2. 28

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

OPTIMIZACION EN ESPACIOS DE FUNCIONES

Tesis que para obtener el título de:
MATEMATICO
Presenta:
FRANCISCO VENEGAS MARTINEZ

México, D.F. 1983

INDICE

CAPITULO UNO:

INTRODUCCION

1. INTRODUCCION:

El principiante... no se desalentaría, si... él encuentra que no tiene los prerrequisitos necesarios para leer los requisitos. P. Halmos.

Todo mundo, sabe lo que es una curva, hasta aquel que ha estudiado suficientes Matemáticas para llegar a confundirse por una cantidad innumerable de posibles excepciones. F. Klein.

1.1 ANTECEDENTES Y MOTIVACION

Durante las últimas décadas, la matemática aplicada y la ingeniería se han dirigido en forma ascendente hacia la resolución de problemas de la teoría de control, programación dinámica, teoría de información y teoría de perturbaciones mediante la aplicación de métodos variacionales. Esta tendencia ha sido motivada primeramente por lo significativo que es caracterizar una solución óptima y por la repetida demostración de que tales problemas pueden ser formulados y analizados matemáticamente de manera realista.
El estudio de la teoría de optimización en espacios de funciones, como un tópico independiente debe, por supuesto, ser considerada como una rama de la matemática aplicada y como tal, esta debe ser conformada por varias areas de la matemática pura por su unificación, por su formalismo y por sus fundamentos generales. Una de tales areas de particular relevancia, lo es sin duda el análisis funcional.

El propósito de esta tesis es el de presentar los elementos necesarios del análisis funcional para dar un tratamiento lo suficientemente moderno a la teoría de optimización en espacios de funciones, de hecho, a los métodos variacionales y a través de este enfoque relacionar estos con la teoría de control, puesto que los métodos variacionales son centrales para el entendimiento de esta última. Así, como el de presentar la relación que en muchas situaciones surge entre la teoría de control y la programación dinámica y finalmente discutir sobre algunas aplicaciones de los métodos variacionales a la estadística.

Mucho de la teoría clásica de optimización en espacios de funciones, ha sido motivada esencialmente por problemas resultantes de sistemas físicos. En el siglo XVII, Newton trataba de seleccionar la coraza de un barco, de tal manera que la resistencia que ofreciera el agua fuera mínima mediante el planteamiento de un problema variacional.
La investigación de una trayectoria mínima para un móvil que debe mantenerse en contacto sobre una superficie determinada, es un problema ya de antiguo ilustrado, tal problema es tratado también por los métodos variacionales y su interés no es solamente histórico puesto que cada vez con más frecuencia el matemático y el ingeniero se encuentran en presencia de problemas de optimización, cuyo estudio revela que son de esta naturaleza o similares, los cuales han sido reexaminados, extendidos, algunas veces redescubiertos y aplicados a problemas que tienen origenes completamente diferentes que aquellos tenían por su temprano desarrollo.

A manera de motivación se presenta la siguiente serie de situaciones de algunas de las múltiples aplicaciones de los métodos variacionales.

Ejemplo. 1.1.1 (Pozo Petrolero). Una compañía tiene localizado un yacimiento petrolífero, con una capacidad c, que se supone conocida. Se desea determinar un programa de extracción a largo plazo, tal que maximice el beneficio total descontado.
El problema puede ser formulado como sigue:
Se desea encontrar una función Riemann-integrable sobre $[0,\infty)$, f, que representa la cantidad extraída, tal que maximice

$$\int_0^\infty F(f(t))g(t)dt,$$

sujeto a

$$\int_0^\infty f(t)dt\le c \quad ; \; f(t)\ge 0,$$

La función F(f), representa el beneficio asociado con la cantidad extraída $f=f(t)$, esta función puede considerarse como estrictamente cóncava, y creciente sobre $[0,\infty)$, con $F(0)=0$, suponemos también que $F\in C^1[0,\infty)$. La función $g=g(t)$ representa el factor de descuento, y puede ser supuesta continua, positiva y estrictamente decreciente hacia 0, e integrable sobre $[0,\infty)$.
Primero, consideremos el problema con la restricción adicional de que, $f\in C_0[0,T]=\{f\in C[0,T]\,;\, f(0)=f(T)=0\}$ y $f(t)=0$ para $t\geq T$, con T un tiempo positivo fijo. Definamos ahora, una norma sobre $C[0,T]$ mediante $\|f\|=\max\limits_{0\le t\le T}|f(t)|$.
Consideremos la funcional

$$J(f,p)=\int_0^T F(f(t))g(t)dt+p\left(c-\int_0^T f(t)dt\right),$$

en donde p es el multiplicador de Lagrange asociado a la restricción $\int_0^T f(t)dt\le c$. En este caso buscamos (f_0,p_0), tal que

$$\min_{p\geq 0}\max_{f\geq 0} J(f,p)=J(f_0,p_0).$$

Sea $h\in C_0[0,T]$ y consideremos la diferencia

$$J(f+h,p_0)-J(f,p_0)=\int_0^T (F((f+h(t))-F(f(t)))g(t)dt-p_0\int_0^T h(t)dt$$

Por el teorema de Taylor, tenemos que

$$J(f+h,p_o)-J(f,p_o)=\int_0^T F_f h(t)g(t)dt-p_o\int_0^T h(t)dt+o(\|h\|)$$
$$=L(f,p_o)(h)+o(\|h\|),$$

en donde

$$L(f,p_o)(h)=\int_0^T(F_f g(t)-p_o)h(t)dt$$

y $o(\|h\|)$ es un infinitésimo de orden superior al primero en $\|h\|$. Si f_o maximiza a $J(f,p_o)$, debemos tener necesariamente que

$L(f_o,p_o)(h)=0$ para toda $h\in C_o[0,T]$.

Notamos primero que $L(f_o,p_o)(h)$ define una funcional lineal sobre el conjunto $C_o[0,T]$, esto es

$L(f_o,p_o)(ah_1+bh_2)=aL(f_o,p_o)(h_1)+bL(f_o,p_o)(h_2)$ para todo $a,b\in\mathbb{R}$ y para todo $h_1,h_2\in C_o[0,T]$.

Ahora bien, para $\|h\|$ suficientemente pequeña, el signo de la expresión $J(f_o+h,p_o)-J(f_o,p_o)$ coincide con el signo de $L(f_o,p_o)(h)$. Supongamos ahora que por el contrario $L(f_o,p_o)(h_o)\neq 0$ para algún $h_o\in C_o[0,T]$, entonces para valores de $c>0$ suficientemente pequeños

$-cL(f_o,p_o)(h_o)=L(f_o,p_o)(-ch_o)$

así

$J(f_o-ch_o,p_o)-J(f_o,p_o)=-cL(f_o,p_o)(h_o)+o(c)$

de donde se implica que $J(f_o-ch_o,p_o)-J(f_o,p_o)$ puede cambiar de signo para valores de $c>0$ suficientemente pequeños y por lo tanto de $\|ch_o\|$, pero esto es imposible, puesto que por hipótesis $J(f_o+h,p_o)-J(f_o,p_o)\leq 0$ para toda $\|h\|$ suficientemente pequeña, lo cual contradice el hecho de que $J(f,p_o)$ tenga un máximo en f_o. La condición $L(f_o,p_o)(h)=0$ para toda $h\in C_o[0,T]$, equivale a que $F_f(f(t))g(t)-p_o=0$ sobre f_o,

en efecto supongamos que $(F_f g(t)-p_o)\big|_{f=f_o}>0$ para algún t_o, entonces debido a la continuidad de $(F_f g(t)-p_o)\big|_{f=f_o}$ se tiene que $(F_f g(t)-p_o)\big|_{f=f_o}>0$ en algún subintervalo $[t_1,t_2]$ de $[0,T]$, sea

$$h(t)=\begin{cases}(t-t_1)(t_2-t) & \text{si } t\in[t_1,t_2].\\ 0 & \text{si } t\in[0,T]-[t_1,t_2].\end{cases}$$

entonces $h\in C_o[0,T]$ y

$\int_0^T (F_f(f(t))g(t)-p_o)h(t)dt>0$

lo cual contradice el hecho de que $L(f_o,p_o)\equiv 0$.

Por lo tanto si f_o maximiza a $J(f,p_o)$, entonces

$(F_f(f(t))g(t)-p_o)\big|_{f=f_o}=0.$

Por consiguiente tenemos que $f_o=f_o(t)$ y p_o satisfacen para toda $t\in[0,T]$,

$f_o(t)\geq 0$; $p_o\geq 0$; $\int_0^T f_o(t)dt=c$; $(F_f(f(t))g(t)-p_o)\big|_{f=f_o}=0.$

Puesto que F_f es continua y estrictamente decreciente, ya que F es cóncava y creciente, esta tiene una inversa F_f^{-1} la cual es continua y estrictamente decreciente, entonces $f_o(t)=F^{-1}(p_o/g(t))$, $t\in[0,T]$ la cual es decreciente puesto que $p_o/g(t)$ es creciente. Se sigue que si f_o maximiza a $J(f,p_o)$, f_o debe tener la forma

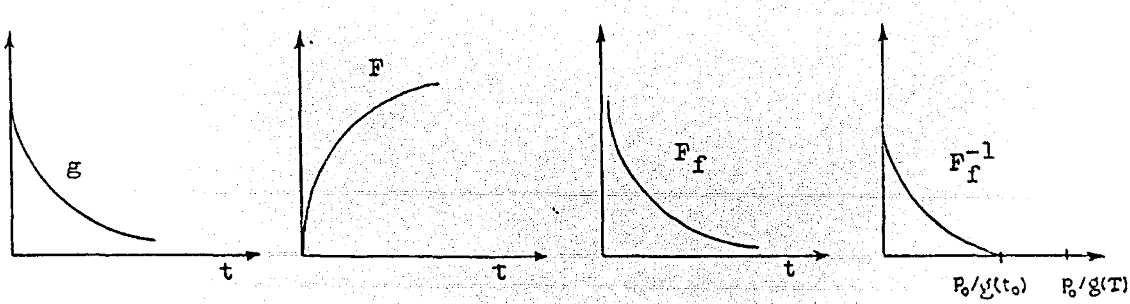
$$f_o(t)=\begin{cases} F_f^{-1}(p_o/g(t)), & 0\leq t\leq t_o \\ 0 & , t_o\leq t\leq T. \end{cases}$$

En esta solución podemos tener $t_o=T$, si $t_o<T$ con el propósito de mantener continuidad en la solución, debemos tener que

$F_f^{-1}(p_o/g(t_o))=0$

y por lo tanto

$p_o=F_f(0)g(t_o)$



Definamos ahora

$K(t_o)=\int_0^{t_o} F_f^{-1}(F_f(0)g(t_o)/g(t))dt$

claramente la función K es continua y $\lim_{t_o\to 0} K(t_o)=0$, $\lim_{t_o\to\infty} K(t_o)=\infty$.

Por lo tanto, podemos escoger t_o tal que $K(t_o)=c$. Por consiguiente si $t_o<T$, tenemos la solución

$$f_o(t)=\begin{cases} F_f^{-1}(F_f(0)g(t_o)/g(t)), & 0\leq t\leq t_o \\ 0 & , t_o\leq t\leq T. \end{cases}$$

Ejemplo. 1.1.2 (Problema de asignación del granjero).

Un granjero produce un sólo cultivo, digamos trigo. Después de la cosecha debe almacenarlo o venderlo y reinvertirlo comprando tierra adicional y equipo para incrementar su producción. El granjero desea maximizar la cantidad total almacenada durante un tiempo T.

Sean: $f_1(t)$=producción total en el tiempo t,

$f_2(t)$=cantidad invertida en el tiempo t y

$f_3(t)$=almacenamiento en el tiempo t,

pongamos:

$\frac{d}{dt}f_1(t)=f_2(t)$, $t\in[0,T]$;

$f_2(t)+f_3(t)=f_1(t)$, $t\in[0,T]$;

$f_1(t),f_2(t),f_3(t)\geq 0$, $t\in[0,T]$; $f_1(0)>0$.

El granjero desea operar de tal forma de maximizar

$\int_0^T f_3(t)dt$,

este problema puede ser expresado completamente en términos de $f_2(t)$ como sigue:

maximizar $\int_0^T(\int_0^t f_2(s)ds-f_2(t))dt+f_1(0)T$

sujeto a

$f_2(t)\geq 0$, $f_1(t)=f_1(0)+\int_0^t f_2(s)ds\geq f_2(t)$, $t\in[0,T]$.

Finalmente planteamos el problema anterior como un problema de control óptimo, sea $u(t)$ la tasa de producción que es reinvertida al tiempo t. El problema es entonces descrito por:

maximizar $\int_0^T (1-u(t))f_1(t)dt$

sujeto a

$\frac{d}{dt}f_1(t)=u(t)f_1(t)$, $0\leq u(t)\leq 1$, $f_1(t)\geq 0$, $t\in[0,T]$, $f_1(0)>0$.

Este problema se resolverá en la sección 3.4 utilizando el principio del máximo de Pontryagin.

Ejemplo. 1.1.3

Consideremos el problema de acotar inferiormente la funcional $J(f)=\int_0^\infty f(x)dF(x)$, $F(x)$ es una función monótona creciente conocida, sujeto a

$\int_0^\infty x dF(x)=m_1$, $\int_0^\infty x^2 dF(x)=m_2>0$, $m_2-m_1^2>0$,

$f\in V$.

Una función f pertenece a V, si satisface:

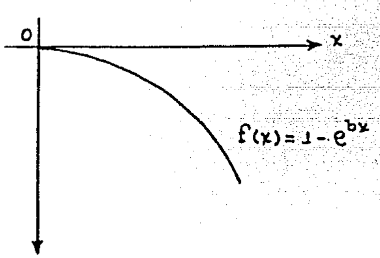
(i). $f\in C^3[0,\infty)$.

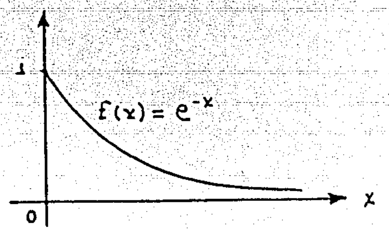
(ii). $f(x)$ es acotada inferior y superiormente por ciertas funciones del tipo $a_1x+a_2x^2$ para toda $x\in[0,\infty)$.

(iii). $f(x)/x$ es estrictamente convexa para toda $x>0$.

Por ejemplo, la función $f(x)=1-e^{bx}$; $x\geq 0$; $b>0$ pertenece a V, la función $f(x)=e^{-x}$, $x\geq 0$ no pertenece a V puesto que no satisface la condición (iii).

Problemas de este tipo se resolverán en la sección 3.6 utilizando el teorema de Hahn-Banach.





Ejemplo. 1.1.4 En el estudio de la teoría de información se presentan los siguientes problemas variacionales:

(I) Maximizar $J(f)=-\int_0^\infty f(x)\ln f(x)dx$

sujeto a

$f(x)\geq 0,\ x>0;\ \int_0^\infty f(x)dx=1;\ \int_0^\infty xf(x)dx=m>0,$

y

(II) Maximizar $J(f)=-\int_{-\infty}^\infty f(x)\ln f(x)dx$

sujeto a

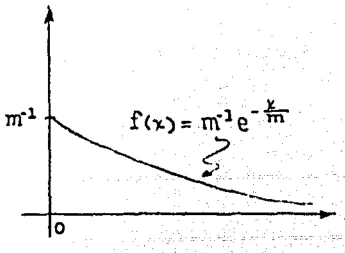
$f(x)\geq 0,\ x\in(-\infty,\infty);\ \int_{-\infty}^\infty f(x)dx=1;\ \int_{-\infty}^\infty x^2f(x)dx=s^2,\ s>0.$

Problemas de este tipo serán discutidos y resueltos en el transcurso de la sección 3.8 . Se verá que la solución de (I) corresponde a la función de densidad de probabilidad exponencial

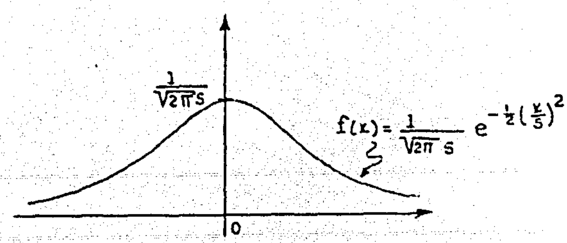
$$f(x)=m^{-1}\exp(-x/m),\ x>0,\ m>0,$$

y la solución de (II) corresponde a la función de densidad de probabilidad Gaussiana

$$f(x)=(\sqrt{2\pi}\ s)^{-1}\exp\{-\tfrac{1}{2}(x/s)^2\},\ x\in(-\infty,\infty),\ s>0.$$



(I)



(II)

Ejemplo. 1.1.5 Sea el estado de un sistema dinámico descrito por $f(t)$ en el tiempo t, $0\leq t\leq T$. Supongamos que el sistema es controlado por $u(t)$ y que la ecuación que gobierna al sistema es

$$f'(t)=af(t)+u(t) \quad 0\leq t\leq T,\ a=cte>0,$$

en donde

$$f'(t)=\frac{d}{dt}f(t),$$

y considere los siguientes restricciones (condiciones de frontera)

Caso 1. Sobre el estado inicial del sistema

$$f(0)=c_1.$$

Caso 2. Sobre los estados inicial y final del sistema

$$f(0)=c_1\ ,\ f(T)=c_2\ .$$

Supongamos finalmente que una medida de la evaluación del sistema está dada por la funcional

$$J(f,u)=\int_0^T(f^2(t)+u^2(t))dt$$

en donde

$$f\in(C^1[0,T],\|f\|=\max_{t\in[0,T]}|f(t)|+\max_{t\in[0,T]}|f'(t)|),\ u,f''\in(PC[0,T],\|u\|=\max_{t\in[0,T]}|u(t)|)$$

y $C^1[0,T]$, $PC[0,T]$ denotan el espacio de funciones con derivada continua sobre $[0,T]$ y el espacio de funciones continuas por pedazos sobre $[0,T]$ con las normas asociadas correspondientes, respectivamente.

Nuestro objetivo es encontrar un control $u(t)$ tal que minimice a la funcional $J(f,u)$ sujeto al conjunto de restricciones antes mencionadas. En otras palabras, deseamos resolver el problema de encontrar un control $u_o(t)$ tal que

$$J(f,u_o)=\min J(f,u)$$

sujeto a

$$f'(t)=af(t)+u(t),\ 0\leq t\leq T,\ a=cte>0,$$
$$f(0)=c_1,\ f(T)=c_2$$
$$f\in C^1[0,T];\ f'',u\in PC[0,T],$$

con las respectivas normas asociadas.

Aunque el propósito de esta tesis es el de mostrar las técnicas de solución a problemas de este tipo, proporcionamos anticipadamente una alternativa para resolver este problema, para ello, procedamos como sigue:

A modo de simplificar cálculos posteriores denotemos al integrando de la funcional $J(f,u)$ por $G(f,u)=f^2(t)+u^2(t)$, $0\leq t\leq T$.
Ahora bien, tomemos $h\in C^1[0,T]\cap C_0[0,T]$, es decir, $h\in C^1[0,T]$ y h de soporte compacto sobre $[0,T]$ $(h(0)=h(T)=0)$ e incrementemos f a f+h.
Por otra parte, notamos que $u(t)=f'(t)-af(t)$, $0\leq t\leq T$ y esta relación define una única función $u=u(f)$, por consiguiente

$J(f,u)=J(f)$, $G(f,u(f))=F(f,f')=f^2(t)+(f'(t)-af(t))^2$. De esta manera, es claro que $F\in C^2(\Omega)$, Ω una región que contiene a (f,f')
Calculemos la diferencia $J(f+h)-J(f)$ empleando el teorema de Taylor

$$
\begin{aligned}
J(f+h)-J(f)&=\int_0^T(F(f+h,f'+h')-F(f,f'))dt\\
&=\int_0^T(F_f(f,f')h(t)+F_{f'}(f,f')h'(t))dt\\
&+\tfrac{1}{2}\int_0^T(F_{ff}(f+\theta h,f'+\theta h')h^2(t)+2F_{ff'}(f+\theta h,f'+\theta h')h(t)h'(t)\\
&+F_{f'f'}(f+\theta h,f'+\theta h')h'^2(t))dt,
\end{aligned}
$$

en donde $\theta\in(0,1)$,
es claro que existe un $M>0$, tal que
$M\geq|F_{ff}|,|F_{ff'}|,|F_{f'f'}|$ para toda $t\in[0,T]$, entonces

$$
\left|\tfrac{1}{2}\int_0^T(F_{ff}(f+\theta h,f'+\theta h')h^2(t)+2F_{ff'}(f+\theta h,f'+\theta h')+F_{f'f'}(f+\theta h,f'+\theta h')h'^2(t))dt\right|\leq \tfrac{1}{2}MT\|h\|^2=o(\|h\|),
$$

en consecuencia

$$
J(f+h)-J(f)=\int_0^T(F_f(f,f')h(t)+F_{f'}(f,f')h'(t))dt+o(\|h\|).
$$

Entonces

$$
\begin{aligned}
J(f+h)-J(f)&=\int_0^T F_f(f,f')h(t)dt+F_{f'}(f,f')h(t)\Big|_0^T-\int_0^T(\tfrac{d}{dt}F_{f'}(f,f'))h(t)dt\\
&+o(\|h\|)\\
&=\int_0^T(F_f(f,f')-\tfrac{d}{dt}F_{f'}(f,f'))h(t)dt+o(\|h\|),
\end{aligned}
$$

denotemos

$$L(f,u)(h)=\int_0^T (F_f(f,f')-\frac{d}{dt}F_{f'}(f,f'))h(t)dt.$$

Observamos que $L(f,u)(h)$ es una funcional lineal sobre $C^1[0,T]\cap C_o[0,T]$, es decir,

$L(f,u)(a_1h_1+a_2h_2)=a_1L(f,u)(h_1)+a_2L(f,u)(h_2)$ para todo $a_1,a_2\in\mathbb{R}$ y $h_1,h_2\in C^1[0,T]\cap C_o[0,T]$.

Notamos que por definición $|o(\|h\|)|/\|h\| \to 0$ cuando $\|h\| \to 0$, es decir $o(\|h\|)$ es un infinitésimo de orden superior al primero con respecto de $\|h\|$. Mostremos ahora que si $u_o=u_o(t)$ minimiza a $J(f,u)$, entonces $L(f_o,u_o)(h)=0$ para toda $h\in C^1[0,T]\cap C_o[0,T]$, en efecto, dado que $o(\|h\|)$ es un infinitésimo de orden superior al primero en $\|h\|$, para $\|h\|$ suficientemente pequeña, el signo de la expresión $J(f_o+h,u_o)-J(f_o,u_o)$ coincide con el signo de $L(f_o,u_o)(h)$. Supongamos que por el contrario $L(f_o,u_o)(h_o)\neq 0$ para algún $h_o\in C^1[0,T]\cap C_o[0,T]$, entonces para valores de $c>0$ suficientemente pequeños se tiene que

$$L(f_o,u_o)(-ch_o)=-cL(f_o,u_o)(h_o)$$

así,

$$\begin{aligned}J(f_o-ch_o,u_o)-J(f_o,u_o)&=-cL(f_o,u_o)(h_o)+o(c\|h_o\|)\\&=-cL(f_o,u_o)(h_o)+o(c),\end{aligned}$$

entonces $J(f_o-ch_o,u_o)-J(f_o,u_o)$ puede cambiar de signo para valores de c arbitrariamente pequeños y por lo tanto de $c\|h_o\|$, pero esto es imposible, puesto que por hipótesis $J(f,u)$ posee un mínimo en u_o, es decir $J(f_o+h,u_o)-J(f_o,u_o)\geq 0$ para toda $\|h\|$ suficientemente pequeña. Por lo tanto $L(f_o,u_o)(h)=0$ para toda $h\in C^1[0,T]\cap C_o[0,T]$. En otras palabras, si u_o minimiza $J(f,u)$ necesariamente

$$L(f_o,u_o)\equiv 0,$$

notamos que esta condición equivale a

$$\int_0^T (F_f(f,f')-\frac{d}{dt}F_{f'}(f,f'))h(t)dt=0 \text{ para toda } h\in C^1[0,T]\cap C_o[0,T]$$

y puesto que

$$F_f(f,f')-\frac{d}{dt}F_{f'}(f,f')=F_f(f,f')-F_{ff'}(f,f')f'-F_{f'f'}(f,f')f''$$

es una función continua por pedazos en $[0,T]$, veamos que $L(f_o,u_o)(h)=0$ para toda $h\in C^1[0,T]\cap C_o[0,T]$ implica que

$F_f(f,f')-\frac{d}{dt}F_{f'}(f,f')=0$ sobre la trayectoria óptima,

en efecto, supongamos que por el contrario

$(F_f(f,f')-\frac{d}{dt}F_{f'}(f,f'))\big|_{f=f_o}>0$ en algún punto $t_o\in[0,T]$, entonces

$(F_f(f,f')-\frac{d}{dt}F_{f'}(f,f'))\big|_{f=f_o}>0$ sobre algún intervalo que contiene a t_o, digamos $[t_1,t_2]$, y definamos

$$h(t)=\begin{cases}(t-t_1)(t_2-t) & t\in[t_1,t_2]\\ 0 & t\in[0,T]-[t_1,t_2].\end{cases}$$

es claro que $h\in C_o[0,T]\cap C^1[0,T]$ y

$\int_0^T(F_f(f,f')-\frac{d}{dt}F_{f'}(f,f'))\big|_{f=f_o}h(t)>0$

lo que contradice nuestra hipótesis.

Ahora bien, consideremos a la funcional $J(f,u)$ que evalúa al sistema. Para el caso 2, se tiene que

$$\begin{aligned}J(f,u)&=\int_0^T(f^2(t)+(f'(t)-af(t))^2)dt\\&=\int_0^T((1+a^2)f^2(t)+f'^2(t))dt-2a\int_0^T f(t)f'(t)dt\\&=\int_0^T((1+a^2)f^2(t)+f'^2(t))dt+a(c_1^2-c_2^2),\end{aligned}$$

entonces

$F_f(f,f')-\frac{d}{dt}F_{f'}(f,f')=0$ sobre la trayectoria óptima

equivale a

$f''(t)-(1+a^2)f(t)=0$ sobre la trayectoria óptima

Sea $b^2=(1+a^2)$, y consideremos el operador diferencial $D:C^1[0,T]\to PC[0,T]$, $D(f)=f''-b^2f$, $b>0$, si $f(t)=e^{rt}$, entonces $D(e^{rt})=r^2e^{rt}-b^2\cdot e^{rt}=(r^2-b^2)e^{rt}$, si $D(e^{rt})=0$, se debe tener $(r^2-b^2)=(r+b)(r-b)=0$, de donde $r=b,-b$, en consecuencia, tenemos que la solución general de la ecuación diferencial de segundo orden $f''-b^2f=0$ es $f_o(t)=k_1e^{bt}+k_2e^{-bt}$, k_1 y k_2 son determinados para el caso 2 por $c_1=k_1+k_2$ y $c_2=k_1e^{bT}+k_2e^{-bT}$, así que

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ e^{bT} & e^{-bT} \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix},$$

entonces

$$\begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} e^{-bT} & -1 \\ -e^{bT} & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

y puesto que

$$\det \begin{bmatrix} e^{-bT} & -1 \\ -e^{bT} & 1 \end{bmatrix} = e^{-bT} - e^{bT} \neq 0 \text{ si } T \neq 0$$

concluimos que la solución es única.

Finalmente, demos a f_o un incremento f_o+h con $h(0)=h(T)=0$, en donde, como se sabe f_o satisface la ecuación

$F_f(f,f')-\frac{d}{dt}F_{f'}(f,f')=0.$

Observamos también que

$(f_o+h)(0)=c_1$, $(f_o+h)(T)=c_2$

y hagamos la siguiente consideración sobre $J(f_o+h,u_o)$

$$\begin{aligned} J(f_o+h,u_o) &= \int_0^T (b^2(f_o(t)+h(t))^2+(f_o'(t)+h'(t))^2)dt + a(c_1^2-c_2^2) \\ &= \int_0^T (b^2f_o^2(t)+f_o'^2(t))dt + a(c_1^2-c_2^2) + \int_0^T (b^2h^2(t)+h'^2(t))dt \\ &\quad +2\int_0^T (b^2f_o(t)h(t)+f_o'(t)h'(t))dt \\ &= J(f_o,u_o)+\int_0^T (b^2h^2(t)+h'^2(t))dt \\ &\quad +2\int_0^T (b^2f_o(t)h(t)+f_o'(t)h'(t))dt, \end{aligned}$$

pero

$0 \leq \int_0^T (b^2h^2(t)+h'^2(t))dt \leq Tb^2\|h\|^2 = o(\|h\|)$ para todo $h \in C^1[0,T] \cap C_o[0,T]$,

además

$\int_0^T (b^2f_o(t)h(t)+f_o'(t)h'(t))dt = (f_o'(t)h(t))\Big|_0^T - \int_0^T (f_o''(t)-b^2f_o(t))h(t)dt = 0$

por consiguiente

$J(f_o+h,u_o)-J(f_o,u_o)=o(\|h\|)\geq 0$ para toda $h \in C^1[0,T] \cap C_o[0,T]$

de donde

$J(f_o+h,u_o)\geq J(f_o,u_o)$ para toda $h\in C^1[0,T]\cap C_o[0,T]$,
el mínimo de $J(f,u)$ es alcanzado cuando $h(t)=0$ para toda $t\in[0,T]$.
El caso 1 y otras situaciones especiales pueden discutirse de manera análoga.

Consideremos ahora, la analogía del problema anterior para el caso discreto, con el propósito de introducir el concepto de ecuación funcional de recursividad, así como el de presentar las técnicas que la programación dinámica proporciona para encontrar soluciones aproximadas.

Ejemplo. 1.1.6 En lugar de determinar funciones $f\in C^1[0,T]$ y $u\in PC[0,T]$, supongamos que es posible considerar valores de f y u,

$f_k=f(t_k)$, $u_k=(t_k)$, $k=0,1,\dots,N$,

en los puntos

$0=t_o<t_1<t_2<\dots<t_N=T$,

en donde

$t_k=kh$, $k=0,1,\dots,N$ y $h=T/N$.

Supongamos que el sistema puede ser descrito por el conjunto discreto de variables $f=(f_o,\dots,f_N)$ y supongamos que el sistema es gobernado por las ecuaciones de diferencias

$h(f_{k+1},f_k)=af_k+u_k \quad k=0,1,\dots,N-1$

con la condición de frontera $f_o=c$, en donde el vector $u=(u_o,u_1,\dots,u_N)$ denota las variables de control.

En lugar de resolver el problema

$$\min_u J(f,u)=\int_0^T(f^2(t)+u^2(t))dt$$

sujeto a

$f'(t)=af(t)+u(t)$, $0\leq t\leq T$, $a=cte>0$

$f(0)=c_1$, $f(T)=c_2$

$f\in C^1[0,T]$; $f'',u\in PC[0,T]$.

deseamos resolver el problema

$$\min_{u} J(f,u)=\sum_{k=0}^{N} (f_k^2+u_k^2)h$$

sujeto a

$$f_{k+1}=f_k+(af_k+u_k)h,\ k=0,1,\ldots,N-1,$$

$$f_o=c,\ \underline{u}_k\leq u_k\leq \bar{u}_k,\ f_k\in S_k \quad (S_k \text{ un conjunto conocido})$$

$\underline{u}_k, \bar{u}_k$ valores conocidos.

Observamos que el valor de f'_k ha sido aproximado por $\frac{f_{k+1}-f_k}{h}$.

Para resolver este problema, introduzcamos la función

$$G_N(c)=\min_{u} J(f,u).$$

Notemos que si u=0, entonces

$$J(f,0)=hc^2 \sum_{k=0}^{N} (1+ah)^{2k}$$

puesto que

$$f_{k+1}=(1+ah)f_k+hu_k,\ k=0,1,\ldots,N$$

y

$$f_k=(1+ah)f_{k-1}=(1+ah)^2 f_{k-2}=\ldots=(1+ah)^k f_o=(1+ah)^k c.$$

Observamos que en general la ecuación

$$G_N(c)=\min_{u} J(f,u)$$

define recursividad. Supongamos que u_o ha sido determinado, entonces

$$f_1=(1+ah)f_o+hu_o=(1+ah)c+hu_o$$

en consecuencia

$$J(f,u)=hc^2+hu_o^2+\sum_{k=1}^{N}(f_k^2+u_k^2)h$$

luego

$$G_{N-1}(f_1)=\min_{(u_1,\ldots u_N)}\sum_{k=1}^{N}(f_k^2+u_k^2)h,\ f_1=(1+ah)c+hu_o,$$

entonces

$$G_N(c)=\min_{u_o}\{(hc^2+hu_o^2)+G_{N-1}((1+ah)c+hu_o)\}$$

De esta manera el problema de minimización de J(f,u) en N+1 variables ha sido reducido a resolver N+1 problemas de una variable. La última ecuación es conocida como ecuación funcional de recursividad de Bellman e introduce las técnicas de la programación dinámica para resolver problemas de optimización del tipo descrito.

CAPITULO DOS:

ESPACIOS DE FUNCIONES

2. ESPACIOS DE FUNCIONES:

Los Matemáticos son como los Franceses: Todas las cosas que usted les dice las traducen a su propio lenguaje, y en ocasiones, algunas de esas cosas parecen ser completamente diferentes. J. W. Goethe.

2.1 INTRODUCCION.

En el presente capítulo se establecen los fundamentos y conceptos básicos del análisis funcional necesarios para dar un tratamiento lo suficientemente moderno a los métodos variacionales y a la teoría de control. El desarrollo del mismo se da como sigue: En la sección 2.2 se introduce la noción de función y se presentan los conceptos fundamentales de la teoría de conjuntos. En las secciones 2.3 y 2.4 se discuten las propiedades de los espacios lineales, álgebras, espacios métricos y espacios lineales métricos. Hacemos énfasis en la sección 2.4 sobre la propiedad de los espacios lineales métricos de poseer una completación. A partir de las propiedades anteriores, creamos una familia de definiciones útiles en el estudio de espacios topológicos y espacios lineales topológicos que se discuten el la sección siguiente, 2.5. Dentro de la teoría de optimización en espacios de funciones el estudio de tales espacios es importante. Por ejemplo, compacidad es una propiedad topológica y en ocasiones se desea optimizar una funcional continua cuyas restricciones forman un conjunto compacto, y como se sabe, ésta alcanza su óptimo dentro de dicho conjunto. A través de la sección 2.6 se establecen las propiedades de los espacios lineales normados, en donde se da especial importancia al teorema de Hahn-Banach y a sus corolarios, ya que permiten la mejor aplicación de los principios de la teoría de optimización en espacios de funciones en situaciones verdaderamente complicadas, tales situaciones surgen en una extensa clase de problemas de optimización al considerar la noción abstracta de integral como una funcional definida sobre algún espacio de funciones y sujeta a ciertas restricciones. Es en esta sección, en donde también se estudian los espacios y álgebras de Banach, que día a día, cobran mayor interés en el campo de la matemática aplicada. Finalmente en la sección 2.7 se estudian los espacios lineales con un producto interior, dando especial importancia a los espacios de Hilbert y a operadores adjuntos que actuan sobre ellos, puesto que en una gran variedad de problemas de óptimización en espacios de funciones las restricciones impuestas por ecuaciones, sean diferenciales, integrales, matriciales, etc., pueden ser representadas por operadores lineales. La resolución de tales problemas, casi invariablemente, llama a la consideración de un operador asociado, el adjunto. La razón de hacer esto es que en ocasiones los operadores adjuntos proporcionan un mecanismo conveniente para describir relaciones de ortogonalidad y dualidad, las cuales permiten todo análisis de optimización.

La causa de llevar a cabo este exhaustivo estudio sobre espacios de funciones, es la siguiente: A primera vista se podría pensar que la teoría de espacios lineales normados es lo suficientemente amplia para abastecer todas las necesidades del análisis funcional. Sin embargo, en el espacio de funciones infinitamente diferenciables definidas sobre algún intervalo cerrado, la topología natural no se puede definir mediante alguna norma y esto obliga a estudiar los espacios topológicos no normados, siendo estos tan importantes, que algunos de ellos representan la generalización de los espacios lineales normados numerables.

2.2 FUNCIONES Y CONJUNTOS.

Desde el comienzo de la matemática moderna. Los conceptos fundamentales sobre los cuales descansa la esencia toda del análisis, ya sea matemático o funcional, son los conceptos de función y límite de una sucesión infinita. Dentro del estudio de estos campos estos conceptos aparecen casi invariablemente ya sea implícita o explícitamente.

En el estudio de cualquier rama de las matemáticas, sea análisis, álgebra o geometría, resulta útil emplear la notación y la terminología de la teoría de conjuntos. Esta teoría, ha tenido una profunda influencia en el desarrollo de la matemática moderna. Ha unificado muchas ideas aparentemente inconexas y ha contribuido a reducir gran número de conceptos matemáticos a sus fundamentos lógicos por un método elegante y sistemático. Por fortuna, las nociones básicas de la teoría son un número reducido y familiarizarse con ellas resulta muy sencillo, por lo que se suponen dadas de antemano.

En esta sección se introduce el concepto de función, y a partir de la sección 2.4 se discutirá el concepto de límite de una sucesión infinita.

Definición. 2.2.1 Diremos que F es una función de un conjunto X a un conjunto Y, lo cual denotaremos por: $F:X \to Y$, o bien por: $f \mapsto F(f)$ con $f\in X$ y $F(f)\in Y$. Si F es una regla, la cual asigna a cada $f\in X$ un único $F(f)\in Y$, con frecuencia F es llamada también un mapeo o una aplicación de X en Y.

Si $A\subset X$, $B\subset Y$ y $F:X \to Y$, se define $F(A)=\{g\in Y ;$ existe $f\in X$ tal que $F(f)=g\}$, y $F^{-1}(B)=\{f\in X ; F(f)\in B\}$.

Definición. 2.2.2 Una función F es llamada inyectiva; si para cada $g\in F(X)$ existe a lo más un $f\in X$, tal que $F(f)=g$.
Una función F es llamada suprayectiva, si $F(X)=Y$.
Una función F inyectiva y suprayectiva, es llamada biyectiva.

Dentro del tratamiento axiomático del conjunto de números reales, $\mathbb{R}$, se introduce el axioma del supremo, debido a que algunos subconjuntos de $\mathbb{R}$ no vacíos están acotados superiormente, pero no poseen un elemento máximo. Para ellos damos un concepto que sustituye al de máximo, el supremo.
Un número S es llamado el supremo de un conjunto no vacío, X, el cual es acotado superiormente, si S es la mínima cota superior de X, equivalentemente S es el supremo de un conjunto $X\neq\emptyset$ acotado superiormente, si para toda $\epsilon>0$ existe un $x_0\in X$, tal que $x_0>S-\epsilon$.
Obviamente, el supremo de un conjunto $X\neq\emptyset$ acotado superiormente es único. Ahora establecemos el siguiente axioma en $\mathbb{R}$: Todo conjunto no vacío X acotado superiormente posee supremo. Insistimos en que el supremo de un conjunto X, no necesariamente pertenece a X.

Finalmente, a través de los siguientes enunciados hacemos un razonamiento muy sencillo, que será utilizado con mucha frecuencia en este capítulo y el siguiente.

<u>Proposición</u>. 2.2.3 Sean $F,H:V\to\mathbb{R}^+$, funciones de V a $\mathbb{R}^+$, tales que si $f\in V$, $f\neq 0$, entonces $H(f)\neq 0$ y existe una constante $C>0$, tal que $F(f)\leq CH(f)$ para toda $f\in V$, si $S'=\min\{C\ ;\ F(f)\leq CH(f)$ para toda $f\in V\}$ y $S=\sup\{F(f)/H(f)\ ;\ f\in V,\ f\neq 0\}$. Entonces $S'=S$.
Prueba: Por definición, para toda $\epsilon>0$, existe $g\in V$, $g\neq 0$, tal que $F(g)/H(g)>S-\epsilon$, esto implica que para toda $\epsilon>0$

$$CH(g)\geq F(g)>(S-\epsilon)H(g)$$

para aquellas C's tales que $F(g)\leq CH(g)$, entonces se tiene que $C>S-\epsilon$, en particular $S'>S-\epsilon$, de donde $S'\geq S$, veamos ahora que la desigualdad $S'>S$ es imposible. Sea $\epsilon>0$ tal que $S'-S=2\epsilon>0$, entonces para toda $f\in V$, $f\neq 0$

$$S'-\epsilon>S\geq F(f)/H(f),$$

de donde, para toda $f\in V$, $f\neq 0$, $H(f)S'>H(f)(S'-\epsilon)>F(f)$ y esto significa que S' no es el elemento mínimo del conjunto $\{C\ ;\ F(f)\leq CH(f)$ para toda $f\in V$, $f\neq 0\}$. En consecuencia se tiene que $S'=S$.

Corolario. 2.2.4 Sean $F,H:V \to \mathbb{R}^+$, funciones de V a $\mathbb{R}^+$, tales que si $f\in V$, $f\neq 0$, entonces $H(f)\neq 0$ y existe una constante $C\geq 0$, tal que $F(f)\leq CH(f)$ para toda $f\in V$. Sean $D=\{F(f)/H(f)\ ;\ f\in V,\ f\neq 0\}$, $S=\sup D$ y M una cota superior de D. Si existe un $g\in V$ $g\neq 0$, tal que $F(g)/H(g)=M$. Entonces $S=M$

Prueba: Dado que $F(f)/H(f)\leq M$ para toda $f\in V$, $f\neq 0$, se tiene que $S\leq M$ puesto que por definición de supremo, S es la mínima cota superior. Si además existe un $g\in V$, $g\neq 0$ para el cual se satisface $F(g)/H(g)=M$, en particular se tendrá que

$S\geq F(g)/H(g)=M$,

es decir $S\geq M$. En consecuencia $S=M$.

A continuación se presentan y discuten los conceptos de relación de equivalencia y orden parcial.

Una relación R sobre un conjunto $X\neq\emptyset$, es un subconjunto R del producto cartesiano $X\times X$; si $(f,g)\in R$ diremos que f esta relacionado con g.

Definición. 2.2.5 Una relación R es llamada de equivalencia, si para todo $f,g,h\in X$, se satisface que:

(i). $(f,f)\in R$.

(ii). $(f,g)\in R$ implica $(g,f)\in R$.

(iii). $(f,g)\in R$ y $(g,h)\in R$ implica $(f,h)\in R$.

El conjunto de elementos en X, que están relacionados a un $f\in X$ dado, es llamado la clase de equivalencia de f y es frecuentemente denotado por $[f]=\{g\in X\ ;\ (f,g)\in R\}$.

Ejemplo. 2.2.6 Sea $X= C[a,b]$, el conjunto de funciones continuas definidas sobre el intervalo cerrado $[a,b]$, decimos que $(f,g)\in R$ si y sólo si $f(x_0)=g(x_0)$ para un x_0 fijo en el intervalo $[a,b]$.

<u>Proposición</u>. 2.2.7 Sea R una relación de equivalencia sobre un conjunto X. Entonces cada $f\in X$ pertenece a una única clase de equivalencia.
Demostración: Supongamos que $f\in[g]$ y $f\in[h]$, es decir $(f,g)\in R$ y $(f,h)\in R$, entonces $(g,f)\in R$ y $(f,h)\in R$, así $(g,h)\in R$. Luego $f\in[g]$ y $(g,h)\in R$ implica $(f,g)\in R$ y $(g,h)\in R$, entonces $(f,h)\in R$, esto es $f\in[h]$; por tanto $[g]\subset[h]$. De manera análoga se prueba que $[h]\subset[g]$, así $[g]=[h]$.

<u>Definición</u>. 2.2.8 Una relación R sobre un conjunto no vacío X que satisface las propiedades (i) y (iii), además de que si $(f,g)\in R,(f,h)\in R$ implica $f=g$, es llamada un orden parcial y es usual denotar $f<g$ en lugar de $(f,g)\in R$.

<u>Ejemplo</u>. 2.2.9 Sea $X=C[a,b]$, definimos un orden parcial $<$ estableciendo que $f<g$ si y sólo si $f(x)\leq g(x)$ para todo $x\in[a,b]$.

Hemos usado la palabra "parcial" en la definición anterior debido a que dos elementos de X no siempre satisfacen o que $f<g$ o $g<f$, (ejemplo anterior).

<u>Definición</u>. 2.2.10 Un conjunto X es totalmente ordenado, si para todo $f,g\in X$ se cumple o que $f<g$ ó $g<f$.

<u>Ejemplo</u>. 2.2.11 Sea $X=\{f\ ;\ f(x)=ax,\ a\in[0,\infty),\ x\in(0,1]\}$, definimos un orden total $<$ sobre X estableciendo que: $f<g$ si y sólo si $f(x_0)\leq g(x_0)$ para un valor fijo $x_0\in(0,1]$.

Observamos que si X es un conjunto totalmente ordenado bajo $<$ y $Y\subset X$, entonces Y es también un conjunto totalmente ordenado bajo $<$.

<u>Definición</u> 2.2.12 Si X es un conjunto parcialmente ordenado bajo $<$ y $Y\subset X$, un elemento $f\in X$ es llamado cota superior para Y si $g<f$ para todo $g\in Y$.
Es importante observar que la existencia de un tal f depende de la naturaleza de los conjuntos X y Y.

<u>Definición</u>. 2.2.13 Si X es un conjunto parcialmente ordenado bajo $<$, decimos que f es un elemento maximal de X si no existe $g\in X$ tal que $g\neq f$ y $f<g$.

De nuevo hacemos notar que la existencia de un tal f depende de la naturaleza del conjunto X, notamos además, que un elemento maximal no necesariamente es una cota superior.

Ejemplos. 2.2.14 Sea X el conjunto del ejemplo 2.2.9, es claro que X no posee elementos maximales.
Sea X un conjunto y consideremos su conjunto potencia $P(X)$, definimos un orden parcial $\prec$ sobre $P(X)$, estableciendo que si $A,B\in P(X)$ entonces $A\prec B$ si y sólo si $A\subset B$. El único elemento maximal de $P(X)$ es X.

Dependiendo del punto de partida, el lema de Zorn es o una suposición básica de la teoría de conjuntos o es derivada de otras suposiciones básicas (este es equivalente al axioma de elección). Tomaremos el lema de Zorn y el resto de la teoría de conjuntos está dada. Estamos especialmente interesados en la aplicación explícita del lema de Zorn, en la demostración del teorema de Hahn-Banach.

Teorema. 2.2.15 (Lema de Zorn.): Sea X un conjunto no vacío parcialmente ordenado bajo $\prec$, tal que todo subconjunto totalmente ordenado tiene una cota superior en X. Entonces cada subconjunto totalmente ordenado tiene una cota superior que es también un elemento maximal de X.

2.3 ESPACIOS LINEALES Y ALGEBRAS.

Definición. 2.3.1 Un conjunto X es llamado un espacio lineal sobre el campo $\mathbb{R}$; si tenemos dos funciones $+:X\times X \to X$ y $\cdot:\mathbb{R}\times X \to X$ las cuales satisfacen las siguientes condiciones para todo $f,g,h\in X$ y $a,b\in\mathbb{R}$:

(i). $f+g=g+f$.
(ii). $(f+g)+h=f+(g+h)$.
(iii). Existe un $0\in X$ tal que $f+0=f$ para todo $f\in X$.
(iv). Para todo $f\in X$ existe un $-f\in X$ tal que $f+(-f)=0$.
(v). $a(f+g)=af+ag$.
(vi). $(a+b)f=af+bf$.
(vii). $a(bf)=(ab)f$.
(viii). $1\cdot f=f$ para todo $f\in X$.

Llamaremos a $+$ adición y $\cdot$ multiplicación por escalares.

Si además tenemos una función $\cdot : X\times X \to X$ que denotamos por: $\cdot(f,g) \mapsto f\cdot g = fg$, la cual satisface para todo $f,g,h\in X$ y $a\in\mathbb{R}$:

(ix). $(fg)h=f(gh)$.

(x). X tiene un elemento unidad e, tal que $ef=fe=f$.

(xi). $(f+g)h=fh+gh$, $f(g+h)=fg+fh$.

(xii). $a(fg)=(af)g=f(ag)$.

Se dice entonces que X forma un álgebra, si en adición se satisface que:

(xiii). $fg=gf$ para todo $f,g\in X$.

Entonces X es llamado un álgebra conmutativa.

Definición. 2.3.2 Si X es un espacio cuyos elementos son funciones. Por adición, multiplicación por escalares y producto entre elementos de X, entenderemos $(f+g)(x)=f(x)+g(x)$, $(af)(x)=af(x)$ y $(f\cdot g)(x)=f(x)\cdot g(x)$, respectivamente, y nos referiremos a ellas como operaciones usuales entre funciones.

Ejemplo. 2.3.3 Sea $X=C[a,b]$, el espacio de todas las funciones continuas definidas sobre el intervalo $[a,b]$, entonces con las operaciones usuales entre funciones, X forma un álgebra conmutativa.

Ejemplo. 2.3.4 Sea $X=C^n[a,b]$, el espacio de todas las funciones continuas que tienen derivada continua hasta el orden n inclusive, y que están definidas sobre el intervalo $[a,b]$. Entonces con las operaciones usuales entre funciones, X forma un espacio lineal.

Sea α una función monótonamente creciente sobre $[a,b]$ (puesto que $\alpha(a)$ y $\alpha(b)$ son finitos, se sigue que la función α es acotada sobre $[a,b]$). Si P es una partición de $[a,b]$, esto es, un conjunto finito de puntos $x_0,x_1,\ldots,x_n$, donde $a=x_0\leq x_1\leq\ldots\leq x_n=b$, correspondiendo a cada partición P de $[a,b]$, escribimos $\Delta\alpha_i=\alpha(x_i)-\alpha(x_{i-1})$, $i=1,\ldots,n$. Para cualquier función f la cual es acotada sobre $[a,b]$ poniendo $U(P,f,\alpha)=\sum_{i=1}^{n} M_i\,\Delta\alpha_i$ y $L(P,f,\alpha)=\sum_{i=1}^{n} m_i\,\Delta\alpha_i$ en donde tenemos que $M_i=\sup\{f(x)\ ;\ x_{i-1}\leq x\leq x_i\}$ y $m_i=\inf\{f(x)\ ;\ x_{i-1}\leq x\leq x_i\}$, definimos

$\overline{\int_a^b} f(x)d\alpha(x)=\inf_P U(P,f,\alpha)$ y $\underline{\int_a^b} f(x)d\alpha(x)=\sup_P L(P,f,\alpha)$, en donde el ínfimo y el supremo son tomados sobre todas las particiones posibles de $[a,b]$. Si $\overline{\int_a^b} f(x)d\alpha(x)$ y $\underline{\int_a^b} f(x)d\alpha(x)$ son iguales, denotamos este valor común por $\int_a^b f(x)d\alpha(x)$. Esta es llamada la integral de Stieltjes de la función f con respecto de α, sobre $[a,b]$. Si $\int_a^b f(x)d\alpha(x)$ existe diremos que f es integrable con respecto de α en el sentido de Riemann. Tenemos las siguientes propiedades: Si f,g son integrables con respecto de α en el sentido de Riemann, entonces $\int_a^b(c_1f(x)+c_2g(x))d\alpha(x)=c_1\int_a^b f(x)d\alpha(x)+c_2\int_a^b g(x)d\alpha(x)$; $f(x)\leq g(x)$ sobre $[a,b]$ implica $\int_a^b f(x)d\alpha(x)\leq\int_a^b g(x)d\alpha(x)$.

<u>Ejemplo</u>. 2.3.5 Sea $X=S^p([a,b],d\alpha(x))$, el conjunto consistente de todas las funciones f definidas sobre el intervalo $[a,b]$, tales que la integral de Stieltjes con respecto de α, $(\int_a^b |f(x)|^p d\alpha(x))^{1/p}$ existe. Vemos que si $f,g\in S^p([a,b],d\alpha(x))$, entonces $f+g\in S^p([a,b], d\alpha(x))$. Puesto que $|f(x)+g(x)|^p\leq 2^p(|f(x)|^p+|g(x)|^p)$ y que si $a\in\mathbb{R}$, entonces $af\in S^p([a,b],d\alpha(x))$ siempre que $f\in S^p([a,b],d\alpha(x))$. Por lo tanto X forma un espacio lineal sobre $\mathbb{R}$.

<u>Definición</u>. 2.3.6 Sea X un espacio lineal y sea Y un subconjunto de X, se dice entonces que Y es un subespacio lineal de X si satisface que:

(i). $f,g\in Y$ implica $f+g\in Y$.

(ii). Si $a\in\mathbb{R}$ y $f\in Y$ implica $af\in Y$.

(iii). 0 de X es también un elemento de Y.

Diremos también que el conjunto Y constituye una subálgebra de un álgebra X, si Y es un subespacio lineal de X, y si además, Y es un álgebra como un subconjunto de X.

<u>Ejemplo</u>. 2.3.7 Sea $X=C[a,b]$, el espacio de funciones continuas sobre $[a,b]$. Sea $Y=\{f\in X; f(a)=f(b)=0\}$, entonces con las operaciones usuales el espacio Y constituye una subálgebra de X. Este espacio es conocido como el espacio de funciones continuas sobre $[a,b]$ de soporte compacto, y es denotado por $C_0[a,b]$.

<u>Ejemplo</u>. 2.3.8 Sea $X=C^n[a,b]$ y sea $Y=\{f\in X; f^{(k)}(a)=f^{(k)}(b)=0, k=0,1,\dots,p\}, p\leq n$, entonces con las operaciones usuales entre funciones

Y forma un subespacio lineal de X, y es denotado por $C^n_{0,p}[a,b]$.

Teorema. 2.3.9 Sea X un espacio lineal y sea $\{Y_\alpha\}_{\alpha\in I}$ una familia de subespacios lineales de X, donde I es una familia de índices. Entonces $\bigcap_{\alpha\in I} Y_\alpha$ es un subespacio lineal de X.

Ejemplo. 2.3.10 Sea $Y_n=C^n[a,b]$, entonces si Y es el conjunto de todas las funciones infinitamente diferenciables, tenemos que $Y=\bigcap_{n=1}^{\infty} Y_n$, y Y es un subespacio lineal de $C[a,b]$.

Definición. 2.3.11 Dos espacios lineales X,Y se llaman isomorfos si existe una función biyectiva $F:X \to Y$, tal que, para todo $f,g\in X$ y $a\in\mathbb{R}$ se satisface que:

(i). $F(f+g)=F(f)+F(g)$.

(ii). $F(af)=aF(f)$.

Consideremos a $C[a,b]$ y $F:C[a,b] \to C[a,b]$ definida por $F(f)=f(x)g_0(x)$, $g_0(x)\in C[a,b]$ es una función fija. Entonces F es un isomorfismo del espacio $C[a,b]$ en sí mismo.

Definición. 2.3.12 Sea X un espacio lineal sobre $\mathbb{R}$. Un subconjunto C de X, es llamado convexo si $cf+(1-c)g\in C$ siempre que $f,g\in C$ y $c\in[0,1]$.

Ejemplo. 2.3.13 Sea $X=C[a,b]$, el conjunto de funciones continuas sobre $[a,b]$, y sea $C=\{f \ ; \ |f(x)|\leq 1 \text{ para toda } x\in[a,b]\}$. Este conjunto es convexo.

Proposición. 2.3.14 La intersección arbitraria de conjuntos convexos es de nuevo un conjunto convexo. Efectivamente, sean $f,g\in\bigcap_{\alpha\in I} C_\alpha$, luego si $c\in[0,1]$ tenemos que $cf+(1-c)g$ pertenece a C_α para toda $\alpha\in I$ y, por consiguiente, a $\bigcap_{\alpha\in I} C_\alpha$.

Definición. 2.3.15 Sea Y un subespacio de un espacio lineal X y consideremos la relación de equivalencia entre elementos de X; $(f,g)\in R$ si y sólo si $f-g\in Y$. El conjunto de todas las clases de equivalencia se llamará espacio cociente de X según Y, que frecuentemente se denota por X/Y.

Las operaciones de adición y multiplicación por escalares en el espacio cociente se introducen de manera natural; si $[f]$ y $[g] \in X/Y$, entonces $[f]+[g]=[f+g]$, $a[f]=[af]$, es fácil ver que X/Y es un espacio lineal.

Sea X un espacio lineal sobre $\mathbb{R}$, sean Y y Z subespacios de X. Se define la suma de Y y Z como el subconjunto de X que consta de todas las sumas $f+g$ con $f \in Y$ y $g \in Z$. Se denota esta suma por $Y+Z$, y es claro que este subconjunto es un subespacio lineal de X. Se dice que un espacio lineal X es suma directa de los subespacios Y y Z, y se denota por $Y \oplus Z$, si para todo elemento $f \in X$, existen elementos únicos $g \in Y$ y $h \in Z$ tales que $f=g+h$.

<u>Proposición</u>. 2.3.16 Sea X un espacio lineal sobre R, y sean Y y Z subespacios lineales de X. Si $X=Y+Z$ y $Y \cap Z=\{0\}$. Entonces X es la suma directa de Y y Z.

Supongamos que X y Y son espacios lineales arbitrarios sobre $\mathbb{R}$. Consideremos el producto cartesiano $X \times Y$, si $(f_1,g_1),(f_2,g_2) \in X \times Y$. Definimos la adición de tales elementos y el producto por un escalar $a \in \mathbb{R}$ como sigue: $(f_1,g_1)+(f_2,g_2)=(f_1+f_2,g_1+g_2)$, $a(f_1,g_1)=(af_1,ag_1)$. Obviamente el producto cartesiano de dos espacios lineales sobre un mismo campo, $\mathbb{R}$, es otra vez un espacio lineal sobre $\mathbb{R}$.

En el estudio de la teoría de espacios lineales de dimensión infinita, al igual que en los espacios de dimensión finita, también se establecen los conceptos de combinación lineal, dependencia e independencia lineal entre elementos.

<u>Definición</u>. 2.3.17 Sea X un espacio lineal y $f_1,f_2,\ldots,f_n \in X$. Sean $a_1,a_2,\ldots,a_n \in \mathbb{R}$. Una combinación lineal de $f_1,f_2,\ldots,f_n \in X$ es una expresión del tipo $a_1f_1+a_2f_2+\cdots+a_nf_n$.

Hacemos notar que el conjunto de todas las combinaciones lineales de $f_1,f_2,\ldots,f_n \in X$, X un espacio lineal, es un subespacio de X.

Definición. 2.3.18 Sea $f_1,f_2,\ldots,f_n \in X$, X un espacio lineal. Se dice que estos elementos son linealmente dependientes sobre $\mathbb{R}$ si existen elementos $a_1,a_2,\ldots,a_n \in \mathbb{R}$ no todos iguales a 0, tales que $a_1f_1+a_2f_2+\cdots+a_nf_n=0$. Si no existen tales números, se dice entonces que $f_1,f_2,\ldots,f_n$ son linealmente independientes. En otras palabras, $f_1,f_2,\ldots,f_n \in X$ son linealmente independientes si y sólo si satisfacen la condición: Siempre que $a_1,a_2,\ldots,a_n \in \mathbb{R}$ y son tales que $a_1f_1+a_2f_2+\cdots+a_nf_n=0$, entonces $a_k=0$ para toda $k=1,2,\ldots,n$.

Ejemplo. 2.3.19 Las funciones $e^t,e^{2t},\ldots,e^{nt} \in C[a,b]$ son linealmente independientes.

Proposición. 2.3.20 Sean $f_1,f_2,\ldots,f_n \in X$, X un espacio lineal. Si $f_1,f_2,\ldots,f_n \in X$ son linealmente independientes, entonces dos combinaciones lineales de estos elementos son iguales.

Un conjunto infinito de elementos f,g,h,... de un espacio lineal X se llama linealmente independiente, cuando todo subconjunto finito suyo es linealmente independiente, y si en X se puede determinar un conjunto infinito tal que cualquier subconjunto finito sea linealmente independiente, se dice que X es un espacio de dimensión infinita. Por ejemplo, el espacio $C[a,b]$ de funciones continuas tiene dimensión infinita al considerar el ejemplo 2.3.19.

2.4 ESPACIOS METRICOS.

Definición. 2.4.1 Un espacio métrico es un par (M,d), formado por un conjunto no vacío M y una función $d:M\times M \to \mathbb{R}^+$ llamada métrica, con las siguientes propiedades para todo $f,g,h \in M$:

(i). $d(f,f)=0$.

(ii). $f\neq g$ implica $d(f,g)>0$.

(iii). $d(f,g)=d(g,f)$.

(iv). $d(f,g)\leq d(f,h)+d(h,g)$.

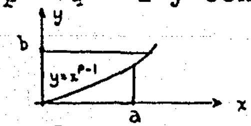
Si la función d satisface sólo (i), (iii) y (iv), entonces d es llamada una pseudométrica.

Ejemplo. 2.4.2 Sea $M=C^n[a,b]$, se define $d_{(n)}(f,g)=\sum_{k=0}^{n}\max_{x\in[a,b]}|f^{(k)}(x)-g^{(k)}(x)|$.
La propiedad (iv) se desprende de la desigualdad $\max_{x\in[a,b]}|f^{(k)}(x)-g^{(k)}(x)|\leq\max_{x\in[a,b]}|f^{(k)}(x)|+\max_{x\in[a,b]}|g^{(k)}(x)|$, $(k=0,\ldots,n)$, al sustituir f-h por f y g-h por g.

Obviamente se puede definir inductivamente una métrica para $C^{n+1}[a,b]$ mediante la métrica $d_{(n)}$ de $C^n[a,b]$, debido a que $C^{n+1}[a,b]\subset C^n[a,b]$. En efecto, sean $f,g\in C^{n+1}[a,b]$, entonces $d_{(n+1)}(f,g)=\max_{x\in[a,b]}|f^{(n+1)}(x)-g^{(n+1)}(x)|+d_{(n)}(f,g)$.

Consideremos ahora el espacio $S^p([a,b],d\alpha(x))$ que ya fue visto con anterioridad, para definir una pseudométrica sobre este espacio es necesario establecer los siguiente teoremas.

