lusin type approximation result for Sobolev-Lorentz spaces: let $k, l \in \{1, \ldots, n\}$, $l \leq k$, $p \in (1, \infty)$ and $f \in W^{k,p}_1(\mathbb{R}^n; \mathbb{R}^m)$, then for all $\varepsilon > 0$ there exists an open set $U \subset \mathbb{R}^n$ and a function $g \in C^l$ such that $\mathcal{H}^{n-(k-l)p}_\infty(U) < \varepsilon$ and $f = g$, $D^j f = D^j g$ in $\mathbb{R}^n \setminus U$ for $j = 1, \ldots, l$, where $\mathcal{H}^t_\infty$ denotes the Hausdorff content. This property is essential to prove the results appearing in [95].

For the special class of convex functions $f : \mathbb{R}^n \to \mathbb{R}$, Alberti and Imonkulov [2, 101] showed that every convex function has the Lusin property of class C^2 (being the approximating function not necessarily convex); see also [1] for a related problem. More recently Azagra and Hajlasz [23] have proved that the approximation in the Lusin sense can be taken of class $C^{1,1}_{loc}$ and convex if and only if f is essentially coercive (in the sense that $\lim_{|x| \to \infty} f(x) - l(x) = \infty$ for some linear function l) or else f is already of class $C^{1,1}_{loc}$ (in such a case taking f itself as the approximation is the only choice).

In Chapter 2 of this thesis we are interested in the possible versions of this result making use of the notion of subdifferential. It will be proven that subdifferentiable functions satisfy Lusin properties of class C^1 or C^2. By subdifferentials we mean the Fréchet subdifferential, the proximal subdifferential, or the viscosity subdifferential of second order. We remind the reader that the Fréchet subdifferential of a function $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ is defined as the set of $\xi \in \mathbb{R}^n$ such that $\liminf_{y\to x} \frac{f(y)-f(x)-\langle \xi, y\rangle}{|y-x|} \geq 0$, and that the proximal subdifferential of f at x is defined as the set of $\xi \in \mathbb{R}^n$ such that $\liminf_{y\to x} \frac{f(y)-f(x)-\langle \xi, y\rangle}{|y-x|^2} > -\infty$. The question in its more general form that we intend to solve is: given $k \in \mathbb{N}$ and a function $f : \mathbb{R}^n \to \mathbb{R}$, suppose that for almost every $x \in \mathbb{R}^n$ there exists a polynomial $p_{k-1}(x; y)$ centered at x, with $p_{k-1}(x; x) = f(x)$ and degree less than or equal to $k-1$ such that

$$\liminf_{y\to x} \frac{f(y) - p_{k-1}(x; y)}{|y - x|^k} > -\infty.$$

Does it follow that f has the Lusin property of class C^k? As we have already said, in Chapter 2 we will give an affirmative answer for the case $k = 1, 2$, but negative for $k \geq 3$. All this could have applications in nonsmooth analysis or in the theory of viscosity solutions to PDE such as Hamilton-Jacobi equations.

The precise statement for $k = 1$ is:

Theorem 3. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, and $f : \Omega \to \mathbb{R}$ a function. Assume that for almost every $x \in \Omega$ we have*

$$\liminf_{y\to x} \frac{f(y) - f(x)}{|y - x|} > -\infty.$$

Then, for all $\varepsilon > 0$ there exists a function $g \in C^1(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is a measurable function and $\Omega = \{x \in \mathbb{R}^n : \partial^- f(x) \neq \emptyset\}$ is the set of points where the Fréchet subdifferential of f is not empty, then it follows from Theorem 3 that for each $\varepsilon > 0$ there exists a function $g \in C^1(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

For the case of C^2 smoothness:

Theorem 4. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, and $f : \Omega \to \mathbb{R}$ be a function such that for almost every $x \in \Omega$ there exists a vector $\xi_x \in \mathbb{R}^n$ such that*

$$\liminf_{y\to x} \frac{f(y) - f(x) - \langle \xi_x, y - x\rangle}{|y - x|^2} > -\infty.$$

Then for all $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is a measurable function and $\Omega = \{x \in \mathbb{R}^n : \partial_P f(x) \neq \emptyset\}$ is the set of points where the proximal subdifferential of f is not empty, then it follows from Theorem 4 that for each $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that

$$\mathcal{L}^n(\{x \in \Omega : f(x) \neq g(x)\}) \leq \varepsilon.$$

In Section 2.4 we offer some examples showing that this kind of results are no longer true for *Taylor subexpansions* of greater orders. One of the main reasons why this does not work for $k \geq 3$ lies in the fact that having nonempty subdifferentials of order 2 implies that the subdifferentials of order $k \geq 3$ are not empty either (see for instance [17, Proposition 1.1]).

All the results of Chapter 2 can be found published in [19].

Let us now drastically change the context. Everything developed in Chapters 3, 4 and 5 is in infinite-dimensional spaces. In Chapter 3 we establish new results about diffeomorphic extraction of some closed sets in infinite-dimensional Banach spaces. These results are a key piece for our proofs of the most important theorems from Chapters 4 and 5 of the thesis. It is preferable to expose first the motivations and the historical context that led us to develop these last two chapters, and leave until the end of this introduction everything concerning the diffeomorphic extraction of closed sets in Banach spaces.

After noting the importance of the classical Morse-Sard theorem in the literature one cannot help but wonder what happens if we look for analogues of the theorem for functions $f : M \to N$ acting between infinite-dimensional manifolds?. Which conditions of smoothness must we impose to the functions for their critical set of points $f(C_f)$ to be small in some sense? Here we define the set of critical points by $C_f := \{x \in M : Df(x) \text{ is not a surjective operator}\}$. In general M and N will denote Banach spaces or else as manifolds modelled in Banach spaces.

This question was firstly studied by Smale in 1965, who in [142] proved that if M and N are separable manifolds, connected and C^∞, modelled in Banach spaces, and $f : M \to N$ is a C^r Fredholm mapping (that is to say, every differential $Df(x)$ is a Fredholm operator between the corresponding tangent spaces) then $f(C_f)$ is a set of first category, and in particular $f(C_f)$ has no interior points, provided that $r > \max\{\text{index}(Df(x)), 0\}$ for all $x \in M$; here $\text{index}(Df(x))$ means the index of the Fredholm operator $Df(x)$, which is the difference between the dimensions of the kernel of $Df(x)$ and the codimension of the image of $Df(x)$, both of which are finite. Of course, these assumptions are very restrictive as, for instance, if E is infinite-dimensional then no function $f : M \to \mathbb{R}$ is Fredholm.

In general, every attempt to adapt the Morse-Sard theorem to infinite dimensions will have to impose vast restrictions because, as shown by a Kupka's counterexample [116], there are functions $f : \ell_2 \to \mathbb{R}$ of class C^∞ such that their set of critical values $f(C_f)$ contain intervals. As Bates and Moreira showed in [32], one can even make f be a polynomial of degree 3. Namely one can take the function

$$f(\sum_{n=1}^{\infty} x_n e_n) = \sum_{n=1}^{\infty} (3 \cdot 2^{-\frac{n}{3}} x_n^2 - 2x_n^3),$$

whose set of critical points is $C_f = \{\sum_{n=1}^{\infty} x_n e_n : x_n \in \{0, 2^{-\frac{n}{3}}\}\}$ and such that $\mathcal{H}^3(C_f) < +\infty$ and $f(C_f) = [0,1]$.

Luckily, for many applications of the Morse-Sard theorem, it is often enough to know that any given continuous mapping can be uniformly approximated by a mapping whose set of critical values is small in some sense; therefore it is natural to ask what mappings between what infinite-dimensional manifolds will at least have such an approximation property. Going in this direction, Eells and McAlpin established the following theorem [69]: If E is a separable Hilbert space, then every continuous function from E into $\mathbb{R}$ can be uniformly approximated by a smooth function f whose set of critical values $f(C_f)$ is of measure zero. This allowed them to deduce a version of this theorem for mappings between smooth manifolds M and N modelled on E and a Banach space F respectively, which they called an *approximate Morse-Sard theorem*: Every continuous mapping from M into N can be uniformly approximated by a smooth mapping $f : M \longrightarrow N$ so that $f(C_f)$ has empty interior. Nevertheless, as observed in [69, Remark 3A], we have $C_f = M$ in the case that F is infinite-dimensional (so, even though the set of critical values of f is relatively small, the set of critical points of f is huge, which is somewhat disappointing). A similar contemporary result gets also the approximations of the derivatives, provided that the initial function is already of class C^1. This is a result of Moulis [128, p. 331] stating that: for every C^1 function $f : E \to F$, where E is a separable infinite-dimensional Hilbert space and F is a separable Hilbert space, and for every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^∞ function $g : E \to F$ such that $||f(x) - g(x)|| \leq \varepsilon(x)$, $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ for all $x \in E$ and such that $g(C_g)$ has empty interior in F.

In [12], a much stronger result was obtained by M. Cepedello-Boiso and D. Azagra: if M is a C^∞ smooth manifold modelled on a separable infinite-dimensional Hilbert space X, then every continuous mapping from M into $\mathbb{R}^m$ can be uniformly approximated by smooth mappings *with no critical points*. This kind of statement is the strongest one can get in the context of approximate Morse-Sard theorems. Unfortunately, since part of the proof requires the use of the good properties of the Hilbertian norm, this cannot be extended in a direct way to other Banach spaces. On the other hand an important part of its proof requires the possibility to find diffeomorphisms $h : E \to E \setminus X$, where X is a closed locally compact set, and that are arbitrarily close to the identity, specifically it is used [150, Theorem 1]. It is convenient to retain this idea, since it will be again a key for the development of Chapters 3, 4 and 5 of this dissertation.

P. Hájek and M. Johanis [91] established a similar result for $m = 1$ in the case that X is a separable Banach space which contains c_0 and admits a C^k-smooth bump function (by a bump function we mean a continuous function $\lambda : E \to [0, \infty)$ whose support supp$(\lambda) = \overline{\{x \in E : \lambda(x) \neq 0\}}$ is bounded). In this case the approximating functions are of class C^k. The method employed by Hájek and Johannis is based in the result that real-valued functions of class C^∞ defined on c_0 and that locally depend on finitely many coordinates [8] are dense in the space of continuous functions defined on c_0 with real values (see [60]). However, the authors already pointed out that their method does not apply when the space E is reflexive, leaving out hence the classical Banach spaces ℓ_p and L^p, $1 < p < \infty$. A bit later, Azagra and Mar Jiménez in 2007 characterized the class of separable Banach spaces E for which for every continuous function $f : E \to \mathbb{R}$ and $\varepsilon : E \to (0, \infty)$ there exists a C^1 function $g : E \to \mathbb{R}$ such that $||f(x) - g(x)|| \leq \varepsilon(x)$ and $g'(x) \neq 0$ for all $x \in E$, as those infinite-dimensional Banach spaces with separable dual (see [25]).

So far, there are some *good* consequences of these results, all of them related somehow with Morse-Sard type theorems. But there also some *bad* consequences. One of them, since the set of smooth functions without critical points is dense in the set of continuous functions, is that there are quite big sets of functions for which no conceivable Morse-Sard theory is valid.

Between the *good* consequences there are some ones interesting as the existence of a nonlinear Hahn Banach theorem or the construction of examples of functions $f : E \to \mathbb{R}$ that do not satisfy Rolle's theorem and have prefixed support (see [25]). It is well known that Rolle's theorem fails in general in infinite dimensions; the first example of such type is due to Shkarin in [141]. The interested reader can also have a look to the papers [77, 22, 24], although we will comment on that theme in more detail in Section 5.6.

Though it is obvious, it is worth mentioning that in finite dimensions this kind of results of approximating continuous functions by smooth ones without critical points cannot take place. Think for example of $f : \mathbb{R}^n \to \mathbb{R}^m$ defined as $f(x) = (f^1(x), \ldots, f^m(x)) = (|x|, 0, \ldots, 0)$, where $|\cdot|$ denotes the usual euclidean norm. If we want to uniformly approximate f by another g of class C^1, we must have that $g^1(x)$ approximates $|x|$ and therefore g^1 must attain a global minimum at some point x_0 for which we should have $\nabla g^1(x_0) = (0, \ldots, 0)$.

Our concern in the development of Chapter 4 was trying to replace the pairs $(\ell_2, \mathbb{R}^m)$ or $(E, \mathbb{R})$, from the results of Azagra and Cepedello and of Azagra and Mar Jiménez, by other pairs of the form (E, F) where E was a Banach space as general as possible and F is a quotient of E, possibly of infinite dimension. The hypothesis about considering F as a quotient of E is mandatory because if there are no surjective operators from E onto F then all points are critical for all differentiable functions.

One of the keys to the results that we present is that the Banach space E has some "composite" structure. In one case by this we mean that E is isomorphic to its square and in another case that E has unconditional basis.

For spaces E which are reflexive and are isomorphic to its square we have the following result.

Theorem 5. *Let E be a separable reflexive Banach space of infinite dimension, and F a Banach space. In the case that F is infinite-dimensional, let us assume moreover that:*

1. *E is isomorphic to $E \oplus E$.*

[8] We say that a function f defined on a Banach space E locally depends on finitely many of coordinates if for all $x \in E$ there exists a natural number l_x, an open neighbourhood U_x of x, some functionals $L_1, \ldots, L_{l_x} \in E^*$ and a function $\gamma : \mathbb{R}^{l_x} \to \mathbb{R}$ such that

$$f(y) = \gamma(L_1(y), \ldots, L_{l_x}(y))$$

for all $y \in U_x$.

2. *There exists a bounded linear operator from E onto F (equivalently, F is a quotient of E).*

Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^1 mapping $g : E \to F$ such that $||f(x) - g(x)|| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

Take into account that there exists some separable and reflexive Banach spaces E that are not isomorphic to $E \oplus E$. The first example of that type was given by Fiegel in 1972 [81].

For spaces which are not necessarily reflexive but have an appropriate Schauder basis we have the following.

Theorem 6. *Let E be an infinite-dimensional Banach space, and F a Banach space such that:*

(i) E has an equivalent locally uniformly convex norm $\| \cdot \|$ that is C^1.

(ii) $E = (E, \| \cdot \|)$ has a (normalized) Schauder basis $\{e_n\}_{n \in \mathbb{N}}$ such that for every $x = \sum_{j=1}^{\infty} x_j e_j$ and every $j_0 \in \mathbb{N}$ we have

$$\left\| \sum_{j \in \mathbb{N},\, j \neq j_0} x_j e_j \right\| \leq \left\| \sum_{j \in \mathbb{N}} x_j e_j \right\| .$$

(iii) In the case that F is infinite-dimensional, there exists a subset $\mathbb{P}$ of $\mathbb{N}$ such that both $\mathbb{P}$ and $\mathbb{N} \setminus \mathbb{P}$ are infinite and, for each infinite subset J de $\mathbb{P}$, there exists a linear bounded operator from $\overline{\text{span}}\{e_j : j \in J\}$ onto F (equivalently, F is a quotient space of $\overline{\text{span}}\{e_j : j \in J\}$).

Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^1 mapping $g : E \to F$ such that $||f(x) - g(x)|| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

The proofs of these theorems, which were a joint work with Azagra and Dobrowolski, can be found in Chapter 4 and are published in [16].

The second assumption (ii) of Theorem 6 is equivalent to the fact that, for every set $A \subset \mathbb{N}$ (equivalently, every finite set $A \subset \mathbb{N}$) we have $\|P_A\| \leq 1$, where P_A represents the projection $P_A(x) = \sum_{j \in A} x_j e_j$. This, in particular, implies that $\{e_n\}_{n \in N}$ is an uncoditional basis, with suppression unconditional basis K_s equal to 1; for more details see [4, p. 53] or [3].

These results include the classical Banach spaces as c_0, ℓ_p, L^p, $1 < p < \infty$, F being a Banach space such that there exist bounded linear operators from E onto F. In particular the result also applies to Sobolev spaces $W^{k,p}(\mathbb{R}^n)$ with $1 < p < \infty$ because they are isomorphic to $L^p(\mathbb{R}^n)$ (see [132, Theorem 11]). Moreover it will also be true for some less classical spaces as $C(K)$, with K a countable metrizable compact set (in general any isometric predual of ℓ_1) or as the James space J. In the latter case we can take as the target space F any quotient of any infinite-dimensional complemented and reflexive subspace of J.

Furthermore, we can make the approximating functions to be smoother than simply C^1. For that it is necessary to use Nicole Moulis' results about C^1-fine approximation in Banach spaces [128], or the more general results of [93, Corollary 7.96]. It is then possible to uniformly approximate continuous functions $f : E \to F$, where $E = c_0, \ell_p, L^p, 1 < p < \infty$, by C^k functions without critical points, where k denotes the order of regularity of the space E in each case. For example for c_0 or ℓ_p, L^p with p even we get C^∞ smoothness.

A direct application of our theorems is that for all pairs of Banach spaces (E, F) for which one can apply either Theorem 5 or 6, every continuous function $f : E \to F$ can be uniformly approximated by open mappings of class C^k (where k denotes the order of smoothness of the space E).

The proof of Theorems 5 and 6 consists of two parts. Firstly, since the result is invariant by diffeomorphisms, it will be enough to prove it for continuous functions $f : S^+ = \{(u, t) : u \in E, t > 0, ||u||^2 + t^2 = 1\} \to F$ defined in the upper unit sphere of the space $Y = E \times \mathbb{R}$, which we endow with the norm $|(u,t)| = (||u||^2 + t^2)^{1/2}$. This way we define a first uniform approximation $\varphi : S^+ \to F$ of class C^1,

$$\varphi(y) = \sum_{n\in\mathbb{N}} \psi_n(y)(f(y_n) + T_n(y)),$$

where $\{\psi_n\}_{n\in\mathbb{N}}$ are smooth partitions of unity specially constructed on S^+, $y_n \in \text{supp}(\psi_n)$ and where we are locally perturbing by certain linear surjective operators $T_n : S^+ \to F$. This construction is made with extremely care so that $||f(x) - \varphi(x)|| \leq \varepsilon(x)/2$ for $x \in S^+$ and so that the set of critical points C_φ lies inside a diffeomorphically extractible set (this will be a closed set that is locally contained in the graph of a continuous function defined in an infinite-codimensional complemented subspace of E and takes values in its linear complement). In this first step we have to use the separability of the space E, the existence of an equivalent norm of class C^1 and the existence of *a lot* of linear surjective operators $S_n : E \to F$. These operators will be defined from complemented subspaces E_n of infinite codimension in E and so that $E_n \cap E_m = \{0\}$ for all $n \neq m$. This is where we need to be able to decompose the space E in enough complemented subspaces of infinite codimension, that is, to be able to write $E = E_1 \oplus E_2$, with E_1 and E_2 isomorphic to E, or else to have an unconditional basis at our disposal in the space E. Once the first step has been performed it simply remains to diffeomorphically extract the set of critical points of this approximating function φ. We must look for a C^1 diffeomorphism, $h : S^+ \to S^+ \setminus C_\varphi$ such that $\{\{x, h(x)\} : x \in S^+\}$ refines $\mathcal{G}$ (in other words, h is *limited by* $\mathcal{G}$), where $\mathcal{G}$ is an open cover of S^+ by open balls $B(x, \delta_z)$ chosen in such a way that if $x, y \in B(z, \delta_z)$ then

$$||\varphi(y) - \varphi(x)|| \leq \frac{\varepsilon(z)}{4} \leq \frac{\varepsilon(x)}{2}.$$

The existence of such diffeomorphisms comes from the results included in Chapter 3. Taking $g := \varphi \circ h$ we are done.

Chapter 5 has a slightly different flavour. Let us suppose that $f : E \to F$ is a function of class C^1 and that we know that its set of critical points C_f is included in some open set U. Given a continuous function $\varepsilon : E \to (0,\infty)$, is it possible to find $\varphi : E \to F$ of class C^1 without critical points, with $||f(x) - \varphi(x)|| \leq \varepsilon(x)$, and so that $f = \varphi$ outside U?

We will answer to this question affirmatively for the case that E is c_0 or ℓ_p, $1 < p < \infty$, and $F = \mathbb{R}^d$. Moreover, in the case of c_0 it is possible to get that $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ for every $x \in c_0$.

The chosen way to attack this problem is the following. First we take a C^1 function $\delta : E \to [0,\infty)$ so that $\delta(x) \leq \varepsilon(x)$ and $\delta^{-1}(0) = E \setminus U$. Then we build up a C^1 function $g : U \to \mathbb{R}^d$ such that $|f(x) - g(x)| \leq \delta(x)/2$ and $||Df(x) - Dg(x)|| \leq \delta(x)$ and such that C_g is either the empty set for the case of c_0, or is locally contained in a finite union of complemented subspaces of infinite codimension in E for the case of ℓ_p. The techniques employed by Nicole Moulis about C^1-fine approximation of [128], which were also used in [20], will be of great help. Secondly we extend the function g to the whole space E by letting it be equal to f outside U. Due to the C^1-fine approximation of the first step this extension is still of class C^1 on E. For the case of c_0 we would have finished. For the case of ℓ_p we must find a C^1 diffeomorphism $h : E \to E \setminus C_g$ that will be the identity outside U and such that $\{\{x, h(x)\} : x \in E\}$ refines $\mathcal{G} = \bigcup_{z\in E} B(z, \delta_z)$, where $\delta_z > 0$ is chosen so that if $x, y \in B(z, \delta_z)$ then $|g(y) - g(x)| \leq \frac{\delta(z)}{4} \leq \frac{\delta(x)}{2}$. The existence of such diffeomorphism f follows again by the results of Chapter 3. Hence the map $\varphi(x) := g(h(x))$ has no critical points, is equal to f outside U and satisfies $|f(x) - g(x)| \leq \varepsilon(x)$ for all $x \in E$.

Let us recall for a moment the already mentioned work of Moulis [128]. In her paper Moulis proves that for every C^1 function $f : E \to F$, where E is an infinite-dimensional separable Hilbert space and F is a separable Hilbert space, and every continuous function $\varepsilon : E \to (0,\infty)$ there exists a C^∞ function $g : E \to F$ such that $||f(x) - g(x)|| \leq \varepsilon(x)$, $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ for every $x \in E$ and such that $g(C_g)$ has empty interior in F. When comparing it with our result it is clear that we strengthen the conclusion by being able to get $C_g = \emptyset$ and by considering other Banach spaces, not necessarily

Hilbertian. Nonetheless for the Hilbert case we cannot write as the target space an infinite-dimensional Banach space as Moulis does and also we do not get the approximation in the derivatives.

Below we present precise statements of the results obtained in Chapter 5. In the first one we deal with the case when E is an infinite-dimensional Banach space with unconditional basis and C^1 equivalent norm that locally depends on finitely many coordinates (in particular c_0). In the second one we consider E as an infinite-dimensional Banach space with C^1 strictly convex equivalent norm and $1-$suppression unconditional basis (in particular ℓ_p, $1 < p < \infty$).

Theorem 7. *Let E be an infinite-dimensional Banach space with unconditional basis and C^1 equivalent norm that locally depends on finitely many coordinates. Let $f : E \to \mathbb{R}^d$ be a C^1 function and $\varepsilon : E \to (0, \infty)$ be a continuous function. Take any open set U such that $C_f \subset U$. Then there exists a C^1 function $\varphi : E \to \mathbb{R}^d$ such that,*

1. *$|f(x) - \varphi(x)| \leq \varepsilon(x)$ for all $x \in E$.*
2. *$f(x) = \varphi(x)$ for all $x \in E \setminus U$.*
3. *$||Df(x) - D\varphi(x)|| \leq \varepsilon(x)$ for all $x \in E$; and*
4. *$D\varphi(x)$ is surjective for all $x \in E$.*

Theorem 8. *Let E be an infinite-dimensional Banach space with C^1 strictly convex equivalent norm and $1-$suppression unconditional basis $\{e_n\}_{n \in \mathbb{N}}$, that is a Schauder basis such that for every $x = \sum_{j=1}^{\infty} x_j e_j$ and every $j_0 \in \mathbb{N}$ we have that*

$$|| \sum_{j \in \mathbb{N}, j \neq j_0} x_j e_j || \leq || \sum_{j \in \mathbb{N}} x_j e_j ||.$$

Let $f : E \to \mathbb{R}^d$ be a C^1 function and $\varepsilon : E \to (0, \infty)$ be a continuous function. Take any open set U such that $C_f \subset U$. Then there exists a C^1 function $\varphi : E \to \mathbb{R}^d$ such that,

1. *$|f(x) - \varphi(x)| \leq \varepsilon(x)$ for all $x \in E$.*
2. *$f(x) = \varphi(x)$ for all $x \in E \setminus U$.*
3. *$D\varphi(x)$ is surjective for all $x \in E$.*

The results of Chapter 5 can be found published in [87].

Now we shall describe the results obtained in Chapter 3, which, as it has already been mentioned, are essential for the development of the approximate Morse-Sard theorems in Banach spaces of infinite dimensions of Chapters 4 and 5.

Let us begin with an historical introduction about the diffeomorphic deletion of closed sets in Banach spaces.

What one could call *negligibility theory* (in occasions we will refer to this as *extraction theory*) in Banach spaces started in 1953 when Victor L. Klee in [110] proved that if E is a non-reflexive Banach space or is a classical L^p space and K is a compact subset of E, there exists a homeomorphism between E and $E \setminus K$ which is the identity outside an open neighbourhood of K. We will say in that case that K is (topologically) negligible or extractible. Klee also proved that for such infinite-dimensional Banach spaces E the unit sphere is homeomorphic to any of their closed hyperplanes, and gave a topological classification of convex bodies in Hilbert spaces. These results were later extended to the class of all infinite-dimensional normed spaces by C. Bessaga and Klee himself in [36, 37]. There is no need to emphasize that these result do not hold in finite dimensions: one cannot obtain homeomorphisms between $\mathbb{R}^n$ and $\mathbb{R}^n \setminus \{0\}$, since the first set is contractible (homotopic to a point) and the second one is not. Then, in what follows, when we talk about negligibility theory we must think in spaces or manifolds of infinite dimensions.

Klee's work was motivated by those of Tychonoff [146] and Kakutani [107]. From Tychonoff's fixed point theorem it follows that, in the weak topology, the unit ball B_E of the Hilbert space E has to have the fixed point property. In the norm topology, however, this is false. S. Kakutani built a homeomorphism without fixed points from B_E to itself. Using this fact he showed that the unit sphere S_E of E is contractible and is a deformation retract of B_E. Kakutani asked: Are E, B_E and S_E homeomorphic?, to which Klee's work answered positively.

Klee's original proofs were of a strong geometrical flavour: very beautiful, but rather difficult to handle in an analytical way. Bessaga found elegant explicit formulas to build deleting homeomorphisms, based on the existence of continuous non-complete (non-equivalent) norms in every infinite-dimensional Banach space. Using this so-called *non-complete norm technique* and thanks to the existence of a C^∞ non-complete norm in the Hilbert space E, in 1966, C. Bessaga proves that there exists a diffeomorphism between E and $E \setminus \{0\}$ being the identity outside a ball (see [34]). Consequently he concluded that the Hilbert space is diffeomorphic to its unit sphere.

After Bessaga's great result, in the late 60s and beginning of the 70s an important amount of papers appeared treating the homeomorphic and diffeomorphic extraction of closed sets, either in manifolds, or in Fréchet spaces (topological spaces locally convex that are complete with respect to a translation invariant metric) or in Banach spaces. The questions that were asked at that time were: Which spaces allow the homeomorphic or diffeomorphic extraction of points or of compact sets (or other closed sets in general)?; If we want to homeomorphically or diffeomorphically extract compact sets from a space, which smoothness conditions must we impose to the space in question in its geometrical structure to make this possible? (observe that the existence of bump functions with the differentiability degree of the diffeomorphism sought is a necessary condition); and in which cases can we also know that the deleting homeomorphism or diffeomorphism is as close to the identity as we want, that is, that it is limited by any open cover?

In the homeomorphic world about negligibility theory a result of R. D. Anderson, D. H. Henderson and J. West ([8], 1969) stands out. They showed that if M was a metrizable manifold modelled in an infinite-dimensional separable Fréchet space, then every open submanifold N, dense and with the property that every open set U, U and $U \cap N$ have the same homotopy type, is homeomorphic to M by a homeomorphism that can be required to be the identity in any closed subset inside N and is limited by any cover. Such submanifolds include the complements of all closed and locally compact sets of M.

But focusing on the diffeomorphic deletion theory, maybe the most celebrated result that nowadays is classical and fundamental is the fact that two infinite-dimensional Hilbert manifolds, separable and homotopic are diffeomorphic (see the essays of Burghelea, Kuiper, Eells, Elworthy and Moulis [47, 127, 67, 70]), for which was key the already cited work of Bessaga.

We highlight now two works that become key for the development of Chapter 3 of this dissertation. In the first one we get that the deleting diffeomorphisms h are as close to the identity as we want, meaning that they refine a given open cover $\mathcal{G}$ (that for every $x \in E$ there exists $U \in \mathcal{G}$ such that $x, h(x) \in U$).

In 1969 James West proved that in the case when E is a separable Hilbert space or even a separable Hilbert manifold, for every subset K locally compact, for every open set U including K, and every open cover $\mathcal{G}$ of E, there exists a C^∞ diffeomorphism $h : E \to E \setminus K$ such that h is the identity outside U and is limited by $\mathcal{G}$.

Peter Renz in his thesis of 1969 proved the following: Let E be an infinite-dimensional Banach space with unconditional basis and with an equivalent C^k norm. Then for every closed and locally compact set K and every open set U there exists a C^k diffeomorphism from $E \setminus K$ onto $E \setminus (K \setminus U)$ which is the identity on $E \setminus (U \setminus K)$. (The same is true for paracompact C^k manifolds modelled on such kind of Banach spaces).

After these years of huge mathematical production in this field we essentially find three authors that have continued to refine the diffeomorphic negligibility theory in Banach spaces since then until the present day; these are T. Dobrowolski, D. Azagra and A. Montesinos. In all cases, their works do not pay attention to getting the deleting diffeomorphisms to be arbitrarily close to the identity, in the sense of refining a given cover. This subtle property have been retaken after almost 50 years in this thesis, because it appears to be indispensable for our theorems about uniform approximation without critical points.

The diffeomorphic negligibility of compact sets in Banach spaces was taken up by Tadeusz Dobrowolski, who developed the non-complete norm technique of Bessaga. In his paper [63] of 1979 Dobrowolski firstly proved that if E is an infinite-dimensional separable Banach space and if K is a weakly compact set or a closed infinite-codimensional subspace, then E and $E \setminus K$ are real-analytic diffeomorphic[9]. Previously the only thing that was known about real-analytic negligibility was that the infinite-dimensional Hilbert space E is real-analytic diffeomorphic to $E \setminus \{0\}$ and the extraction of the point can happen at the end of a C^∞ isotopy. This was proven by Burghelea and Kuiper [47]. In second place Dobrowolski showed that for every infinite-dimensional Banach space E having a non-complete C^k smooth norm and for every compact set K in E, the space E is C^k diffeomorphic $E \setminus K$. In particular the Banach spaces for which there exist continuous and linear embeddings into some $c_0(\Gamma)$ (weakly compactly generated spaces WCG[10] satisfy this, see [60, p. 246], hence also all separable and all reflexive spaces) have non-complete C^∞ smooth norms. If, moreover, E has an equivalent norm $||\cdot||$ of class C^k then one can deduce that $S = \{x \in E : ||x|| = 1\}$ is C^k diffeomorphic to any hyperplane of E. Dobrowolski also employed his results of diffeomorphic negligibility to give a classification of C^k smooth bodies in WCG Banach spaces (see [62]).

We remark that there are examples of spaces with equivalent C^∞ smooth norms that do not linearly embed into any $c_0(\Gamma)$. An example of that kind (non-separable) is given in [60, Ex VI 8.8] and can be chosen to be a $C(K)$ space for some compact set K. Then, when one wants to generalize these results to any infinite-dimensional Banach space that has a C^k smooth norm, one faces the following problem: Does every infinite-dimensional Banach space with an equivalent C^k smooth norm admit a non-complete C^k smooth norm too? For $k \geq 2$, this intriguing question was solved very recently by D'Alessandro and Hájek [57], but the C^1 case remains unsolved.

Without proving the existence of non-complete smooth norms, by introducing a kind of non-complete asymmetric and convex function (called asymmetric norm) D. Azagra showed in 1997 [10] that every Banach space E with a (not necessarily equivalent) C^k smooth norm is C^k diffeomorphic to $E \setminus \{0\}$ and, moreover, that every closed hyperplane H in E is C^k diffeomorphic to the sphere $\{x \in E : ||x|| = 1\}$. Nowadays, for the case of $k \geq 2$, by using D'Alessandro and Hajek's result this would be simple by a direct appplication of [63].

In 1998 Azagra and Dobrowolski joined efforts to strengthen the asymmetric norm technique of deleting points introduced in [10] so as to generalize some results on smooth negligibility of compact sets and subspaces to the class of all Banach spaces having a (not necessarily equivalent) C^k smooth norm. They also gave a full smooth classification of the convex bodies of every Banach space. In particular, they showed that every smooth convex body containing no linear subspaces in an infinite-dimensional Banach space is diffeomorphic to a half-space.

These results enabled them to enlarge the class of spaces for which some other applications of negligibility are valid. A sample of such applications includes Garay's theorems [85, 86] concerning the existence of solutions to ordinary differential equations and cross-sections of solution funnels in Banach spaces, as well as sharper statements of Klee's results [110] on periodic homeomorphisms without fixed points.

In the intrinsic geometry and structure of Banach spaces an important role is played by the existence of smooth equivalent norms, smooth bump functions and smooth partitions of unity. After the works of Azagra and Dobrowolski in which the main hypothesis was the existence of a norm (not necessarily equivalent) of class C^k there arises the question of which results about diffeomorphic negligibility are possible assuming existence of bump functions or partitions of unity of certain class C^k. These initial assumptions are natural since they are weaker than the mere existence of an equivalent norm of class C^k in the Banach space, as Haydon's example showed in [99]. This was the problem studied in Montesinos'

[9] Such results are not valid in the non-separable case: if Γ is an uncountable set, then $c_0(\Gamma)$ is not real-analytic diffeomorphic to $c_0(\Gamma) \setminus \{0\}$ (see [63, Proposition 4.7]).

[10] A Banach space E is called weakly compactly generated if there is a weakly compact set K such that $\overline{\mathrm{span}}(K) = E$.

thesis, a student of Azagra and Jaramillo.

As a result of this work Azagra and Montesinos added two new results to the diffeomorphic negligibility theory in 2003. On the one hand they proved that if E is an infinite-dimensional Banach space with C^k partitions of unity then, for each weakly compact set K and every starlike body A so that $\operatorname{dist}(K, E \setminus A) > 0$, there exists a C^k diffeomorphism $h : E \to E \setminus K$ such that h is the identity outside A. On the other hand if one assumes that E is an infinite-dimensional Banach space with Schauder basis then E has C^k smooth bumps if and only if for every compact set K and for every open set U of E that includes K, there exists a C^k diffeomorphism $h : E \to E \setminus K$ such that h is the identity on $E \setminus U$.

This kind of results about topological negligibility have found many interesting applications in various fields of mathematics, which include fixed point theory, topological classification of convex bodies, strange phenomena related with ordinary differential equations and dynamical systems in infinite dimensions, the failure if Rolle's theorem in infinite-dimensions and many other things. See [13, 14, 15, 25, 35, 12, 16, 87] and the references therein.

Although maybe the application that we are more interested in, and is worth mentioning for our purposes, was the application of West's theorem [150] to prove the result of Azagra and Cepedello saying that continuous functions defined in a separable Hilbert space and taking values in $\mathbb{R}^m$ can be uniformly approximated by C^∞ fucntions without any critical point (see [12]). Actually here lies the idea to use results about diffeomorphic negligibility of closed sets in Banach spaces to prove our approximate Morse-Sard theorems from Chapters 4 and 5.

Let us go back now to the pioneering work of this dissertation. We, as it was already said, want to retake the property that the deleting diffeomorphisms are limited by any given cover, that was ignored by Azagra, Dobrowolski and Montesinos. In Chapters 4 and 5 we will prove the following.

Theorem 9. *Let E be a Banach space, $p \in \mathbb{N} \cup \{\infty\}$, and $X \subset E$ a closed set with the property that, for each $x \in X$, there exists a neighbourhood U_x of x in E, Banach spaces $E_{(1,x)}$ and $E_{(2,x)}$, and a continuous mapping $f_x : C_x \to E_{(2,x)}$, where C_x is a closed subset of $E_{(1,x)}$, such that:*

1. *$E = E_{(1,x)} \oplus E_{(2,x)}$;*
2. *$E_{(1,x)}$ has C^k partitions of unity;*
3. *$E_{(2,x)}$ is infinite-dimensional and has a (not necessarily equivalent) C^k smooth norm;*
4. *$X \cap U_x \subset G(f_x)$, where*

$$G(f_x) = \{y = (y_1, y_2) \in E_{(1,x)} \oplus E_{(2,x)} : y_2 = f_x(y_1), y_1 \in C_x\}.$$

Then, for every open cover $\mathcal{G}$ of E and every open subset U of E, there exists a C^k diffeomorphism h of $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$. Moreover, the same conclusion is true if we replace E by an open subset of E.

For its proof we will combine techniques and methods of the works of West [150], Renz [134, 135] and Azagra and Dobrowolski [10, 14]. Observe that compact sets in Banach spaces that possess unconditional basis can be locally seen as graphs of closed sets defined in an infinite-codimensional subspace and taking values in its linear complement. The proof of this fact, in which a theorem due to Corson from [55] is key, is included in Section 3.6 and makes our Theorem 9 generalize West and Renz's theorems.

Lastly, to be used in Theorem 8 of Chapter 5, we will also need the following variant of the previous theorem, in which we ask the extractible set to be locally contained in a finite union of closed subspaces of infinite-codimension.

Theorem 10. *Let E be a Banach space with a norm of class C^k. Take an open cover $\mathcal{G}$ of an open set U and a closed set $X \subset U$ such that for each $x \in X$, there exist an open neighbourhood U_x of x and some* *closed subspaces* $E_1, \ldots, E_{n_x} \subset E$ *complemented in* E *and of infinite codimension such that*

$$X \cap U_x \subset \bigcup_{j=1}^{n_x} E_j.$$

Then there exists a C^k *diffeomorphism* $h : E \to E \setminus X$ *which is the identity outside* U *and is limited by* $\mathcal{G}$.

Theorem 9 appears in [16, Section 2], which is a joint work with Azagra and Dobrowolski, while Theorem 10 can be found in [87, Section 2].

Chapter 1

Some remarks about the Morse-Sard theorem and approximate differentiability

Throughout this dissertation, given a C^k smooth mapping $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, C_f stands for the set of critical points of f (that is, the points $x \in \mathbb{R}^n$ at which the differential $Df(x)$ is not surjective), and $f(C_f)$ is thus the set of critical values of f; the same terminology applies to smooth mappings between manifolds, both finite and infinite-dimensional. The Morse-Sard theorem [125, 140] states that if $k \geq \max\{n-m+1, 1\}$ then $f(C_f)$ is of Lebesgue measure zero in $\mathbb{R}^m$. This result also holds true for C^k smooth mappings $f : N \longrightarrow M$ between two smooth manifolds of dimensions n and m respectively. A celebrated example of Whitney's [152] shows that this classical result is sharp within the classes of functions C^j. Given the crucial applications of the Morse-Sard theorem in several branches of mathematics, it is nonetheless natural and useful to try and refine the Morse-Sard theorem for other classes of functions, and by now there is a rich literature in this line of work. We already mentioned and commented on all of the important contributions generated by this problem in the introduction. Here we shall content ourselves with referring the reader to [154, 130, 31, 58, 82, 43, 44, 28, 113, 114, 29, 96, 95, 17, 18, 75, 76] and the references therein. Undoubtedly it is a fundamental result in Differential Geometry and Analysis.

In this chapter we will show how, by combining some of the strategies and tools which are common to several of these with the idea of the proof of [121, Theorem 1] and an induction argument, one can obtain the following result: let $n \geq m$ and $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Borel function. Suppose that f is approximately differentiable of order 1 at $\mathcal{H}^m-$almost every point and satisfies

(a) $\operatorname{ap}\limsup_{y\to x} \frac{|f(y)-f(x)|}{|y-x|} < +\infty$ for all $x \in \mathbb{R}^n \setminus N_0$, where N_0 is a countable set, and

(b) $\operatorname{ap}\lim_{y\to x} \frac{|f(y)-f(x)-\cdots\frac{F_i(x)}{i!}(y-x)^i|}{|y-x|^i} = 0$ for all $i = 2, \ldots, n-m+1$ and for all $x \in \mathbb{R}^n \setminus N_i$, where each set N_i is $(i+m-2)$-σ-finite and the coefficients $F_i(x)$ are Borel functions,

then f has the Morse-Sard property (that is to say, the image of the critical set of f is null with respect to the Lebesgue measure in $\mathbb{R}^m$). See Theorem 1.25 in Section 1.3 below for a precise statement and proof. In Theorem 1.21 we are able to dispense with the condition about the Borel measurability of the functions but we must strengthen conditions (b) above by replacing the (s)-σ-finite exceptional sets with countably $(\mathcal{H}^s, s)$ rectifiable sets of certain classes C^k. In Theorem 1.23 we provide an interesting variant of this result. See Sections 1.2 and 1.3 for auxiliary results and definitions. Theorem 1.21 and Theorem 1.25 generalize the versions of the Morse-Sard theorem provided by Bates's theorem and the Appendix of [17], and are not stronger, nor weaker, than the versions of [44, 113, 114] for BV_n or Sobolev functions with smaller exponents; see Section 1.4 below for examples and further comments.

1.1 Preliminaries

1.1.1 Lebesgue and Hausdorff measure

We denote by $\mathcal{L}^n(E)$ the outer Lebesgue measure of a set $E \subseteq \mathbb{R}^n$. For $s \geq 0$, the $s-$dimensional Hausdorff measure is denoted by $\mathcal{H}^s$, and the $s-$dimensional Hausdorff content by $\mathcal{H}^s_\infty$. Recall that for any subset E of $\mathbb{R}^n$ we have, by definition,

$$\mathcal{H}^s(E) = \lim_{\delta \searrow 0} \mathcal{H}^s_\delta(E) = \sup_{\delta > 0} \mathcal{H}^s_\delta(E),$$

where for each $0 < \delta \leq \infty$,

$$\mathcal{H}^s_\delta(E) = \inf \left\{ \sum_{i=1}^{\infty} (\text{diam}\, F_i)^s : \text{diam}\, F_i \leq \delta,\, E \subseteq \bigcup_{i=1}^{\infty} F_i \right\}.$$

It is well known that the measures $\mathcal{H}^n$, $\mathcal{H}^n_\infty$ and $\mathcal{L}^n$ are equivalent on $\mathbb{R}^n$, in the sense that there exists a constant $C(n)$ such that $\mathcal{L}^n(E) = C(n)\mathcal{H}^n(E)$ and $\mathcal{H}^n(E) = \mathcal{H}^n_\infty(E)$ for every $E \subset \mathbb{R}^n$ (see for instance Section 2.2 of [71] for the equivalence between $\mathcal{H}^n$ and $\mathcal{L}^n$). Moreover $\mathcal{H}^s$ and $\mathcal{H}^s_\infty$ have the same null sets. Another important fact about the Hausdorff measure $\mathcal{H}^s$ is that it is Borel regular (see e.g. [71, Theorem 1 (p. 61)]).

Definition 1.1. We say that a set is *s-σ-finite* if it can be written as a countable union of sets of finite $\mathcal{H}^s-$measure.

Also a set $N \subseteq \mathbb{R}^n$ is called *countably $(\mathcal{H}^s, s)$ rectifiable of class C^k,* where $s \leq n$, provided that there exist countably many $s-$dimensional submanifolds A_j of class C^k such that $\mathcal{H}^s(N \setminus \bigcup_{j=1}^{\infty} A_j) = 0$. Specifically there exist C^k functions $\phi_j : \mathbb{R}^s \to \mathbb{R}^n$ so that $\phi_j(\mathbb{R}^s) = A_j \subset \mathbb{R}^n$. This notion, for $k > 1$, has been first introduced in [9].

1.1.2 Whitney's Extension Theorem

Recall that a modulus of continuity is a concave, increasing function $\omega : [0, \infty) \to [0, \infty)$ such that $\omega(0^+) = 0$. Given a positive integer k, for a fixed modulus of continuity ω the class $C^{k,\omega}(\mathbb{R}^n; \mathbb{R})$ is defined as the set of functions which are k times differentiable and its partial derivatives of order k are uniformly continuous with modulus of continuity ω. In the particular case $\omega(s) = s^t$ for some $t \in (0, 1]$ we will write $C^{k,t}(\mathbb{R}^n; \mathbb{R})$.

A fundamental tool in our proofs of Theorem 1.21 and Theorem 1.23 will be the following version of the Whitney extension theorem, see [88, 123, 144].

Theorem 1.2 (Uniform version of Whitney's Extension Theorem)**.** *Let ω be a modulus of continuity. Let C be a subset of $\mathbb{R}^n$ and $\{f_\alpha\}_{|\alpha| \leq k}$ be a family of real valued functions defined on C satisfying*

$$f_\alpha(x) = \sum_{|\beta| \leq k - |\alpha|} \frac{f_{\alpha+\beta}(y)}{\beta!}(x - y)^\beta + R_\alpha(x, y) \tag{1.1.1}$$

for all $x, y \in C$ and all multi-indices α with $|\alpha| \leq k$. Suppose that for some constant $M > 0$ we have

$$|f_\alpha(x)| \leq M, \text{ and } \quad |R_\alpha(x, y)| \leq M|x - y|^{k-|\alpha|}\omega(|x - y|) \quad \text{for all} \quad x, y \in C \text{ and all } |\alpha| \leq k, \tag{1.1.2}$$

where the term R_α is defined by equation (1.1.1)*. Then there exists a function $F : \mathbb{R}^n \to \mathbb{R}$ such that:*

(i) $F \in C^{k,\omega}(\mathbb{R}^n; \mathbb{R})$.

(ii) $D^\alpha F = f_\alpha$ *on C for all $|\alpha| \leq k$.*

Our notation with multi-indices is standard (see e.g. [155, p. 2]). Given a muti-index $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N} \cup \{0\})^n$ we write $\alpha! = \alpha_1 \cdots \alpha_n$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and if $x = (x_1 \dots, x_n) \in \mathbb{R}^n$ then $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. If $\alpha = (\alpha_1, \dots, \alpha_n)$ then $D^\alpha f$ denotes the partial derivative $\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$. Also if we write $\sum_{|\alpha| \leq k}$ we mean that we are summing over all possible choices of multi-indices $\alpha = (\alpha_1, \dots, \alpha_n)$ satisfying $|\alpha| \leq k$.

The previous version of the Whitney extension theorem is usually stated for *closed* subsets C of $\mathbb{R}^n$, but it is easily checked that Theorem 1.2 also holds for arbitrary subsets $C \subset \mathbb{R}^n$, because a modification of the usual argument showing that an uniformly continuous function defined on a set D has a unique uniformly continuous extension (with the same modulus of continuity) to the closure $\overline{D}$ of D, together with conditions (1.1.1) and (1.1.2), imply that if C is not closed then the functions f_α have unique extensions to $\overline{C}$ that also satisfy (1.1.1) and (1.1.2) on $\overline{C}$. The theorem also remains true if we replace the target space $\mathbb{R}$ with $\mathbb{R}^m$, as one can apply the above result to the coordinate functions of $f = (f^1, \dots, f^m)$. In our proofs of Theorem 1.21 and Theorem 1.23 we will use this version of the Whitney extension theorem in the particular instances of $\omega(s) = s$ (thus obtaining extensions of class $C^{k,1}$), or $\omega(s) = s^t$, with $0 < t < 1$ (in which case we will have extensions belonging to the Hölder differentiability classes $C^{k,t}$).

We will also use Whitney's original theorem for C^k, which we next restate for the reader's convenience.

Theorem 1.3 (Whitney Extension Theorem)**.** *Let $C \subset \mathbb{R}^n$ be closed. A necessary and sufficient condition, for a function $f : C \to \mathbb{R}$ and a family of functions $\{f_\alpha\}_{|\alpha| \leq k}$ defined on C satisfying $f = f_0$ and*

$$f_\alpha(x) = \sum_{|\beta| \leq k - |\alpha|} \frac{f_{\alpha+\beta}(y)}{\beta!} (x-y)^\beta + R_\alpha(x,y)$$

for all $x, y \in C$ and all multi-indices α with $|\alpha| \leq k$, to admit a C^k extension F to all of $\mathbb{R}^n$ such that $D^\alpha F = f_\alpha$ on C for all $|\alpha| \leq k$, is that

$$\lim_{|x-y| \to 0} \frac{R_\alpha(x,y)}{|x-y|^{k-|\alpha|}} = 0 \qquad (W^k)$$

uniformly on compact subsets of C, for every $|\alpha| \leq k$.

1.1.3 Approximate differentiability

Let us begin by introducing the concept of approximate limit.

Definition 1.4. For a function $f : \mathbb{R}^n \to \mathbb{R}^m$ we write $\operatorname{ap}\lim_{y \to x} f(y) = L$ to mean that for every $\varepsilon > 0$,

$$\lim_{r \to 0} \frac{\mathcal{L}^n \left(B(x,r) \cap \{ y \in \mathbb{R}^n : \, |f(y) - L| \geq \varepsilon \} \right)}{\mathcal{L}^n(B(x,r))} = 0,$$

and, if f is real-valued, $\operatorname{ap}\limsup_{y \to x} f(y)$ is defined to be the infimum of all $\lambda \in \mathbb{R}$ such that

$$\lim_{r \to 0} \frac{\mathcal{L}^n \left(B(x,r) \cap \{ y \in \mathbb{R}^n : \, f(y) \geq \lambda \} \right)}{\mathcal{L}^n(B(x,r))} = 0.$$

Equivalently, we could have defined the approximate limit $ap \lim_{y \to x} f(y)$ to be equal to $L \in \mathbb{R}^m$ whenever there exists a set A_x with Lebesgue density one at x so that $\lim_{\substack{y \to x \\ y \in A_x}} f(y) = L$. It is an easy exercise to check that these two definitions are in fact the same. We next show the proof for the interested reader.

Proposition 1.5. *A function $f : \mathbb{R}^n \to \mathbb{R}^m$ satisfies $ap \lim_{y \to x} f(y) = L$ if and only if there exists a set $A_x \subseteq \mathbb{R}^n$ which has Lebesgue density one at x and such that $\lim_{\substack{y \to x \\ y \in A_x}} f(y) = L$.*

Proof. For the sufficiency first recall that x is a point of density one for A_x if

$$\lim_{r\to 0}\frac{\mathcal{L}^n(B(x,r)\cap A_x)}{\mathcal{L}^n(B(x,r))}=1.$$

Take $\varepsilon>0$. Note that by hypothesis there is some $r_0>0$ small enough such that if $0<r<r_0$ then

$$[B(x.r)\cap\{y\in\mathbb{R}^n:\,|f(y)-L|\geq\varepsilon\}]\subseteq[B(x,r)\setminus A_x].$$

Therefore we can write

$$\begin{aligned}\lim_{r\to 0}\frac{\mathcal{L}^n(B(x,r)\cap\{y\in\mathbb{R}^n:\,|f(y)-L|\geq\varepsilon\})}{\mathcal{L}^n(B(x,r))}&\leq\lim_{r\to 0}\frac{\mathcal{L}^n(B(x,r)\setminus A_x)}{\mathcal{L}^n(B(x,r))}\\&=1-\lim_{r\to 0}\frac{\mathcal{L}^n(B(x,r)\cap A_x)}{\mathcal{L}^n(B(x,r))}=0.\end{aligned}$$

To prove the necessity, by definition of approximate limit we have that for every $j\in\mathbb{N}$ there exists $r_j>0$ such that

$$\mathcal{L}^n\left(B(x,r)\cap\{y\in\mathbb{R}:\,|f(y)-L|\geq\frac{1}{j}\}\right)\leq\frac{\mathcal{L}^n(B(x,r))}{2^j}.$$

for every $r\leq r_j$. We may assume without loss of generality that $r_{j+1}<r_j$ for all $j\in\mathbb{N}$. Define

$$A_x=\bigcup_{j=1}^{\infty}(B(x,r_j)\setminus B(x,r_{j+1}))\cap\{y\in\mathbb{R}^n:\,|f(y)-L|<\frac{1}{j}\}.$$

We will see first that A_x has density one at x, that is

$$\lim_{r\to 0}\frac{\mathcal{L}^n(B(x,r)\cap(\mathbb{R}^n\setminus A_x))}{\mathcal{L}^n(B(x,r))}=0.$$

Take $\delta>0$ and choose $j_0\in\mathbb{N}$ such that $\sum_{j=j_0}^{\infty}\frac{1}{2^j}<\delta$. Let also $0<r<r_{j_0}$ and call $j_1\geq j_0$ the natural number satisfying $r_{j_1+1}\leq r<r_{j_1}$. Then

$$\begin{aligned}\mathcal{L}^n(B(x,r)\cap(\mathbb{R}^n\setminus A_x))&\leq\sum_{j=j_1}^{\infty}\mathcal{L}^n\left((B(x,r_j)\setminus B(x,r_{j+1}))\cap\{y\in\mathbb{R}^n:\,|f(y)-L|\geq\frac{1}{j}\}\right)\\&\leq\sum_{j=j_1}^{\infty}\mathcal{L}^n\left(B(x,r_j)\cap\{y\in\mathbb{R}^n:\,|f(y)-L|\geq\frac{1}{j}\}\right)\\&\leq\frac{\mathcal{L}^n(B(x,r))}{2^{j_1}}+\sum_{j=j_1+1}^{\infty}\frac{\mathcal{L}^n(B(x,r_j))}{2^j}\\&\leq\mathcal{L}^n(B(x,r))\sum_{j=j_1}^{\infty}\frac{1}{2^j}<\mathcal{L}^n(B(x,r))\delta.\end{aligned}$$

Finally let us check that $\lim_{\substack{y\to x\\ y\in A_x}}f(y)=L$. For a given $\varepsilon>0$ we fix $\frac{1}{j}<\varepsilon$ and therefore for every $0<r\leq r_j$ we get

$$B(x,r_j)\cap A_x\subset\{y\in\mathbb{R}^n:\,|f(y)-L|<\frac{1}{j}\}\subset\{y\in\mathbb{R}^n:\,|f(y)-L|<\varepsilon\}.$$

□

If the approximate limit exists, then it is unique (see for instance [71, p. 46]). We also have that if $\lim_{y\to x} f(y) = l$ then $\operatorname{ap}\lim_{y\to x} f(y) = l$. And we say that f is approximately continuous at x if $\operatorname{ap}\lim_{y\to x} f(y) = f(x)$. Observe that we have defined the approximate continuity without requiring f to be Lebesgue measurable. In fact a function is measurable if and only if it is approximately continuous almost everywhere, as shown for example in [74, Theorem 2.9.13].

Definition 1.6. A function $p : \mathbb{R}^n \to \mathbb{R}^m$ is said to be a *polynomial of degree k centered at the point* $x \in \mathbb{R}^n$ if it is written in the form

$$p(x; y) = \sum_{|\alpha|\le k} \frac{p_\alpha(x)}{\alpha!}(y-x)^\alpha,$$

where each $p_\alpha(x) = (p^1_\alpha(x), \dots, p^m_\alpha(x)) \in \mathbb{R}^m$.

Definition 1.7. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *approximately differentiable* of order k at $x \in \mathbb{R}^n$ if there is a polynomial $p_k(x; y)$, centered at x, of degree at most k and where $p_k(x; x) = f(x)$, such that

$$\operatorname{ap}\lim_{y\to x} \frac{|f(y) - p_k(x; y)|}{|y-x|^k} = 0.$$

Definition 1.8. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to have an *approximate* $(k-1)-$*Taylor polynomial* at x if there is a polynomial $p_{k-1}(x; y)$, centered at x, of degree at most $k-1$ and where $p_{k-1}(x; x) = f(x)$ such that

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y) - p_{k-1}(x; y)|}{|y-x|^k} < +\infty.$$

In particular if a function is approximately differentiable or has an approximate Taylor polynomial of any order at x, then it will be approximately continuous at x as well. We denote $p_0(x; y) = f(x)$.

By a straightforward application of Proposition 1.5, $f : \mathbb{R}^n \to \mathbb{R}^m$ is approximately differentiable of order k at $x \in \mathbb{R}^n$ if and only if there exits a set $A_x \subset \mathbb{R}^n$ which has Lebesgue density 1 at x and such that $f|_{A_x}$ has a Taylor polynomial of order k in the classical sense at x.

Therefore if a function f is of class C^k then in particular f has a Taylor expansion of order k at x, and hence f is approximately differentiable of order k at x, with corresponding Taylor polynomial

$$p_k(x; y) = \sum_{|\alpha|\le k} \frac{D^\alpha f(x)}{\alpha!}(y-x)^\alpha.$$

($D^\alpha f(x) = (D^\alpha f^1(x), \dots, D^\alpha f^m(x)) \in \mathbb{R}^m$).

It should be noted that if f is approximately differentiable of order k (or has an approximate $(k-1)-$Taylor polynomial) at x then the corresponding polynomial $p_k(x; y)$ (or $p_{k-1}(x; y)$) of degree at most k (or $k-1$) is unique. Actually all the usual rules about differentiability of sums, products and quotients of functions apply to approximate differentiable functions as well.

As an example let us comment how one gets the uniqueness of the approximate derivative. Suppose there exists two distinct polynomials of order 1 centered at x, $p_1(x; y), q_1(x; y)$ such that

$$\operatorname{ap}\lim_{y\to x} \frac{|f(y) - p_1(x; y)|}{|y-x|} = \operatorname{ap}\lim_{y\to x} \frac{|f(y) - q_1(x; y)|}{|y-x|} = 0.$$

We can write $p_1(x; y) = f(x) + T_1(x-y)$ and $q_1(x; y) = f(x) - T_2(x-y)$, where T_1 and T_2 are two continuous linear operators from $\mathbb{R}^n$ to $\mathbb{R}^m$. We have

$$\operatorname{ap}\lim_{y\to x} \frac{|p_1(x; y) - q_1(x; y)|}{|y-x|} = \operatorname{ap}\lim_{y\to x} \frac{|(T_1 - T_2)(y-x)|}{|y-x|} = \operatorname{ap}\lim_{h\to 0} \frac{|(T_1 - T_2)(h)|}{|h|} = 0.$$

Let $0 < \varepsilon < 1$. Then there exists $r > 0$ such that

$$\mathcal{L}^n(B(0, r) \cap \{h \in \mathbb{R}^m : |(T_1 - T_2)(h)| > \varepsilon|h|\}) < w_n r^n \varepsilon^n,$$

where w_n is the volume of the unit ball in $\mathbb{R}^n$. Hence for every $u \in \mathbb{R}^n$ with $|u| = r - \varepsilon r$ there exists $h \in B(u, \varepsilon r) \subset B(0, r)$ with $|(T_1 - T_1)(h)| \leq \varepsilon |h|$, which implies

$$|(T_1 - T_2)(u)| \leq |(T_1 - T_2)(u - h)| + |(T_1 - T_2)(h)| \leq ||T_1 - T_2||\varepsilon r + \varepsilon r = \frac{(||T_1 - T_2|| + 1)\varepsilon}{1 - \varepsilon}|u|.$$

Therefore $||T_1 - T_2|| \leq \frac{(||T_1 - T_2|| + 1)\varepsilon}{1 - \varepsilon}$, and since this holds for every $0 < \varepsilon < 1$ we get that $T_1 = T_2$.

In view of the uniqueness of the approximate Taylor polynomial it is reasonable now to change our notation and express the polynomial $p_k(x; y)$ by

$$p_k(x; y) = \sum_{|\alpha| \leq k} \frac{f_\alpha(x)}{\alpha!}(y - x)^\alpha. \tag{1.1.3}$$

In particular $f_0(x) = f(x)$. Throughout the chapter we will use the same notation for the coefficients of the polynomials regardless of whether f is approximately differentiable or has an approximate Taylor polynomial. This will cause no problem because if both things are true then the coefficients of the corresponding polynomials are the same (this assertion is a consequence of Lemma 1.10 below).

Observe that using the notation f_α does not by any means imply that there exists a derivative $D^\alpha f$ in the usual sense, nor that $f_\alpha(x) = D^\alpha f(x)$ even if $D^\alpha f(x)$ exists.

From now on, every time we say a function or a set is measurable, and unless we specify the measure, we will mean it with respect to the Lebesgue measure.

1.2 From approximate differentiability to C^k regularity

Let us state first the following important characterization of approximate differentiability, due to the independent works of Lui and Tai [121] on the one hand, and Isakov [102] on the other hand.

Theorem 1.9 (Liu and Tai, 1994)**.** *For a measurable set $\Omega \subseteq \mathbb{R}^n$ and a measurable function $f : \Omega \to \mathbb{R}^m$ the following statements are equivalent:*

1. *For every $\varepsilon > 0$, there exists $g : \mathbb{R}^n \to \mathbb{R}^m$ of class C^k such that $\mathcal{L}(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}) \leq \varepsilon$.*
2. *f has an approximate $(k-1)-$ Taylor polynomial for almost every $x \in \Omega$.*
3. *f is approximately differentiable of order k for almost every $x \in \Omega$.*

In the proofs of Theorem 1.21 and Theorem 1.23 we will make heavy use of the following Lemma, which is an easy consequence of the argument that Liu and Tai used in their proof of the Theorem above. For completeness, and for the reader's convenience, we provide a detailed argument.

Lemma 1.10. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a measurable function, k a positive integer and N a subset of $\mathbb{R}^n$. Consider the following statements.*

(i) f is approximately differentiable of order k for all $x \in \mathbb{R}^n \setminus N$.

(ii) f has an approximate $(k-1)-$Taylor polynomial for all $x \in \mathbb{R}^n \setminus N$.

Then we have that $(i) \Rightarrow (ii)$. Furthermore assume that (ii) holds and let $f_\alpha(x)$ denote the coefficients of the existing approximate $(k-1)-$Taylor polynomial. Then:

(iii) *there exists a decomposition*

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_j \cup N,$$

such that for each $j \in \mathbb{N}$ there is a function $g_j \in C^{k-1,1}(\mathbb{R}^n; \mathbb{R}^m)$ with $f_\alpha(x) = D^\alpha g_j(x)$ for all $x \in B_j$ and $|\alpha| \leq k-1$.

We will need to use the following lemma due to De Giorgi for which we present the proof of Campanato [50, Lemma 2.1].

Lemma 1.11 (De Giorgi)**.** *Let E be a measurable subset of the ball $B(x,r) \subset \mathbb{R}^n$ such that $\mathcal{L}^n(E) \geq Ar^n$ for some constant $A > 0$. Then for each positive integer k there is a positive constant C, depending only on n, k and A, such that*

$$|D^\alpha p(x)| \leq \frac{C}{r^{n+|\alpha|}} \int_E |p(y)|\, dy$$

for all polynomials $p : \mathbb{R}^n \to \mathbb{R}^m$ of degree at most k and all multi-indices $|\alpha| \leq k$.

Proof. Without loss of generality one may assume that the polynomials are real-valued, that is $p : \mathbb{R}^n \to \mathbb{R}$.
We denote by $\mathcal{P}_k$ the space of polynomials of degree at most k and by $\mathcal{T}_k \subseteq \mathcal{P}_k$ the set of polynomials

$$p(y) = \sum_{|\alpha| \leq k} a_\alpha y^\alpha$$

satisfying the relation

$$\sum_{|\alpha| \leq k} |a_\alpha|^2 = 1. \tag{1.2.1}$$

Let $\mathcal{F}$ be the class of measurable functions f on $\mathbb{R}^n$ with compact support in $B(0,1)$ such that

$$\begin{cases} 0 \leq f(y) \leq 1 \\ \int_{\mathbb{R}^n} f(y)\, dy \geq A. \end{cases} \tag{1.2.2}$$

Write

$$\gamma(A) = \inf_{p \in \mathcal{T}_k, f \in \mathcal{F}} \int_{B(0,1)} |p(y)| f(y)\, dy \tag{1.2.3}$$

and we will prove that

$$\gamma(A) = \min_{p \in \mathcal{T}_k, f \in \mathcal{F}} \int_{B(0,1)} |p(y)| f(y)\, dy. \tag{1.2.4}$$

By definition, for each positive integer m there exists a polynomial $p_m \in \mathcal{T}_k$ and a function $f_m \in \mathcal{F}$ such that

$$\gamma(A) \leq \int_{B(0,1)} |p_m(y)| f_m(y)\, dy < \gamma(A) + \frac{1}{m}. \tag{1.2.5}$$

Thanks to condition (1.2.1) we can take a subsequence $|p_\nu|$, uniformly convergent on each compact set of $\mathbb{R}^n$ to a polynomial $p^* \in \mathcal{T}_k$. Similarly, by (1.2.2), from $|f_\nu|$ we take a subsequence $|f_\mu|$ that converges weakly in $L^2(B(0,1))$ to a function $f^* \in \mathcal{F}$.
On the other hand, using (1.2.5),

$$\gamma(A) \leq \int_{B(0,1)} |p_\mu(y)| f_\mu(y)\, dy < \gamma(A) + \frac{1}{\mu}.$$

We let μ tend to infinity and we get

$$\gamma(A) = \int_{B(0,1)} |p^*(y)| f^*(y)\, dy.$$

This proves our assertion (1.2.4). An important consequence is that $\gamma(A) > 0$.
Now, if E is any measurable subset of $B(0,1)$ with $\mathcal{L}^n(E) \geq A$ and $p \in \mathcal{T}_k$, taking $f = \mathcal{X}_E \in \mathcal{F}$ we have that

$$\int_E |p(y)|\, dy \geq \gamma(A).$$

Observe that if $p \in \mathcal{P}_k$ then $p \cdot \left(\sum_{|\alpha|\le k} |a_\alpha|^2\right)^{-\frac{1}{2}} \in \mathcal{T}_k$ and using the last inequality,

$$\left(\sum_{|\alpha|\le k} |a_\alpha|^2\right)^{\frac{1}{2}} \le \frac{1}{\gamma(A)} \int_E |p(y)|\, dy,$$

and also

$$|a_\alpha| \le \frac{1}{\gamma(A)} \int_E |p(y)|\, dy, \quad \text{for all } |\alpha| \le k.$$

Finally let $p \in \mathcal{P}_k$ and E be a measurable subset of $B(x,r)$ with $\mathcal{L}^n(E) \ge Ar^n$. We denote by $z = T(y)$ the transformation

$$z = \frac{y-x}{r},$$

hence

$$\int_E |p(y)|\, dy = r^n \int_{T(E)} |p(x+rz)|\, dz.$$

On the other hand $T(E) \subseteq B(0,1)$, $\mathcal{L}^n(T(E)) \ge A$ and $p(x+rz) = \sum_{|\alpha|\le k} \frac{r^{|\alpha|} D^\alpha p(x)}{\alpha!} z^\alpha$. So putting everything together we conclude

$$|D^\alpha p(x)| \le \frac{\alpha!}{r^{n+|\alpha|}\gamma(A)} \int_E |p(y)|\, dy \le C(n,k,A) \frac{1}{r^{n+|\alpha|}} \int_E |p(y)|\, dy$$

□

Proof of Lemma 1.10.

