

CÁLCULO DE VARIACIONES

Jorge Ize¹

¹Departamento de Matemáticas y Mecánica. IIMAS-FENOMECC, Universidad Nacional Autónoma de México.

A

Teresa

Pablo

Felipe

Andrés

Contenido

Prefacio a la Segunda Edición	5
Capítulo 0. A modo de introducción	7
Capítulo 1. Ejemplos y problemas	15
1.1 Funcionales	16
1.2 Algunas soluciones clásicas	21
1.3 Existencia, unicidad	30
Capítulo 2. Cálculo en Espacio de Funciones	35
2.1 Espacios de Funciones	35
2.2 Continuidad	40
2.3 Derivadas	44
2.4 Extremos de un funcional	48
Capítulo 3. Lemas Fundamentales	55
Capítulo 4. Condiciones necesarias	67
4.1 Ecuación de Euler-Lagrange	67
4.2 Integral primera	72
4.3 Varias funciones	79
4.4 Varias variables	87
4.5 Derivadas de orden más alto	93
4.6 Frontera libre: condiciones naturales	99
4.7 Discontinuidades, condición de Weierstrass	106
4.8 Multiplicadores de Lagrange	115
4.9 Problemas isoperimétricos	134
4.9.1 Problema isoperimétrico clásico	134
4.9.2 Condiciones holonómicas	137
4.9.3 condiciones no holonómicas	139
Capítulo 5. Cambios de variables y Teoría de Hamilton-Jacobi	143
5.1 Ecuaciones en forma paramétrica	143
5.2 Invariancia de las ecuaciones de Euler	145
5.3 Coordenadas Lagrangianas	155

5.4	Sistema canónico y transformada de Legendre	157
5.5	Teorema de Noether y leyes de conservación	167
5.6	La ecuación de Hamilton-Jacobi	173
5.6.1	Función distancia geodésica	173
5.6.2	Condiciones de transversalidad	176
5.6.3	Soluciones de ecuaciones en derivadas parciales de primer orden . . .	180
5.6.4	Aplicación a la ecuación de Hamilton-Jacobi	187
5.6.5	Campos centrales	191
5.6.6	Campos exactos	200
Capítulo 6. Condiciones suficientes		205
6.1	Extremos de funcionales	206
6.2	Condiciones suficientes para extremos en la norma C^1	212
6.3	Condiciones suficientes para varias funciones	225
6.4	Puntos conjugados y campos centrales.	238
6.5	Condiciones suficientes en la norma C^0	242
Apéndice. Notas sobre métodos directos		253
Bibliografía		269
Índice		269

Prefacio a la Segunda Edición

El material de este libro fue originalmente elaborado para uno de los cursos en el V Coloquio del Departamento de Matemáticas, coordinado por el Dr. Cristóbal Vargas, que se llevó a cabo del 2 al 22 de agosto de 1987 en Pátzcuaro, Michoacán.

En la introducción a las notas el Dr. Vargas apuntaba *“Esperamos que con las sugerencias y la ayuda de los participantes a los cursos, evolucionen en un texto que sea de utilidad para profesores y estudiantes y que enriquezcan el ambiente académico de las escuelas de ingeniería y ciencias del país”*.

Estas notas han sido usadas en varios cursos de la Facultad de Ciencias de la UNAM con ligeras mejoras cada vez. En particular Enrique Loubet y Ramón Plaza dedicaron su servicio social a estas notas, con el resultado de una mejor presentación y figuras más profesionales. Por otra parte, la ayuda informática de Ana Cecilia López fue muy valiosa. Finalmente, Alma Rosa Rodríguez, con su eficiencia de costumbre, hizo que esta edición llegara a su versión final. Mis agradecimientos a los cuatro.

México, D.F. octubre del 2002.

Capítulo 0

A modo de introducción

La critique est aisée
Mais l'art est difficile
Boileau

En 1985 escribí, en *Miscelánea Matemática* No. 15, la reseña del muy buen libro de J. L. Troutman: *Variational Calculus with Elementary Convexity*. En ese momento no entraba en mis planes el estar obligado a escribir un texto, aún un conjunto de notas, sobre la materia. Por lo tanto, en esa crítica, expresé libremente mis ideas y deseos sobre un texto ideal de Cálculo de Variaciones. Por honestidad y quizás para el divertimento del lector, me permitiré reproducir integralmente el texto de la reseña:

“El Cálculo de Variaciones es uno de los campos con más tradición dentro de las matemáticas y de sus aplicaciones. Si bien la idea de minimización como fundamento de muchas leyes físicas tuvo su origen entre los griegos, no es sino hasta los últimos años del siglo XVII que, con la ayuda del recientemente formalizado cálculo diferencial, se formuló por primera vez el problema de encontrar el mínimo de una función definida sobre un conjunto de funciones. A partir del desafío público de Johann Bernoulli, en 1696, a los mejores matemáticos de su tiempo, de resolver el problema de la curva dando el mínimo tiempo de recorrido entre dos puntos del plano, la mayoría de los grandes nombres de la Matemática anterior a este siglo ha sido asociada al Cálculo de Variaciones. Los principios variacionales, empezando por el principio de mínima acción y la teoría de Hamilton-Jacobi, han sido el motor del desarrollo conceptual de la Física y siguen contribuyendo poderosamente en la modelación matemática de muchos otros campos de la Ciencia. Los métodos variacionales directos, iniciados por Hilbert, han permitido el crecimiento espectacular del estudio de las ecuaciones en derivadas parciales, de un gran número de los métodos del Análisis Numérico y, más recientemente, de una buena parte del Análisis No lineal.

Hasta hace algunos años, un curso de Cálculo de Variaciones era obligatorio para los estudiantes de las carreras de Matemático, Físico o Ingeniero. En la mayoría de los casos este curso era una pesadilla para los alumnos con inquietudes matemáticas y un deslumbramiento por los Misterios de la Matemática para los otros. En general uno se pasaba en blanco la materia y se quedaba con malos recuerdos de ese tema tan complicado. Primero había unos bichos extraños, llamados variaciones, a veces reales, a veces virtuales, a veces de primera y a veces de segunda. Según su categoría tenían un símbolo diferente, deltas todos pero con caligrafía gótica. Estos se aplicaban a integrales en virtud de principios de minimalidad de entes también góticos pero sagrados y por lo tanto incuestionables. El llegar a las ecuaciones de Euler-Lagrange era un alivio: por fin algo con deltas derechas! y una decepción ya que invariablemente el profesor trataba de ignorantes a los alumnos que pedían ¿“cómo se resuelve eso?”. Después seguían acumulándose las dudas: ¿Por qué $1 + u^2 = 1$ y u no es cero? ¿Por qué multiplicadores góticos? ¿Por qué una solución a una condición necesaria se volvía un mínimo del terrible ente gótico? pregunta sacrílega por excelencia. Después? más misterios: las temibles ecuaciones canónicas (¿Serán o no serán equivalentes a las ecuaciones de Euler-Lagrange?), las gloriosas transformaciones canónicas (¿Que tantas cosas se pueden hacer con ellas! ¿No serán simplemente la regla de la cadena?), la ubicua transformación de Legendre (¿funciona o no?) y, el misterio de los misterios, la santa ecuación de Hamilton-Jacobi (un minuto de silencio admirador, por favor!). . . . De implicación dudosa en implicación dudosa, de renunciamento en renunciamento, el estudiante más iconoclasta y más crítico había perdido la fe y pocos eran los que se atrevían a preguntar el por qué y el para qué de todo esto ni a pedir un ejemplo “trivial”. El profesor, más tranquilo, podía concluir su curso viendo algunas otras perlas sobre la “suficiencia” de la teoría, en general más deltas góticas, esta vez elevadas al cuadrado, y algunos otras \mathcal{E} y \mathcal{W} con sabor a Hessiano.

Quizás estos recuerdos sean demasiado apocalípticos, (aunque los lectores tendrán sus propios cuentos de terror), pero reflejan la mezcla de principios meta-físicos mal digeridos, de notaciones complicadas y de métodos matemáticos confusos y mal establecidos, por parte, seguramente, de los alumnos y, muy posiblemente, por parte de los profesores quienes no poseían la magnífica intuición física de los grandes fundadores de la materia ni la humildad necesaria para reconocer que tampoco ellos entendían.

Para la fortuna de muchos, el curso de Cálculo de Variaciones fue desplazado poco a poco hacia pequeños segmentos de cursos de aplicaciones y, en la práctica de muchos matemáticos, el vacío completo. Responsables de esta situación lo son los problemas que ya describí y los hechos que la Física moderna se volvió “lineal”, que los modelos de la Física clásica ya estaban bien establecidos con su formulación en términos de ecuaciones diferenciales y que muchos matemáticos se dedicaron a estudiar, clasificar y resolver esas ecuaciones.

Sin embargo, desde hace ya algunas décadas, varios matemáticos y físicos empezaron a cuestionar la validez de los modelos que se les presentaban y de las simplificaciones que hasta ese momento se les había vendido como pan sagrado. Surgieron entonces nuevas versiones de las distintas áreas de la Física, todas caracterizadas por la palabra no lineal, y nuevos métodos matemáticos. Gracias a la herramienta del Análisis Funcional, pulida por su uso en el estudio

de las ecuaciones en derivadas parciales lineales, los métodos variacionales han permitido analizar los problemas no lineales que plantean una variedad cada vez más grande de campos (a veces un poco esotéricos). Limitándose al área y a los métodos de las matemáticas, podría citar teoría de control, optimización no lineal, desigualdades variacionales, geometría (una de las áreas más afines), sistemas hamiltonianos infinitos, ecuaciones elípticas no lineales, análisis global, etc., es decir, una buena parte de la investigación que se realiza actualmente en análisis tiene algo de Cálculo de Variaciones.

Al estar trabajando en esos problemas no lineales o al tomar buenos cursos de graduados, uno se da cuenta que, al final de cuentas, el curso de Cálculo de Variaciones, que uno sufrió en licenciatura, no es más que un caso muy particular de todas estas nuevas técnicas y que bastaría de un poco de modernización para hacerlo inteligible. Pero como la materia ya no es obligatoria, como los textos son obsoletos, como la importancia de un curso introductorio no se ve muy claramente y por otras múltiples razones, uno deja el asunto para después.

A pesar de lo anterior, creo muy firmemente que un curso de Cálculo de Variaciones tiene su lugar, y un muy buen lugar, entre las materias más importantes de una licenciatura en Matemáticas. En efecto, permite: 1) Introducir nuevos conceptos de manera natural y a un nivel de poca dificultad, 2) Ilustrar la teoría con ejemplos accesibles y realistas, 3) Hacer ver que la materia tiene sus logros y sus limitaciones, que es una materia viva y en formación y, en este caso, que es una materia que se puede exponer desmistificando los misterios asociados a su nombre. Quisiera ser más explícito sobre cada uno de los puntos anteriores y revisar los temas más importantes de un curso en Cálculo de Variaciones.

1) En Cálculo diferencial e integral, el alumno aprende, entre otras cosas, a encontrar puntos críticos y puntos extremos de funciones de una o varias variables. El paso siguiente, el mismo estudio para funcionales, es tan natural que los mismos fundadores del Cálculo, Newton, Leibnitz, lo dieron intuitivamente y que la siguiente generación de Matemáticos trató de formalizarlo. Sin embargo, si la noción de infinitesimal fue rápidamente reemplazada en la definición formal de la derivada (y tardó siglos en encontrar una construcción matemática rigurosa), la noción equivalente de variación tuvo que esperar el principio de este siglo para encontrar su substituta en la definición de una derivada definida en un espacio de dimensión infinita. Al definir unos cuantos espacios normados (en este caso esencialmente los espacios C^n y quizás L^2 de manera informal), las nociones de derivadas direccionales y de diferencial son tan naturales que uno no entiende muy bien por qué llevan nombres de matemáticos famosos. Basta entonces saber si estas derivadas están definidas, ver que en un extremo la derivada es cero, una integración por partes (un poco delicada técnicamente) y unos lemas muy simples de cálculo usual, para llegar sin traumas a las ecuaciones de Euler-Lagrange (haré algunos comentarios sobre este punto más adelante). Al agrandar el espacio de trabajo uno puede tratar sin mayor dificultad los problemas de condiciones de frontera naturales (no especificados en la frontera), de fronteras móviles sobre curvas o superficies dadas (condiciones de transversalidad) y de discontinuidades en las funciones (condiciones de Weierstrass). Los problemas isoperimétricos y de constricciones de Lagrange (condicionados) no requieren nada más que en el caso de dimensión finita: el teorema de la función implícita

y la regla de la cadena. Los multiplicadores de Lagrange tienen también su interpretación geométrica pero esta vez con curvas de nivel en espacios de dimensión infinita. Los problemas en forma paramétrica, el teorema de Noether y la invariabilidad de las ecuaciones de Euler, no son más que simples aplicaciones de la regla de la cadena. (Trataré las condiciones suficientes en el tercer punto).

2) El Cálculo de Variaciones ofrece una variedad muy rica de ejemplos: tanto de problemas de geometría (geodésicas, tiempos de recorrido, áreas mínimas) como de optimización y teoría de control y de Física de medios continuos. Estos ejemplos dan la oportunidad de estudiar como se desarrolla un modelo con hipótesis de comportamiento global y no microscópico y partícula por partícula como en los modelos derivados de la ley de Newton. En este sentido es importante describir el desarrollo histórico del principio de mínima acción, ver bajo que condiciones la transformación de Legendre es invertible, cuál es su interpretación geométrica y que las extremales de la acción en las formulaciones Lagrangianas y Hamiltonianas son las mismas (todo esto no es más que la función implícita). Una vez sentadas las bases matemáticas de manera rigurosa, la interpretación física se vuelve más fácil. Del mismo modo las transformaciones canónicas no son más que un simple truco y las leyes de conservación una consecuencia inmediata del teorema de Noether. En cuanto a la derivación de la ecuación de Hamilton-Jacobi, es preferible partir de la noción de funciones geodésicas (o de campos exactos) y ver que bajo esas condiciones las ecuaciones de Euler implican la de Hamilton-Jacobi (la derivación clásica de Jacobi tiene demasiados puntos oscuros para una Lagrangiana arbitraria). Además es imprescindible estudiar cuándo es válida la implicación inversa (es decir el teorema de Jacobi), ya que al resolver una ecuación en derivadas parciales de primer orden con el método de características, su interpretación geométrica permite de manera muy natural estudiar flujos Hamiltonianos y las teorías de campos. En este punto es necesario usar todo el rigor posible (en particular con respecto a las superficies iniciales) y dar una gran cantidad de ejemplos para quitar al método su carácter de misterio y ver que en la práctica encontrar soluciones a una formulación u otra no es más fácil.

3) No es una coincidencia que el punto de los logros y limitaciones de la teoría esté íntimamente relacionado con la discusión de las condiciones suficientes. En efecto si la derivación de las ecuaciones de Euler-Lagrange y de sus “equivalentes” tiene un contenido conceptual reducido, una vez que el cálculo diferencial sobre los espacios de dimensión infinita ha sido establecido, el panorama para las condiciones suficientes es más confuso, a pesar de los esfuerzos de grandes matemáticos como Weierstrass y Hilbert. Parte de la confusión proviene de la poca claridad sobre cuál es el espacio de trabajo. De hecho si uno busca extremos en el espacio natural (funciones continuamente diferenciables, con su norma asociada) las condiciones sobre el Hessiano (Legendre y Weierstrass) y de convexidad local no son más que la extensión de las condiciones correspondientes en el Cálculo clásico. El estudio de la segunda derivada permite introducir los puntos conjugados, la relación con la teoría de Sturm-Liouville, los métodos directos para ecuaciones en derivadas parciales y su teoría espectral. Se hace así más clara la construcción de campos centrales (teoría de Jacobi) y de las soluciones a la ecuación de Hamilton-Jacobi. Esta sería la parte de logros que abre perspectivas muy interesantes en geometría y en teoría de Morse (ver para esto el muy elegante,

pero más difícil, libro de Milnor). Al buscar extremos fuertes, es decir donde la norma es únicamente la de funciones continuas (por lo tanto dos funciones cercanas en esa norma pueden estar muy alejadas en la norma natural), la única esperanza es una condición de convexidad global sobre la derivada o sea sobre la función de exceso de Weierstrass. Estas condiciones son muy restrictivas en la práctica, en particular para problemas multidimensionales, y son más accesibles, ya en otro language, en la teoría de operadores monótonos.

Después de esta larga introducción, creo que es obvio que lo anterior refleja mis opiniones sobre el contenido y la presentación de un curso introductorio al Cálculo de Variaciones. Antes de examinar en más detalle el libro de Troutman, es necesario ver cuáles son los otros libros que uno puede usar para dar ese curso. Los libros sobre Cálculo de Variaciones se dividen en dos grandes grupos. El primero consiste en las extensiones de la teoría, con la herramienta del Análisis Funcional, escritas recientemente para el nivel de posgrado. Entre estos libros, los textos de Ekeland-Temam, Ioffe-Tihomirov, Kinderlehrer-Stampacchia y Vainberg presentan un buen panorama de la investigación reciente. Para funcionales cuadráticos (y ecuaciones lineales) hay muchos libros; a nivel introductorio recomendaría el texto de Rectorys. El segundo grupo es el de los libros introductorios, la mayoría escritos hace más de 20 años, reflejando quizás la baja de “popularidad” de la materia. Entre ellos están los clásicos, obras de matemáticos de principios de este siglo (Bolza, Bliss, Caratheodory) y por lo tanto muy cuidadosos sobre la diferenciabilidad de las funciones pero sin estructura algebraica y con pocos ejemplos. Están también los libros serios de “métodos matemáticos de la Física” como el Courant-Hilbert y el libro de Mikhlin (en general la parte de Cálculo de Variaciones se reduce a las condiciones necesarias y se tratan más bien los problemas multidimensionales y de preferencia cuadráticos). La literatura soviética tiene textos muy completos y con muchos ejemplos (El 'sgol'c, Gelfand-Fomin) pero quizás la necesidad de producir libros para ingenieros les impidió usar francamente el lenguaje del Análisis Funcional. Sin embargo, estos últimos eran, hasta hace algunos años y mezclados con las notas de Courant y con algunos elementos de espacios normados, los textos más fáciles de entender y de presentar a los alumnos. (Prefiero no mencionar los libros de magia ingenieril).

El libro de Troutman es, sin duda alguna para un curso introductorio, muy superior a todos los textos mencionados: llena un gran hueco y podría ser, salvo algunas reservas, óptimo. El libro consta de dos partes, cada una diseñada por el autor para un semestre (creo que de hecho en un curso para físicos y matemáticos uno puede ver todo el material del libro y aún más en un semestre). La primera parte cubre las condiciones necesarias y es la mejor lograda. En un primer capítulo plantea los problemas clásicos de optimización, con una muy buena discusión histórica y una gran variedad de ejemplos. En el segundo capítulo introduce, de manera muy clara, espacios de funciones (solamente C^1) y la derivada de Gâteaux. En el tercer capítulo plantea el problema de suficiencia para integrandos y constricciones convexas: este no es el orden tradicional y uno puede alegar que estos son casos particulares y que uno no llega directamente a la parte más substancial de las ecuaciones de Euler-Lagrange. El orden es una cuestión de gusto personal (puede cambiarse) y el de Troutman permite plantear todos los problemas clásicos al mismo tiempo (frontera libre, condiciones de transversalidad y problemas condicionados) y tratar completamente una serie

de ejemplos clásicos, sin maquinaria pesada. Además puede dar buenos resultados para alumnos que no tienen muy afianzado el concepto de diferencial. En el cuarto capítulo, el autor prueba los lemas de Lagrange y de Du Bois-Reymond: en este capítulo hubiera sido útil agregar el uso de los alisadores de Friedrichs ya que el trabajar con funciones de clase C_0^∞ simplifica mucho la derivación y la aplicación de las ecuaciones de Euler-Lagrange y permite dar un primer ejemplo de derivadas débiles de distribuciones (otro de los “misterios” de la Matemática). En el quinto capítulo Troutman trata los espacios normados y la derivada de Fréchet. Aquí, como en todo el libro, el autor es extremadamente cuidadoso acerca de la existencia de las distintas derivadas y sobre la regularidad de las funciones.

El último capítulo de la primera parte es sobre las ecuaciones de Euler-Lagrange y sus extensiones para los diversos problemas de optimización. Son omitidos los problemas en forma paramétrica y los integrandos homogéneos, la regla de la cadena y el teorema de Noether (este último es una pena por las aplicaciones a las leyes de conservación y a la derivación de integrales primeras, como $\frac{d}{dx}(F - y'F_{y'}) = F_x$, sobre todo dado que el autor prueba esta última identidad tratando un caso particular del teorema de Noether; el esfuerzo adicional para probar el teorema en general no hubiera sido muy grande). A parte de estas omisiones, la presentación de los conceptos, de los ejemplos con sus soluciones parciales y su discusión muy detallada, del desarrollo histórico, de las pocas dificultades técnicas y de la selección de los ejercicios, es difícil de mejorar.

La segunda parte es la más corta, a pesar del material presentado, y quizás la menos pensada. En el primer capítulo se ve muy cuidadosamente el caso de funciones diferenciables por pedazos, las constricciones internas y las condiciones necesarias para un mínimo local, usando fuertemente la convexidad. Troutman dedica el segundo capítulo a los principios variacionales de la Mecánica: el desarrollo histórico está muy bien documentado y explicado en términos matemáticos; las formas canónicas son limpiamente derivadas, con la ayuda del teorema de la función implícita, y los ejemplos de medios continuos son tratados de manera adecuada. Sin embargo, el autor no ve con suficiente detalle la interpretación geométrica de la transformación de Legendre y olvida las transformaciones canónicas y las leyes de conservación. Como este capítulo tiene un carácter histórico, le falta rigor en las posibles equivalencias de las distintas formulaciones; en particular me parece de suma importancia, para la ecuación de Hamilton-Jacobi, ver como se resuelve una ecuación en derivadas parciales de primer orden. Hay en este capítulo demasiados calificativos de la forma “en general”, “uno podría”, “se podría demostrar” y otras vaguedades que contrastan mucho con el rigor de la primera parte. Algunos de estos puntos son aclarados en el último capítulo pero algo de la continuidad del estilo se ha perdido. El último capítulo es un compendio de las condiciones suficientes. Este es el capítulo más incompleto y menos satisfactorio del libro. Por alguna razón el autor no usa la segunda derivada, sea para no sobrecargar la parte de Análisis Funcional sea para demostrar que en lo esencial la convexidad de una función puede definirse sin recurrir a la segunda derivada. No cabe duda que este punto de vista es correcto matemáticamente pero implica un razonamiento más pesado e induce a la confusión entre extremos fuertes y débiles. El capítulo está basado en la teoría de campos en sus diversas versiones. Como el autor no discutió con el rigor debido la ecuación de Hamilton-Jacobi, la

construcción de un campo resulta un poco artificial. La construcción de Jacobi, vista desde un punto puramente algebraico, no refleja las interpretaciones geométricas de la teoría, por parte de los Físicos en relatividad general o por parte de los Matemáticos en geometría. Como el autor no introduce los puntos conjugados, tampoco se ve la relación con los métodos variacionales directos (una pequeña sección sobre ellos, aún de modo puramente informal, hubiera sido muy apreciada y daría más ganas de emprender su estudio que una serie de referencias). Del mismo modo algunos comentarios sobre la teoría L^2 y su versatilidad hubieran sido deseables, pero también aquí el autor no quiso profundizar un poco en el Análisis Funcional. En resumen esta segunda parte merece ser complementada por algún otro texto (por ejemplo el libro de Gelfand-Fomin) y clarificada en el espíritu de todos los comentarios previos.

En conclusión, el libro de Troutman es un excelente libro introductorio al Cálculo de Variaciones. Los alumnos y los profesores tendrán un texto claro, bien documentado, bien escrito, con buenos ejemplos, con muchos ejercicios de toda clase y con espíritu matemático de rigor muy necesario en este tema. Quizás su defecto más grande es el de la falta de completez de la segunda parte. Espero, de todos modos, que al usarlo se podrá modificar substancialmente la imagen de la materia y regresar el tema a un lugar de primera importancia entre los cursos de Cálculo y los cursos de Ecuaciones en Derivadas Parciales.”

FIN DE LA CITA

El lector podrá darse cuenta que me puse en una situación difícil, en la cual la cita de Boileau se justifica plenamente. Como único argumento en mi defensa diré que estas son NOTAS (es decir algo incompleto) y no un libro de texto pulido.

Estas notas están basadas en un curso que dí en la Facultad de Ciencias de la UNAM en 1979 y en 1984. Los prerrequisitos son un buen manejo del Cálculo de varias variables y un sólido primer curso de Ecuaciones Diferenciales Ordinarias. En la medida de lo posible he tratado de no usar los ejemplos y los resultados del libro de Troutman. De este modo el lector podrá complementar estas notas por los textos listados en las referencias. En particular, dado el nivel de esta presentación, no fue posible incluir material sobre los métodos directos en varias variables, ni sobre temas más avanzados como las desigualdades variacionales o los funcionales conjugados, los cuales han tenido mucho éxito en la investigación reciente.

Con respecto a la organización de las notas, se han incluido algunos ejercicios pero no en cantidad suficiente. Pedimos al lector que complete los ejemplos por su cuenta. En cuanto a notación se usan las distintas formas clásicas de escribir las derivadas parciales, diferenciales, etc. Las numeraciones de los resultados (Teoremas, Lemas, etc.) ejemplos y ejercicios son independientes unas de otras dentro de cada sección. La numeración: “Ejemplo 4.1”, se refiere al ejemplo 1 de la sección 4 del capítulo donde se hace ese comentario. La numeración “Ejemplo IV-4.1” es para el ejemplo 1 de la sección 4 del capítulo IV: en ese caso se trata de una referencia a algo que está fuera del capítulo que el lector está estudiando.

Deseo agradecer a Marta Cerilla y a Elizabeth Derbez la magnífica tipografía, a Eugenio Durand, a Clara Garza y a Pablo Padilla el trabajo editorial sin el cual estas notas no hubieran podido aparecer.

MEXICO,D.F.
ABRIL DE 1987.

Capítulo 1

Ejemplos y problemas

Según el poeta Virgilio, el Cálculo de Variaciones nació en 850 a.c. cuando la reina Dido fundó la ciudad de Cartago: su ciudad sería tan grande como pudiera encerrar una piel de toro. Dido cortó la piel en tiras muy finas y logró hacer una cuerda de longitud l . Usando la costa del mar Mediterráneo como un lado de la ciudad, la reina tenía que encontrar la curva, con longitud l , encerrando la mayor área y empezando y terminando sobre la costa.

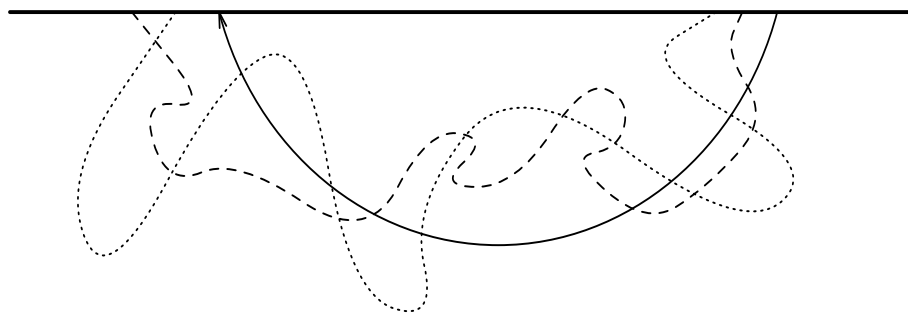


Fig. 0.1

Sin embargo, el nacimiento moderno del Cálculo de Variaciones fue con fanfarrias poco usadas en el mundo matemático: en 1696, Johann Bernoulli desafió públicamente a los mejores matemáticos de su tiempo, proponiéndoles el problema siguiente: Si una partícula se desliza sobre una curva que une a dos puntos en el plano, bajo la acción de la gravedad, ésta tardará un cierto tiempo que depende de la curva. ¿Cuál es la curva que da el tiempo mínimo? ¿La braquistocrona? Para Galileo (60 años antes) era claro que no era la línea recta y opinaba que debía ser un arco de círculo (ver Figura 0.2).

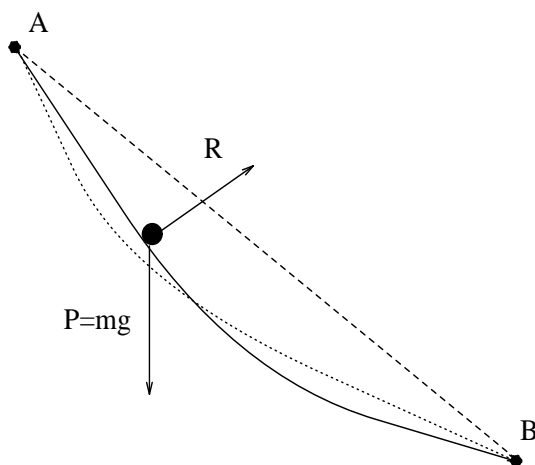


Fig. 0.2

Otra categoría de problemas que han intrigado (y siguen haciéndolo) a los Matemáticos, son las geodésicas, es decir ¿cuál es la curva, sobre una superficie, de longitud mínima?

Finalmente, mencionaremos en esta breve introducción, los problemas de área mínima, o problemas de Plateau (químico belga del siglo pasado), es decir, dada una curva en el espacio, ¿cuál es la superficie de área mínima que tiene a esa curva como frontera?. Uno puede materializar esas superficies con películas de jabón (el área mínima se traduce en tensión mínima) y por lo tanto hacer experimentos. El nombre de Courant está íntimamente ligado a este tipo de problemas, pero la investigación no ha concluido.

1.1 Funcionales

En la breve introducción a este capítulo, hemos indicado algunos de los problemas asociados al Cálculo de Variaciones. En todos ellos se trata de encontrar máximos o mínimos de cantidades donde las variables son curvas, superficies, etc..., es decir, extremos de funciones de funciones. Para esclarecer este concepto daremos algunos ejemplos.

Ejemplo 1.1 : Sea $A(f)$, el área entre la gráfica de la función $f(x)$, $a \leq x \leq b$, $f(x) \geq 0$, y el eje x , es decir

$$A(f) = \int_a^b f(x) dx.$$

Si $f(a) = A$, $f(b) = B$ son dadas y f es continua, ¿cuál es el mínimo de $A(f)$ sobre todas las funciones que satisfacen esas condiciones?

De la figura 1.1(b), parece ser que si tomamos funciones casi 0, tendremos $\inf A(f) = 0$, pero el límite correspondiente a la función $f(x) \equiv 0$, $a < x < b$, $f(a) = A$, $f(b) = B$, no es

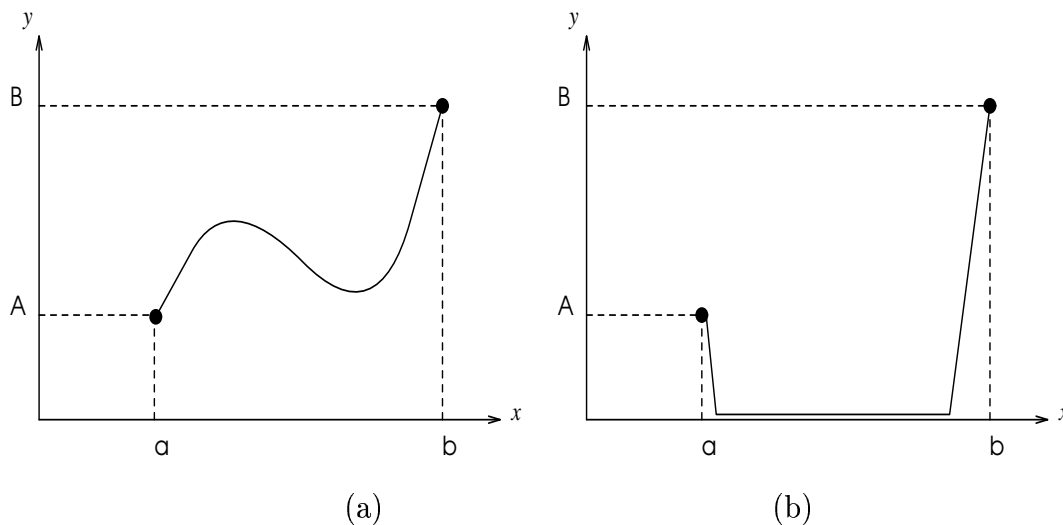


Fig. 1.1

continua, es decir, no pertenece a la clase de funciones estudiadas, o no hay un mínimo.

Ejemplo 1.2 : Geodésicas en el plano.

La longitud de una curva dada por la gráfica de una función es

$$\ell(f) = \int_a^b (1 + f'(x)^2)^{1/2} dx.$$

Por lo tanto al buscar la curva de longitud mínima que une el punto (a, A) a (b, B) , podemos buscar el mínimo de $\ell(f)$, si descartamos curvas que no sean gráficas de funciones. Si no queremos hacerlo, tendremos que considerar

$$\ell(x, y) = \int_0^1 (x'^2 + y'^2)^{1/2} dt,$$

donde parametrizamos la curva por t (el tiempo de recorrido se puede tomar como 1, después de un cambio de variables). Se pide $x(0) = a$, $x(1) = b$, $y(0) = A$, $y(1) = B$. ¿Cambia el resultado si tomamos curvas más generales?

Ejemplo 1.3 : Geodésicas en una región plana.

Si sabemos cuál es la respuesta en el ejemplo anterior, ¿qué pasa si pedimos que las curvas estén en cierta región? La línea recta no es necesariamente la respuesta. Aquí tendremos que controlar la curva.

Ejemplo 1.4 : Geodésicas sobre una superficie en el espacio.

Si una superficie está dada en forma paramétrica por $M(u, v)$, con (u, v) en una región del plano, toda curva $M(t)$ sobre la superficie proviene de una curva $(u(t), v(t))$ en el espacio de los parámetros.

