

Métodos III - Derivadas

February 12, 2019

6 Cambio de coordenadas y regla de la cadena

Jose A. Hernando

Departamento de Física de Partículas. Universidade de Santiago de Compostela
Enero 2019

```
In [1]: import time
        print(' Last version ', time.asctime() )
```

Last version Tue Feb 12 20:37:46 2019

6.1 Objetivos

Aprender a hacer cambios de coordenadas

Mostrar la regla de la cadena y algunos usos.

```
In [2]: # general imports
        # general imports
        %matplotlib inline
        %reload_ext autoreload
        %autoreload 2

        # numpy and matplotlib
        import numpy as np
        import matplotlib
        import matplotlib.pyplot as plt
        from mpl_toolkits.mplot3d import Axes3D
        matplotlib.style.use('ggplot')
        import graph_utils as gf
```

6.2 Cambio de variables

La siguiente función es el potencial de una carga eléctrica, q en el espacio:

$$V(x, y, z) = k \frac{q}{\sqrt{x^2 + y^2 + z^2}}$$

El potencial sabemos que es radial, que solo depende de la distancia a la carga, $r = \sqrt{x^2 + y^2 + z^2}$.

Podemos expresar el potencial en *coordenadas cartesianas*, como en la anterior ecuación, pero también, de forma más sencilla, en *coordenadas esféricas*, donde la variable r es la distancia al origen, así:

$$V(r) = k \frac{q}{r}$$

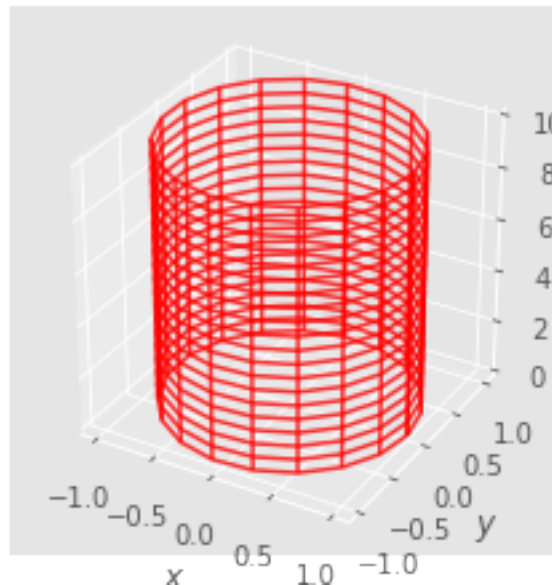
En este caso solo depende de r

En la sección de parametrización de superficies, vimos que podíamos dar la pared de un cilindro de radio r , y altura infinita como:

$$\sigma(\phi, z) = (r \cos \phi, r \sin \phi, z)$$

Donde el radio, r , era una constante. Si consideramos ahora que r puede tomar todos los valores de $[0, \infty)$. El conjunto de todos los cilindros *cube*, mapea, el espacio tridimensional.

```
In [3]: phi_range = 0., 2.*np.pi
        z_range   = 0., 10.
        r = 1.
        funx = lambda phi, z: r * np.cos(phi)
        funy = lambda phi, z: r * np.sin(phi)
        funz = lambda phi, z: z
        gf.wfsurface(phi_range, z_range, funx, funy, funz, nbins = 20);
```



Podemos dar todos los puntos, (x, y, z) , de un espacio R^3 en *coordenadas cilíndricas*, donde expresamos las coordenadas (x, y) por, (r, ϕ) donde r es la distancia al origen y ϕ el ángulo con el eje x . Mientras que mantenemos la coordenada z . Esto es:

$$(r, \phi, z) \rightarrow (x, y, z)$$

Donde:

$$x = r \cos \phi, \quad y = r \sin \phi, \quad z = z$$

El dominio en cilíndricas es $r \in [0, \infty]$, $\phi \in [0, 2\pi)$, mientras que z es todo \mathbb{R} .

Cuestión: ¿Cuál es la transformación inversa, de cartesianas a cilíndricas?