$(i) \Rightarrow (ii)$: This implication is straightforward. The same points for which (i) holds make (ii) true. Indeed, if a polynomial that gives (i) centered at some x is

$$p_k(x;y) = \sum_{|\alpha|\le k} \frac{p_\alpha(x)}{\alpha!}(y-x)^\alpha,$$

we take $p_{k-1}(x;y) = \sum_{|\alpha|\le k-1} \frac{p_\alpha(x)}{\alpha!}(y-x)^\alpha$ and let $\lambda = 1 + \sum_{|\alpha|=k} \frac{|p_\alpha(x)|}{\alpha!} < \infty$. Since

$$\frac{|f(y) - p_{k-1}(x;y)|}{|y-x|^k} \le \frac{|f(y) - p_k(x;y)|}{|y-x|^k} + \sum_{|\alpha|=k} \frac{|p_\alpha(x)|}{\alpha!},$$

we have that

$$\left\{y \in \mathbb{R}^n : |f(y) - p_k(x;y)| \le |y-x|^k\right\} \subseteq \left\{y \in \mathbb{R}^n : |f(y) - p_{k-1}(x;y)| \le \lambda |y-x|^k\right\},$$

hence

$$\frac{\mathcal{L}^n\left(B(x,r) \cap \left\{y \in \mathbb{R}^n : \frac{|f(y)-p_{k-1}(x;y)|}{|y-x|^k} \le \lambda\right\}\right)}{\mathcal{L}^n(B(x,r))} \ge \frac{\mathcal{L}^n\left(B(x,r) \cap \left\{y \in \mathbb{R}^n : \frac{|f(y)-p_k(x;y)|}{|y-x|^k} \le 1\right\}\right)}{\mathcal{L}^n(B(x,r))}.$$

By our hypothesis (i),

$$\lim_{r\to 0} \frac{\mathcal{L}^n\left(B(x,r) \cap \left\{y \in \mathbb{R}^n : \frac{|f(y)-p_k(x;y)|}{|y-x|^k} \le 1\right\}\right)}{\mathcal{L}^n(B(x,r))} = 1,$$

which implies that

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y) - p_{k-1}(x;y)|}{|y-x|^k} \le \lambda < +\infty.$$

(iii) : Recall that the approximate $(k-1)-$Taylor polynomials are unique so we may use the notation of (1.1.3). The idea of the proof consists in splitting $\mathbb{R}^n \setminus N$ into a countable union of sets $\{B_j\}_{j\geq 1}$, on each of which, with the help of De Giorgi's lemma, we can apply Theorem 1.2. We will show that for each $j \in \mathbb{N}$,

$$\begin{cases} |f_\alpha(y) - D^\alpha p_{k-1}(x;y)| \leq M|x-y|^{k-|\alpha|}, \;\; \forall x,y \in B_j, \; |y-x| \leq \frac{1}{j}, \; |\alpha| \leq k-1 \\ |f_\alpha(x)| \leq j, \;\; \forall x \in B_j \end{cases}$$

where $p_{k-1}(x;y)$ is the polynomial of degree at most $k-1$ that gives (ii) and M is a constant (to be fixed later on) that depends only on n, k and j.
Let us define

$$\rho := \frac{\mathcal{L}^n(B(x,|y-x|) \cap B(y,|y-x|))}{|y-x|^n}, \;\; x,y \in \mathbb{R}^n, \, x \neq y \, ;$$

$$W_j(x;r) := B(x,r) \setminus \left\{ y \in \mathbb{R}^n : \, |f(y) - p_{k-1}(x;y)| \leq j|y-x|^k \right\}, \;\; x \in \mathbb{R}^n, \; r > 0, \; j \in \mathbb{N}, \text{ and}$$

$$B_j := \left\{ x \in \mathbb{R}^n : \, \mathcal{L}^n(W_j(x;r)) \leq \rho \frac{r^n}{4} \text{ for all } r \leq \frac{1}{j} \right\} \cap \{ x \in \mathbb{R}^n : \, |f_\alpha(x)| \leq j \text{ for all } |\alpha| \leq k-1 \} \, .$$

Note that ρ only depends on n. Since f is measurable we have that $W_j(x,r)$ are measurable sets. It is immediately checked that B_j is an increasing sequence of sets and

$$\bigcup_{j=1}^{\infty} B_j = \mathbb{R}^n \setminus N.$$

For us it will not be important that the sets B_j and the coefficients f_α are measurable, although they are indeed so (see Liu and Tai's proof for the delicate induction argument that allows one to show this).
Now, given $j \in \mathbb{N}$, consider two different points $x, y \in B_j$ with $|y-x| \leq \frac{1}{j}$, and for $r = |y-x|$ let

$$S(x,y;r,j) := [B(x,r) \cap B(y,r)] \setminus [W_j(x;r) \cup W_j(y;r)] \, ,$$

which is measurable. Moreover,

$$\mathcal{L}^n(S(x,y;r,j)) \geq \mathcal{L}^n(B(x,r) \cap B(y,r)) - \mathcal{L}^n(W_j(x;r)) - \mathcal{L}^n(W_j(y;r)) \geq \rho \frac{r^n}{2} > 0.$$

If we take $z \in S(x,y;r,j)$ then we have for $q(z) = p_{k-1}(y;z) - p_{k-1}(x;z)$ the estimate

$$|q(z)| \leq |p_{k-1}(x;z) - f(z)| + |f(z) - p_{k-1}(y;z)| \leq j(|z-x|^k + |y-z|^k) \leq 2jr^k.$$

We now apply Lemma 1.11 with $E = S(x,y;r,j)$ and the constant $A = \rho/2$, which depends only on n, to obtain that for each multi-index $|\alpha| \leq k-1$,

$$|D^\alpha q(y)| = |f_\alpha(y) - D^\alpha p_{k-1}(x;y)| \leq \frac{C}{r^{n+|\alpha|}} \int_{S(x,y;r,j)} |q(z)| \, dz \leq 2jw_n C r^{k-|\alpha|},$$

where w_n is the volume of the unit ball in $\mathbb{R}^n$ and C is the constant in Lemma 1.11, which depends only on n and k (because $A = \rho/2$ only depends on n).
We see now that for x, y in B_j with $|y-x| \leq \frac{1}{j}$, the last estimate implies

$$\begin{cases} |f_\alpha(y) - D^\alpha p_{k-1}(x;y)| \leq M(n,k,j)|x-y|^{k-|\alpha|}, \;\; \forall \, |\alpha| \leq k-1 \\ |f_\alpha(x)| \leq j \end{cases}$$

and by applying the Uniform Whitney Extension Theorem 1.2 we are done. □

We will now present a variant of Lemma 1.10 where we allow the exponents of the denominator $|y-x|$ of the approximate limits to be real numbers, not necessarily integers. This change will allow us to get decompositions with $C^{k-1,t}$ functions, $t \in (0,1]$.

Lemma 1.12. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a measurable function, k a positive integer, $t \in (0,1]$ and N a subset of $\mathbb{R}^n$. Suppose that*

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y)-p_{k-1}(x;y)|}{|y-x|^{k-1+t}} < +\infty \text{ for all } x \in \mathbb{R}^n \setminus N. \tag{1.2.6}$$

Then there exists a decomposition

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_j \cup N$$

such that for each $j \in \mathbb{N}$ there exists $g_j \in C^{k-1,t}(\mathbb{R}^n;\mathbb{R}^m)$ with $f_\alpha(x) = D^\alpha g_j(x)$ for all $x \in B_j$ and $|\alpha| \le k-1$.

Proof. The proof is exactly the same as that of Lemma 1.10, until the point where we use the Whitney Extension Theorem 1.2. In this case we have that for each $j \in \mathbb{N}$ and for all $x, y \in B_j$ with $|y-x| \le \frac{1}{j}$,

$$\begin{cases} |f_\alpha(y) - D^\alpha p_{k-1}(x;y)| \le M(n,k,j)|x-y|^{k-1+t-|\alpha|}, \ \forall\, |\alpha| \le k-1 \\ |f_\alpha(x)| \le j. \end{cases}$$

At this point we use Theorem 1.2 with $\omega(s) = s^t$ and we conclude similarly. □

Lemma 1.10 and Lemma 1.12 will be useful for the proof of our versions of the Morse-Sard theorem (Theorem 1.21 and 1.23) where the exceptional sets are countably $(\mathcal{H}^s, s)$ rectifiable of certain class C^k (see the statements of the results for details). However, in order to achieve a result where we are allowed to work with s-σ-finite exceptional sets (Theorem 1.25), it will also be necessary to have the following lemma at our disposal. We use once again the ideas of Liu and Tai [121], together with those of Whitney [153, Theorem 1].

Lemma 1.13. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Borel function, k a positive integer and $s > 0$. Suppose that f is approximately differentiable of order k at $\mathcal{H}^s$−almost every point $x \in B$, where $B \subseteq \mathbb{R}^n$ is a Borel set and $\mathcal{H}^s(B) < \infty$. Suppose also that the coefficients $f_\alpha(x)$, $|\alpha| \le k$, are Borel functions. Then there exists a decomposition*

$$B = \bigcup_{j=1}^{\infty} B_j \cup N,$$

where for each $j \in \mathbb{N}$ there exists $g_j \in C^k(\mathbb{R}^n;\mathbb{R}^m)$ with $f_\alpha(x) = D^\alpha g_j(x)$ for all $x \in B_j$, $|\alpha| \le k$, and $\mathcal{H}^s(N) = 0$.

Proof. Without loss of generality let us suppose that f is approximately differentiable of order k for all $x \in B$.

Let $\rho > 0$ be as in the proof of Lemma 1.11; recall that ρ only depends on n. For each $\eta > 0$, $i \in \mathbb{N}$, $x \in B$, define

$$W_\eta(x;i) := B(x;\frac{1}{i}) \setminus \left\{ y \in B : |f(y) - p_k(x;y)| \le \eta |y-x|^k \right\}.$$

Since f is Borel, these sets are Borel measurable. Consider the set

$$T = \left\{ (x,y) \in B \times B : |y-x| < \frac{1}{i}, |f(y) - p_k(x;y)| > \eta |y-x|^k \right\}.$$

All the f_α are Borel functions, so T is a Borel measurable set in $B \times B$. It is clear that $W_\eta(x;i) = \{y \in B : (x,y) \in T\}$, hence from Fubini's Theorem (see e.g. [54, Proposition 5.1.3.]) it follows that $\mathcal{L}^n(W_\eta(x;i))$ is a Borel measurable function of x.
We know that for all $\eta > 0$ and $x \in B$,

$$\lim_{i\to\infty} \frac{\mathcal{L}^n(W_\eta(x;i))}{\mathcal{L}^n(B(x;\frac{1}{i}))} = 0. \tag{1.2.7}$$

Define

$$\phi_i(x) := \inf\left\{\eta > 0 : \mathcal{L}^n(W_\eta(x;i)) < \frac{\rho}{4}\left(\frac{1}{i}\right)^n\right\}.$$

for each $i \in \mathbb{N}$ and $x \in B$. Observe that for fixed x and i, $\mathcal{L}^n(W_\eta(x;i))$ is Borel measurable, decreasing in η and continuous on the left. Thus

$$\phi_i(x) < \eta \text{ if and only if } \mathcal{L}^n(W_\eta(x;i)) < \frac{\rho}{4}\left(\frac{1}{i}\right)^n \tag{1.2.8}$$

and we have that $\phi_i(x)$ is a Borel measurable function.
From (1.2.7) and (1.2.8) it also follows that

$$\lim_{i\to\infty} \phi_i(x) = 0 \text{ for every } x \in B.$$

Now, with the goal of getting uniform convergence (up to a small enough set) in the previous limit, we want to apply Egorov's theorem for the measure $\mathcal{H}^s$. Notice that we are allowed to do so because the functions ϕ_i are Borel, the set B is $\mathcal{H}^s-$finite, and $\mathcal{H}^s$ is a Borel measure. We thus obtain, for each $j \in \mathbb{N}$, a closed[1] set $B_j \subseteq B$ such that $\mathcal{H}^s(B \setminus B_j) < \frac{1}{j}$ and $\lim_{i\to\infty} \phi_i(x) = 0$ uniformly on B_j.
Observe that $\mathcal{H}^s\left(B \setminus \bigcup_{j=1}^\infty B_j\right) = 0$. Let us call $N = B \setminus \bigcup_{j=1}^\infty B_j$.
Now we just have to see that for each of these sets B_j we can apply Theorem 1.3 in order to get a C^k extension to the whole space. Fix $j \in \mathbb{N}$ and a multi-index $|\alpha| \le k$. We have to prove that for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f_\alpha(y) - D^\alpha p_k(x;y)| \le \varepsilon |y-x|^{k-|\alpha|} \text{ if } x, y \in B_j,\ |y-x| < \delta.$$

Let then $\varepsilon > 0$. If $C > 0$ denotes the constant of De Giorgi's Lemma 1.11, applied with $A = \rho/2$, by the uniform convergence of ϕ_i on B_j we choose $i_0 \in \mathbb{N}$ such that

$$\phi_i(x) < \frac{\varepsilon}{2w_n C(1+\varepsilon)} := \varepsilon_0 \text{ for all } x \in B_j \text{ and all } i \ge i_0.$$

Also take $i_1 \ge i_0$ sufficiently large such that $\left(1 + \frac{1}{i_1}\right)^k \le 1 + \varepsilon$. Using (1.2.8) we have

$$\mathcal{L}^n\left(W_{\varepsilon_0}(x;i)\right) < \frac{\rho}{4}\left(\frac{1}{i}\right)^n \quad \text{for all } x \in B_j \text{ and all } i \ge i_1.$$

Take $\delta < \frac{1}{i_1}$ and $x, y \in B_j$ with $|y-x| < \delta$. There exists $i_2 \ge i_1$ such that $\frac{1}{i_2+1} \le |y-x| \le \frac{1}{i_2}$. Now consider the set $S\left(x,y;i_2,\varepsilon_0\right) = \left[B(x,\frac{1}{i_2}) \cap B(y,\frac{1}{i_2})\right] \setminus [W_{\varepsilon_0}(x;i_2) \cup W_{\varepsilon_0}(y;i_2)]$. One can check that the following inclusion holds

$$B\left(x, \frac{1}{i_2}\right) \cap B\left(x + \frac{y-x}{i_2|y-x|}, \frac{1}{i_2}\right) \subseteq B(x,\frac{1}{i_2}) \cap B(y,\frac{1}{i_2}),$$

[1] Egorov's theorem in general would give us Borel sets B_j, but the Hausdorff measures $\mathcal{H}^s$ are Borel regular measures, so it is well known that for every Borel set A, if $\mathcal{H}^s(A) < \infty$ there exists for each $\varepsilon > 0$ a closed set C such that $C \subseteq A$ and $\mathcal{H}^s(A \setminus C) < \varepsilon$. This fact cannot be overlooked because our use of Theorem 1.3 forces us to work with closed sets.

and by the definition of ρ we have that

$$\mathcal{L}^n\left(S\left(x,y;i_2,\varepsilon_0\right)\right) \geq \mathcal{L}^n\left(B\left(x,\frac{1}{i_2}\right)\cap B\left(x+\frac{y-x}{i_2|y-x|},\frac{1}{i_2}\right)\right)-\frac{\rho}{2}\left(\frac{1}{i_2}\right)^n=\frac{\rho}{2}\left(\frac{1}{i_2}\right)^n>0.$$

Observe that for every $z\in S\left(x,y;i_2,\varepsilon_0\right)$,

$$|z-x|^k,|z-y|^k\leq\left(\frac{1}{i_2}\right)^k\leq\left(\left(1+\frac{1}{i_2}\right)|y-x|\right)^k\leq\left(1+\frac{1}{i_1}\right)^k|y-x|^k\leq(1+\varepsilon)|y-x|^k,$$

where in the second inequality we have used that $\frac{1}{i_2+1}\leq|y-x|$. Using De Giorgi's Lemma 1.11 for the set $S(x,y;i_2,\varepsilon_0)$, the constant $A=\rho/2$ and the polynomial $q(z)=p_k(y;z)-p_k(x;z)$, we conclude that

$$\begin{aligned}|D^\alpha q(y)|=|f_\alpha(y)-D^\alpha p_k(x;y)|&\leq\frac{C}{(1/i_2)^{n+|\alpha|}}\int_{S(x,y;i_2,\varepsilon_0)}|q(z)|\,dz\\&\leq\frac{C}{(1/i_2)^{n+|\alpha|}}\int_{S(x,y;i_2,\varepsilon_0)}\varepsilon_0\left(|z-y|^k+|z-x|^k\right)\,dz\\&\leq C\varepsilon_0\left(\frac{\mathcal{L}^n\left(S\left(x,y;i_2,\varepsilon_0\right)\right)}{(1/i_2)^n}\right)2(1+\varepsilon)\left(\frac{|y-x|^k}{(1/i_2)^{|\alpha|}}\right)\\&\leq C\varepsilon_0 w_n 2(1+\varepsilon)|y-x|^{k-|\alpha|}=\varepsilon|x-y|^{k-|\alpha|}.\end{aligned}$$

By applying the Whitney Extension Theorem 1.3 the proof is complete. □

1.3 A Morse-Sard theorem for approximate differentiable functions

Our aim is to prove a Morse-Sard theorem for functions that only are approximately differentiable of order k or that have approximate $(k-1)-$Taylor polynomials on some sets. Consequently we will need to deal with weaker notions of derivatives and critical sets.

Definition 1.14. Let $f:\mathbb{R}^n\to\mathbb{R}^m$ be a measurable function that is approximately differentiable of order k at x, with unique approximate polynomial

$$p_k(x)=\sum_{|\alpha|\leq k}\frac{f_\alpha(x)}{\alpha!}(y-x)^\alpha.$$

For each multi-index α, $|\alpha|\leq k$, we define the *$\alpha-$th differential coefficient* of f at x as $f_\alpha(x)$. If f has an approximate $(k-1)-$Taylor polynomial at x we can only define the $\alpha-$th differential coefficient $f_\alpha(x)$ for $|\alpha|\leq k-1$.

Recall that we do not necessarily have $f_\alpha(x)=D^\alpha f(x)$ in any usual sense, and $D^\alpha f(x)$ may not even exist.

From now on we will use the following notation

$$p_k(x;y)=\sum_{|\alpha|\leq k}\frac{f_\alpha(x)}{\alpha!}(y-x)^\alpha=f(x)+F_1(x)(y-x)+\cdots\frac{F_k(x)}{(k)!}(y-x)^k.$$

where the $F_j(x):(\mathbb{R}^n)^j\to\mathbb{R}^m$ are the $j-$multilinear and symmetric maps whose coefficients with respect to the standard basis of $\mathbb{R}^n$ are given by $f_\alpha(x)$, $|\alpha|=j$ $(j=1,\dots,k)$. Namely, given $z\in\mathbb{R}^n$, $F_j(x)z^j$ means

$$F_j(x)z^j=F_j(x)(z,\overbrace{\dots}^{j)},z)=\sum_{|\alpha|=j}f_\alpha(x)z^\alpha.$$

Again we stress that we do not necessarily have $F_j(x)=D^jf(x)$, and the latter may not exist. However, if a function is one time differentiable at x in the usual sense, we do have $Df(x)=F_1(x)$.

We are now in a position to introduce a generalized notion of critical set.

Definition 1.15. Let AppDiff(f) denote the set of points where a function f is approximately differentiable of order 1. We define

$$C_f := \{x \in \text{AppDiff}(f) : \text{rank}\,(F_1(x)) \text{ is not maximum}\}.$$

Remark 1.16. If a function f only has an approximate $(0)-$Taylor polynomial at almost every point of $\mathbb{R}^n$ we can still define the set of critical points up to an $\mathcal{L}^n-$null set. According to Liu and Tai's result [121, Theorem 1], f is approximately differentiable of order 1 almost everywhere, so we consider the coefficients of the linear part of the corresponding polynomial.

Definition 1.17 ($(N_0)-$property)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be measurable and suppose AppDiff(f) is nonempty. We say that f satisfies the $(N_0)-$property with respect to the Hausdorff measure $\mathcal{H}^s$, $s \in (0, n]$, if and only if

$$E \subseteq C_f, \;\; \mathcal{H}^s(E) = 0 \Rightarrow \; \mathcal{L}^m(f(E)) = 0.$$

Definition 1.18 (Lusin's $N-$property)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be measurable. We say that f satisfies Lusin's $N-$property with respect to the Hausdorff measure $\mathcal{H}^s$, $s > 0$, if and only if

$$E \subseteq \mathbb{R}^n, \;\; \mathcal{H}^s(E) = 0 \Rightarrow \; \mathcal{H}^s(f(E)) = 0.$$

The following theorem, due to Norton [130, Theorem 2], will also be an important ingredient in our proofs of Theorem 1.21, 1.23 and 1.25.

Theorem 1.19 (Norton)**.** *Let k be a positive integer, $t \in (0, 1]$ and $f : \mathbb{R}^n \to \mathbb{R}^m$.*

(i) If $f \in C^{k,t}$ and $E \subseteq C_f$ is $\mathcal{H}^{k+t+m-1}-$null, then $\mathcal{L}^m(f(E)) = 0$. That is to say, f has the $(N_0)-$property with respect to the measure $\mathcal{H}^{k+t+m-1}$.

(ii) If $f \in C^k$ and $E \subseteq C_f$ is $(k + m - 1)$-σ-finite, then $\mathcal{L}^m(f(E)) = 0$

We will also need to use Bates's version of the Morse-Sard theorem for $C^{n-m,1}$ (see [31, Theorem 2]).

Theorem 1.20 (Bates)**.** *Let n, m be positive integers with $m \leq n$ and $f : \mathbb{R}^n \to \mathbb{R}^m$. If $f \in C^{n-m,1}(\mathbb{R}^n; \mathbb{R}^m)$, then the set of critical values of f has $\mathcal{L}^m-$measure zero.*

The first of our main results is as follows.

Theorem 1.21. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$. Suppose that for every $x \in \mathbb{R}^n$ and $j = 1 \ldots, n - m$ there exist $j-$multilinear and symmetric mappings $F_j(x)$ such that*

(a)

$$\text{ap}\limsup_{y \to x} \frac{|f(y) - f(x)|}{|y - x|} < +\infty \;\; \textit{for all } x \in \mathbb{R}^n \setminus N_0,$$

where N_0 is a countable set.

(b) The polynomial $p_1(x; y) = f(x) + F_1(x)(y - x)$ centered at x satisfies that the set

$$N_1 := \{x \in \mathbb{R}^n : \text{ap}\limsup_{y \to x} \frac{|f(y) - p_1(x; y)|}{|y - x|^2} = +\infty\}$$

is $\mathcal{H}^m-$null.

(c) For each $i = 2, \ldots, n - m$ the polynomial $p_i(x; y) = f(x) + \sum_{j=1}^{i} \frac{F_j(x)}{j!}(y - x)^j$ centered at x satisfies that the set

$$N_i := \{x \in \mathbb{R}^n : \text{ap}\limsup_{y \to x} \frac{|f(y) - p_i(x; y)|}{|y - x|^{i+1}} = +\infty\}$$

is countably $(\mathcal{H}^{i+m-1}, i + m - 1)$ rectifiable of class C^i.

Then $\mathcal{L}^m(f(C_f)) = 0$.
The same is true if we replace $\mathbb{R}^n$ *with an open subset* U *of* $\mathbb{R}^n$.

[Observe that if $n = m$ we only have (a) and if $n = m + 1$ we only have (a) and (b).]

Proof. Note that (a) tells us that f is approximately continuous $\mathcal{L}^n-$almost everywhere so the function f is measurable.
Let us also set some notation by writing each exceptional set N_i $(i = 1, \dots, n - m)$ as

$$N_i = A_i \cup \bigcup_{k=1}^{\infty} A_{i,k}, \tag{1.3.1}$$

where $\mathcal{H}^{i+m-1}(A_i) = 0$ and each subset $A_{i,k} \subseteq \mathbb{R}^n$ is an $(i + m - 1)-$dimensional submanifold of class C^i.
First of all let us show that condition (a) implies Lusin's $N-$condition with respect to the Hausdorff measure $\mathcal{H}^s$, $s \in (0, m]$. In fact we will see that there exists a collection $\{B_{0,j}\}_{j=1}^{\infty}$ with $\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_{0,j} \cup N_0$ and such that each restriction $f|_{B_{0,j}}$ is locally Lipschitz with constant $2j$. Since N_0 is countable, $\mathcal{H}^s(f(N_0)) = 0$, and this readily implies that the image of sets of $s-$Hausdorff measure zero $(s \leq m)$ has $s-$Hausdorff measure zero. The following argument is again inspired by Liu and Tai's result ([121, Theorem 1]). Consider the sets

$$W_{0,j}(x; r) = B(x, r) \setminus \{y \in \mathbb{R}^n : |f(y) - f(x)| \leq j|y - x|\}, \quad x \in \mathbb{R}^n, \ r > 0, \ j \in \mathbb{N}$$
$$B_{0,j} = \left\{x \in \mathbb{R}^n : \mathcal{L}^n(W_{0,j}(x; r)) \leq \rho\frac{r^n}{4} \text{ for all } r \leq \frac{1}{j}\right\} \cap \{x \in \mathbb{R}^n : |f(x)| \leq j\}$$
$$S(x, y; r, j) = [B(x, r) \cap B(y, r)] \setminus [W_{0,j}(x; r) \cup W_{0,j}(y; r)]$$

similarly to the proof of Lemma 1.10 above. We take $x, y \in B_{0,j}$, $|y - x| \leq \frac{1}{j}$, $r = |y - x|$. Using that $\mathcal{L}^n(S(x, y; r, j)) > 0$ it is possible to take $z \in S(x, y; r, j)$ and then

$$|f(y) - f(x)| \leq |f(y) - f(z)| + |f(z) - f(x)| \leq j|z - y| + j|z - x| \leq 2j|y - x|.$$

We have just shown that $f|_{B_{0,j}}$ is locally $2j-$Lipschitz. Now it is clear that, for every $s \in (0, m]$, if $\mathcal{H}^s(A) = 0$, $A \subseteq \mathbb{R}^n$, then $\mathcal{H}^s(f(A)) = 0$. Therefore the set of points where f is not approximately differentiable (hence C_f cannot be defined), which belongs to N_1 and has $\mathcal{H}^m-$measure zero, has $\mathcal{L}^m-$null image.
Let us make a pause to comment on the special case $n = m$ (we only have condition (a)). In this case we also have the critical set of points defined up to a set of $\mathcal{L}^n-$measure zero (see Remark 1.16). Moreover Liu and Tai's result [121, Theorem 1] asserts in particular that

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} D_j \cup M,$$

where $\mathcal{L}^n(M) = 0$ and such that for each $j \in \mathbb{N}$ there is a function $g_j \in C^1(\mathbb{R}^n; \mathbb{R}^n)$ with

$$D_j \subseteq \{x \in \mathbb{R}^n : f(x) = g_j(x), \, F_1(x) = Dg_j(x)\}.$$

Using the classical Morse-Sard theorem we have that for every $j \in \mathbb{N}$,

$$\mathcal{L}^n(f(C_f \cap D_j)) = \mathcal{L}^n(f|_{D_j}(C_f \cap D_j)) = \mathcal{L}^n(g_j(C_{g_j} \cap D_j)) = 0.$$

Consequently

$$\mathcal{L}^n(f(C_f)) \leq \sum_{j=1}^{\infty} \mathcal{L}^n(f(C_f \cap D_j)) + \mathcal{L}^n(f(M)) = 0.$$

► Step 1: Condition (c) with $i = n - m$ tells us that f has an approximate $(n - m)-$Taylor polynomial for every $x \in \mathbb{R}^n \setminus N_{n-m}$, so we can use Lemma 1.10 and write

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_{n-m,j} \cup N_{n-m},$$

in such a way that for each $j \in \mathbb{N}$ there is a function $g_j \in C^{n-m,1}$ with

$$B_{n-m,j} \subseteq \{x \in \mathbb{R}^n : \ f(x) = g_j(x), \ F_1(x) = Dg_j(x)\}.$$

We decompose C_f as

$$C_f = \left(\bigcup_{j=1}^{\infty} C_f \cap B_{n-m,j} \right) \cup (C_f \cap N_{n-m}).$$

Using Bates's result (Theorem 1.20) we have that for every $j \in \mathbb{N}$,

$$\mathcal{L}^m(f(C_f \cap B_{n-m,j})) = \mathcal{L}^m(f|_{B_{n-m,j}}(C_f \cap B_{n-m,j})) = \mathcal{L}^m(g_j(C_{g_j} \cap B_{n-m,j})) = 0.$$

By the subadditivity of the Lebesgue measure we have thus reduced our problem to showing that $\mathcal{L}^m(f(C_f \cap N_{n-m})) = 0$.

► Step 2: We now work with the condition (c) but for the case $i = n - m - 1$. By applying Lemma 1.10 again, we obtain a decomposition

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_{n-m-1,j} \cup N_{n-m-1},$$

where for each $j \in \mathbb{N}$ there is a function $g_j \in C^{n-m-1,1}$ with

$$B_{n-m-1,j} \subseteq \{x \in \mathbb{R}^n : \ f(x) = g_j(x), \ F_1(x) = Dg_j(x)\}.$$

Recall that $N_{n-m} = A_{n-m} \cup \bigcup_{k=1}^{\infty} A_{n-m,k}$ where $\mathcal{H}^{n-1}(A_{n-m}) = 0$ and such that there exist maps

$$\phi_{n-m,k} : \mathbb{R}^{n-1} \longrightarrow A_{n-m,k} \subseteq \mathbb{R}^n$$

of class C^{n-m} for each $k \in \mathbb{N}$.
We now write $C_f \cap N_{n-m}$ as

$$C_f \cap N_{n-m} = \left(\bigcup_{j=1}^{\infty} C_f \cap A_{n-m} \cap B_{n-m-1,j} \right) \cup \left(\bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} C_f \cap A_{n-m,k} \cap B_{n-m-1,j} \right) \cup (C_f \cap N_{n-m-1}). \tag{1.3.2}$$

Remember that $\mathcal{H}^{n-1}(A_{n-m}) = 0$, so by using Norton Theorem 1.19 (i), for every $j \in \mathbb{N}$ we get

$$\begin{aligned}\mathcal{L}^m(f(C_f \cap A_{n-m} \cap B_{n-m-1,j})) &= \mathcal{L}^m(f|_{B_{n-m-1,j}}(C_f \cap A_{n-m} \cap B_{n-m-1,j}) \\ &= \mathcal{L}^m(g_j(C_{g_j} \cap A_{n-m} \cap B_{n-m-1,j})) = 0.\end{aligned}$$

Fix $j, k \in \mathbb{N}$. We now see that $\mathcal{L}^m(f(C_f \cap A_{n-m,k} \cap B_{n-m-1,j})) = 0$. It is easy to check that $C_{g_j} \cap A_{n-m,k} \subseteq \phi_{n-m,k}(C_{g_j \circ \phi_{n-m,k}})$, so

$$\begin{aligned}\mathcal{L}^m(f(C_f \cap A_{n-m,k} \cap B_{n-m-1,j})) &= \mathcal{L}^m(g_j(C_{g_j} \cap A_{n-m,k} \cap B_{n-m-1,j})) \\ &\leq \mathcal{L}^m(g_j(\phi_{n-m,k}(C_{g_j \circ \phi_{n-m,k}}) \cap B_{n-m-1,j})) \\ &= \mathcal{L}^m(g_j \circ \phi_{n-m,k}|_{\phi^{-1}_{n-m,k}(B_{n-m-1,j} \cap A_{n-m,k})}(C_{g_j \circ \phi_{n-m,k}})) = 0,\end{aligned}$$

where in the last equality we have used Bates's theorem 1.20) applied to the function

$$g_j \circ \phi_{n-m,k}|_{\phi^{-1}_{n-m,k}(B_{n-m-1,j}\cap A_{n-m,k})} : \mathbb{R}^{n-1} \longrightarrow \mathbb{R}^m,$$

which is of class $C^{n-m-1,1}_{\text{loc}}(\mathbb{R}^{n-1};\mathbb{R}^m)$.
Therefore, by the subadditivity of the Lebesgue measure and (1.3.2), our problem reduces to checking that $\mathcal{L}^m(f(C_f \cap N_{n-m-1})) = 0$.

► Final step: Reasoning in the same way for the cases $i = n-m-2, \dots, i = 1$, we inductively arrive to the conclusion that it is enough to prove that $\mathcal{L}^m(f(C_f \cap N_1)) = 0$, which now follows by using Lusin's $N-$condition with respect to the measure $\mathcal{H}^m$. □

Remark 1.22. It is clear that the above proof can be adapted to get a similar result in which condition (b) is dropped and condition (c) now holds for $i = 1, 2, \dots, n-m$. In principle this result is more general than Theorem 1.21. The reason for our statement of Theorem 1.21 is that one of the typical applications of Morse-Sard-type theorems is ensuring that for almost every $y \in \mathbb{R}^m$ the set $f^{-1}(y)$ is regular enough (for instance, it is a C^1 manifold if f is assumed to be C^1), a property that we would lose if we do not require condition (b).

Let us consider now the simpler case where the exceptional sets N_i are $\mathcal{H}^{i+m-1}-$null for $i = 1, \dots, n-m$ (i.e $A_{i,k} = \emptyset$ for all $k \geq 1$ in equation (1.3.1)). We establish an alternate version of the preceding result, in which we let the exponents of the denominators $|y-x|$ be smaller, and not integers. In return we must ask these limits to be finite in larger sets in order to achieve the Morse-Sard property. The arguments will be the same, but for our use of Lemma 1.12 instead of Lemma 1.10.

Theorem 1.23. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $m \leq n$.*

If $n > m+1$, for each $i = 0, \dots, n-m-1$ choose numbers $s(i) \in (i, i+1]$ and suppose that for every $x \in \mathbb{R}^n$ and $j = 1 \dots, n-m$ there exist $j-$multilinear and symmetric mappings $F_j(x)$ such that

(a) For $i = 0$,

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y)-f(x)|}{|y-x|^{s(0)}} < +\infty \quad \textit{for all } x \in \mathbb{R}^n \setminus N_0, \textit{ where } N_0 \textit{ is countable.}$$

(b) For $i = 1$, the polynomial $p_1(x;y) = f(x) + F_1(x)(y-x)$ centered at x satisfies that

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y)-p_1(x;y)|}{|y-x|^{s(1)}} < +\infty \quad \textit{for } \mathcal{H}^{s(0)m} - \textit{almost every } x \in \mathbb{R}^n.$$

(c) For each $i = 2, \dots, n-m-1$, the polynomial $p_i(x;y) = f(x) + \sum_{j=1}^{i} \frac{F_j(x)}{j!}(y-x)^j$ centered at x satisfies that

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y)-p_i(x;y)|}{|y-x|^{s(i)}} < +\infty \quad \textit{for } \mathcal{H}^{s(i-1)+m-1} - \textit{almost every } x \in \mathbb{R}^n.$$

(If $n = m+2$ we do not have this condition).

(d) The polynomial $p_{n-m}(x;y) = f(x) + \sum_{j=1}^{n-m} \frac{F_j(x)}{j!}(y-x)^j$ centered at x satisfies that

$$\operatorname{ap}\limsup_{y\to x} \frac{|f(y)-p_{n-m}(x;y)|}{|y-x|^{n-m+1}} < +\infty \quad \textit{for } \mathcal{H}^{s(n-m-1)+m-1} - \textit{almost every } x \in \mathbb{R}^n.$$

If $n = m+1$, after choosing $s(0) \in (0,1]$, suppose only (a) and (d), where the exceptional set in (d) must be $\mathcal{H}^{s(n-m-1)m} = \mathcal{H}^{s(0)m}-$null.

If $n = m$ suppose only that condition (a) with $s(0) = 1$ holds everywhere except perhaps on a countable set N_0.

Then $\mathcal{L}^m(f(C_f)) = 0$.

Proof. The case $n = m$ is exactly the same as in Theorem 1.21. Suppose then $n > m$.
We will first see that (a) implies that $\mathcal{L}^m(f(A)) = 0$ for every set $A \subseteq \mathbb{R}^n$ with $\mathcal{H}^{s(0)m}(A) = 0$. This implies that the points where f is not approximately differentiable, and hence C_f cannot be defined, is mapped into a set of $\mathcal{L}^m-$measure zero.
Again we employ arguments similar to previous proofs. We take the same decomposition of $\mathbb{R}^n$ as in Lemma 1.10, except that in this case we set

$$W_j(x;r) = B(x,r) \setminus \left\{y \in \mathbb{R}^n : |f(y) - f(x)| \le j|y-x|^{s(0)}\right\}, \ \ x \in \mathbb{R}^n, \ r > 0, \ j \in \mathbb{N}$$
$$B_j = \left\{x \in \mathbb{R}^n : \mathcal{L}^n(W_j(x;r)) \le \rho\frac{r^n}{4} \ \text{ for all } r \le \frac{1}{j}\right\} \cap \{x \in \mathbb{R}^n : |f(x)| \le j\}.$$

We have

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_j. \tag{1.3.3}$$

For $j \in \mathbb{N}$, we consider two different points $x, y \in B_j$ with $|y-x| \le \frac{1}{j}$, $r = |y-x|$. Using that $\mathcal{L}^n(S(x,y;r,j)) > 0$ we can take $z \in S(x,y;r,j)$ and then

$$|f(y)-f(x)| \le |f(y)-f(z)| + |f(z)-f(x)| \le j|z-y|^{s(0)} + j|z-x|^{s(0)} \le 2j|y-x|^{s(0)}.$$

So we obtain

$$|f(y)-f(x)|^m \le 2j|y-x|^{s(0)m}.$$

This immediately implies that $\mathcal{L}^m(f(A \cap B_j)) = 0$ for all j and all $A \subseteq \mathbb{R}^n$ with $\mathcal{H}^{s(0)m}(A) = 0$, hence, recalling (1.3.3), also that $\mathcal{L}^m(f(A)) = 0$ for all $A \subseteq \mathbb{R}^n$ with $\mathcal{H}^{s(0)m}(A) = 0$.

► Step 1: Condition (d) allows us to use Lemma 1.10 and find a decomposition

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_j \cup N_{n-m}, \ \ \mathcal{H}^{s(n-m-1)+m-1}(N_{n-m}) = 0,$$

such that for each $j \in \mathbb{N}$ there is a function $g_j \in C^{n-m,1}$ with

$$B_j \subseteq \{x \in \mathbb{R}^n : f(x) = g_j(x), \, F_1(x) = Dg_j(x)\}.$$

Hence by using Bates's result (Theorem 1.20) we have that for every $j \in \mathbb{N}$,

$$\mathcal{L}^m(f(C_f \cap B_j)) = \mathcal{L}^m(f|_{B_j}(C_f \cap B_j)) = \mathcal{L}^m(g_j(C_{g_j} \cap B_j)) = 0.$$

So we have reduced our problem to prove that all sets $E \subseteq C_f$ with $\mathcal{H}^{s(n-m-1)+m-1}-$measure zero satisfy $\mathcal{L}^m(f(E)) = 0$.

► Step 2: We work with the condition (c) for the case $i = n-m-1$. We apply Lemma 1.12 for the case $k-1+t = s(n-m-1)$ and $\mu = \mathcal{H}^{s(n-m-2)+m-1}$ to find a decomposition

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_j \cup N_{n-m-1}, \ \ \mathcal{H}^{s(n-m-2)+m-1}(N_{n-m-1}) = 0,$$

where for each $j \in \mathbb{N}$ there is a function $g_j \in C^{n-m-1, s(n-m-1)-n+m+1}$ with

$$B_j \subseteq \{x \in \mathbb{R}^n : f(x) = g_j(x), \, F_1(x) = Dg_j(x)\}.$$

We write a given set $E \subseteq C_f$, $\mathcal{H}^{s(n-m-1)+m-1}(E) = 0$ as

$$E = \bigcup_{j=1}^{\infty} (E \cap B_j) \cup (N \cap E)$$

where $\mathcal{H}^{s(n-m-2)+m-1}(N) = 0$. Now we use Norton's result (Theorem 1.19 (i)) and for every $j \in \mathbb{N}$,

$$\mathcal{L}^m(f(E \cap B_j)) = \mathcal{L}^m(f|_{B_j}(E \cap B_j)) = \mathcal{L}^m(g_j(C_{g_j} \cap E \cap B_j)) = 0.$$

Therefore we must only consider sets $E \subseteq C_f$ such that $\mathcal{H}^{s(n-m-2)+m-1}(E) = 0$ and check that $\mathcal{L}^m(f(E)) = 0$.

► Final step: Reasoning in the same way for the cases $i = n - m - 2, \dots, i = 1$, we arrive to the conclusion that it is enough to prove that sets $E \subseteq C_f$ with $\mathcal{H}^{s(0)m}$−measure zero satisfy $\mathcal{L}^m(f(E)) = 0$. But this follows from (a), as we have already seen. □

Remark 1.24. If we choose $s(i) = i + 1$ for each $i = 0, \dots, n - m - 1$ we get exactly Theorem 1.21 in the particular case that $\mathcal{H}^{i+m-1}(N_i) = 0$, $i = 1, \dots, n - m$.

Notice that in Theorem 1.23 a selection of numbers $s(j)$ make it nor stronger, neither weaker than any other possible choice.

It is a natural question to ask whether or not we can change our exceptional sets N_i $(i = 1, \dots, n - m)$ to be $(i + m - 1)$-σ-finite. This is the purpose of our next main result, in which we must work with the stronger notion of approximate differentiability instead of the property of having approximate $(k - 1)$−Taylor polynomials. The result will generalize Theorem 1.21 but in addition we will have to require that f and its differential coefficients f_α are Borel functions in order to use Lemma 1.13.

Theorem 1.25. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a Borel function, $m \le n$. Suppose that*

(a) $\operatorname{ap}\limsup_{y \to x} \frac{|f(y)-f(x)|}{|y-x|} < +\infty$ *for all $x \in \mathbb{R}^n \setminus N_0$, where N_0 is a countable set.*

(b) f is approximately differentiable of order 1 for all $x \in \mathbb{R}^n \setminus N_1$, where N_1 is $\mathcal{H}^m$−null.

(c) For each $i = 2, \dots, n - m$, f is approximately differentiable of order i for all $x \in \mathbb{R}^n \setminus N_i$, where N_i is $(i + m - 2)$-σ-finite. (If $n \le m + 1$ we do not have this condition).

(d) f has an approximate $(n - m)$−Taylor polynomial for all $x \in \mathbb{R}^n \setminus N_{n-m+1}$, where N_{n-m+1} is $(n - 1)$-σ-finite.[2]

Suppose also that the coefficients $f_\alpha(x)$, $|\alpha| \le n - m$, are Borel functions.
Then $\mathcal{L}^m(f(C_f)) = 0$.

Proof. The case $n = m$ is exactly the same as in Theorem 1.21 and Theorem 1.23.
Recall that condition (a) gives us the Lusin's N−condition with respect to the Hausdorff measure $\mathcal{H}^m$. In particular the set of points where C_f cannot be defined (that is N_1, which has $\mathcal{H}^m$−measure zero) has $\mathcal{L}^m$−null image.
► Step 1: Condition (d) allows us to reduce our problem to showing that $\mathcal{L}^m(f(C_f \cap N_{n-m+1})) = 0$ (here we use the same arguments as in Step 1 of Theorem 1.21).
Since the set N_{n-m+1} is $(n - 1)$-σ-finite, by the subadditivity of the Lebesgue measure, without loss of generality we can assume that $\mathcal{H}^{n-1}(N_{n-m+1}) < \infty$. Consequently we can focus on studying sets $A \subseteq C_f$ with $\mathcal{H}^{n-1}(A) < \infty$.

► Step 2: We now work with the condition (c) but for the case $i = n - m$.
Let us call $B = A \cap (\mathbb{R}^n \setminus N_{n-m})$. If we prove that $\mathcal{L}^m(f(B)) = 0$ it will only be needed to see that $\mathcal{L}^m(f(C_f \cap N_{n-m+1} \cap N_{n-m})) = 0$, or what is the same, $\mathcal{L}^m(f(C_f \cap N_{n-m})) = 0$ (note that we can suppose $N_{n-m} \subset N_{n-m+1}$).

[2]Observe that if $n = m$ we only have (a) (however this implies (b) using Liu and Tai's result [121, Theorem 1]) and if $n = m + 1$ we only have (a), (b) and (d).

We have that f is approximately differentiable of order $n-m$ everywhere on B and $\mathcal{H}^{n-1}(B) < \infty$. We can apply Lemma 1.13 and write

$$B = \bigcup_{j=1}^{\infty} B_j \cup N, \tag{1.3.4}$$

where for each $j \in \mathbb{N}$ there exists $g_j \in C^{n-m}(\mathbb{R}^n;\mathbb{R}^m)$ with $f_\alpha(x) = D^\alpha g_j(x)$ for all $x \in B_j$ ($|\alpha| \leq n-m$), and $\mathcal{H}^{n-1}(N) = 0$.
With the set N we proceed as in Theorem 1.21 when we had to deal with $\mathcal{H}^{n-1}-$null sets and we had to apply Lemma 1.10 together with Norton's Theorem 1.19 (i), and we conclude that $\mathcal{L}^m(f(N)) = 0$.
For the sets $B_j \subseteq C_f$ we use Norton's Theorem 1.19 (ii) with C^{n-m} regularity and we have

$$\mathcal{L}^m(f(B_j)) = \mathcal{L}^m(g_j(B_j)) = 0.$$

Therefore, by the subadditivity of the Lebesgue measure and (1.3.4), our problem reduces to checking if $\mathcal{L}^m(f(C_f \cap N_{n-m})) = 0$.

► Final step: Reasoning in the same way for the cases $i = n-m-2, \ldots, i = 1$, we inductively arrive to the conclusion that it is enough to prove that $\mathcal{L}^m(f(C_f \cap N_1)) = 0$, where $\mathcal{H}^m(N_1) = 0$. But this follows by (a).

□

1.4 Final considerations and examples

A key point in the above arguments is obtaining a nice splitting of the space $\mathbb{R}^n$ into a countable union of sets (plus a small enough exceptional set) such that our function has enough regularity on each of those sets. Following the same strategy, there is another well known property that allows a similar decomposition.

Definition 1.26. A measurable function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to have the Lusin property of order k with respect to the measure μ if for every $\varepsilon > 0$ there is a function $g \in C^k(\mathbb{R}^n;\mathbb{R}^m)$ such that

$$\mu\left(\{x \in \mathbb{R}^n : f(x) \neq g(x)\}\right) < \varepsilon.$$

It is clear that for such a function there always exist decompositions of the form

$$\mathbb{R}^n = \bigcup_{j=1}^{\infty} B_j \cup N,$$

where for each $j \in \mathbb{N}$ there is a function $g_j \in C^k(\mathbb{R}^n;\mathbb{R}^m)$ with $f_\alpha(x) = D^\alpha g_j(x)$ for all $x \in B_j$ ($|\alpha| \leq k$), and $\mu(N) = 0$. Therefore, if, instead of approximate differentiability in our conditions of Theorem 1.21, Theorem 1.23 and Theorem 1.25, we consider Lusin properties of the corresponding orders, with respect to the same Hausdorff measures, we may obtain the same conclusions. For example the analogue of Theorem 1.21 would be the following.

Theorem 1.27. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$, $m < n$, be locally Lipschitz. Suppose that for each $i = 2, \ldots, n-m+1$, f has the Lusin property of order i with respect to the measure $\mathcal{H}^{i+m-2}$. Then we have $\mathcal{L}^m(f(C_f)) = 0$.*

For the proof we just mention that local Lipschitzness gives us the Lusin's $N-$property with respect to the measure $\mathcal{H}^m$, and that we must use the classical Morse-Sard theorem instead of Bates's result.

However, we cannot deduce Theorems 1.21, 1.23 and 1.25 from Theorem 1.27, because, to the best of our knowledge, the problem whether an $\mathcal{H}^s-$almost everywhere approximately differentiable function of order j must have the Lusin property of order j (or a $C^{j-1,1}$ Lusin type property) with respect to the measure $\mathcal{H}^s_\infty$ or $\mathcal{H}^s$ is open. The proof of [121] cannot be adapted to the measures $\mathcal{H}^s_\infty$ or $\mathcal{H}^s$ $(s < n)$.

We will finally comment on three examples that illustrate how Theorem 1.21 covers functions for which none of the previous Morse-Sard type results that exist in the literature can be applied to, and also how Theorem 1.21 and Theorem 1.25 are sharp in the following sense: for each $t \in (0, 1]$ we can always find a function $f : \mathbb{R}^n \to \mathbb{R}$ of class C^{n-1} which has an approximate $(n-1)-$Taylor polynomial everywhere on $\mathbb{R}^n$ except on a set N of Hausdorff dimension $n - 1 + t$, but which does not satisfy the Morse-Sard theorem.

1. We first note that Theorem 1.21 is not weaker, nor stronger than the recent Bourgain-Korobkov-Kristensen generalizations [44, 113, 114] of the Morse-Sard theorem in the case of real-valued functions for the spaces $W^{n,1}_{loc}(\mathbb{R}^n; \mathbb{R})$ and $BV_{n,loc}(\mathbb{R}^n)$. The following example, taken from [18, p. 18],
$$f(x, y) = \begin{cases} x^4 \sin\left(\frac{1}{x^2}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$
shows that there are functions $f : \mathbb{R}^2 \to \mathbb{R}$ satisfying the conditions of Theorem 1.21 and such that $f \notin BV_2(\mathbb{R}^2)$.
On the other hand there are functions $f : [0, 1] \to \mathbb{R}$ which are in $W^{1,1}(\mathbb{R}; \mathbb{R})$ (and therefore have the Morse-Sard property) and which do not satisfy assumption (a) of Theorem 1.21 because they
$$\lim_{y \to x} \frac{f(y) - f(x)}{y - x} = \infty$$
for all x in an uncountable set of measure zero. Those examples are well known, but we have not found any appropriate reference, so let us briefly recall a possible construction. Let us take the ternary Cantor set C on $[0, 1]$. For each $i = 1, 2, \ldots$ choose a sequence of closed and disjoint intervals $\{I_{ij}\}_{j \in \mathbb{N}}$ such that $C \subseteq \bigcup_{j=1}^{\infty} I_{ij}$ and $\sum_{j=1}^{\infty} \ell(I_{ij}) \leq (2/3)^i$. Define then
$$f(x) := \int_0^x \sum_{i,j=1}^{\infty} \mathcal{X}_{I_{ij}}(t)\, dt = \sum_{i,j=1}^{\infty} \ell(I_{ij} \cap [0, x]) \ , \quad x \in [0, 1].$$
Since the function $\sum_{i,j=1}^{\infty} \mathcal{X}_{I_{ij}}(t)$ is in $L^1[0, 1]$, it is clear that f is absolutely continuous. On the other hand, it is not difficult to check that $\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = +\infty$ for every $x \in C$.

2. It is also worth noting that Theorem 1.21 extends ([17, Theorem 3.7]): if $f \in C^{n-m}(\mathbb{R}^n; \mathbb{R}^m)$ is such that for all $x \in \mathbb{R}^n$ it has an $(n - m - 1)-$Taylor polynomial at x then f has the Morse-Sard property. It is enough to take a function f in the conditions of [17, Theorem 3.7] and for example change its value in all the points with rational coordinates (call this set N). This new function stops being of class C^{n-m} but it still satisfies the assumptions of Theorem 1.21. Recall that we require condition (a) in Theorem 1.21 to hold everywhere except perhaps on a countable set, and that $\mathcal{H}^s(N) = 0$ for all $s > 0$.

3. A subset γ of $\mathbb{R}^n$ is said to be an *arc* if it is the image of a continuous injection defined on the closed unit interval. For $x, y \in \gamma$, let $\gamma(x, y)$ denote the subarc of γ lying between x and y. An arc γ is a *quasi-arc* if there is some $K > 0$ such that for every $x, y \in \gamma$, $\gamma(x, y)$ is contained in some ball of radius $K|x - y|$. A function f is said to be critical on a set A if $A \subseteq C_f$. Let $A \subseteq \mathbb{R}^n$, $k \geq 1$ an integer number and $t \in (0, 1)$. We say that A is $(k + t)-$*critical* if there exists a real-valued function $f \in C^{k,t}$ which is critical but not constant on A.
To provide a wide range of examples of functions $f : \mathbb{R}^n \to \mathbb{R}$ of class C^{n-1} that have approximate $(n - 1)-$Taylor polynomials everywhere on $\mathbb{R}^n$ except on at most a set N with $\mathcal{H}^{n-1+t}(N) > 0$, but that do not satisfy the Morse-Sard property, we state the following theorem from Norton ([131, Theorem 2]).

Theorem 1.28 (Norton)**.** *Let $k \geq 1$ be an integer number and $t \in (0,1)$. If γ is a quasi-arc with $\mathcal{H}^{k+t}(\gamma) > 0$, then γ is $(k+t)-$critical.*

In the same paper Norton noted that such arcs "are in plentiful supply (e.g. as Julia sets for certain rational maps in the plane)". Hence, building such a quasi-arc γ with $k = n-1$, $t \in (0,1)$ and $\infty > \mathcal{H}^{n-1+t}(\gamma) > 0$, we can get a function $f : \mathbb{R}^n \to \mathbb{R}$ that is of class $C^{n-1,t}$ and does not satisfy the Morse-Sard theorem. Note that we have

$$\operatorname{ap}\limsup_{y\to x} \frac{\left|f(y) - f(x) - \cdots - \frac{D^{n-1}f(x)}{(n-1)!}(y-x)^{n-1}\right|}{|y-x|^n} < +\infty$$

for all $x \in \mathbb{R}^n \setminus \gamma$, since the construction of f comes from an application of the Whitney extension theorem and consequently $f \in C^\infty(\mathbb{R}^n \setminus \gamma; \mathbb{R})$.
This last example shows in particular the existence of C^{n-1} functions $f : \mathbb{R}^n \to \mathbb{R}$ that satisfy the Lusin property of class C^n but do not have the Morse-Sard property. What is nonetheless true is that if $f \in C^{n-1}(\mathbb{R}^n; \mathbb{R})$ has an approximate $(n-1)-$Taylor polynomial $\mathcal{H}^{n-1}$ almost everywhere then it has the Morse-Sard property (apply for instance Theorem 1.21).

Chapter 2

Subdifferentiable functions satisfy Lusin properties of class C^1 or C^2

The classical theorem of Lusin [122] states that for every Lebesgue measurable function $f : \mathbb{R}^n \to \mathbb{R}$ and every $\varepsilon > 0$ there exists a continuous function $g : \mathbb{R}^n \to \mathbb{R}$ such that

$$\mathcal{L}^n \left(\{ x \in \mathbb{R}^n : f(x) \neq g(x) \} \right) \leq \varepsilon. \tag{2.0.1}$$

Here, as in the previous chapter, $\mathcal{L}^n$ denotes the Lebesgue measure in $\mathbb{R}^n$.

Several authors have shown that one can take g of class C^k, provided that f has some regularity properties of order k (for instance, locally bounded distributional derivatives up to the order k, or Taylor expansions of order k almost everywhere). If, given a differentiability class $\mathcal{C}$ and a function $f : \mathbb{R}^n \to \mathbb{R}$ we can find, for each $\varepsilon > 0$, a function $g \in \mathcal{C}$ satisfying (2.0.1), we will say that *f has the Lusin property of class $\mathcal{C}$*.

The first of such results was discovered by Federer [73, p. 442], who showed that almost everywhere differentiable functions (and in particular locally Lipschitz functions) have the Lusin property of class C^1. H. Whitney [153] improved this result by showing that a function $f : \mathbb{R}^n \to \mathbb{R}$ has approximate partial derivatives of first order almost everywhere if and only if f has the Lusin property of class C^1.

In [49, Theorem 13] Calderon and Zygmund established analogous results of order k for the classes of Sobolev functions $W^{k,p}(\mathbb{R}^n)$. Other authors, including Liu [120], Bagby, Michael and Ziemer [27, 124, 155], Bojarski, Hajłasz and Strzelecki [38, 39], and Bourgain, Korobkov and Kristensen [44] have improved Calderon and Zygmund's result in different ways, by obtaining additional estimates for $f - g$ in the Sobolev norms, as well as the Bessel capacities or the Hausdorff contents of the exceptional sets where $f \neq g$. In [44] some Lusin properties of the class $BV_k(\mathbb{R}^n)$ (of integrable functions whose distributional derivatives of order up to k are Radon measures) are also established, as well as in [114], where the Sobolev-Lorentz spaces $W_1^{k,p}(\mathbb{R}^n)$ are considered and the excepcional sets have small Hausdorff content. The Whitney extension technique [151], and some related techniques as the *Whitney smoothing* introduced in [39], play a key role in the proofs of all of these results.

For the special class of convex functions $f : \mathbb{R}^n \to \mathbb{R}$, Alberti and Imomkulov [2, 101] showed that every convex function has the Lusin property of class C^2 (with g not necessarily convex in (2.0.1)); see also [1] for a related problem. More recently Azagra and Hajłasz [23] have proved that g can be taken to be $C^{1,1}_{loc}$ and convex in (2.0.1) if and only if either f is essentially coercive (meaning that f is coercive up to a linear perturbation) or else f is already $C^{1,1}_{loc}$ (in which case taking $g = f$ is the only possible option).

On the other hand, generalizing Whitney's result [153] to higher orders of differentiability, Isakov [102] and Liu and Tai [121] independently established that a function $f : \mathbb{R}^n \to \mathbb{R}$ has the Lusin property of class C^k if and only if f is approximately differentiable of order k almost everywhere (and if and only if f has an approximate $(k-1)$-Taylor polynomial at almost every point). See the statement of this result in Theorem 1.9 of the previous Chapter 1.

In this chapter we will answer the following question (which we think may be quite natural for people working on nonsmooth analysis or viscosity solutions to PDE such as Hamilton-Jacobi equations): do functions with nonempty subdifferentials almost everywhere have Lusin properties of order C^1 or C^2? By subdifferentials we mean the Fréchet subdifferential, or the proximal subdifferential, or the second order viscosity subdifferential. As we will see the answer is positive: Fréchet subdifferentiable functions have the Lusin property of class C^1, and functions with nonempty proximal subdifferentials almost everywhere (in particular functions with almost everywhere nonempty viscosity subdifferentials of order 2) have the Lusin property of class C^2.

This question can be formulated in a more general form (perhaps appealing to a wider audience) as a problem about *Taylor subexpansions*: given $k \in \mathbb{N}$ and a function $f : \mathbb{R}^n \to \mathbb{R}$, assume that for almost every $x \in \mathbb{R}^n$ there exists a polynomial $p_{k-1}(x; y)$ of degree less than or equal to $k-1$ such that

$$\liminf_{y \to x} \frac{f(y) - p_{k-1}(x; y)}{|y - x|^k} > -\infty.$$

Is it then true that f has the Lusin property of order k?

The results of this chapter will show that the answer to this question is positive for $k = 1, 2$, but negative for $k \geq 3$. For the case $k = 2$ our main theorem proves that if $f : \mathbb{R}^n \to \mathbb{R}$ is a function, Ω a measurable set and for almost every $x \in \Omega$ there exists a vector $\xi_x \in \mathbb{R}^n$ such that

$$\liminf_{h \to 0} \frac{f(x+h) - f(x) - \langle \xi_x, h \rangle}{|h|^2} > -\infty,$$

then f satisfies a Lusin-type property of class C^2 in Ω. In particular every function which has a nonempty proximal subdifferential almost everywhere also has the Lusin property of class C^2.

In Section 2.1 we present the definitions of subdifferentiable functions. In Sections 2.2 and 2.3 we prove that subdifferentiable functions satisfy Lusin properties of class C^1 and C^2, for the Fréchet subdifferentiable and the proximal subdifferential respectively. Section 2.4 is devoted to make some remarks and exhibit a counterexample showing that these kind of results are no longer true for *Taylor subexpansions* of higher order. And finally Section 2.5 is intended to be a link between subdifferentiable functions and the Morse-Sard theorem, where we collect results from [17].

2.1 Subdifferentiability

Let us make a brief introduction to subdifferentiable functions.

By studying nonsmooth analysis we attempt to extend differentiability and, more specifically, calculus to a broader setting. This theory experienced a great growth in the seventies. One reason for its development was the recognition that non-differentiability phenomena are more widespread, and play an important role, than had been thought.

The concept of subdifferential on convex functions was first introduced by Jean Jacques Moreau and R. Tyrrell Rockafellar in the early 1960's. The subdifferential of a non-convex function has been introduced by F. Clarke in his PhD thesis (1973) (under the supervision of T. Rockafellar). Rockafellar published in 1980 a paper on generalized directional derivatives (now called the Clarke-Rockafellar directional derivative) for a lower semicontinuous function that recoves analytically the Clarke subdifferential. (For Lipschitz functions, Clarke, in his thesis, had already given an analytic formula of a generalized directional derivative).

We do not intend to restrict ourselves to convex functions so we assume $f : \mathbb{R}^n \to \mathbb{R}$ is any function and present the following different concepts of subdifferentials:

1. The *Fréchet subdifferential* of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$, denoted by $\partial^- f(x)$, is defined as the set of vectors $\xi \in \mathbb{R}^n$ such that

$$\liminf_{h \to 0} \frac{f(x+h) - f(x) - \langle \xi, h \rangle}{|h|} \geq 0. \tag{2.1.1}$$

This is equivalent to the existence of a C^1 function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $\varphi(x) = f(x)$, $\nabla\varphi(x) = \xi$ and $\varphi \le f$ in a neighbourhood of x (or even in the whole $\mathbb{R}^n$).

In addition one can consider the Fréchet superdifferential defined as

$$\partial^+ f(x) = \{\xi \in \mathbb{R}^n : \limsup_{h\to 0} \frac{f(x+h) - f(x) - \langle \xi, h\rangle}{|h|} \le 0\}.$$

We have $\partial^-(-f)(x) = -\partial^+ f(x)$. Furthermore if we assume that f is differentiable at $x \in \mathbb{R}^n$ then $\partial^-(f)(x) = \partial^+ f(x) = \{\nabla f(x)\}$. Conversely, if $\partial^- f(x) \ne \emptyset \ne \partial^+ f(x)$ then f is differentiable.

2. The *proximal subdifferential* of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$, written $\partial_P f(x)$, is defined as the set of vectors $\xi \in \mathbb{R}^n$ such that

$$\liminf_{h\to 0} \frac{f(x+h) - f(x) - \langle \xi, h\rangle}{|h|^2} > -\infty. \tag{2.1.2}$$

This condition is equivalent to the existence of $\sigma, \eta > 0$ such that

$$f(y) \ge f(x) + \langle \xi, y - x\rangle - \sigma|y-x|^2 \text{ for all } y \in B(x,\eta),$$

and also is equivalent to the existence of a C^2 function $\varphi : \mathbb{R}^n \to \mathbb{R}$ such that $\varphi(x) = f(x)$, $\nabla\varphi(x) = \xi$ and $\varphi \le f$ in a neighbourhood of x.

3. The *viscosity subdifferential* of second order of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ is defined to be the set

$$\partial_V f(x) = \{(\nabla\varphi(x), D^2\varphi(x)) \in \mathbb{R}^n \times \mathbb{R}^{2n} : \varphi \in C^2(\mathbb{R}^n),\ f - \varphi \text{ attains a local minimum at } x\}.$$

Note that if the viscosity subdifferential of second order of f at x is not empty then so it is the proximal subdifferential.

4. The *limiting subdifferential* of $f : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$, written $\partial_L f(x)$, is defined as the set of vectors $\xi \in \mathbb{R}^n$ such that there exist sequences $\{x_n\}$ converging to x and $\{\xi_n\}$ converging to ξ so that $\xi_n \in \partial^- f(x_n)$ and $\lim_{n\to\infty} f(x_n) = f(x)$.

Observe that $\partial_P f(x) \subseteq \partial^- f(x) \subseteq \partial_L f(x)$. Moreover any of the conditions (2.1.1) or (2.1.2) implies that f is a lower semicontinuous function and in particular Lebesgue measurable.

We remark finally that all these previous definitions, and the consequent theory that derives from them, can be made in the more general setting of Hilbert spaces or Banach spaces with smooth norms. The interested reader can have a look at [53, 56, 78] and the references therein for more information about subdifferentials and their applications.

2.2 Lusin property of class C^1 for Fréchet subdifferentiable functions

In the case $k = 1$ the proof is very simple and natural.

Theorem 2.1. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, and $f : \Omega \to \mathbb{R}$ a function. Assume that for almost every $x \in \Omega$ we have*

$$\liminf_{y\to x} \frac{f(y) - f(x)}{|y-x|} > -\infty. \tag{2.2.1}$$

Then, for every $\varepsilon > 0$ there exists a function $g \in C^1(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \ne g(x)\}\right) \le \varepsilon.$$

In order to facilitate the proof of Theorem 2.1, as well as that of Theorem 2.6, let us state the following technical lemma, which is standard. We include its proof for the readers' convenience.