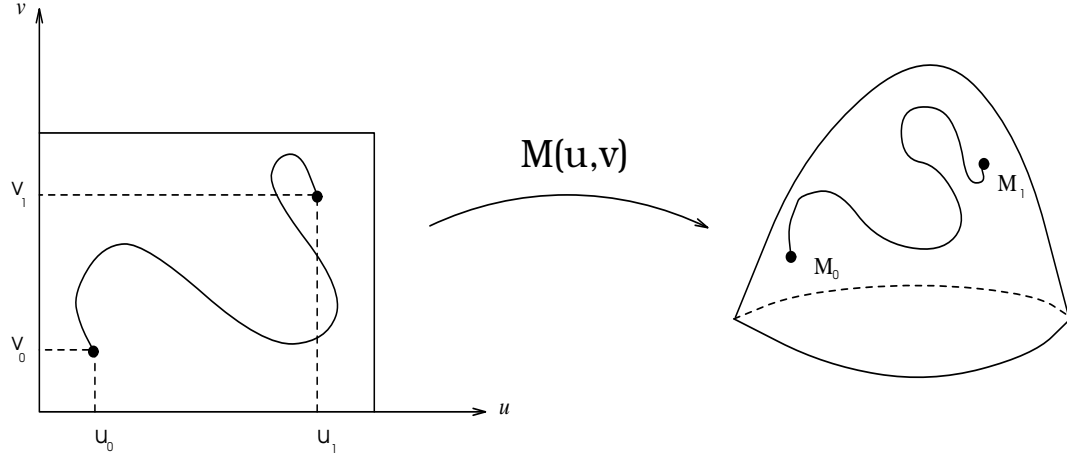


Fig. 1.2

Su longitud será:

$$\int_0^1 \|dM(t)/dt\| dt.$$

Tenemos una cantidad $L(u, v)$ definida por

$$L(u, v) = \int_0^1 \|M_u u' + M_v v'\| dt,$$

con la condición $M(0) = M_0$, $M(1) = M_1$. “Sabemos” que, para una esfera, la curva de longitud mínima es un círculo principal. ¿Por qué? ¿Y en otras superficies?

Ejemplo 1.5 : Problema isoperimétrico.

¿De todas las curvas cerradas, planas y simples de longitud dada, cuál es la que encierra una área máxima? Si denotamos por $(x(t), y(t))$ la curva (t en $[0, 2\pi]$, cambiando el tiempo de recorrido), tenemos que encontrar el máximo de

$$A(x, y) = \int_0^{2\pi} (xy' - x'y) dt/2,$$

con la condición $L = \int_0^{2\pi} (x'^2 + y'^2)^{1/2} dt$.

Este es el prototipo de toda una serie de problemas condicionados. En este caso, la reina Dido hubiera podido concluir que la solución es el círculo, sin embargo la solución rigurosa fue dada solamente a principios de este siglo.

Ejemplo 1.6 : ¿De todas las curvas que unen a dos puntos y de longitud dada, cuál es la que alcanza una altura máxima? Aquí se trata de encontrar la cantidad

$$\max_{f \in C} \max_{x \in [a, b]} f(x),$$

donde C es el conjunto de funciones $f(x)$, con $f(a) = A$, $f(b) = B$,

$$\int_a^b (1 + f'^2(x))^{1/2} dx = L.$$

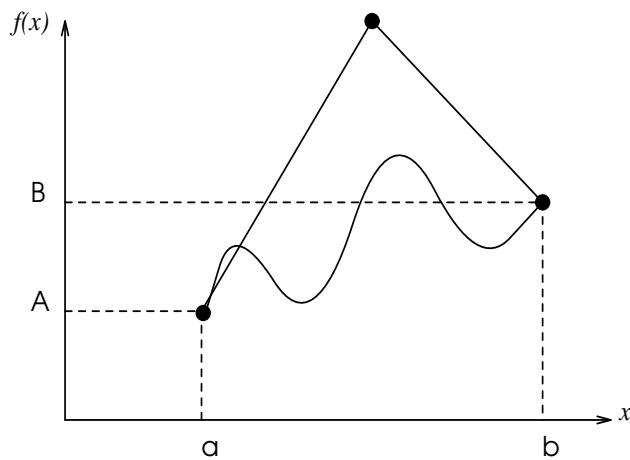


Fig. 1.3

El lector se podrá convencer que la solución corresponde a una curva formada por dos segmentos de recta, por lo tanto una curva que no es diferenciable en el intervalo $[a, b]$.

Ejemplo 1.7 : Superficie de revolución de área mínima.

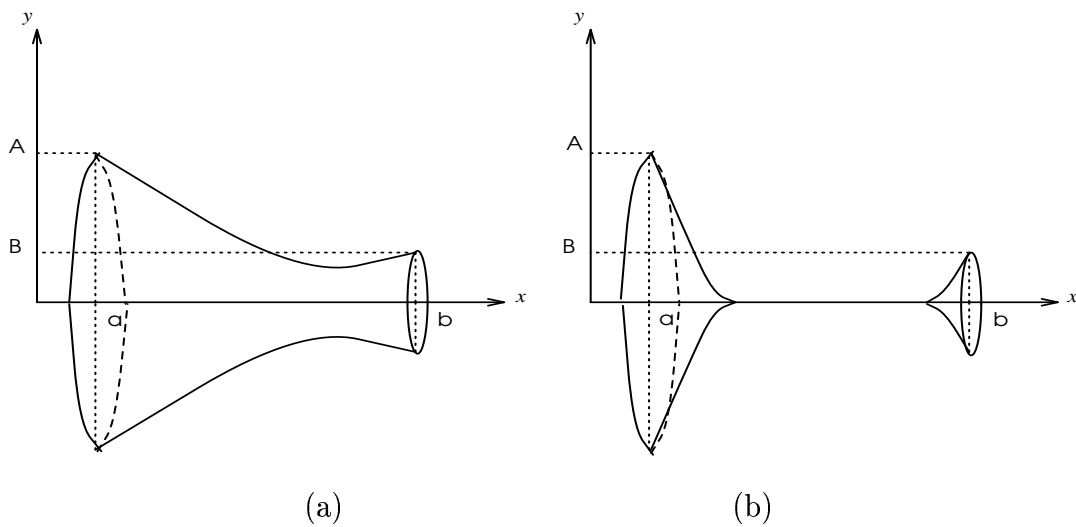


Fig. 1.4

El área de la superficie de revolución generada al girar la curva $y = f(x)$ alrededor del eje x está dada por:

$$S(f) = 2\pi \int_a^b f(x)(1 + f'(x)^2)^{1/2} dx,$$

donde se pide a la función $f(x)$ que cumpla con las condiciones $f(a) = A$, $f(b) = B$. Este es un caso particular del problema de Plateau, donde la frontera de la superficie está dada por dos anillos y la superficie debe ser de revolución. Si alejamos los dos anillos, como en la Figura 1.4 (b), uno puede ver que la pompa de jabón se divide en dos discos llenando los aros, por lo tanto no será una superficie de revolución generada por una función continua.

Ejemplo 1.8 : Si tenemos un campo vectorial $(X(x, y), Y(x, y))$ en el plano, ¿para qué curva, uniendo dos puntos dados, el promedio de la norma del campo sobre la curva será mínimo? Es decir, si la curva es la gráfica de una función, tendremos que considerar la cantidad

$$V(f) = \int_a^b (X(x, f(x))^2 + Y(x, f(x))^2)^{1/2} dx.$$

Ejemplo 1.9 : Membrana elástica.

La posición de equilibrio de una membrana elástica, fija en su frontera, se puede encontrar, minimizando su energía:

$$J(u) = \frac{1}{2} \iint_{\Omega} (u_x^2 + u_y^2) dx dy$$

sobre todas las funciones $u(x, y)$ definidas en Ω , con $u(x, y)$ dada sobre $\partial\Omega$.

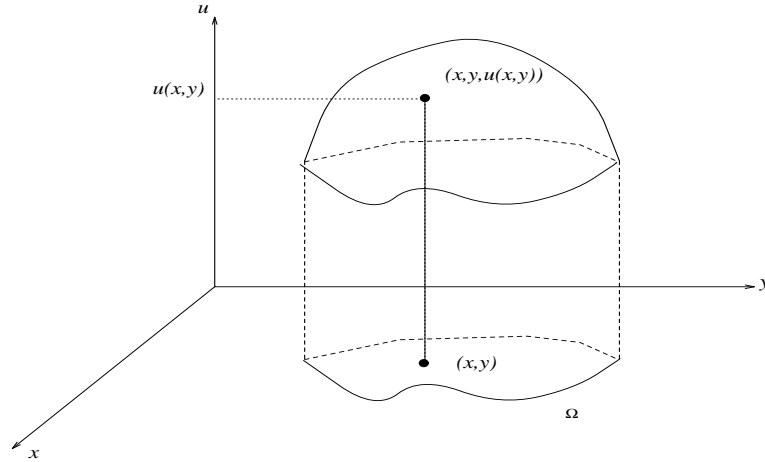


Fig. 1.5

Este problema lleva a la ecuación de Laplace.

En cada uno de los ejemplos anteriores, se trata de encontrar el extremo de una cantidad ($A(f)$, $\ell(x, y)$, $L(u, v)$ etc.) que asocia a cada función, dentro de cierta clase, un número real.

A esta asociación se le llama un *funcional*. En el Cálculo de Variaciones se estudian los extremos de funcionales, muchos de ellos en forma de integral (aunque no exclusivamente, como en el ejemplo 1.6). Muchas de las técnicas que desarrollaremos son aplicables no solamente a funcionales, sino también a cantidades que asocian a una función, otra función, es decir a *operadores* (no lineales). Por ejemplo, en el ejemplo 1.8 podríamos haber estudiado, no solamente el promedio de la norma del campo, sino el campo completo, es decir:

$$f \in C \longrightarrow (X(x, f(x)), Y(x, f(x))),$$

donde $C = \{f : [a, b] \rightarrow \mathbb{R}, f \text{ continua}, f(a) = A, f(b) = B\}$.

En todos estos problemas es muy importante definir con mucha precisión el dominio del funcional. Muchas de las dificultades del Cálculo de Variaciones provienen de la falta de cuidado en esta definición.

1.2 Algunas soluciones clásicas

Ejemplo 2.1 : La Braquistocrona ¹.

Plantearemos primero el problema en términos matemáticos y después daremos dos “soluciones” al problema. Tomando los ejes como en la figura siguiente, y suponiendo que la partícula se deja sin velocidad en el origen, obtenemos:

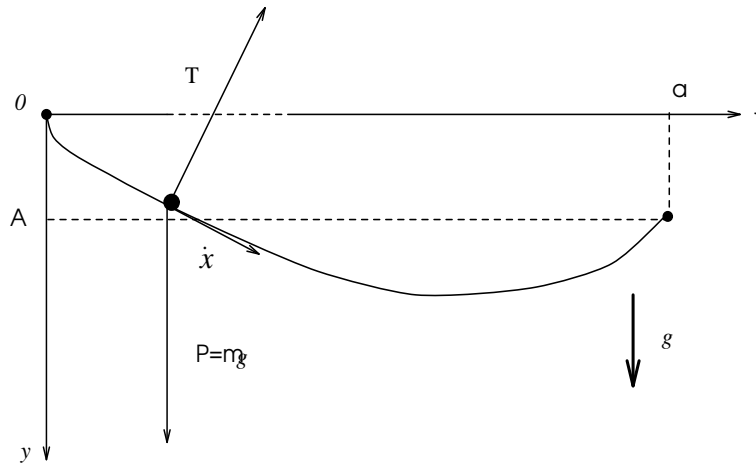


Fig. 2.1

Si representamos el camino por la curva $y = f(x)$ (pedimos al lector que piense si esto no es una restricción) con $f(x) \geq 0$, el tiempo que se tarda la partícula es:

¹Tiempo más corto en griego.

$$T(f) = \int_0^\ell \frac{ds}{v},$$

donde ℓ es la longitud de la curva, $v = ds/dt$ es la velocidad a la distancia s sobre la curva. Como $ds/dt = (\dot{x}^2 + \dot{y}^2)^{1/2} = |\dot{x}|(1 + f'(x)^2)^{1/2}$ y hemos supuesto que $\dot{x} > 0$ (la partícula va siempre hacia la derecha), tenemos: $ds = (1 + f'(x)^2)^{1/2} dx$ y

$$T(f) = \int_0^a \frac{(1 + f'(x)^2)^{1/2}}{v} dx.$$

Para poder expresar v como función de x y de $f(x)$, usamos la ley de Newton:

$$m\ddot{X} = mG + T$$

donde $X = (x(t), y(t))$ denota el vector de posición de la partícula, $G = (0, g)$ es el vector de aceleración de la gravedad, y T representa la reacción de la curva sobre la partícula; es un vector normal a la curva si suponemos que no hay fricción. Proyectando esta ecuación sobre el vector tangencial \dot{X} , obtenemos

$$\dot{X} \cdot \ddot{X} = g\dot{y}.$$

Integrando, obtenemos la conservación de energía:

$$v^2 = |\dot{X}|^2 = 2gy(t) + \text{constante}.$$

Suponiendo que la partícula parte del reposo, evaluando la constante en $t = 0$, con $y(0) = 0, v(0) = 0$, obtenemos que

$$v = (2gy)^{1/2} = (2gf(x))^{1/2}.$$

(Notemos que, de esta ley de conservación de energía, $y(t)$ no puede ser negativa.) Por lo tanto el funcional que debemos estudiar es

$$T(f) = \int_0^a ((1 + f'(x)^2)/(2gf(x)))^{1/2} dx$$

con las condiciones $f(0) = 0, f(a) = A$.

Ejercicio 2.1 :

- (a) Calcular $T(f)$ para una línea recta.
- (b) Calcular $T(f)$ para un arco de círculo. (Esto requiere de cálculo, se puede tomar el radio del círculo como parámetro.)

Solución de Bernoulli.

El propio Bernoulli encontró una solución usando un símil óptico (su hermano Jakob, Euler y otros propusieron otras soluciones más cercanas a la teoría moderna).

Si se aproxima f por una curva poligonal y se supone v como constante en cada intervalo, podemos interpretar el camino de la partícula, como el camino óptico de un rayo de luz atravesando un medio estratificado con índice variando de estrato a estrato.

Según el principio de Fermat, la luz tomará el camino que minimice el tiempo (esta es la analogía con el problema de la braquistocrona, sin embargo la justificación no es evidente) y por lo tanto seguirá la ley de Snell:

$$\frac{\text{sen } i}{v_i} = \frac{\text{sen } r}{v_r}.$$

Si tomamos el límite cuando el ancho de los estratos tiende a 0, obtenemos

$$\frac{\text{sen } \theta}{v} = \text{constante}.$$

Usando $\text{sen}^2 \theta = \cos^2 \phi = (1 + \tan^2 \phi)^{-1} = (1 + f'^2(x))^{-1}$, tenemos

$$f(x)(1 + f'^2(x)) = k > 0.$$

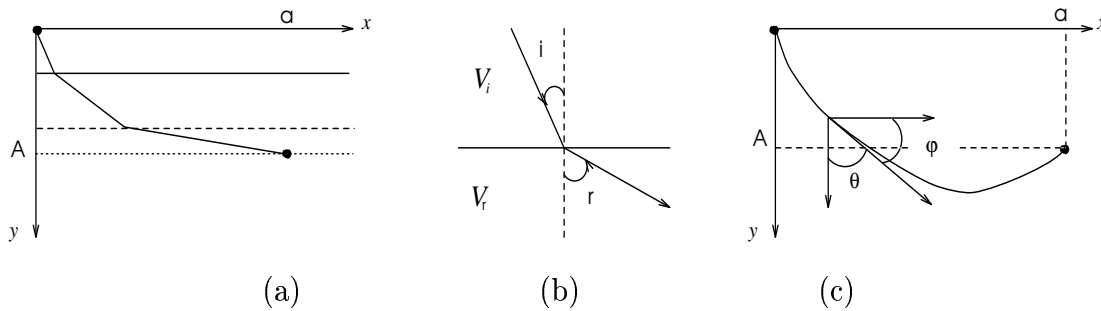


Fig. 2.2

Ejercicio 2.2 :

1. Probar que la cicloide

$$x(\theta) = k(\theta - \text{sen } \theta)/2, \quad y(\theta) = k(1 - \cos \theta)/2$$

es solución de esta ecuación, con $f(x) = y(\theta(x))$, $f(0) = 0$.

2. Probar que las únicas soluciones a la ecuación

$$f'(x) = \pm(k/f(x) - 1)^{1/2}$$

son cicloides o constantes.

3. Probar que si se pide que la solución sea dos veces continuamente diferenciable, entonces existe una y una sola cicloide solución de la ecuación diferencial. (Se sugiere estudiar aparte el caso $a/A \leq \pi/2$).

Esta “solución” de Bernoulli tiene varios problemas:

1. El análogo óptico no está muy bien justificado, en particular la velocidad $(2gf(x))^{1/2}$ no es constante para una trayectoria recta.
2. El argumento del límite debe justificarse.
3. Como se ve del ejercicio 2.2, hay problemas de unicidad según la clase de soluciones que consideramos.
4. No se ha probado que la solución a la ecuación es efectivamente un mínimo del funcional.
5. Se ha restringido el estudio a curvas parametrizadas por x (es decir, a gráficas de funciones).
6. Hay un problema técnico de convergencia de la integral ya que $y(0) = 0$. Esto se debe tomar en cuenta en el momento de definir la clase de funciones estudiadas.

Solución de diferencias finitas.

En esta “solución”, mantendremos la idea de aproximación poligonal, pero trabajaremos con el funcional.

Dividiendo el intervalo $[0, a]$ en n subintervalos $[x_i, x_{i+1}]$, $i = 0, \dots, n-1$, de longitud $h = a/n$, y aproximando la función f por $m_i(x - x_i) + y_i$, con $y_i = f(x_i)$, $m_i = (y_{i+1} - y_i)/h$ entonces $T(f)$ es aproximado por

$$\sum_{i=0}^{n-1} t_i = (2g)^{-1/2} \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} ((1 + m_i^2)/(m_i(x - x_i) + y_i))^{1/2} dx.$$

Es decir, integrando :

$$T(y_1, \dots, y_{n-1}) = (2/g)^{1/2} h \sum_{i=0}^{n-1} (1 + (y_{i+1} - y_i)^2/h^2)^{1/2} / (y_{i+1}^{1/2} + y_i^{1/2}),$$

recordando que $y_0 = 0, y_n = A$.

Al igualar a 0 las parciales con respecto a y_1, \dots, y_{n-1} , se obtienen las ecuaciones:

$$\begin{aligned} & -\frac{1}{h}(m_i(1+m_i^2)^{-1/2} - m_{i-1}(1+m_{i-1}^2)^{-1/2}) = \\ & = (1+m_{i-1}^2)^{-1/2}(y_{i-1}^{1/2} + y_i^{1/2})^{-2} + (1+m_i^2)^{-1/2}(y_i^{1/2} + y_{i+1}^{1/2})^{-2} \end{aligned}$$

para $i = 1, \dots, n-1$.

En lugar de tratar de resolver este sistema de ecuaciones, tomemos el límite cuando h tiende a 0, reemplazando m_i por $f'(x)$ y y_i por $f(x)$. (Evidentemente este paso no está justificado):

$$-(f'(x)(1+f'(x)^2)^{-1/2})' = 1/(2f(x)(1+f'(x)^2)^{1/2}).$$

Calculando la derivada, simplificando y multiplicando por $f'(x)$ tenemos:

$$f'(x)(2f(x)f''(x) + 1 + f'(x)^2) = 0,$$

o sea $(f(x)(1+f'(x)^2))' = 0$.

Llegamos así a la misma ecuación que con el análogo óptico. Sin embargo hemos dado un argumento más cercano a la manera clásica de abordar problemas de extremos en Cálculo de varias variables. Notemos que para poder aplicar el argumento de cambio del tiempo total por la suma de los tiempos en cada intervalo $[x_i, x_{i+1}]$, usamos el hecho que el funcional $T(f)$ se expresa como una integral. Esta forma particular del funcional es tan útil que gran parte del Cálculo de Variaciones se centra sobre este tipo de funcionales (el ejemplo 1.6 no es de ese tipo).

Ejemplo 2.2 : Problema isoperimétrico.

Damos aquí una solución casi satisfactoria al ejemplo 1.5. Encontrar el máximo del área encerrada por una curva:

$$A(x, y) = \int_0^{2\pi} (xy' - x'y)/2 dt,$$

con la condición que la longitud sea dada:

$$L = \int_0^{2\pi} (x'^2 + y'^2)^{1/2} dt.$$

En el caso del círculo $A = \pi r^2, L = 2\pi r$, por lo tanto $A = L^2/4\pi$. Aquí $t = 2\pi s/L$. Queremos probar que la solución al problema es un círculo, es decir, probar que:

$$I = L^2/4\pi - A > 0$$

para toda curva que no sea un círculo; entonces A será máxima para el círculo: $I = 0$.

Si escogemos como parametrización de la curva, la variable $t = 2\pi s/L$, s longitud de arco, tenemos:

$$(ds/dt)^2 = x'^2 + y'^2 = L^2/4\pi^2,$$

y por lo tanto:

$$\int_0^{2\pi} (x'^2 + y'^2) dt = L^2/2\pi$$

Podemos entonces escribir

$$\begin{aligned} I &= (1/2) \int_0^{2\pi} (x'^2 + y'^2 - xy' + x'y) dt \\ &= 1/4 \int_0^{2\pi} ((x' + y)^2 + (y' - x)^2 + x'^2 - x^2 + y'^2 - y^2) dt \end{aligned}$$

Ahora bien, las funciones $x(t)$, $y(t)$ son periódicas, con período 2π , por lo tanto tienen una expansión en series de Fourier:

$$\begin{aligned} x(t) &= \sum_0^\infty (a_n \cos nt + b_n \sin nt) \\ y(t) &= \sum_0^\infty (a'_n \cos nt + b'_n \sin nt) \end{aligned}$$

Podemos, sin cambiar área y longitud, trasladar la curva de tal modo que el centro de gravedad de la figura esté en el origen, es decir, suponer que los promedios de $x(t)$ y $y(t)$ sean 0: $a_0 = a'_0 = 0$.

Es entonces fácil ver que:

$$\int_0^{2\pi} (x'^2 - x^2) dt = \pi \sum_1^\infty (a_n^2 + b_n^2)(n^2 - 1),$$

(usar $\int_0^{2\pi} \cos nt \sin mt dt = 0$, $\int_0^{2\pi} \cos nt \cos mt dt = \pi \delta_{nm}$, $\int_0^{2\pi} \sin nt \sin mt dt = \pi \delta_{nm}$).

Esta expresión es estrictamente positiva, al menos que $a_n = b_n = 0$, para $n > 1$. Por lo tanto: $I > 0$, al menos que sólo permanezcan los términos de orden 1. En ese caso

$$\begin{aligned} I &= \pi/2((b'_1 - a_1)^2 + (a'_1 + b_1)^2) \geq 0 \\ I &= 0 \text{ sólo si } a'_1 = -b_1, b'_1 = a_1, \text{ es decir,} \\ x(t) &= a_1 \cos t + b_1 \sin t, y(t) = -b_1 \cos t + a_1 \sin t \end{aligned}$$

o sea la curva es un círculo.

Aquí se ve claramente que el número de variables (a_n, b_n, a'_n, b'_n) en los problemas del Cálculo de Variaciones es infinito. Evidentemente deberíamos justificar la expansión en series de Fourier y su manipulación.

Ejemplo 2.3 : Geodésica en el plano.

Daremos ahora la respuesta al ejemplo 1.2. Si la curva que une el punto $A = (x_0, y_0)$ con el punto $B = (x_1, y_1)$ está parametrizada por $(x(t), y(t))$, t en $[0, 1]$, entonces

$$L = \int_0^1 (x'^2 + y'^2)^{1/2} dt.$$

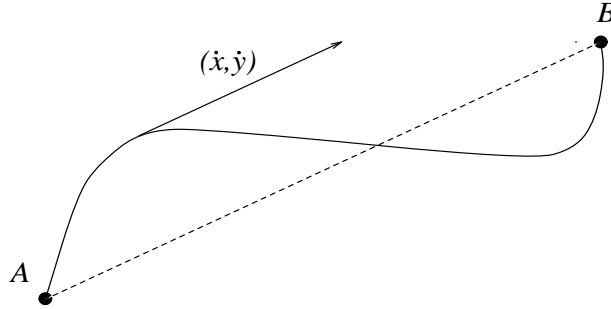


Fig. 2.3

Ahora, interpretando a $(x_1 - x_0)x' + (y_1 - y_0)y'$ como el producto escalar de dos vectores tenemos, por la desigualdad de Cauchy-Schwarz :

$$(x_1 - x_0)x' + (y_1 - y_0)y' \leq ((x_1 - x_0)^2 + (y_1 - y_0)^2)^{1/2} (x'^2 + y'^2)^{1/2}$$

con igualdad sólo si $(x_1 - x_0)/x' = (y_1 - y_0)/y'$.

Integrando para t entre 0 y 1:

$$(x_1 - x_0)^2 + (y_1 - y_0)^2 \leq ((x_1 - x_0)^2 + (y_1 - y_0)^2)^{1/2} L,$$

es decir:

$$\text{Longitud de la secante } AB \leq L$$

con igualdad sólo si y'/x' es la pendiente de la secante AB , para todo t , o sea si es una recta.

Ejemplo 2.4 : Geodésica en una región plana.

Veamos ahora el caso de una curva que debe permanecer dentro de una región Ω del plano. Supondremos que la frontera de Ω es lisa (tiene tangente continua).

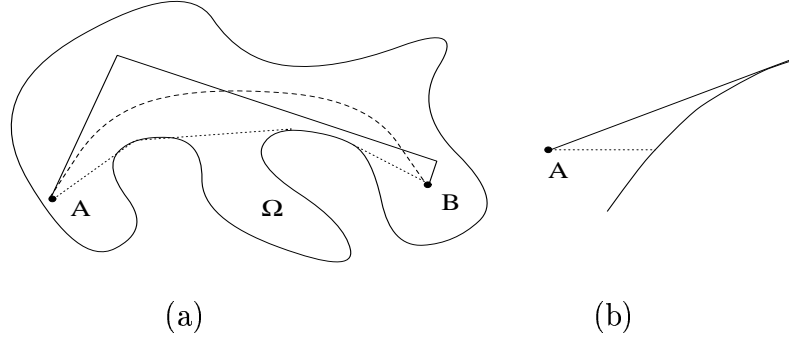


Fig. 2.4

Sea C la solución (si hay una). Claramente C debe consistir de segmentos de líneas rectas en Ω (por el ejemplo anterior) y pedazos de la frontera. Además es fácil ver que si un segmento toca la frontera, esto debe ser tangencialmente (ver la figura 2.4(b)). Por lo tanto hay sólo un número finito de posibilidades a partir de A y B . Es suficiente escoger el camino de menor longitud.

Ejemplo 2.5 : Geodésica en una esfera.

Sea $M(t)$ el camino sobre la esfera, t en $[0,1]$ (es decir $(\phi(t), \theta(t))$ un camino en $[0, 2\pi] \times [0, \pi]$ da $M(t)$ en coordenadas esféricas). Como las rotaciones conservan las distancias (y por lo tanto la longitud de los caminos) podemos suponer que el punto A es el polo norte.

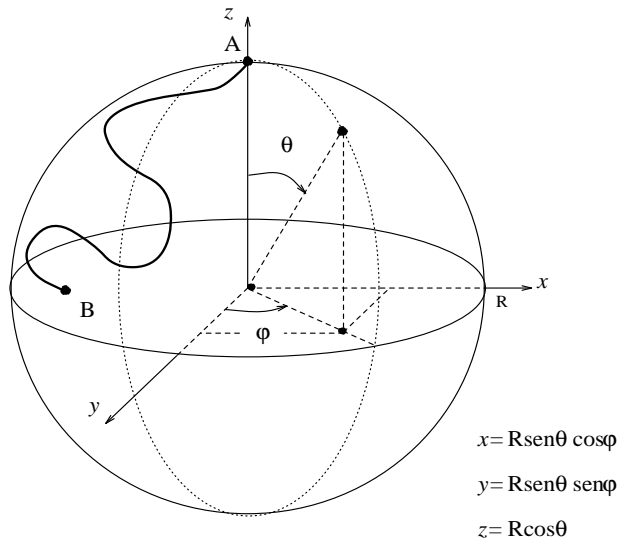


Fig. 2.5

$$L = \int_0^1 \|M'(t)\| dt = R \int_0^1 (\sin^2 \theta \phi'^2 + \theta'^2)^{1/2} dt.$$

Entonces,

$$L \geq R \int_0^1 \theta'(t) dt = R\theta(1).$$

Pero $R\theta(1)$ corresponde a la longitud de la curva

$$\phi(t) = \phi(B), \quad \theta(t) = \theta(B)t,$$

es decir, el círculo principal que une A a B en el mismo hemisferio.

Ejemplo 2.6 : Cómo llegar a la cima de una montaña en un tiempo mínimo..

Supongamos que tenemos una montaña dada por la curva $z = g(r)$, girada alrededor del eje z . La velocidad a la cual uno puede subir depende de la pendiente del camino, es decir $v = f(\alpha)$, donde α es el ángulo con la horizontal. ¿Cuál es el camino que hay que tomar para llegar a la cima lo antes posible?

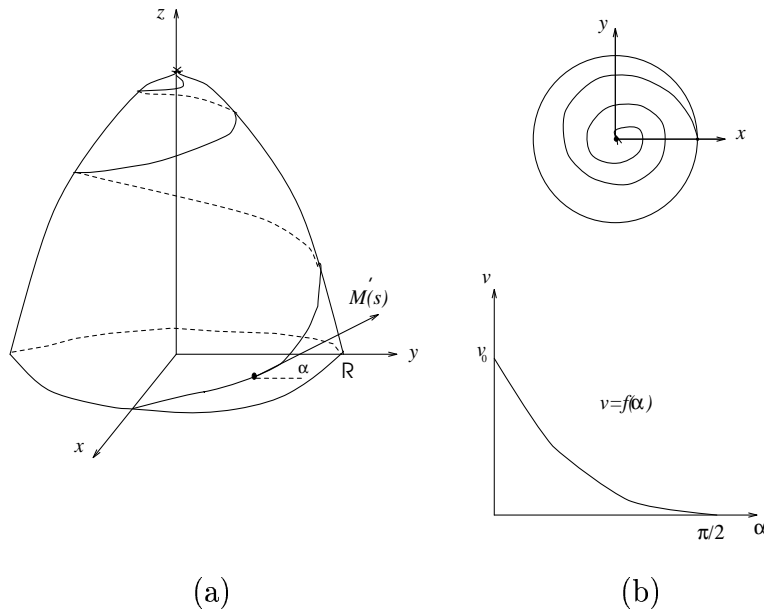


Fig. 2.6

Es fácil ver que el camino estará determinado por la función $\alpha(r)$. Además si s es la longitud de arco, entonces

$$\frac{dz}{ds} = \|M'(s)\| \sin \alpha = g'(r) \frac{dr}{ds}.$$

Como $\frac{ds}{dt} = f(\alpha)$, si $f(\alpha) > 0$, se puede invertir t en función de s y obtener:

$$T(\alpha) = \int_0^L \frac{ds}{f(\alpha)} = \int_0^R \frac{-g'(r)dr}{\operatorname{sen}\alpha f(\alpha)}.$$

Ahora bien, es razonable suponer que $f(\alpha)$ es una función decreciente y que $f(\pi/2) = 0$ (velocidad vertical en ese caso). Supondremos además que $g'(r) < 0$ (sólo subida). Como α debe ser menor que $\operatorname{Arctan}(-g'(r))$ (camino radial), para tener el mínimo de $T(\alpha)$ debemos escoger el camino de tal modo que, para cada r , $f(\alpha)\operatorname{sen}(\alpha)$ sea máximo dentro de los límites, para α entre 0 y $\operatorname{Arctan}(-g'(r))$. Por ejemplo, si la gráfica de $f(\alpha)\operatorname{sen}\alpha$ tiene un solo máximo en $\bar{\alpha}$ y $g'(r)$ es decreciente, tendremos que tomar el camino con $\alpha = \bar{\alpha}$ (hay dos posibilidades si pedimos que este camino sea diferenciable) y al llegar a la altura $g(r)$, tal que $\bar{\alpha} = \operatorname{Arctan}(-g'(r))$, seguir la vía directa hasta la cima (esto sucederá si la cima es plana, es decir $g'(0) = 0$).

1.3 Existencia, unicidad y otras preguntas

Los ejemplos anteriores son sólo un botón de muestra de la variedad de aplicaciones del Cálculo de Variaciones. Algunos eran tan simples que el Cálculo no era necesario. Sin embargo, otros no fueron resueltos o resueltos a medias. En esta sección queremos convencer al lector que no hay que hacer de lado las cuestiones de existencia, unicidad, convergencia, etc., bajo pretexto que lo que cuenta es la respuesta concreta al problema planteado. En efecto, una de las “dificultades” asociadas al Cálculo de Variaciones proviene de la falta de rigor en el planteamiento inicial (clase de funciones, tipo de mínimo, etc.), el no considerar el estudio preliminar de la existencia, unicidad, y (mucho más grave) confundir condiciones necesarias con suficientes.

Veamos algunos ejemplos de falta de existencia:

Ejemplo 3.1 : Encontrar la curva de menor longitud que sea ortogonal en A y B al segmento que une A a B .

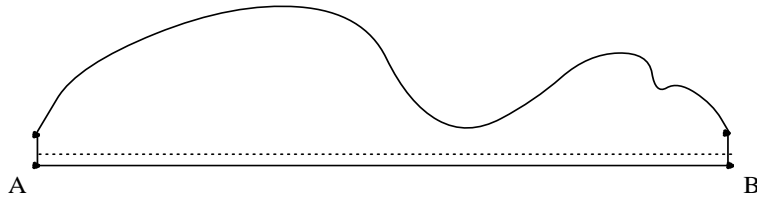


Fig. 3.1

Es evidente que la longitud es mayor que la longitud $|AB|$ del segmento AB , y $|AB|$ es el ínfimo de esas longitudes. Sin embargo, por el ejemplo 2.3, este ínfimo no puede ser un mínimo, es decir que la única curva que alcanza el ínfimo es el segmento, el cual no tiene la propiedad de ortogonalidad: No hay solución.

Ejemplo 3.2 : Encontrar la curva de menor longitud, que una A a B pasando por un punto C es fácil. Sin embargo, el análogo en tres dimensiones, encontrar la superficie de área mínima pasando por un círculo y un punto C que no está en el plano del círculo, no tiene solución.

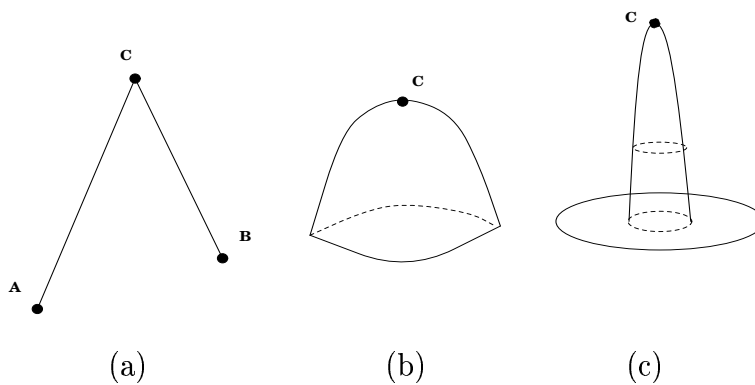


Fig.3.2

En efecto, el área de la superficie tiene como ínfimo el área del círculo (Fig. 3.2 (c)) esco-
giendo el pico de área muy pequeña. Sin embargo, el ínfimo (el círculo y una recta que une
 C al disco) no es una superficie.