Teorema. 2.4.3 (Desigualdad de Hölder) Si p y q son tales que $p^{-1}+q^{-1}=1$ con $1<p,q<\infty$, y si $f\in S^p([a,b],d\alpha(x))$ y $g\in S^q([a,b],d\alpha(x))$. Entonces $fg\in S^1([a,b],d\alpha(x))$ y
$\int_a^b|f(x)g(x)|d\alpha(x)\leq(\int_a^b|f(x)|^pd\alpha(x))^{1/p}(\int_a^b|g(x)|^qd\alpha(x))^{1/q}$.
Demostración: Consideremos en el plano xy la curva dada por $y=x^{p-1}$, entonces $x=y^{q-1}$ si $p^{-1}+q^{-1}=1$ y consideremos la siguiente gráfica con $a,b\geq0$.



Entonces si $\int_0^a x^{p-1}dx$ y $\int_0^b y^{q-1}dy$ representan las areas de la función y sobre el eje x y de la función x sobre el eje y, entonces para todo $a,b\geq0$, se tiene $\int_0^a x^{p-1}dx+\int_0^b y^{q-1}dy\geq ab$, en donde $p^{-1}+q^{-1}=1$ y $1<p,q<\infty$, ahora bien, esto equivale a que $(a^p)/p+(b^q)/q\geq ab$, sustituyendo c por p^{-1}, a^c por a y b^{1-c} por b tenemos $ca+(1-c)b\geq a^cb^{(1-c)}$, $0<c<1$.
Supongamos $(\int_a^b|f(x)|^pd\alpha(x))^{1/p}=A_p=1$ y $(\int_a^b|g(x)|^qd\alpha(x))^{1/q}=A_q=1$, y si $a=|f(x)|^p$, $b=|g(x)|^q$, obtenemos $c|f(x)|^p+(1-c)|g(x)|^q\geq|f(x)g(x)|$, e integrando tenemos que $\int_a^b|f(x)g(x)|d\alpha(x)\leq cA_p+(1-c)A_q=1$, después al sustituir $f(x)/A_p$ por f(x) y $g(x)/A_q$ por g(x) concluimos la demostración.

Teorema. 2.4.4 (Desigualdad de Minkowski) Si $f,g\in S^p([a,b],d\alpha(x))$. Entonces $f+g\in S^p([a,b],d\alpha(x))$ y

$$\left(\int_a^b |f(x)+g(x)|^p d\alpha(x)\right)^{1/p} \leq \left(\int_a^b |f(x)|^p d\alpha(x)\right)^{1/p} + \left(\int_a^b |g(x)|^p d\alpha(x)\right)^{1/p}.$$

Demostración: Puesto que $|f(x)+g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p)$ es válida para todo $p>1$, tenemos que $f+g\in S^p([a,b],d\alpha(x))$ siempre que $f,g\in S^p([a,b],d\alpha(x))$.

Por otra parte $|f(x)+g(x)|^p \leq |f(x)+g(x)|^{p-1}|f(x)| + |f(x)+g(x)|^{p-1}|g(x)|$, e integrando obtenemos:

$$\int_a^b |f(x)+g(x)|^p d\alpha(x) \leq \int_a^b |f(x)+g(x)|^{p-1}|f(x)| d\alpha(x) + \int_a^b |f(x)+g(x)|^{p-1}|g(x)| d\alpha(x).$$

Aplicando ahora la desigualdad de Hölder a los sumandos del lado derecho obtenemos:

$$\int_a^b |f(x)+g(x)|^{p-1}|f(x)| d\alpha(x) \leq \left(\int_a^b |f(x)+g(x)|^p d\alpha(x)\right)^{1/q}\left(\int_a^b |f(x)|^p d\alpha(x)\right)^{1/p}$$

$$\int_a^b |f(x)+g(x)|^{p-1}|g(x)| d\alpha(x) \leq \left(\int_a^b |f(x)+g(x)|^p d\alpha(x)\right)^{1/q}\left(\int_a^b |g(x)|^p d\alpha(x)\right)^{1/p},$$

puesto que $p=(p-1)q$ siempre que $p^{-1}+q^{-1}=1$. Entonces

$$\int_a^b |f(x)+g(x)|^p d\alpha(x) \leq \left(\int_a^b |f(x)+g(x)|^p d\alpha(x)\right)^{1/q}\left(\left(\int_a^b |f(x)|^p d\alpha(x)\right)^{1/p} + \left(\int_a^b |g(x)|^p d\alpha(x)\right)^{1/p}\right),$$

con lo que concluimos la demostración.

Si $f,g\in S^p([a,b],d\alpha(x))$ se define una pseudométrica sobre este espacio como sigue:

$$d(f,g)=\left(\int_a^b |f(x)-g(x)|^p d\alpha(x)\right)^{1/p}.$$

La propiedad (iv) se desprende del teorema 2.4.4 sustituyendo $f-h$ por f y $g-h$ por g.

Proposición. 2.4.5 Si (M,d) es un espacio métrico con $d:M\times M \to \mathbb{R}^+$, entonces se puede construir a partir de la métrica d, una métrica d' para el espacio M tal que $d':M\times M \to [0,1)$.

La demostración es obvia a partir de la siguiente definición de d'

$d'(f,g)=\dfrac{d(f,g)}{1+d(f,g)}$. Notamos que $0\leq d'(f,g)<1$ para todo $f,g\in M$.

Ahora trataremos sobre la construcción de métricas para el producto cartesiano de una colección finita de espacios métricos.

Dada una colección finita de espacios métricos $\{(M_k,d_k)\}_{k=1}^n$ se pueden construir métricas para el producto cartesiano $M=\bigtimes_k M_k$, a

partir de la colección de métricas $\{d_k\}_{k=1}^{n}$ como sigue: Si $f=(f_1,\ldots,f_n)$, $g=(g_1,\ldots,g_n)\in M$, una manera natural de construir una métrica d para M, es que: $d(f,g)=\sum_{k=1}^{n} d_k(f_k,g_k)$.
Ahora bien, si consideramos a $(d_1(f_1,g_1),\ldots,d_n(f_n,g_n))$ como un punto de un espacio euclídeo, entonces definimos $d(f,g)$ como la distancia del punto al origen, esto es: $d(f,g)=\left[\sum_{k=1}^{n} d_k(f_k,g_k)^2\right]^{1/2}$.

Definición. 2.4.6 Dos métricas d_1 y d_2 de un espacio métrico M, son llamadas equivalentes, si existen dos constantes $a,b>0$, tales que $ad_1(f,g)\leq d_2(f,g)\leq bd_1(f,g)$.

En lo que sigue emplearemos la notación $\max_{x,k}\{|f^{(k)}(x)|\}$ en lugar de $\max_{\substack{x\in[a,b]\\ 0\leq k\leq n}}\{|f^{(k)}(x)|\}$ cuando $f\in C^n[a,b]$.

Ejemplo. 2.4.7 Sea $M=C^n[a,b]$, sean $d_1(f,g)=\max_{x,k}\{|f^{(k)}(x)-g^{(k)}(x)|\}$ y $d_2(f,g)=d_{(n)}(f,g)=\sum_{k=1}^{n}\max_{x\in[a,b]}|f^{(k)}(x)-g^{(k)}(x)|$, con n fija, entonces es claro que: $1d_1(f,g)\leq d_2(f,g)\leq nd_1(f,g)$.

Definición. 2.4.8 Una métrica d sobre un espacio métrico lineal M, es llamada invariante bajo traslaciones, si $d(f+h,g+h)=d(f,g)$ para todo $f,g,h\in M$.

Proposición. 2.4.9 Si d es una métrica invariante bajo traslaciones sobre un espacio métrico lineal M. Entonces $d(nf,0)\leq nd(f,0)$ para todo $f\in M$ y para toda $n=1,2,\ldots$
La demostración se sigue de la desigualdad
$d(nf,0)\leq\sum_{k=1}^{n} d(kf,(k-1)f)=nd(f,0)$.

Una sucesión infinita de elementos de un espacio M, es una función del conjunto de números naturales en M.
A partir de esta definición, una sucesión infinita será designada por una letra como f, y los valores particulares por: $f(1),f(2),\ldots$ Sin embargo la notación con subíndices es usada con mayor frecuencia, esto es, $f_1,f_2,\ldots$; La sucesión como conjunto es usualmente denotada por $\{f_n\}_{n=1}^{\infty}$.

Definición. 2.4.10 Una sucesión de elementos $\{f_n\}_{n=1}^{\infty}$ de un espacio métrico (M,d), se dice que converge a un elemento $f\in M$, si $d(f_n,f)\to 0$ cuando $n\to\infty$, o bien que $\lim_{n\to\infty} d(f_n,f)=0$. Frecuentemente este hecho se denota por $f_n\overset{d}{\to} f$ cuando $n\to\infty$ o por $\lim_{n\to\infty} f_n=f$.

Definición. 2.4.11 Una sucesión de elementos $\{f_n\}_{n=1}^{\infty}$ de un espacio métrico (M,d), es llamada una sucesión de Cauchy si para todo $\epsilon>0$, existe un N tal que para $n,m>N$ implica $d(f_n,f_m)<\epsilon$. Es usual denotar este hecho por $d(f_n,f_m) \to 0$ cuando $n,m \to \infty$.

Proposición. 2.4.12 El límite de una sucesión convergente en un espacio métrico (M,d) es único. Dem. Si $d(f_n,g) \to 0$ y $d(f_n,h) \to 0$ cuando $n \to \infty$, entonces $d(g,h) \leq d(g,f_n)+d(f_n,h) \to 0$ cuando $n \to \infty$. así $g=h$.

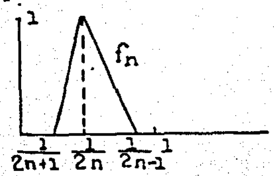
Proposición. 2.4.13 Cualquier sucesión convergente en un espacio métrico (M,d) es de Cauchy. Dem. Si $d(f_n,f) \to 0$ y $d(f_m,f) \to 0$ cuando $n,m \to \infty$, entonces $d(f_n,f_m) \leq d(f_n,f)+d(f,f_m) \to 0$ cuando $n,m \to \infty$.

El recíproco de la proposición anterior no es cierto. Sea $M=\mathbb{Q}$ y $d(x,y)=|x-y|$, sea $x^* \in \mathbb{I}=\mathbb{R}/\mathbb{Q}$ y $\{x_n\}_{n=1}^{\infty} \subset \mathbb{Q}$ tal que $x_n \to x^*$, entonces $\{x_n\}_{n=1}^{\infty}$ es de Cauchy en $\mathbb{Q}$ pero no converge a algún y en $\mathbb{Q}$.
Puesto que si $x_n \to y$ en $\mathbb{Q}$ entonces $x_n \to y$ en $\mathbb{R}$, entonces $x^*=y$.
Por contrapuesta de la proposición 2.4.13 obtenemos la siguiente:

Proposición. 2.4.14 Si en un espacio métrico (M,d) una sucesión $\{f_n\}_{n=1}$ no es de Cauchy, entonces la sucesión no es convergente en (M,d).

Ejemplo. 2.4.15 Sea $M=C[0,1]$, sea $d(f,g)=\max_{x\in[a,b]}|f(x)-g(x)|$ y consideremos la sucesión de funciones definida por:

$$f_n(x)=\begin{cases} 0 \text{ si } 0\leq x\leq 1/(2n+1). \\ 2n(2n+1)x-2n \text{ si } 1/(2n+1)\leq x\leq 1/(2n). \\ -2n(2n-1)x+2n \text{ si } 1/(2n)\leq x\leq 1/(2n-1). \\ 0 \text{ si } 1/(2n-1)\leq x\leq 1. \end{cases}$$



Es fácil ver que si $n\neq m$, entonces $d(f_n,f_m)=1$, por lo que $\{f_n\}_{n=1}^{\infty}$ no es una sucesión de Cauchy con esta métrica: por lo tanto esta sucesión no converge en el espacio.
Sin embargo $\{f_n\}_{n=1}^{\infty}$ es una sucesión convergente en la métrica $d(f,g)=\int_0^1 |f(x)-g(x)|\, dx$.

Definición. 2.4.16 Un espacio métrico en el cual toda sucesión de Cauchy converge es llamado un espacio métrico completo.

Teorema. 2.4.17 Sea $M=C[a,b]$, con la métrica $d(f,g)=\max_{x\in[a,b]}|f(x)-g(x)|$. Entonces (M,d) es un espacio métrico completo.

Demostración: Sea $\{f_n\}_{n=1}^{\infty}\subset M$ una sucesión de Cauchy. Entonces, para cualquier x fijo en $[a,b]$, $|f_n(x)-f_m(x)|\leq d(f_n,f_m)\to 0$ cuando $n,m\to\infty$, así $\{f_n(x)\}_{n=1}^{\infty}$ es una sucesión de Cauchy de números reales. Puesto que los reales son completos, para cada x existe un número, $f(x)$, con $f_n(x)\to f(x)$ cuando $n\to\infty$. Dado $\epsilon>0$, encontramos N, tal que $n,m>N$ implica $d(f_n,f_m)\leq\epsilon$. Entonces

$|f(x)-f_N(x)|\leq \max_{x\in[a,b]}\lim_{n\to\infty}|f_n(x)-f_N(x)|\leq \max_{x\in[a,b]}\sup_{n\geq N}|f_n(x)-f_N(x)|=$

$\sup_{n\geq N} d(f_n,f_m)\leq\epsilon$.

Así, si podemos demostrar que $f\in C[a,b]$, entonces podemos concluir que $d(f_n,f)\to 0$ cuando $f_n\to f$ en $C[a,b]$. Para esto vamos a demostrar que " el límite uniforme de funciones continuas es una función continua". Fijamos $x\in[a,b]$ y $\epsilon>0$. Queremos encontrar $\delta>0$ tal que $|x-y|<\delta$ implica $|f(x)-f(y)|<\epsilon$. Escojemos n tal que $d(f_n,f)<\epsilon/3$, puesto que f_n es continua, escojemos $\delta>0$ tal que $|x-y|<\delta$ implica $|f_n(x)-f_n(y)|<\epsilon/3$. Entonces $|x-y|<\delta$ implica que

$|f(x)-f(y)|\leq|f(x)-f_n(x)|+|f_n(x)-f_n(y)|+|f_n(y)-f(y)|<\epsilon/3+\epsilon/3+\epsilon/3=\epsilon$.

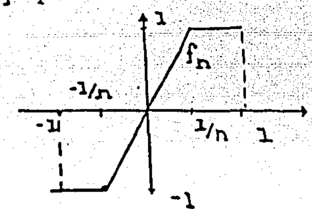
Así f es continua.

Corolario. 2.4.18 Sea $M=C^n[a,b]$, con la métrica $d(f,g)=\max_{x,k}\{|f^{(k)}(x)-g^{(k)}(x)|\}$. Entonces (M,d) es un espacio métrico completo.

La demostración se sigue de que $f\in C^n[a,b]$ si, y sólo si $f^{(k)}\in C[a,b]$ para cada k $(k=0,\ldots n)$.

Ejemplo. 2.4.19 Sea $M=C[-1,1]$ con la métrica $d(f,g)=\left(\int_{-1}^{1}|f(x)-g(x)|^2dx\right)^{1/2}$. Este espacio es llamado, el espacio de funciones continuas sobre $[-1,1]$ con métrica cuadrática. Entonces (M,d) no es un espacio métrico completo. En efecto, consideremos la sucesión de funciones definidas en el intervalo $[-1,1]$ por:

$$f_n(x)=\begin{cases}-1 & \text{si } -1\leq x\leq -1/n.\\ nx & \text{si } -1/n\leq x\leq 1/n.\\ 1 & \text{si } 1/n\leq x\leq 1.\end{cases}$$



Esta sucesión es de Cauchy, puesto que
$\int_{-1}^{1}|f_n(x)-f_m(x)|^2dx < 4/\min(m,n)$, entonces
$d(f_n,f_m)=(\int_{-1}^{1}|f_n(x)-f_m(x)|^2dx)^{1/2} < 2/(\min(m,n))^{1/2}$.
Sin embargo esta sucesión no converge a ninguna función continua sobre $[-1,1]$.

<u>Definición</u>. 2.4.20 Una funcion F de un espacio metrico (M_1,d_1) a otro, (M_2,d_2), es llamada continua en el punto f, si $F(f_n) \xrightarrow{d_2} F(f)$ cuando $f_n \xrightarrow{d_1} f$.

<u>Ejemplo</u>. 2.4.21 Sea $M_1=M_2=C[0,1]$ y $d_1(f,g)=\max_{x\in[0,1]}|f(x)-g(x)|$, $d_2(f,g)=\int_0^1|f(x)-g(x)|dx$. De esta manera la función identidad de (M_1,d_1) a (M_2,d_2) es continua, puesto que $\int_0^1|f(x)-g(x)|dx \leq \max_{x\in[0,1]}|f(x)-g(x)|$.

<u>Definición</u>. 2.4.22 Una función F biyectiva de (M_1,d_1) a (M_2,d_2) la cual preserva métrica, esto es, $d_2(F(f),F(g))=d_1(f,g)$, es llamada una isometría, y los espacios (M_1,d_1), (M_2,d_2) son llamados isométricos. Obviamente una isometría F es una función continua.

<u>Ejemplo</u>. 2.4.23 Sea (M,d) un espacio métrico lineal, tal que d es invariante bajo traslaciones. Así, si $F:M \to M$; $F(f)=f+h$ con h fijo, entonces $d(F(f),F(g))=d(f+h,g+h)=d(f,g)$, entonces la función F es una isometría del espacio en sí mismo.

<u>Definición</u>. 2.4.24 Sea (M,d) un espacio métrico. Un subconjunto Y de M es llamado denso, si todo $f\in M$ es límite de elementos de Y.

<u>Ejemplo</u>. 2.4.25 El conjunto de polinomios con coeficientes racionales definidos en el intervalo $[a,b]$; $P_Q[a,b]$ es denso en $C[a,b]$, en la métrica $d(f,g)=\max_{x\in[a,b]}|f(x)-g(x)|$.

<u>Definición</u>. 2.4.26 Sea (M,d) un espacio métrico. Un espacio métrico (M^*,d^*) se llama completación del espacio métrico (M,d), si:
(i). (M,d) es un subespacio de (M^*,d^*).
(ii). El conjunto M es denso en M^*.

Ejemplo. 2.4.27 De acuerdo con el ejemplo 2.4.25. El espacio $C[a,b]$ es la completación del espacio de polinomios con coeficientes racionales $P_{Q}[a,b]$, en la métrica $d(f,g)=\max_{x\in[a,b]}|f(x)-g(x)|$.

Teorema. 2.4.28 Todo espacio métrico (M,d) posee una completación, y esta completación es única salvo isometrías que transforman los puntos de M en sí mismos.

Demostración:

Unicidad: Veamos que existe una isometría $F:(M^*,d^*) \to (M^{**},d^{**})$ donde (M^*,d^*) y (M^{**},d^{**}) son completaciones del espacio (M,d), tales que:

(i). $F(f)=f$ para todo $f\in M$, así $d^*(f,g)=d^{**}(f,F(g))$.

(ii). Si $F(f^*)=f^{**}$ y $F(g^*)=g^{**}$, entonces $d^*(f^*,g^*)=d^{**}(F(f^*),F(g^*))=d^{**}(f^{**},g^{**})$.

Dem.(i). Sea $f^*\in M$, puesto que M es denso en M^*, existe una sucesión $\{f_n\}_{n=1}^{\infty}\subset M$, tal que $f_n \to f^*$ en la métrica d^* cuando $n\to\infty$, además puesto que $\{f_n\}_{n=1}^{\infty}\subset M\subset M^{**}$ y dado que M^{**} es también completo, $\{f_n\}_{n=1}^{\infty}$ converge en M^{**} a un punto f^{**} en la métrica d^{**}. Es obvio que f^{**} no depende de la elección de la sucesión $\{f_n\}_{n=1}^{\infty}$. Sea $F(f^*)=f^{**}$ y de esta manera es claro que $F(f)=f$ para todo $f\in M$, puesto que si $\{f_n\}_{n=1}^{\infty}$ converge a un $f\in M$, entonces $f=f^*=f^{**}$.

Dem.(ii). Sea $f_n \to f^*$ y $g_n \to g^*$ cuando $n\to\infty$ en la métrica d^* sobre el espacio M^*, y $f_n \to f^{**}$, $g_n \to g^{**}$ cuando $n\to\infty$ en la métrica d^{**} sobre el espacio M^{**}. Entonces

$d^*(f_n,f^*) \to 0$, $d^*(g_n,g^*) \to 0$ cuando $n\to\infty$, de esta manera

$$|d^*(f_n,g_n)-d^*(f^*,g^*)|\le|d^*(f_n,g_n)-d^*(g_n,f^*)|+|d^*(g_n,f^*)-d^*(f^*,g^*)|$$

$$\le d^*(f_n,f^*)+d^*(g_n,g^*) \to 0 \text{ cuando } n\to\infty.$$

De donde $\lim_{n\to\infty} d^*(f_n,g_n)=d^*(f^*,g^*)$.

De manera análoga se prueba que en el espacio M^{**}, $\lim_{n\to\infty} d^{**}(f_n,g_n)=d^{**}(f^{**},g^{**})$. Entonces

$$d^*(f^*,g^*)=\lim_{n\to\infty} d^*(f_n,g_n)=\lim_{n\to\infty} d(f_n,g_n)=\lim_{n\to\infty} d^{**}(f_n,g_n)=d^{**}(f^{**},g^{**}).$$

Sea (M,d) un espacio métrico y sea X el conjunto de sucesiones de Cauchy en (M,d), definimos una relación de equivalencia en X, como sigue: $(\{f_n\}_{n=1},\{f'_n\}_{n=1})\in R$ si y sólo si $\lim_{n\to\infty} d(f_n,f'_n)=0$.

Definimos el espacio (M^*,d^*) como el espacio de todas las clases de equivalencias, definiendo la métrica d^* como sigue:

Sean $[\{f_n\}_{n=1}^{\infty}]$ y $[\{g_n\}_{n=1}^{\infty}]$ dos clases de equivalencia, sean $\{f_n\}_{n=1}^{\infty} \in [\{f_n\}_{n=1}^{\infty}]$ y $\{g_n\}_{n=1}^{\infty} \in [\{g_n\}_{n=1}^{\infty}]$. Pongamos $d^*([\{f_n\}_{n=1}^{\infty}], [\{g_n\}_{n=1}^{\infty}]) = \lim_{n\to\infty} d(f_n,g_n)$.

Para demostrar que la función d^* esta bien definida, debemos ver:

(iii). $\lim_{n\to\infty} d(f_n,g_n)$ existe.

(iv). d^* no depende de la elección de $\{f_n\}_{n=1}^{\infty} \in [\{f_n\}_{n=1}^{\infty}]$ y de $\{g_n\}_{n=1}^{\infty} \in [\{g_n\}_{n=1}^{\infty}]$, esto es, si $\{f_n\}_{n=1}^{\infty}$, $\{f'_n\}_{n=1}^{\infty} \in [\{f_n\}_{n=1}^{\infty}]$ y $\{g_n\}_{n=1}^{\infty}, \{g'_n\}_{n=1}^{\infty} \in [\{g_n\}_{n=1}^{\infty}]$, entonces $\lim_{n\to\infty} d(f_n,g_n) = \lim_{n\to\infty} d(f'_n,g'_n)$.

Dem.(iii). Puesto que la sucesiones $\{f_n\}_{n=1}^{\infty}$ y $\{g_n\}_{n=1}^{\infty}$ son de Cauchy y

$$\begin{aligned}|d(f_n,g_n)-d(f_m,g_m)| &\le |d(f_n,g_n)-d(g_n,f_m)| + |d(g_n,f_m)-d(f_m,g_m)| \\ &\le d(f_n,f_m)+d(g_n,g_m) \to 0 \text{ cuando } n,m \to \infty.\end{aligned}$$

De donde la sucesión $\{d(f_n,g_n)\}_{n=1}^{\infty}$ de números reales es de Cauchy, y por lo tanto tiene un límite.

Dem.(iv). Sean $\{f_n\}_{n=1}^{\infty}, \{f'_n\}_{n=1}^{\infty} \in [\{f_n\}_{n=1}^{\infty}]$ y $\{g_n\}_{n=1}^{\infty}, \{g'_n\}_{n=1}^{\infty} \in [\{g_n\}_{n=1}^{\infty}]$ y como

$$\begin{aligned}|d(f_n,g_n)-d(f'_n,g'_n)| &\le |d(f_n,g_n)-d(g_n,f'_n)| + |d(g_n,f'_n)-d(f'_n,g'_n)| \\ &\le d(f_n,f'_n)+d(g_n,g'_n) \to 0 \text{ cuando } n \to \infty.\end{aligned}$$

Puesto que $(\{f_n\}_{n=1}^{\infty}, \{f'_n\}_{n=1}^{\infty}) \in R$ y $(\{g_n\}_{n=1}^{\infty}, \{g'_n\}_{n=1}^{\infty}) \in R$, así $\lim_{n\to\infty} d(f_n,g_n) = \lim_{n\to\infty} d(f'_n,g'_n)$.

Definimos una isometría $F:(M,d) \to (M^*,d^*)$ si a cada punto $f \in M$ le hacemos corresponder la clase de equivalencia de sucesiones de Cauchy convergentes a f, de acuerdo a la relación de equivalencia antes definida. Puesto que si $f_n \to f$ y $f'_n \to f$ en la métrica d,

$$\begin{aligned}|d(f_n,f'_n)| &= |d(f_n,f'_n)-d(f'_n,f)+d(f'_n,f)| \\ &\le |d(f_n,f'_n)-d(f'_n,f)| + |d(f'_n,f)| \\ &\le d(f_n,f)+d(f'_n,f) \to 0 \text{ cuando } n \to \infty.\end{aligned}$$

Así $F(f) = [\{f_n\}_{n=1}^{\infty}]$ y $d(f,g) = d^*(F(f),F(g)) = d^*([\{f_n\}_{n=1}^{\infty}], [\{g_n\}_{n=1}^{\infty}])$.

En lo que sigue identificamos a (M,d) con su imagen $F((M,d))$ y por lo tanto como un subconjunto de (M^*,d^*) y a d^* como una extensión de d, esto es, $d=d^*$ en (M,d). Finalmente demostraremos que el espacio (M^*,d^*) es completación del espacio (M,d), para esto es necesario demostrar que:

(v). M es denso en M^*.
(vi). El espacio (M^*,d^*) es completo.
Dem.(v). Sea $\{f_n\}_{n=1}^{\infty}\in f^*=[\{f_n\}_{n=1}^{\infty}]\in(M^*,d^*)$. Puesto que $\{f_n\}_{n=1}^{\infty}$ es de Cauchy $d^*(f_n,f_m)\to 0$ cuando $n,m\to\infty$. En donde hemos identificado a f_n y f_m con las clases de equivalencia de sucesiones de Cauchy convergentes a f_n y f_m respectivamente para cada n y m, así tendremos que $d^*(f_n,f^*)=\lim_{m\to\infty} d^*(f_n,f_m)\to 0$ cuando $n\to\infty$.
Dem.(vi). Observamos que el espacio (M^*,d^*) ha sido construido de tal manera que toda sucesión de Cauchy $\{f_n\}_{n=1}^{\infty}\subset(M,d)$ converge en (M^*,d^*) a un punto, a saber, al punto $[\{f_n\}_{n=1}^{\infty}]$ determinado por la sucesión $\{f_n\}_{n=1}^{\infty}$. Además, puesto que (M,d) es denso en (M^*,d^*), para toda sucesión de Cauchy $\{f_n^*\}_{n=1}^{\infty}\subset(M^*,d^*)$ es posible construir una sucesión equivalente $\{f_n\}_{n=1}^{\infty}\subset(M,d)$ de Cauchy, poniendo $d^*(f_n^*,f_n)<1/n$. Entonces $\{f_n\}_{n=1}^{\infty}$ converge a un punto $f^*\in(M^*,d^*)$; por lo tanto $d^*(f_n^*,f_n)\to 0$ y $d^*(f_n,f^*)\to 0$ cuando $n\to\infty$, esto implica que $d^*(f_n^*,f^*)\leq d^*(f_n^*,f_n)+d^*(f_n,f^*)\to 0$ cuando $n\to\infty$, lo que concluye la demostración.

<u>Definición</u>. 2.4.29 Sea (M,d) un espacio métrico:
(i). El conjunto $\{f\ ;\ f\in M,\ d(f,g)<r\}$ es llamado la bola abierta $B(f,r)$, de radio r y centro en el punto g.
(ii). Un conjunto $O\subset M$ es llamado abierto, si para todo $f\in O$ existe $r>0$ tal que $B(f,r)\subset O$.
(iii). Un conjunto $N\subset M$ es llamado una vecindad de $f\in N$, si $B(f,r)\subset N$ para algún $r>0$.
(iv). Sea $E\subset M$. Un punto f es llamado un punto límite de E, si para todo $r>0$ $B(f,r)\cap(E-\{f\})\neq\emptyset$, esto es, f es un punto límite de E, si E contiene puntos distintos de f arbitrariamente cercanos a f.
(v). Un conjunto $F\subset M$ es llamado cerrado, si F contiene todos sus puntos límite.
(vi). Si $G\subset M$, $f\in G$ es llamado un punto interior de G, si G es una vecindad de f.

Es importante mencionar los siguientes resultados del análisis elemental, puesto que de ellos será tomada directamente una familia completa de definiciones, para el estudio de espacios topológicos.

<u>Teorema</u>. 2.4.30 Sea (M,d) un espacio métrico:
(i). Un conjunto O es abierto si y sólo si $M-O$ es cerrado.
(ii). Toda vecindad N es un conjunto abierto.
(iii). Si f es un punto límite de un conjunto E, entonces toda vecindad de f contiene infinidad de puntos de E.
(iv). Para cualquier colección $\{O_\alpha\}_{\alpha\in I}$ de conjuntos abiertos, se tiene que $\bigcup_{\alpha\in I} O_\alpha$ es abierto.
(v). Para cualquier colección $\{F_\alpha\}_{\alpha\in I}$ de conjuntos cerrados, se tiene que $\bigcap_{\alpha\in I} F_\alpha$ es cerrado.
(vi). Para cualquier colección finita $\{G_1,\ldots,G_n\}$ de conjuntos abiertos se tiene que $\bigcap_{k=1}^{n} G_k$ es abierto.
(vii). Para cualquier colección finita $\{F_1,\ldots,F_n\}$ de conjuntos cerrados, se tiene que $\bigcup_{k=1}^{n} F_k$ es cerrado.

Demostración: (i). Primero supongamos que $M-O$ es cerrado. Sea $f\in O$, entonces $f\notin M-O$, y f no es punto límite de $M-O$. Por lo tanto existe una vecindad N de f tal que $(M-O)\cap N=\emptyset$, esto es, $f\in N\subset O$, así f es un punto interior de O, y dado que $f\in O$ es arbitrario, tenemos que O es abierto.

Ahora supongamos que O es abierto, y sea f un punto límite de $N-O$, entonces toda vecindad de f contiene puntos de $N-O$, así que f no es un punto interior de O, y puesto que O es abierto $f\in N-O$ y, por consiguiente como f es arbitrario, $N-O$ contiene todos sus puntos límite y por lo tanto $N-O$ es cerrado.

(ii). Sea $N=\{g \ ; \ d(f,g)<r_1\}$ una vecindad de f, y sea $g\in N$. Entonces existe $r_2>0$ tal que $d(f,g)=r_1-r_2$. Para todos los puntos h tales que $d(g,h)<r_2$, tenemos que $d(f,h)\leq d(f,g)+d(g,h)<r_1-r_2+r_2=r_1$, así que $h\in N$, y por lo tanto g es un punto interior de N, puesto que $g\in\{h \ ; \ d(g,h)<r_2\}\subset N$.

(iii). Supongamos que existe una vecindad N de f la cual contiene sólo un número finito de puntos de E distintos de f. Sean $g_1,\ldots,g_n \in N\cap E$ dichos puntos, y pongamos $r=\min_{1\leq k\leq n}\{d(f,g_k)\}>0$. Entonces el conjunto $\{g \ ; \ d(f,g)<r\}$ no contiene puntos de E diferentes de f, de donde f no es un punto límite de E. Esta contradicción concluye la demostración.

(iv). Sea $f\in\bigcup_{\alpha\in I}O_\alpha$, entonces para algún $\alpha\in I, f\in O_\alpha$ con O_α abierto, luego existe $r>0$ tal que $f\in B(f,r)\subset O_\alpha\subset\bigcup_{\alpha\in I}O_\alpha$, así $\bigcup_{\alpha\in I}O_\alpha$ es un conjunto abierto, puesto que todos sus puntos son puntos interiores.
(v). La demostración se sigue de que $M-\bigcup_{\alpha\in I}O_\alpha=\bigcap_{\alpha\in I}(M-O_\alpha)$ y del inciso (i).
(vi). Sea $f\in\bigcap_{k=1}^{n}G_k$, entonces $f\in G_k$ para $k=1,\ldots,n$, y dado que cada G_k es un conjunto abierto, existen $r_1,r_2,\ldots,r_k$ tales que $f\in B(f,r_k)$ para $k=1,\ldots,n$, así $f\in B(f,\min_{1\le k\le n}\{r_k\})\subset\bigcap_{k=1}^{n}G_k$.
(vii). La demostracion se sigue de que $M-\bigcap_{k=1}^{n}G_k=\bigcup_{k=1}^{n}(M-G_k)$ y del inciso (i).

A través de la definición 2.4.20 se había definido continuidad de una función F de un espacio métrico (M_1,d_1) a un espacio métrico (M_2,d_2) en términos de sucesiones, debido a que no se habían introducido los conceptos de conjuntos abiertos y cerrados en un espacio métrico (M,d). Ahora daremos una definición equivalente en términos de distancias, que después traduciremos a términos de bolas. Una función $F:(M_1,d_1)\to(M_2,d_2)$ es llamada continua en el punto $f\in M_1$, si para todo $\epsilon>0$, existe una $\delta>0$ tal que $d_2(F(f),F(g))<\epsilon$ siempre que $d_1(f,g)<\delta$. Si F es continua en todo punto de un subconjunto A de M_1, decimos entonces que F es continua en A. Esta definición refleja la idea intuitiva de que puntos cercanos de f son mapeados en puntos cercanos de $F(f)$, esto puede ser establecido en términos de bolas. Una función F es continua en el punto $f\in M_1$ si y sólo si para toda $\epsilon>0$, existe $\delta>0$ tal que $F(B_1(f,\delta))\subset B_2(F(f),\epsilon)$, donde $B_1(f,\delta)$ y $B_2(F(f),\epsilon)$ son bolas abiertas en M_1 y M_2, respectivamente.

<u>Teorema</u>. 2.4.31 Una función F de un espacio métrico (M_1,d_1) a un espacio métrico (M_2,d_2) es continua en M_1 si y sólo si para todo conjunto abierto $O\subset M_2$, $F^{-1}(O)\subset M_1$ es abierto, esto es, la imagen inversa de cualquier conjunto abierto $O\subset M_2$ es un conjunto abierto.

Demostración: Supongamos que F es continua sobre M_1. Sea O un conjunto abierto en M_2 y sea f un punto arbitrario de $F^{-1}(O)$, probaremos que f es un punto interior de $F^{-1}(O)$. Sea $g=F(f)$, puesto que O es abierto tenemos que $B_2(g,\epsilon)\subset O$ para algún $\epsilon>0$. Ya que F es continua en f, existe un $\delta>0$ tal que $F(B_1(f,\delta))\subset B_2(g,\epsilon)\subset O$, por lo tanto $f\in B_1(f,\delta)\subset F^{-1}(F(B_1(f,\delta)))\subset F^{-1}(B_2(g,\epsilon))\subset F^{-1}(O)$, de donde f es un punto interior de $F^{-1}(O)$, por lo tanto $F^{-1}(O)$ es abierto. Inversamente, supongamos que $F^{-1}(O)$ es abierto para todo abierto $O\subset M_2$. Sea $f\in M_1$ y sea $g=F(f)$. Probaremos que F es continua en el punto f. Para todo $\epsilon>0$ la bola $B_2(g,\epsilon)$ es abierta en M_2, así $F^{-1}(B_2(g,\epsilon))$ es abierta en M_1. Ahora bien, $f\in F^{-1}(B_2(g,\epsilon))$ así existe un $\delta>0$ tal que $B_1(f,\delta)\subset F^{-1}(B_2(g,\epsilon))$ de donde $F(B_1(f,\delta))\subset B_2(g,\epsilon)$, y dado que $f\in M_1$ es un punto arbitrario, tenemos que F es continua en M_1.

<u>Corolario</u>. 2.4.32 Una función F de un espacio métrico (M_1,d_1) a un espacio métrico (M_2,d_2) es continua en M_1 si y sólo si para todo conjunto cerrado $C\subset M_2$, $F^{-1}(C)$ es un conjunto cerrado.
La demostración se sigue del teorema anterior, puesto que un conjunto es cerrado si y sólo si su complemento es abierto y $F^{-1}(E^c)=(F^{-1}(E))^c$ para todo $E\subset M_2$.

<u>Definición</u>. 2.4.33 Por una cubierta abierta de un conjunto E de un espacio métrico (M,d), entenderemos una colección $\{O_\alpha\}_{\alpha\in I}$ de subconjuntos abiertos de (M,d) tales que $E\subset\bigcup_{\alpha\in I}O_\alpha$.

<u>Definición</u>. 2.4.34 Un subconjunto E de un espacio métrico (M,d) es llamado compacto, si toda cubierta abierta de E contiene una subcubierta finita. Más explícitamente, si $\{G_\alpha\}_{\alpha\in I}$ es una cubierta abierta de E, $E\subset\bigcup_{\alpha\in I}G_\alpha$, entonces existe un número finito de índices $\alpha_1,\ldots,\alpha_n$ tales que $E\subset G_{\alpha_1}\cup\cdots\cup G_{\alpha_n}$.

<u>Teorema</u>. 2.4.35
(i). Subconjuntos compactos de espacios métricos son cerrados.
(ii). Subconjuntos cerrados de conjuntos compactos son compactos.
La demostración de este teorema es esencialmente la misma de la proposición 2.5.29 de la sección siguiente.

2.5 ESPACIOS TOPOLOGICOS.

<u>Definición</u>. 2.5.1 Un espacio topológico es un conjunto S junto con una familia distinguida de subconjuntos $\mathcal{T}$ llamados conjuntos abiertos, con las siguientes propiedades:
(i). $\mathcal{T}$ es cerrado bajo intersecciones finitas, esto es si $A,B\in\mathcal{T}$, entonces $A\cap B\in\mathcal{T}$.
(ii). $\mathcal{T}$ es cerrado bajo uniones arbitrarias, esto es, si $\{A_\alpha\}_{\alpha\in I}$ donde I es una familia de índices, entonces $\bigcup_{\alpha\in I} A_\alpha\in\mathcal{T}$.
(iii). $\emptyset\in\mathcal{T}$ y $S\in\mathcal{T}$.
$\mathcal{T}$ es llamada una topología para S, denotaremos un espacio topológico como $(S,\mathcal{T})$.

<u>Ejemplo</u>. 2.5.2 Todo espacio métrico (M,d), es un espacio topológico. Los conjuntos abiertos son aquellos conjuntos, $N\subset M$, con la propiedad de que: para todo $f\in N$ existe $r>0$ tal que $\{g\ ;\ d(f,g)<r\}\subset N$. En el caso en que la topología $\mathcal{T}$ sea inducida por una métrica diremos que $\mathcal{T}$ y d son compatibles.

<u>Definición</u>. 2.5.3 Un espacio topológico se llama metrizable cuando su topología, puede ser introducida mediante una métrica.

<u>Definición</u>. 2.5.4 Un subconjunto C de un espacio topológico $(S,\mathcal{T})$ es llamado cerrado si y sólo si su complemento S–C es abierto.

<u>Proposición</u>. 2.5.5 Sea $(S,\mathcal{T})$ un espacio topológico:
(i). Los conjuntos $\emptyset$ y S son cerrados.
(ii). La intersección arbitraria de cerrados es de nuevo un conjunto cerrado.
(iii). La unión finita de cerrados es de nuevo un conjunto cerrado.
La demostración se sigue de la Def. 2.5.4 y de que
$S-\bigcup_{\alpha\in I} A_\alpha=\bigcap_{\alpha\in I}(S-A_\alpha)$, $S-\bigcap_{\alpha\in I} A_\alpha=\bigcup_{\alpha\in I}(S-A_\alpha)$.

La familia de todas las topologías sobre un conjunto S, es ordenada parcialmente en forma natural $\mathcal{T}_1<\mathcal{T}_2$ si y sólo si $\mathcal{T}_1\subset\mathcal{T}_2$, y diremos que $\mathcal{T}_1$ es una topología más débil que $\mathcal{T}_2$. (el término más débil significa que más sucesiones convergen en $\mathcal{T}_1$ que en $\mathcal{T}_2$).

Proposición. 2.5.6 La intersección arbitraria de topologías de un conjunto cualquiera S, es una topología de S, y esta topología es la topología más débil de S. Efectivamente, es claro que $S,\emptyset\in\bigcap_{\alpha\in I}\mathcal{T}_\alpha$. Además cada familia $\mathcal{T}_\alpha$ es cerrada respecto a uniones arbitrarias e intersecciones finitas; de aqui se deduce que $\bigcap_{\alpha\in I}\mathcal{T}_\alpha$ también posee esta propiedad. Y $\bigcap_{\alpha\in I}\mathcal{T}_\alpha\subset\mathcal{T}_\beta$ para toda $\beta\in I$.

Definición. 2.5.7 Una familia $B\subset\mathcal{T}$ es llamada una base si y sólo si cualquier $A\in\mathcal{T}$ es de la forma $A=\bigcup_{\alpha\in I}B_\alpha$, para alguna subfamilia $\{B_\alpha\}_{\alpha\in I}\subset B$.

Ejemplo. 2.5.8 Las bolas en un espacio métrico (M,d) son una base.

Ahora daremos una familia completa de definiciones tomadas directamente de las propiedades de los espacios métricos.

Definición. 2.5.9 Un conjunto N es llamado una vecindad de un punto $f\in S$, S un espacio topológico, si existe un abierto O tal que $f\in O\subset N$.

Notamos que las vecindades así definidas no necesariamente son conjuntos abiertos, para evitar confusiones, nos referiremos explícitamente a vecindades abiertas, o vecindades cerradas dentro del estudio de los espacios topológicos, según sea el caso.

Definición. 2.5.10 Una familia X de subconjuntos de un espacio topológico $(S,\mathcal{T})$, es llamada una base de vecindades en el punto f, si cada $N\in X$ es una vecindad de f, y si dada cualquier vecindad M de f, existe una $N\in X$, tal que $N\subset M$. Equivalentemente X es una base de vecindades en f, si y sólo si $\{M\ ;\ N\subset M$ para algún $N\in X\}$ es la familia de todas las vecindades de f.

Ejemplo. 2.5.11 En un espacio métrico (M,d) las bolas cerradas de radio positivo con centro en f, son una base de vecindades en f.

Definición. 2.5.12 Sea A un subconjunto de un espacio topológico $(S,\mathcal{T})$. La cerradura de A, denotada por $\bar{A}$, es el mínimo conjunto cerrado que contiene a A. El interior de A, denotado por A^o, es el más grande conjunto abierto contenido en A. La frontera de A, denotada por ∂A, es el conjunto $\bar{A}-A^o=\bar{A}\cap\overline{(S-A)}$. Así $\bar{A}$ es la intersección de todos los cerrados que contienen a A y A^o es la unión de todos los subconjuntos abiertos de A. Vemos que $f\in\bar{A}$ si y sólo si toda vecindad de f intersecta a A.

Definición. 2.5.13 Sea $(S_1,\mathcal{T}_1)$ y $(S_2,\mathcal{T}_2)$ dos espacios topológicos. Una función $F:(S_1,\mathcal{T}_1)\to(S_2,\mathcal{T}_2)$ es llamada continua si $F^{-1}(A)\in\mathcal{T}_1$ es abierto para todo abierto $A\in\mathcal{T}_2$; esto es, la imagen inversa de todo conjunto abierto es también un conjunto abierto. Si además $F(B)\in\mathcal{T}_2$ es abierto para cada abierto $B\in\mathcal{T}_1$, F es llamada bicontinua. Una biyección bicontinua es llamada un homeomorfismo.

Ejemplo. 2.5.14 Los intervalos $(-\infty,\infty)$ y $(-1,1)$ son homeomorfos bajo $x\mapsto x/(1+x^2)$, es obvio que $(-\infty,\infty)$ y $(-1,1)$ no son isométricos en la métrica usual, de hecho, sólo uno de ellos es completo. Esto muestra que la completación no es una noción topológica.

Definición. 2.5.15 Sea K una familia de funciones de un conjunto X a un espacio topológico $(S,\mathcal{T})$. La topología K-débil sobre X es la topología más débil, para la cual todas las funciones $E\in K$ son continuas.

Ejemplo. 2.5.16 Consideremos $C[a,b]$. La topología de convergencia puntual sobre $C[a,b]$, es la topología débil dada por la familia de funciones $f\mapsto f(x)$. Esto es, para cada $x\in[a,b]$, sea $E_x(f)=f(x)$ así las $E_x(\cdot)$ son funciones de $C[a,b]$ a $\mathbb{R}$, la topología de convergencia puntual sobre $C[a,b]$ es tal que $f_n\to f$ si y sólo si $f_n(x)\to f(x)$ para cada $x\in[a,b]$.

Definición. 2.5.17

(i). Un espacio topológico es llamado un T_1-espacio si y sólo si para toda f,g con $f\neq g$, existe un conjunto abierto O con $g\in O$ y $f\notin O$.

(ii). Un espacio topológico es llamado un T_2-espacio o Hausdorff si y sólo si para todo f,g con $f\neq g$, existen conjuntos abiertos O_1 y O_2 tales que $f\in O_1$ y $g\in O_2$ con $O_1\cap O_2=\emptyset$.

(iii). Un espacio topológico es llamado regular o un T_3-espacio si y sólo si es un T_1-espacio y para todo f y C un conjunto cerrado, con $f\notin C$, existen conjuntos abiertos O_1 y O_2 tales que $f\in O_1$, $C\subset O_2$ y $O_1\cap O_2=\emptyset$.

(iv). Un espacio topológico es llamado normal o un T_4-espacio si y sólo si este es un T_1-espacio y para todo C_1,C_2, conjuntos cerrados, tales que $C_1\cap C_2=\emptyset$, existen conjuntos abiertos O_1 y O_2, tales que $C_1\subset O_1$ y $C_2\subset O_2$ con $O_1\cap O_2=\emptyset$.

Observamos que un espacio topológico S es un T_1-espacio si y sólo si el conjunto $\{f\}$ es cerrado para cada f. En efecto, si cada $\{f\}$ es cerrado, y $f,g\in S$ con $f\neq g$, entonces $S-\{f\}$ es abierto y $g\in O=S-\{f\}$ y $f\notin O$, inversamente, cada $g\in S-\{f\}$ esta contenido en un abierto O_g tal que $g\in O_g\subset S-\{f\}$. Así $S-\{f\}=\bigcup_g O_g$ es abierto, luego $\{f\}$ es cerrado. Obviamente:

Proposición. 2.5.18 $T_4\Rightarrow T_3\Rightarrow T_2\Rightarrow T_1$.

Definición. 2.5.19

(i). Un espacio topológico $(S,\mathcal{T})$ es llamado separable si y sólo si S tiene un subconjunto denso numerable.

(ii). Se dice que un espacio topológico $(S,\mathcal{T})$ satisface el primer axioma de numerabilidad si y sólo si cada punto $f\in S$, tiene una base de vecindades numerable.

(iii). Se dice que un espacio topológico $(S,\mathcal{T})$ satisface el segundo axioma de numerabilidad si y sólo si S tiene una base numerable.

Proposición. 2.5.20

(i). Todo espacio topológico $(S,\mathcal{T})$ cuya topología $\mathcal{T}$ es compatible con una métrica d satisface el primer axioma de numerabilidad.

(ii). Un espacio topológico $(S,\mathcal{T})$ cuya topología $\mathcal{T}$ es compatible con una métrica d satisface el segundo axioma de numerabilidad si y sólo si este es separable.

(iii). Cualquier espacio topológico $(S,\mathcal{T})$ que satisface el segundo axioma de numerabilidad es separable.

Dem.(i). A cada punto $f\in S$ le asignamos la familia numerable de bolas cerradas de radio m/n con n,m en el conjunto de números naturales y centro en el punto f, denotadas por $\bar{B}(f,m/n)$. Es inmediato que esta familia constituye una base de vecindades numerable en f, puesto que cada $\bar{B}(f,m/n)$ es una vecindad cerrada de f, y dada cualquier vecindad N de f, existe una vecindad cerrada $\bar{B}(f,m/n)$ tal que $\bar{B}(f,m/n)\subset N$.

Dem.(ii). Supongamos que $\{G_n\}_{n=1}^{\infty}$ es una base numerable de S. Tomemos de cada G_n un punto arbitrario g_n. El conjunto $\{g_n\}_{n=1}^{\infty}$ es denso en S, ya que, de lo contrario el conjunto abierto $S-(\{g_n\}_{n=1}^{\infty})$ no contendría ningún punto de $\{g_n\}_{n=1}^{\infty}$ y esto no puede ocurrir, puesto que $S-(\{g_n\}_{n=1}^{\infty})$ es unión de determinados conjuntos del sistema $\{G_n\}_{n=1}^{\infty}$.

Inversamente, sea $\{g_n\}_{n=1}$ un subconjunto denso numerable de S y consideremos la familia numerable constituida por las bolas abiertas de la forma $B(g_n,1/m)$, donde n,m recorren todos los valores naturales. Claramente esta familia forma una base numerable de S.

Definición. 2.5.21 Un espacio topológico $(S,\mathcal{T})$ es llamado un espacio de Lindelöf, si toda cubierta abierta de S contiene una subcubierta abierta numerable de S.

Proposición. 2.5.22 Todo espacio topológico $(S,\mathcal{T})$ que satisface el segundo axioma de numerabilidad es un espacio de Lindelöf.
Demostración: Sea $\{G_n\}_{n=1}^{\infty}$ una base numerable de S, de cada G_n tomemos un g_n arbitrario, como hemos visto $\{g_n\}_{n=1}^{\infty}$ es un subconjunto denso numerable de S. Sea $\{O_\alpha\}_{\alpha\in I}$ una cubierta abierta de S. Entonces para cada g_n tomamos cualquier O_α tal que $g_n \in O_\alpha$, que denotamos por O_n. Veamos que $\{O_n\}_{n=1}^{\infty}$ cubre a S. En efecto, supongamos que $\{O_n\}_{n=1}^{\infty}$ no cubre a S, entonces el abierto $S-\bigcup_{n=1}^{\infty} O_n$ no contendría ningún g_n, lo que contradice el hecho de que $\{g_n\}_{n=1}^{\infty}$ es denso en S.

Un espacio topológico $(S,\mathcal{T})$ es llamado disconexo si y sólo si este contiene un subconjunto propio no vacío, el cual es a la vez abierto y cerrado; equivalentemente. Un espacio topológico es llamado conexo si y sólo si no existen dos subconjuntos de S abiertos, ajenos y no vacios tales que su unión es igual a S.
Sea F una función continua suprayectiva de un espacio topológico conexo $(S_1,\mathcal{T}_1)$ a un espacio topológico $(S_2,\mathcal{T}_2)$. Entonces S_2 es conexo. Efectivamente, supongamos que S_2 es disconexo, entonces existen O_1 y O_2 abiertos, ajenos y no vacios tales que $S_2=O_1\cup O_2$. Entonces $F^{-1}(O_1)$ y $F^{-1}(O_2)$ son abiertos y ajenos en S_1, y dado que F es suprayectiva, entonces también $F^{-1}(O_1)$, $F^{-1}(O_2)$ son no vacios.
Sea $(S,\mathcal{T})$ un espacio topológico y sea $A\subset S$. La topología relativa sobre A es la familia de conjuntos $\mathcal{T}_A=\{O\cap A \ ; \ O\in\mathcal{T}\}$. Un subconjunto $B\subset A$ es llamado abierto relativo de A si $B\in\mathcal{T}_A$ y es llamado cerrado relativo de A si $A-B\in\mathcal{T}_A$.
Si $(S_1,\mathcal{T}_1)$, $(S_2,\mathcal{T}_2)$ son dos espacios topológicos definimos una topología sobre el producto $S_1\times S_2$, tomando como una base la colección de todos los conjuntos de la forma $O_1\times O_2$, donde $O_1\in\mathcal{T}_1$ y $O_2\in\mathcal{T}_2$. Esta topología es llamada la topología producto para $S_1\times S_2$.

Definición. 2.5.23 Un sistema dirigido es un conjunto de índices I junto con un orden el cual satisface:
(i). Si $\alpha,\beta\in I$, entonces existe $\gamma\in I$ tal que $\alpha<\gamma$, $\beta<\gamma$.
(ii). $<$ es un orden parcial.

Daremos ahora una generalización de la noción de sucesión.

Definición. 2.5.24 Una red en un espacio topológico $(S,\mathcal{T})$, es una función del sistema dirigido a S; denotamos esto por $\{f_\alpha\}_{\alpha\in I}$.

Definición. 2.5.25 Una red en un espacio topológico $(S,\mathcal{T})$, se dice que converge a un punto $f\in S$, denotado $f_\alpha\to f$ si para cualquier vecindad N de f existe $\beta\in I$ tal que $f_\alpha\in N$ si $\alpha>\beta$.

Teorema. 2.5.26 Sea A un subconjunto de un espacio topológico $(S,\mathcal{T})$. Entonces un punto f está en la cerradura de A si y sólo si existe una red $\{f_\alpha\}_{\alpha\in I}$, tal que $f_\alpha\to f$.
Demostración: Primero vemos que $\bar{A}$ es el conjunto de puntos f tales que cualquier vecindad de f contiene un punto de A. Este conjunto $\bar{A}$ es el mínimo conjunto cerrado que contiene a A y su complemento es el más grande conjunto abierto que no contiene puntos de A. Ahora supongamos que $f_\alpha\to f$ donde cada $f_\alpha\in A$. Entonces cualquier vecindad de f contiene algún f_α y por lo tanto algún punto de A, esto es, f es un punto límite de A, así $f\in\bar{A}$. Inversamente, supongamos que $f\in\bar{A}$. Sea I la colección de vecindades de f con el orden $N_1<N_2$ si y sólo si $N_1\supset N_2$. Para cada $N\in I$. Sea f_N un punto en $A\cap N$. Entonces $\{f_N\}_{N\in I}$ es una red y $f_N\to f$.

Proposición. 2.5.27 Sea S un espacio de Hausdorff. Entonces una red $\{f_\alpha\}_{\alpha\in I}$ en S tiene a lo más un límite, esto es, si $f_\alpha\to f$ y $f_\alpha\to g$. Entonces $f=g$.
Demostración: Supongamos que $f\neq g$, y que $f_\alpha\to f$, $f_\alpha\to g$, entonces para cualesquier vecindades N_f y N_g de f y g, respectivamente, existen β y $\beta'\in I$ tales que, $f_\alpha\in N_f$ siempre que $\alpha>\beta$ y $f_\alpha\in N_g$ siempre que $\alpha>\beta'$. Si $\gamma=\max(\beta,\beta')$, entonces $f_\alpha\in N_f$ y $f_\alpha\in N_g$ si $\alpha>\gamma$ y este hecho de

muestra que $N_f \cap N_g \neq \emptyset$ para cualesquier vecindades de f y g con $f \neq g$, lo que contradice que S es un espacio de Hausdorff.

Definición. 2.5.28 Un espacio topológico es llamado compacto, si toda cubierta abierta del espacio tiene una subcubierta finita. Más explícitamente, si S es un espacio topológico y $\{G_\alpha\}_{\alpha \in I}$ es una cubierta abierta de S, entonces existe un número finito de índices $\alpha_1, \ldots, \alpha_n$ tal que $S \subset G_{\alpha_1} \cup \cdots \cup G_{\alpha_n}$.

Proposición. 2.5.29

(i). Un subconjunto cerrado de un espacio topológico compacto es compacto.

(ii). Un subconjunto compacto de un espacio de Hausdorff es cerrado.

Demostración: Sea S compacto, $F \subset S$, cerrado y U una cubierta abierta de F. Entonces $U \cup (S-F)$ es una cubierta abierta de S, y por lo tanto S tiene una subcubierta finita $\{S-F, O_1, \ldots O_N\}$. Entonces los conjuntos $\{O_1, \ldots, O_N\}$ cubren F, así U tiene una subcubierta finita. Supongamos ahora que S es Hausdorff y K un subconjunto compacto de S, mostraremos que S-K es abierto, de donde concluiremos que K es cerrado. Sea $f \in S-K$, ya que S es Hausdorff para cada $g \in K$, existen conjuntos abiertos O_g y N_g ajenos con $g \in O_g$ y $f \in N_g$. Los conjuntos $\{O_g \ ; \ g \in K\}$ forman una cubierta abierta de K, como K es compacto existe una subcolección finita $\{O_{g1}, \ldots, O_{gN}\}$ la cual cubre a K. Sea $N = \bigcap_{i=1}^{N} N_{gi}$. Entonces N es un conjunto que contiene a f y es ajeno a cualquier O_{gi}. Puesto que $K \subset \bigcup_{i=1}^{N} O_{gi}$, N es ajeno con K y así $f \in N \subset S-K$, de donde f es un punto interior de S-K, y como f es un punto arbitrario se sigue que S-K es abierto; por lo tanto K es cerrado.

Definición. 2.5.30 Supongamos que $\mathcal{T}$ es una topología sobre un espacio lineal X tal que:

(i). Todo punto $f \in X$ es un conjunto cerrado.

(ii). La operaciones del espacio lineal son continuas con respecto a $\mathcal{T}$.

Bajo estas condiciones $\mathcal{T}$ es llamada una topología lineal sobre X, y X es un espacio lineal topológico.

Decir que la adición de elementos de X es continua, con respecto a $\mathcal{T}$ significa que $+: X\times X \to X$, $(f,g) \mapsto f+g$ es continua, esto es, si V es una vecindad de f+g, entonces existirán V_f y V_g tales que $V_f+V_g \subset V$ donde $V_f+V_g=\{f+g \ ; \ f\in V_f \text{ y } g\in V_g\}$.
Similarmente, la suposición de que la multiplicación por escalares es continua con respecto a $\mathcal{T}$ significa que la función $\cdot: \mathbb{R}\times X \to X$ $(a,f) \mapsto af$ es continua, esto es, si V es una vecindad de af, entonces para algún $r>0$ y alguna vecindad W de f tenemos que $bW\subset V$ siempre que $|b-a|<r$.

Sea $C^{\infty}[a,b]$ el espacio de funciones infinitamente diferenciables definidas sobre $[a,b]$. Definimos una topología en $C^{\infty}[a,b]$ por medio del siguiente sistema de vecindades abiertas de cero. Cada una de estas vecindades $N_{n,\epsilon}$ se determinan por los números n y ϵ y consta de todas las funciones f que verifican las desigualdades $|f^{(k)}(x)|<\epsilon$ para toda $x\in[a,b]$ y $k=0,1,\dots,n$. Así $C^{\infty}[a,b]$ es un espacio lineal topológico.

<u>Definición</u>. 2.5.31 Una base local de un espacio topológico lineal es una colección B de vecindades abiertas de 0 (cero) tal que toda vecindad abierta de 0 contiene un elemento de B. Los conjuntos abiertos de X son aquellos que son uniones de traslaciones de elementos de B.

Sea $(X,\mathcal{T})$ un espacio topológico lineal. La noción de sucesión de Cauchy puede ser definida sin hacer referencia a métricas sobre X, de la siguiente manera. Fijamos una base local B para $\mathcal{T}$, una sucesión $\{f_n\}_{n=1}^{\infty}\subset X$ es entonces llamada de Cauchy si para todo $V\in B$ corresponde un N tal que $f_n-f_m\in V$ si $n,m>N$.
Supongamos ahora que $(X,\mathcal{T})$ es un espacio topológico lineal cuya topología $\mathcal{T}$ es compatible con una métrica invariante bajo traslaciones d, y ya que tenemos que una sucesión es $\mathcal{T}$-Cauchy si y sólo si es d-Cauchy, puesto que las d-bolas de radio positivo centradas en el origen forman una base local para $\mathcal{T}$ y $d(f_n,f_m)=d(f_n-f_m,f_m-f_m)=d(f_n-f_m,0)$. De donde se desprende la siguiente proposición.

Proposición. 2.5.32 Si $(X,\mathcal{T})$ es un espacio lineal topológico tal que su topología $\mathcal{T}$ es compatible con una métrica d invariante bajo traslaciones. Entonces una sucesión en X es $\mathcal{T}$-Cauchy si y sólo si es d-Cauchy.(La dem. se sigue de que $d(f_n-f_m,0)=d(f_n,f_m)\to 0$ si $n,m\to\infty$)

Proposición. 2.5.33 Todo espacio topológico $(X,\mathcal{T})$ tal que su topología $\mathcal{T}$ es compatible con una métrica d es un T_4-espacio o normal. Demostración: Sean Y, Z dos subconjuntos de X cerrados de intersección vacía. Todo punto $f\in Y$ tiene una vecindad N_f que es ajena con Z y por consiguiente está a una distancia positiva r_f de Z. De la misma forma todo $g\in Z$ esta a una distancia positiva r_g de Y. Consideremos los conjuntos abiertos $\bigcup_{f\in Y} B(f,\frac{1}{2}r_f)$, $\bigcup_{g\in Z} B(g,\frac{1}{2}r_g)$ que contienen a Y y Z respectivamente, y veamos que la intersección de estos conjuntos es vacía. Supongamos que $h\in B(f,\frac{1}{2}r_f)\cap B(g,\frac{1}{2}r_g)$. En este caso existe un punto $v\in Y$, tal que $d(v,h)<\frac{1}{2}r_f$, y en Z un punto w tal que $d(w,h)<\frac{1}{2}r_g$. Podemos suponer sin pérdida de generalidad que $r_g>r_f$. Entonces $d(v,w)\leq d(v,h)+d(h,w)<\frac{1}{2}r_f+\frac{1}{2}r_g<r_g$, es decir $v\in B(w,r_g)$ lo que contradice a la forma en que se definió r_g. Lo que concluye nuestra afirmación.