Lemma 2.2. *Let Ω be a Lebesgue measurable subset of $\mathbb{R}^n$, $k \in \mathbb{N}$, and $f : \Omega \to \mathbb{R}$ be measurable. Then f has the Lusin property of class C^k (meaning that for every $\varepsilon > 0$ there exists $g \in C^k(\mathbb{R}^n)$ such that $\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon$) if and only if the restriction of f to each compact subset of Ω has the Lusin property of class C^k.*

Proof. It is obvious that if $f : \Omega \to \mathbb{R}$ has the Lusin property of class C^k then, for every compact subset K of Ω, the function $f_{|_K} : K \to \mathbb{R}$ has the Lusin property of class C^k. Let us prove the converse. Assume first that Ω is bounded. By the regularity of the measure $\mathcal{L}^n$, for every $\varepsilon > 0$ we may find K_ε, a compact subset of Ω, such that $\mathcal{L}^n\left(\Omega \setminus K_\varepsilon\right) \leq \varepsilon/2$. By assumption, there exists a function $g = g_{K_\varepsilon} \in C^k(\mathbb{R}^n)$ such that $\mathcal{L}^n\left(\{x \in K_\varepsilon : f(x) \neq g(x)\}\right) \leq \varepsilon/2$. Then we have

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \mathcal{L}^n\left(\Omega \setminus K_\varepsilon\right) + \mathcal{L}^n\left(\{x \in K_\varepsilon : f(x) \neq g(x)\}\right) \leq \varepsilon,$$

and therefore $f : \Omega \to \mathbb{R}$ has the Lusin property of class C^k.
Now let us consider the general case that Ω is not necessarily bounded. We can write

$$\Omega = \bigcup_{j=1}^{\infty} \Omega_j, \text{ where } \Omega_1 = \Omega \cap B(0,1), \text{ and } \Omega_{j+1} := \Omega \cap B(0, j+1) \setminus \overline{B(0,j)}.$$

According to the previous argument, for each $j \in \mathbb{N}$ there exists a function $g_j \in C^k(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x \in \Omega_j : g_j(x) \neq f(x)\}\right) \leq \frac{\varepsilon}{6^j}.$$

Let $(\psi_j)_{j=1}^{\infty}$ be a C^∞ smooth partition of unity subordinated to the covering $\{B(0,j+1) \setminus \overline{B(0,j-1)}\}_{j=1}^{\infty} \cup \{B(0,1)\}$ of $\mathbb{R}^n$ (see for instance [100, Chapter 2, Theorem 2.1]), and let us define

$$g(x) = \sum_{j=1}^{\infty} \psi_j(x) g_j(x).$$

Notice that

$$\{x \in \Omega_j : f(x) \neq g(x)\} \subseteq \bigcup_{i=j-1}^{j} \{x \in \Omega_j : f(x) \neq g_i(x)\}.$$

This implies that

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq 2 \sum_{j=1}^{\infty} \mathcal{L}^n\left(\{x \in \Omega_j : f(x) \neq g_j(x)\}\right) \leq 2 \sum_{j=1}^{\infty} \frac{\varepsilon}{6^j} \leq \varepsilon,$$

and concludes the proof of the Lemma. □

Proof of Theorem 2.1. Let us call $N \subset \Omega$ the set of points for which (2.2.1) does not hold. Since N has measure zero, proving the Lusin property of class C^1 for the restriction of f to $\Omega \setminus N$ would immediately lead to the Lusin property of class C^1 for f. So we may and do assume in what follows that $N = \emptyset$, and in particular that

$$\liminf_{y \to x} \frac{f(y) - f(x)}{|y - x|} > -\infty$$

for every $x \in \Omega$. Note that this inequality implies that f is lower semicontinuous on Ω, and in particular f is measurable. Now, according to Lemma 2.2, it is enough to check that the restriction of f to every compact subset of Ω has the Lusin property of class C^1, and therefore we may also assume without loss of generality that Ω is compact. Define for each $j \in \mathbb{N}$,

$$E_j := \left\{ x \in \Omega : \ f(y) - f(x) \geq -j|y - x| \text{ for all } y \in B\left(x, \frac{1}{j}\right) \cap \Omega \right\} \cap \{x \in \Omega : \ |f(x)| \leq j\}.$$

Because f is lower semicontinuous the sets

$$\left\{x \in \Omega : \; f(y) - f(x) \geq -j|y-x| \text{ for all } y \in B\left(x, \frac{1}{j}\right) \cap \Omega\right\}$$

are closed, and by using the measurability of f this implies that each set E_j is measurable. These sets form an increasing sequence such that

$$\Omega = \bigcup_{j=1}^{\infty} E_j,$$

so we have

$$\lim_{j\to\infty} \mathcal{L}^n\left(\Omega \setminus E_j\right) = 0,$$

and therefore, for a given $\varepsilon > 0$ we may find $j_0 \in \mathbb{N}$ large enough such that $\mathcal{L}^n(\Omega \setminus E_{j_0}) < \frac{\varepsilon}{2}$.
Take now $x, y \in E_{j_0}$. If $|y-x| \leq \frac{1}{j_0}$ then we have

$$|f(y) - f(x)| \leq j_0|y-x| \text{ and } |f(x)| \leq j_0.$$

On the other hand, if $x, y \in E_{j_0}$ and $|y-x| > 1/j_0$ then we trivially get

$$|f(y) - f(x)| \leq 2 \sup_{z\in E_{j_0}} |f(z)| \leq M_0|y-x|,$$

where $M_0 := 2j_0\left(1 + \sup_{z\in\Omega}|f(z)|\right)$.
Observe that $M_0 \geq j_0$. Thus in either case we see that

$$|f(y) - f(x)| \leq M_0|y-x| \text{ and } |f(x)| \leq M_0, \text{ for all } x, y \in E_{j_0}.$$

That is, f is bounded and M_0-Lipschitz on E_{j_0}. Then we can extend f to a Lipschitz function F on $\mathbb{R}^n$, for instance by using the McShane-Whitney formula

$$F(x) = \inf_{y\in E_{j_0}} \{f(y) + M_0|x-y|\},$$

which defines an M_0-Lipschitz function on $\mathbb{R}^n$ that coincides with f on E_{j_0}. Obviously we have

$$\mathcal{L}^n(\{x \in \Omega : \; f(x) \neq F(x)\}) \leq \mathcal{L}^n(\Omega \setminus E_{j_0}) < \frac{\varepsilon}{2}.$$

But according to the result of Federer's that we mentioned above (see also [71, Theorem 6.11]), Lipschitz functions have the C^1 Lusin property, so we may find another function $g \in C^1(\mathbb{R}^n)$ such that $\mathcal{L}^n(\{x \in \Omega : \; F(x) \neq g(x)\}) < \frac{\varepsilon}{2}$. Thus we conclude that

$$\begin{aligned}\mathcal{L}^n(\{x \in \Omega : \; f(x) \neq g(x)\}) &= \mathcal{L}^n(\{x \in E_{j_0} : \; F(x) \neq g(x)\} \cup \{x \in \Omega \setminus E_{j_0} : \; f(x) \neq g(x)\}) \\ &\leq \mathcal{L}^n(\{x \in E_{j_0} : \; F(x) \neq g(x)\}) + \mathcal{L}^n(\Omega \setminus E_{j_0}) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

□

Corollary 2.3. *Let U be a measurable subset of $\mathbb{R}^n$, $f : U \to \mathbb{R}$ be a measurable function, and define $\Omega = \{x \in U : \partial^- f(x) \neq \emptyset\}$ to be the set of points where the Fréchet subdifferential is nonempty. Then for every $\varepsilon > 0$ there exists a function $g \in C^1(\mathbb{R}^n)$ such that*

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

Remark 2.4. *In the above corollary we also have $\partial^- f(x) = \{\nabla g(x)\}$ for almost every $x \in \Omega$ with $f(x) = g(x)$.*

Proof. Almost every point of the set $A = \{x \in \Omega : f(x) = g(x)\}$ is a point of density 1 of A, and for every such point x and every $\xi_x \in \partial^- f(x)$ we have

$$0 \le \liminf_{y\to x, y\in A} \frac{f(y) - f(x) - \langle \xi_x, y-x\rangle}{|y-x|} = \liminf_{y\to x, y\in A} \frac{g(y) - g(x) - \langle \xi_x, y-x\rangle}{|y-x|},$$

and

$$\lim_{y\to x, y\in A} \frac{g(y) - g(x) - \langle \nabla g(x), y-x\rangle}{|y-x|} = 0,$$

hence also

$$\liminf_{y\to x, y\in A} \frac{\langle \nabla g(x) - \xi_x, y-x\rangle}{|y-x|} \ge 0, \tag{2.2.2}$$

which, because x is a point of density 1 of A and $h \mapsto \langle \nabla g(x) - \xi_x, h\rangle$ is linear, implies that $\nabla g(x) = \xi_x$. Indeed, we have

$$\lim_{r\to 0^+} \frac{\mathcal{L}^n (A \cap B(x,r))}{\mathcal{L}^n (B(x,r))} = 1. \tag{2.2.3}$$

Assume we had $\zeta := \nabla g(x) - \xi_x \neq 0$, and consider the sets

$$S_\zeta := \{v \in \mathbb{R}^n : |v| = 1, \langle \zeta, v\rangle \le -\frac{1}{2}|\zeta|\},$$

which determine a region of positive surface measure in the unit sphere, and the associated cone

$$C_{x,\zeta} = \{x + tv : v \in S_\zeta, t > 0\},$$

of which x is thus a point of positive density. Hence $C_{x,\zeta}$ also satisfies, in view of (2.2.3), that

$$\liminf_{r\to 0^+} \frac{\mathcal{L}^n (A \cap C_{x,\zeta} \cap B(x,r))}{\mathcal{L}^n (B(x,r))} > 0.$$

In particular there exists a sequence $(y_k) = (x + t_k v_k) \subset A \cap C_{x,\zeta}$ (with $t_k > 0$ and $v_k \in S_\zeta$, $k \in \mathbb{N}$) such that $\lim_{k\to\infty} y_k = x$. For this sequence we have, because of the definition of $C_{x,\zeta}$, that

$$\frac{\langle \nabla g(x) - \xi_x, y_k - x\rangle}{|y_k - x|} = \frac{\langle \zeta, t_k v_k\rangle}{t_k} \le -\frac{1}{2}|\zeta| < 0$$

for all $k \in \mathbb{N}$, which contradicts (2.2.2). □

A natural question at this point is the following. Does Corollary 2.3 hold true if we replace the Frechet subdifferential by the *limiting subdifferential*? Let us recall that the limiting subdifferential $\partial_L f(x)$ of a (lower semicontinuous) function $f : \mathbb{R}^n \to \mathbb{R}$ at a point x consists of all vectors of the form $\zeta = \lim_n \zeta_n$, where $\zeta_n \in \partial^- f(x_n)$, for sequences $\{x_n\}$ satisfying $\lim_n x_n = x$, and $\lim_n f(x_n) = f(x)$; see [53, 78], for instance, for elementary properties of this subdifferential. The question is whether or not the assumption that $\partial_L f(x) \neq \emptyset$ for every $x \in \mathbb{R}^n$ implies that f satisfies the Lusin property of order C^1. Since one trivially has that $\partial^- f(x) \subset \partial_L f(x)$, such a result would be much stronger than Corollary 2.3. The following example shows that the answer is negative.

Example 2.5. We consider the classical Takagi function $T : \mathbb{R} \to \mathbb{R}$ defined as follows. If D_n denotes the set of real numbers $\{\frac{k}{2^n} : k \in \mathbb{Z}\}$, and $d(x, D_n)$ is the distance of x to D_n, then

$$T(x) = \sum_{n=1}^{\infty} d(x, D_n)$$

This function was introduced by Takagi, [145], as an easy example of a continuous function which is nowhere differentiable. In [45, Theorem 2] it is proved that T does not agree with any C^1 function on any set of positive measure, and in particular T does not satisfy the Lusin property of order C^1. However, in [80, Corollary 1.4], and also implicitly in [89], it is proved that $\partial_L T(x) = \mathbb{R}$ for every $x \in \mathbb{R}$.

2.3 Lusin property of class C^2 for proximal subdifferentiable functions

Concerning the Lusin property of class C^2 we have the following result.

Theorem 2.6. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, and $f : \Omega \to \mathbb{R}$ be a function such that for almost every $x \in \Omega$ there exists a vector $\xi_x \in \mathbb{R}^n$ such that*

$$\liminf_{y \to x} \frac{f(y) - f(x) - \langle \xi_x, y - x \rangle}{|y - x|^2} > -\infty. \tag{2.3.1}$$

Then for every $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that

$$\mathcal{L}^n \left(\{ x \in \Omega : f(x) \neq g(x) \} \right) \leq \varepsilon.$$

Proof. Let N be the subset of points for which (2.3.1) does not hold, and put $\Omega_1 = \Omega \setminus N$. Since N has measure zero, it will be enough to show that the restriction f_1 of f to Ω_1 has the Lusin property of class C^2. Since (2.3.1) holds for every $x \in \Omega_1$, it follows that f is lower semicontinuous on Ω_1, and in particular f_1 is measurable (hence so is f, since N has measure zero). Now, according to Lemma 2.2, if we take an arbitrary compact subset Ω_2 of Ω_1, it will be enough for us to check that the restriction f_2 of f_1 to Ω_2 has the Lusin property of class C^2.
Because (2.3.1) holds for every $x \in \Omega_2$ and this implies

$$\liminf_{y \to x} \frac{f_2(y) - f_2(x)}{|y - x|} > -\infty$$

for all $x \in \Omega_2$, given $\varepsilon > 0$, we may apply Theorem 2.1 to get a function $g \in C^1(\mathbb{R}^n)$ such that

$$\mathcal{L}^n(\{x \in \Omega_2 \, : \, f_2(x) \neq g(x)\}) \leq \frac{\varepsilon}{4}.$$

Observe also that the set $A = \{x \in \Omega_2 : f_2(x) = g(x)\}$ is measurable and bounded, and according to the preceding remark we have $\xi_x = \nabla g(x)$ for almost every $x \in A$, so we can find a compact subset Ω_3 of A such that $\mathcal{L}^n \left(A \setminus \Omega_3 \right) \leq \varepsilon/4$ and $\xi_x = \nabla g(x)$ for all $x \in \Omega_3$. Then we have that

$$\liminf_{y \to x, y \in \Omega_3} \frac{g(y) - g(x) - \langle \nabla g(x), y - x \rangle}{|y - x|^2} > -\infty \tag{2.3.2}$$

for every $x \in \Omega_3$. Now let us define for each $j \in \mathbb{N}$

$$E_j := \left\{ x \in \Omega_3 \, : \, g(y) - \langle \nabla g(x), y \rangle \geq g(x) - \langle \nabla g(x), x \rangle - j|y - x|^2 \text{ for all } y \in \Omega_3 \right\},$$

and note that the sets E_j are measurable and increasing to Ω_3. There exists $j_0 \in \mathbb{N}$ such that

$$\mathcal{L}^n(\Omega_3 \setminus E_{j_0}) \leq \frac{\varepsilon}{4}.$$

It will be enough for us to prove the following:

Claim 2.7. *We have that*

$$\limsup_{y \to x, \, y \in E_{j_0}} \frac{|g(y) - g(x) - \langle \nabla g(x), y - x \rangle|}{|y - x|^2} < +\infty$$

for almost every $x \in E_{j_0}$.

Assume for a moment that the Claim is true, that is, the restriction of g to E_{j_0} has an approximate $(2-1)$-Taylor polynomial at every $x \in E_{j_0}$. By [121, Theorem 1] this is equivalent to saying that the restriction of g to E_{j_0} has the Lusin property of class C^2. So we may find a function $h \in C^2(\mathbb{R}^n; \mathbb{R})$ such that

$$\mathcal{L}^n(\{x \in E_{j_0} \, ; \, g(x) \neq h(x)\}) \leq \frac{\varepsilon}{4},$$

and we easily conclude that

$$\mathcal{L}^n(\{x \in \Omega_2 :\ f_2(x) \neq h(x)\}) \leq \varepsilon,$$

as we wanted to show.
In order to prove Claim (2.7) we will borrow some ideas from [111]. We define new functions $\widetilde{g} : \mathbb{R}^n \to \mathbb{R}$ and $\hat{g} : \mathbb{R}^n \to \mathbb{R}$ by

$$\begin{aligned} &\widetilde{g}(x) = g(x) + j_0|x|^2, && x \in \mathbb{R}^n \\ &\hat{g}(x) = \sup\{p(x) :\ p \text{ affine and } p \leq \widetilde{g} \text{ on } \Omega_3\ \}, && x \in \mathbb{R}^n \end{aligned}$$

An affine function $p : \mathbb{R}^n \to \mathbb{R}$ is a function of the form $p(x_1, \dots, x_n) = a_1x_1 + \cdots + a_nx_n + a$ where $a, a_1, \dots, a_n \in \mathbb{R}$.
By definition of E_{j_0} we have $\widetilde{g}(y) \geq \widetilde{g}(x) + \langle \nabla\widetilde{g}(x), y - x\rangle$ for all $y \in \Omega_3, x \in E_{j_0}$, and by using this inequality it is easy to see that

$$\widetilde{g}(x) = \hat{g}(x)$$

for all $x \in E_{j_0}$. On the other hand, since Ω_3 is compact and g is continuous on Ω_3, it is easy to see that $\hat{g}$ is everywhere finite. Moreover, as a supremum of affine functions, $\hat{g}$ is convex. Therefore $\hat{g}$ is locally Lipschitz on Ω_3. Also g is of class C^1, hence so is $\widetilde{g}$. Since the functions $\widetilde{g}$ and $\hat{g}$ agree on E_{j_0}, we then also have that

$$\nabla\hat{g}(x) = \nabla\widetilde{g}(x)$$

for almost every $x \in E_{j_0}$ (see [71, Theorem 3.3(i)] for instance).
Next, by applying Alexandroff's theorem [5] (see also [48] in dimension 2) with the convex function $\hat{g}$, we obtain that $\hat{g}$ is twice differentiable almost everywhere in Ω_3. This implies that

$$\begin{aligned} &\limsup_{y\to x,\ y\in E_{j_0}} \frac{|\widetilde{g}(y) - \widetilde{g}(x) - \langle\nabla\widetilde{g}(x), y - x\rangle|}{|y - x|^2} \\ &\qquad = \limsup_{y\to x,\ y\in E_{j_0}} \frac{|\hat{g}(y) - \hat{g}(x) - \langle\nabla\hat{g}(x), y - x\rangle|}{|y - x|^2} < +\infty \end{aligned} \tag{2.3.3}$$

for almost every $x \in E_{j_0}$. However, by the definition of $\widetilde{g}(x) = g(x) + j_0|x|^2$, we have

$$\begin{aligned} &\frac{|g(y) - g(x) - \langle\nabla g(x), y - x\rangle|}{|y - x|^2} \\ &\qquad \leq \frac{|g(y) - g(x) - \langle\nabla g(x), y - x\rangle + (j_0(|y|^2 + |x|^2 - 2\langle x, y\rangle))|}{|y - x|^2} + j_0 \\ &\qquad\qquad = \frac{|\widetilde{g}(y) - \widetilde{g}(x) - \langle\nabla\widetilde{g}(x), y - x\rangle|}{|y - x|^2} + j_0, \end{aligned}$$

and by combining with (2.3.3) we immediately obtain Claim (2.7). □

Corollary 2.8. *Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, and $f : \Omega \to \mathbb{R}$ be a function such that for almost every $x \in \Omega$ there exists a vector $\xi_x \in \mathbb{R}^n$ such that*

$$\limsup_{y\to x} \frac{f(y) - f(x) - \langle\xi_x, y - x\rangle}{|y - x|^2} < +\infty. \tag{2.3.4}$$

Then for every $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

This is of course an immediate consequence of Theorem 2.6 applied to $-f$.

According to Remark 2.4, we also have that

$$\xi_x = \nabla g(x)$$

for almost every $x \in \Omega$ with $f(x) = g(x)$.

Corollary 2.9. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a measurable function, and define $\Omega = \{x \in \mathbb{R}^n : \partial_P f(x) \neq \emptyset\}$ to be the set of points where the proximal subdifferential is nonempty. Then for every $\varepsilon > 0$ there exists a function $g \in C^2(\mathbb{R}^n)$ such that*

$$\mathcal{L}^n\left(\{x \in \Omega : f(x) \neq g(x)\}\right) \leq \varepsilon.$$

It is clear that the above corollary is an immediate consequence of Theorem 2.6. Notice also that this corollary allows us to recover, with a different proof, the mentioned result for convex functions established independently by Alberti [2] and Imomkulov [101].

2.4 Some counterexamples

Let us finally present two examples.

2.4.1 Almost everywhere subdifferentiable functions need not be almost everywhere superdifferentiable

The first one concerns the following matter: one could erroneously think that if a function f satisfies (2.3.1) then f will automatically satisfy

$$\limsup_{y \to x} \frac{f(y) - f(x) - \langle \xi_x, y - x \rangle}{|y - x|^2} < +\infty \tag{2.4.1}$$

for almost every $x \in \Omega$ as well, and then one could immediately apply Liu and Tai's theorem [121] to conclude the proof of Theorem 2.6. This is not feasible.

Example 2.10. Let us first consider a Cantor set of positive measure, $C \subset [0, 1]$. More precisely,

$$C = [0, 1] \setminus \bigcup_n J_n$$

where each J_n is the union of 2^{n-1} disjoint intervals of length $\frac{1}{4^n}$ and $J_n \cap J_m = \emptyset$ for $n \neq m$. We have

$$J_n = \bigcup_{k=1}^{2^{n-1}} (a_n^k, b_n^k),$$

where $b_n^k < a_n^{k+1}$ for $k < 2^{n-1}$. Let us inductively construct the sets J_n. Setting $J_1 = (\frac{3}{8}, \frac{5}{8})$, if $n \geq 1$, we assume that $J_1, \ldots, J_n$ satisfy that

$$[0, 1] \setminus \bigcup_{k=1}^{n} J_k$$

consists in 2^n disjoint intervals of length $\frac{1}{2^{n+1}} + \frac{1}{2^{2n+1}}$, because

$$\mathcal{L}\big([0, 1] \setminus \bigcup_{k=1}^{n} J_k\big) = 1 - \sum_{k=1}^{n} \frac{2^{k-1}}{4^k} = 1 - \frac{1}{2}(1 - \frac{1}{2^n}) = \frac{1}{2} + \frac{1}{2^{n+1}}.$$

For each of these intervals composing $[0, 1] \setminus \bigcup_{k=1}^{n} J_k$, we consider a subinterval, centered at the corresponding middle point, of length $\frac{1}{4^{n+1}}$. Then J_{n+1} will be the union of these subintervals. It is clear that $\mathcal{L}(C) = \frac{1}{2}$.

Now let us define a function f in the following way: we set

$$f(x) = 0 \text{ for every } x \in C,$$

while for every $n \in \mathbb{N}$ and $k = 1, \dots, 2^{n-1}$, $f : [a_n^k, b_n^k] \to \mathbb{R}$ will be a non negative continuous function such that $f : (a_n^k, b_n^k) \to \mathbb{R}$ is C^∞,

$$\max_{x \in I_n^k} f(x) = f(a_n^k + \frac{1}{2}(b_n^k - a_n^k)) = \frac{1}{2^n},$$

and such that f, as well as all its one-sided derivatives, equals 0 at a_n^k and at b_n^k. It is clear that f is continuous. Let us denote

$$\Delta_x(y) = \frac{f(y) - f(x) - \xi_x(y-x)}{|y-x|^2}.$$

If $x \notin C$ then, taking $\xi_x = f'(x)$, we have $\lim_{y \to x} \Delta_x(y) = \frac{1}{2} f''(x)$. If $x \in C$, then

$$\frac{f(y) - f(x)}{|y-x|^2} \geq 0.$$

Hence for every x there exists ξ_x such that

$$\liminf_{y \to x} \Delta_x(y) > -\infty.$$

Let us observe that f also satisfies conditions of the form

$$\liminf_{y \to x} \frac{f(y) - p_{k-1}(x; y)}{|y-x|^k} > -\infty,$$

where $p_{k-1}(x; \cdot)$ is a polynomial of degree $k-1$ for every k, centered at x, with $p_{k-1}(x; x) = f(x)$.

Now let $\tilde{C} = C \setminus (\{0, 1\} \cup \{a_n^k, b_n^k\}_{n,k})$. We claim that

$$\limsup_{y \to x} \Delta_x(y) = +\infty$$

for every $x \in \tilde{C}$ and every ξ_x. Let us prove this. If $x \in \tilde{C}$ there exist subsequences $\{a_{m_j}^{r_j}\}_j$ and $\{b_{n_j}^{k_j}\}_j$, decreasing and increasing respectively, such that

$$\lim_j a_{m_j}^{r_j} = \lim_j b_{n_j}^{k_j} = x.$$

More precisely, we chose $a_{m_j}^{r_j}$ such that

$$0 < a_{m_j}^{r_j} - x \leq \frac{1}{2^{m_j+1}} + \frac{1}{2^{2m_j+1}},$$

and $b_{n_j}^{k_j}$ such that

$$0 < x - b_{n_j}^{k_j} \leq \frac{1}{2^{n_j+1}} + \frac{1}{2^{2n_j+1}}.$$

Let us consider the case that $\xi_x \geq 0$. We take $y_j = b_{n_j}^{k_j} - \frac{1}{2}(b_{n_j}^{k_j} - a_{n_j}^{k_j})$. We have

$$\Delta_x(y_j) \geq \frac{f(y_j)}{|y_j - x|^2} = \frac{1}{2^{n_j}} \frac{1}{|y_j - x|^2} \geq 2^{n_j}$$

since $|y_j - x| \leq \frac{1}{2^{n_j}}$. In particular we obtain that $\limsup_{y \to x} \Delta_x(y) = +\infty$.

The case $\xi_x \leq 0$ can be dealt with similarly by considering $y_j = a_{m_j}^{r_j} + \frac{1}{2}(b_{m_j}^{r_j} - a_{m_j}^{r_j})$. □

What is true, by Liu and Tai's result, is that for almost every point $x \in [0, 1]$, there exists ξ_x and a set A_x with density 1 at x such that

$$\limsup_{\substack{y \to x \\ y \in A_x}} \frac{f(y) - f(x) - \xi_x(y-x)}{|y-x|^2} < +\infty.$$

Furthermore the property that $\liminf_{y \to x} \frac{f(y)-f(x)}{|y-x|} > -\infty$ almost everywhere is neither enough to ensure that $\limsup_{y \to x} \frac{f(y)-f(x)}{|y-x|} < +\infty$ almost everywhere. Simply take the previous Example 2.10 but now let the maximum of f in each interval I_n^k, which is supposed to be attained at the middle point, to be equal to $2^{n/2}$. Another way of saying this is that functions with nonempty Fréchet subdifferential almost everywhere need not be differentiable almost everywhere.

2.4.2 Failure of Lusin property for classes C^k, $k \geq 3$

Our second example shows that there are no analogues of Theorem 2.6 for higher order of differentiability.

Example 2.11. Let $f : \mathbb{R} \to \mathbb{R}$ be the function given by

$$f(x) = \frac{1}{\pi^2} \sum_{n=1}^{\infty} 2^{-3n} \cos\left(2^n \pi x\right).$$

This is a C^2 function such that f'' is not differentiable at any point (see [97]) and

$$\limsup_{|y| \to 0} \frac{|f''(x+y) + f''(x-y) - 2f''(x)|}{|y|} < +\infty$$

for every $x \in \mathbb{R}$ (see [144, p. 148]). By [121, Theorem 4] f'' is not approximately differentiable on a set of positive measure.

For every x, we have that

$$\lim_{y \to x} \frac{f(y) - f(x) - f'(x)(y-x) - \frac{1}{2} f''(x)(y-x)^2}{|y-x|^2} = 0.$$

If $a > 0$ we have

$$\liminf_{y \to x} \frac{f(y) - f(x) - f'(x)(y-x) - (\frac{1}{2} f''(x) - a)(y-x)^2}{|y-x|^2} > 0,$$

hence

$$\liminf_{y \to x} \frac{f(y) - f(x) - f'(x)(y-x) - (\frac{1}{2} f''(x) - a)(y-x)^2}{|y-x|^k} = \infty > -\infty$$

for every $k > 2$. If an analogue of Theorem 2.6 for some order $k > 2$ were true for this function f, then, according to Liu and Tai's characterization of Lusin properties and approximate differentiability of higher order [121], we would have that f is approximately differentiable of order k. However, in [121, p. 194] it is shown that the coefficients of order j of the Taylor expansion of an approximately differentiable function of order k coincide, up to sets of arbitrarily small measure, with derivatives of order j of C^k functions; in particular those coefficients have the Lusin property of class C^{k-j} and therefore, again by [121, Theorem 1], they are almost everywhere approximately differentiable of order $k - j$. This would imply that f'' is approximately differentiable almost everywhere, which we know to be false.

Another example can be given by taking $g : \mathbb{R} \to \mathbb{R}$ to be a continuous function which is nowhere approximately differentiable (see [109, Chapter 6]), setting

$$f(x) = \int_0^x \left(\int_0^t g(s) ds \right) dt,$$

and repeating the preceding argument word by word. One could also use as g the Takagi function of Example 2.5, which by [45, Theorem 2] and [121] is not approximately differentiable on any set of positive measure. In fact is has been proven very recently by Ferrera and Gómez-Gil that the Takagi function is nowhere approximately differentiable (see [79]).

2.5 Subdifferentiable functions and the Morse-Sard theorem

We include this section to comment the relation between subdifferentiable functions and the Morse-Sard property, just for completeness, in accordance with all the discussion in the first chapter. The main theorems that we state below about this issue can be found in the work of Azagra, Ferrera and Gómez-Gil [17].

For simplicity we will consider functions $f : \mathbb{R}^n \to \mathbb{R}$ that are differentiable everywhere. This is because we want to deal with critical points, which forces the existence of a derivative. Note that if $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable everywhere this means that

$$\lim_{y \to x} \frac{f(y) - f(x) - \langle \nabla f(x), (y - x) \rangle}{|y - x|} = 0$$

for every $x \in \mathbb{R}^n$, and hence f admits a Taylor expansion of order one at every point. Moreover f satisfies the Lusin's N-property with respect to the Hausdorff measure $\mathcal{H}^1$. That is if $\mathcal{H}^1(E) = 0$ then $\mathcal{L}(f(E)) = 0$. This fact can be derived by an argument similar to that at the beginning of the proof of Theorem 1.21. Basically, everywhere differentiability of a function $f : \mathbb{R}^n \to \mathbb{R}$ implies that $\mathbb{R}^n$ can be decomposed into some sets $\bigcup_{j \geq 1} B_j = \mathbb{R}^n$ such that the restrictions $f|_{B_j}$ are Lipschitz. Therefore $\mathcal{H}^1-$null sets are sent to $\mathcal{H}^1-$null sets.

Theorem 2.12. *If $f : \mathbb{R} \to \mathbb{R}$ is differentiable everywhere then f satisfies the Morse-Sard property (that is, $\mathcal{L}(f(C_f)) = 0$).*

Theorem 2.13. *If $f : \mathbb{R}^2 \to \mathbb{R}$ is differentiable everywhere and*

$$\liminf_{y \to x} \frac{f(y) - f(x) - \langle \nabla f(x), (y - x) \rangle}{|y - x|^2} > -\infty$$

for $\mathcal{H}^1$-almost every $x \in \mathbb{R}^2$, then f satisfies the Morse-Sard property.

Theorem 2.14. *If $f : \mathbb{R}^n \to \mathbb{R}$, $n \geq 3$, is differentiable everywhere and admits a Taylor expansion $p_{n-1}(x; y)$ of order $n - 1$ at every point $x \in \mathbb{R}^n$ such that*

$$\liminf_{y \to x} \frac{f(y) - p_{n-1}(x; y)}{|y - x|^n} > -\infty$$

for $\mathcal{H}^1$-almost every $x \in \mathbb{R}^n$, then f has the Morse-Sard property.

We remark that we can replace $\mathbb{R}^n$ with any other open set $\Omega \subseteq \mathbb{R}^n$ in all the previous results.

Observe also that thanks to our results of the previous sections we know than under the hypothesis of any of the previous Theorems 2.12 and 2.13 we also get that f satisfies the Lusin property of class C^1 and C^2 respectively.

Chapter 3

Diffeomorphic extraction of closed sets in Banach spaces

The results of this chapter generalize important theorems on diffeomorphic *extractions* of some kind of sets. Although it is well known (see [47, 127, 67, 70] and the references therein) that every two separable, homotopy equivalent, infinite-dimensional Hilbert manifolds M, N are in fact diffeomorphic, a diffeomorphism $h : M \to N$ provided by this deep result has not been (and, in general, cannot be) shown to be limited by an arbitrary open cover $\mathcal{G}$ of M.

Definition 3.1. We say that a diffeomorphism $h : M \to \tilde{M} \subseteq M$ is *limited* by an open cover $\mathcal{G}$ of M provided that the set $\{\{x, h(x)\} : x \in M\}$ *refines* $\mathcal{G}$; that is, for every $x \in M$, we may find a $G_x \in \mathcal{G}$ such that both x and $h(x)$ are in G_x.

This property is essential in the development of Chapters 4 and 5. The finest result we know of which provides a diffeomorphism $h : E \to E \setminus X$ limited by a given open cover $\mathcal{G}$ of E, where E is a separable infinite-dimensional Hilbert space E and X is a closed subset of E, is a theorem of J. E. West [150] in which X is assumed to be locally compact. There is another important result that provides diffeomorphisms extracting locally compact sets. It is a result of Renz, see [134, Corollary 7 and 8] where he works with not necessarily Hilbert spaces but also with certain Banach spaces as c_0 and ℓ_p, $1 < p < \infty$. However the constructed diffeomorphisms are not limited by any open cover.

In the proof of Theorem 4.2, where we do not work necessarily with Hilbert spaces, we need to diffeomorphically extract a closed set X which is not necessarily locally compact, but merely locally contained in the graph of a continuous mapping defined on a complemented subspace of infinite codimension in E and taking values in its linear complement (for a precise explanation of this terminology, see the statement of Theorem 3.3), and also we want the diffeomorphism to be limited by a given open cover. As it has been already explained, no result in the literature suits our purposes. In this chapter we will construct diffeomorphisms h which extract such closed sets X and that are limited by a given open cover.

3.1 Diffeomorphic negligibility result

For the case of Hilbert spaces, the main result of this chapter is the following.

Theorem 3.2. *Let E be an infinite-dimensional Hilbert space, X a closed subset of E which is locally contained in the graph of a continuous function defined on a subspace of infinite codimension in E and taking values in its orthogonal complement, $\mathcal{G}$ an open cover of E, and U an open subset of E. Then, there exists a C^∞ diffeomorphism h of $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$.*

Theorem 3.2 is a straightforward consequence of the following much more general result, which is true for many Banach spaces not necessarily Hilbertian.

Theorem 3.3. *Let E be a Banach space, $p \in \mathbb{N} \cup \{\infty\}$, and $X \subset E$ be a closed set with the property that, for each $x \in X$, there exist a neighbourhood U_x of x in E, Banach spaces $E_{(1,x)}$ and $E_{(2,x)}$, and a continuous mapping $f_x : C_x \to E_{(2,x)}$, where C_x is a closed subset of $E_{(1,x)}$, such that:*

1. $E = E_{(1,x)} \oplus E_{(2,x)}$;
2. $E_{(1,x)}$ *has C^p smooth partitions of unity;*
3. $E_{(2,x)}$ *is infinite-dimensional and has a (not necessarily equivalent) norm of class C^p;*
4. $X \cap U_x \subset G(f_x)$, *where*

$$G(f_x) = \{y = (y_1, y_2) \in E_{(1,x)} \oplus E_{(2,x)} \,:\, y_2 = f_x(y_1), y_1 \in C_x\}.$$

Then, for every open cover $\mathcal{G}$ of E and every open subset U of E, there exists a C^p diffeomorphism h from $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$ (see Definition 3.1). Moreover, the same conclusion is true if we replace E with an open subset of E.

By a C^p smooth norm on $E_{(2,x)}$ we mean a (possibly nonequivalent) norm on $E_{(2,x)}$ which is of class C^p on $E_{(2,x)} \setminus \{0\}$.

Observe also that we do not assume separability of the Banach space E, in contrast with West and Renz's results. The separability of the space implies that any open cover of the space has a star-finite open refinement, meaning that every point of the space has a neighbourhood that only meets at most a finite number of sets of the cover, and this fact helps a lot in the proof of West's result when one wants the extracting diffeomorphism to be limited by any open cover. In absence of separability, we will use sigma-discrete refinements similar to [139] (see also [104, Lemma 3.3]).

When addressing this new *technical* result, the reader must observe the following: for Banach spaces with unconditional basis every compact subset can be seen as the graph of a continuous function defined from a closed set belonging to an infinite-codimensional subspace and taking values in its linear complement. This fact will be shown in the last Section 3.6, following Renz's thesis [134], which uses an important theorem due to Corson [55]. A consequence is that Theorem 3.2 generalizes West's theorem [150] and also the results from Renz's Thesis [134].

It is then not surprising that the proof of Theorem 3.3 combines ideas and techniques from Peter Renz's Ph.D. thesis [134], James West's paper [150], and also some of the work of D. Azagra and T. Dobrowolski [10, 14].

For more information about diffeomorphic extraction of closed sets in Banach spaces see for instance [34, 150, 134, 135, 63, 10, 14, 16].

3.2 The ε-strong C^p extraction property

Throughout this section E will be an infinite-dimensional Banach space. We will introduce an abstract concept of when a set X is *diffeomorphically extractible* from E. Later we will state a general result about extractibility for this kind of sets, this will be Theorem 3.9, from which our proof of Theorem 3.3 will rely on.

First let us introduce the following definitions.

Definition 3.4. We will say that a subset X of E has the strong C^p extraction property with respect to an open set U if $X \subseteq U$, X is relatively closed in U, and for every open set $V \subseteq U$, every subset $Y \subseteq X$ relatively closed in U there exists a C^p diffeomorphism φ from $U \setminus Y$ onto $U \setminus (Y \setminus V)$ which is the identity on $(U \setminus V) \setminus Y$. If in addition for any $\varepsilon > 0$ we can ask the diffeomorphism not to move points more than ε (that is to say, $||\varphi(x) - x|| \leq \varepsilon$ for all x), we will say that X has the ε-strong C^p extraction property with respect to U.
We will also say that such a closed set X has locally the strong (or ε-strong) C^p extraction property if for

every point $x \in X$ there exists an open neighbourhood U_x of x such that $X \cap \overline{U_x}$ has the strong (ε-strong respectively) C^p extraction property with respect to every open set U with $X \cap \overline{U_x} \subseteq U$ (equivalently, there exists an open neighbourhood U_x of x such that $X \cap U_x$ has the strong (ε-strong respectively) C^p extraction property with respect to every open set U for which $X \cap U_x$ is a relatively closed subset of U).

Remark 3.5. Let (U, W), $W \subset U$, be a pair of open sets in a Banach space E. We say that (U, W) has the strong C^p expansion property if, for every open subsets V and U' of U, $W \subset U'$, there exists a C^p diffeomorphism $U' \cap V \to V$ which, by letting $\varphi(x) = x$ for $x \in U' \setminus V$, extends to a C^p diffeomorphism $\varphi : U' \to U' \cup V$.

In particular, letting $U' = W$, there exists a C^p diffeomorphism $W \cap V \to V$ which extends to a C^p diffeomorphism of W onto $W \cup V$ via the identity off $W \cap V$. Hence, W is smoothly extended to $W \cup V$; this justifies the term of C^p expansion. Should this expansion be valid for all open sets U', $W \subset U'$, then we have the strong C^p expansion property.

Notice that a relatively closed subset X has the strong C^p extraction property with respect to U if and only if $(U, W) = (U, U \setminus X)$ has the strong C^p expansion property.

We say that an open subset W of E has locally the strong C^p expansion property if every $x \in E \setminus W$ has an open neighbourhood U_x such that $(U, U_x \cap W)$ has the strong expansion property for every open set $U \supset U_x \cap W$.

Notice that a closed set X has locally the strong C^p extraction property if and only if $W = E \setminus X$ has locally the strong C^p expansion property.

Some basic properties that can be derived from Definition 3.4 are listed in the following lemma.

Lemma 3.6. *Let us suppose that $X, X_1, X_2 \subset E$ have the ε-strong C^p extraction property with respect to an open set U of E. Then*

(1) *For every set $Y \subseteq X$, relatively closed in U, Y has the ε-strong C^p extraction property with respect to U;*

(2) *For every open subset $U' \subseteq U$, $X \cap U'$ has the ε-strong C^p extraction property with respect to U'.*

(3) *$X_1 \cup X_2$ has the ε-strong C^k extraction property with respect to U.*

(4) *If h is a C^p diffeomorphism defined on U and such that $h(U)$ is open, then $h(X)$ has the strong C^p extraction property with respect to $h(U)$.*

Proof.
(1) This follows directly from the definition.

(2) Take an open subset $V' \subseteq U'$, a subset $Y \subseteq X \cap U'$ relatively closed in U'. Since X has the strong C^p extraction property with respect to U there exists a C^p diffeomorphism φ from $U \setminus \overline{Y}$ onto $U \setminus (\overline{Y} \setminus V')$ which is the identity on $(U \setminus V') \setminus \overline{Y}$. When restricting φ to $U' \setminus Y$ we actually get a C^p diffeomorphism from $U' \setminus Y$ onto $U' \setminus (Y \setminus V')$ which is the identity on $(U' \setminus V') \setminus Y$.

(3) Take $Y \subseteq X_1 \cup X_2$ relatively closed in U and an open set $V \subseteq U$. We want to find a C^p diffeomorphism φ from $U \setminus Y$ onto $U \setminus (Y \setminus V)$ which is the identity on $(U \setminus V) \setminus Y$ and does not move points more than ε.
Define the sets $Y_1 = Y \cap X_1$ and $Y_2 = Y \cap X_2$, which are relatively closed in U and satisfy $Y_1 \cup Y_2 = Y$. In particular by (1) they have the ε-strong C^p extraction property with respect to U.
Firstly, there exists a C^p diffeomorphism $\varphi_1 : U \setminus Y_1 \to U \setminus (Y_1 \setminus V)$ which is the identity on $(U \setminus V) \setminus Y_1$ and does not move points more than $\varepsilon/2$.
Secondly, for the open set $U \setminus Y_1$, using (2) we know that $Y_2 \cap (U \setminus Y_1) = Y_2 \setminus Y_1$ has the ε-strong C^p extraction property with respect to $U \setminus Y_1$. Hence there exists a C^p diffeomorphism $\varphi_2 : U \setminus (Y_1 \cup Y_2) \to (U \setminus Y_1) \setminus ((Y_2 \setminus Y_1) \setminus V)$, which is the identity on $((U \setminus Y_1) \setminus V) \setminus (Y_2 \setminus Y_1)$ and does not move points

more than $\varepsilon/2$.
Observe that

$$\begin{aligned}\varphi_1((U \setminus Y_1) \setminus ((Y_2 \setminus Y_1) \setminus V)) &= [U \setminus (Y_1 \setminus V)] \setminus [\varphi_1((Y_2 \setminus Y_1) \setminus V)] \\ &= [U \setminus (Y_1 \setminus V)] \setminus [(Y_2 \setminus Y_1) \setminus V] \\ &= U \setminus (Y_1 \cup Y_2) \setminus V).\end{aligned}$$

Hence we can define a C^p diffeomorphism

$$\varphi := \varphi_1 \circ \varphi_2 : U \setminus (Y_1 \cup Y_2) \to U \setminus ((Y_1 \cup Y_2) \setminus V,$$

which is the identity on $(U \setminus V) \setminus (Y_1 \cup Y_2)$ and does not move points more than ε.

(4) Take an open subset V of $h(U)$, a subset $Y \subseteq h(X)$ relatively closed in $h(U)$. Since X has the strong C^p extraction property with respect to U and $h^{-1}(Y)$ is relatively closed in U, there exists a C^p diffeomorphism φ from $U \setminus h^{-1}(Y)$ onto $U \setminus (h^{-1}(Y) \setminus h^{-1}(V))$ which is the identity on $(U \setminus h^{-1}(V)) \setminus h^{-1}(Y)$. Then the mapping

$$g := h \circ \varphi \circ h^{-1} : h(U) \setminus Y \longrightarrow h(U) \setminus (Y \setminus V)$$

is a surjective C^p diffeomorphism which restricts to the identity on $(h(U) \setminus V) \setminus Y$. □

Remark 3.7. In Lemma 3.6 (4) we will not have in general the ε-strong C^p extraction property of $h(X)$ with respect to $h(U)$, but we still have the following: suppose h does not move the points more than some $\varepsilon > 0$. For every $\eta > 0$ in the proof of Lemma 3.6 (4) we can assume that φ does not move points more than η. Hence $g \circ h = (h \circ \varphi \circ h^{-1}) \circ h = h \circ \varphi$ does not move points more than $\varepsilon + \eta$.

Let us state the following two results which are crucial in the proof of Theorem 3.3 provided below.

Theorem 3.8. *Let $E = E_1 \times E_2$ be a product of Banach spaces such that E_1 admits C^p smooth partitions of unity and E_2 admits a C^p (not necessarily equivalent) norm. Assume that X_1 is a closed subset of E_1, that $f : E_1 \to E_2$ is a continuous mapping, and that E_2 is infinite-dimensional. Define*

$$X = \{(x_1, x_2) \in E_1 \times E_2 \, : \, x_1 \in X_1, x_2 = f(x_1)\}.$$

Let U be an open subset of E and $\varepsilon > 0$. Then there exists a C^p diffeomorphism g from $E \setminus X$ onto $E \setminus (X \setminus U)$ such that g is the identity on $(E \setminus U) \setminus X$ and moves no point more than ε.

Theorem 3.9. *Let E be a Banach space and X be a closed subset of E which has locally the ε-strong C^p extraction property. Let U be an open subset of E and $\mathcal{G} = \{G_r\}_{r \in \Omega}$ be an open cover of E. Then there exists a C^p diffeomorphism g from $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$.*

Proof of Theorem 3.3. Let us show that X from the statement of Theorem 3.3 has locally the ε-strong C^p extraction property.
To this end, fix $x \in X$ and choose a neighbourhood U_x such that $X \cap U_x \subset G(f_x)$; we can assume that $U_x = \overline{U_x}$. Further, we can assume that f_x is defined and continuous on the whole $E_{(1,x)}$. We will show that $X' := X \cap \overline{U_x}$ has the ε-strong C^p extraction property with respect to every open set U with $X' \subseteq U$. Notice that $X' = G(f_x|X_1')$ for a certain closed $X_1' \subset E_{(1,x)}$. Furthermore, if $Y' \subset X'$ is relatively closed in U, then Y' is closed in X'. Hence, $Y' = G(f_x|Y_1')$ for a certain closed $Y_1' \subset X_1' \subset E_{(1,x)}$. Let V be an open subset of U. Take now $\varepsilon > 0$ and apply Theorem 3.8 to $E_1 := E_{(1,x)}$, $E_2 := E_{(2,x)}$, $f := f_x$, $X := Y'$, and V (in place of U) to obtain a C^p diffeomorphism g from $E \setminus Y'$ onto $E \setminus (Y' \setminus V)$ such that g is the identity on $(E \setminus V) \setminus Y'$ and moves no point more than ε. Then $\varphi = g|U$ is as required in the definition of the ε-strong C^p extraction of X' with respect to U.
Now, an application of Theorem 3.9 concludes our proof. □

So the following two sections of the chapter are devoted to two objectives:

1. To give a proof of Theorem 3.8 in Section 3.3.
2. To give a proof of the abstract extractibility result, Theorem 3.9, in Section 3.4.

3.3 Extracting closed sets which are contained in graphs of infinite codimension

We will split the proof of Theorem 3.8 into three subsections. First, in Subsection 3.3.1, we will see that X can be flattened by means of two homeomorphisms $h, \varphi : E \to E$ which are sufficiently close to each other, and whose restrictions to $E \setminus X$ and $E \setminus (X \setminus U)$ are diffeomorphisms, respectively. Next, in Subsections 3.3.2 and 3.3.3, we will show that there exists a diffeomorphism $g : E \setminus (X_1 \times \{0\}) \to E \setminus ((X_1 \times \{0\}) \setminus h(U))$ which is the identity on $(E \setminus h(U)) \setminus (X_1 \times \{0\})$ and moves no point more than a fixed small number ε. Then, the composition $\varphi^{-1} \circ g \circ h$ will extract the local chunk of graph $U \cap X$ and will move no point too much.

In Subsection 3.3.1, we will closely follow Peter Renz's results from [134, 135]. In Subsections 3.3.2 and 3.3.3, we will combine ideas and techniques from [134, 10, 14].

3.3.1 Flattening graphs

Here we will prove the following.

Theorem 3.10. *Let E_1 be a Banach space with C^p smooth partitions of unity and E_2 be a Banach space which admits a (not necessarily equivalent) C^p norm. Let $(E = E_1 \times E_2, \|\cdot\|)$ and $\pi_1 : E \to E_1$ be the natural projection, i.e., $\pi_1(x_1, x_2) = x_1$, $(x_1, x_2) \in E$. Let $X_1 \subset E_1$ be a closed set, $f : X_1 \to E_2$ a continuous mapping, $U \subset E$ an open set, and $\varepsilon > 0$. Write $G(f) = \{(x_1, x_2) \in E : x_2 = f(x_1),\ x_1 \in X_1\}$. Then there exist a couple of homeomorphisms $h, \varphi : E \to E$ such that:*

(1) *$h(G(f)) \subset E_1 \times \{0\}$ and $\varphi(G(f) \setminus U) \subset (E_1 \times \{0\}) \setminus h(U)$;*

(2) *$h = \varphi$ off of U;*

(3) *$\pi_1 \circ h = \pi_1 = \pi_1 \circ \varphi$;*

(4) *h restricted to $E \setminus G(f)$ is a C^p diffeomorphism of $E \setminus G(f)$ onto $E \setminus (X_1 \times \{0\})$;*

(5) *φ restricted to $E \setminus (G(f) \setminus U)$ is a C^p diffeomorphism of $E \setminus (G(f) \setminus U)$ onto $E \setminus ((X_1 \times \{0\}) \setminus h(U))$;*

(6) *$\|h^{-1}(x) - \varphi^{-1}(x)\| \leq \varepsilon$ for every $x \in E$; and*

(7) *$h^{-1}(x_1, x_2)$ is uniformly continuous with respect to the second coordinate x_2, meaning that for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $\|x_2 - x_2'\| < \delta$ then $\|h^{-1}(x_1, x_2) - h^{-1}(x_1, x_2')\| < \varepsilon$ for all x_1.*

We will assume without loss of generality that $\varepsilon \leq 1$.

In what follows, slightly abusing notation, we will indistinctly use the symbol $\|\cdot\|$ to denote the norms $\|\cdot\|_{E_1}$, $\|\cdot\|_{E_2}$, and $\|\cdot\|$ with which the Banach spaces E_1, E_2 or $E_1 \times E_2$ are endowed. We may and do assume that $\|x_1\|_{E_1} = \|(x_1, 0)\|$ and $\|x_2\|_{E_2} = \|(0, x_2)\|$ for all $(x_1, x_2) \in E_1 \times E_2$.

Now, we state and prove a sequence of lemmas that will be employed in proving the above theorem. The most important are Lemmas 3.12 and 3.14. Basically, we follow the ideas of Renz's paper [135] and Ph.D. thesis [134], with some minor but very important changes.

The proof of our first lemma is a consequence of the existence of C^p smooth partitions of unity on E_1. We follow [60, Chapter VIII, section 3] for the proof of the next result.

Lemma 3.11. *The function $f : X_1 \to E_2$ extends to a continuous function $\bar{f} : E_1 \to E_2$ such that $\bar{f}|E_1 \setminus X_1$ is C^p smooth.*

Proof. Given the continuous function $f : X_1 \to E_2$, by the vector-valued Tietze's extension theorem [66] there is a continuous function $g_1 : E_1 \to E_2$ with $g_1|_{X_1} = f$. Using the existence of C^p-smooth partitions of unity on E_1 we find $h_1 : E_1 \to E_2$ of class C^p such that $||g_1(x_1) - h_1(x_1)|| \leq 1$ for every $x_1 \in E_1$. We can as well define $u_1 : E_1 \to [0, 1]$ of class C^p such that $u_1 = 1$ on X_1 and $u_1 = 0$ on $\{x_1 \in E_1 : \operatorname{dist}(x_1, X_1) \geq 1\}$ (for a proof of these facts see [60, Theorem VIII.3.2]). We have

1. $\sup_{x_1 \in X_1} ||f(x_1) - h_1(x_1)u_1(x_1)|| = \sup_{x_1 \in X_1} ||g_1(x_1) - h_1(x_1)u_1(x_1)|| \leq 1$.
2. $h_1(x_1)u_1(x_1) = 0$ if $\text{dist}(x_1, X_1) \geq 1$.

Consider the function $f - h_1 u_1 : X_1 \to E_2$, which is continuous, and we extend it to another continuous function $g_2 : E_1 \to E_2$ so that $g_2|_{X_1} = f - h_1 u_1$ and also $\sup_{x_1 \in E_1} ||g_2(x_1)|| \leq \sup_{x_1 \in X_1} ||f(x_1) - h_1(x_1)u_1(x_1)|| \leq 1$. Then there exists $h_2 : E_1 \to E_2$ of class C^p so that $||h_2(x_1) - g_2(x_1)|| \leq 1/2$ for all $x_1 \in E_1$. We as well define $u_2 : E_1 \to [0,1]$ of class C^p such that $u_2 = 1$ on X_1 and $u_2 = 0$ on $\{x_1 \in E_1 : \text{dist}(x_1, X_1) \geq 1/2\}$. We have

1. $\sup_{x_1 \in X_1} ||f(x_1) - h_1(x_1)u_1(x_1) - h_2(x_1)u_2(x_1)|| = \sup_{x_1 \in X_1} ||g_2(x_1) - h_2(x_1)|| \leq 1/2$.
2. $\sup_{x_1 \in E_1} ||h_2(x_1)u_2(x_1)|| \leq \sup_{x_1 \in E_1} ||g_2(x_1) - h_2(x_1)|| + ||g_2(x_1)|| \leq 3/2$.
3. $h_2(x_1)u_2(x_1) = 0$ if $\text{dist}(x_1, X_1) \geq 1/2$.

By induction, for $n \geq 2$ we find a sequence of mappings $h_n : E_1 \to E_2$ and $u_n : E_1 \to [0,1]$ of class C^p with $u_n = 1$ on X_1 and $u_n = 0$ on $\{x_1 \in E_1 : \text{dist}(x_1, X_1) \geq 2^{-n}\}$ so that

1. $\sup_{x_1 \in X_1} ||f(x_1) - \sum_{i=1}^{n} h_i(x_1)u_i(x_1)|| \leq 2^{-n}$.
2. $\sup_{x_1 \in E_1} ||h_n(x_1)u_n(x_1)|| \leq 3 \cdot 2^{-n}$.
3. $h_n(x_1)u_n(x_1) = 0$ if $\text{dist}(x_1, X_1) \geq 2^{-n}$.

Finally define $\bar{f}(x_1) = \sum_{i=1}^{\infty} h_i(x_1)u_i(x_1)$, which is a continuous function because the series $\sum_{i=2}^{\infty} h_i u_i$ is absolutely and uniformly convergent. If $x_1 \in X_1$ we have $f(x_1) = \bar{f}(x_1)$ and if $x_1 \notin X_1$ there is a neighbourhood on which all but finitely many of the f_i's are identically zero, therefore $\bar{f}$ is of class C^p outside X_1. □

Lemma 3.12. *Let E_1 be a Banach space with C^p smooth partitions of unity, X_1 be a closed subset of E_1, and $f : X_1 \to E_2$ be a continuous mapping. For every $n \in \mathbb{N}$, write*

$$W_n = \left\{ x_1 \in E_1 : \text{dist}(x_1, X_1) \leq \frac{1}{n} \right\}.$$

Assume $\bar{f} : E_1 \to E_2$ is a continuous extension of f such that $\bar{f}|E_1 \setminus X_1$ is C^p smooth. Then, there is a continuous mapping

$$F : \mathbb{R} \times E_1 \to E_2$$

such that

(1) *$F(r, x_1) = \bar{f}(x_1)$ for all $(r, x_1) \in (r_n, \infty) \times E_1 \setminus W_n$ and some $0 < r_n < 1$; in particular, $F(r, x_1) = \bar{f}(x_1)$ for all (r, x_1) in some neighbourhood of the set $\{1\} \times (E_1 \setminus X_1)$ in $\mathbb{R} \times (E_1 \setminus X_1)$;*

(2) *$F|\mathbb{R} \times (E_1 \setminus X_1) \cup (-\infty, 1) \times E_1$ is C^p smooth;*

(3) *$F(r, x_1) = f(x_1)$ for $r \geq 1$ and $x_1 \in X_1$;*

(4) *$\|D_1 F(r, x_1)\| \leq \frac{1}{2}$ for all $r \in \mathbb{R}$, $x_1 \in E_1$.*

Proof. For every $n \in \mathbb{N}$ we can find a sequence of C^p functions $\bar{f}_n : E_1 \to E_2$ such that

$$||\bar{f}(x_1) - \bar{f}_n(x_1)|| \leq 2^{-2n-4}$$

for every $x_1 \in E_1$. The existence of such a sequence is again guaranteed by the existence of C^p partitions of unity in E_1 (see again for instance [60, Theorem VIII.3.2]). We will now improve the sequence $\{\bar{f}_n\}_{n \geq 1}$ to $\{f_n\}_{n \geq 1}$ so that the sequence $\{f_n|E_1 \setminus X_1\}$ locally stabilizes with respect to n. To achieve this, we use the existence of C^p partitions of unity to find a C^p function $\lambda_n : E_1 \to [0,1]$ which is 1 on $E \setminus W_n$ and 0 on W_{n+1}. Define

$$f_n(x_1) = \lambda_n(x_1)\bar{f}(x_1) + (1-\lambda_n(x_1))\bar{f}_n(x_1)$$

for all $x_1 \in E_1$. It follows that $\|f_n(x_1) - f_{n+1}(x_1)\| \leq 2^{-2n-3}$ for all $x_1 \in E_1$.
For $n \in \mathbb{N}$, pick a nondecreasing C^∞ function $h_n : \mathbb{R} \to [0,1]$ such that $h_n(r) = 0$ for $r \leq 1 - 2^{1-n}$, $h_n(r) = 1$ for $r \geq 1 - 2^{-n}$, and $h'_n(r) \leq 2^{n+1}$. One can check that

$$F(r, x_1) = f_1(x_1) + \sum_{n=1}^{\infty} h_{n+1}(r)(f_{n+1}(x_1) - f_n(x_1))$$

defines a required mapping. Let us check it.

(1) Given $n \in \mathbb{N}$ take $r_n \geq 1 - 2^{-n}$, then if $r \in (r_n, \infty)$ we have $h_k(r) = 1$ for all $k \leq n$. On the other hand if $x_1 \in E_1 \setminus W_n$, then $\lambda_k(x_1) = 1$ and hence $f_k(x_1) = \bar{f}(x_1)$ for all $k \geq n$. Therefore if $(r, x_1) \in (r_n, \infty) \times E_1 \setminus W_n$ we have $F(r, x_1) = f_n(x_1) = \bar{f}(x_1)$.

(2) The functions h_n are C^∞ everywhere. If $x_1 \in E \setminus X_1$ there exists n_0 such that $x_1 \notin W_n$ for all $n \geq n_0$ and therefore $f_n(x_1) = \bar{f}(x_1)$ for all $n \geq n_0$. We then have a finite sum

$$F(r, x_1) = f_1(x_1) + \sum_{n=1}^{n_0-1} h_{n+1}(r)(f_{n+1}(x_1) - f_n(x_1)),$$

where all the summands are C^p functions on $\mathbb{R} \times (E_1 \setminus X_1)$.

Suppose now that $(r, x_1) \in (-\infty, 1) \times E_1$. Then there is a neighbourhood of r and some $n_0 \in \mathbb{N}$ for which $h_n(r) = 0$ for all $n \geq n_0$ and all r in that neighbourhood. Hence $F(r, x_1)$ is defined by a finite sum of C^p functions and will be C^p smooth as well.

(3) Let $r \geq 1$. Then $h_n(r) = 1$ for all $n \in \mathbb{N}$ and $F(r, x_1) = \lim_{n\to\infty} f_n(x_1)$ for every $x_1 \in E_1$. If we also let $x_1 \in X_1$ we have $x_1 \in W_n$ for every $n \in \mathbb{N}$ and so $f_n(x_1) = \bar{f}_n(x_1)$ for all $n \in \mathbb{N}$. We can then write $F(r, x_1) = \lim_{n\to\infty} \bar{f}_n(x_1) = \bar{f}(x_1)$.

(4) Finally for every $(r, x_1) \in \mathbb{R} \times E_1$,

$$\|D_1 F(r, x_1)\| \leq \sum_{n=1}^{\infty} |h'_{n+1}(r)| \, ||f_{n+1}(x_1) - f_n(x_1)|| \leq \sum_{n=1}^{\infty} 2^{n+2} \cdot 2^{-2n-3} = \sum_{n=1}^{\infty} 2^{-n-1} = \frac{1}{2}.$$

□

Observe that in fact $F(r, x_1)$ is Lipschitz with constant 1 with respect to the first variable $r \in \mathbb{R}$. That is,

$$\begin{aligned} ||F(r, x_1) - F(r', x_1)|| &\leq \sum_{n=1}^{\infty} |h_{n+1}(r) - h_{n+1}(r')| ||f_{n+1}(x_1) - f_n(x_1)|| \\ &\leq \sum_{n=1}^{\infty} 2^{n+2} |r - r'| 2^{-2n-3} \leq |r - r'| \end{aligned}$$

for every $x_1 \in E_1$.
We will write

$$U_0 = \pi_1(G(\bar{f}) \cap U),$$

which is an open set in E_1, and also

$$Y_1 = X_1 \setminus U_0 = X_1 \setminus \pi_1(G(\bar{f}) \cap U) = \pi_1(G(f) \setminus U),$$

which is a closed subset of E_1. By replacing U with $U \cap \pi_1^{-1}(U_0)$, we can assume that

$$U_0 = \pi_1(U).$$

Lemma 3.13. *With the above notation, take a decreasing sequence of positive numbers $\{\delta_n\}_{n\geq 1}$ converging to zero. Then there exists an increasing sequence of open subsets in E_1*

$$V_1 \subseteq \overline{V_1} \subseteq V_2 \subseteq \overline{V_2} \subseteq \cdots \subseteq \overline{V_n} \subseteq V_{n+1} \subseteq \cdots \subseteq U_0$$

such that $\bigcup_{n=1}^{\infty} V_n = U_0$ and the sets

$$U_n := \{(x_1, x_2) \in E : \|x_2 - \bar{f}(x_1)\| < \delta_n, \, x_1 \in V_n\}$$

are contained in U.

Proof. To be able to get the required inclusions between the sets V_n, we first take an auxiliary sequence of open sets W_n in U_0 such that $\overline{W_n} \subseteq W_{n+1}$ for every $n \in N$ and $\bigcup_{n=1}^{\infty} W_n = U_0$.
Then we define

$$V_n' = \{x_1 \in U_0 : \{(x_1, x_2) \in E : \|x_2 - \bar{f}(x_1)\| < \delta_n\} \subseteq U\}$$

for every $n \in \mathbb{N}$. Observe that we have $V_n' \subseteq V_{n+1}'$ and $\bigcup_{n=1}^{\infty} V_n' = U_0$, but we cannot assure that $\overline{V_n'} \subseteq V_{n+1}'$ for every $n \in N$. So now we mix these sets with the previous W_n, that is, we let $V_n = W_n \cap V_n'$. Obviously, by definition, for every $n \in \mathbb{N}$ the set $U_n = \{(x_1, x_2) \in E : \|x_2 - \bar{f}(x_1)\| < \delta_n, \, x_1 \in V_n\}$ is contained in U. Now, we have that $\overline{V_n} \subseteq V_{n+1}$ for every $n \in N$; also $\bigcup_{n=1}^{\infty} V_n = U_0$. □

The following lemma resembles [135, Lemma 2.2] and [134, Lemma 2] (in which only one function ϕ is considered). However, Theorem 3.10 requires constructing two homeomorphisms h and φ which are identical outside U. The building block in constructing those homeomorphisms are two functions ϕ and $\tilde{\phi}$ whose existence is claimed in the lemma below. The existence of $\tilde{\phi}$ is crucial. Incidentally, let us note that Renz's proof of [134, Theorem 4] is flawed (and this is the reason why we must deal with two functions ϕ and $\tilde{\phi}$ instead of just the function ϕ), but can be corrected by using Theorem 3.10.

Lemma 3.14. *Let $\bar{f} : E_1 \to E_2$ be the uniform limit of C^p functions, where E_1 has C^p smooth partitions of unity and E_2 has a (not necessarily equivalent) C^p smooth norm. Then there are two continuous functions ϕ, $\tilde{\phi} : E \to [0, 1]$ such that*

(1) $\phi^{-1}(1) = G(\bar{f})$ *and* $\tilde{\phi}^{-1}(1) = G(\bar{f}) \setminus U$;

(2) $\phi|E \setminus G(\bar{f})$ *and* $\tilde{\phi}|E \setminus (G(\bar{f}) \setminus U)$ *are C^p smooth;*

(3) $\|D_2\phi(x_1, x_2)\| \leq \frac{1}{2}$ *for all* $(x_1, x_2) \in E \setminus G(\bar{f})$, *and* $\|D_2\tilde{\phi}(x_1, x_2)\| \leq \frac{1}{2}$ *for all* $(x_1, x_2) \in E \setminus (G(\bar{f}) \setminus U)$;

(4) $\phi = \tilde{\phi}$ *outside U.*

Proof. To construct ϕ we will follow [135, Lemma 2.2]. A similar argument will be used to construct $\tilde{\phi}$; however, we have to make sure that $\tilde{\phi}|G(\bar{f}) \cap U < 1$.
For $n \in \mathbb{N}$, let $a_n, b_n, c_n, d_n, \varepsilon_n$ be positive numbers with the following properties:

(1) they tend to zero as n tends to infinity;

(2) $a_n < b_n$ for all n;

(3) $\varepsilon_{n+1} + b_{n+1} < a_n - \varepsilon_n$ for all n;

(4) $\sum_{n=1}^{\infty} c_n \leq \frac{\epsilon}{2} \leq 1$ (here $\varepsilon > 0$ is the one given by the statement of Theorem 3.10);

(5) $\sum_{n=1}^{\infty} d_n \leq \frac{1}{2}$

(for instance, let us set $a_n = \epsilon 2^{-2n}$, $b_n = 2a_n$, $c_n = \epsilon 2^{-5n}$, $d_n = 2^{-2n}$ and $\varepsilon_n = \epsilon 2^{-4(n+1)}$). Let h_n be a nonincreasing C^p function from $\mathbb{R}$ to $\mathbb{R}$ satisfying

$$\begin{aligned} &c_n = h_n(r) = h_n(0) > 0 && \text{whenever } r \le a_n, \\ &h_n(r) = 0 && \text{whenever } r \ge b_n, \\ &|h_n'(r)| \le d_n && \text{for all } r \text{ in } \mathbb{R}, \end{aligned}$$

and $g_n : E_1 \to E_2$ be a C^p mapping such that $\|g_n(x_1) - \bar{f}(x_1)\| \le \varepsilon_n$ for every $x_1 \in E_1$. Then

$$\psi_n(x_1, x_2) = h_n(\|x_2 - g_n(x_1)\|)$$

defines a nonnegative C^p function on $E_1 \times E_2 = E$ satisfying

$$\begin{aligned} &c_n = \psi_n(x_1, x_2) = h_n(0) > 0 && \text{if } \|x_2 - \bar{f}(x_1)\| \le a_n - \varepsilon_n, \\ &\psi_n(x_1, x_2) = 0 && \text{if } \|x_2 - \bar{f}(x_1)\| \ge b_n + \varepsilon_n, \\ &\|D_2\psi_n(x_1, x_2)\| \le d_n && \text{for all } (x_1, x_2) \in E_1 \times E_2. \end{aligned}$$

The nonnegativity and first two properties of ψ_n are evident, and it is easy to see that ψ_n is C^p on E. The bound on the norm of the derivative $D_2\psi_n$ is established by using the chain rule and the fact that the operator norm of the derivative of the norm of any Banach space is less than or equal to one.
Define

$$\psi(x_1, x_2) = \sum_{n=1}^{\infty} \psi_n(x_1, x_2) \tag{3.3.1}$$

for all $(x_1, x_2) \in E$.
Similarly, define

$$\tilde{\psi}(x_1, x_2) = \sum_{n=1}^{\infty} \lambda_n(x_1)\psi_n(x_1, x_2) \tag{3.3.2}$$

for all $(x_1, x_2) \in E$, where $\lambda_n : E_1 \to [0,1]$ is a C^p smooth function such that $\lambda_n(x_1) = 1$ if $x_1 \notin V_n$ and $\lambda_n(x_1) = 0$ if $x \in V_{n-1}$ (here we are using that E_1 has C^p smooth partitions of unity). Here, the sets V_n are provided by Lemma 3.13 for the sequence $\delta_n := \varepsilon_n + b_n$ (let $V_0 = \emptyset$ and assume $V_1 \neq \emptyset$). In particular, observe that since $\bigcup_{n=1}^{\infty} V_n = U_0$ then $\lambda_n(x_1) = 1$ for every $x_1 \notin U_0$; hence, $\psi(x_1, x_2) = \tilde{\psi}(x_1, x_2)$ for $x_1 \notin U_0$.
Since the functions ψ and $\tilde{\psi}$ are defined via absolutely and uniformly convergent series of continuous functions, they are continuous.
If $(x_1, x_2) \in E \setminus G(\bar{f})$, then $\|x_2 - \bar{f}(x_1)\| > b_n + \varepsilon_n$ for some $n \in \mathbb{N}$. By continuity, the inequality holds in a neighbourhood of (x_1, x_2) so ψ_k vanishes for $k \ge n$. Hence, ψ is locally a finite sum of C^p functions, and in particular is C^p on $E \setminus G(\bar{f})$.
Also, if $(x_1, x_2) \in G(\bar{f}) \cap U$, then $x_1 \in U_0$ and $x_1 \in V_n$ for some $n \in \mathbb{N}$. So $\lambda_k(x_1) = 0$ for all $k \ge n+1$. This means that $\tilde{\psi}$ is locally a finite sum of C^p functions and, thus, is of class C^p on $E \setminus (G(\bar{f}) \setminus U)$.
The derived series for $D_2\psi$ and $D_2\tilde{\psi}$ are absolutely and uniformly convergent in view of the bounds on $\|D_2\psi_n\|$ and the fact that $D_2\lambda_n = 0$. Then differentiation term by term is justified and $\|D_2\psi(x_1, x_2)\| \le \frac{1}{2}$ for all $(x_1, x_2) \in E \setminus G(\bar{f})$ and $\|D_2\tilde{\psi}(x_1, x_2)\| \le \frac{1}{2}$ for all $(x_1, x_2) \in E \setminus (G(\bar{f}) \setminus U)$.
Each point $(x_1, x_2) \in G(\bar{f})$ satisfies $0 = \|x_2 - \bar{f}(x_1)\| < a_n - \varepsilon_n$ for all $n \in \mathbb{N}$, consequently ψ equals the constant

$$d^* = \sum_{n=1}^{\infty} \psi_n(x_1, \bar{f}(x_1)) = \sum_{n=1}^{\infty} h_n(0) = \sum_{n=1}^{\infty} c_n \le 1.$$

On the other hand, if $(x_1, x_2) \in G(\bar{f}) \setminus U$, then $0 = \|x_2 - \bar{f}(x_1)\| < a_n + \varepsilon_n$ and $\lambda_n(x_1) = 1$ for all $n \in \mathbb{N}$, so $\tilde{\psi}$ equals again the constant d^*. In fact d^* is the supremum of ψ and of $\tilde{\psi}$, and is easily seen to be attained in $G(\bar{f})$ and $G(\bar{f}) \setminus U$, respectively.
To show that ψ and $\tilde{\psi}$ are equal outside U take $(x_1, x_2) \in E$. By a remark after the definition of ψ and $\tilde{\psi}$, we can assume $x_1 \in U_0$.