Ejemplo 3.3 : Para las funciones $u(x)$ continuamente diferenciables por pedazos, con $u(-1) = -1$, $u(1) = 1$, encontrar el mínimo de

$$J(u) = \int_{-1}^1 x^2 u'(x)^2 dx.$$

$J(u) \geq 0$, y si hay un mínimo de J en la clase de funciones propuestas, entonces $J(u)$ es estrictamente positivo: si $J(0) = 0$, entonces $u' = 0$, u es constante por pedazos y no puede ser continua con esas condiciones en ± 1 . Sin embargo, el ínfimo de $J(u)$ es 0.

Sea

$$u_\varepsilon(x) = \begin{cases} -1 & \text{si } -1 \leq x \leq -\varepsilon \\ x/\varepsilon & \text{si } -\varepsilon \leq x \leq \varepsilon \\ 1 & \text{si } \varepsilon \leq x \leq 1 \end{cases}$$

Entonces $J(u_\varepsilon) = 2\varepsilon/3$ tiende a 0 si ε tiende a 0. No hay mínimo.

Ejemplo 3.4 : Para las funciones $u(x)$ continuamente diferenciables por pedazos, con $u(0) = 0$, $u(1) = 1$, encontrar el máximo y el mínimo de:

$$I(u) = \int_0^1 \frac{dx}{1 + u'^2}.$$

Claramente $0 < I(u) < 1$ y los extremos no son alcanzables por funciones en la clase considerada. Sin embargo, la función

$$u_\varepsilon(x) = \begin{cases} x/\varepsilon & \text{para } 0 \leq x \leq (1+\varepsilon)/2 \\ -x/\varepsilon + (1+\varepsilon)/\varepsilon & \text{para } (1+\varepsilon)/2 \leq x \leq 1 \end{cases}$$

tiene $I(u_\varepsilon) = 1/(1+\varepsilon^{-2})$, que tiende a 0 cuando ε tiende a 0. No hay mínimo.

Del mismo modo si

$$v_\varepsilon(x) = \begin{cases} x/\varepsilon & \text{para } 0 \leq x \leq \varepsilon \\ 1 & \text{para } \varepsilon \leq x \leq 1 \end{cases}$$

entonces $I(v_\varepsilon) = 1 - \varepsilon/(1+\varepsilon^2)$ tiende a 1 si ε tiende a 0. No hay máximo.

De hecho dada una función u en la clase, existen otras funciones v_ε y w_ε en la misma clase, con $I(v_\varepsilon)$ vecino a 0 e $I(w_\varepsilon)$ vecino a 1, que aproximan a la función u :

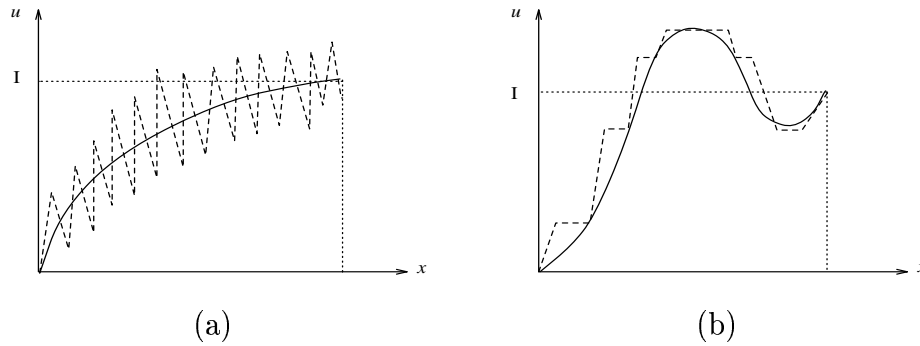


Fig.3.3

v_ε es la curva en dientes de sierra con pendiente $\pm 1/\varepsilon$ y $I(v_\varepsilon) = \varepsilon^2/(1+\varepsilon^2)$, w_ε es la curva en escalones de pendiente $\pm 1/\varepsilon$ y 0. Si u es monótona $I(w_\varepsilon) = 1 - \varepsilon/(1+\varepsilon^2)$. Dejamos al lector escribir en detalle estas funciones en el caso general. Notemos que se tomaron funciones con discontinuidades en las derivadas porque son más fáciles de escribir; redondeando los ángulos obtendríamos funciones lisas. Cuando ε tiende a 0, v_ε y w_ε tienden a u uniformemente (más adelante veremos que esta convergencia se llama uniforme o C^0), pero $I(u)$ no es continuo para la clase de funciones estudiadas provistas de la norma de convergencia uniforme.

Ejemplo 3.5 : Encontrar la curva simple y cerrada, que encierre un área mínima, tal que un segmento de longitud ℓ , pueda girar sobre él mismo dentro de la curva. Si la región encerrada por la curva es convexa, el lector podrá probar fácilmente que la curva que da el área mínima es un triángulo equilátero de altura ℓ .

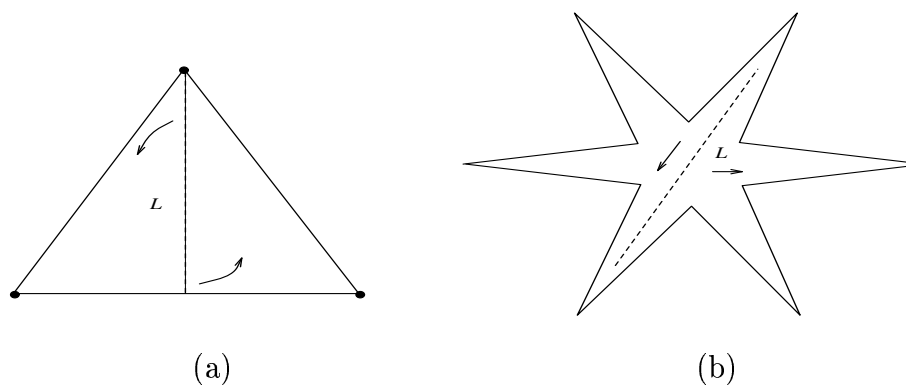


Fig. 3.4

Por otra parte si tomamos una “estrella” con n rayos de longitud ℓ y parte central de radio ε , el área será menor que

$$\pi\varepsilon^2 + n\frac{\pi\varepsilon\ell}{n} = \pi\varepsilon(\varepsilon + \ell),$$

(la longitud de cada arco es $2\pi\varepsilon/n$). Tomando n muy grande (según qué tan chico sea ε), el círculo interno será suficiente para girar un poco el segmento al meterlo en el rayo opuesto. Dejamos los detalles al lector. Esto implica que el ínfimo de las áreas es 0 y no es alcanzado.

Estos pocos ejemplos indican que, antes de abordar un problema de Cálculo de Variaciones, es sumamente importante plantearlo con la mayor precisión. En particular se debe definir la clase de funciones en la cual se estudiará el problema, es decir, la clase de trabajo. Ahí se tendrá que decir qué se entiende por convergencia de funciones, distancia entre funciones, es decir, cuál es la topología de trabajo. También se deberá estudiar la existencia (y unicidad si la hay) de soluciones al problema, en la clase de trabajo, o en otra, etc. Finalmente uno tiene que escoger la clase de trabajo, de tal modo que sea manejable técnicamente pero que sea la más cercana posible al contexto físico, geométrico u otro que dio origen al problema.

Al hacer este ejercicio previo, el usuario del Cálculo de Variaciones no tendrá sorpresas desagradables en el momento de interpretar sus resultados.

Capítulo 2

Cálculo en Espacio de Funciones

En este capítulo reuniremos los elementos básicos de Análisis Funcional que necesitaremos en el resto de estas notas. El lector que tenga familiaridad con este material puede omitirlo y sólo fijarse en la notación que usamos. Veremos en particular que casi todo el cálculo clásico de varias variables se extiende al caso de espacios de funciones, si se toman algunas precauciones al momento de aplicarlo. Definiremos primero los espacios de trabajo, estudiaremos después las funciones entre estos espacios, con sus propiedades de continuidad y de diferenciabilidad, terminando por las condiciones necesarias para tener un extremo.

2.1 Espacios de Funciones

Para un problema con n variables, el espacio de trabajo es naturalmente \mathbb{R}^n , la distancia entre dos puntos es la distancia euclidiana o cualquier otra equivalente a ella. Sin embargo, para un espacio de funciones hay varias maneras de definir la “distancia” entre dos funciones, y no son equivalentes. Tampoco hay un único espacio natural donde plantear todos los problemas, el espacio de trabajo dependerá del problema.

Por ejemplo, si se quiere estudiar el funcional

$$J(y) = \int_a^b F(x, y(x), y'(x)) dx,$$

queda claro que necesitamos que las funciones $y(x)$ tengan derivadas en algún sentido en el intervalo $[a, b]$, por ejemplo derivadas continuas, derivadas continuas por pedazos, o derivadas en casi todo punto y de cuadrado integrable (esto dependerá de la forma de F : continuidad, diferenciabilidad y crecimiento etc.), pero uno puede imaginar más clases de funciones para las cuales la expresión $J(y)$ tenga sentido.

Definición 1.1 : (Espacio lineal (sobre \mathbb{R} ó \mathbb{C})).

Sea X un conjunto con elementos, f, g, \dots , sobre el cual se define una adición y una multiplicación por un escalar (real o complejo) tal que:

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

Existe un elemento 0 tal que: $f + 0 = 0$.

Existe un elemento $-f$ tal que: $f + (-f) = 0$, para todo f en X .

Además $1 \cdot f = f$, $\alpha(\beta f) = (\alpha\beta)f$, $(\alpha + \beta)f = \alpha f + \beta f$,

$\alpha(f + g) = \alpha f + \alpha g$, donde 1 , α , β son escalares.

Nota importante. En estas notas sólo se considerarán espacios lineales reales.

Ejemplo 1.1 : Espacios lineales.

1. $X = \{\text{Funciones continuas en } [a, b]\} \equiv C^0[a, b]$
2. $X = \{\text{Funciones continuas en } [a, b] \text{ con } f(a) = f(b)\}.$
3. $X = \{u \in C^0[a, b], u(a) = u(b) = 0\}.$
4. $X \equiv C^\infty[a, b] = \{\text{funciones con derivadas de cualquier orden.}\}$
5. $X = C_0^\infty[a, b] = \{u \in C^\infty[a, b], u^{(k)}(a) = u^{(k)}(b) = 0, \forall k\}.$

Ejercicio 1.1 : Verificar que los ejemplos anteriores son espacios lineales.

En los espacios lineales, es importante definir una distancia entre dos elementos del espacio. Por ejemplo ¿ Cuándo se puede decir que dos funciones son vecinas?

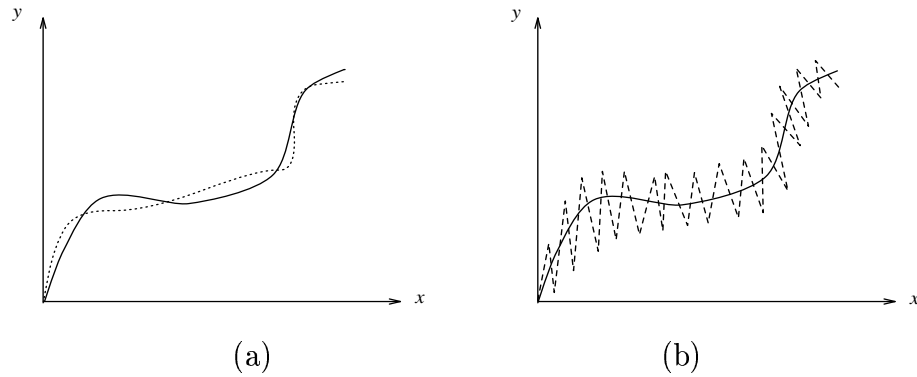


Fig. 1.1

En estas figuras tenemos dos maneras de aproximar la misma función: en valores de las funciones, parecen ser equivalentes, sin embargo, si consideramos las derivadas, las curvas de la figura 1.1 (b) no son vecinas.

Definición 1.2 : (Espacio normado).

Un espacio lineal X es normado si para todo f en X existe un número real único, no negativo $\|f\|$, con las propiedades :

$$\begin{aligned}\|f\| &= 0 \Leftrightarrow f = 0 \\ \|\alpha f\| &= |\alpha| \|f\| \\ \|f + g\| &\leq \|f\| + \|g\|.\end{aligned}$$

Se define la distancia entre f y g como $\|f - g\|$. A partir de esto podemos definir una topología (vecindades) sobre X .

Ejemplo 1.2 : Espacios normados.

1. $C^0[a, b]$ con $\|f\| = |f|_0 = \max_{x \in [a, b]} |f(x)|$.
Aquí si $|f - g|_0 \leq \varepsilon$, quiere decir $|f(x) - g(x)| \leq \varepsilon$ para todo x en $[a, b]$. Esta norma se llama de convergencia uniforme.
2. $C^1[a, b] = \{f : f \text{ es continuamente diferenciable en } [a, b]\}$, con
$$\|f\| = |f|_1 \equiv \max_{x \in [a, b]} |f(x)| + \max_{x \in [a, b]} |f'(x)|.$$
3. $C^0[a, b]$ con $\|f\| \equiv \|f\|_0 = (\int_a^b f(x)^2 dx)^{1/2}$.
4. $C^1[a, b]$ con $\|f\| \equiv \|f\|_1 = (\int_a^b (f(x)^2 + f'(x)^2) dx)^{1/2}$.
5. $L^2[a, b] = \{f(x) \text{ con } \int_a^b f(x)^2 dx < \infty\}$ y $\|f\| = \|f\|_0$, aquí se trata de integral de Lebesgue.
6. $H^1[a, b] = \{f(x) \text{ con } \|f\|_1 < \infty\}$,

Ejercicio 1.2 : Verificar que las normas propuestas satisfacen las propiedades de la definición 1.2.

Es importante especificar el espacio lineal y su norma. Cuando no lo hagamos será porque la norma es la natural para ese espacio. Por ejemplo $|f|_0$ es la norma natural para $C^0[a, b]$, $\|f\|_0$ para $L^2[a, b]$, etc.

En un espacio normado se puede hablar de sucesiones convergentes, como en cálculo de varias variables, y de sucesiones de Cauchy :

Definición 1.3 : Una sucesión $\{f_n\}$ de un espacio normado X es de Cauchy si dado $\varepsilon > 0$, existe $N(\varepsilon)$ tal que $\|f_n - f_m\| \leq \varepsilon$ si $n, m \geq N(\varepsilon)$.

Definición 1.4 : Un espacio normado es completo si toda sucesión de Cauchy converge a un elemento del espacio. En ese caso se dice que el espacio es un espacio de Banach.

Ejemplo 1.3 :

1. $C^0[a, b]$ con la norma $\|f\|_0$, $C^1[a, b]$ con la norma $\|f\|_1$ son espacios de Banach: si $\{f_n\}$ es una sucesión de Cauchy en $C^0[a, b]$, entonces $\{f_n(x)\}$ es una sucesión de Cauchy en \mathbb{R} (nuestros espacios son reales) y converge a una función $f(x)$; además el límite uniforme de funciones continuas es continuo.
2. $C^0[a, b]$ con la norma $\|f\|_0$, no es completo: la sucesión f_n de la figura 1.2 es de Cauchy, pero no converge a un elemento de $C^0[a, b]$.

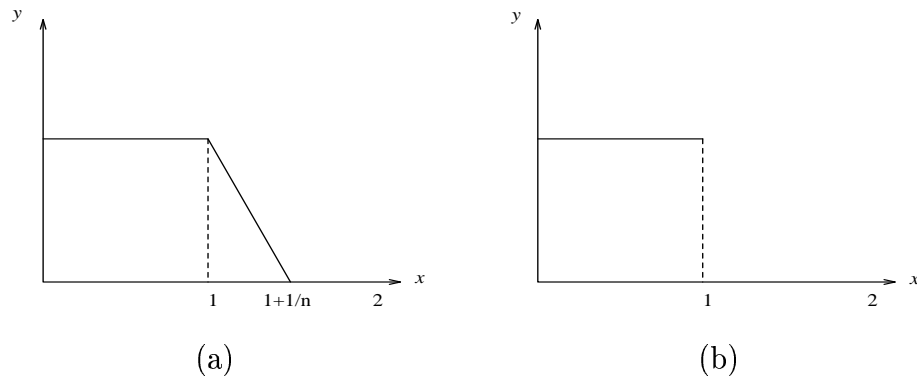


Fig. 1.2

Aquí $\|f_n - f_m\|_0^2 = \int_1^{1+1/n} |f_n(x) - f_m(x)|^2 dx \leq 1/n$, si $m \geq n$.

Si agregamos a un espacio normado todos los límites de sucesiones de Cauchy, se dice que se *completó* el espacio (del mismo modo que se completan los racionales para obtener los reales). Sobre el espacio de sucesiones de Cauchy de elementos del espacio, se define una relación de equivalencia: dos sucesiones de Cauchy $\{u_n\}$ y $\{v_n\}$ están en la misma clase si la sucesión de números reales $\|u_n - v_n\|$ converge a 0. Siempre converge a algo ya que

$$|||u_n - v_n|| - ||u_m - v_m||| \leq ||u_n - u_m|| + ||v_n - v_m||$$

implica que $\|u_n - v_n\|$ es una sucesión de Cauchy en \mathbb{R} . El espacio está contenido en ese espacio de sucesiones de Cauchy a través de sucesiones constantes. En muchos casos se puede identificar el espacio de sucesiones con algo más fácil de manejar. Por ejemplo, usando teoría de la medida, la “completación” de $C^0[a, b]$ bajo la norma $\|f\|_0$ es el espacio $L^2[a, b]$. En estas notas trabajaremos siempre en espacios completos fáciles de identificar. El lector interesado por el problema de completación puede consultar cualquier libro de análisis avanzado.

Usaremos finalmente otra estructura sobre espacios de funciones, la de producto escalar.

Definición 1.5 : Un producto escalar sobre un espacio lineal X es una operación bilineal (f, g) de $X \times X$ en el espacio de escalares (\mathbb{R} o \mathbb{C} , aquí solamente \mathbb{R}) tal que :

$$\begin{aligned}(f, g) &= \overline{(g, f)} \quad (\text{en particular } (f, f) \text{ es real}) \\ (f, f) &\geq 0 \quad \text{y} \quad (f, f) = 0 \quad \text{implica} \quad f = 0 \\ (\alpha f, g) &= \alpha(f, g) \\ (f + g, h) &= (f, h) + (g, h)\end{aligned}$$

Ejercicio 1.3 :

1. Probar la desigualdad de Cauchy-Schwarz:

$$|(f, g)| \leq \|f\| \|g\|$$

(usar los mismos argumentos que en \mathbb{R} o en \mathbb{C}).

2. Probar que $\|f\| \equiv (f, f)^{1/2}$ tiene las propiedades de una norma.

Ejemplo 1.4 :

1. $L^2[a, b]$ es un espacio con producto escalar

$$(f, g) = \int_a^b f(x)g(x) dx.$$

2. $H^1[a, b]$ es un espacio con producto escalar

$$(f, g)_1 = \int_a^b (f(x)g(x) + f'(x)g'(x)) dx.$$

(Verificar las propiedades).

Si el espacio con producto escalar (f, g) y norma asociada $\|f\| = (f, f)^{1/2}$ es completo, se dice de Hilbert (si no, es pre-Hilbert). Por ejemplo $L^2[a, b]$, $H^1[a, b]$ son espacios de Hilbert. Completando $C^0[a, b]$ con la norma $\|f\|_0$ se obtiene el espacio de Hilbert $L^2[a, b]$.

Terminaremos esta primera sección con algunos elementos de la topología de espacios normados.

Definición 1.6 : 1) Si f es un elemento de un espacio normado X , entonces una vecindad elemental de f está dada por el conjunto

$$V_\varepsilon \equiv \{g \text{ en } X \text{ con } \|g - f\| < \varepsilon\}.$$

- 2) Un subconjunto A de X es abierto si para todo f en A , existe una vecindad V_ε de f contenida en A .
- 3) f es un punto de acumulación de A si existe una sucesión $\{f_n\}$ de elementos de A que converge a f .
- 4) La cerradura de A es la unión de A con el conjunto de los puntos de acumulación de A . A es cerrado si coincide con su cerradura, o si su complemento es abierto (Ejercicio: probarlo).
- 5) A es denso en X , si su cerradura es X .

Ejemplo 1.5 :

1. $C^0[a, b]$ es denso en $L^2[a, b]$ con la norma $\| \cdot \|_0$, por construcción.
2. Los polinomios son densos en $C^0[a, b]$ (con la norma $\| \cdot \|_0$): esto es el teorema de Stone-Weierstrass.
3. $C_0^\infty[a, b]$ es denso en $L^2[a, b]$. (Verificar esto como ejercicio).

2.2 Continuidad

Sean X y Y dos espacios normados, en esta sección veremos algunos resultados simples sobre funciones de X a Y . (Aquí consideramos espacios reales).

Definición 2.1 : Sea J una función de X a Y . J es continua en f si, para todo $\varepsilon > 0$, existe $\delta > 0$ tal que:

$$\|J(g) - J(f)\|_Y \leq \varepsilon \quad \text{si} \quad \|g - f\|_X \leq \delta.$$

Ejemplo 2.1 : Si la función $F(x, y, z)$ es C^1 sobre $[a, b] \times \mathbb{R} \times \mathbb{R}$, entonces el funcional

$$J(y) = \int_a^b F(x, y(x), y'(x)) dx$$

del espacio $X = C^1[a, b]$, con norma $|y|_1$, es continuo en cualquier elemento y_0 de X ; de hecho, para cada x ,

$$\begin{aligned} F(x, y(x), y'(x)) &- F(x, y_0(x), y'_0(x)) = \\ &= F_y(x, \theta, \phi)(y(x) - y_0(x)) + F'_y(x, \theta, \phi)(y'(x) - y'_0(x)) \end{aligned}$$

por el teorema del valor intermedio (el dominio de diferenciabilidad de F es convexo), donde θ está entre $y_0(x)$ y $y(x)$, ϕ entre $y'_0(x)$ y $y'(x)$.

Si tomamos una primera vecindad de y_0 en $C^1[a, b]$, consistente de todas las funciones $y(x)$ con $|y - y_0|_1 \leq \delta$, dando una banda de ancho δ alrededor de $y_0(x)$ y de $y'_0(x)$, entonces (x, θ, ϕ) estarán acotadas ($[a, b]$ es compacto) y por lo tanto

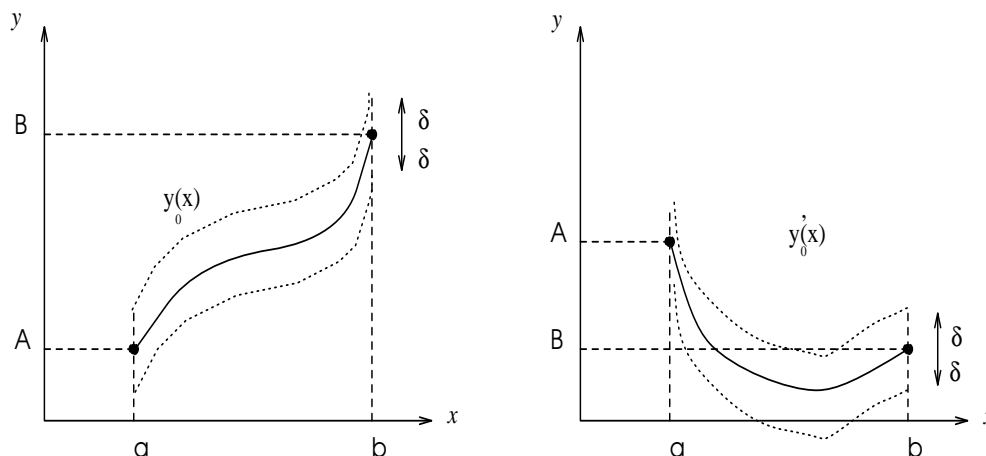


Fig. 2.1

$$|F_y|, |F_{y'}| \leq M \quad \text{para } (x, \theta, \phi) \text{ en esa banda.}$$

Tendremos entonces:

$$|J(y) - J(y_0)| \leq M|b - a||y - y_0|_1$$

si $|y - y_0|_1 \leq \delta$; el funcional es Lipschitz continuo. Nótese que $J(y)$ no es continuo en y_0 si sólo se da a $C^1[a, b]$, la norma $|\cdot|_0$. Ver el ejemplo I. 3.4.

Definición 2.2 : (Semicontinuidad de funcionales). . Sea $J(y) : X \rightarrow \mathbb{R}$ un funcional. Diremos que :

a) J es semicontinuo inferiormente en y_0 si dado $\varepsilon > 0$, existe $\delta > 0$, tal que:

$$J(y) - J(y_0) > -\varepsilon \quad \text{para } \|y - y_0\| \leq \delta.$$

b) J es semicontinuo superiormente en y_0 si dado $\varepsilon > 0$, existe $\delta > 0$, tal que:

$$J(y) - J(y_0) < \varepsilon \quad \text{para } \|y - y_0\| \leq \delta.$$

Nótese que si un funcional es semicontinuo para ambos lados entonces es continuo. Advertencia: Este material no será usado en estas notas. Se incluye sólo para dar una idea de algunas de las herramientas más avanzadas.

Ejercicio 2.1 : Sea $J(y) = \int_0^1 (1 + y'^2)^{1/2} dx$ definido sobre $C^1[0, 1]$.

a. Probar que $J(y)$ es continuo si la norma es $|\cdot|_1$.

b. Probar que $J(y)$ es semicontinuo inferiormente si la norma es $|\cdot|_0$: para ésto aproximar $J(y_0)$ por la longitud de una curva poligonal y estimar que tan chica puede ser la longitud de una curva vecina a $y(x)$.

c. Probar que $J(y)$ no es semicontinua superiormente con la norma $|\cdot|_0$.

Proposición 2.1 :

- 1) Un funcional semicontinuo inferiormente sobre un compacto alcanza su mínimo.
- 2) Un funcional semicontinuo superiormente sobre un compacto alcanza su máximo.

Prueba. Sea $J(f)$ el funcional y K el compacto. Sea $\varepsilon > 0$. Para cada f_0 de K , hay una vecindad de f_0 tal que si f pertenece a esa vecindad, entonces $J(f) \geq J(f_0) - \varepsilon$. Usando la compacidad de K , se ve que J es acotado inferiormente sobre K . Si $\{f_n\}$ es una sucesión minimizante, entonces hay una subsucesión convergente (por la compacidad) a algún f_0 . Tendremos, por la semicontinuidad en f_0 y la sucesión minimizante, para n_j suficientemente grande:

$$J(f_0) - \varepsilon \leq J(f_{n_j}) \leq \inf_K J(f) + \varepsilon.$$

Por lo tanto

$$\inf_K J(f) \leq J(f_0) \leq \inf_K J(f) + 2\varepsilon$$

para cualquier $\varepsilon : J(f_0) = \inf_K J(f)$.

Para (2) basta considerar $-J$.

Q.E.D.

Este resultado bien conocido en Cálculo (por lo menos conjuntando las dos semicontinuidades) no es inmediatamente aplicable en espacios de funciones: la razón es que si en dimensión finita los conjuntos compactos son abundantes (cualquier conjunto acotado y cerrado), esto no es el caso en dimensión infinita. Por ejemplo el conjunto $\{sen n\pi x\}$ es acotado en $C^0[0, 1]$, pero si tuviese un punto límite $u(x)$, entonces deberíamos tener:

$$\begin{aligned} |u|_0 &= \lim_{n \rightarrow \infty} |sen n\pi x|_0 = 1 \\ \int_0^1 u^2(x) dx &= \lim_{n \rightarrow \infty} \int_0^1 sen n\pi x sen(n+1)\pi x dx = 0 \end{aligned}$$

(recordar que la convergencia es uniforme). De hecho si la bola unitaria es compacta en un espacio de Banach, entonces el espacio es de dimensión finita.

Sin embargo, este resultado puede ser modificado para que pueda aplicarse en espacios de funciones. La compacidad se usó en dos puntos: para probar que el funcional es acotado por abajo y para probar que la sucesión minimizante tenía una subsucesión convergente. El primer punto puede obviarse pidiendo que el funcional mande conjuntos acotados en conjuntos acotados y que tienda a $+\infty$ si $\|f\|$ tiende a ∞ (hipótesis de coercividad); esto es fácil comprobar en la práctica. Para el segundo punto se pide que el funcional sea semicontinuo

en la topología débil: para un espacio de Hilbert toda sucesión acotada $\{f_n\}$ tiene la propiedad que tiene una subsucesión $\{f_{n_j}\}$ tal que (f_{n_j}, g) converge a (f_0, g) para todo g en el espacio (esto se llama convergencia débil); en este caso se pide que $J(f_{n_j}) \geq J(f_0) - \varepsilon$, para n_j grande.

Este tipo de resultados permiten probar existencia de mínimos en muchas aplicaciones. Sin embargo, requiere mucho más Análisis Funcional de la que deseamos emplear en estas notas.

Terminaremos esta sección con el equivalente de matrices:

Definición 2.3 : Un operador lineal (funcional lineal si el rango es \mathbb{R}) entre un espacio X y un espacio Y es una función $J(f)$ con las propiedades:

$$\begin{aligned} J(\alpha f) &= \alpha J(f), \forall \alpha, \forall f \\ J(f + g) &= J(f) + J(g), \quad \forall f, \forall g. \end{aligned}$$

Proposición 2.2 : Sea J un operador lineal definido sobre todo X , con valores en Y , espacios normados; entonces, las siguientes propiedades son equivalentes:

- 1) J es continuo
- 2) J es continuo en 0.
- 3) $\|J\| \equiv \sup_{\|f\|=1} \|J(f)\| < \infty$.

Prueba. Si J es continuo en 0, entonces, dado ε , existe $\delta > 0$, tal que

$$\|J(f) - J(0)\| \leq \varepsilon \quad \text{si } \|f\| \leq \delta.$$

Como $J(0) = 0$, si $\|f\| = 1$, tenemos:

$$\|J(f)\| = \|J(\delta f)\|/\delta \leq \varepsilon/\delta$$

(usando la homogeneidad de J) y por lo tanto

$$\|J\| \leq \varepsilon/\delta.$$

Así, (2) implica (3).

Por otra parte, sea f_0 en X y sea $\varepsilon > 0$. Sea f , diferente de f_0 , entonces:

$$\|J(f) - J(f_0)\| = \|J(f - f_0)\| = \|f - f_0\| \|J((f - f_0)/(f - f_0))\| \leq \|J\| \|f - f_0\|.$$

Por lo tanto si $\|f - f_0\| \leq \varepsilon/\|J\|$, tendremos el resultado.

Notemos que

$$\sup_{\|f\| \leq 1} \|J(f)\| = \sup_{\|f\|=1} \|J(f)\|$$

ya que el lado derecho no es más grande que el lado izquierdo (conjunto más grande) y que si $f \neq 0$, $\|f\| \leq 1$

$$\|J(f)\| = \|f\| \|J(f/\|f\|)\| \leq \|f\| \sup_{\|g\|=1} \|J(g)\| \leq \sup_{\|g\|=1} \|J(g)\|.$$

Q.E.D.

Ejemplo 2.2 :

1. Los siguientes funcionales son lineales y continuos sobre $C^0[a, b]$, con norma $\|\cdot\|_0$:

a. $J(f) = f(a)$.

b. $J(f) = \int_a^b f(x) dx$.

c. $J(f) = \int_a^b \alpha(x) f(x) dx$ con $|\alpha(x)|$ integrable en $[a, b]$.

2. Sobre $C^n[a, b]$, con norma $\|f\|_n = \sum_0^n \max |f^{(k)}(x)|$, el funcional :

$$J(f) = \int_a^b \sum_0^n \alpha_k(x) f^{(k)}(x) dx$$

es lineal y continuo, si $|\alpha_k(x)|$ es integrable en $[a, b]$.

2.3 Derivadas

El éxito del Cálculo viene en gran parte del hecho que en un extremo de una función la derivada (o las derivadas parciales) vale 0. La misma situación se presenta en el Cálculo de Variaciones al tratar de localizar los posibles candidatos a extremos de un funcional. Para ésto definiremos las nociones correspondientes a derivadas direccionales y a la diferencial.

Definición 3.1 : (Derivada de Gâteaux de un funcional).

Sea J un funcional definido sobre un espacio normado X . Diremos que J tiene una derivada de Gâteaux en f_0 en la dirección h si:

1) $J(f_0 + th)$ está definido para $|t|$ pequeño.

2) $\lim_{t \rightarrow 0} \frac{J(f_0 + th) - J(f_0)}{t} = \frac{d}{dt}(J(f_0 + th)) \big|_{t=0}$ existe.

En este caso se denota el límite como $D_h J(f_0)$.

Notemos que $J(f_0 + th)$ es una función de una sola variable t .

Ejemplo 3.1 : La Braquistocrona.

Según el ejemplo I.2.1 el funcional es

$$J(f) = \int_0^a ((1 + f'(x)^2)/2g f(x))^{1/2} dx$$

con las condiciones de frontera $f(0) = 0$, $f(a) = A$. En particular la integral es impropia cada vez que $f(x) = 0$ ó $\|f'(x)\| = \infty$. Por ejemplo en el caso de la cicloide

$$x(\theta) = c(\theta - \operatorname{sen}\theta), \quad y(\theta) = c(1 - \cos\theta)$$

para θ cerca de 0, $x(\theta) \simeq c\theta^3/6$, $y(\theta) \simeq c\theta^2/2 \sim x^{2/3}$, $f'(x) = \frac{y}{x} \simeq 2/\theta \sim x^{-1/3}$, el integrando se comporta como $x^{-2/3}$ cerca de $x = 0$.

Asimismo, si $f(x) \sim x^\alpha$, para x pequeño, es necesario que $\alpha < 2$.

Por lo tanto, para que $J(f_0 + th)$ esté definido se necesitará que h sea 0 en los ceros de f_0 y lo sea de tal forma que preserve la integrabilidad. Por otra parte, para calcular la derivada, en $t = 0$, de $J(f_0 + th)$, nos gustaría poder intercambiar la integral con la derivada y derivar el integrando, es decir :

$$\left. \frac{d}{dt} \left(\frac{1 + (f'_0 + th')^2}{2g(f_0 + th)} \right)^{1/2} \right|_{t=0} = (2g)^{-1/2} \left(\frac{f'_0 h'}{f_0^{1/2} (1 + f_0'^2)^{1/2}} - \frac{(1 + f_0'^2)^{1/2} h}{2f_0^{3/2}} \right).$$

El intercambio es válido si la integral de esta última expresión es uniformemente convergente. Por ejemplo si f_0 se comporta como x^α , para x pequeño, y h como x^β , será necesario tomar $\beta > 1 - \alpha/2$, si $1 \leq \alpha < 2$, y $\beta > \alpha/2$, si $0 < \alpha \leq 1$, para tener esa convergencia. Esto será el caso si se toma h en $C_0^\infty[0, a]$, es decir si todas las derivadas de h valen 0 en 0 y en a .