Corolario. 2.5.34 Todo espacio topológico $(X,\mathcal{T})$ tal que su topología $\mathcal{T}$ es compatible con una métrica d es un T_2-espacio o de Hausdorff.

Todo subespacio de un espacio métrico es por sí mismo un espacio métrico y por esto también posee la propiedad de normalidad.

2.6 ESPACIOS NORMADOS.

Definición. 2.6.1 Un espacio lineal normado, es un espacio lineal V sobre $\mathbb{R}$ y una función $\|\cdot\|$ de V a $\mathbb{R}^+$, tal que para todo $f,g\in V$ y $a\in\mathbb{R}$ satisface que:

(i). $\|f\|\geq 0$.

(ii). $\|f\|=0$ si y sólo si $f=0$.

(iii). $\|af\|=|a|\cdot\|f\|$.

(iv). $\|f+g\|\leq\|f\|+\|g\|$.

La función $\|\cdot\|$ es llamada una norma para V.

Denotamos un espacio lineal normado por $(V,\|\cdot\|)$.

Ejemplo. 2.6.2 Sea $C[a,b]$, definimos $\|f\|_1=\max_{x\in[a,b]}|f(x)|$ o $\|f\|_2=\int_a^b|f(x)|dx$, entonces $(C[a,b],\|\cdot\|_1)$, $(C[a,b],\|\cdot\|_2)$ son espacios lineales normados.

Definición. 2.6.3 Una transformación lineal acotada (operador lineal acotado) de un espacio lineal normado $(V_1,\|\cdot\|_1)$ a otro $(V_2,\|\cdot\|_2)$, es una función $T:V_1 \to V_2$ la cual satisface para todo $f,g\in V_1$ y $a,b\in\mathbb{R}$ que:

(i). $T(af+bg)=aT(f)+bT(g)$.

(ii). Existe $C\geq 0$ tal que $\|T(f)\|_2\leq C\,\|f\|_1$ para toda $f\in V_1$.

El mínimo de tales C's, es llamado la norma de T, denotada $\|T\|$. Así $\|T\|=\sup_{\|f\|\leq 1}\|T(f)\|=\sup_{f\neq 0}\|T(f)\|/\|f\|$, puesto que $\|T(f)\|/\|f\|=\|T(f/\|f\|)\|$ y por definición de supremo, tenemos que para todo $\epsilon>0$ existe $g\in V_1$ tal que $\|T(g)\|/\|g\|>\|T\|-\epsilon$, esto implica que para todo $\epsilon>0$ existe $g\in V$ tal que $\|T(g)\|>(\|T\|-\epsilon)\|g\|$. (ver Prop. 2.2.5)

Ejemplo. 2.6.4 Sea $V=C[a,b]$ y sea $\|f\|=\max_{x\in[a,b]}|f(x)|$, si $T:V\to V$; $T(f)=f(x)g_0(x)$ donde $g_0(x)$ es una función fija en $C[a,b]$. Entonces T es un operador lineal acotado de V en sí mismo. En efecto, es claro T es lineal y que $\|T(f)\|\leq\|f\|\|g_0\|$ para todo $f\in V$, si $f\neq 0$, entonces $\|T(f)\|/\|f\|\leq\|g_0\|$, luego $\|T\|\leq\|g_0\|$. Por otra parte, tomemos $f(x)=1$ para toda $x\in[a,b]$, así $\|T(f)\|\geq\|f\||g_0(x)|$ es válida para $f(x)=1$ y para $x\in[a,b]$, luego $\|T\|\geq\|T(f)\|/\|f\|\geq|g_0(x)|$ es válida para toda $x\in[a,b]$, entonces $\|T\|\geq\|g_0\|$; Por lo que $\|T\|=\|g_0\|$.

Observamos que cualquier espacio lineal normado $(V,\|\cdot\|)$ es un espacio lineal métrico, definiendo $d(f,g)=\|f-g\|_V$. De donde tenemos las nociones de convergencia, completación y continuidad en espacios lineales normados y, por consiguiente $(V,\|\cdot\|)$ es también un espacio lineal topológico, donde sólo se conservan, como se sabe las nociones de convergencia y continuidad.

Ahora daremos una familia completa de definiciones y proposiciones, tomadas directamente de las propiedades de espacios métricos.

Definición. 2.6.5 Una sucesión $\{f_n\}_{n=1}^{\infty}$ en un espacio lineal normado V, se dice que converge a un elemento f en el espacio V, si dado $\epsilon>0$, existe una N tal que para toda $n>N$ tenemos $\|f_n-f\|<\epsilon$. Denotaremos que $\{f_n\}_{n=1}^{\infty}$ converge a f en V por $f_n\overset{\|\cdot\|}{\to}f$ o por $\lim_{n\to\infty}f_n=f$.

Frecuentemente este hecho se denota por: $\| f_n - f \| \to 0$ cuando $n \to \infty$.

Proposición. 2.6.6 En un espacio lineal normado $(V, \|\cdot\|)$ el límite de una sucesión es único. Dem. Si $\| f_n - g \| \to 0$ y $\| f_n - h \| \to 0$ cuando $n \to \infty$, entonces $\| g-h \| \leq \| g-f_n \| + \| f_n - h \| \to 0$ cuando $n \to \infty$. así $g=h$.

Definición. 2.6.7 Una sucesión en un espacio lineal normado $(V, \|\cdot\|)$ se llama de Cauchy, si dado $\epsilon>0$, existe un N tal que para todo $n,m>N$ tenemos $\| f_n - f_m \| < \epsilon$. Frecuentemente, este hecho se denota por $\| f_n - f_m \| \to 0$ cuando $n,m \to \infty$.

Proposición. 2.6.8 En un espacio lineal normado $(V, \|\cdot\|)$, cualquier sucesión convergente es de Cauchy. Dem. Si $\| f_n - f \| \to 0$ y $\| f_m - f \| \to 0$ cuando $n,m \to \infty$, entonces $\| f_n - f_m \| \leq \| f_n - f \| + \| f - f_m \| \to 0$ cuando $n,m \to \infty$.

Definición. 2.6.9 Sea $(V, \|\cdot\|)$ un espacio lineal normado. El conjunto $\{g \; ; \| f-g \| \leq r\}$ es llamado una bola cerrada con centro en f y radio r. El conjunto $\{g \; ; \| g-f \| < r\}$ es llamado una bola abierta con centro en f y radio r.

Definición. 2.6.10 Sea C un subconjunto de un espacio lineal normado V, C es llamado un conjunto convexo si $cf+(1-c)g \in C$ siempre que $f,g \in C$ y $c \in [0,1]$.

Por la proposición 2.3.14 sabemos que la intersección arbitraria de conjuntos convexos es de nuevo un conjunto convexo.

Al considerar la Def. 2.5.12. se puede ver que C^o es también un conjunto convexo siempre que C lo sea, puesto que $C^o \subseteq C$ y si $c \in (0,1)$, entonces $cC^o + (1-c)C^o \subseteq C$ y cC^o, $(1-c)C^o$ son abiertos, luego también lo es su suma y como todo abierto de C pertenece a C^o, se tiene C^o convexo.

Proposición. 2.6.11 En un espacio lineal normado $(V, \|\cdot\|)$ las bolas son convexas.

Demostración: Sea $B(g,r)$ una bola abierta de radio r y centro en g, entonces si $f,h \in B(g,r)$, $\| f-g \| < r$ y $\| h-g \| < r$, así $\| cf+(1-c)h-g \| = \| cf+(1-c)h-cg-(1-c)g \| \leq \| cf-cg \| + \| (1-c)h-(1-c)g \| = c \| f-g \| + (1-c) \| h-g \| < cr+(1-c)r = r$.

Definición. 2.6.12 Diremos que un espacio lineal normado $(V, \|\cdot\|)$ es completo, si éste es completo como un espacio métrico en la métrica inducida.

Proposición. 2.6.13 Si $(V,\|\cdot\|)$ es un espacio lineal normado, entonces V tiene una completación como un espacio métrico.
La demostración se sigue del teorema 2.4.28 y de que $d(f,g)=\|f-g\|$ es una métrica natural para espacios lineales normados.

Teorema. 2.6.14 (Teorema de la transformación lineal acotada). Supongamos que T es una transformación lineal acotada de un espacio lineal normado $(V_1,\|\cdot\|_1)$ a un espacio lineal normado completo $(V_2,\|\cdot\|_2)$. Entonces T puede ser extendida en forma única, a una transformación lineal acotada (con la misma cota), T', de la completación de $(V_1, \|\cdot\|_1)$ a $(V_2,\|\cdot\|_2)$.
Demostración: Sea V_1' la completación de V_1. Para cada $f\in V_1'$, existe una sucesión de elementos $\{f_n\}_{n=1}^{\infty}\subset V_1$ con $f_n \xrightarrow{\|\cdot\|_1} f$ cuando $n\to\infty$. Puesto que $\{f_n\}_{n=1}^{\infty}$ converge, esta sucesión es de Cauchy, así dado $\epsilon>0$, podemos encontrar N tal que para $n,m>N$ implica $\|f_n-f_m\|\leq \epsilon/C$, supongamos que existe un C tal que $\|T(h)\|\leq C\|h\|$ para toda $h\in V_1$. Entonces $\|T(f_n)-T(f_m)\|_2 = \|T(f_n-f_m)\|_2\leq \|T\|\|f_n-f_m\|_1\leq C\|f_n-f_m\|_1<\epsilon$ lo cual prueba que $\{T(f_n)\}_{n=1}^{\infty}$ es de Cauchy en V_2. Puesto que V_2 es completo $T(f_n)\xrightarrow{\|\cdot\|_2} g$ para algún $g\in V_2$. Pongamos $T'(f)=g$. Debemos ver que esta definición es independiente de la sucesión $f_n \to f$ elegida. Para esto supongamos que $f_n\to f$ y $f_n'\to f$ cuando $n\to\infty$, entonces la sucesión $f_1,f_1',f_2,f_2',\dots \to f$, así $T(f_1),T(f_1'),T(f_2),T(f_2'),\dots \to g'$ para algún $g'\in V_2$, y puesto que V es completo, $\lim_{n\to\infty} T(f_n')=g'=\lim_{n\to\infty} T(f_n)=g$. Además T' es acotada, puesto que $\|T'(f)\|_2=\lim_{n\to\infty}\|T(f_n)\|_2\leq\lim_{n\to\infty} C\|f_n\|_1= C\lim_{n\to\infty}\|f_n\|_1=C\|f\|$. La unicidad se desprende de la unicidad del límite, puesto que hemos definido $T'(f)=\lim_{n\to\infty} T(f_n)$, es obvio que T' es lineal.

Definición. 2.6.15 Un espacio lineal normado completo, es llamado un espacio de Banach.

Ahora consideremos a T un operador o transformación lineal acotada de un espacio lineal normado $(V_1,\|\cdot\|_1)$ a un espacio lineal normado $(V_2,\|\cdot\|_2)$. Denotemos al conjunto de todos los operadores lineales acotados de $(V_1,\|\cdot\|_1)$ a $(V_2,\|\cdot\|_2)$ por $\mathfrak{L}(V_1,V_2)$, e introduzcamos una norma sobre $\mathfrak{L}(V_1,V_2)$ definiendo

$$\|T\|=\sup_{f\neq 0}\frac{\|T(f)\|}{\|f\|}.$$

Teorema. 2.6.16 Si V_1 es un espacio lineal normado y V_2 es un espacio lineal normado completo (de Banach). Entonces $\mathcal{L}(V_1,V_2)$ es también un espacio de Banach.

Demostración: Puesto que cualquier combinación lineal finita de operadores lineales acotados, es otra vez un operador lineal acotado. $\mathcal{L}(V_1,V_2)$ es un espacio lineal, es fácil ver que $\|\cdot\|$ es una norma: por ejemplo la desigualdad del triángulo es probada como sigue: Sean $T_1,T_2\in\mathcal{L}(V_1,V_2)$, entonces

$$\|T_1+T_2\| = \sup_{f\neq 0}\frac{\|(T_1+T_2)(f)\|}{\|f\|} \leq \sup_{f\neq 0}\frac{\|T_1(f)\| + \|T_2(f)\|}{\|f\|}$$
$$\leq \sup_{f\neq 0}\frac{\|T_1(f)\|}{\|f\|} + \sup_{f\neq 0}\frac{\|T_2(f)\|}{\|f\|} = \|T_1\| + \|T_2\| .$$

Para demostrar que $\mathcal{L}(V_1,V_2)$ es completo, debemos probar que si $\{T_n\}_{n=1}^{\infty}$ es una sucesión de Cauchy en la norma operador, entonces existe un operador lineal acotado T, tal que $\|T_n-T\| \to 0$ cuando $n\to\infty$. Vemos que

$\|T_n(f)-T_m(f)\| = \|(T_n-T_m)(f)\| \leq \|T_n-T_m\|\,\|f\| \to 0$ cuando $n,m\to\infty$, de donde para cada $f\in V_1$, la sucesión $\{T_n(f)\}_{n=1}^{\infty}$ es de Cauchy en V_2.

Puesto que V_2 es completo $T_n(f)$ converge a un elemento $g\in V_2$. Definimos $T(f)=g$. Veamos que T es un operador lineal acotado. Puesto que $\big|\|T_n\| - \|T_m\|\big| \leq \|T_n-T_m\| \to 0$ cuando $n,m\to\infty$, así $\{\|T_n\|\}_{n=1}^{\infty}$ es una sucesión de Cauchy de números reales que converge a algún número C, y $\big|\|T_n(f)\| - \|T_m(f)\|\big| \leq \|T_n(f)-T_m(f)\| \to 0$ cuando $n,m\to\infty$. De donde $\|T(f)\| = \lim_{n\to\infty}\|T_n(f)\| \leq \lim_{n\to\infty}\|T_n\|\,\|f\| = C\,\|f\|$, así T es un operador lineal acotado. Debemos ahora mostrar que $T_n\to T$ en la norma operador. Puesto que

$\|(T-T_n)(f)\| = \lim_{m\to\infty}\|(T_m-T_n)(f)\| \leq \lim_{m\to\infty}\|T_m-T_n\|\,\|f\|$, si $f\neq 0$ tenemos

$$\frac{\|(T-T_n)(f)\|}{\|f\|} \leq \lim_{m\to\infty}\|T_m-T_n\| .$$

De esta manera $\lim_{m\to\infty}\|T_m-T_n\|$ es cota superior del conjunto de números $\|(T-T_n)(f)\| / \|f\|$ y puesto que el supremo de estos números es la mínima cota superior tenemos que

$\|T-T_n\| = \sup_{f\neq 0}\dfrac{\|(T-T_n)(f)\|}{\|f\|} \leq \lim_{m\to\infty}\|T_m-T_n\| \to 0$ cuando $n\to\infty$. Finalmente vemos que $\|T\|=C$, en efecto

$$\big|\|T\| - C\big| \leq \big|\|T\| - \|T_n\|\big| + \big|\|T_n\| - C\big|$$
$$\leq \|T-T_n\| + \big|\|T_n\| - C\big| \to 0 \text{ cuando } n\to\infty .$$

<u>Definición</u>. 2.6.17 Una sucesión de elementos $\{f_n\}_{n=1}^{\infty}$ en un espacio lineal normado $(V,\|\cdot\|)$, es llamada absolutamente sumable, si $\sum_{n=1}^{\infty}\|f_n\|<\infty$, y es llamada sumable si $\sum_{n=1}^{N} f_n$ converge cuando $N\to\infty$ a un $f\in V$.

<u>Teorema</u>. 2.6.18 Un espacio lineal normado $(V,\|\cdot\|)$ es completo si y sólo si toda sucesión absolutamente sumable, es sumable.

Demostración: Supongamos que $(V,\|\cdot\|)$ es completo, y sea $\{f_n\}_{n=1}^{\infty}$ una sucesión absolutamente sumable, y que $N>M$, entonces

$$\left\|\sum_{n=1}^{N} f_n - \sum_{n=1}^{M} f_n\right\| = \left\|\sum_{n=M+1}^{N} f_n\right\| \le \sum_{n=M+1}^{N}\|f_n\| = \left|\sum_{n=1}^{N}\|f_n\| - \sum_{n=1}^{M}\|f_n\|\right| \to 0$$

cuando $N,M\to\infty$, de donde la sucesión $\left\{\sum_{n=1}^{N} f_n\right\}_{N=1}^{\infty}$ es de Cauchy en V, y por ser V completo, esta converge a un $f\in V$.

Inversamente supongamos que $\{f_n\}_{n=1}^{\infty}$ es una sucesión de Cauchy. Entonces para todo entero k existe N_k tal que $\|f_n-f_m\|<2^{k}$ si $n,m>N_k$; escojamos N_k's tales que $N_{k+1}>N_k$. Entonces la sucesión $\{f_{N_k}\}_{k=1}^{\infty}$ es una subsucesión de $\{f_n\}_{n=1}^{\infty}$. Sea $g_1=f_{N1}$ y $g_k=f_{N_k}-f_{N(k-1)}$, de esta manera $\sum_{k=2}^{m} g_k=f_{Nm}$ y $\|g_k\|\le 2^{-k-1}$ si $k>1$; por lo tanto $\sum_{k=2}^{\infty}\|g_k\|\le\|g_1\|+\sum_{k=2}^{\infty}2^{-k+1}=\|g_1\|+1<\infty$, de donde $\{g_k\}_{k=1}^{\infty}$ es absolutamente sumable. Entonces por hipótesis $\{g_k\}_{k=1}^{\infty}$ es sumable, así $\sum_{k=2}^{m} g_k=f_{Nm}\to f$ para algún $f\in V$, y como $\|f_n-f\|\le\lim_{m\to\infty}\|f_n-f_{Nm}\|+\lim_{m\to\infty}\|f_{Nm}-f\|\to 0$ cuando $n\to\infty$, entonces $f_n\to f$ cuando $n\to\infty$ para algún $f\in V$.

<u>Definición</u>. 2.6.19 Dos normas, $\|\cdot\|_1$ y $\|\cdot\|_2$, sobre un espacio lineal normado V, son llamadas equivalentes, si existen constantes positivas C y C', tales que, para toda $f\in V$, $C\|f\|_1\le\|f\|_2\le C'\|f\|_1$.

<u>Ejemplo</u>. 2.6.20 Sea $V=C^n[a,b]$ y sean $\|f\|_1=\max_{x,k}|f^{(k)}(x)|$ y $\|f\|_2=\sum_{k=0}^{n}\max_{x\in[a,b]}|f^{(k)}(x)|$, entonces $(1/n)\|f\|_1\le\|f\|_1\le\|f\|_2=n\|f\|_2$.

<u>Definición</u>. 2.6.21 Una transformación T entre dos espacios lineales normados $T:(V_1,\|\cdot\|_1)\to(V_2,\|\cdot\|_2)$, es llamada continua, si $T(f_n)\xrightarrow{\|\cdot\|_2}T(f)$ siempre que $f_n\xrightarrow{\|\cdot\|_1}f$.

La definición anterior se puede plantear en una definición equivalente en términos de módulos de continuidad ϵ y δ, como sigue: Sea $E\subset(V_1,\|\cdot\|_1)$. Un operador lineal $T:E\subset(V_1,\|\cdot\|_1)\to(V_2,\|\cdot\|_2)$ es llamado continuo en E, si para todo $\epsilon>0$, existe $\delta>0$ tal que para todo $f,g\in E$ con $\|f-g\|<\delta$ tenemos $\|T(f)-T(g)\|<\epsilon$.

<u>Teorema</u>. 2.6.22 Sea T una transformación lineal (operador lineal). $T:(V_1,\|\cdot\|_1)\to(V_2,\|\cdot\|_2)$. Entonces las siguientes proposiciones son equivalentes.

(i). T es continua en todo punto.
(ii). T es continua en un punto.
(iii). T es acotada.

Demostración: (i) $\Rightarrow$ (ii) es obvia. (ii) $\Rightarrow$ (iii). Supongamos que T es continua en f. Entonces existe un $\delta>0$ tal que $\|T(g)-T(f)\|<1$ para todo $g\in V_1$ tal que $\|g-f\|<\delta$. Para cualquier $h\in V_1$, pongamos $w=\eta h/\|h\|$, donde $0<\eta<\delta$. Entonces $(\eta/\|h\|)T(h)=T(w)=T(w+f)-T(f)$ y $(\eta/\|h\|)\|T(h)\|=\|T(w+f)-T(f)\|<1$, puesto que $\|w+f-f\|=\|w\|=\eta<\delta$. Consecuentemente, $\|T(h)\|<\eta^{-1}\|h\|$ para todo $h\in V_1$, así T es acotado.
(iii) $\Rightarrow$ (i). Supongamos que T es acotado, esto es, existe $C>0$ tal que $\|T(h)\|\leq C\|h\|$ para todo $h\in V_1$. Entonces $\|T(f)-T(g)\|=\|T(f-g)\|\leq\|T\|\|f-g\|\leq C\|f-g\|<\epsilon$ para todo $f,g\in V_1$ tales que $\|f-g\|<\epsilon/C$. Siendo f y g puntos arbitrarios tenemos que T es continuo en V_1.
(ii) $\Rightarrow$ (i). Sea T continuo en el punto $f\in V_1$, y sea g cualquier punto de V_1, tenemos $\{g_n\}_{n=1}^{\infty}$ tal que $g_n\to g$ cuando $n\to\infty$, entonces $\|T(g_n)-T(g)\|=\|T((g_n-g)+f)-T(f)\|$, y como $g_n\to g$ cuando $n\to\infty$, entonces $g_n-g\to 0$ cuando $n\to\infty$, por lo tanto $(g_n-g)+f\to f$ cuando $n\to\infty$. Por la continuidad de T en f tenemos que $T((g_n-g)+f)\to T(f)$ cuando $n\to\infty$. Por lo tanto $\|T((g_n-g)+f)-T(f)\|\to 0$ cuando $n\to\infty$. Equivalentemente $\|T(g_n)-T(g)\|\to 0$ cuando $n\to\infty$.

<u>Ejemplo</u>. 2.6.23 El operador de diferenciación puede ser considerado como un operador que actúa del espacio $C^1[a,b]$ con la norma $\|f\|=\max_{x\in[a,b]}|f(x)|+\max_{x\in[a,b]}|f'(x)|$ en el espacio $C[a,b]$. En este caso el operador diferenciación es lineal y continuo y transforma todo el espacio $C^1[a,b]$ en todo el espacio $C[a,b]$. No resulta conveniente considerar el operador diferenciación como un operador el cual

actúa de $C^1[a,b]$ en $C[a,b]$, ya que, a pesar de obtener en este caso tal operador continuo definido sobre todo el espacio, este operador no se puede aplicar dos veces a cualquier función de $C^1[a,b]$. Es más cómodo considerar el operador de diferenciación en un espacio mucho más reducido que $C^1[a,b]$, en el espacio $C^\infty[a,b]$ de funciones infinitamente diferenciables sobre $[a,b]$ cuya topología se define mediante un sistema numerable de normas $\|f\|_n = \max_{x,k} |f^{(k)}(x)|$. El operador diferenciación transforma este espacio en sí mismo, además es continuo sobre este espacio.

Definición. 2.6.24 Un operador lineal acotado T de un espacio lineal normado $(V_1,\|\cdot\|_1)$ a un espacio lineal normado $(V_2,\|\cdot\|_2)$, es llamado un isomorfismo, si este es una biyección, la cual es continua y tiene inversa continua. Si esta preserva norma es llamada una isometría.

Ejemplo. 2.6.25 Supongamos que dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ sobre un espacio de Banach V son equivalentes, entonces la función identidad es un isomorfismo de $(V_1,\|\cdot\|_1)$ en $(V_2,\|\cdot\|_2)$. Más concretamente, sea $V=C^n[a,b]$, $\|f\|_1 = \max_{x,k} |f^{(k)}(x)|$ y $\|f\|_2 = \sum_{k=0}^{n} \max_{x\in[a,b]} |f^{(k)}(x)|$.

Definición. 2.6.26 El conjunto de todas las transformaciones lineales acotadas de un espacio de Banach V a $\mathbb{R}$, $\mathcal{L}(V,\mathbb{R})$ con la norma $\|F\| = \sup_{f\neq 0} |F(f)|/\|f\|$, es denotado por $(V^*,\|\cdot\|)$ o por V^*, y es llamado el espacio dual de V. Los elementos de V^* son llamados funcionales lineales acotados sobre V. La topología en V^*, correspondiente a la norma introducida se llama topología fuerte en V^*.

Proposición. 2.6.27 El espacio dual $(V^*,\|\cdot\|)$ de un espacio lineal normado V, es un espacio de Banach. (La Dem. se sigue del Teo. 2.6.16).

Ejemplos. 2.6.28 Sea $V=C[a,b]$, sea g una función fija de $C[a,b]$ y pongamos para toda $f\in V$, $G(f)=\int_a^b f(x)g(x)dx$, vemos que G es lineal y acotada, entonces $G\in(V^*,\|\cdot\|)$.

Sea $V=C^n[a,b]$, sean $g_0,g_1,\ldots,g_n$ funciones fijas de $C[a,b]$ para f en V definimos $G(f)=\int_a^b g_0(x)f(x)dx+\int_a^b g_1(x)f'(x)dx+\cdots+\int_a^b g_n(x)f^{(n)}(x)dx$, vemos también que G es una funcional lineal y acotada, que por lo tanto $G\in(V^*,\|\cdot\|)$.

A modo de discusión y como antecedente para la demostración del Teorema de Representación de Riesz, consideremos los siguientes resultados:

Sea y una función definida sobre $[a,b]$. Si $P=\{s_0,\ldots,s_n\}$ es una partición de $[a,b]$, ponemos $\Delta_i y(s)=y(s_i)-y(s_{i-1})$, para $i=1,\ldots,n$. Si existe una constante positiva C, tal que $\sum_{i=1}^{n}|\Delta_i y(s)|\leq C$ para todas las particiones de $[a,b]$, entonces y es llamada de variación acotada sobre $[a,b]$.

Las funciones monótonas sobre $[a,b]$ son de variación acotada sobre $[a,b]$, puesto que si y es una función creciente $\sum_{i=1}^{n}|\Delta_i y(s)| = \sum_{i=1}^{n}\Delta_i y(s)=y(b)-y(a)$, para el caso en que y es una función decreciente basta considerar que $|\Delta_i y(s)|=-\Delta_i y(s)$.

Las funciones que pertenecen a $C^1[a,b]$, son también de variación acotada sobre $[a,b]$, puesto que para cada $y\in C^1[a,b]$, existe una constante positiva C, tal que $|y'(s)|\leq C$ para todo $s\in[a,b]$, y $\sum_{i=1}^{n}|\Delta_i y(s)|\leq \sum_{i=1}^{n}|y'(t_i)|\Delta s_i\leq C(b-a)$, donde $t_i\in[s_{i-1},s_i]$.

Si una función y es de variación acotada sobre $[a,b]$, entonces y se puede expresar como diferencia de dos funciones monótonas definidas en $[a,b]$. En efecto. Sea y de variación acotada sobre $[a,b]$, y consideremos las funciones:

$$r^+ = \begin{cases} r \text{ si } r>0. \\ 0 \text{ si } r\leq 0. \end{cases}$$

y $r^-=|r|-r^+$, definimos $G(s)=\sup(\sum_{i=1}^{n}[\Delta_i y(s)]^+)$ y $H(s)=\sup(\sum_{i=1}^{n}[\Delta_i y(s)]^-)$, donde el supremo es tomado sobre todas las particiones de $[a,s]$. Claramente G y H son funciones monótonas y como $y(s)-y(a)=G(s)-H(s)$, así $y(s)=G(s)-(H(s)-y(a))$, y puesto que también $H(s)-y(a)$ es monótona, se concluye la validez del enunciado anterior.

Si una función y es de variación acotada sobre $[a,b]$, entonces y es una integral indefinida.

En efecto. Supongamos que y es de variación acotada sobre $[a,b]$, entonces por la proposición anterior $y(s)=y_1(s)-y_2(s)$, donde las funciones y_i son monótonas crecientes, así $y'(s)=y_1'(s)-y_2'(s)$, de donde $y'(s)$ existe en casi todo punto de $[a,b]$, además $|y'(s)|\leq |y_1'(s)|+|y_2'(s)|=y_1'(s)+y_2'(s)$ y $\left|\int_a^b y'(s)ds\right|\leq\int_a^b|y'(s)|ds\leq y_1(b)+y_2(b)-y_1(a)-y_2(a)$, de donde $y'(s)$ es integrable. Pongamos $y(s)=\int_a^s y'(s)ds + y(a)$.

<u>Teorema</u>. 2.6.29 (Teorema de Representación de Riesz). Sea F una funcional lineal acotada definida en el espacio de funciones acotadas sobre el intervalo $[0,1]$ e integrables en el sentido de Riemann, con la norma $\|f\| = \max_{x\in[0,1]} |f(x)|$, al cual denotaremos por $BR[0,1]$. Entonces existe una función fija $g \in BR[0,1]$, tal que $F(f) = \int_0^1 f(x)g(x)dx$, $f \in BR[0,1]$ y $\|F\| = \int_0^1 |g(x)|dx$.

Demostración: Sea $C>0$ tal que $|F(f)| \leq C\|f\|$ para todo $f \in BR[0,1]$. Para cualquier $s \in [0,1]$, consideremos la función característica χ_s, entonces para cada $s \in [0,1]$ el valor $F(\chi_s)$ es un número real $y(s)$, y esto define una función sobre $[0,1]$. Veamos primero que la función y es de variación acotada. Sea $\{s_i\}_{i=0}^{n}$ una partición cualquiera de $[0,1]$. Entonces

$\sum_{i=1}^{n} |y(s_i) - y(s_{i-1})| = \sum_{i=1}^{n} (y(s_i) - y(s_{i-1}))sgn(y(s_i) - y(s_{i-1})) =$
$\sum_{i=1}^{n} (F(\chi_{si}) - F(\chi_{s(i-1)}))sgn(y(s_i) - y(s_{i-1})) =$
$F(\sum_{i=1}^{n} (\chi_{si} - \chi_{s(i-1)})sgn(y(s_i) - y(s_{i-1})) \leq$
$\|F\| \|\sum_{i=1}^{n} (\chi_{si} - \chi_{s(i-1)})\| \leq C \|\sum_{i=1}^{n} (\chi_{si} - \chi_{s(i-1)})\| = C$

lo que muestra que la función y es de variación acotada sobre $[0,1]$, entonces existe una función acotada g definida en $[0,1]$, tal que $y(s) = \int_0^s g(x)dx$, así $F(\chi_s) = \int_0^1 g(x)\chi_s(x)dx$, y como toda función simple es combinación lineal de funciones características. Si $z = \sum_{i=1}^{n} c_i \chi_{si}$, entonces $F(z) = \int_0^1 g(x)z(x)dx$. Sea f una función acotada sobre $[0,1]$ e integrable en el sentido de Riemann, entonces existe una sucesión $\{z_n\}_{n=1}^{\infty}$ de funciones simples definidas sobre $[0,1]$ tales que $z_n(x) \to f(x)$ cuando $n \to \infty$ para toda $x \in [0,1]$, y puesto que f es acotada en $[0,1]$, entonces la convergencia es uniforme, así $\|f - z_n\| = \sup_{x\in[0,1]} |f(x) - z_n(x)| \to 0$ cuando $n \to \infty$. Y $|F(f) - F(z_n)| = |F(f - z_n)| \leq \|F\|\|f - z_n\| \leq C \|f - z_n\| \to 0$ cuando $n \to \infty$, así $F(f) = \lim_{n\to\infty} F(z_n) = \lim_{n\to\infty} \int_0^1 g(x)z_n(x)dx$, entonces calculando

$|F(f) - \int_0^1 g(x)f(x)dx|$
$= |F(f) - F(z_n) + \int_0^1 g(x)z_n(x)dx - \int_0^1 g(x)f(x)dx|$
$\leq |F(f - z_n)| + \int_0^1 |g(x)|dx (\max_{x\in[0,1]} |f(x) - z_n(x)|)$
$\leq \|F\|\|f - z_n\| + \int_0^1 |g(x)| dx \|f - z_n\|$
$\leq C \|f - z_n\| + (\int_0^1 |g(x)| dx)\|f - z_n\| \to 0$ cuando $n \to \infty$. Entonces
$F(f) = \int_0^1 f(x)g(x)dx$ para $f \in BR[0,1]$.

Ahora veamos que $\|F\|=\int_0^1 |g(x)|dx$. Para todo $f\in BR[0,1]$, tenemos que $|F(f)|=\left|\int_0^1 g(x)f(x)dx\right|\leq \int_0^1 |g(x)||f(x)|dx\leq\left(\int_0^1 |g(x)|dx\right)\max_{x\in[0,1]}|f(x)|$

$=\left(\int_0^1 |g(x)|dx\right)\|f\|$, si $f\neq 0$, entonces $|F(f)|/\|f\|\leq$ $\int_0^1 |g(x)|dx$, así $\|F\|\leq\int_0^1 |g(x)|dx$.

Por otra parte, tomemos $f(x)=sgn(g(x))$, claramente esta función es acotada e integrable sobre $[0,1]$, entonces $|F(f)|=\int_0^1 g(x)sgn(g(x))dx=\left(\int_0^1 |g(x)|dx\right)\|f\|$, supuesto que $\|f\|=1$, así $\|F\|\geq|F(f)|/\|f\|=\int_0^1|g(x)|dx$, y por lo tanto $\|F\|=\int_0^1|g(x)|dx$.

<u>Definición</u>. 2.6.30 Una funcional no negativa p, definida sobre un espacio lineal V, es llamada convexa, si satisface para todo $f,g\in V$ y $c\in[0,1]$: $p(cf+(1-c)g)\leq cp(f)+(1-c)p(g)$.

<u>Ejemplo</u>. 2.6.31 Sea V el conjunto de las funciones acotadas sobre $[a,b]$ y sea x_0 un punto fijo en $[a,b]$. Entonces $p_{x_0}(f)=|f(x_0)|$ es una funcional convexa sobre V.

<u>Definición</u>. 2.6.32 Sea V un espacio lineal real, y sea V_0 un subespacio lineal suyo. Supongamos además, que sobre el subespacio V_0 se define una funcional lineal G. Una funcional lineal F definida sobre todo el espacio V, se llama extensión de la funcional G, cuando $F(f)=G(f)$ para todo $f\in V_0$.

<u>Lema</u>. 2.6.33 Sea V un espacio lineal real, y sea V_0 un subespacio lineal suyo. Supongamos además que G es una funcional lineal definida sobre V_0 la cual satisface $G(f)\leq p(f)$ donde p es una funcional convexa y $f\in V_0$. Entonces existe un número real a tal que:

$$\sup_{\substack{b>0\\ f\in V_0}}(1/b(-p(f-bg)+G(f)))\leq a\leq \inf_{\substack{b>0\\ f\in V_0}}(1/b(p(f+bg)-G(f))).$$

Supongamos que $f_1,f_2\in V_0$ y $b,c>0$. Entonces

$$cG(f_1)+bG(f_2)=G(cf_1+bf_2)=(c+b)G\left(\frac{c}{c+b}f_1+\frac{b}{c+b}f_2\right)\leq$$

$$(c+b)p\left(\frac{c}{c+b}f_1+\frac{b}{c+b}f_2\right)=(c+b)p\left(\frac{c}{c+b}(f_1-bg)+\frac{b}{c+b}(f_2+cg)\right)\leq$$

$$cp(f_1-bg)+bp(f_2+cg).$$

Así, para todo $f_1,f_2\in V_0$ y $b,c>0$, tenemos

$$1/b(-p(f_1-bg)+G(f_1))\leq 1/c(p(f_2+cg)-G(f_2)).$$

De donde existe un número a. tal que

$$\sup_{\substack{b>0\\ f\in V_0}}(1/b(-p(f-bg)+G(f)))\leq a\leq \inf_{\substack{b>0\\ f\in V_0}}(1/b(p(f+bg)-G(f))).$$

Teorema. 2.6.34 (Teorema de Hahn-Banach). Sea V un espacio lineal real y sea p una funcional convexa finita. Supongamos que G es una funcional lineal definida sobre un subespacio lineal V_0 de V, la cual satisface $G(f)\leq p(f)$ para todo $f\in V_0$. Entonces existe una funcional lineal F, definida sobre todo V, que satisface:
$F(f)\leq p(f)$ para todo $f\in V$ y $F(f)=G(f)$ para todo $f\in V_0$.
Demostración: Supongamos que $V_0\neq V$, así existe $g\in V$ tal que $g\notin V_0$. Sea V_g el espacio generado por V_0 y g. La extensión de G a V_g será llamada G_g y esta será especificada en cuanto definamos $G_g(g)$. puesto que todo elemento de V_g es de la forma $f+bg$ con $b\in R$ y $f\in V_0$, y $G_g(f+bg)=G(f)+bG_g(g)$. Ahora bien del lema anterior, definimos $G_g(g)=a$, y verifiquemos que la extensión resultante satisface que $G_g(f)\leq p(f)$ para todo $f\in V_g$, puesto que
$a\leq \inf_{\substack{b>0\\ f\in V_0}}(1/b(p(f+bg)-G(f)))\leq 1/b(p(f+bg)-G(f))$, implica para $f\in V_0$ y $b>0$, $G(f)+ba\leq p(f+bg)$, o bien $G_g(f+bg)=G(f)+bG_g(g)=G(f)+ba\leq p(f+bg)$ para todo $f+bg\in V_g$, además es obvio que $G_g(f)=G(f)$ para todo $f\in V_0$.
Ahora sea E la colección de extensiones e de G las cuales satisfacen $e(f)\leq p(f)$ sobre el subespacio donde están definidas. Definimos un orden parcial $<$ sobre E poniendo $e_1<e_2$ si e_2 esta definida sobre un conjunto más grande que e_1 y $e_1(f)=e_2(f)$ donde ambas están definidas. Sea $\{e_\alpha\}_{\alpha\in I}$ un subconjunto totalmente ordenado de E; sea V_α el subespacio donde e_α es definida. Definimos e sobre $\bigcup_{\alpha\in I}V_\alpha$ poniendo $e(f)=e_\alpha(f)$ si $f\in V_\alpha$, así para cada subconjunto totalmente ordenado de E, se tiene una cota superior. Por el lema de Zorn, E tiene un elemento maximal F, definido sobre algún conjunto V', el cual satisface $F(f)\leq p(f)$ para todo $f\in V'$. Pero V' debe ser todo V, ya que de otra manera F podría extenderse a F^* sobre un espacio más grande por una dimensión como se hizo antes, pero esto contradice que F es un elemento maximal, de donde tenemos que $V'=V$. Así la extensión F esta definida en todas partes.

Corolario. 2.6.35 Sea V un espacio lineal normado. V_o un subespacio de V y G un elemento de V_o^*. Entonces existe $F\in V^*$ una extensión de G, la cual satisface: $\|F\|_{V^*}=\|G\|_{V_o^*}$.

Demostración: Considerando el teorema anterior 2.6.34 y poniendo $|F(f)|\leq p(f)=\|G\|_{V_o^*}\|f\|_V$, en donde esta desigualdad es válida para todo $f\in V$, y si $f\neq 0$, entonces $|F(f)|/\|f\|_V\leq\|G\|_{V_o^*}$, así $\|F\|_{V^*}\leq\|G\|_{V_o^*}$. Por otro lado $\|F\|_{V^*}\geq|F(f)|/\|f\|=|G(f)|/\|f\|$ es válida para toda $f\in V_o$, $f\neq 0$, de donde $\|F\|_{V^*}\geq\|G\|_{V_o^*}$. Por lo tanto $\|F\|_{V^*}=\|G\|_{V_o^*}$.

Corolario. 2.6.36 Sea $g\neq 0$ un elemento de un espacio lineal normado V. Entonces existe una funcional no nula $F\in V^*$, tal que $F(g)=\|F\|_{V^*}\|g\|_V$.

Demostración: Sea V_o el subespacio generado por el elemento g, y definimos $G(ag)=a\|g\|$ con $a\in\mathbb{R}$, entonces haciendo uso del corolario anterior 2.6.35 podemos construir F con $\|F\|_{V^*}=\|G\|_{V_o^*}$ extendiendo G a todo V: Pero como $F(g)=G(g)=\|g\|$ en V_o se tiene que $\|F\|_{V^*}=\sup_{f\neq 0}|F(f)|/\|f\|\geq|F(g)|/\|g\|=1$, esto es $\|F\|_{V^*}\geq 1$, y dado que $\|F\|_{V^*}$ es la mínima constante C tal que $|F(h)|\leq C\|h\|$ para toda $h\in V$, entonces se tiene que $\|F\|_{V^*}=1$, puesto que $F(g)=\|g\|$ en V_o; Por lo tanto $F(g)=\|F\|_{V^*}\|g\|_V$.

Sea I una familia de índices, y supongamos que para cada $\alpha\in I$, V_α es un espacio de Banach. Sea $V=\{\{f_\alpha\}_{\alpha\in I}\ ;\ f_\alpha\in V_\alpha,\sum_{\alpha}\|f\|_{V_\alpha}<\infty\}$. Entonces el espacio V con la norma $\|\{f_\alpha\}_{\alpha\in I}\|=\sum_\alpha\|f\|_{V_\alpha}$ es un espacio de Banach. Y este espacio es llamado suma directa de los espacios V_α y es frecuentemente denotado por $V=\bigoplus_{\alpha\in I}V_\alpha$.

Sea V un espacio de Banach y M un subespacio cerrado de V. Si $f,g\in V$ definimos una relación de equivalencia entre elementos de V como sigue: $(f,g)\in R$ si y sólo si $f-g\in M$. Denotamos al conjunto de clases de equivalencia por V/M, y como antes definimos la adición y multiplicación por elementos de $\mathbb{R}$ de estas clases como sigue: $a[f]+b[g]=[af+bg]$. De esta manera este espacio es lineal, en donde $M=[0]$.

Definimos $\|[f]\|_{V/M}= \inf_{m\in M} \|f-m\|_V$. Podemos dar una interpretación un tanto geométrica de esta definición para dilucidar sobre ella, sabemos que en todo espacio lineal normado V podemos definir una métrica natural poniendo $d(f,m)= \|f-m\|_V$. Ahora bien, la distancia de un punto $f\in V$ a un conjunto M se puede definir como $d(f,M) = \inf_{m\in M} d(f,m)= \inf_{m\in M} \|f-m\|_V$. De esta manera $d(f,M)= \|[f]\|_{V/M}$. Es fácil ver que $\|[\cdot]\|$ es una norma: Por ejemplo $\|[f]\|_{V/M}=0$ implica que $[f]=M=[0]$. En efecto, sea $\{m_n\}_{n=1}^{\infty}$ una sucesión de elementos de V tal que $\|f-m_n\|_V\to 0$ cuando $n\to\infty$, entonces $\|m_n-m_m\|_V\leq \|m_n-f\|_V+\|f-m_m\|_V \to 0$ cuando $n,m\to\infty$, así $\{m_n\}_{n=1}^{\infty}$ es una sucesión de Cauchy en el espacio M, y puesto que M es un subespacio cerrado, entonces M es un subespacio completo de V, puesto que cada sucesión de Cauchy de M tiene un límite en V, pero por ser M cerrado, este límite debe necesariamente pertenecer a M, De donde, existe $m\in M$ tal que $m_n\to m$ cuando $n\to\infty$, así $\|f-m\|_V\leq \|f-m_n\|_V+\|m_n-m\|_V \to 0$ cuando $n\to\infty$; Por lo tanto $\|f-m\|_V=0$, esto implica que $f=m\in M$: Por lo tanto $[f]=M=[0]$.

La desigualdad del triángulo es probada como sigue: $\|[f]+[g]\|_{V/M}= \|[f+g]\|_{V/M}= \inf_{m\in M} \|(f+g)-m\|_V\leq \|f-\frac{1}{2}m\|_V+\|g-\frac{1}{2}m\|_V$, entonces $\inf_{m\in M} \|(f+g)-m\|_V$ es cota inferior del conjunto de números $\|f-\frac{1}{2}m\|_V+\|g-\frac{1}{2}m\|_V$ con $m\in M$, pero $\inf_{m\in M} \{\|f-\frac{1}{2}m\|_V+\|g-\frac{1}{2}m\|_V\}$ es la máxima cota inferior de este conjunto, entonces $\inf_{m\in M} \|(f+g)-m\|_V\leq \inf_{m\in M} \{\|f-\frac{1}{2}m\|_V+\|g-\frac{1}{2}m\|_V\}$ y por las propiedades de ínfimo tenemos que $\inf_{m\in M}\{\|f-\frac{1}{2}m\|_V+\|g-\frac{1}{2}m\|_V\}= \inf_{m\in M} \|f-\frac{1}{2}m\|_V + \inf_{m\in M} \|g-\frac{1}{2}m\|_V$, y dado que también $\frac{1}{2}m\in M$, tenemos que $\|[f]+[g]\|_{V/M}\leq \|[f]\|_{V/M} + \|[g]\|_{V/M}$. La propiedad $\|[af]\|_{V/M}=a\|[f]\|_{V/M}$ con $a\neq 0$ es obvia.

<u>Teorema</u>. 2.6.37 Sea V un espacio de Banach y sea M un subespacio cerrado suyo, entonces el espacio cociente de V según M, V/M, es un espacio de Banach.

Demostración: Sea $\{[f_n]\}_{n=1}^{\infty}$ una sucesión absolutamente sumable en el espacio cociente V/M. Esto significa que la serie:

$\sum_{n=1}^{\infty} \inf_{m\in M} \|f_n-m\|_V<\infty.$

Para cada n, escojamos $m_n \in M$ tal que $\|f_n-m_n\| \leq 2 \inf_{m\in M}\|f_n-m\|$, entonces $\{f_n-m_n\}_{n=1}^{\infty}$ es una sucesión absolutamente sumable en V. Sea $g=\lim_{N\to\infty}\sum_{n=1}^{N}(f_n-m_n)$, entonces $\|\sum_{n=1}^{N}[f_n]-[g]\|_{V/M} \leq \|\sum_{n=1}^{N} f_n-g-\sum_{n=1}^{N} m_n\| \to 0$ cuando $N\to\infty$. Esto prueba que $\{[f_n]\}_{n=1}^{\infty}$ es una sucesión sumable. Entonces por el teorema 2.6.18, tenemos que V/M es un espacio lineal normado completo.

Ejemplo. 2.6.38 Sea $V=C[a,b]$ y $M=\{f\ ;\ f(0)=0\}$. Entonces $V/M=\mathbb{R}$.

Definición. 2.6.39 Sean V_1 y V_2 espacios de Banach, T un operador lineal acotado de V_1 a V_2. El operador adjunto de T, denotado por T', es el operador lineal acotado de V_2^* a V_1^* definido por $T'(G(f))=G(T(f))$ para todo $G\in V_2^*$, y para todo $f\in V_1$.

Teorema. 2.6.40 Sean V_1 y V_2 dos espacios de Banach. La función $T \to T'$ es un isomorfismo isométrico de $\mathcal{L}(V_1,V_2)$ a $\mathcal{L}(V_2^*,V_1^*)$.

Demostración: La función $T\mapsto T'$ es lineal, puesto que $G(aT_1+T_2)(f)=(aGT_1+GT_2)(f)=aGT_1(f)+GT_2(f)=aT_1'G(f)+T_2'G(f)=(aT_1'G+T_2'G)(f)=(aT_1'+T_2')G(f)$. EL hecho de que T' es acotado y que la función es una isometría se sigue de:

$$\|T\|_{\mathcal{L}(V_1,V_2)} = \sup_{\|f\|\leq 1}\|T(f)\|_{V_2} = \sup_{\|f\|\leq 1}\Big(\sup_{\substack{\|G\|\leq 1\\ G\in V_2^*}}|G(T(f))|\Big)$$
$$= \sup_{\substack{\|G\|\leq 1\\ G\in V_2^*}}\Big(\sup_{\|f\|\leq 1}|T'(G(f))|\Big) = \sup_{\substack{\|G\|\leq 1\\ G\in V_2^*}}\|T'G\| = \|T'\|_{\mathcal{L}(V_2^*,V_1^*)}.$$

La segunda igualdad se obtiene de corolario al teorema de Hahn-Banach 2.6.36. Haciendo $|G(T(f))|=\|G\|_{V_2^*}\|T(f)\|_{V_2}$ donde $G\in V_2^*$ es no nula, entonces $\|T(f)\|_{V_2}=|G(T(f))|/\|G\|$, así $\sup_{\|f\|\leq 1}\|T(f)\|_{V_2}$ $= \sup_{\|f\|\leq 1}\Big(\sup_{\substack{\|G\|\leq 1\\ G\in V_2^*}}|G(T(f))|\Big)$.

Un espacio lineal normado numerable es un espacio lineal V provisto de un sistema numerable de normas $\|\cdot\|_n$ equivalentes entre sí. Todo espacio lineal numerable se convierte en un espacio lineal topológico, si definimos un sistema de vecindades abiertas de cero $N_{n,\epsilon}$, cada una de las cuales esta definida por los números n y ϵ, y consta de los $f\in V$ tales que $\|f\|_1<\epsilon$, $\|f\|_2<\epsilon,\ldots,\|f\|_n<\epsilon$. Vemos también que un espacio lineal normado numerable satisface el primer axioma de numerabilidad haciendo que ϵ tome los valores $1,\frac{1}{2},\ldots,1/n$. Es más la topología puede ser introducida por una métrica invariante bajo traslaciones $d(f,g)=\sum_{n=1}^{\infty} 2^{-n}(\|f-g\|_n/(1+\|f-g\|_n))$, $f,g\in V$. Como ejemplo tomamos a $C^{\infty}[a,b]$ con las normas $\|f\|_n=\max_{\substack{x\in[a,b]\\ 0\leq k\leq n}}\{|f^{(k)}(x)|\}$, $n=0,1,2\ldots$

Teorema. 2.6.41 (Teorema de Banach sobre el operador inverso). Sea U un operador lineal acotado, que efectúa una transformación inyectiva de un espacio de Banach V_1 sobre un espacio de Banach V_2. Entonces el operador inverso U^{-1} es un operador acotado.

Demostración: Observamos primero que, U^{-1} es lineal, en efecto, basta comprobar que se cumple la igualdad $U^{-1}(af+bg)=aU^{-1}(f)+bU^{-1}(g)$. Para esto, pongamos $U(f_1)=f$ y $U(g_1)=g$. Debido a la linealidad de U tenemos

$U(af_1+bg_1)=aU(f_1)+bU(g_1)$,

de acuerdo con la definición de operador inverso $U^{-1}(f)=f_1$ y $U^{-1}(g)=g_1$, de donde

$aU^{-1}(f)+bU^{-1}(g)=af_1+bg_1=U^{-1}(U(af_1+bg_1))=U^{-1}(aU(f_1)+bU(g_1))=U^{-1}(af+bg)$.

Notamos también, que si M es un subconjunto denso de V_2. Entonces todo elemento no nulo de V_2, puede expresarse como:

$f=f_1+f_2+\dots+f_n+\dots$, donde $f_n\in M$ y $\|f_n\|\leq(3/2^n)\|f\|$.

Para esto, construyamos la sucesión $\{f_n\}_{n=1}^{\infty}$ como sigue: escogemos $f_1\in M$, de tal manera que $\|f-f_1\|<\frac{1}{2}\|f\|$, lo cual es siempre posible ya que esta desigualdad define una bola abierta de radio $\frac{1}{2}\|f\|$ y centro en el punto f; $B(f,\frac{1}{2}\|f\|)$, dentro de la cual debe existir un elemento de M, puesto que M es un subconjunto denso de V_2, escojamos $f_2\in M$, de manera que $\|f-f_1-f_2\|\leq(\frac{1}{2})^2\|f\|$, en general escojamos f_n, tal que $\|f-f_1-f_2-\dots-f_n\|\leq(\frac{1}{2})^n\|f\|$. Tal selección es posible, ya que M es denso en V_2, de acuerdo con esta selección de elementos de M, tenemos

$\|f-\sum_{n=1}^{\infty} f_n\| \to 0$ cuando $n\to\infty$,

esto es, la serie $\sum_{n=1}^{\infty} f_n$ converge hacia f. Estimemos ahora, las normas de los elementos f_n:

$\|f_1\|=\|f-f_1+f\|\leq\|f-f_1\|+\|f\|\leq\frac{1}{2}\|f\|+\|f\|=(3/2)\|f\|$,

$\|f_2\|=\|f_2+f_1-f+f-f_1\|\leq\|f-f_1-f_2\|+\|f-f_1\|\leq\frac{1}{4}\|f\|+\frac{1}{2}\|f\|=(3/4)\|f\|$.

Finalmente

$$\|f_n\|=\|f_n+f_{n-1}+\dots+f_1-f+f-f_1-f_2-\dots-f_{n-1}\|$$
$$\leq\|f-f_1-\dots-f_n\|+\|f-f_1-\dots-f_{n-1}\|\leq(\tfrac{1}{2})^n\|f\|+(\tfrac{1}{2})^{n-1}\|f\|=(3/2^n)\|f\|.$$

Ahora, consideremos en el espacio V_2 los conjuntos M_k, donde M_k consiste de elementos $f\in V_2$, tales que satisfacen la desigualdad $\|U^{-1}f\|\leq k\|f\|$. Todo elemento del espacio V_2, se encuentra en cierto M_k, es decir $V_2=\bigcup_{k=1}^{\infty} M_k$. De acuerdo al Teorema de Baire sobre categorías. Todo espacio métrico completo, no puede representarse como la unión numerable de conjuntos nunca densos, es decir, de conjuntos que no son densos en ninguna bola. Luego, al menos uno de los conjuntos M_k, digamos M_n, es denso en una bola abierta $B(f_o,\epsilon)$. consideremos dentro de esta bola el conjunto S de puntos g, para los cuales se cumple la desigualdad $0<b<\|g-f_o\|<a<\epsilon$, donde $f_o\in M_n$. Hagamos una traslación de $B(f_o,\epsilon)$ al origen, y denotemos al conjunto resultante de la traslación de S, como S_o. Probemos que en S_o existe un conjunto denso M_N. Sea $g\in S\cap M_n$; entonces $g-f_o\in S_o$ y se cumple que

$$\|U^{-1}(g-f_o)\|\leq n(\|g-f_o\|)\leq n(\|g\|-\|f_o\|)$$
$$=n(\|g-f_o+f_o\|+\|f_o\|)\leq n(\|g-f_o\|+2\|f_o\|)=n\|g-f_o\|(1+(2\|f_o\|/\|g-f_o\|)).$$
$$\leq n(\|g-f_o\|(1+(2\|f_o\|/b)).$$

La magnitud $n(1+(2\|f_o\|/b))$ no depende de g. Tomemos $N=1+(1+(2\|f_o\|/b))$ Entonces $g-f_o\in S_o$ y, como M_n es denso en s, M_N será denso en S_o con $\|U^{-1}(g-f_o)\|\leq N\|g-f_o\|$. Consideremos un elemento f no nulo en V_2, entonces siempre se puede escoger $c\in R$, tal que $b<\|cf\|<a$, es decir que $cf\in S_o$. Puesto que M_N es denso en S_o, se puede construir una sucesión de elementos $f_k\in M_N$, convergente hacia cf. Entonces la sucesión $\{(1/c)f_k\}_{k=1}^{\infty}$ converge hacia f. Es evidente que si $f_k\in M_N$, entonces también $(1/c)f_k\in M_N$, para todo $c\neq 0$, $c\in R$, puesto que

$$\|U^{-1}(f_k)\|\leq N\|f_k\| \quad \text{implica} \quad \|U^{-1}((1/c)f_k)\|\leq N\|(1/c)f_k\| \quad \text{para } c\neq 0.$$

Por consiguiente, M_N es denso en $V_2-\{0\}$ y, por esto, también lo es en V_2.

Consideremos un elemento no nulo $f\in V_2$; entonces f puede expresarse como: $f=f_1+f_2+\ldots+f_k+\ldots$, siendo $\|f_k\|\leq(3/2^k)\|f\|$, $f_k\in M_N$.
Consideremos en V_1, la serie formada por las imágenes recíprocas de los elementos f_k, es decir, de los elementos $g_k=U^{-1}(f_k)$. Esta serie converge a un elemento g, ya que tiene lugar la desigualdad

$\|g_k\|=\|U^{-1}(f_k)\|\leq N\|f_k\|=N(3/2^k)\|f\|$, para toda k, y si denotamos las sumas parciales de $\sum_{k=1} g_k$ por s_n, entonces para $n>m$, tenemos

$$\|s_n-s_m\|=\|\sum_{k=m+1}^{n} g_k\|\leq\sum_{k=m+1}^{n}\|g_k\|\leq 3N\|f\|\sum_{k=m+1}^{n}(\tfrac{1}{2})^k$$
$$\leq 3N\|f\|(\tfrac{1}{2})^m\sum_{k=1}^{n}(\tfrac{1}{2})^k \to 0 \text{ cuando } n,m\to\infty,$$

Entonces la sucesión $\{s_n\}_{n=1}^{\infty}$ de sumas parciales de la serie $\sum_{k=1}^{n} g_k$, es una sucesión de Cauchy, por tanto la serie $\sum_{k=1}^{\infty} g_k$ converge a un elemento $g\in V_1$, además

$$\|g\|\leq\sum_{k=1}^{\infty}\|g_k\|\leq 3N\|f\|\sum_{k=1}^{\infty}(\tfrac{1}{2})^k=3N\|f\|.$$

Como la serie $\sum_{k=1}^{\infty} g_k$ converge a un elemento $g\in V_1$, y U es un operador continuo, se tiene que U se puede aplicar a esta serie término a término, ya que al denotar por $s_n=\sum_{k=1}^{n} g_k$ a la n-ésima suma parcial de la serie $\sum_{k=1}^{\infty} g_k$, tenemos que si $\lim_{n\to\infty} s_n=g$. y U es un operador continuo

$$\lim_{n\to\infty}(U(s_n))=U(\lim_{n\to\infty} s_n),$$

entonces

$$U(g)=Ug_1+Ug_2+\ldots=f_1+f_2+\ldots$$

de donde $g=U^{-1}(f)$. Además

$$\|U^{-1}(f)\|=\|g\|\leq 3N\|f\|,$$

y como esta es una estimación válida para cualquier $f\neq 0$, el operador lineal U^{-1}, es un operador lineal acotado. Lo que concluye la demostración del Teorema.

A continuación hacemos un breve estudio, sobre álgebras de Banach, como una extensión de espacios de Banach, es decir, estructuras que además de ser espacios de Banach, satisfacen las siguientes propiedades: $\|fg\|\leq\|f\|\|g\|$ para todo f,g y la existencia de un elemento unidad, tal que, su norma es igual a uno. De igual manera extendemos el concepto de homomorfismos a homomorfismos de carácter multiplicativo.

Definición. 2.6.42 Si V es un álgebra de acuerdo con la def. 2.3.1, y si en adición V es un espacio de Banach, cuya norma satisface:
(i). $\|fg\| \leq \|f\|\|g\|$ para todo $f,g\in V$.
(ii). El elemento unidad e de V es tal que $\|e\| = 1$.
Entonces V es llamada un álgebra de Banach.
Notamos que no se requiere que V sea conmutativo con respecto al producto.

Ejemplo. 2.6.43 El espacio $C[a,b]$ con las operaciones usuales entre funciones y con la norma $\|f\| = \max_{x\in[a,b]}|f(x)|$, es un álgebra de Banach.

Ejemplo. 2.6.44 Sea V un espacio de Banach, entonces el álgebra $\mathcal{L}(V,V)$ de todos los operadores lineales acotados sobre V con la norma operador es un álgebra de Banach.

Proposición. 2.6.45 Sea V un álgebra de Banach, entonces la multiplicación es una operación continua.
Demostración: Supongamos que $f_n \to f$ y $g_n \to g$ cuando $n \to \infty$, entonces $f_ng_n - fg = (f_n - f)g_n + f(g_n - g) \to 0$ cuando $n \to \infty$. Equivalentemente $f_ng_n \to fg$ cuando $n \to \infty$ siempre que $f_n \to f$ y $g_n \to g$ cuando $n \to \infty$.

Definición. 2.6.46 Una función F de un álgebra de Banach V_1 a un álgebra de Banach V_2, es llamada un homomorfismo si:
(i). $F(af+bg)=aF(f)+bF(g)$ para todo $f,g\in V_1$ y $a,b\in\mathbb{R}$.
(ii). $F(fg)=F(f)F(g)$ para todo $f,g\in V_1$.

Definición. 2.6.47 Sea V un álgebra, sea $f\in V$, el elemento f es llamado invertible si este tiene un inverso multiplicativo en V, esto es, existe un elemento $f^{-1}\in V$ tal que $f^{-1}f=ff^{-1}=e$.

Proposición. 2.6.48 Sea V un álgebra, sea $f\in V$, si f es invertible, entonces el inverso es único.

Proposición. 2.6.49 Si F es un homomorfismo sobre un álgebra V. Entonces $F(e)=1$, y $F(f)\neq 0$ para todo f invertible en V.
Demostración: Sea $g\in V$ tal que $F(g)\neq 0$. Puesto que $F(g)=F(ge)=F(g)F(e)$ se sigue que $F(e)=1$.
Si f es invertible, entonces $F(f)F(f^{-1})=F(ff^{-1})=F(e)=1$, así que $F(f)\neq 0$.

Teorema. 2.6.50 Supongamos que V es un álgebra de Banach, $f \in V$, $\| f \| < 1$. Entonces

(i). e-f es invertible.

(ii). $\| (e-f)^{-1} - e - f \| \leq \| f \|^2/(1-\|f\|)$.

(iii). $|F(f)| < 1$ para todo homomorfismo F sobre V.

Demostración: (i). Puesto que $\|f^n\| \leq \|f\|^n < 1$, los elementos $g_n = e+f+f^2+\cdots+f^n$ forman una sucesión de Cauchy en V, en efecto supongamos que $n > m$, entonces $\| g_n - g_m \| = \| f^{m+1} + \cdots + f^n \| \leq \| f^{m+1} \| + \cdots + \| f^n \| \leq \|f\|^{m+1} + \cdots + \|f\|^n \to 0$ cuando $n,m \to \infty$.