La transformación inversa, de *cilíndricas* a *cartesianas* viene dada por:

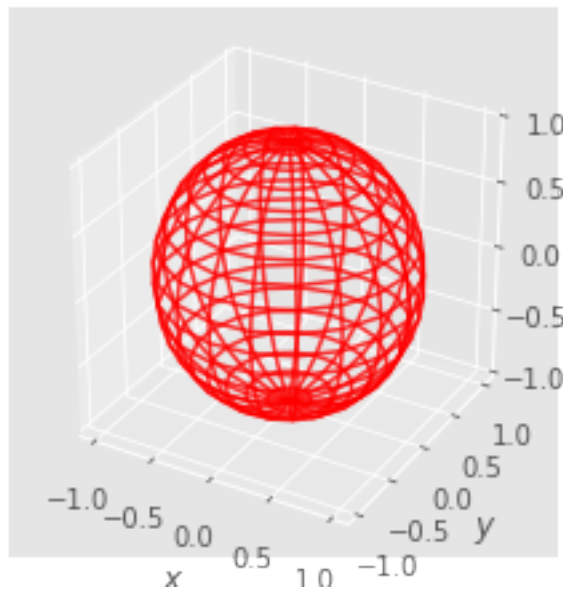
$$r = \sqrt{x^2 + y^2}, \quad \cos \phi = \frac{x}{\sqrt{x^2 + y^2}}, \quad z = z$$

Si recuerdas, también parametrizamos una esfera de radio r :

$$\sigma(\theta, \phi) = (r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta)$$

Considerando ahora todas las posibles esferas, esto es variando r de $[0, \infty]$ podemos cubrir de nuevo todo el espacio tri-dimensional, \mathbb{R}^3

```
In [4]: theta_range = 0., np.pi
        phi_range   = 0., 2.*np.pi
        r = 1.
        funx = lambda theta, phi : r * np.cos(phi) * np.sin(theta)
        funy = lambda theta, phi : r * np.sin(phi) * np.sin(theta)
        funz = lambda theta, phi : r * np.cos(theta)
        gf.wfsurface(theta_range, phi_range, funx, funy, funz);
```



Las *coordenadas esféricas* nos permiten asociar cualquier punto del espacio, (x, y, z) , de \mathbb{R}^3 , mediante el radio, r , o la distancia del punto al origen, el ángulo, θ , que forma el punto con el eje z , y el ángulo que forma (x, y) en el eje x .

El cambio en esféricas viene dado por:

$$x = r \cos \phi \sin \theta, \quad y = r \sin \phi \sin \theta, \quad z = r \cos \theta$$

Donde si recuerdas el range de θ es $[0, \pi]$ y el de ϕ de $[0, 2\pi)$.

Con $\theta = 0, \pi$ damos los polos de la esfera, con $\theta = \pi/2$ su ecuador. Situados a una altura de la esfera, el ángulo ϕ , nos permite recorrerla en una circunferencia de radio $r \sin \theta$ en sentido antihorario (visto desde arriba). En el ecuador, claro está, con radio r .

Cuestión: ¿Cuál es el cambio inverso que nos pasa de cartesianas a esféricas?

El cambio inverso de cartesianas a esféricas, viene dado por:

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \cos \phi = \frac{x}{\sqrt{x^2 + y^2}}, \quad \cos \theta = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$$

Cambios de coordenadas Los cambios de coordenadas son por lo tanto funciones vectoriales, $\mathbf{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, diferenciables, biyectivas (a excepción de algún punto singular) y que admiten una función inversa también diferenciable.

Nota: Una función biyectiva asociada a cada punto de espacio inicial un punto distinto en el espacio imagen, y todo punto del espacio imagen tiene un punto en el inicial. Fíjate que hay algunos puntos singulares que no cumplen esto, como los polos en el cambio a esféricas y cuando $r = 0$ para las dos. Sin embargo esto no nos impide cubrir completamente el espacio \mathbb{R}^3

Matriz jacobiana del cambio de coordenadas Vamos a calcular la matriz jacobiana de los cambios de coordenadas a cilíndricas.

la función vectorial del cambio de coordenadas a cilíndricas es:

$$\mathbf{x}(r, \phi, z) = (r \cos \phi, r \sin \phi, z)$$

Y su matriz jacobiana:

$$\mathbf{D}\mathbf{x}(r, \phi, z) = \begin{pmatrix} \cos \phi & -r \sin \phi & 0 \\ \sin \phi & r \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

la función vectorial del cambio de coordenadas a esféricas es:

$$\mathbf{x}(r, \phi, \theta) = (r \cos \phi \sin \theta, r \sin \phi \sin \theta, r \cos \theta)$$