Ejemplo 3.2 : Sea $F(x, y, z)$ una función continua, con derivadas parciales F_y, F_z continuas en $[a, b] \times \mathbb{R}^2$. Sea

$$J(y) = \int_a^b F(x, y(x), y'(x)) dx$$

definido sobre un subconjunto de funciones en $C^1[a, b]$.

Entonces

$$D_h J(y) = \int_a^b (F_y(x, y(x), y'(x))h(x) + F_z(x, y(x), y'(x))h'(x)) dx$$

ya que aquí no hay problemas de convergencia de las integrales y que

$$\left. \frac{d}{dt} (F(x, y + th, y' + th')) \right|_{t=0} = F_y(x, y, y')h + F_z(x, y, y')h'.$$

Del mismo modo que en Cálculo de varias variables se definen derivadas direccionales (en particular parciales) y diferenciales, en el Cálculo de Variaciones, existe una diferencial.

Definición 3.2 : (Derivada de Fréchet de un funcional).

El funcional $J(y)$ tiene una derivada de Fréchet en y_0 , si existe un funcional lineal continuo, $DJ(y_0)$, tal que

$$J(y_0 + h) = J(y_0) + DJ(y_0)h + o(\|h\|)$$

para todo h en el espacio normado X .

Recordamos que $o(\|h\|)$ quiere decir que:

$$\lim_{\|h\| \rightarrow 0} \frac{o(\|h\|)}{\|h\|} = 0.$$

Ejemplo 3.3 : Sea $F(x, y, z)$ con F_y, F_z continuas, y sea

$$J(y) = \int_a^b F(x, y(x), y'(x)) dx$$

como en el ejemplo anterior, sobre un subespacio de $C^1[a, b]$, con la norma $\|\cdot\|_1$. Podemos entonces escribir:

$$\begin{aligned} J(y+h) - J(y) &= \int_a^b (F_y(x, y, y')h + F_z(x, y, y')h') dx + \\ &+ \int_a^b \left[\int_0^1 (F_y(x, y + \tau h, y' + \tau h') - F_y(x, y, y')) h d\tau \right] dx \\ &+ \int_a^b \left[\int_0^1 (F_z(x, y + \tau h, y' + \tau h') - F_z(x, y, y')) h' d\tau \right] dx. \end{aligned}$$

Como $[a, b]$ es compacto, F_y, F_z son uniformemente continuas en $(x, y(x), y'(x))$ para x en $[a, b]$. Por lo tanto F_y, F_z son acotadas por una constante M , para todo x en $[a, b]$, y las diferencias son de orden $O(|h| + |h'|)$ uniformemente en $[a, b]$.

Si denotamos por:

$$DJ(y)h = \int_a^b (F_y(x, y, y')h + F_z(x, y, y')h') dx,$$

entonces $|DJ(y)h| \leq M|b-a||h|_1$ y el residuo es acotado por $O(|h|_1)|h|_1 = o(|h|_1)$.

Ejemplo 3.4 : Sea $F(x, y_1, z_1, y_2, z_2)$ continua en $[a, b] \times \mathbb{R}^4$, con $F_{y_1}, F_{z_1}, F_{y_2}, F_{z_2}$ continuas en el mismo dominio, entonces

$$J(y_1, y_2) = \int_a^b F(x, y_1(x), y_1'(x), y_2(x), y_2'(x)) dx$$

tiene como derivada de Fréchet, sobre $C^1[a, b] \times C^1[a, b]$, con la norma $\|(y_1, y_2)\| = |y_1|_1 + |y_2|_1$:

$$DJ(y_1, y_2)(h_1, h_2) = \int_a^b (F_{y_1}h_1 + F_{z_1}h_1' + F_{y_2}h_2 + F_{z_2}h_2') dx$$

donde los argumentos en $F_{y_1}, F_{z_1}, F_{y_2}, F_{z_2}$ son $(x, y_1(x), y_1'(x), y_2(x), y_2'(x))$.

Dejamos como ejercicio verificar esto.

Ejemplo 3.5 : Sea $F(x, y, z, u, v)$ continua con F_z, F_u, F_v continuas sobre $\bar{\Omega} \times \mathbb{R}^3$, donde Ω es un dominio acotado del plano. Entonces, para una función $f(x, y)$ en $C^1(\bar{\Omega})$, con norma

$$||f|| = \max_{\Omega} |f(x, y)| + \max_{\Omega} |f_x(x, y)| + \max_{\Omega} |f_y(x, y)|$$

el funcional,

$$J(f) = \iint_{\Omega} F(x, y, f(x, y), f_x(x, y), f_y(x, y)) \, dx dy$$

tiene como derivada de Fréchet:

$$DJ(f)h = \iint_{\Omega} (F_z h + F_u h_x + F_v h_y) \, dx dy$$

donde los argumentos de F_z, F_u, F_v son $(x, y, f(x, y), f_x(x, y), f_y(x, y))$.
Dejamos también como ejercicio el verificar esto.

Proposición 3.1 : (Unicidad de la derivada de Fréchet).

Si existe un funcional lineal continuo, ϕ , tal que la aproximación lineal siguiente es válida:

$$J(y_0 + h) = J(y_0) + \phi h + o(||h||)$$

entonces $\phi h = DJ(y_0)h$ es único.

Prueba. Si hay dos funcionales ϕ_1, ϕ_2 que realizan la aproximación, entonces:

$$\phi_1 h - \phi_2 h = o_2(||h||) - o_1(||h||) = o(||h||).$$

Por lo tanto

$$|(\phi_1 - \phi_2)(h/||h||)| = o(||h||)/||h|| = O(||h||).$$

Por la proposición 2.2 : $||\phi_1 - \phi_2|| = 0$, es decir $\phi_1 = \phi_2$.

Q.E.D.

Proposición 3.2 : *Si un funcional J es Fréchet-diferenciable en y_0 , en un espacio normado X , entonces J tiene derivada de Gâteaux en y_0 en toda dirección h de X , con*

$$D_h J(y_0) = DJ(y_0)h.$$

Prueba. $J(y_0 + th) - J(y_0) = DJ(y_0)th + o(th)$

Q.E.D.

Proposición 3.3 : *Si un funcional J tiene derivada de Gâteaux en y para y en una vecindad de y_0 , en toda dirección h del espacio normado X , y si $D_h J(y_0)$ es un funcional lineal continuo en h y $D_h J(y)$ es continuo en y_0 , es decir*

$$||D_h J(y) - D_h J(y_0)|| \leq O(||y - y_0||)||h||,$$

entonces J tiene derivada de Fréchet, con $DJ(y_0)h = D_h J(y_0)$.

Prueba. Para $\|y - y_0\| \leq \delta$, la vecindad de y_0 , podemos escribir $y = y_0 + t_0 h$, con $0 \leq t_0 \leq \delta$, $\|h\| = 1$. Consideremos la función de una variable $f(t) = J(y_0 + th)$. Por hipótesis $f(t)$ tiene derivada

$$f'(t) = D_h J(y_0 + th)$$

para t entre $-\delta$ y δ . Además

$$|f'(t_1) - f'(t_2)| \leq O(|t_1 - t_2| \|h\|) \|h\| = O(|t_1 - t_2|)$$

es decir $f'(t)$ es continua para $|t| \leq \delta$.

Por el teorema del valor medio (en una variable), existe θ , $0 < \theta < t_0$, tal que: $f(t_0) = f(0) + f'(\theta)t_0$, es decir:

$$J(y) = J(y_0) + t_0 D_h J(y_0 + \theta h) = J(y_0) + t_0 D_h J(y_0) + t_0 (D_h J(y_0 + \theta h) - D_h J(y_0)).$$

Usando la linealidad, en h , de $D_h J(y_0)$, tenemos

$$J(y) = J(y_0) + D_{y-y_0} J(y_0) + O(\theta)t_0$$

donde $O(\theta)$ no depende de la dirección h , con la propiedad que $O(\theta)$ tiende a 0 si t_0 tiende a 0.

Q.E.D.

Como se puede notar, estas tres propiedades de las derivadas son similares a las que se encuentran en el Cálculo de varias variables. Es importante destacar que para la derivada de Gâteaux no se necesita la topología del espacio (sólo la de \mathbb{R}) mientras que para la derivada de Fréchet ésta es esencial. Si esto puede verse como una ventaja (que será explotada para encontrar condiciones necesarias para los extremos) ya no lo es para las condiciones suficientes: en Cálculo de varias variables, las condiciones suficientes están basadas en la expansión de Taylor a segundo orden, y por lo tanto en la diferencial. Algo muy semejante pasará en el Cálculo de Variaciones.

2.4 Extremos de un funcional

Definición 4.1 : El funcional $J(y)$ tiene un máximo (mínimo) local en y_0 si $J(y) - J(y_0) \leq 0$ (≥ 0) para todo y vecino a y_0 , para el cual $J(y)$ esté definido.

Notemos que en esta definición la topología del espacio entra en la noción de vecindad: y puede ser vecino a y_0 en una norma pero no en otra (ver la Figura 1.1). Cuando la topología no está explícitamente definida esto es fuente de confusiones; en la última parte

de estas notas diferenciaremos los extremos en fuertes o débiles según la norma empleada en el mismo espacio lineal.

Teorema 4.1 : *Si $J(y)$ tiene un extremo local en y_0 , entonces $D_h J(y_0) = 0$ en todas las direcciones h donde la derivada de Gâteaux esté definida. En particular si J tiene una derivada de Fréchet en y_0 , entonces $DJ(y_0) \equiv 0$.*

Prueba. Sea h una dirección admisible, entonces la función

$$f(t) = J(y_0 + th)$$

está definida para t pequeño y tiene un extremo en $t = 0$, por lo tanto $f'(0) = 0$.

Q.E.D.

Es muy importante tener en mente el hecho que este resultado sólo es una *condición necesaria*: en cada dirección h uno puede tener un extremo o un punto de inflexión. Al considerar varias direcciones uno puede obtener un punto silla. Por otra parte, en un espacio de funciones hay un número infinito de direcciones linealmente independientes. Por lo tanto esta condición dista mucho de ser suficiente.

Varios de los problemas que hemos planteado en el primer capítulo son condicionados por un funcional: por ejemplo en los problemas isoperimétricos. El lector ya habrá pensado que estos problemas deben ser tratados por el método de multiplicadores de Lagrange o algo semejante. Recordaremos primero el resultado en dimensión finita.

Teorema 4.2 : (Multiplicadores de Lagrange en \mathbb{R}^n).

Sean $f(X), g_1(X), \dots, g_m(X)$, funciones C^1 en una vecindad de X_0 en \mathbb{R}^n , a valores en \mathbb{R} . Si $f(X)$ tiene un extremo local en X_0 condicionado por $g_i(X) = 0$, $i = 1, \dots, m$, entonces existen $m + 1$ constantes $\nu, \lambda_1, \dots, \lambda_m$ tales que:

$$\nu \nabla f(X_0) + \sum_{i=1}^m \lambda_i \nabla g_i(X_0) = 0$$

con $|\nu| + |\lambda_1| + \dots + |\lambda_m| \neq 0$.

Notemos que no se excluye la opción $\nu = 0$: en ese caso los gradientes de g_i en X_0 son linealmente dependientes; esto será el caso si $m > n$. En cada situación concreta hay que estudiar esa posibilidad. Recordemos que este resultado se demuestra suponiendo que los vectores $\nabla g_i(X_0)$ son linealmente independientes, lo cual permite, a través del teorema de la función implícita, despejar m de las variables en función de las otras, usando las condiciones. Después se usa la regla de la cadena sobre las variables libres para dar la condición necesaria de un extremo. Geométricamente, para $n = 2$, $m = 1$, $g(x, y) = 0$ (con $\nabla g(x_0, y_0) \neq 0$) genera una curva regular en el plano, con normal ∇g . Representando la superficie $z = f(x, y)$ por sus curvas de nivel, es fácil ver que si el camino, $g(x, y) = 0$, atraviesa una curva de nivel de f , no hay posibilidad de un extremo. Los extremos estarán en los puntos donde el camino es tangente a la curva de nivel y de un solo lado de ella, o si f tiene un punto crítico (por ejemplo en el paso de una montaña, en ese caso $\nabla f = 0$). En el primer caso

las normales al camino y a la curva de nivel deben ser proporcionales. Para cubrir todos los casos, incluyendo $\nabla g = 0$, obtenemos la fórmula, con $|\nu| + |\lambda| \neq 0$:

$$\nu \nabla f(X_0) + \lambda \nabla g(X_0) = 0.$$

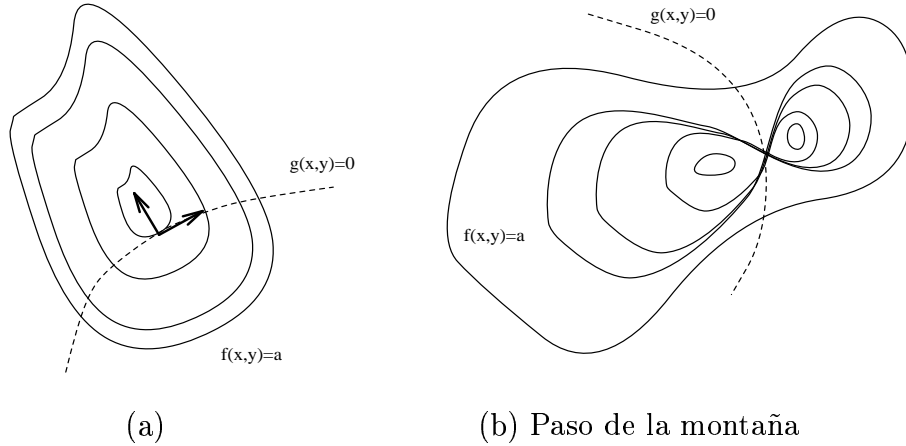


Fig. 4.1

Usaremos este resultado para probar el teorema de multiplicadores en un espacio normado X .

Teorema 4.3 : (Multiplicadores de Lagrange en X).

Sean $J(y), G_1(y), \dots, G_m(y)$ funcionales (a valores en \mathbb{R}) definidos en una vecindad, en X , de y_0 , con derivadas de Gâteaux $D_h J(y), D_h G_i(y)$, en esa vecindad, para toda dirección h en X . Suponemos que cada una de esas derivadas es continua en y para cada h fijo. Entonces, si J tiene un extremo local en y_0 , condicionado por $G_i(y) = 0$, existen constantes, $\nu, \lambda_1, \dots, \lambda_m$, con $|\nu| + |\lambda_1| + \dots + |\lambda_m| \neq 0$, tales que:

$$\nu D_h J(y_0) + \sum_{i=1}^m \lambda_i D_h G_i(y_0) = 0$$

para todo h en X .

En particular si los funcionales tienen derivadas de Fréchet, continuas en y , tendremos:

$$\nu DJ(y_0) + \sum_{i=1}^m \lambda_i DG_i(y_0) = 0.$$

Prueba. Sean h_1, \dots, h_n , n elementos de X , con $\|h_i\| = 1$, consideremos las funciones:

$$\begin{aligned} f(x_1, \dots, x_n) &= J(y_0 + x_1 h_1 + \dots + x_n h_n) \\ g_i(x_1, \dots, x_n) &= G_i(y_0 + x_1 h_1 + \dots + x_n h_n). \end{aligned}$$

Claramente si x_1, \dots, x_n son pequeños, $y = y_0 + x_1 h_1 + \dots + x_n h_n$ pertenece a la vecindad de y_0 donde se encuentra el extremo condicionado. En particular f tiene un extremo en 0, condicionado por $g_i(x_1, \dots, x_n) = 0$, $i = 1, \dots, m$.

Por otra parte

$$f_{x_j}(x_1, \dots, x_n) = D_{h_j} J(y_0 + x_1 h_1 + \dots + x_n h_n)$$

$$g_{ix_j}(x_1, \dots, x_n) = D_{h_j} G_i(y_0 + x_1 h_1 + \dots + x_n h_n)$$

son continuas en x_1, \dots, x_n . Por lo tanto (resultado de Cálculo de varias variables), f y g_i son continuamente diferenciables en una vecindad de 0, y se puede aplicar el teorema 4.2 : existen constantes $\nu, \lambda_1, \dots, \lambda_m$ (*a priori* dependientes de h_1, \dots, h_m) tales que:

$$\nu \nabla f(0) + \sum_{i=1}^m \lambda_i \nabla g_i(0) = 0.$$

Es decir

$$\nu D_{h_j} J(y_0) + \sum_{i=1}^m \lambda_i D_{h_j} G_i(y_0) = 0 \quad \text{para } j = 1, \dots, n.$$

Ahora bien, si tomamos $n = m + 1$, con $h_{m+1} = h$ cualquiera, y si consideramos la matriz:

$$\begin{pmatrix} D_h J(y_0) & D_{h_1} J(y_0) & \dots & D_{h_m} J(y_0) \\ D_h G_1(y_0) & D_{h_1} G_1(y_0) & \dots & D_{h_m} G_1(y_0) \\ \dots & \dots & \dots & \dots \\ D_h G_m(y_0) & D_{h_1} G_m(y_0) & \dots & D_{h_m} G_m(y_0) \end{pmatrix}$$

la relación anterior dice que los renglones son linealmente dependientes, por lo tanto el determinante de esta matriz es 0; desarrollando el determinante con respecto a la primera columna, tendremos:

$$\nu D_h J(y_0) + \sum_{i=1}^m \lambda_i D_h G_i(y_0) = 0$$

donde ν es el cofactor de $D_h J(y_0)$, λ_i el cofactor de $D_h G_i(y_0)$: ν y λ_i no dependen por lo tanto de h , sólo de h_1, \dots, h_m .

El teorema queda probado si alguno de esos cofactores es diferente de 0, para algún m -tuplo (h_1, \dots, h_m) . De lo contrario, tendremos que si escogemos $m - 1$ direcciones h_1, \dots, h_{m-1} , entonces el determinante de la matriz :

$$\begin{pmatrix} D_h G_1(y_0) & D_{h_1} G_1(y_0) & \dots & D_{h_{m-1}} G_1(y_0) \\ D_h G_2(y_0) & D_{h_1} G_2(y_0) & \dots & D_{h_{m-1}} G_2(y_0) \\ \dots & \dots & \dots & \dots \\ D_h G_m(y_0) & D_{h_1} G_m(y_0) & \dots & D_{h_{m-1}} G_m(y_0) \end{pmatrix}$$

debe ser siempre 0, para todo h (cada h da un $\nu = 0$). Desarrollando el determinante con respecto a la primera columna, tendremos

$$\sum_{i=1}^m \mu_i D_h G_i(y_0) = 0$$

con μ_i el cofactor de $D_h G_i(y_0)$, es independiente de h . De nuevo si todos los μ_i posibles, para todos los $(m-1)$ -tuplos (h_1, \dots, h_{m-1}) son 0, podemos repetir el argumento eliminando h_{m-1} y G_1 (por ejemplo). Si en ningún momento tenemos un cofactor diferente de 0, llegaremos a la situación con dos funcionales : la siguiente matriz tendría determinante 0, para todo h :

$$\begin{pmatrix} D_h G_{m-1}(y_0) & D_{h_1} G_{m-1}(y_0) \\ D_h G_m(y_0) & D_{h_1} G_m(y_0) \end{pmatrix}$$

Si para ningún h_1 los cofactores son diferentes de 0, entonces se tiene $D_h G_m(y_0) = 0$ y $D_{h_1} G_{m-1}(y_0) = 0$, pero entonces hubiéramos podido tomar $\nu = 0$, $\lambda_i = 0$, $i = 1, \dots, m-2$, λ_{m-1} , λ_m arbitrarios.

Q.E.D.

Ejemplo 4.1 : Problema isoperimétrico.

Si pensamos en el problema, del ejemplo I.2.2, de la curva de longitud dada que encierra una área máxima, el funcional de área es:

$$A(x, y) = \int_0^{2\pi} (xy' - yx')/2 dt$$

con la condición

$$G(x, y) = \int_0^{2\pi} (x'^2 + y'^2)^{1/2} dt - L = 0.$$

El espacio de trabajo es el conjunto de pares de funciones (x, y) en $C^1[0, 2\pi]$ (con la norma $\|\cdot\|_1$), y la propiedad de ser periódicas: $x(0) = x(2\pi)$, $y(0) = y(2\pi)$.

Como los integrandos son C^1 (si $x'^2 + y'^2 \neq 0$) podemos calcular las derivadas de Fréchet de los funcionales, usando la fórmula del ejemplo 3.2 (supondremos que el candidato a máximo es un par de funciones (x, y) regulares); tendremos entonces la relación:

$$\nu \int_0^{2\pi} ((y'h - yh') + (xk' - x'k))/2 dt + \lambda \int_0^{2\pi} (x'h' + y'k')(x'^2 + y'^2)^{-1/2} dt = 0$$

para todo (h, k) en el espacio (es decir en $C^1[0, 2\pi]$, con mismo valor en los extremos).

Ejemplo 4.2 : Si tenemos que encontrar extremos del funcional

$$J(y) = \int_a^b F(x, y, y') dx$$

con F en C^1 , y las condiciones $y(a) = A$, $y(b) = B$, $y(x)$ en C^1 , podemos plantear el problema de dos maneras:

1) Si y_0 es el candidato a extremo, entonces todo y en $C^1[a, b]$ puede escribirse como $y = y_0 + h$.

Como $y(x)$ y $y_0(x)$ deben satisfacer las condiciones de frontera, tenemos $h(a) = 0$, $h(b) = 0$. Por lo tanto $J(y)$ puede escribirse

$$I(h) \equiv J(y_0 + h) = \int_a^b F(x, y_0 + h, y'_0 + h') dx$$

que se estudia en el espacio de las funciones $h(x)$ en $C^1[a, b]$, con $h(a) = h(b) = 0$ y norma $\|\cdot\|_1$ (es un espacio lineal). Debemos estudiar la derivada de Fréchet de I en 0:

$$DI(0)h = \int_a^b (F_y(x, y_0, y'_0)h + F_{y'}(x, y_0, y'_0)h') dx.$$

2) También podríamos trabajar en el espacio $C^1[a, b]$ y tratar las condiciones de frontera $y(a) - A = 0$, $y(b) - B = 0$, como dos funcionales condicionando al extremo: el funcional

$$G_a(y) = y(a) - A$$

es claramente Fréchet diferenciable con derivada:

$$DG_a(y_0)h = h(a).$$

El teorema de multiplicadores de Lagrange dará el siguiente resultado:

$$\nu \int_a^b (F_y h + F_{y'} h') dx + \lambda_a h(a) + \lambda_b h(b) = 0$$

para todo h en $C^1[a, b]$.

Ahora si $\nu = 0$, podemos tomar $h(x)$ con $h(a) = 0$, $h(b) = 1$, por lo tanto $\lambda_b = 0$. Escogiendo $h(x)$ con $h(a) = 1$, $h(b) = 0$, tendríamos $\lambda_a = 0$: las tres constantes serían 0. Por lo tanto $\nu \neq 0$ y, dividiendo la relación por ν , podemos tomarla como 1. Tomando $h(x)$ con $h(a) = h(b) = 0$, tendremos

$$\int_a^b (F_y h + F_{y'} h') dx = 0.$$

Veremos (en el capítulo siguiente) que si esto es cierto, entonces $F_{y'}$ tiene derivada y esa derivada vale F_y .

Podemos por lo tanto integrar por partes la integral y obtener:

$$\int_a^b (F_y - \frac{d}{dx}(F_{y'}))h dx + F_{y'}h|_a^b + \lambda_a h(a) + \lambda_b h(b) = 0$$

es decir:

$$(-F_{y'}(a, A, y'_0(a)) + \lambda_a)h(a) + (F_{y'}(b, B, y'_0(b)) + \lambda_b)h(b) = 0.$$

Como $h(a)$ y $h(b)$ son arbitrarios, tendremos los valores de λ_a y λ_b . Evidentemente los dos métodos son equivalentes y, a estas alturas, el segundo parece una complicación innecesaria. Sin embargo, en otros casos vale la pena esa complicación.

Hasta ahora los funcionales estudiados eran generales. En el Cálculo de Variaciones, la mayoría de ellos son de la forma integral, como en el ejemplo anterior. Para estos se puede pasar

de la relación $DJ(y_0)h = 0$ a una ecuación diferencial. Para poder hacerlo, necesitaremos algunos resultados de Cálculo que, dada su importancia, hemos agrupado en un capítulo aparte.

Capítulo 3

Lemas Fundamentales

En este breve capítulo se verá cómo pasar de la derivada de un funcional a una ecuación diferencial. Si el funcional

$$J(y) = \int_a^b F(x, y, y') dx$$

definido sobre $C^1[a, b]$, con F, F_y, F_z continuas, tiene un extremo en $y_0(x)$ condicionado por $y(a) = A, y(b) = B$, entonces $I(h) = J(y_0 + h)$ tiene un extremo en 0, para las funciones h en $C^1[a, b]$, con norma $|h|_1$, y $h(a) = h(b) = 0$. Tenemos entonces:

$$DI(0)h = \int_a^b (F_y(x, y_0, y'_0)h + F_z(x, y_0, y'_0)h') dx = 0$$

para todo h .

Si denotamos por

$$\alpha(x) = F_y(x, y_0(x), y'_0(x)),$$

$$\beta(x) = F_z(x, y_0(x), y'_0(x)),$$

$\alpha(x)$ y $\beta(x)$ son continuas en $[a, b]$.

Lema 1 : (Lagrange (1736-1813)). Sea $\alpha(x)$ una función continua en $[a, b]$ tal que:

$$\int_a^b \alpha(x)h(x)dx = 0$$

para todo h en $C_0[a, b] = \{h(x) \text{ continua en } [a, b], \text{ con } h(a) = h(b) = 0\}$, entonces $\alpha(x) \equiv 0$.

Prueba. Si $\alpha(x_0) \neq 0$ en algún punto x_0 entre a y b , entonces, por continuidad, $\alpha(x)$ será del mismo signo que $\alpha(x_0)$ en una vecindad $[x_1, x_2]$ alrededor de x_0 , contenida en $[a, b]$. Sea $h(x)$ definida como:

$$h(x) = \begin{cases} (x - x_1)(x_2 - x) & \text{si } x \in [x_1, x_2] \\ 0 & \text{si } x \notin [x_1, x_2] \end{cases}$$

h está en $C_0[a, b]$, positiva en $[x_1, x_2]$. Por lo tanto:

$$\int_a^b \alpha(x)h(x) dx = \int_{x_1}^{x_2} \alpha(x)h(x) dx$$

tiene el signo de $\alpha(x_0)$, por lo tanto tenemos una contradicción. $\alpha(x)$ debe ser 0 en el interior de $[a, b]$ y, por continuidad, en a y b .

Q.E.D.

Lema 2 : *El resultado del lema 1 es válido si sólo se pide que*

$$\int_a^b \alpha(x)h(x) dx = 0,$$

para todo h en $C_0^n[a, b] = \{h(x) \text{ con derivadas continuas en } [a, b], h^{(j)}(a) = h^{(j)}(b) = 0, j = 0, \dots, n\}$ o para h en $C_0^\infty[a, b] = \{h(x) \text{ en } C^\infty[a, b] \text{ con todas las derivadas 0 en } a \text{ y en } b\}$.

Prueba. Basta modificar la prueba del lema 1 tomando

$$h(x) = \begin{cases} ((x - x_1)(x_2 - x))^{n+1}, & \text{en } [x_1, x_2]; \\ 0, & \text{fuera de } [x_1, x_2] \end{cases}$$

o bien,

$$h(x) = \begin{cases} \exp(-1/(x - x_1)^2(x - x_2)^2); & \text{en } [x_1, x_2] \\ 0, & \text{fuera de } [x_1, x_2] \end{cases}$$

Q.E.D.

Notemos que si queremos usar estos lemas, es más general el resultado con h en $C_0^\infty[a, b]$, una clase más chica que $C_0[a, b]$.

Ejemplo 1 : Supongamos que y_0 es un extremo local del funcional

$$J(y) = \int_a^b F(x, y(x)) dx$$

definido sobre $C^1[a, b]$ con la norma $\|\cdot\|_1$, con las condiciones $y(a) = A$, $y(b) = B$. Considerando el funcional

$$I(h) = \int_a^b F(x, y_0(x) + h(x)) dx$$

sobre $C^1[a, b] \cap C_0[a, b]$, con norma $\|\cdot\|_1$, $h = 0$ es un extremo local. Si F, F_y son continuas en (a, b) , entonces $I(h)$ tiene derivada de Gâteaux en 0, en toda dirección h en $C_0^\infty[a, b]$:

$$D_h I(0) = \int_a^b F_y(x, y_0(x))h(x) dx = 0.$$

Por lo tanto $F_y(x, y_0(x)) = 0$ en (a, b) : uno toma $h \equiv 0$ en vecindades de a y b si no hay continuidad de F_y en (a, A) o en (b, B) . Notemos que si F_y es continuamente diferenciable, con $F_y(a, A) = 0$, y $F_{yy}|_{(a,A)} \neq 0$ entonces por el teorema de la función implícita, la ecuación:

$$F_y(x, y) = 0, \quad y(a) = A$$

tiene una solución, $y(x)$, única en $[a, b]$. Como $y_0(b) = B$ este tipo de problemas tiene demasiadas condiciones.

Ejercicio 1 : Tratar las funcionales: $J(y) = \int_a^b y^2 dx$, $I(y) = \int_a^b (x - y)^2 y^2 dx$.

Lema 3 : Sea $\alpha(x)$ una función continua en $[a, b]$ tal que

$$\int_a^b \alpha(x) h'(x) dx = 0$$

para todo h en $C^1[a, b] \cap C_0[a, b]$, entonces $\alpha(x) = \text{constante}$.

Prueba. Notemos que si $\alpha(x)$ fuese diferenciable, integrando por partes tendríamos:

$$\int_a^b \alpha(x) h'(x) dx = - \int_a^b \alpha'(x) h(x) dx + \alpha h|_a^b = 0,$$

por lo tanto, como $h(a) = h(b) = 0$,

$$\int_a^b \alpha'(x) h(x) dx = 0.$$

Por el lema 2, tenemos $\alpha'(x) = 0$. Sin embargo, $\alpha(x)$ no es necesariamente diferenciable. La idea es entonces usar integración por partes en el otro sentido. Sea $c = \int_a^b \alpha(x) dx / (b - a)$, el promedio de α . Sea

$$h(x) = \int_a^x (\alpha(\xi) - c) d\xi.$$

Entonces $h(a) = 0$, $h(b) = 0$, por la definición de c , $h'(x) = \alpha(x) - c$. h pertenece a la clase del lema. Entonces

$$0 = \int_a^b \alpha h' dx = \int_a^b \alpha(x) (\alpha(x) - c) dx = \int_a^b (\alpha(x) - c)^2 dx$$

ya que $\int_a^b c(\alpha(x) - c) dx = ch(b) = 0$. Por lo tanto $\alpha(x) = c$.

Q.E.D.

Ejemplo 2 : Sea

$$I(h) = \int_a^b F(x, y_0'(x) + h'(x)) dx,$$

con un extremo en $C^1[a, b] \cap C_0[a, b]$, en $h = 0$, y norma $|\cdot|_1$. Si F, F_z son continuas en $(a, b) \times \mathbb{R}$, entonces, como en el ejemplo 1, tenemos:

$$D_h I(0) = \int_a^b F_z(x, y'_0(x)) h'(x) dx = 0$$

para todo h en $C_0^\infty[a, b]$. Por el lema 3, con los mismos argumentos, tenemos:

$$F_z(x, y'_0(x)) = c \quad \text{en } (a, b).$$

Si F_z es C^1 , con $F_{zz}|_{(x, y'_0(x))} \neq 0$, podemos resolver esta ecuación para $z = z(x, c)$. Entonces

$$y_0(x) = A + \int_a^x z(\xi, c) d\xi,$$

donde c se determina a partir de la condición $y_0(b) = B$.

Ejercicio 2 : Tratar el funcional

$$J(y) = \int_a^b f(x)(1 + y'(x)^2)^{1/2} dx,$$

$y(a) = A, y(b) = B$. Ver que $y'f(x)$ tiene un solo signo, así como y' . Como ejemplo resolver los casos:

1. $f(x) = 1/x$ $a = 1, b = 2, A = 0, B = 1$.
2. $f(x) = 1, y(a) = A, y(b) = B$.

Ejercicio 3 : Sea $J(y) = \int_0^1 (1 + y'^2(x))^{-1} dx$ definido para las funciones $C^1[0, 1]$ con $y(0) = 0, y(1) = 1$. Sea $u(x) = y(x) - x$, con u en $C^1[0, 1] \cap C_0[0, 1]$ (con norma $|\cdot|_1$).

- a) Probar que $I(u) = J(x + u(x))$ es Fréchet diferenciable.
- b) Resolver la ecuación $DI(u)h = 0$ para todo h , y ver que $u \equiv 0$ es el único punto crítico.
- c) Probar que $u = 0$ es un mínimo local.

Lema 4 : Sea $\alpha(x)$ una función continua tal que

$$\int_a^b \alpha(x) h''(x) dx = 0$$

para todo h en $C^2[a, b] \cap C_0^1[a, b]$, entonces $\alpha(x)$ es un polinomio de grado a lo más 1.

Prueba. Si supiéramos que $\alpha(x)$ es C^2 , entonces integrando dos veces por partes, tendríamos:

$$0 = \int_a^b \alpha h'' dx = \int_a^b \alpha'' h dx - \alpha' h|_a^b + \alpha h'|_a^b,$$

usando el lema 2 y las condiciones de frontera, tendríamos que $\alpha''(x) = 0$.

Sean c_0 y c_1 dos constantes tales que:

$$\int_a^b (\alpha(x) - c_0 - c_1(x - a)) dx = 0$$

$$\int_a^b \int_a^x (\alpha(\zeta) - c_0 - c_1(\zeta - a)) d\zeta dx = 0,$$

es decir:

$$(b - a)^2(c_0/2 + c_1(b - a)/3!) = \int_a^b \int_a^x \alpha(\zeta) d\zeta dx,$$

$$c_0(b - a) + c_1(b - a)^2/2 = \int_a^b \alpha(x) dx.$$

Si $p(x - a)$ denota al polinomio

$$\int_a^x \int_a^\zeta (c_0 + c_1(\zeta_1 - a)) d\zeta_1 d\zeta = c_0(x - a)^2/2 + c_1(x - a)^3/3!,$$

entonces $p(x - a) = (x - a)^2 q(x - a)$ con

$$q(x - a) = c_0/2 + c_1(x - a)/3! = \sum_0^1 q^{(i)}(b - a)(x - b)^i/i!$$

expandiendo en serie de Taylor.