Y puesto que V es completo, entonces existe $g \in V$ tal que $g_n \to g$ cuando $n \to \infty$, y como además $f^n \to 0$ cuando $n \to \infty$ y $g_n(e-f) = e - f^{n+1} = (e-f)g_n$, entonces la continuidad de la multiplicación implica que g es inverso de e-f, puesto que $g_n(e-f) = (e-f)g_n \to e$ cuando $n \to \infty$.

(ii). Vemos que $\| g-e-f \| = \| f^2 + f^3 + \cdots \| \leq \sum_{n=2}^{\infty} \| f \|^n = \| f \|^2/(1 - \|f\|)$ pero $g = (e-f)^{-1}$.

(iii). Sea $a \in \mathbb{R}$ tal que $|a| \geq 1$, entonces por (i) $e - a^{-1}f$ es invertible y por la proposición 2.6.48 tenemos que $1 - a^{-1}F(f) = F(e) - a^{-1}F(f) = F(e - a^{-1}f) \neq 0$. Por lo tanto $F(f) \neq a$, entonces $|F(f)| < 1$ para toda $f \in V$ tal que $\| f \| \leq 1$ y para todo F un homomorfismo sobre V.

2.7 ESPACIOS CON UN PRODUCTO INTERIOR.

Definición. 2.7.1 Un espacio lineal V sobre el campo $\mathbb{R}$, es llamado un espacio con un producto interior, si existe una función $(\cdot,\cdot)$: $V \times V \to \mathbb{R}$ que satisface las siguientes condiciones para todo $f,g,h \in V$ y $a \in \mathbb{R}$:

(i). $(f,f) \geq 0$ y $(f,f) = 0$ si y sólo si $f = 0$.

(ii). $(f,g+h) = (f,g) + (f,h)$.

(iii). $(f,ag) = a(f,g)$.

(iv). $(f,g) = (g,f)$.

Notamos que (ii), (iii) y (iv) implican que $(f,ag+bh) = a(f,g) + b(f,h)$, $(af,g) = a(f,g)$. La función $(\cdot,\cdot)$ es llamada un producto interior para V.

Ejemplo. 2.7.2 Sea $V = C[a,b]$ para $f,g \in V$ definimos
$(f,g) = \int_a^b f(x)g(x)dx$.

Ejemplo. 2.7.3 Sea $V = C^n[a,b]$ para $f,g \in V$ definimos
$(f,g) = \int_a^b f(x)g(x)dx + \int_a^b f'(x)g'(x)dx + \cdots + \int_a^b f^{(n)}(x)g^{(n)}(x)dx$.

Definición. 2.7.4 Dos elementos f y g de un espacio con un producto interior V, son llamados ortogonales, si $(f,g)=0$. Una colección $\{f_i\}$ de elementos de V es llamada ortonormal, si $(f_i,f_j)=\delta_{ij}$.

Por el momento denotaremos $\|f\|=((f,f))^{1/2}$, veremos más adelante que en efecto $\|\cdot\|$ es una norma.
Con lo que se concluye que toda funcional lineal acotada sobre $C[a,b]$, se puede expresar como $(\cdot,g)$ con una única $g\in C[a,b]$ al considerar el teorema 2.6.29 y el ejemplo 2.7.2.

La afirmación anterior se formalizará con el Lema de Riesz 2.7.15

Teorema. 2.7.5 (Teorema de Pitágoras). Sea $\{f_n\}_{n=1}^{N}$ un conjunto ortonormal de un espacio con un producto interior V. Entonces para todo $f\in V$, tenemos que
$\|f\|^2=\sum_{n=1}^{N}|(f_n,f)|^2+\|f-\sum_{n=1}^{N}(f_n,f)f_n\|^2$.
La demostración se sigue de que si $f=\sum_{n=1}^{N}(f_n,f)f_n+(f-\sum_{n=1}^{N}(f_n,f)f_n)$. Entonces $(\sum_{n=1}^{N}(f_n,f)f_n,f-\sum_{n=1}^{N}(f_n,f)f_n)=\sum_{n=1}^{N}(f_n,f)^2-\sum_{n=1}^{N}(f_n,f)^2=0$.

Corolario. 2.7.6 (Desigualdad de Bessel). Sea $\{f_n\}_{n=1}^{N}$ un conjunto ortonormal de un espacio con producto interior V. Entonces para toda $f\in V$.
$\|f\|^2\geq\sum_{n=1}^{N}|(f_n,f)|^2$.

Corolario. 2.7. 7 (Desigualdad de Schwarz-Buniakovski) Sea V un espacio con producto interior, entonces $|(f,g)|\leq\|f\|\|g\|$ para todo f,g V.
Demostración: Si $g=0$ la prueba es trivial, en otro caso el vector $g/\|g\|$ forma él mismo un conjunto ortonormal, así haciendo uso del corolario 2.7.6, tenemos que
$\|f\|^2\geq|(f,g/\|g\|)|^2=(f,g)^2/\|g\|^2$.

Corolario. 2.7.8 Todo espacio V con un producto interior, es un espacio lineal normado con $\|f\|=((f,f))^{1/2}$.
Demostración: Sólo basta ver que la desigualdad del triángulo se sigue de la desigualdad de Schwarz.
Este corolario muestra que se tiene una métrica natural $d(f,g)=((f-g,f-g))^{1/2}$. De esta manera, tenemos la nociones de convergencia, completación y densidad definidas para espacios métricos.

Observamos que en un espacio lineal con un producto interior, V, la función $(\cdot,\cdot)$ producto interior es continua. En efecto, supongamos que $f_n \to f$ y $g_n \to g$ cuando $n \to \infty$ en el sentido de convergencia en norma. Primero vemos que si $f_n \to f$, entonces para $\epsilon=1$ existe N tal que para $n>N$ $\|f_n\|-\|f\| \leq \|f_n-f\| < 1$, así $\|f_n\|<1+\|f\|$ es válida para $n>N$. Sea $M=\max\{\|f_1\|,\|f_2\|,\ldots,\|f_N\|,1+\|f\|\}$ luego $\|f_n\|\leq M$ para toda n. Ahora bien $|(f_n,g_n)-(f,g)| \leq |(f_n,g_n)+(f_n,-g)| + |(f_n,g)+(-f,g)| = |(f_n,g_n-g)| + |(f_n-f,g)| \leq \|f_n\|\|g_n-g\|+\|f_n-f\|\|g\| \leq M\|g_n-g\|+\|f_n-f\|\|g\| \to 0$ cuando $n \to \infty$, la segunda desigualdad es la de Schwarz-Buniakovski.

<u>Definición</u>. 2.7.9 Un espacio con producto interior que es completo en la métrica inducida por $\|f-g\|=((f-g,f-g))^{1/2}$, es llamado de Hilbert.

<u>Definición</u>. 2.7.10 Dos espacios de Hilbert H_1 y H_2, son llamados isomorfos, si existe un operador $U:H_1 \to H_2$ tal que $(U(f),U(g))_{H_2} = (f,g)_{H_1}$ para todo $f,g\in H_1$, el operador U es llamado unitario.

Sea $(C[a,b])^2=H_1=H_2$ y sea T una matriz ortogonal de 2X2, digamos, T tiene vectores fila $(\cos\theta,\operatorname{sen}\theta)$ y $(-\operatorname{sen}\theta,\cos\theta)$, así $T^t=T^{-1}$. Si $f=(f_1,f_2),g=(g_1,g_2)\in H_1$, definimos $(f,g)=f^tTg$, entonces $(f,g)=f^tTg=f^t(T^{-1}T)Tg=(f^tT^t)T(Tg)=(Tf)^tT(Tg)=(Tf,Tg)$, T es unitario.

<u>Proposición</u>. 2.7.11 (Ley del paralelogramo). Si $f,g\in V$, V un espacio con producto interior. Entonces
$\|f+g\|^2+\|f-g\|^2=2\|f\|^2+2\|g\|^2$.

Además, si en un espacio lineal normado, V, se satisface esta igualdad para todo $f,g\in V$, entonces V es también un espacio lineal con un producto interior, con $(f,g)=\frac{1}{4}(\|f+g\|^2-\|f-g\|^2)$.

Veamos que esta condición es suficiente. Para f=g tenemos $(f,f)=\frac{1}{4}(\|2f\|^2-\|f-f\|^2)=\|f\|^2$, este será precisamente aquel producto escalar que induce en el espacio lineal normado V la norma definida en él. Ante todo, se ve inmediatamente que $(f,g)=(g,f)$. Ahora bien, consideremos $(f+g,h)-(f,h)-(g,h)$ y veamos que esta relación es idénticamente cero, esto es, $(f+g,h)-(f,h)-(g,h)=0$ para todo $f,g,h\in V$. Primero observamos que $4((f+g,h)-(f,h)-(g,h))=\|f+g+h\|^2-\|f+g-h\|^2-\|f+h\|^2+\|f-h\|^2-\|g+h\|^2+\|g-h\|^2$. De acuerdo con la ley del paralelogramo tenemos que $\|f+g+h\|^2=2\|f+g\|^2+2\|g\|^2-\|f+h-g\|^2$ y $\|f+g-h\|^2=2\|f-h\|^2+2\|g\|^2-\|f-h-g\|^2$. Sustituyendo estas expresiones en $4((f+g,h)-(f,h)-(g,h))=\|f+g+h\|^2-\|f+g-h\|^2-\|f+h\|^2+\|f-h\|^2-\|g+h\|^2+\|g-h\|^2$, tenemos que $4((f+g,h)-(f,h)-(g,h))= -\|f+h-g\|^2+\|f-h-g\|^2+\|f+h\|^2-\|f-h\|^2-\|g+h\|^2+\|g-h\|^2$.

Luego tomando la suma media de las dos relaciones que se tienen para $4((f+g,h)-(f,h)-(g,h))$, tenemos que $4((f+g,h)-(f,h)-(g,h))=\frac{1}{2}(\|g+h+f\|^2+\|g+h-f\|^2)-\frac{1}{2}(\|g-h+f\|^2+\|g-h-f\|^2)-\|g+h\|^2+\|g-h\|^2=\|g+h\|^2+\|f\|^2-\|g-h\|^2-\|f\|^2-\|g+h\|^2+\|g-h\|^2=0$ para todo $f,g,h\in V$.

Consideremos ahora, para cualesquiera f y g fijos la relación $(cf,g)-c(f,g)$, con $c\in\mathbb{R}$, y veamos que esta relación es idénticamente cero, esto es, $(cf,g)-c(f,g)=0$ para todo $c\in\mathbb{R}$, observamos que $(-f,g)=-(f,g)$. Por eso para cualquier entero n, tenemos que $(nf,g)=(\operatorname{sgn}(n)(f+f+\cdots+f),g)=\operatorname{sgn}(n)((f,g)+(f,g)+\cdots+(f,g))=n\operatorname{sgn}(n)(f,g)=n(f,g)$. Luego para m y n enteros con $n\neq 0$, se tiene que, $((m/n)f,g)=m((1/n)f,g)=(m/n)n((1/n)f,g)=(m/n)(f,g)$. Pero además $(cf,g)-c(f,g)$ es una función continua de c. Por tanto si $((m/n)f,g)-(m/n)(f,g)=0$, entonces $(cf,g)-c(f,g)=0$ para toda $c\in\mathbb{R}$.

<u>Lema</u>. 2.7.12 Sea H un espacio de Hilbert. M un subespacio cerrado de H, y supongamos que $f\in H$. Entonces existe en M un único elemento $h\in M$ tal que $\inf_{g\in M}\|f-g\|=\|f-h\|$.

Demostración: Sea $d=\inf_{g\in M}\|f-g\|$. Sea $\{g_n\}_{n=1}^{\infty}$ una sucesión de elementos de M tal que $\|f-g_n\|\to d$ cuando $n\to\infty$. Entonces

$$\begin{aligned}\|g_n-g_m\|^2 &= \|(g_n-f)-(g_m-f)\|^2\\ &=2\|g_n-f\|^2+2\|g_m-f\|^2-\|-2f+g_n+g_m\|^2\\ &=2\|g_n-f\|^2+2\|g_m-f\|^2-4\|f-(1/2)(g_n-g_m)\|^2\\ &\leq 2\|g_n-f\|^2+2\|g_m-f\|^2-4d^2\\ &\to 2d^2+2d^2-4d^2=0 \quad \text{cuando } m,n\to\infty.\end{aligned}$$

La segunda igualdad se sigue de la Ley del paralelogramo (proposición anterior); la desigualdad se sigue del hecho de que $1/2(g_n+g_m)\in M$, así $\{g_n\}_{n=1}^{\infty}$ es una sucesión de Cauchy y ya que M es cerrado, $\{g_n\}_{n=1}^{\infty}$ converge a un elemento $h\in M$. Así $d=\|f-h\|$

puesto que $\|f-g_n\| \to d$ y $\|g_n-h\| \to 0$ cuando $n \to \infty$, entonces se tiene que $\|f-h\| \leq \|f-g_n\| + \|g_n-h\| \to d$ cuando $n \to \infty$. La unicidad se obtiene como sigue, supongamos que $g_n \to h$ y $g_n \to h'$ cuando $n \to \infty$. Entonces $\|h-h'\| \leq \|h-g_n\| + \|g_n-h'\| \to 0$ cuando $n \to \infty$. Por lo tanto $h=h'$.

Teorema. 2.7.13 (Teorema de Proyección). Sea H un espacio de Hilbert. M un subespacio cerrado. Entonces todo $f \in H$ puede expresarse en forma única como $f=g+h$ con $h \in M$ y $g \in M^{\perp} = \{w \ ; \ (w,v)=0, \ v \in M\}$.
Demostración: Sea $f \in H$, entonces por el lema anterior, 2.7.12, existe un único $h \in M$ tal que $\inf_{w \in M} \|f-w\| = \|f-h\|$. Definimos $g=f-h$, entonces $f=g+h$. Sea $v \in M$ y $t \in \mathbb{R}$. Si $d = \|f-h\| = \inf_{w \in M} \|f-w\|$, entonces para toda $t \in \mathbb{R}$, tenemos que $d^2 \leq \|f-(h+tv)\|^2 = \|(f-h)+tv\|^2 = \|g-tv\|^2 = \|g\|^2 - 2(g,tv) + t^2\|v\|^2 = d^2 - 2t(g,v) + t^2\|v\|^2$, lo cual implica que $-2t(g,v)+t^2\|v\|^2 \geq 0$ para toda $t \in \mathbb{R}$, así $-2t(g,v)=0$, y esto implica que $g \in M^{\perp}$. La unicidad se sigue de que, si $f=g+h$ y $f=g'+h'$, entonces $h'-h=g-g'$, entonces $h'-h \in M$, $h'-h \in M^{\perp}$ y $g-g' \in M$, $g-g' \in M^{\perp}$. Pero claramente $M \cap M^{\perp} = \{0\}$, de donde $h'-h=0$ y $g-g'=0$ o equivalentemente $h'=h$ y $g=g'$.

Denotaremos por $\mathcal{L}(H_1,H_2)$ el conjunto de transformaciones lineales de un espacio de Hilbert H_1 a otro H_2.

Definición. 2.7.14 El espacio $\mathcal{L}(H,\mathbb{R})$ es llamado el espacio dual del espacio de Hilbert H, y es denotado por H^*. Los elementos de H^* son llamados funcionales lineales continuos.

Teorema. 2.7.15 (Lema de Riesz). Para cada $T \in H^*$, existe un único $g_T \in H$, tal que $T(f)=(g_T,f)$ para todo $f \in H$, y $\|g_T\|_H = \|T\|_{H^*}$.
Demostración: Sea N el conjunto de los $f \in H$, tales que $T(f)=0$, entonces por la continuidad de T, N es un subespacio cerrado. Si $N=H$. Entonces $T(f)=0=(0,f)$ para toda $f \in H$. Ahora supongamos que N no es todo H. Entonces por el teorema de proyección, 2,7,13 existe un $f_0 \neq 0$, $f_0 \in N^{\perp}$. Definimos $g_T = T(f_0)\|f_0\|^{-2} f_0$. Vamos a verificar que g_T tiene las propiedades mencionadas arriba.
Sea $f \in N$, entonces $T(f)=0=(g_T,f)$, además si $f=af_0$, entonces

$T(f)=T(af_o)=aT(f_o)=aT(f_o)(\|f_o\|^{-2}f_o,f_o)=(T(f_o)\|f_o\|^{-2}f_o,af_o)=(g_T,af_o)$.

Puesto que las funciones $T(\cdot)$ y $(g_T,\cdot)$ son lineales y coinciden sobre N y f_o, ellas deben coincidir sobre el espacio generado por N y f_o. Pero N y f_o generan H, puesto que todo elemento $g\in H$ puede ser expresado como $g=(g-(T(g)/T(f_o))f_o)+(T(g)/T(f_o))f_o$. Así $T(f)=(g_T,f)$ para toda $f\in H$. Para probar la unicidad supongamos que $T(f)=(g',f)$ para toda $f\in H$, entonces $\|g'-g_T\|^2=(g'-g_T,g'-g_T)=(g',g'-g_T)-(g_T,g'-g_T)=T(g'-g_T)-T(g'-g_T)=0$, así $g'=g_T$.

Para probar que $\|T\|_{H^*}=\|g_T\|_H$, observamos que

$\|T\|=\sup\limits_{\|f\|\le 1}|T(f)|=\sup\limits_{\|f\|\le 1}|(g_T,f)|\le\sup\limits_{\|f\|\le 1}\|g_T\|\,\|f\|=\|g_T\|$.

$\|T\|=\sup\limits_{\|f\|\le 1}|T(f)|\ge|T(g_T/\|g_T\|)|=(g_T,g_T/\|g_T\|)=\|g_T\|$.

De donde $\|T\|_{H^*}=\|g_T\|_H$.

<u>Corolario</u>. 2.7.16 Sea $B(\cdot,\cdot)$ una función de $H\times H$ en $\mathbb{R}$, H un espacio de Hilbert. Si esta función satisface para todo $f,g,h\in H$ y $a,b\in\mathbb{R}$

(i). $B(f,ag+bh)=aB(f,g)+bB(f,h)$.

(ii). $B(af+bg,h)=aB(f,h)+bB(g,h)$.

(iii). $|B(f,g)|\le C\|f\|\|g\|$.

Entonces existe una única transformación lineal acotada T de H a H, tal que $B(f,g)=(T(f),g)$ para todo $f,g\in H$.

La norma de T es la mínima constante C tal que (iii) es válida.

Demostración: Fijamos $f\in H$. Entonces $B(f,\cdot)$ es una funcional lineal continua, entonces por el teorema anterior, 2.7.15, existe $h\in H$ tal que $B(f,g)=(h,g)$ para toda $g\in H$. Definimos $T(f)=h$. Así para $f,g\ne 0$, $|B(f,g)|=|(T(f),g)|\le\|T(f)\|\|g\|\le\|T\|\|f\|\|g\|$, de donde $\|B\|\le\|T\|$ y $\|B\|\ge\frac{|B(f,g)|}{\|f\|\|g\|}=\frac{|(T(f),g)|}{\|f\|\|g\|}=|(T(\frac{f}{\|f\|}),\frac{g}{\|g\|})|$ y puesto que $\|T\|=\sup\limits_{\substack{\|f\|\le 1\\ \|g\|\le 1}}\{(T(f),g)\}$, $\|T\|\le\|B\|$. Por lo tanto $\|B\|=\|T\|$.

<u>Definición</u>. 2.7.17 Sea H un espacio de Hilbert, T un operador lineal acotado de H en sí mismo, definimos T' un operador lineal acotado de H^* en sí mismo definido por $T'(G(f))=G(T(f))$ para todo $G\in H^*$, $f\in H$. Sea $F:H\to H^*$ la función la cual asigna a cada $g\in H$, la funcional lineal acotada $(g,\cdot)$ en H^*. Por el lema de Riesz, 2.7.15, la función F es una isometría suprayectiva. Ahora definimos un mapeo $T^*:H\to H$ por $T^*=F^{-1}T'F$. Entonces T^* satisface

$(f,T(g))=(F(f))(T(g))=FF^{-1}(T'F(f))(g)=(F^{-1}T'F(f),g)=(T^*(f),g)$.

T^* es llamado el operador adjunto de T.

<u>Definición</u>. 2.7.18 Un operador lineal acotado T sobre un espacio de Hilbert H, es llamado semiadjunto si $T^*=T$.

<u>Proposición</u>. 2.7.19 Sea T un operador lineal acotado sobre un espacio de Hilbert H, y sea T^* el operador adjunto de T. Entonces
(i). La función $T \mapsto T^*$ es un isomorfismo isométrico de $\mathfrak{L}(H,H)$ en sí mismo.
(ii). $(TS)^*=S^*T^*$.
(iii). $(T^*)^*=T$.
(iv). Si T tiene inversa acotada T^{-1}, entonces T^* tiene inversa y $(T^*)^{-1}=(T^{-1})^*$.
(v). $\|T^*T\| = \|T\|^2$.

Demostración: La prueba de la parte (i), es esencialmente la misma que la del teorema 2.6.40. Así, $\|T\|=\sup\{|(T(f),g)| \;;\; \|f\|\leq 1, \|g\|\leq 1\}=\sup\{|(f,T^*(g))| \;;\; \|f\|\leq 1, \|g\|\leq 1\}=\sup\{\|T^*(g)\| \;;\; \|g\|\leq 1\}$ $=\|T^*\|$. Y obviamente la función $T \mapsto T^*$ es lineal, puesto que
$(f,(aT_1+T_2)(g))=(F(f))(aT_1+T_2)(g)=(F(f))(aT_1(g)+T_2(g))=a(F(f))(T_1(g))$ $+(F(f))(T_2(g))=a(T_1'F(f))(g)+(T_2'F(f))(g)=(aF^{-1}T_1'F(f),g)+(F^{-1}T_2'F(f),g)=$ $(aF^{-1}T_1'F(f)+F^{-1}T_2'F(f),g)=((aF^{-1}T_1'F+F^{-1}T_2'F)(f),g)=((aT_1^*+T_2^*)(f),g)$.

Las pruebas de las partes (ii), (iii) y (iv) son triviales.
La prueba de (v) es como sigue:
$$\|T^*T\| = \sup_{\|f\|\leq 1}|(T^*T(f),f)| = \sup_{\|f\|\leq 1}|(T(f),T(f))| = \sup_{\|f\|\leq 1}\|T(f)\|^2=\|T\|^2.$$

<u>Definición</u>. 2.7.20 Sea H un espacio de Hilbert y $T\in\mathfrak{L}(H,H)$, el espacio nulo y el rango de T, denotados por $N(T)$ y $R(T)$, respectivamente son los conjuntos: $N(T)=\{f \;;\; T(f)=0\}$, $R(T)=\{g \;;\; T(f)=g$ para algún $f\}$.

<u>Proposición</u>. 2.7.21 Supongamos que $T\in\mathfrak{L}(H,H)$, H un espacio de Hilbert y T^* es el adjunto de T. Entonces $N(T^*)=R(T)^{\perp}$ y $N(T)=R(T^*)^{\perp}$.

Demostración: $g\in N(T^*)$ si, y sólo si $T^*(g)=0$ si, y sólo si $(f,T^*(g))=0$ para toda f si, y sólo si $(T(f),g)=0$ para toda f si, y sólo si $g\in R(T)^{\perp}$.

$f\in N(T)$ si, y sólo si $T(f)=0$ si, y sólo si $(T(f),g)=0$ para todo g si, y sólo si $(f,T^*(g))=0$ para todo g si, y sólo si $f\in R(T^*)^{\perp}$.

Bases Ortonormales.

Ya sabemos lo que significa que un conjunto de vectores sea ortonormal, es importante extender la noción de base ortonormal útil en el estudio de espacios lineales de dimensión finita, a espacios lineales con un producto interior completos. Si S es un conjunto ortonormal en un espacio de Hilbert H y ningún otro conjunto ortonormal contiene a S como un subconjunto propio, entonces S es llamado una base ortonormal (o un sistema ortonormal completo) para H.

Ejemplo.2.7.22 Sea $H=C[a,b]$ con el producto interior $(f,g)=\int_a^b f(x)g(x)dx$. Entre diferentes bases ortogonales que se pueden señalar para H es de importancia principal el sistema trigonométrico, compuesto por las funciones $\frac{1}{2}$, $\cos n(\frac{2\pi x}{b-a})$, $\operatorname{sen} n(\frac{2\pi x}{b-a})$ $(n=1,\ldots,n)$. Sobre el proceso de ortonormalización se discutirá más adelante.

Teorema. 2.7.23 Todo espacio de Hilbert H tiene una base ortonormal.
Demostración: Consideremos la colección C de conjuntos ortonormales en V. Ordenamos C por inclusión; esto es $S_1 < S_2$ si $S_1 \subset S_2$, con esta definición de $<$, C es un conjunto parcialmente ordenado; C es también un conjunto no vacío, puesto que si $f \in H$, el conjunto consistente sólo de $f/\|f\|$, él mismo es un conjunto ortonormal. Ahora, sea $\{S_\alpha\}_{\alpha\in I}$ cualquier subconjunto totalmente (o linealmente) ordenado. Entonces $\bigcup_{\alpha\in I} S_\alpha$ es un conjunto ortonormal el cual contiene a S_α para toda $\alpha\in I$, y así $\bigcup_{\alpha\in I} S_\alpha$ es cota superior de $\{S_\alpha\}_{\alpha\in I}$. Puesto que todo conjunto C totalmente ordenado tiene una cota superior, se sigue del Lema de Zorn que C tiene un elemento maximal; esto es un sistema ortonormal contenido no propiamente en cualquier otro sistema ortonormal.

El siguiente Teorema muestra que todo elemento de un espacio de Hilbert H puede ser expresado como una combinación lineal (posiblemente infinita) de elementos básicos, a similitud de como sucede en el caso de dimensión finita.

Teorema. 2.7.24 (Riesz-Fisher). Sea H un espacio de Hilbert y $S=\{f_\alpha\}_{\alpha\in I}$ una base ortonormal. Entonces para cada $f\in H$,
$f=\sum_{\alpha\in I}(f_\alpha,f)f_\alpha$ y $\|f\|^2=\sum_{\alpha\in I}|(f_\alpha,f)|^2$, esta relación es conocida como la igualdad de Parseval. y los (f_α,f) son frecuentemente conocidos como los coeficientes de Fourier de f con respecto a la base $\{f_\alpha\}_{\alpha\in I}$.
La igualdad de Parseval significa que la suma del lado derecho converge independientemente del orden a $f\in H$. Inversamente, si $c_\alpha=(f_\alpha,f)$ y $\sum_{\alpha\in I}|c_\alpha|^2<\infty$, entonces $\sum_{\alpha\in I}c_\alpha f_\alpha$ converge a un elemento de H.
Demostración: Ya hemos visto que para cualquier subconjunto finito $J\subset I$ tenemos que $\sum_{\alpha\in J}|(f_\alpha,f)|^2\leq\|f\|^2$ (Desigualdad de Bessel). Así $(f_\alpha,f)\neq 0$ para a lo más un conjunto numerable de α's en I los cuales son numerados de alguna manera como: $\alpha_1,\ldots,\alpha_n,\ldots$ Además, ya que $\sum_{j=1}^{N}|(f_{\alpha_j},f)|^2$ es monótona creciente y acotada, ésta converge a un límite finito cuando $N\to\infty$. Sea $f_n=\sum_{j=1}^{n}(f_{\alpha_j},f)f_{\alpha_j}$. Entonces para $n>m$, tenemos que

$$\|f_n-f_m\|^2=\|\sum_{j=m+1}^{n}(f_{\alpha_j},f)f_{\alpha_j}\|^2=\sum_{j=m+1}^{n}|(f_{\alpha_j},f)|^2.$$

De donde $\{f_n\}_{n=1}^{\infty}$ es una sucesión de Cauchy, y puesto que H es de Hilbert, ésta converge a un elemento $f'\in H$.
Observamos que

$$(f-f',f_{\alpha_\lambda})=\lim_{n\to\infty}(f-\sum_{j=1}^{n}(f_{\alpha_j},f)f_{\alpha_j},f_{\alpha_\lambda})=(f,f_{\alpha_\lambda})-(f,f_{\alpha_\lambda})=0.$$

Y si $\alpha\neq\alpha_\lambda$ para algún λ tenemos que

$$(f-f',f_\alpha)=\lim_{n\to\infty}(f-\sum_{j=1}^{n}(f_{\alpha_j},f)f_{\alpha_j},f_\alpha)=0.$$

De donde $f-f'$ es ortogonal a todo $f_\alpha\in S$, y puesto que S es un sistema ortogonal completo, debemos tener que $f-f'=0$, así $f=\lim_{n\to\infty}\sum_{j=1}^{n}(f_{\alpha_j},f)f_{\alpha_j}$.
Además

$$0=\lim_{n\to\infty}\|f-\sum_{j=1}^{n}(f_{\alpha_j},f)f_{\alpha_j}\|^2=\lim_{n\to\infty}(\|f\|^2-\sum_{j=1}^{n}|(f_{\alpha_j},f)|^2)=\|f\|^2-\sum_{\alpha\in I}|(f_{\alpha_j},f)|^2.$$

Con lo que se concluye la demostración del Teorema.

Ahora describimos un procedimiento útil, para construir un conjunto ortonormal a partir de una sucesión arbitraria de vectores independientes, llamado el proceso de ortogonalización de Gram-Schmidt.
Supongamos que los vectores independientes $g_1,g_2,\ldots$ son dados y definimos

$h_1=g_1,$

$$h_2=g_2-\frac{(h_1,g_2)}{(h_1,h_1)}h_1,$$

$$h_3=g_3-\frac{(h_2,g_3)}{(h_2,h_2)}h_2-\frac{(h_1,g_3)}{(h_1,h_1)}h_1,$$

$\vdots$

$$h_n=g_n-\frac{(h_{n-1},g_n)}{(h_{n-1},h_{n-1})}h_{n-1}-\ldots-\frac{(h_1,g_n)}{(h_1,h_1)}h_1,$$

$\vdots$

Entonces el conjunto $\{h_1/\|h_1\|, h_2/\|h_2\|,\ldots,h_n/\|h_n\|,\ldots\}$ es un conjunto ortonormal.

Ejemplo.2.7.25 (Polinomios de Lagrange). Los Polinomios de Lagrange, son obtenidos al aplicar el proceso de Gram-Schmidt a las funciones $1,x,x^2,x^3,\ldots$, definidas sobre el intervalo $[0,1]$, con el producto interior $(f,g)=\int_{-1}^{1} f(x)g(x)dx$.

CAPITULO TRES:

METODOS VARIACIONALES

3. METODOS VARIACIONALES:

Reducción al absurdo, es una de las más finas armas matemáticas. Es un gambito más fino que cualquier gambito de ajedrez: Un jugador de ajedrez, puede ofrecer el sacrificio de un peón o de otra pieza para obtener una posición favorable. Sin embargo, un matemático ofrece el juego. G. H. Hardy.

3.1 INTRODUCCION.

En el transcurso del capítulo anterior, mientras se desarrollaban los elementos del análisis funcional, frecuentemente se consideraron problemas de optimización de norma mínima, en los cuales la disponibilidad de una variedad de normas diferentes proporcionaba suficiente flexibilidad para su estudio. Otro tipo de problemas de optimización en espacios de funciones que surgieron de manera natural fueron el teorema de Hahn-Banach y el teorema de proyección.

En el presente capítulo, trataremos con funcionales de un tipo especial: funcionales representadas por integrales, cuyos integrandos contienen funciones conocidas pertenecientes a algún espacio y mucho de la teoría del capítulo anterior es de beneficio directo en el estudio de problemas de optimización que involucran dichas funcionales, las cuales en general son no lineales.

En el capítulo anterior, se discutieron los conceptos de funcional y operador lineal. Sin embargo, muchos de los problemas que surgen en la teoría de optimización en espacios de funciones tienen un carácter sustancialmente no lineal, le pertenecen a ésta, de hecho, los métodos variacionales, los cuales en su conjunto constituyen una de las ramas principales del análisis funcional no lineal.

El desarrollo del presente capítulo se da como sigue: En la sección 3.2 se estudian los conceptos de la teoría de diferenciación en espacios de funciones lineales normados y su aplicación en la obtención de condiciones necesarias y/o suficientes de extremo de una funcional no necesariamente lineal, en esta sección se discuten además los conceptos de funcionales bilineales y cuadráticas como una generalización de las formas bilineales y cuadráticas en espacios de dimensión finita. En la sección 3.3 se desarrollan los métodos variacionales clásicos y sus aplicaciones a la mecánica y geometría. Aquí mismo se plantea desde el problema variacional mas simple hasta el más generalizado y complejo, estableciendo para cada uno de ellos condiciones necesarias y/o suficientes de extremo. Para la sección 3.4 el estudio de los métodos variacionales clásicos es central para el entendimiento de muchos problemas de la teoría de control y procesos óptimos. Se plantea además el problema de control en el caso general y se presenta el principio del máximo como un conjunto de condiciones necesarias de extremo. En la sección 3.5, se presenta la relación que en muchas situaciones surge entre la teoría de control y la programación dinámica cuando es posible discretizar. Se discute también aquí el problema básico de la programación dinámica, el principio de optimalidad y la relación de éste con los métodos variacionales. En la sección 3.6 se presentan algunas aplicaciones del teorema de Hahn-Banach a problemas de tipo variacional. Por último, en la sección 3.7 se dan algunas aplicaciones de los métodos variacionales a la estadística, en particular a la teoría de información sobre el principio de máxima entropía.

3.2 DIFERENCIACION EN ESPACIOS LINEALES NORMADOS.

Una vez que se ha visto la utilidad de emplear un lenguaje geométrico en el estudio de operadores y funcionales lineales, viendo a cada función f como un punto de algún espacio. Nos interesa ahora extender esta utilidad hacia la teoría de diferenciación sobre espacios lineales normados, dentro del estudio del análisis funcional no lineal.

Sea $(V,\|\cdot\|)$ un espacio lineal normado, en lo que sigue $B(f;\epsilon)$, al igual que antes, denotará a la bola abierta de V, $\{h;\|f-h\|<\epsilon\}$, con centro en el punto $f\in V$ y radio $\epsilon>0$. Observemos que $\|h\|<\epsilon$ si y sólo si $f+h\in B(f;\epsilon)$, puesto que $\|f-(f+h)\|=\|h\|<\epsilon$.
Todos los conjuntos abiertos que se traten serán bolas, y recordamos también que las bolas son convexas, este hecho se empleará posteriormente.

<u>Definición</u>. 3.2.1 Sea $(V,\|\cdot\|)$ un espacio lineal normado, $O=B(f;\epsilon)$ un subconjunto abierto de V y $F:O\subset V\to\mathbb{R}$ una funcional (no necesariamente lineal). Se dice que la funcional F es diferenciable en el punto $f\in O$, si existe una funcional lineal acotada $dF(f)\in\mathcal{L}(V,\mathbb{R})=V^*$, tal que para $\|h\|<\epsilon$, $F(f+h)-F(f)=dF(f)(h)+o(\|h\|)$, donde $o(\|h\|)$ significa que $|o(\|h\|)|/\|h\|=|F(f+h)-F(f)-dF(f)(h)|/\|h\|\to 0$ cuando $\|h\|\to 0$, ($o(\|h\|)$ es frecuentemente llamado un infinitésimo de orden superior al primero respecto a $\|h\|$). $dF(f)(h)$ es llamada la diferencial de Fréchet de la funcional F en el punto $f\in O$. Obviamente este concepto es local. Por lo que es conveniente definir a una funcional F como diferenciable en un conjunto $O\subset V$, si es diferenciable para cada $f\in O$, esto es, para cada $f\in O$ existe $dF(f)\in\mathcal{L}(V,\mathbb{R})=V^*$, tal que cumple con la propiedad antes mencionada.

Observamos que F sólo está definida en $B(f;\epsilon)$, sin embargo, $dF(f)$ está definida en todo V, en efecto sea $h\in V$, $h\neq 0$, entonces $f+\frac{\epsilon}{2\|h\|}h\in B(f;\epsilon)$ puesto que $\|f-(f+\frac{\epsilon}{2\|h\|}h)\|=\|\frac{\epsilon}{2\|h\|}h\|=\frac{\epsilon}{2}<\epsilon$.
Si F es diferenciable en f, entonces
$F(f+\frac{\epsilon}{2\|h\|}h)-F(f)=\frac{\epsilon}{2\|h\|}dF(f)(h)+o(\frac{1}{2}\epsilon)$.

<u>Ejemplos</u>. 3.2.2 Sea $V=C[a,b]$ con la norma $\|h\|=\max_{x\in[a,b]}|h(x)|$, definimos $F(f)=\frac{1}{2}\int_a^b f^2(x)g_0(x)dx$, donde $g_0(x)$ es una función fija en $C[a,b]$,

entonces $F(f+h)-F(f)=\int_a^b f(x)h(x)g_0(x)dx+\frac{1}{2}\int_a^b h^2(x)g_0(x)dx$. Además tenemos que existe un $M\geq 0$, tal que $|g_0(x)|\leq M$ para toda $x\in[a,b]$, así $|\frac{1}{2}\int_a^b h^2(x)g_0(x)dx|\leq\frac{1}{2}(b-a)M\|h\|^2=o(\|h\|)$, de donde $dF(f)(h)=\int_a^b f(x)\cdot g_0(x)h(x)dx$, claramente $dF(f)$ es lineal y como hemos visto $\|dF(f)\|=\int_a^b|f(x)g_0(x)|dx$, por lo que también es acotada.

Veamos ahora otro ejemplo. Sea $V=C[a,b]$ con la norma $\|h\|=\max|h(x)|$, definimos $F(f)=\frac{1}{2}\int_a^b\int_a^b K(x,y)f(x)f(y)dxdy$, en donde $K\in C([a,b]\times[a,b])$ y satisface la condición $K(x,y)=K(y,x)$ para todo $(x,y)\in[a,b]\times[a,b]$. Entonces $F(f+h)-F(f)=\int_a^b\int_a^b K(x,y)f(x)h(y)dxdy+\frac{1}{2}\int_a^b\int_a^b K(x,y)h(x)h(y)dxdy$, Ahora bien, sabemos que existe un $M\geq 0$ tal que $|K(x,y)|\leq M$ para toda $(x,y)\in[a,b]\times[a,b]$, así $|\frac{1}{2}\int_a^b\int_a^b K(x,y)h(x)h(y)dxdy|\leq\frac{1}{2}M(b-a)^2\|h\|^2=o(\|h\|)$, entonces $dF(f)(h)=\int_a^b\int_a^b K(x,y)f(x)h(y)dxdy$, y es claro que $dF(f)$ es lineal, para ver que también es acotada basta considerar que $|dF(f)(h)|=(\|f\|\int_a^b\int_a^b|K(x,y)|dxdy)\|h\|$ para toda $h\in V$, y dado que existe un $C\geq 0$, tal que $|f(x)|\leq C$ para todo $x\in[a,b]$, entonces $|dF(f)(h)|\leq MC(b-a)^2\|h\|$ para toda $h\in V$. Si tenemos que $dF(f)(h)=0$ para toda $h\in V=C[a,b]$, tendremos entonces que $\int_a^b K(x,y)f(x)dx=0$ para todo $y\in[a,b]$.

Notamos que $\int_a^b K(x,y)f(x)dx=0$ es una ecuación integral homogenea de Fredholm de primera especie con núcleo K respecto a f. Claramente la funcional $T(f)=\int_a^b K(x,y)f(x)dx$ es lineal y acotada. Una de las soluciones de esta ecuación para $T(f)=0$, es la función $f(x)=0$ para toda $x\in[a,b]$. La respuesta a la pregunta de si existe extremo en este punto y si existen otros puntos en los que es posible un extremo obviamente depende de la forma del núcleo K.

<u>Proposición</u>. 3.2.3 Sea $(V,\|\cdot\|)$ un espacio lineal normado, $O=B(f;\varepsilon)$ un conjunto abierto de V y $F:O\to\mathbb{R}$ una funcional que es diferenciable en el sentido de Fréchet en el punto $f\in O$. Entonces $dF(f)$ es única.

Demostración: Supongamos que $f+h\in O$, es decir $\|h\|<\varepsilon$ y que $F(f+h)-F(f)=dF(f)(h)+o_1(\|h\|)=eF(f)(h)+o_2(\|h\|)$.

Entonces $dF(f)(h)-eF(f)(h)=o_2(\|h\|)-o_1(\|h\|)=o(\|h\|)$, y por lo tanto $dF(f)(h)-eF(f)(h)$ es un infinitésimo de orden superior al primero respecto a $\|h\|$, y como ya sabemos, de acuerdo a la definición 3.2.1,

esto significa que $|dF(f)(h)-eF(f)(h)|/\|h\| \to 0$ cuando $\|h\| \to 0$.
Para $h\neq 0$, tenemos que $th \to 0$ cuando $t \to 0$, entonces si $f+th\in O$
$0=\lim_{t\to 0} |dF(f)(th)-eF(f)(th)|/\|th\|=|dF(f)(h)-eF(f)(h)|/\|h\|$.
Por tanto $dF(f)(h)=eF(f)(h)$ para toda $h\in V$.

<u>Teorema</u>. 3.2.4 Sea $(V,\|\cdot\|)$ un espacio lineal normado.
(i). Si $F:V \to \mathbb{R}$ es una funcional constante, digamos, $F(f)=C$ para toda $f\in V$, entonces $dF(f)(h)=0$ para todo $h\in V$.
(ii). Si $F:V \to \mathbb{R}$ es una funcional lineal continua (y por lo tanto acotada), entonces $dF(f)(h)=F(h)$ para toda $h\in V$.
Demostración. (i). $\lim_{\|h\|\to 0} |F(f+h)-F(f)-0|/\|h\|=\lim_{\|h\|\to 0} |C-C|/\|h\|=0$.
(ii). $\lim_{\|h\|\to 0} |F(f+h)-F(f)-F(h)|/\|h\|=\lim_{\|h\|\to 0} |F(f+h)-F(f+h)|/\|h\|=0$.

<u>Proposición</u>. 3.2.5 Sea $(V,\|\cdot\|)$ un espacio lineal normado, $O=B(f;\varepsilon)$ un conjunto abierto de V y $F:O\subset V \to \mathbb{R}$, $G:O\subset V \to \mathbb{R}$ dos funcionales diferenciables en el sentido de Fréchet en el punto $f\in O$, entonces $d(F+G)(f)(h)=dF(f)(h)+dG(f)(h)$ para todo $h\in V$.
La demostración se sigue de que si $f+h\in O$, entonces
$|(F+G)(f+h)-(F+G)(f)-(dF(f)(h)+dG(f)(h))|/\|h\|\leq$
$|F(f+h)-F(f)-dF(f)(h)|/\|h\|+ |G(f+h)-G(f)-dG(f)(h)|/\|h\| \to 0$ cuando $\|h\| \to 0$.

<u>Definición</u>. 3.2.6 Sea $(V,\|\cdot\|)$ un espacio lineal normado, $O=B(f;\varepsilon)$ un conjunto abierto de V y $F:O\subset V \to \mathbb{R}$ una funcional. Se llama diferencial de Gateaux de la funcional F en el punto $f\in O$ al límite
$\lim_{t\to 0} (F(f+th)-F(f))/t= \frac{d}{dt}F(f+th)\Big|_{t=0}$, con $f+th\in O$.

<u>Proposición</u>. 3.2.7 Sea $(V,\|\cdot\|)$ un espacio lineal normado, $O=B(f;\varepsilon)$ un conjunto abierto de V y $F:O\subset V \to \mathbb{R}$ una funcional. Si F es diferenciable en el sentido de Fréchet en el punto $f\in O$, entonces F es también diferenciable en el sentido de Gateaux en el punto $f\in O$, y además $dF(f)(h)= \frac{d}{dt}F(f+th)\Big|_{t=0}$.
Demostración; Supongamos que $f+th\in O$ y que F es diferenciable en el sentido de Fréchet, entonces $F(f+th)-F(f)=dF(f)(th)+o(th)=t\,dF(f)(h)+o(th)$, entonces $(F(f+th)-F(f))/t=dF(f)(h)+(o(th)/t) \to dF(f)(h)$ cuando $t \to 0$, así $\frac{d}{dt}F(f+th)\Big|_{t=0}=dF(f)(h)$.

<u>Definición</u>. 3.2.8 Sea $(V,\|\cdot\|)$ un espacio lineal normado. Se dice que $B:V\times V\to\mathbb{R}$ es una funcional bilineal acotada, si para todo f, $g,h\in V$ y $a,b\in\mathbb{R}$ satisface que:
(i). $B(af+bg,h)=aB(f,h)+bB(g,h)$.
(ii). $B(f,ag+bh)=aB(f,g)+bB(f,h)$.
(iii). Existe un número $C\geq 0$ tal que $|B(f,g)|\leq C\|f\|\|g\|$ para todo f, $g\in V$.
El mínimo de tales C's es llamado la norma de B, denotado $\|B\|$.
De esta manera $\|B\|=\sup\limits_{f,g\neq 0}|B(f,g)|/(\|f\|\|g\|)=\sup\limits_{\|f\|,\|g\|\leq 1}|B(f,g)|$, puesto que $|B(f,g)|/(\|f\|\|g\|)=B(f/\|f\|,g/\|g\|)$, y para todo $\epsilon>0$ existen $h_1,h_2\in V$ tales que $|B(h_1,h_2)|/(\|h_1\|\|h_2\|)>\|B\|-\epsilon$, esto implica que para todo $\epsilon>0$ existen $h_1,h_2\in V$ tales que $|B(h_1,h_2)|>(\|B\|-\epsilon)\|h_1\|\|h_2\|$ (ver Prop. 2.2.3.).
Además se define $(B_1+B_2)(f,g)=B_1(f,g)+B_2(f,g)$.
Es así como el espacio de todas las funcionales bilineales acotadas del espacio V en $\mathbb{R}$, constituye un espacio lineal normado que denotaremos por $B(V\times V,\mathbb{R})$.

<u>Ejemplo</u>. 3.2.9 Consideremos el álgebra de Banach $V=C[a,b]$ con la norma $\|f\|=\max\limits_{x\in[a,b]}|f(x)|$, definimos $B(f,g)=\int_a^b f(x)g(x)dx$, obviamente B es una funcional bilineal. Ahora veamos que $\|B\|=b-a$, en efecto, para todo $f,g\in V$ tenemos que $|B(f,g)|\leq\int_a^b|f(x)g(x)|dx\leq(b-a)\|f\|\|g\|$, si $f,g\neq 0$, entonces $|B(f,g)|/(\|f\|\|g\|)\leq b-a$, luego $\|B\|\leq b-a$. Por otra parte, si $f(x)=g(x)=1$ para toda $x\in[a,b]$, entonces $|B(f,g)|=(b-a)\|f\|\|g\|$, así $\|B\|\geq|B(f,g)|/(\|f\|\|g\|)=b-a$. Por tanto $\|B\|=b-a$.

<u>Ejemplo</u>. 3.2.10 Sea $V=BR[a,b]$, el espacio consistente de todas las funciones acotadas e integrables en el sentido de Riemann sobre $[a,b]$, con la norma $\|f\|=\max\limits_{x\in[a,b]}|f(x)|$, definimos $B(f,g)=\int_a^b f(x)g(x)h_o(x)dx$, donde $h_o(x)$ es una función fija en $BR[a,b]$. Entonces B es una funcional bilineal acotada, sabemos que para toda $f,g\in V$ $|B(f,g)|\leq\int_a^b|f(x)g(x)||h_o(x)|dx\leq(\int_a^b|h_o(x)|dx)\|f\|\|g\|$, si $f,g\neq 0$, tenemos que $|B(f,g)|/(\|f\|\|g\|)\leq\int_a^b|h_o(x)|dx$, entonces $\|B\|\leq\int_a^b|h_o(x)|dx$. Mientras que por otra parte si $f(x)=1$ para toda $x\in[a,b]$ y $g(x)=\operatorname{sgn}(h_o(x))$, entonces $\|f\|=\|g\|=1$ y $|B(f,g)|=(\int_a^b|h_o(x)|dx)\|f\|\|g\|$, así $\|B\|\geq|B(f,g)|/(\|f\|\|g\|)=\int_a^b|h_o(x)|dx$. Por tanto $\|B\|=\int_a^b|h_o(x)|dx$.

Definición. 3.2.11 Sea $(V,\|\cdot\|)$ un espacio lineal normado, y $B\in \mathcal{L}(V\times V,\mathbb{R})$, diremos que T es una funcional cuadrática si $f=g$ en $B(f,g)$, esto es, $T(f)=B(f,f)$. Una funcional cuadrática T es llamada definida positiva si $T(f)>0$ para toda $f\neq 0$, y es llamada fuertemente positiva, si existe una constante $C>0$ tal que $T(h)\geq C\|h\|^2$ para todo $h\in V$.

Ejemplo. 3.2.12 Sea $V=C^1[a,b]$ con la norma $\|f\|=\max_{a\le x\le b}|f(x)|+\max_{a\le x\le b}|f'(x)|$, entonces $T(f)=\int_a^b (a_1(x)f^2(x)+2a_2(x)f(x)f'(x)+a_3(x)(f'(x))^2)dx$, donde $a_1(x), a_2(x)$ y $a_3(x)$ son funciones fijas en $C^1[a,b]$, es una funcional cuadrática acotada que se obtiene de la funcional bilineal $B(f,g)=\int_a^b(a_1(x)f(x)g(x)+a_2(x)(f'(x)g(x)+f(x)g'(x))+a_3(x)f'(x)g'(x))dx$. Veamos que B es acotada, sabemos que existe una $M\geq 0$, $M=\max\{M_i$; $|a_i(x)|\leq M_i$, $x\in[a,b]$, $i=1,2,3\}$, así $|B(f,g)|\leq(b-a)M\|f\|\|g\|$ es válida para todo $f,g\in V$.

Ejemplo. 3.2.13 Sea $V=C^1[a,b]$ con la norma $\|f\|=\max_{a\le x\le b}|f(x)|+\max_{a\le x\le b}|f'(x)|$, entonces $T(f)=\int_a^b (a_1(x)f^2(x)+a_2(x)(f'(x))^2)dx$, donde $a_1(x)$ y $a_2(x)$ son funciones fijas en $C^1[a,b]$, es una funcional cuadrática acotada que se obtiene de la funcional bilineal acotada $B(f,g)=\int_a^b (a_1(x)f(x)g(x)+a_2(x)f'(x)g'(x))dx$. Es claro que B es acotada, sea $M=\max(M_1,M_2)$, donde $|a_1(x)|\leq M_1$ y $|a_2(x)|\leq M_2$ para toda $x\in[a,b]$, así $|B(f,g)|\leq(b-a)M\|f\|\|g\|$ para todo $f,g\in V$.

Proposición. 3.2.14 Sean $\mathcal{L}(V,\mathcal{L}(V,\mathbb{R}))$ y $B(V\times V,\mathbb{R})$, entonces a cada $T\in\mathcal{L}(V,\mathcal{L}(V,\mathbb{R}))$ se le puede poner en correspondencia un elemento $B\in B(V\times V,\mathbb{R})$ mediante $B(f,g)=(Tf)(g)$. Esta correspondencia es un isomorfismo isométrico de $\mathcal{L}(V,\mathcal{L}(V,\mathbb{R}))$ en $B(V\times V,\mathbb{R})$.

Demostración: Es claro que esta correspondencia es lineal, en efecto, sea $a\in\mathbb{R}$, T_1, $T_2\in\mathcal{L}(V,\mathcal{L}(V,\mathbb{R}))$ y B_1, $B_2\in B(V\times V,\mathbb{R})$ tales que $(T_1(f))(g)=B_1(f,g)$, $(T_2(f))(g)=B_2(f,g)$, entonces $((aT_1+T_2)(f))(g)=(aT_1(f)+T_2(f))(g)=(aT_1(f))(g)+(T_2(f))(g)=aB_1(f,g)+B_2(f,g)=(aB_1+B_2)(f,g)$. Veamos ahora que esta correspondencia es una isometría, para todo $f,g\in V$ tenemos que $|B(f,g)|=|(Tf)(g)|\leq\|Tf\|\|g\|\leq\|T\|\|f\|\|g\|$, de donde $\|B\|\leq\|T\|$. Por otro lado la aplicación $g\mapsto(Tf)(g)=B(f,g)$ es para cada f fijo una funcional lineal sobre V, entonces $\|T(f)\|=\sup_{\|g\|\leq 1}|(Tf)(g)|=\sup_{\|g\|\leq 1}|B(f,g)|\leq\|B\|\|f\|$, así $\|T\|\leq\|B\|$. Por tanto $\|T\|=\|B\|$. Además esta correspondencia es suprayectiva.

<u>Definición</u>. 3.2.15 Sea $(V,\|\cdot\|)$ un espacio lineal normado, $O=B(f,\ )$ un conjunto abierto de V y $F:O\subset V\to\mathbb{R}$ una funcional. Se dice que la funcional F es dos veces diferenciable en el punto $f\in O$, si existe una funcional lineal acotada $dF(f)\in\mathcal{L}(V,\mathbb{R})$ y una funcional cuadrática acotada $d^2F(f)\in B(V\times V,\mathbb{R})$, tales que si $f+h\in O$, $F(f+h)-F(f)=dF(f)(h)+d^2F(f)(h,h)+o(\|h\|^2)$, donde $o(\|h\|^2)$ significa que $|o(\|h\|^2)|/\|h\|^2=|F(f+h)-F(h)-dF(f)(h)-d^2F(f)(h,h)|/\|h\|^2\to 0$ cuando $\|h\|\to 0$, $d^2F(f)(h,h)$ es llamada la diferencial segunda de Fréchet de la funcional F en el punto $f\in O$.

<u>Ejemplo</u>. 3.2.16 Sea $V=C[a,b]$ con la norma $\|f\|=\max_{x\in[a,b]}|f(x)|$, definimos $F(f)=\frac{1}{2}\int_a^b f^2(x)g_0(x)dx$, donde $g_0(x)$ es una función fija en $C[a,b]$, entonces $F(f+h)-F(f)=\int_a^b f(x)h(x)g_0(x)dx+\frac{1}{2}\int_a^b h^2(x)g_0(x)dx$, aquí obviamente $dF(f)(h)=\int_a^b f(x)g_0(x)h(x)dx$ y $d^2F(f)(h,h)=\frac{1}{2}\int_a^b h^2(x)g_0(x)dx$, con $\|dF(f)\|=\int_a^b|f(x)g_0(x)|dx$ y $\|d^2F(f)\|=\frac{1}{2}\int_a^b|g_0(x)|dx$.

Una de las partes más elaboradas del análisis funcional no lineal son sin duda los métodos variacionales, cuyo estudio esencialmente trata de la búsqueda de condiciones necesarias y/o suficientes para que una funcional F definida sobre un conjunto abierto de un espacio de Banach alcance su óptimo en un punto $f\in O$. Para estar de acuerdo con la terminología de los métodos variacionales un óptimo será llamado un extremo.

<u>Definición</u>. 3.2.17 Sea $(V,\|\cdot\|)$ un espacio Banach, O un subconjunto abierto de V y $F:O\subset V\to\mathbb{R}$ una funcional. Se dice que F tiene un mínimo (local o relativo) en el punto $f\in O$, cuando para todos los $f+h\in O$, suficientemente próximos a f, se cumple la desigualdad $F(f+h)-F(f)\geq 0$, esto es, existe un número $\varepsilon>0$ tal que $F(f+h)-F(f)\geq 0$ se cumple para todos los $f+h\in O$ tales que $\|f+h-f\|=\|h\|<\varepsilon$.
De manera análoga se define un máximo.

<u>Teorema</u>. 3.2.18 Sea $(V,\|\cdot\|)$ un espacio de Banach, $O=B(f,\varepsilon)$ un conjunto abierto de V y $F:O\subset V\to\mathbb{R}$ una funcional diferenciable en el punto $f\in O$ en el sentido de Fréchet. Para que F alcance un extremo en el punto $f\in O$ es necesario que su diferencial en ese punto sea igual a cero

para todo h, esto es, $dF(f)(h)=0$ para todo $h\in V$.
Demostración: Supongamos que $f+h\in O$, entonces $F(f+h)-F(f)=dF(f)(h)+o(\|h\|)$, y que $dF(f)(h_o)\neq 0$ para algún h_o, entonces para valores reales suficientemente pequeños de $a\neq 0$, si $f+ah_o\in O$ y si denotamos a $h=ah_o$, entonces el signo de la expresión $F(f+h)-F(f)=dF(f)(h)+o(\|h\|)$ coincide con el signo de $dF(f)(h)$. Pero $dF(f)$ es una funcional lineal por lo que $dF(f)(ah_o)=adF(f)(h_o)$. De manera que siempre $dF(f)(h_o)\neq 0$, se tiene que $F(f+h)-F(f)$ puede tomar para h arbitrariamente pequeño, tanto valores positivos, como negativos, dependiendo del signo de a, es decir, no puede haber extremo en el punto $f\in O$.

Ejemplo. 3.2.19 Sea $V=C[a,b]$ con la norma $\|h\|=\max_{x\in[a,b]}|h(x)|$, definimos $F(f)=\int_a^b H(x,f(x))dx$, donde H es una función con derivada continua, entonces haciendo uso del Teorema de Taylor tenemos que $F(f+h)-F(f)=\int_a^b (H(x,f(x)+h(x))-H(x,f(x))dx=\int_a^b \frac{\partial H}{\partial f}(x,f(x))h(x)dx+o(\|h\|)$. Así la funcional F es diferenciable en el punto f, con $dF(f)(h)=\int_a^b \frac{\partial H}{\partial f}(x,f(x))$ $h(x)dx$. Si suponemos que $dF(f)(h)=0$ para toda $h\in V$, entonces $\frac{\partial H}{\partial f}(x,f(x))=0$. Esta ecuación determina en general, una curva en la cual la funcional F puede alcanzar un extremo.

Teorema. 3.2.20 Sea $(V,\|\cdot\|)$ un espacio de Banach, $O=B(f,\epsilon)$ un abierto de V y $F:O\subset V\to\mathbb{R}$ una funcional con diferencial segunda continua de Fréchet en una vecindad $N\subset O$ de f. Si esta funcional alcanza un mínimo en el punto $f\in N$, entonces $d^2F(f)(h,h)\geq 0$ para toda $h\in V$.
Demostración: Supongamos que $f+h\in N$ y que $F(f+h)-F(f)=dF(f)(h)+d^2F(f)(h,h)+o(\|h\|^2)$, si F tiene un mínimo en $f\in N$, entonces $dF(f)(h)=0$ para todo h, de donde $F(f+h)-F(f)=d^2F(f)(h,h)+o(\|h\|^2)$. Supongamos ahora que $d^2F(f)(h_o,h_o)<0$ para algún h_o. Y como para todo $a\neq 0$, se cumple que $d^2F(f)(ah_o,ah_o)=a^2d^2F(f)(h_o,h_o)<0$, si $f+ah_o\in N$ y si denotamos a $h=ah_o$, entonces existen elementos h para a arbitrariamente pequeño, para los cuales se cumple que $d^2F(f)(h,h)<0$, pero el signo de la expresión $F(f+h)-F(f)=d^2F(f)(h,h)+o(\|h\|^2)$ coincide con el signo de $d^2F(f)(h,h)$ para h suficientemente pequeños y por lo tanto $F(f+h)-F(f)<0$ para h suficientemente pequeños, es decir, no hay mínimo en $f\in N$.
De manera análoga se considera el caso de un máximo.

La condición $d^2F(f)(h,h)\geq 0$ para todo h es necesaria pero no suficiente para que una funcional $F:O\subset V \to \mathbb{R}$ tenga un mínimo, al contrario de como sucede en el caso en que V es un espacio de dimensión finita. En efecto consideremos un espacio de Banach de dimensión infinita, digamos, el espacio de Hilbert de sucesiones infinitas de números reales tales que $\sum_{n=1}^{\infty} x_n^2<\infty$, con el producto interior $(x,y)=\sum_{n=1}^{\infty} x_n y_n$, donde $x=\{x_n\}_{n=1}^{\infty}$ y $y=\{y_n\}_{n=1}^{\infty}$. Definimos la funcional

$F(x)=\sum_{n=1}^{\infty} x_n^2/n^3-\sum_{n=1}^{\infty} x_n^4$.

Entonces en el punto 0, la primera diferencial de F en el sentido de Fréchet es igual a cero, esto es, si $h=\{h_n\}_{n=1}^{\infty}$, $dF(0)(h)=0$ para todo h, y la diferencial segunda de F en el punto 0, en el sentido de Fréchet $d^2F(0)(h,h)=\sum_{n=1}^{\infty} h_n^2/n^3$ para todo h, es decir $d^2F(0)$ representa una funcional cuadrática definida positiva. Sin embargo, en el punto 0 no hay mínimo, ya que $F(0)=0$ y $F(0,..,1/n,0,...)=(1/n^5)-(1/n^4)<0$. Por consiguiente en cualquier vecindad del punto 0 existen puntos x para los cuales $F(x)<F(0)$.

<u>Teorema</u>. 3.2.21 Sea $(V,\|\cdot\|)$ un espacio de Banach, $C=B(f,\epsilon)$ un conjunto abierto de V y $F:O\subset V \to \mathbb{R}$ una funcional con diferencial segunda continua de Fréchet en una vecindad $N\subset O$ de f. Una condición suficiente para que la funcional F alcance un mínimo en el punto $f\in N$, dado que $dF(f)(h)=0$ para todo h, es que $d^2F(f)$ sea una funcional cuadrática fuertemente positiva, esto es, existe una constante $C>0$ tal que $d^2F(f)(h,h)\geq C\|h\|^2$ para toda $h\in V$.

Demostración: Supongamos que $f+h\in N$ y que $F(f+h)-F(h)=dF(f)(h)+d^2F(f)(h,h)+o(\|h\|^2)$, si $dF(f)(h)=0$ para todo h, entonces $F(f+h)-F(f)=d^2F(f)(h,h)+o(\|h\|^2)$. Ahora tomemos un número $\epsilon>0$ lo suficientemente pequeño tal que para $\|h\|<\epsilon$ se verifique que $|o(\|h\|^2)|<(C/2)\|h\|^2$, (Por continuidad. Dado $\frac{C}{2}$ existe $\epsilon>0$, tal que $\left|\frac{o(\|h\|)}{\|h\|^2}-0\right|<\frac{C}{2}$ para $\|h\|<\epsilon$).