Y su matriz jacobiana:

$$\mathbf{D}\mathbf{x}(r, \phi, \theta) = \begin{pmatrix} \cos \phi \sin \theta & -r \sin \phi \sin \theta & r \cos \phi \cos \theta \\ \sin \phi \sin \theta & r \cos \phi \sin \theta & r \sin \phi \cos \theta \\ \cos \theta & 0 & -r \sin \theta \end{pmatrix}$$

6.3 La regla de la cadena

En determinados casos tenemos una función vectorial, $\mathbf{f}(\mathbf{x})$, de \mathbb{R}^3 , de la que conocemos su derivada, $\mathbf{D}\mathbf{f}(\mathbf{x})$, y nos preguntamos cual será su derivada en coordenadas cilíndricas o esféricas.

¿Existe una relación sencilla entre la derivada de la función en las nuevas coordenadas y la derivada en cartesianas?

La respuesta es la regla de la cadena, que nos dice, que la nueva derivada es el producto de la derivada de la primera función y la derivada, la matriz jacobiana, del cambio de variables:

Sea $\mathbf{f}(\mathbf{x})$ con derivada, $\mathbf{Df}(\mathbf{x})$ en cartesianas y $\mathbf{x}(\mathbf{u})$ un cambio de coordenadas, con matriz jacobiana, $\mathbf{Dx}(\mathbf{u})$. La función compuesta, $\mathbf{h}(\mathbf{u}) = \mathbf{f} \circ \mathbf{x}(\mathbf{u}) = \mathbf{f}(\mathbf{x}(\mathbf{u}))$, tiene como derivada:

$$\mathbf{Dh}(\mathbf{u}) = \mathbf{Df}(\mathbf{x}(\mathbf{u})) \mathbf{Dx}(\mathbf{u})$$

¡El producto de las derivadas!

Teorema de la regla de la cadena *Teorema:* Sea $\mathbf{g} : A \subset \mathbf{R}^n \rightarrow \mathbf{R}^m$, A un conjunto abierto de \mathbf{R}^n , y $\mathbf{f} : B \subset \mathbf{R}^m \rightarrow \mathbf{R}^p$, B un conjunto abierto de \mathbf{R}^m , donde el dominio de \mathbf{f} esta incluido en el rango de \mathbf{g} . Si \mathbf{g} es diferenciable en un punto \mathbf{x} interior a A y \mathbf{f} lo es en el punto $\mathbf{g}(\mathbf{x})$, entonces, la función composición $\mathbf{h} = \mathbf{f} \circ \mathbf{g} : A \subset \mathbf{R}^n \rightarrow \mathbf{R}^p$, definida como: $\mathbf{h}(\mathbf{x}) = \mathbf{f} \circ \mathbf{g}(\mathbf{x}) = \mathbf{f}(\mathbf{g}(\mathbf{x}))$, es diferenciable en \mathbf{x} y su diferencial es:

$$\mathbf{D}(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) = \mathbf{Df}(\mathbf{g}(\mathbf{x})) \mathbf{Dg}(\mathbf{x})$$

Ejemplo: Con las funciones: $f(x, y) = x^2 + y^2$ y $\mathbf{g}(r, \theta) = (r \cos \theta, r \sin \theta)$, verifica la regla de la cadena para la función compuesta, $h(r, \theta) = (f \circ \mathbf{g})(r, \theta)$.

solución: Como producto de derivadas:

$$\nabla f(x, y) = (2x, 2y) \Rightarrow \nabla f(r, \theta) = (2r \cos \theta, 2r \sin \theta)$$

$$\mathbf{Dg}(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

$$\nabla f(r, \theta) \mathbf{Dg}(r, \theta) = (2r \cos \theta, 2r \sin \theta) \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix} = (2r, 0)$$

Ahora expresando la función inicial en polares:

$$h(r, \theta) = r^2$$

$$\nabla h(r, \theta) = (2r, 0)$$

Ejercicio: Verifica la regla de la cadena para la función compuesta $h = f \circ \mathbf{g} : \mathbf{R}^3 \rightarrow \mathbf{R}$ dada por $\mathbf{g} : \mathbf{R}^3 \rightarrow \mathbf{R}^3$, $\mathbf{g}(x, y, z) = (x^2y, y^2, e^{-xz})$ y $f : \mathbf{R}^3 \rightarrow \mathbf{R}$, $f(u, v, w) = u^2 + v^2 - w$.