Queremos determinar $q(x - a)$ de tal modo que $p(b - a)$ y $p'(b - a)$ sean dados por las integrales de $\alpha(x)$. Como

$$p'(b - a) = (b - a)^2 q'(b - a) + 2(b - a)q(b - a)$$

podemos escribir las condiciones en la forma:

$$\begin{pmatrix} p(b - a) \\ p'(b - a) \end{pmatrix} = \begin{pmatrix} (b - a)^2 & 0 \\ 2(b - a) & (b - a)^2 \end{pmatrix} \begin{pmatrix} q(b - a) \\ q'(b - a) \end{pmatrix}$$

Invirtiendo esta matriz (válido pues es triangular), obtenemos los coeficientes de $q(x - a)$, es decir c_0 y c_1 . (La razón por no haber resuelto directamente el par de ecuaciones es para dar una indicación de como hacer el ejercicio 4.)

Sea entonces $h(x) = \int_a^x \int_a^\xi (\alpha(\xi_1) - c_0 - c_1(\xi_1 - a)) d\xi_1 d\xi$. Claramente $h(x)$ es C^2 , con $h(a) = h(b) = h'(a) = h'(b) = 0$, y $h'' = \alpha(x) - c_0 - c_1(x - a)$.

Además $\int_a^b (c_0 + c_1(x - a)) h''(x) dx = (c_0 + c_1(x - a)) h'(x) - c_1 h(x)|_a^b = 0$, por lo tanto:

$$\int_a^b (\alpha(x) - c_0 - c_1(x - a))^2 dx = \int_a^b (\alpha(x) - c_0 - c_1(x - a)) h''(x) dx$$

$$= \int_a^b \alpha(x) h''(x) dx - \int_a^b (c_0 + c_1(x-a)) h''(x) dx.$$

El primer término vale 0 por hipótesis, y el segundo también por el argumento anterior. Concluimos que $\alpha(x) = c_0 + c_1(x-a)$.

Q.E.D.

Lema 5 : Sea $\alpha(x)$ una función continua tal que

$$\int_a^b \alpha(x) h^{(n)}(x) dx = 0$$

para todo h en $C^n[a, b] \cap C_0^{n-1}[a, b]$, entonces $\alpha(x)$ es un polinomio de grado a lo más $n-1$.

Prueba. Sea

$$h(x) = \int_a^x \int_a^{\xi_1} \cdots \int_a^{\xi_{n-1}} (\alpha(\xi_n) - c_0 - c_1(\xi_n - a) - \cdots - c_{n-1}(\xi_n - a)^{n-1}) d\xi_n \cdots d\xi_1$$

donde c_0, \dots, c_{n-1} son n constantes determinadas por las condiciones $h^{(i)}(b) = 0$, $i = 0, \dots, n-1$. Por construcción $h^{(i)}(a) = 0$, $i = 0, \dots, n-1$, y

$$h^{(n)}(x) = \alpha(x) - c_0 - c_1(x-a) - \cdots - c_{n-1}(x-a)^{n-1}.$$

Integrando por partes, es fácil ver que

$$\int_a^b h^{(n)}(x) (c_0 + \cdots + c_{n-1}(x-a)^{n-1}) dx = 0.$$

Entonces $\int_a^b h^{(n)}(x)^2 dx = 0$, implicando el resultado.

Q.E.D.

Ejercicio 4 : Probar que las constantes c_0, \dots, c_{n-1} existen. Para esto escribir el polinomio

$$p(x-a) = \int_a^x \int_a^{\xi_1} \cdots \int_a^{\xi_{n-1}} (c_0 + c_1(\xi_n - a) + \cdots + c_{n-1}(\xi_n - a)^{n-1}) d\xi_n \cdots d\xi_1$$

como $(x-a)^n q(x-a)$, con

$$q(x-a) = \sum_{i=0}^{n-1} i! c_i (x-a)^i / (i+n)! = \sum_{i=0}^{n-1} q^{(i)}(b-a) (x-b)^i / i!$$

Escribir $(p(b-a), \dots, p^{(n-1)}(b-a))$ en función de $(q(b-a), \dots, q^{(n-1)}(b-a))$ y una matriz triangular, para expresar $q^{(i)}(b-a)$ y por lo tanto las c_i .

Lema 6 : (Du Bois-Reymond (1831-1889)).

Sean $\alpha(x)$ y $\beta(x)$ dos funciones continuas tales que:

$$\int_a^b (\alpha(x)h(x) + \beta(x)h'(x)) dx = 0$$

para todo h en $C^1[a, b]$, con $h(a) = h(b) = 0$. Entonces β es C^1 en $[a, b]$ con derivada $\beta'(x) = \alpha(x)$.

Prueba. Notemos que si β es C^1 , integrando por partes, usando las condiciones de frontera, el resultado es una consecuencia del lema 1. Para llegar al resultado daremos el siguiente rodeo: sea $A(x) = \int_a^x \alpha(\xi) d\xi$, entonces $A'(x) = \alpha(x)$,

$$\int_a^b \alpha(x)h(x) dx = \int_a^b A'(x)h(x) dx = - \int_a^b A(x)h'(x) dx + A(x)h(x) \Big|_a^b.$$

La relación del enunciado del lema se escribe entonces :

$$\int_a^b (\beta(x) - A(x))h'(x) dx = 0$$

para todo $h \in C^1[a, b] \cap C_0[a, b]$. Por el lema 3:

$$\beta(x) - A(x) = \text{constante}.$$

Como $A(x)$ es C^1 , también lo es $\beta(x)$, con $\beta'(x) = A'(x) = \alpha(x)$.

Q.E.D.

En estas notas, usaremos más frecuentemente el siguiente resultado, el cual generaliza todos los lemas previos:

Teorema 1 : (Meyers-Serrin, 1964).

1) Sea $\alpha(x)$ continua en $[a, b]$ tal que :

$$\int_a^b \alpha(x)h^{(n)}(x) dx = 0$$

para todo h en $C_0^\infty[a, b]$. Entonces $\alpha(x)$ es un polinomio de grado a lo más $n - 1$.

2) Sean $\alpha(x), \beta(x)$ continuas en $[a, b]$ tales que :

$$\int_a^b (\alpha(x)h(x) + \beta(x)h^{(n)}(x)) dx = 0$$

para todo h en $C_0^\infty[a, b]$. Entonces $\beta \in C^n$ con

$$\beta^{(n)}(x) = (-1)^{n-1}\alpha(x).$$

Prueba.

1) Sea $\rho(x) > 0$ en $] -1, 1[$, con ρ en $C_0^\infty(\mathbb{R})$ y

$$\int_{\mathbb{R}} \rho(x) dx = 1.$$

Por ejemplo

$$\rho(x) = \begin{cases} e^{-1/(1-x^2)} / \int_{-1}^1 e^{-1/(1-x^2)} dx, & \text{para } |x| \leq 1 \\ 0, & \text{para } |x| \geq 1 \end{cases}.$$

Sea $J_\varepsilon \alpha(x) = \varepsilon^{-1} \int_a^b \rho((y-x)/\varepsilon) \alpha(y) dy$, el “alisador” de Friedrichs de α .

Claramente $J_\varepsilon \alpha(x)$ es C^∞ para x en (a, b) y, como $\rho((y-x)/\varepsilon) = 0$ para $|y-x| \geq \varepsilon$, si $|x-a| \geq \varepsilon_0 \geq \varepsilon$, $|x-b| \geq \varepsilon_0 \geq \varepsilon$, entonces el intervalo de integración se reduce a $[x-\varepsilon, x+\varepsilon]$. Entonces, el cambio de variable $y = x + \varepsilon z$, permite escribir:

$$J_\varepsilon \alpha(x) = \int_{-1}^1 \rho(z) \alpha(x + \varepsilon z) dz,$$

$$J_\varepsilon \alpha(x) - \alpha(x) = \int_{-1}^1 \rho(z) (\alpha(x + \varepsilon z) - \alpha(x)) dz.$$

Dado que $\alpha(x)$ es uniformemente continua en $[a, b]$, $J_\varepsilon \alpha(x)$ tiende uniformemente a $\alpha(x)$, para todo x en $[a + \varepsilon_0, b - \varepsilon_0]$. Por otra parte :

$$\frac{d^n}{dx^n} (J_\varepsilon \alpha(x)) = \varepsilon^{-1-n} \int_a^b (-1)^n \rho^{(n)}((y-x)/\varepsilon) \alpha(y) dy$$

con $\rho^{(n)}((y-x)/\varepsilon)$ en $C_0^\infty[a, b]$ como función de y , si x pertenece al intervalo $[a + \varepsilon_0, b - \varepsilon_0]$, $\varepsilon < \varepsilon_0$, es decir, esta función pertenece a la clase para la cual la integral es 0; en consecuencia $J_\varepsilon \alpha(x)$ es un polinomio de grado a lo más $n-1$, para x en $[a + \varepsilon_0, b - \varepsilon_0]$:

$$J_\varepsilon \alpha(x) = P_\varepsilon(x) = \sum_0^{n-1} a_{i\varepsilon} x^i$$

donde las $a_{i\varepsilon}$ pueden depender de ε .

Tomemos x_1, \dots, x_n , n puntos distintos en $[a + \varepsilon_0, b - \varepsilon_0]$. Tenemos

$$\begin{pmatrix} P_\varepsilon(x_1) \\ \vdots \\ P_\varepsilon(x_n) \end{pmatrix} = \begin{pmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{pmatrix} \begin{pmatrix} a_{0\varepsilon} \\ \vdots \\ a_{n-1\varepsilon} \end{pmatrix}.$$

Si A es la matriz, A es invertible con determinante $\prod_{i \neq j} (x_i - x_j)$: determinante de Vandermonde.

Por lo tanto:

$$\begin{pmatrix} a_{0\varepsilon} \\ \vdots \\ a_{n-1\varepsilon} \end{pmatrix} = A^{-1} \begin{pmatrix} P_\varepsilon(x_1) \\ \vdots \\ P_\varepsilon(x_n) \end{pmatrix}$$

converge a:

$$A^{-1} \begin{pmatrix} \alpha(x_1) \\ \vdots \\ \alpha(x_n) \end{pmatrix} \equiv \begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix}.$$

Invirtiendo de nuevo, tendremos

$$\alpha(x_j) = \sum_0^{n-1} a_i x_j^i,$$

para $j = 1, \dots, n$. Cuando ε tiende a 0, $J_\varepsilon \alpha(x)$ tiende a α uniformemente, lo que permite escoger n valores x_i arbitrarios sin perder generalidad. Como $a_i = \lim_{\varepsilon \rightarrow 0} a_{i\varepsilon}$ no depende de x_1, \dots, x_n por la convergencia uniforme, no puede haber más de un límite, y como x_1, \dots, x_n son arbitrarios, $\alpha(x)$ es ese polinomio en $[a + \varepsilon_0, b - \varepsilon_0]$. ε_0 es arbitrario y $\alpha(x)$ es continuo en $[a, b]$, por lo tanto el primer punto queda probado.

Para (2), sea $A(x) = \int_a^x \int_a^{x_1} \dots \int_a^{x_{n-1}} \alpha(x_n) dx_n \dots dx_1$. Entonces, $A(x)$ es C^n con

$$A^{(n)}(x) = \alpha(x).$$

$$\int_a^b \alpha(x) h(x) dx = \int_a^b A^{(n)}(x) h(x) dx = (-1)^n \int_a^b A(x) h^{(n)}(x) dx$$

ya que h está en $C_0^\infty[a, b]$.

Si la relación de (2) es cierta, entonces, por la parte (1), $\beta(x) + (-1)^n A(x)$ es un polinomio de grado a lo más $n - 1$. De ahí la conclusión.

Q.E.D.

Lema 7 : Sean α, β, γ funciones continuas en $[a, b]$ tales que:

$$\int_a^b (\alpha(x)h(x) + \beta(x)h'(x) + \gamma(x)h''(x)) dx = 0$$

para todo h en $C_0^\infty[a, b]$. Entonces $\gamma(x)$ es C^1 , $\beta(x) - \gamma'(x)$ es C^1 con derivada $\alpha(x)$.

Prueba. Sean

$$A(x) = \int_a^x \int_a^{x_1} \alpha(x_2) dx_2 dx_1$$

$$B(x) = \int_a^x \beta(x_1) dx_1.$$

A es C^2 , con $A'' = \alpha$, B es C^1 , con $B' = \beta$. Integrando por partes tenemos:

$$\int_a^b (A(x) - B(x) + \gamma(x))h''(x) dx = 0$$

para todo h en $C_0^\infty[a, b]$. Por lo tanto $A - B + \gamma$ es un polinomio de grado a lo más 1. Como $A - B$ es C^1 , también lo es γ , y $\beta - \gamma'$ tiene como derivada a α .

Q.E.D.

Para más variables la situación es menos clara, con las herramientas que hemos desarrollado hasta ahora.

Lema 8 : (Haar). Sea Ω un dominio de \mathbb{R}^2 , para el cual el teorema de la divergencia es válido.

1) Sea $\alpha(x, y)$ una función continua en $\bar{\Omega}$, tal que:

$$\iint_{\Omega} \alpha(x, y) h(x, y) dx dy = 0,$$

para todo h en $C_0^\infty(\bar{\Omega})$. Entonces $\alpha(x, y) \equiv 0$.

2) Sean $\alpha(x, y)$ continua, $\beta(x, y)$, $\gamma(x, y)$ C^1 en $\bar{\Omega}$, tales que:

$$\iint_{\Omega} (\alpha h + \beta h_x + \gamma h_y) dx dy = 0,$$

para todo h en $C_0^\infty(\bar{\Omega})$. Entonces

$$\alpha = \beta_x + \gamma_y.$$

3) Si α , β , γ satisfacen las mismas propiedades que en (2) excepto que β y γ son solamente continuas, entonces para todo $\Omega' \subset \Omega$:

$$\iint_{\Omega'} \alpha dx dy = \int_{\partial\Omega'} \beta dy - \gamma dx.$$

Prueba.

1) Usando los mismos argumentos que en el lema 2, dejaremos esta parte como ejercicio.

2) Ejercicio: usar el teorema de la divergencia y (1).

3) Es suficiente probar el resultado cuando Ω' es un rectángulo $[a, b] \times [c, d]$, ya que un dominio general puede ser cubierto por rectángulos de esa forma.

Sean

$$A(x, y) = \int_a^x \int_c^y \alpha(\xi, \eta) d\xi d\eta$$

$$B(x, y) = \int_c^y \beta(x, \eta) d\eta$$

$$C(x, y) = \int_a^x \gamma(\xi, y) d\xi.$$

Tomemos $h(x, y) = v(x)w(y)$ con v en $C_0^\infty[a, b]$, w en $C_0^\infty[c, d]$.

Tenemos

$$\int_a^b \int_c^d (A_{xy}vw + B_yv'w + C_xvw') dx dy = 0$$

$$= \int_c^d \left(\int_a^b -A_x v - Bv' + C_x v \right) dx w'(y) dy = 0,$$

usando el hecho que $w(c) = w(d) = 0$ e integrando por partes en y . Por el lema 3 tenemos que $\int_a^b ((C_x - A_x)v - Bv') dx$ es constante. Integrando por partes y evaluando la constante en $y = c$, $y = d$:

$$\int_a^b ((A - B - C)|_{y=d} - (A - B - C)|_{y=c}) v'(x) dx = 0,$$

para todo v en $C_0^\infty[a, b]$. Por el teorema 1:

$(A - B - C)|_{y=d} - (A - B - C)|_{y=c}$ es una constante en x .

Evaluando esta expresión en $x = a$ y $x = b$:

$$A(b, d) - B(b, d) - C(b, d) + C(b, c) = -B(a, d).$$

Expresando cada uno de esos términos como integral, tenemos el resultado.

Q.E.D.

Nota 1: Si $\alpha \equiv 0$, entonces $\beta dy - \gamma dx$ es una forma exacta: existe un potencial $V(x, y)$ tal que $\beta = -V_y$, $\gamma = V_x$.

Nota 2: Usando derivadas en el sentido de distribuciones se puede extender la noción de divergencia de (β, γ) . Referimos a los textos más avanzados para estudiar estas consideraciones.

Capítulo 4

Condiciones necesarias

En el primer capítulo vimos como los Bernoulli plantearon y resolvieron los primeros problemas variacionales. Sin embargo fueron Euler (1707-1783) y Lagrange (1736-1813) quienes aportaron los primeros elementos técnicos del Cálculo de Variaciones, principalmente en cuanto a condiciones necesarias se refiere. En este capítulo veremos cómo usar las nociones de derivadas y los lemas fundamentales para encontrar las condiciones que debe satisfacer un extremo de un funcional, en varios contextos.

4.1 Ecuación de Euler-Lagrange

En esta primera sección estudiaremos el funcional

$$J(y) = \int_a^b F(x, y(x), y'(x)) dx,$$

donde $F(x, y, z)$, F_y , F_z son continuas en $[a, b] \times \mathbb{R}^2$. Aquí la clase de funciones es el subconjunto de $C^1[a, b]$, con norma $\|\cdot\|_1$, que satisfacen las condiciones $y(a) = A$, $y(b) = B$, como en el caso de la braquistocrona.

Supongamos que $y_0(x)$ es un extremo local, entonces podemos escribir $y(x) = y_0(x) + h(x)$, con $h(a) = h(b) = 0$, y estudiar el funcional

$$I(h) = \int_a^b F(x, y_0(x) + h(x), y'_0(x) + h'(x)) dx$$

sobre $C^1[a, b] \cap C_0[a, b]$, con norma $\|\cdot\|_1$.

Dado que 0 es un extremo local de $I(h)$, la derivada de Fréchet de I en 0 debe anularse (Ejemplo II.3.2 y Teorema II.4.1), es decir,

$$DI(0)h = \int_a^b (F_y(x, y_0, y'_0)h + F_z(x, y_0, y'_0)h') dx = 0,$$

para todo h en $C^1[a, b] \cap C_0[a, b]$.

Por el teorema III.1, tomando $\alpha(x) = F_y(x, y_0(x), y'_0(x))$, $\beta(x) = F_z(x, y_0(x), y'_0(x))$, tenemos que $\beta(x)$ es C^1 y su derivada es $\alpha(x)$, obteniendo así :

Teorema 1.1 : (Ecuación de Euler-Lagrange).

Si y_0 es un extremo local de $J(y)$ entonces, $F_z(x, y_0(x), y'_0(x))$ es C^1 y se cumple:

$$\frac{d}{dx}(F_z(x, y_0, y'_0)) = F_y(x, y_0, y'_0).$$

Pensando en el ejemplo de la braquistocrona, donde la función F no es C^1 , nos gustaría relajar las condiciones del teorema 1.1.

Teorema 1.2 : Sea y_0 un punto crítico del funcional $J(y)$, es decir:

1) $F_y(x, y_0 + th, y'_0 + th')$, $F_z(x, y_0 + th, y'_0 + th')$ son continuas en (a, b) para todo t pequeño y h en $C_0^\infty[a, b]$, con $|h|_1 = 1$.

2) $D_h(J(y))|_{y_0} = \int_a^b (F_y h + F_z h') dx = 0$ para todo h en $C_0^\infty[a, b]$,

entonces:

$F_z(x, y_0, y'_0)$ es C^1 en (a, b) y

$$\frac{d}{dx}(F_z(x, y_0, y'_0)) = F_y(x, y_0, y'_0).$$

Prueba. Basta notar que la prueba del teorema 1.1 es válida en este caso, ya que no hay problemas en los límites de integración.

Q.E.D.

Ejemplo 1.1 : La braquistocrona.

En este caso $F(x, y, y') = ((1 + y'^2)/y)^{1/2}$ es C^1 excepto que hay que restringirse a $y > 0$. Como se vió en el ejemplo II.3.1, cumple con las condiciones del Teorema 1.2. Por lo tanto tenemos que:

$$\frac{d}{dx} \left(\frac{y'}{(y(1 + y'^2))^{1/2}} \right) + \frac{(1 + y'^2)^{1/2}}{2y^{3/2}} = 0.$$

Aunque sabemos que $F_{y'}$ tiene derivada, no es claro que podemos derivar y' (lo veremos más adelante). Pero si multiplicamos la ecuación de Euler por $F_{y'}$, tenemos:

$$\frac{d}{dx}(F_{y'})F_{y'} + \frac{y'}{2y^2} = 0.$$

Por lo tanto $F_{y'}^2 - 1/y = c = -1/(y(1 + y'^2))$, es decir:

$$y(1 + y'^2) = \text{constante},$$

relación que había encontrado Bernoulli.

Como siempre cada término en el enunciado del teorema tiene su razón de ser. Por ejemplo uno puede tener un extremo pero sin que sea solución de la ecuación de Euler-Lagrange (evidentemente en ese caso no se cumplen las condiciones del teorema).

Ejemplo 1.2 : Sea $J(y) = \int_0^1 (y - x)^{1/2} dx$ con $y(0) = 0$, $y(1) = 1$. $J(y)$ es C^1 sólo si $y(x) > x$. Aquí la ecuación de Euler se reduce a:

$$F_y = 1/(2(y - x)^{1/2}) = 0$$

que no tiene solución.

Sin embargo $y(x) = x$ es claramente un mínimo. Lo que pasa es que $J(x + th) = \int_0^1 (th)^{1/2} dx$ está definido sólo si $h = 0$, ya que t debe cambiar de signo. En ese caso la derivada de Gâteaux, $D_0 J(x) = 0$, está definida sólo para $h = 0$ y el lema de Du Bois-Reymond no se puede aplicar.

Proposición 1.1 : Si $F(x, y, z)$ es C^2 en $[a, b] \times \mathbb{R}^2$ y y_0 es un punto crítico de $J(y)$, entonces $y_0(x)$ es C^2 en todos los puntos x donde $F_{zz}(x, y_0(x), y'_0(x)) \neq 0$.

Prueba.

$$\begin{aligned} & (F_z(x + \Delta x, y_0(x + \Delta x), y'_0(x + \Delta x)) - F_z(x, y_0(x), y'_0(x)))/\Delta x \\ &= F_{xz}(\bar{x}, \bar{y}, \bar{z}) + F_{yz}(\bar{x}, \bar{y}, \bar{z})(y_0(x + \Delta x) - y_0(x))/\Delta x + \\ & \quad F_{zz}(\bar{x}, \bar{y}, \bar{z})(y'_0(x + \Delta x) - y'_0(x))/\Delta x \end{aligned}$$

donde $(\bar{x}, \bar{y}, \bar{z})$ es un punto intermedio entre $(x, y_0(x), y'_0(x))$ y $(x + \Delta x, y_0(x + \Delta x), y'_0(x + \Delta x))$.

Cuando Δx tiende a 0, el lado izquierdo tiene un límite, por ser y_0 un punto crítico; por lo tanto :

$$\frac{d}{dx}(F_z(x, y_0(x), y'_0(x))) = F_{xz} + F_{yz}y'_0(x) + F_{zz} \lim_{\Delta x \rightarrow 0} \frac{y'_0(x + \Delta x) - y'_0(x)}{\Delta x}.$$

Por lo tanto si $F_{zz} \neq 0$, uno puede despejar $y''(x)$ y ver que es continuo.

Q.E.D.

Ejemplo 1.3 : En el caso de la braquistocrona, la cicloide no puede conectarse con $y = \text{constante}$, ya que $F_{zz} = y^{-1/2}(1 + y'^2)^{-3/2}$ para $F = (1 + y'^2)^{1/2}y^{-1/2}$: la solución al problema del mínimo, si existe, debe ser C^2 para $y > 0$.

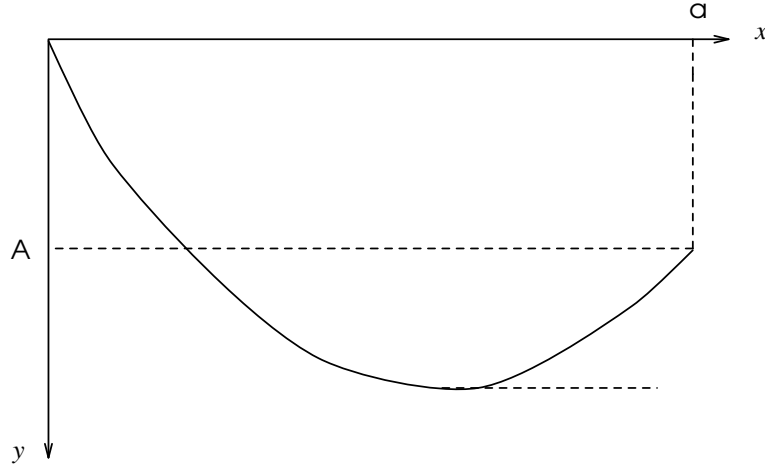


Fig. 1.1

Ejemplo 1.4 : Si $F_{zz} = 0$, el punto crítico puede no ser C^2 . Por ejemplo si tenemos el problema:

$$J(y) = \int_{-1}^1 y^2(2x - y')^2 dx$$

con $y(-1) = 0$, $y(1) = 1$, entonces $J(y) \geq 0$. Tomando

$$y_0 = \begin{cases} 0, & \text{para } -1 \leq x \leq 0 \\ x^2, & \text{para } 0 \leq x \leq 1 \end{cases}$$

entonces $J(y_0) = 0$. y_0 es C^1 pero no C^2 en $x = 0$. Aquí $F_{zz} = 2y^2 \neq 0$ sólo si $y \neq 0$.

Nota 1.1 : Si F es C^3 y $F_{zz} \neq 0$, la ecuación de Euler se escribe:

$$F_{zz}y'' + F_{yz}y' + F_{xz} = F_y$$

y podemos despejar y'' ,

$$y'' = g(x, y, y')$$

donde g es una función C^1 . Sabemos entonces que este problema con las condiciones *iniciales* $y(a) = A$, $y'(a) = \alpha$ tiene una solución única para x vecino a a . Sin embargo aquí se trata de

resolver esa ecuación con las condiciones de *frontera* $y(a) = A$, $y(b) = B$. Podríamos pensar en variar el ángulo α en la figura 1.2, hasta “pegarle” a B en $x = b$.

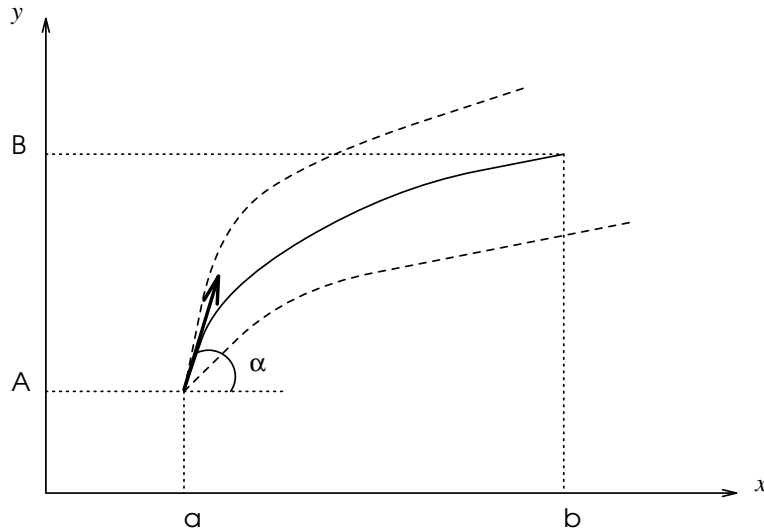


Fig. 1.2

Este es de hecho un método que se usa en el tratamiento de ecuaciones diferenciales, pero es claro que no necesariamente hay una solución o que sea única: el lector puede experimentar con tiros al blanco de piedras.

Ejemplo 1.5 : En el problema de la superficie de revolución (ejemplo I.1.7) el funcional a considerar era :

$$A(y) = 2\pi \int_a^b y(1 + y'^2)^{1/2} dx$$

con $y(a) = A$, $y(b) = B$.

En este caso $F_{zz} = 2\pi y/(1 + y'^2)^{3/2}$ por lo tanto el mínimo, si existe, es C^2 cuando $y \neq 0$ y satisface la ecuación:

$$yy'' - (1 + y'^2) = 0$$

(dejamos al lector hacer los cálculos a partir de la ecuación de Euler). Notemos que la ecuación de Euler se escribe:

$$[yy'(1 + y'^2)^{-1/2}]' - (1 + y'^2)^{1/2} = 0.$$

Multiplicando por F_z , tenemos:

$$(F_z)'F_z - yy' = 0 = \frac{d}{dx}\left(\frac{1}{2}F_z^2 - \frac{1}{2}y^2\right).$$

Por lo tanto:

$$y^2/(1 + y'^2) = \text{constante}.$$

Derivando esta relación obtenemos de nuevo la ecuación de Euler. Veremos más adelante cómo resolver esta integral primera.

4.2 Integral primera

En esta sección veremos que la ecuación de Euler tiene una integral primera si F no depende explícitamente de x . De hecho, si no nos preocupamos de la diferenciabilidad, la ecuación de Euler puede escribirse:

$$\frac{d}{dx}(F_z(y, y')) = F_y(y, y').$$

Pero $\frac{d}{dx}(F(y, y')) = F_y y' + F_z y''$. Por lo tanto, multiplicando la ecuación de Euler por y' :

$$y'(F_z)' + y''F_z = F' = (y'F_z)'.$$

Es decir, $F - y'F_z = \text{constante}$.

Para que este argumento funcione necesitamos que se suponga que y es C^2 ; esto no es necesario si y es un extremo local del funcional.

Teorema 2.1 : (Segunda ecuación de Euler).

Si y_0 es un extremo local del funcional

$$J(y) = \int_a^b F(x, y, y') dx$$

con $y(a) = A$, $y(b) = B$, F es C^1 en $[a, b] \times \mathbb{R}^2$, entonces

$$F(x, y_0(x), y_0'(x)) - y_0'(x)F_z(x, y_0(x), y_0'(x))$$

es C^1 , con derivada $F_x(x, y_0(x), y'_0(x))$.

En particular si F no depende explícitamente de x , entonces $F_x = 0$, y

$$F - y'_0 F_z = \text{constante}.$$

Prueba. Otra vez si no nos preocupamos de la diferenciabilidad de y_0 , tendremos:

$$\frac{d}{dx}(F) = F_x + F_y y' + F_z y''.$$

Usando la ecuación de Euler: $F_y = (F_z)'$, obtenemos

$$F_x = (F - y' F_z)'$$

Sea ε tan pequeño que bajo la transformación:

$$x = \xi + \varepsilon \eta$$

$$y = \eta$$

la curva $y = y_0(x)$ puede ser parametrizada por ξ y dar la curva $\eta = \eta_0(\xi)$ para ξ entre $\alpha = a - \varepsilon A$ y $\beta = b - \varepsilon B$.

Asimismo cualquier curva $y = y(x)$ vecina de $y_0(x)$ dará una curva $\eta = \eta(\xi)$ e inversamente, con

$$\eta(\xi) = y(\xi + \varepsilon \eta(\xi)) \quad , \quad \eta'(\xi) = y'(\xi + \varepsilon \eta(\xi))(1 + \varepsilon \eta'(\xi)),$$

es decir, $y'(x) = \eta'/(1 + \varepsilon \eta')$.

Ahora, bajo el cambio de variables $x = \xi + \varepsilon \eta(\xi)$, tenemos

$$\begin{aligned} J(y) &= \int_a^b F(x, y(x), y'(x)) dx \\ &= \int_\alpha^\beta F(\xi + \varepsilon \eta(\xi), \eta(\xi), \eta'(\xi)/(1 + \varepsilon \eta'(\xi)))(1 + \varepsilon \eta'(\xi)) d\xi \\ &= \int_\alpha^\beta \tilde{F}(\xi, \eta, \eta') d\xi = \tilde{J}(\eta). \end{aligned}$$

Claramente si $J(y)$ tiene un extremo local en y_0 para la clase de y en $C^1[a, b]$ con $y(a) = A$, $y(b) = B$, entonces $\tilde{J}(\eta)$ tendrá un extremo local en η_0 para la clase de η en $C^1[\alpha, \beta]$ con $\eta(\alpha) = A$, $\eta(\beta) = B$. Por lo tanto \tilde{F}_η , tiene como derivada (con respecto a ξ) a \tilde{F}_η . Pero

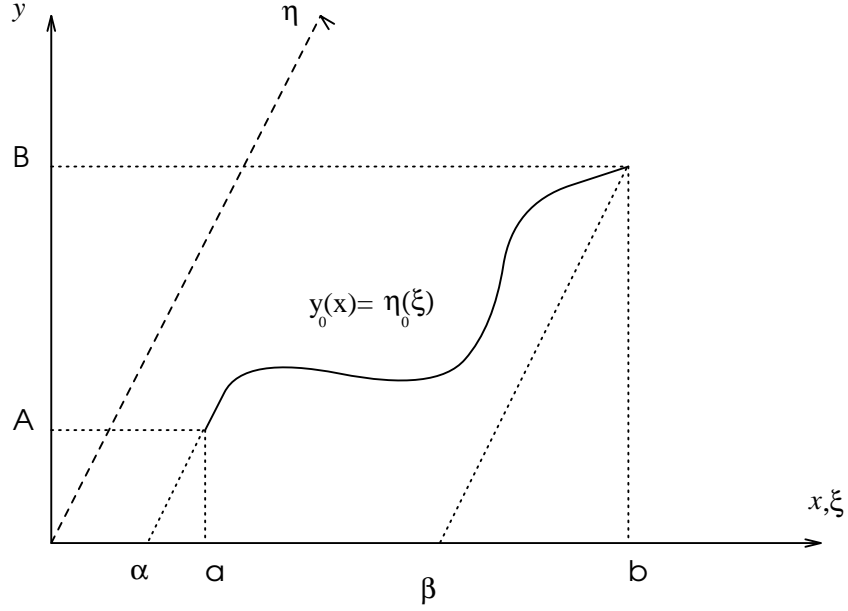


Fig. 2.1

$$\begin{aligned}\tilde{F}_\eta &= [\varepsilon F_x(\xi + \varepsilon \eta_0(\xi), \eta_0(\xi), \eta'_0(\xi)/(1 + \varepsilon \eta'_0(\xi))) + F_y](1 + \varepsilon \eta'_0(\xi)) \\ \frac{d}{d\xi}(\tilde{F}_{\eta'}) &= \frac{d}{d\xi}(F_{y'}/(1 + \varepsilon \eta'_0) + \varepsilon F) = \frac{d}{dx}(\tilde{F}_{\eta'})(1 + \varepsilon \eta'_0(\xi)).\end{aligned}$$

Por lo tanto:

$$\begin{aligned}F_y + \varepsilon F_x(x, y_0, y'_0) &= \frac{d}{dx}(F_{y'}(1 - \varepsilon \eta'_0(1 + \varepsilon \eta'_0)^{-1}) + \varepsilon F) \\ &= \frac{d}{dx}(F_{y'} - \varepsilon y' F_{y'} + \varepsilon F).\end{aligned}$$

Como $F_{y'}$ tiene derivada igual a F_y , obtenemos el resultado.