Entonces $F(f+h)-F(f)=d^2F(f)(h,h)+o(\|h\|^2)>(C/2)\|h\|^2>0$ para $\|h\|<\epsilon$. Por tanto $F(f+h)-F(f)<0$ para $\|h\|<\epsilon$, es decir la funcional F tiene un mínimo en el punto f.

Note que la funcional cuadrática $d^2J(0)(h,h)=\sum_{n=1}^{\infty}h_n^2/n^3$ no es fuertemente positiva, puesto que $\sum_{n=1}^{\infty}h_n^2/n^3\geq C\sum_{n=1}^{\infty}h_n^2$ para toda h implica $\sum_{n=1}^{\infty}h_n^2(n^{-3}-C)\geq 0$ para toda h y esto a su vez implica que $C=0$. En efecto, si $C>0$, entonces por la propiedad Arquimediana $n^{-3}-C<0$ para n suficientemente grande ($n\geq N$). Sea $h_n=0$ si $n<N$ y $h_n=n^{-1}$ si $n\geq N$, así $\sum_{n=1}^{\infty}n^{-2}(n^{-3}-C)<0$. Por lo tanto $C=0$

3.3 METODOS VARIACIONALES CLASICOS.

Una vez que hemos introducido el concepto de diferencial de una funcional definida sobre un conjunto abierto de un espacio normado, centramos ahora nuestra atención, al estudio de un tipo especial de funcionales, integrales cuyos integrandos son funciones conocidas de algún espacio, y como antes, estamos interesados en la búsqueda de condiciones necesarias y/o suficientes para que una funcional de este tipo alcance su extremo.

Definición. 3.3.1
(i). Sea $V=C^1[a,b]$ con la norma $\|f\|_1=\max_{x\in[a,b]}|f(x)|+\max_{x\in[a,b]}|f'(x)|$, O un subconjunto abierto de V y $F:O\subset V\to\mathbb{R}$ una funcional. Se dice que la funcional F tiene un extremo débil en el punto f, si existe un número $\epsilon>0$ tal que $F(f+h)-F(f)$ tiene el mismo signo para todo $f+h\in O$ tal que $\|f+h-f\|=\|h\|_1<\epsilon$.
(ii). Sea $V=C[a,b]$ con la norma $\|f\|_0=\max_{x\in[a,b]}|f(x)|$, O un subconjunto abierto de V y $F:O\subset V\to\mathbb{R}$ una funcional. Se dice que la funcional F tiene un extremo fuerte en el punto f, si existe un número $\epsilon>0$ tal que $F(f+h)-F(f)$ tiene el mismo signo para todo $f+h\in O$ tal que $\|f+h-f\|_0=\|h\|_0<\epsilon$.

Note que si J tiene un extremo fuerte en f, entonces existe un $\epsilon>0$, tal que $J(f)-J(g)$ tiene el mismo signo para todo g tal que $\|f-g\|_0<\epsilon$, de donde f es extremo también para todo $g\in C^1$, tal que $\|f-g\|_1<\epsilon$, ya que esta desigualdad implica que $\|f-g\|_0<\epsilon$. Así, todo extremo fuerte es también un extremo débil.

Definición. 3.3.2 Una función f definida sobre el intervalo cerrado $[a,b]$ es llamada continua por pedazos sobre $[a,b]$ si satisface que:
(i). $f(x)$ es acotada sobre $[a,b]$.
(ii). $\lim_{x\to x_0^+} f(x)$ existe para toda $x_0\in[a,b)$.
(iii). $\lim_{x\to x_0^-} f(x)$ existe para toda $x_0\in(a,b]$.
(iv). $f(x)$ es continua sobre (a,b).
El conjunto consistente de todas las funciones continuas por pedazos sobre $[a,b]$ será denotado por $PC[a,b]$.

<u>Lema I</u>. (3.3.3) Si $g_0\in PC[a,b]$ es una función fija, y si $\int_a^b g_0(x)h(x)dx=0$ para toda $h\in C_0[a,b]$, entonces $g_0(x)=0$ para toda $x\in[a,b]$.
Demostración: Supongamos que g_0 es no nula en algún punto de $[a,b]$, digamos $g_0(x_0)>0$ para algún $x_0\in[a,b]$. Entonces $g_0(x_0)>0$ para algún subintervalo $[x_1,x_2]\subset[a,b]$ que contiene a x_0. Pongamos $h(x)=(x-x_1)\cdot(x_2-x)$ para $x\in[x_1,x_2]$ y $h(x)=0$ para $x\in[a,b]-[x_1,x_2]$, obviamente $h\in C_0[a,b]$. Sin embargo $0=\int_a^b g_0(x)h(x)dx=\int_{x_1}^{x_2} g_0(x)((x-x_1)(x_2-x))dx>0$. Esta contradicción prueba el lema.

<u>Corolario</u>. 3.3.4 Si $g_0\in PC[a,b]$ es una función fija, y si $\int_a^b g_0(x)h(x)=0$ para toda $h\in C^n[a,b]\cap C_0[a,b]$, entonces $g_0(x)=0$ para todo $x\in[a,b]$.
La demostración es la misma del lema anterior excepto que definimos $h(x)=((x-x_1)(x_2-x))^{n+1}$ para $x\in[x_1,x_2]$ y $h(x)=0$ para $x\in[a,b]-[x_1,x_2]$.

<u>Lema II</u>. (3.3.5) Si $g_0\in PC[a,b]$ es una función fija, y si $\int_a^b g_0(x)h'(x)dx=0$ para toda $h\in C_0[a,b]$ con $h'\in PC[a,b]$, entonces $g_0(x)=c$ para toda $x\in[a,b]$, donde c es una constante.
Demostración: Sea c la constante definida por la condición $\int_a^b(g_0(x)-c)dx=0$ y sea $h(x)=\int_a^x(g_0(y)-c)dy$, así $h\in C_0[a,b]$ y $h'\in PC[a,b]$ automáticamente. Entonces, por una parte tenemos que $\int_a^b(g_0(x)-c)h'(x)dx=\int_a^b g_0(x)h'(x)dx-c(h(b)-h(a))=0$.
Mientras que por otra parte,
$\int_a^b(g_0(x)-c)h'(x)dx=\int_a^b(g_0(x)-c)^2dx$, se sigue entonces que $g_0(x)-c=0$ para toda $x\in[a,b]$.

<u>Lema III</u>. (3.3.6) Si $g_0\in PC[a,b]$ es una función fija, y si $\int_a^b g_0(x)h''(x)dx=0$ para toda $h\in C^1_{0,1}[a,b]$ con $h''\in PC[a,b]$, entonces $g_0(x)=c_0+c_1x$ para toda $x\in[a,b]$.
Demostración: Sean c_0 y c_1 definidas por las condiciones $\int_a^b(g_0(x)-c_0-c_1x)dx=0$, $\int_a^b dx\int_a^x(g_0(y)-c_0-c_1y)dy=0$, y sea $h(x)=\int_a^x dy\int_a^y(g_0(t)-c_0-c_1t)dt$, así $h\in C^1_{0,1}[a,b]$ y $h''\in PC[a,b]$ automáticamente.
La existencia de c_0 y c_1 se sigue de que

$$h'(x)=\int_a^x(g_0(y)-c_0-c_1y)dy=\int_a^x g_0(y)dy-c_0(x-a)-(c_1/2)(x^2-a^2), \text{ y}$$

$$h(x)=\int_a^x dy\int_a^y(g_0(t)-c_0-c_1t)dt=\int_a^x dy\int_a^y g_0(t)dt-c_0\left(\frac{x^2-a^2}{2}-a(x-a)\right)-$$

$-(c_1/2)(\frac{x^3-a^2}{3}-a^2(x-a))$, y si $h(b)=h'(b)=0$, entonces

$\int_a^b g_o(x)dx-c_o(b-a)-(c_1/2)(b^2-a^2)=0,$

$\int_a^b dy\int_a^y g_o(x)dx-c_o(\frac{b^2-a^2}{2}-a(b-a))-(c_1/2)(\frac{b^3-a^2}{3}-a^2(b-a))=0,$

y esto proporciona un sistema de ecuaciones simultáneas,

$((b-a)(b+a)/2)c_1+(b-a)c_o=\int_a^b g_o(x)dx,$

$((b-a)^2(b+2a)/6)c_1+((b-a)^2/2)c_o=\int_a^b dy\int_a^y g_o(x)dx.$

El determinante de los coeficientes c_o y c_1 está dado por

$((b-a)(b+a)/2)((b-a)^2/2)-(b-a)((b-a)^2(b+2a)/6)=(b-a)^4/12\neq 0.$

Por tanto existen c_o y c_1, definidos por las condiciones antes mencionadas.

Ahora bien, por una parte tenemos que

$\int_a^b(g_o(x)-c_o-c_1x)h''(x)dx=\int_a^b g_o(x)h''(x)dx-c_o(h'(b)-h'(a))-c_1\int_a^b xh''(x)dx$

$=-c_o(h'(b)-h'(a))-c_1(bh'(b)-ah'(a))+c_1(h(b)-h(a))=0.$

Mientras que por otra parte,

$\int_a^b(g_o(x)-c_o-c_1x)h''(x)dx=\int_a^b(g_o(x)-c_o-c_1x)^2dx=0$, se sigue que $g_o(x)-c_o-c_1x=0$ para toda $x\in[a,b]$, es decir $g_o(x)=c_o+c_1x$ para toda $x\in[a,b]$.

Corolario. 3.3.7 Si $g_o\in PC[a,b]$ es una función fija, y si $\int_a^b g_o(x)h^{(n)}(x)dx=0$ para toda $h\in C_{o,n-1}^{(n-1)}[a,b]$ con $h^{(n)}\in PC[a,b]$, entonces $g_o(x)=\sum_{k=0}^{n-1} c_kx^k$, esto es, $g_o(x)$ es un polinomio de grado n-1.

La demostración es por inducción sobre n.

Lema IV. (3.3.8) Si $g_1,g_o\in PC[a,b]$ son funciones fijas, y si $\int_a^b(g_1(x)h(x)-g_o(x)h'(x))dx=0$ para toda $h\in C_o[a,b]$ con $h'\in PC[a,b]$, entonces $g_o'\in PC[a,b]$ y $g_o'(x)=g_1(x)$ para toda $x\in[a,b]$.

Demostración: Pongamos $A(x)=\int_a^x g_1(y)dy$, e integrando por partes a $\int_a^b g_1(x)h(x)dx$ obtenemos que $\int_a^b g_1(x)h(x)dx=$

$A(x)h(x)\Big|_a^b-\int_a^b A(x)h'(x)dx=-\int_a^b A(x)h'(x)dx$, entonces si $\int_a^b(g_1(x)h(x)-g_o(x)h'(x))dx=0$, tenemos que $\int_a^b g_1(x)h(x)dx=\int_a^b g_o(x)$ $h'(x)dx$, de donde $\int_a^b A(x)h'(x)dx=\int_a^b g_o(x)h'(x)dx$ y esto implica que $\int_a^b(A(x)-g_o(x))h'(x)dx=0$ y por el Lema II tenemos que $A(x)-g_o(x)=c$, para toda $x\in[a,b]$ y por lo tanto $g_o'(x)=g_1(x)$ para toda $x\in[a,b]$, enfatizamos que la diferenciabilidad de g_o no había sido supuesta de antemano.

Planteamos ahora, un problema concreto de los métodos variacionales. Sea $F(x,f,g)$ una función con primera y segunda derivadas parciales continuas con respecto a todos sus argumentos sobre el conjunto $[a,b]\times O\times O'$, con $x\in[a,b]$ y O, O' son dos conjuntos abiertos de dos espacios de funciones V, V', tales que $f\in O\subset V$ y $g\in O'\subset V'$, esto es $F\in C^2([a,b]\times O\times O')$. Si $f\in O=B(f,\varepsilon)\subset C^1[a,b]$ y satisface las condiciones de frontera $f(a)=A$, $f(b)=B$, y $J:O\subset C^1[a,b]\to\mathbb{R}$ es una funcional de la forma $J(f)=\int_a^b F(x,f,f')dx$, deseamos encontrar una función f para la cual la funcional J tenga un extremo débil.

<u>Teorema</u>. 3.3.9 Sea $(V,\|\cdot\|)$ el espacio de Banach constituido por el espacio lineal $V=C^1[a,b]$ con la norma $\|f\|=\max_{x\in[a,b]}|f(x)|+\max|f'(x)|$, y $J:O=B(f,\varepsilon)\subset C^1[a,b]\to\mathbb{R}$ una funcional de la forma $J(f)=\int_a^b F(x,f,f')dx$, donde $F\in C^2([a,b]\times O\times O')$, $O=B(f,\varepsilon)\subset C^1[a,b]$ y $f'\in O'\subset C[a,b]$, si f satisface las condiciones de frontera $f(a)=A$, $f(b)=B$. Entonces una condición necesaria para que la funcional J tenga un extremo en una función dada f, es que f satisfaga la ecuación de Euler.

$\frac{\partial F}{\partial f}(x,f,f')-\frac{d}{dx}\left(\frac{\partial F}{\partial f'}(x,f,f')\right)=0.$

Demostración: Demos a la función f un incremento h, $\|h\|<\varepsilon$, de tal manera que para $f+h$ se sigan satisfaciendo las condiciones de frontera, entonces $h(a)=h(b)=0$, esto es $h\in C_0[a,b]\cap C^1[a,b]\cap B(0,\varepsilon)$, en la terminología de los métodos variacionales una tal función h es llamada una función admisible, entonces calculando la diferencia $J(f+h)-J(f)$ tenemos que $J(f+h)-J(f)=\int_a^b (F(x,f+h,f'+h')-F(x,f,f'))dx$, se sigue del teorema de Taylor que

$J(f+h)-J(f)=\int_a^b \left(\frac{\partial F}{\partial f}(x,f,f')h(x)+\frac{\partial F}{\partial f'}(x,f,f')h'(x)\right)dx+$
$\frac{1}{2}\int_a^b \left(\frac{\partial^2 F}{\partial f^2}(x,f+\theta h,f'+\theta h')h^2(x)+2\frac{\partial^2 F}{\partial f\partial f'}(x,f+\theta h,f'+\theta h')h(x)h'(x)\right.$
$\left.+\frac{\partial^2 F}{\partial f'^2}(x,f+\theta h,f'+\theta h')(h'(x))^2\right)dx$, donde $\theta\in[0,1]$.

Además sabemos que existen M_1, M_2 y M_3 tales que $\left|\frac{\partial^2 F}{\partial f^2}(x,f,f')\right|\le M_1$, $\left|\frac{\partial^2 F}{\partial f\partial f'}(x,f,f')\right|\le M_2$ y $\left|\frac{\partial^2 F}{\partial f'^2}(x,f,f')\right|\le M_3$ para todo (x,f,f'), pues F tiene segundas derivadas continuas con respecto a todos sus argumentos. Tomemos $M=\max\{M_i\}$, así en valor absoluto la segunda integral es menor o igual que $M(b-a)\|h\|^2=o(\|h\|)$. Entonces $\delta J(f)(h)=\int_a^b\left(\frac{\partial F}{\partial f}(x,f,f')h(x)\right.$ $\left.+\frac{\partial F}{\partial f'}(x,f,f')h'(x)\right)dx$, y una condición necesaria para que la funcional J tenga un extremo en el punto f es que $\delta J(f)(h)=0$ para toda

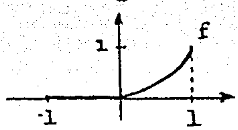
h admisible en $C_0[a,b]\cap C^1[a,b]\cap B(0,\epsilon)$, entonces
$\int_a^b(\frac{\partial F}{\partial f}(x,f,f')h(x)-\frac{\partial F}{\partial f'}(x,f,f')h'(x))dx=0$ para toda h admisible,
pero integrando por partes el segundo sumando del integrando tenemos que $\int_a^b \frac{\partial F}{\partial f'}(x,f,f')h'(x)dx= \frac{\partial F}{\partial f'}(x,f,f')h(x)\Big|_a^b-\int_a^b \frac{d}{dx}\frac{\partial F}{\partial f'}(x,f,f'))h(x)dx$
$=\int_a^b \frac{d}{dx}(\frac{\partial F}{\partial f'}(x,f,f'))h(x)dx$, así
$\int_a^b(\frac{\partial F}{\partial f}(x,f,f')+\frac{d}{dx}(\frac{\partial F}{\partial f'}(x,f,f'))h(x)dx=0$ para toda h admisible, luego
por el lema I se tiene que $\frac{\partial F}{\partial f}(x,f,f')-\frac{d}{dx}(\frac{\partial F}{\partial f'}(x,f,f'))=0$,
la existencia de la derivada $\frac{d}{dx}(\frac{\partial F}{\partial f'})$ se sigue del Lema IV.

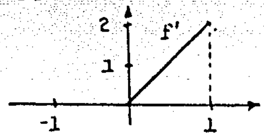
La curvas integrales de la ecuación de Euler son llamadas curvas extremales. Claramente esta ecuación diferencial es de segundo orden, ya que $\frac{d}{dx}F_{f'}-F_f=F_{f'f'}f''+F_{ff'}f'+F_{xf'}-F_f=0$, entonces sus soluciones dependen de dos constantes arbitrarias, las cuales son determinadas de las condiciones de frontera $f(a)=A$, $f(b)=B$.
La ecuación de Euler proporciona una condición necesaria para un extremo, sin embargo, en muchos casos ella es suficiente para ser una solución completa del problema variacional. De hecho, la existencia de un extremo se determina frecuentemente por las característica de el problema. Si en tales casos sólo existe una curva extremal la cual satisface las condiciones de frontera del problema, entonces esta curva extremal debe ser aquella para la cual el extremo es alcanzado.

Para una funcional $J:C^1[a,b]\to\mathbb{R}$ de la forma $J(f)=\int_a^b F(x,f,f')dx$.
La ecuación de Euler es en general una ecuación diferencial de segundo orden. Pero puede ser que para una curva dada la funcional tenga un extremo y dicha curva no sea dos veces diferenciable y sin embargo, satisface la ecuación de Euler, como lo muestra el siguiente ejemplo.

Ejemplo. 3.3.10 Consideremos la funcional $J(f)=\int_{-1}^{1} f^2(2x-f')^2dx$, definida sobre $C^1[-1,1]$, donde las condiciones de frontera están dadas por $f(-1)=0$, $f(1)=1$.
El mínimo de la funcional J es igual a cero y éste es alcanzado para la función

$$f=f(x)=\begin{cases}0 & \text{si } -1\le x\le 0\\ x^2 & \text{si } 0\le x\le 1.\end{cases}$$





la cual no tiene segunda derivada para $x=0$. Sin embargo, $f(x)$ satisface la ecuación de Euler. En efecto, ya que en este caso $F(x,f,f')=f^2(2x-f')^2$, y de esto se sigue que las funciones $F_f=2f(2x-f')^2$, $F_{f'}=-2f^2(2x-f')$ y $\frac{d}{dx}(F_{f'})$ se anulan idénticamente para $x\in[-1,1]$. Así, a pesar de que la ecuación de Euler es una ecuación diferencial de segundo orden y $f''(x)$ no existe en todas partes en $[-1,1]$, la sustitución de $f(x)$ en la ecuación de Euler la convierte en una identidad.

Ahora daremos condiciones las cuales garantizan que una solución de la Ecuación de Euler tenga una segunda derivada.

<u>Teorema</u>. 3.3.11 Supongamos que $f\in O=B(f,\epsilon)\subset C^1[a,b]$ y que satisface la ecuación de Euler $F_f-\frac{d}{dx}F_{f'}=0$. Entonces si $F\in C^2([a,b]\times O\times O')$, $O\subset C^1[a,b]$ y $f'\in O'\subset C[a,b]$, $f=f(x)$ tiene una segunda derivada continua en todos los puntos $(x,f(x))$ en donde $F_{f'f'}(x,f,f')\neq 0$.

Demostración: Supongamos que $f+\Delta f\in O$, calculemos la diferencia

$$\Delta F_{f'}=F_{f'}(x+\Delta x,f+\Delta f,f'+\Delta f')-F_{f'}(x,f,f')=\Delta x F_{f'x}(x+\theta\Delta x,f+\theta\Delta f,f'+\theta\Delta f')$$
$$+\Delta f F_{f'f}(x+\theta\Delta x,f+\theta\Delta f,f'+\theta\Delta f')+\Delta f' F_{f'f'}(x+\theta\Delta x,f+\theta\Delta f,f'+\theta\Delta f')$$ en donde

$\theta\in[0,1]$. Ahora dividamos esta diferencia entre Δx, y consideremos el límite cuando $\Delta x\to 0$ de la expresión resultante, así obtenemos

$$\lim_{\Delta x\to 0}\frac{F_{f'}}{\Delta x}=\lim_{\Delta x\to 0}\left(F_{f'x}(x+\theta\Delta x,f+\theta\Delta f,f'+\theta\Delta f'+\theta\Delta f')+\frac{\Delta f}{\Delta x}F_{f'f}(x+\theta\Delta x,f+\theta\Delta f,f'+\theta\Delta f')\right)$$
$$+\lim_{\Delta x\to 0}\left(\frac{\Delta f'}{\Delta x}F_{f'f'}(x+\theta\Delta x,f+\theta\Delta f,f'+\theta\Delta f')\right).$$

Este límite existe, puesto que $F_{f'}$ tiene derivada con respecto a x de acuerdo con la ecuación de Euler, ya que $\frac{d}{dx}(F_{f'})=F_f$. Por hipótesis las segundas derivadas de F son continuas, y debido a la existencia de $f'(x)$, se sigue que el límite de el primer paréntesis existe. Por lo tanto el límite del segundo paréntesis existe, ya que el límite de la suma de los otros dos términos existe, y como $F_{f'f'}\neq 0$, se tiene que $f''(x)$ existe. Finalmente de la ecuación

$\frac{d}{dx}F_{f'}-F_f=0$ es posible encontrar una expresión para f'', y de esta expresión es claro que f'' es continua si $F_{f'f'}\neq 0$, puesto que

$\frac{d}{dx}F_{f'}-F_f=F_{f'f'}f''+F_{ff'}f'+F_{xf'}-F_f=0$. Esto prueba el Teorema.

Ahora indicaremos algunos casos especiales, en donde la ecuación de Euler puede ser reducida a una ecuación diferencial de primer orden, y por lo tanto sus soluciones pueden ser obtenidas completamente en términos de integrales.

Caso I. (3.3.12) Sea $J:C^1[a,b]\to\mathbb{R}$ una funcional de la forma $J(f)=\int_a^b F(x,f')dx$, en donde $F(x,f')$ tiene primera y segunda derivadas continuas para todos sus argumentos, en este caso la ecuación de Euler es $\frac{d}{dx}F_{f'}=0$, entonces $F_{f'}=C$, donde C es una constante. Esta es una ecuación diferencial de primer orden la cual no contiene a f, resolviendola para f' se obtiene una ecuacion de la forma $f'=g(x,C)$, de donde f puede ser encontrada en términos de una integral.

Caso II. (3.3.13) Sea $J:C^1[a,b]\to\mathbb{R}$ una funcional de la forma $J(f)=\int_a^b F(f,f')dx$, en donde $F(f,f')$ tiene primera y segunda derivadas continuas con respecto a todos sus argumentos, entonces

$F_f-\frac{d}{dx}F_{f'}=F_f-F_{f'f}f'-F_{f'f'}f''$ y multiplicando por f', obtenemos que

$F_ff'-(\frac{d}{dx}F_{f'})f'=F_ff'-F_{f'f}f'^2-F_{f'f'}f'f''=\frac{d}{dx}(F-f'F_{f'})$;

de donde la ecuación de Euler se convierte en $\frac{d}{dx}(F-f'F_{f'})=0$, entonces $F-f'F_{f'}=C$, donde C es una constante.

En el siguiente caso la ecuación de Euler toma la forma de una ecuación algebraica, cuyas soluciones consisten de una o más curvas $f=f(x)$.

Caso III. (3.3.14) Sea $J:C^1[a,b]\to\mathbb{R}$ una funcional de la forma $J(f)=\int_a^b F(x,f)dx$, en donde $F(x,f)$ tiene primera y segunda derivadas continuas con respecto a todos sus argumentos, en este caso la ecuación de Euler toma la forma $F_f(x,f)=0$, que es una ecuación algebraica, cuyas soluciones son una o más curvas.

En una gran variedad de problemas encontraremos funcionales de la forma $J(f)=\int_a^b g(x,f)(1+f'^2)^{1/2}dx$, la cual se analiza en el siguiente caso.

Caso IV. (3.3.15) Sea $J:C^1[a,b] \to \mathbb{R}$ una funcional de la forma $J(f)=\int_a^b g(x,f)(1+f'^2)^{\frac{1}{2}}dx$, en donde la función $g(x,f)$ tiene primera y segunda derivadas continuas con respecto a todos sus argumentos, en este caso la ecuación de Euler puede ser transformada en

$$\frac{\partial F}{\partial f}-\frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right)=g_f(x,f)(1+f'^2)^{\frac{1}{2}}-\frac{d}{dx}\left(g(x,f)f'(1+f'^2)^{-\frac{1}{2}}\right)$$

$$=g_f(x,f)(1+f'^2)^{\frac{1}{2}}-g_x(x,f)f'(1+f'^2)^{-\frac{1}{2}}-g_f(x,f)f'^2(1+f'^2)^{-\frac{1}{2}}$$

$$-g(x,f)f''(1+f'^2)^{-3/2}$$

$$=(1+f'^2)^{-\frac{1}{2}}\left(g_f(x,f)-g_x(x,f)f'-g(x,f)f''(1+f'^2)^{-1}\right)=0,$$

es decir, $g_f(x,f)-g_x(x,f)f'-g(x,f)f''(1+f'^2)^{-1}=0$.

Ejemplo. 3.3.16 (Problema de la Braquistocrona) Uno de los primeros problemas que impulsaron el desarrollo de los métodos variacionales es el siguiente: Una partícula se mueve bajo la influencia de la gravedad a lo largo de una curva que une un punto dado A con un punto dado B, deseamos encontrar una curva tal que el tiempo que toma la partícula de ir de A a B sea mínimo.

Sea m la masa de la partícula y sea v su velocidad, y g la constante de gravitación. La ley de conservación de la energía da $\frac{1}{2}mv^2-mgf=0$, entonces $v=(2gf)^{\frac{1}{2}}$ ó $\frac{ds}{dt}=(2gf)^{\frac{1}{2}}$, donde s es la distancia viajada en el tiempo t, el tiempo tomado es obtenido de $dt=(ds)/v$, en donde ds es el elemento de arco recorrido en el tiempo dt. Sea A representado por $x=0$ y $B=f(b)$ para $x=b$. Entonces el tiempo total tomado, el cual es función de la trayectoria $f(x)$, está dado por $J(f)=\int_0^b (ds)/v=\int_0^b \frac{(1+f'^2)^{\frac{1}{2}}}{(2gf)^{1/2}}dx$, en donde hemos usado la fórmula $\left(\frac{ds}{dx}\right)^2=1+\left(\frac{df}{dx}\right)^2$. El problema es entonces encontrar una función diferenciable definida sobre $[0,b]$ que minimice a la funcional J.

Observamos que en este caso $F(x,f,f')=\frac{(1+f'^2)^{\frac{1}{2}}}{(2gf)^{1/2}}$, y que $\frac{\partial F}{\partial x}=0$. entonces una condición necesaria para que una función f sea extremo es que satisfaga la ecuación de Euler, así

$$f'\left(\frac{\partial F}{\partial f}-\frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right)\right)=\frac{\partial F}{\partial x}+\frac{\partial F}{\partial f}f'+\frac{\partial F}{\partial f'}f''-f''\frac{\partial F}{\partial f'}-f'\frac{d}{dx}\left(\frac{\partial F}{\partial f'}\right)=\frac{d}{dx}\left(F-f'\frac{\partial F}{\partial f'}\right)=0, \text{ entonces}$$

$F-f'\frac{\partial F}{\partial f'}=C$, donde C es una constante, así

$(\frac{1+f'^2}{f})^{\frac{1}{2}} - \frac{f'^2}{(1+f'^2)^{1/2}} f^{-\frac{1}{2}} = C$ ó $(f(1+f'^2))^{-\frac{1}{2}} = C'^{-\frac{1}{2}}$, así $f' = ((C'-f)/f)^{\frac{1}{2}}$

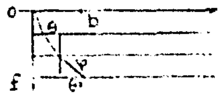
o bien $\frac{df}{dx} = (\frac{C'}{f} - 1)^{\frac{1}{2}}$, entonces $\frac{dx}{df} = (\frac{C'}{f} - 1)^{-\frac{1}{2}}$, hagamos la sustitución

$f = C' \operatorname{sen}^2 \theta/2$, e integrando obtenemos que

$x = \frac{1}{2}C' \int (1-\cos\theta)\, d\theta$, puesto que $\operatorname{sen}^2\theta/2 = \frac{1}{2}(1-\cos\theta)$, de donde

$x = \frac{1}{2}C'(\theta - \operatorname{sen}\theta) + C_1$ y $f = \frac{1}{2}C'(1-\cos\theta)$, y dado que la curva parte del origen, $C_1 = 0$, frecuentemente se utiliza $f = y$. C' es determinada por $f(b) = B$. La curva así definida paramétricamente es una cicloide.

Ahora consideremos el siguiente razonamiento que sigue una línea completamente diferente, primero aproximamos la trayectoria $f=f(x)$ por una serie de segmentos de recta que dividen la distancia de caída en partes iguales, como lo muestra la figura.

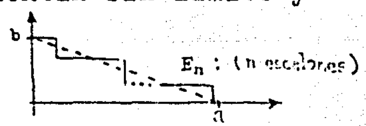


La velocidad de la partícula es supuesta constante a través de cada segmento de recta. Consideremos ahora la ley de refracción de Snell $\operatorname{sen}\theta/v = \operatorname{sen}\theta'/v' =$ constante. Esta es una condición para que la trayectoria de tiempo mínimo cruce una discontinuidad. Tomando el límite cuando estos segmentos se hacen cada vez más pequeños, inferimos que la curva estará dada por $\operatorname{sen}\theta/(2gf)^{\frac{1}{2}} =$ constante, entonces

$\frac{\operatorname{sen}\theta}{\sqrt{f}} = C$ o $\frac{\cos\varphi}{\sqrt{f}} = C$ o $\frac{1}{\sec\varphi\sqrt{f}} = C$, entonces $(f(1+f'^2))^{\frac{1}{2}} = \frac{1}{C}$, de donde

$f' = (\frac{1}{fC^2} - 1)^{\frac{1}{2}}$, y poniendo $C' = 1/C^2$ obtenemos $\frac{df}{dx} = (\frac{C'}{f} - 1)^{-\frac{1}{2}}$, que es el resultado previamente obtenido por medio de la Ecuación de Euler.

Hacemos notar también que en nuestro primer razonamiento partimos de una condición necesaria, y a pesar de esto, obtuvimos la solución completa del problema, y que en nuestro segundo razonamiento se planteó una sucesión de problemas $\{P_n\}$ y se dió su respectiva sucesión de soluciones $\{S_n\}$. En este caso, si $P_n \to P$ y $S_n \to S$ se tiene que S es solución de P. En general lo anterior no es válido. En efecto, sea P_n: Determinar la longitud de la escalera E_n mostrada abajo. Así $S_n = \text{long}(E_n) = a+b$. Sin embargo, si los escalones se incrementan sin límite y se hacen cada vez más pequeños $\lim_{n\to\infty} S_n = (a^2+b^2)^{\frac{1}{2}}$.



Ejemplo. 3.3.17 (El problema de la superficie de revolución de área mínima).
De entre todas las curvas las cuales unen dos puntos dados (a,A) y (b,B), deseamos encontrar una, la cual genere la superficie de área mínima cuando esta rota sobre el eje x.
Es bien conocido, que el área de una superficie de revolución generada al hacer rotar la curva $f=f(x)\geq 0, f\in C^1[a,b]$, sobre el eje x viene dada por la funcional
$J(f)=2\pi\int_a^b f(1+f'^2)^{\frac{1}{2}}dx.$
Puesto que el integrando no depende explícitamente de x. La ecuación de Euler tiene la forma $F-f'F_{f'}=C$, en donde $F(x,f,f')=2\pi f(1+f'^2)^{\frac{1}{2}}$ (Caso II y problema anterior), esto da que

$f(1+f'^2)^{\frac{1}{2}}-ff'^2(1+f'^2)^{-\frac{1}{2}}=C$ ó $f=C(1+f'^2)^{\frac{1}{2}}$, así que

$f'=(\frac{f^2-C^2}{C^2})^{\frac{1}{2}}$, separando variables obtenemos $dx=\frac{Cdf}{(f^2-C^2)^{\frac{1}{2}}}$, e integrando

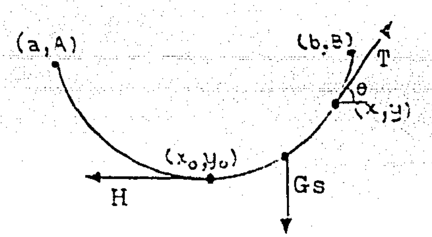
tenemos que $x+C'=C \operatorname{arc\,cosh}\frac{f}{C}$ ó

$f=C\cosh(\frac{x+C'}{C})$. Los valores de las constantes arbitrarias C y C'

son determinados por las condiciones $f(a)=A$ y $f(b)=B$.
La curva así definida es una catenaria.

Consideremos ahora el siguiente razonamiento.
El nombre de catenaria proviene del hecho de que una cadena suspendida de sus dos extremos adopta la forma de esta curva. Es curioso que la misma curva surja en una aplicación física completamente diferente. Una película de jabón limitada por dos círculos en el espacio que estan en planos paralelos y con sus centros sobre la misma perpendicular a los planos, tiene la forma de una superficie de revolución obtenida al hacer girar la catenaria alrededor del eje x,

Consideremos la siguiente figura:



Si tal superficie ha de ser de área mínima, entonces la tensión en cada punto de la superficie ha de ser mínima (como en la película de jabón). Tomemos primero una curva sobre la superficie, tal que su proyección esté sobre el eje x, y sea (x,y) un punto arbitrario sobre la curva, y (x_0,y_0) es el punto más bajo sobre la curva, entonces sobre el segmento s determinado por los puntos (x_0,y_0) y (x,y) actúan tres fuerzas en equilibrio. La tensión H constante en el punto más bajo de la curva, la tensión T en el punto (x,y) en la dirección de la tangente y el peso Gs del segmento s determinado por los puntos (x_0,y_0) y (x,y), en donde G es el peso específico lineal del segmento de curva s, entonces

$T\cos\theta=H$ y $T\text{sen}\theta=Gs$, así $\tan\theta=\frac{G}{H}s$ y dado que la tensión H es constante ponemos $\frac{G}{H}=\frac{1}{C}$, entonces

$\frac{df}{dx}=\frac{1}{C}s$, y $\frac{d^2f}{dx^2}=\frac{1}{C}\frac{ds}{dx}=\frac{1}{C}(1+f'^2)^{\frac{1}{2}}$. Si $\frac{df}{dx}=p$, entonces $\frac{d^2f}{dx^2}=\frac{dp}{dx}$, así

$\frac{dp}{dx}=\frac{1}{C}(1+p^2)^{\frac{1}{2}}$ y separando variables tenemos que $\frac{1}{C}dx=(1+p^2)^{-\frac{1}{2}}$,

e integrando obtenemos $\frac{x}{C}+\frac{C'}{C}=\text{arc senh } p$ o $p=\text{senh}(\frac{x+C'}{C})$ y dado que

$p=\frac{df}{dx}$ integrando de nuevo tenemos que $f=C\cosh(\frac{x+C'}{C})$.

Este resultado coincide con el resultado previamente obtenido.

La superficie de revolución obtenida al hacer girar la catenaria alrededor del eje x es llamada catenoide.

Como antes, hacemos notar, que en nuestro primer razonamiento partimos de una condición necesaria (de extremo), y sin embargo, se obtuvo la solución completa del problema, y que en el segundo razonamiento se hizo uso de un hecho físico.

Hasta aquí, se han considerado funcionales cuyo dominio de definición consiste en un conjunto abierto de algún espacio de funciones de una variable independiente, es decir curvas. En muchos problemas, sin embargo, se encuentran funcionales definidas sobre funciones de varias variables independientes, digamos, sobre superficies. En lo que sigue, tales problemas de carácter multidimensional, serán considerados en detalle.

Planteamos ahora, el problema variacional anterior para una función $F(x,y,f,g,h)$ con primera y segunda derivadas parciales continuas con respecto a todos sus argumentos sobre el conjunto $\Omega\times O\times O'$, en donde $(x,y)\in\Omega\subset\mathbb{R}^2$, Ω un conjunto cerrado y conexo y O, O' dos conjuntos abiertos de dos espacios de funciones V, V', tales que $f\in O\subset V$ y $(g,h)\in O'\subset C(\Omega)\times C(\Omega) = (C(\Omega))^2$, esto es, $F\in C^2(\Omega\times O\times O')$. Si $f:\Omega\to\mathbb{R}$, $f\in O=B(f,\epsilon)\subset C^1(\Omega)$ y satisface las condiciones de frontera $f(x,y)=w(x,y)$, $(x,y)\in\partial\Omega$, siendo w una función conocida. Si $J:O=B(f,\epsilon)\subset C^1(\Omega)\to\mathbb{R}$ es una funcional de la forma $J(f)=\iint_\Omega F(x,y,f,f_x,f_y)\,dxdy$, deseamos encontrar una función f para la cual la funcional J alcance un extremo (débil).

<u>Lema</u> 3.3.18 Si $g_0\in C(\Omega)$ es una función fija, donde Ω es un conjunto cerrado y conexo, y si $\iint_\Omega g_0(x,y)h(x,y)\,dxdy=0$ para todo $h\in C^1(\Omega)\cap C_0(\Omega)$ ($C_0(\Omega)$ es el conjunto de funciones definidas en Ω las cuales se anulan en $\partial\Omega$, la frontera de Ω), entonces $g_0(x,y)=0$ para todo $(x,y)\in\Omega$.

Demostración: Supongamos que g_0 es no nula en algún punto de Ω, digamos $g_0(x_0,y_0)>0$ para algún $(x_0,y_0)\in\Omega$, entonces $g_0(x,y)>0$ para alguna bola cerrada $\bar{B}((x_0,y_0),\epsilon)=\{(x,y);(x-x_0)^2+(y-y_0)^2\leq\epsilon^2\}$ contenida en Ω, con centro en el punto (x_0,y_0) y radio ϵ. Pongamos $h(x,y)=0$ para todo $(x,y)\in\Omega-\bar{B}((x_0,y_0),\epsilon)$, y $h(x,y)=((x-x_0)^2+(y-y_0)^2-\epsilon^2)^2$ para todo $(x,y)\in\bar{B}((x_0,y_0),\epsilon)$, así h satisface automáticamente las condiciones del Lema, es decir $h\in C^1(\Omega)\cap C_0(\Omega)$. Sin embargo

$$0=\iint_\Omega g_0(x,y)h(x,y)\,dxdy=\iint_\Omega g_0(x,y)((x-x_0)^2+(y-y_0)^2-\epsilon^2)^2\,dxdy =$$
$$\iint_{\bar{B}} g_0(x,y)((x-x_0)^2+(y-y_0)^2-\epsilon^2)^2>0.$$ Esta contradicción prueba el lema.

<u>Teorema</u>. 3.3.19 Sea $(V,\|\cdot\|)$ el espacio de Banach constituido por el espacio lineal $V=C^1(\Omega)$, $\Omega\subset\mathbb{R}^2$ un conjunto cerrado y conexo, con la norma $\|f\|=\max_{(x,y)\in\Omega}|f(x,y)|+\max_{(x,y)\in\Omega}|f_x(x,y)|+\max_{(x,y)\in\Omega}|f_y(x,y)|$, y $J:O=B(f,\epsilon)\subset C^1(\Omega)\to\mathbb{R}$ una funcional de la forma $J(f)=\iint_\Omega F(x,y,f,f_x,f_y)\,dxdy$, en donde $F\in C^2(\Omega\times O\times O')$, $f\in O=B(f,\epsilon)\subset C^1(\Omega)$ y $(f_x,f_y)\in O'\subset C(\Omega)\times C(\Omega)$, si f satisface la condición de frontera $f(x,y)=w(x,y)$, con $(x,y)\in\partial\Omega$, siendo w una función conocida. Entonces una condición necesaria para que la funcional J alcance un extremo en una función dada f, es que f satisfaga la ecuación de Euler $\frac{\partial F}{\partial f}-\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial f_x}\right)-\frac{\partial}{\partial y}\left(\frac{\partial F}{\partial f_y}\right)=0$.

Demostración:

Demos a la función f un incremento h, $\|h\|<\epsilon$, de tal manera que para $f+h$ se siga satisfaciendo la condición de frontera, entonces $h(x,y)=0$ para toda $(x,y)\in\partial\Omega$, esto es, $h\in C^1(\Omega)\cap C_0(\Omega)\cap B(0,\epsilon)$.

Calculemos ahora la diferencia

$J(f+h)-J(f)=\iint_\Omega (F(x,y,f+h,f_x+h_x,f_y+h_y)-F(x,y,f,f_x,f_y))\,dxdy$, se sigue del teorema de Taylor que

$J(f+h)-J(f)=\iint_\Omega (F_f h(x,y)+F_{f_x}h_x(x,y)+F_{f_y}h_y(x,y))\,dxdy+o(\|h\|)$. Y por lo tanto la diferencial de J es igual a $dJ(f)(h)=\iint_\Omega (F_f h(x,y)+F_{f_x}h_x(x,y)+F_{f_y}h_y(x,y))\,dxdy$.

Ahora bien, usando el Teorema de Green tenemos que

$$\iint_\Omega (F_{f_x}h_x+F_{f_y}h_y)\,dxdy=\iint_\Omega \left(\frac{\partial}{\partial x}(F_{f_x}h)-\frac{\partial}{\partial x}F_{f_x}h+\left(\frac{\partial}{\partial y}(F_{f_y}h)-\frac{\partial}{\partial y}F_{f_y}h\right)\right)dxdy$$

$$=\iint_\Omega \left(\frac{\partial}{\partial x}(F_{f_x}h)+\frac{\partial}{\partial y}(F_{f_y}h)\right)dxdy-\iint_\Omega \left(\frac{\partial}{\partial x}F_{f_x}+\frac{\partial}{\partial y}F_{f_y}\right)h\,dxdy$$

$$=\int_{\partial\Omega} F_{f_x}h\,dy-F_{f_y}h\,dx-\iint_\Omega \left(\frac{\partial}{\partial x}F_{f_x}+\frac{\partial}{\partial y}F_{f_y}\right)h\,dxdy,$$ y puesto que $h(x,y)=0$ para toda $(x,y)\in\partial\Omega$, tenemos que

$\iint_\Omega (F_{f_x}h_x+F_{f_y}h_y)\,dxdy=-\iint_\Omega \left(\frac{\partial}{\partial x}F_{f_x}+\frac{\partial}{\partial y}F_{f_y}\right)h\,dxdy$, entonces

$dJ(f)(h)=\iint_\Omega \left(F_f-\frac{\partial}{\partial x}F_{f_x}-\frac{\partial}{\partial y}F_{f_y}\right)h(x,y)\,dxdy$, y una condición necesaria para que la funcional J tenga un extremo en el punto f es que $dJ(f)(h)=0$ para toda h admisible ($h\in C^1(\Omega)\cap C_0(\Omega)\cap B(0,\epsilon)$), así por el lema anterior, tenemos que $F_f-\frac{\partial}{\partial x}F_{f_x}-\frac{\partial}{\partial y}F_{f_y}=0$. Lo que concluye la prueba del teorema.

Ejemplo. 3.3.20 Superficies mínimas (Problema de Plateau)
De entre todas las superficies parametrizadas por $H(x,y)=(x,y,f(x,y))$ tales que la diferencial dH es inyectiva (tal condición, es conocida como regularidad), donde $(x,y)\in\Omega$, Ω un conjunto cerrado y conexo. deseamos encontrar $f=f(x,y)$, $f\in C^2(\Omega)$ que minimice a la funcional $J(f)=\iint_\Omega (1+f_x^2+f_y^2)^{\frac{1}{2}}dxdy$.

Primero vemos que el área de una superficie regular, con la parametrización $H(x,y)=(x,y,f(x,y))$, se define como el número positivo $\iint_\Omega \|H_x\times H_y\|dxdy$. Los coeficientes de la primera forma cuadrática fundamental están dados por $E=(H_x,H_x)$, $F=(H_x,H_y)$ y $G=(H_y,H_y)$ y observamos que $\|H_x\times H_y\|^2+(H_x,H_y)^2=\|H_x\|^2\|H_y\|^2$. Por lo tanto $\|H_x\times H_y\|=(EG-F^2)^{\frac{1}{2}}$. Así $E=(H_x,H_x)=(1,0,f_x)\cdot(1,0,f_x)=1+f_x^2$, $F=f_xf_y$ y $G=1+f_y^2$, entonces $J(f)=\iint_\Omega \|H_x\times H_y\|\,dxdy=\iint_\Omega (EG-F^2)^{\frac{1}{2}}dxdy=\iint_\Omega (1+f_x^2+f_y^2)^{\frac{1}{2}}dxdy$.
En este caso la ecuación de Euler tiene la forma
$f_{xx}(1+f_y^2)-2f_{xy}f_xf_y+f_{yy}(1+f_x^2)=0$, esta ecuación tiene un significado geométrico simple, sabemos que la curvatura media de una superficie viene dada por la fórmula
$M=\dfrac{Eg-2Fk+Ge}{2(EG-F^2)}$, en donde e,k,g son los coeficientes de la segunda forma cuadrática fundamental de la superficie, $e=(H_{xx},N)$, $k=(H_{xy},N)$ y $g=(H_{yy},N)$ con
$N=\dfrac{H_x\times H_y}{\|H_x\times H_y\|}=(-f_x,-f_y,1)/(1+f_x^2+f_y^2)^{\frac{1}{2}}$,
entonces $e=\dfrac{f_{xx}}{(1+f_x^2+f_y^2)^{1/2}}$, $k=\dfrac{f_{xy}}{(1+f_x^2+f_y^2)^{1/2}}$ y $g=\dfrac{f_{yy}}{(1+f_x^2+f_y^2)^{1/2}}$
y por lo tanto $M=\dfrac{f_{yy}(1+f_x)-2f_{xy}f_xf_y+f_{xx}(1+f_y)}{(1+f_x+f_y^2)^{\frac{3}{2}}}$,
esto implica que la curvatura media de la superficie que minimice a J debe ser igual a cero. En geometría diferencial tales superficies son llamadas superficies mínimas.

Mencionamos aquí, el problema de Dirichlet, el cual consiste en minimizar $\iint_\Omega (f_x^2+f_y^2)dxdy$ sujeto a $f=g$ en $\partial\Omega$ (g conocida, Ω una región). La ecuación de Euler Lagrange del problema variacional es $f_{xx}+g_{yy}=0$, la cual es conocida como ecuación de Laplace y cuyas soluciones son funciones armónicas.

Generalización del problema variacional.

Sea $F(\bar{x},f,g_1,\ldots,g_n)$ una función con primera y segunda derivadas parciales continuas con respecto a todos sus argumentos sobre el conjunto $\Omega\times O\times O'$, en donde $\bar{x}=(x_1,\ldots,x_n)\in\Omega\subset\mathbb{R}^n$, Ω un conjunto cerrado y conexo y O, O' son dos conjuntos abiertos de dos espacios de funciones V, V', tales que $f\in O\subset V$ y $(g_1,\ldots,g_n)\in O'\subset(C(\Omega))^n$, esto es $F\in C^2(\Omega\times O\times O')$. Si $f:\Omega\to\mathbb{R}$, $f\in O=B(f,\epsilon)\subset C^1(\Omega)$ y satisface la condición de frontera $f(\bar{x})=w(\bar{x})$, $\bar{x}\in\partial\Omega$, siendo w una función conocida. Si $J:O=B(f,\epsilon)\subset C^1(\Omega)\to\mathbb{R}$ es una funcional de la forma $J(f)=\int\cdots\int_\Omega F(\bar{x},f,(f_{x_1},f_{x_2},\ldots,f_{x_n}))dx_1\ldots dx_n=\int\cdots\int_\Omega F(\bar{x},f,\mathrm{grad}(f))dx_1\ldots dx_n$, entonces deseamos encontrar una función f, para la cual la funcional J alcance un extremo (débil).

<u>Lema.</u> 3.3.21 Si $g_o\in C(\Omega)$ es una función fija, donde Ω es un conjunto cerrado y conexo de $\mathbb{R}^n$, y si $\int\cdots\int_\Omega g_o(\bar{x})h(\bar{x})dx_1\ldots dx_n=0$ $(\bar{x}\in\mathbb{R}^n)$ para toda $h\in C^1(\Omega)\cap C_o(\Omega)$, entonces $g_o(\bar{x})=0$ para todo $\bar{x}\in\Omega$.

La demostración no es esencialmente diferente a la del Lema 3.3.16.

<u>Teorema</u>. 3.3.22 Sea $(V,\|\cdot\|)$ el espacio de Banach constituido por el espacio lineal $V=C^1(\Omega)$, $\Omega\subset\mathbb{R}^n$ un conjunto cerrado y conexo, con la norma $\|f\|=\sum_{k=0}^{n}\max_{\bar{x}\in\Omega}|f_{x_k}(\bar{x})|$ y $J:O=B(f,\epsilon)\subset C^1(\Omega)\to\mathbb{R}$ una funcional de la forma $J(f)=\int\cdots\int_\Omega F(x,f,\mathrm{grad}(f))dx_1\ldots dx_n$, en donde $F\in C^2(\Omega\times O\times O')$, $f\in O=B(f,\epsilon)\subset C^1(\Omega)$ y $\mathrm{grad}(f)\in O'\in(C(\Omega))^n$, si f satisface la condición de frontera $f(\bar{x})=w(\bar{x})$, $\bar{x}\in\partial\Omega$, siendo w una función conocida. Entonces una condición necesaria para que la funcional J alcance un extremo para una función dada f, es que f satisfaga la ecuación de Euler $\frac{\partial F}{\partial f}-\sum_{k=1}^{n}\frac{\partial}{\partial x_k}\left(\frac{\partial F}{\partial f_{x_k}}\right)=0$.

Demostración: Sea $h\in C^1(\Omega)\cap C_o(\Omega)\cap B(0,\epsilon)$, calculando la diferencia ΔJ y haciendo uso del teorema de Taylor, encontramos que

$$J(f+h)-J(f)=\int\cdots\int_\Omega(F(x,f+h,\mathrm{grad}(f)+\mathrm{grad}(h))-F(x,f,\mathrm{grad}(f)))dx_1dx_2\ldots dx_n=\int\cdots\int_\Omega(F_fh+F_{f_{x_1}}h_{x_1}+\ldots+F_{f_{x_n}}h_{x_n})dx_1dx_2\ldots dx_n,$$ entonces calculando

$$\int\cdots\int_\Omega(F_{f_{x_1}}h_{x_1}+\ldots+F_{f_{x_n}}h_{x_n})dx_1dx_2\ldots dx_n=\int\cdots\int_\Omega\left(\frac{\partial}{\partial x_1}(F_{f_{x_1}}h)+\ldots+\frac{\partial}{\partial x_n}(F_{f_{x_n}}h)\right)dx_1dx_2\ldots dx_n-\int\cdots\int_\Omega\left(\frac{\partial}{\partial x_1}F_{f_{x_1}}+\ldots+\frac{\partial}{\partial x_n}F_{f_{x_n}}\right)hdx_1\ldots dx_n$$

Ahora bien, consideremos la siguiente generalización del teorema de Gauss

$$\int\cdots\int_{\Omega}\left(\frac{\partial}{\partial x_1}(F_{f_{x_1}}h)+\cdots+\frac{\partial}{\partial x_n}(F_{f_{x_n}}h)\right)dx_1dx_2\cdots dx_n=$$

$$\int\cdots\int_{\partial\Omega}\left(F_{f_{x_1}}h\,\frac{dx_1}{dn}+\cdots+F_{f_{x_n}}h\,\frac{dx_n}{dn}\right)dS$$, en donde el vector $(\frac{dx_1}{dn},\ldots,\frac{dx_n}{dn})$ es unitario y normal exterior a Ω un subconjunto cerrado y convexo de $\mathbb{R}^n$, $\partial\Omega$ es la frontera de Ω y dS el elemento diferencial de superficie.

Una extensa discusión de esta generalización al teorema de Gauss, pertenece, de hecho al Calculo Exterior. Sin embargo, esta discusión no la damos aquí, y sólo tomaremos este resultado.
Y puesto que $h(\bar{x})=0$ para toda $\bar{x}\in\partial\Omega$, tenemos que

$$\int\cdots\int_{\Omega}(F_{f_{x_1}}h_{x_1}+\cdots+F_{f_{x_n}}h_{x_n})dx_1dx_2\cdots dx_n=\int\cdots\int_{\Omega}\left(\frac{\partial}{\partial x_1}F_{f_{x_1}}+\cdots+\frac{\partial}{\partial x_n}F_{f_{x_n}}\right)h\,dx_1dx_2\cdots dx_n,$$

entonces $dJ(f)(h)=\int\cdots\int_{\Omega}\left(F_f-\frac{\partial}{\partial x_1}F_{f_{x_1}}-\cdots-\frac{\partial}{\partial x_n}F_{f_{x_n}}\right)h\,dx_1dx_2\cdots dx_n$, una condición necesaria para que la funcional J tenga extremo en el punto f es que $dJ(f)(h)=0$ para toda h admisible en $C^1(\Omega)\cap C_0(\Omega)\cap B(0,\epsilon)$. Luego por el lema anterior (3.3.18), tenemos que $\frac{\partial F}{\partial f}-\sum_{k=1}^{n}\frac{\partial}{\partial x_k}\left(\frac{\partial F}{\partial f_{x_k}}\right)=0$. lo que concluye la prueba del teorema.

Existen, por supuesto, muchas otras clases de problemas variacionales. Planteamos ahora el problema anterior con puntos finales variables para el siguiente caso particular: De entre todas las funciones (curvas) definidas sobre una bola abierta $O=B(f,\epsilon)$ cuyos puntos finales están sobre las dos rectas verticales $x=a$ y $x=b$, deseamos encontrar una función f, para la cual la funcional $F:O\subset C^1[a,b]\to\mathbb{R}$ de la forma $J(f)=\int_a^b F(x,f,f')dx$ alcance un extremo, en donde $F\in C^2([a,b]\times O\times O')$, con $x\in[a,b]$ y O, O' dos conjuntos abiertos tales que $f\in O\subset C^1[a,b]$ y $f'\in O'\subset C[a,b]$.

Comenzamos calculando la diferencial $dJ(f)$ de la funcional J, primero notamos que

$\Delta J=J(f+h)-J(f)=\int_a^b(F(x,f+h,f'+h')-F(x,f,f'))dx$, en donde $f+h\in O$, es decir, $\|h\|<\epsilon$.
Usando el teorema de Taylor para extender el integrando, tenemos que

$\int_a^b (F_f h + F_{f'} h')dx + o(\|h\|)$, en donde, como antes $\|h\| = \max_{x\in[a,b]} |h(x)| + \max_{x\in[a,b]} |h'(x)|$ y $o(\|h\|)$ es un infinitésimo de orden superior al primero respecto de $\|h\|$. Por tanto

$dJ(f)(h) = \int_a^b (F_f h + F_{f'} h')dx.$

Observamos que por la naturaleza del problema, h no necesariamente se anula en los puntos finales a y b, e integrando por partes el segundo sumando del integrando de $dJ(f)(h)$ tendremos que

$\int_a^b F_{f'} h'(x)dx = F_{f'} h(x)\Big|_{x=a}^{x=b} - \int_a^b \frac{d}{dx} F_{f'} h(x)dx$, entonces

$dJ(f)(h) = \int_a^b (F_f - \frac{d}{dx} F_{f'})h(x)dx + F_{f'}(x,f,f')\big|_{x=b}\, h(b) - F_{f'}(x,f,f')\big|_{x=a}\, h(a).$

Consideremos primero funciones $h \in C^1[a,b] \cap C_0[a,b] \cap B(0,\epsilon)$, es decir $h(a) = h(b) = 0$, entonces la condición $dJ(f)(h) = 0$ para toda $h \in C^1[a,b] \cap C_0[a,b] \cap B(0,\epsilon)$, implica $F_f - \frac{d}{dx} F_{f'} = 0$. Puesto que si una función $f = f(x)$ es extremal comparada con las curvas admisibles con puntos variables, entonces $f = f(x)$ es extremal comparada con las curvas admisibles con puntos fijos $f(a) = A$ y $f(b) = B$. Pero si $f = f(x)$ es extremal, $\int_a^b (F_f - \frac{d}{dx} F_{f'})h(x)dx = 0$, la condición $dJ(f)(h) = 0$ para toda $h \in C^1[a,b] \cap B(0,\epsilon)$ toma la forma $F_{f'}(x,f,f')\big|_{x=b}\, h(b) - F_{f'}(x,f,f')\big|_{x=a}\, h(a) = 0$, y puesto que h es arbitraria, se sigue que $F_{f'}(x,f,f')\big|_{x=a} = 0$ y $F_{f'}(x,f,f')\big|_{x=b} = 0$.

Así, para resolver el problema con puntos finales variables , debemos encontrar primero una solución general de la ecuación de Euler $F_f - \frac{d}{dx} F_{f'} = 0$, y entonces usar las condiciones $F_{f'}\big|_{x=a} = 0$ y $F_{f'}\big|_{x=b} = 0$ para determinar los valores de las constantes arbitrarias.

Ejemplo. 3.3.23 (Una variante del problema de la braquistocrona). Partiendo del punto (0,0) una partícula se desliza sobre una curva que está en un plano vertical. Se busca la curva tal que la partícula alcanza la línea vertical $x = b$ $(b \neq 0)$ en un tiempo mínimo. Como ya se había visto en el problema 3.3.16. La solución general de la correspondiente ecuación de Euler consiste de la familia de cicloides $x = \frac{1}{2}C'(\theta - \operatorname{sen}\theta)$, $f = \frac{1}{2}C'(1 - \cos\theta)$ $(f = y)$. Para determinar C', usamos la condición

$F_{f'} = \dfrac{f'}{(2gf(1+f'^2))^{\frac{1}{2}}} = 0$ para $x = b$, entonces $f' = 0$ para $x = b$, es decir

$f'(b) = 0$, se sigue que $\frac{1}{2}C' = b/\pi$ y por lo tanto $x = (b/\pi)(\theta - \operatorname{sen}\theta)$, $f = (b/\pi)(1 - \cos\theta)$.

Definición. 3.3.24 Sea $(V, \|\cdot\|)$ el espacio de Banach constituido por el espacio lineal $V=C^1[a,b]$ con la norma $\|f\|=\max|f(x)|+\max|f'(x)|$, y $J:O=B(f,\epsilon)\subset C^1[a,b] \to \mathbb{R}$ una funcional de la forma $J(f)=\int_a^b F(x,f,f')dx$, donde $F\in C^2([a,b]\times O\times O')$, $O=B(f,\epsilon)\subset C^1[a,b]$ y $f'\in O'\subset C[a,b]$. Supongamos que $\|h\|<\epsilon$ y que $h(x)\neq 0$ en una vecindad N_{x_0} de x_0, si $\Delta\sigma_h=\int_a^b h(x)dx$ y si el límite $\lim\limits_{\Delta\sigma_h\to 0}\frac{J(f+h)-J(f)}{\Delta\sigma_h}$ existe, éste es llamado la derivada de la funcional J en el punto x_0 (para la función f), y es denotado por $\left.\frac{dJ}{df}\right|_{x=x_0}$.

Observamos que si $h(x)\neq 0$ en una vecindad N_{x_0}, entonces $J(f+h)-J(f)=\left.\frac{dJ}{df}\right|_{x=x_0}\Delta\sigma_h+o(\|h\|)$, así $dJ(f)(h)=\left.\frac{dJ}{df}\right|_{x=x_0}\Delta\sigma_h$. Además a partir de la definición es claro que $\left.\frac{d}{df}\right|_{x=x_0}$ es un operador lineal sobre el espacio de funcionales definidas sobre el abierto O.

Ahora bien, sabemos que $J(f+h)-J(f)=\int_a^b(F_f-\frac{d}{dx}F_{f'})h(x)dx+o(\|h\|)$, entonces por el teorema del valor medio para integrales, tenemos que existe un $x_0\in[a,b]$ tal que $(F_f-\frac{d}{dx}F_{f'})\Big|_{x=x_0}\Delta\sigma_h=\int_a^b(F_f-\frac{d}{dx}F_{f'})h(x)dx$, entonces $\left.\frac{dJ}{df}\right|_{x=x_0}=(F_f-\frac{d}{dx}F_{f'})\Big|_{x=x_0}$. Esta definición y sus consecuencias serán de gran utilidad cuando tratemos con problemas variacionales con condiciones subsidiarias.

Recordamos que se hizo una generalización para un determinado problema variacional, en donde consideramos integrandos del tipo $F(\bar{x},f,grad(f))$ en donde $\bar{x}\in\Omega\subset\mathbb{R}^n$ y $F\in C^2(\Omega\times O\times O')$, $f\in O=B(f,\epsilon)$ y $grad(f)\subset O'\subset(C(\Omega))^n$, bajo la norma $\|f\|=\sum_{k=0}^n \max\limits_{\bar{x}\in\Omega}|f_{x_k}(\bar{x})|$, $f\in C^1(\Omega)$.