Solución:

Sustituyendo:

$$h(x, y, z) = (x^2y)^2 + (y^2)^2 - e^{-xz}$$

$$\nabla h(x, y, z) = (4x^3y^2 + ze^{-xz}, 2x^4y + 4y^3, xe^{-xz})$$

Aplicando la regla de la cadena:

$$\mathbf{Dg}(x, y, z) = \begin{pmatrix} 2xy & x^2 & 0 \\ 0 & 2y & 0 \\ -ze^{-xz} & 0 & -xe^{-xz} \end{pmatrix}$$

$$\nabla f(u, v, w) = (2u, 2v, -1) \Rightarrow \nabla f(x, y, z) = (2x^2y, 2y^2, -1)$$

$$\nabla f(x, y, z) \mathbf{Dg}(x, y, z) = (2x^2y, 2y^2, -1) \begin{pmatrix} 2xy & x^2 & 0 \\ 0 & 2y & 0 \\ -ze^{-xz} & 0 & -xe^{-xz} \end{pmatrix} = (4x^3y^2 + ze^{-xz}, 2x^4y + 4y^3, xe^{-xz})$$

Ejemplo del uso de la regla de la cadena:

Plano tangente a un conjunto de nivel de una función $f(x, y, z)$.

Sea un función escalar, $f(x, y, z)$ de $\mathbb{R}^3 \rightarrow \mathbb{R}$, y un conjunto de nivel, c , $f(x, y, z) = c$, y una línea, una trayectoria, parametrizada, de puntos de ese conjunto de nivel $\mathbf{r}(t)$, con $t \in [t_0, t_e]$, donde en $t = t_0$, nos da $\mathbf{r}(t_0) = (x_0, y_0, z_0)$.

Si hacemos la función compuesta $h = f \circ \mathbf{r}(t)$, que fíjate, es simplemente una función real de una dimensión, $h(t) = f(\mathbf{r}(t))$, $\mathbb{R} \rightarrow \mathbb{R}$.

En este caso dado que los puntos de $\mathbf{r}(t)$ están en el conjunto de nivel de $f(x, y, z) = c$, la función siempre vale c , y por lo tanto su derivada es nula.

$$\frac{dh(t_0)}{dt} = 0.$$

Que por la regla de la cadena es lo mismo que:

$$\nabla f(\mathbf{r}(t_0)) \mathbf{r}'(t_0) = 0$$

Que corresponde a dos vectores ortogonales. El primero es el gradiente de $f(x, y, z)$ en (x_0, y_0, z_0) , esto es, $\nabla f(x_0, y_0, z_0)$. El segundo es el vector “velocidad” de la trayectoria en t_0 .

En un espacio \mathbb{R}^3 , dado un punto $\mathbf{x}_0 = (x_0, y_0, z_0)$ y vector \mathbf{n} . Podemos dar la ecuación del plano normal a \mathbf{n} , que pasa por \mathbf{x}_0 , como aquellos puntos, $\mathbf{x} = (x, y, z)$ cuyo vector, $\mathbf{x} - \mathbf{x}_0$, $(x - x_0, y - y_0, z - z_0)$, es ortogonal a \mathbf{n} , esto es:

$$(\mathbf{x} - \mathbf{x}_0) \mathbf{n} = 0$$

Si nos fijamos en el caso anterior, la ‘velocidad’ de la trayectoria, $\mathbf{r}'(t_0)$, debe estar contenida entonces en el plano definido por el punto (x_0, y_0, z_0) y el vector $\nabla f(x_0, y_0, z_0)$.

Ese plano, que es el plano tangente a al conjunto de nivel $f(x, y, z) = c$ en (x_0, y_0, z_0) , viene dado por la ecuación:

$$(\mathbf{x} - \mathbf{x}_0) \nabla f(\mathbf{x}_0) = 0$$

$$f'_x(x_0, y_0, z_0)(x - x_0) + f'_y(x_0, y_0, z_0)(y - y_0) + f'_z(x_0, y_0, z_0)(z - z_0) = 0.$$

Date cuenta que esta forma de dar el plano, solo sirve para funciones, $f(x, y, z)$, cuyo gradiente en un punto, $\nabla f(x_0, y_0, z_0)$, sea distinto de cero.