Q. E. D.

En el caso particular de F independiente de x usaremos el siguiente resultado.

Teorema 2.2 : *Si F no depende explícitamente de x y es C^1 , entonces el conjunto de soluciones C^2 y no constantes de la ecuación de Euler, coincide con el conjunto de soluciones C^2 y no constantes de la ecuación $F(y, y') - y'F_z(y, y') = c$ para alguna constante c .*

Prueba. Una de las implicaciones ya fue probada. Inversamente derivando $F - y'F_z = c$ y usando la relación $(F)' = F_y y' + F_z y''$ se llega a:

$$y'((F_z)' - F_y) = 0.$$

Notemos que $y = d$ no es solución de la ecuación de Euler, al menos que $F_y(d, 0) = 0$ (ya que $(F_{y'}(d, 0))' = 0$), mientras que $y = d$ es siempre solución de $F - y'F_{y'} = c$ con $c = F(d, 0)$.

Q. E. D.

Ejemplo 2.1 : La braquistocrona.

$$F(y, y') = (1 + y'^2)^{1/2} y^{-1/2}.$$

Ya sabemos que los puntos críticos son C^2 si $y \neq 0$.

$$F_y = -(1 + y'^2)^{1/2} y^{-3/2} / 2$$

$F_y(d, 0) \neq 0$: las constantes no son soluciones de la ecuación de Euler. $F - y'F_{y'} = c = y^{-1/2}(1 + y'^2)^{-1/2}$, por lo tanto:

$$y(1 + y'^2) = 1/c^2 = \text{constante}.$$

Además, de la ecuación de Euler:

$$2yy'' + (1 + y'^2) = 0$$

vemos que $y'' < 0$, y' es monótona decreciente: podemos tomar a y' como parámetro; si $y' = y'(x)$ entonces $x = x(y')$. Tomando $y' = \cotan \theta/2$

$$y = 1/(c^2(1 + y'^2)) = (\sin^2(\theta/2))/c^2 = (1 - \cos \theta)/2c^2$$

$$\frac{dx}{d\theta} = \frac{dy}{d\theta} \frac{1}{y'} = (\sin^2(\theta/2))/c^2 = (1 - \cos \theta)/2c^2.$$

Así $x(\theta) = (\theta - \sin \theta)/2c^2$ justificando los resultados del primer capítulo. (Todavía no se ha probado que la cicloide es un mínimo.)

Ejemplo 2.2 : Superficie de revolución.

$$A(y) = 2\pi \int_0^a y(1 + y'^2)^{1/2} dx.$$

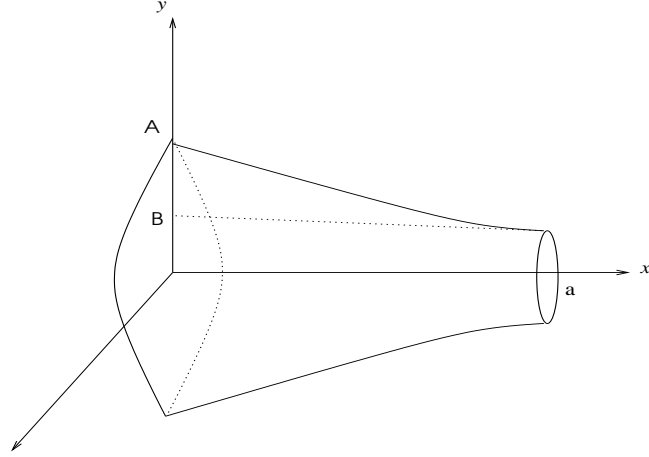


Fig. 2.2

Ya vimos que cualquier punto crítico es C^2 si $y \neq 0$. Como $F_y = (1 + y'^2)^{1/2}$, las constantes no son soluciones.

$$F - y'F_z = c = y(1 + y'^2)^{-1/2}.$$

Por lo tanto en un extremo y es de un solo signo: positivo si A es positivo. De la ecuación de Euler

$$yy'' - (1 + y'^2) = 0,$$

vemos que y es C^2 , cuando $y \neq 0$, y $y'' > 0$. Por lo tanto y' es creciente y puede ser tomada como parámetro. Definiendo a θ por:

$$y' = \sinh \theta,$$

entonces $y = c(1 + y'^2)^{1/2} = c \cosh \theta$

$$\frac{dx}{d\theta} = \frac{dy}{d\theta} / y' = c$$

$$x(\theta) = c\theta + \alpha = c(\theta - \beta)$$

$$y(x) = c \cosh(x/c + \beta).$$

La solución, si existe, es una catenaria.

Debemos todavía satisfacer las condiciones de frontera $y(0) = A$, $y(a) = B$. De la primera condición se obtiene $1/c = (\cosh \beta)/A$, para la segunda hay que encontrar β real tal que:

$$(B \cosh \beta)/A = \cosh((a \cosh \beta)/A + \beta).$$

Ejercicio 2.1 :

1. Probar que si $a \geq A + B$, no hay solución y dar la interpretación física. (Probar que $\cosh x \geq |x|$.)
2. Desarrollar la expresión anterior y escribirla en la forma:

$$B/A = \cosh(ax/A) \pm (\sinh ax/A)(x^2 - 1)/x^2)^{1/2} \equiv f_{\pm}(x)$$

con $x = \cosh \beta$, \pm según el signo de β . Estudiando con cuidado el comportamiento de $f_{\pm}(x)$ probar que si $a \leq A \operatorname{Log} (B/A + (B^2/A^2 - 1)^{1/2})$, $B \geq A$, hay por lo menos dos soluciones.

3. Si $B = A$, hacer el estudio completo y probar que hay 2, 1 ó 0 soluciones según el valor de a .

Ejemplo 2.3 : Un problema isoperimétrico.

Si tenemos una cuerda de longitud L , amarrada en el origen, ¿cuál es el área máxima entre la cuerda y el eje x ? El extremo libre se desliza sobre el eje.

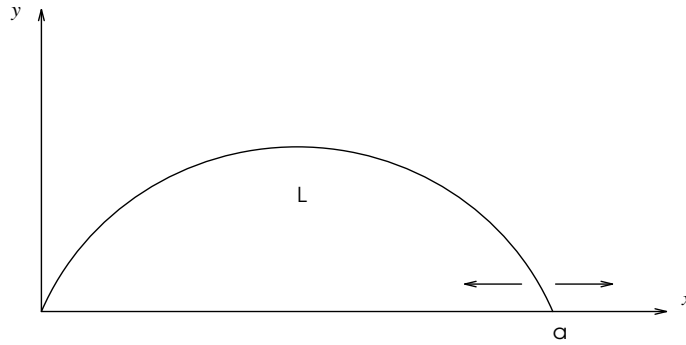


Fig. 2.3

Aquí el área es

$$A = \int_0^a y(x) dx,$$

donde a no se conoce, con la condición $L = \int_0^a (1 + y'^2)^{1/2} dx$.

Tomando como parámetro la longitud de arco, la curva se representa como $(x(s), y(s))$ con $\dot{x}(s)^2 + \dot{y}(s)^2 = 1$. Si suponemos que la curva puede ser parametrizada por x entonces $\dot{x}(s) \geq 0$, $\dot{x} = (1 - \dot{y}^2)^{1/2}$, $dx = \dot{x} ds$

$$A = \int_0^L y(s) \dot{x}(s) ds = \int_0^L y(s) (1 - \dot{y}^2)^{1/2} ds$$

con $y(0) = y(L) = 0$.

El integrando es C^2 si $|\dot{y}|_0 < 1$, $F_y(d, 0) = 1$, por lo tanto las constantes no son solución de la ecuación de Euler. Como $F_{zz} = -y(1 - \dot{y}^2)^{-3/2}$, los puntos críticos son C^2 cuando $y \neq 0$. La ecuación de Euler se reduce a

$$y\ddot{y} + (1 - \dot{y}^2) = 0$$

por lo tanto $\ddot{y} < 0$ si $y > 0$.

La integral primera $F - \dot{y}F_z = c$ da:

$$y(1 - \dot{y}^2)^{-1/2} = c$$

o sea

$$y = c(1 - \dot{y}^2)^{1/2}.$$

Por lo tanto y tiene un solo signo, \dot{y} es monótono: podemos usar a \dot{y} como parámetro y escribir $\dot{y} = \cos \theta$, con θ en $[0, \pi]$. Por lo tanto $y = c \sen \theta$.

$$\frac{ds}{d\theta} = \frac{dy}{d\theta} / \dot{y} = c$$

$$s = c\theta + \alpha, \theta = s/c + \beta.$$

$$y(s) = c \sen(s/c + \beta).$$

La condición $y(0) = 0$ implica $\beta = 0$, $y(L) = 0$ y $y(s) \geq 0$ implica $c = L/\pi$,

$$y(s) = L \sen(\pi s/L)/\pi.$$

Notemos que aquí $|\dot{y}(0)| = |\dot{y}(L)| = 1$: la condición $|\dot{y}| < 1$ no se cumple en los extremos. Sin embargo, al tomar h en $C_0^\infty[0, L]$, la derivada de Gâteaux en la dirección h está perfectamente definida.

Finalmente de $\dot{x}(s) = (1 - \dot{y}^2)^{1/2} = \sen(\pi s/L)$ y la condición $x(0) = 0$, tenemos

$$x(s) = (L/\pi)(1 - \cos(\pi s/L)).$$

Es decir, la solución es un semicírculo. (Todavía no hemos probado que es un máximo.)

Notemos que en el problema isoperimétrico del ejemplo I.1.5, podríamos argumentar de la manera siguiente: si la curva que da el área máxima es parametrizada por $(x(s), y(s))$, escogemos un segmento transversal tal que en la parte superior $x'(s) > 0$.

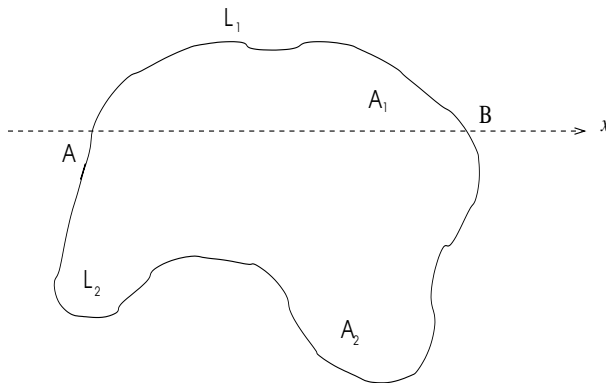


Fig. 2.4

Claramente A_1 es máximo para todas las curvas que tienen longitud L_1 y que pasan por A y B : si no fuera el caso se reemplaza la porción de longitud L_1 por otra curva dando un área mayor que A_1 . Hay que estudiar el problema de encontrar el máximo de

$$A(y) = \int_0^{L_1} y(s)(1 - y'^2)^{1/2} ds$$

con $y(0) = 0$, $y(L_1) = 0$,

$$\dot{x}^2 + \dot{y}^2 = 1 \quad \text{con} \quad x(0) = 0, x(L_1) = a.$$

Si no fuese por esta última condición $x(L_1) = a$, tendríamos el problema anterior cuya respuesta, el semi-círculo, no puede ser la solución al problema, ya que subiendo y bajando la transversal AB tendríamos siempre como respuesta un semi-círculo. Por lo tanto no hay que olvidar condiciones en el momento de plantear un problema; aquí la condición $x(L_1) = a$ es crucial: en este caso hay que usar multiplicadores de Lagrange sobre $x(s)$ y $y(s)$.

4.3 Varias funciones

Si el funcional depende de varias funciones y_1, \dots, y_n con y_i en $C^1[a, b]$, $y_i(a) = A_i$, $y_i(b) = B_i$,

$$J(y_1, \dots, y_n) = \int_a^b F(x, y_1, \dots, y_n, y_1', \dots, y_n') dx$$

con F en $C^1([a, b] \times \mathbb{R}^{2n})$, el problema no es mucho más complicado.

Aunque ya vimos algunos ejemplos, bastará recordar el problema de encontrar las geodésicas sobre una superficie dada en la forma $z = f(x, y)$. La longitud de una curva es entonces:

$$J(x, y) = \int_a^b (x'^2 + y'^2 + z'^2)^{1/2} dt = \int_a^b (x'^2 + y'^2 + (f_x x' + f_y y')^2)^{1/2} dt$$

ejemplo que estudiaremos más adelante en mayor generalidad.

En este caso las derivadas de Gâteaux (o de Fréchet) del funcional están dadas por (ver el ejemplo II.3.4):

$$D_{(h_1, \dots, h_n)} J(y_1, \dots, y_n) = \int_a^b \sum_{i=1}^n (F_{y_i} h_i + F_{z_i} h'_i) dx$$

donde h_i pertenece a $C^1[a, b] \cap C_0[a, b]$, ó a $C_0^\infty[a, b]$ si hay problemas de continuidad en a y b .

Si (y_1, \dots, y_n) es un punto crítico de J entonces podemos tomar variaciones (h_1, \dots, h_n) de la forma $(0, \dots, 0, h_i, 0, \dots, 0)$. En ese caso la derivada de Gâteaux se reduce a:

$$D_{(0, \dots, 0, h_i, 0, \dots, 0)} J(y_1, \dots, y_n) = \int_a^b (F_{y_i} h_i + F_{z_i} h'_i) dx = 0$$

para todo h_i en $C_0^\infty[a, b]$. Por el teorema III.1, obtenemos:

Teorema 3.1 :

1) En un punto crítico de $J(y_1, \dots, y_n)$, F_{z_i} es C^1 y satisface la ecuación de Euler:

$$\frac{d}{dx} (F_{z_i}(x, y_1(x), \dots, y_n(x), y'_1(x), \dots, y'_n(x))) = F_{y_i}(x, \dots, y'_n(x)),$$

para $i = 1, \dots, n$.

2) En un extremo local, si F no depende de x , entonces:

$$F(y_1(x), \dots, y'_n(x)) - \sum_{i=1}^n y'_i F_{z_i}(y_1(x), \dots, y'_n(x)) = c,$$

para alguna constante c .

Prueba. Sólo hay que probar (2). Notemos que sólo hay una ecuación integral y no n . Si y_i es C^2 , entonces:

$$\frac{d}{dx} (F) = \sum_1^n (y'_i F_{y_i} + y''_i F_{z_i}) = \sum_1^n (y'_i F_{z_i})'$$

al usar las ecuaciones de Euler. Dejamos al lector la prueba rigurosa en el siguiente :

Ejercicio 3.1 : Modificar la prueba del Teorema 2.1 para probar que, en general, en un extremo local:

$$(F - \sum_1^n y'_i F_{z_i})' = F_x.$$

Usar la transformación $x = \xi + \varepsilon \sum_1^n \eta_i$, $y_i = \eta_i$ (cuidado: η'_i aparece en los términos y'_j , $j \neq i$).

Ejercicio 3.2 : Modificar la prueba de la proposición 1.1, para demostrar que si F es C^2 , entonces un punto crítico (y_1, \dots, y_n) es C^2 en todos los puntos donde la matriz $(F_{z_i z_j})_{ij}$ es invertible.

Ejemplo 3.1 : Principio de mínima acción.

Aunque trataremos mucho más extensamente, en el siguiente capítulo, los principios variacionales de la Mecánica, podemos ya adelantar algunos de los fundamentos, en particular el principio de mínima acción.

Si tenemos un sistema con n “grados de libertad”, es decir descrito por n coordenadas Lagrangianas, $q_1(t), \dots, q_n(t)$ (por ejemplo para un sistema de N partículas, $n = 3N$ y q_i representan las posiciones de las partículas), la acción del sistema está dada por:

$$I(q_1, \dots, q_n) = \int_{t_0}^{t_1} L(t, q_1, \dots, q_n, q'_1, \dots, q'_n) dt$$

donde L es la Lagrangiana del sistema, la cual se escribe en general como:

$$L = T - U$$

donde T es la parte de *energía cinética*, T en general no depende de t . U es la parte de *energía potencial*. Si el sistema es *conservativo*, U no depende de q'_1, \dots, q'_n . El *principio de la mínima acción* decreta que el sistema físico, conocido en los tiempos t_0 y t_1 (por $q_i(t_0)$ y $q_i(t_1)$) evoluciona entre estos tiempos de tal modo que la acción sea mínima (en general se entiende más bien como un punto crítico). Por lo tanto las ecuaciones de Euler-Lagrange deben satisfacerse:

$$\frac{d}{dt}(L_{q'_i}) - L_{q_i} = 0, \quad i = 1, \dots, n.$$

Esto da lugar a un sistema de n ecuaciones no lineales del segundo orden. Veremos como usar ciertas transformaciones para dar en algunos casos una solución parcial.

Por ejemplo si $n = 1$, $T = mq'^2/2$, $U = U(q)$, la ecuación de Euler es:

$$(mq')' + U_q = 0$$

es decir la ecuación de Newton: $mq'' = -U_q$.

Si hay N partículas con

$$T = \sum_1^N m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2)/2,$$

$$U = U(x_1, y_1, z_1, \dots, x_n, y_n, z_n),$$

entonces:

$$m_i \ddot{x}_i = -U_{x_i}, \quad m_i \ddot{y}_i = -U_{y_i}, \quad m_i \ddot{z}_i = -U_{z_i}.$$

Notemos que si T y U no dependen del tiempo y el sistema es conservativo, tendremos una integral primera:

$$L - \sum q'_i L_{q'_i} = c = T - U - \sum q'_i T_{q'_i}.$$

Si además T es homogéneo de grado 2 en los q'_i :

$$T = \sum a_{ij}(q_1, \dots, q_n) q'_i q'_j$$

entonces, por la fórmula de Euler:

$$\sum q'_i T_{q'_i} = 2T.$$

La integral primera se escribe entonces:

$$T + U = -c,$$

es decir, la energía *total* se *conserva*.

Ejemplo 3.2 : Propagación de luz en un medio no homogéneo.

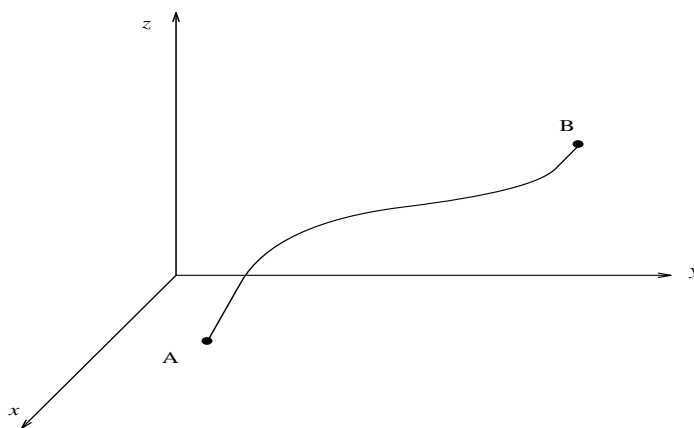


Fig. 3.1

Si un rayo de luz sale del punto A y llega al punto B , el recorrido del rayo de luz es tal que el principio de Fermat es satisfecho, es decir el tiempo de recorrido es mínimo. En cada punto (x, y, z) del medio, la luz tiene una velocidad de propagación $v(x, y, z)$. Supondremos que la trayectoria $(x(t), y(t), z(t))$ puede parametrizarse por x , es decir $x'(t) > 0$. En tal caso:

$$v(x, y, z) = \frac{ds}{dt} = (\dot{x}^2 + \dot{y}^2 + \dot{z}^2)^{1/2} = (1 + y'^2(x) + z'^2(x))^{1/2} \frac{dx}{dt},$$

por lo tanto el tiempo de recorrido es:

$$T(y(x), z(x)) = \int_{x_0}^{x_1} (1 + y'^2 + z'^2)^{1/2} v^{-1}(x, y, z) dx$$

con $(y(x), z(x))$ dadas en x_0 y x_1 .

Si $v \neq 0$, el funcional tiene una derivada de Fréchet y el recorrido de tiempo mínimo satisface:

$$\begin{aligned} \frac{d}{dx} \left(\frac{y'(1 + y'^2 + z'^2)^{-1/2}}{v} \right) + v_y \frac{(1 + y'^2 + z'^2)^{1/2}}{v^2} &= 0 \\ \frac{d}{dx} \left(\frac{z'(1 + y'^2 + z'^2)^{-1/2}}{v} \right) + v_z \frac{(1 + y'^2 + z'^2)^{1/2}}{v^2} &= 0. \end{aligned}$$

Estas ecuaciones son demasiado complicadas para poder resolverse en general. Notemos que si v no depende de x , la integral primera se reduce a:

$$v(1 + y'^2 + z'^2)^{1/2} = c,$$

mientras que si v depende sólo de x , es decir si el medio está estratificado en x , entonces $v_y = v_z = 0$, tendremos:

$$\begin{aligned} y'(1 + y'^2 + z'^2)^{-1/2} v^{-1} &= c_1 \\ z'(1 + y'^2 + z'^2)^{-1/2} v^{-1} &= c_2 \end{aligned}$$

y' y z' tienen un solo signo.

Por lo tanto $y'/z' = c_1/c_2$, es decir,

$$c_2 y(x) = c_1 z(x) + c.$$

El rayo de luz está en el plano que une los puntos A y B y que es paralelo al eje x . Si escogemos los ejes tales que A esté en el origen y $z(x_1) = 0$, entonces hay un punto entre x_0 y x_1 tal que $z'(x) = 0$, por lo tanto $c_2 = 0$, $z(x) \equiv 0$.

La ecuación en y se escribe:

$$y'(x) = c_1 v(x)(1 - c_1^2 v^2)^{-1/2}$$

(ya que y' tiene el signo de c_1), la cual es integrable si v es conocido. Por ejemplo, si v es constante, y' es constante y la propagación es en línea recta.

Ejemplo 3.3 : Geodésicas sobre una superficie $M(u, v)$.

Como se vió en el ejemplo I.1.4, si la superficie está parametrizada por (u, v) , dando puntos $M(u, v)$ en \mathbb{R}^3 , la geodésica se escribe como $M(u(t), v(t)) \equiv M(t)$.

Su longitud es entonces:

$$\begin{aligned} L(u, v) &= \int_{t_0}^{t_1} \|M'(t)\| dt = \int_{t_0}^{t_1} \|M_u u' + M_v v'\| dt \\ &= \int_{t_0}^{t_1} (Eu'^2 + 2Fu'v' + Gv'^2)^{1/2} dt \end{aligned}$$

donde $E = \|M_u\|^2$, $F = (M_u, M_v)$, $G = \|M_v\|^2$.

Supondremos que $M(u, v)$ es C^2 , que M_u no es paralelo a M_v (es decir que la normal $N(u, v) = M_u \times M_v \neq 0$, la superficie es regular) y que la curva es regular: $(u', v') \neq (0, 0)$. En tal caso, $M' \neq 0$ y el integrando es C^1 . Por lo tanto las ecuaciones de Euler son:

$$\frac{d}{dt} \left(\frac{Eu' + Fv'}{\|M'\|} \right) - \frac{E_u u'^2 + 2F_u u'v' + G_u v'^2}{2\|M'\|} = 0 \quad (1)$$

$$\frac{d}{dt} \left(\frac{Fu' + Gv'}{\|M'\|} \right) - \frac{E_v u'^2 + 2F_v u'v' + G_v v'^2}{2\|M'\|} = 0 \quad (2)$$

Evidentemente estas ecuaciones son difíciles de resolver, sin embargo veremos algunos casos particulares y una interpretación geométrica.

Caso 1: Cilindro. Si $M(\phi, z) = (a \cos \phi, a \sin \phi, z)$ entonces $E = a^2$, $F = 0$, $G = 1$, las ecuaciones se integran:

$$\frac{a^2 \phi'}{\|M'\|} = c_1, \quad \frac{z'}{\|M'\|} = c_2$$

ϕ' y z' tienen uno solo signo, además $z'/\phi' = c$, es decir $z = c\phi + d$.

El candidato a geodésica es una hélice. Si $z(0) = \phi(0) = 0$, entonces $d = 0$, sin embargo, como ϕ puede cambiarse por $\phi + 2k\pi$, las posibles c no son únicas: hay que escoger la hélice que llegue al punto final en menos de medio paso para tener la geodésica.

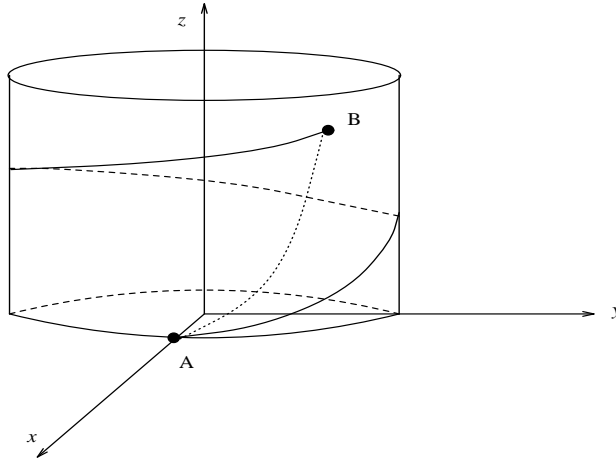


Fig. 3.2

Caso 2: Geodésica en una esfera. Dejaremos este caso al lector como ejercicio :

Ejercicio 3.3 : Escribir las ecuaciones (1) y (2) en coordenadas esféricas y en coordenadas cartesianas y resolverlas.

Interpretación geométrica : Si (1) y (2) tienen solución, sea s la longitud de arco de la geodésica: $ds/dt = \|M'\|$

$$\dot{u} = du/ds = u'/\|M'\|$$

$$\dot{v} = dv/ds = v'/\|M'\|$$

$$\frac{d}{dt}\left(\frac{u'}{\|M'\|}\right) = \frac{d}{ds}(\dot{u})\|M'\| = \ddot{u}\|M'\|.$$

La ecuación (1) se escribe:

$$\begin{aligned} \frac{d}{ds}(E\dot{u} + F\dot{v}) \|M'\| &= (E_u\dot{u}^2 + (E_v + F_u)\dot{u}\dot{v} + F_v\dot{v}^2 + E\ddot{u} + F\ddot{v}) \|M'\| \\ &= \|M'\|(E_u\dot{u}^2 + 2F_u\dot{u}\dot{v} + G_u\dot{v}^2)/2. \end{aligned}$$

Por lo tanto:

$$E\ddot{u} + F\ddot{v} + E_u\dot{u}^2/2 + E_v\dot{u}\dot{v} + (F_v - G_u/2)\dot{v}^2 = 0.$$

Pero

$$E_u = (M_u, M_u)_u = 2(M_{uu}, M_u)$$

$$E_v = 2(M_{uv}, M_u)$$

$$F_v - G_u/2 = (M_{vv}, M_u).$$

Por lo tanto (1) se reduce a:

$$(M_u \ddot{u} + M_v \ddot{v} + M_{uu} \dot{u}^2 + 2M_{uv} \dot{u} \dot{v} + M_{vv} \dot{v}^2, M_u) = 0,$$

es decir

$$(\ddot{M}, M_u) = 0.$$

Del mismo modo (2) se escribe: $(\ddot{M}, M_v) = 0$. Por lo tanto las dos ecuaciones de Euler se interpretan como el hecho que la normal principal a la curva, \ddot{M} , es ortogonal a M_u y a M_v , por lo tanto es colineal a la normal a la superficie.

En el caso de la esfera, \ddot{M} es radial. Derivando la relación $(M, M) = r^2$, obtenemos $(M, \dot{M}) = 0$; $(M, \ddot{M}) = -(\dot{M}, \dot{M}) = -1$ por lo tanto $\|\ddot{M}\| = 1/r$, la curvatura es constante. Como $M \times \ddot{M} = 0$, derivando tenemos $(\dot{M}, \ddot{M}, \ddot{\ddot{M}}) = 0$, es decir la torsión de la curva es 0. La curva está en un plano, con radio de curvatura igual al radio de la esfera: es un círculo principal.

Nota 3.1: Quizás el lector se haya extrañado que no hayamos usado en el ejemplo anterior la integral primera, ya que F no depende explícitamente de t . Un cálculo rápido lo convencerá que $F - u'F_{u'} - v'F_{v'} \equiv 0$.

En efecto la función F es homogénea de grado 1 en u' y v' , por lo tanto la relación anterior es únicamente la identidad de Euler. De hecho supongamos que:

$$F(y_1, \dots, y_n, y'_1, \dots, y'_n)$$

es homogénea de grado 1 (positivamente) en y'_1, \dots, y'_n , entonces la identidad de Euler es justamente: $\sum y'_i F_{y'_i} = F$.

En este caso no solamente la integral primera se cumple independientemente de las ecuaciones de Euler, sino las ecuaciones de Euler no son independientes: llamemos en efecto

$$Euler(i) = (F_{z_i})' - F_{y_i}$$

entonces si y_i es C^2 :

$$\sum (y'_i (F_{z_i})' - y'_i F_{y_i}) = (\sum y'_i F_{z_i})' - \sum (y'_i F_{y_i} + y''_i F_{z_i})$$

$$= F' - F' = 0,$$

donde se usó la identidad de Euler y se derivó $F(y_1, \dots, y'_n)$.

Por lo tanto

$$\sum y'_i \text{Euler}(i) = 0,$$

aun si no se cumplen las ecuaciones de Euler. Veremos más adelante otras consecuencias de esta falta de independencia.

4.4 Varias variables

Como ya lo hemos visto, en problemas de medios continuos, el funcional puede depender de varias variables. Por ejemplo

$$J(\phi) = \iint_{\Omega} F(x, y, \phi, \phi_x, \phi_y) dx dy$$

donde $\Omega \subset \mathbb{R}^2$ es un dominio acotado para el cual el teorema de la divergencia es válido. Aquí ϕ está en $C^1(\overline{\Omega})$, con norma $|\cdot|_1$, ϕ está dado sobre $\partial\Omega$ y F es C^1 . En esas condiciones, vimos en el ejemplo II.3.4, que $J(\phi)$ tiene una diferencial de Fréchet:

$$DJ(\phi)h = \iint_{\Omega} (F_{\phi}h + F_{\phi_x} h_x + F_{\phi_y} h_y) dx dy$$

para h en $C_0^{\infty}(\Omega)$.

En un punto crítico esa derivada vale 0, y, sea usando las consideraciones después del lema de Haar, en el capítulo III, sea suponiendo que ϕ es C^2 así como F , tenemos:

$$F_{\phi_x} h_x + F_{\phi_y} h_y = (F_{\phi_x} h)_x + (F_{\phi_y} h)_y - h((F_{\phi_x})_x + (F_{\phi_y})_y).$$

Usando el teorema de la divergencia:

$$\iint_{\Omega} (F_{\phi} - (F_{\phi_x})_x - (F_{\phi_y})_y) h dx dy + \int_{\partial\Omega} h \begin{pmatrix} F_{\phi_x} \\ F_{\phi_y} \end{pmatrix} \cdot n ds = 0.$$

Como $h = 0$ en $\partial\Omega$, podemos concluir, a partir del lema II.8,

Teorema 4.1 : Si ϕ es un punto crítico de $J(\phi)$, con ϕ y F en C^2 , ϕ dado en $\partial\Omega$, entonces la ecuación de Euler es:

$$F_{\phi} - \frac{\partial}{\partial x}(F_{\phi_x}) - \frac{\partial}{\partial y}(F_{\phi_y}) = 0 \text{ en } \Omega.$$

Ejemplo 4.1 : Membrana elástica.

Retomemos el ejemplo I.1.8, donde la posición de equilibrio de una membrana elástica, fija en la frontera, se puede encontrar minimizando la siguiente aproximación de la energía potencial:

$$I(\phi) = \iint_{\Omega} (\phi_x^2 + \phi_y^2) dx dy / 2,$$

con $\phi(x, y) = \phi_0(x, y)$ en $\partial\Omega$.

La ecuación de Euler en este caso es:

$$\Delta\phi = 0.$$

Notemos que cualquier punto crítico es un *mínimo*. De hecho

$$I(\phi + h) - I(\phi) = I(h) + \iint_{\Omega} (\phi_x h_x + \phi_y h_y) dx dy.$$

Como la derivada de Fréchet es 0 en un punto crítico, la diferencia es $I(h)$. Además, si $I(\phi + h) = I(\phi)$ para algún h , con $h = 0$ en $\partial\Omega$, entonces $\nabla h = 0$, $h = \text{constante}$ ($= 0$ en la frontera). El mínimo, si existe, es único.

Notemos que si usamos el lema de Haar obtendremos un potencial $V(x, y)$ tal que:

$$\begin{aligned}\phi_x &= -V_y \\ \phi_y &= V_x\end{aligned}$$

es decir las ecuaciones de Cauchy-Riemann.

Finalmente hay que señalar que falta probar la existencia del mínimo. Este punto, obvio para muchos matemáticos del siglo pasado, empezando por Dirichlet, fue criticado por Weierstrass de tal forma que los métodos variacionales para el problema de Laplace fueron desacreditados hasta que Hilbert, a principios de este siglo, les dió una base rigurosa y satisfactoria. En la actualidad no podemos ya ignorar las críticas fingiendo que no sabemos que a veces no hay una solución. En estos casos hay dos posibles actitudes: sea probar lo que se dice o, reconociendo que pueda haber una dificultad, asumir el punto (en este caso la existencia) explícitamente.

Ejemplo 4.2 : Superficie de área mínima, con frontera dada.

Veremos ahora otro caso particular del problema de Plateau: ¿cuál es la superficie de la forma $z = f(x, y)$, definida para (x, y) en Ω , un dominio acotado en \mathbb{R}^2 , con $f(x, y) = f_0(x, y)$ en $\partial\Omega$, y de área mínima?

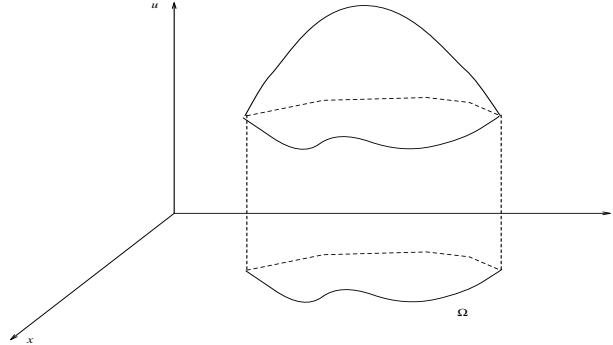


Fig. 4.1

El área está dada por:

$$A(f) = \iint_{\Omega} (1 + f_x^2 + f_y^2)^{1/2} dx dy$$

el cual es un funcional Fréchet diferenciable sobre $C^1(\bar{\Omega})$. La ecuación de Euler, lleva a la siguiente relación (después de simplificaciones y si f es C^2):

$$f_{xx}(1 + f_y^2) - 2f_{xy}f_xf_y + f_{yy}(1 + f_x^2) = 0.$$

Geométricamente, esto significa que la curvatura media de la superficie es 0. Esta ecuación es cuasilineal (i.e. es lineal en las segundas derivadas) de tipo elíptico (es decir, de la forma $a f_{xx} + 2b f_{xy} + c f_{yy} = 0$ con $b^2 - ac < 0$). Aquí $b^2 - ac = -(1 + f_x^2 + f_y^2) \leq -1$.