Claim 3.15. *For every $n \in \mathbb{N}$, if $x_1 \in V_n \setminus V_{n-1}$ and if $x_2 \in E_2$ is such that $\|x_2 - \bar{f}(x_1)\| \geq b_n + \varepsilon_n$, then $\psi(x_1, x_2) = \tilde{\psi}(x_1, x_2)$.*

Proof of Claim. If $\|x_2 - \bar{f}(x_1)\| \geq b_n + \varepsilon_n$ we have that $\psi_k(x_1, x_2) = 0$ for all $k \geq n$. So we have to see that $\lambda_k(x_1) = 1$ for $k = 1, \dots, n-1$. But this is clear since $x_1 \notin V_{n-1}$ and hence $x_1 \notin V_k$ for any $k = 1, \dots, n-1$. □

Now, we can conclude that for each $n \in \mathbb{N}$, $\psi = \tilde{\psi}$ on the set

$$((V_n \setminus V_{n-1}) \times E_2) \setminus U_n \supseteq ((V_n \setminus V_{n-1}) \times E_2) \setminus U.$$

Since $\bigcup_{n=1}^{\infty} V_n \setminus V_{n-1} = U_0$ and $\bigcup_{n=1}^{\infty} U_n \subseteq U$, it follows that ψ is equal to $\tilde{\psi}$ outside U.
Finally, to obtain functions ϕ and $\tilde{\phi}$ with the desired properties it is sufficient to set

$$\phi(x_1, x_2) = \psi(x_1, x_2) + 1 - d^*$$
$$\tilde{\phi}(x_1, x_2) = \tilde{\psi}(x_1, x_2) + 1 - d^*$$

for all $(x_1, x_2) \in E$. This ensures that the supremum, which is attained precisely on $G(\bar{f})$ for ϕ and precisely on $G(\bar{f}) \setminus U$ for $\tilde{\phi}$, is equal to 1.

□

In the proof of Theorem 3.10, we will employ a well known fact stating that the identity mapping perturbed by a contracting mapping is a homeomorphism (even a diffeomorphism provided that the contracting mapping is smooth). This fact is stated and proved in Lemma 3 of Renz's Ph.D. thesis [134].

Lemma 3.16. *Let E_1 be a normed linear space and E_2 be a Banach space. Let $E = E_1 \times E_2$ and let $d : E \to E_2$ be a continuous mapping satisfying the following condition*

$$\|d(x_1, x_2) - d(x_1, x_2')\| \leq \frac{1}{2}\|x_2 - x_2'\|$$

for all $x_1 \in E_1$ and $x_2, x_2' \in E_2$. Then the mapping defined by $h(x_1, x_2) = (x_1, x_2 - d(x_1, x_2))$ is a homeomorphism of E onto itself. Moreover, h is a C^p diffeomorphism when restricted to any open set (onto its image) on which d is C^p smooth.

Proof. To show that h is a bijection it suffices to show that h restricted to $\{x_1\} \times E_2$ is a bijection of $\{x_1\} \times E_2$ with itself for each $x_1 \in E_1$.
Let $x_1 \in E_1$ and $y_2 \in E_2$ be arbitrary but fixed during the following paragraph. Notice that $h(x_1, x_2) = (x_1, y_2)$ if and only if $x_2 - d(x_1, x_2) = y_2$ or equivalently $x_2 = y_2 + d(x_1, x_2)$. However the map $x_2 \mapsto y_2 + d(x_1, x_2)$ is a contraction map and so by Banach's fixed point theorem[1] there is one and only one fixed point $a_2 \in E_2$ such that $a_2 = y_2 + d(x_1, a_2)$, or equivalently such that $h(x_1, a_2) = (x_1, y_2)$. Hence we have proved that h is bijection.
Since d is continuous it is clear that h is continuous. We show now that h is locally open, hence it will be open, and therefore a homeomorphism. For a fixed point $(y_1, y_2) \in E$ and for any open neighbourhood of it, there exists an open set contained in it $U_1 \times U_2$ and some $\varepsilon > 0$ such that

1. $U_2 = \{x_2 \in E_2 : ||x_2 - y_2|| < \varepsilon\}$ and
2. $||d(x_1, y_2) - d(y_1, y_2)|| < \varepsilon/4$ for every $x_1 \in U_1$.

We assert that

$$h(U_1 \times U_2) \supset W := U_1 \times \{x_2 \in E_2 : ||x_2 - (y_2 - d(y_1, y_2))|| < \varepsilon/4\}. \quad (3.3.3)$$

[1] It is mandatory to assume completeness of the space E_2 where we apply the theorem.

Since W is an open neighbourhood of $h(y_1, y_2)$ this implies that h is an open mapping. To check the inclusion (3.3.3) let us show that if $(x_1, x_2) \notin U_1 \times U_2$ then $h(x_1, x_2) \notin W$. Clearly if $x_1 \notin U_1$, $h(x_1, x_2) \notin W$. On the other hand if $x_1 \in U_1$ and $x_2 \notin U_2$,

$$\begin{aligned}||\pi_2(h(x_1,x_2)) - (y_2 - d(y_1,y_2))|| =& ||x_2 - d(x_1,x_2) - y_2 + d(y_1,y_2)|| \geq ||x_2 - y_2|| \\ & - (||d(x_1,x_2) - d(x_1,y_2)|| + ||d(x_1,y_2) - d(y_1,y_2)||) \\ \geq & ||x_2 - y_2|| - \left(\frac{1}{2}||x_2 - y_2|| + \frac{\varepsilon}{4}\right) = \frac{1}{2}||x_2 - y_2|| - \frac{\varepsilon}{4} \\ \geq & \frac{\varepsilon}{2} - \frac{\varepsilon}{4} \geq \frac{\varepsilon}{4},\end{aligned}$$

and consequently $h(x_1, x_2) \notin W$. Using the surjectivity of h it follows that $W \subset h(U_1 \times U_2)$.
The above arguments show that h is an open continuous bijection from E onto itself, hence a homeomorphism. By the inverse function theorem (see for instance [117, Chapter 1, Theorem 5.2]) in order to complete the proof it suffices to see that $Dh(x_1, x_2)$ is a linear isomorphism for every point (x_1, x_2) at which h is of class C^p. First notice that if d has a derivative at (x_1, x_2) then h has a derivative at (x_1, x_2) too. So at the points (x_1, x_2) where d is C^p smooth, $Dh(x_1, x_2)$ will be a continuous linear mapping of the form

$$Dh(x_1,x_2)(a_1,a_2) = (a_1, a_2 - D_1d(x_1,x_2)(a_1) - D_2d(x_1,x_2)(a_2))$$

for every $(a_1, a_2) \in E$. It is enough to show that $Dh(x_1, x_2)$ is a bijection and then an application of the open mapping theorem finishes the proof. We can proceed as at the beginning of this proof and check that $Dh(x_1, x_2)$ is a bijection from $\{a_1\} \times E_2$ into itself for every $a_1 \in E_1$. Fix $a_1 \in E_1$ and $b_2 \in E_2$. We have that $Dh(x_1, x_2)(a_1, a_2) = (a_1, b_2)$ if and only if $a_2 = D_1d(x_1, x_2)(a_1) + D_2d(x_1, x_2)(a_2) + b_2$. But the map $a_2 \mapsto D_1d(x_1, x_2)(a_1) + D_2d(x_1, x_2)(a_2) + b_2$ is a contraction map since $||D_2d(x_1, x_2)|| \leq \frac{1}{2}$, so by Banach's fixed point theorem there exists a unique c_2 such that $c_2 = D_1d(x_1, x_2)(a_1) + D_2d(x_1, x_2)(c_2) + b_2$. We have then showed the surjectivity of $Dh(x_1, x_2)$ and the proof is complete.

□

Let us now present the proof of Theorem 3.10.

Proof of Theorem 3.10. Basically, we will follow the proof of Theorem 1 of Renz's Ph.D. thesis [134]. First, we apply Lemma 3.11 to $f : X_1 \to E_2$ to obtain a continuous mapping $\bar{f} : E_1 \to E_2$ such that $\bar{f}|X_1 = f$ and $\bar{f}|E_1 \setminus X_1$ is C^p smooth. Then, we apply Lemma 3.12 to the mapping $\bar{f}$ to obtain a mapping $F : \mathbb{R} \times E_1 \to E_2$ satisfying conditions (1)–(4) of Lemma 3.12. Next, we apply Lemma 3.14 to $\bar{f}$ to obtain functions ϕ and $\tilde{\phi}$ satisfying conditions (1)–(4) of Lemma 3.14. Now, we define

$$\begin{aligned}d(x_1,x_2) &= F(\phi(x_1,x_2),x_1) \\ \tilde{d}(x_1,x_2) &= F(\tilde{\phi}(x_1,x_2),x_1).\end{aligned}$$

Let us check that Lemma 3.16 is applicable to d and $\tilde{d}$ so that

$$h(x_1,x_2) = (x_1, x_2 - d(x_1,x_2))$$

and

$$\varphi(x_1,x_2) = (x_1, x_2 - \tilde{d}(x_1,x_2))$$

are homeomorphisms (which, additionally, will satisfy the conditions enumerated in Theorem 3.10).
Both functions d and $\tilde{d}$ are continuous as compositions of continuous functions. We compute D_2d and $D_2\tilde{d}$ to obtain

$$\begin{aligned}D_2d(x_1,x_2) &= D_1F(\phi(x_1,x_2),x_1) \circ D_2\phi(x_1,x_2) \\ D_2\tilde{d}(x_1,x_2) &= D_1F(\tilde{\phi}(x_1,x_2),x_1) \circ D_2\tilde{\phi}(x_1,x_2).\end{aligned}$$

The estimates of the norms of D_1F, $D_2\phi$, and $D_2\tilde{\phi}$ yields $\|D_2d(x_1,x_2)\|$, $\|D_2\tilde{d}(x_1,x_2)\| \leq \frac{1}{4}$ when $(x_1,x_2) \notin G(\bar{f})$. Since d and $\tilde{d}$ are continuous and $E \setminus G(\bar{f})$ is dense in E, by the mean value theorem, we can write

$$\|d(x_1,x_2) - d(x_1,x_2')\| \leq \frac{1}{4}\|x_2 - x_2'\|$$
$$\|\tilde{d}(x_1,x_2) - \tilde{d}(x_1,x_2')\| \leq \frac{1}{4}\|x_2 - x_2'\|$$

for all $x_1 \in E_1$ and all $x_2, x_2' \in E_2$. Hence, Lemma 3.16 applies and yields that h and φ are homeomorphisms.

Let us show conditions (1)–(7) of Theorem 3.10.

First, we will verify condition (1). If $(x_1,x_2) \in G(f)$, then $d(x_1,x_2) = F(1,x_1) = f(x_1) = x_2$ and

$$h(x_1,x_2) = (x_1, x_2 - d(x_1,x_2)) = (x_1, 0). \tag{3.3.4}$$

If $(x_1,x_2) \in G(f) \setminus U$, then $\tilde{d}(x_1,x_2) = F(1,x_1) = x_2$ and

$$\varphi(x_1,x_2) = (x_1, x_2 - \tilde{d}(x_1,x_2)) = (x_1, 0). \tag{3.3.5}$$

Condition (3) is obvious. Since ϕ and $\tilde{\phi}$ are equal outside U we obtain (2).

Let us see that $d|E \setminus G(f)$ and $\tilde{d}|E \setminus (G(f) \setminus U)$ are C^p diffeomorphisms. If $(x_1,x_2) \notin G(\bar{f})$ then ϕ and $\tilde{\phi}$ are C^p smooth and $\phi(x_1,x_2)$, $\tilde{\phi}(x_1,x_2) < 1$ by condition (2) and (1) of Lemma 3.14. It follows that $F(\phi(x_1,x_2),x_1)$ and $F(\tilde{\phi}(x_1,x_2),x_1)$ are C^p smooth in a neighbourhood of (x_1,x_2). Thus h and φ are C^p on $E \setminus G(\bar{f})$.

On the other hand, we have $\phi|G(\bar{f}) \setminus G(f) = 1$. Then, by continuity of ϕ and condition (1) of Lemma 3.12, we infer that $F(\phi(x_1,x_2),x_1) = \bar{f}(x_1)$ and, consequently, $h(x_1,x_2) = (x_1, x_2 - \bar{f}(x_1))$ in a neighbourhood of $G(\bar{f}) \setminus G(f)$. We have proved that d is C^p smooth on $E \setminus G(f)$.

It remains to show that $\varphi|U$ is C^p smooth. By condition (1) of Lemma 3.14, we have $\tilde{\phi}|U < 1$; by condition (2) of Lemma 3.12, $\tilde{d}|U$ is C^p smooth. The proof that $\tilde{d}$ and, therefore, φ restricted to $E \setminus (G(f) \setminus U)$ is C^p smooth is complete.

Now, Lemma 3.16 tells us that h and φ are C^p diffeomorphisms of $E \setminus G(f)$ onto $h(E \setminus G(f))$ and $E \setminus (G(f) \setminus U)$ onto $\varphi(E \setminus (G(f) \setminus U))$. So to get (4) and (5) it is sufficient to show that $h(G(f)) = X_1 \times \{0\}$ and $\varphi(G(f) \setminus U) = (X_1 \times \{0\}) \setminus h(U)$. The first equality is clear from equation (3.3.4). For the second one, observe that (3.3.5) tells us that $\varphi(G(f) \setminus U) = (X_1 \setminus U_0) \times \{0\} = Y_1 \times \{0\}$. So, we must check that

$$(X_1 \times \{0\}) \setminus h(U) = (X_1 \setminus U_0) \times \{0\},$$

or, what is the same, that $\pi_1(h(U)) = U_0$. The latter follows from condition (3) and the fact that $\pi_1(U) = U_0$.

Let us finish the proof by showing (6) and (7). Firstly let us check that $||h^{-1}(x_1,x_2) - \varphi^{-1}(x_1,x_2)|| \leq \varepsilon$ for every $(x_1,x_2) \in E_1 \times E_2$. Since h^{-1} preserves the first coordinate, we can write $h^{-1}(x_1,x_2) = (x_1,y_2)$ and $\varphi^{-1}(x_1,x_2) = (x_1,z_2)$ where $y_2, z_2 \in E_2$ are such that

$$y_2 - d(x_1,y_2) = x_2$$
$$z_2 - \tilde{d}(x_1,z_2) = x_2.$$

We then have that $||h^{-1}(x_1,x_2) - \varphi^{-1}(x_1,x_2)|| \leq \varepsilon$ if and only if $||y_2 - z_2|| \leq \varepsilon$ and if and only if $||d(x_1,y_2) - \tilde{d}(x_1,z_2)|| \leq \varepsilon$. Since $r \mapsto F(r,x_1)$ is 1-Lipschitz, this is true if $|\phi(x_1,y_2) - \tilde{\phi}(x_1,z_2)| \leq \epsilon$, or what is the same if $|\psi(x_1,y_2) - \tilde{\psi}(x_1,z_2)| \leq \epsilon$. And this is the case because

$$|\psi(x_1,y_2) - \tilde{\psi}(x_1,z_2)| \leq |\psi(x_1,y_2)| + |\tilde{\psi}(x_1,z_2)| = |\sum_{n=1}^{\infty} \psi_n(x_1,y_2)| + |\sum_{n=1}^{\infty} \lambda_n(x_1)\psi_n(x_1,z_2)|$$
$$\leq \sum_{n=1}^{\infty} c_n + \sum_{n=1}^{\infty} c_n \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Secondly, let us see that for every $\eta > 0$ there exists $\delta > 0$ such that if $||(x_1, x_2) - (x_1, x_2')|| = ||x_2 - x_2'|| \leq \delta$ then $||h^{-1}(x_1, x_2) - h^{-1}(x_1, x_2')|| \leq \eta$. It will be enough to set $\delta = \frac{\eta}{2}$. Indeed, take $(x_1, x_2), (x_1, x_2') \in E_1 \times E_2$ such that $||(x_1, x_2) - (x_1, x_2')|| = ||x_2 - x_2'|| \leq \frac{\eta}{2}$. Write $h^{-1}(x_1, x_2) = (x_1, y_2)$ and $h^{-1}(x_1, x_2') = (x_1, y_2')$ where $y_2, y_2' \in E_2$ are such that

$$y_2 - d(x_1, y_2) = x_2$$
$$y_2' - \tilde{d}(x_1, y_2') = x_2'.$$

Then we have that

$$||h^{-1}(x_1, x_2) - h^{-1}(x_1, x_2')|| = ||y_2 - y_2'|| \leq ||x_2 - x_2'|| + ||d(x_1, y_2) - d(x_1, y_2')|| \leq ||x_2 - x_2'|| + \frac{1}{2}||y_2 - y_2'||,$$

which implies that $||y_2 - y_2'|| \leq 2||x_2 - x_2'|| = \eta$, and the proof is complete. □

3.3.2 An extracting scheme tailored for closed subsets of a subspace of infinite codimension

Our next step is to establish the following.

Theorem 3.17. *Let E_1 and E_2 be Banach spaces such that E_2 is infinite-dimensional and admits a (not necessarily equivalent) C^p smooth norm, where $p \in \mathbb{N} \cup \{\infty\}$. Define $E = E_1 \times E_2$ and, for $i = 1, 2$, write $\pi_i : E \to E_i$ for the natural projections, that is, $\pi_i(x_1, x_2) = x_i$ for $(x_1, x_2) \in E$. Let W_1 be an open subset of E_1, and $\psi : W_1 \to [0, \infty)$ be a continuous function such that ψ is of class C^p on $\psi^{-1}(0, \infty)$. Denote $K = \psi^{-1}(0) \times \{0\}$. Then, there exists a C^p diffeomorphism h from $(W_1 \times E_2) \setminus K$ onto $W_1 \times E_2$ which satisfies $\pi_1 \circ h = h$ and is the identity off of a certain open subset U of $W_1 \times E_2$.*

Specifically, the set U is defined as follows

$$U := \{x = (x_1, x_2) \in W_1 \times E_2 \, : \, S(\psi(x_1), \omega(x_2)) < 1\},$$

where S is a certain C^∞ norm on $\mathbb{R}^2$ and $\omega : E_2 \to [0, \infty)$ is a certain (not necessarily symmetric) subadditive and positive-homogeneous functional of class C^p on $E_2 \setminus \{0\}$; see Lemmas 3.19 and 3.22 for precise definitions.

We will need to use the following three auxiliary results from [10, 14].

Lemma 3.18. *Let $F : (0, \infty) \longrightarrow [0, \infty)$ be a continuous function such that, for every $\beta \geq \alpha > 0$,*

$$F(\beta) - F(\alpha) \leq \frac{1}{2}(\beta - \alpha), \quad \textit{and} \quad \limsup_{t \to 0^+} F(t) > 0.$$

Then there exists a unique $\alpha > 0$ such that $F(\alpha) = \alpha$.

Proof. We follow [10, Lemma 2]. Note that

$$\lim_{\beta \to \infty} [F(\beta) - \beta] \leq \lim_{\beta \to \infty} \left[F(1) + \frac{1}{2}(\beta - 1) - \beta \right] = -\infty,$$

while

$$\limsup_{\beta \to 0^+} [F(\beta) - \beta] > 0.$$

Applying Bolzano's theorem there exists $\alpha > 0$ such that $F(\alpha) = \alpha$. Finally, observing that the function $\beta \to F(\beta) - \beta$ is strictly decreasing, this yields the uniqueness of α. □

It is not known whether every infinite-dimensional Banach space with a C^1 equivalent norm possesses a C^1 smooth non-complete norm.[2] The following lemma shows that for every Banach space with a C^p smooth norm there exists a kind of C^p *asymmetric non-complete subadditive functional* which successfully replaces the smooth non-complete norm in Bessaga's technique [34] for extracting points.

Lemma 3.19. *Let $(E_2, \|\cdot\|)$ be an infinite-dimensional Banach space which admits a (not necessarily equivalent) C^p smooth norm, where $p \in \mathbb{N} \cup \{\infty\}$. Then there exists a continuous function $\omega : E_2 \longrightarrow [0, \infty)$ which is C^p smooth on $E_2 \setminus \{0\}$ and satisfies the following properties:*

(1) $\omega(x + y) \leq \omega(x) + \omega(y)$, *and, consequently,* $\omega(x) - \omega(y) \leq \omega(x - y)$, *for every* $x, y \in E_2$;

(2) $\omega(rx) = r\omega(x)$ *for every* $x \in E_2$, *and* $r \geq 0$;

(3) $\omega(x) = 0$ *if and only if* $x = 0$;

(4) $\omega(\sum_{k=1}^{\infty} z_k) \leq \sum_{k=1}^{\infty} \omega(z_k)$ *for every convergent series* $\sum_{k=1}^{\infty} z_k$ *in* $(E_2, \|\cdot\|)$; *and*

(5) *for every* $\varepsilon > 0$, *there exists a sequence of vectors* $(y_k) \subset E_2$ *such that*

$$\omega(y_k) \leq \frac{\varepsilon}{4^{k+1}}, \quad ;$$

for every $k \in \mathbb{N}$, *and*

$$\liminf_{n\to\infty} \omega(y - \sum_{j=1}^{n} y_j) > 0$$

for every $y \in E_2$.

Notice that ω need not be a norm in E_2, as in general we have $\omega(x) \neq \omega(-x)$.

Proof. The reader can find a detailed argument in [14, Lemma 2.3]. Here we sketch some ideas. Let us called by ρ the C^p smooth norm of E_2. Three cases are considered.

- Case 1: The norm ρ is complete and E_2 is not reflexive.

  By the open mapping theorem ρ is a C^p smooth equivalent norm in E_2 and we can therefore assume $||\cdot|| = \rho(\cdot)$. Since E_2 is not reflexive, using James' theorem [103] there exists a continuous linear functional $T : E_2 \to \mathbb{R}$ which does not attain its norm and $||T|| = 1$. We define $\omega : E_2 \to [0, \infty)$ by $\omega(x) = ||x|| - T(x)$. Properties $(1) - (4)$ are easy. For property (5), given $\varepsilon > 0$, since $||T|| = \sup\{T(x) : ||x|| = 1\} = 1$, there exists a sequence of linearly independent vectors (y_k) with $||y_k|| = 1$ and $\omega(y_k) = ||y_k|| - T(y_k) \leq \frac{\varepsilon}{4^{k+1}}$ for every $k \in \mathbb{N}$. One can check that this sequence satisfies (5).

- Case 2: The norm ρ is not complete.

  Define $\omega = \rho$ and we have properties $(1) - (4)$ straightforwardly. Since ω is not complete, for every $\varepsilon > 0$ we can find a sequence (y_k) of linearly independent vectors such that $\omega(y_k) \leq \frac{\varepsilon}{4^{k+1}}$ for each k, and a point $\hat{y}$ in the completion of (E_2, ω), denoted by $(\hat{E_2}, \hat{\omega})$, such that $\hat{y} \notin E_2$ and $\lim_{n\to\infty} \hat{\omega}(\hat{y} - \sum_{k=1}^{n} y_k) = 0$. This such a choice of a sequence makes property (5) hold as well.

- Case 3: The norm ρ is complete and E_2 is reflexive.

  This case is reduced to Case 2. Reflexive spaces are weakly compact generated WCG spaces and by [60, p. 246] they can be linearly injected into some $c_0(\Gamma)$. And now, applying [63, Proposition 5.1] we get that E_2 has a non-complete C^∞ smooth norm ω.

□

[2]For C^k with $k \geq 2$ in place of C^1, the answer to this question is positive; see [57].

Using the properties of the functional ω we can construct an *extracting curve* as follows.

Lemma 3.20. *Let $(E_2, \|\cdot\|)$ be a Banach space, and let ω be a functional satisfying conditions (1), (2), and (5) of Lemma 3.19. Then there exists a C^∞ curve $\gamma : (0,\infty) \longrightarrow E_2$ such that*

(1) $\omega(\gamma(\alpha)-\gamma(\beta)) \le \frac{1}{2}(\beta-\alpha)$ *if* $\beta \ge \alpha > 0$;

(2) $\limsup_{t\to 0^+} \omega(y-\gamma(t)) > 0$ *for every* $y \in E_2$; *and*

(3) $\gamma(t) = 0$ *if* $t \ge 1$.

Proof. Let $\theta : [0,\infty) \longrightarrow [0,1]$ be a non-increasing C^∞ function such that $\theta = 1$ on $[0,1/2]$, $\theta = 0$ on $[1,\infty)$ and $\sup\{|\theta'(t)| : t \in [0,\infty)\} \le 4$. Let us choose a sequence of vectors $(y_k) \subset E_2$ which satisfies condition (5) of Lemma 3.19 for $\varepsilon = 1$, and define $\gamma : (0,\infty) \longrightarrow E_2$ by the following formula

$$\gamma(t) = \sum_{k=1}^{\infty} \theta(2^{k-1}t) y_k.$$

It is not difficult to check that this curve satisfies the properties of the statement. See [14, Lemma 2.5] for details. □

We will also need a technical tool (see, for instance, [26, Lemmas 2.27 and 2.28]) that allows us to obtain, on the product space $E_1 \times E_2$, a norm which preserves the smoothness properties that the corresponding norms of the factors may have. Notice that the natural formula $\left(\|x_1\|_1^2 + \|x_2\|_2^2\right)^{1/2}$ defines a C^1 norm in $E_1 \times E_2 \setminus \{0\}$ but, in general, this norm will not be C^2 on this set, even if $\|\cdot\|_1$ and $\|\cdot\|_2$ are C^∞ on $E_1 \setminus \{0\}$ and $E_2 \setminus \{0\}$, respectively, because the function $x_2 \mapsto \|x_2\|_2^2$ may not be C^2 smooth on all of E_2 even though it is C^∞ smooth on $E_2 \setminus \{0\}$. As a matter of fact, it is not difficult to show that, for every Banach space $(E, \|\cdot\|)$, if $\|\cdot\|^2$ is twice Fréchet differentiable at 0, then E is isomorphic to a Hilbert space; see, for instance [72, Exercise 10.4, pp. 475–476].

Definition 3.21. *We will say that a subset $\mathcal{S}$ of the plane $\mathbb{R}^2$ is a smooth square provided that:*

(i) $\mathcal{S} \subset \mathbb{R}^2$ *is a bounded, symmetric convex body with* $0 \in$ *int*$(\mathcal{S})$, *and whose boundary* $\partial\mathcal{S}$ *is* C^∞ *smooth.*

(ii) $(x,y) \in \partial\mathcal{S} \Leftrightarrow (\epsilon_1 x, \epsilon_2 y) \in \partial\mathcal{S}$ *for each couple* $(\epsilon_1, \epsilon_2) \in \{-1,1\}^2$ *(that is, $\mathcal{S}$ is symmetric about the coordinate axes).*

(iii) $[-\frac{1}{2}, \frac{1}{2}] \times \{-1,1\} \cup \{-1,1\} \times [-\frac{1}{2}, \frac{1}{2}] \subset \partial\mathcal{S}$.

(iv) $\mathcal{S} \subset [-1,1] \times [-1,1]$.

Of course, it is elementary to produce smooth squares in $\mathbb{R}^2$.

The following lemma enumerates the essential properties of a smooth square. Recall that the Minkowski functional of a convex body $\mathcal{A}$ such that $0 \in \text{int}(\mathcal{A})$ is defined by

$$\mu_{\mathcal{A}}(x) = \inf\{t > 0 : \frac{1}{t}x \in \mathcal{A}\}.$$

Lemma 3.22. *Let $\mathcal{S} \subset \mathbb{R}^2$ be a smooth square. Then its Minkowski functional $\mu_{\mathcal{S}} : \mathbb{R}^2 \to \mathbb{R}$ is a C^∞ smooth norm on $\mathbb{R}^2$ such that, for every $(x,y) \in \mathbb{R}^2$, we have*

(1) $\mu_{\mathcal{S}}(x,y) \le |x| + |y| \le 2\mu_{\mathcal{S}}(x,y)$;

(2) $\max(|x|,|y|) \le \mu_{\mathcal{S}}(x,y) \le 2\max(|x|,|y|)$;

(3) $\mu_{\mathcal{S}}(0,y) = |y|$, $\mu_{\mathcal{S}}(x,0) = |x|$;

(4) *For every* $(x_0, y_0) \in \mathbb{R}^2 \setminus \{(0,0)\}$, *there exists* $\sigma > 0$ *so that*

$$\mu_{\mathcal{S}}(x,y) = |x| \ \text{ if } \ \max(|x - x_0|, |y|) \leq \sigma \quad \text{and} \quad \mu_{\mathcal{S}}(x,y) = |y| \ \text{ if } \ \max(|x|, |y - y_0|) \leq \sigma.$$

(5) *The functions* $(0,\infty) \ni t \to \mu_{\mathcal{S}}(x, ty)$ *and* $(0,\infty) \ni t \to \mu_{\mathcal{S}}(tx, y)$ *are both nondecreasing.*

Note that property (4) (which is related to properties (iii) and (iv) of Definition 3.21) means that every sphere of $(\mathbb{R}^2, \mu_{\mathcal{S}})$ centered at the origin, which coincides with $\lambda(\partial\mathcal{S})$ for some $\lambda > 0$, is orthogonal to the coordinate axes and is locally flat on a neighbourhood of the intersection of $\lambda(\partial\mathcal{S})$ with the lines $\{x = 0\} \cup \{y = 0\}$. By using this property it is easy to show that, for any couple of Banach spaces $(E_1, \|\cdot\|_1)$ and $(E_2, \|\cdot\|_2)$ with C^p smooth norms, the expression

$$\mu_{\mathcal{S}}\left(\|x_1\|_1, \|x_2\|_2\right)$$

defines an equivalent norm of class C^p in $E_1 \times E_2$.

Proof. Properties $(1) - (3)$ and (5) are easy to show. Let us prove (4). Assume for instance that $x_0 \neq 0$, and set $\sigma = |x_0|/4$. If $\max(|x - x_0|, |y|) \leq \sigma$, then we have

$$\frac{|y|}{|x|} \leq \frac{\sigma}{|x_0| - \sigma} = \frac{|x_0|/4}{|x_0| - |x_0|/4} = \frac{1}{3} < \frac{1}{2},$$

hence

$$\left(\frac{x}{|x|}, \frac{y}{|x|}\right) \in \{-1, 1\} \times \left[-\frac{1}{2}, \frac{1}{2}\right] \subset \partial\mathcal{S},$$

and it follows that $\mu_{\mathcal{S}}(x,y) = |x|$. □

The lemma below shows how, with the help of a smooth square, we can combine the given C^p smooth function $\psi : W_1 \to [0,\infty)$ together with the C^p smooth functional $\omega : E_2 \to [0,\infty)$ obtained in Lemma 3.19, in order to obtain a C^p smooth function on $W_1 \times E_2$ which behaves more or less like $\psi(x_1) + \omega(x_2)$ (or, equivalently, like $\left(\psi(x_1)^2 + \omega(x_2)^2\right)^{1/2}$).

Lemma 3.23. *Let* $E = E_1 \times E_2$ *be a Banach space,* $\rho_1 : E_1 \to [0,\infty)$ *and* $\rho_2 : E_2 \to [0,\infty)$ *continuous functions which are of class* C^p *on* $E_1 \setminus \rho_1^{-1}(0)$ *and* $E_2 \setminus \rho_1^{-1}(0)$, *respectively. Then, for any smooth square* $\mathcal{S}$ *of* $\mathbb{R}^2$, *the function* $\rho : E_1 \times E_2 \to [0,\infty)$ *defined by*

$$\rho(x) = \rho(x_1, x_2) = \mu_{\mathcal{S}}(\rho_1(x_1), \rho_2(x_2)), \;\; x = (x_1, x_2) \in E_1 \times E_2,$$

is continuous on E *and of class* C^p *on* $E \setminus \left(\rho_1^{-1}(0) \times \rho_2^{-1}(0)\right)$.
The same is true if we replace E_1 *with an open subset* W_1 *of* E_1.

Proof. It is clear that ρ is continuous on E, and that it is C^p smooth on $\{(x_1, x_2) \in E : \rho_1(x_1) \neq 0 \neq \rho_2(x_2)\}$. Let us see that ρ is also C^p smooth on a neighbourhood of the set

$$\left(\{(x_1,x_2) \in E : \rho_1(x_1) = 0\} \cup \{(x_1,x_2) \in E : \rho_2(x_2) = 0\}\right) \setminus \left(\rho_1^{-1}(0) \times \rho_2^{-1}(0)\right).$$

Suppose for instance that $\rho_1(x_1) \neq 0 = \rho_2(x_2)$. Then, by continuity of ρ_1, ρ_2 and by property (4) of Lemma 3.22, there exist a neighbourhood U of the point (x_1, x_2) such that $U \subset \{(y_1, y_2) \in E : \rho_1(y_1) \neq 0\}$ and

$$\rho(y_1, y_2) = \rho_1(y_1)$$

for all $(y_1, y_2) \in U$. It follows that ρ is of class C^p on U. The case $\rho_1(x_1) = 0 \neq \rho_2(x_2)$ can be treated similarly. □

Now we are ready to prove Theorem 3.17.

Proof of Theorem 3.17. From now on we will fix a smooth square $\mathcal{S}$ on $\mathbb{R}^2$, and we will denote

$$S = \mu_{\mathcal{S}}.$$

Thus, by Lemma 3.23 applied to $\rho_1 = \psi$ and $\rho_2 = \omega$ (recall that ω was constructed in Lemma 3.19), the function

$$\rho(x_1, x_2) := S(\psi(x_1), \omega(x_2))$$

is continuous on $W_1 \times E_2$ and of class C^p on $(W_1 \times E_2) \setminus \big(\psi^{-1}(0) \times \{0\}\big)$.
Let us define $h : (W_1 \times E_2) \setminus K \to E$ by

$$h(x_1, x_2) = (x_1,\, x_2 + \gamma \circ \rho(x_1, x_2)) = (x_1,\, x_2 + \gamma\left(S(\psi(x_1), \omega(x_2))\right)), \;\; (x_1, x_2) \in (W_1 \times E_2) \setminus K,$$

where γ is provided by Lemma 3.20. Note that

$$K = \psi^{-1}(0) \times \{0\} = \rho^{-1}(0).$$

Let (y_1, y_2) be an arbitrary point of $W_1 \times E_2$, and let $F_{y_1,y_2} : (0, \infty) \longrightarrow [0, \infty)$ be defined by

$$F_{y_1,y_2}(\alpha) = \rho(y_1, y_2 - \gamma(\alpha)) = S\left(\psi(y_1), \omega(y_2 - \gamma(\alpha))\right) \tag{3.3.6}$$

for $\alpha > 0$. Let us see that $F_{y_1,y_2}(\alpha)$ satisfies the conditions of Lemma 3.18. As for the first condition, we consider two cases: if $\omega(y_2 - \gamma(\beta)) \leq \omega(y_2 - \gamma(\alpha))$ then, since the function $(0, \infty) \ni t \mapsto S(\psi(y_1), t)$ is increasing (see condition (5) of Lemma 3.22), we have that

$$S(\psi(y_1), \omega(y_2 - \gamma(\beta))) \leq S(\psi(y_1), \omega(y_2 - \gamma(\alpha))),$$

and therefore

$$F_{y_1,y_2}(\beta) - F_{y_1,y_2}(\alpha) \leq 0 \leq \frac{1}{2}(\beta - \alpha)$$

trivially for all $\beta \geq \alpha > 0$. Otherwise, we have $\omega(y_2 - \gamma(\beta)) - \omega(y_2 - \gamma(\alpha)) > 0$, and therefore, using the fact that S is a norm in $\mathbb{R}^2$, condition (3) of Lemma 3.22, the properties of the functional ω, and condition (1) of Lemma 3.20, we obtain

$$\begin{aligned} &S(\psi(y_1), \omega(y_2 - \gamma(\beta))) - S(\psi(y_1), \omega(y_2 - \gamma(\alpha))) \leq S\left(0, \omega(y_2 - \gamma(\beta)) - \omega(y_2 - \gamma(\alpha))\right) \\ &= \omega(y_2 - \gamma(\beta)) - \omega(y_2 - \gamma(\alpha)) \leq \omega\big(y_2 - \gamma(\beta) - (y_2 - \gamma(\alpha))\big) = \omega(\gamma(\alpha) - \gamma(\beta)) \leq \frac{1}{2}(\beta - \alpha) \end{aligned}$$

for every $\beta \geq \alpha > 0$. In either case we have that

$$F_{y_1,y_2}(\beta) - F_{y_1,y_2}(\alpha) \leq \frac{1}{2}(\beta - \alpha) \tag{3.3.7}$$

for all $\beta \geq \alpha > 0$.
On the other hand, by condition (2) of Lemma 3.20 we know that

$$\limsup_{\alpha \to 0^+} \omega(y_2 - \gamma(\alpha)) > 0,$$

and therefore, by condition (2) of Lemma 3.22, we have

$$\limsup_{\alpha \to 0^+} F_{y_1,y_2}(\alpha) = \limsup_{\alpha \to 0^+} S(\psi(y_1), \omega(y_2 - \gamma(\alpha))) \geq \limsup_{\alpha \to 0^+} \omega(y_2 - \gamma(\alpha)) > 0,$$

so that F_{y_1,y_2} also satisfies the second condition of Lemma 3.18.
Then, applying Lemma 3.18, we deduce that the equation $F_{y_1,y_2}(\alpha) = \alpha$ has a unique solution. This means that, for each $(y_1, y_2) \in W_1 \times E_2$, a number $\alpha(y_1, y_2) > 0$ with the property

$$S\left(\psi(y_1), \omega\big(y_2 - \gamma(\alpha(y_1, y_2))\big)\right) = \rho\left(y_1, y_2 - \gamma(\alpha(y_1, y_2))\right) = \alpha(y_1, y_2), \tag{3.3.8}$$

is uniquely determined.

Let us see why these facts imply that h is a C^p diffeomorphism from $W_1 \times E_2 \setminus K$ onto $W_1 \times E_2$. Assume first that $h(x_1, x_2) = (y_1, y_2) = h(z_1, z_2)$, that is to say $x_1 = y_1 = z_1$, and

$$x_2 + \gamma(\rho(y_1, x_2)) = y_2 = z_2 + \gamma(\rho(y_1, z_2)), \tag{3.3.9}$$

or equivalently

$$(y_1, x_2) = (y_1, y_2 - \gamma(\rho(y_1, x_2)) \text{ and } (y_1, z_2) = (y_1, y_2 - \gamma(\rho(y_1, z_2)).$$

Applying ρ to all sides of the above equations and using (3.3.6), we obtain

$$\rho(y_1, x_2) = \rho(y_1, y_2 - \gamma(\rho(y_1, x_2))) = F(\rho(y_1, x_2))$$

and

$$\rho(y_1, z_2) = \rho(y_1, y_2 - \gamma(\rho(y_1, z_2))) = F(\rho(y_1, z_2)).$$

It follows that both $\rho(y_1, x_2)$ and $\rho(y_1, z_2)$ are fixed points of F_{y_1,y_2}. By the uniqueness of the fixed point, we conclude that

$$\alpha(y_1, y_2) = \rho(y_1, x_2) = \rho(y_1, z_2).$$

Now applying (3.3.9), we have

$$x_2 = z_2 \text{ and } x_2 = y_2 - \gamma(\alpha(y_1, y_2)).$$

This shows that h is one to one, and also that, given $(y_1, y_2) \in W_1 \times E_2$ we have

$$h(y_1, y_2 - \gamma(\alpha(y_1, y_2))) = (y_1, y_2).$$

Hence h is also onto, and $h^{-1} : W_1 \times E_2 \to W_1 \times E_2 \setminus K$ is given by

$$h^{-1}(y_1, y_2) = (y_1, y_2 - \gamma(\alpha(y_1, y_2))).$$

It is clear that h is of class C^p. In order to see that h^{-1} is C^p as well, let us define $\Phi : W_1 \times E_2 \times (0, \infty) \longrightarrow \mathbb{R}$ by

$$\Phi(y_1, y_2, \alpha) = \alpha - S(\psi(y_1), \omega(y_2 - \gamma(\alpha))) = \alpha - \rho(y_1, y_2 - \gamma(\alpha)).$$

On the one hand, according to (3.3.8) and the fact that S is a norm in $\mathbb{R}^2$, we have

$$(\psi(y_1), \omega(y_2 - \gamma(\alpha(y_1, y_2))) \neq (0, 0)$$

for every $(y_1, y_2) \in W_1 \times E_2$. Since S is C^∞ smooth away from $(0, 0)$, this implies that Φ is C^p smooth on a neighbourhood of every point $(y_1, y_2, \alpha(y_1, y_2))$ in $W_1 \times E_2 \times (0, \infty)$. On the other hand, we know that $F_{y_1,y_2}(\beta) - F_{y_1,y_2}(\alpha) \leq \frac{1}{2}(\beta - \alpha)$ for $\beta \geq \alpha > 0$, which implies that $F'_{y_1,y_2}(\alpha) \leq \frac{1}{2}$ for every α in a neighbourhood of $\alpha(y_1, y_2)$, and therefore

$$\frac{\partial \Phi(y_1, y_2, \alpha)}{\partial \alpha} = 1 - F'_{y_1,y_2}(\alpha) \geq 1 - 1/2 > 0.$$

Hence, by the implicit function theorem, the mapping $(y_1, y_2) \to \alpha(y_1, y_2)$ is of class C^p on $W_1 \times E_2$, and, since γ is C^p smooth, so is h^{-1}.

Finally, it is obvious that $\pi_1 \circ h = \pi_1$, and the fact that $\gamma(t) = 0$ whenever $t \geq 1$ (see property (3) of Lemma 3.20) implies that h is the identity off of the set $\{x = (x_1, x_2) \in W_1 \times E_2 : S(\psi(x_1), \omega(x_2)) < 1\}$. □

3.3.3 Extracting pieces of continuous graphs of infinite codimension

We will finally give in this subsection the proof of Theorem 3.8. Let us recall its statement.

Theorem 3.8. *Let $E = E_1 \times E_2$ be a product of Banach spaces such that E_1 admits C^p smooth partitions of unity and E_2 admits a C^p (not necessarily equivalent) norm. Assume that X_1 is a closed subset of E_1, that $f : E_1 \to E_2$ is a continuous mapping, and that E_2 is infinite-dimensional. Define*

$$X = \{(x_1, x_2) \in E_1 \times E_2 \,:\, x_1 \in X_1, x_2 = f(x_1)\}.$$

Let U be an open subset of E and $\varepsilon > 0$. Then there exists a C^p diffeomorphism g from $E \setminus X$ onto $E \setminus (X \setminus U)$ such that g is the identity on $(E \setminus U) \setminus X$ and moves no point more than ε.

Proof. We may of course assume $U \cap X \neq \emptyset$ (as otherwise the result holds trivially with g equal to the identity map).

Claim 3.23. *It is sufficient to prove the result for $f = 0$ and such that the extracting diffeomorphism preserves the first coordinate, that is $g(x_1, x_2) = (x_1, \pi_2(g(x_1, x_2)))$.*

Proof. Let h and φ be homeomorphisms given by Theorem 3.10 such that $||\varphi^{-1}(x) - h^{-1}(x)|| \leq \frac{\varepsilon}{2}$ for every $x \in E$. By the uniform continuity of $h^{-1}(x_1, x_2)$ with respect to the second variable $x_2 \in E_2$, we may choose $\delta > 0$ such that if $||(x_1, x_2) - (x_1, x_2')|| \leq \delta$ then

$$||h^{-1}(x_1, x_2) - h^{-1}(x_1, x_2')|| \leq \frac{\varepsilon}{2}.$$

Assuming the result is true for $f = 0$ we can find a C^p diffeomorphism $g : E \setminus (X_1 \times \{0\}) \to E \setminus ((X_1 \times \{0\}) \setminus h(U))$ such that g is the identity on $(E \setminus h(U)) \setminus (X_1 \times \{0\})$, moves no point more than δ and preserves the first coordinate. Then the composition

$$\varphi^{-1} \circ g \circ h : E \setminus X \to (E \setminus (X \setminus U)$$

defines a C^p diffeomorphism with the required properties. Observe that

$$||g(h(x)) - h(x)|| = ||(x_1, \pi_2(g(h(x)))) - (x_1, \pi_2(h(x)))|| \leq \delta,$$

hence

$$||\varphi^{-1}(g(h(x))) - x|| \leq ||\varphi^{-1}(g(h(x))) - h^{-1}(g(h(x)))|| + ||h^{-1}(g(h(x))) - x|| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for every $x \in E \setminus X$. □

So it will be enough to see that if X_1 is a closed subset of E_1 and W is an open subset of E such that $W \cap X_1 \times \{0\} \neq \emptyset$ then there exists a C^p diffeomorphism g from $E \setminus (X_1 \times \{0\})$ onto $E \setminus ((X_1 \times \{0\}) \setminus W)$ such that g is the identity on $(E \setminus W) \setminus (X_1 \times \{0\})$ and moves no point more than δ.

To this end we next construct some auxiliary functions following Renz's strategy [134, pp. 54–59]. In what follows ω will denote the smooth asymmetric subadditive functional on E_2 given by Lemma 3.19.

Lemma 3.24. *There exists a continuous function $\varphi : E_1 \to [0, \frac{\delta}{2}]$ such that:*

(1) *φ is of class C^p on $E_1 \setminus \partial\varphi^{-1}(0)$.*

(2) *$W \cap (X_1 \times \{0\}) \subset \{(x_1, x_2) \in E \,:\, \|x_2\| < \varphi(x_1)\} \subset W$.*

Proof. Let $\pi_1 : E \to E_1$ denote the canonical projection defined by $\pi_1(x_1, x_2) = x_1$. The set

$$W_1 := \pi_1 \left(W \cap (E_1 \times \{0\})\right)$$

is open in E_1, and the function $G : E_1 \to [0, \infty)$ defined by

$$G(x_1) = \min\{\frac{\delta}{2}, \text{dist}\left((x_1, 0), E \setminus W\right)\}$$

is continuous and satisfies that $G > 0$ on W_1 and $G = 0$ on $\pi_1\left((E_1 \times \{0\}) \setminus W\right)$. Since E_1 has C^p smooth partitions of unity and G is continuous and strictly positive on W_1, we can find a C^p smooth function F on W_1 such that

$$0 < \frac{1}{4}G(x_1) < F(x_1) < \frac{1}{2}G(x_1)$$

for every $x_1 \in W_1$. Now let us define $\varphi : E_1 \to [0, 1]$ by

$$\varphi(x_1) = \begin{cases} F(x_1) & \text{if } x_1 \in W_1 \\ 0 & \text{if } x_1 \in E_1 \setminus W_1. \end{cases}$$

It is immediately seen that φ is continuous, and of course φ is of class C^p on $E_1 \setminus \partial\varphi^{-1}(0) = W_1 \cup \text{int}(E_1 \setminus W_1)$. Since $\varphi(x_1) = F(x_1) > 0$ for all $x_1 \in W_1$, it is obvious that

$$W \cap (X_1 \times \{0\}) \subset W_1 \times \{0\} \subset \{(x_1, x_2) \in E : \|x_2\| < \varphi(x_1)\}.$$

On the other hand, if $\|x_2\| < \varphi(x_1)$ then observe first that $x_1 \in W_1$ (because if $\varphi(x_1) = 0$ the inequality is impossible). We then must have $(x_1, x_2) \in W$, as otherwise we would get

$$\begin{aligned} &\text{dist}\left((x_1, 0), E \setminus W\right) \leq \text{dist}\left((x_1, 0), (x_1, x_2)\right) + \text{dist}\left((x_1, x_2), E \setminus W\right) = \\ &\|x_2\| + 0 = \|x_2\| < \varphi(x_1) < \frac{1}{2}G(x_1) \leq \frac{1}{2}\text{dist}\left((x_1, 0), E \setminus W\right), \end{aligned}$$

which is absurd. This shows that $\{(x_1, x_2) \in E : \|x_2\| < \varphi(x_1)\} \subset W$ and concludes the proof of the lemma. □

We will also need to use a diffeomorphism h_2 of E_2 onto itself which carries the unit ball of E_2 onto the convex body $\{x_2 \in E_2 : \omega(x) \leq 1\}$ and such that $h_2(0) = 0$. The existence of h_2 is ensured by the next lemma, that appears a few lines below. Recall that a convex body is a closed and convex set with nonempty interior. The fact that the set $\{x_2 \in E_2 : \omega(x) \leq 1\}$ is a convex body may not seem clear. The convexity of the set is by properties (1) and (2) of Lemma 3.19. We have $\omega(0) = 0$, so the origin belongs to the set and using the continuity of $\omega : E_2 \to [0, \infty)$ there exists some $r > 0$ such that for every $x_2 \in B(0, r)$, $||\omega(x_2)|| \leq 1$. This shows that the set has not empty interior. Finally the set is closed by using again the continuity of ω.

We say that a convex body U which contains 0 as an interior point is *radially bounded* provided that for every $x \in U$ the set $\{tx : t \in [0, \infty)\} \cap U$ is bounded. And we say that an open set U of a Banach space X is C^p smooth if its boundary ∂U is a C^p smooth one-codimensional submanifold of X.

Lemma 3.25. *Let X be a Banach space, and let U_1, U_2 be radially bounded, C^p smooth convex such that the origin is an interior point of both U_1 and U_2. Then there exists a C^p diffeomorphism $g : X \longrightarrow X$ such that $g(U_1) = U_2$, $g(0) = 0$, and $g(\partial U_1) = \partial U_2$.*

Proof. If U and V are C^p smooth, radially bounded convex bodies such that the origin is an interior point of both U and V, and we additionally assume that $U \subseteq V$, such a diffeomorphism can be constructed as follows: let $\theta(t)$ be a non-decreasing real function of class C^∞ defined for $t > 0$, such that $\theta(t) = 0$ for $t \leq 1/2$ and $\theta(t) = 1$ for $t \geq 1$, and define

$$g(x) = \left(\theta(\mu_U(x))\frac{\mu_U(x)}{\mu_V(x)} + 1 - \theta(\mu_U(x))\right) x$$

for $x \neq 0$, and $g(0) = 0$. Here μ_A denotes the Minkowski functional of A.
In the general case, let $U = \{x \in X \ : \ \mu_{U_1}(x) + \mu_{U_2}(x) \leq 1\}$, then $U \subseteq U_j$, for $j = 1, 2$, and there exist diffeomorphisms $g_1, g_2 : X \to X$ such that $g_j(U) = U_j$ and $g_j(\partial U) = \partial U_j$, $j = 1, 2$. Then $g = g_2 \circ g_1^{-1}$ does the job. See [62] for details. □

The following lemma is an immediate consequence of the existence of partitions of unity in E_1.

Lemma 3.26. *Suppose that E_1 is a Banach space with C^p smooth partitions of unity, and let X_1 be a closed subset of E_1. Then there exists a continuous function $\eta : E_1 \to [0, \infty)$ such that:*

(1) $X_1 = \eta^{-1}(0)$*;*

(2) *η is of class C^p on $E_1 \setminus X_1$.*

Proof. Consider the continuous function $f : E_1 \to [0, \infty)$, $f(x) = \operatorname{dist}(x, X_1)$, and also $\varepsilon : E_1 \setminus X_1 \to (0, \infty)$, $\varepsilon(x) = \frac{1}{2} \operatorname{dist}(x, X_1)$. Recalling that the existence of C^p partitions of unity implies the uniform approximation by C^p functions, explained at the beginning of the proof of Lemma 3.11, there exists $\eta : E_1 \setminus X_1 \to (0, \infty)$ of class C^p such that $|g(x) - f(x)| \leq \varepsilon(x)$ for every $x \in E_1 \setminus X_1$. Now extend η to the whole space E_1 by letting it be equal to f in X_1. One can check that η is continuous, $\eta^{-1}(0) = X_1$ and $\eta|_{E_1 \setminus X_1}$ is of class C^p.

□

We are ready to proceed with the proof of Theorem 3.8. Let η be a function as in the statement of Lemma 3.26, and pick a diffeomorphism $h_2 : E_2 \to E_2$ such that $h_2(0) = 0$ and

$$h_2 \left(\{x_2 \in E_2 : \|x_2\| \leq 1\} \right) = \{x_2 \in E_2 \ : \ \omega(x_2) \leq 1\}.$$

Let us define

$$A := \varphi^{-1} \left((0, 1] \right) \times E_2 = W_1 \times E_2,$$

and

$$\Phi(x_1, x_2) = \left(x_1, h_2 \left(\frac{1}{\varphi(x_1)} x_2 \right) \right), \quad (x_1, x_2) \in A.$$

It is clear that $\Phi : A \to A$ is a C^p diffeomorphism, with inverse

$$\Phi^{-1}(y_1, y_2) = \left(y_1, \varphi(y_1) h_2^{-1}(y_2) \right),$$

and also that

$$\Phi \left((X_1 \times \{0\}) \cap A \right) = (X_1 \times \{0\}) \cap A.$$

Next let us define $\psi : \varphi^{-1} \left((0, 1] \right) = \pi_1(A) = W_1 \to [0, \infty)$ by

$$\psi(x_1) = \frac{\eta(x_1)}{\varphi(x_1)},$$

and notice that ψ is continuous, that

$$\psi^{-1}(0) = \pi_1 \left((X_1 \times \{0\}) \cap A \right) = X_1 \cap W_1,$$

and that ψ is of class C^p outside $\psi^{-1}(0)$. By Theorem 3.17 we can find a C^p diffeomorphism H from $A \setminus (X_1 \times \{0\})$ onto A such that H is the identity outside $\{(x_1, x_2) \in A \setminus (X_1 \times \{0\}) \ : \ S(\psi(x_1), \omega(x_2)) < 1\}$, where S is a smooth square. Since $\Phi : A \to A$ is a C^p diffeomorphism which takes $(X_1 \times \{0\}) \cap A$ onto itself, we have that the composition $\Phi^{-1} \circ H \circ \Phi$ defines a C^p diffeomorphism from $A \setminus (X_1 \times \{0\})$ onto A. Now we extend this diffeomorphism outside $A \setminus (X_1 \times \{0\})$ by defining $g : E \setminus (X_1 \times \{0\}) \to E$ by

$$g(x) = \begin{cases} \Phi^{-1} \circ H \circ \Phi(x) & \text{if } x \in A \setminus (X_1 \times \{0\}) \\ x & \text{if } x \in E \setminus (A \cup (X_1 \times \{0\})). \end{cases}$$

This mapping is clearly a bijection. Thus, in order to see that g is a C^p diffeomorphism, it is enough to see that g is locally a C^p diffeomorphism. We already know this is so for all points of $E \setminus \big(\partial(A \setminus (X_1 \times \{0\})) \cup (X_1 \times \{0\})\big)$. Let us show that this is also true for every point (x_1, x_2) of $\partial(A \setminus (X_1 \times \{0\})) \setminus (X_1 \times \{0\})$. We have $\varphi(x_1) = 0$, and also either $\eta(x_1) > 0$ or $\|x_2\| > 0$. Then

$$\lim_{A \ni (y_1, y_2) \to (x_1, x_2)} \max \left\{ \frac{\eta(y_1)}{\varphi(y_1)}, \frac{\|y_2\|}{\varphi(y_1)} \right\} = \infty,$$

hence there exists a neighbourhood V of (x_1, x_2) in $E \setminus (X_1 \times \{0\})$ such that

$$\max \left\{ \frac{\eta(y_1)}{\varphi(y_1)}, \frac{\|y_2\|}{\varphi(y_1)} \right\} > 1 \quad \text{for all } (y_1, y_2) \in V \cap A.$$

By Lemma 3.22(2) it follows that

$$S\left(\psi \circ \pi_1 \circ \Phi(y), \omega \circ \pi_2 \circ \Phi(y)\right) > 1 \quad \text{for all } y \in V \cap A,$$

hence that H is the identity on $\Phi(V \cap A)$, and consequently that g is the identity on V, and in particular a C^p diffeomorphism locally at (x_1, x_2). Thus $g : E \setminus (X_1 \times \{0\}) \to E$ is a C^p diffeomorphism.
Furthermore, if $\|x_2\| \geq \varphi(x_1)$, $(x_1, x_2) \in A \setminus (X_1 \times \{0\})$, then we have $\omega(\pi_2 \circ \Phi(x_1, x_2)) \geq 1$. Hence, as above by Lemma 3.22(2), we conclude that $H(\Phi(x_1, x_2)) = \Phi(x_1, x_2)$, and it follows that $g(x_1, x_2) = (x_1, x_2)$. Thus g is the identity off of the set $\{(x_1, x_2) \in E \setminus (X_1 \times \{0\}) \,:\, \|x_2\| < \varphi(x_1)\}$. Since

$$\{(x_1, x_2) \in E \setminus (X_1 \times \{0\}) \,:\, \|x_2\| < \varphi(x_1)\} \subset (E \setminus W) \setminus (X_1 \times \{0\}),$$

g is the identity off of the set $(E \setminus W) \setminus (X_1 \times \{0\})$ as well.
Finally let us check that g does not move any point more than δ. We know that if g moves a point $(x_1, x_2) \in E \setminus (X_1 \times \{0\})$ then $||x_2|| < \varphi(x_1)$, and also that g_2 only moves the second coordinate x_2, that is $g(x_1, x_2) = (x_1, \pi_2(g(x_1, x_2)))$. Hence

$$||g(x_1, x_2) - (x_1, x_2)|| \leq ||\pi_2(g(x_1, x_2)) - x_2|| \leq ||\pi_2(g(x_1, x_2))|| + ||x_2|| \leq \varphi(x_1) + \varphi(x_1) \leq \delta.$$

The proof of Theorem 3.8 is complete. □

3.4 An abstract extractibility result

This section is devoted to proving Theorem 3.9, for whose proof we will borrow a technique of James West's [150, pp. 288–290].

Theorem 3.9. *Let E be a Banach space and X be a closed subset of E which has locally the ε-strong C^p extraction property. Let U be an open subset of E and $\mathcal{G} = \{G_r\}_{r \in \Omega}$ be an open cover of E. Then there exists a C^p diffeomorphism g from $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$.*

Firstly, the fact that we are working with a set X that has locally the strong C^p extraction property, and the requirement that our final C^p diffeomorphism must be limited by a given open cover $\mathcal{G}$ forces us to employ *good* refinements of covers of the Banach space E. In the separable case star-finite refinements provide an adequate tool to face the problem (see West [150]). Recall that a cover is said to be *star-finite* provided that each element of the cover intersects at most finitely many others. In 1947, [106, Theorem 1], Kaplan proved that in a separable metric space, every open cover admits a countable star-finite refinement. However, in the nonseparable case, getting a star-finite refinement of an open cover, in general, is not possible. We will use sigma-discrete refinements as shown in the following. Compare with [139, 104].

Lemma 3.27. *Let E be a Banach space and X be a closed subset of E which has locally the ε-strong C^p extraction property. Let $\mathcal{G} = \{G_r\}_{r\in\Omega}$ be an open cover of E, where the cardinality of the indexing set Ω is the density of E. Then there exist countable collections $\{X_i\}_{i\geq 1}$, $\{W_i\}_{i\geq 1}$, $\{V_i\}_{i\geq 1}$, such that:*

(1) $X_i \subseteq W_i \subseteq \overline{W_i} \subseteq V_i$ *for all* $i \in \mathbb{N}$ *;*

(2) $\{V_i\}_{i\geq 1}$ *and* $\{W_i\}_{i\geq 1}$ *are star-finite open covers of E;*

(3) $\{X_i\}_{i\geq 1}$ *is a cover of X by closed subsets of X;*

(4) *Each W_i and V_i admits an open discrete cover $\{W_{i,r}\}_{r\in\Omega}$ and $\{V_{i,r}\}_{r\in\Omega}$, respectively; more precisely,*

$$W_i = \bigcup_{r\in\Omega} W_{i,r} \ \text{ and } \ V_i = \bigcup_{r\in\Omega} V_{i,r},$$

$$\overline{W_{i,r}} \subseteq V_{i,r} \ \text{ for every } \ r \in \Omega,$$

and

$$\operatorname{dist}(V_{i,r}, V_{i,r'}) \geq \frac{1}{2^{i+1}} \ \text{ for every } \ r, r' \in \Omega, r \neq r';$$

(5) $\{W_{i,r}\}_{i\geq 1, r\in\Omega}$ *and* $\{V_{i,r}\}_{i\geq 1, r\in\Omega}$ *are open refinements of $\mathcal{G}$;*

(6) *Each X_i can be written as $X_i = \bigcup_{r\in\Omega} X_{i,r}$, where $X_{i,r}$ is a closed subset of X satisfying the following requirements*

$$X_{i,r} \subseteq W_{i,r} \subseteq \overline{W_{i,r}} \subseteq V_{i,r}$$

and

$$X_{i,r} \ \text{ has the } \varepsilon - \text{strong } \ C^p \ \text{ extraction property with respect to } \ V_{i,r}.$$

Proof. For each $x \in E$, let U_x be an open neighbourhood of x such that $\overline{U_x} \subseteq G$ for some $G \in \mathcal{G}$ and also satisfying that

1. if $x \in X$ then $X \cap \overline{U_x}$ has the ε-strong C^p extraction property with respect to every open set U with $X \cap \overline{U_x} \subseteq U$, and
2. if $x \notin X$ then $X \cap \overline{U_x} = \emptyset$.

Since the cardinality of Ω is the density of E, we can extract a subcover $\mathcal{U} = \{U_r : r \in \Omega\}$ from $\{U_x : x \in E\}$. Now we use a result of Rudin [139] (see also [93, p. 390]) to obtain two open refinements $\{A_{j,r}\}_{j\geq 1, r\in\Omega}$ and $\{B_{j,r}\}_{j\geq 1, r\in\Omega}$ of $\mathcal{U}$ such that

1. $A_{j,r} \subseteq B_{j,r} \subseteq U_r$ for all $j \in \mathbb{N}$ and $r \in \Omega$;
2. $\operatorname{dist}(A_{j,r}, E \setminus B_{j,r}) \geq \frac{1}{2^j}$ for all $j \in \mathbb{N}$ and $r \in \Omega$;
3. $\operatorname{dist}(B_{j,r}, B_{j,r'}) \geq \frac{1}{2^{j+1}}$ for all $j \in \mathbb{N}$ and $r, r' \in \Omega$, $r \neq r'$;
4. Letting $A_j = \bigcup_{r\in\Omega} A_{j,r}$ and $B_j = \bigcup_{j\in\Omega} B_{j,r}$ each collection $\{A_j\}_{j\geq 1}$, $\{B_j\}_{j\geq 1}$ forms a locally finite open cover of E.

Observe that $\overline{A_j} \subseteq B_j$ for every $j \in \mathbb{N}$.
For every j, there exists a sequence of open sets B_j^n, $n \geq j$, so that

$$\overline{A_j} \subset B_j^j \subset \overline{B_j^j} \subset B_j^{j+1} \subset \overline{B_j^{j+1}} \subset \cdots \subset B_j^n \subset \overline{B_j^n} \subset B_j^{n+1} \subset \cdots \subset B_j.$$

For each j, write $\mathcal{B}_j = \left\{B_j^n : n \geq j\right\}$. Clearly, $\mathcal{B} = \bigcup_{j=1}^{\infty} \mathcal{B}_j$ is an open cover of E; likewise, the family $\left\{\overline{B} \cap X : B \in \mathcal{B}\right\}$ is a closed cover of X.
Defining for each $n \in \mathbb{N}$: $Y_n := \bigcup_{j=1}^{n} B_j^n$, $H_n := Y_n \setminus \overline{Y_{n-3}}$ and $K_n := \overline{Y_n} \setminus Y_{n-1}$ (let $Y_{-2} = Y_{-1} = Y_0 = \emptyset$), we have the following properties:

- $E = \bigcup_{n=1}^{\infty} Y_n$;
- $\overline{Y_n} \subseteq Y_{n+1}$ for all $n \in \mathbb{N}$;
- $K_n \subseteq H_{n+1}$ for all $n \in \mathbb{N}$;
- $E = \bigcup_{n=1}^{\infty} K_n$;
- $H_m \cap H_n = \emptyset$ for all m, n with $|m - n| \geq 3$.