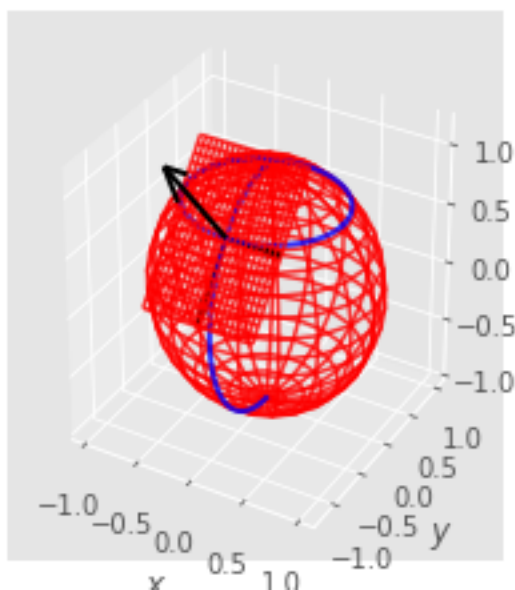
La siguiente celda muestra el plano tangente en un punto al conjunto de nivel de la función $f(x, y, z) = x^2 + y^2 + z^2$ de valor $c = 1$, que corresponde a la esfera de radio unidad.

También se muestra el gradiente de la función en el punto de la esfera dado por $\theta = \pi/4$, $\phi = 3\pi/2$, y dos trayectorias dentro de la esfera, una vertical, un meridiano, y otra horizontal, un paralelo, que pasan por ese punto. Puedes ver también los vectores “velocidad” de esas trayectorias, que están, claro, dentro del plano.

Explora y cambia el punto de la esfera.

```
In [13]: theta0, phi0 = 1.*np.pi/4., 3.*np.pi/2.
_, _, grad, vphi, vtheta = gf.sphere_plane(theta0, phi0, 0.5);
print('gradiente   :', grad)
print('v-paralelo   :', vphi)
print('v-meridiano  :', vtheta)

gradiente   : (-2.5978681687064796e-16, -1.414213562373095, 1.4142135623730951)
v-paralelo   : (0.7071067811865475, -1.2989340843532398e-16, 0.0)
v-meridiano  : (-1.29893408435324e-16, -0.7071067811865476, -0.7071067811865475)
```



Ejercicio: Sea la función $f(x, y, z) = x^2 + y^2 + z^2$, y el punto $\mathbf{x}_0 = \left(0, -1/\sqrt{2}, 1/\sqrt{2}\right)$, calcula el plano tangente del conjunto de nivel de la función a la que pertenece el punto.

El valor de la función en el punto es $f(\mathbf{x}_0) = 1$, luego el punto pertenece al conjunto de nivel con valor $c = 1$.

El conjunto de nivel $f(x, y, z) = 1$ es la esfera de radio unidad: $x^2 + y^2 + z^2 = 1$.

El punto \mathbf{x}_0 , en coordenadas esféricas es $r = 1, \theta = \pi/4, \phi = 3\pi/2$.

El gradiente de la función es: $\nabla f(x, y, z) = (2x, 2y, 2z)$.

Y en el punto \mathbf{x}_0 vale: $\nabla f(\mathbf{x}_0) = (0, -\sqrt{2}, \sqrt{2})$.

El plano tangente, en un punto \mathbf{x}_0 de la esfera cumple la ecuación:

$$2x(x - x_0) + 2y(y - y_0) + 2z(z - z_0) = 0$$

Y en el punto \mathbf{x}_0 es:

$$-\sqrt{2} \left(y + \frac{1}{\sqrt{2}} \right) + \sqrt{2} \left(z - \frac{1}{\sqrt{2}} \right) = 0$$

Por lo tanto el plano tangente tiene como ecuación:

$$z = y + \sqrt{2}$$

¡Aún hay más!

Las simetrías en Física son fundamentales. Fue la gran física Emmy Noether quien formuló un teorema que afirma que detrás de una simetría en un problema físico hay una cantidad conservada.

Por ejemplo, si el problema tiene simetría esférica, la cantidad conservada es el momento angular.

Emmy Noether sufrió desafortunadamente los males de su tiempo (y que aún perdura en nuestros), el machismo de sus compañeros de Universidad y el antisemitismo nazi.

Aquí tienes su entrada en la Wikipedia: https://es.wikipedia.org/wiki/Emmy_Noether

¡Esto es todo por ahora!