Es conveniente ahora, generalizar el problema variacional, en donde la funcional J esté definida sobre algún espacio de dimensión n. Sea $F(x,f_1,\ldots,f_n,g_1,\ldots,g_n)$ una función con primera y segunda derivadas parciales continuas con respecto a todos sus argumentos sobre el conjunto $[a,b]\times O\times O'$, en donde O y O' son dos conjuntos abiertos de dos espacios de funciones V y V', tales que $(f_1,\ldots,f_n)\in O\subset V$ y $(g_1,\ldots g_n)\in O'\subset V'$, esto es $F\in C^2([a,b]\times O\times O')$, si $\|(f_1,\ldots,f_n)\|=\sum_{k=1}^n\|f_k\|_{V_k}$ en donde $f_k\in V_k$. V_k un espacio de Banach, $f_k:[a,b]\to\mathbb{R}$ $(k=1,\ldots,n)$ y se satisfacen las condiciones de frontera $f_k(a)=A_k$, $f_k(b)=B_k$ $(k=1,\ldots,n)$, y si $J:O\to\mathbb{R}$ es una funcional de la forma $J(f_1,\ldots,f_n)=\int_a^b F(x,f_1,\ldots,f_n,f_1',\ldots,f_n')dx$, entonces deseamos encontrar un vector $(f_1,f_2,\ldots,f_n)$, para el cual la funcional J alcance un extremo (débil), entendiendo por esto que $(f_1,f_2,\ldots,f_n)$ es extremal con respecto a la norma $\|(f_1,\ldots,f_n)\|=\sum_{k=1}^n(\max\limits_{x\in[a,b]}|f_k(x)|+\max\limits_{x\in[a,b]}|f_k'(x)|)$ al considerar la definición 3.3.1.

<u>Teorema</u>. 3.3.25 Sea $(V,\|\cdot\|)$ el espacio de Banach constituido por el espacio lineal $V=(C^1[a,b])^n$, con la norma $\|(f_1,\ldots,f_n)\|=\sum_{k=1}^n \|f_k\|=\sum_{k=1}^n (\max_{x\in[a,b]}|f_k(x)|+\max_{x\in[a,b]}|f_k'(x)|)$ y $J:O=B((f_1,\ldots,f_n),\epsilon)\to\mathbb{R}$ una funcional de la forma $J(f_1,\ldots,f_n)=\int_a^b F(x,f_1,\ldots,f_n,f_1',\ldots,f_n')dx$, en donde $F\in C^2([a,b]\times O\times O')$, $(f_1,\ldots,f_n)\in O\subset(C^1[a,b])^n$, $(f_1',\ldots,f_n')\in O'\subset(C[a,b])^n$, si el vector $(f_1,\ldots,f_n)$ satisface la condición de frontera $f_k(a)=A_k$, $f_k(b)=B_k$ $(k=1,..,n)$. Entonces una condición necesaria para que la funcional J alcance un extremo en un vector dado $(f_1,\ldots,f_n)$, es que este vector satisfaga las ecuaciones de Euler

$$F_{f_k}-\frac{d}{dx}F_{f_k'}=0 \quad (k=1,\ldots,n).$$

Demostración: Demos al vector $(f_1,\ldots,f_n)$ un incremento $(h_1,\ldots,h_n)$, $\|(h_1,\ldots,h_n)\|<\epsilon$, de tal manera que se sigan satisfaciendo las condiciones de frontera, entonces $(h_1,\ldots,h_n)\in(C_0[a,b])^n\cap(C_1[a,b])^n\cap B((0,\ldots,0),\epsilon)$, entonces calculando la diferencia $J(f_1+h_1,\ldots,f_n+h_n)-J(f_1,\ldots,f_n)$ y usando el teorema de Taylor tenemos que

$$J(f_1+h_1,\ldots,f_n+h_n)-J(f_1,\ldots,f_n)=\int_a^b\sum_{k=1}^n (F_fh_k-F_{f_k'}h_k')dx+o(\|(h_1,\ldots h_n)\|).$$

Por lo tanto $dF(f_1,\ldots,f_n)(h_1,\ldots,h_n)=\int_a^b\sum_{k=1}^n (F_fh_k-F_{f_k'}h_k')dx.$

La condición necesaria

$dJ(f_1,\ldots f_n)(h_1,\ldots,h_n)=0$, $(h_1,\ldots,h_n)\in(C_0[a,b])^n\cap(C^1[a,b])^n\cap B((0,\ldots,0),\epsilon)$ implica que $\int_a^b (F_fh_k-F_{f_k'}h_k')dx=0$ $(k=1,\ldots,n)$, puesto que todos los incrementos $h_k(x)$ son independientes, entonces de acuerdo con el teorema 3.3.9 tenemos el siguiente sistema de ecuaciones de Euler $F_{f_k}-\frac{d}{dx}F_{f_k'}=0$ $(k=1,\ldots,n)$.

Observamos que $F_{f_k}-\frac{d}{dx}F_{f_k'}=0$ $(k=1,\ldots,n)$ es un sistema de n ecuaciones diferenciales de segundo orden, y su solución general contiene 2n constantes arbitrarias, las cuales son determinadas de las condiciones de frontera $f_k(a)=A_k$ y $f_k(b)=B_k$.

<u>Ejemplo</u>. 3.3.26 (Propagación de la luz en un medio no homogeneo) Supongamos que el espacio de tres dimensiones está ocupado por un medio ópticamente no homogeneo, tal que la velocidad de propagación de la luz $v(x,f_1,f_2)$ en cada punto es función de las coordenadas del punto (x,f_1,f_2).

De acuerdo al principio de Fermat, la luz viaja de un punto a otro a través de la trayectoria para la cual el tiempo de tránsito es mínimo. Si la curva que une los puntos A y B es especificada por las ecuaciones $f_1=f_1(x)$, $f_2=f_2(x)$, el tiempo que tarda la luz para transitar sobre la curva es igual a $t=\int_a^b(1+f_1'^2+f_2'^2)^{\frac{1}{2}}/v(x,f_1,f_2)dx$, y el sistema de ecuaciones de Euler para esta funcional es

$$\frac{\partial v}{\partial f_1}(1+f_1'^2+f_2'^2)^{\frac{1}{2}}v^{-2}+\frac{d}{dx}\left(\frac{f_1'}{v}(1+f_1'^2+f_2'^2)^{-\frac{1}{2}}\right)=0,$$

$$\frac{\partial v}{\partial f_2}(1+f_1'^2+f_2'^2)^{\frac{1}{2}}v^{-2}+\frac{d}{dx}\left(\frac{f_2'}{v}(1+f_1'^2+f_2'^2)^{-\frac{1}{2}}\right)=0.$$

Y estas son las ecuaciones diferenciales para la curva a través de la cual la luz se propaga.

<u>Ejemplo</u>. 3.3.27 (Geodésicas).
Las geodésicas representan uno de los campos de interés en el estudio de la geometría diferencial.
Supongamos que tenemos una superficie S especificada por la parametrización $H=H(f_1,f_2)$. La curva de menor longitud sobre S que conecta dos puntos de S, es llamada la geodésica que conecta los dos puntos. Claramente las ecuaciones para las geodésicas de S son las ecuaciones de Euler para los correspondientes problemas variacionales.
Sabemos que la longitud de arco de una curva parametrizada por $H(t)=(f_1(t),f_2(t))$ correspondiente a los valores t_o y t_1 es igual a $s(t)=\int_{t_o}^{t_1}\|H'(t)\|\,dt=\int_{t_o}^{t_1}(H'(t),H'(t))^{\frac{1}{2}}dt$

$$=\int_{t_o}^{t_1}(H_{f_1}f_1'+H_{f_2}f_2',H_{f_1}f_1'+H_{f_2}f_2')^{\frac{1}{2}}dt=\int_{t_o}^{t_1}((H_{f_1},H_{f_1})f_1'^2+2(H_{f_1},H_{f_2})f_1'f_2'+(H_{f_2},H_{f_2})f_1'^2)^{\frac{1}{2}}dt,$$

si recordamos que los coeficientes de la primera forma fundamental de la superficie están dados por $E=(H_{f_1},H_{f_1})$, $F=(H_{f_1},H_{f_2})$ y $G=(H_{f_2},H_{f_2})$,

entonces $s(t)=\int_{t_o}^{t_1}(Ef_1'^2+2Ff_1'f_2'+Gf_2'^2)^{\frac{1}{2}}dt$, tendremos así que $J(f_1,f_2)=\int_{t_o}^{t_1}(Ef_1'^2+2Ff_1'f_2'+Gf_2'^2)^{\frac{1}{2}}dt$, y las ecuaciones de Euler para esta funcional son determinadas como sigue:
Primero denotemos a $W(f_1,f_2)=Ef_1'^2+2Ff_1'f_2'+Gf_2'^2$, entonces

$$\frac{E_{f_1}f_1'^2+2F_{f_1}f_1'f_2'+G_{f_1}f_2'^2}{W^{\frac{1}{2}}} - \frac{d}{dt}\left(\frac{2(Ef_1'+Ff_2')}{W^{\frac{1}{2}}}\right)=0$$

$$\frac{E_{f_2}f_1'^2+2F_{f_2}f_1'f_2'+G_{f_2}f_2'^2}{W^{\frac{1}{2}}} - \frac{d}{dt}\left(\frac{2(Ff_1'+Gf_2')}{W^{\frac{1}{2}}}\right)=0$$

Ejemplo. 3.3.28 Deseamos encontrar las geodésicas de un cilindro circular, $H(\theta,z)=(a\cos\theta, a\,\mathrm{sen}\,\theta, z)$. En este caso los coeficientes de la primera forma fundamental del cilindro son $E=a^2$, $F=0$, $G=1$. Las geodésicas del cilindro circular tienen las ecuaciones (de Euler)

$$\frac{d}{dt}\left(\frac{a^2\theta'}{(a^2\theta'^2+z'^2)^{\frac{1}{2}}}\right)=0,\quad \frac{d}{dt}\left(\frac{z'}{(a^2\theta'^2+z'^2)^{\frac{1}{2}}}\right)=0$$, es decir

$$\frac{a^2\theta'}{(a^2\theta'^2+z'^2)^{\frac{1}{2}}}=C_1,\quad \frac{z'}{(a^2\theta'^2+z'^2)^{\frac{1}{2}}}=C_2.$$ Dividiendo la segunda ecuación entre la primera obtenemos $\frac{dz}{d\theta}=c_1$, la solución de esta ecuación es $z=c_1\theta+c_2$, la cual representa una familia de hélices de dos parámetros sobre el cilindro $(a\cos\theta, a\,\mathrm{sen}\,\theta, c_1\theta+c_2)$.

El concepto de geodésica puede ser definido no sólo para superficies, sino también para variedades de dimensiones superiores.

En lo que sigue consideraremos una generalización sobre el orden de las derivadas involucradas en la forma de la funcional.

Sea $F(x,f,g_1,g_2,\ldots,g_n)$ una función con primera y segunda derivadas parciales continuas con respecto a todos sus argumentos sobre el conjunto $[a,b]\times O_n\times O_{n-1}\times\cdots\times O_0$, en donde O_k $(k=0,\ldots,n)$ son conjuntos abiertos de $n+1$ espacios de funciones, V_k $(k=0,\ldots,n)$, tales que $f\in O_n\subset V_n$ y $g_k\in O_{n-k}\subset V_{n-k}$ $(k=1..,n)$, esto es $F\in C^2([a,b]\times O_n\times O_{n-1}\times\cdots\times O_0)$. Si $f\in C^n[a,b]$ y satisface las condiciones de frontera $f^{(k)}(a)=A_k$, $f^{(k)}(b)=B_k$ $(k=0,\ldots,n-1)$, considerando la norma $\|f\|=\sum_{k=0}^{n}\max_{x\in[a,b]}|f^{(k)}(x)|$, y si $J:O_n=B(f,\varepsilon)\subset C^n[a,b]\to\mathbb{R}$ es una funcional de la forma $J(f)=\int_a^b F(x,f,f',f'',\ldots,f^{(n)})dx$, entonces deseamos encontrar una función f, para la cual la funcional J alcance un extremo.

<u>Teorema</u>. 3.3.29 Sea $(V,\|\cdot\|)$ el espacio de Banach constituido por el espacio lineal $V=C^n[a,b]$ con la norma $\|f\|=\sum_{k=0}^{n}\max|f^{(k)}(x)|$, $J:O_n=B(f,\epsilon)\subset C^n[a,b]\to\mathbb{R}$ una funcional de la forma $J(f)=\int_a^b F(x,f,f'',\ldots,f^{(n)})dx$, en donde $F\in C^2([a,b]\times O_n\times O_{n-1}\times\ldots\times O_1)$, $f\in O_n\subset C^n[a,b]$, $f^{(k)}\in O_{n-k}\subset C^{n-k}[a,b]$, $(k=1,\ldots,n)$ y f satisface la condición de frontera $f^{(k)}(a)=A_k$, $f^{(k)}(b)=B_k$ $(k=0,\ldots,n-1)$. Entonces una condición nece saria para que la funcional J alcance un extremo para una función dada f, es que f satisfaga la ecuación de Euler
$F_f-\frac{d}{dx}F_{f'}+\frac{d^2}{dx^2}F_{f''}-\ldots+(-1)^n\frac{d^n}{dx^n}F_{f^{(n)}}=0$.

Demostración: Demos a la función f un incremento h, $\|h\|<\epsilon$, de tal manera que para $f+h$ se sigan satisfaciendo las condiciones de fron tera, entonces $h^{(k)}(a)=0$, $h^{(k)}(b)=0$ $(k=0,\ldots,n-1)$, esto es $h\in C^{n-1}_{0,n-1}[a,b]\cap C^n[a,b]\cap B(0,\epsilon)$. Entonces calculando la diferencia $J(f+h)-J(f)$ y haciendo uso del teorema de Taylor, obtenemos que
$J(f+h)-J(f)=\int_a^b(F(x,f+h,f'+h',f''+h'',\ldots,f^{(n)}+h^{(n)})-F(x,f,f',f'',\ldots,f^{(n)}))dx=\int_a^b(F_fh+F_{f'}h'+F_{f''}h''+\ldots+F_{f^{(n)}}h^{(n)})dx+o(\|h\|)$. Por lo tanto $dJ(f)(h)=\int_a^b(F_fh+F_{f'}h'+F_{f''}h''+\ldots+F_{f^{(n)}}h^{(n)})dx$ +

Luego una condición necesaria para que la funcional J tenga un extremo en una función dada f, es que $dJ(f)(h)=0$ para toda $h\in C^{n-1}_{0,n-1}[a,b]\cap C^n[a,b]\cap B(0,\epsilon)$ (h admisible), esto es
$\int_a^b(F_fh+F_{f'}h'+F_{f''}h''+\ldots+F_{f^{(n)}}h^{(n)})dx=0$, e integrando por partes repetidamente y usando las condiciones de frontera, encontramos que $\int_a^b(F_f-\frac{d}{dx}F_{f'}+\frac{d^2}{dx^2}F_{f''}-\ldots+(-1)^n\frac{d^n}{dx^n}F_{f^{(n)}})h(x)dx=0$, luego por el Lema I se tiene que $F_f-\frac{d}{dx}F_{f'}+\frac{d^2}{dx^2}F_{f''}-\ldots+(-1)^n\frac{d^n}{dx^n}F_{f^{(n)}}=0$.
La existencia de las derivadas $\frac{d}{dx}F_{f'}$, $\frac{d^2}{dx^2}F_{f''}\ldots,\frac{d^n}{dx^n}F_{f^{(n)}}$ se sigue por un argumento similar al del Lema IV.

Hasta aquí, sólo hemos tratado con problemas variacionales, en don de la clase de funciones admisibles quedaba especificada por condi ciones impuestas sobre los puntos finales de las curvas. Sin embar go, muchas aplicaciones de los Métodos Variacionales nos llevan a problemas en los cuales no sólo hay condiciones de frontera, sino condiciones de tipos completamente diferentes, conocidas como con diciones subsidiarias o restricciones que son impuestas sobre las curvas (funciones) admisibles.

En la definición 3.2.1 habíamos establecido que una funcional J definida en un conjunto abierto $O \subset V$, V un espacio lineal normado, es diferenciable en el sentido de Fréchet en el punto f, si existe una funcional lineal acotada $dJ(f) \in V' = \mathcal{L}(V, \mathbb{R})$, tal que $J(f+h)-J(f) = dJ(f)(h)+o(\|h\|)$, en donde $f+h \in O$. A la funcional acotada $dJ(f)$ evaluada en $h \in V$, $dJ(f)(h)$, se le llamó la diferencial de Fréchet de la funcional J en el punto $f \in V$, frecuentemente la funcional acotada $dJ(f)$ es llamada la derivada de Fréchet en la funcional J en f. Ahora generalizamos esta definición a transformaciones definidas en un conjunto abierto O de un espacio de Banach $(V_1, \|\cdot\|_1)$ en un espacio de Banach $(V_2, \|\cdot\|_2)$.

Definición. 3.3.30 Sea T una transformación de V_1 en V_2, V_1 y V_2 espacios de Banach. Si para f fijo en O y para toda $h \in V_1$ existe una transformación $dT(f)(h) \in V_2$ la cual es lineal y acotada para todo $h \in V_1$, tal que

$$\lim_{\|h\| \to 0} \frac{\|T(f+h)-T(f)-dT(f)(h)\|}{\|h\|} = 0 .$$

Entonces la transformación T es llamada diferenciable en el sentido de Fréchet en $f \in V_1$, $dT(f)(h)$ es llamada la diferencial de Fréchet de la transformación T en f y $dT(f)$ es llamada la derivada de Fréchet de la transformación T en f.

El objetivo de extender esta definición, es el de probar el teorema generalizado de la función inversa, que nos será útil para establecer la existencia de multiplicadores de Lagrange.

Definición 3.3.31 Sea T una transformación con derivada continua de Fréchet en un conjunto abierto O de un espacio de Banach V_1 en un espacio de Banach V_2, si $f \in O$ es tal que $dT(f)$ es un mapeo suprayectivo de V_1 en V_2 el punto f es llamado un punto regular de la transformación T.

Ejemplo. 3.3.32 Sea la transformación $T: O \subset \mathbb{R}^n \to \mathbb{R}^m$, $\bar{x} \in \mathbb{R}^n$ es un punto regular si la matriz Jacobiana de T tiene rango m.

Lema 3.3.33 (Teorema Generalizado del valor medio). Sea T una transformación diferenciable de Fréchet definida en un conjunto abierto O, de un espacio de Banach V_1 en un espacio de Banach V_2, y supongamos que $f+ch\in O$ para todo c, $0\leq c\leq 1$. Entonces

$$\|T(f+h)-T(f)\|\leq\|h\| \sup_{1<c<0}\|dT(f+h)\|.$$

Demostración: Tomemos $G\in V_2^*=\mathcal{L}(V_2,\mathbb{R})$ tal que $\|G\|\|T(f+h)-T(f)\|=|G(T(f+h)-T(f))|$, esto significa que G es alineado con $T(f+h)-T(f)$, tal funcional lineal continua existe, consideremos el corolario al teorema de Hahn-Banach 2.6.36, si $T(f+h)-T(f)\neq 0$ entonces existe G en V_2^* tal que $G(T(f+h)-T(f))=\|G\|\|T(f+h)-T(f)\|$.

Definamos la función $v:[0,1]\to\mathbb{R}$ como $v(c)=G(T(f+ch))$, entonces por definición

$$v'(c)=\lim_{\Delta c\to 0}\frac{G(T(f+ch+\Delta ch))-G(T(f+ch))}{\Delta c}$$

$$=\lim_{\Delta c\to 0}\frac{G(T(f+ch+\Delta ch)-T(f+ch))}{\Delta c}$$

$$=G\left(\lim_{\Delta c\to 0}\frac{T(f+ch+\Delta ch)-T(f+ch)}{\Delta c}\right)=G\left(\frac{d}{dt}T(f+ch+\Delta ch)\Big|_{\Delta c=0}\right)=G(dT(f+ch)(h)).$$

Ya que T es diferenciable en el sentido de Fréchet, entonces por la proposición 3.2.4 T es también diferenciable en el sentido de Gateaux, notamos también que si $G\in V_2^*$, entonces G es una funcional lineal y continua.

Ahora bien, por el teorema del valor medio para funciones reales de variable real, tenemos que

$v(1)-v(0)=v'(c_0)$, $1<c_0<0$, y por tanto

$G(T(f+h))-G(T(f))=G(dT(f+c_0h)(h))$, $1<c_0<0$, así

$|G(T(f+h)-T(f))|=|G(dT(f+c_0h)(h))|\leq\|G\|\|dT(f+c_0h)\|\|h\|$, $1<c_0<0$, y como escogimos G alineada con $T(f+h)-T(f)$ tenemos que

$\|G\|\|T(f+h)-T(f)\|=\|G\|\|h\|\|dT(f+c_0h)\|$, $1<c_0<0$, por consiguiente

$\|T(f+h)-T(f)\|\leq\|h\|\|dT(f+c_0h)\|$, $1<c_0<0$, de donde

$$\|T(f+h)-T(f)\|\leq\|h\| \sup_{0<c<1}\|dT(f+ch)\|.$$

Proposición. 3.3.34 Sea T una transformación dos veces diferenciable en el sentido de Fréchet definida en un conjunto abierto O de un espacio de Banach V_1 en un espacio de Banach V_2, y supongamos que $f+ch\in O$ para toda c, $0\leq c\leq 1$. Entonces

$$\|T(f+h)-T(f)-dT(f)(h)\|\leq\tfrac{1}{2}\|h\|^2 \sup_{0<c<1}\|d^2T(f+ch)\|.$$

La demostración es análoga a la del Lema anterior, tomando G alineada con $T(f+h)-T(f)-dT(f)(h)$ y empleando el Teorema de Taylor para extender $v(1)-v(0)=v'(0)+\frac{1}{2}v''(c)$, $0<c<1$.

<u>Teorema</u>. 3.3.35 (Teorema Generalizado de la Función Inversa).
Sea f_o un punto regular de la transformación T, la cual mapea un espacio de Banach V_1 en un espacio de Banach V_2. Entonces existe una vecindad N_{g_o} del punto $g_o=T(f_o)$, es decir una bola centrada en g_o, y una constante $K>0$ tal que la ecuación $T(f)=g$ tiene solución para todo $g\in N_{g_o}$, y la solución satisface $\|f-f_o\|\leq K\|g-g_o\|$.
Demostración: Sea $L_o=\{h\in V_1;\ dT(f_o)(h)=0\}$, puesto que L_o es un subespacio cerrado, ya que L_o contiene todos sus puntos de acumulación, en efecto, sea h_o un punto de acumulación de L_o, entonces cualquier vecindad de h_o contiene una infinidad de puntos de L_o. Sea $B(h_o,\epsilon)=\{h;\ \|h_o-h\|<\epsilon\}$, dado que $B(h_o,\epsilon)$ es un conjunto convexo tenemos que h se puede expresar como combinación lineal convexa de elementos de $B(h_o,\epsilon)\cap L_o$, entonces $h_o=th_1+(1-t)h_2$, $h_1,h_2\in B(h_o,\epsilon)\cap L_o$, así $dT(f_o)(h_o)=tdT(f_o)(h_1)+(1-t)dT(f_o)(h_2)=0$ de donde $h_o\in L_o$. Luego al considerar el Teorema 2.6.37, tenemos que el espacio cociente V_1/L_o es un espacio de Banach. Definimos un operador $U:V_1/L_o\to V_2$; $U([f])=dT(f_o)(f)$, en donde $[f]$ denota la clase de elementos equivalentes a f módulo L_o. El operador U está bien definido, puesto que si $f_1,f_2\in[f]$, entonces $dT(f_o)(h_1)=U([f])=dT(f_o)(h_2)$, es decir, elementos equivalentes entre sí son mapeados en elementos idénticos en V_2, obviamente U es un operador lineal por la linealidad de $dT(f_o)$, también U es un operador continuo, puesto que $dT(f_o)$ es continua en V_1. El operador U es inyectivo, en efecto, supongamos $U([f_1])=U([f_2])$, entonces $dT(f_o)(f_1)=dT(f_o)(f_2)$, así $dT(f_o)(f_1-f_2)=0$, luego $f_1-f_2\in L_o=[0]$ ó $[f_1-f_2]=[0]$ ó $[f_1]-[f_2]=[0]$, entonces $[f_1]=[f_2]$, y por último U es un operador suprayectivo, puesto que $dT(f_o)$ es suprayectiva, ya que f_o es un punto regular de la transformación T.
Por el Teorema de Banach sobre el operador inverso 2.6.41. El operador U tiene inversa continua. Recordamos también que por el Teorema 2.6.22, en espacios lineales normados continuidad de un operador equivale a su acotación, por lo tanto existe $M\geq 0$ tal que $\|U^{-1}(g)\|\leq M\|g\|$ para todo $g\in V_2$.

Dada $g\in V_2$ suficientemente próxima a g_o (próximo en el sentido de la norma en V_2), construimos una sucesión de elementos de V_1/L_o, $\{[f_n]\}_{n=1}^{\infty}$ y la correspondiente sucesión $\{f_n\}_{n=1}^{\infty}$ con $f_n\in[f_n]$, tal que f_o+f_n converge a la solución $T(f)=g$. Para $g\in V_2$ fija, sea $f_o\in[0]=L_o$ y definamos las sucesiones $\{[f_n]\}_{n=1}^{\infty}$ y $\{f_n\}_{n=1}^{\infty}$ recursivamente por $[f_n]-[f_{n-1}]=U^{-1}(g-T(f_o+f_{n-1})$, y de la clase $[f_n]$ escojamos f_n tal que $\|f_n-f_{n-1}\|\leq 2\|[f_n]-[f_{n-1}]\|$, lo cual es posible puesto que $\|[f_n]-[f_{n-1}]\|=\inf_{f\in[f_n]}\|f-f_{n-1}\|$, entonces

$2\|[f_n]-[f_{n-1}]\|>\inf_{f\in[f_n]}\|f-f_{n-1}\|$. Luego

$$
\begin{aligned}
[f_n] &= U^{-1}(g-T(f_o+f_{n-1}))+U^{-1}U([f_{n-1}])\\
&= U^{-1}(g-T(f_o+f_{n-1})+U([f_{n-1}]))\\
&= U^{-1}(g-T(f_o+f_{n-1})+dT(f_o)(f_{n-1})),
\end{aligned}
$$

similarmente tenemos que

$[f_{n-1}]=U^{-1}(g-T(f_o+f_{n-2})+dT(f_o)(f_{n-2}))$, de donde

$$
\begin{aligned}
[f_n]-[f_{n-1}] &= U^{-1}(g-T(f_o+f_{n-1})+dT(f_o)(f_{n-1}))-U^{-1}(g-T(f_o+f_{n-2})+dT(f_o)(f_{n-2})\\
&\quad +dT(f_o)(f_{n-2}))\\
&= -U^{-1}(T(f_o+f_{n-1})-T(f_o+f_{n-2})-dT(f_o)(f_{n-1}-f_{n-2})),
\end{aligned}
$$

así tomando $f_t=tf_{n-1}+(1-t)f_{n-2}$ con $0\leq t\leq 1$ y aplicando el Lema 3.3.33, tenemos que

$$
\begin{aligned}
\|[f_n]-[f_{n-1}]\| &\leq \|U^{-1}\|\,\|f_{n-1}-f_{n-2}\|\sup_{0<t<1}\|dT(f_o+f_t)-dT(f_o)\|\\
&\leq M\|f_{n-1}-f_{n-2}\|\sup_{0<t<1}\|dT(f_o+f_t)-dT(f_o)\|,
\end{aligned}
$$

y puesto que por definición $dT(f_o)$ es una transformación lineal continua en f_o. Dado $\epsilon>0$, existe $\delta>0$ tal que

$\|dT(f_o+f_t)-dT(f_o)\|<\epsilon$ para $\|f_t\|<\delta$, siempre que $\|f_{n-1}\|<\delta$ y $\|f_{n-2}\|<\delta$, puesto que $\|f_t\|\leq t\|f_{n-1}\|+(1-t)\|f_{n-2}\|<t\delta+(1-t)\delta=\delta$.

Así $\|[f_n]-[f_{n-1}]\|\leq M\|f_{n-1}-f_{n-2}\|$. Además, habíamos escogido $f_n\in[f_n]$, tal que $\|f_n-f_{n-1}\|\leq 2\|[f_n]-[f_{n-1}]\|\leq 2\epsilon M\|f_{n-1}-f_{n-2}\|$ y para $\epsilon>0$ suficientemente pequeña tenemos que

$\|f_n-f_{n-1}\|\leq\frac{1}{2}\|f_{n-1}-f_{n-2}\|$. Y dado que $\|f_1\|\leq 2\|[f_1]\|=2\|U^{-1}(g-T(f_o))\|=2\|U^{-1}(g-g_o)\|\leq 2M\|g-g_o\|$, se sigue que para $\|g-g_o\|$ suficientemente pequeño, se tiene que $\|f_1\|<\frac{1}{2}\delta$, así

$\|f_1\|<\frac{1}{2}\delta$, $\|f_2-f_1\|\leq\frac{1}{2}\|f_1\|<(\frac{1}{2})^2\delta$, $\|f_3-f_2\|=\frac{1}{2}\|f_2-f_1\|=(\frac{1}{2})^2\|f_1\|<(\frac{1}{2})^3\delta,\ldots.$
$\|f_n-f_{n-1}\|<(\frac{1}{2})^{n-1}\delta$, entonces

$\|f_n\|=\|f_1+(f_2-f_1)+\ldots+(f_n-f_{n-1})\|\leq\|f_1\|+\|f_2-f_1\|+\ldots+\|f_n-f_{n-1}\|$
$\leq(1+\frac{1}{2}+(\frac{1}{2})^2+\ldots+(\frac{1}{2})^{n-1})\|f_1\|\leq 2\|f_1\|<\delta$ para toda $n\in\mathbb{N}$, de donde
$\|f_n-f_{n-1}\|\leq\frac{1}{2}\|f_{n-1}-f_{n-2}\|\leq\frac{1}{2}(\|f_{n-1}\|+\|f_{n-2}\|)<\frac{1}{2}(2\delta)=\delta.$

Y ya que $\|f_n-f_{n-1}\|\leq\frac{1}{2}\|f_{n-1}-f_{n-2}\|$ es válida para toda n, si $n>m$,

$\|f_m-f_n\|\leq\|f_m-f_{m+1}\|+\|f_{m+1}-f_{m+2}\|+\ldots+\|f_{n-2}-f_{n-1}\|+\|f_{n-1}-f_n\|$
$\leq((\frac{1}{2})^m+(\frac{1}{2})^{m+1}+\ldots+(\frac{1}{2})^{n-2}+(\frac{1}{2})^{n-1})\|f_1\|$
$=((\frac{1}{2})^m(1+\frac{1}{2}+(\frac{1}{2})^2+\ldots+(\frac{1}{2})^{n-m-1})\|f_1\|\leq(\frac{1}{2})^m(1+\frac{1}{2}+\ldots+(\frac{1}{2})^n)\|f_1\|$
$=(\frac{1}{2})^m(\sum_{k=0}^{n}(\frac{1}{2})^k)\|f_1\|\to 0$ cuando $n,m\to\infty$. Por tanto la sucesión $\{f_n\}_{n=1}^{\infty}$ es de Cauchy, y como V_1 es de Banach, tenemos que esta sucesión converge a un elemento f en V_1, correspondiendo a este hecho la sucesión $\{[f_n]\}_{n=1}^{\infty}$ converge a la clase de f, puesto que

$\|[f_n]-[f_m]\|=\inf_{f\in[f_n]}\|f-f_m\|\leq\|f_n-f_m\|\to 0$ cuando $n,m\to\infty$; esto es $\{[f_n]\}_{n=1}^{\infty}$ es también una sucesión de Cauchy, y puesto que V_1/L_o es por el Teorema 2.6.37 un espacio de Banach, ya que L_o es un subespacio cerrado de V_1, la sucesión $\{[f_n]\}_{n=1}$ converge a una clase $[f]\in V_1/L_o$, y es claro que $f\in[f]$, puesto que los representantes de $[f_n]$ tienden a f, así $[f_n-f_{n-1}]=[f_n]-[f_{n-1}]=U^{-1}(g+T(f_o+f_{n-1}))$, entonces tomando el límite cuando $n\to\infty$, tenemos $[0]=U^{-1}(g+T(f_o+f))$, en donde recordamos que la transformación T tiene derivada continua de Fréchet en una vecindad O del punto f_o, f_o un punto regular de T, entonces T es continua en f_o+f_n para n suficientemente grande, o equivalentemente para $f_o+f_n\in O$. De donde

$T(f_o-f)=g$ y $\|f\|\leq 2\|f_1\|\leq 4\|U^{-1}\|\,\|g-g_o\|\leq 4M\|g-g_o\|$, tomando $K=4M$ se tiene que $\|(f+f_o)-f_o\|\leq K\|g-g_o\|$, lo que concluye la prueba del teorema.

Planteamiento del Problema Isoperimétrico.

Sea $F(x,f,g)$ un función con primera y segunda derivadas parciales continuas con respecto a todos sus argumentos sobre el conjunto $[a,b]\times O\times O'$, con $x\in[a,b]$ y O, O' dos conjuntos abiertos de dos espacios de funciones V, V', tales que $f\in O\subset V$ y $g\in O'\subset V'$, esto es $F\in C^2([a,b]\times O\times O')$. Si $f\in C^1[a,b]$, $\|f\|=\max_{x\in[a,b]}|f(x)|+\max_{x\in[a,b]}|f'(x)|$ $f\in O=B(f,\epsilon)$ y satisface las condiciones de frontera $f(a)=A$, $f(b)=B$, y es tal que para una funcional $K(f)=\int_a^b G(x,f,f')dx$ conocida toma un valor fijo Q, $K(f)=Q$, Si $J:O\subset C^1[a,b]\to\mathbb{R}$ es una funcional de la forma $J(f)=\int_a^b F(x,f,f')dx$, deseamos encontrar entonces una función para la cual la funcional J tenga un extremo.

Teorema. 3.3.36 Sea $(V,\|\cdot\|)$ el espacio de Banach formado por el espacio lineal $V=C^1[a,b]$ con la norma $\|f\|=\max_{x\in[a,b]}|f(x)|+\max_{x\in[a,b]}|f'(x)|$, $K,J:O=B(f,\epsilon)\subset C^1[a,b]\to\mathbb{R}$ dos funcionales de la forma $K(f)=\int_a^b G(x,f,f')dx$ y $J(f)=\int_a^b F(x,f,f')dx$, en donde $K,F\in C^2([a,b]\times O\times O')$, $x\in[a,b]$ $O=B(f,\epsilon)\subset C^1[a,b]$, $f'\in O'\subset C[a,b]$, si f satisface las condiciones (restricciones) $f(a)=A$, $f(b)=B$ y $K(f)=\int_a^b G(x,f,f')dx=Q$, y si $f=f(x)$ es extremal de J pero no de K, entonces existe una constante p, tal que f es extremal de la funcional $J+pK$, y por lo tanto f satisface la ecuación diferencial

$F_f-\frac{d}{dx}F_{f'}+p(G_f-\frac{d}{dx}G_{f'})=0.$

Demostración: Supongamos que J tiene un extremo para $f=f(x)$ sujeta a las condiciones $f(a)=A$, $f(b)=B$ y $K(f)=\int_a^b G(x,f,f')dx=Q$, sean $x_1,x_2\in[a,b]$ puntos arbitrarios. Entonces demos a $f=f(x)$ un incremento $h_1(x)+h_2(x)$, en donde $h_1(x)\neq 0$ en una vecindad N_{x_1} de x_1 y $h_1(x)=0$ en $[a,b]-N_{x_1}$; y $h_2(x)\neq 0$ en una vecindad N_{x_2} de x_2 y $h_2(x)=0$ en $[a,b]-N_{x_2}$; Si $h(x)=h_1(x)+h_2(x)$, $\|h\|<\epsilon$, entonces de acuerdo con la definición 3.3.34. Tenemos que

$J(f+h)-J(f)=\frac{dJ}{df}\Big|_{x=x_1}\Delta\sigma_1+\frac{dJ}{df}\Big|_{x=x_2}\Delta\sigma_2+o(\|h\|)$, en donde

$\Delta\sigma_1=\int_a^b h_1(x)dx$ y $\Delta\sigma_2=\int_a^b h_2(x)dx$.

Además, por una parte tenemos que $K(f+h)-K(f)=0$, mientras que por otra parte $K(f+h)-K(f)=\frac{dK}{df}\Big|_{x=x_1}\Delta\sigma_1+\frac{dK}{df}\Big|_{x=x_2}\Delta\sigma_2+o(\|h\|)$,

Preconsideremos x_2 tal que $\frac{dK}{df}\Big|_{x=x_2}\neq 0$, tal punto existe, puesto que por hipótesis $f=f(x)$ no es extremal de la funcional K, entonces

$$\Delta\sigma_2=-\left(\frac{\frac{dK}{df}\Big|_{x=x_1}}{\frac{dK}{df}\Big|_{x=x_2}}\right)\Delta\sigma_1+o(\|h\|) \quad \text{y pongamos } p=-\frac{\frac{dJ}{df}\Big|_{x=x_2}}{\frac{dK}{df}\Big|_{x=x_2}},$$

entonces sustituyendo p en $J(f+h)-J(f)=\frac{dJ}{df}\Big|_{x=x_1}\Delta\sigma_1+\frac{dJ}{df}\Big|_{x=x_2}\Delta\sigma_2+o(\|h\|)$ tenemos que

$J(f+h)-J(f)=\left(\frac{dJ}{df}\Big|_{x=x_1}+p\frac{dK}{df}\Big|_{x=x_1}\right)\Delta\sigma_1+o(\|h\|)$, puesto que las derivadas sólo están evaluadas en $x=x_1$, y el incremento de "compensación" $h_2(x)$ es tomado automáticamente al considerar la condición $K(f+h)-K(f)=0$, Así $\left(\frac{dJ}{df}\Big|_{x=x_1}+p\frac{dK}{df}\Big|_{x=x_1}\right)\Delta\sigma_1=\frac{d}{df}(J-pK)\Big|_{x=x_1}\Delta\sigma_1$ representa la parte lineal de $J(f+h)-J(f)$, y puesto que una condición necesaria para un extremo es que esta parte lineal sea identicamente cero (para toda h admisible), y ya que $\Delta\sigma_1$ no es cero, puesto que $h_1(x)\neq 0$ en una vecindad N_{x_1} de x_1, tenemos que

$\frac{d}{df}(J+pK)\Big|_{x=x_1}=0$ con x_1 arbitrario en $[a,b]$, así $\frac{d}{df}(J+pK)=0$. Lo que concluye la demostración del teorema.

<u>Ejemplo</u>. 3.3.37 (Problema Isoperimétrico).
De entre todas las curvas de una longitud determinada Q en el semiplano superior que pasan a través de los puntos $(-a,0)$ y $(a,0)$, deseamos encontrar una, tal que junto con el intervalo $[-a,a]$ encierre el área más grande. Entonces buscamos una función $f=f(x)$ para la cual la integral $\int_{-a}^{a}f\,dx$ tome el valor más grande posible, sujeta a las condiciones $f(-a)=f(a)=0$, $K(f)=\int_{-a}^{a}(1+f'^2)^{\frac{1}{2}}dx=Q$.
Consideremos primero a la integral $J(f)+pK(f)=\int_{-a}^{a}(f+p(1+f'^2)^{\frac{1}{2}})dx$, entonces la correspondiente ecuación de Euler es

$1+p\frac{d}{dx}\frac{f'}{(1+f'^2)^{\frac{1}{2}}}=0$ lo cual implica que $x+p\frac{f'}{(1+f'^2)^{\frac{1}{2}}}=C_1$,

entonces

$\frac{df}{dx}=\frac{x-C_1}{(p^2-(x-C_1)^2)^{\frac{1}{2}}}$, separando variables e integrando tenemos que

$f+C_2=\int(p^2-(x-C_1)^2)^{-\frac{1}{2}}(x-C_1)dx=(p^2-(x-C_1)^2)^{\frac{1}{2}}$. Por lo tanto $(x-C_1)^2+(f-C_2)^2=p^2$. Los valores de C_1, C_2 y p son determinados de las condiciones $f(a)=f(-a)=0$ y $K(f)=Q$. Esta ecuación representa una familia de círculos.

Generalizamos inmediatamente el problema anterior a funcionales que dependen de varias funciones de una variable independiente sujetas a condiciones subsidiarias.

Teorema. 3.3.33 Sea $(V,\|\cdot\|)$ el espacio de Banach formado por el espacio lineal $V=(C^1[a,b])^n$, con la norma $\|(f_1,..,f_n)\|=\sum_{k=1}^{n}\|f_k\|=\sum_{k=1}^{n}(\max|f(x)|+\max|f'(x)|)$ y $J,K_1,..,K_s:O=R((f_1,..,f_n),\varepsilon)\to\mathbb{R}$ s+1 funcionales de la forma $J(f_1,..,f_n)=\int_a^b F(x,f_1,..,f_n,f'_1,..,f'_n)dx$, $K_j(f_1,..,f_n)=\int_a^b G_j(x,f_1,..,f_n,f'_1,..,f'_n)dx$ $(j=1,..,s)$, en donde $F\in C^2([a,b]\times O\times O')$ y $G_j\in C^2([a,b]\times O\times O')$ para $j=1,..,s$, y $(f_1,..,f_n)\in O\subset(C^1[a,b])^n$, $(f'_1,..,f'_n)\in O'\subset(C[a,b])^n$, si el vector $(f_1,..,f_n)$ satisface las condiciones (restricciones) $f_k(a)=A_k$, $f_k(b)=B_k$ $(k=1,..,n)$ y $K_j(f_1,..,f_n)=\int_a^b G_j(x,f_1,..,f_n,f'_1,..,f'_n)dx=Q_j$ $(j=1,..,s)$, en donde los valores Q_j son valores fijos para $j=1,..,s$, y si el vector $(f_1,..,f_n)$ es extremal de la funcional J pero no de las funcionales K_j $(j=1,..,s)$, entonces existen s constantes p_j $(j=1,..,s)$ (parámetros), frecuentemente llamados multiplicadores de Lagrange, tales que el vector $(f_1,..,f_n)$ es extremal de la funcional $J+\sum_{j=1}^{s}p_jK_j$, y por lo tanto el vector $(f_1,..,f_n)$ satisface el sistema de ecuaciones diferenciales (de Euler)

$$\frac{\partial}{\partial f_k}(F+\sum_{j=1}^{s}p_jG_j)-\frac{d}{dx}\left(\frac{\partial}{\partial f'_k}(F+\sum_{j=1}^{s}p_jG_j)\right)=0\ (k=1,..n).$$

La demostración de este teorema no es esencialmente diferente a la del Teorema 3.3.30.

Las 2n constantes que aparecen en la solución del sistema de ecuaciones diferenciales (de Euler) y los valores de los s parámetros $p_1,..,p_s$ (multiplicadores de Lagrange) son determinados de las condiciones de frontera $f_k(a)=A_k$, $f_k(b)=B_k$ $(k=1,..,n)$ y de las condiciones subsidiarias $K_k(f_1,..,f_n)=\int_a^b G_j(x,f_1,...,f_n,f'_1,..,f'_n)dx=Q_j$ $(j=1,..s)$.

Consideramos ahora, un problema de tipo diferente. El problema variacional con condiciones subsidiarias finitas, como se plantea en el siguiente teorema.

Teorema. 3.3.39 Sea $(V,\|\cdot\|)$ el espacio de Banach formado por el espacio lineal $V=(C^1[a,b])^2$, con la norma $\|(f_1,f_2)\|=\|f_1\|+\|f_2\|=\max_{x\in[a,b]}|f_1(x)|+\max_{x\in[a,b]}|f_1'(x)|+\max_{x\in[a,b]}|f_2(x)|+\max_{x\in[a,b]}|f_2'(x)|$ y $J:O=P((f_1,f_2),\varepsilon)\to\mathbb{R}$ es una funcional de la forma $J(f_1,f_2)=\int_a^b F(x,f_1,f_2,f_1',f_2')dx$, en donde $F\in C^2([a,b]\times O\times O')$, $(f_1,f_2)\in O\subset(C^1[a,b])^2$, $(f_1',f_2')\in O'$ $(C[a,b])^2$, si el vector (f_1,f_2) satisface las condiciones de frontera $f_1(a)=A_1$, $f_1(b)=B_1$, $f_2(a)=A_2$ y $f_2(b)=B_2$ y, la condición subsidiaria finita, de que las curvas admisibles estén sobre la superficie $G(x,f_1,f_2)=0$, con $G\in C^1([a,b]\times O\times O')$ y G_{f_1}, G_{f_2} no se anulan simultaneamente en todos los puntos sobre la superficie, y J tiene un extremo para la curva $f_1=f_1(x)$, $f_2=f_2(x)$. Entonces existe una función $p(x)$ definida en $[a,b]$, tal que $f_1=f_1(x)$, $f_2=f_2(x)$ es extremal de la funcional $\int_a^b(F+p(x)G)dx$ y por lo tanto $f_1=f_1(x)$, $f_2=f_2(x)$ satisface el sistema de ecuaciones diferenciales (de Euler)

$$F_{f_1}+pG_{f_1}-\frac{d}{dx}F_{f_1'}=0,\quad F_{f_2}+pG_{f_2}-\frac{d}{dx}F_{f_2'}=0.$$

Demostración: Supongamos que J tiene un extremo para $f_1=f_1(x)$, $f_2=f_2(x)$, y sea $x_1\in[a,b]$. Entonces demos a $f_1(x)$ un incremento $h_1(x)$ y a $f_2(x)$ un incremento $h_2(x)$, en donde $h_1(x)$ y $h_2(x)$ son no nulas sólo en una vecindad N_{x_1} de x_1, si $\|(h_1,h_2)\|<\varepsilon$, entonces

$$J((f_1,f_2)+(h_1,h_2))-J(f_1,f_2)=\frac{dJ}{df_1}\Big|_{x=x_1}\Delta\sigma_1+\frac{dJ}{df_2}\Big|_{x=x_1}\Delta\sigma_2+o(\|(h_1,0)\|)$$
$$+o(\|(0,h_2)\|),$$ en donde $\Delta\sigma_1=\int_a^b h_1(x)dx$ y $\Delta\sigma_2=\int_a^b h_2(x)dx$, además se requiere que $G(x,f_1+h_1,f_2+h_2)=0$, y por lo tanto

$$0=\int_a^b(G(x,f_1+h_1,f_2+h_2)-G(x,f_1,f_2))dx=\int_{N_{x_1}}(G_{f_1}h_1+G_{f_2}h_2)dx+o(\|(h_1,0)\|)$$
$$+o(\|(0,h_2)\|)=(G_{f_1}\big|_{x=x_1}\Delta\sigma_1+G_{f_2}\big|_{x=x_1}\Delta\sigma_2)+o(\|(h_1,0)\|)+o(\|(0,h_2)\|).$$

La última igualdad se debe al teorema del valor medio para integrales, en donde reconsideramos a $x_1\in[a,b]$ con esta propiedad. Observamos también que la existencia de curvas admisibles $f_1(x)+h_1(x)$, $f_2(x)+h_2(x)$ en una vecindad O de la curva $f_1=f_1(x)$, $f_2=f_2(x)$, que satisfagan la condición $G(x,f_1+h_1,f_2+h_2)=0$, se sigue del Teorema de la función implícita: Si la ecuación $G(x,f_1,f_2)=0$, tiene una solución para (x_1,f_1,f_2), y si $G\in C^1(N)$, N una vecindad de (x_1,f_1,f_2), $N\subset O$ y si $G_{f_2}(x_1,f_1,f_2)\neq0$, entonces $G(x,f_1,f_2)=0$ define una única función $f_2(x,f_1)$ la cual es $C^1(M)$, M una vecindad

de (x_1,f_1). (Podemos también hacer este planteamiento para $G_{f_1}(x_1,f_1,f_2)\neq 0$).
Así, si $G_{f_2}(x,f_1(x),f_2(x))\neq 0$ en una vecindad del punto x_1, N_{x_1}.
Es posible cambiar la curva $f_1=f_1(x)$ por $f_1+h_1=(f_1+h_1)(x)$ en esta vecindad y entonces determinar $f_2(x)+h_2(x)$ a partir de la relación $f_2(x)+h_2(x)=f_2(x,f_1(x)+h_1(x))$.

Ahora bien, por hipótesis, ó $G_{f_1}|_{x=x_1}$ ó $G_{f_2}|_{x=x_1}$ son no nulas.
Si $G_{f_2}|_{x=x_1}\neq 0$, entonces a partir de
$0=(G_{f_1}|_{x=x_1}\Delta\sigma_1+G_{f_2}|_{x=x_1}\Delta\sigma_2)+o(\|(h_1,0)\|)+o(\|(0,h_2)\|)$, tenemos

$\Delta\sigma_2=-\left(\dfrac{G_{f_1}|_{x=x_1}}{G_{f_2}|_{x=x_1}}\Delta\sigma_1+o(\|(h_1,0)\|)\right)$, entonces sustituyendo en
$J((f_1,f_2)+(h_1,h_2))-J(f_1,f_2)=\dfrac{dJ}{df_1}\Big|_{x=x_1}\Delta\sigma_1+\dfrac{dJ}{df_2}\Big|_{x=x_2}\Delta\sigma_2-o(\|(h_1,0)\|)$
$+o(\|(0,h_2)\|)$, tenemos que

$$J((f_1,f_2)+(h_1,h_2))-J(f_1,f_2)=\left(\frac{dJ}{df_1}\Big|_{x=x_1}-\left(\frac{G_{f_1}}{G_{f_2}}\frac{dJ}{df_2}\right)\Big|_{x=x_1}\right)\Delta\sigma_1+o(\|(h_1,0)\|),$$

puesto que el incremento de "compensación" $h_2(x)$ es tomado automáticamente al considerar la condición $0=\int_a^b(G(x,f_1+h_1,f_2+h_2)-G(x,f_1,f_2))dx$.

El primer sumado del lado derecho constituye la parte lineal de $J((f_1,f_2)+(h_1,h_2))-J(f_1,f_2)$.

Puesto que una condición necesaria para un extremo es que esta parte lineal sea idénticamente cero para toda (h_1,h_2) admisible, y puesto que $\Delta\sigma_1\neq 0$ y x_1 es arbitrario, tenemos que

$$\frac{dJ}{df_1}-\frac{G_{f_1}}{G_{f_2}}\frac{dJ}{df_2}=F_{f_1}-\frac{d}{dx}F_{f_1'}-\frac{G_{f_1}}{G_{f_2}}\left(F_{f_2}-\frac{d}{dx}F_{f_2'}\right)=0$$

ó

$$\frac{F_{f_1}-\frac{d}{dx}F_{f_1'}}{G_{f_1}}=\frac{F_{f_2}-\frac{d}{dx}F_{f_2'}}{G_{f_2}}$$

sobre la curva $f_1=f_1(x)$, $f_2=f_2(x)$, el valor común de estos radios (cocientes) es alguna función de x, la cual denotamos por $p(x)$.
Lo que concluye la demostración del teorema.

El Problema Variacional con puntos finales variables.

Consideremos la funcional $J(f)=\int_V F(x,f,f')dx$, en donde $F\in C^2(V\times O\times O')$, $O=B(f,\epsilon)\subset C^1(V)$, $f'\in O'\subset C(V)$, O' una bola abierta.

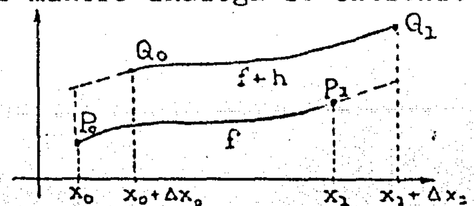
Supongamos que los puntos finales de las curvas para las cuales la funcional J es definida pueden moverse en forma "arbitraria".

Si J está definida para las curvas $f=f(x)$ y $f+h=(f+h)(x)$, definimos la distancia entre estas curvas por

$d_*(f,f+h)=\max_{x\in V}|f-(f+h)|+\max_{x\in V}|f'-(f'+h')|+d(P_0,Q_0)+d(P_1,Q_1)$,

así, haciendo $d_*(f,f+h)=\|f-(f+h)\|_*=\|h\|+d(P_0,Q_0)+d(P_1,Q_1)$, en donde $\|h\|$ es la norma usual en $C^1(V)$ y d es la distancia euclidiana, en donde P_0, Q_0 representan los puntos finales derechos, P_1, Q_1 representan los puntos finales izquierdos de las curvas $f=f(x)$ y $f+h=(f+h)(x)$, respectivamente. En general las funciones $f=f(x)$ y $f+h=(f+h)(x)$ son definidas sobre diferentes intervalos I e I', así para que $d_*(f,f+h)$ tenga sentido, extendemos f y f+h a un intervalo que contenga I e I', digamos, $V=I\cup I'$.

Si $P_0=(x_0,y_0)$, $P_1=(x_1,y_1)$, $Q_0=(x_0+\Delta x_0,y_0+\Delta y_0)$ y $Q_1=(x_1+\Delta x_1,y_1+\Delta y_1)$, y si $I=[x_0,x_1]$ e $I'=[x_0+\Delta x_0,x_1+\Delta x_1]$, como lo muestra la figura, entonces $V=[x_0,x_1+\Delta x_1]$ y extendemos linealmente a $f=f(x)$, como el segmento de recta que parte del punto $(x_1,f(x_1))$ con pendiente $f'(x_1)$ sobre el intervalo $[x_1,x_1+\Delta x_1]$, de manera análoga se extiende f+h sobre el intervalo $[x_0,x_0+\Delta x_0]$.



Entonces usando el teorema de Taylor al calcular $J(f+h)-J(f)$ tenemos que

$$J(f+h)-J(f)=\int_{x_0+\Delta x_0}^{x_1+\Delta x_1}F(x,f+h,f'+h')dx-\int_{x_0}^{x_1}F(x,f,f')dx=$$

$$\int_{x_0}^{x_1}(F(x,f+h,f'+h')-F(x,f,f'))dx+\int_{x_1}^{x_1+\Delta x_1}F(x,f+h,f'+h')dx$$

$$-\int_{x_0}^{x_0+\Delta x_0}F(x,f+h,f'+h')dx$$

$$=\int_{x_0}^{x_1}(F_fh+F_{f'}h')dx+F(x,f,f')\Big|_{x=x_1}\Delta x_1-F(x,f,f')\Big|_{x=x_0}\Delta x_0+o(\|h\|_*)$$

$$=\int_{x_0}^{x_1}(F_f-\frac{d}{dx}F_{f'})h(x)dx+F\Big|_{x=x_1}\Delta x_1+F_{f'}h\Big|_{x=x_1}-F\Big|_{x=x_0}\Delta x_0-F_{f'}h\Big|_{x=x_0}+o(\|h\|_*),$$

en donde el término que contenía h' ha sido integrado por partes. De la figura es claro que $h(x_0)=\Delta y_0+f'(x_0)\Delta x_0+o(\|h\|_*)$, $h(x_1)=\Delta y_1-f'(x_1)\Delta x_1+o(\|h\|_*)$.
Observamos también que de acuerdo a la figura $h(x)\neq 0$ en $[x_0+\Delta x_0, x_1+\Delta x_1]$ y $h(x)=0$ en otra parte antes de ser extendida $f+h$, y que

la funcional J no sólo depende de f sino también de P_0 y P_1, así

$$dJ(f,P_1,Q_1)(h,\Delta x_0,\Delta x_1,\Delta y_0,\Delta y_1)=\int_{x_0}^{x_1}(F_f-\frac{d}{dx}F_{f'})h(x)dx+F_{f'}\Big|_{x=x_1}\Delta y_1+$$

$$(F-F_{f'}f')\Big|_{x=x_1}\Delta x_1-F_{f'}\Big|_{x=x_0}\Delta y_0-(F-F_{f'}f')\Big|_{x=x_0}\Delta x_0,$$

en donde remarcamos que la diferencial de la funcional J es definida como una expresión la cual es lineal en h y su derivada, y en $\Delta x_0,\Delta x_1,\Delta y_0,\Delta y_1$.

Observamos además que para el caso particular en que las curvas están restringidas a tener sus puntos finales sobre las rectas verticales $x=x_0$, $x=x_1$, tenemos que $\Delta x_0=\Delta x_1=0$. Mientras que para el caso del problema variacional con puntos fijos, tenemos $\Delta x_0=\Delta x_1=0$ y $\Delta y_0=\Delta y_1=0$.

Este problema se generaliza para funcionales $J(f_1,..,f_n)=\int_V F(x,f_1,..,f_n,f_1',..,f_n')dx$, en donde $F\in C^2(V\times O\times O')$, $B((f_1,..,f_n),\varepsilon)=O\subset(C^1[a,b])^n$ $(f_1',..,f_n')\in O'\subset(C[a,b])^n$, O' un conjunto abierto, recordemos que $(f_1,...,f_n)$ puede ser interpretada como una curva en $\mathbb{R}^{n+1}$.

Tomamos incrementos f_k+h_k $(k=1,...,n)$ extendidos linealmente sobre V, y $P_0=(x_0,y_1^0,...,y_n^0)$, $P_1=(x_1,y_1^1,...,y_n^1)$ denotan los puntos finales de la curva $(f_1,...,f_n)$, y $Q_0=(x_0+\Delta x_0,y_1^0+\Delta y_1^0,...,y_n^0+\Delta y_n^0)$, $Q_1=(x_1+\Delta x_1,y_1^1+\Delta y_1^1,...,y_n^1+\Delta y_n^1)$ denotan los puntos finales de la curva $(f_1+h_1,...,f_n+h_n)$, y definimos $d^*((f_1,...,f_n),(f_1+h_1,..., f_n+h_n))=d_*(f_1,f_1+h_1)+...+d_*(f_n,f_n+h_n)$, ponemos $d^*((f_1,...,f_n),(f_1+h_1,...,f_n+h_n))=\|(f_1,...,f_n)-(f_1+h_1,...,f_n+h_n)\|^*=\|(h_1,...,h_n)\|^*$.

Entonces

$$J((f_1,...,f_n)-(h_1,...,h_n))-J(f_1,...,f_n)=\int_{x_0+\Delta x_0}^{x_1+\Delta x_1}F(x,f_1+h_1,...,f_n+h_n, f_1'+h_1',...,f_n'+h_n')dx-\int_{x_0}^{x_1}F(x,f_1,...,f_n,f_1',...,f_n')dx=$$

$$\int_{x_0}^{x_1}(F(x,f_1+h_1,...,f_n+h_n,f_1'+h_1',...,f_n'+h_n')-F(x,f_1,...,f_n,f_1',...,f_n'))dx$$

$$+\int_{x_1}^{x_1+\Delta x_1}F(x,f_1+h_1,...,f_n+h_n,f_1'+h_1',...,f_n'+h_n')dx$$

$$-\int_{x_0}^{x_0+\Delta x_0}F(x,f_1+h_1,...,f_n+h_n,f_1'+h_1',...,f_n'+h_n')dx.$$

Usando el teorema de Taylor, tenemos

$$\Delta J=\int_{x_o}^{x_1}\sum_{k=1}^{n}(F_{f_k}h_k+F_{f'_k}h'_k)dx-F|_{x=x_1}\Delta x_1-F|_{x=x_o}\Delta x_o+o(\|(h_1,\dots,h_n)\|^*)$$

$$=\int_{x_o}^{x_1}\sum_{k=1}^{n}(F_{f_k}-\tfrac{d}{dx}F_{f'_k})h_k(x)dx+F|_{x=x_1}\Delta x_1+\sum_{k=1}^{n}F_{f'_k}h_k|_{x=x_1}-F|_{x=x_o}\Delta x_o$$

$-\sum_{k=1}^{n}F_{f'_k}h_k|_{x=x_o}$, en donde los términos que contienen h'_k han sido integrados por partes, luego

$$dJ((f_1,\dots,f_n),P_o,P_1)((h_1,\dots,h_n),\Delta x_o,\Delta x_1,\Delta y_1^o,\Delta y_1^1,\dots,\Delta y_n^o,\Delta y_n^1)$$

$$=\int_{x_o}^{x_1}\sum_{k=1}^{n}(F_{f_k}-\tfrac{d}{dx}F_{f'_k})h_k(x)dx+F|_{x=x_1}\Delta x_1+\sum_{k=1}^{n}(F_{f'_k}(\Delta y_k^1-f'(x_1)\Delta x_1))|_{x=x_1}$$

$$-F|_{x=x_o}\Delta x_o-\sum_{k=1}^{n}(F_{f'_k}(\Delta y_k^o-f'_k(x_o)\Delta x_o))|_{x=x_o}$$

$$=\int_{x_o}^{x_1}\sum_{k=1}^{n}(F_{f_k}-\tfrac{d}{dx}F_{f'_k})h_k(x)dx+\sum_{k=1}^{n}F_{f'_k}|_{x=x_1}\Delta y_k^1+(F-\sum_{k=1}^{n}f'_kF_{f'_k})|_{x=x_1}\Delta x_1$$

$$-\sum_{k=1}^{n}F_{f'_k}|_{x=x_o}\Delta y_k^o+(F-\sum_{k=1}^{n}f'_kF_{f'_k})|_{x=x_o}\Delta x_o.$$

Daremos ahora una fórmula más concisa para dJ.