Hence, the collection

$$\bigcup_{n=1}^{\infty} \{K_n \cap \overline{B_j^n} : j = 1, \dots, n\}$$

is a closed cover of E and therefore

$$\bigcup_{n=1}^{\infty} \{H_{n+1} \cap B_j^{n+1} : j = 1, \dots, n\}$$

is an open cover of E. Both covers are countable and star-finite, and they are refinements of $\{B_j\}_{j\geq 1}$. We call the first one $\{T_i\}_{i\geq 1}$ and the second one $\{V_i\}_{i\geq 1}$, that is, for every i there corresponds a unique pair (j, n), $n \geq j$, with $T_i = K_n \cap \overline{B_j^n}$ and $V_i = H_{n+1} \cap B_j^{n+1}$. Consequently, we have $T_i \subseteq V_i$ for every $i \in \mathbb{N}$.
Now for each $i \in \mathbb{N}$ we take $j = j(i) \in \mathbb{N}$ such that $T_i \subseteq V_i \subseteq B_j$. Let us assume without loss of generality that $j(i) \leq i$. We can write

$$T_i = \bigcup_{r\in\Omega} T_i \cap B_{j,r} \quad \text{and} \quad V_i = \bigcup_{r\in\Omega} V_i \cap B_{j,r}$$

and we define $T_{i,r} = T_i \cap B_{j,r}$ and $V_{i,r} = V_i \cap B_{j,r}$ for every $i \in \mathbb{N}$ and $r \in \Omega$. Clearly we have that $T_{i,r} \subseteq V_{i,r}$ for all $i \in \mathbb{N}$ and $r \in \Omega$. Also $\operatorname{dist}(V_{i,r}, V_{i,r'}) \geq \frac{1}{2^{j+1}} \geq \frac{1}{2^{i+1}}$ for all $i \in \mathbb{N}$ and $r, r' \in \Omega$, $r \neq r'$.
Finally let us define $X_{i,r} = X \cap T_{i,r}$. Bearing in mind that $T_{i,r} \subset V_{i,r} \subset B_{j,r} \subset U_x$ for some $x \in X$ and that $T_{i,r}$ is closed, we obtain $X_{i,r} = X \cap T_{i,r} \subset X \cap \overline{U_x}$. Since $X \cap \overline{U_x}$ has the ε-strong C^p extraction property with respect to every open set U with $X \cap \overline{U_x} \subset U$, applying Lemma 3.6(1), we get that $X_{i,r}$ has the ε-strong C^p extraction property with respect to every such an open set U. Finally, applying Lemma 3.6(2), $X_{i,r}$ has the strong C^p extraction property with respect to every open set U' with $X_{i,r} \subset U'$. In particular, $X_{i,r}$ has the ε-strong C^p extraction property with respect to $V_{i,r}$.
Let us consider now for each $i \in \mathbb{N}$ and $r \in \Omega$ an open set $W_{i,r}$ with $T_{i,r} \subseteq W_{i,r} \subseteq \overline{W_{i,r}} \subseteq V_{i,r}$ and call $W_i = \bigcup_{r\in\Omega} W_{i,r}$. We still have that $\{W_{i,r}\}_{i\geq 1, r\in\Omega}$ is a refinement of $\mathcal{G}$, and that $\{W_i\}_{i\geq 1}$ is a star-finite open cover of E.
Then the collections $\{X_{i,r}\}_{i\geq 1, r\in\Omega}$, $\{W_{i,r}\}_{i\geq 1, r\in\Omega}$, $\{V_{i,r}\}_{i\geq 1, r\in\Omega}$ have the required properties. Note that each $X_{i,r}$ has the strong C^p extraction property with respect to every open set containing it. □

Lemma 3.28. *Let X_i, V_i, and $V_{i,r}$ be as in Lemma 3.27. Then, for every $i, j \in \mathbb{N}$, $X_i \cap V_j$, has the strong C^p extraction property with respect to V_j. Moreover, if h is a C^p diffeomorphism satisfying the definition of the ε-strong extraction property for these sets, then $h(V_{i,r}) \subset V_{i,r}$ and $h(V_{j,r}) \subset V_{j,r}$ for every $r \in \Omega$.*

Proof. Take $V \subseteq V_j$ an open set. Let $r \in \Omega$ and consider the open set $V_{j,r}$. For each $s \in \Omega$ the set $X_{i,s} \cap V_{j,r}$ has the ε-strong C^p extraction property with respect to the open set $V_{i,s} \cap V_{j,r}$. We have

$$X_{i,s} \cap V_{j,r} \subseteq W_{i,s} \cap V_{j,r} \subseteq \overline{W_{i,s}} \cap V_{j,r} \subseteq V_{i,s} \cap V_{j,r}.$$

There exists a C^p diffeomorphism $h_{r,s} : (V_{i,s} \cap V_{j,r}) \setminus X_{i,s} \to V_{i,s} \cap V_{j,r} \setminus (X_{i,s} \setminus (W_{i,s} \cap V))$ which is the identity on $((V_{i,s} \cap V_{j,r} \setminus (W_{i,s} \cap V)) \setminus X_{i,s}$. Outside $V_{i,s} \cap V_{j,r}$ we define $h_{r,s}$ to be the identity. Since $\overline{W_{i,s}} \subseteq V_{i,s}$ we have a well-defined C^p diffeomorphism

$$h_{r,s} : V_{j,r} \setminus X_{i,s} \to V_{j,r} \setminus (X_{i,s} \setminus (W_{i,s} \cap V))$$

which is the identity on $(V_{j,r} \setminus (W_{i,s} \cap V)) \setminus X_{i,s}$. In particular $h_{r,s}$ is the identity on $V_{i,s'} \cap V_{j,r}$ for every $s' \in \Omega$, $s \neq s'$.
Having defined $h_{r,s}$ for each $s \in \Omega$ in the way described above, we finally define

$$h_r = \bigcirc_{s\in\Omega} h_{r,s}$$

as an infinite composition of $h_{r,s}$, $s \in \Omega$. It is easy to see that h_r well-defines a C^p diffeomorphism of $V_{j,r} \setminus X_i$ onto $V_{j,r} \setminus (X_i \setminus V)$ which is the identity on $(V_{j,r} \setminus V) \setminus X_i$.
By the discreteness of the family $\{V_{j,r}\}_{r\in\Omega}$, the formula $h(x) = h_r(x)$, $x \in V_{j,r} \setminus X_i$, defines a C^p diffeomorphism of $\bigcup_{r\in\Omega} V_{j,r} \setminus X_i = V_j \setminus X_i$ onto $\bigcup_{r\in\Omega} (V_{j,r} \setminus (X_i \setminus V)) = V_j \setminus (X_i \setminus V)$ which is the identity on $\bigcup_{r\in\Omega} ((V_{j,r} \setminus V) \setminus X_i) = (V_j \setminus V) \setminus X_i$.
To end the proof observe that each $h_{r,s}$ is the identity outside $V_{i,s} \cap V_{j,r}$, so h sends $V_{i,r}$ into $V_{i,r}$ and $V_{j,r}$ into $V_{j,r}$ for every $r \in \Omega$. □

Notice in the previous two Lemmas 3.28 and 3.27 one can replace the ε-strong C^p extraction property with just the strong C^p extraction property.

The last tool that we need to introduce before going into the proof of Theorem 3.9 is the next lemma (see Statement A in West's paper [150, p. 289]).

Lemma 3.29. *Let $V_0, \dots, V_n$ be open sets of E, and $X_0, \dots, X_n$ be subsets of $V_0, \dots, V_n$. Take also an open set U. Suppose each X_i is relatively closed in $V = \bigcup_{i=0}^n V_i$ and has the ε-strong C^p extraction property with respect to V. Then for every $\varepsilon > 0$ there exists a C^p diffeomorphism of $V \setminus X_0$ onto $V \setminus (X_0 \setminus U)$ which is the identity outside $V_0 \cap U$, carries $X_i \setminus X_0$ into V_i for each $i = 1, \dots, n$, and moves no point more than ε.*

Proof. We will divide the set X_0 that we want to extract as follows. For each $j = 0, \dots, n$, let

$$Q_j = \left\{ Z : Z = \bigcap_{i=0}^{n-j} X_{p(i)} \setminus \bigcup_{i=n-j+1}^{n} X_{p(i)} \text{ for some permutation } p \text{ of } \{0,\dots,n\} \text{ carrying } 0 \text{ to } 0 \right\}$$

be a family of subsets of X_0. Let $Q = \bigcup_{j=0}^n Q_j$. The family Q is a pair-wise disjoint cover of X_0 with cardinality $\leq 2^n$. Order Q in such a manner that if $j < k$ then all elements of Q_j precede those of Q_k. We then will list the elements of Q as $Z_1, Z_2, \dots, Z_{2^n}$ bearing in mind that some Z_m's may repeat in that listing; that is, $Q = \{Z_m\}_{m=1}^{2^n}$ and $\bigcup_{i=1}^{2^n} Z_i = X_0$. Note that $Q_0 = \{Z_1\}$ and $Q_n = \{Z_{2^n}\}$, where

$$Z_1 = \bigcap_{i=0}^{n} X_i \quad \text{and} \quad Z_{2^n} = X_0 \setminus \bigcup_{i=1}^{n} X_i.$$

Likewise, for each $m = 1, \dots, 2^n$, if $Z_m \in Q_j$ and p is a permutation for which $Z_m = \bigcap_{i=0}^{n-j} X_{p(i)} \setminus \bigcup_{i=n-j+1}^{n} X_{p(i)}$, we define

$$N_m = \bigcap_{i=0}^{n-j} V_{p(i)} \subset V_0.$$

The family $\{N_m\}_{m=1}^{2^n}$ is an open cover of V_0; $N_1 = \bigcap_{i=0}^n V_i$ and $N_{2^n} = V_0$. Denote by Q_k^* the union of all the elements of Q_k, $k = 0, 1, \dots, n$; note that $Q_0^* = Z_1$ and $Q_n^* = Z_{2^n}$. For each $j > 0$, the elements of Q_j form a pairwise disjoint family of relatively closed subsets of the open set $V \setminus \bigcup_{k=0}^{j-1} Q_k^*$. Also each Z_m in Q_j lies in N_m. Therefore, for each $j > 0$, there exists a collection of pairwise disjoint open sets M_m in $V \setminus \bigcup_{k=0}^{j-1} Q_k^*$, one for each Z_m in Q_j (that is, if $Z_m = Z_{m'} \in Q_j$, $m \neq m'$, then $M_m = M_{m'}$), such that

$$Z_m \subseteq M_m \subseteq N_m \setminus \bigcup_{i=n-j+1}^{n} X_{p(i)},$$

where j is such that Z_m is in Q_j and p is a permutation defining Z_m as above.
The set $Z_1 = \bigcap_{i=0}^n X_i \subseteq N_1 = \bigcap_{i=0}^n V_i$ is relatively closed in V and has the ε-strong C^p extraction property with respect to V, so there is a C^p diffeomorphism

$$h_1 : V \setminus Z_1 \to V \setminus (Z_1 \setminus U)$$

which is the identity outside $(N_1 \cap U) \cup Z_1$ and moves no point more than $\frac{\varepsilon}{2^n}$; in particular, $h_1(Z_2) \setminus U = Z_2 \setminus U$ should $Z_2 \neq Z_1$.
We will apply induction to prove that for $1 < m \leq 2^n$ there exists a C^p diffeomorphism

$$h_m : (V \setminus (\bigcup_{i=1}^{m-1} Z_i \setminus U)) \setminus h_{m-1} \circ \cdots \circ h_1(Z_m) \to V \setminus (\bigcup_{i=1}^{m} Z_i \setminus U)$$

which is the identity outside $(h_{m-1} \circ \cdots \circ h_1(M_m) \cap N_m \cap U) \cup h_{m-1} \circ \cdots \circ h_1(Z_m)$, satisfies

$$h_m \circ h_{m-1} \circ \cdots \circ h_1(V \setminus \bigcup_{i=1}^{m} Z_i) = V \setminus (\bigcup_{i=1}^{m} Z_i \setminus U) \text{ and } h_m \circ h_{m-1} \circ \cdots \circ h_1(Z_{m+1}) \setminus U = Z_{m+1} \setminus U,$$

and such that $h_m \circ \cdots \circ h_1$ moves no point more than $\frac{m\varepsilon}{2^n}$.
Suppose this is true for every $1 \leq k \leq m-1$ and let us check so it is for m. Assume $Z_m \in Q_j$; additionally, we can assume that $Z_{m-1} \neq Z_m$, otherwise, the identity in place of h_m will do. By the definition of the ε-strong C^p extractibility and Lemma 3.6(2), Z_m has the ε-strong C^p extraction property with respect to $V \setminus \bigcup_{k=0}^{j-1} Q_k^*$. Furthermore, one more application of Lemma 3.6(2) yields that $Z_m \subset V \setminus \bigcup_{i=1}^{m-1} Z_i$ is relatively closed and has the ε-strong C^p extraction property with respect to $V \setminus \bigcup_{i=1}^{m-1} Z_i$. By Lemma 3.6 (4), $h_{m-1} \circ \cdots \circ h_1(Z_m) \subset h_{m-1} \circ \cdots \circ h_1(V \setminus \bigcup_{i=1}^{m-1} Z_i)$ is relatively closed and has the strong C^p extraction property with respect to the open set

$$h_{m-1} \circ \cdots \circ h_1(V \setminus \bigcup_{i=1}^{m-1} Z_i) = V \setminus (\bigcup_{i=1}^{m-1} Z_i \setminus U).$$

Considering the open set $h_{m-1} \circ \cdots \circ h_1(M_m) \cap N_m \cap U$, there exists a C^p diffeomorphism h_m from

$$(V \setminus (\bigcup_{i=1}^{m-1} Z_i \setminus U)) \setminus h_{m-1} \circ \cdots \circ h_1(Z_m)$$

onto

$$(V \setminus (\bigcup_{i=1}^{m-1} Z_i \setminus U)) \setminus (h_{m-1} \circ \cdots \circ h_1(Z_m) \setminus (h_{m-1} \circ \cdots \circ h_1(M_m) \cap N_m \cap U) \tag{3.4.1}$$

which is the identity outside $(h_{m-1} \circ \cdots \circ h_1(M_m) \cap N_m \cap U) \cup h_{m-1} \circ \cdots \circ h_1(Z_m)$. Furthermore by Remark 3.7 and the fact that $h_{m-1} \circ \cdots \circ h_1$ moves no point more than $\frac{(m-1)\varepsilon}{2^n}$ (induction hypothesis), we have that $h_m \circ (h_{m-1} \circ \cdots \circ h_1)$ moves no point more than $\frac{\varepsilon}{2^n} + \frac{(m-1)\varepsilon}{2^n} = \frac{m\varepsilon}{2^n}$.
Finally note that the expression (3.4.1) is equal to $V \setminus (\bigcup_{i=1}^m Z_i \setminus U)$ because $h_{m-1} \circ \cdots \circ h_1(Z_m) \subseteq h_{m-1} \circ \cdots \circ h_1(M_m) \cap N_m$ and the fact that $h_{m-1} \circ \cdots \circ h_1(Z_m) \setminus U = Z_m \setminus U$.
To conclude define

$$h = h_{2^n} \circ \cdots \circ h_1 : V \setminus X_0 \to V \setminus (X_0 \setminus U).$$

This is a C^p diffeomorphism of $V \setminus X_0$ onto $V \setminus (X_0 \setminus U)$ which is the identity outside $(V_0 \cap U) \cup X_0$ and moves no point more than ε.
To end the proof we will show that h carries $X_i \setminus X_0$ into V_i for each $i = 1, \ldots, n$. Take $x \in X_i \setminus X_0$ for some $i = 1, \ldots, n$ and let us see that $h(x) \in V_i$. If $h_1(x) \neq x$ then $h_1(x) \in N_1$, so because $N_1 \subseteq V_i$ we have that $h_1(x) \in V_i$. Suppose now that $h_{m-1} \circ \cdots \circ h_1(x) \in V_i$. If $h_m \circ \cdots \circ h_1(x) \neq$

$h_{m-1} \circ \cdots \circ h_1(x)$ then we have that $x \in M_m$ and $h_m \circ \cdots \circ h_1(x) \in N_m$. We have that $x \in X_i \cap M_m$ and $M_m \subseteq N_m \setminus \bigcup_{i=n-j+1}^{n} X_{p(i)}$ where j is such that Z_m is in Q_j and p is a permutation that defines Z_m. Obviously we must have that $i = p(i_0)$ where $i_0 = \{1, \dots, n-j\}$. Also $N_m = \bigcap_{k=0}^{n-j} V_{p(k)} \subseteq V_i$, hence $h_m \circ \cdots \circ h_1(x)$ must also lie in V_i. Applying induction we have that $h(x) \in V_i$.

□

Remark 3.30. Notice that if we only have the strong C^p extraction property of the sets X_i with respect to V, we get the same result except that for any $\varepsilon > 0$ we cannot assure that the final extracting diffeomorphism moves points less than ε. In such a case we will say that we have a *weak version* of Lemma 3.29.
However, imagine we have that $X_0 \subseteq V_0, X_1 \subseteq V_1, \dots, X_n \subseteq V_n$ are of the form $g(X_i) \subseteq g(V_i)$, $i = 0, 1, \dots, n$, where g is a C^p diffeomorphism that moves points less than some $\delta_1 > 0$. In such a case, using Remark 3.7, for any $\delta_2 > 0$ we can make the final extracting diffeomorphism h of the proof of Lemma 3.29 satisfy that $h \circ g$ moves no point more than $\delta_1 + \delta_2$. In particular if $g(x) = x$ then we have that $||h(g(x)) - x|| \leq \delta_2$.

Remark 3.31. Assume, additionally, that the set V of Lemma 3.29 is of the form $V = \bigcup_{r \in \Omega} V_r$, where Ω is a set of indexes, each V_r is an open set , and $V_r \cap V_{r'} = \emptyset$ for every $r, r' \in \Omega$, $r \neq r'$. Then we can also require that the extracting C^p diffeomorphism of $V \setminus X_0$ onto $V \setminus (X_0 \setminus U)$ sends each set $V_r \setminus X_0$ into V_r for every $r \in \Omega$. To prove this, fix $r \in \Omega$ and replace the sequences $V_0, V_1, \dots, V_n$ and $X_0, X_1, \dots X_n$ with $V_0 \cap V_r, V_1 \cap V_r \dots, V_n \cap V_r$ and $X_0 \cap V_r, X_1 \cap V_r, \dots, V_n \cap V_r$, respectively. Further, observe that $\bigcup_{i=0}^{n} V_i \cap V_r = V_r$ and that each $X_i \cap V_r$ has the strong C^p extraction property with respect to V_r by Lemma 3.6 (2). According to the assertion of Lemma 3.29, we conclude that there exists a C^p diffeomorphism $h_r : V_r \setminus X_0 \to V_r \setminus (X_0 \setminus U)$ satisfying the suitable conditions. Finally, it is enough to set $h : V \setminus X_0 \to V \setminus (X_0 \setminus U)$ by letting $h(x) = h_r(x)$ for $x \in V_r \setminus X_0$.

The rest of the proof of Theorem 3.9 goes as in [150, Theorem 1], with some modifications due to the facts that we work here with an open set U not necessarily containing X, and that E is not necessarily separable.

Proof of Theorem 3.9. Apply Lemma 3.27 to the given cover $\mathcal{G}$ to find collections $\{X_i\}_{i\geq 1}$, $\{W_i\}_{i\geq 1}$ and $\{V_i\}_{i\geq 1}$ of subsets of E satisfying conditions (1)–(6) of that lemma. By Lemma 3.28, for all $i, j \in \mathbb{N}$, $X_i \cap V_j$ has the ε-strong C^p extraction property with respect to V_j. Moreover, if φ is a C^p diffeomorphism with this property, then $\varphi(V_{i,r}) \subset V_{i,r}$ and $\varphi(V_{j,r}) \subset V_{j,r}$ for every $r \in \Omega$.
Let us now define the required C^p diffeomorphism $g : E \setminus X \to E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$.

1. For a given V_1, define $I_1 = \{1_1 = 1, 1_2, \dots, 1_{n(1)}\} \subset \mathbb{N}$ to be the finite set of natural numbers such that $W_{1_1}, W_{1_2}, \dots, W_{1_{n(1)}}$ are the only $W_i's$ sets for which $V_1 \cap W_i \neq \emptyset$ (if there were infinitely many such $W_i's$, for infinitely many i, then $V_1 \cap V_i's \neq \emptyset$ which would contradict the star-finiteness of $\{V_i\}_{i\geq 1}$; obviously, we assume that $V_i \neq V_{i'}$ for $i \neq i'$). Since $\{W_i\}_{i\geq 1}$ is a cover we have $\bigcup_{i\in I_1} W_i \cap V_1 = V_1$. (A priori V_1 can be covered by a proper subfamily of $\{V_2, \dots, V_n\}$). Assuming that i_1 is the greatest number in I_1 (in particular $i_1 \geq 1$) we set

$$\varepsilon_1 = \frac{1}{2} \cdot \frac{1}{2^{i_1+1}} > 0.$$

We want to apply Lemma 3.29 for the sets

$$X_{1_1} = X_1 = X_1 \cap V_1, X_{1_2} \cap V_1, \dots, X_{1_{n(1)}} \cap V_1$$

which play the role of $X_0, \dots, X_n$ in the statement of the Lemma 3.29, and for the sets

$$W_{1_1} = W_1 = W_1 \cap V_1, W_{1_2} \cap V_1, \dots, W_{1_{n(1)}} \cap V_1,$$

which play the role of $V_0, \dots, V_n$ respectively, and for the positive number $\varepsilon_1 > 0$. Observe that each $X_i \cap V_1$, $i \in I_1$, has the strong C^p extraction property with respect to V_1. Hence, applying Lemma 3.29, we find a C^p diffeomorphism g_1 of $V_1 \setminus X_1$ onto $V_1 \setminus (X_1 \setminus U)$ which is the identity outside $(W_1 \cap U) \cup X_1$, carries $(X_l \setminus X_1) \cap V_1$ into $W_l \cap V_1$ for each $l > 1$ and moves no point more than ε_1.

By Remark 3.31 we may also assume that g_1 sends each set $V_{1,r} \setminus X_1$ into $V_{1,r}$. This means in particular that g_1 refines $\mathcal{G}$. Also, since g_1 moves no point more than ε_1 we cannot have that $x \in V_{i,r}$ and $g_1(x) \in V_{i,r'}$ for some $i \in I_1$ and different $r, r' \in \Omega$ (recall that $\operatorname{dist}(V_{i,r}, V_{i,r'}) \geq \frac{1}{2^{i+1}} > \varepsilon_1$).

Since $W_1 \cap U \subseteq \overline{W_1} \cap U \subseteq V_1 \cap U$ by making g_1 to be the identity outside $V_1 \setminus X_1$ there exists a well-defined natural extension of g_1 from $V_1 \setminus X_1$ to $E \setminus X_1$. Now we have a C^p diffeomorphism g_1 such that

(a) g_1 acts from $E \setminus X_1$ onto $E \setminus (X_1 \setminus U)$.

(b) g_1 is the identity on $E \setminus [(W_1 \cap U) \cup X_1]$. In particular it is the identity on $(E \setminus U) \setminus X_1$.

(c) g_1 carries $(X_l \setminus X_1) \cap V_1$ into $W_l \cap V_1$ for each $l > 1$.

(d) We require that if $X_1 = \emptyset$ or $X_{1,r} = \emptyset$ then g_1 is the identity on V_1 or $V_{1,r}$, respectively.

(e) g_1 moves no point more than $\varepsilon_1 = \frac{1}{2} \cdot \frac{1}{2^{i_1+1}}$.

(f) g_1 sends each set $V_{1,r} \setminus X_1$ into $V_{1,r}$ for every $r \in \Omega$, so g_1 refines $\mathcal{G}$.

2. Consider now the set V_2 and define $I_2 = \{2_1 = 2, 2_2, \dots, 2_{n(2)}\} \subset \mathbb{N}$ to be the finite set of natural numbers such that $W_{2_1}, W_{2_2}, \dots, W_{2_{n(2)}}$ are the only $W_i's$ sets for which $V_2 \cap W_i \neq \emptyset$ (if there were infinitely many such $W_i's$, for infinitely many i, then $V_2 \cap V_i's \neq \emptyset$ which would contradict the star-finiteness of $\{V_i\}_{i \geq 1}$). Assume that i_2 is the greatest number in I_2 (in particular $i_2 \geq 2$), and set
$$\varepsilon_2 = \frac{1}{2^2} \cdot \frac{1}{2^{i_2+1}} > 0.$$
Since $\{W_i\}_{i \geq 1}$ is a cover we have
$$\bigcup_{i \in I_2} g_1(V_2 \setminus X_1) \cap W_i \cap V_2 = g_1(V_2 \setminus X_1) \cap V_2.$$
Again, we want to apply Lemma 2.26 for the sets
$$\{g_1((X_i \setminus X_1) \cap V_2) \cap V_2 : \, i \in I_2\}$$
playing the role of $X_0, \dots, X_n$ in the statement of the lemma, for
$$\{g_1(V_2 \setminus X_1) \cap W_i \cap V_2 : \, i \in I_2\}$$
playing the role of the sets $V_0 \dots, V_n$ respectively. Here we should recall that $g_1(X_i \setminus X_1) \subseteq W_i$. Observe that by Lemma 3.6 (4), each $g_1((X_i \setminus X_1) \cap V_2) \cap V_2$ has the strong C^p extraction property with respect to the open set $g_1(V_2 \setminus X_1) \cap V_2$. Applying the weak version of Lemma 3.29 to these sets we get a C^p diffeomorphism g_2 of
$$[g_1(V_2 \setminus X_1) \cap V_2] \setminus [g_1(X_2 \setminus X_1) \cap V_2] = g_1(V_2 \setminus (X_1 \cup X_2)) \cap V_2$$
onto
$$[g_1(V_2 \setminus X_1) \cap V_2] \setminus [(g_1(X_2 \setminus X_1) \cap V_2) \setminus U]$$
which is the identity outside
$$(g_1(V_2 \setminus (X_1 \cup X_2)) \cap W_2 \cap V_2 \cap U) \cup (g_1(X_2 \setminus X_1) \cap V_2)$$

and carries

$$g_1(X_k \setminus (X_1 \cup X_2) \cap V_2) \cap V_2$$

into

$$g_1(V_2 \setminus X_1) \cap W_k \cap V_2$$

for each $k > 2$. Moreover using Remark 3.30 one can also assume that $g_2 \circ g_1$ moves points less than $\varepsilon_1 + \varepsilon_2$.
Because $\overline{g_1(V_2 \setminus (X_1 \cup X_2)) \cap W_2} \cap U \subseteq g_1(V_2 \setminus (X_1 \cup X_2)) \cap V_2 \cap U$, by letting g_2 be the identity outside $g_1(V_2 \setminus (X_1 \cup X_2)) \cap V_2$ there exists a well-defined natural extension of g_2 to $g_1(E \setminus (X_1 \cup X_2))$. To sum up we have the following properties:

(a) g_2 acts from

$$g_1(E \setminus (X_1 \cup X_2))$$

onto

$$\begin{aligned} g_1(E \setminus X_1) &\setminus [g_1(X_2 \setminus X_1) \setminus U] \\ &= E \setminus (X_1 \setminus U) \setminus [(g_1((X_2 \setminus X_1) \cap U) \cup g_1((X_2 \setminus X_1)) \setminus U) \setminus U] \\ &= E \setminus (X_1 \setminus U) \setminus [(X_2 \setminus X_1) \setminus U] \\ &= E \setminus ((X_1 \cup X_2) \setminus U). \end{aligned}$$

(Here we are using the fact that $g_1((X_2 \setminus X_1) \cap U) \subseteq U$ and that g_1 is the identity outside U).

(b) g_2 is the identity on $E \setminus [(g_1(W_2 \setminus (X_1 \cup X_2)) \cap W_2 \cap U) \cup (g_1(X_2 \setminus X_1) \cap V_2)]$. Since

$$(g_1(W_2 \setminus (X_1 \cup X_2)) \cap W_2 \cap U) \cup (g_1(X_2 \setminus X_1) \cap V_2) \subseteq U \cup X_1 \cup X_2,$$

in particular g_2 is the identity on $E \setminus (U \cup (X_1 \cup X_2))$. Because g_1 is the identity outside U then g_2 is the identity on $g_1(E \setminus (U \cup (X_1 \cup X_2)) = g_1((E \setminus U) \setminus (X_1 \cup X_2))$.

(c) g_2 carries

$$g_1(X_l \setminus (X_1 \cup X_2)) \cap V_2$$

into

$$W_l \cap V_2$$

for each $l > 2$.

(d) If $X_{2,r} = \emptyset$ then we require that g_2 is the identity on $g_1(V_{2,r} \setminus X_1) \cap V_2$ and in $g_1(V_2 \setminus X_1) \cap V_{2,r}$.

(e) We have

$$\begin{cases} \|g_2(g_1(x)) - x\| \leq \varepsilon_1 + \varepsilon_2 = \frac{1}{2} \cdot \frac{1}{2^{i_1+1}} + \frac{1}{4} \cdot \frac{1}{2^{i_2+1}} & \text{for every } x \in E \setminus (X_1 \cup X_2) \\ \|g_2(g_1(x)) - x\| < \frac{1}{2^{2+1}} & \text{for every } x \in V_2 \setminus (X_1 \cup X_2). \end{cases} \tag{3.4.2}$$

The first inequality is clear. For the second one, when $x \notin V_2 \setminus V_1$ we have $g_1(x) = x$, hence

$$\|g_2(g_1(x)) - x\| = \|g_2(x) - x\| \leq \varepsilon_2 = \frac{1}{4} \cdot \frac{1}{2^{i_2+1}} < \frac{1}{2^{2+1}},$$

and if $x \in V_1 \cap V_2$ then $V_1 \cap V_2 \neq \emptyset$, so $i_1 \geq 2$ and

$$\|g_2(g_1(x)) - x\| \leq \frac{1}{2} \cdot \frac{1}{2^{i_1+1}} + \frac{1}{4} \cdot \frac{1}{2^{i_2+1}} < \left(\frac{1}{2} + \frac{1}{4}\right) \cdot \frac{1}{2^{2+1}}.$$

(f) The composition $g_2 \circ g_1$ is a mapping from $E \setminus (X_1 \cup X_2)$ onto $E \setminus ((X_1 \cup X_2) \setminus U)$ and it is the identity outside $(E \setminus U) \setminus (X_1 \cup X_2)$. Let us check that it refines $\mathcal{G}$. Take $x \in E \setminus (X_1 \cup X_2)$. If $g_2(g_1(x)) = g_1(x)$, since g_1 refines $\mathcal{G}$ we are done. Otherwise we have $g_2(g_1(x)) \neq g_1(x)$, and then $g_1(x), g_2(g_1(x)) \in g_1(V_2 \setminus (X_1 \cup X_2)) \cap V_2$. We have that $x, g_2(g_1(x)) \in V_2$. We must have $x, g_2(g_1(x)) \in V_{2,r}$ for some $r \in \Omega$ (as otherwise $x \in V_{2,r}$ and $g_2(g_1(x)) \in V_{2,r'}$ a contradiction with (3.4.2) since $\text{dist}(V_{2,r}, V_{2,r'}) \geq \frac{1}{2^{2+1}}$).

3. We go on doing this process by successive applications of Lemma 3.29. We want to apply induction to prove that for each $j \geq 3$ we can find a C^p diffeomorphism g_j such that

(a) g_j acts from

$$g_{j-1} \circ \cdots \circ g_1 \left(E \setminus \bigcup_{k \leq j} X_k \right)$$

onto

$$E \setminus \left(\bigcup_{k \leq j} X_k \setminus U \right).$$

(b) g_j is the identity on $g_{j-1} \circ \cdots \circ g_1 \Big((E \setminus U) \setminus (\bigcup_{k \leq j} X_k) \Big)$.

(c) g_j carries

$$g_{j-1} \circ \cdots \circ g_1 (X_l \setminus \bigcup_{k \leq j} X_k) \cap V_j$$

into

$$W_l \cap V_j$$

for each $l > j$.

(d) If $X_{j,r} = \emptyset$ then g_j is the identity on $g_{j-1} \circ \cdots \circ g_1(V_{j,r} \setminus \bigcup_{k<j} X_k) \cap V_j$ and on $g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap V_{j,r}$.

(e) We have

$$\begin{cases} ||g_j \circ \cdots \circ g_1(x) - x|| \leq \sum_{k=1}^{j} \frac{1}{2^k} \cdot \frac{1}{2^{i_k+1}} & \text{for every } x \in E \setminus (\bigcup_{k \leq j} X_k) \\ ||g_j \circ \cdots \circ g_1(x) - x|| < \frac{1}{2^{j+1}} & \text{for every } x \in V_j \setminus (\bigcup_{k \leq j} X_k), \end{cases} \tag{3.4.3}$$

where i_k is the greatest number such that $V_{i_k} \cap V_k \neq \emptyset$. Let us call for every $k = 1, \ldots, j$,

$$\varepsilon_k = \frac{1}{2^k} \cdot \frac{1}{2^{i_k+1}} > 0.$$

(f) $g_j \circ \cdots \circ g_1$ refines $\mathcal{G}$.

Suppose this is true for $j - 1$ and let us check this is so for j.

The idea is the same as in steps (1) and (2). We first find the set $I_j = \{j_1 = j, j_2, \ldots, j_{n(j)}\} \subset \mathbb{N}$ such that $W_{j_1}, W_{j_2}, \ldots, W_{j_{n(j)}}$ are the only $W_i's$ sets for which $V_j \cap W_i \neq \emptyset$. Assume i_j is the greatest number in I_j (in particular $i_j \geq j$) and set

$$\varepsilon_j = \frac{1}{2^j} \cdot \frac{1}{2^{i_j+1}} > 0.$$

We have that

$$\bigcup_{i \in I_j} g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap W_i \cap V_j = g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap V_j.$$

We want to apply Lemma 2.26 for the sets

$$\{g_{j-1} \circ \cdots \circ g_1((X_i \setminus \bigcup_{k<j} X_k) \cap V_j) \cap V_j : i \in I_j\}$$

playing the role of $X_0, \ldots, X_n$ in the statement of the lemma, for

$$\{g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap W_i \cap V_j : i \in I_j\}$$

playing the role of the sets $V_0 \ldots, V_n$ respectively.
Observe that by Lemma 3.6 (4), each $g_{j-1} \circ \cdots \circ g_1((X_i \setminus \bigcup_{k<j} X_k) \cap V_j) \cap V_j$ has the strong C^p extraction property with respect to the open set $g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap V_j$. Applying the weak version of Lemma 3.29 to these sets we get a C^p diffeomorphism g_j from

$$\left[g_{i-1} \circ \cdots \circ g_1(V_i \setminus \bigcup_{j<i} X_j) \cap V_i\right] \setminus \left[g_{j-1} \circ \cdots \circ g_1(X_j \setminus \bigcup_{k<j} X_k) \cap V_j\right]$$
$$= g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k\le j} X_k) \cap V_j$$

onto

$$\left[g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap V_j\right] \setminus \left[g_{j-1} \circ \cdots \circ g_1(X_j \setminus \bigcup_{k<j} X_k) \cap V_j \setminus U\right]$$

which is the identity outside

$$\left((g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k\le j} X_k) \cap W_j \cap U\right) \cup \left(g_{j-1} \circ \cdots \circ g_1(X_j \setminus \bigcup_{k<j} X_k) \cap V_j\right)$$

and carries

$$g_{j-1} \circ \cdots \circ g_1(X_l \setminus \bigcup_{k\le j} X_k) \cap V_j$$

into

$$W_l \cap V_j$$

for each $l > j$. This last property establishes (c).
Define g_j to be the natural extension of $g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k\le j} X_k) \cap V_j$ to $g_{j-1} \circ \cdots \circ g_1(E \setminus \bigcup_{k\le j} X_k)$, so now g_j is defined from

$$g_{j-1} \circ \cdots \circ g_1 \left(E \setminus \bigcup_{k\le j} X_k)\right)$$

onto

$$\left[g_{j-1} \circ \cdots \circ g_1(E \setminus \bigcup_{k<j} X_k)\right] \setminus \left[g_{j-1} \circ \cdots \circ g_1(X_j \setminus \bigcup_{k<j} X_k) \setminus U\right] = E \setminus (\bigcup_{k<j} X_j \setminus U)$$
$$\setminus \left[\left(g_{j-1} \circ \cdots \circ g_1((X_j \setminus \bigcup_{k<j} X_k) \cap U) \cup g_{j-1} \circ \cdots \circ g_1((X_j \setminus \bigcup_{k<j} X_k) \setminus U)\right) \setminus U\right]$$
$$= E \setminus (\bigcup_{k<j} X_k \setminus U) \setminus \left[(X_j \setminus \bigcup_{k<j} X_k) \setminus U\right] = E \setminus \left(\bigcup_{k\le j} X_k \setminus U\right).$$

Here we are using that $g_{j-1} \circ \cdots \circ g_1((X_j \setminus \bigcup_{k<j} X_k) \cap U \subseteq U$ and also (a) from the induction hypothesis. This establishes (a).

We also have that g_j is the identity on $(E \setminus U) \setminus (\bigcup_{k\leq j} X_k) = g_{j-1} \circ \cdots \circ g_1 \left((E \setminus U) \setminus (\bigcup_{k\leq j} X_k) \right)$, which establishes (b).

If $X_{j,r} = \emptyset$ then we let g_j be the identity on $g_{j-1} \circ \cdots \circ g_1(V_{j,r} \setminus \bigcup_{k<j} X_k) \cap V_j$ and on $g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap V_{j,r}$. This last property implies (d).

For the property (e), using Remark 3.30 one can assume that $g_j \circ \cdots \circ g_1$ moves points less than $\sum_{k=1}^{j} \varepsilon_k$, because by the induction hypothesis $g_{j-1} \circ \cdots g_1$ moves points less than $\sum_{k=1}^{j-1} \varepsilon_k$. Let us check then that the second property in (3.4.3) is satisfied. Take $x \in V_j \setminus (\bigcup_{k\leq j} X_j)$ and observe that if $g_k \circ \cdots \circ g_1(x) \neq g_{k-1} \circ \cdots \circ g_1(x)$ for some $k = 1, \cdots, j$ means that $x \in V_k$, hence $k \in I_j$ and $i_k \geq j$. We can write

$$||g_j \circ \cdots \circ g_1(x) - x|| \leq \sum_{k \in I_j, k \leq j} \left(\frac{1}{2^k} \cdot \frac{1}{2^{i_k+1}} \right) \leq \sum_{k=1}^{j} \frac{1}{2^k} \cdot \frac{1}{2^{j+1}} < \frac{1}{2^{j+1}}.$$

Now, using our induction hypothesis (f) that $g_{j-1} \circ \cdots \circ g_1$ refines $\mathcal{G}$, we can prove that $g_j \circ \cdots \circ g_1$ still refines $\mathcal{G}$. If we take an $x \in E \setminus (\bigcup_{k\leq j} X_k)$ and $g_j \circ \cdots \circ g_1(x) = g_{j-1} \circ \cdots \circ g_1(x)$, since $g_{j-1} \circ \cdots g_1$ refines $\mathcal{G}$ we are done. Otherwise $g_{j-1} \circ \cdots g_1(x), g_j \circ \cdots g_1(x) \in g_{j-1} \circ \cdots \circ g_1(V_j \setminus \bigcup_{k<j} X_k) \cap V_j$, so $x, g_j \circ \cdots \circ g_1(x) \in V_j$. We must have $x, g_j \circ \cdots \circ g_1(x) \in V_{j,r}$ for some $r \in \Omega$. Indeed, otherwise $x \in V_{j,r}$ and $g_j \circ \cdots \circ g_1(x) \in V_{j,r'}$ for different $r, r' \in \Omega$, a contradiction with (3.4.3) since $\text{dist}(V_{j,r}, V_{j,r'}) \geq \frac{1}{2^{j+1}}$.

Hence we have finished our induction process.

To conclude, note that $g_j \circ \cdots \circ g_1(x) \neq g_{j-1} \circ \cdots \circ g_1(x)$ implies $x \in V_j$. This fact ensures the existence of a well-defined C^p diffeomorphism $g(x) = \lim_{j\to\infty} g_j \circ \cdots \circ g_1(x)$ from $E \setminus X$ onto $E \setminus (X \setminus U)$. The mapping g is the identity on $(E \setminus U) \setminus X$ and because $\{V_{j,r}\}_{j\geq 1, r\in\Omega}$ is a refinement of $\mathcal{G}$, g is limited by $\mathcal{G}$. ☐

3.5 Some corollaries

We present in this section three easy corollaries. The first one will be used later on in Chapter 5 for the proof of Theorem 5.3, and the last two ones are versions of Theorem 3.3 where we further assume the separability of the Banach space E.

When we say that a closed set $X \subset E$ is locally contained in a finite union of complemented subspaces of infinite codimension we mean that for every $x \in X$ there exists an open neighbourhood U_x of x and some closed subspaces $E_1, \ldots, E_{n_x} \subset E$ complemented in E and of infinite codimension such that

$$X \cap U_x \subset \bigcup_{j=1}^{n_x} E_j.$$

Corollary 3.32. *Let E be a Banach space with a C^p smooth norm. Take an open cover $\mathcal{G}$ of an open set U and a closed set $X \subset U$ that is locally contained in a finite union of complemented subspaces of infinite codimension in E. Then there exists a C^p-diffeomorphism $h : E \to E \setminus X$ which is the identity outside U and is limited by $\mathcal{G}$.*

Proof. For every $x \in X$ choose an open neighbourhood U_x of x and some closed subspaces $E_1, \ldots, E_{n_x}$ complemented in E and of infinite codimension such that

$$X \cap U_x \subseteq \bigcup_{j=1}^{n_x} E_j.$$

If E admits an equivalent C^p smooth norm it is known, by applying Theorem 3.3, that given a complemented subspace $H \subset E$ of infinite codimension and the open set U_x, then $H \cap U_x$ has the ε-strong C^p extraction property with respect to any open set U' for which $H \cap U_x$ is a relatively closed subset of U'. Therefore thanks to Lemma 3.6 (3) the set $\bigcup_{j=1}^{n_x} E_j \cap U_x$ has the ε-strong C^p extraction property with respect to any open set U' for which $\bigcup_{j=1}^{n_x} E_j \cap U_x$ is relatively closed on U'.
Now, using Lemma 3.6 (1), the set $X \cap U_x \subseteq \bigcup_{j=1}^{n_x} E_j \cap U_x$ has the ε-strong C^p strong extraction property with respect to any open set U' for which $X \cap U_x$ is relatively closed on U'. And this means that X has locally the ε-strong C^p strong extraction property. To conclude the proof apply Theorem 3.9 to find a C^p diffeomorphism $g : E \setminus X \to E$ which is the identity on $E \setminus U$ and is limited by $\mathcal{G}$ (note that we have $X \subset U$ and hence $X \setminus U = \emptyset$). Finally define $h = g^{-1}$. □

We say that a closed set $X \subset E$ is locally contained in subspaces of finite dimension if for every $x \in X$ there exists an open set U_x and a finite-dimensional subspace E_x of E such that $U_x \cap X \subset E_x$. In particular such kind of sets are locally compact.

Corollary 3.33. *Let E be an infinite-dimensional separable Banach space and $X \subset E$ be a closed set that is locally contained in finite-dimensional subspaces of E. Then for every open cover $\mathcal{G}$ of E and every open subset U of E, there exists a C^∞ diffeomorphism h from $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$.*

Proof. We know that the separability of E implies that E embeds linearly into some $c_0(\Gamma)$ (see [60, P. 246]) and hence by [63, Proposition 5.1], E has a C^∞ smooth non-complete norm (obviously not equivalent). On the other hand since any finite-dimensional space is complemented and admits C^∞ partitions of unity we can straightforwardly apply Theorem 3.3 to conclude the proof. □

Similarly, using the fact that the separability of the Banach space together with the existence of C^p smooth bump functions imply the existence of C^p smooth partitions of unity we can get:

Corollary 3.34. *Let E be an infinite-dimensional separable Banach space with C^p smooth bump functions, $p \in \mathbb{N} \cup \{\infty\}$. Let $X \subset E$ be a closed set with the property that, for each $x \in X$, there exist a neighbourhood U_x of x in E, Banach spaces $E_{(1,x)}$ and $E_{(2,x)}$, and a continuous mapping $f_x : C_x \to E_{(2,x)}$, where C_x is a closed subset of $E_{(1,x)}$, such that:*

1. *$E = E_{(1,x)} \oplus E_{(2,x)}$, where $E_{(2,x)}$ is infinite-dimensional.*
2. *$X \cap U_x \subset G(f_x)$, where*

$$G(f_x) = \{y = (y_1, y_2) \in E_{(1,x)} \oplus E_{(2,x)} \ : \ y_2 = f_x(y_1), y_1 \in C_x\}.$$

Then, for every open cover $\mathcal{G}$ of E and every open subset U of E, there exists a C^p diffeomorphism h from $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$.

3.6 Compact sets are graphs of continuous functions

In this last section of Chapter 3 we wonder for which Banach spaces E, if we are given a compact subset $K \subset E$, we can write $E = E_1 \oplus E_2$, with E_2 infinite-dimensional, and we can find a continuous function $f : X_1 \to E_2$, where X_1 is closed subset of E_1, such that $K = G(f) = \{x_1 + f(x_1) : x_1 \in X_1\}$. It will be shown in this section that for spaces with unconditional basis this fact is true.

So for Banach spaces with unconditional basis Theorem 3.3 applies for compact sets. In particular it becomes clear that we are extending Renz and West's results, which focused on the diffeomorphic extraction of compact sets.

Although we work in the framework of Banach spaces, the results of this section are still true in the more general context of Fréchet spaces.

It will be mandatory to use a result due to Corson from 1971 ([55]).

Theorem 3.35 (Corson, 1971). *Let E be a Banach space and let S be a σ-compact (i.e. a countable union of compact sets) linear subspace. Let $\{V_n\}_{n\in\mathbb{N}}$ be a family of infinite-dimensional closed subspaces of E. Then there exists a sequence of linearly independent vectors $\{v_n\}_{n\in\mathbb{N}}$ such that $v_n \in V_n$ for every $n \in \mathbb{N}$ and such that $\overline{\operatorname{span}}\{v_n : n \in \mathbb{N}\} \cap S = \{0\}$.*

Proof. Write $S = \bigcup_{i=1}^{\infty} S_i$, where each S_i is a compact set. Write $-S_i = \{x : -x \in S_i\}$. Because E is a Banach space the closed convex hull of each $S_i \cup (-S_i)$, that is

$$\overline{co}(S_i \cup (-S_i)) = \overline{\{\sum_{j=1}^{n} \lambda_j x_j : x_j \in S_i \cup (-S_i),\, \lambda_j \in \mathbb{R},\, \lambda_j \geq 0,\, \sum_{j=1}^{n} \lambda_j = 1\}},$$

is a compact convex set [3] . Define $S^* = \bigcup_{n,i=1}^{\infty} n \cdot \overline{co}(S_i \cup (-S_i)) \supseteq S$, which is a σ-compact linear subspace.
Thanks to the separability of the space we can find a countable number of closed balls $\{B_j\}_{j\geq 1}$ without including the origin such that $\bigcup_{j=1}^{\infty} B_j \supseteq S^* \setminus \{0\}$. Noting that for each $i, j \in \mathbb{N}$, the set $\overline{co}(S_i \cup (-S_i)) \cap B_j$ is compact and convex, and that

$$\bigcup_{j,n,i=1}^{\infty} (n \cdot \overline{co}(S_i \cup (-S_i))) \cap B_j = \bigcup_{j=1}^{\infty} S^* \cap B_j = S^* \setminus \{0\},$$

we conclude that we can write $S^* \setminus \{0\} = \bigcup_{n=1}^{\infty} K_n$ as a countable union of compact convex sets not including the origin. If we prove the result for this new σ-compact subspace S^* we get it also for S, hence there is no problem in assuming that $S \setminus \{0\} = \bigcup_{n=1}^{\infty} K_n$, where each K_n is compact and convex. By induction we will prove that there exist closed subspaces W_n of finite-codimension in E and points $v_n \in W_n \cap V_n$ such that:

(a) $v_1, \dots, v_n$ are linearly independent.

(b) If $n \geq 2$, $W_n \subseteq W_{n-1} \setminus \{v_{n-1}\}$.

(c) $\operatorname{span}\{v_1, \dots, v_n\} \cap S = \{0\}$.

(d) $(K_n + \operatorname{span}\{v_1, \dots, v_{n-1}\}) \cap W_n = \emptyset$

Observe that K_1 is closed and convex and does not touch the origin, hence by applying the Hahn-Banach separation theorem there exists an hyperplane W_1, which is the kernel of a continuous linear functional, such that $W_1 \cap K_1 = \emptyset$. Then $W_1 \cap V_1$ is closed and infinite-dimensional, so there exists a non-null vector $v_1 \in (W_1 \cap V_1) \setminus S$. If this were not true we would get that $W_1 \cap V_1$ is a σ-compact subspace, which is a contradiction with the fact that it is an infinite-dimensional and closed subspace. Clearly conditions (a) and (d) are satisfied. For condition (c) observe that $v_1 \notin S$ and that S is a subspace, then $\operatorname{span}\{v_1\} \cap S = \{0\}$.
Let us suppose now that $v_1, \dots, v_m$ and $W_1, \dots, W_m$ have been chosen satisfying (a) ,(b), (c) and (d). The sets $\operatorname{span}\{v_1, \dots, v_m\}$ and K_{m+1} are closed and convex and so it is $\operatorname{span}\{v_1, \dots, v_m\} + K_{m+1}$. Note that we have that $0 \notin \operatorname{span}\{v_1, \dots, v_m\} + K_{m+1}$ because otherwise we would be contradicting (c) from the induction hypothesis. As before by the separation theorem one may pick H_{m+1} a hyperplane which does not touch $\operatorname{span}\{v_1, \dots, v_m\} + K_{m+1}$. Since W_m was infinite-codimensional we may also take some infinite-codimensional subspace $W_{m+1} \subset H_{m+1}$ satisfying $W_{m+1} \subset W_m \setminus \{v_m\}$. Now again we can choose another non-null element $v_{m+1} \in (W_{m+1} \cap V_{m+1}) \setminus (S + \operatorname{span}\{v_1, \dots, v_n\})$. Clearly (b) and (d) are valid. For (a) just observe that we are imposing $v_{m+1} \notin \operatorname{span}\{v_1 \dots, v_m\}$. And to prove (c) assume by reduction to the absurd that there exists $a_1, \dots, a_{m+1} \in \mathbb{R}$ such that

$$\sum_{n=1}^{m} a_n v_n + a_{m+1} v_{m+1} \in S \setminus \{0\}.$$

[3]This is even true in Fréchet spaces, see [6, Corollary 5.34, 5.35].

Then we must have $a_{m+1} \neq 0$ as otherwise we get a contradiction with (c) from the induction hypothesis. But if $a_{m+1} \neq 0$ then v_{m+1} belongs to $S + \operatorname{span}\{v_1, \dots, v_n\}$, which is a contradiction.
Let us now complete the proof. We have already said that the points $\{v_n\}_{n\in\mathbb{N}}$ are linearly independent. It remains to check that $\overline{\operatorname{span}}\{v_n : n \in \mathbb{N}\} \cap S = \{0\}$. By contradiction suppose this is not true. Then for some $i \in \mathbb{N}$ there exists an element $k_i \in K_i$ such that

$$k_i = \lim_{n\to\infty} x_j = \lim_{j\to\infty} \sum_{n=1}^{\infty} a_{nj} v_n,$$

where $x_j = \sum_{n=1}^{\infty} a_{nj} v_n \in \operatorname{span}\{v_n : n \in \mathbb{N}\}$, so in fact each element x_j is defined by a finite sum. The space W_i is closed and therefore $\operatorname{span}\{v_1, \dots, v_{i-1}\} + W_i$ is also closed. And now noting that for every $n \geq i$, $v_n \in W_i$,

$$x_j = \sum_{n=1}^{i-1} a_{nj} v_n + \sum_{n=i}^{\infty} a_{nj} v_n \;\in\; \operatorname{span}\{v_1, \dots, v_{i-1}\} + W_i,$$

and hence $k_i \in \{v_1, \dots, v_{i-1}\} + W_i$ which is a contradiction with property (d) above.

□

Corollary 3.36. *Let E be a Banach space with an unconditional Schauder basis $B = \{e_i\}_{i\in\mathbb{N}}$ and let S be a σ-compact linear subspace. Then there exists a closed complemented infinite-dimensional subspace F such that $F \cap S = \{0\}$.*

Proof. Write the set of natural numbers as a disjoint union of infinite subsets $\mathbb{N} = \bigcup_{n\in\mathbb{N}} I_n$ and define the closed infinite-dimensional subspaces $V_n = \overline{\operatorname{span}}\{e_j : j \in I_n\}$, which are complemented in E thanks to the unconditionality of the basis. Applying the previous Theorem 3.35 we get a sequence of linearly independent vectors $\{v_n\}_{n\in\mathbb{N}}$ such that $v_n \in V_n$ for every $n \in \mathbb{N}$ and such that $\overline{\operatorname{span}}\{v_n : n \in \mathbb{N}\} \cap S = \{0\}$.
For each $n \in \mathbb{N}$ we can write $v_n \in V_n$ as

$$v_n = \sum_{j\in I_n} a_j e_j,$$

and since $v_n \neq 0$ there exists some $a_{j(n)} \neq 0$. Then we have that

$$V_n = \overline{\operatorname{span}}\,\{e_j : j \in I_n \setminus \{j(n)\}\} \oplus \operatorname{span}\{v_n\}.$$

Thanks to the unconditionality of the basis we also have that

$$E = \overline{\operatorname{span}}\{e_j : j \neq j(n) \text{ for every } n \in \mathbb{N}\} \oplus \overline{\operatorname{span}}\{v_n : n \in \mathbb{N}\}.$$

Set $F := \overline{\operatorname{span}}\{v_n : n \in \mathbb{N}\}$ and the proof is complete. □

Theorem 3.37. *Let E be a Banach space with an unconditional Schauder basis $B = \{e_i\}_{i\in\mathbb{N}}$ and let K be a compact set. Then we can write $E = E_1 \oplus E_2$ as a complemented sum of subspaces where E_2 is infinite-dimensional and $E_2 \cap K = \{0\}$. If $\pi_1 : E \to E_1$ denotes the orthogonal projection onto E_1, there exists a continuous function $f : \pi_1(K) \to E_2$ such that*

$$K = \{x_1 + f(x_1) : x_1 \in \pi_1(K)\}.$$

Proof. Let $S = \operatorname{span}(K)$ which is a σ-compact linear subspace because we can write

$$\operatorname{span}(K) = \bigcup_{n=1}^{\infty} \{a_1 x_1 + \cdots + a_n x_n : x_1, \dots, x_n \in K,\, a_1, \dots, a_n \in [-n, n]\}.$$

Then by Corollary 3.36 there are complementary closed subspaces E_1 and E_2 of E such that E_2 is infinite-dimensional and $E_2 \cap S = \{0\}$. It follows that for each fixed $x_1 \in E_1$ the set $(x_1 + E_2) \cap K$ consists of at most one point. Hence there is a function $f : \pi_1(K) \to E_2$ such that

$$K = \{x_1 + f(x_1) : x_1 \in \pi_1(K)\} = G(f).$$

Let us show that the preimage of any closed set $C \subset E_2$ by f is a closed set to get the continuity of f. Since $\pi_1(K)$ is compact, $(\pi_1(K) \oplus C) \cap K$ is closed and then $f^{-1}(C) = \pi_1((\pi_1(K) \oplus C) \cap K)$ is closed because $\pi_1 : G(f) = K \to \pi_1(K)$ is a continuous and closed map.

□

We therefore achieve the following result, which makes it clear why our Theorem 3.3 extends the results of West and Renz.

Corollary 3.38. *Let E be an infinite-dimensional Banach space with an unconditional Schauder basis and a C^p smooth bump function, and $X \subset E$ be a closed set that is locally compact. Then for every open cover $\mathcal{G}$ of E and every open subset U of E, there exists a C^p diffeomorphism h from $E \setminus X$ onto $E \setminus (X \setminus U)$ which is the identity on $(E \setminus U) \setminus X$ and is limited by $\mathcal{G}$. Moreover, the same conclusion is true if we replace E with an open subset of E.*

Proof. Note that the separability of the space and the existence of C^p smooth bump functions imply the existence of C^p smooth partitions of unity. Moreover we have the existence of noncomplete C^p smooth norms on any subspace of E. These facts together with Theorem 3.37 allow us to apply Theorem 3.3 and conclude the proof. □

Chapter 4

Smooth approximation without critical points of continuous functions between Banach spaces

4.1 Introduction and main results

The main purpose of this chapter is to show the following two results.

Theorem 4.1. *Let E, F be separable Hilbert spaces, and assume that E is infinite-dimensional. Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^∞ mapping $g : E \to F$ such that $\|f(x) - g(x)\| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.*

Theorem 4.2. *Let E be one of the classical Banach spaces c_0, ℓ_p or L^p, $1 < p < \infty$. Let F be a Banach space, and assume that there exists a bounded linear operator from E onto F. Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^k mapping $g : E \to F$ such that $\|f(x) - g(x)\| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.*

Here k denotes the order of smoothness of the space E, defined as follows: $k = \infty$ if $E \in \{c_0\} \cup \{\ell_{2n} : n \in \mathbb{N}\} \cup \{L^{2n} : n \in \mathbb{N}\}$; $k = 2n$ if $E \in \{\ell_{2n+1} : n \in \mathbb{N}\} \cup \{L^{2n+1} : n \in \mathbb{N}\}$, and k is equal to the integer part of p if $E \in \{\ell_p\} \cup \{L^p\}$ and $p \notin \mathbb{N}$. The Sobolev spaces $W^{k,p}(\mathbb{R}^n)$ with $1 < p < \infty$ are also included in Theorem 4.2 since they are isomorphic to $L^p(\mathbb{R}^n)$ (see [132, Theorem 11]).

Notice that the assumption that there exists a bounded linear operator from E onto F is necessary, as otherwise all points of E are critical for all functions $g \in C^1(E, F)$.

Of course Theorem 4.1 is a particular case of Theorem 4.2 (also note that if E is a separable Hilbert space, F is a Banach space, and there exists a continuous linear surjection $T : E \to F$, then F must be isomorphic to $\mathbb{R}^n$ or to E). In general, note that a continuous linear surjection $T : E \to F$ between Banach spaces exists if and only if F is isomorphic to a quotient space of E.

We will also establish more technical results (see Theorems 4.4 and 4.5 below) that generalize the preceding theorems to much larger classes of Banach spaces (especially in the case that E is reflexive).

Part of the motivation for this kind of results is in their connection with the Morse-Sard theorem [125, 140] which states that if $k \geq \max\{n - m + 1, 1\}$ then $f(C_f)$ is of Lebesgue measure zero in $\mathbb{R}^m$. This result also holds true for C^k smooth mappings $f : N \longrightarrow M$ between two smooth manifolds of dimensions n and m respectively.

Given the crucial applications of the Morse-Sard theorem in several branches of mathematics, it is natural both to try to extend this result for other classes of mappings, and also to ask what happens in the case that M and N are infinite-dimensional manifolds. Regarding the first issue, many refinements of the Morse-Sard theorem for other classes of mappings (notably Hölder, Sobolev, and BV mappings) have appeared in the literature; see for instance [154, 130, 31, 32, 126, 58, 82, 43, 44, 113, 96, 95, 18, 21] and the references therein.

As for the second issue, which in this chapter (and in the following one) is of our concern, let us mention the results of several authors who have studied the question as to what extent one can obtain results similar to the Morse-Sard theorem for mappings between infinite-dimensional Banach spaces or manifolds modelled on such spaces.

S. Smale [142] proved that if X and Y are separable connected smooth manifolds modelled on Banach spaces and $f : X \longrightarrow Y$ is a C^r Fredholm mapping (that is, every differential $Df(x)$ is a Fredholm operator between the corresponding tangent spaces) then $f(C_f)$ is meager, and in particular $f(C_f)$ has no interior points, provided that $r > \max\{\text{index}(Df(x)), 0\}$ for all $x \in X$; here index$(Df(x))$ stands for the index of the Fredholm operator $Df(x)$, that is, the difference between the dimension of the kernel of $Df(x)$ and the codimension of the image of $Df(x)$, both of which are finite. Of course, these assumptions are very restrictive as, for instance, if X is infinite-dimensional then no function $f : X \longrightarrow \mathbb{R}$ is Fredholm.

In general, every attempt to adapt the Morse-Sard theorem to infinite dimensions will have to impose vast restrictions because, as shown by Kupka's counterexample [116], there are C^∞ smooth functions $f : \ell_2 \longrightarrow \mathbb{R}$ so that their sets of critical values $f(C_f)$ contain intervals. Furthermore, as shown by Bates and Moreira in [32], one can take f to be a polynomial of degree 3.

Nevertheless, for many applications of the Morse-Sard theorem, it is often enough to know that any given continuous mapping can be uniformly approximated by a mapping whose set of critical values is small in some sense; therefore it is natural to ask what mappings between infinite-dimensional manifolds will at least have such an approximation property. Going in this direction, Eells and McAlpin established the following theorem [69]: If E is a separable Hilbert space, then every continuous function from E into $\mathbb{R}$ can be uniformly approximated by a smooth function f whose set of critical values $f(C_f)$ is of measure zero. This allowed them to deduce a version of this theorem for mappings between smooth manifolds M and N modelled on E and a Banach space F respectively, which they called an *approximate Morse-Sard theorem*: Every continuous mapping from M into N can be uniformly approximated by a smooth mapping $f : M \longrightarrow N$ so that $f(C_f)$ has empty interior. However, as observed in [69, Remark 3A], we have $C_f = M$ in the case that F is infinite-dimensional (so, even though the set of critical values of f is relatively small, the set of critical points of f is huge, which is somewhat disappointing).

In [12], a much stronger result was obtained by Daniel Azagra and Manuel Cepedello-Boiso: if M is a C^∞ smooth manifold modelled on a separable infinite-dimensional Hilbert space X, then every continuous mapping from M into $\mathbb{R}^m$ can be uniformly approximated by smooth mappings *with no critical points*. P. Hájek and M. Johanis [91] established a similar result for $m = 1$ in the case that X is a separable Banach space which contains c_0 and admits a C^p-smooth bump function. Finally, in the case that $m = 1$, these results were extended by Daniel Azagra and Mar Jiménez-Sevilla and the first-named author [25] for continuous functions $f : X \to \mathbb{R}$, where X is a separable Banach space admitting an equivalent smooth and locally uniformly rotund norm.

In what follows, we will improve these results by showing that the pairs $(\ell_2, \mathbb{R}^m)$ or $(X, \mathbb{R})$ can be replaced with pairs of the form (E, F), where E is a Banach space from a large class (including all the classical spaces with smooth norms such as c_0, ℓ_p or L^p, $1 < p < \infty$), and F can be taken to be any quotient space of E. So we may say that even though an exact Morse-Sard theorem for mappings between classical Banach spaces is false, a stronger approximate version of the Morse-Sard theorem is nonetheless true.

The general plan of the proof of Theorem 4.2 consists in following these steps:

- Step 1: We construct a smooth mapping $\varphi : E \to F$ such that $\|\varphi(x) - f(x)\| \le \varepsilon(x)/2$ and C_φ, the critical set of φ, is locally contained in the graph of a continuous mapping defined on a complemented subspace of infinite codimension in E and taking values in its linear complement.

- Step 2: We find a diffeomorphism $h : E \to E \setminus C_\varphi$ such that h is sufficiently close to the identity, in the sense that $\{\{x, h(x)\} : x \in E\}$ refines $\mathcal{G}$ (in other words, h is *limited by* $\mathcal{G}$), where $\mathcal{G}$ is an open cover of E by open balls $B(z,\delta_z)$ chosen in such a way that if $x,y\in B(z,\delta_z)$ then

$$\|\varphi(y)-\varphi(x)\|\leq\frac{\varepsilon(z)}{4}\leq\frac{\varepsilon(x)}{2}.$$

The existence of such a diffeomorphism h follows by the results of Chapter 3, namely Theorem 3.3.

- Step 3: Then, the mapping $g(x):=\varphi(h(x))$ has no critical point and satisfies $\|f(x)-g(x)\|\leq\varepsilon(x)$ for all $x\in E$.

The proof of Theorem 4.1 will show that E and F can be replaced with open subsets U and V of E and F respectively. Then, by combining such an equivalent statement of Theorem 4.1 with the well known result [67, 115] stating that every separable infinite-dimensional Hilbert manifold is diffeomorphic to an open subset of ℓ_2, one may easily deduce the following.

Theorem 4.3. *Let M, N be separable infinite-dimensional Hilbert manifolds. For every continuous mapping $f:M\to N$ and every open cover $\mathcal{U}$ of N, there exists a C^∞ mapping $g:M\to N$ such that g has no critical point and $\{\{f(x),g(x)\}:x\in M\}$ refines $\mathcal{U}$.*

Alternatively, one can also adjust the proof of Theorem 4.1 to obtain a direct proof of Theorem 4.3.

It is worth noting that Theorems 4.1 and 4.2 follow from the following more general (but also more technical) results. For spaces E which are reflexive and have a certain "composite" structure, we have the following.

Theorem 4.4. *Let E be a separable reflexive Banach space of infinite dimension, and F be a Banach space. In the case that F is infinite-dimensional, let us assume furthermore that:*

1. *E is isomorphic to $E\oplus E$.*
2. *There exists a linear bounded operator from E onto F (equivalently, F is a quotient space of E).*

Then, for every continuous mapping $f:E\to F$ and every continuous function $\varepsilon:E\to(0,\infty)$ there exists a C^1 mapping $g:E\to F$ such that $\|f(x)-g(x)\|\leq\varepsilon(x)$ and $Dg(x):E\to F$ is a surjective linear operator for every $x\in E$.

Note that there exists separable, reflexive Banach spaces E such that E is not isomorphic to $E\oplus E$. The first example of such a space was given by Figiel in 1972 [81].

See also Theorem 4.22 and Theorem 4.23 below for more general variants of this result.

For spaces which are not necessarily reflexive but have an appropriate Schauder basis we have the following.

Theorem 4.5. *Let E be an infinite-dimensional Banach space, and F be a Banach space such that:*

1. *E has an equivalent locally uniformly convex norm $\|\cdot\|$ which is C^1 smooth.*
2. *$E=(E,\|\cdot\|)$ has a (normalized) Schauder basis $\{e_n\}_{n\in\mathbb{N}}$ such that for every $x=\sum_{j=1}^\infty x_je_j$ and every $j_0\in\mathbb{N}$ we have that*

$$\left\|\sum_{j\in\mathbb{N},\,j\neq j_0}x_je_j\right\|\leq\left\|\sum_{j\in\mathbb{N}}x_je_j\right\|.$$

3. *In the case that F is infinite-dimensional, there exists a subset $\mathbb{P}$ of $\mathbb{N}$ such that both $\mathbb{P}$ and $\mathbb{N}\setminus\mathbb{P}$ are infinite and, for every infinite subset J of $\mathbb{P}$, there exists a linear bounded operator from $\overline{\operatorname{span}}\{e_j:j\in J\}$ onto F (equivalently, F is a quotient space of $\overline{\operatorname{span}}\{e_j:j\in J\}$).*

Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^1 mapping $g : E \to F$ such that $\|f(x) - g(x)\| \le \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

Recall that a norm $\|\cdot\|$ in a Banach space E is said to be locally uniformly convex (LUC) (or locally uniformly rotund (LUR)) provided that, for every sequence $(x_n) \subset E$ and every point x_0 in E, we have that

$$\lim_{n\to\infty} 2\left(\|x_0\|^2 + \|x_n\|^2\right) - \|x_0 + x_n\|^2 = 0 \implies \lim_{n\to\infty} \|x_n - x_0\| = 0.$$

The second condition in Theorem 4.5 is equivalent to the fact that, for every (equivalently, finite) set $A \subset \mathbb{N}$, $\|P_A\| \le 1$, where P_A stands for the projection $P_A(x) = \sum_{j\in A} x_j e_j$. This, in particular, implies that $\{e_n\}_{n\in N}$ is an unconditional basis with suppression unconditional constant equal to one, what we will call a 1-suppression undonctional basis; for more details see [4, p. 53] or [3].

The proofs of these theorems will be provided in Sections 4.2 and 4.3. These results combine to yield Theorem 4.2 for $k = 1$ (see also Remark 4.26 in Section 4.4 for an explanation of why the space c_0 satisfies the assumptions of Theorem 4.5). In order to deduce Theorem 4.2 in the cases of higher order smoothness, we just have to use Nicole Moulis's results on C^1-fine approximation in Banach spaces [128] or the more general results of [93, Corollary 7.96], together with the following fact.

Proposition 4.6. *Assume that the Banach spaces E, F satisfy the following properties:*

1. *For every continuous mapping $f : E \to F$ and every continuous function $\delta : E \to (0, \infty)$ there exists a C^1 smooth mapping $\varphi : E \to F$ such that $\|f(x) - \varphi(x)\| \le \delta(x)$ and $D\varphi(x) : E \to F$ is surjective for all $x \in E$.*

2. *For every C^1 mapping $\varphi : E \to F$ and every continuous function $\eta : E \to (0, \infty)$ there exists a C^k mapping $g : E \to F$ such that $\|f(x) - \varphi(x)\| \le \eta(x)$ and $\|D\varphi(x) - Dg(x)\| \le \eta(x)$ for all $x \in E$.*

Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^k smooth mapping $g : E \to F$ such that $\|f(x) - g(x)\| \le \varepsilon(x)$ and $Dg(x) : E \to F$ is surjective for every $x \in E$.

Nevertheless, it should be noted that our proof of Theorem 4.4 directly provides C^∞ approximations without critical points in the case that E is a separable Hilbert space; see Remark 4.24 in Section 4.4 below. An easy proof of Proposition 4.6, together with some examples, remarks and more technical variants of our results, is given in Section 4.4.