6.4 Apéndice

Demostración de la regla de la cadena

Las dos funciones son diferenciables:

$$\mathbf{g}(\mathbf{x} + \mathbf{v}) - \mathbf{g}(\mathbf{x}) = \mathbf{Dg}(\mathbf{x}) \mathbf{v} + \|\mathbf{v}\| \mathbf{E}_g(\mathbf{x}, \mathbf{v})$$

$$\mathbf{f}(\mathbf{y} + \mathbf{u}) - \mathbf{f}(\mathbf{y}) = \mathbf{Df}(\mathbf{y}) \mathbf{u} + \|\mathbf{u}\| \mathbf{E}_f(\mathbf{y}, \mathbf{u})$$

Como $\mathbf{u} = \mathbf{g}(\mathbf{x} + \mathbf{v}) - \mathbf{g}(\mathbf{x})$, tenemos:

$$\begin{aligned} \mathbf{f}(\mathbf{y} + \mathbf{u}) - \mathbf{f}(\mathbf{y}) &= \mathbf{Df}(\mathbf{y}) (\mathbf{Dg}(\mathbf{x}) \mathbf{v} + \|\mathbf{v}\| \mathbf{E}_g(\mathbf{x}, \mathbf{v})) + \|\mathbf{u}\| \mathbf{E}_f(\mathbf{y}, \mathbf{u}) \\ &= \mathbf{Df}(\mathbf{y}) \mathbf{Dg}(\mathbf{x}) \mathbf{v} + \|\mathbf{v}\| \left(\mathbf{Df}(\mathbf{y}) \mathbf{E}_g(\mathbf{x}, \mathbf{v}) + \frac{\|\mathbf{u}\|}{\|\mathbf{v}\|} \mathbf{E}_f(\mathbf{y}, \mathbf{u}) \right) \end{aligned}$$

Que podemos asociar a:

$$\mathbf{h}(\mathbf{x} + \mathbf{v}) - \mathbf{h}(\mathbf{x}) = \mathbf{Dh}(\mathbf{x}) \mathbf{v} + \|\mathbf{v}\| \mathbf{E}_h(\mathbf{x}, \mathbf{v})$$

Donde

$$\mathbf{Dh}(\mathbf{x}) = \mathbf{Df}(\mathbf{y}) \mathbf{Dg}(\mathbf{x})$$

$$\mathbf{E}_h(\mathbf{x}, \mathbf{v}) = \mathbf{Df}(\mathbf{y}) \mathbf{E}_g(\mathbf{x}, \mathbf{v}) + \frac{\|\mathbf{u}\|}{\|\mathbf{v}\|} \mathbf{E}_f(\mathbf{y}, \mathbf{v})$$

La derivada es el producto de las derivadas.

Solo falta comprobar que: $\lim_{\mathbf{v} \rightarrow \mathbf{0}} \mathbf{E}_h(\mathbf{x}, \mathbf{v}) \rightarrow \mathbf{0}$.

El primer termino es directo $\mathbf{E}_g(\mathbf{x}, \mathbf{v}) \rightarrow \mathbf{0}$.

En el segundo podemos acotar, $\frac{\|\mathbf{u}\|}{\|\mathbf{v}\|} \leq M_g(\mathbf{x}) + \|\mathbf{E}_g(\mathbf{x}, \mathbf{v})\|$, donde $M_g(\mathbf{x}) = \sum_{k=1}^n \|\nabla g_k(\mathbf{x})\|$.

ya que $\mathbf{u} = \mathbf{g}(\mathbf{x} + \mathbf{v}) - \mathbf{g}(\mathbf{x}) = \mathbf{Dg}(\mathbf{x})\mathbf{v} + \|\mathbf{v}\| \mathbf{E}_g(\mathbf{x}, \mathbf{v})$ y por lo tanto $\|\mathbf{u}\| \leq (\sum_{i=1}^n \|\nabla g_k\| + \|\mathbf{E}_g(\mathbf{x}, \mathbf{v})\|) \|\mathbf{v}\|$.

Cuando $\mathbf{v} \rightarrow \mathbf{0}$ también $\mathbf{u} \rightarrow \mathbf{0}$ y $\mathbf{E}_f(\mathbf{y}, \mathbf{u}) \rightarrow \mathbf{0}$.