Sea $p_k=F_{f'_k}$ $(k=1,\dots,n)$ y supongamos que el jacobiano

$$\frac{\partial(p_1,\dots,p_n)}{\partial(f'_1,\dots,f'_n)}=\det(F_{f'_jf'_k})\neq 0.$$

Entonces podemos resolver localmente $p_k=F_{f'_k}$ para $f'_1,\dots,f'_n$ como funciones de las variables $x,f_1,\dots,f_n,p_1,\dots,p_n$. Definimos una nueva función $H=H(x,f_1,\dots,f_n,p_1,\dots,p_n)$ como

$H=-F+\sum_{k=1}^{n}f'_kF_{f'_k}=-F+\sum_{k=1}^{n}f'_kp_k$, en donde las funciones f'_k son vistas como funciones de $x,f_1,\dots,f_n,p_1,\dots,p_n$. La función H es llamada el Hamiltoniano correspondiente a la funcional $J(f_1,\dots,f_n)$. Las cantidades $x,f_1,\dots,f_n,p_1,\dots,p_n,H$ son llamadas variables canónicas, entonces

$$dJ(f_1,\dots,f_n),P_o,P_1)((h_1,\dots,h_n),\Delta x_o,\Delta x_1,\Delta y_1^o,\Delta y_1^1,\dots,\Delta y_n^o,\Delta y_n^1)$$

$$=\int_{x_o}^{x_1}\sum_{k=1}^{n}(F_{f_k}-\tfrac{d}{dx}p_k)h_k(x)dx+\sum_{k=1}^{n}p_k|_{x=x_1}\Delta y_k^1-H|_{x=x_1}\Delta x_1$$

$$-\sum_{k=1}^{n}p_k|_{x=x_o}\Delta y_k^o+H|_{x=x_o}\Delta x_o.$$

Si J tiene un extremo para el vector $(f_1,\ldots,f_n)$ comparado con todas las curvas admisibles, entonces J tiene un extremo para $(f_1,\ldots,f_n)$ comparada con todas las curvas con puntos fijos, de donde si $(f_1,\ldots,f_n)$ es extremal, entonces $F_{f_k}-\frac{d}{dx}F_{f'_k}=0$ $(k=1,\ldots,n)$. Entonces dJ toma la forma

$$dJ((f_1,\ldots,f_n),P_o,P_1)((h_1,\ldots,h_n),\Delta x_o,\Delta x_1,\Delta y_1^o,\Delta y_1^1,\ldots,\Delta y_n^o,\Delta y_n^1)=$$

$$=\sum_{k=1}^{n} p_k\big|_{x=x_1}\Delta y_k^1-H\big|_{x=x_1}\Delta x_1-\sum_{k=1}^{n} p_k\big|_{x=x_o}\Delta y_k^o+H\big|_{x=x_o}\Delta x_o.$$

Ahora bien, si $H=-F+\sum_{k=1}^{n}f'_k p_k$, entonces por definición de H tenemos que

$dH=-dF+\sum_{k=1}^{n}p_k df'_k+\sum_{k=1}^{n}f'_k dp_k$, así que

$$dH=-\frac{\partial F}{\partial x}dx-\sum_{k=1}^{n}\frac{\partial F}{\partial f_k}df_k-\sum_{k=1}^{n}\frac{\partial F}{\partial f'_k}df'_k+\sum_{k=1}^{n}p_k df'_k+\sum_{k=1}^{n}f'_k dp_k,$$

podríamos ahora expresar a df'_k en términos de $x,f_1,\ldots,f_n,p_1,\ldots,p_n$; pero $\frac{\partial F}{\partial f'_k}=p_k$ $(k=1,\ldots,n)$, entonces los términos que contienen df'_k son cancelados, de donde

$dH=-\frac{\partial F}{\partial x}dx-\sum_{k=1}^{n}\frac{\partial F}{\partial f_k}df_k+\sum_{k=1}^{n}f'_k dp_k$, y por lo tanto

$$\frac{\partial H}{\partial x}=-\frac{\partial F}{\partial x}\quad,\quad\frac{\partial H}{\partial f_k}=-\frac{\partial F}{\partial f_k}\quad y\quad\frac{\partial H}{\partial p_k}=f'_k.$$

En otras palabras F_{f_k} y f'_k estan relacionadas por medio de las derivadas parciales del Hamiltoniano H, finalmente podemos escribir las ecuaciones de Euler (Euler-Lagrange)

$$\frac{df_k}{dx}=f'_k,\quad F_{f_k}-\frac{d}{dx}F_{f'_k}=0\ (k=1,\ldots,n)\ \text{como}$$

$$\frac{df_k}{dx}=\frac{\partial H}{\partial p_k},\quad\frac{dp_k}{dx}=-\frac{\partial H}{\partial f_k}\qquad(k=1,\ldots,n).$$

Ejemplo. 3.3.40 De entre todas las curvas diferenciables cuyos puntos finales P_o y P_1 están sobre dos curvas dadas $v=v(x)$ y $w=w(x)$, encontrar una curva para la cual la funcional $J(f)=\int_{x_o}^{x_1}F(x,f,f')dx$ tenga un extremo.

Supongamos primero que la funcional J tiene un extremo para la curva $f=f(x)$, entonces $f=f(x)$ es solución a la ecuación de Euler correspondiente, y por lo tanto $\int_{x_0}^{x_1}(F_f-\frac{d}{dx}F_{f'})h(x)dx=0$, y dJ tiene la forma

$$dJ(f,P_0,P_1)(h,\Delta x_0,\Delta x_1,\Delta y_1,\Delta y_0)=F_{f'}|_{x=x_1}\Delta y_1+(F-F_{f'}f')|_{x=x_1}\Delta x_1$$
$$-F_{f'}|_{x=x_0}\Delta y_0-(F-F_{f'}f')|_{x=x_0}\Delta x_0,$$

la cual se debe anular si J tiene un extremo para $f=f(x)$.
De acuerdo a la figura observamos que $\Delta y_0=(v'(x)+\epsilon_0)\Delta x_0$, $\Delta y_1=(w'(x)+\epsilon_1)\Delta x_1$, en donde $\epsilon_0 \to 0$ cuando $x_0 \to 0$ y $\epsilon_1 \to 0$ cuando $\Delta x_1 \to 0$. Así la condición $dJ(f,P_0,P_1)=0$ toma la forma

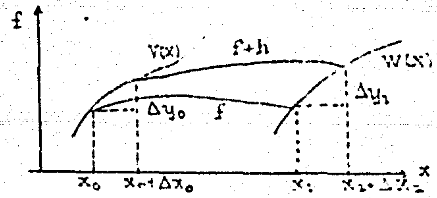
$$dJ(f,P_0,P_1)(h,\Delta x_0,\Delta x_1,\Delta y_0,\Delta y_1)=(F_{f'}w'+F-F_{f'}f')|_{x=x_1}\Delta x_1-$$
$$(F_{f'}v'+F-F_{f'}f')|_{x=x_0}\Delta x_0=0.$$

Y puesto que los incrementos Δx_0 y Δx_1 son independientes, entonces

$$(F_{f'}v'+F-F_{f'}f')|_{x=x_0}=0,\quad (F_{f'}w'+F-F_{f'}f')|_{x=x_1}=0$$

ó $(F+(v'-f')F_{f'})=0$, $(F+(w'-f')F_{f'})=0$,

estas condiciones de frontera son frecuentemente llamadas condiciones de transversalidad.



Ejemplo. 3.3.41 De entre todas las curvas diferenciables cuyos puntos finales P_0 y P_1 están sobre dos curvas dadas $v=v(x)$ y $w=w(x)$, deseamos encontrar una curva que sea extremal de la funcional $J(f)=\int_{x_0}^{x_1} f(x,y)(1+f'^2)^{\frac{1}{2}}dx$.
Para este caso las condiciones de transversalidad tienen una apariencia particularmente simple. De hecho en este caso

$$F_{f'}=f(x,y)f'(1+f'^2)^{-\frac{1}{2}}=f'F(1+f'^2)^{-1},$$

así que las condiciones de transversalidad tienen la forma

$$F+(v'-f')F_{f'}=\frac{(1+f'v')F}{1+f'^2}=0,\quad F+(w'-f')F_{f'}=\frac{(1+f'w')F}{1+f'^2}=0,$$

lo cual implica que $f'=(-v')^{-1}$ en el punto final izquierdo, mientras que $f'=(-w')^{-1}$ en el punto final derecho.

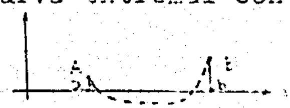
<u>Ejemplo</u>. 3.3.42 Deseamos encontrar la curva de menor longitud que conecta a dos punto dados (a,A) y (b,B).
En este caso deseamos minimizar la funcional $J(f)=\int_a^b (1+f'^2)^{\frac{1}{2}}dx$.
El objetivo de este ejercicio es ilustrar el empleo del Hamiltoniano.
Consideremos la funcional $J(f)=\int_a^b F(x,f,f')dx$, y pongamos $p=F_{f'}(x,f,f')$ y $H(x,f,p)=-F+pf'$, supongamos que $F_{f'f'}\neq 0$, entonces f' es una función de x,f,p. Entonces se introduce la siguiente funcional $J(f,p)=\int_a^b(-H(x,f,f')+pf')dx$, en donde f y p son vistas como variables independientes. Las ecuaciones de Euler para la funcional $J(f,p)$ son

$-\frac{\partial H}{\partial f}-\frac{dp}{dx}=0$, $-\frac{\partial H}{\partial p}+\frac{df}{dx}=0$, además $dH=-\frac{\partial F}{\partial x}dx-\frac{\partial F}{\partial f}df+f'dp$, entonces $\frac{\partial H}{\partial p}=f'$,
y por lo tanto $-H+p\frac{\partial H}{\partial p}=F-pf'+pf'=F$, entonces las funcionales $J(f)$ y $J(f,p)$ tienen extremos para las mismas curvas.
Regresando ahora a nuestro problema original encontramos que $p=F_{f'}=f'(1+f'^2)^{-\frac{1}{2}}$ ó $p^2=f'^2(1+f'^2)^{-1}$, de donde $f'^2=p^2(1-p^2)^{-1}$,
el Hamiltoniano esta dado por
$H=-(1+f'^2)^{\frac{1}{2}}+f'^2(1+f'^2)^{-\frac{1}{2}}=-(1+f'^2)^{-\frac{1}{2}}=-(1-p^2)^{\frac{1}{2}}$, así $\frac{dp}{dx}=-\frac{\partial H}{\partial f}=0$, entonces $p=C$, C constante, y

$\frac{df}{dx}=\frac{\partial H}{\partial p}=p(1-p^2)^{-\frac{1}{2}}$, entonces $f=C_1x+C_2$.

Extremales con esquinas

Consideremos el ejemplo 3.3.36, en donde se desea encontrar una curva que une dos puntos dados (a,A) y (b,B) la cual genera la superficie de área mínima cuando ésta se rota alrededor del eje x. Es claro que si A y B son suficientemente pequeños comparados con $b-a$, la solución del problema está dada por la curva extremal con esquinas $AabB$, como se muestra en la figura.



Este extremal consiste de tres segmentos de línea (dos verticales y uno horizontal) y puede ser incluído en la clase de funciones $PC[a,b]$ (el espacio de funciones continuas por pedazos definidas sobre $[a,b]$).
Tomando en cuenta las consideraciones anteriores, planteamos el siguiente problema variacional.

De entre todas las funciones en $PC[a,b]$ que además satisfacen tener derivada continua excepto posiblemente en algún punto c ($a<c<b$), las cuales satisfacen las condiciones de frontera $f(a)=A$, $f(b)=B$, deseamos encontrar una función f para la cual la funcional $J(f)=\int_a^b F(x,f,f')dx$ tenga un extremo débil, en donde $F\in C^2([a,c]\times O_c\times O_c')$ y $F\in C^2([c,b]\times O_c\times O')$, $O_c=B(f|_{[a,c]},\epsilon)\subset C^1[a,c]$, $f'|_{[a,c]}\in O_c'$, O_c' una bola abierta de $C[a,c]$ con centro en f', consideraciones análogas se hacen para $f|_{[c,b]}$. Es claro que sobre cada uno de los intervalos $[a,c]$ y $[c,b]$ la función para la cual J alcanza un extremo debe satisfacer la ecuación de Euler $F_f-\frac{d}{dx}F_{f'}=0$. Consideremos a J como la suma de dos funcionales J_1 y J_2,

$J(f)=\int_a^b F(x,f,f')dx=\int_a^c F(x,f,f')dx+\int_c^b F(x,f,f')dx\equiv J_1(f)+J_2(f)$ y calculemos las diferenciales $dJ_1(f)$ y $dJ_2(f)$ de las funcionales J_1 y J_2 separadamente. Los puntos finales para $x=a$, $x=b$ son fijos y buscamos dos pedazos de la función $f=f(x)$ que se unen continuamente en $x=c$, en donde el punto para el cual $x=c$ puede moverse arbitrariamente, entonces

$$dJ_1(f)(h,\Delta x_1,\Delta y_1)=F_{f'}|_{x=c-0}\Delta y_1+(F-F_{f'}f')|_{x=c-0}\Delta x_1,$$
$$dJ_2(f)(h,\Delta x_1,\Delta y_1)=F_{f'}|_{x=c+0}\Delta y_1+(F-F_{f'}f')|_{x=c+0}\Delta x_1.$$

La condición de que $f=f(x)$ sea continua en $x=c$ implica $dJ(f)=dJ_1(f)+dJ_2(f)\equiv 0$ y por lo tanto

$$(F_{f'}|_{x=c-0}-F_{f'}|_{x=c+0})\Delta y_1+((F-f'F_{f'})|_{x=c-0}-(F-f'F_{f'})|_{x=c+0})\Delta x_1=0,$$

y ya que Δx_1 y Δx_1 son arbitrarios tenemos que

$$F_{f'}|_{x=c-0}=F_{f'}|_{x=c+0},\ (F-f'F_{f'})|_{x=c-0}=(F-f'F_{f'})|_{x=c+0}.$$

Estas condiciones de esquina son llamadas las condiciones de Weierstrass-Erdmann.

Condiciones necesarias de extremo considerando la diferencial segunda

Teorema. 3.3.43 (Lagrange) Sea $(V,\|\cdot\|)$ el espacio de Banach formado por el espacio lineal $V=C^1[a,b]$ con la norma $\|f\|=\max_{x\in[a,b]}|f(x)|+\max_{x\in[a,b]}|f'(x)|$, y $J:O=B(f,\epsilon)\subset C^1[a,b]\to\mathbb{R}$ una funcional de la forma $J(f)=\int_a^b F(x,f,f')dx$, en donde $F\in C^3([a,b]\times O\times O')$, $O=B(f,\epsilon)\subset C^1[a,b]$ y $f'\in O'\subset C[a,b]$, O' una bola abierta, si f satisface las condiciones de frontera $f(a)=A$, $f(b)=B$. Entonces una condición necesaria para que la funcional J tenga un mínimo para la curva $f=f(x)$ es que la desigualdad $F_{f'f'}\geq 0$ se verifique en todos los puntos sobre la curva $f=f(x)$.

Demostración: Sea $h\in C_0[a,b]\cap C^1[a,b]$ tal que $\|h\|<\varepsilon$, entonces calculando $J(f+h)-J(f)$ y haciendo uso del teorema de Taylor tenemos que

$$J(f+h)-J(f)=\int_a^b(F_f h+F_{f'}h')dx+\tfrac{1}{2}\int_a^b(F_{ff}h^2+2F_{ff'}hh'+F_{f'f'}(h')^2)dx+o(\|h\|^2).$$

Así $d^2J(f)(h,h)=\frac{1}{2}\int_a^b(F_{ff}h^2+2F_{ff'}hh'+F_{f'f'}(h')^2)dx$.

Notamos además que

$$\int_a^b 2F_{ff'}hh'dx=2F_{ff'}h^2\Big|_a^b-\int_a^b\frac{d}{dx}(2F_{ff'}h)hdx=-2\int_a^b(\frac{d}{dx}F_{f'f'})h^2dx-2\int_a^b F_{f'f'}hh'dx$$

lo cual implica que $\int_a^b 2F_{f'f'}hh'dx=-\int_a^b(\frac{d}{dx}F_{f'f'})h^2dx$, entonces haciendo $R_1=R_1(x)=\frac{1}{2}F_{f'f'}$ y $R_2=R_2(x)=\frac{1}{2}(F_{ff}\quad-\frac{d}{dx}F_{ff'})$, tenemos que
$d^2J(f)(h,h)=\int_a^b(R_1h'^2+R_2h^2)dx$.

Sabemos que si una función $h(x)$ toma valores pequeños en $[a,b]$, si su derivada $h'(x)$ toma valores pequeños en $[a,b]$. Sin embargo, el inverso no es válido. Esto implica que $R_1h'^2$ juega un papel predominante en $d^2J(f)$, en el sentido de que $R_1h'^2$ puede ser mucho más grande que R_2h^2 y no puede ser mucho más pequeño que R_2h^2, de donde se puede esperar que $R_1(x)$ determine si $d^2J(f)$ toma valores con un sólo signo o con ambos. Ahora precisamos este argumento.

Una condición necesaria para que la funcional cuadrática $d^2J(f)(h,h)=\int_a^b(R_1h'^2+R_2h^2)dx$ definida para $h\in C_0[a,b]\cap C^1[a,b]$ sea no negativa, es que $R_1(x)\geq 0$ para todo $x\in[a,b]$, y en consecuencia $F_{f'f'}\geq 0$, entonces haciendo uso del Teorema 3.2.20 concluiremos la demostración del teorema.

Ahora bien, supongamos que $R_1(x_0)<0$ para algún $x_0\in[a,b]$, digamos $R_1(x_0)=2b$, $b<0$, para algún $x_0\in[a,b]$, entonces puesto que $R_1(x)$ es una función continua sobre $[a,b]$ existe un $c>0$ tal que $R_1(x)<-b$ para $x\in[x_0-c,x_0+c]$. Ahora construimos $h(x)\in C_0[a,b]\cap C^1[a,b]$ como

$$h(x)=\begin{cases}\operatorname{sen}^2(\frac{\pi(x-x_0)}{c}) & ,x_0-c\leq x\leq x_0+c.\\ 0 & ,x\in[b,a]-[x_0-c,x_0+c].\end{cases}$$

Entonces $\int_a^b(R_1h'^2+R_2h^2)dx=\int_{x_0-c}^{x_0+c}R_1\frac{\pi^2}{c^2}(2\cos(\frac{\pi(x-x_0)}{c})\operatorname{sen}(\frac{\pi(x-x_0)}{c}))^2dx$
$+\int_{x_0-c}^{x_0+c}R_2\operatorname{sen}^4(\frac{\pi(x-x_0)}{c})dx=\int_{x_0-c}^{x_0+c}R_1\frac{\pi^2}{c^2}\operatorname{sen}^2(\frac{2\ (x-x\ )}{c})dx+\int_{x_0-c}^{x_0+c}R_2\operatorname{sen}^4(\frac{\pi(x-x\ )}{c})dx$
$<-(2b\pi)/c)+2Mc$, en donde $M=\max_{a\leq x\leq b}|R_2(x)|$, para c suficientemente pequeña $-((2b\pi)/c)+2Mc<0$. Esto prueba el teorema.

Discutiremos ahora, sobre la existencia de extremos globales considerando la forma de la función F.

Teorema. 3.3.44 Supongamos que $F(x,f,f')$ es tal que $F\in C^2([a,b]\times O\times O')$ $(f,f')\in O\times O'$ un conjunto abierto y convexo, y que F es una función convexa, para x en $[a,b]$. Sea f una función admisible que satisface la ecuación de Euler-Lagrange correspondiente. Entonces la funcional $J(f)=\int_a^b F(x,f,f')dx$ tiene un mínimo global.

Demostración: Sea f+h cualquier otra función admisible, $h\in C_o[a,b]\cap C^1[a,b]$. Y puesto que F es convexa, tenemos que $F(x,f+h,f'+h)\geq F(x,f,f')+hF_f+h'F_{f'}$, entonces integrando ambas partes de la desigualdad tenemos que

$J(f+h)\geq J(f)+\int_a^b(F_fh+F_{f'}h')dx=J(f)+\int_a^b(F_f+\frac{d}{dx}F_{f'})hdx$, y puesto que f satisface la ecuación de Euler se tiene que $F_f-\frac{d}{dx}F_{f'}=0$, por tanto $J(f+h)\geq J(f)$ para toda f+h admisible, puesto que f+h era arbitraria.

El caso en que F es una función concava y J tiene un máximo global se sigue por un razonamiento análogo al Teorema anterior.

Teorema. 3.3.45 Bajo las hipótesis del Teorema anterior (3.3.37). Si F es extrictamente convexa en (f,f') sobre $O\times O'$, entonces la función $f=f(x)$ obtenida a partir de la ecuación de Euler es única.

Demostración: Supongamos que $f=f(x)$ no es única y que existe otra función $g=g(x)$ tal que maximiza a la funcional J.

Ahora sea $h(x)=cf(x)+(1-c)g(x)$, $c\in[0,1]$, es claro que h satisface también la ecuación de Euler, y puesto que la funcion F es estrictamente convexa, tenemos que

$F(x,h,h')<cF(x,f,f')+(1-c)F(x,g,g')$, integrando ambos lados de la desigualdad tenemos que $J(h)<cJ(f)+(1-c)J(g)$ ó $J(h)<J(f)$ lo que contradice el hecho de que $J(f)$ fuera mínimo.

Hasta aquí, para resolver un problema variacional determinado, este se ha reducido a un problema que envuelve ecuaciones diferenciales. Sin embargo, esto no es siempre efectivo, y generalmente es muy complicado, especialmente cuando resultan varias variables independientes obteniendo de esta manera ecuaciones diferenciales parciales. Por lo que es necesario buscar métodos variacionales de un tipo diferente, conocidos como métodos directos.

Consideremos el problema de encontrar el mínimo de una funcional J definida sobre un conjunto abierto O de un espacio lineal normado V, en donde los elementos del conjunto O son llamados funciones admisibles. Para que el problema tenga sentido, supongamos que existen funciones $g\in O$ para las cuales se satisface que $J(g)<+\infty$, y más aún que $\inf_{f\in O} J(f)=m>-\infty$, entonces por definición de m, existe una sucesión infinita de funciones $\{f_n\}_{n=1}^{\infty}$, llamada una sucesión minimizante, tal que $\lim_{n\to\infty} J(f_n)=m$.

Si la sucesión $\{f_n\}_{n=1}^{\infty}$ tiene un límite g en el sentido de la norma en V, y si $J(g)=\lim_{n\to\infty} J(f_n)$, es decir $J(\lim_{n\to\infty} f_n)=\lim_{n\to\infty} J(f_n)$, entonces $J(g)=m$ y g es la solución del problema variacional. Más aún, las funciones de la sucesión minimizante pueden ser consideradas como soluciones aproximadas del problema variacional.

Así, para resolver un problema variacional dado, por métodos directos, debemos:

(i). Construir una sucesión minimizante $(\lim_{n\to\infty} J(f_n)=\inf_{f\in O} J(f)=m)$.

(ii). Probar que $\{f_n\}_{n=1}$ tiene un límite, en el sentido de convergencia en la norma de V, $f_n \xrightarrow{\|\cdot\|_V} g$.

(iii). Probar que $J(g)=\lim_{n\to\infty} J(f_n)$.

<u>Ejemplo</u>. 3.3.46 Consideremos la funcional $J(f)=\int_{-1}^{1} x^2 f'^2 dx$, en donde f debe satisfacer las condiciones de frontera $f(-1)=-1$, $f(1)=1$.

Obviamente la funcional J toma sólo valores positivos, además $\inf_f J(f)=0$. Escojamos la sucesión minimizante $f_n(x)=\dfrac{\arctan nx}{\arctan n}$ $n=1,2,\ldots$, $x\in[-1,1]$. En efecto

$$J(f_n)=\int_{-1}^{1} \frac{n^2x^2dx}{(\arctan n)^2(1+n^2x^2)^2} < \frac{1}{(\arctan n)^2}\int_{-1}^{1}\frac{dx}{1+n^2x^2} = \frac{2}{n \arctan n}$$

$\to 0$ cuando $n\to\infty$. Sin embargo, la sucesión $\{f_n\}_{n=1}^{\infty}$ no converge en el sentido de la norma $\|f\|=\max_{x\in[-1,1]}|f(x)|$ en la clase de funciones continuas que satisfacen las condiciones de frontera.

En conclusión, este ejemplo muestra que existe una sucesión minimizante para un problema variacional dado, y sin embargo, la sucesión $\{f_n\}_{n=1}^{\infty}$ no converge en el sentido de la norma usual de $C[a,b]$.

Recordemos que una funcional J es llamada continua en un punto $g\in O$, O un conjunto abierto de un espacio lineal normado V, si para todo $\epsilon>0$, existe $\delta>0$ tal que $\|f-g\|<\delta$, $f\in O$, implica $|J(f)-J(g)|<\epsilon$.

<u>Definición</u>. 3.3.47 Una funcional J es llamada semicontinua inferiormente en el punto $g\in O$, O un subconjunto abierto de un espacio lineal normado V, si para todo $\epsilon>0$, existe $\delta>0$ tal que $\|f-g\|<\delta$, $f\in O$ implica $J(f)-J(g)>-\epsilon$.
De manera análoga se define una funcional semicontinua superiormente, al considerar $J(f)-J(g)<\epsilon$.

<u>Ejemplo</u>. 3.3.48 Consideremos la funcional $J(f)=\int_a^b (1+f'^2)^{\frac{1}{2}}dx$, denotemos $L(C)=J(f)$ la longitud de la curva C definida por la función f, denotemos también por C a la recta que une los puntos $(a,f(a))$ y $(b,f(b))$. Entonces J es semicontinua inferiormente en f, puesto que si una sucesión $\{C_n\}_{n=1}^{\infty}$ de curvas converge a C, entonces la sucesión de longitudes $\{L(C_n)\}_{n=1}^{\infty}$ debe aproximarse a $L(C)$ y pueden ser mucho más grandes si las curvas C_n oscilan lo suficiente.

<u>Teorema</u>. 3.3.49 Si $\{f_n\}_{n=1}^{\infty}$ es una sucesión minimizante de la funcional J, con una función límite g, y si la funcional J es semicontinua inferiormente en g, entonces $J(g)=\lim_{n\to\infty} J(f_n)$.
Demostración: Por una parte tenemos que por definición de ínfimo $J(g)\geq\lim_{n\to\infty} J(f_n)=\inf J(f)$.
Mientras que por otro lado, puesto que J es semicontinua inferiormente en g, tenemos $J(f_n)-J(g)>-\epsilon$, si n es suficientemente grande, entonces $J(g)\leq\lim_{n\to\infty} J(f_n)+\epsilon$ o $J(g)\leq\lim_{n\to\infty} J(f_n)$ puesto que $\epsilon>0$ es arbitraria, así $J(g)\geq\lim_{n\to\infty} J(f_n)$ y $J(g)\leq\lim_{n\to\infty} J(f_n)$. Por tanto $J(g)=\lim_{n\to\infty} J(f_n)$.

Ahora describimos el método de Ritz, uno de los métodos directos más comunes. Supongamos que la funcional J esta definida en algún conjunto abierto O de funciones admisibles de un espacio lineal normado V, y supongamos además que J tiene un mínimo para f y $O=V$. Sea $\{g_n\}_{n=1}^{\infty}$ una sucesión de funciones en V y sea V_n el subespacio lineal n-dimensional generado por las primeras n funciones de $\{g_n\}_{n=1}^{\infty}$, esto es $V_n=\{g;g=a_1g_1+\ldots+a_ng_n,\ a_k\in\mathbb{R}\ (k=1,\ldots,n)\}$. Entonces sobre cada subespacio V_n la funcional J induce la función de n variables $U(a_1,\ldots,a_n)=J(a_1g_1+,\ldots,a_ng_n)$, escojamos $a_1,\ldots,a_n$ de tal manera que se minimice $J(a_1g_1+\ldots+a_ng_n)$, denotando el mínimo por m_n, y el elemento de V_n para el cual J toma el mínimo por f_n, esto es $m_n=J(f_n)$. Claramente la sucesión m_n es no creciente, puesto que cualquier combinación lineal de $g_1,\ldots,g_n$ es automáticamente una combinación lineal de $g_1,\ldots,g_n,g_{n+1}$, pues $V_n\subset V_{n+1}$ y $m_{n+1}=J(f_{n+1})\leq J(a_1g_1+\cdots+a_ng_n+0\cdot g_{n+1})=J(a_1g_1+\cdots+a_ng_n)$ y por definición de ínfimo, se tiene que $m_{n+1}\leq m_n$.

<u>Definición</u>. 3.3.50 Una sucesión $\{g_n\}_{n=1}$ es llamada completa en el espacio lineal normado V, si dado cualquier $f\in V$ y $\epsilon>0$, existe una combinación lineal $a_1g_1+\ldots+a_ng_n$ tal que $\|(a_1g_1+\ldots+a_ng_n)-f\|<\epsilon$, en donde $\epsilon=\epsilon(n)$.

<u>Teorema</u>. 3.3.51 Si una funcional J es continua (continua en la norma de V), y si la sucesión $\{g_n\}_{n=1}^{\infty}$ es completa, entonces $\lim\limits_{n\to\infty} m_n=m$, en donde $m=\inf\limits_f J(f)$.

Demostración: Dado $\epsilon>0$, sea g tal que $J(g)<m+\epsilon$, tal g existe para cualquier $\epsilon>0$ por definición de ínfimo. Puesto que J es continua $|J(f)-J(g)|<\epsilon$ siempre que $\|f-g\|<\delta=\delta(\epsilon)$. Sea $a_1g_1+\ldots+a_ng_n$ una combinación lineal tal que $\|(a_1g_1+\ldots+a_ng_n)-g\|<\delta$ tal combinación lineal existe para n suficientemente grande, puesto que por hipótesis la sucesión $\{g_n\}_{n=1}^{\infty}$ es completa en el espacio. Sea f_n la combinación lineal para la cual J alcanza su mínimo, entonces $m\leq m_n=J(f_n)\leq J(a_1g_1+\ldots.a_ng_n)<J(g)+\epsilon<m+2\epsilon$ y puesto que $\epsilon>0$ es arbitrario tenemos que $\lim\limits_{n\to\infty} J(f_n)=\lim\limits_{n\to\infty} m_n=m$.

Finalmente describimos el método de diferencias finitas.

Consideremos el problema de encontrar un extremo de la funcional $J(f)=\int_a^b F(x,f,f')dx$, $f(a)=A$, $f(b)=B$, este puede ser aproximado por el problema de encontrar un extremo de la función de n variables, que se obtiene como sigue: Primero dividimos $[a,b]$ en $n+1$ subintervalos iguales introduciendo los puntos $x_0=a,x_1,\ldots,x_n,x_{n+1}=b$, $\Delta x_{i+1}=x_{i+1}-x_i$, $(i=0,1,\ldots,n)$, y reemplacemos la función $f=f(x)$ por la línea poligonal con vértices $(x_0,y_0),(x_1,y_1),\ldots,(x_{n+1},y_{n+1})$, en donde $y_0=A$ y $y_{n+1}=B$, denotemos $y_i=f(x_i)$, $(i=0,1\ldots,n)$. Entonces la fun cional J puede ser aproximada por la suma

$$J(v_1,\ldots,v_n)=\sum_{i=0}^{n} F(x_i,v_i,\ \frac{v_{i+1}-v_i}{\Delta x_{i+1}})\Delta x_{i+1},$$

la cual es una función de n variables. Si para cada n, es posible encontrar una línea poligonal que minimice a $J(v_1,\ldots,v_n)$, obtendremos una sucesión de soluciones aproximadas al problema variacional original.

Ejemplo. 3.3.52 Consideremos el problema de la braquistocrona, que consiste en determinar la trayectoria que describe una partícula al moverse de un punto dado a otro cuando está sujeta al efecto de la gravedad. A través del ejemplo 3.3.16 este problema fue planteado como sigue: Minimizar

$$J(f)=(2g)^{-\frac{1}{2}}\int_0^b \sqrt{\frac{1+f'^2}{f}}\,dx$$

sujeto a

$$f(0)=A, f(b)=B,\ f\in C^1[0,b]\ \text{ y }\ f''\in PC[0,b].$$

Empleando el método de diferencias finitas encontramos que la funcional J se transforma en una función de N variables y presenta el siguiente aspecto:

$$J(f_1,\ldots,f_N)=(2g)^{-\frac{1}{2}}\sum_{k=1}^{N-1}\left(\frac{1+\frac{(f_{k+1}-f_k)^2}{h^2}}{f_k}\right)^{\frac{1}{2}}\cdot h\ ;\ f_k=f(t_k),\ \ t_k=kh,\ k=1,\ldots N,\ h=\frac{b}{N}$$

Para calcular las derivadas parciales de J, notamos que f_k aparece justamente en dos términos de J, luego

$$\frac{\partial J}{\partial f_k}=\frac{\partial}{\partial f_k}\left(\left(\frac{h^2+(f_{k+1}-f_k)^2}{f_k}\right)^{\frac{1}{2}}+\left(\frac{h^2+(f_k-f_{k-1})^2}{f_{k-1}}\right)^{\frac{1}{2}}\right)$$

después de algunas manipulaciones algebraicas encontramos que

$$\frac{\partial J}{\partial f_k}=\left\{\frac{h^{-2}((f_{k+1}-f_k)-(f_k-f_{k-1}))}{\left(f_k\left[\frac{h^2+(f_{k+1}-f_k)^2}{h^2}\right]\right)^{1/2}}+\frac{\frac{h^2+(f_{k+1}-f_k)^2}{h^2}}{2f_k\left(f_k\left[\frac{h^2+(f_{k+1}-f_k)^2}{h^2}\right]\right)^{1/2}}\right\}\cdot h$$

$$+\frac{(f_k-f_{k-1})}{h}\left[\ (f_k(\frac{h^2+(f_{k+1}-f_k)^2}{h^2}))^{\frac{1}{2}}-(f_{k-1}(\frac{h^2+(f_k-f_{k-1})^2}{h^2}))^{-\frac{1}{2}}\right]$$

cuando $h \to 0$ el número de puntos de la partición se incrementa sin límite, obviamente cuando $h \to 0$ el primer sumando de la última expresión tiende a cero, puesto que es una cantidad de orden h, el segundo sumando se anula por sí mismo como una diferencia de dos términos idénticos. En general la expresión determinada en el primer sumando

$$\frac{h^{-2}((f_{k+1}-f_k)-(f_k-f_{k-1}))}{\left(f_k\left[\frac{h^2+(f_{k+1}-f_k)^2}{h^2}\right]\right)^{1/2}}+\frac{\frac{h^2+(f_{k+1}-f_k)^2}{h^2}}{2f_k\left(f_k\left[\frac{h^2+(f_{k+1}-f_k)^2}{h^2}\right]\right)^{1/2}}$$

es distinta de cero. Sabemos que una condición necesaria de extremo es que

$\frac{\partial J}{\partial f_k}=0$ para toda k $(k=1,\ldots,N)$, si además $h \to 0$, entonces

$$\lim_{h\to 0}\left(\frac{h^{-2}(f_{k+1}-f_k)-(f_k-f_{k-1}))}{\left(f_k\left[\frac{h^2+(f_{k+1}-f_k)^2}{h^2}\right]\right)^{1/2}}+\frac{\frac{h^2+(f_{k+1}-f_k)^2}{h^2}}{2f_k\left(f_k\left[\frac{h^2+(f_{k+1}-f_k)^2}{h^2}\right]\right)^{1/2}}\right)=0\ ,$$

lo cual implica que

$$\frac{f''}{(f(1+f'^2))^{\frac{1}{2}}}+\frac{(1+f'^2)}{2f(f(1+f'^2))^{\frac{1}{2}}}=0$$

ó

$$\frac{2ff'f''+f'(1+f'^2)}{2(f(1+f'^2))^{\frac{1}{2}}}=0$$

equivalentemente

$\frac{d}{dx}((f(1+f'^2))^{\frac{1}{2}})=0$, entonces $(f(1+f'^2))^{-\frac{1}{2}}=c^{-\frac{1}{2}}$,

de donde

$$f'=\left(\frac{c}{f}-1\right)^{\frac{1}{2}}$$

expresión que es idéntica a la obtenida en el ejemplo 3.3.16.

3.4 METODOS VARIACIONALES Y TEORIA DE CONTROL.

El estudio de los métodos variacionales es central para el entendimiento de muchos problemas de la teoría de procesos óptimos y de la teoria de control.

Antes de entrar a la discusión y formalización de estas teorías, consideramos de importancia mencionar unos cuantos ejemplos especialmente seleccionados de la teoría de control, con la finalidad de entender algunos de los conceptos con los que esta teoría trata a manera de motivación.

Ejemplo. 3.4.1 (Bellman) Sea el estado de un sistema descrito por $f(t)$ en el tiempo t, $0\leq t\leq T$, y sea $u(t)$ un control del sistema. Supongamos que la ecuación del comportamiento del sistema está dada por $f'(t)=af(t)+u(t)$, a=cte, en donde

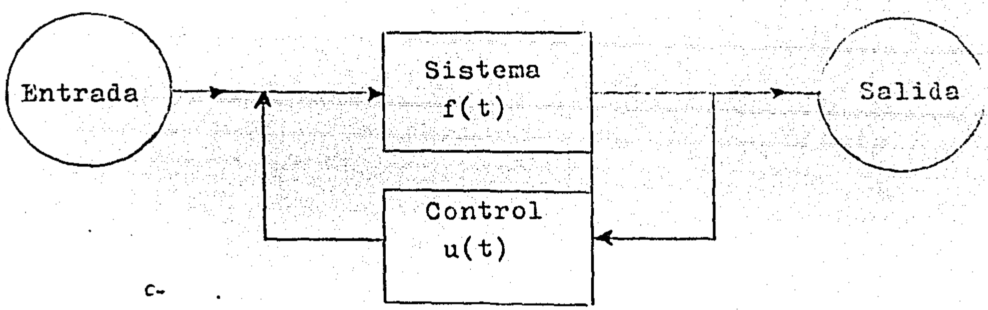
$f'(t)=\frac{d}{dt}f(t)$, y consideremos los siguientes casos de estudio

Caso 1. $f(0)=c$.

Caso 2. $f(0)=c_1$ y $f(T)=c_2$.

En el caso 1, tenemos restringido el estado inicial, mientras que para el caso 2. tanto el estado inicial como el final son restringidos
Supongamos también, que una medida de evaluación del sistema es dada por la funcional

$$J(f,u)=\int_0^T (f^2(t)+u^2(t))dt.$$



Consideremos ahora, el espacio de Banach formado por el espacio lineal $C^1[0,T]\times C[0,T]$, con la norma $\|(f,u)\|=\|f\|+\|u\|$, en donde $\|f\|$ es la norma usual de $C^1[0,T]$, $\|f\|=\max_{t\in[0,T]}|f(t)|+\max_{t\in[0,T]}|f'(t)|$ y $\|u\|$ es la norma usual de $C[0,T]$, $\|u\|=\max_{t\in[0,T]}|u(t)|$.
Nuestro objetivo es ahora, encontrar un control u(t) que minimice J(f,u). Supongamos que la clase de pares admisibles de funciones es un subconjunto abierto de $C^1[0,T]\times C[0,T]$. condición necesaria de extremo puede ser obtenida a partir de las ecuaciones de Euler-Lagrange correspondientes, si estas ecuaciones diferenciales tienen una solución única, y esta conduce a un mínimo, entonces tendremos resuelto completamente el problema.
Ahora bien, sustituyendo $u(t)=f'(t)-af(t)$ en J(f,u) tenemos

$$J(f,u)=\int_0^T(f^2(t)+(f'(t)-af(t))^2)dt=$$
$$\int_0^T((1+a^2)f^2(t)+f'^2(t))dt-2a\int_0^T f(t)f'(t)dt=$$
$$\int_0^T((1+a^2)f^2(t)+f'^2(t))dt+a(c_1^2-c_2^2),\text{(considerando el caso 2).}$$

En este caso tenemos la siguiente ecuación de Euler-Lagrange

$$f''(t)-(1+a^2)f(t)=0.$$

La solución general tiene la forma $f(t)=k_1e^{bt}+k_2e^{-bt}$, en donde $b=(1+a^2)^{\frac{1}{2}}$ y k_1, k_2 son determinadas en el caso 2 por $c_1=k_1+k_2$, $c_2=k_1e^{bT}+k_2e^{-bT}$, así que

$$\begin{bmatrix}c_1\\c_2\end{bmatrix}=\begin{bmatrix}1 & 1\\ e^{bT} & e^{-bT}\end{bmatrix}\begin{bmatrix}k_1\\k_2\end{bmatrix},\text{ dando }\begin{bmatrix}k_1\\k_2\end{bmatrix}=\begin{bmatrix}e^{-bT} & -1\\ -e^{bT} & 1\end{bmatrix}\begin{bmatrix}c_1\\c_2\end{bmatrix}.$$

y puesto que

$$\det\begin{bmatrix}e^{-bT} & -1\\ -e^{bT} & 1\end{bmatrix}=e^{-bT}-e^{bT}\neq 0 \text{ si } T\neq 0 \text{ , tenemos que la solución es}$$

única.

Ahora supongamos que damos un incremento f+h a la función f para la cual u es extremal de J, con h(0)=0, h(T)=0, en donde como se sabe f satisface la ecuación de Euler correspondiente, notamos

también que $(f+h)(0)=c_1$ y $(f+h)(T)=c_2$.
Consideremos $J(f+h,u)$ con el control extremal de $J(f,u)$, entonces

$$J(f+h,u)=\int_0^T (b^2(f(t)+h(t))^2+(f'(t)+h'(t))dt+a(c_1^2-c_2^2)=$$
$$\int_0^T (b^2f^2(t)+f'^2(t))dt+a(c_1^2-c_2^2)+\int_0^T (b^2h^2(t)+h'^2(t))dt+$$
$$2\int_0^T (b^2f(t)h(t)+f'(t)h'(t))dt$$
$$=J(f,u)+\int_0^T (b^2h^2(t)+h'^2(t))dt+2\int_0^T (b^2f(t)h(t)+f'(t)h'(t))dt,$$

pero

$$\int_0^T (b^2h^2(t)+h'^2(t))dt\leq b^2\max_{x\in[0,T]} h^2(t)+\max_{x\in[0,T]} h'^2(t)\leq b^2\max_{x\in[0,T]} h^2(t)+b^2\max_{x\in[0,T]} h'^2(t)$$

puesto que $b^2=(1+a^2)$, así

$$\int_0^T (b^2h^2(t)+h'^2(t))dt\leq b^2\|h^2\|\leq b^2\|h\|^2=o(\|(h,0)\|)=|o(\|(h,0)\|)|=0,$$

entonces $J(f+h,u)-J(f,u)=2\int_0^T (b^2f(t)h(t)+f'(t)h'(t))dt+o(\|(h,0)\|)$,

de esta manera la parte lineal de $J(f+h,u)-J(f,u)$ esta constituida por la integral de la derecha, por tanto la diferencial de Fréchet de la funcional J, $dJ(f,u)\in\mathcal{L}(V,\mathbb{R})$, en donde $V=C^1[0,T]\times C[0,T]$, es dada por $dJ(f,u)(h,0)=2\int_0^T (b^2f(t)h(t)+f'(t)h'(t))dt$, e integrando por partes tenemos

$$\int_0^T (b^2f(t)h(t)+f'(t)h'(t))dt=(f'(t)h(t))\Big|_0^T -\int_0^T (f''(t)-b^2f(t))h(t)dt,$$

y una condición necesaria de extremo es que $dJ(f,u)(h,0)=0$ para todo $(h,0)$ admisible, luego
$J(f+h,u)-J(f,u)=o(\|(h,0)\|)\geq 0$, y por tanto $J(f+h,u)-J(f,u)\geq 0$, así $J(f+h,u)\geq J(f,u)$ para todo $(h,0)$ admisible, y por lo tanto el control u para el cual J es extremo, minimiza a la funcional J, o bien, observamos que la segunda diferencial de Fréchet de la funcional J esta dada por $d^2J(f,u)((h,0),(h,0))=\int_0^T (b^2h^2(t)-h'(t))dt$, $d^2J(f,u)\in B(V\times V,\mathbb{R})\cong\mathcal{L}(V,\mathcal{L}(V,\mathbb{R}))$, en donde $V=C^1[0,T]\times C[0,T]$.
Así $J(f+h,u)-J(f,u)=dJ(f,u)(h,0)+d^2J(f,u)((h,0),(h,0))$, si f es la función para la cual u es extremal de la funcional J, tenemos que $d^2J(f,u)((h,0),(h,0))\geq 0$ para toda $(h,0)$ admisible es una condición necesaria de mínimo.
La igualdad en $J(f+h,u)\geq J(f,u)$ para el control u que minimiza a J, se obtiene para $h(t)=0$ para toda $t\in[0,T]$. De esta manera, hemos probado la existencia y unicidad de la solución, mostrando cual es ésta. El caso 1 y otras situaciones especiales pueden discutirse de manera semejante.

Planteamiento del Problema.

En muchos casos, encontrar el régimen operante óptimo para un sistema (digamos, un sistema físico), con un criterio conveniente de optimalidad, nos conduce al siguiente problema matemático.

Supongamos que el estado de un sistema es caracterizado por n números reales $f_1(t), f_2(t), \ldots, f_n(t)$, los cuales forman el vector $f(t)=(f_1(t),\ldots,f_n(t))\in\mathbb{R}^n$, y supongamos que el estado del sistema varía con el tiempo en la forma descrita por el sistema de ecuaciones diferenciales

$$\frac{df_k}{dt}=F_k(f_1(t),\ldots,f_n(t),u_1(t),\ldots,u_m(t)) \quad (k=1,\ldots,n),$$

en donde los números reales $u(t)=(u_1(t),\ldots,u_m(t))$ forman un vector perteneciente a una región de "control" fija $\Omega\subset R^m$, Ω un conjunto cerrado y conexo, y las funciones $F_k(f(t),u(t))\in C(0\times\Omega)$, O un subconjunto abierto de $\mathbb{R}^n$, y $u_j(t)\in PC[0,T]$ (conjunto de funciones continuas por pedazos) $(j=1,\ldots,m)$.

Ahora supongamos que la función vectorial de variable real $u(t)$, $0\leq t\leq T$, es llamada la función de control. El agregado $U(u(t),0,T,f(0))$ consistente de la función de control $u(t)$, el intervalo $[0,T]$ y $f(0)$ un valor inicial, será llamado un proceso de control, así todo proceso tiene asociado una trayectoria, esto es una solución de

$$\frac{df_k}{dt}=F_k(f_1(t),\ldots,f_n(t),u_1(t),\ldots,u_m(t)) \quad (k=1,\ldots,n),$$

Ahora, sea $F(f_1(t),\ldots,f_n(t),u_1(t),\ldots,u_m(t))$ una función la cual es definida, junto con sus derivadas parciales $\frac{\partial F}{\partial f_k}$ $(k=1,\ldots,n)$ para todo $(f(t),u(t))\in O\times\Omega$.

A todo proceso de control U, le asignamos el número $J(U)=\int_0^T F(f,u)dt$, esto es J es una funcional definida sobre el conjunto de procesos de control W. Entonces un proceso de control $U^*\in W$ es óptimo (mínimo) si la desigualdad $J(U^*)\leq J(U)$ se verifica para toda $U\in W$ tal que lleva el punto $f(0)$ a un punto $f(T)$, es decir, tal que la correspondiente trayectoria óptima $f^*(t)$ satisface $f^*(T)=f(T)$. Por una trayectoria óptima entenderemos, aquella trayectoria correspondiente al proceso de control óptimo.

Nuestro objetivo es encontrar condiciones necesarias para caracteri zar a los procesos de control óptimo y a las trayectorias óptimas, es importante puntualizar que en el estudio de tales procesos, estos son especificados de antemano($u\in\Omega$, Ω conocido).
Un caso especial importante del problema de control óptimo, corresponde a la situación en que J(U) se reduce a una integral de la forma $J(U)=\int_0^T dt$, la cual representa el tiempo de ir de un punto f(0) al punto f(T). En este caso optimalidad significa tomar el tiempo mínimo de ir de f(0) a f(T).

Relación con los Métodos Variacionales Clásicos.
El problema de control óptimo está íntimamente ligado a ciertos problemas tradicionales de los métodos variacionales. De hecho la integral $\int_0^T F(f,u)dt=\int_0^T F(f_1,\ldots,f_n,u_1,\ldots,u_m)dt$ puede se vista como una funcional dependiente de n+m funciones, es decir, una funcional definida sobre alguna clase de curvas en R^{n+m+1}. Y puesto que las funciones $f_1,\ldots,f_n,u_1,\ldots,u_m$ están conectadas por medio de las ecuaciones

$$\frac{df_k}{dt}=F_k(f_1,\ldots,f_n,u_1,\ldots,u_m)\ (k=1,\ldots,n),$$

y ya que las condiciones de frontera son equivalentes a los requerimientos de que la trayectoria óptima deseada comience en el punto f(0) y termine en el punto f(T), los punto finales de las curvas admisibles estan en el espacio R^{n+m+1}, estos han de estar sobre dos hiper planos de dimensión m+1, determinados por los valores fijos $(f_1(0),f_2(0),\ldots,f_n(0))$ y $(f_1(T),f_2(T),\ldots,f_n(T))$.

De esta manera vemos que el problema de control óptimo es una variante de el problema de encontrar un mínimo sujeto a condiciones subsidiarias. El problema de control óptimo tiene una característica especial que consiste en especificar de antemano una clase definida de procesos de control admisibles, en donde las funciones $u_1(t),\ldots,$ $u_n(t)$ toman valores en una región fija $\Omega\subset R^n$, pero en general no se requiere que sean continuas.

Se puede mostrar que el problema variacional en donde se consideran funcionales de la forma $J(f_1,\ldots,f_n)=\int_0^T F(f_1,\ldots,f_n,f_1',\ldots,f_n')dt$ en donde el integrando no depende explícitamente de t, es un caso particular del problema de control óptimo. Para ver esto supongamos que de entre todas las curvas que pasan a través de dos puntos fijos $(f_1(0),\ldots,f_n(0))$ y $(f_1(T),\ldots,f_n(T))$, deseamos encontrar una curva $f(t)=(f_1(t),\ldots,f_n(t))$ para la cual la funcional $J(f_1,\ldots,f_n)=\int_0^T F(f_1,\ldots,f_n,f_1',\ldots,f_n')dt$ alcance un mínimo, entonces basta escribir $F(f_1,\ldots,f_n,u_1,\ldots,u_n)$, en donde $\frac{df_k}{dt}=u_k$ $(k=1,\ldots,n)$

Condiciones necesarias para optimalidad.

Para encontrar condiciones necesarias de optimalidad para un proceso de control y su correspondiente trayectoria, complementamos el sistema de ecuaciones diferenciales $\frac{df_k}{dt}=F_k(f,u)$ $(k=1,\ldots,n)$ con la ecuación diferencial extra $\frac{dg}{dt}=F(f,u)$. Al mismo tiempo complementamos las condiciones iniciales $f_k(0)=c_k$, c_k valores dados. $(k=1,\ldots,n)$ con la condición extra $g(0)=0$.

Por conveniencia, introducimos la función vectorial $(n+1)$-dimensional de variable real $\bar{f}(t)=(g(t),f_1(t),\ldots,f_n(t))$.

Es claro que si U es un proceso de control admisible y si $\bar{f}=\bar{f}(t)$ es solución del sistema $\frac{df_k}{dt}=F_k(f,u)$ $(k=1,\ldots,n)$, $\frac{dg}{dt}=F(f,u)$ correspondiente al proceso de control U y a las condiciones iniciales $f_k(0)=c_k$ $(k=1,\ldots,n)$, $g(0)=0$, entonces

$J(U)=\int_0^T F(f,u)dt=g(T)-g(0)=g(T)$.

Así, el problema de control óptimo puede ser establecido como sigue: Deseamos encontrar el proceso de control admisible U para el cual la solución $\bar{f}(t)$ del sistema $\frac{df_k}{dt}=F_k(f,u)$ $(k=1,\ldots,n)$, $\frac{dg}{dt}=F(f,u)$, la cual también satisface las condiciones iniciales $f_k(0)=c_k$ $g(0)=0$, tenga el mínimo valor posible de $g(T)$.

Ahora, además de las variables $g,f_1,\ldots,f_n$, introducimos nuevas variables $v_g,v_1,\ldots,v_n$, las cuales satisfacen el siguiente sistema de ecuaciones diferenciales

$$\frac{dv_i}{dt}=-\sum_{k=1}^{n} \frac{F_k(f,u)}{\partial f_i} v_k-\frac{\partial F(f,u)}{\partial f_i}v_g, (i=1,\ldots,n),$$

$$\frac{dv_g}{dt}=-\sum_{k=1}^{n} \frac{F_k(f,u)}{\partial g} v_k-\frac{\partial F(f,u)}{\partial g}v_g \ .$$

La introducción del sistema $v_g, v_1, \ldots, v_n$ tiene la siguiente interpretación geométrica: Consideremos el espacio vectorial V' de puntos $\bar{f}=(g,f_1,\ldots,f_n)$. Sea T una funcional lineal acotada definida sobre este espacio. Entonces por el Teorema de Representación de Riesz, existe un único punto $v=(v_g,v_1,\ldots,v_n)$, tal que $T(\bar{f})=(v,\bar{f})$ para toda $\bar{f}$ en V', en donde $(\cdot,\cdot)$ denota el producto escalar usual, esto es $(v,\bar{f})=\sum_{k=1}^{n} v_k f_k + v_g g$. Inversamente cualquier producto escalar define una funcional lineal sobre V', consideremos el conjunto de todos los vectores v, denotado por V", el cual es isomorfo al espacio dual $(V')^*=\mathcal{L}(V',\mathbb{R})$, el espacio V" es frecuentemente llamado el espacio conjugado a V', si definimos $\|v\|=\sup_{\bar{f}\neq 0}\frac{|(v,\bar{f})|}{\|\bar{f}\|}$ tendremos entonces un isomorfismo isométrico de V" en $(V')^*$, en efecto la correspondencia $v \mapsto T_v$ es lineal, puesto que si $a\in\mathbb{R}$, entonces

$$T_{v_1}(\bar{f})+aT_{v_2}(\bar{f})=(v_1,\bar{f})+(av_2,\bar{f})=(v_1+av_2,\bar{f})=T_{v_1+av_2}(\bar{f}) \text{ para toda } \bar{f}\in V',$$

además

$$\|T_v\|=\sup_{\|\bar{f}\|\leq 1}|T_v(\bar{f})| \geq |T_v(v/\|v\|)| = (v,v/\|v\|) = \|v\|, \text{ y}$$

$$\|T_v\|=\sup_{\|\bar{f}\|\leq 1}|T_v(\bar{f})| = \sup_{\|\bar{f}\|\leq 1}|(v,\bar{f})| \leq \sup_{\|\bar{f}\|\leq 1}\|v\|\|\bar{f}\|=\|v\|, \text{ por tanto } \|T_v\|=\|v\|.$$

Es usual llamar al vector $(v_g,v_1,\ldots,v_n)$, el vector conjugado a $(g,f_1,\ldots,f_n)$. En el espacio de vectores $(v_g,v_1,\ldots,v_n)$ conjugado a el espacio de vectores $(g,f_1,\ldots,f_n)$, consideramos el hiperplano $\sum_{k=1}^{n} v_k f_k + v_g g = C = \text{cte.}$ que pasa a través del punto $(0,f_1(0),\ldots,f_n(0)) =(0,c_1,\ldots,c_n)$. entonces el sistema de ecuaciones diferenciales conjugadas

$$\frac{dv_i}{dt}=-\sum_{k=1}^{n}\frac{\partial F_k(f,u)}{\partial f_i}v_k - \frac{\partial F(f,u)}{\partial f_i}v_g \quad (i=1,\ldots,n),$$

$$\frac{dv_g}{dt}=-\sum_{k=1}^{n}\frac{\partial F_k(f,u)}{\partial g}v_k - \frac{\partial F(f,u)}{\partial g}v_g ,$$

describe la traslación de este hiperplano a través de las trayectorias correspondientes a las soluciones del sistema de ecuaciones diferenciales $\frac{df_k}{dt}=F_k(f,u)$ $(k=1,\ldots,n)$, $\frac{dg}{dt}=F(u,v)$.

Para dilucidar más sobre esta interpretación, consideremos el siguiente caso particular, en el cual tenemos el sistema de ecuaciones diferenciales $\frac{df}{dt}=G(f,u)$ y $\frac{dg}{dt}=F(f,u)$, con las condiciones iniciales $g(0)=0$, $f(0)=0$. Sean v_g y v las variables conjugadas a g y f, y consideremos el hiperplano $vf+v_g g=C$ que pasa a través del punto $(g(0),f(0))=(0,c)$, derivando con respecto a t tenemos que

$$\frac{dv}{dt}f+v\frac{df}{dt}+\frac{dv_g}{dt}g+v_g\frac{dg}{dt}=\frac{dv}{dt}f+vG+\frac{dv_g}{dt}g+v_gF=0$$

luego derivando parcialmente primero con respecto a f, y luego con respecto a g obtenemos

$\frac{dv}{dt}+vG_f+v_gF_f=0$ y $\frac{dv_g}{dt}+vG_g+v_gF_g=0$, o equivalentemente

$$\frac{dv}{dt}=-vG_f-v_gF_f \text{ y } \frac{dv_g}{dt}=-vG_g-v_gF_g$$

y este es el sistema de ecuaciones diferenciales conjugadas.

Sea $v(t)=(v_g(t),v_1(t),\ldots,v_n(t))$ definimos la siguiente función $H'(v,f,u)=\sum_{k=1}^{n} v_kF_k+v_gF$, entonces podemos escribir las ecuaciones $\frac{df_k}{dt}=F_k(f,u)$ $(k=1,\ldots,n)$, $\frac{dg}{dt}=F(f,u)$ y $\frac{dv_k}{dt}=-\sum_{k=1}^{n}\frac{\partial F_k(f,u)}{\partial f_i}v_k-\frac{\partial F(f,u)}{\partial f_i}v_g$ $(i=1,\ldots,n)$, $\frac{dv_g}{dt}=-\sum_{k=1}^{n}\frac{\partial F_k(f,u)}{\partial g}v_k-\frac{\partial F(f,u)}{\partial g}v_g$, en la forma

$$\frac{df}{dt}=\frac{\partial H'}{\partial v_k}\ (k=1,\ldots,n),\ \frac{dg}{dt}=\frac{\partial H'}{\partial v_g}\ \text{ y } \frac{dv}{dt}=-\frac{\partial H'}{\partial f_k},\ (k=1,\ldots,n),\ \frac{dv_g}{dt}=-\frac{\partial H'}{\partial g}.$$

Estas ecuaciones presentan cierta similitud, en el uso de variables canónicas para el sistema de ecuaciones de Euler, correspondientes al problema variacional en donde tratamos funcionales de la forma $J(f_1,\ldots,f_n)=\int_a^b F(x,f_1,\ldots,f_n,f_1',\ldots,f_n')dx$. Sin embargo, estas tienen un significado diferente, puesto que para el caso en donde $J(f_1,\ldots,f_n)=\int_a^b F(x,f_1,\ldots,f_n,f_1',\ldots,f_n')dx$ el número de ecuaciones canónicas es igual al número de funciones desconocidas, mientras que en el problema de control aparecen además funciones desconocidas $u_1,\ldots,u_n$. De hecho para emplear variables canónicas en las ecuaciones correspondientes al problema de control, debemos hablar de la función $H(v,f)=\sup_{u\in\Omega} H'(v,f,u)$.

El Principio del Mínimo de Pontryagin.

El Principio del mínimo proporciona un conjunto de condiciones necesarias para problemas de control, en las cuales el control $u(t)$ es restringido a pertenecer a un conjunto dado. En esta sección desarrollaremos el principio de mínimo desde un punto de vista abstracto empleando el concepto de operador adjunto.

Sean V_1 y V_2 dos espacios lineales normados, y $J(f,u)$ una funcional (eficiencia) definida sobre $V_1 \times V_2$, y consideremos una restricción de la forma $K(f,u)=0$, donde K es un mapeo de $V_1 \times V_2$ en V_1. La transformación K describe un sistema de ecuaciones, y puede representar un conjunto de ecuaciones, ya sean diferenciales, integrales, de diferencias, etc.

Supongamos que $K(f,u)=0$ define una única función implícita $f=f(u)$, además supongamos que J y K son diferenciables en el sentido de Fréchet con respecto al argumento f, y que las derivadas parciales K_f, y J_f son continuas sobre $V_1 \times V_2$. Finalmente suponemos que la función implícita $f(u)$ satisface la condición de Lipschitz

$\|f(u)-f(v)\| \leq M\|u-v\|$.

El problema de control consiste en encontrar (f,u) que minimicen $J(f,u)$ de tal manera que se satisfagan las restricciones $K(f,u)=0$ y $u \in \Omega$, en donde Ω es un conjunto fijo de V_2. Puesto que f es determinada en forma única a partir de u, es bien entendido que $(J(f,u))(u)=J(f(u),u)$ o si se prefiere que $J(u)=J(f(u),u)$.

Ahora consideremos la funcional (el Lagrangiano) de nuestro problema de optimización bajo restricciónes

$L(f,u,p^*)=p^*K(f,u)+J(f,u)$, $f \in V_1$, $u \in V_2$,

en donde p^* es el multiplicador de Lagrange asociado con la restricción del problema de optimización.

Proposición. 3.4.2 Para cualquier $u \in \Omega$ sea p^* una solución de la ecuación

$p^*K_f(f(u),u)+J_f(f(u),u)=0$.

Entonces para toda $v \in \Omega$,

$J(u)-J(v)=L(f(u),u,p^*)-L(f(u),v,p^*)+o(\|u-v\|)$.

Demostración: Por definición

$$J(u)-J(v)=J(f(u),u)-J(f(v),v)$$

$$=J(f(u),u)-J(f(u),v)+J(f(u),v)-J(f(v),v)$$

$$=J(f(u),u)-J(f(u),v)+J_f(f(v),v)(f(u)-f(v))+o(\|f(u)-f(v)\|)$$

$$=J(f(u),u)-J(f(u),v)+J_f(f(u),u)(f(u)-f(v))$$

$$+(J_f(f(v),v)-J_f(f(u),u))(f(u)-f(v))+o(\|f(u)-f(v)\|)$$

Observamos que

$$|(J_f(f(v),v)-J_f(f(u),u))(f(u)-f(v))|/\|u-v\|$$

$$\leq |J_f(f(v),v)-J_f(f(u),u)|\ \frac{\|f(u)-f(v)\|}{\|u-v\|}$$

$$\leq |J_f(f(v),v)-J_f(f(u),u)|\ M \to 0 \text{ cuando } \|(f(v),v)-(f(u),u)\| \to 0$$

debido a la continuidad de J_f. Pero cuando $\|v-u\| \to 0$ se tiene que $\|(f(v),v)-(f(u),u)\| \to 0$, en efecto, ya que

$$\|(f(v),v)-(f(u),u)\| = \max_{0\le t\le T} \|f(v)-f(u)\| + \max_{0\le t\le T} \|v-u\| \leq M\|v-u\| + \max_{0\le t\le T} \|v-u\|,$$

de donde

$$J_f(f(v),v)-J_f(f(u),u))(f(u)-f(v))=o(\|v-u\|) .$$

Además, al considerar $o(\|f(u)-f(v)\|)$, tenemos

$$\frac{\|o(\|f(u)-f(v)\|)\|}{M\|u-v\|} \leq \frac{\|o(\|f(u)-f(v)\|)\|}{\|f(u)-f(v)\|} \to 0 \text{ cuando } \|f(u)-f(v)\| \to 0,$$

pero cuando $\|u-v\| \to 0$, se tiene que $\|f(u)-f(v)\| \to 0$ por la condición de Lipschitz, por lo tanto

$$\frac{\|o(\|f(u)-f(v)\|)\|}{M\|u-v\|} \to 0 \text{ cuando } M\|u-v\| \to 0, \text{ entonces}$$

$$o(\|f(u)-f(v)\|)=o(M\|u-v\|)=o(\|u-v\|).$$

Por consiguiente

$$J(u)-J(v)=J(f(u),u)-J(f(u),v)+J_f(f(u),u)(f(u)-f(v))+o(\|u-v\|).$$

De la misma forma

$$0=K(f(u),u)-K(f(u),v)+K_f(f(u),u)(f(u)-f(v))+o(\|u-v\|).$$

Después de multiplicar ésta última expresión por p^*, y sumarla a la expresión para $J(u)-J(v)$, al aplicar la hipótesis

$p^*K_f(f(u),u)+J_f(f(u),u)=0$, en consecuencia

$$J(u)-J(v)=L(f(u),u,p^*)-L(f(u),v,p^*)+o(\|u-v\|).$$

La importancia de la proposición anterior, radica en lo siguiente: Proporciona una forma de determinar un cambio en J debido a un cambio en u, sin reevaluar la función implícita $f=f(u)$.