Finally, let us mention that as a straightforward application of Theorem 4.2, we obtain that, for all Banach spaces E and F appearing in Theorem 4.2, every continuous mapping $f : E \to F$ can be uniformly approximated by *open* mappings of class C^k. For a more general statement, see Remark 4.30. Obviously, the latter result is false in the case that E is finite-dimensional.

4.2 Approximation in reflexive Banach spaces isomorphic to their squares

Proof of Theorem 4.4. First of all notice that since the result we want to establish is invariant by diffeomorphisms, it is enough to prove it for a C^1 manifold M diffeomorphic to E in place of E. It will be very convenient for us to do so with $M = S^+$, the upper sphere of $E \times \mathbb{R}$.

Let $\|\cdot\|$ denote an equivalent norm in E which is LUR and C^1 (we will also denote by $\|\cdot\|$ the norm of F; this will do not do any harm because there will be no risk of confusion). Since E^* is separable there always exists such a norm; see [60, Corollary II.4.3] for instance.

Let us define $Y = E \times \mathbb{R}$, with norm

$$|(u,t)| = \left(\|u\|^2 + t^2\right)^{1/2},$$

and let us denote the upper sphere of Y by

$$S^+ := \{(u,t) : u \in E, t > 0, \|u\|^2 + t^2 = 1\}.$$

Observe that S^+ is the graph of

$$s(u) = \sqrt{1 - \|u\|^2},$$

which is a C^1 function[1] defined on the open unit ball B_E of E; hence,

$$d(u) = (u, s(u)) = (u, \sqrt{1 - \|u\|^2}),$$

$u \in B_E$, defines a C^1 diffeomorphism of B_E onto S^+. Since B_E is obviously C^1 diffeomorphic with E, the upper sphere S^+, which is a C^1 submanifold of codimension 1 of Y, is also diffeomorphic to E. Therefore it will be enough to prove that every continuous mapping $f : S^+ \to F$ can be ε-approximated by a mapping $\varphi : S^+ \to F$ which is of class C^1 and has no critical points. As explained in the introduction, this will be done in three steps, the first of which consists in finding a smooth approximation of f whose critical set is a set that we can extract with the help of Theorem 3.3.

In order to find such a smooth approximation, as in [25] we will have to use a partition of unity $\{\psi_n : n \in \mathbb{N}\}$ in S^+ made out of slices of the unit ball of Y by linear functionals $g_k \in Y^*$, so that the derivative at $y \in S^+$ of a local sum $\sum \psi_k$ will belong to the span of the restrictions to $T_y S^+$ (the tangent space to S^+ at y) of a finite collection of g_k. However, the construction of the partition of unity, the technical properties that we will require, and the use that we will make of it, will be much simpler than in that paper.

To construct our partition of unity $\{\psi_n\}_{n\in\mathbb{N}}$ we next translate an old standard argument (going back to Eells in the Hilbert space case, and probably first appearing in the pages 28–30 of the first edition of [117], later generalized by Bonic and Frampton [40] for separable Banach spaces with smooth bump functions; we follow [72, Theorem 8.25]) to the upper sphere, as in [147, 83, 25].

Let us denote $S := S_{|\cdot|}$, the unit sphere of $(Y, |\cdot|)$, and $S^* := S_{|\cdot|^*}$, the unit sphere of $(Y^*, |\cdot|^*)$. The duality mapping of the norm $|\cdot|$, defined as

$$D : S \longrightarrow S^*$$
$$D(x) = |\cdot|'(x),$$

is $|\cdot| - |\cdot|^*$ continuous since the norm $|\cdot|$ is of class C^1.

Since the norm $|\cdot|$ is locally uniformly convex we can find, for every $x \in S^+$, open slices $R_x = \{y \in S : g_x(y) > \delta_x\} \subset S^+$ and $P_x = \{y \in S : g_x(y) > \delta_x^2\} \subset S^+$, where $g_x = D(x) \in Y^*$, $0 < \delta_x < 1$, and $|g_x|^* = 1 = g_x(x)$, so that the oscillation of the functions f and ε on every P_x is less than $\varepsilon(x)/16$. We also assume, with no loss of generality, that dist$(P_x,\, E \times \{0\}) > 0$.

Since Y is separable we can select a countable subfamily of $\{R_x\}_{x\in S^+}$ which covers S^+. Let us denote this countable subfamily by $\{R_n\}_n$, where $R_n = R_{x_n} = \{y \in S : g_n(y) > \delta_n\}$ and $g_n(x_n) = 1$. Recall that the oscillation of the functions f and ε on every $P_n = P_{x_n} = \{y \in S : g_n(y) > \delta_n^2\}$ is less than $\varepsilon(x_n)/16$; this implies that

$$\frac{15}{16}\,\varepsilon(x_n) \le \varepsilon(x) \le \frac{17}{16}\,\varepsilon(x_n) \text{ and } \|f(x) - f(y)\| \le \frac{\varepsilon(x_n)}{16}$$

for every $x, y \in P_n$. Note also that $\{P_n\}_{n\in\mathbb{N}}$ is an open cover of S^+.

For each $k \in \mathbb{N}$, let $\theta_k : \mathbb{R} \to [0,1]$ be a C^∞ function such that $\theta_k(t) = 1$ if and only if $t \ge \delta_k$, and $\theta(t) = 0$ if and only if $t \le \delta_k^2$. Next, for each $k \in \mathbb{N}$ we define $\varphi_k : S^+ \to [0,1]$ by

$$\varphi_k(x) = \theta_k(g_k(x)),$$

[1] Smoothness of $s(u)$ at $u = 0$ is a consequence of the facts that $\|\cdot\|^2$ is trivially differentiable at 0, and that an everywhere differentiable convex function is always of class C^1; see for instance [41, Corollary 4.2.12].

and note that the interior of the support of φ_k, which coincides with $\varphi_k^{-1}((0,1])$, is the open slice $P_k = \{y \in S : g_k(x) > \delta_k^2\}$.
Now, for $k=1$, define $h_1 : S^+ \to \mathbb{R}$ by

$$h_1(x) = \varphi_1(x).$$

Notice that the interior of the support of h_1 is the open set $U_1 := P_1$.
For $k \geq 2$ let us define $h_k : S^+ \to \mathbb{R}$ by

$$h_k(x) = \varphi_k(x) \prod_{j<k} (1 - \varphi_j(x)),$$

and notice that the interior of the support of h_k is the set

$$U_k := \{y \in S^+ : g_k(y) > \delta_k^2 \text{ and } g_j(y) < \delta_j \text{ for all } j < k\}.$$

Claim 4.7. *The family $\{U_k\}_{k\in\mathbb{N}}$ is a locally finite open covering of S^+ that refines $\{P_k\}_{k\in\mathbb{N}}$. Therefore the functions*

$$\psi_n := \frac{h_n}{\sum_{k=1}^{\infty} h_k}, \quad n \in \mathbb{N},$$

define a C^1 partition of unity in S^+ subordinate to $\{P_k\}_{k\in\mathbb{N}}$.

Proof. Given $j \in \mathbb{N}$, if $x, y \in \{y : g_j(y) > \delta_j\}$ and $k > j$ then $\varphi_j(y) = 1$, hence $h_k(y) = 0$. Since $\{y : g_j(y) > \delta_j\}$ is a neighbourhood of x in S^+ this implies that the family of supports of the h_k is locally finite. On the other hand, if $h_k(x) > 0$ for some x, k then $\varphi_k(x) > 0$, hence $x \in P_k$, and this shows that the family of the open supports of the functions h_k, which coincides with $\{U_k\}_{k\in\mathbb{N}}$, refines $\{P_k\}_{k\in\mathbb{N}}$. It only remains to prove that the family of the open supports of the functions h_k is indeed a cover, that is, for every $x \in S^+$ there exists some n_x such that $h_{n_x}(x) > 0$. We argue by contradiction: assume we had $h_k(x) = 0$ for all $k \in \mathbb{N}$, then we can show by induction that $\varphi_n(x) = 0$ for all $n \in \mathbb{N}$, which implies that $x \notin P_n$ for all n and contradicts the fact that $\{P_n\}_{n\in\mathbb{N}}$ covers S^+. Indeed, for $n = 1$ we have $0 = h_1(x) = \varphi_1(x)$. Now suppose that we have $\varphi_1(x) = \varphi_2(x) = ... = \varphi_n(x) = 0$. Then

$$0 = h_{n+1}(x) = \varphi_{n+1}(x) \prod_{j<n+1} (1 - \varphi_j(x)) = \varphi_{n+1}(x) = 0,$$

so it is also true that $\varphi_{n+1}(x) = 0$. □

We will employ the following remarkable fact.

Claim 4.8. *For every $k \in \mathbb{N}$ and every $y \in S^+$, we have*

$$h_k(y) = 0 \Longrightarrow Dh_k(y) = 0.$$

Proof. Let us assume that $h_k(y) = 0$. Suppose first that $\varphi_k(y) = 0$. Computing the derivative of h_k at y we get that

$$Dh_k(y) = D\varphi_k(y) \prod_{j<k} (1 - \varphi_j(y)) + \varphi_k(y) D \left(\prod_{j<k} (1 - \varphi_j(y)) \right) = D\varphi_k(y) \prod_{j<k} (1 - \varphi_j(y)).$$

But $\varphi_k(y) = \theta_k(g_k(y)) = 0$ implies that $D\varphi_k(y) = 0$, hence $Dh_k(y) = 0$.
If $\varphi_k(y) \neq 0$ we must have $\varphi_j(y) = 1$ for some $j < k$. But again it follows that $D\varphi_j(y) = 0$, so one can check that also

$$Dh_k(y) = \varphi_k(y) D \left(\prod_{j<k} (1 - \varphi_j(y)) \right) = 0.$$

□

Notice that, according to Claim 4.7, for every $x \in S^+$ there exist a number $n = n_x$ and an open neighbourhood V_x of x in S^+ such that $\sum_{j\leq n} h_j(y) > 0$ for every $y \in V_x$, and $h_k(y) = 0$ for every $k > n$. This means that $\psi_k(y) = 0$ for every $k > n$ and every $y \in V_x$. More precisely, n_x can be chosen as the first such j so that $x \in \{y \in S^+ : g_j(y) > \delta_j\} = R_j$ and $V_x = R_j$.
Let us also call $m = m_y$ the largest j such that $h_j(y) \neq 0$. Note that m_y is also the largest j for which $\psi_j(y) \neq 0$. Thus, for every $y \in V_x$, we have

$$m_y = \max\{j : y \in \psi_j^{-1}((0,1])\} \leq n_x.$$

In order to calculate the derivatives of this partition of unity in this neighbourhood V_x of x, let us introduce the functions

$$H_k(y) = \frac{h_k(y)}{\sum_{j=1}^n h_j(y)}, \qquad y \in V_x, \;\; k = 1, ..., n,$$

which are well defined on V_x and in fact can be extended as C^1 smooth functions to an open subset $\mathcal{S}_n$ of Y containing V_x. Specifically, noting that each h_k is well-defined and C^∞ smooth on the whole Y, one can choose $\mathcal{S}_n = \bigcup_{i=1}^n h_i^{-1}((0,1])$. Slightly abusing notation, we will keep denoting these extensions by H_k, and we will also think of the functions h_j, ψ_j, g_j, $j \leq n$, as being C^1 smooth functions defined on this open set $\mathcal{S}_n$. Therefore, to calculate the derivative of ψ_k on V_x for $k = 1, ..., n$, we only have to calculate the derivative of H_k at each $y \in V_x$ for $k = 1, ..., n$ and then restrict it to the tangent spaces T_yS^+. The exact expression for the derivative of the functions $H_1, ..., H_n$ on V_x will not be particularly interesting or useful to us. The only thing we need to know is that there are C^1 smooth functions $\sigma_{k,j}$, $1 \leq k, j \leq n$, (actually, $\sigma_{k,j}$ will be C^∞ smooth) defined on $\mathcal{S}_n$ so that

$$DH_k(y) = \sum_{j=1}^n \sigma_{k,j}(y)\, g_j$$

for $k = 1, ..., n$, $y \in V_x$, and that, in fact, as an immediate consequence of Claim 4.8 and the definition of m_y we have

$$DH_k(y) = \sum_{j=1}^{m_y} \sigma_{k,j}(y)\, g_j.$$

for $k = 1, \ldots, m_y$, $y \in V_x$.
Note that even though each H_k is C^∞ smooth on an open subset $\mathcal{S}_n$ of Y, we cannot say that ψ_k is C^∞ smooth too, because S^+ has not a C^∞ smooth submanifold structure. The functions ψ_k are just C^1 because S^+ is just a C^1 manifold modelled on E. Now, if we want to know what the derivative of the functions ψ_k, $k = 1, ..., n$, on V_x looks like, because $H_k = \psi_k$ on V_x and V_x is open in S^+, we only have to restrict DH_k to the tangent spaces T_yS^+ for each $y \in V_x$. Hence we have

$$D\psi_k(y)(v) = \sum_{j=1}^n \sigma_{k,j}(y)\, g_j(v) = \sum_{j=1}^{m_y} \sigma_{k,j}(y)\, g_j(v)$$

for each $v \in T_yS^+$.
This expression can be somewhat misleading at first sight, because one might think that, for $k = 1, ..., n$, $D\psi_k(y)$ is just a linear combination of the functionals $g_1, ..., g_n$, and this is not exactly so. It is a linear combination of the restrictions of $g_1, ..., g_n$ to T_yS^+, and therefore for every $y \in V_x$ it is a different linear combination of different linear functionals $g_1|_{T_yS^+}, \ldots, g_n|_{T_yS^+}$, each of them defined on a space depending on y.
In order to fully clarify this important point, let us calculate the tangent space T_yS^+ at $y = (u_y, t_y)$. Since S^+ is the graph of the function $s(u) = \sqrt{1 - \|u\|^2}$, the most natural representation of T_yS^+ is given by

$$T_yS^+ = \{(u,t) \in Y = E \times \mathbb{R} \,:\, t = L_y(u), u \in E\} = \{(u, L_y(u)) : u \in E\} \subset Y, \qquad (4.2.1)$$

where L_y is the derivative $Ds(u_y)$ of the function s evaluated at the point $u_y = d^{-1}(y) \in B_E$ (recall that $d(u) = (u, s(u))$); in other words, if $y \neq (0,1)$,

$$L_y(w) = -\frac{\|u_y\|}{\sqrt{1-\|u_y\|^2}} D\|\cdot\|(u_y)(w) = -\frac{\sqrt{1-t_y^2}}{t_y} D\|\cdot\|(u_y)(w)$$

for each $w \in E$. Of course we have $L_{(0,1)}(w) = Ds(0)(w) = 0$ for each $w \in E$, and $T_{(0,1)}S^+ = E \times \{0\}$.
This is the vectorial tangent hyperplane to S^+ at y, as opposed to the affine tangent hyperplane to S^+, which is just $y + T_yS^+$. Since derivatives of mappings act on vectorial tangent hyperplanes we may forget the affine hyperplanes $y + T_yS^+$ in what follows.
Now, recall that $g_j \in Y^*$ and therefore these functionals are of the form

$$g_j(u,t) = g_j^1(u) + g_j^2 t, \qquad (u,t) \in E \times \mathbb{R} = Y,$$

where $g_j^1 \in E^*$ and $g_j^2 \in \mathbb{R}$.
Therefore the derivative of $g_{j|_{S^+}}$ at $y \in S^+$ is given by

$$Dg_j(y)(u, L_y(u)) = g_j(u, L_y(u)) = g_j^1(u) + g_j^2 L_y(u) \tag{4.2.2}$$

for every $v = (u, L_y(u)) \in T_yS^+$. Thus, for every $y \in V_x$ and every $k = 1, ..., n$, we have

$$D\psi_k(y)(v) = D\psi_k(y)(u, L_y(u)) = \sum_{j=1}^{n} \sigma_{k,j}(y)\,(g_j^1(u) + g_j^2 L_y(u)) = \sum_{j=1}^{m_y} \sigma_{k,j}(y)\,(g_j^1(u) + g_j^2 L_y(u))$$

for every $v = (u, L_y(u)) \in T_yS^+$.
Finally, let us note that the points $x_n \in R_n$ satisfy that

$$\frac{15}{16}\,\varepsilon(x_n) \leq \varepsilon(y) \leq \frac{17}{16}\,\varepsilon(x_n) \quad \text{for every } y \in P_n, \quad \text{and} \quad \sup_{x,y \in P_n} \|f(x) - f(y)\| \leq \frac{\varepsilon(x_n)}{16}.$$

The following lemma summarizes the properties of the partition of unity $\{\psi_n\}_{n\in\mathbb{N}}$ which will be most useful to us.

Lemma 4.9. *Given two continuous functions $f : S^+ \to F$ and $\varepsilon : S^+ \to (0,\infty)$, there exists a collection of norm-one linear functionals $\{g_k\}_{k\in\mathbb{N}} \subset Y^*$, an open covering $\{P_n\}_{n\in\mathbb{N}}$ of S^+, and a C^1 partition of unity $\{\psi_n\}_{n\in\mathbb{N}}$ on S^+ such that:*

1. *$\{\psi_n\}_{n\in\mathbb{N}}$ is subordinate to $\{P_n\}_{n\in\mathbb{N}}$.*

2. *For every $x \in S^+$ there exist a neighbourhood V_x of x in S^+ and a number $n = n_x \in \mathbb{N}$ such that $\psi_m = 0$ on V_x for all $m > n$, and the derivatives of the functions $\psi_1, ..., \psi_n$ on V_x are of the form*

$$D\psi_k(y)(v) = \sum_{j=1}^{n} \sigma_{k,j}(y)\, g_j(v) = \sum_{j=1}^{m_y} \sigma_{k,j}(y)\, g_j(v),$$

for $v \in T_yS^+$, the tangent hyperplane to S^+ at $y \in S^+ \cap V_x$, and where $m_y \leq n$ is the largest number such that $\psi_{m_y}(y) \neq 0$. More precisely, if L_y denotes the derivative of the function $s(u) = \sqrt{1-\|u\|^2}$ evaluated at the point $d^{-1}(y)$, where $d(u) = (u, s(u))$, we have

$$D\psi_k(y)(v) = D\psi_k(y)(u, L_y(u)) = \sum_{j=1}^{n} \sigma_{k,j}(y)\,(g_j^1(u)+g_j^2 L_y(u)) = \sum_{j=1}^{m_y} \sigma_{k,j}(y)\,(g_j^1(u)+g_j^2 L_y(u))$$

for every $k = 1, ..., n$, and for every $v = (u, L_y(u)) \in T_yS^+$, where the functions $\sigma_{k,j} : V_x \to \mathbb{R}$ are of class C^1, and $g_j^1 \in E^*$, $g_j^2 \in \mathbb{R}$, $j = 1, ..., n$.

3. *For every $n \in \mathbb{N}$ there exist a point $y_n := x_n \in P_n$ such that*

$$\frac{15}{16}\varepsilon(y_n) \leq \varepsilon(y) \leq \frac{17}{16}\varepsilon(y_n) \quad \text{for every } y \in P_n, \quad \text{and} \quad \sup_{x,y\in P_n} \|f(x) - f(y)\| \leq \frac{\varepsilon(y_n)}{16}.$$

Now we are ready to start the construction of our approximating function $\varphi : S^+ \to F$, which will be of the form

$$\varphi(x) = \sum_{n=1}^{\infty} (f(y_n) + T_n(x))\, \psi_n(x),$$

where the y_n are the points given by the third condition of the preceding lemma, and the operators $T_n : Y \to F$ will be carefully defined below.

Case 1: Assume that F is infinite-dimensional.
We will have to make repeated use of the following fact.

Lemma 4.10. *If E is a Banach space which is isomorphic to $E \oplus E$ then for every finite-codimensional closed subspace V of E, there exists a decomposition*

$$E = E_1 \oplus E_2 \oplus E_3,$$

with factors E_1, E_2, E_3 isomorphic to E, such that $E_1 \oplus E_2 \subset V$.

Proof. Since V is of finite codimension, say $\text{codim}(V) = n$, we can write $E = V \oplus \mathbb{R}^n$.
On the other hand we have

$$E = E_1' \oplus E_2'$$

with factors E_1', E_2' isomorphic to E. Considering $E_1' \oplus \{0\}$ we can find n linearly independent vectors x_j and n corresponding linear functionals x_j^* that separate $E_1' \oplus \{0\}$ from x_j respectively. We get

$$E_1' \subset \bigcap_{j=1}^{n} \text{Ker}(x_j^*)$$

and $\text{codim} \bigcap_{j=1}^n \text{Ker}(x_j^*) = n$, so we can write $E = \bigcap_{j=1}^n \text{Ker}(x_j^*) \oplus \mathbb{R}^n$.
Now, using the fact that every two closed subspaces with the same finite codimension are isomorphic, there exists an isomorphism

$$T : E = \bigcap_{j=1}^{n} \text{Ker}(x_j^*) \oplus \mathbb{R}^n \to V \oplus \mathbb{R}^n = E,$$

where $T(\bigcap_{j=1}^n \text{Ker}(x_j^*)) = V$. We have that $E_1 := T(E_1')$, $E_2 := T(E_2')$ are isomorphic to E and that $E_1 \subset V$. We can finally conclude that

$$E = T(E) = T(E_1' \oplus E_2') = T(E_1') \oplus T(E_2') = E_1 \oplus E_2$$

with factors isomorphic to E and such that $E_1 \subset V$. □

We start considering a decomposition

$$E = E_{1,1} \oplus E_{1,2},$$

with infinite-dimensional factors isomorphic to E, and we define a continuous linear surjection $S_1 : E \to F$ such that $S_1 = 0$ on $E_{1,2}$. This can be done by taking a continuous linear surjection $R_1 : E_{1,1} \to F$ (which exists because by assumption there exists such an operator from E onto F and $E_{1,1}$ is isomorphic to E), and setting $S_1 = R_1 \circ P_{1,1}$, where $P_{1,1} : E_{1,1} \oplus E_{1,2} \to E_{1,1}$ is the projection onto the first factor associated to this decomposition of E. Next, recall that the linear functionals $g_n \in Y^* = (E \times \mathbb{R})^*$ are of the form

$$g_j(u,t) = g_j^1(u) + g_j^2 t,$$

where $g_j^1 \in E^*$ and $g_j^2 \in \mathbb{R}$. Of course, $\bigcap_{j=1}^2 \text{Ker } g_j^1$ is a finite-codimensional subspace of E, so by using Lemma 4.10 with $V = E_{1,2} \cap \bigcap_{j=1}^2 \text{Ker}(g_j^1)$, and with $E_{1,2}$ in place of E, we may find a decomposition of the second factor $E_{1,2}$,

$$E_{1,2} = E_{2,1} \oplus E_{2,2} \oplus E_{2,3},$$

with factors $E_{2,1}$, $E_{2,2}$ and $E_{2,3}$ isomorphic to E, and

$$E_{2,1} \oplus E_{2,2} \subset \bigcap_{j=1}^{2} \text{Ker } g_j^1,$$

and we may easily define a bounded linear operator S_2 from E onto F such that $S_2 = 0$ on $E_{1,1} \oplus E_{2,2} \oplus E_{2,3}$ (this can be done by taking a surjective operator $R_2 : E_{2,1} \to F$ and defining $S_2 = R_2 \circ P_{2,1} \circ P_{1,2}$, where $P_{1,2} : E_{1,1} \oplus E_{1,2} \to E_{1,2}$ and $P_{2,1} : E_{2,1} \oplus E_{2,2} \oplus E_{2,3} \to E_{2,1}$ are the projections associated to the corresponding decompositions).
We continue this process by induction: assuming that we have already defined decompositions $E = E_{1,1} \oplus E_{1,2}$, $E_{1,2} = E_{2,1} \oplus E_{2,2} \oplus E_{2,3}$, $E_{2,3} = E_{3,1} \oplus E_{3,2} \oplus E_{3,3}$, ..., $E_{n-1,3} = E_{n,1} \oplus E_{n,2} \oplus E_{n,3}$, and surjective operators

$$S_k : E = E_{1,1} \oplus (E_{2,1} \oplus E_{2,2}) \oplus ... \oplus (E_{k-1,1} \oplus E_{k-1,2}) \oplus (E_{k,1} \oplus E_{k,2} \oplus E_{k,3}) \to F, \;\; k = 2, ..., n,$$

so that S_k is zero on all the factors of this decomposition except $E_{k,1}$, and

$$E_{k,1} \oplus E_{k,2} \subset \bigcap_{j=1}^{k} \text{Ker}(g_j^1),$$

we again apply Lemma 4.10 to write

$$E_{n,3} = E_{n+1,1} \oplus E_{n+1,2} \oplus E_{n+1,3},$$

with factors isomorphic to E and

$$E_{n+1,1} \oplus E_{n+1,2} \subset \bigcap_{j=1}^{n+1} \text{Ker}(g_j^1),$$

and we define a continuous linear surjection

$$S_{n+1} : E = E_{1,1} \oplus (E_{2,1} \oplus E_{2,2}) \oplus ... \oplus (E_{n,1} \oplus E_{n,2}) \oplus (E_{n+1,1} \oplus E_{n+1,2} \oplus E_{n+1,3}) \to F,$$

by setting it equal to 0 on all the factors of this decomposition except $E_{n+1,1}$, which is mapped onto F.
Having this collection of surjective operators $S_n : E \to F$ at our disposal, we finally define $T_n : Y \to F$ by

$$T_n(u,t) = \frac{\varepsilon(y_n)}{4\|S_n\|} S_n(u),$$

and $\varphi : S^+ \to F$ by

$$\varphi(x) = \sum_{n=1}^{\infty} (f(y_n) + T_n(x)) \, \psi_n(x).$$

It is clear that φ is well defined and of class C^1.
In the rest of the proof we will check that this mapping ε-approximates f on S^+, and that the set of critical points of φ is a set which can be diffeomorphically extracted by using Theorem 3.3. Then the proof will be completed by setting $g = \varphi \circ h$, where $h : S^+ \to S^+ \setminus C_\varphi$ is a C^1 diffeomorphism which is close enough to the identity.

Claim 4.11. *The mapping φ approximates f.*

Proof. By condition 3. of Lemma 4.9, we know that the oscillation of f in P_n is less that $\varepsilon(y_n)/16$, and by definition of T_n, we have $\|T_n\| \le \varepsilon(y_n)/4$, hence $\|T_n(x)\| \le \varepsilon(y_n)/4$ for all $x \in S^+$ too, because $\|x\| = 1$. Now, if $\psi_n(x) \neq 0$, then $x \in P_n$ and

$$\begin{aligned}\|f(y_n) + T_n(x) - f(x)\| &\le \|f(y_n) - f(x)\| + \|T_n(x)\| \\ &\le \varepsilon(y_n)/16 + \varepsilon(y_n)/4 = \frac{5}{16}\varepsilon(y_n) < \frac{15}{32}\varepsilon(y_n) \le \varepsilon(x)/2.\end{aligned} \tag{4.2.3}$$

Therefore

$$\|\varphi(x) - f(x)\| = \Big\|\sum_{n=1}^{\infty}(f(y_n) + T_n(x) - f(x))\,\psi_n(x)\Big\| \le (\varepsilon(x)/2)\sum_{n=1}^{\infty}\psi_n(x) \le \varepsilon(x)/2.$$

□

Let us now consider the question as to how big the critical set C_φ can be. We need to calculate the derivative of our function φ. To this end we first have to examine the expressions for the derivatives of the operators T_n restricted to S^+. These are simpler than those of the g_n's on S^+, because of the way the operators $T_n : Y \to F$ have been defined. Indeed, for every $v = (u, L_y(u)) \in T_yS^+$ we have that

$$T_n(v) = T_n(u, L_y(u)) = \frac{\varepsilon(y_n)}{4\|S_n\|}\, S_n(u), \tag{4.2.4}$$

and we have that $D(T_{n|_{S^+}})(y)$ is the restriction of $DT_n(y) = T_n$ to T_yS^+, that is to say, if $v = (u, L_y(u)) \in T_yS^+$ then

$$DT_n(y)(u, L_y(u)) = \frac{\varepsilon(y_n)}{4\|S_n\|}\, S_n(u). \tag{4.2.5}$$

Now, recall that, by the second condition of Lemma 4.9, for every $x \in S^+$ there is a neighbourhood V_x of x in S^+ and a number $n = n_x \in \mathbb{N}$ such that

$$\varphi(y) = \sum_{j=1}^{n}(f(y_j) + T_j(y))\,\psi_j(y)$$

for every $y \in V_x$. Fix $y \in V_x$ and recall that for $m = m_y$, the largest number j for which $h_j(y) \neq 0$ (or $\psi_j(y) \neq 0$), we have

$$\varphi(y) = \sum_{j=1}^{m}(f(y_j) + T_j(y))\,\psi_j(y).$$

By using (4.2.5) and the expression for $D\psi_j(y)$ given in Lemma 4.9, we easily see that

$$\begin{aligned}D\varphi(y)(u, L_y(u)) &= \sum_{j=1}^{n}\psi_j(y)\,\frac{\varepsilon(y_j)}{4\|S_j\|}\, S_j(u) + \sum_{j=1}^{n}\big(g_j^1(u) + g_j^2 L_y(u)\big)\,\alpha_{n,j}(y) \\ &= \sum_{j=1}^{m}\psi_j(y)\,\frac{\varepsilon(y_j)}{4\|S_j\|}\, S_j(u) + \sum_{j=1}^{m}\big(g_j^1(u) + g_j^2 L_y(u)\big)\,\alpha_{m,j}(y)\end{aligned} \tag{4.2.6}$$

for every $(u, L_y(u)) \in T_yS^+$, $y \in V_x$, where the functions $\alpha_{n,j} : V_x \subset S^+ \to F$ are of class C^1 because $\alpha_{n,j}(y) = \sum_{i=1}^{n}(f(y_i) + T_i(y))\sigma_{i,j}(y)$.

Let us now show that the critical set of φ is relatively small.

When $n_x = 1$ we have $\varphi(y) = f(y_1) + T_1(y)$ on V_x, so $D\varphi(y)$ is the restriction of T_1 to T_yS^+, and equation (4.2.4) for $n = 1$ implies that this restriction is a surjective operator. So it is clear that φ has no critical point on V_x.

Claim 4.12. *If $n_x \geq 2$ then $C_\varphi \cap V_x$ is contained in the set*

$$A_x := \left\{ y \in S^+ : E_{n,2} \subset \text{Ker } L_y \right\}.$$

Recall that $L_y = Ds(u_y)$, where $s(u) = \sqrt{1 - \|u\|^2}$, $d(u) = (u, s(u))$, and $u_y = d^{-1}(y)$.

Proof. Let us see that, if $y \in V_x \setminus A_x$ then $D\varphi(y) : T_y S^+ \to F$ is surjective, that is, for every $w \in F$ there exists $v \in T_y S^+$ such that $D\varphi(y)(v) = w$. Let $m = m_y$ be the largest number such that $\psi_m(y) \neq 0$. Recall that $m \leq n$. Since the operator

$$S_m : E = E_{1,1}(\oplus E_{2,1} \oplus E_{2,2}) \oplus ... \oplus (E_{m-1,1} \oplus E_{m-1,2}) \oplus (E_{m,1} \oplus E_{m,2} \oplus E_{m,3}) \to F$$

is surjective and equal to zero on all the factors of this decomposition except $E_{m,1}$ (which is mapped onto F), we may find $u_{m,1} \in E_{m,1}$ so that

$$S_m(u_{m,1}) = 4\varepsilon(y_m)^{-1} \psi_m(y)^{-1} \|S_m\| \, w.$$

Now, since $y \notin A_x$ there exists some $e_{n,2} \in E_{n,2} \setminus \text{Ker } L_y$, that is to say, $L_y(e_{n,2}) \neq 0$, and this implies that, if we put

$$t_0 := -\frac{L_y(u_{m,1})}{L_y(e_{n,2})},$$

then the vector

$$u := u_{m,1} + t_0 e_{n,2},$$

satisfies that

$$L_y(u) = 0.$$

But recall that, for every $k \leq n$, we have $E_{k,1} \oplus E_{k,2} \subset \bigcap_{j=1}^k \text{Ker}(g_j^1)$; in particular, $E_{m,1} \oplus E_{m,2} \subset \bigcap_{j=1}^m \text{Ker}(g_j^1)$ because $m \leq n$. Hence, $g_j^1(u) = 0$ for every $1 \leq j \leq m$. It follows that

$$\sum_{j=1}^m \left(g_j^1(u) + g_j^2 L_y(u) \right) \alpha_{m,j}(y) = 0.$$

The rest of the operators $S_1, ..., S_{m-1}$ are zero on $E_{m,1} \oplus E_{n,2}$, so we have

$$S_j(u) = 0 \text{ for every } j = 1, ..., m-1,$$

and since S_m is zero on $E_{n,2} \subset E_{m,3}$ we also have $S_m(t_0 e_{n,2}) = 0$. Therefore, by combining these equalities with equation (4.2.6), we obtain that

$$D\varphi(y)(u, L_y(u)) = w,$$

and the proof of the claim is complete. □

Lemma 4.13. *For $x \in S^+$ with $n := n_x \geq 2$, the set A_x of Claim 4.12 is of the form*

$$A_x = d(G(f_x) \cap B_E),$$

where $G(f_x)$ is the graph of a continuous mapping $f_x : E_{1,1} \oplus (E_{2,1} \oplus E_{2,2}) \oplus \cdots \oplus (E_{n,1} \oplus E_{n,3}) \to E_{n,2}$.

Proof. Note that, by Claim 4.12,

$$A_x = \{d(u) = \left(u, \sqrt{1 - \|u\|^2}\right) : \|u\| < 1, u \in \mathcal{A}_x\},$$

where

$$\mathcal{A}_x := \bigcap_{e \in E_{n,2}} \{u \in E \setminus \{0\} \, : \, \langle D\| \cdot \|(u), e \rangle = 0\} \cup \{0\}.$$

Let us denote $E'_{n,2} = E_{1,1} \oplus (E_{2,1} \oplus E_{2,2}) \oplus \cdots \oplus (E_{n,1} \oplus E_{n,3})$, and let us see that there exists a mapping $f_x : E'_{n,2} \to E_{n,2}$ such that $\mathcal{A}_x = G(f_x) = \{w + f_x(w) : w \in E'_{n,2}\}$.
Pick a point $w \in E'_{n,2}$. Note that the function $E_{n,2} \ni v \mapsto \psi_w(v) := \|w+v\|^2$ is convex and continuous, and satisfies $\lim_{\|v\|\to\infty} \psi_w(v) = \infty$, hence, since $E_{n,2}$ is reflexive, ψ_w attains a minimum at some point $v_w \in E_{n,2}$; in fact this minimum point v_w is unique because the norm $\|\cdot\|$ is strictly convex. Let us denote

$$f_x(w) := v_w.$$

Note that the critical points of ψ_w, with $w \neq 0$, are exactly the points $v \in E_{n,2}$ such that

$$\frac{d}{dt}\|w + v + te\|^2 \,|_{t=0} = 0 \text{ for every } e \in E_{n,2}$$

or equivalently

$$\|w + v\| \,\langle D\| \cdot \|(w + v), e\rangle = 0 \text{ for every } e \in E_{n,2},$$

which in turn is equivalent to saying that $w + v \in \mathcal{A}_x$; we let $f_x(0) = v_0 = 0$.
Therefore the unique point $v \in E_{n,2}$ so that $w + v \in \mathcal{A}_x$ is the point $v = f_x(w)$. This shows that $\mathcal{A}_x$ is the graph of the function f_x.
Now let us see that the function $f_x : E'_{n,2} \to E_{n,2}$ is continuous. Suppose f_x is discontinuous at w_0 and let $v_0 := f_x(w_0)$. Then there exist sequences $w_k \to w_0$ in $E'_{n,2}$ and $v_k := f_x(w_k)$ in $E_{n,2}$ and a number $\varepsilon_0 > 0$ so that

$$\|v_k - v_0\| \geq \varepsilon_0 \text{ for all } k \in \mathbb{N}. \tag{4.2.7}$$

From the previous argument we know that the point v_k is characterized as being the unique point $v_k \in E_{n,2}$ for which we have

$$\|w_k + v_k\| \leq \|w_k + v_k + e\| \text{ for all } e \in E_{n,2}, \tag{4.2.8}$$

and similarly v_0 is the unique point $v_0 \in E_{n,2}$ for which

$$\|w_0 + v_0\| \leq \|w_0 + v_0 + e\| \text{ for all } e \in E_{n,2}. \tag{4.2.9}$$

By taking $e = -v_k$ in (4.2.8) we learn that

$$\|v_k\| - \|w_k\| \leq \|w_k + v_k\| \leq \|w_k\|,$$

hence $\|v_k\| \leq 2\|w_k\|$, and because $\|w_k\|$ converges to $\|w_0\|$ we deduce that (v_k) is bounded. Since $E_{n,2}$ is reflexive, this implies that (v_k) has a subsequence that weakly converges to a point $\xi_0 \in E_{n,2}$. We keep denoting this subsequence by (v_k).
Now, if we take $e = -v_k + e'$ in (4.2.8), with $e' \in E_{n,2}$, we obtain

$$\|w_k + v_k\| \leq \|w_k + e'\| \text{ for all } e' \in E_{n,2}.$$

This implies (using the facts that $v_k \rightharpoonup \xi_0$ and $w_k \to w_0$, and the weak lower semicontinuity of the norm) that

$$\|w_0 + \xi_0\| \leq \liminf_{k\to\infty} \|w_k + v_k\| \leq \liminf_{k\to\infty} \|w_k + e'\| = \|w_0 + e'\| \text{ for all } e' \in E_{n,2}. \tag{4.2.10}$$

That is, we have shown that

$$\|w_0 + \xi_0\| \leq \|w_0 + e'\| \text{ for all } e' \in E_{n,2}. \tag{4.2.11}$$

By taking $e' = \xi_0 + \xi$ with $\xi \in E_{n,2}$ we conclude that

$$\|w_0 + \xi_0\| \leq \|w_0 + \xi_0 + \xi\| \text{ for all } \xi \in E_{n,2}.$$

According to (4.2.9), v_0 is the only point which can satisfy this inequality. Hence $\xi_0 = v_0$.
But (4.2.10) tells us even more: by taking $e' = \xi_0$ we also learn that there exists a subsequence (w_{k_j}) of (w_k) such that

$$\|w_{k_j} + v_{k_j}\| \to \|w_0 + \xi_0\|.$$

Since we also know that $w_{k_j} + v_{k_j}$ converges to $w_0 + \xi_0$ weakly and the norm $\|\cdot\|$ is locally uniformly convex (hence $\|\cdot\|$ has the Kadec-Klee property), this implies that $w_{k_j} + v_{k_j}$ converges to $w_0 + \xi_0 = w_0 + v_0$ in the norm topology as well. As we also have $\lim_{j\to\infty} w_{k_j} = w_0$ in norm, we deduce that $\lim_{j\to\infty} \|v_{k_j} - v_0\| = 0$, which contradicts (4.2.7). □

Now we can easily finish the proof of Theorem 4.4. By Claim 4.12 and Lemma 4.13, we see that C_φ is a diffeomorphic image in S^+ of a relatively closed set Z of the open unit ball B_E of E which has the property of being locally contained in the graph of a continuous function defined on a complemented subspace of infinite codimension in E. Indeed, let $Z := d^{-1}(C_\varphi) \subset B_E$. Since C_φ is closed in S^+, Z is relatively closed in B_E. Also if we take $z \in Z$ then, according to Lemma 4.13 applied to $x = d(z) \in S^+$, $n = n_x$, and V_x, for a neighbourhood $U_z := d^{-1}(V_x)$ of z, we have $Z \cap U_z \subseteq G(f_x)$, where

$$G(f_x) = \{u = (w, v) \in E'_{n,2} \oplus E_{n,2} = E : v = f_x(w)\}.$$

Observe that E has C^1 smooth partitions of unity since E has a separable dual. Therefore we may apply Theorem 3.3 to find a C^1 diffeomorphism which extracts C_φ from S^+; more precisely, there exists a diffeomorphism $h : S^+ \to S^+ \setminus C_\varphi$ which, in addition, is limited by the open cover $\mathcal{G}$ that we next define. Recall that we have

$$\|\varphi(x) - f(x)\| \leq \varepsilon(x)/2 \tag{4.2.12}$$

for all $x \in S^+$. Since φ and ε are continuous, for every $z \in S^+$ there exists $\delta_z > 0$ so that if $x, y \in B(z, \delta_z)$ then $\|\varphi(y) - \varphi(x)\| \leq \varepsilon(z)/4 \leq \varepsilon(x)/2$. We set $\mathcal{G} = \{B(x, \delta_x) : x \in S^+\}$.
Finally, let us define

$$g = \varphi \circ h.$$

Since h is limited by $\mathcal{G}$ we have that, for any given $x \in S^+$, there exists $z \in S^+$ such that $x, h(x) \in B(z, \delta_z)$, and therefore $\|\varphi(h(x)) - \varphi(x)\| \leq \varepsilon(z)/4$, that is, we have that

$$\|g(x) - \varphi(x)\| \leq \varepsilon(z)/4 \leq \varepsilon(x)/2.$$

By combining this inequality with (4.2.12), we obtain that

$$\|g(x) - f(x)\| \leq \varepsilon(x)$$

for all $x \in S^+$. Besides, it is clear that g does not have any critical point: since $h(x) \notin C_\varphi$, we have that the linear map $D\varphi(h(x)) : T_{h(x)}S^+ \to F$ is surjective, and $Dh(x) : T_xS^+ \to T_{h(x)}S^+$ is a linear isomorphism, so $Dg(x) = D\varphi(h(x)) \circ Dh(x)$ is a linear surjection from T_xS^+ onto F for every $x \in S^+$.

Case 2: Assume that $F = \mathbb{R}^m$. The main idea of the proof is very similar to that of Case 1. The fact that F is finite dimensional will allow us dispense with the hypothesis that $E = E \oplus E$. We will use the same partition of unity $\{\psi_n\}_{n\in\mathbb{N}}$ provided by Lemma 4.9. We will decompose E inductively as follows. Since $\operatorname{Ker} g_1^1$ has infinite dimension we can write

$$E = E_1 \oplus G_1,$$

where $E_1 = \mathbb{R}^m$ and $G_1 \subseteq \operatorname{Ker} g_1^1$. Then $G_1 \cap \bigcap_{j=1}^2 \operatorname{Ker} g_j^1$ has codimension 0 or 1 in G_1, which is infinite-dimensional, and we can write

$$E = E_1 \oplus (E_{2,1} \oplus E_{2,2} \oplus G_2),$$

where $E_{2,1} = \mathbb{R}^m$, $E_{2,2} = \{0\}$ or $E_{2,2} = \mathbb{R}$, $G_1 = E_{2,1} \oplus E_{2,2} \oplus G_2$ for some G_2 with $\dim G_2 = \infty$, and

$$E_{2,1} \oplus G_2 = \bigcap_{j=1}^{2} \operatorname{Ker} g_j^1 \cap G_1 \subseteq \bigcap_{j=1}^{2} \operatorname{Ker} g_j^1.$$

Inductively, we can write

$$E = E_1 \oplus (E_{2,1} \oplus E_{2,2}) \oplus \cdots \oplus (E_{n-1,1} \oplus E_{n-1,2}) \oplus (E_{n,1} \oplus E_{n,2} \oplus G_n), \tag{4.2.13}$$

where $E_1, E_{2,1} \ldots, E_{n,1} = \mathbb{R}^m$, $E_{2,2}, \ldots, E_{n,2}$ are subspaces of dimension 0 or 1,

$$E_{n,1} \oplus G_n \subseteq \bigcap_{j=1}^{n} \operatorname{Ker} g_j^1,$$

and

$$G_k = (E_{k+1,1} \oplus E_{k+1,2} \oplus G_{k+1})$$

for every $k = 1, \ldots, n$.
Now, for each $n \in \mathbb{N}$, we define a continuous linear surjection $S_n : E \to F$ by setting it to be 0 on all the factors of the decomposition (4.2.13) except on $E_{n,1}$, which is mapped onto $F = \mathbb{R}^m$, and we construct our approximating function φ exactly as in the proof of Theorem 1.6. At this point, we only need to show the following variant of Claim 4.12 (in which G_n replaces the subspace $E_{n,2}$ of the previous proof).

Claim 4.14. *If $n_x \geq 2$ then $C_\varphi \cap V_x$ is contained in the set*

$$A_x := \left\{ y \in S^+ : G_n \subset \operatorname{Ker} L_y \right\}.$$

Recall that $L_y = Ds(u_y)$, where $s(u) = \sqrt{1 - \|u\|^2}$, $d(u) = (u, s(u))$, and $u_y = d^{-1}(y)$.

Proof. Let us see that, if $y \in V_x \setminus A_x$ then $D\varphi(y) : T_y S^+ \to F$ is surjective, that is, for every $w \in F$ there exists $v \in T_y S^+$ such that $D\varphi(y)(v) = w$. Let $m = m_y$ be the largest number such that $\psi_m(y) \neq 0$. Recall that $m \leq n$. Since the operator

$$S_m : E = E_1 \oplus (E_{2,1} \oplus E_{2,2}) \oplus ... \oplus (E_{m-1,1} \oplus E_{m-1,2}) \oplus (E_{m,1} \oplus E_{m,2} \oplus G_m) \to F$$

is surjective and equal to zero on all the factors of the decomposition (4.2.13) except on $E_{m,1}$ (which is mapped onto F), we may find $u_{m,1} \in E_{m,1}$ so that

$$S_m(u_{m,1}) = 4\varepsilon(y_m)^{-1}\psi_m(y)^{-1}\|S_m\|\, w.$$

Now, since $y \notin A_x$ there exists $e_n \in G_n \setminus \operatorname{Ker} L_y$. If we set

$$t_0 := -\frac{L_y(u_{m,1})}{L_y(e_n)},$$

then the vector

$$u := u_{m,1} + t_0 e_n,$$

satisfies that

$$L_y(u) = 0.$$

But recall that, for every $k \leq n$, we have $E_{k,1} \oplus G_k \subset \bigcap_{j=1}^{k} \operatorname{Ker}(g_j^1)$; in particular, $E_{m,1} \oplus G_m \subset \bigcap_{j=1}^{m} \operatorname{Ker}(g_j^1)$ because $m \leq n$. Hence, $g_j^1(u) = 0$ for every $1 \leq j \leq m$. It follows that

$$\sum_{j=1}^{m} \left(g_j^1(u) + g_j^2 L_y(u)\right) \alpha_{m,j}(y) = 0.$$

The rest of the operators $S_1, ..., S_{m-1}$ are zero on $E_{m,1} \oplus G_n$, so we have

$$S_j(u) = 0 \text{ for every } j = 1, ..., m-1,$$

and since S_m is zero on $G_n \subset G_m$ we also have $S_m(t_0 e_n) = 0$. Therefore, by combining these equalities with equation (4.2.6), we obtain that

$$D\varphi(y)(u, L_y(u)) = w,$$

and the proof of the claim is complete. □

Then we also have the following.

Lemma 4.15. *For $x \in S^+$ with $n := n_x \geq 2$, the set A_x of Claim 4.14 is of the form*

$$A_x = d(G(f_x) \cap B_E),$$

where $G(f_x)$ is the graph of a continuous mapping $f_x : E_{1,1} \oplus (E_{2,1} \oplus E_{2,2}) \oplus \cdots \oplus (E_{n,1} \oplus E_{n,2}) \to G_n$.

Proof. Repeat the proof Lemma 4.13, just replacing $E_{n,2}$ with G_n. □

Let $Z := d^{-1}(C_\varphi) \subset B_E$. According to Lemma 4.15 applied to $x = d(z) \in S^+$, $n = n_x$, and V_x, for a neighbourhood $U_z := d^{-1}(V_x)$ of z, we have $Z \cap U_z \subseteq G(f_x)$, where

$$G(f_x) = \{u = (w, v) \in G'_n \oplus G_n = E : v = f_x(w)\},$$

with G'_n denoting $E_1 \oplus (E_{2,1} \oplus E_{2,2}) \oplus \cdots \oplus (E_{n-1,1} \oplus E_{n-1,2}) \oplus (E_{n,1} \oplus E_{n,2})$. Since G_n is infinite-dimensional, we may use Theorem 3.3, and the rest of the proof goes exactly as in Case 1. □

Remark 4.16. Observe that in the infinite-dimensional case we could have asked A_x to be

$$A_x := \{y \in S^+ : E_{n,3} \subset \operatorname{Ker} L_y\},$$

using $E_{n,3}$ instead of $E_{n,2}$ and requiring that $E_{n,1} \oplus E_{n,3} \subseteq \bigcap_{j=1}^n \operatorname{Ker} g_j^1$.

4.3 Approximation in Banach spaces with 1-suppression unconditional basis

Proof of Theorem 4.5. First of all let us note that since E admits a C^1 equivalent norm, E cannot contain a closed subspace isomorphic to ℓ_1. Furthermore, as noted following the statement of Theorem 4.5, the second condition implies that the basis $\{e_n\}_{n\in\mathbb{N}}$ is unconditional, and therefore by [119, Theorem 1.c.9], is *shrinking*, that is, we have that $E^* = \overline{\operatorname{span}}\{e_n^* : n \in \mathbb{N}\}$, where $\{e_n^*\}_{n\in\mathbb{N}}$ are the biorthogonal functionals associated to $\{e_n\}_{n\in\mathbb{N}}$, that is to say, $e_n^* : E \to \mathbb{R}$ is defined by $e_n^*(e_k) = 1$ if $k = 1$ and $e_n^*(e_k) = 0$ otherwise. In particular, E^* is separable.
We keep using the notations $Y = E \times \mathbb{R}$ and S^+ from the proof of Theorem 4.4. As in the case of Theorem 4.4, it will be enough to prove Theorem 4.5 with S^+ in place of E. We define $e_0 = (0, 1) \in Y$, $e_0^* : Y = E \times \mathbb{R} \to \mathbb{R}$ by

$$e_0^*(u, t) = t,$$

and by slightly abusing notation we identify $e_n \in E$ to $(e_n, 0) \in Y$ and also extend the $e_n^* \in E^*$ to $e_n^* : Y = E \times \mathbb{R} \to \mathbb{R}$ by

$$e_n^*(u, t) = e_n^*(u) = u_n \quad \text{for all} \quad u = \sum_{j=1}^{\infty} u_j e_j, \in E,\ t \in \mathbb{R}.$$

Then we may consider $\{e_n\}_{n\in\mathbb{N}} \cup \{e_0\}$ as a basis of $E \times \mathbb{R} = Y$ with associated coordinate functionals $\{e_n^*\}_{n\in\mathbb{N}} \cup \{e_0^*\}$, and we have that this basis is also shrinking.
Next we are going to construct a partition of unity in S^+, quite similar but not identical to that of the proof of Theorem 4.4.
Since the norm $|\cdot|$ is locally uniformly convex we can find, for every $x \in S^+$, open slices $R_x = \{y \in S : f_x(y) > \delta_x\} \subset S^+$ and $P_x = \{y \in S : f_x(y) > \delta_x^4\} \subset S^+$, where $f_x \in Y^*$, $0 < \delta_x < 1$, and $|f_x|^* = 1 = f_x(x)$, so that the oscillation of the functions f and ε on every P_x is less than $\varepsilon(x)/16$. We also assume, with no loss of generality, that $\mathrm{dist}(P_x,\, E \times \{0\}\,) > 0$.
Since Y is separable we can select a countable subfamily of $\{R_x\}_{x\in S^+}$, which covers S^+. Let us denote this countable subfamily by $\{R_n\}_n$, where $R_n = R_{x_n} = \{y \in S : f_n(y) > \delta_n\}$ and $f_n(x_n) = 1$. Recall that the oscillation of the functions f and ε on every $P_n = P_{x_n} = \{y \in S : f_n(y) > \delta_n^4\}$ is less than $\varepsilon(x_n)/16$, and this implies that

$$\frac{15}{16}\,\varepsilon(x_n) \le \varepsilon(x) \le \frac{17}{16}\,\varepsilon(x_n) \ \text{ and } \ \|f(x) - f(y)\| \le \frac{\varepsilon(x_n)}{16}$$

for every $x, y \in P_n$. Note that $\{P_n\}_{n\in\mathbb{N}}$ is an open cover of S^+.
• **For $\mathbf{k = 1}$**, since $\mathrm{span}\{e_n^* : n \in \mathbb{N}\}$ is dense in E^*, we may find numbers $N_1 \in \mathbb{N}$, $\epsilon_1, \gamma_1 \in (0,1)$ with $\epsilon_1 > \gamma_1$, and $\beta_{1,0}, ..., \beta_{1,N_1} \in \mathbb{R}$ with $\beta_{1,0} > 0$ so that the functional g_1 defined by

$$g_1 := \sum_{j=0}^{N_1} \beta_{1,j} e_j^*$$

has norm 1 and satisfies

$$\{x \in S : f_1(x) > \delta_1^2\} \subset \{x \in S : g_1(x) > \epsilon_1\} \subset \{x \in S : g_1(x) > \gamma_1\} \subset \{x \in S : f_1(x) > \delta_1^3\}.$$

Let us define

$$h_1 : S^+ \longrightarrow \mathbb{R}$$
$$h_1 = \theta_1(g_1),$$

where $\theta_1 : \mathbb{R} \to [0,1]$ is a C^∞ function satisfying

$$\theta_1(t) = 0 \ \text{ if and only if } t \le \gamma_1$$
$$\theta_1(t) = 1 \ \text{ if and only if } t \ge \epsilon_1.$$

Note that the interior of the support of h_1 is the open set $U_1 := \{x \in S^+ : g_1(x) > \gamma_1\}$.
• **For $\mathbf{k = 2}$**. We may again use the density of $\mathrm{span}\{e_n^* : n \in \mathbb{N}\}$ in E^*, in order to find numbers $N_2 \in \mathbb{N}$, $\gamma_2, \epsilon_2 \in (0,1)$ with $\gamma_2 < \epsilon_2$, and $\beta_{2,0}, ..., \beta_{2,N_2} \in \mathbb{R}$ so that the linear functional

$$g_2 := \sum_{j=0}^{N_2} \beta_{2,j} e_j^*$$

has norm 1 and satisfies

$$\{x \in S : f_2(x) > \delta_2^2\} \subset \{x \in S : g_2(x) > \epsilon_2\} \subset \{x \in S : g_2(x) > \gamma_2\} \subset \{x \in S : f_2(x) > \delta_2^3\}.$$

We may assume without loss of generality that $N_1 \le N_2$ (otherwise we may set $\beta_{2,j} = 0$ for $N_2 < j \le N_1$ and take a new N_2 equal to N_1).
Now we define

$$h_2 : S^+ \longrightarrow \mathbb{R}$$
$$h_2 = \theta_2(g_2)\,(1 - \theta_1(g_1))\,,$$

where $\theta_2 : \mathbb{R} \to [0,1]$ is a C^∞ function satisfying:

$$\theta_2(t) = 0 \ \text{ if and only if } \ t \leq \gamma_2$$
$$\theta_2(t) = 1 \ \text{ if and only if } \ t \geq \epsilon_2.$$

Notice that the interior of the support of h_2 is the open set

$$U_2 = \{x \in S^+ : g_1(x) < \epsilon_1 \, , \ g_2(x) > \gamma_2\}.$$

• **For $\mathbf{k = 3}$,** By density of $\text{span}\{e_n^* : n \in \mathbb{N}\}$ in E^* we may pick numbers $N_3 \in \mathbb{N}$, $\gamma_3, \epsilon_3 \in (0,1)$ with $\epsilon_3 > \gamma_3$, and $\beta_{3,0}, ..., \beta_{3,N_3} \in \mathbb{R}$ so that, for

$$g_3 := \sum_{j=0}^{N_3} \beta_{3,j} e_j^*$$

we have that g_3 has norm 1 and

$$\{x \in S : \ f_3(x) > \delta_3^2\} \subset \{x \in S : \ g_3(x) > \epsilon_3\} \subset \{x \in S : \ g_3(x) > \gamma_3\} \subset \{x \in S : f_3(x) > \delta_3^3\}$$

Again we may assume without loss of generality that $N_2 \leq N_3$.
We define

$$h_3 : \ S^+ \longrightarrow \mathbb{R}$$
$$h_3 = \theta_3(g_3) \prod_{j=1}^{2} (1 - \theta_j(g_j)) \, ,$$

where $\theta_3 : \mathbb{R} \to [0,1]$ is a C^∞ function satisfying

$$\theta_3(t) = 0 \quad \text{if and only if } t \leq \gamma_3$$
$$\theta_3(1) = 1 \quad \text{if and only if } t \geq \epsilon_3.$$

Clearly the interior of the support of h_3 is the set

$$U_3 = \{x \in S^+ : \ g_1(x) < \epsilon_1 \, , \ g_2(x) < \epsilon_2 \ \text{ and } \ g_3(x) > \gamma_3\}.$$

We continue this process by induction.
• Assume that, in the steps $j = 2, ..., k$, with $k \geq 2$, we have selected points $y_j \in S^+$, positive integers $N_1 \leq N_2 \leq ... \leq N_k$, and constants $\gamma_j, \epsilon_j \in (0,1)$, $\beta_{j,i} \in \mathbb{R}$ so that the functionals

$$g_j := \sum_{i=0}^{N_j} \beta_{j,i} e_i^*$$

have norm 1 and satisfy

$$\{x \in S : \ f_j(x) > \delta_j^2\} \subset \{x \in S : \ g_j(x) > \epsilon_j\} \subset \{x \in S : \ g_j(x) > \gamma_j\} \subset \{x \in S : \ f_j(x) > \delta_j^4\}, \tag{4.3.1}$$

for all $j = 2, ..., k$. Assume also that we have defined numbers γ_j and functions

$$h_j = \theta_j(g_j) \prod_{i<j} (1 - \theta_i(g_i)) \, ,$$

where $\theta_j : \mathbb{R} \to [0,1]$ are C^∞ functions satisfying

$$\theta_j(t) = 0 \quad \text{if and only } t \leq \gamma_j$$
$$\theta_j(t) = 1 \quad \text{if and only } t \geq \epsilon_j.$$

The interior of the support of h_j is the set

$$U_j = \{x \in S^+ : \ g_1(x) < \epsilon_1 \, , ..., \, g_{j-1}(x) < \epsilon_{j-1} \ \text{ and } \ g_j(x) > \gamma_j\}.$$

Then we may again use the density of $\text{span}\{e_n^* : n \in \mathbb{N}\}$ in E^*, in order to find a positive integer $N_{k+1} \geq N_k$, and constants $\gamma_{k+1}, \epsilon_{k+1} \in (0,1)$, and $\beta_{k+1,0}, ..., \beta_{k+1,N_{k+1}} \in \mathbb{R}$ so that, for

$$g_{k+1} := \sum_{j=0}^{N_{k+1}} \beta_{k+1,j} e_j^*$$

we have that g_{k+1} has norm 1 and

$$\{x \in S : \ f_{k+1}(x) > \delta_{k+1}^2\} \subset \{x \in S : \ g_{k+1}(x) > \epsilon_{k+1}\}$$
$$\subset \{x \in S : \ g_{k+1}(x) > \gamma_{k+1}\} \subset \{x \in S : \ f_{k+1}(x) > \delta_{k+1}^3\}.$$

We now set

$$U_{k+1} := \{x \in S^+ : \ g_1(x) < \epsilon_1 \, , ..., \, g_k(x) < \epsilon_k \ \text{ and } \ g_{k+1}(x) > \gamma_{k+1}\}, \tag{4.3.2}$$

and define

$$h_{k+1} : S^+ \longrightarrow \mathbb{R}$$
$$h_{k+1} = \theta_{k+1}(g_{k+1}) \prod_{j<k+1} (1 - \theta_j(g_j)) \, .$$

where $\theta_{k+1} : \mathbb{R} \to [0,1]$ is a C^∞ function such that

$$\theta_{k+1}(t) = 0 \quad \text{if and only if } t \leq \gamma_{k+1}$$
$$\theta_{k+1}(t) = 1 \quad \text{if and only if } t \geq \epsilon_{k+1}.$$

Clearly the interior of the support of h_{k+1} is the set U_{k+1}.
Thus a sequence $\{h_n\}_{n\in\mathbb{N}}$ of C^1 smooth functions with the above properties is well defined by induction. As in the proof of Theorem 4.4 it is not difficult to check that the family $\{U_k\}_{k\in\mathbb{N}}$ is a locally finite open covering of S^+ refining $\{P_k\}_{k\in\mathbb{N}}$. Therefore the functions

$$\psi_n := \frac{h_n}{\sum_{k=1}^{\infty} h_k}, \ \ n \in \mathbb{N},$$

define a C^1 partition of unity in S^+ subordinate to $\{P_k\}_{k\in\mathbb{N}}$.
We will also need the following fact.

Claim 4.17. *For every $k \in \mathbb{N}$ and every $y \in S^+$, we have*

$$h_k(y) = 0 \Longrightarrow Dh_k(y) = 0.$$

Proof. Proceed as in the proof of Claim 4.8. □

Again we have that for every $x \in S^+$ there exist a number $n = n_x$ and an open neighbourhood V_x of x in S^+ such that $\psi_k(y) = 0$ for every $k > n$ and every $y \in V_x$. Let us also call $m = m_y$ the largest j such that $h_j(y) \neq 0$; this is also the largest j for which $\psi_j(y) \neq 0$. Thus, for every $y \in V_x$, we have

$$m_y = \max\{j : y \in \psi_j^{-1}((0,1])\} \leq n_x.$$

The derivatives of the functions ψ_n can be calculated as in the proof of Theorem 4.4. We have

$$D\psi_k(y)(v) = \sum_{j=1}^{n} \lambda_{k,j}(y) \, g_j(v) = \sum_{j=1}^{m_y} \lambda_{k,j}(y) \, g_j(v)$$

for each $v \in T_y S^+$, where the functions $\lambda_{k,j} : V_x \to \mathbb{R}$ are of class C^1.
The following lemma summarizes the properties of the partition of unity $\{\psi_n\}_{n\in\mathbb{N}}$ which will be most useful to us.

Lemma 4.18. *Given two continuous functions $f : S^+ \to F$ and $\varepsilon : S^+ \to (0,\infty)$, there exists a collection of norm-one linear functionals $\{g_k\}_{k\in\mathbb{N}} \subset Y^*$ of the form*

$$g_k = \sum_{j=0}^{N_k} \beta_{k,j} e_j^*,$$

where $N_1 \leq N_2 \leq N_3 \leq ...$, an open covering $\{P_n\}_{n\in\mathbb{N}}$ of S^+, and a C^1 partition of unity $\{\psi_n\}_{n\in\mathbb{N}}$ in S^+ such that:

1. *$\{\psi_n\}_{n\in\mathbb{N}}$ is subordinate to $\{P_n\}_{n\in\mathbb{N}}$.*
2. *For every $x \in S^+$ there exist a neighbourhood V_x of x in S^+ and a number $n = n_x \in \mathbb{N}$ such that $\psi_m = 0$ on V_x for all $m > n$, and the derivatives of the functions $\psi_1, ..., \psi_n$ on V_x are of the form*

$$D\psi_k(y)(v) = \sum_{j=1}^{n} \lambda_{k,j}(y)\, g_j(v) = \sum_{j=1}^{m_y} \lambda_{k,j}(y)\, g_j(v),$$

   for $v \in T_y S^+$, the tangent hyperplane to S^+ at $y \in S^+ \cap V_x$, where $m_y \leq n$ is the largest number such that $\psi_{m_y}(y) \neq 0$. More precisely, if L_y denotes the derivative of the function $u \mapsto \sqrt{1-\|u\|^2}$ evaluated at the point u_y such that $y = \left(u_y, \sqrt{1-\|u_y\|^2}\right)$, we have

$$\begin{aligned} D\psi_k(y)(v) = D\psi_k(y)(u, L_y(u)) &= L_y(u)\mu_{k,0}(y) + \sum_{j=1}^{N_n} \mu_{k,j}(y) e_j^*(u) \\ &= L_y(u)\mu_{k,0}(y) + \sum_{j=1}^{N_{m_y}} \mu_{k,j}(y) e_j^*(u) \end{aligned}$$

   for every $k = 1, ..., n$, and for every $v = (u, L_y(u)) \in T_y S^+$, where the functions $\mu_{k,j} : V_x \to \mathbb{R}$ are of class C^1, $j = 1, ..., n$.
3. *For every $n \in \mathbb{N}$ there exist a point $y_n := x_n \in P_n$ such that*

$$\frac{15}{16}\,\varepsilon(y_n) \leq \varepsilon(y) \leq \frac{17}{16}\,\varepsilon(y_n) \;\; \textit{for every } y \in P_n, \quad \textit{and} \quad \sup_{x,y\in P_n} \|f(x) - f(y)\| \leq \frac{\varepsilon(y_n)}{16}.$$

Note also that the integer n_x can be chosen as the first such j so that $x \in \{y \in S^+ : g_j(y) > \epsilon_j\}$. Then V_x can be chosen as $\{y \in S^+ : g_j(y) > \epsilon_j\}$ and, hence, we have $R_j \subset V_x \subset P_j$ for such V_x.
Now we are ready to start the construction of our approximating function $\varphi : S^+ \to F$, which will be of the form

$$\varphi(x) = \sum_{n=1}^{\infty} \left(f(y_n) + T_n(x)\right) \psi_n(x),$$

where the y_n are the points given by condition 3. of the preceding lemma, and the operators $T_n : Y \to F$ will be defined below. We have to distinguish two cases.

Case 1: Assume that F is infinite-dimensional.

In order to define the operators T_n, we work with the infinite subset $\mathbb{P}$ of $\mathbb{N}$ given by assumption 3. of Theorem 4.5, and we take a countable pairwise disjoint family of infinite subsets of $\mathbb{P}$ which *goes to infinity*. More precisely, we write

$$\bigcup_{n=1}^{\infty} I_n \subseteq \mathbb{P},$$

in such a way that:

1. $I_n := \{n_i : i \in \mathbb{N}\}$ is infinite for each $n \in \mathbb{N}$;
2. $I_n \cap I_m = \emptyset$ for all $n \neq m$; and
3. $\{1, \ldots, N_n\} \cap I_n = \emptyset$ for all $n \in \mathbb{N}$.

Here $\{N_n\}_{n\in\mathbb{N}}$ is the non-decreasing sequence of positive integers that appears in the construction of the functionals g_n of Lemma 4.18.
Now, by using assumption 3. of the statement, we can find, for each number $n \in \mathbb{N}$, a linear continuous surjection $S_n : E \to F$ of the form

$$S_n = A_n \circ P_n,$$

where A_n is a bounded linear operator from $\overline{\text{span}}\{e_{n_k} : k \in \mathbb{N}\} = \overline{\text{span}}\{e_m : m \in I_n\}$ onto F, and $P_n : E \to \overline{\text{span}}\{e_{n_k} : k \in \mathbb{N}\}$ is the natural projection associated to the unconditional basis $\{e_j\}_{j\in\mathbb{N}}$.
Now we finally define $T_n : Y \to F$ by

$$T_n(u,t) = \frac{\varepsilon(y_n)}{4\|S_n\|} S_n(u),$$

and $\varphi : S^+ \to F$ by

$$\varphi(x) = \sum_{n=1}^{\infty} (f(y_n) + T_n(x))\, \psi_n(x).$$

It is clear that φ is well defined and of class C^1.

Claim 4.19. *We have that* $\|\varphi(x) - f(x)\| \leq \varepsilon(x)$ *for every* $x \in S^+$.