El estudio de esta ecuación rebasa las pretensiones de estas notas. Lo único que podemos decir es que si pensamos que $|f_x|, |f_y|$ son muy pequeñas comparadas a las segundas derivadas, esta ecuación puede ser aproximada por la ecuación: $\Delta f = 0$.

El paso siguiente sería entonces justificar que las soluciones a la ecuación inicial existen y son vecinas a la función armónica correspondiente. Esto se puede hacer (de hecho la demostración es una aplicación de un teorema de punto fijo). Notemos que con este ejemplo podemos ver más claro el por qué de la analogía entre superficies de área mínima y de mínima tensión, es decir las pompas de jabón.

Ejemplo 4.3 : Cuerda vibrante.

Consideremos una cuerda de densidad constante ρ y longitud en reposo ℓ . Si se estira la cuerda, tendremos la posición de la figura 4.2. El elemento de longitud $\Delta s =$

$(1 + u_x^2)^{1/2} \Delta x$ y masa $\rho \Delta s$, contribuye a la energía cinética: $\rho \Delta s u_t^2 / 2$, si suponemos que no hay desplazamiento horizontal (es decir $x_t = 0$ o muy pequeño comparado a u_t).

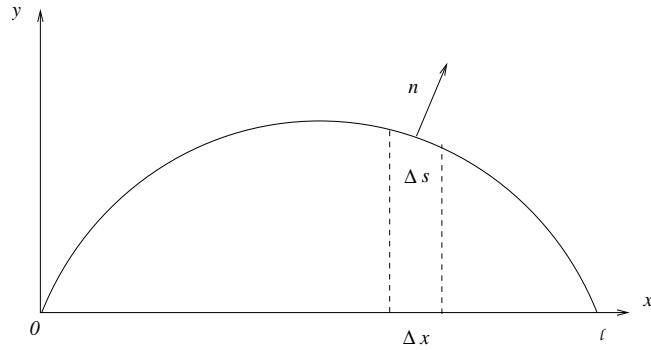


Fig. 4.2

Del mismo modo, si suponemos que la energía potencial es debida únicamente a la elongación de la cuerda (no hay deflexión o torsión), entonces la ley de Hooke indica que el incremento de energía potencial debido al elemento Δs es proporcional a Δs .

Por lo tanto, la energía cinética de la cuerda es:

$$\int_0^\ell \frac{1}{2} \rho u_t^2 (1 + u_x^2)^{1/2} dx$$

y la energía potencial:

$$\int_0^\ell k (1 + u_x^2)^{1/2} dx$$

en cada instante t .

Si aplicamos el principio de mínima acción en el dominio $[0, \ell] \times [0, T]$, suponiendo que los extremos de la cuerda son fijos, $u(0, t) = u(\ell, t) = 0$ y que las posiciones de la cuerda al tiempo $t = 0$ y al tiempo $t = T$ son conocidas:

$$J(u) = \int_0^T \int_0^\ell (\rho u_t^2 / 2 - k) (1 + u_x^2)^{1/2} dx dt$$

debe ser mínimo en la posición de “equilibrio” (es decir en x y en t). Después de simplificar, la ecuación de Euler es:

$$\rho u_{tt} (1 + u_x^2)^2 + 2 \rho u_t u_x u_{tx} (1 + u_x^2) + (\rho u_t^2 / 2 - k) u_{xx} = 0$$

la cual es otra vez una ecuación cuasilineal. Aquí $b^2 - ac = \rho(1 + u_x^2)^2 (k + \rho u_t^2 u_x^2 - \rho u_t^2 / 2)$ es positivo si k es grande comparado a u_t y a u_x . En este caso la ecuación es de tipo hiperbólico.

Si suponemos que u_t, u_x son muy pequeños comparados a las constantes, tenemos la aproximación (que se debe justificar)

$$\rho u_{tt} - k u_{xx} = 0$$

es decir la ecuación de onda, con velocidad de propagación $(k/\rho)^{1/2}$.

Notemos que hay una manera más elegante y más clara de hacer esta aproximación: si consideramos que u_x será pequeña, la densidad de la energía cinética

$$\rho u_t^2 (1 + u_x^2)^{1/2} / 2$$

puede aproximarse por $\rho u_t^2 / 2$, mientras que la densidad de energía potencial, quitándole la densidad de una energía potencial en reposo k :

$$k((1 + u_x^2)^{1/2} - 1)$$

puede aproximarse por $ku_x^2 / 2$. La nueva Lagrangiana será $\rho u_t^2 / 2 - ku_x^2 / 2$, con ecuación de Euler: $\rho u_{tt} - k u_{xx} = 0$.

El saber cual término descartar y cual conservar es un arte, sobre todo si se quiere conservar una ecuación no lineal. Esto depende mucho de la experiencia previa y de algunas consideraciones técnicas (demasiado pocas). En la práctica muy pocas de estas aproximaciones son (o pueden ser) justificadas. Sin embargo, el hacer la aproximación antes de usar la ecuación de Euler ayuda a no descartar términos importantes.

Para resolver la ecuación de onda

$$u_{tt} - u_{xx} = 0$$

$$u(0, t) = u(1, t) = 0$$

(pondremos $k/\rho = 1, \ell = 1$), podemos usar el método de separación de variables, buscando soluciones de la forma $u(x, t) = v(x)w(t)$ o sumas de expresiones de esta forma.

Tendremos $vw'' - v''w = 0$ o sea:

$$\frac{v''(x)}{v(x)} = \frac{w''(t)}{w(t)} = -k$$

ya que tenemos dos funciones, dependiendo de variables distintas, y que son iguales. La solución al problema

$$v'' + kv(x) = 0 \quad v(0) = v(1) = 0$$

es entonces: $v(x) = A_n \sin n\pi x$, con $k = n^2\pi^2$. Para $w(t) = B_n \sin(n\pi t + \phi_n)$.

Si conocemos $u(x, 0)$ y $u(x, T)$ podemos determinar las constantes A_n, B_n, ϕ_n (no de manera única) buscando una solución de la forma:

$$u(x, t) = \sum_1^{\infty} C_n \operatorname{sen} n\pi x \operatorname{sen} (n\pi t + \phi_n).$$

Supongamos que $u(x, 0) = 0$, $T = 1$ y $u(x, 1) = 0$ entonces $w(0) = w(1) = 0$, $w(t) = B_n \operatorname{sen} n\pi t$. Notemos que en ese caso $I(u) = \int_0^1 \int_0^1 (u_t^2 - u_x^2) dx dt = 0$ y que ningún punto crítico es un mínimo.

De hecho usando la ecuación de Euler:

$$I(u + h) - I(u) = \int_0^1 \int_0^1 (h_t^2 - h_x^2) dx dt = \varepsilon^2 \pi^2 (n^2 - m^2)/4$$

si tomamos $h(x, t) = \varepsilon \operatorname{sen} n\pi t \operatorname{sen} m\pi x$. Esta cantidad es positiva si $n > m$, negativa si $n < m$. Ahora bien si separamos variables antes de aplicar el principio variacional y consideramos el funcional:

$$I(v, w) = \int_0^1 \int_0^1 (v^2 w'^2 - v'^2 w^2) dx dt$$

tendremos como ecuaciones de Euler:

$$\begin{aligned} v''w^2 + vw'^2 &= 0, & v(0) &= v(1) = 0 \\ wv'^2 + v^2w'' &= 0, & w(0) &= w(1) = 0. \end{aligned}$$

Por lo tanto:

$$\begin{aligned} \frac{v''(x)}{v(x)} &= -\frac{w'^2(t)}{w^2(t)} = -k_1 \\ \frac{w''(t)}{w(t)} &= -\frac{v'^2(x)}{v^2(x)} = -k_2. \end{aligned}$$

Tendremos

$$\begin{aligned} v(x) &= A \operatorname{sen} n\pi x & \text{con } k_1 &= n^2 \pi^2 \\ w(t) &= B \operatorname{sen} m\pi t & \text{con } k_2 &= m^2 \pi^2 \end{aligned}$$

pero entonces $w'^2/w^2 = m^2 \pi^2 (\cot m\pi t)^2 = k_1$ no puede ser cierto al menos que $m = 0$, $k_1 = 0$, $n = 0$. Es decir el único punto crítico en este caso es $u = 0$.

Esto quiere decir que separar variables antes o después de aplicar el principio variacional no lleva al mismo resultado.

Notemos finalmente que si escribimos

$$\begin{aligned} -I(u) &= \int_0^1 \int_0^1 (v'^2 w^2 - v^2 w'^2) dx dt \\ &= \int_0^1 \left(\left(\int_0^1 v'^2 dx \right) w^2 - \left(\int_0^1 v^2 dx \right) w'^2 \right) dt = \int_0^1 (Aw^2 - Bw'^2) dt \end{aligned}$$

donde A y B son constantes si v es dado. Entonces la ecuación de Euler para w es:

$$Aw + Bw'' = 0, \quad w(0) = w(1) = 0.$$

Por lo tanto $w(t) = c \sin n\pi t$ con $A/B = n^2\pi^2$. Notemos que si escribimos $v(x) = \sum_1^\infty a_m \sin m\pi x$, entonces $A = \sum_1^\infty m^2\pi^2 a_m^2/2$, $B = \sum_1^\infty a_m^2/2$.

Si tomamos $a_m = 0$, para $m \neq m_0, m_1$, con $m_0 < n < m_1$, $a_{m_0} = (m_1^2 - n^2)^{1/2}$, $a_{m_1} = (n^2 - m_0^2)^{1/2}$, entonces se cumple la relación $A = n^2\pi^2 B$.

Si integramos primero con respecto t , obtenemos $Cv + Dw'' = 0$, $v = d \sin m\pi x$ con $C/D = m^2\pi^2$, en este caso $n = m$. Es decir si, después de integrar con respecto a la otra variable, buscamos los puntos críticos, obtenemos el mismo resultado que separando las variables en la ecuación de Euler para u .

Nota 4.1 : Si tenemos un funcional que depende de varias funciones, por ejemplo:

$$J(\phi_1, \phi_2) = \iint_{\Omega} F(x, y, \phi_1, \phi_2, \phi_{1x}, \phi_{1y}, \phi_{2x}, \phi_{2y}) dx dy$$

con ϕ_1, ϕ_2 dadas en $\partial\Omega$ y F es C^1 , y si ϕ_1, ϕ_2 son C^2 , tendremos dos ecuaciones de Euler (tomando incrementos h_1, h_2):

$$\frac{\partial}{\partial x}(F_{\phi_{ix}}) + \frac{\partial}{\partial y}(F_{\phi_{iy}}) - F_{\phi_i} = 0, \quad i = 1, 2.$$

4.5 Derivadas de orden más alto

En algunos problemas, el funcional depende de derivadas de orden más alto, por ejemplo:

$$J(y) = \int_a^b F(x, y(x), y'(x), y''(x)) dx.$$

Aquí el espacio de trabajo es un subconjunto de $C^2[a, b]$, con norma $\|\cdot\|_2$. Supondremos que $y(a) = A$, $y(b) = B$, $y'(a) = \alpha$, $y'(b) = \beta$ son dadas. Si F es C^1 en $[a, b] \times \mathbb{R}^3$, y_0 es el candidato a extremo o a punto crítico, todo y se escribe

$$y(x) = y_0(x) + h(x)$$

con $h(x)$ en $C^2[a, b] \cap C_0^1[a, b]$.

$$J(y_0 + h) - J(y_0) = \int_a^b (F_y h + F_{y'} h' + F_{y''} h'') dx + o(\|h\|_2).$$

En un punto crítico:

$$DJ(y_0)h = \int_a^b (F_y h + F_{y'} h' + F_{y''} h'') dx$$

para todo h en $C^2[a, b] \cap C_0^1[a, b]$, o si hay problemas de integrabilidad en los extremos, para todo h en $C_0^\infty[a, b]$.

Por el lema III.7, tendremos que:

Teorema 5.1 : *En un punto crítico, $F_{y''}$ es C^1 , $\frac{d}{dx}(F_{y''}) - F_{y'}$ es C^1 y:*

$$\frac{d}{dx} \left[\frac{d}{dx} (F_{y''}) - F_{y'} \right] = -F_y.$$

Si el punto crítico es C^4 y F es C^3 , la ecuación de Euler se escribe:

$$(F_{y''})'' - (F_{y'})' + F_y = 0.$$

Ejercicio 5.1 :

1. Probar que si F es C^2 , entonces y_0 es C^3 en los puntos donde $F_{y''y''} \neq 0$ (usar el hecho que $F_{y''}$ es C^1 y el mismo argumento que en la proposición 1.1).
2. Probar que si F es C^3 , entonces y_0 es C^4 en todos los puntos donde $F_{y''y''} \neq 0$.

Ejemplo 5.1 : Barra elástica pesada.

Si a una barra de longitud ℓ en reposo y empotrada en dos paredes, le aplicamos una fuerza vertical, esta se deformará hasta guardar un equilibrio.

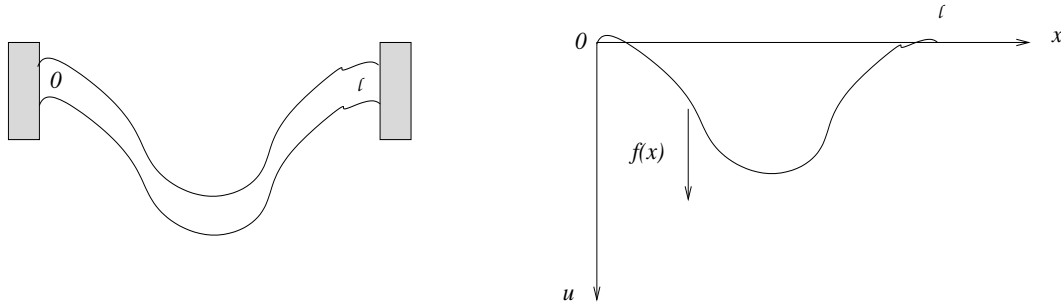


Fig. 5.1

Idealizamos la barra y suponemos que la línea media representa a la barra. (Este pasaje de un problema tridimensional a uno de una dimensión no está aún completamente justificado. En varios libros de elasticidad, las hipótesis que se usan son de hecho contradictorias.)

Si suponemos como antes que no hay desplazamiento horizontal ($x_t = 0$), la energía cinética queda dada por:

$$\int_0^\ell \rho u_t^2 (1 + u_x^2)^{1/2} dx / 2$$

en el tiempo t .

La energía potencial tiene dos componentes: la energía potencial interna, proporcional al cuadrado de la curvatura, con densidad $\frac{k}{2}(u_{xx}(1 + u_x^2)^{-3/2})^2 ds$

$$\int_0^\ell \frac{k}{2} u_{xx}^2 (1 + u_x^2)^{-5/2} dx / 2$$

y la energía potencial externa, debida al trabajo de la fuerza por unidad de longitud, $f(x)$:

$$- \int_0^\ell u f (1 + u_x^2)^{1/2} dx$$

(tomando en cuenta el desplazamiento del elemento Δs).

Si suponemos que la barra es fija en los extremos, tendremos

$$u(0, t) = u_x(0, t) = u(\ell, t) = u_x(\ell, t) = 0$$

(las condiciones sobre u_x implican que la barra es ortogonal a las paredes) y conociendo a $u(x, 0)$ y $u(x, T)$, el principio de mínima acción, requiere estudiar los puntos críticos del funcional:

$$2J(u) = \int_0^T \int_0^\ell [(\rho u_t^2 + 2uf)(1 + u_x^2)^{1/2} - k u_{xx}^2 (1 + u_x^2)^{-5/2}] dx dt.$$

Antes de abordar el problema dependiente del tiempo, veremos primero el caso estacionario, donde u no depende del tiempo. Entonces

$$J(u) = T \int_0^\ell (uf(1+u'^2)^{1/2} - ku''^2(1+u'^2)^{-5/2}/2) dx$$

con ecuación de Euler:

$$\begin{aligned} f(1+u'^2)^{1/2} &- (uu'f(1+u'^2)^{-1/2})' \\ &- (5ku'u^{1/2}(1+u'^2)^{-7/2}/2)' - (ku''(1+u'^2)^{5/2})'' = 0. \end{aligned}$$

Como aún esta ecuación es demasiado complicada, haremos la simplificación de reemplazar $1+u_x^2$ por 1 en la Lagrangiana:

$$J(u) = T \int_0^\ell (uf - ku''^2/2) dx$$

con ecuación de Euler:

$$f - ku^{(4)}(x) = 0,$$

por lo tanto,

$$u(x) = \int_0^x \int_0^{x_1} \int_0^{x_2} \int_0^{x_3} f(x_4)/k dx_4 dx_3 dx_2 dx_1 + \alpha x^3 + \beta x^2 + \gamma x + \delta.$$

Con las condiciones $u(0) = 0$, $u'(0) = 0$ tenemos $\delta = \gamma = 0$. Con las condiciones $u(\ell) = u'(\ell) = 0$, obtendremos α y β . Por ejemplo si la fuerza es la gravedad, $f(x) = \rho g$, entonces:

$$u(x) = \frac{\rho g}{k4!}(x - \ell)^2 x^2.$$

Para el caso no estacionario, veremos primero cual es la ecuación de Euler para un problema de dimensión dos y con derivadas segundas. Sea F en C^1 y veamos el funcional:

$$J(u) = \iint_{\Omega} F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) dx dy$$

donde Ω es acotado, u en $C^2(\overline{\Omega})$.

Entonces

$$J(u+h) - J(u) = \iint_{\Omega} (F_u h + F_{u_x} h_x + F_{u_y} h_y + F_{u_{xx}} h_{xx} + F_{u_{xy}} h_{xy} + F_{u_{yy}} h_{yy}) dx dy + o(|h|_2)$$

ya que Ω es acotado.

La derivada de Fréchet de J es:

$$DJ(u)h = \iint_{\Omega} (F_u h + F_{u_x} h_x + F_{u_y} h_y + F_{u_{xx}} h_{xx} + F_{u_{xy}} h_{xy} + F_{u_{yy}} h_{yy}) dx dy$$

Si u es C^4 , F es C^3 y se puede aplicar el teorema de la divergencia, podemos escribir el integrando en la forma:

$$\begin{aligned} & \frac{\partial}{\partial x}(F_{u_x} h + F_{u_{xx}} h_x + F_{u_{xy}} h_y / 2) + \frac{\partial}{\partial y}(F_{u_y} h + F_{u_{xy}} h_x / 2 + F_{u_{yy}} h_y) - \\ & \frac{\partial}{\partial x}(F_{u_x}) h - \frac{\partial}{\partial y}(F_{u_y}) h - \frac{\partial}{\partial x} \left[\left(\frac{\partial}{\partial x}(F_{u_{xx}}) + \frac{\partial}{\partial y}(F_{u_{xy}} / 2) \right) h \right] - \\ & \frac{\partial}{\partial y} \left[\left(\frac{\partial}{\partial y}(F_{u_{xy}}) + \frac{\partial}{\partial x}(F_{u_{yy}} / 2) \right) h \right] + \\ & \left[\frac{\partial^2}{\partial x^2}(F_{u_{xx}}) + \frac{\partial^2}{\partial x \partial y}(F_{u_{xy}}) + \frac{\partial^2}{\partial y^2}(F_{u_{yy}}) \right] h. \end{aligned}$$

(Dejamos al lector verificar estas agrupaciones.)

Usando el teorema de la divergencia obtenemos

$$\begin{aligned} DJ(u)h = & \iint_{\Omega} \left[F_u - \frac{\partial}{\partial x}(F_{u_x}) - \frac{\partial}{\partial y}(F_{u_y}) + \frac{\partial^2}{\partial x^2}(F_{u_{xx}}) + \frac{\partial^2}{\partial x \partial y}(F_{u_{xy}}) + \frac{\partial^2}{\partial y^2}(F_{u_{yy}}) \right] h dx dy + \\ & \int_{\partial \Omega} \left\{ h \left(F_{u_x} - \frac{\partial}{\partial x}(F_{u_{xx}}) - \frac{\partial}{\partial y}(F_{u_{xy}}) / 2 \right) \cdot \vec{n} + \left(F_{u_{xx}} \quad F_{u_{xy}} / 2 \right) \begin{pmatrix} h_x \\ h_y \end{pmatrix} \cdot \vec{n} \right\} ds. \end{aligned}$$

Si u es un punto crítico de J para una clase de funciones h que contiene a $C_0^\infty(\Omega)$, entonces h , h_x , h_y son 0 en $\partial \Omega$, los términos de frontera desaparecen y el lema III-8 da la ecuación de Euler:

$$F_u - \frac{\partial}{\partial x}(F_{u_x}) - \frac{\partial}{\partial y}(F_{u_y}) + \frac{\partial^2}{\partial x^2}(F_{u_{xx}}) + \frac{\partial^2}{\partial x \partial y}(F_{u_{xy}}) + \frac{\partial^2}{\partial y^2}(F_{u_{yy}}) = 0.$$

Ahora bien, para uso futuro, podemos expresar los términos (h_x, h_y) en la frontera en términos de las derivadas normales y tangenciales, h_n y h_T , de la forma siguiente:

$$h_n = \nabla h \cdot n, \quad h_T = \nabla h \cdot T = \frac{d}{ds}(h(x(s), y(s)))$$

donde $n = (n_1, n_2)$, $T = (-n_2, n_1)$, con $n_1^2 + n_2^2 = 1$.

$$\begin{pmatrix} h_T \\ h_n \end{pmatrix} = \begin{pmatrix} -n_2 & n_1 \\ n_1 & n_2 \end{pmatrix} \begin{pmatrix} h_x \\ h_y \end{pmatrix}$$

y por lo tanto,

$$\begin{pmatrix} h_x \\ h_y \end{pmatrix} = \begin{pmatrix} -n_2 & n_1 \\ n_1 & n_2 \end{pmatrix} \begin{pmatrix} h_T \\ h_n \end{pmatrix}.$$

Integrando por partes el término $h_T = \frac{d}{ds}(h)$ y notando que no hay términos de frontera ya que $\partial\Omega$ es una curva cerrada, tenemos que la integral sobre $\partial\Omega$ es:

$$\begin{aligned} \int_{\partial\Omega} \{ (F_{u_x} - \frac{\partial}{\partial x}(F_{u_{xx}}) - \frac{\partial}{\partial y}(F_{u_{xy}})/2)n_1 + (F_{u_y} - \frac{\partial}{\partial y}(F_{u_{yy}}) - \frac{\partial}{\partial x}(F_{u_{xy}})/2)n_2 \\ + \frac{d}{ds}[(F_{u_{xx}} - F_{u_{yy}})n_1n_2 - F_{u_{xy}}(n_1^2 - n_2^2)/2] \} h \, ds + \\ \int_{\partial\Omega} (F_{u_{xx}}n_1^2 + F_{u_{xy}}n_1n_2 + F_{u_{yy}}n_2^2)h_n \, ds. \end{aligned}$$

Vemos que si $h = 0$ en $\partial\Omega$ (por lo tanto $h_T = 0$) y $h_n = 0$ entonces esta integral sobre $\partial\Omega$ desaparece.

Ejercicio 5.2 : En el caso de la barra elástica pesada no estacionaria, con las condiciones: $u(0, t) = u(\ell, t) = u_x(0, t) = u_x(\ell, t) = 0$, $u(x, 0)$, $u(x, T)$ dados:

1. Encontrar la ecuación de Euler y verificar que los términos de frontera no aparecen.
2. Si se hace la sustitución $(1 + u_x^2)$ por 1, encontrar la ecuación de Euler lineal correspondiente.

Nota 5.1 : Si tenemos un funcional que depende de $y(x)$ y de sus derivadas hasta orden n :

$$J(y) = \int_a^b F(x, y, y', \dots, y^{(n)}(x)) \, dx$$

donde $y(x)$ está en $C^n[a, b]$, y sus derivadas hasta orden $n-1$ son dadas en a y en b , entonces si F es C^1 en $[a, b] \times \mathbb{R}^{n+1}$, $J(y)$ tiene como derivada de Fréchet

$$DJ(y)h = \int_a^b (F_y h + F_{y'} h' + \dots + F_{y^{(n)}} h^{(n)}) \, dx.$$

En un punto crítico y , tendremos la ecuación de Euler:

$$F_y - (F_{y'})' + (F_{y''})'' + \dots + (-1)^n (F_{y^{(n)}})^{(n)} = 0$$

si y está en $C^{2n}[a, b]$ y, si $y(x)$ está en C^n , una combinación de derivadas, como en el Teorema 5.1, será 0.

Ejercicio 5.3 : Segunda ecuación de Euler.

Usando la transformación del Teorema 2.1, ver que tenemos la relación $y'' = \eta''(1 + \varepsilon\eta')^{-3}$ y un problema en $[\alpha, \beta]$ con $\eta(\alpha)$, $\eta(\beta)$, $\eta'(\alpha)$, $\eta'(\beta)$ dados. Usando la ecuación de Euler para este problema transformado ver que se obtiene la ecuación:

$$\frac{d}{dx}[y''F_{y''} - y'(\frac{d}{dx}(F_{y''}) - F_{y'}) - F] = -F_x$$

en un extremo local del problema para el funcional

$$\int_a^b F(x, y, y', y'') dx.$$

4.6 Frontera libre: condiciones naturales

Hasta ahora los funcionales que hemos visto se aplicaban a funciones con valores dados en la frontera. Sin embargo, esto no cubre todas las aplicaciones. Por ejemplo en el ejemplo 5.1 de la barra elástica podríamos dejar la barra sostenida en los extremos y no empotrada: en ese caso $u(0, t)$ y $u(\ell, t)$ son dadas pero no u_x . O, en el ejemplo 2.2 de la superficie de revolución, podemos pedir que la superficie se apoye en el lado derecho en una pared, el cual sería el caso de una pompa de jabón pegada a esa pared. Sea entonces:

$$J(y) = \int_a^b F(x, y, y') dx$$

donde y es C^1 pero sin condiciones de frontera. El espacio de trabajo será entonces $C^1[a, b]$ y, si F es C^1 , $J(y)$ tendrá una derivada de Fréchet:

$$DJ(y)h = \int_a^b (F_y h + F_{y'} h') dx.$$

En un punto crítico, $DJ(y)h = 0$, para todo h en $C^1[a, b]$. Por lo tanto si tomamos h en la subclase de $C^1[a, b]$, $C_0^\infty[a, b]$, tendremos la ecuación de Euler:

$$F_y - \frac{d}{dx}(F_{y'}) = 0$$

y $F_{y'}$ es C^1 . Por lo tanto podemos usar integración por partes en el cálculo de la derivada de Fréchet:

$$DJ(y)h = 0 = \int_a^b (F_y - \frac{d}{dx}(F_{y'}))h dx + F_{y'}h|_a^b = F_{y'}h|_a^b.$$

Tomando h en $C^1[a, b]$, con $h(a) = 0$, $h(b) = 1$, es necesario que $F_{y'}|_b = 0$. Del mismo modo, con $h(a) = 1$, $h(b) = 0$, tenemos $F_{y'}|_a = 0$.

Teorema 6.1 : *Si no hay condiciones de frontera para el funcional $J(y)$, entonces en un punto crítico se obtiene:*

1) *La ecuación de Euler:*

$$F_y - \frac{d}{dx}(F_{y'}) = 0.$$

2) *Las condiciones de frontera “naturales”:*

$$F_{y'}(a, y(a), y'(a)) = 0, \quad F_{y'}(b, y(b), y'(b)) = 0.$$

La palabra “natural” viene del hecho que un punto crítico debe satisfacer esta condición de frontera.

Nota 6.1 : Si y está dada en un punto, por ejemplo a , entonces queda claro que obtendremos solamente la condición de frontera “natural” en b .

Nota 6.2 : Si tenemos

$$J(y_1, y_2) = \int_a^b F(x, y_1, y_2, y'_1, y'_2) dx$$

entonces tendremos las condiciones naturales:

$$F_{y'_i}(a, y_1(a), \dots, y'_2(a)) = 0, \quad F_{y'_i}(b, y_1(b), \dots, y'_2(b)) = 0$$

o alguna de éstas si y_i está dada en a o en b . El número total de condiciones debe ser 4.

Ejercicio 6.1 : Encontrar las condiciones naturales para el funcional

$$J(y) = \int_a^b F(x, y, y', y'') dx.$$

Usar el teorema 5.1 para poder integrar por partes, escoger $h(x)$ en $C^2[a, b]$ tal que sólo una de las 4 cantidades, $h(a)$, $h(b)$, $h'(a)$, $h'(b)$, sea diferente de 0.

Ejemplo 6.1 : Superficie de revolución con mínima área.

Supongamos que sólo $f(0)$ está dado (ver figura 6.1). Del ejemplo 2.2, sabemos que el funcional es:

$$A(y) = \int_0^a 2\pi y(1 + y'^2)^{1/2} dx$$

con su respectiva ecuación de Euler. La condición de frontera natural en a será: $F_{y'} = 0$, lo cual implica $y(a)y'(a) = 0$.

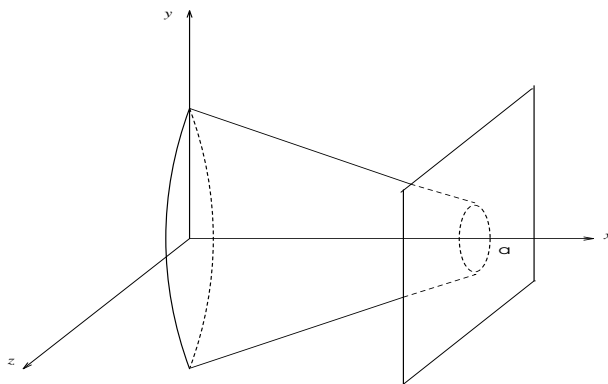


Fig. 6.1

Como ya lo hemos visto, la condición $y(a) = 0$ implica una posible singularidad de $y(x)$ en a . Por lo tanto la condición más frecuente será $y'(a) = 0$: la superficie es ortogonal a la pared.

Ejemplo 6.2 : Barra pesada estacionaria.

Si usamos la aproximación del ejemplo 5.1, la acción está dada por:

$$J(u) = T \int_0^\ell (uf - ku'^2/2) dx$$

con derivada de Fréchet:

$$DJ(u)h = T \int_0^\ell (fh - ku''h'') dx.$$

En un punto crítico $DJ(u)h = 0$ para toda h en $C^2[a, b]$ (si no hay condiciones) o en un subespacio de $C^2[a, b]$, pero seguramente para todo h en $C_0^\infty[a, b]$. Por el teorema III.1, el punto crítico es C^4 con $u^{(4)}(x) = f/k$. Podemos por lo tanto integrar por partes:

$$DJ(u)h = 0 = \int_0^\ell (f - ku^{(4)})h dx - (ku''h' - ku'''h) \Big|_0^\ell.$$

Todas las posibles condiciones de frontera (dadas o naturales) resultarán de $(u''h' - u'''h) \Big|_0^\ell = 0$. Por ejemplo:

1) Si u, u' están dadas en 0 y en ℓ , entonces $h(0) = h'(0) = 0, h(\ell) = h'(\ell) = 0$. La barra está fija en 0 y en ℓ .

2) Si u, u' están dadas en 0, entonces $h(0) = h'(0) = 0$, pero u y u' están libres en ℓ , entonces para esta barra, fija en 0 y libre en ℓ , podemos tomar h con $h(0) = h'(0) = h'(\ell) = 0, h(\ell) = 1$ y obtener la condición natural $u'''(\ell) = 0$. Tomando $h(x)$ con $h(0) = h'(0) = h(\ell) = 0, h'(\ell) = 1$, tendremos $u''(\ell) = 0$: la curvatura en ℓ es 0.

3) Si u, u' están dadas en 0, u está dada en ℓ pero $u'(\ell)$ está libre, tendremos una barra fija en 0 y soportada en ℓ . Entonces $h(0) = h'(0) = h(\ell) = 0$. Podemos tomar una función h (hacer dibujos) con estas condiciones y $h'(\ell) = 1$.

La condición natural será $u''(\ell) = 0$.

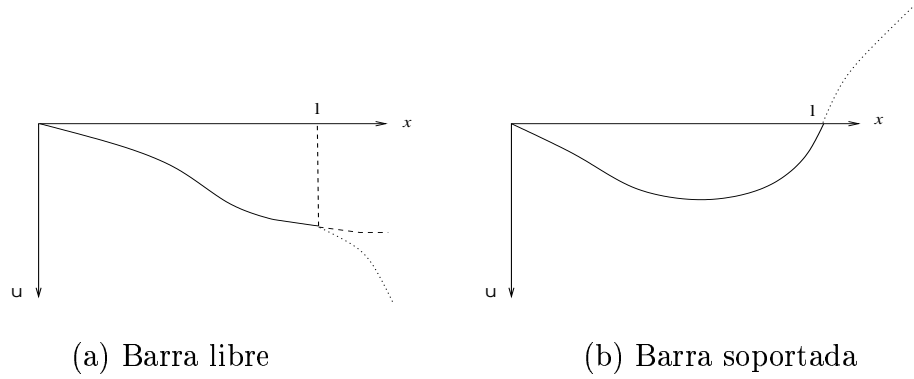


Fig. 6.2

Ejemplo 6.3 : Problema mixto en dos dimensiones.

Si tenemos el funcional

$$J(u) = \iint_{\Omega} F(x, y, u, u_x, u_y) dx dy$$

con u dado en Γ_1 , parte de $\partial\Omega$, y u libre en el resto $\Gamma_2 = \partial\Omega \setminus \Gamma_1$. Entonces

$$DJ(u)h = \iint_{\Omega} (F_u h + F_{u_x} h_x + F_{u_y} h_y) dx dy$$

para h en $C^1(\bar{\Omega})$ con $h(x, y) = 0$ en Γ_1 .