Lema. 3.4.3 (Bellman-Gronwall). Sea $c\geq 0$ una constante. Sean $f(t)$ y $g(t)$ elementos de $PC[0,T]$, supongamos que $g(t)\geq 0$ para toda $t\in [0,T]$. Si

$f(t)\leq c+\int_0^t f(s)g(s)ds$ para toda $t\in[0,T]$, entonces

$f(t)\leq c\ \exp(\int_0^t g(s)ds)$ para toda $t\in[0,T]$.

Demostración: Denotemos por $F(t)=c+\int_0^t f(s)g(s)ds$, entonces

$f(t)\leq F(t)$,

y dado que $g(t)\exp(-\int_0^t g(s)ds)\geq 0$, tenemos

$f(t)g(t)\exp(-\int_0^t g(s)ds)\leq F(t)g(t)\exp(-\int_0^t g(s)ds)$.

Observamos que

$F'(t)=f(t)g(t)$ para las t's en donde f y g son continuas,

además

$$\frac{d}{dt}(F(t)\exp(-\int_0^t g(s)ds))=F'(t)\exp(-\int_0^t g(s)ds)-F(t)g(t)\exp(-\int_0^t g(s)ds)$$
$$=f(t)g(t)\exp(-\int_0^t g(s)ds)-F(t)g(t)\exp(-\int_0^t g(s)ds)\leq 0,$$

por lo tanto,

$\frac{d}{dt}(F(t)\exp(-\int_0^t g(s)ds))\leq 0$ casi dondequiera.

De donde $F(t)\exp(-\int_0^t g(s)ds$ es una función decreciente, en consecuencia

$F(t)\exp(-\int_0^t g(s)ds)\leq F(0)=c$ para toda $t\in[0,T]$, es decir

$f(t)\leq F(t)\leq c\ \exp(\int_0^t g(s)ds)$ para toda $t\in[0,T]$.

Lo que concluye la demostración del lema.

Lema. 3.4.4 Sean f_o y u_o óptimos para el problema minimizar $J(f,u)=\int_0^T F(f,f',u)dt$

sujeto a

$f'(t)=G(f,u)$, $f(0)=c_1$, $f(T)=c_2$, $u(t)\in\Omega\subset PC[0,T]$.

en donde $f\in C^1[0,T]$.

Supongamos que

$F_{f'f'}(f,f',u)\neq 0$ y $f=f(u)$

y definamos el Hamiltoniano

$H(f,u,p)= F(f,f',u)+pf'$.

Entonces la ecuación de Euler-Lagrange asociada al problema anterior en la forma canónica corresponde a las ecuaciones

$H_f=-\frac{dp}{dt}$, $H_p=\frac{df}{dt}$.

Demostración: Si la trayectoria $f_o=f_o(u_o)$ es óptima, entonces sobre ella se satisface que

$F_f(f,f',u)-\frac{d}{dt}F_{f'}(f,f',u)=0.$

Ahora si f' es vista como una nueva función $\frac{df}{dt}=f'$ y si definimos $-p=F_{f'}(f,f',u)$ tenemos que

$-\frac{dp}{df'}=F_{f'f'}(f,f',u)\neq 0$, entonces localmente $f'=f'(f,u,p)$, luego si las funciones f y p son vistas como variables independientes y f' como una función de f y p, la ecuacion de Euler-Lagrange se transforma en las ecuaciones

$\frac{df}{dt}=f'$, $-p=F_{f'}$, $F_f+\frac{dp}{dt}=0$.

Ahora bien si introducimos la función

$H(f,u,p)= F(f,f',u)+pG(f,u)$.

Las nuevas variables f,u,p y H son conocidas como variables canónicas, luego al calcular la diferencial de H encontramos que

$$\begin{aligned}
dH&=F_f df+F_{f'}df'+G(f,u)dp+pdG(f,u)\\
&=F_f df-pdf'+G(f,u)dp+pdf'\\
&=F_f df+G(f,u)dp,
\end{aligned}$$

de donde

$H_f=F_f$ y $H_p=G(f,u)=f'$

en otras palabras las cantidades F_f y f' están relacionadas por las derivadas parciales de H.
Finalmente, las ecuaciones de Euler-Lagrange asociadas al problema planteado se convierten en

$$H_p=\frac{df}{dt}\ ,\ H_f=-\frac{dp}{dt}\ ,$$

lo que se deseaba demostrar.

Teorema. 3.4.5 (Principio de Pontryagin). Sean f_0 y u_0 óptimos para el problema

minimizar $J(f,u)=\int_0^T F(f,u)dt$,

sujeto a $f'(t)=G(f,u)$, $f(0)=c$, $u(t)\in\Omega\subset PC[0,T]$,

en donde $f\in C^1[0,T]$ y la función F satisface la condición de Lipschitz

$$|F(f,u)-F(h,v)|\leq M(|f-h|+|u-v|).$$

F y G con derivadas parciales continuas de Fréchet con respecto del argumento f.

Sea p la solución de la ecuación

$$-p'(t)=F_f+p(t)G_f,\quad p(0)=0,\quad p(T)=0$$

en donde las derivadas parciales son evaluadas a lo largo de la trayectoria óptima. Definamos ahora el Hamiltoniano

$$H(f,u,p,t)=F(f,u)+p(t)G(f,u).$$

Entonces para todo $t\in[0,T]$,

$$H(f_0(t),u_0(t),p(t))\leq H(f_0(t),u(t),p(t))\text{ para todo }u\in\Omega.$$

Demostración: Tomemos $V_1=C^1[0,T]$ con la norma $\|f\|=\max_{t\in[0,T]}|f(t)|+\max_{t\in[0,T]}|f'(t)|$ y $V_2=PC[0,T]$ con la norma $\|u\|=\max_{t\in[0,T]}|u(t)|$.
Definamos

$$K(f,u)=f(t)-f(0)-\int_0^t G(f(s),u(s))ds,$$

$$J(f,u)=\int_0^T F(f,u)dt.$$

Si f y f+h corresponden a u y u+v, respectivamente en $\{(f,u);K(f,u)=0\}$ y si $h(0)=0$, entonces

$$0=(f(t)-f(0))+h(t)-\int_0^t F(f+h,u+v)ds$$

$$=\int_0^t F(f,u)ds+h(t)-\int_0^t F(f+h,u+v)ds,$$

así

$h(t)=\int_0^t (F(f+h,u+v)-F(f,u))ds,$

luego al aplicar el Lema 3.4.3 tenemos

$|h(t)|\leq\int_0^t |F(f+h,u+v)-F(f,u)|ds\leq\int_0^t M(|h(s)|+|v(s)|)ds\leq TMe^{tM}\|v\|$

para toda $t\in[0,T]$, por tanto al tomar $K\geq TMe^{TM}$, se verifica que $|h(t)|\leq TMe^{tM}\leq TMe^{TM}\leq K$ para toda $t\in[0,T]$, entonces $\|h\|\leq K\|v\|$.

En consecuencia la transformación $K(f,u)$ satisface la condición de Lipschitz requerida en la proposición 3.4.3 que más adelante será usada, si $f_0=f_0(u_0)$ es óptimo (trayectoria óptima), entonces por la definición 3.3.24 y por los resultados que de ella se desprenden, tenemos que

$J_f+pK_f=(F_f-\frac{d}{dt}F_{f'}+pG_f-\frac{d}{dt}G_{f'})=0$ sobre la trayectoria óptima,

pero $G=f'$, luego $G_{f'}=1$, entonces $\frac{d}{dt}G_{f'}=0$, así

$J_f+pK_f=(F_f-\frac{d}{dt}F_{f'}+pG_f)=0,$

introduciendo la variable canónica $-p=F_{f'}$, tenemos que

$-p'(t)=F_f+p(t)G_f$, sobre la trayectoria óptima,

También por el Lema 3.3.4 , si consideramos la ecuación de Euler-Lagrange en la forma canónica asociada al problema planteado, con variables canónicas f,p y H tenemos que

$-p'(t)=H_f=F_f+p(t)G_f$ es válida sobre la trayectoria óptima

Por otra parte la funcional $\int_0^T H(f,u,p)dt$ satisface

$\int_0^T H(f,u,p)dt=L(f,u,p)-\int_0^T f(t)p'(t)dt$, en efecto

$$\begin{aligned}\int_0^T H(f,u,p)dt&=\int_0^T p(t)G(f,u)dt+\int_0^T F(f,u)dt\\&=p(t)\int_0^t G(f(s),u(s))ds\Big|_0^T-\int_0^T\Big(\int_0^t G(f(s),u(s))ds\Big)p'(t)dt\\&\quad+J(f,u)\\&=-\int_0^T\Big(\int_0^t G(f(s),u(s))ds\Big)p'(t)dt+J(f,u)\end{aligned}$$

luego

$$\int_0^T H(f,u,p,t)dt=\int_0^T((f(0)-f(t))+(f(t)-f(0)-\int_0^T G(f(s),u(s))ds))\cdot$$
$$\cdot p'(t)dt+J(f,u)$$
$$=\int_0^T(f(0)-f(t))p'(t)dt+\int_0^T K(f,u)p'(t)dt+J(f,u)$$
$$=f(0)\int_0^T p'(t)dt+\int_0^T f(t)p'(t)dt+p(t)K(f,u)$$
$$-\int_0^T p(t)\frac{d}{dt}K(f,u)dt+J(f,u)$$
$$=\int_0^T f(t)p'(t)dt+p(t)K(f,u)+J(f,u)$$

puesto que $p(0)=p(T)=0$ y

$\frac{d}{dt}K(f,u)=f'(t)-G(f,u)=0$ para toda $t\in[0,T]$,

pero $\int_0^T f(t)p'(t)dt$ no depende explícitamente de u, entonces por la proposición 3.4.2 , tenemos que

$$J(u_0)-J(u)=L(f_0,u_0,p)-L(f,u,p)+o(\|u-u_0\|)$$
$$=\int_0^T H(f_0,u_0,p,t)dt-\int_0^T H(f,u,p,t)dt+o(\|u-u_0\|).$$

Para probar que

$H(f_0(t),u_0(t),p(t))\leq H(f_0(t),u(t),p(t))$ para todo $t\in[0,T]$ y para toda $u\in\Omega$. Supongamos que por el contrario existen $t_0\in[0,T]$ y $\bar{u}\in\Omega$ tales que

$H(f(t_0),u_0(t_0),p(t_0))>H(f(t_0),\bar{u}(t_0),p(t_0))$,

puesto que $u\in PC[0,T]$, $f,p\in C^1[0,T]$ y $F,G\in C[0,T]$, tenemos que $H\in PC[0,T]$, entonces existe un subintervalo $[t_1,t_2]\subset[0,T]$ que contiene a t_0 y $\epsilon>0$ tales que

$H(f(t),u_0(t),p(t))-H(f(t),\bar{u}(t),p(t))>\epsilon$ para toda $t\in[t_1,t_2]$.

Sea ahora, $v(t)\in PC[0,T]$ definida por

$$v(t)=\begin{cases} u_0(t) & t\in[0,T]-[t_1,t_2], \\ \bar{u}(t) & t\in[t_1,t_2], \end{cases}$$

entonces

$$J(u_0)-J(v)=\int_0^T(H(f(t),u_0(t),p(t))-H(f(t),v(t),p(t)))dt+o(\|u-u_0\|)$$
$$>\int_{t_1}^{t_2}\epsilon\, dt+o(\|u-u_0\|)=\epsilon(t_2-t_1)+\max_{t_1\leq t\leq t_2}|u(t)-u_0(t)|=\epsilon(t_2-t_1)+o(t_2-t_1)$$

si $t_1\to t_2$, entonces $J(u_0)>J(v)$, lo que contradice la optimalidad de u_0.

Discusión sobre el principio de Pontriagin.

En principio, se había planteado el problema de minimizar $J(U)=\int_0^T F(f,u)dt$, $U=\{u(t),0,T,f(0)=c\}$ sujeto a

$f(0)=c$, $f'(t)=G(f,u)$, $u(t)\in\Omega$, $f\in C^1[0,T]$,

y F,G con derivadas parciales continuas con respecto del argumento f.

Después la función $g\in C^1[0,T]$ tal que

$g'(t)=F(f,u)$, $g(0)=0$ y $J(U)=g(T)$,

fue introducida, asimismo las funciones conjugadas v_f y v_g fuerón definidas mediante las relaciones

$$\frac{d}{dt}(v_g)=-F_g v_g - G_g v_f,$$

$$\frac{d}{dt}(v_f)=-F_f v_g - G_f v_f$$

que fueron obtenidas mediante la traslación del hiperplano

$$g v_g + f v_f = cte$$

a través de las soluciones de

$f'(t)=G(f,u)$ y $g'(t)=F(f,u)$.

Si definimos el Hamiltoniano

$H(v_g,v_f,f,u)=v_g F+v_f G$,

para el cual

$H_{v_g}=F=g'(t)$, $H_{v_f}=G=f'(t)$,

$H_g=F_g v_g+G_g v_f=-\frac{d}{dt}(v_g)$,

$H_f=F_f v_g+G_f v_f=-\frac{d}{dt}(v_f)$,

entonces si $v_g=1$ y $v_f=p(t)$ para toda $t\in[0,T]$, tenemos

$$-p'(t)=F_f+p(t)G_f$$

ecuación que corresponde a la condición pedida en el Teorema 3.4.4.

Ejemplo. 3.4.6 (Problema de asignación del granjero).
Un granjero produce un sólo cultivo, digamos, trigo. Después de la cosecha debe almacenarlo o venderlo y reinvertirlo comprando tierra adicional y equipo para incrementar su producción. El granjero desea maximizar la cantidad total almacenada al tiempo T. De acuerdo con el planteamiento de este problema en el ejemplo 1.1.4. Si u(t) es la tasa de producción que es reinvertida al tiempo t. El problema es entonces descrito por:

maximizar $J(f,u)=\int_0^T(1-u(t))f(t)dt$

sujeto a

$\frac{d}{dt}f(t)=u(t)f(t)$, $0\leq u(t)\leq 1$, $f(t)\geq 0$, $f(0)>0$,

en donde f(t) es la producción al tiempo t.
Usando el Teorema 3.4.4 tenemos que

$F(f,u)=(1-u(t))f(t)$, $G(f,u)=u(t)f(t)$.

Una solución óptima f_o,u_o y p debe satisfacer

$-p'(t)=F_f+p(t)G_f$ equivalentemente que

$-p'(t)=u_o(t)(p(t)-1)+1$

y debe ser tal que maximice el Hamiltoniano

$H(f,u,p)=F+p(t)G$ que equivale a

$H(f,u,p)=f_o(t)u_o(t)(p(t)-1)+f_o(t)$

puesto que $f_o(t)\geq 0$ para toda $t\in[0,T]$ debemos tener que

$$u_o(t)=\begin{cases}1 \text{ si } p(t)\geq 1,\\ 0 \text{ si } p(t)<1,\end{cases}$$

puesto que $p(T)=0$, se tiene que $u_o(T)=0$, además integrando la expresión

$-p'(t)=u_o(t)(p(t)-1)+1$ sobre el intervalo $[T-t,T]$, $t>0$,

vemos que

$$-\int_{T-t}^{T}p'(t)dt=\int_{T-t}^{T}(u_o(t)(p(t)-1)+1)dt,$$

lo cual implica que

$p(T-t)=\int_{T-t}^{T} u_o(t)(p(t)-1)dt+t$

si $t=1$, entonces

$p(T-1)=\int_{T-1}^{T} u_o(t)(p(t)-1)dt+1\geq 1$

entonces $u_o(T-1)=1$. Veamos que el control óptimo $u_o(t)$ queda determinado por

$$u_o(t)=\begin{cases}1 \text{ si } 0\leq t\leq T-1,\\ 0 \text{ si } T-1<t\leq T.\end{cases}$$

En efecto, si $u_o(t)=0$ en el intervalo $[T-1,T]$, entonces $p(T-1)=1$ y

$-\int_{T-1}^{T}p'(t)dt=\int_{T-1}^{T}dt,$

de donde

$p'(t)=-1$ $(p(t)=-t+T)$ en el intervalo $[T-1,T]$, esto es $p(t)$ es una función no creciente sobre $[T-1,T]$, así

$p(t)<p(T-1)=1$ para toda $t\in(T-1,T]$.

Si $u_o(t)=1$ en el intervalo $[0,T-1]$, entonces

$-\int_{0}^{T-1}p'(t)dt=\int_{0}^{T-1}p(t)dt,$

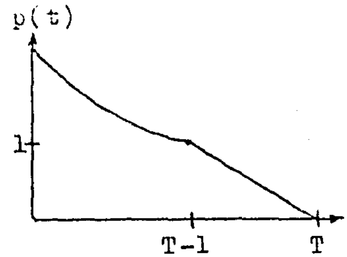
de donde $-p'(t)=p(t)$ sobre el intervalo $[0,T-1]$, si se ha de satisfacer que $p(T-1)=1$, entonces

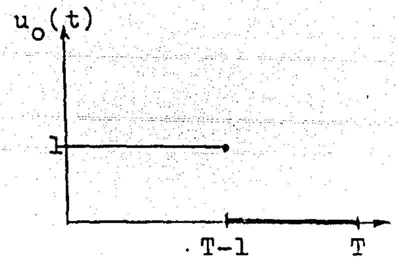
$p(t)=e^{T-1}e^{-t}$ en el intervalo $[0,T-1]$ y dado que $p(t)$ es una función no creciente sobre el intervalo $[0,T-1]$,

$p(t)>p(T-1)=1$ para toda $t\in[0,T-1)$. Por lo tanto

$$p(t)=\begin{cases}e^{T-1}e^{-t} & \text{si } 0\leq t\leq T-1,\\ -t+T & \text{si } T-1<t\leq T.\end{cases}$$

En consecuencia, el granjero no debe almacenar reinvirtiendo toda la producción de 0 a T-1 y almacenar toda la producción de T-1 hasta T.





3.5 METODOS VARIACIONALES Y PROGRAMACION DINAMICA.

En la sección anterior se discutieron algunos ejemplos de la teoría de control y se indicaron algunas técnicas para resolverlos. La teoría de la programación dinámica ha probado su eficiencia en la resolución de una extensa clase de problemas de la teoría de control en situaciones más o menos complicadas.
Supongamos que un sistema dinámico es estudiado a través de varios estados y que el sistema es caracterizado en cualquier tiempo t por ciertas variables de estado. En cada estado es posible escoger un número de decisiones que transforman a las variables de estado, en donde además suponemos que el pasado del sistema no influye en la determinación de acciones futuras, el estudio de tales procesos con técnicas proporcionadas por la programación dinámica será nuestro objetivo. Los procesos son estudiados hasta que se optimiza cierta función (objetivo) de las variables de estado, tal función representa un criterio de evaluación del sistema, el problema es entonces encontrar una secuencia de decisiones óptimas sobre los estados, tal secuencia de decisiones es llamada una política óptima. Para la solución de tales problemas es necesario el principio de optimalidad, establecido por Bellman, este principio consiste en lo siguiente: Una política óptima tiene la propiedad de que cualesquiera que sean el estado y la decisión iniciales, las decisiones restantes deben constituir una política óptima.
Antes de entrar a la formalización de la teoría de la programación dinámica, consideramos de gran utilidad hacer mención a algún problema que surja de una aplicación de la teoría de control, en donde sea posible aproximar una solución por medio de técnicas de la programación dinámica, así como el de familiarizarse con algunos de los conceptos que esta teoría maneja.

En el ejemplo 3.4.1 de la sección anterior fue planteado el siguiente problema:
Consideremos un sistema descrito por la función de estados $f(t)$ en el tiempo t, $0 \leq t \leq T$, y sea $u(t)$ un control del sistema. Supon-

gamos que la ecuación que describe la evolución del sistema está dada por $f'(t)=af(t)+u(t)$, $0=t=T$, $a=$cte 0 y consideremos ahora su analogía para el caso discreto en el siguiente ejemplo.

<u>Ejemplo</u>. 3.5.1 Consideremos un sistema descrito por un conjunto discreto de variables $f=(f_o,f_1,\ldots,f_N)$ y gobernado por las ecuaciones de diferencias

$f_{k+1}=af_k+u_k$ $k=0,1,\ldots,N-1$

con $f_o=c$, en donde el vector $u=(u_o,u_1,\ldots,u_N)$ denota las variables de control del sistema.

Supongamos que la función criterio (evaluación del sistema) está dada por

$J_N(f,u)=\sum_{k=0}^{N}(f_k^2+u_k^2)$.

Supongamos que estamos interesados en encontrar valores u_k tales que la función criterio sea minimizada.

Introduzcamos la función

$G_N(c)=\min\limits_{u} J_N(f,u)$.

Notamos que si $u_k=0$ $k=0,1,\ldots,N$, entonces las f_k's pueden ser fácilmente determinadas, puesto que

$f_k^2=a^2f_{k-1}^2=a^2(a^2f_{k-2})=\ldots=a^{2k}f_o^2=a^{2k}c^2$,

de donde

$J_N(f,0)=\sum_{k=0}^{N} a^{2k}c^2=c^2\sum_{k=0}^{N} a^{2k}$.

Notamos que en general $G_N(c)=\min J_N(f,u)$ define recursividad.

Si u_o ha sido determinado, entonces $f_1=ac+u_o$, por consiguiente

$J_N(f,u)=(c^2+u_o^2)+\sum_{k=1}^{N}(f_k^2+u_k^2)$ y $G_{N-1}(f_1)=\min\limits_{u_1,\ldots,u_N}\sum_{k=1}^{N}(f_k^2+u_k^2)$,

pero $G_{N-1}(f_1)=G_{N-1}(ac+u_o)$, por consiguiente

$G_N(c)=\min\limits_{u_o}\left\{(c^2+u_o^2)+G_{N-1}(ac+u_o)\right\}$.

Observamos que la minimización de J sobre un vector de dimensión N, ha sido reducido a N problemas de carácter unidimensional.

La última ecuación establecida es conocida como la ecuación funcional de recursividad de Bellman.

El problema básico de la programación dinámica.

Considere el sistema dinámico de parámetro discreto

$f_{i+1}=g_i(f_i,u_i)$, $i=1,..,N$

donde f_i se denomina la variable de estado (i.e., resume toda la información del sistema dinámico) en el tiempo i y se supone que $f_i \in S_i$ (S_i conocido). Asímismo, u_i se denomina la variable de decisión y se supone también que $u_i \in C_i$ (C_i conocido). La función g_i es llamada la función de transformación de estados y permite determinar el estado en que se encuentra el sistema en el tiempo i+1 dado que se conoce el estado inicial en el tiempo i y que se efectuó la decisión u_i.

Dado un estado inicial f_o se desea determinar la política de decisiones $P=(u_1^*,u_2^*,...,u_N^*)$ tal que minimice (o maximice)

$J(f_1,..,f_N,f_{N+1},u_1,..,u_N)=\sum_{i=1}^{N} F_i(f_i,f_{i+1},u_i)$ (Evaluación del sistema),

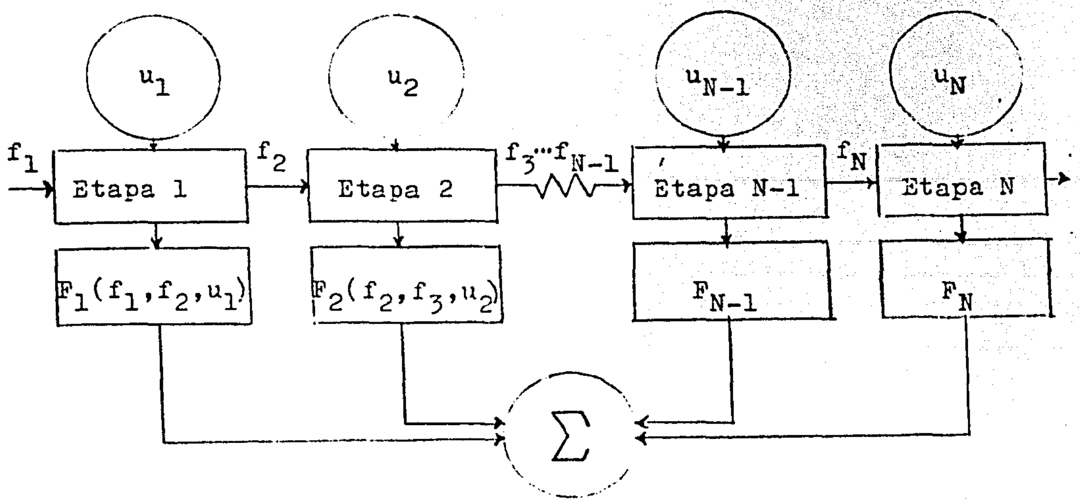
sujeto a

$f_{i+1}=g_i(f_i,u_i)$ $i=1,..,N$

$f_i \in S_i$; $u_i \in C_i$ $i=1,..,N$ (S_i y C_i conocidos)

en donde las funciones g_i y F_i son conocidas.

Una descripción esquemática del desarrollo del sistema dinámico de parámetro discreto con su correspondiente proceso de evaluación se muestra en la siguiente figura



Una explicación de la manera en que se toman las decisiones, se transforman los estados y se evalúan las decisiones, se tiene a continuación:

Etapa 1. El decisor observa el estado inicial f_1 y basado en esta información efectúa la decisión u_1 (que debería escribirse $u_1(f_1)$). Como resultado de ésto se obtiene el estado f_2 de acuerdo a la transformación $f_2=g_1(f_1,u_1)$ y el correspondiente valor asociado a tales estados y decisiones, denotado

$F_1(f_1,f_2,u_1)$.

Etapa k. El decisor observa el estado f_k y basado en esta información toma la decisión u_k. Entonces la transformación $g_k(f_k,u_k)$ proporciona el nuevo estado f_{k+1} y se contabiliza el correspondiente valor

$F_k(f_k,f_{k+1},u_k)$.

Etapa N. El decisor observa el estado f_N y efectúa la decisión u_N. Se obtiene el nuevo estado f_{N+1} y el correspondiente costo o beneficio $F_N(f_N,f_{N+1},u_N)$.

Lo que se desea en este proceso de decisiones secuenciales es determinar la política $P=(u_1^*,u_2^*,\ldots,u_N^*)$ que minimice la suma total de los valores $F_i(f_i,f_{i+1},u_i)$ de cada etapa, esto es,

$$J(f_1,\ldots,f_N,f_{N+1},u_1,\ldots,u_N)=\sum_{i=1}^{N} F_i(f_i,f_{i+1},u_i) .$$

Una manera de resolver el problema de N etapas de decisión es proceder a descomponerlo en N problemas cuya solución es equivalente al problema original. A grandes rasgos lo que se pretende es resolver la última etapa y usar dichos resultados para resolver la penúltima etapa y así sucesivamente. Este proceso es justificado mediante el siguiente Teorema.

<u>Teorema</u>. 3.5.2 (Principio de Optimalidad de Bellman).
Toda subpolítica $P'=(u_i^*,u_{i+1}^*,\ldots,u_j^*)$ extraída de una política óptima $P=(u_1^*,u_2^*,\ldots,u_N^*)$, es óptima de f_i a f_j.

Demostración: Que la subpolítica $P'=(u_i^*,\ldots,u_j^*)$ sea óptima significa que

$\min J(f_i,..,f_j,f_{j+1},u_i,...,u_j)=J(f_i,..,f_j,f_{j+1},u_i^*,..,u_j^*)$

sujeto a

$f_{k+1}=g_k(f_k,u_k) \quad k=i,...,j$

$f_k\in S_k \; ; \; u_k\in C_k \quad k=i,...,j.$

Supongamos que por el contrario, que la subpolítica $P'=(u_i^*,..,u_j^*)$ no es óptima, entonces existiría una subpolítica de f_i a f_j que tuviera un valor menor y sería posible mejorar $P=(u_1^*,...,u_N^*)$ reemplazando la parte $\{u_i^*,...,u_j^*\}$ por esta nueva subpolítica, lo que contradice el hecho de que P sea una política óptima.

Hagamos las siguientes observaciones con respecto a los métodos variacionales. Consideremos una funcional de la forma

$J(f)=\int_a^b F(x,f,f')dx$

sujeta a

$f(a)=A$; $f(b)=B$; $f\in S$ (S un subconjunto abierto de $C^1[a,b]$)

$F\in C^2(R)$; $(f,f')\in R$.

Aquí, como ya se sabe, cada curva es asignada a cierto número. Dividamos el intervalo $[a,b]$ en n+1 partes iguales, por medio de los puntos

$a=x_o,x_1,...,x_{n-1},x_n=b,$

reemplacemos a la curva $f=f(x)$ por la línea poligonal con vértices

$(a,A),\ (x_1,f(x_1)),...,\ (x_{n-1},f(x_{n-1})),\ (x_n,f(x_n))=(b,B)$

y aproximemos la funcional J por la suma

$J(f_1,..,f_n)=\sum_{i=1}^{n+1} F(x_i,f_i,(f_i-f_{i-1})/h),$

en donde $f_i=f(x_i)$, $h=x_i-x_{i-1}$.

Ahora bien la funcional $J(f)=\int_a^b F(x,f,f')dx$ tiene la propiedad de "localización", es decir, si dividimos la curva $f=f(x)$ y calculamos el valor de la funcional para cada una de las partes, la suma de los valores de la funcional de estas partes es igual al valor de la funcional si consideramos toda la curva, entonces al aproximarnos a la funcional por líneas poligonales esta propiedad debe ser heredada.

De acuerdo a la propiedad de "localización". Si f es extremo de J, entonces al dividir f en dos partes, digamos, en el punto $x=c$. $a<c<b$, tenemos que $f|_{[a,c]}$ es extremo de $J|_{[a,c]}$ y $f|_{[c,b]}$ es extremo de $J|_{[c,b]}$. Esto puede ser planteado más formalmente como sigue: De entre todas las funciones $f\in C^1[a,b]\cap PC[a,b]$ (digamos $f\in C^1[a,b]$ excepto en $x=c$), las cuales satisfacen las condiciones de frontera

$f(a)=A$; $f(b)=B$,

deseamos encontrar una función para la cual la funcional

$$J(f)=\int_a^b F(x,f,f')dx$$

tenga un extremo débil. Es claro que sobre cada uno de los intervalos $[a,c]$ y $[c,b]$ la función para la cual J tiene un extremo débil, debe satisfacer la ecuación de Euler-Lagrange

$$F_f-\frac{d}{dx}F_{f'}=0.$$

Al igual que antes esta propiedad debe ser heredada para aproximaciones por medio de líneas poligonales, en concordancia con el principio de optimalidad (o por lo menos con una de sus formas).

Ahora bien, regresando a nuestra discusión sobre el problema básico de la programación dinámica y a la forma en que éste se resolvería. Específicamente el procedimiento que se realiza es como sigue:

En la última etapa la decisión u_N depende únicamente del vector de estados $f_N=(f_{N_1},f_{N_2},\ldots,f_{N_S})$ y para cada valor (componente) de f_N lo que se desea es determinar u_N^* tal que resuelva el problema

$$G_N(f_N)=\min\left\{F_N(f_N,f_{N+1},u_N)\ ;\ u_N\in C_N\right\},$$

donde

$f_{N+1}=g_N(f_N,u_N)$. Aquí se supone que $G_N(f_N)$ es un valor finito y que existe la decisión u_N^* óptima.

En la etapa N-1, las decisiones (u_N,u_{N-1}) que se realicen dependen únicamente del vector de estados f_{N-1}. Dichas decisiones serán óptimas si resuelven el problema

$$G_{N-1}(f_{N-1}) = \min_{\{u_i\}} \{F_{N-1}(f_{N-1}, f_N, u_{N-1}) + F_N(f_N, f_{N+1}, u_N)\}$$

sujeto a

$$f_{i+1} = g_i(f_i, u_i) \ ; \ f_i \in S_i \ ; \ u_i \in C_i \quad i = N-1, N.$$

Sin embargo, esta función objetivo es equivalente a

$$G_{N-1}(f_{N-1}) = \min_{u_{N-1}} \{F_{N-1}(f_{N-1}, f_N, u_{N-1}) + \min_{u_N}\{F_N(f_N, f_{N+1}, u_N)\}\}$$

$$= \min_{u_{N-1}} \{F_{n-1}(f_{N-1}, f_N, u_{N-1}) + G_N(f_N)\}$$

y se observa que la solución de la etapa N-1 es equivalente a resolver un problema de una sóla decisión (u_{N-1}) y que dicho problema incorpora los valores óptimos asociados con los estados f_N calculados en la etapa N.

En general, para la etapa k donde k=1,2,....,N-1, las decisiones a realizar ($u_k, u_{k+1}, ..., u_N$) dependen únicamente del vector de estados f_K. Dichas decisiones serán óptimas si resuelven el problema

$$G_k(f_k) = \min_{\{u_i\}} \{F_k(f_k, f_{k+1}, u_k) + \sum_{i=k+1}^{N} F_i(f_i, f_{i+1}, u_i)\}$$

sujeto a

$$f_{i+1} = g_i(f_i, u_i) \ ; \ f_i \in S_i \ ; \ u_i \in C_i \quad i = k, k+1, ..., N.$$

Sin embargo, la función objetivo es equivalente a

$$G_k(f_k) = \min_{u_k} \{F_k(f_k, f_{k+1}, u_k) + \min_{u_i} \sum_{i=k+1}^{N} F_i(f_i, f_{i+1}, u_i)\}$$

$$= \min_{u_k} \{F_k(f_k, f_{k+1}, u_k) + G_{k+1}(f_{k+1})\},$$

de donde la solución de la etapa k es equivalente a resolver un problema de una sóla decisión (u_k) y dicho problema incorpora los valores óptimos asociados con los estados f_{k+1} calculados en la etapa k+1.

Como resultado de esta discusión se tiene la siguiente proposición.

<u>Proposición</u>. 3.5.3 La política de decisión y el valor de la función objetivo óptimos asociados con el problema básico de la programa-- ción dinámica pueden obtenerse como sigue:

$$G_N(f_N)=\min_{u_N}\left\{F_N(f_N,f_{N+1},u_N)\right\}$$

$$G_K(f_K)=\min_{u_k}\left\{F_k(f_k,f_{k+1},u_k)+G_{k+1}(f_{k+1})\right\}$$

en donde $k=1,2,\ldots,N-1$. La primera relación se denomina condición de frontera y la segunda relación se denomina ecuaciones recursivas de la programación dinámica. El vector de decisiones $h=(u_1^*,\ldots,u_N^*)$ que resuelve las relaciones antes mencionadas es la solución óptima del problema básico y $G_1(f_1)$ el correspondiente valor óptimo, es decir $\min J(f_1,\ldots,f_N,f_{N+1},u_1,\ldots,u_N)=G_1(f_1)$ sujeto a

$f_{i+1}=g_i(f_i,u_i)$ $i=1,\ldots,N$

$f_i\in S_i$; $u_i\in C_i$ $i=1,\ldots,N.$

Aproximaciones Discretas.

Si deseamos hacer uso de computadoras digitales en la determinación de soluciones numéricas a problemas variacionales, el siguiente método hace uso de las técnicas de la programación dinámica para aproximar tales soluciones.

En lugar de seleccionar una función $f=f(t)$ definida sobre el intervalo $[0,T]$, supongamos que es permitido escoger valores de $f(t)$

$f_k=f(t_k)$ $k=0,1,\ldots,N$

en los puntos

$0<t_o<t_1<\ldots<t_N,$

en donde

$t_k=kh,$ $k=0,1,\ldots,N$ y $h=T/N,$

y en lugar de minimizar

$$J(f)=\int_0^T F(t,f,f')dt$$

sujeto a

$f'(t)=G(t,f,f'),$ $f(0)=c$

supongamos que deseamos minimizar la suma finita

$$J(f_1,\dots,f_N)=\sum_{k=0}^{N-1} F(kh,f_k,\frac{f_{k+1}-f_k}{h})h$$

sujeto a

$$f_0=c,\ f_{k+1}=f_k+G(hk,f_k,\frac{f_{k+1}-f_k}{h})h,\ k=0,1,\dots,N-1.$$

Pongamos ahora

$$g_{N-j}(f_j)=\min_{f_j}\left\{\sum_{k=j}^{N-1} F(kh,f_k,\frac{f_{k+1}-f_k}{h})h\right\},$$

entonces en general $g_{N-j}(f_j)$ define recursividad, es decir

$$g_N(c)=\min_{f_0}\left\{F(0,c,\frac{f_1-c}{h})h+g_{N-1}(c+G(0,f_o,\frac{f_1-f_o}{h})h)\right\},$$

$$g_{N-1}(f_1)=\min_{f_1}\left\{F(h,f_1,\frac{f_2-f_1}{h})h+g_{N-2}(f_2)\right\},\ \text{etc.}$$

Ejemplo. 3.5.4 Considere el problema de minimizar la funcional

$$J(f)=\int_0^T f^2(t)dt$$

sujeta a las restricciones

$$H(f)=\int_0^T f(t)dt=b>0;\ f(t)\geq 0;\ f\in(L^2[0,T],\ \|f\|=\left\{\int_0^T |f(t)|^2 dt\right\}^{1/2}).$$

En este caso el Lagrangiano está dado por

$F(f)=f^2(t)+cf(t)$,

donde c es el multiplicador de Lagrange asociado a la restricción del problema variacional y este es tal que $c\leq 0$.

Si f minimiza a J sujeta a las restricciones dadas, entonces $d(J+cH)(f)=0$, equivalentemente $F_f=0$, lo cual conduce a que $f(t)=-c/2$ y c se determina a partir de la relación

$$\int_0^T -c/2dt=b \text{ ó } -c/2=b/T,$$

entonces $f(t)=b/T$. $t\in[0,T]$.

Veamos ahora que efectivamente $f(t)=b/T$ es tal que minimiza a J sujeta a las restricciones dadas. Una condición suficiente de mínimo es que la diferencial segunda de Fréchet $d^2(J+cH)$ sea fuertemente positiva. Primero notamos que el Lagrangiano $F\in C^3(\Omega)$, Ω una región que contiene a f, luego

$$d^2(J+cH)(f,f)=\tfrac{1}{2}\int_0^T F_{ff}(f(t))h^2(t)dt=\int_0^T h^2(t)dt=\|h\|^2 \text{ para toda } h\in L^2[0,T],$$

es decir, la funcional cuadrática $d^2(J+cH)$ es fuertemente positiva.

Suponga ahora que en lugar de determinar una función $f \in L^2[0,T]$ que minimice al problema anterior, es posible considerar valores f_k, $k=1,\ldots,N$ dados por

$f_k=f(t_k)$, $t_k=hk$, $k=1,\ldots,N$; y $h=T/N>0$

de tal manera que sean solución del problema

minimizar $J(f_1,\ldots,f_N)=\sum_{k=1}^{N} f_k^2 h$

sujeto a

$\sum_{k=1}^{N} f_k=b/h$; $f_k \geq 0$, $k=1,\ldots,N$.

Observe primero que determinar el mínimo de $J(f_1,\ldots,f_N)$ sujeto a las restricciones dadas, es equivalente a determinar el mínimo de $\sum_{k=1}^{N} f_k^2$ sujeto a las restricciones dadas puesto que $h=T/N>0$.

Empleando técnicas de la programación dinámica, definimos las ecuaciones recursivas

$$J_k(b_{k-1})=\min\left\{ f_k^2 + J_{k+1}(b_{k-1}-f_k) \;;\; 0 \leq f_k \leq b_{k-1} \right\}$$

donde

$b_k=(b/h)-f_1-\cdots-f_k=b_{k-1}-f_k$, $k=1,\ldots,N$; $b_0=b/h$; $b_N=0$.

La condición de frontera en este caso es $J_{N+1}(u)=0$ para todo u.

Ahora bien, se puede probar por inducción que las relaciones

$$J_{N-k}(b_{N-k-1})=\frac{b_{N-k-1}^2}{k+1} \quad , \quad f_{N-k}^*=\frac{b_{N-k-1}}{k+1}$$

son válidas para toda $k \in \mathbb{N}$.

Así para $k=N-1$

$J_1(b_0)=J_1(b/h)=b^2/(Th)$ y $f_1^*=b/(hN)=b/T$

de la misma manera para $k=N-2$, tenemos que

$f_2^*=\frac{b_1}{N-1}=((b/h)-f^*)/(N-1)=b/T$,

en $N-1$ pasos obtendremos la solución

$f_1^*=\ldots=f_N^*=b/T$.

Empleando ahora el método de diferencias finitas para resolver el problema anterior se hace el siguiente planteamiento. Se desea resolver el problema de minimizar la función $J:\mathbb{R}^N \to \mathbb{R}$ definida por

$J(f_1,\dots,f_N)=\sum_{k=0}^{N} f_k^2$

sujeta a las restricciones

$\sum_{k=1}^{N} f_k=b/h;\ f_k\geq 0,\ k=1,\dots,N.$

El Lagrangiano de la función real de variable vectorial J sujeta a las restricciones dadas, está dado por

$F(f_1,\dots,f_N)=\sum_{k=1}^{N} f_k^2 + c\sum_{k=1}^{N} f_k$,

en donde c es el multiplicador de Lagrange asociado a la restricción $\sum_{k=1}^{N} f_k=b/h$ y este es tal que $c\leq 0$.

Una condición necesaria de un mínimo $(f_1,\dots,f_N)$ es que

$dF(f_1,\dots,f_N)(g_1,\dots,g_N)=(\operatorname{grad}F(f_1,\dots,f_N),(g_1,\dots,g_N))$

$\sum_{k=1}^{N}\frac{\partial F}{\partial f_k}g_k=0$ para todo $(g_1,\dots,g_N)\in R^N$

equivalentemente

$\operatorname{grad}F(f_1,\dots,f_N)=0$ ó $\frac{\partial F}{\partial f_k}=0,\ k=1,\dots,N$ ó $f_k=-c/2,\ k=1,\dots,N$

y el multiplicador de Lagrange c es determinado mediante

$-\sum_{k=1}^{N}\frac{1}{2}c=b/h$ ó $-\frac{1}{2}c=b/(Nh)=b/T$,

en consecuencia

$f_k=b/T,\ k=1,\dots,N$

y una condición suficiente de mínimo es que la matriz simétrica

$$d^2F(f_1,\dots,f_N)=\begin{bmatrix} \frac{\partial^2 F}{\partial f_1^2} & \cdots & \frac{\partial^2 F}{\partial f_1\partial f_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial f_N\partial f_1} & \cdots & \frac{\partial^2 F}{\partial f_N^2}\end{bmatrix},$$

sea definida positiva, es decir

$((g_1,\dots,g_N),d^2F(f_1,\dots,f_N)(g_1,\dots,g_N))=$

$(d^2F(f_1,\dots,f_N)(g_1,\dots,g_N),(g_1,\dots,g_N))>0$ para todo $(g_1,\dots,g_N)\neq 0$

pero

$\frac{\partial^2 F}{\partial f_k \partial f_m}=0$, $\frac{\partial^2 F}{\partial f_k^2}=2$, $k,m=1,\ldots,N$
en consecuencia, la matriz $d^2F(f_1,\ldots,f_N)$ es definida positiva. Finalmente, observe que el problema anterior se puede plantear como el problema de control (discretizado).
Minimizar $J(u_1,\ldots,u_{N+1})=\sum_{k=1}^{N}(u_k-u_{k+1})^2$,
sujeto a
$0\leq u_k-u_{k+1}=f_k$, $k=1,\ldots,N$; $u_{N+1}=0$; $u_1=b/h$,
puesto que
$0=u_{N+1}=u_N-f_N=u_{N-1}-(f_{N-1}+f_N)=\ldots=u_1-\sum_{k=1}^{N}f_k=b/h-\sum_{k=1}^{N}f_k$
y como antes $f_k^*=b/T$, $k=1,\ldots,N$, y los controles óptimos son
$u_k^*=(b/T)(N-(k-1))$, $k=1,\ldots,N+1$.

3.6 METODOS VARIACIONALES Y EL TEOREMA DE HAHN-BANACH.

Una vez que se han introducido los conceptos de funcional y convexidad a través del capítulo 2. Consideramos ahora, la noción abstracta de integral como una funcional lineal definida sobre algún espacio de funciones. Toca ahora su turno a la formulación de algunas aplicaciones del teorema de Hahn-Banach para la determinación de soluciones a problemas de tipo variacional.

Consideremos el problema de acotar inferiormente la funcional
$J(f)=\int_0^\infty f(x)dF(x)$, $F(x)$ una función conocida, monótona creciente y acotada en $[0,\infty)$, sujeta a las condiciones
$\int_0^\infty x dF(x)=m_1$, $\int_0^\infty x^2 dF(x)=m_2>0$, $m_2-m_1^2>0$
y $f\in V$,
en donde el conjunto V es definido como sigue:
Una función f pertenece a V, si satisface:
(i). $f\in C^3[0,\infty)$.
(ii). $f(x)$ es acotada inferior y superiormente por ciertas funciones del tipo $a_1x+a_2x^2$ para toda $x\in[0,\infty)$.
(iii). $f(x)/x$ es estrictamente convexa para toda $x>0$.

Por ejemplo, la función $f(x)=1-e^{bx}$; $x\geq 0$; $b>0$ pertenece a V, la función $f(x)=e^{-x}$, $x\geq 0$ no pertenece a V puesto que no satisface la condición (iii).

Sea V_o el conjunto de funciones de la forma $g(x)=a_1x+a_2x^2$ definidas sobre $[0,\infty)$.

Sea $G(g)=a_1m_1+a_2m_2$ definida para $g\in V_o$ y sea

$$p(f)=\inf_{f\leq g} G(g),\ f\in V.$$

Observamos que:

V es un espacio lineal, en efecto, si $f_1,f_2\in V$, entonces existen funciones tales que las desigualdades

$$a_1x+a_2x^2\leq f_1(x)\leq a_3x+a_4x^2,\quad b_1x+b_2x^2\leq f_2(x)\leq b_3x+b_4x^2$$

son válidas para $x\geq 0$, entonces la desigualdad

$$(a_1+b_1)x+(a_2+b_2)x^2\leq (f_1+f_2)(x)\leq (a_3+b_2)x+(a_4+b_4)x^2$$

es válida para $x\geq 0$. Si $f_1(x)/x$ y $f_2(x)/x$ son funciones convexas, entonces si $c\in[0,1]$

$$f_1(cx+(1-c)y)/(cx+(1-c)y)\leq cf_1(x)/x + (1-c)f_1(y)/y,$$

$$f_2(cx+(1-c)y)/(cx+(1-c)y)\leq cf_2(x)/x + (1-c)f_2(y)/y,$$

lo cual implica que

$$(f_1+f_2)(cx+(1-c)y)/(cx+(1-c)y)\leq c(f_1+f_2)(x)/x + (1-c)(f_1+f_2)(y)/y,$$

esto es la función $(f_1+f_2)(x)/x$ es también una función convexa.
Además V_o es obviamente un subespacio lineal de V.
Es claro también que

$$p(af)=ap(f) \text{ para toda } f\in V,\ a\geq 0,$$

$$\text{y } p(f_1+f_2)\leq p(f_1)+p(f_2) \text{ para toda } f_1,f_2\in V,$$

puesto que si $f_1,f_2\in V$, $g_1,h_2\in V_o$, $f_1\leq g_1$ y $f_2\leq h_2$, entonces $f_1+f_2\in V$, $g_1+h_2\in V_o$, $f_1+f_2\leq g_1+h_2$ y

$$p(f_1+f_2)=\inf_{\{f_1+f_2=g+h\}} G(g+h)\leq G(g+h)=G(g)+G(h) \text{ para toda } g+h\in V_o \text{ tales}$$

que $f_1+f_2\leq g+h$ y puesto que $p(f_1)+p(f_2)$ es la máxima cota inferior del conjunto de números $G(g)+G(h)$ para los cuales $f_1+f_2\leq g+h$, $g,h\in V_0$, entonces

$$p(f_1+f_2)\leq p(f_1)+p(f_2).$$

Finalmente, se observa que si $f,g\in V_0$ y

$f(x)=a_1x+a_2x^2$, $g(x)=b_1x+b_2x^2$ y $f\leq g$, entonces integrando ambos lados de la desigualdad se tiene que

$$a_1\int_0^\infty x\,dF(x)+a_2\int_0^\infty x^2dF(x)\leq b_1\int_0^\infty x\,dF(x)+b_2\int_0^\infty x^2dF(x),$$

esto es

$$G(f)=a_1m_1+a_2m_2\leq b_1m_1+b_2m_2=G(g),$$

pero $p(f)$ es la máxima cota inferior del conjunto de números $G(g)$ para los cuales $f\leq g$, entonces

$$G(f)\leq p(f) \text{ para todo } f\in V_0.$$

En consecuencia V_0,V,G y p satisfacen las condiciones del teorema de Hahn-Banach 2.6.34. Construiremos la extensión H de G mediante la maximización de $G(g)$ para $g\in V_0$.

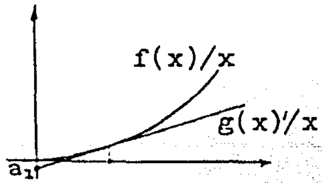
Sean $g\in V_0$ y $f\in V$ tales que $g\leq f$, entonces

$$g(x)=a_1x+a_2x^2\leq f(x),$$

luego si $x>0$

$$g(x)/x=a_1+a_2x\leq f(x)/x,$$

por hipótesis la función $f(x)/x$ es estrictamente convexa para $x>0$. Por consiguiente para maximizar G, consideramos sólo aquellas funciones $g\in V_0$ para las cuales $g(x)/x$ es tangente a $f(x)/x$ en algún punto $x>0$.



Dado $c>0$, si $a_1(c)+a_2(c)x$ es tangente a $f(x)/x$ en c, debemos tener que

(i). $a_1(c)+a_2(c)c=f(c)/c$.

(ii). $a_2(c)=f'(c)/c-f(c)/c^2$.

resolviendo estas ecuaciones con respecto a $a_1(c)$ y $a_2(c)$, encontramos que

$a_1(c)=(-cf'(c)+2f(c))/c$,

$a_2(c)=(cf'(c)-f(c))/c^2$.

Entonces

$$G(a_1(c)x+a_2(c)x^2)=a_1(c)m_1+a_2(c)m_2$$
$$=c^{-2}((-c^2f'(c)+2cf(c))m_1+(cf'(c)-f(c))m_2).$$

Ahora para obtener la extensión H de G maximizamos la expresión anterior con respecto a c. Derivando G con respecto a c dos veces y simplificando encontramos que

$$G'=-c^3(m_1c-m_2)(c^2f''(c)-2cf'(c)+2f(c)),$$
$$G''=-c^4\,[cm_1(c^2f''(c)-2cf'(c)+2f(c))+(m_1c-m_2)(c^3f'''(c)-3c^2f''(c)+$$
$$+6cf'(c)-6f(c)].$$

Si hacemos $G'=0$, obtenemos $c=m_2/m_1$ y notamos que para este valor $G''<0$, en consecuencia G tiene un máximo para $c=m_2/m_1$ para $f\in V$, tal que $g\leq f$ $g\in V_o$.

Sea $H(f)=\sup_{g\leq f} G(g)=(m_1^2/m_2)f(m_2/m_1)$ para toda $f\in V$, observamos también que $0<m_1^2/m_2<1$.

Ahora veamos que $J(f)\leq p(f)$ para toda $f\in V$. En efecto, notamos primero que $J(f)$ es una funcional lineal para toda $f\in V$, también notamos que para $f\in V$, existen $g_1,g_2\in V_o$ tales que

$$g_1(x)=a_1x+a_2x^2\leq f(x)\leq b_1x+b_2x^2=g_2(x),$$

integrando ambos lados de la desigualdad, encontramos que

$$\int_o^\infty(a_1x+a_2x^2)dF(x)\leq\int_o^\infty f(x)dF(x)\leq\int_o^\infty(b_1x+b_2x^2)dF(x),$$

equivalentemente

$$a_1m_1+a_2m_2\leq J(f)\leq b_1m_1+b_2m_2,$$

de donde $J(f)$ existe para $f\in V$. Tomemos $f\in V$ y $g\in V_o$ tales que $f\leq g$, es decir

$f(x)\leq g(x)=a_1x+a_2x^2$, entonces

$J(f) \leq J(g) = \int_0^\infty (a_1x + a_2x^2)\,dF(x) = a_1m_1 + a_2m_2 = G(g),$

pero

$p(f) = \inf_{f \leq g} G(g)$ para $f \in V,$

por lo tanto

$J(f) \leq p(f)$ para $f \in V,$

puesto que $p(f)$ es la máxima cota inferior del conjunto de números $G(g)$ para los cuales $f \leq g$, $g \in V_0$.

Finalmente, si $f \in V$, $g \in V_0$ y $g \leq f$, entonces

$G(g) = J(g) \leq J(f),$

luego

$J(f) \geq \sup_{g \leq f} G(g) = H(f),$

puesto que $\sup_{g \leq f} G(g)$ es la mínima cota superior del conjunto de números $G(g)$ que satisfacen $g \leq f$, $g \in V_0$. En consecuencia

$J(f) \geq (m_1^2/m_2) f(m_2/m_1)$ para toda $f \in V$.

3.7 METODOS VARIACIONALES Y TEORIA DE INFORMACION.

Considere un espacio de probabilidad $(\Omega, A, P(\cdot))$, donde Ω es un espacio muestral, A una σ-álgebra de eventos de Ω y $P(\cdot)$ una medida de probabilidad sobre A. Sea X una variable aleatoria definida en Ω, la cual tiene función de densidad de probabilidad $f(x)$. Se define la entropía o información de Shannon $H(X)$ de la variable aleatoria X como

$H(X) = -\int_{-\infty}^{\infty} f(x) \ln f(x)\,dx = -E(\ln f(x)),$

donde $E(r)$ es la esperanza matemática de la variable aleatoria r.

El principio de máxima entropía de Shannon, proporciona un método general de inferencia sobre una función de densidad de probabilidad desconocida, la cual debe además satisfacer un conjunto de restricciones en términos de valores esperados. El principio de máxima entropía, establece que de todas las densidades a priori que satisfacen las restricciones, se debe de escoger la función de densidad a posteriori f, de tal manera que la entropía H(X) sea máxima, esto es,

$$-\int_{-\infty}^{\infty} f(x)\ln f(x)dx=\max_g\{-E(\ln g(x))\}=\max_g\left\{-\int_{-\infty}^{\infty} g(x)\ln g(x)dx\right\},$$

además de que las restricciones impuestas sean satisfechas.

El principio de máxima entropía resuelve el siguiente problema: Si se desea asignar de manera justa una función de densidad de probabilidad f(x) a una variable aleatoria X cuando sólo se conocen los valores E(X) y/o $E(X^2)$, se debe tomar f(x) de tal manera que la entropía H(X) sea máxima y se satisfagan las restricciones

$$f(x)\geq 0,\ x\in(-\infty,\infty);\ \int_{-\infty}^{\infty} f(x)dx=1;\ E(X)=\int_{-\infty}^{\infty} xf(x)dx=m$$

y/ó

$$E(X^2)=\int_{-\infty}^{\infty} x^2 f(x)dx=s^2.$$

Ejemplo. 3.7.1 Considere el problema variacional

Maximizar $J(f)=-\int_0^{\infty} f(x)\ln f(x)dx$

sujeto a

$$f(x)\geq 0,\ x>0;\ \int_0^{\infty} f(x)dx=1;\ \int_0^{\infty} xf(x)dx=m>0.$$

En este caso el Lagrangiano del problema variacional está dado por

$$F(f)=-f(x)\ln f(x)+c_1 f(x)+c_2 xf(x),$$

donde c_1 y c_2 son los multiplicadores de Lagrange asociados a las restricciones dadas. Si f maximiza a J sujeta a las restricciones dadas, entonces $F_f=0$, lo cual conduce a que

$$f(x)=Ce^{c_2x} \text{ con } C=e^{c_1-1}.$$

Los valores de C y c_2 son obtenidos a partir de las ecuaciones

$$\int_0^{\infty} Ce^{c_2x}\,dx=1 \quad y \quad \int_0^{\infty} Cxe^{c_2x}\,dx=m$$

En este caso C>0 puesto que $e^{c_1-1}>0$ y $c_2<0$, ya que en caso contrario $\int_0^{\infty} Ce^{c_2x}\,dx>\infty$. Haciendo $c_2=-a$, a>0, tenemos que

$$1=\int_0^{\infty} Ce^{-ax}dx=C/a \text{ ó } C=a, \text{ y } \quad m=\int_{-\infty}^{\infty} axe^{-ax}dx=a^{-1} \text{ ó } a=m^{-1}$$

en consecuencia

$$f(x)=m^{-1}\exp(-x/m)\ ,\ x>0,\ m>0,$$

esta función es conocida como función de densidad de probabilidad exponencial.

Ejemplo. 3.7.2 Considere el problema variacional

Maximizar $J(f)=-\int_{-\infty}^{\infty} f(x)\ln f(x)dx$

sujeto a

$f(x)\geq 0,\ x\in(-\infty,\infty);\ \int_{-\infty}^{\infty} f(x)dx=1;\ \int_{-\infty}^{\infty} x^2 f(x)dx=s^2,\ s>0.$

En este caso el Lagrangiano del problema variacional está dado por

$F(f)=-f(x)\ln f(x)+c_1 f(x)+c_2 x^2 f(x),$

donde c_1 y c_2 son los multiplicadores de Lagrange asociados a las restricciones dadas. Si f maximiza a J sujeta a las restricciones dadas, entonces $F_f=0$, lo cual implica que

$f(x)=Ce^{c_2x^2}$ con $C=e^{c_1-1}$.

Los valores de C y c_2 son obtenidos a partir de las ecuaciones

$\int_{-\infty}^{\infty} Ce^{c_2x^2}dx=1$ y $\int_{-\infty}^{\infty} Cx^2e^{c_2x^2}dx=s^2$

Observamos primero que $C>0$ puesto que $C=e^{c_1-1}>0$ y $c_2<0$, ya que en caso contrario $\int_{-\infty}^{\infty} Ce^{c_2x^2}dx>\infty$. Haciendo $C_2=-a^2$, $a>0$, tenemos

$1=\int_{-\infty}^{\infty} Ce^{-a^2x^2}dx=C\sqrt{\pi}/a$ ó $C=a/\sqrt{\pi}$

y

$s^2=\int_{-\infty}^{\infty} (a/\sqrt{\pi})x^2e^{-a^2x^2}dx=\frac{1}{2}a^{-2}$ ó $a^2=\frac{1}{2}s^{-2}$,

en consecuencia

$f(x)=\frac{1}{\sqrt{2\pi}\,s}\exp\{-\frac{1}{2}(x/s)^2\},\ x\in(-\infty,\infty);\ s>0,$

esta función es conocida como función de densidad de probabilidad normal o Gaussiana.

REFERENCIAS BIBLIOGRAFICAS.

1. Athans, Michael., Falb, L. Peter., OPTIMAL CONTROL, McGraw Hill, 1966.
2. Bellman, Richard., Dreyfus, S. E., APPLIED DYNAMIC PROGRAMMING, Princeton Univ. Press, 1957.
3. Bellman, Richard., DYNAMIC PROGRAMMING, Princeton Univ. Press, 1957.
4. Courant, Richard., CALCULUS OF VARIATIONS AND SUPPLEMENTARY NOTES AND EXERCISES, Courant Institute of Mathematical Sciences, N. Y. Univ., 1968.
5. Dreyfus, S. E., DYNAMIC PROGRAMMING AND THE CALCULUS OF VARIATIONS, Academic Press, 1965.
6. Gelfand, Israil. M., Fomin, S. V., CALCULUS OF VARIATIONS, Prentice Hall., 1963.
7. Halmos, P. R., A HILBERT SPACE PROBLEM BOOK, D. Van Nostrand Co., 1967.
8. Hewitt, E., Stromberg, K., REAL AND ABSTRACT ANALYSIS, Springer Verlag, 1965.
9. Kolmogorov, Andrei. N., Fomin, S. V., ELEMENTS OF THE THEORY OF FUNCTIONAL ANALYSIS, Graylock Press, 1957.
10. Luenberger, David. G., OPTIMIZATION BY VECTOR SPACE METHODS, John Wiley, 1969.
11. Pontryagin, L. S., THE MATHEMATICAL THEORY OF OPTIMAL PROCESSES, John Wiley, 1962.
12. Proceedings of a Symposium, CALCULUS OF VARIATIONS AND CONTROL THEORY, Academic Press, 1976.
13. Reed, Mike., Simon, Barry., FUNCTIONAL ANALYSIS, Academic Press, 1972.
14. Riesz, F., Sz Nagy., FUNCTIONAL ANALYSIS, Frederick Ungar, 1955.
15. Royden, H. L., REAL ANALYSIS, Macmillan, 1963.

16. Rudin, Walter., FUNCTIONAL ANALYSIS, McGraw Hill, 1964.

17. Rustagi, Jagdish. S., VARIATIONAL METHODS IN STATISTICS, Academic Press, 1976.

18. Yosida, K., FUNCTIONAL ANALYSIS, Springer Verlag, 1965.