Proof. This is shown exactly as in Claim 4.11. □

Let us now calculate the derivative of our function φ. For every $v = (u, L_y(u)) \in T_yS^+$ we have that

$$T_n(v) = T_n(u, L_y(u)) = \frac{\varepsilon(y_n)}{4\|S_n\|} S_n(u), \tag{4.3.3}$$

and we have that $D(T_{n|_{S^+}})(y)$ is the restriction of $DT_n(y) = T_n$ to T_yS^+, that is to say, if $v = (u, L_y(u)) \in T_yS^+$ then

$$DT_n(y)(u, L_y(u)) = \frac{\varepsilon(y_n)}{4\|S_n\|} S_n(u). \tag{4.3.4}$$

We can now compute the derivative of φ on S^+. Recall that, by condition 2. of Lemma 4.18, for every $x \in S^+$ there is a neighbourhood V_x of x in S^+ and a number $n = n_x \in \mathbb{N}$ such that

$$\varphi(y) = \sum_{j=1}^{n} (f(y_j) + T_j(y))\, \psi_j(y)$$

for every $y \in V_x$. Fix $y \in V_x$ and recall that for $m = m_y$ (the largest number j for which $\psi_j(y) \neq 0$), we have

$$\varphi(y) = \sum_{j=1}^{m} (f(y_j) + T_j(y))\, \psi_j(y).$$

By using (4.3.4) and the expression for $D\psi_j(y)$ given in Lemma 4.18, we see that

$$\begin{aligned} D\varphi(y)(u, L_y(u)) &= \left(\sum_{j=1}^{n} \psi_j(y) \frac{\varepsilon(y_j)}{4\|S_j\|} S_j(u) \right) + L_y(u)\alpha_{N_n,0}(y) + \sum_{j=1}^{N_n} \alpha_{N_n,j}(y) e_j^*(u) \\ &= \left(\sum_{j=1}^{m} \psi_j(y) \frac{\varepsilon(y_j)}{4\|S_j\|} S_j(u) \right) + L_y(u)\alpha_{N_m,0}(y) + \sum_{j=1}^{N_m} \alpha_{N_m,j}(y) e_j^*(u) \qquad (4.3.5) \end{aligned}$$

for every $(u, L_y(u)) \in T_yS^+$, $y \in V_x$, where the functions $\alpha_{N_n,j} : V_x \subset S^+ \to F$ are of class C^1 (because we have $\alpha_{N_n,j}(y) = \sum_{i=1}^n (f(y_i) + T_i(y))\, \mu_{i,j}(y)$ and $\alpha_{N_n,0}(y) = \sum_{i=1}^n (f(y_i) + T_i(y))\, \mu_{i,0}(y)$, where $\mu_{i,j}$ are as in Lemma 4.18).
Let us now prove that the critical set of φ is relatively small.

Lemma 4.20. *The set $C_\varphi := \{x \in S^+ \,:\, D\varphi(x)$ is not surjective$\}$ is of the form*

$$C_\varphi = \{\left(w, \sqrt{1 - \|w\|^2}\right) : w \in A\},$$

where $A \subset E$ is a relatively closed subset of the open unit ball of E that is locally contained in a complemented subspace of infinite codimension in E.

Proof. Observe that if $n = n_x = 1$ then $\varphi(y) = f(y_1) + T_1(y)$ for every $y \in V_x$, and because T_1 is surjective φ does not have any critical point in V_x. Now let us assume that $n = n_x \geq 2$. Let $y = (w, \sqrt{1 - \|w\|^2})$ be point of V_x. Let $m = m_y$ be the largest number such that $\psi_m(y) \neq 0$. Recall that $m \leq n$. We only need to show that if

$$w \notin \overline{\text{span}}\,\{e_j \,:\, j \in \mathbb{P} \text{ or } j = 1, \dots, N_n\}$$

then for every $v \in F$ there exists $u \in E$ such that

$$D\varphi(w, \sqrt{1 - \|w\|^2})(u, L_y(u)) = v,$$

since this will mean that the set

$$A := \{w \in E : (w, \sqrt{1 - \|w\|^2}) \in C_\varphi\}$$

will be locally contained in subspaces of the form $\overline{\text{span}}\,\{e_i \,:\, i \in \mathbb{P} \text{ or } i = 1, \dots, N_n\}$, which are complemented, and of infinite codimension, in E.
We will need to use the following.

Fact 4.21. *For every $w = \sum_{j=1}^\infty w_j e_j \in E \setminus \{0\}$ and every $j_0 \in \mathbb{N}$ we have that*

$$w_{j_0} \neq 0 \implies \langle J(w), e_{j_0} \rangle \neq 0,$$

where $J(w)$ denotes $D\|\cdot\|(w)$, and $\langle J(w), u\rangle := J(w)(u)$.

Proof. If $w_{j_0} \neq 0$ then, by assumption 2. of the statement of Theorem 4.5, which is the 1-suppression unconditionality of the basis $\{e_n\}_{n\in\mathbb{N}}$, we have that

$$\|\sum_{j=1,\, j\neq j_0}^\infty w_j e_j\| \leq \|\sum_{j=1}^\infty w_j e_j\|.$$

This means that the convex function $\theta : \mathbb{R} \to \mathbb{R}$ defined by

$$\theta(t) = \|w + te_{j_0}\|$$

has a minimum at $t = -w_{j_0}$. On the other hand, if we had $\langle J(w), e_{j_0}\rangle = 0$, then the same function θ would have another minimum at the point $t = 0$. But since $\|\cdot\|$ is strictly convex the function θ can only attain its minimum at a unique point. Therefore we must have $\langle J(w), e_{j_0}\rangle \neq 0$. □

So let us pick a point $w \in E \setminus \overline{\text{span}}\,\{e_j \,:\, j \in \mathbb{P} \text{ or } j = 1, \dots, N_n\}$ and a vector $v \in F$, and let us construct a vector $u \in E$ such that $D\varphi(w, \sqrt{1 - \|w\|^2})(u, L_y(u)) = v$, where $y = (w, \sqrt{1 - \|w\|^2})$. By assumption, there exists $j_0 \in \mathbb{N}$, $j_0 \in \mathbb{N} \setminus \mathbb{P}$, such that $j_0 > N_n \geq N_m$ and $w_{j_0} \neq 0$. According to the fact just shown, we have $\langle J(w), e_{j_0}\rangle \neq 0$. Now, since S_m is surjective and $\psi_m(y) \neq 0$, we may find a sequence $(u_{m_i})_{i\in\mathbb{N}}$ (indexed by the subsequence $(m_i)_{i\in\mathbb{N}}$ defined by I_m) such that

$$\psi_m(y)\frac{\varepsilon(y_m)}{4\|S_m\|}\, S_m \left(\sum_{i=1}^{\infty} u_{m_i} e_{m_i}\right) = v.$$

Note that $j_0 \notin I_m = \{m_i : i \in \mathbb{N}\}$, because $j_0 \in \mathbb{N}\setminus\mathbb{P}$ and $I_m \subset \mathbb{P}$. Then we can set

$$u_{j_0} := -\frac{\langle J(w),\, \sum_{i=1}^{\infty} u_{m_i} e_{m_i}\rangle}{\langle J(w), e_{j_0}\rangle},$$

so that we have

$$\langle J(w),\, u_{j_0} e_{j_0} + \sum_{i=1}^{\infty} u_{m_i} e_{m_i}\rangle = 0,$$

which bearing in mind that

$$L_y = -\frac{\|w\|}{\sqrt{1-\|w\|^2}}\langle J(w),\, \cdot\rangle$$

also implies that

$$L_y\left(u_{j_0} e_{j_0} + \sum_{i=1}^{\infty} u_{m_i} e_{m_i}\right) = 0.$$

So if we set $u_j = 0$ for all $j \notin I_m \cup \{j_0\}$ and we define

$$u := \sum_{j=1}^{\infty} u_j e_j$$

then we have that

$$\psi_m(y)\frac{\varepsilon(y_m)}{4\|S_m\|}\, S_m\,(u) = v,\ \ L_y(u) = 0,\ \ \sum_{j=1}^{N_m} \alpha_{N_m,j}(y) e_j^*(u) = 0, \quad \text{and also } S_j(u) = 0 \text{ for } j < m,$$

because $j_0 > N_n \geq N_m$, $I_m \cap \{1, 2, ..., N_m\} = \emptyset$, and the sets I_j are pairwise disjoint. In view of (4.3.5) these equalities imply that $D\varphi(y)(u) = v$. □

Now, according to Lemma 4.20 and Theorem 3.3, we can extract the set C_φ, since it is C^1 diffeomorphic (via the projection of the graph S^+ of the function $w \mapsto \sqrt{1-\|w\|^2}$ onto the open unit ball of E) to a subset which can be extracted. Therefore we can finish the proof of Theorem 4.5 exactly as we did with Theorem 4.4.

Case 2: Assume that $F = \mathbb{R}^m$. The proof is almost identical, but with the following important difference: now the set $\mathbb{P}$ is by definition the set of *even* positive integers, and the sets I_n are *finite* subsets of $\mathbb{P}$ such that:

1. $\sharp I_n = m$ for each $n \in \mathbb{N}$;
2. $I_n \cap I_j = \emptyset$ for all $n \neq j$; and
3. $\{1, \ldots, N_n\} \cap I_n = \emptyset$ for all $n \in \mathbb{N}$.

Here $\{N_n\}_{n\in\mathbb{N}}$ is the non-decreasing sequence of positive integers that appears in the construction of the functionals g_n of Lemma 4.18.
Of course in this case we can always find linear surjections $A_n : \mathrm{span}\{e_i\,:\, i \in I_n\} \to \mathbb{R}^m$. □

4.4 Some technical versions, examples and remarks

In this section we will give some examples, make some remarks and establish more technical variants of our Theorems 4.4 and 4.5 which follow by the same method of proof. We will also prove Theorem 4.3 and Proposition 4.6.

The proof of Theorem 4.4 can be easily adjusted to obtain more general results with more complicated statements. Namely, the following two results are true.

Theorem 4.22. *Let E and F be Banach spaces. Assume that:*

1. *E is infinite-dimensional, with a separable dual E^*.*
2. *There exist three sequences $\{E_{n,1}\}_{n\geq 1}$, $\{E_{n,2}\}_{n\geq 1}$, $\{E_{n,3}\}_{n\geq 2}$ of subspaces of E such that*

$$E = E_{1,1} \oplus E_{1,2},$$
$$E_{1,2} = (E_{2,1} \oplus E_{2,2}) \oplus ... \oplus (E_{n,1} \oplus E_{n,2} \oplus E_{n,3}),$$
$$E_{n,3} = E_{n+1,1} \oplus E_{n+1,2} \oplus E_{n+1,3},$$

with either $E_{n,3}$ being infinite-dimensional and reflexive and $\dim E_{n,2} \geq 1$, or else $E_{n,2}$ being infinite-dimensional and reflexive for all $n \geq 2$. Suppose also that there exists a bounded linear operator from $E_{n,1}$ onto F for every $n \in \mathbb{N}$.

Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^1 mapping $g : E \to F$ such that $\|f(x) - g(x)\| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

Observe that if $E_{n,3}$ is infinite-dimensional and reflexive, the spaces $E_{n,2}$ can be taken to be of dimension 1 for every $n \in \mathbb{N}$. The proof is almost the same as that of Theorem 4.4. Here, up to finite-dimensional perturbations of the subspaces, $E_{k,j}$, we can arrange that $E_{1,2} \subset \operatorname{Ker} g_1^1$ and that $E_{n,1} \oplus E_{n,3} \subseteq \bigcap_{j=1}^{n} \operatorname{Ker} g_j^1$, and we may set $A_x = \{y \in S^+ : E_{n,3} \subset \operatorname{Ker} L_y\}$.

Theorem 4.23. *Let E, X, and F be Banach spaces. Assume either that E is infinite-dimensional, separable, and reflexive, and F is finite-dimensional, or that:*

1. *E is infinite-dimensional, with a separable dual E^*.*
2. *There exists a decomposition of E,*

$$E = G \oplus E_1 \oplus X,$$

such that G is infinite-dimensional and reflexive, and E_1 is isomorphic to E.

3. *There exists a bounded linear operator from $G \oplus X$ onto F (equivalently, F is a quotient of $G \oplus X$).*

Then, for every continuous mapping $f : E \to F$ and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^1 mapping $g : E \to F$ such that $\|f(x) - g(x)\| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

If, additionally, X is isomorphic to E, then F can taken as a quotient of E.

Observe that each of these two results imply Theorem 4.4, with Theorem 4.22 being the most general one.

For instance, Theorem 4.23 can be applied to the James space J and to its dual J^*. Indeed, both spaces have separable dual. It is known that J has many reflexive infinite-dimensional complemented subspaces G [52]. Since J is prime [51], for each such G, we can write $J = G \oplus J$ (for instance we have $J = l_2 \oplus J$). Now, recalling the fact that J has a separable dual, apply Theorem 4.23 to $E = E_1 = J$ and $X = \{0\}$ to see that every continuous function $f : J = G \oplus J \to F$, where F is a quotient of G, can be uniformly approximated by C^1 smooth mappings without critical points. Similar arguments work for the dual of the James space J^*.

It also follows that the conclusion of Theorem 4.4 is true for *composite* spaces of the form $c_0 \oplus \ell_p$ or $c_0 \oplus L^p$, $1 < p < \infty$, with k being the order of smoothness of ℓ_p or L^p. More generally, if E is any finite direct sum of the classical Banach spaces c_0, ℓ_p or L^p, $1 < p < \infty$, and there is a bounded linear operator from E onto F then the conclusion of Theorem 4.2 is true with k being the minimum of the orders of smoothness of the spaces appearing in this decomposition of E.

Remark 4.24. Notice that that in the case that E is a separable Hilbert space, we have that the function $w \mapsto \|w\|^2$ is of class C^∞, hence all the mappings appearing in the proof of Theorem 4.4 are of class C^∞, and we directly obtain an approximating function g of class C^∞ with no critical points.

In fact, in the Hilbertian case we do not need to use a partition of unity in the upper sphere S^+. We can directly construct a partition of unity $\{\psi_n\}_{n\in\mathbb{N}}$ in E subordinated to an open covering by open balls with linearly independent centers $\{y_j\}$, as in [12]. Then, choosing an orthonormal basis $\{e_j\}$ for which $\text{span}\{y_1, \ldots, y_n\} = \text{span}\{e_1, \ldots, e_n\}$ for every $n \in \mathbb{N}$, we define operators $T_n : E \to F$ as in the proof of Theorem 4.5, where $\mathbb{P}$ can be any infinite subset of $\mathbb{N}$ such that $\mathbb{N} \setminus \mathbb{P}$ is also infinite. Then one can easily check that the function

$$\varphi(y) = \sum_{n=1}^{\infty} \left(f(y_n) + T_n(y - y_n)\right) \psi_n(y)$$

approximates f and the set C_φ of its critical points is locally contained in a subspace of infinite codimension in E, specifically in subspaces of the form $\overline{\text{span}}\{e_j : j \in \mathbb{P} \text{ or } j = 1, \ldots, n\}$. Then one can extract C_φ by means of a C^∞ diffeomorphism $h : E \to E \setminus C_\varphi$ which is sufficiently close to the identity, and conclude that the function $g := \varphi \circ h$ approximates f and has no critical points.

The same proof as that of Theorem 4.5, with obvious adjustments, allows us to obtain a more general (and also more technical) result as follows.

Theorem 4.25. *Let E be an infinite-dimensional Banach space, and F be a Banach space such that:*

1. *E has an equivalent norm $\|\cdot\|$ which is C^1 and locally uniformly convex.*
2. *E has a (normalized) Schauder basis $\{e_n\}_{n\in\mathbb{N}}$ which is shrinking.*
3. *There exists an infinite subset $\mathbb{I}$ of $\mathbb{N}$ such that the subspace $\overline{\text{span}}\{e_j : j \in \mathbb{N} \setminus \mathbb{I}\}$ is complemented in E, and for every $x = \sum_{j=}^{\infty} x_j e_j$ and every $j_0 \in \mathbb{I}$ we have that*

$$\left\| \sum_{j\in\mathbb{N},\, j\neq j_0} x_j e_j \right\| \leq \left\| \sum_{j\in\mathbb{N}} x_j e_j \right\|.$$

4. *In the case that F is infinite-dimensional, there exists an infinite subset $\mathbb{P}$ of $\mathbb{N}$ such that $\mathbb{I} \setminus \mathbb{P}$ is infinite and for every infinite subset J of $\mathbb{P}$ the subspace $E' = \overline{\text{span}}\{e_j : j \in J \cup (\mathbb{N} \setminus \mathbb{I})\}$ is complemented in E, and there exists a linear bounded operator from E' onto F.*

Then, for every continuous mapping $f : E \to F$ and for every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^1 mapping $g : E \to F$ such that $\|f(x) - g(x)\| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

Proof. The only important difference with the proof of Theorem 4.5 is that now we have to use an analogue of Fact 5.10 which is true if we just pick $j_0 \in \mathbb{I}$. Therefore, if we take $w = \sum_{j=1}^{\infty} w_j e_j \notin \overline{\text{span}}\{e_j : j \in \mathbb{P} \cup (\mathbb{N} \setminus \mathbb{I}) \text{ or } j = 1, \ldots, N_n\}$, since $\mathbb{I} \setminus \mathbb{P}$ is infinite, there will exist $j_0 \in \mathbb{I}$, $j_0 > N_n$ such that $w_{j_0} \neq 0$ and thus $\langle J(w), e_{j_0}\rangle \neq 0$. The operators S_n have supports in complemented subspaces of the form $\overline{\text{span}}\{e_j : j \in I_n \cup (\mathbb{N} \setminus \mathbb{I})\}$, where the sets $I_n \subset \mathbb{P}$ are defined as in the proof of Theorem 4.5. □

Remark 4.26. It is clear that the spaces ℓ_p and L^p, $1 < p < \infty$ satisfy the assumptions of Theorem 4.4. It may not be so obvious why the space c_0 satisfy the assumptions of Theorem 4.5; let us clarify this point. If we repeat the proof of [60, Theorem V.1.5] in the particular case that $\Gamma = \mathbb{N}$, since all the operations that are made in this proof are coordinate-wise monotone, we see that the C^1 and LUR renorming $\|\cdot\|$ that we obtain for c_0 has the property that

$$\left\| \sum_{j\in\mathbb{N},\, j\neq j_0} x_j e_j \right\| \leq \left\| \sum_{j\in\mathbb{N}} x_j e_j \right\|$$

for every $j_0 \in \mathbb{N}$ and every $x = (x_1, x_2, x_3, ...) \in c_0$, where $\{e_n\}$ is the canonical basis of c_0. This shows that this norm $\|\cdot\|$ satisfies assumptions 1. and 2. of Theorem 4.5. On the other hand, for every infinite subset J of $\mathbb{N}$ we have that $\overline{\text{span}}\{e_j : j \in J\}$ is isomorphic to c_0, so it is clear that assumption 3. is satisfied as well, provided that there exists a continuous linear operator from c_0 onto F. Therefore $E = c_0$ satisfies the conclusion of Theorem 4.2.

The latter fact can be generalized to Banach spaces with a shrinking basis which contain copies of c_0.

Theorem 4.27. *Let E be a Banach space that contains the space c_0 and admits a shrinking Schauder basis. Let F be a quotient of E.*

Then, for every continuous mapping $f : E \to F$ and for every continuous function $\varepsilon : E \to (0,\infty)$ there exists a C^1 mapping $g : E \to F$ such that $\|f(x) - g(x)\| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

Proof. We will show that E satisfies the assumptions of Theorem 4.25.
First, by Sobczyk's Theorem [143], c_0 is complemented in E, that is, E is isomorphic to $G \oplus c_0$, for a certain Banach space G. Since c_0 is isomorphic to $c_0 \oplus c_0$, G may be taken as E. So, we can and will assume that

$$E = c_0 \oplus E.$$

Let $\{e_j\}_{j\in N}$ be the canonical Schauder basis in c_0. Equip c_0 with the C^1 and LUR norm $||\cdot||$ which was described in Remark 4.26. That is, for every $j_0 \in \mathbb{N}$, we have

$$|| \sum_{j=1, j\neq j_0}^{\infty} \alpha_j e_j || \leq || \sum_{j=1}^{\infty} \alpha_j e_j ||,$$

for every $x = \sum_{j=1}^{\infty} \alpha_j e_j \in c_0$.
Similarly, let $\{d_n\}_{n\in\mathbb{N}}$ be a shrinking Schauder basis in E. Equip E with a C^1 and LUR norm $|\cdot|$. Define a new norm in $c_0 \oplus E$, by letting

$$|||x + y||| = \sqrt{||x||^2 + |y|^2},$$

for every $x + y \in c_0 \oplus E$. This norm is C^1 and LUR as well. Define $\{f_k\}_{k\in\mathbb{N}} \subset c_0 \oplus E$, where $f_{2j-1} = e_j + 0$ and $f_{2n} = 0 + d_n$ for every $j, n \in \mathbb{N}$. It is also easy to check that $\{f_k\}_{k\in\mathbb{N}}$ is a shrinking Schauder basis for $c_0 \oplus E$.
For $x + y \in c_0 \oplus E$, let $x = \sum_{j=1}^{\infty} \alpha_j e_j \in c_0$ and $y = \sum_{n=1}^{\infty} \beta_n d_n \in E$ be their basis expansions. Then, writing $z_{2j-1} = \alpha_j$ and $z_{2n} = \beta_n$, we obtain the expansion of $z = x + y = \sum_{k=1}^{\infty} z_k f_k =$

$\sum_{j=1}^{\infty} z_{2j-1}e_j + \sum_{n=1}^{\infty} z_{2n}d_n \in c_0 \oplus E$. For every $j_0 \in \mathbb{N}$, we have

$$|||\sum_{k=1,k\neq 2j_0}^{\infty} z_k f_k||| = |||\sum_{j=1,j\neq j_0}^{\infty} z_{2j-1}e_j + \sum_{n=1}^{\infty} z_{2n}d_n||| \leq |||\sum_{j=1,j\neq j_0}^{\infty} \alpha_j e_j + \sum_{n=1}^{\infty} \beta_n d_n|||$$
$$\leq |||\left(\sum_{j=1,j\neq j_0}^{\infty} \alpha_j e_j\right) + y||| = \left(||\sum_{j=1,j\neq j_0}^{\infty} \alpha_j e_j||^2 + |y|^2\right)^{\frac{1}{2}} \leq \sqrt{||x||^2 + |y|^2}$$
$$= |||x + y||| = |||\sum_{k=1}^{\infty} z_k f_k|||,$$

where in the second line we have used the fact that $||\sum_{j=1,j\neq j_0}^{\infty} \alpha_j e_j||^2 \leq ||\sum_{j=1}^{\infty} \alpha_j e_j||^2 = ||x||^2$. Now, we are in a position to apply Theorem 4.25. Namely, let $\mathbb{I} = \{2n : n \in \mathbb{N}\}$ and $\mathbb{P} = \{4n : n \in \mathbb{N}\}$. □

Corollary 4.28. *Let $C(K)$ be the Banach space of continuous functions, where K is a metrizable countable compactum and F be a quotient of $C(K)$.*

Then, for every continuous mapping $f : C(K) \to F$ and for every continuous function $\varepsilon : C(K) \to (0, \infty)$ there exists a C^∞ mapping $g : E \to F$ such that $\|f(x) - g(x)\| \leq \varepsilon(x)$ and $Dg(x) : E \to F$ is a surjective linear operator for every $x \in E$.

Proof. By an application of [105, Theorem 1.4], which states that a Banach space has a shrinking basis provided its dual has a Schauder basis, we obtain that $C(K)$ has a shrinking basis (because $C(K)^* = l_1$). Moreover, using the fact that c_0 is a subspace of $C(K)$, we infer that $C(K)$ is isomorphic to $c_0 \oplus G$ for some Banach space G, which yields (as in the above proof) that $C(K)$ is isomorphic to $c_0 \oplus C(K)$. Hence, by Theorem 4.27, the C^1 version of our assertion holds. The C^∞ version requires the fact that $C(K)$ has an equivalent C^∞ norm, which is due to Haydon [98], and Proposition 4.6. □

For more information about the spaces $C(K)$ we refer the reader to [138]. The space $C(K)$ is an example of *isometric* predual of ℓ_1 (meaning a Banach space E with an equivalent norm $\|\cdot\|$ such that the dual $(E^*, \|\cdot\|^*)$ is isometric to ℓ_1). The class of isomorphic predual spaces for ℓ_1 is larger that the class of isometric predual spaces (the space constructed by Bourgain and Delbaen [42] is such an example), which in turn is smaller than the class of $C(K)$ spaces for metrizable countable compactum K, see [33].

Remark 4.29. Since every isometric predual space E of ℓ_1 contains c_0 (see for instance [156, Corollary 1]) and admits an equivalent real-analytic norm [59, Corollary 3.3], the above corollary is valid for E. Even more, the corollary is valid for any infinite-dimensional separable Banach space E which has a shrinking basis and which admits an equivalent polyhedral norm (equivalently, with a countable James boundary). This follows from the facts that, being polyhedral, E must contain c_0, and that a space with a countable James boundary admits an equivalent real-analytic norm (see [59] or [93, Chapter 5, section 6] for reference).

As we noted in the introduction our main results imply that continuous functions between many Banach spaces can be arbitrarily well approximated by smooth open mappings.

Remark 4.30. *Let (E, F) be a pair of Banach spaces with the property that for every continuous mapping $f : E \to F$ and for every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^k mapping $g : E \to F$ with no critical points such that $\|f(x) - g(x)\| \leq \varepsilon(x)$, $x \in E$. Then the pair (E, F) also has the following property: for every continuous mapping $f : E \to F$ and for every continuous function $\varepsilon : E \to (0, \infty)$ there exists and open mapping $g : E \to F$ of class C^k such that $\|f(x) - g(x)\| \leq \varepsilon(x)$, $x \in E$.*

This follows trivially from [118, Theorem XV.3.5]. Recall that $g : E \to F$ is said to be open if for every open subset U of E we have that $g(U)$ is open in F. Notice that the approximation of arbitrary continuous maps by smooth (or even merely continuous) open maps is impossible for $E = \mathbb{R}^n$: for instance, if $E = \mathbb{R}^n$, $F = \mathbb{R}$, $f(x) = e^{-\|x\|^2}$, $\varepsilon(x) = 1/3$, every continuous function g which ε-approximates f must attain a global maximum in $\mathbb{R}^n$, hence $g(\mathbb{R}^n)$ is not open in $\mathbb{R}$.

For the analogous result in manifolds we have Theorem 4.3 whose proof is as follows.

Proof of Theorem 4.3. As said in the introduction, by the results of [67, 115], it is sufficient to show Theorem 4.3 for functions $f : U \to V$, where $U \subset E$ and $V \subset F$ are open subsets of two separable Hilbert spaces E, F, respectively. Observe that we can assume $V = F$. Indeed, if $f : U \to V \subset F$, $\varepsilon : U \to (0, \infty)$ are continuous functions then, by taking $\widetilde{\varepsilon}(x) = \frac{1}{2}\min\{\varepsilon(x), \text{dist}(f(x), F \setminus V)\}$, if we are able to $\widetilde{\varepsilon}$-approximate $f : U \to F$ by a smooth function $g : U \to F$ with no critical points, then we also have that $\|g(x) - f(x)\| < \text{dist}(f(x), F \setminus V)$, which implies that $g(x) \in V$ for every $x \in U$; that is, we really have $g : U \to V$. On the other hand, showing the result for $f : U \to F$ is not more difficult than proving it in the case $U = E$ (though it does encumber the notation). For example, it requires a version of the extractibility fact (a counterpart of Theorem 3.2) where the whole space E, its closed subset X, and an open cover $\mathcal{G}$ of E must be replaced with an open subset U (of E), a closed subset of U, and an open cover of U, respectively. Such a fact can be proved by mimicking the technique of the proof of Theorem 3.2; one just has to make some easy adjustments in the appropriate places. We leave the details to the interested reader. □

Throughout the chapter the "limiting" function $\varepsilon(x)$ is assumed to be positive. The following remark explains what can be said if we merely require that $\varepsilon(x) \geq 0$.

Remark 4.31. Let H be a separable, infinite-dimensional Hilbert space and $f : H \to H$ be a continuous mapping. Then, for every continuous function $\varepsilon : H \to [0, \infty)$, there exists a continuous mapping $g : H \to H$ such that the restriction $g_{|_{H\setminus\varepsilon^{-1}(0)}}$ is C^∞ smooth and has no critical points, and $\|f(x)-g(x)\| \leq \varepsilon(x)$ for every $x \in H$ (hence, $f(x) = g(x)$ provided $\varepsilon(x) = 0$). This a consequence of Theorem 4.3 applied to $U = H \setminus \varepsilon^{-1}(0)$ and $\varepsilon_{|_U}$.

Let us conclude this chapter with the proof of Proposition 4.6.

Proof of Proposition 4.6. Let $f : E \to F$ and $\varepsilon : E \to (0, \infty)$ be continuous. By the first assumption there exists a C^1 function $\varphi : E \to F$ without critical points so that

$$\|f(x) - \varphi(x)\| \leq \varepsilon(x)/2.$$

It is well known that the set of continuous linear surjections from a Banach space E onto a Banach space F is open; see [118, Theorem XV.3.4] for instance. Therefore, for each $x \in E$ there exists $r_x > 0$ such that if $S : E \to F$ is a bounded linear operator then

$$\|S - D\varphi(x)\| < 2r_x \implies S \text{ is surjective.} \tag{4.4.1}$$

By continuity of $D\varphi$, for every x we may find a number $s_x \in (0, r_x)$ such that if $y \in B(x, s_x)$ then

$$\|D\varphi(y) - D\varphi(x)\| < r_x.$$

Since E is separable, we can extract a countable subcovering

$$E = \bigcup_{n=1}^{\infty} B(x_n, s_n),$$

where $s_n := s_{x_n}$. Let us also denote $r_n := r_{x_n}$, and define $\eta : E \to (0, 1)$ by

$$\eta(y) = \min\left\{\frac{\varepsilon(y)}{2}, \sum_{n=1}^{\infty} \frac{s_n}{2}\psi_n(y)\right\},$$

where $\{\psi_n\}$ is a partition of unity such that the open support of ψ_n is contained in $B(x_n, s_n)$. Now we may apply the second assumption to find a C^k function $g : E \to F$ such that

$$\|\varphi(y) - g(y)\| \leq \eta(y), \text{ and } \|D\varphi(y) - Dg(y)\| \leq \eta(y)$$

for all $y \in E$. Then for every $y \in E$ there exists $n = n_y \in \mathbb{N}$ such that $y \in B(x_n, s_n)$ and $\eta(y) \leq s_n/2$. It follows that $\|Dg(y) - D\varphi(y)\| \leq s_n/2 < r_n$ and $\|D\varphi(y) - D\varphi(x_n)\| < r_n$, hence $\|Dg(y) - D\varphi(x_n)\| < 2r_n$, and according to (4.4.1) this implies that $Dg(y)$ is surjective. This shows that g has no critical points. On the other hand, since $\eta \leq \varepsilon/2$, it is clear that

$$\|f(y) - g(y)\| \leq \|f(y) - \varphi(y)\| + \|\varphi(y) - g(y)\| \leq \varepsilon(y)/2 + \varepsilon(y)/2 = \varepsilon(y),$$

so g also approximates f as required. □

Chapter 5

Extraction of critical points of smooth functions on Banach spaces

Recall that Kupka's counterexample [116] showed that no analogue of the Morse-Sard theorem is possible in the framework of infinite-dimensional Banach spaces. Approximated Morse-Sard results then appeared, where the goal is uniformly approximate any given continuous mapping by a smooth mapping whose set of critical values is small in some sense, even empty. And this was the direction of research taken in the previous Chapter 4.

In the present chapter we consider a different approach to this problem. Suppose that our given continuous function $f : E \to F$ is already of class C^1 and we know that its set of critical points C_f is included in some open set U. The question is, are we able not only to uniformly approximate f by another C^1 function φ without critical points but also to make φ be equal to f outside U? We answer this question in the affirmative for some classical Banach spaces E and the case $F = \mathbb{R}^d$. Namely we prove that if E is an infinite-dimensional separable Hilbert space, then for every C^1 function $f : E \to \mathbb{R}^d$, every open set U with $C_f := \{x \in E : Df(x) \text{ is not surjective}\} \subset U$ and every continuous function $\varepsilon : E \to (0,\infty)$ there exists a C^1 mapping $\varphi : E \to \mathbb{R}^d$ such that $|f(x) - \varphi(x)| \leq \varepsilon(x)$ for every $x \in E$, $f = \varphi$ outside U and φ has no critical points ($C_\varphi = \emptyset$). This result can be generalized to the case where $E = c_0$ or $E = \ell_p$, $1 < p < \infty$. In the case $E = c_0$ it is also possible to get that $||Df(x) - D\varphi(x)|| \leq \varepsilon(x)$ for every $x \in E$.

5.1 Main results

Our main goal is to prove the following result:

Theorem 5.1. *Let E be one of the classical infinite-dimensional Banach spaces c_0 or ℓ_p with $1 < p < \infty$. Let $f : E \to \mathbb{R}^d$ be a C^1 function and $\varepsilon : E \to (0,\infty)$ a continuous function. Take any open set U containing the critical set of points of f, that is $C_f := \{x \in E : Df(x) \text{ is not surjective}\}$. Then there exists a C^1 function $\varphi : E \to \mathbb{R}^d$ such that,*

(1) *$|f(x) - \varphi(x)| \leq \varepsilon(x)$ for all $x \in E$;*

(2) *$f(x) = \varphi(x)$ for all $x \in E \setminus U$;*

(3) *$D\varphi(x)$ is surjective for all $x \in E$, i.e. φ has no critical points; and*

(4) *in the case that $E = c_0$ we also have that $||Df(x) - D\varphi(x)|| \leq \varepsilon(x)$ for all $x \in E$.*

We can make φ be of class C^k inside the open set U, where k denotes the order of smoothness of the space ℓ_p, $1 < p < \infty$ or c_0. A brief explanation of this fact can be found in Remark 5.11.

This theorem is a particular case of the following two more technical results.

Theorem 5.2. *Let E be an infinite-dimensional Banach space with an unconditional basis and with a C^1 equivalent norm $||\cdot||$ that locally depends on finitely many coordinates. Let $f : E \to \mathbb{R}^d$ be a C^1 function and $\varepsilon : E \to (0,\infty)$ a continuous function. Take any open set U such that $C_f \subset U$. Then there exists a C^1 function $\varphi : E \to \mathbb{R}^d$ such that,*

(1) $|f(x) - \varphi(x)| \leq \varepsilon(x)$ *for all* $x \in E$;

(2) $f(x) = \varphi(x)$ *for all* $x \in E \setminus U$;

(3) $||Df(x) - D\varphi(x)|| \leq \varepsilon(x)$ *for all* $x \in E$; *and*

(4) $D\varphi(x)$ *is surjective for all* $x \in E$.

Theorem 5.3. *Let E be an infinite-dimensional Banach space with a C^1 strictly convex equivalent norm $||\cdot||$ and with a 1-suppression unconditional basis $\{e_n\}_{n\in\mathbb{N}}$, that is a Schauder basis such that for every $x = \sum_{j=1}^{\infty} x_j e_j$ and every $j_0 \in \mathbb{N}$ we have that*

$$\left\| \sum_{j\in\mathbb{N},\, j\neq j_0} x_j e_j \right\| \leq \left\| \sum_{j\in\mathbb{N}} x_j e_j \right\|.$$

Let $f : E \to \mathbb{R}^d$ be a C^1 function and $\varepsilon : E \to (0,\infty)$ a continuous function. Then for every open set U such that $C_f \subset U$ there exists a C^1 function $\varphi : E \to \mathbb{R}^d$ such that,

(1) $|f(x) - \varphi(x)| \leq \varepsilon(x)$ *for every* $x \in U$.

(2) $f(x) = \varphi(x)$ *for all* $x \in E \setminus U$.

(3) $D\varphi(x)$ *is surjective for all* $x \in E$.

The case c_0 and ℓ_p, $1 < p < \infty$ in Theorem 5.1 follow from Theorem 5.2 and Theorem 5.3 respectively. The reader can find the details of why this is so in Remark 5.12.

Note that the approximating function that we build does not have any critical point, hence it is an open mapping.

The proof of both Theorems 5.2 and 5.3 will follow these two steps:

- **Step 1**: Firstly we take a C^1 function $\delta : E \to [0,\infty)$ so that $\delta(x) \leq \varepsilon(x)$ and $\delta^{-1}(0) = E \setminus U$. Then we construct a C^1 function $g : U \to \mathbb{R}^d$ such that $|f(x) - g(x)| \leq \delta(x)/2$ and $||Df(x) - Dg(x)|| \leq \delta(x)$ and such that C_g either is the empty set for the case of Theorem 5.2, or is locally contained in a finite union of complemented subspaces of infinite codimension in E for the case of Theorem 5.3. We follow here [128, 20]'s arguments.

- **Step 2**: We extend the function g to the whole space E by letting it be equal to f outside U. Because of the approximation in the C^1-topology of Step 1 this extension is still of class C^1 on E. For the case of Theorem 5.2 we are done. For the case of Theorem 5.3 we must find a C^1-diffeomorphism $h : E \to E \setminus C_g$ which will be the identity outside U and such that h is limited by $\mathcal{G}$, where $\mathcal{G}$ is an open cover of E by open balls $B(z,\delta_z)$ chosen in such a way that if $x, y \in B(z,\delta_z)$ then
$$\|\varphi(y) - \varphi(x)\| \leq \frac{\delta(z)}{4} \leq \frac{\delta(x)}{2}.$$
The existence of such a diffeomorphism h follows by Corollary 3.32 from Chapter 3. Then, the mapping $\varphi(x) := g(h(x))$ has no critical point, is equal to f outside U and satisfies $\|f(x) - \varphi(x)\| \leq \varepsilon(x)$ for all $x \in E$.

Let us comment how the rest of the chapter is structured. In Section 5.2 we present some notations and definitions. In Section 5.3 we show a key lemma taken from [128, 20]. In Sections 5.4, 5.5 we prove two results that correspond to Step 1 above and are in correspondence with the hypothesis of Theorem 5.2 and 5.3 respectively. Finally in Section 5.6 we give the proof of our main Theorems 5.2 and 5.3, concluding Step 2 commented above, and in the last Section 5.7 we present some easy corollaries in relation with the failure of Rolle's theorem in infinite dimensions.

5.2 Some notations and definitions

Before starting the next sections let us fix now some notations and definitions.

We call $\{e_n\}_{n\in\mathbb{N}}$ the unconditional basis of E and $\{e_n^*\}_{n\in\mathbb{N}}$ the associated biorthogonal functionals. Let also $P_n : E \to \text{span}\{e_1, \ldots, e_n\}$ be the natural projections defined as $P_n(\sum_{j=1}^{\infty} x_j e_j) = \sum_{j=1}^{n} x_j e_j$ and let K_u be the unconditional constant for the basis. This constant is defined to be the least number such that for every $\{\varepsilon_j\}_{j=1}^n \subset \{-1, +1\}$ and every $\sum_{j=1}^n x_j e_j \in E$,

$$||\sum_{j=1}^{n} \varepsilon_j x_j e_j|| \leq K_u ||\sum_{j=1}^{n} x_j e_j||.$$

Note that $||P_n|| \leq K_u$ for every $n \in \mathbb{N}$. We shall not confuse K_u with the suppression unconditional constant K_s, defined as the least number such that for all (equivalent finite) set $A \subset \mathbb{N}$, $\|P_A\| \leq K_s$, where P_A represents the projection $P_A(x) = \sum_{j\in A} x_j e_j$. We have the relation $K_s \leq K_u \leq 2K_s$. Observe also that in the statement of Theorem 5.3 it is required that $K_s = 1$.

We say that the norm $|| \cdot ||$ locally depends on finitely many coordinates (LFC) if for every $x \in E$ there exists a natural number l_x, an open neighbourhood U_x of x, some functionals $L_1, \ldots, L_{l_x} \in E^*$ and a function $\gamma : \mathbb{R}^{l_x} \to \mathbb{R}$ such that

$$||y|| = \gamma(L_1(y), \ldots, L_{l_x}(y))$$

for every $y \in U_x$. In particular we will make use of the fact that if the norm is of class C^1 and we take $v \in \bigcap_{j=1}^{l_x} \text{Ker}\, L_j$, then

$$D|| \cdot ||(y)(v) = \lim_{t\to 0} \frac{||y + tv|| - ||y||}{t} = 0,$$

for every $y \in U_x \setminus \{0\}$.

Reacall from previous chapters that a function $h : E \to E$ is said to be limited by an open cover $\mathcal{G}$ if the set $\{\{x, h(x)\} : x \in E\}$ refines $\mathcal{G}$; that is, for every $x \in E$, we may find a $G_x \in \mathcal{G}$ such that both x and $h(x)$ are in G_x.

Also recall, from the end of Section 3.4, that when we say that a closed set $X \subset E$ is locally contained in a finite union of complemented subspaces of infinite codimension we mean that for every $x \in X$ there exists an open neighbourhood U_x of x and some closed subspaces $E_1, \ldots, E_{n_x} \subset E$ complemented in E and of infinite codimension such that

$$X \cap U_x \subset \bigcup_{j=1}^{n_x} E_j.$$

Finally for a C^1 function $f : E \to \mathbb{R}^d$, where $f = (f^1, \ldots f^d)$, we write its Fréchet derivative at a point $x \in E$ by $Df(x) = (Df^1(x), \ldots, Df^d(x)) : E \to \mathbb{R}^d$, where each $Df^i(x)$ is a continuous linear functional on E. If f is $\mathbb{R}$-valued we sometimes simply write $f'(x)$ for its derivative.

We will also use indistinctly the symbol $|| \cdot ||$ to denote the norm in E and E^*, and we reserve the notation $| \cdot |$ for the euclidean norm in $\mathbb{R}^d$.

5.3 Key tool for C^1-fine approximation results

The key for the proofs of our main theorems is to apply a C^1-fine approximating result for the function $f|_U : U \to \mathbb{R}^d$, and this is provided by the results of [128, 20]. If $f : E \to F$ is a C^k function between Banach spaces, and $\varepsilon : E \to (0, \infty)$ is a continuous map, then we say that f is C^k-fine approximated by a C^j function $g : E \to F$, where $j \geq k$, if $||D^{i)} f(x) - D^{i)} g(x)|| \leq \varepsilon(x)$ holds for every $x \in E$ and $i = 0, \ldots, k$. We also say that in this case f can be uniformly approximated in the C^k-fine topology by a C^j function. In the infinite-dimensional case this general problem is far from being solved. If one just want to uniformly approximate a continuous function by a C^k function in the C^0-topology the principal tool to work with is partitions of unity. However if the initial function is C^1 and we want a C^1-fine approximation, this technique fails, because we would need to find a common bound for all the derivatives of the functions of the family composing the partition of unity, which is in general not possible.

The first approach to solve this problem has been done in 1971 by Moulis [128]. She first splits the space in finite-dimensional spaces to apply integral convolution and get good approximations preserving the Lipschitz constants, and then she uses an unconditional basis for gluing together the finite-dimensional approximations. For functions $f : E \to F$ of class C^1 she gets C^1-fine approximations by C^k functions where $E = \ell_p$, $1 < p < \infty$, or $E = c_0$, F is a Banach space and k is the order of smoothness of the space E. In 2003 the paper [20] exploited Moulis' technique to get C^1-fine approximations of functions $f : E \to F$ by C^k functions where E is a Banach space having unconditional basis and a C^k-smooth Lipschitz bump function, and F is a Banach space. One year after Robb Fry [84] introduced the technique of *sup-partitions of unity* (see [93, p. 423] for a precise definition) to show that in a separable Banach space with a C^1 norm, any real-valued, bounded and uniform continuous function can be uniformly approximated by C^1 functions with bounded derivative. A characterization of this new concept by means of componentwise C^k-smooth and bi-Lipschitz embeddings into $c_0(\Gamma)$ was given by Hajek and Johannis in [92]. In that paper they also showed that every separable Banach space admitting a Lipschitz and C^k-smooth bump admits C^k smooth sup-partitions of unity, and for such cases it is established the existence of C^k smooth and Lipschitz approximations of a given Lipschitz function, preserving its Lipschitz constant. They deduce then some C^1-fine approximation results for functions $f : E \to F$ by C^∞ functions for the cases when E is separable and has a C^∞ smooth bump and F is a certain Banach space, as for example a Banach space with unconditional Schauder basis and with a separable dual, and also when $E = c_0(\Gamma)$ for some arbitrary set Γ and F is any Banach space.

The theory of fine approximation for higher order classes is practically non-existent. We just mention that in Moulis' paper [128] it is proved that C^{2k-1} functions in Hilbert spaces are uniformly approximated by C^∞ functions in the C^k-fine topology. Moreover it is also known that in the space c_0 we cannot approximate C^2 functions by C^∞ functions in the C^2-fine topology (see [149]).

To finish this brief introduction about C^k-fine approximations in infinite-dimensional Banach spaces let us comment the following result [128, p. 331] due to Moulis, relating C^1-fine approximations with approximate Morse-Sard type results: for every C^1 function $f : E \to F$, where E is an infinite-dimensional Hilbert space and F is a Hilbert space, and every continuous function $\varepsilon : E \to (0, \infty)$ there exists a C^∞ function $g : E \to F$ such that $||f(x) - g(x)|| \leq \varepsilon(x)$, $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ for every $x \in E$ and such that $g(C_g)$ has empty interior in F.

Here, for our purposes, our aim is not to gain regularity in the approximation function but to have a method of approximating a function and its derivative, of which we know nothing, by another one for which we can understand and study better its set of critical points. In fact it will be made clear that the derivative that provides Moulis' approximation technique is a good one, in which the set of critical points will be inside a diffeomorphically extractible set, for which our results of Chapter 3 will apply.

Therefore in this section we present the following lemma that is an easier and slightly different version of [20, Lemma 5] (the arguments will be similar to those of [128] as well).

Lemma 5.4. *Let E and F be a Banach spaces. Suppose that E is infinite-dimensional and has a K_u-unconditional basis and a C^1 equivalent norm. Take an open set U of E. For every open ball*

$B_0 = B(z_0, r_0)$ with $B(z_0, 2r_0) \subseteq U$, and for every C^1 function $f_1 : U \to F$ and numbers $\varepsilon, \eta > 0$ with $\sup_{x \in B(z_0, 2r_0)} ||Df_1(x)|| < \eta$, there exists a C^1 function $\Psi : E \to E$ such that for $f_2 := f_1 \circ \Psi$, we have

(1) $\sup_{x \in B_0} ||f_1(x) - f_2(x)|| < \varepsilon$.

(2) $\sup_{x \in B_0} ||Df_2(x)|| < (K_u)^2 8\eta$.

(3) *For every $x \in E$ there exists $n_0 \in \mathbb{N}$ and a neighbourhood V_0 of x such that*

$$D\Psi(y)(v) = \sum_{n=1}^{n_0} \left[a_n(y) D|| \cdot ||(y - P_{n-1}(y))(v - P_{n-1}(v)) y_n + \xi_n(y) v_n\right] e_n$$

for every $v = \sum_{n=1}^{\infty} v_n e_n \in E$ and $y \in V_0$, where $\xi_n, a_n : V_0 \to \mathbb{R}$ are C^1 functions.

Proof. Choose $0 < r < \min\{\frac{\varepsilon}{K_u \eta}, \frac{r_0}{K_u}\}$. Let $\varphi : \mathbb{R} \to [0,1]$ be a C^∞ smooth function such that $\varphi(t) = 1$ if $|t| < \frac{1}{2}$, $\varphi(t) = 0$ if $|t| > 1$ and $\varphi'(\mathbb{R}) \subseteq [-3, 0]$.
For every $n \in \mathbb{N}$ we define the functions $\xi_n : E \to \mathbb{R}$ and $\Psi : E \to E$,

$$\xi_n(x) = 1 - \varphi\left(\frac{||x - P_{n-1}(x)||}{r}\right),$$
$$\Psi(x) = \sum_{n=1}^{\infty} \xi_n(x) x_n e_n,$$

where $x = \sum_{n=1}^{\infty} x_n e_n \in E$. We denote by P_0 the zero operator.

Fact 5.5. *The mapping $\Psi : E \to \operatorname{span}\{e_n : n \in \mathbb{N}\}$ is well-defined, C^1 smooth on E, and has the following properties:*

(i) $||D\Psi(x)|| \leq (K_u)^2 8$ *for all $x \in E$;*

(ii) $||x - \Psi(x)|| \leq K_u r$ *for all $x \in E$;*

(iii) $\Psi(B_0) \subseteq B(z_0, 2r_0)$.

Proof. For any $x \in E$, because $P_n(x) \to x$ and the $||P_n||$ are uniformly bounded, there exists a neighbourhood V_0 of x and an $n_0 \in \mathbb{N}$ such that $\xi_n(y) = 0$ for all $y \in V_0$ and $n > n_0$, and so $\Psi(V_0) \subset \operatorname{span}\{e_1, \dots, e_{n_0}\}$. Thus $\Psi : E \to \bigcup_{n=1}^{\infty} \operatorname{span}\{e_1, \dots, e_n\}$ is a well-defined C^1 smooth map. We next compute and estimate its derivative.
We have that

$$(\xi_n(y) y_n)' = \xi_n'(y) y_n + \xi_n(y) e_n^*.$$

If $v \in E$ and $y \in V_0$

$$\begin{aligned}\xi_n'(y)(v) &= -\varphi'\left(\frac{||y - P_{n-1}(y)||}{r}\right) D|| \cdot ||(y - P_{n-1}(y))(v - P_{n-1}(v)) r^{-1} \\ &= a_n(y) D|| \cdot ||(y - P_{n-1}(y))(v - P_{n-1}(v)),\end{aligned}$$

where $a_n : E \to \mathbb{R}$ are C^1 functions, defined by $a_n(y) = -\varphi'\left(\frac{||y - P_{n-1}(y)||}{r}\right) r^{-1}$.
Looking at the expression of Ψ we compute its derivative for every $y \in V_0$,

$$\begin{aligned}D\Psi(y)(v) &= \sum_{n=1}^{n_0} \left[\xi_n'(y)(v) y_n + \xi_n(y) v_n\right] e_n \\ &= \sum_{n=1}^{n_0} \left[a_n(y) D|| \cdot ||(y - P_{n-1}(y))(v - P_{n-1}(v)) y_n + \xi_n(y) v_n\right] e_n.\end{aligned}$$

Observe that we have proved (3) of Lemma 5.4.
Now since $|\varphi'(t)| \leq 3$, $||(I - P_{n-1})'(y)|| \leq 1 + K_u$ and the derivative of the norm always has norm one, for all y and all n we get that

$$||\xi_n'(y)|| \leq \left|\varphi'\left(\frac{||y - P_{n-1}(y)||}{r}\right)\right| r^{-1}||(I - P_{n-1})'(y)|| \leq 3(1 + K_u)r^{-1}.$$

For a fixed x, define $n_1 = n_1(x)$ to be the smallest integer with $||x - P_{n_1-1}(x)|| \leq r$. Then for any $m < n_1$, $\xi_m(x) = 1$ and $\xi_m'(x) = 0$, and so, for every $v \in B(0,1)$,

$$\begin{aligned}
||D\Psi(x)(v)|| &\leq ||\sum_{n=n_1}^{\infty} \xi_n'(x)(v)x_ne_n|| + ||\sum_{n=1}^{\infty} \xi_n(x)v_ne_n|| \\
&\leq K_u \sup_{n_1 \leq n} |\xi_n'(x)(v)|\,||\sum_{n=n_1}^{\infty} x_ne_n|| + K_u \sup_{n} |\xi_n(x)|\,||\sum_{n=1}^{\infty} v_ne_n|| \\
&\leq 3K_u(1 + K_u)r^{-1}||\sum_{n=n_1}^{\infty} x_ne_n|| + K_u \leq (3 + 4K_u)K_u < 8(K_u)^2,
\end{aligned}$$

proving (i).
We next estimate $||x - \Psi(x)||$.

$$||x - \Psi(x)|| = ||\sum_{n \geq n_1} x_n(1 - \xi_n(x))e_n|| \leq K_u \sup_n |1 - \xi_n(x)|\,||\sum_{n \geq n_1} x_ne_n|| \leq K_u r \leq r_0,$$

which proves (ii). Lastly, property (iii) is immediate from (ii) and the choice of r. □

Going back to the proof of the Lemma 5.4 define

$$f_2(x) := f_1(\Psi(x)),$$

which is a C^1 function. Firstly we have that for every $x \in B_0$,

$$||f_1(x) - f_2(x)|| \leq \eta||x - \Psi(x)|| \leq \eta K_u r < \varepsilon,$$

using the Lipschitzness of f_1 in $B(z_0, 2r_0)$.
Secondly for every $x \in B_0$,

$$||Df_2(x)|| \leq ||Df_1(\Psi(x))||\,||D\Psi(x)|| \leq \eta(K_u)^2 8.$$

The proof of the lemma is now complete.

□

5.4 *Special* C^1-fine approximation on Banach spaces with a C^1, LFC equivalent norm and unconditional basis

We intend to prove the following theorem, following the ideas of the papers [128, 20].

Theorem 5.6. *Let E be an infinite-dimensional Banach space with an unconditional basis and with a C^1 equivalent norm that locally depends on finitely many coordinates. Let U be an open subset of E, $f : U \to \mathbb{R}^d$ a C^1 function and $\varepsilon : U \to (0, \infty)$ a continuous function. Then there exists a C^1 function $g : U \to \mathbb{R}^d$ such that*

(1) $|f(x) - g(x)| \leq \varepsilon(x)$ *for every* $x \in U$.

(2) $||Df(x) - Dg(x)|| \leq \varepsilon(x)$ *for every* $x \in U$.

(3) $C_g = \emptyset$, *i.e. g has no critical points.*

Proof of Theorem 5.6. Using the openness of U, the continuity of ε and f', the separability of E and the assumption that the norm $||\cdot||$ locally depends on finitely many coordinates, we find a covering

$$\bigcup_{j=1}^{\infty} B(x^j, r_j) = U$$

of U such that

(i) $B(x^j, 4r_j) \subset U$ with $r_j \leq 1$ for every $j \in \mathbb{N}$.

(ii) $\varepsilon(x) \geq \frac{\varepsilon(x^j)}{2}$ for every $x \in B(x^j, 2r_j)$.

(iii) $||Df(x) - Df(x^j)|| \leq \frac{\varepsilon(x^j)}{(K_u)^2 72}$ for every $x \in B(x^j, 4r_j)$.

(iv) For every $j \in \mathbb{N}$ there exist a number $l_j \in \mathbb{N}$, some linear functionals $L_{j(1)}, \dots, L_{j(l_j)}$, and a C^1 function $\gamma_j : \mathbb{R}^{l_j} \to \mathbb{R}$ such that

$$||y|| = \gamma_j(L_{j(1)}(y), \dots, L_{j(l_j)}(y))$$

for every $y \in B(x^j, 2r_j)$.

Now for every $j \in \mathbb{N}$ choose functions $\varphi_j \in C^\infty(E; [0,1])$ with bounded derivative so that $\varphi_j(x) = 1$ for $x \in B(x^j, r_j)$ and $\varphi_j(x) = 0$ for $x \notin B(x^j, 2r_j)$. We precisely take $\varphi_j(x) = \theta_j(||x - x^j||)$ where $\theta_j : \mathbb{R} \to [0,1]$ is C^∞ and $\theta_j^{-1}(1) = (-\infty, r_j]$ and $\theta_j^{-1}(0) = [2r_j, \infty)$. It must be noted here that despite the fact that the norm $||\cdot||$ is not differentiable at the origin, the functions φ_j are C^1 for every $x \in E$ because in a neighbourhood of x^j they are constantly one.
We introduce the following constants,

$$\tilde{M}_k = \sup_{x \in B(x^k, 2r_k)} ||\varphi_k'(x)||,$$

$$M_j = \max\{1, \sum_{k=1}^{j} \tilde{M}_k\}.$$

Next define for every $j \in \mathbb{N}$,

$$h_j = \varphi_j \prod_{k<j} (1 - \varphi_k).$$

One can easily check that we have the following properties:

- For every $x \in U$ there exists $n_x = \min\{m \in \mathbb{N} : x \in B(x^m, r_m)\}$ such that $1 - \varphi_{n_x}(x) = 0$ and hence $h_m(y) = 0$ for every $m > n_x$ and $y \in B(x^{n_x}, r_{n_x})$.
- $\sum_{j=1}^{\infty} h_j(x) = 1$ for every $x \in U$.
- $||h_j'(x)|| \leq M_j$ for every $j \in \mathbb{N}$ and $x \in B(x^j, 2r_j)$.

In particular $\{h_j\}_{j \in \mathbb{N}}$ is a C^1 partition of unity which is subordinate to $\{B(x^j, 2r_j)\}_{j \in \mathbb{N}}$.
For every $j \in \mathbb{N}$ we apply the previous Lemma 5.4 for each ball $B(x^j, 2r_j)$, the function $f_1(x) = f(x^j) + Df(x^j)(x - x^j) - f(x)$ and the constants $\frac{\varepsilon(x^j)}{2^{j+3} M_j}$ and $\frac{\varepsilon(x^j)}{(K_u)^2 72}$ for ε and η respectively. Note that we can apply the Lemma 5.4 because

$$\sup_{x \in B(x^j, 4r_j)} ||Df_1(x)|| = \sup_{x \in B(x^j, 4r_j)} ||Df(x^j) - Df(x)|| \leq \frac{\varepsilon(x^j)}{(K_u)^2 72}.$$

The resulting functions from the proof of the lemma will be called $\delta_j = f_1 \circ \Psi_j$. In particular we have

$$||f(x^j) + Df(x^j)(x - x^j) - \delta_j(x) - f(x)|| \leq \frac{\varepsilon(x^j)}{2^{j+3}M_j} \tag{5.4.1}$$

and

$$||D\delta_j(x)|| \leq 8\frac{\varepsilon(x^j)}{72}, \tag{5.4.2}$$

for every $x \in B(x^j, 2r_j)$. Let us define finally

$$g(x) := \sum_{j=1}^{\infty} h_j(x)(f(x^j) + Df(x^j)(x - x^j) - \delta_j(x) + T_j(x - x^j)), \tag{5.4.3}$$

where $T_j : E \to \mathbb{R}^d$ is a continuous linear surjective operator which we next construct. Define $T_j = (T_j^1, \ldots, T_j^d)$ inductively such that for each $i = 1, \ldots, d$, T_j^i is a non-null element of E^* satisfying that

$$T_j^i \notin \text{span}\{e_n^*, Df^k(x^n), L_{n(1)}, \ldots, L_{n(l_n)}, T_1^k, \ldots, T_{j-1}^k, T_j^1, \ldots, T_j^{i-1} : n \in \mathbb{N},\ 1 \leq k \leq d\}$$

(note that it is the span, not the closed span); which can never fill the whole space E^* because Banach spaces of infinite dimension cannot have a countable Hamel basis. We also impose that their norms are small enough, more precisely,

$$||T_j|| \leq \varepsilon(x^j)M_j^{-1}2^{-j-4} \leq \frac{\varepsilon(x^j)}{8}. \tag{5.4.4}$$

An important property that derives from this definition of T_j is that the set $\{T_j^1, \ldots, T_j^d\}$ is linearly independent and hence $T_j : E \to \mathbb{R}^d$ is a surjective linear operator. We also have that

$$T_j^i \notin \text{span}\{e_n^*, Df^k(x^n), L_{n(1)}, \ldots, L_{n(l_n)}, T_1^k, \ldots, T_{j-1}^k, T_j^p : n \in \mathbb{N},\ 1 \leq k \leq d,\ 1 \leq p \leq d, p \neq i\}.$$

Using the expression (5.4.3) let us check that properties $(1), (2)$ and (3) of the statement of the main theorem are satisfied for this choice of T_j^i.
Firstly if $h_j(x) \neq 0$, then $x \in B(x^j, 2r_j)$ and

$$\begin{aligned}|f(x^j) + Df(x^j)(x - x^j) &- \delta_j(x) + T_j(x - x^j) - f(x)| \\ &\leq |f(x^j) + Df(x^j)(x - x^j) - \delta_j(x) - f(x)| + |T_j(x - x^j)|) \\ &\leq \frac{\varepsilon(x^j)}{2^{j+3}M_j} + \frac{\varepsilon(x^j)2r_j}{8} \leq \frac{\varepsilon(x^j)}{2} \leq \varepsilon(x).\end{aligned}$$

Therefore for every $x \in U$,

$$\begin{aligned}|g(x) - f(x)| &= |\sum_{j=1}^{\infty} h_j(x)(f(x^j) + Df(x^j)(x - x^j) - \delta_j(x) + T_j(x - x^j) - f(x))| \\ &\leq \varepsilon(x)\sum_{j=1}^{n} h_j(x) = \varepsilon(x).\end{aligned}$$

We have proved (1).
In order to show (2) and (3), let us analyse what the derivative of g looks like, and inspect its critical set.

Claim 5.7. *For every $x \in U$ there exist $n, k_1, \ldots, k_n \in \mathbb{N}$ and a neighbourhood $V_x = V \subset B(x^n, r_n)$ of x such that:*

(i) For every $y \in B(x^n, r_n)$,

$$g(y) = \sum_{j=1}^{n} h_j(y)(f(x^j) + Df(x^j)(y - x^j) - \delta_j(y) + T_j(y - x^j)), \tag{5.4.5}$$

and

$$\begin{aligned} Dg(y) = &\sum_{j=1}^{n} h_j'(y) \left[f(x^j) + Df(x^j)(y - x^j) - \delta_j(y) + T_j(y - x^j)\right] \\ &+ \sum_{j=1}^{n} h_j(y) \left[Df(x^j) - D\delta_j(y) + T_j\right]. \end{aligned} \tag{5.4.6}$$

(ii) For every $y \in V$ and $1 \le j \le n$, $D\delta_j(y)(v) = Df(\Psi_j(y)) \circ (D\Psi_j(y)(v))$ has the form

$$Df(\Psi_j(y)) \circ \left\{ \sum_{n=1}^{k_j} \left[a_n^j(y) D|| \cdot ||(y - P_{n-1}(y))(v - P_{n-1}(v)) y_n + \xi_n^j(y) v_n\right] e_n \right\}. \tag{5.4.7}$$

Proof. Recall that for every $x \in U$ there is $n_x = n = \min\{m \in \mathbb{N} : x \in B(x^m, r_m)\}$ such that $h_m(y) = 0$ for every $m > n$ and every $y \in B(x^n, r_n)$. So expression (5.4.3) becomes

$$g(y) = \sum_{j=1}^{n} h_j(y)(f(x^j) + Df(x^j)(y - x^j) - \delta_j(y) + T_j(y - x^j))$$

for all $y \in B(x^n, r_n)$. Computing the derivative we get

$$Dg(y) = \sum_{j=1}^{n} h_j'(y) \left[f(x^j) + Df(x^j)(y - x^j) - \delta_j(y) + T_j(y - x^j)\right] + h_j(y) \left[Df(x^j) - D\delta_j(y) + T_j\right]$$

for every $y \in B(x^n, r_n)$.
For every $j = 1, \dots, n$, by (3) of Lemma 5.4, we can find a neighbourhood $V_{x,j} \subset B(x^n, r_n)$ of x and a number k_j such that for every $y \in V_{x,j}$,

$$\begin{aligned} &D\delta_j(y)(v) = Df(\Psi_j(y)) \circ (D\Psi_j(y)(v)) \\ &Df(\Psi_j(y)) \circ \left\{ \textstyle\sum_{n=1}^{k_j} \left[a_n^j(y) D|| \cdot ||(y - P_{n-1}(y))(v - P_{n-1}(v)) y_n + \xi_n^j(y) v_n\right] e_n \right\}. \end{aligned}$$

Define then $V_x := \bigcap_{j=1}^{n} V_{x,j} \subset B(x^n, r_n)$. □

Using equation (5.4.6) of Claim 5.7, we can write

$$\begin{aligned} ||Dg(x) - Df(x)|| \le &|| \sum_{j=1}^{n} h_j'(x) \left(f(x^j) + Df(x^j)(x - x^j) - \delta_j(x) + T_j(x - x^j) - f(x)\right) || + \\ &+ || \sum_{j=1}^{n} h_j(x)(Df(x^j) - D\delta_j(x) + T_j - Df(x)) || \le \\ \le &\sum_{j=1}^{n} ||h_j'(x)|| \left(||f(x^j) + Df(x^j)(x - x^j) - \delta_j(x) - f(x)|| + ||T_j(x - x^j)||\right) + \\ &+ \sum_{j=1}^{n} h_j(x) \left(||Df(x^j) - Df(x)|| + ||D\delta_j(x)|| + ||T_j||\right) \end{aligned}$$

for every $x \in U$. Let us try to estimate all these quantities. Applying inequality (5.4.1) and the bound of $||T_j||$ given by (5.4.4) we get

$$||f(x^j) + Df(x^j)(x - x^j) - \delta_j(x) - f(x)|| + ||T_j(x - x^j)|| \leq \frac{\varepsilon(x^j)}{2^{j+3}M_j} + \frac{\varepsilon(x^j)2r_j}{2^{j+4}M_j}$$

for every $x \in B(x^j, 2r_j)$. On the other hand $||Df(x^j) - Df(x)|| \leq \frac{\varepsilon(x^j)}{(K_u)^2 72} \leq \frac{\varepsilon(x^j)}{72}$ by our choice of the partition of unity, and using (5.4.2) and again (5.4.4) we have that for every $x \in B(x^j, 2r_j)$,

$$||Df(x^j) - Df(x)|| + ||D\delta_j(x)|| + ||T_j|| \leq \frac{\varepsilon(x^j)}{72} + 8\frac{\varepsilon(x^j)}{72} + \frac{\varepsilon(x^j)}{8} = \frac{\varepsilon(x^j)}{4}.$$

We also know that the norm of $h_j'(x)$ is bounded by M_j as was indicated when stating the properties of the partition of unity. This fact together with these previous computations allow us to conclude that

$$\begin{aligned} ||Dg(x) - Df(x)|| \leq & \sum_{j=1}^{n} M_j \left(\frac{\varepsilon(x^j)}{2^{j+3}M_j} + \frac{\varepsilon(x^j)2r_j}{2^{j+4}M_j} \right) + \sum_{j=1}^{n} h_j(x) \left(\frac{\varepsilon(x^j)}{4} \right) \leq \\ \leq & \sum_{j=1}^{n} \frac{\varepsilon(x^j)}{2^{j+2}} + \sum_{j=1}^{n} h_j(x)\frac{\varepsilon(x^j)}{4} \leq \frac{\varepsilon(x)}{2} + \frac{\varepsilon(x)}{2} = \varepsilon(x) \end{aligned}$$

for every $x \in U$. We have then proved (2) of Theorem 5.6.

Let us focus now on studying the critical set of points of g.
Use Claim 5.7 to choose a vector $x \in U$ for which there exist numbers $n, k_1, \ldots, k_n$ and a neighbourhood $V = V_x \subset B(x_n, r_n)$ such that (i) and (ii) of the claim hold. Define also

$$\tilde{n} := \max\{n, k_1, \ldots, k_n\}.$$

Take $(t_1, \ldots, t_d) \in \mathbb{R}^d$ and $y \in V$. Our goal is to find a vector $v \in E$ such that $Dg(y)(v) = (t_1, \ldots, t_d)$. Once we prove this we will get (3) of Theorem 5.6.

With $y \in V$ fixed, looking at the formula (5.4.7) of Claim 5.7, we are interested in the expression of the bounded linear operators $h_j'(y), D\delta_j(y), Df(x^j), T_j$ for $j = 1, \ldots, n$. Let $m = m_y$ be the least number such that $y \in B(x^m, r_m)$, hence $h_m(y) \neq 0$ (observe that necessarily $m \leq n$), then we write equation (5.4.6) as

$$Dg(y) = \sum_{j=1}^{m} h_j'(y) \left[f(x^j) + Df(x^j)(y - x^j) - \delta_j(y) + T_j(y - x^j) \right] + h_j(y) \left[Df(x^j) - D\delta_j(y) + T_j \right].$$

We want to find a vector $v \in E$ for which

$$\begin{cases} h_j'(y)(v) = 0 \quad \text{for every } \ 1 \leq j \leq m, \\ D\delta_j(y)(v) = (0, \ldots, 0); \ \text{ for every } \ 1 \leq j \leq m, \\ Df(x^j)(v) = (Df^1(x^j)(v), \ldots, Df^d(x^j)(v)) = (0, \ldots, 0), \quad \text{for every } \ 1 \leq j \leq m, \\ T_j(v) = (0, \ldots, 0) \ \text{ for every } \ 1 \leq j < m, \\ h_m(y)T_m(v) = (t_1, \ldots, t_d). \end{cases}$$

Let us pay attention to the vectors $y - x^j$ and $y - P_{i-1}(y)$, for $1 \leq j \leq m$ and $1 \leq i \leq \tilde{n}$. For simplicity let us rename these vectors as $\{z^1, \ldots, z^{k_0}\}$. Each of these elements z^k, $1 \leq k \leq k_0$, belongs to some ball $B(x^{k'}, 2r_{k'})$ (for each k we associate a unique k', not necessarily equal to k). So by using property (iv) from the beginning of the proof there exists a finite number of continuous linear functionals $\{L_{k'(1)}, \ldots, L_{k'(l_{k'})}\}$ and a C^1 function $\gamma_{k'} : \mathbb{R}^{l_{k'}} \to \mathbb{R}$ such that

$$||z^k|| = \gamma_{k'}(L_{k'(1)}(y), \ldots, L_{k'(l_{k'})}(y)).$$

We intend to take a vector $v \in \bigcap_{j=1}^{l_{k'}} \text{Ker } L_{k'(j)}$, so that $D||\cdot||(z^k)(v) = 0$ for every $k = 1, \dots, k_0$. For every $i = 1, \dots, d$, let us introduce the finite set of functionals

$$A_i := \{e_1^*, \dots, e_{\tilde{n}}^*\} \cup \{Df^j(x^1), \dots, Df^j(x^m) : 1 \le j \le d\} \cup \{L_{k'(1)}, \dots, L_{k'(l_{k'})} : 1 \le k \le k_0\}$$
$$\cup \{T_1^j, \dots, T_{m-1}^j : 1 \le j \le d\} \cup \{T_m^j : 1 \le j \le d, j \ne i\}.$$

By the definition of T_m^i we have that $T_m^i \notin \text{span}\,(A_i)$, which is equivalent to $\bigcap_{a^* \in A_i} \text{Ker } a^* \subsetneq \text{Ker } T_m^i$. Therefore there exists an element $w^i \in E$ such that $T_m^i(w^i) \ne 0$ and $a^*(w^i) = 0$ for every $a^* \in A_i$. For every $i = 1, \dots, d$, take $v^i = \frac{t_i w^i}{h_m(y) T_m^i(w^i)}$ and define $v := v^1 + \dots + v^d$, so we have

$$h_m(y)T_m(v) = h_m(y)(T_m^1(v), \dots T_m^d(v)) = (h_m(y)T_m^1(v^1), \dots, h_m(y)T_m^d(v^d)) = (t_1, \dots, t_d).$$

Moreover, $D||\cdot||(y - x^j)(v) = 0$ for every $1 \le j \le m$, $D||\cdot||(y - P_{i-1}(y))(v) = 0$ for every $1 \le i \le \tilde{n}$, and $Df(x^j)(v) = (Df^1(x^j)(v), \dots, Df^d(x^j)(v)) = (0, \dots, 0)$ for every $1 \le j \le m$. Furthermore, writing v in coordinates, $v = \sum_{j=1}^{\infty} v_j e_j$ we have that $v_1 = \dots = v_{\tilde{n}} = 0$.
Recall that $h_j(y) = \theta_j(||y - x^j||) \prod_{k<j}(1 - \theta_k(||y - x^k||))$, so

$$h_j'(y)(\cdot) = \sum_{k=1}^{j} \gamma_{k,j}(y) D||\cdot||(y - x^k)(\cdot),$$

where $\gamma_{k,j} : E \to \mathbb{R}$ are C^1 functions. Hence with our choice of v we have $h_j'(y)(v) = 0$ for every $1 \le j \le m$.
On the other hand, looking at formula (5.4.7) of Claim 5.7, we also get $D\delta_j(v) = 0$ for every $1 \le j \le m$. Finally we also have $T_j(v) = (T_j^1, \dots, T_j^d(v)) = (0 \dots, 0)$ for every $j < m$, because $T_j^1, \dots, T_j^d \in \bigcap_{i=1}^{d} A_i$ for every $j < m$.
Putting all these facts together, we have proved that $Dg(y)(v) = (t_1, \dots, t_d)$ and consequently the critical set of points of g is empty. □

5.5 *Special* C^1-fine approximation on Banach spaces with 1-suppression unconditional basis

Here we will prove the following.

Theorem 5.8. *Let E be an infinite-dimensional Banach space with a C^1 strictly convex equivalent norm and with a* 1*-suppression unconditional basis (in particular K_u-unconditional with $1 \le K_u \le 2$). Let U be an open subset of E, $f : U \to \mathbb{R}^m$ a C^1 function and $\varepsilon : U \to (0, \infty)$ a continuous function. Then there exists a C^1 function $g : U \to \mathbb{R}^m$ such that:*

(1) $|f(x) - g(x)| \le \varepsilon(x)$ *for every* $x \in U$.

(2) $||Df(x) - Dg(x)|| \le \varepsilon(x)$ *for every* $x \in U$.

(3) C_g *is locally contained in a finite union of complemented subspaces of infinite codimension in* E.

The essence of the proof will be close to the one of the previous section, hence following [128, 20] as well. However there are some important changes. Here we do not rely on a norm that locally depends on finitely many coordinates, but on the property of the basis of being 1-suppression unconditional, which will provide us with the necessary tools to approximate the function f and its derivative f' by another function with a *small* critical set of points.

Proof of Theorem 5.8. E has a separable dual, so it does not contain copies of l_1 and since it has an unconditional basis, by [119, Theorem 1.c.9] we know that the basis is also shrinking, that is, $\overline{\text{span}}\{e_n^* : n \in \mathbb{N}\} = E^*$.

Using the openness of U, the continuity of ε and Df, and the facts that $\overline{\text{span}}\{e_n : n \in \mathbb{N}\} = E$ and $\overline{\text{span}}\{e_n^* : n \in \mathbb{N}\} = E^*$, we find a covering

$$\bigcup_{j=1} B(x^j, r_j) = U$$

of U and continuous linear functionals $F_j : E \to \mathbb{R}^d$ for every $j \in \mathbb{N}$ such that:

(i) $B(x^j, 4r_j) \subset U$ with $r_j \leq 1$ for every $j \in \mathbb{N}$.

(ii) $\varepsilon(x) \geq \frac{\varepsilon(x^j)}{2}$ for all $x \in B(x^j, 2r_j)$.

(iii) $||Df(x) - Df(x^j)|| \leq \frac{\varepsilon(x^j)}{(K_u)^2 144}$ for every $x \in B(x^j, 4r_j)$.

(iv) $||F_j - Df(x^j)|| \leq \frac{\varepsilon(x^j)}{(K_u)^2 144}$.

(v) For every $j \in \mathbb{N}$,

$$\begin{cases} x^j = \sum_{i=1}^{N_j} \alpha_{i,j} e_i, \\ F_j = (F_j^1, \ldots, F_j^d) = (\sum_{i=1}^{N_j} \beta_{i,j}^1 e_i^*, \ldots, \sum_{i=1}^{N_j} \beta_{i,j}^d e_i^*). \end{cases}$$

for some $\alpha_{1,j}, \ldots, \alpha_{N_j,j}, \beta_{1,j}^q, \ldots, \beta_{N_j,j}^q \in \mathbb{R}$, $1 \leq q \leq d$, where $N_1 \leq N_2 \leq \ldots$ is an increasing sequence of natural numbers. Note that we allow some $\alpha_{i,j}$ or $\beta_{i,j}^q$ to be null.

At this point we proceed exactly as in the previous section, defining the C^1 partition of unity $\{h_j\}_{j\geq 1}$ subordinate to $\{B(x^j, 2r_j)\}_{j\geq 1}$, and also the constants $\widetilde{M}_k$ and M_k. We also apply Lemma 5.4, exactly in the same way as before, but now to the function $f_1(x) = f(x^j) + F_j(x - x^j) - f(x)$ and the constants $\frac{\varepsilon(x^j)}{2^{j+3} M_j}$ and $\frac{\varepsilon(x^j)}{(K_u)^2 72}$ for ε and η respectively, obtaining $\delta_j = f \circ \Psi_j$.
We define finally

$$g(x) := \sum_{j=1}^{\infty} h_j(x)(f(x^j) + F_j(x - x^j) - \delta_j(x) + T_j(x - x^j)), \tag{5.5.1}$$

where $T_j : E \to \mathbb{R}^d$ is a continuous linear surjective operator that will be defined in the following paragraph.
Choose a family of pairwise disjoint subsets $\{I_n\}_{n\geq 1}$ of natural numbers such that each $I_n \subset \mathbb{N}$ has infinite elements and, if we denote $\mathbb{I} = \bigcup_{n\geq 1} I_n$, then $\mathbb{N} \setminus \mathbb{I}$ is infinite. Write also $I_n = I_n^1 \cup \cdots \cup I_n^d$ as a pairwise disjoint union of sets, each of them having again infinite elements. For every $j \in \mathbb{N}$ and $i = 1, \ldots, d$ we choose $T_j^i \in E^*$ satisfying that

$$T_j^i \in \overline{\text{span}}\{e_n^* : n \in I_j^i\} \setminus \text{span}\{e_n^* : n \in I_j^i\}.$$

Define $T_j := (T_j^1, \ldots, T_j^d)$ and also assume with no loss of generality that

$$||T_j|| \leq \varepsilon(x^j) M_j^{-1} 2^{-j-4} \leq \frac{\varepsilon(x^j)}{8}.$$

Following the computation made for proving Theorem 5.6 (1) in the previous subsection, we can check that for every $x \in U$,

$$|g(x) - f(x)| = |\sum_{j=1}^{\infty} h_j(x)(f(x^j) - F_j(x - x^j) - \delta_j(x) + T_j(x - x^j) - f(x))| \leq \varepsilon(x) \sum_{j=1}^{n} h_j(x) = \varepsilon(x),$$

which proves (1).
To analyse the derivative of g and its set of critical points in order to show (2) and (3) we also have at our disposal the following.

Claim 5.9. *For every $x \in U$ there exist $n, k_1, \dots, k_n \in \mathbb{N}$ and a neighbourhood $V_x = V \subset B(x^n, r_n)$ of x such that:*

(i) For every $y \in B(x^n, r_n)$,

$$g(y) := \sum_{j=1}^{n} h_j(y)(f(x^j) + F_j(y - x^j) - \delta_j(y) + T_j(y - x^j)), \tag{5.5.2}$$

and

$$\begin{aligned} Dg(y) = &\sum_{j=1}^{n} h_j'(y) \left[f(x^j) + F_j(y - x^j) - \delta_j(y) + T_j(y - x^j)\right] \\ &+ \sum_{j=1}^{n} h_j(y) \left[F_j - D\delta_j(y) + T_j\right]. \end{aligned} \tag{5.5.3}$$

(ii) For every $y \in V$ and $1 \le j \le n$, $D\delta_j(y)(v) = Df(\Psi_j(y)) \circ (D\Psi_j(y)(v))$ has the form

$$Df(\Psi_j(y)) \circ \left\{ \sum_{n=1}^{k_j} \left[a_n^j(y) D|| \cdot ||(y - P_{n-1}(y))(v - P_{n-1}(v)) y_n + \xi_n^j(y) v_n\right] e_n \right\}. \tag{5.5.4}$$

Proof. Follow the proof of Claim 5.7. □

Using equation (5.5.3) of Claim 5.9, a straightforward calculation as in the previous section gives

$$\begin{aligned} ||Dg(x) - Df(x)|| \le &\sum_{j=1}^{n} ||h_j'(x)|| \left(||f(x^j) + F_j(x - x^j) - \delta_j(x) - f(x)|| + ||T_j(x - x^j)||\right) \\ &+ \sum_{j=1}^{n} h_j(x) \left(||F_j - Df(x^j)|| + ||Df(x^j) - Df(x)|| + ||D\delta_j(x)|| + ||T_j||\right) \\ \le &\varepsilon(x) \end{aligned}$$

for every $x \in U$. We have thus proved (2) of Theorem 5.8.
It remains to study the critical set of g.
Take a vector $x \in U$. By Claim 5.9 there exist numbers $n, k_1, \dots, k_n$ and a neighbourhood $V = V_x \subset B(x_n, r_n)$ such that (i) and (ii) of the claim hold. Define also

$$\tilde{n} := \max\{n, N_n, k_1, \dots, k_n\}.$$

Let us divide the set $\mathbb{N} \setminus \mathbb{I} = \mathbb{J}$ in another disjoint infinite family of subsets $\{J_n\}_{n \ge 1}$, each of them having infinite elements. Consider also the set

$$A = \left\{y - x^j, y - P_{i-1}(y) : j = 1, \dots, n \; i = 1 \dots, \tilde{n}\right\}, \tag{5.5.5}$$

and define $k_0 := \dim(\{\mathrm{span}(A)\}) \le n + \tilde{n}$.
In order to establish (3) of Theorem 5.8 our goal is to show that if

$$y \in V \setminus \left(\bigcup_{k=1}^{k_0} \overline{\mathrm{span}}\{e_j : j = 1, \dots, \tilde{n} \text{ or } j \in \mathbb{N} \setminus J_k\} \right),$$

and $t = (t_1, \dots, t_d) \in \mathbb{R}$ then there exists a vector $v \in E$ such that $Dg(y)(v) = t$. Indeed for every $x \in U$ we would have found a neighbourhood $V_x = V$ such that

$$C_g \cap V \subseteq \left(\bigcup_{k=1}^{k_0} \overline{\mathrm{span}}\{e_j : j = 1, \dots, \tilde{n} \text{ or } j \in \mathbb{N} \setminus J_k\} \right).$$

Fix $y \in V \setminus \left(\bigcup_{k=1}^{k_0} \overline{\operatorname{span}}\{e_j : j = 1, \ldots, \tilde{n} \text{ or } j \in \mathbb{N} \setminus J_k\}\right)$ and look at the formula of $Dg(y)$ given by property (i) of Claim 5.9. We are interested in the expression of the continuous linear operators $h_j'(y), F_j, D\delta_j(y), T_j$ for $j = 1, \ldots, n$. Let $m = m_y$ be the least number such that $y \in B(x^m, r_m)$, hence $h_m(y) \neq 0$ (observe that necessarily $m \leq n$), then we may write equation (5.5.3) as

$$Dg(y) = \sum_{j=1}^{m} h_j'(y) \left[f(x^j) + F_j(y - x^j) - \delta_j(y) + T_j(y - x^j)\right] + \sum_{j=1}^{m} h_j(y) \left[F_j - D\delta_j(y) + T_j\right].$$

We need to find a vector $v \in E$ for which

$$\begin{cases} h_j'(y)(v) = 0 \quad \text{for every } \ 1 \leq j \leq m, \\ D\delta_j(y)(v) = (0, \ldots, 0); \ \text{ for every } \ 1 \leq j \leq m, \\ F_j(v) = (F_j^1(v), \ldots, F_j^d(v)) = (0, \ldots, 0), \quad \text{for every } \ 1 \leq j \leq m, \\ T_j(v) = (0, \ldots, 0) \ \text{ for every } \ 1 \leq j < m, \\ h_m(y) T_m(v) = (t_1, \ldots, t_d). \end{cases}$$

By definition of y there exist $j(1), \ldots, j(k_0) > \tilde{n}$ such that $j(1) \in J_1, \ldots, j(k_0) \in J_{k_0}$ and $y_{j(1)}, \ldots, y_{j(k_0)} \neq 0$. Furthermore the vectors $y - x^j$ have their $j(1), \ldots, j(k_0)$th coordinates non-null because we had $x^j \in \operatorname{span}\{e_1, \ldots, e_{N_j}\} \subseteq \operatorname{span}\{e_1, \ldots, e_{N_n}\} \subseteq \operatorname{span}\{e_1, \ldots, e_{\tilde{n}}\}$. This implies that the $j(1), \ldots, j(k_0)$th coordinates of all the vectors in the set A (see expression (5.5.5)) are non-null.

We will need the following:

Fact 5.10. *For every $w = \sum_{j=1}^{\infty} w_j e_j \in E \setminus \{0\}$ and every $j_0 \in \mathbb{N}$ we have that*

$$w_{j_0} \neq 0 \implies D\|\cdot\|(w)(e_{j_0}) \neq 0.$$

Proof. This is a consequence of the facts that the norm is strictly convex and the basis $\{e_n\}_{n \in \mathbb{N}}$ is 1-suppression unconditional. For details see for example [16, Fact 4.5]. □

Consequently we can assure that

$$e_{j(k)} \notin \bigcap_{a \in A} \operatorname{Ker}(D\|\cdot\|(a))$$

for every $1 \leq k \leq k_0$. For every $i = 1, \ldots, d$, let us define $E^i_{(m,\tilde{n})} = \overline{\operatorname{span}}\{e_n : n > \tilde{n} \text{ and } n \in \mathbb{J} \cup I^i_m\}$. Since $k_0 = \mathit{codim}\left(\bigcap_{a \in A} \operatorname{Ker} D\|\cdot\|(a)\right)$, we can write

$$E = \left(\bigcap_{a \in A} \operatorname{Ker} D\|\cdot\|(a)\right) \oplus \operatorname{span}\{e_{j(1)}, \ldots, e_{j(k_0)}\},$$

so

$$E^i_{(m,\tilde{n})} = \left(\bigcap_{a \in A} \operatorname{Ker} D\|\cdot\|(a) \cap E^i_{(m,\tilde{n})}\right) \oplus \operatorname{span}\{e_{j(1)}, \ldots, e_{j(k_0)}\}.$$

On the other hand $e_{j(1)}, \ldots, e_{j(k_0)} \in \operatorname{Ker} T^i_m$ for every $i = 1, \ldots, d$. In particular we can find an element

$$w^i \in \left(\bigcap_{a \in A} \operatorname{Ker} D\|\cdot\|(a) \cap E^i_{(m,\tilde{n})}\right) \setminus \left(\operatorname{Ker} T^i_m\right).$$

Otherwise we would have $\left(\bigcap_{a \in A} \operatorname{Ker} D\|\cdot\|(a) \cap E^i_{(m,\tilde{n})}\right) \subset \operatorname{Ker} T^i_m$ which implies that $T^i_m(w) = 0$ for every $w \in E^i_{(m,\tilde{n})}$, a contradiction with the definition of T^i_m.
Let us now mix all these previous ingredients together. The vector v we are looking for is

$$v := \sum_{i=1}^{d} \frac{t_i w^i}{h_m(y) T^i_m(w^i)}.$$

We obviously have $h_m(y)T_m(v) = h_m(y)(T_m^1(v), \dots, T_m^d(v)) = (t_1, \dots, t_d)$, so it remains to check that $h_j'(v) = 0$, that $D\delta_j(v) = F_j(v) = (0, \dots, 0)$ for every $j = 1, \dots, m$ and that $T_j(v) = (0, \dots, 0)$ for every $j < m$.
For the h_j', recall that $h_j(y) = \theta_j(||y - x^j||) \prod_{k<j}(1 - \theta_k(||y - x^k||))$. So we have that

$$h_j'(y)(\cdot) = \sum_{k=1}^{j} \gamma_{k,j}(y) D||\cdot||(y - x^k)(\cdot),$$

where $\gamma_{k,j} : E \to \mathbb{R}$ are C^1 functions. The elements $y - x^j$ belong to the set A so it is clear that $h_j'(v) = 0$ for every $1 \leq j \leq m$.
For the $D\delta_j$, using (5.5.4) and the facts that the elements $y - P_{i-1}(y)$ belong to the set A and that the coordinates $v_1, \dots, v_{\tilde{n}} = 0$, we conclude that $D\delta_j(v) = 0$ for every $1 \leq j \leq m$.
The fact that $F_j(v) = (0, \dots, 0)$ is clear since

$$F_j = (F_j^1, \dots, F_j^d) = (\sum_{i=1}^{N_j} \beta_{i,j}^1 e_i^*, \dots, \sum_{i=1}^{N_j} \beta_{i,j}^d e_i^*),$$

$N_j \leq N_n \leq \tilde{n}$ for every $j = 1 \dots, m$ and $v_1, \dots, v_{\tilde{n}} = 0$.
Finally we also have $T_j(v) = (0, \dots, 0)$ for every $j < m$, because $v \in \overline{\text{span}}\{e_n : n \in \mathbb{J} \cup I_m\}$ and $(\mathbb{J} \cup I_m) \cap I_j = \emptyset$ for every $j < m$.
We have proved that $Dg(y)(v) = (t_1 \dots, t_d)$ and consequently the critical set of points of g is locally contained in a finite union of complemented subspaces of infinite codimension in E. □

5.6 Proof of the main result about extraction of critical points of C^1 functions

Theorems 5.6 and 5.8 above give us an approximation of a C^1 function $f : E \to \mathbb{R}^d$ and of its derivative by another function $g : E \to \mathbb{R}^d$ which has a *nice* critical set of points C_g. In the case of Theorem 5.6 the term *nice* means we are in the best situation where $C_g = \emptyset$. And in the case of Theorem 5.8 the term *nice* will mean for us that the closed set $C_g \subseteq U$ has the ε-strong C^1 extraction property with respect to E, that is, there exists a C^1 diffeomorphism $h : E \to E \setminus C_g$ such that h is the identity outside U and h refines a given open cover $\mathcal{G}$ of E. With these functions at our disposal, and with the help of Corollary 3.32 we can prove our main Theorems 5.2 and 5.3.

Proofs of Theorems 5.2 and 5.3. Firstly we choose another C^1 function $\delta : E \to [0, \infty)$ such that $\delta^{-1}(0) = E \setminus U$ and $\delta(x) \leq \varepsilon(x)$ for every $x \in E$. This is doable because in every separable Banach space with a C^1 equivalent norm, every closed set is the zero set of a C^1 function [1].
By Theorems 5.6 or 5.8 there exists a C^1 function $g : U \to \mathbb{R}^d$ such that

(1) $|f(x) - g(x)| \leq \frac{\delta(x)}{2}$ for every $x \in U$;

(2) $||Df(x) - Dg(x)|| \leq \frac{\delta(x)}{2}$ for every $x \in U$;

(3) $C_g = \emptyset$ in the case of Theorem 5.2, or C_g is locally contained in subspaces of infinite codimension in E in the case of Theorem 5.3.

Let us extend now this function $g : U \to \mathbb{R}^d$ to the whole space E by letting it be equal to f outside U. We keep calling this extension by g and it is important to note that this function is still of class C^1. The

[1] Wells proved in his thesis [148] that if a separable Banach space E admits a C^1 smooth Lipschitz bump function, that is a C^1 non-null function $\lambda : E \to [0, \infty)$ with bounded derivative and bounded support, then every closed set X of E is the zero set of some C^1 function. Since a Banach space admitting an equivalent C^1 norm has a C^1 smooth Lipschitz bump function our statement follows.

only points where this fact could not be clear are those from the boundary of U. However the Fréchet derivative of g at those points $x \in \partial U$ exists and is $Df(x)$ because

$$\begin{aligned}\limsup_{h\to 0} \frac{|g(x+h)-g(x)-Df(x)(h)|}{||h||} &\leq \limsup_{h\to 0} \frac{|g(x+h)-f(x+h)+f(x)-g(x)|}{||h||} \\ &\quad + \limsup_{h\to 0} \frac{|f(x+h)-f(x)-Df(x)(h)|}{||h||} \\ &= \limsup_{h\to 0} \frac{|g(x+h)-f(x+h)|}{||h||} + 0 \\ &\leq \lim_{h\to 0} \frac{\delta(x+h)-\delta(x)}{||h||} = 0.\end{aligned}$$

Here we are using the facts that f is Fréchet differentiable in ∂U and that $f(x) = g(x)$ and $\delta(x) = \delta'(x) = 0$ for every $x \in \partial U$.
We have just shown that g is Fréchet differentiable on E, but it remains to show that it is C^1. Straightforwardly for every $x \in \partial U$,

$$\lim_{y\to x, y\notin U} ||Dg(y)-Df(x)|| = \lim_{y\to x, y\notin U} ||Df(y)-Df(x)|| = 0$$

and

$$\limsup_{y\to x, y\in U} ||Dg(y)-Df(x)|| \leq \lim_{y\to x, y\in U} (||Dg(y)-Df(y)|| + ||Df(y)-Df(x)|| \leq \lim_{y\to x, y\in U} \delta(y) = 0,$$

by the continuity of Df, property (2) of Theorems 5.6 and 5.8 and because $\delta^{-1}(0) = E \setminus U$.

1. **Case of Theorem 5.2:** Define $\varphi = g$ and we obtain that

$$|\varphi(x) - f(x)| \leq \delta(x) \leq \varepsilon(x)$$

for all $x \in E$ and $\varphi(x) = f(x)$ for every $x \in E \setminus U$. Besides, it is clear that φ does not have any critical point.

2. **Case of Theorem 5.3:** We will extract the critical set C_g in the following way. Observe that C_g is a closed set included in U (note that $C_g \cap \partial U = \emptyset$ because $Dg(x) = Df(x)$ is surjective for every $x \in \partial U$), and by (3) of Theorem 5.8 is locally contained in a finite union of complemented subspaces of infinite codimension. Using Corollary 3.32, there exists a C^1-diffeomorphism $h : E \to E \setminus C_g$ which is the identity outside U and is limited by the open cover $\mathcal{G}$ that we next define. Recall that we have

$$|f(x) - g(x)| \leq \delta(x)/2$$

for all $x \in E$. Since g and δ are continuous, for every $z \in E$ there exists $\eta_z > 0$ so that if $x, y \in B(z, \eta_z)$ then $|g(y) - g(x)| \leq \delta(z)/4 \leq \delta(x)/2$. We set $\mathcal{G} = \{B(x, \eta_x) : x \in E\}$.
Finally, let us define

$$\varphi = g \circ h.$$

Since h is limited by $\mathcal{G}$ we have that, for any given $x \in E$, there exists $z \in E$ such that $x, h(x) \in B(z, \eta_z)$, and therefore $|g(h(x)) - g(x)| \leq \delta(z)/4$, that is, we have that

$$|g(x) - \varphi(x)| \leq \delta(z)/4 \leq \delta(x)/2.$$

We obtain that

$$|f(x) - \varphi(x)| \leq \delta(x) \leq \varepsilon(x)$$

for all $x \in E$. Furthermore h is the identity outside U so $\varphi(x) = g(x) = f(x)$ for every $x \in E \setminus U$. Besides, it is clear that φ does not have any critical point: since $h(x) \notin C_g$, we have that the linear map $Dg(h(x))$ is surjective for every $x \in E$, and $Dh(x) : E \to E$ is a linear isomorphism, so $D\varphi(x) = Dg(h(x)) \circ Dh(x)$ is surjective for every $x \in E$.

□

Remark 5.11. We could have gotten that the approximating function φ is of class C^k (where k is the order of smoothness of the space E) inside the open set U. To achieve this one should get a version of Lemma 5.4 exactly as in [20, Lemma 5]. Doing this we would get from that lemma that the functions $\delta_j(x)$ are of class C^k. Hence the approximating function g from Theorems 5.6 and 5.8,

$$g(x) = \sum_{j=1}^{\infty} h_j(x)(f(x^j) + F_j(x - x^j) - \delta_j(x) + T_j(x - x^j))$$

is a function of class C^k on U.
Moreover, we can find an extracting diffeomorphism $h : E \to E \setminus C_g$ of class C^k by Corollary 3.32, hence $\varphi = g \circ h$ will be a C^k mapping on U.

Remark 5.12.

1. The space c_0 satisfies the conditions of Theorem 5.2. The supremum norm in c_0 locally depends on finitely many coordinates, so applying [90, Theorem 1] one gets the existence of an equivalent C^∞ smooth norm on c_0 that locally depends on finitely many coordinates. Namely, it was Kuiper in [40] the first one who found a C^∞ smooth and LFC renorming of c_0. The space $C(K)$, with K a metrizable countable compactum also satisfies the conditions of Theorem 5.2.

2. The space ℓ_p satisfies the conditions of Theorem 5.3. For every $1 < p < \infty$ the canonical norm of ℓ_p is
$$||x|| = ||\sum_{n=1}^{\infty} x_n e_n|| = \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p}.$$
With this expression it is easy to check that the basis is in fact 1-suppression unconditional with unconditional constant $K_u = 1$. It is also a strictly convex norm of class C^k, where k is defined as follows: $k = \infty$ if $p = 2n$, $n \in \mathbb{N}$; $k = 2n + 1$ if $p = 2n + 1$, $n \in \mathbb{N}$, and k is equal to the integer part of p if $p \notin \mathbb{N}$.

3. **Open question:** Does $L^p[0, 1]$, $1 < p < \infty$ admit an equivalent norm of class C^1 and a 1-suppression unconditional basis $\{e_n\}$? The canonical norm on $L^p[0, 1]$ is of class C^1 and for this norm the Haar basis is unconditional, but is it 1-suppression unconditional? A necessary condition is that the unconditional constant K_u satisfies $1 \leq K_u \leq 2$. In view of [46, Corollary 1] which asserts that $K_u = \max\{p, p'\}$, with $1/p + 1/p' = 1$, our second question only has sense for the cases $\frac{3}{2} \leq p \leq 3$.

   What is true and known from 2011 (see [61, Corollary 2.3]) is that we can find a renorming of $L^p[0, 1]$, $1 < p < \infty$, such that the Haar basis becomes 1-unconditional, hence $K_u = 1$ and also 1-suppression unconditional. However this equivalent norm is not C^1.

5.7 Application to counterexamples of Rolle's theorem with prescribed support

The following corollary is related with the failure of Rolle's theorem in infinite-dimensional Banach spaces. Rolle's theorem in finite-dimensional spaces ensures that for every bounded connected open subset $U \subset \mathbb{R}^d$, and every continuous function $f : \overline{U} \to \mathbb{R}$, differentiable in U and constant on ∂U, there exists a point $x_0 \in U$ such that $f'(x_0) = 0$. Unfortunately Rolle's theorem is no longer true in infinite dimensions. The first one who showed this was S. A. Shkarin in 1992 [141], for superreflexive spaces and non-reflexive spaces with smooth norms. Nowadays the class of Banach spaces for which the Rolle's theorem fails has been greatly enlarged, and includes in particular all infinite-dimensional Banach spaces with a (not necessarily equivalent) Fréchet differentiable norm (see [11, p. 23]) or with a C^1 smooth and Lipschitz bump function (see [24, Theorem 1.1]). However in theses examples the functions are just smooth inside the open set U. We also want these counterexamples to be smooth in the whole Banach space we are working on.

The next result focus on the question of finding C^1 bump functions with a prefixed support violating the Rolle's theorem. Obviously if we want a precise bounded support for our bump function a necessary condition, by the results of Wells [148], is that the Banach space has C^1 smooth and Lipschitz bumps.

Corollary 5.13. *Let E be a Banach space satisfying the conditions of Theorem 5.2. Then for every open set U there exists a C^1 function $\lambda : E \to [0,\infty)$ whose support is the closure of U and does not have any critical point in U.*

Proof. As we have already mentioned, by Wells' thesis [148], since E is separable and has an equivalent C^1 smooth norm, we can find a C^1 function $\mu : E \to [0,\infty)$ such that $\mu^{-1}(0) = E \setminus U$. Now consider $\varepsilon : U \to (0,\infty)$ defined by $\varepsilon(x) = \mu(x)/2$, and apply our result Theorem 5.6 to get a C^1 function $\lambda : U \to [0,\infty)$ such that $|\mu(x) - \lambda(x)| \leq \varepsilon(x)$, $||\mu'(x) - \lambda'(x)|| \leq \varepsilon(x)$ and $\lambda'(x) \neq 0$ for every $x \in U$. Now let λ be equal to zero outside U. It is clear that λ is of class C^1, that $\lambda^{-1}(0) = E \setminus U$ and has no critical points inside U. □

Compare also with [24, Theorem 1.1], [12, Theorem 1.5] or [25, Corollary 8]. Similarly we can obtain the following.

Corollary 5.14. *Let E be a Banach space satisfying the conditions of Theorem 5.2. Then for every disjoint closed sets $A, B \subset E$ there exists a C^1 function $\lambda : E \to [0,1]$ such that $\lambda^{-1}(0) = A$, $\lambda^{-1}(1) = B$ and with no critical points inside $\lambda^{-1}((0,1))$.*

Proof. Again we can find two C^1 functions $\mu_A : E \to [0,\infty)$ such that $\mu^{-1}(0) = A$ and $\mu_B : E \to [0,\infty)$ such that $\mu^{-1}(0) = B$. Define $\mu(x) := \mu_A(x)/(\mu_A(x) + \mu_B(x))$, and it is clear that $\mu^{-1}(0) = A$ and $\mu^{-1}(1) = B$. Let us denote $U = E \setminus (A \cup B)$. Now consider $\varepsilon : U \to (0,\infty)$ defined by $\varepsilon(x) = \mu(x)(1-\mu(x))/2$, and apply our Theorem 5.6 to get a C^1 function $\lambda : U \to [0,\infty)$ such that $|\mu(x) - \lambda(x)| \leq \varepsilon(x)$, $||\mu'(x) - \lambda'(x)|| \leq \varepsilon(x)$ and $\lambda'(x) \neq 0$ for every $x \in U$. Now let λ be equal to zero on A and equal to one on B. It is clear that λ is of class C^1, that $\lambda^{-1}(0) = A$, $\lambda^{-1}(1) = B$ and it has no critical points inside $U = \lambda^{-1}((0,1))$. □

Conclusiones

El Teorema de Morse-Sard (1942) es un resultado importantísimo dentro del análisis matemático y con numerosas aplicaciones en otros campos. En su forma clásica establece que si una función $f : \mathbb{R}^n \to \mathbb{R}^m$ tiene suficiente regularidad C^k entonces su conjunto de valores críticos $f(C_f)$ tendrá medida de Lebesgue pequeña. En la literatura uno puede encontrar multitud de refinamientos del Teorema de Morse-Sard para diversas clases de funciones. Nosotros obtenemos la misma conclusión bajo condiciones más débiles en las que sólo pedimos que la función sea aproximadamente diferenciable de ciertos órdenes k en diversos conjuntos, que cada vez van siendo más pequeños en medida cuanto mayor sea k.

Otra pregunta natural que se han planteado los matemáticos en el último siglo es determinar qué funciones $f : \mathbb{R}^n \to \mathbb{R}$ coinciden con funciones de clase C^k salvo en conjuntos de medida de Lebesgue $\mathcal{L}^n$ tan pequeños como queramos. Se dirá que tales funciones satisfacen la propiedad de Lusin de clase C^k. Una caracterización de esta propiedad viene dada por los resultados indepedientes de Isakov y de Liu y Tai (1994). En esta tesis se aporta un nuevo resultado en esta dirección en el que trabajamos con funciones subdiferenciables. En concreto demostramos que si una función tiene una subdiferencial Fréchet (proximal) no vacía en $\mathcal{L}^n$-casi todo punto entonces ésta satisface la propiedad de Lusin de clase C^1 (C^2 respectivamente). Además, como a menudo sucede cuando se trabaja con subdiferenciales, los resultados análogos para órdenes mayores $k \geq 3$ se demuestran falsos.

El Teorema de Morse-Sard no es cierto en dimensión infinita. Esto se mostró por primera vez en el 1965 gracias al ejemplo dado por Kupka de una función $f : \ell_2 \to \mathbb{R}$ de clase C^∞ tal que $\mathcal{L}(f(C_f)) > 0$. Sin embargo en posteriores trabajos se ha estudiado lo que se conoce como *teoremas aproximados de Morse-Sard*. En ellos el objetivo es ser capaz de aproximar uniformemente cualquier función continua $f : E \to F$ entre espacios de Banach por otra diferenciable con un conjunto *pequeño* de valores críticos. Los resultados más fuertes en este sentido consiguen aproximaciones diferenciables sin puntos críticos para funciones $f : E \to \mathbb{R}^m$. Uno puede preguntarse si se puede obtener el mismo tipo de aproximación para funciones con valores en otro espacio de Banach infinito-dimensional F. En este trabajo damos una respuesta afirmativa a esta pregunta. Mostramos que para $E = c_0, \ell_p$ ó L_p, $1 < p < \infty$ y F un cociente de E, entonces cualquier función continua $f : E \to F$ puede aproximarse uniformemente por funciones de clase C^k sin ningún punto crítico, y donde k denota el orden de regularidad de la norma del espacio E en cuestión. Además un resultado algo diferente pero con la misma esencia sería el siguiente: para el caso de $E = c_0, \ell_p$, $1 < p < \infty$ y funciones $f : E \to \mathbb{R}^m$ de clase C^1, uno puede aislar el conjunto de puntos críticos $C_f \subset U$ por un abierto y de alguna forma *extraerlos* sin perturbar demasiado a la función f; esto es, encontrar $g : E \to \mathbb{R}^m$ de clase C^1 sin puntos críticos que aproxima a f y que coincide con ella fuera de U.

Resumiendo, se ha conseguido generalizar y extender enormemente la clase de espacios de Banach (E, F) para los que los teoremas de Morse-Sard aproximados en sentido fuerte (aproximaciones sin puntos críticos) son válidos. Con este trabajo se ha contribuido a mejorar el entendimiento de la geometría y estructura de los espacios de Banach infinito-dimensionales.

Finalizamos explicando una técnica clave usada en las demostraciones anteriores. En cierto momento nosotros necesitamos extraer difeomórficamente ciertos conjuntos cerrados X del espacio E. En general en la literatura se encuentran difeomorfismos $h : E \to E \setminus X$ para el caso de espacios de Banach con normas diferenciables E y donde X es un conjunto localmente compacto. Tuvimos que refinar toda esta teoría para ser capaces de extraer difeomórficamente el tipo de conjuntos cerrados que queríamos en los espacios de Banach adecuados. En concreto para nosotros X son conjuntos cerrados que localmente se ven como gráficas de funciones continuas definidas desde subespacios infinito-codimensionales y tomando valores en sus complementarios lineales en E. Y además, muy importante, podemos hacer que el difeomorfismo extractor h esté tan cerca de la identidad como deseemos. Cabe mencionar que las demostraciones que presentamos se vuelven altamente técnicas en ciertos momentos, pero no en balde conseguimos generalizaciones muy finas de toda la teoría de extracción difeomorfa existente.

Conclusions

The Morse-Sard Theorem is a very important result in mathematical analysis with numerous applications in other fields. In its classical form it asserts that if a function $f : \mathbb{R}^n \to \mathbb{R}^m$ has enough regularity C^k then its set of critical values $f(C_f)$ will have small Lebesgue measure. In the literature one can find plenty of refinements of the Morse-Sard Theorem for several classes of functions. We obtain the same conclusion under weaker conditions by requiring only that the function is approximately differentiable of order k in certain subsets, which are smaller in measure as k grows.

Another natural question that mathematicians have asked themselves in the last century is to determine which functions $f : \mathbb{R}^n \to \mathbb{R}$ coincide with functions of class C^k except on sets of arbitrarily small Lebesque measure. We say that such functions satisfy the Lusin property of class C^k. A characterization of this property comes after the independent works of Isakov, and of Liu and Tai (1995). In this dissertation we contribute with a new result in this direction, working with subdifferentiable functions. Namely we prove that if a function has nonempty Fréchet (proximal) subdifferential $\mathcal{L}^n$-almost everywhere then it has the Lusin property of class C^1 (C^2 respectively). Furthermore, as it often happens when dealing with subdifferentials, the analogous results for higher orders $k \geq 3$ are shown to be false.

The Morse-Sard Theorem is not true in infinite dimensions. This fact was shown firstly in 1965 with an example due to Kukpa of a function $f : \ell_2 \to \mathbb{R}$ of class C^∞ so that $\mathcal{L}(C_f) > 0$. However in subsequent works it has been studied what is known as *approximated Morse-Sard results*. In those, the objective is to uniformly approximate any continuous function $f : E \to F$ between Banach spaces by smooth functions with a *small* set of critical values. The strongest results in this context get smooth approximations without any critical point for functions $f : E \to \mathbb{R}^m$. One can wonder whether it is possible to obtain the same kind of approximation for functions with values in another infinite-dimensional Banach space F. In the present work we give an affirmative answer. We show that for $E = c_0, \ell_p$ or L_p, $1 < p < \infty$ and F a quotient of E, then any continuous function $f : E \to F$ can be uniformly approximated by C^k smooth functions with no critical points, and where k denotes the order of smoothness of the norm of the space E. Moreover a slightly different result with a different flavour reads as follows: for the case $E = c_0, \ell_p$, $1 < p < \infty$ and functions $f : E \to \mathbb{R}^m$ of class C^1, one can isolate the critical set $C_f \subset U$ by an open set and somehow *extract it* without perturbing too much the function f; that is, find $g : E \to \mathbb{R}^m$ of class C^1 without critical points that approximates f and that coincides with it outside U.

Summing up, we have achieved a great generalization of the class of Banach spaces (E, F) for which the approximated Morse-Sard theorems in a strong sense (approximations without critical point) are valid. With this work we have contributed to improving the knowledge of the geometry and structure of infinite-dimensional Banach spaces.

We end by explaining a key technique used in the proofs of the previous results. In some moments we need to diffeomorphically extract certain closed subsets X of the space E. In general in the literature we find diffeomorphisms $h : E \to E \setminus X$ for the case of Banach spaces with smooth norms and where X is a locally compact set. We had to refine all this theory to be able to diffeomorphically extract the kind of closed subsets that we wanted to in the adequate Banach spaces. Namely for us X is a closed set which can be viewed locally as a graph of a continuous function defined from an infinite-codimensional subspace and taking values in its orthogonal complement in E. Moreover, and very importantly, we can make the extracting diffeomorphism to be as close to the identity as we want. It is worth mentioning that the proofs that we present become highly technical at some points, but nevertheless we get very fine generalizations of all the existing theory of diffeomorphic extraction.

Bibliography

[1] G. Alberti, *A Lusin type theorem for gradients*, J. Funct. Anal. 100 (1991), no. 1, 110–118.

[2] G. Alberti, *On the structure of singular sets of convex functions*, Calc. Var. Partial Differential Equations 2 (1994), no. 1, 17–27.

[3] F. Albiac and J. L. Ansorena, *Characterization of* 1*-almost greedy bases*, Rev. Mat. Complut. 30 (2017) no. 1, 13–24.

[4] F. Albiac and N. Kalton, *Topics in Banach space theory*, Graduate Texts in Mathematics, 233, Springer, 2016.

[5] A. D. Alexandroff, *Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it.* (Russian) Leningrad State Univ. Annals [Uchenye Zapiski] Math. Ser. 6, (1939). 3–35.

[6] C. D. Aliprantis and K. C. Border, *Ininite dimensional analysis: A Hitchhiker's guide*, Third Edition, Springer-Verlag Berlin Heidelberg, 2006.

[7] D. E. Alspach and Y. Benyamini, *Primariness of spaces of continuous functions on ordinals*, Israel J. Math. 27 (1977), 64–92.

[8] R. D. Anderson, D. W. Henderson and J. E. West, *Negligible subsets of infinite-dimensional manifolds*, Compositio Math. Vol. 21, Fasc. 2, 1969, pp. 143–150.

[9] G. Anzellotti and R. Serapioni, C^k-*rectifiable sets*, J. Reine Angew. Math. 453: 1–20, 1994.

[10] D. Azagra, *Diffeomorphisms between spheres and hyperplanes in infinite-dimensional Banach spaces*, Studia Math. 125 (1997) no. 2, 179–186.

[11] D. Azagra, *Smooth negligibility and subdifferential calculus in Banach spaces, with applications*, doctoral dissertation, Universidad Complutense de Madrid, 1997.

[12] D. Azagra and M. Cepedello, *Uniform approximation of continuous mappings by smooth mappings with no critical points on Hilbert manifolds*, Duke Math. J. 124 (2004) no. 1, 47–66.

[13] D. Azagra and R. Deville, *James' theorem fails for starlike bodies*, J. Funct. Anal. 180 (2001), 328–346.

[14] D. Azagra and T. Dobrowolski, *Smooth negligibility of compact sets in infinite-dimensional Banach spaces, with applications*, Math. Annalen 312 (1998), no. 3, 445–463.

[15] D. Azagra and T. Dobrowolski, *Real-analytic negligibility of points and subspaces in infinite-dimensional Banach spaces*, with applications, Canadian Math. Bull. 45 (2002) no.1, 3–11.

[16] D. Azagra, T. Dobrowolski and M. García Bravo *Smooth approximations without critical points of continuous mappings between Banach spaces, and diffeomorphic extractions of sets*, Adv. Math. 354 (October 2019).

[17] D. Azagra, J. Ferrera and J. Gómez-Gil, *Nonsmooth Morse-Sard theorems*, Nonlinear Analysis, 160 (2017), 53–69.

[18] D. Azagra, J. Ferrera and J. Gómez-Gil, *The Morse-Sard theorem revisited*, The Quarterly Journal of Mathematics, Volume 69, Issue 3 (2018), 887–913.

[19] D. Azagra, J. Ferrera, M. García-Bravo and J. Gómez-Gil, *Subdifferentiable functions satisfy Lusin properties of class C^1 or C^2*, J. Approx. Theory 230 (2018), 1–12.

[20] D. Azagra, R. Fry, J. Gómez-Gil, J.A. Jaramillo and M. Lovo, *C^1-fine approximation of functions on Banach spaces with unconditional basis*, Quarterly J. Math. 56 (2005) no. 1, 13–20.

[21] D. Azagra and M. García-Bravo, *Some remarks about the Morse-Sard theorem and approximate differentiability*, Revista Matemática Complutense 33 (2020) 161–185.

[22] D. Azagra, J. Gómez-Gil and J. A. Jaramillo, *Rolle's theorem and negligibility of points in infinite-dimensional Banach spaces*, J. Math. Anal. Appl. 213 (1997), no. 2, 487–495.

[23] D. Azagra and P. Hajłasz *Lusin-type properties of convex functions*, preprint, 2019.

[24] D. Azagra, M. Jiménez-Sevilla, *The failure of Rolle's theorem in infinite-dimensional Banach spaces*, J. Funct. Anal. 182 (2001) 207–226.

[25] D. Azagra and M. Jiménez-Sevilla, *Approximation by smooth functions with no critical points on separable infinite-dimensional Banach spaces*, J. Funct. Anal. 242 (2007), 1–36.

[26] D. Azagra and A. Montesinos, *Starlike bodies and deleting diffeomorphisms in Banach spaces*, Extracta Math. 19 (2004), no. 2, 171–213.

[27] T. Bagby and W.P. Ziemer, *Pointwise differentiability and absolute continuity*, Trans. Amer. Math. Soc. 191 (1974), 129–148.

[28] L. Barbet, M. Dambrine, A. Daniilidis and L. Rifford *The Morse-Sard theorem for Clarke critical values*, Adv. Math. 242 (2013), 217–227.

[29] L. Barbet, M. Dambrine, A. Daniilidis and L. Rifford *Sard theorems for Lipschitz functions and applications in optimization*, Israel J. Math. 212 (2016), 757–790.

[30] S. M. Bates, *On the image size of singular maps II*, Duke Mathematical Journal 68 1992, 463–476.

[31] S. M. Bates, *Toward a precise smoothness hypothesis in Sard's theorem*, Proc. Amer. Math. Soc. 117 (1993), no. 1, 279–283.

[32] S. M. Bates and C. G. Moreira, *De nouvelles perspectives sur le théorème de Morse-Sard*, C.R. Acad. Sci. Paris, t. 332, Série I (2001), 13–17.

[33] Y. Benyamini and J. Lindenstrauss, *A predual of l_1 which is not isomorphic to a $C(K)$ space*, Israel J. Math. 13 (1972), 246–254.

[34] C. Bessaga, *Every infinite-dimensional Hilbert space is diffeomorphic with its unit sphere*, Bull. Acad. Polon. Sci., Sér. Sci. Math. Astr. et Phys.14 (1966), 27–31.

[35] C. Bessaga, *Interplay Between Infinite-Dimensional Topology and Functional Analysis. Mappings Defined by Explicit Formulas and Their Applications*, Topology Proceedings, 19 (1994).

[36] C. Bessaga and V. L. Klee, *Two topological properties of topological linear spaces*, Israel J. Math. 2 (1964), pp. 211–220.

[37] C. Bessaga and V. L. Klee, *Every non-normable Fréchet space is homeomorphic with all of its closed convex bodies*, Math. Ann. 163 (1966), pp. 161–166.

[38] B. Bojarski and P. Hajłasz, *Pointwise inequalities for Sobolev functions*, Studia Math. 106 (1993), 77–92.

[39] B. Bojarski, P. Hajłasz and P. Strzelecki, *Improved* $C^{k,\lambda}$ *approximation of higher order Sobolev functions in norm and capacity*. Indiana Univ. Math. J. 51 (2002), 507–540.

[40] R. Bonic and J. Frampton, *Smooth functions on Banach manifolds*, J. Math. Mech. 15 (1966), 877–898.

[41] J. M. Borwein and J. D. Vannderwerff, *Convex Functions. Constructions, Characterizations and Counterexamples.* Encyclopedia of Mathematics and Its Applications 109. Cambridge University Press, 2010.

[42] J. Bourgain and F. Delbaen, *A class of special* $\mathcal{L}_\infty$ *spaces*, Acta Math. 145 (1980) 155–176.

[43] J. Bourgain, M. V. Korobkov and J. Kristensen, *On the Morse-Sard property and level sets of Sobolev and BV functions*, Rev. Mat. Iberoam. 29 (2013), no. 1, 1–23.

[44] J. Bourgain, M. V. Korobkov and J. Kristensen, *On the Morse-Sard property and level sets of* $W^{n,1}$ *Sobolev functions on* $\mathbb{R}^n$, J. Reine Angew. Math. 700 (2015), 93–112.

[45] J. B. Brown and G. Kozlowski, *Smooth interpolation, Hölder continuity, and the Takagi-van der Waerden function.* Amer. Math. Monthly 110 (2003), no. 2, 142–147.

[46] D. L. Burkholder, *A nonlinear partial differential equation and the unconditional constant of the Haar system in* L^p, Bull. Amer. Math. Soc. 7 (1982), 591–595.

[47] D. Burghelea and N. H. Kuiper, *Hilbert manifolds*, Annals of Math. 90 (1969), 379–417.

[48] H. Busemann and W. Feller, *Krümmungseigenschaften Konvexer Flächen.* (German) Acta Math. 66 (1936), no. 1, 1–47.

[49] A. P. Calderón and A. Zygmund, *Local properties of solutions of elliptic partial differential equations*, Studia Math. 20 (1961), 171–225.

[50] S. Campanato, *Propietà di una famiglia di spazi funzionali*, Ann. Scuola Norm. Sup. Pisa 18 (1964), 137–160.

[51] P. G. Cassaza, *James quasi-reflexive space is primary*, Israel J. Math. 26 (1977), 294?305.

[52] P. G. Cassaza, B. L. Lin and R. H. Lohman, *On James' quasi-reflexive Banach space*, Proc. Amer. Math. Soc. 67 (1977), 265–271.

[53] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern and P.R. Wolenski, *Nonsmooth Analysis and Control Theory.* Grad. Texts in Math. 178, Springer, 1998.

[54] D. L. Cohn, *Measure Theory*, Birkhäuser Advanced Texts Basler Lehrbücher, Second Edition, 2013

[55] H. H. Corson, *Thin sets in Fréchet spaces*, Sympos. on Infinite Dimensional Topology (Baton Rouge, 1967), Ann. of Math. Studies, Princeton Univ. Press, Princeton, N. J., 1971.

[56] M. G. Crandall, H. Ishii, P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. 27 (1992) 1–67.

[57] S. D'Alessandro and P. Hájek, *Polynomial algebras and smooth functions in Banach spaces*, J. Funct. Anal. 266 (2014), 1627–1646.

[58] L. De Pascale, *The Morse-Sard theorem in Sobolev spaces*, Indiana Univ. Math. J. 50 (2001), 1371–1386.

[59] R. Deville, V. Fonf and P. Hájek, *Analytic and polyhedral approximation of convex bodies in separable polyhedral Banach spaces*, Israel J. Math. 105 (1998), 139–154.

[60] R. Deville, G. Godefroy and V. Zizler, *Smoothness and renormings in Banach spaces*, vol. 64, Pitman Monographs and Surveys in Pure and Applied Mathematics, Longman Scientific & Technical, Harlow, 1993.

[61] S. J. Dilworth, E. Odell, Th. Schlumprecht and A. Zsák, *Renormings and symmetry properties of* 1*-greedy bases*, J. Approx. Theory 163 (2011) 1049–1075.

[62] T. Dobrowolski, *Relative Classification of Smooth Convex Bodies*, Bull. Acad. Polon. Sci., Sér. Sci. Math. 25 (1977), 309–312.

[63] T. Dobrowolski, *Smooth and R-analytic negligibility of subsets and extension of homeomorphism in Banach spaces*, Studia Math. 65 (1979), 115–139.

[64] T. Dobrowolski, *Every Infinite-Dimensional Hilbert Space is Real-Analytically Isomorphic with Its Unit Sphere*, Journal of Functional Analysis, 134 (1995), 350–362.

[65] J. R. Dorronsoro, *Differentiability properties of functions with bounded variation*, Indiana Univ. Math. J. 38 (1989), no. 4, 1027–1045.

[66] J. Dugundji, *An extension of Tietze's theorem*, Pacific J. Math. 1 (1951), no. 3, 353–367.

[67] J. Eells and K. D. Elworthy, *Open embeddings of certain Banach manifolds*, Ann. of Math. 91 (1970), 465–485.

[68] J. Eells and N. H. Kuiper, *Homotopy negligible subsets*, Comp. Math. 21 (1969) 155–161.

[69] J. Eells and J. McAlpin, *An approximate Morse-Sard theorem*, J. Math. Mech. 17 (1967/1968), 1055–1064.

[70] K. D. Elworthy, *Embeddings, isotopy and stability of Banach manifolds*, Compositio Math. 24 (1972), 175–226.

[71] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[72] M. Fabian, P. Habala, P, Hájek, V. Montesinos and V. Zizler, *Banach space theory. The basis for linear and nonlinear analysis*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011.

[73] H. Federer, *Surface area. II*, Trans. Amer. Math. Soc. 55, (1944), 438–456.

[74] H. Federer, *Geometric measure theory*, Springer-Verlag Berlin Heidelberg, 1996.

[75] A. Ferone, M. V. Korobkov, A. Roviello, *On Luzin $N-$property and uncertainty principle for the Sobolev mappings*, Anal. PDE 12 (2019), no. 5, 1149–1175.

[76] A. Ferone, M. V. Korobkov, A. Roviello, *Morse-Sard theorem and Luzin $N-$property: a new synthesis result for Sobolev spaces*, Sib. Math. J. 60 (2019), 916–926.

[77] J. Ferrer, *Rolle's theorem fails in ℓ_2*, Am. Math. Monthly, vol. 103, n. 2 (1996), 161–165.

[78] J. Ferrera, *An introduction to nonsmooth analysis*. Elsevier/Academic Press, Amsterdam, 2014.

[79] J. Ferrera and J. Gomez-Gil, *Non approximate derivability of the Takagi function*, Journal of Convex Analysis 27 (2020), No. 1, 307–312.

[80] J. Ferrera and J. Gomez-Gil, *Generalized Takagi-Van der Waerden functions and their subdifferentials*, J. Convex Anal. 25 (4) (2018) 1355–1369.

[81] T. Fiegel, *An example of infinite dimensional reflexive space non-isomorphic to its cartesian square*, Mémoires de la S. M. F., 31-32 (1972), 165–167.

[82] A. Figalli, *A simple proof of the Morse-Sard theorem in Sobolev spaces*, Proc.Amer. Math. Soc. 136 (2008), no. 10, 3675–3681.

[83] J. Frontisi, *Smooth partitions of unity in Banach spaces*, Rocky Mountain J. Math. 25 (1995) 1295–1304.

[84] R. Fry, *Approximation by functions with bounded derivative on Banach spaces*, Bull. Austral. Math. Soc. 69 (2004), no. 1, 125–131.

[85] B. M. Garay, *Cross sections of solution funnels in Banach spaces*, Studia Math. 97, 13–26, (1990).

[86] B. M. Garay, *Deleting Homeomorphims and the Failure of Peano?s Existence Theorem in Infinite-Dimensional Banach Spaces*, Funkcialaj Ekvacioj 34, 85–93, (1991).

[87] M. García-Bravo, *Extraction of critical points of smooth functions on Banach spaces*, J. Math. Anal. Appl. 482 (February 2020).

[88] G. Glaeser, *Etudes de quelques algèbres tayloriennes*, J. d'Analyse 6 (1958), 1–124.

[89] P. Gora and R. J. Stern, *Subdifferential analysis of the Van der Waerden function*, J. Convex Anal. 18 (3) (2011), 699–705.

[90] P. Hajek, *Smooth norms that depend locally on finitely many coordinates*, Proc. Amer. Math. Soc. 123 (1995), 3817–3821.

[91] P. Hajek and M. Johanis, *Smooth approximations without critical points*, Cent. Eur. J. Math. 1 (2003), no. 3, 284–291.

[92] P. Hájek and M. Johanis, *Smooth approximations*, J. Funct. Anal. 259 (2010), no. 3, 561–582.

[93] P. Hajek and M. Johanis, *Smooth analysis in Banach spaces*, De Gruyter Series in Nonlinear Analysis and Applications, Berlin/Boston 2014.

[94] P. Hajek, V. Montesinos, J. Vanderwerff and V. Zizler, *Biorthogonal systems in Banach spaces*, CMS Books in Mathematics, Canadian Mathematical Society, Springer, 2007.

[95] P. Hajlasz, M. V. Korobkov and J. Kristensen, *A bridge between Dubovitskiĭ-Federer theorems and the coarea formula*, Journal of Functional Analysis, 2017, vol 272, no 3, p. 1265–1295.

[96] P. Hajlasz and S. Zimmerman, *Dubovitskiĭ-Sard theorem for Sobolev mappings*, Indiana University Mathematics Journal, 66 (2017), 705–723.

[97] G. H. Hardy, *Weierstrass's non-differentiable function.* Trans. Amer. Math. Soc. 17 (1916), no. 3, 301–325.

[98] R. Haydon, *Normes indéfiniment différentiables sur certains espaces de Banach*, C. R. Acad. Sci. Paris 315 (1992), 1175–1178.

[99] R. Haydon, *Trees in renorming theory*, Proc. London Math. Soc. 78(3), 541–584, (1999).

[100] M. W. Hirsch, *Differential topology*. Graduate Texts in Mathematics, 33. Springer-Verlag, New York, 1976.

[101] S. A. Imomkulov, *Twice differentiability of subharmonic functions*. (Russian) Izv. Ross. Akad. Nauk Ser. Mat. 56 (1992), 877–888; translation in Russian Acad. Sci. Izv. Math. 41 (1993), 157–167.

[102] N. M. Isakov, *A global property of approximately differentiable functions*, Mathematical Notes of the Academy of Sciences of the USSR 41 (1987), 280–285.

[103] R. C. James, *Weakly compact sets*, Trans. Amer. Math. Soc. 113 (1964), 129–140.

[104] M. Jiménez-Sevilla and L. F. Sánchez-González Francisco, *On some problems on smooth approximation and smooth extension of Lipschitz functions on Banach-Finsler manifolds*, Nonlinear Analysis: Theory, Methods and Applications, volume 74 (11), 2011, pp. 3487–3500.

[105] W. B. Johnson, H. P. Rosenthal and M. Zippin, *On bases, finite-dimensional decompositions and weaker structures in Banach spaces*, Israel J. Math. 9 (1971), 488–506.

[106] S. Kaplan, *Homology properties of arbitrary subsets of Euclidean spaces* Trans. Amer. Math. Soc. 62 (1947), 248–271.

[107] S. Kakutani, *Topological properties of the unit sphere of a Hilbert space*, Proc. Imp. Acad. Tokyo, 19, 269–271, (1943).

[108] R. Kaufman, *A singular map of a cube onto a square*, J. Differential Geom. 14 (1979), no. 4, 593–594 (1981).

[109] A. B. Kharazishvili, *Strange functions in real analysis*. Second edition. Pure and Applied Mathematics (Boca Raton), 272. Chapman & Hall/CRC, Boca Raton, FL, 2006.

[110] V. L. Klee, *Convex bodies and periodic homeomorphisms in Hilbert space*, Trans. Amer. Math. Soc. 74 (1953), 10–43.

[111] M. Kocan and X.-J. Wang, *On the generalized Stepanov theorem*, Proc. Amer. Math. Soc. 125 (1997), 2347–2352.

[112] M. V. Korobkov, K. Pileckas and R. Russo, *Solution of Leray's problem for stationary Navier-Stokes equations in plane and axially symmetric spatial domains*, Annals of Mathematics 181 (2015), 769–807.

[113] M. V. Korobkov and J. Kristensen, *On the Morse-Sard theorem for the sharp case of Sobolev mappigs*, Indiana Univ. Math. J. 63 (2014), no. 6, 1703–1724.

[114] M.V. Korobkov and J. Kristensen, *The Trace theorem, the Luzin N-and More-Sard properties for the sharp case of Sobolev-Lorentz mappings*, The Journal of Geometric Analysis, Volume 28, Issue 3 (2018), 2834–2856.

[115] N. Kuiper, *The homotopy type of the unitary group of Hilbert spaces*, Topology 3 (1965) 19–30.

[116] I. Kupka, *Counterexample to the Morse-Sard theorem in the case of infinite-dimensional manifolds*, Proc. Amer. Math. Soc. 16 (1965), 954–957.

[117] S. Lang, *Introduction to differentiable manifolds*, Springer-Verlag New York, Second edition, 2002.

[118] S. Lang, *Real and Functional Analysis*, Third Edition. Graduate Texts in Mathematics 142. Springer, New York, 1993.

[119] J. Lindenstrauss and L. Tzafriri, *Classical Banach spaces. I. Sequence spaces*. Ergebnisse der Mathematik und ihrer Grenzgebiete, Vol. 92. Springer-Verlag, Berlin-New York, 1977.

[120] F-C. Liu, *A Lusin type property of Sobolev functions*, Indiana Univ. Math. J. 26 (1977), 645–651.

[121] F.-C. Liu and W.-S. Tai, *Approximate Taylor polynomials and differentiation of functions*, Topol. Methods Nonlinear Anal. 3 (1994), no. 1, 189–196.

[122] N. Lusin, *Sur les propiétés des fonctions mesurables*, Comptes Rendus Acad. Sci. Paris 154 (1912), 1688–1690.

[123] B. Malgrange, *Ideals of differentiable functions*, Tata Institute of Fundamental Research (1966).

[124] J. Michael and W. P. Ziemer, *A Lusin type approximation of Sobolev functions by smooth functions*, Contemp. Math. 42 (1985), 135–167.

[125] A. Morse, *The behavior of a function on its critical set*, Annals of Math. 40 (1939), 62–70.

[126] C. G. Moreira, *Hausdorff measures and the Morse-Sard theorem*, Publ. Mat. 45 (2001), 149–162.

[127] N. Moulis, *Sur les variétés hibertiennes et les fonctions non-dégénérées*, Indagationes Math. 30 (1968), 497–511.

[128] N. Moulis, *Approximation de fonctions différentiables sur certains espaces de Banach*, Ann. Inst. Fourier, Grenoble 21 (1971), 293–345.

[129] O. Nikodym, *Sur une classe de fonctions considérées dans le probleme de Dirichlet*, Fundam. Math. 21 (1933), 129–150.

[130] A. Norton, *A critical set with nonnull image has large Hausdorff dimension*, Trans. Amer. Math. Soc. 296 (1986), 367–276.

[131] A. Norton, *Functions not constant on fractal quasi-arcs of critical points*, Proc. Amer. Math. Soc. 106 (1989) no. 2, 397–405.

[132] A. Pelczyński and M. Wojciechowski, *Sobolev spaces*, Handbook of the Geometry of Banach spaces Vol. 2, W.B. Johnson and J. Lindenstrauss editors, Elsevier Amsterdam (2003), 1361–1423.

[133] R. R. Phelps, *Uniqueness of Hahn-Banach Extensions and Unique Best Approximation*, Transactions of the Amer. Math. Soc. 95 (1960) no. 2, 238–255.

[134] P. Renz, *Smooth extensions and extractions in infinite dimensional Banach spaces*. Thesis (Ph.D.)–University of Washington, 1969.

[135] P. Renz, *Smooth extensions in infinite dimensional Banach spaces*, Trans. Amer. Math. Soc. 168 (1972), 121–132.

[136] L. Rifford, *A Morse-Sard theorem for the distance function on Riemannian manifolds*, Manuscripta Math. 113 (2004), 251–265.

[137] L. Rifford, *On viscosity solutions of certain Hamilton-Jacobi equations: Regularity results and generalized Sard's Theorems*, Comm. Partial Differential Equations, 33:3 (2008), 517–559.

[138] H. P. Rosenthal, *The Banach spaces C(K)*, Handbook of the Geometry of Banach spaces Vol. 2, W.B. Johnson and J. Lindenstrauss editors, Elsevier Amsterdam (2003), 1547–1602.

[139] M. E. Rudin, *A new proof that metric spaces are paracompact*, Proc. Amer. Math. Soc. 20, no. 2 (1969), 603.

[140] A. Sard, *The measure of the critical values of differentiable maps*, Bull. Amer. Math. Soc. 48 (1942), 883–890.

[141] S. A. Shkarin, *On Rolle's theorem in infinite-dimensional Banach spaces*, Mat. Zametki 51 (1992), 128–136.

[142] S. Smale, *An infinite dimensional version of Sard's theorem*, Am. J. Math. 87 (1965), 861–866.

[143] A. Sobczyk, *Projection of the space (m) on its subspace c_0*, Bull. Amer. Math. Soc. 47 (1941), 938–947.

[144] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, Princeton, 1970.

[145] T. Takagi, *A simple example of the continuous function without derivative*, Proc. Phys. Math. Soc. Tokio Ser. II 1 (1903), 176–177.

[146] A. Tychonoff, *Ein Fixpunktsatz*, Math. Ann. 111 (1935), 767–776.

[147] J. Vanderwerff, *Smooth approximations in Banach spaces*, Proc. Amer. Math. Soc. 115 (1992) 113–120.

[148] J. C. Wells, *Smooth Banach spaces and approximations*. Thesis (Ph.D.)–California Institute of Technology, 1969.

[149] J. C. Wells, *Differentiable functions on c_0*, Bull. Amer. Math. Soc. 75 (1969), no. 1, 117–118.

[150] J. E. West, *The diffeomorphic excision of closed local compacta from infinite-dimensional Hilbert manifolds*, Compositio Math. 21 (1969), 271–291.

[151] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. 36 (1934), 63–89.

[152] H. Whitney, *A function not constant on a connected set of critical points*, Duke Math. J. 1 (1935), 514–517.

[153] H. Whitney, *On totally differentiable and smooth functions*, Pacific J. Math. 1 (1951), 143–159.

[154] Y. Yomdin, *The geometry of critical and near-critical values of differentiable mappings*, Math. Ann. 264 (1983), 495–515.

[155] W. P. Ziemer, *Weakly differentiable functions. Sobolev spaces and functions of bounded variation*, Graduate Texts in Mathematics, 120. Springer-Verlag, New York, 1989.

[156] M. Zippin, *On some subspaces of Banach spaces whose duals are L_1 spaces*, Proc. Amer. Math. Soc. 23 (1969), 378–385.