Si F es C^2 , entonces, en un punto crítico

$$DJ(u)h = \iint_{\Omega} \left(F_u - \frac{\partial}{\partial x}(F_{u_x}) - \frac{\partial}{\partial y}(F_{u_y}) \right) h dx dy + \int_{\partial\Omega} h \left(\frac{F_{u_x}}{F_{u_y}} \right) \cdot n ds = 0$$

para todo h . Tomando h en $C_0^\infty(\Omega)$, la integral doble desaparece, dando la ecuación de Euler $F_u - \frac{\partial}{\partial x}(F_{u_x}) - \frac{\partial}{\partial y}(F_{u_y}) = 0$. Pero como la clase de funciones h es más amplia que $C_0^\infty(\Omega)$, tenemos

$$\int_{\Gamma_2} h \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} \cdot n \, ds = 0.$$

Ahora bien, si φ es una función C^1 en $\partial\Omega$, con $\varphi = 0$ en Γ_1 , podemos construir h en $C^1(\overline{\Omega})$, con $h = 0$ en Γ_1 , de manera siguiente:

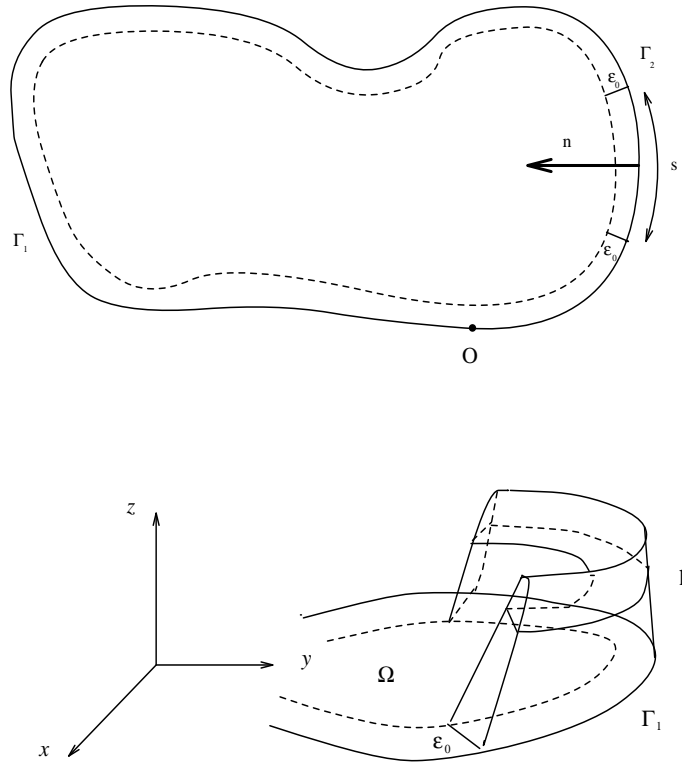


Fig. 6.3

Si $\partial\Omega$ es suficientemente lisa, sea n la normal interior. Cada punto en una vecindad de $\partial\Omega$, puede representarse como $(s, \epsilon n)$ donde s corresponde a la longitud de arco sobre $\partial\Omega$. Sea entonces $h(s, \epsilon n) = \rho(\epsilon)\varphi(s)$, con $\rho(0) = 1$, $\rho(\epsilon_0) = 0$, $\rho'(\epsilon_0) = 0$, ρ en C^∞ , y extendiendo h como 0 fuera de la vecindad.

Entonces

$$\int_{\Gamma_2} \varphi \begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} \cdot n \, ds = 0$$

para todo φ en $C^1(\Gamma_2) \cap C_0(\Gamma_2)$. Por el lema III.2, tenemos la condición natural:

$$\begin{pmatrix} F_{u_x} \\ F_{u_y} \end{pmatrix} \cdot n = 0 \text{ en } \Gamma_2.$$

Por ejemplo si $F = (u_x^2 + u_y^2)/2$, como en el caso de la membrana, la condición natural será $\frac{\partial u}{\partial n} = 0$, es decir la condición de Neumann.

En el caso de una frontera no lisa, con esquina por ejemplo, hay que tomar una vecindad de la esquina donde la extensión no es continua. Como esa vecindad es arbitrariamente pequeña, el resultado es el mismo. No insistiremos más en este punto pero dejamos al lector la tarea de convencerse de que este procedimiento puede hacerse tomando las precauciones necesarias.

Ejemplo 6.4 : Problema mixto con más derivadas.

En el ejemplo 5.1 vimos que el funcional

$$J(u) = \iint_{\Omega} F(x, y, u, u_x, u_y, u_{xy}, u_{yx}, u_{yy}) dx dy$$

tiene, en su derivada de Frechet, una integral doble que lleva a la ecuación de Euler, y una integral de línea:

$$\begin{aligned} & \int_{\partial\Omega} \left\{ (F_{u_x} - \frac{\partial}{\partial x}(F_{u_{xx}}) - \frac{\partial}{\partial y}(F_{u_{xy}})/2)n_1 + (F_{u_y} - \frac{\partial}{\partial x}(F_{u_{xy}})/2)n_2 + \right. \\ & + \frac{d}{ds}((F_{u_{xx}} - F_{u_{yy}})n_1n_2 - F_{u_{xy}}(n_1^2 - n_2^2)/2) \Big\} h ds + \\ & \int_{\partial\Omega} (F_{u_{xx}}n_1^2 + F_{u_{xy}}n_1n_2 + F_{u_{yy}}n_2^2) h_n ds. \end{aligned}$$

Si en parte de $\partial\Omega$, h y h_n no son 0, entonces si $\varphi(s)$ es $C^2(\partial\Omega)$, podemos construir dos extensiones a una vecindad de $\partial\Omega$, del modo siguiente:

$$h_1(x, y) = \rho_1(\epsilon)\varphi(s)$$

$$h_2(x, y) = \rho_2(\epsilon)\varphi(s)$$

donde

$$(x, y) = (s, \epsilon n), \quad \rho_i \text{ es } C^\infty, \quad \rho_i(\epsilon_0) = \rho'_i(\epsilon_0) = \rho''_i(\epsilon_0) = 0,$$

$$\rho_1(0) = 1, \quad \rho'_1(0) = 0, \quad \rho_2(0) = 0, \quad \rho'_2(0) = 1.$$

(Esto requiere que $\partial\Omega$ sea C^2 .) Entonces $h_1 = \varphi$ sobre $\partial\Omega$, con $\partial h_1/\partial n = 0$, $h_2 = 0$ sobre $\partial\Omega$, con $\partial h_2/\partial n = \varphi$. Por lo tanto en las partes donde u no está dado, tendremos la condición natural:

$$\begin{aligned} & \left(F_{u_x} - \frac{\partial}{\partial x}(F_{u_{xx}}) - \frac{\partial}{\partial y}(F_{u_{xy}})/2 \right) n_1 + \left(F_{u_y} - \frac{\partial}{\partial y}(F_{u_{yy}}) - \frac{\partial}{\partial x}(F_{u_{xy}})/2 \right) n_2 + \\ & + \frac{d}{ds} \left((F_{u_{xx}} - F_{u_{yy}}) n_1 n_2 - F_{u_{xy}} (n_1^2 - n_2^2)/2 \right) = 0 \end{aligned}$$

y en las partes donde $\frac{\partial u}{\partial n}$ no está dado:

$$F_{u_{xx}} n_1^2 + F_{u_{xy}} n_1 n_2 + F_{u_{yy}} n_2^2 = 0.$$

Ejercicio 6.2 : Estudiar las condiciones de frontera para la barra no estacionaria, con acción:

$$2J(u) = \int_0^T \int_0^\ell (\rho u_t^2 + 2fu)(1 + u_x^2)^{1/2} - k u_{xx}^2 (1 + u_x^2)^{-5/2} dx dt$$

en los siguientes casos:

1. $h = 0$ en $\partial\Omega$, $h_x = 0$ en $x = 0$, $x = \ell$. Barra sujeta con posiciones iniciales y finales conocidas.
2. h arbitraria.
3. $h = 0$, $h_x = 0$ en $x = 0$, $x = \ell$; $h = 0$ en $t = 0$. Barra sujeta, con posición inicial conocida.
4. $h = 0$ en $x = 0$, $x = \ell$, $t = 0$; $h_x = 0$ en $x = 0$. Barra fija en $x = 0$, soportada en $x = \ell$.
5. $h = 0$ en $x = 0$, $t = 0$; $h_x = 0$ en $x = 0$. Barra fija en $x = 0$, libre en $x = \ell$.

Notemos que si tomamos la aproximación:

$$2J(u) = \int_0^T \int_0^\ell (\rho u_t^2 + 2fu - k u_{xx}^2) dx dt$$

entonces la primera condición natural se reduce a:

$$k u_{xxx} n_1 + \rho u_t n_2 = 0$$

($n_1 n_2 = 0$ en un rectángulo) si h es libre, mientras que la segunda, si h_n es libre, es $k u_{xxx} n_1^2 = 0$. En particular, si la barra es sujeta en $x = 0$ y $x = \ell$, si la posición inicial es conocida pero no la final (lo cual es el problema natural ya que conocer la posición en el tiempo final, o en cada instante, es tener el problema resuelto de antemano), entonces la única condición natural es $u_t(x, T) = 0$.

Como T es arbitraria, el principio de mínima acción lleva a la conclusión que u no depende de t . Es por lo tanto muy importante tener cuidado de no aplicar principios variacionales

de manera indiscriminada. Se puede llegar así a situaciones contrarias a la realidad. Para problemas dependientes del tiempo, se puede aplicar el principio de mínima acción para derivar la ecuación pero hay que pensar con cuidado las condiciones de frontera.

Ejemplo 6.5 : Problema de Robin.

Sea $J(u) = \iint_{\Omega} (u_x^2 + u_y^2) dx dy + \int_{\partial\Omega} r(s) u^2 ds$. Entonces, sobre $C^1(\overline{\Omega})$:

$$J(u+h) - J(u) = J(h) + 2 \iint_{\Omega} -h \Delta u + 2 \int_{\partial\Omega} \left(\frac{\partial u}{\partial n} + r(s)u \right) h ds$$

si u es C^2 . En un punto crítico tendremos: $\Delta u = 0$ en Ω (tomando h en $C_0^\infty(\Omega)$) $\frac{\partial u}{\partial n} + r(s)u = 0$ en $\partial\Omega$: condición de Robin.

Notemos que u es un punto crítico y $r(s) \geq 0$ y no idénticamente 0, entonces el punto crítico es un mínimo y es único: de hecho $J(u+h) - J(u) - J(h) \geq 0$ y $J(h) = 0$ sólo si $h = \text{constante}$ con $r(s)h = 0$ en $\partial\Omega$.

4.7 Problemas con discontinuidades, condición de Weierstrass

En algunos problemas el integrando puede ser discontinuo en algún punto o se desea usar funciones con derivadas discontinuas, es decir

$$J(y) = \int_a^b F(x, y, y') dx$$

donde se permite que y' sea discontinua en algún punto intermedio c .

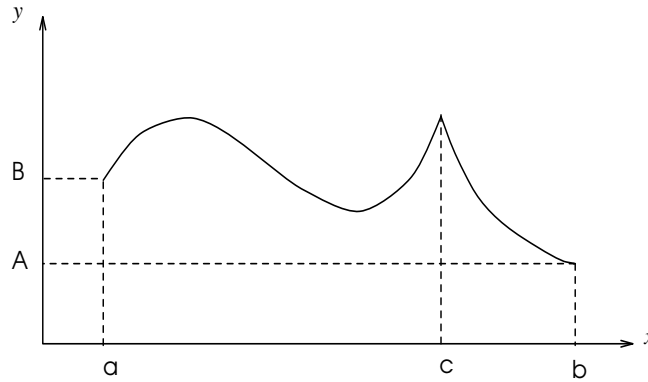


Fig. 7.1

Podemos en ambos casos escribir:

$$J(y) = \int_a^c F(x, y, y') dx + \int_c^b F(x, y, y') dx$$

donde $y(x)$ debe ser C^1 en $[a, c]$ y en $[c, b]$ y continua en c :

$$y(c_-) - y(c_+) \equiv [y]_c = 0,$$

definiendo así el salto de una función en el punto c . Si $y(a) = A$, $y(b) = B$, podemos usar como espacio de trabajo, el espacio:

$$B = \{h \text{ es } C^1 \text{ en } [a, c] \text{ y en } [c, b], h(a) = 0, h(b) = 0, [h]_c = 0\}$$

con la norma:

$$\|h\| = |h|_{C^1[a, c]} + |h|_{C^1[c, b]}.$$

Es claro que B es un espacio de Banach.

Escribiendo $y = y_0 + h$, si F es C^1 en $[a, c] \times \mathbb{R}^2$ y $[c, b] \times \mathbb{R}^2$, entonces $J(y)$ tendrá como derivada de Frechet:

$$DJ(y_0)h = \int_a^c (F_y h + F_{y'} h') dx + \int_c^b (F_y h + F_{y'} h') dx.$$

Esta derivada será 0 para todo h en B , si y_0 es un punto crítico. En particular podemos tomar h en $C_0^\infty[a, c]$ y $h(x) \equiv 0$ en $[c, b]$: tendremos que $F_{y'}$ es C^1 en $[a, c]$, con derivada $F_{y'}(x, y_0(x), y'_0(x))$. Tomando $h \equiv 0$ en $[a, c]$ y en $C_0^\infty[c, b]$, tendremos que la ecuación de Euler es válida también en $[c, b]$. Por lo tanto podemos integrar por partes:

$$DJ(y_0)h = hF_{y'} \Big|_a^c + hF_{y'} \Big|_c^b = [hF_{y'}]_c = 0.$$

Como $[h] = 0$, podemos tomar h en B con $h(c) \neq 0$. Tendremos:

Teorema 7.1 : *Si F es discontinua en c , o si y' puede ser discontinua en c , entonces en un punto crítico y_0 :*

$$\frac{d}{dx}(F_{y'} - F_y(x, y_0(x), y'_0(x))) = 0$$

en (a, c) y en (c, b) y tenemos la condición de Weierstrass:

$$F_{y'}(c, y_0(c), y'_0(c_-)) = F_{y'}(c, y_0(c), y'_0(c_+)).$$

Ejemplo 7.1 : Ley de Snell.

Si tenemos dos medios con índices de refracción $v_1^{-1}(x, y)$ y $v_2^{-1}(x, y)$, entonces el tiempo de recorrido de la luz del punto (a, A) al punto (b, B) está dado por:

$$T(y) = \int_a^c (1 + y'^2)^{1/2} v_1^{-1} dx + \int_c^b (1 + y'^2)^{1/2} v_2^{-1} dx$$

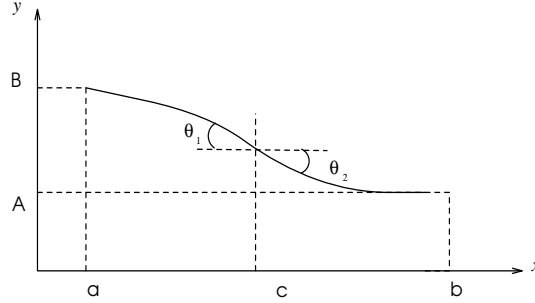


Fig. 7.2

Las ecuaciones de Euler en (a, c) y (c, b) darán la trayectoria en cada medio, recta si v_1 y v_2 son constantes (ver el ejemplo 3.2). La condición de continuidad de Weierstrass es:

$$\frac{y'(c_-)}{v_1(1 + y'^2(c_-))^{1/2}} = \frac{y'(c_+)}{v_2(1 + y'^2(c_+))^{1/2}}$$

es decir:

$$\frac{\text{sen } \theta_1}{v_1} = \frac{\text{sen } \theta_2}{v_2},$$

la ley de Snell.

Ejemplo 7.2 : Potencial electrostático.

Consideremos un cuerpo Ω_1 , con densidad de carga eléctrica ρ , y dentro de un ambiente aislado del exterior por una superficie puesta a la tierra.

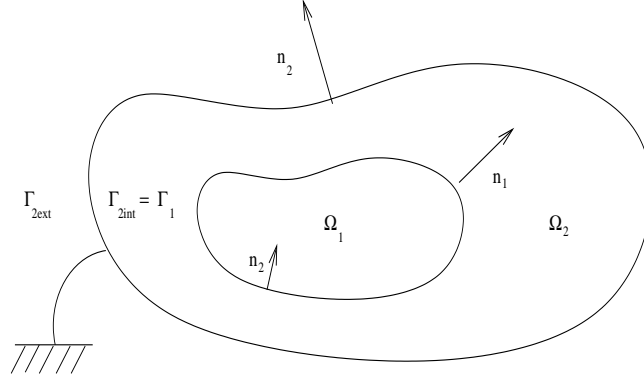


Fig. 7.3

El potencial electrostático u corresponde al mínimo de la energía

$$J(u) = \iiint_{\Omega_1} (|\nabla u|^2/2 + \rho u) dx + \iiint_{\Omega_2} |\nabla u|^2/2 dx$$

sobre el espacio de funciones u que son C^1 en $\overline{\Omega}_1$ y en $\overline{\Omega}_2$ con u continua a través de $\partial\Omega_1$, $u = 0$ en $\partial\Omega_{2\text{Ext}}$, la frontera exterior de Ω_2 .

$$\begin{aligned} DJ(u)h &= \iiint_{\Omega_1} \nabla u \cdot \nabla h + \rho h + \iiint_{\Omega_2} \nabla u \cdot \nabla h \\ &= \iiint_{\Omega_1} (\rho - \Delta u)h + \iint_{\partial\Omega_1} h \frac{\partial u}{\partial n_1} ds + \iiint_{\Omega_2} -h\Delta u + \iint_{\partial\Omega_2} h \frac{\partial u}{\partial n_2} ds \end{aligned}$$

n_1 es la normal exterior a $\partial\Omega_1$, n_2 es la normal exterior a $\partial\Omega_2$. Tomando h en $C_0^\infty(\Omega_1)$, $h \equiv 0$ en Ω_2 , obtenemos

$$\rho - \Delta u = 0 \text{ en } \Omega_1.$$

Tomando h en $C_0^\infty(\Omega_2)$, $h \equiv 0$ en Ω_1 , tenemos

$$-\Delta u = 0 \text{ en } \Omega_2.$$

De esta forma tenemos

$$DJ(u)h = \iint_{\partial\Omega_1} h \left(\frac{\partial u}{\partial n_1} - \frac{\partial u}{\partial n_2} \right) ds$$

ya que $h = 0$ en $\partial\Omega_{2\text{Ext}}$, y h es continuo a través de $\partial\Omega_1$. Ahora, para toda función φ , en $C^1(\partial\Omega_1)$ podemos encontrar h en el espacio de trabajo con $\varphi = h|_{\partial\Omega_1}$. Por el lema III.2:

$$\left[\frac{\partial u}{\partial n} \right]_{\partial \Omega_1} = 0.$$

Notemos que $[u]_{\partial \Omega_1} = 0$, por construcción. Vemos que en este ejemplo u es $C^1(\overline{\Omega}_1 \cup \overline{\Omega}_2)$ pero no C^2 (debido a las ecuaciones de Euler).

Ejemplo 7.3 : Función con segundas derivadas discontinuas.

Sea $J(y) = \int_a^b F(x, y, y', y'') dx$ donde F tiene una discontinuidad en c o y'' tiene una discontinuidad en c . y, y' son dadas en a y b . El espacio de trabajo será

$$B = \left\{ h \text{ en } C^2[a, c] \cap C^2[c, b] \cap C^1[a, b], \text{ con } h(a) = h'(a) = h(b) = h'(b) = 0 \right\}$$

es decir h y h' son continuas en c .

$$DJ(y)h = \int_a^c (F_y h + F_{y'} h' + F_{y''} h'') dx + \int_c^b (F_y h + F_{y'} h' + F_{y''} h'') dx$$

para h en B . Si y es un punto crítico, tomando $h \equiv 0$ en uno de los intervalos, las ecuaciones de Euler serán válidas en (a, c) y (c, b) y podemos integrar por partes:

$$DJ(y)h = 0 = \left[F_{y''} h' + \left(F_{y'} - \frac{d}{dx}(F_{y''}) \right) h \right]_c = 0$$

ya que $h(a) = h(b) = h'(a) = h'(b) = 0$. Podemos tomar h en B tal que $h'(c) = 1$, $h(c) = 0$ ó inversamente y obtenemos:

$$[F_{y''}]_c = 0, \quad \left[F_{y'} - \frac{d}{dx}(F_{y''}) \right]_c = 0.$$

Ejercicio 7.1 : Encontrar las condiciones de continuidad para

$$J(u) = \iint_{\Omega_1 \cup \Omega_2} F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) dx dy$$

donde Ω_1 y Ω_2 sea como en la figura 7.3. u y $\frac{\partial u}{\partial n}$ son dados sobre $\partial \Omega_{2 \text{ Ext}}$, u es C^1 en $\overline{\Omega}_1 \cup \overline{\Omega}_2$. Usar el ejemplo 6.4.

Ejemplo 7.4 : Problemas con discontinuidades no fijas. Condiciones de Weiersstrass I.

En algunos problemas se considera el funcional

$$J(y) = \int_a^b F(x, y, y') dx$$

donde y' puede ser discontinua (pero con límites laterales) en un punto intermedio c que no se conoce de antemano (por ejemplo problemas de confinamiento de plasmas), o en varios puntos. En este caso se interpreta la integral como:

$$J(y) = \int_a^c F(x, y, y') dx + \int_c^b F(x, y, y') dx$$

para una discontinuidad, o como

$$\sum_0^{n-1} \int_{x_i}^{x_{i+1}} F(x, y, y') dx$$

para más discontinuidades.

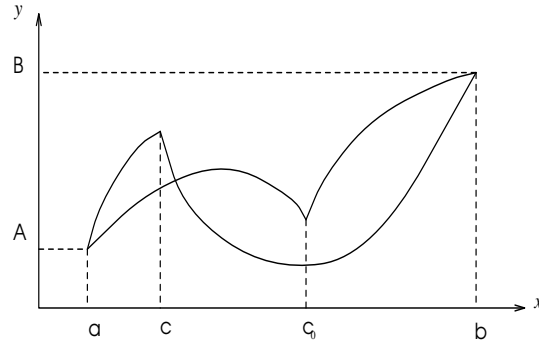


Fig. 7.4

Notemos que si y'_0 tiene una discontinuidad en c_0 , y y' tiene una discontinuidad en c , entonces $h' = y' - y'_0$ tendrá en general dos discontinuidades. Por lo tanto si queremos usar una estructura de espacio lineal no podemos considerar las funciones con una sola discontinuidad.

De hecho tomaremos como espacio de trabajo a

$$X = \{h \text{ con un número finito de saltos en } h' \text{ con } h(a) = h(b) = 0\}.$$

Pondremos como norma sobre X

$$\|h\| = \sum_j |h|_{C^1[I_j]}$$

donde $I_1, \dots, I_{n(h)}$, son los intervalos de continuidad de h y de h' (dependen de h').

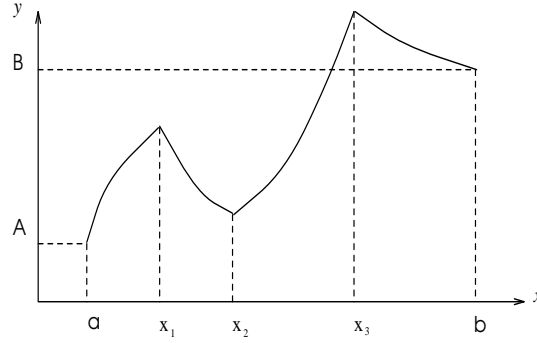


Fig. 7.5

X es un espacio normado pero no completo (límites de funciones en X podrían tener un número infinito de discontinuidades). Recordemos solamente que las condiciones necesarias son válidas en un espacio normado (Teorema II.4.1).

Si y en X es un punto crítico de J , entonces

$$DJ(y)h = \sum \int_{x_i}^{x_{i+1}} (F_y h + F_{y'} h') dx = 0$$

para todo h en X con $h(a) = h(b) = 0$.

En efecto, para cada h en X los intervalos de discontinuidades de $y + th$ dependen únicamente de y y de h (no de t). Si F es C^1 podemos aplicar el razonamiento usual en cada intervalo.

En particular si tomamos h en $C_0^\infty[x_i, x_{i+1}]$, donde $[x_i, x_{i+1}]$ es un intervalo de continuidad de y' , tendremos que $F_{y'}$ tiene derivada en (x_i, x_{i+1}) y la ecuación de Euler es válida en el intervalo.

Además si y es un extremo local, también es válida la segunda ecuación de Euler: $F - y'F_{y'}$ tiene como derivada F_x en (x_i, x_{i+1}) . Ahora si tomamos h en $C_0^\infty[x_{i-1}, x_{i+1}]$, con $h(x_i) \neq 0$, podemos integrar por partes las dos integrales, sobre $[x_{i-1}, x_i]$ y sobre $[x_i, x_{i+1}]$, como en el Teorema 7.1, y obtener:

$$[F_{y'}]_{x_i} = 0.$$

Finalmente $F_x(x, y(x), y'(x))$ es una función continua por pedazos, por lo tanto su integral es continua, es decir

$$[F - y'F_{y'}]_{x_i} = 0.$$

Tenemos pues:

Teorema 7.2 : Si F es C^1 y $y(x)$, con posibles discontinuidades en x_1, \dots, x_N en y' , es un extremo local de $J(y) = \int_a^b F(x, y, y') dx$, con $y(a) = a$, $y(b) = B$, sobre el espacio X , entonces:

- 1) $\frac{d}{dx}(F_{y'}) = F_y$ para $x \neq x_i$
- 2) $[F_{y'}]_{x_i} = 0$
- 3) $[F - y'F_{y'}]_{x_i} = 0$
- 4) En cada x_i , si $[y']_{x_i} \neq 0$, entonces $F_{zz}(x_i, y(x_i), z) = 0$ para algún z entre $y'(x_{i-})$ y $y'(x_{i+})$.

Prueba. Para probar el punto (4), basta usar el teorema de Rolle para la función $F_z(x_i, y(x_i), z)$ en el intervalo definido por $y'(x_{i-})$ y $y'(x_{i+})$.

Q.E.D.

Ejercicio 7.2 : Si F depende de varias funciones, probar que:

1. $\frac{d}{dx}(F_{y'_j}) = F_{y_j}$ para $x \neq x_i$
2. $[F_{y'_j}]_{x_i} = 0$
3. $[F - \sum y'_j F_{y'_j}]_{x_i} = 0$ (usar el ejercicio 3.1).

Nota 7.1: Hay otra manera de probar este resultado, pero requiere de hipótesis más fuertes sobre la y , pero no usa la segunda ecuación de Euler. Sea \bar{y} el candidato a punto crítico. Sea $\varphi_\alpha(x) : [x_{i-1}, x_{i+1}] \rightarrow [x_{i-1}, x_{i+1}]$ una función C^∞ en x y diferenciable en α , con la forma siguiente:

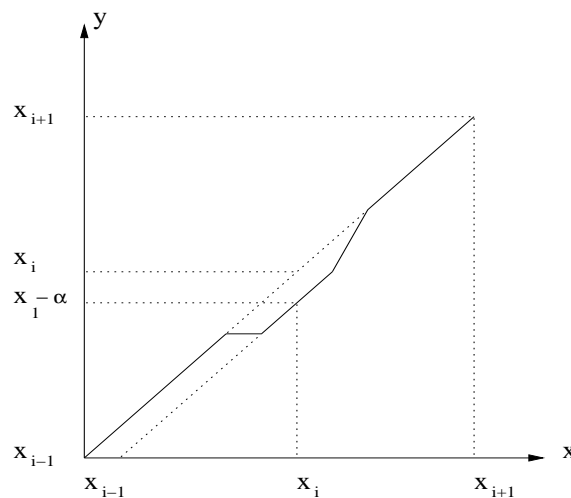


Fig. 7.6

Sea $y(x, \alpha) = \bar{y}(\varphi_\alpha(x))$. Por lo tanto $y(x, \alpha)$ coincide con \bar{y} excepto cerca de x_i donde $y(x, \alpha) = \bar{y}(x - \alpha)$, en particular $y(x, \alpha)$ tiene discontinuidad en x_j , $j \neq i$ y en $x_i + \alpha$.

Si consideramos $I(\alpha) = \int_a^b F(x, y(x, \alpha), \frac{\partial}{\partial x} y(x, \alpha)) dx$, $I(\alpha)$ debe tener un punto crítico en $\alpha = 0$, por lo tanto

$$\left. \frac{dI}{d\alpha} \right|_{\alpha=0} = 0 = \left. \frac{d}{d\alpha} \left(\int_{x_{i-1}}^{x_i+\alpha} F dx + \int_{x_i+\alpha}^{x_{i+1}} F dx \right) \right|_{\alpha=0}$$

ya que sobre el resto la integral no depende de α .

$$\left. \frac{dI}{d\alpha} \right|_{\alpha=0} = [F]_{x_i} + \int_{x_{i-1}}^{x_i} \left(F_y \frac{\partial y}{\partial \alpha} + F_{y'} \frac{\partial^2 y}{\partial \alpha \partial x} \right) \Big|_{\alpha=0} dx + \int_{x_i}^{x_{i+1}} \left(F_y \frac{\partial y}{\partial \alpha} + F_{y'} \frac{\partial^2 y}{\partial \alpha \partial x} \right) \Big|_{\alpha=0} dx.$$

Si suponemos que \bar{y} es C^2 para $x \neq x_j$, podemos intercambiar las derivadas, usar la ecuación de Euler en cada intervalo para integrar por partes y obtener:

$$\left. \frac{dI}{d\alpha} \right|_{\alpha=0} = [F]_{x_i} + F_{y'} \frac{\partial y}{\partial \alpha} \Big|_{\alpha=0} \Big|_{x_{i-1}}^{x_i} + F_{y'} \frac{\partial y}{\partial \alpha} \Big|_{\alpha=0} \Big|_{x_i}^{x_{i+1}}.$$

Como $y(x, \alpha) = \bar{y}(x)$ para x vecino a x_{i-1} y a x_{i+1} , $\frac{\partial y}{\partial \alpha} = 0$ en esos puntos; $\frac{\partial y}{\partial \alpha} \Big|_{x_i} = -\bar{y}'(x_i - \alpha)$. Por lo tanto: $[F]_{x_i} - [F_{y'} y']_{x_i} = 0$.

La primera condición se obtiene como antes.

Nota 7.2 : Veremos en el Ejemplo 8.5 como usar los multiplicadores de Lagrange para obtener el mismo resultado.

Ejercicio 7.3 : Extremos con discontinuidades en las segundas derivadas.

Si en un problema para el funcional

$$\int_a^b F(x, y, y', y'') dx$$

se admiten discontinuidades en las segundas derivadas en puntos no fijos entre a y b , probar que además de las dos condiciones del ejemplo 7.3, hay una tercera condición de continuidad dada por:

$$[y'' F_{y''} - F]_c = 0.$$

Probar esto de las dos maneras siguientes:

a) Con la segunda ecuación de Euler del ejercicio 5.3.

b) Usando la transformación de la nota 7.1.

Se verán dos maneras más en el ejercicio 8.7.

4.8 Multiplicadores de Lagrange

Como lo vimos en el Teorema II.4.3, si $J(y)$, $G_1(y)$, \dots , $G_m(y)$ son funcionales con derivadas de Fréchet (o derivadas de Gâteaux en una vecindad de y_0 , continuas en y para cada h fijo) entonces en un punto crítico y_0 de J condicionado por $G_i(y) = 0$, existen constantes ν , $\lambda_1, \dots, \lambda_m$, no todas 0, tales que:

$$\nu DJ(y_0)h + \sum \lambda_i DG_i(y_0)h = 0.$$

Veremos en esta sección varias aplicaciones de este resultado.

Ejemplo 8.1 : Valores propios del laplaciano.

Sea $J(u) = \iint_{\Omega} (u_x^2 + u_y^2) dx dy$, $u = 0$ en $\partial\Omega$, condicionado por $G(u) = \iint_{\Omega} u^2 dx dy - 1 = 0$. Aquí:

$$\nu DJ(u)h + \lambda DG(u)h = 2 \iint_{\Omega} (\nu \nabla u \cdot \nabla h + \lambda u h) dx dy = 2 \iint_{\Omega} (-\nu \Delta u + \lambda u) h dx dy$$

ya que $h = 0$ en $\partial\Omega$.

Tenemos $-\nu \Delta u + \lambda u = 0$ en Ω . Si $\nu = 0$, entonces $u = 0$, en Ω , y no se puede cumplir la condición de normalización. Por lo tanto u debe satisfacer:

$$\begin{aligned} -\Delta u + \mu u &= 0 \text{ en } \Omega \\ u &= 0 \text{ en } \partial\Omega \end{aligned}$$

μ es el valor propio de Δ , con condición de Dirichlet en $\partial\Omega$. Notemos que

$$0 = \iint_{\Omega} (-\Delta u + \mu u) u dx dy = \iint_{\Omega} (|\nabla u|^2 + \mu u^2) dx dy,$$

por lo tanto necesitamos $\mu < 0$ para tener una solución al problema con $u \not\equiv 0$.

Ejemplo 8.2 : Condiciones naturales.

Podemos pensar el problema de frontera libre como el siguiente problema condicionado (de manera un poco artificial). Consideremos el funcional

$$J(y, x_1) = \int_a^{x_1} F(x, y, y') dx$$

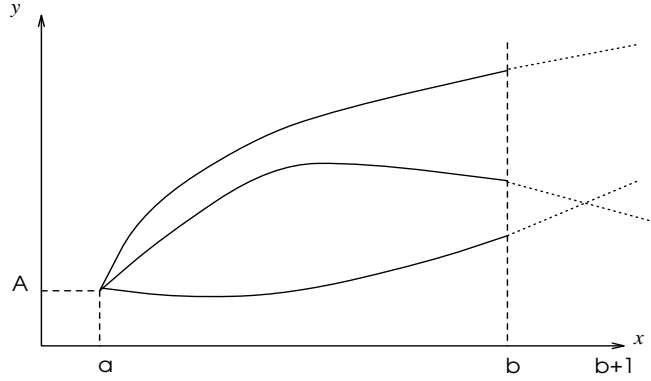


Fig. 8.1

donde y está en $C^1[a, b+1]$, con norma $\|\cdot\|_1$. Si tenemos una función definida en $[a, b]$, se extiende linealmente a una función en $[a, b+1]$: $y(x) = y(b) + y'(b)(x-b)$, para $x > b$. x_1 está en $[a, b+1]$, con norma $\|x_1\|$. $F(x, y, z)$ se extiende a $[a, b+1] \times \mathbb{R}^2$ linealmente:

$$F(x, y, z) = F(b, y, z) + F_x(b, y, z)(x-b), \text{ para } x > b.$$

(Como se verá en el resultado estas extensiones no influyen, por lo tanto se puede tomar cualquier tipo de extensión C^1 .)

Pondremos la condición

$$G(y, x_1) = x_1 - b = 0.$$

Los dos funcionales son entonces diferenciables de Fréchet en $C^1[a, b+1] \times \mathbb{R}$, si se supone que $F_x(b, y, z)$ es C^1 .

En un punto crítico condicionado tendremos:

$$\nu DJ(y, x_1)(h, \alpha) + \lambda DG(y, x_1)(h, \alpha) = 0$$

para todo (h, α) en $C^1[a, b+1] \times \mathbb{R}$, con $h(a) = 0$. Es decir:

$$\nu \left[\int_a^{x_1} (F_y h + F_{y'} h') dx + F \alpha \big|_{x_1} \right] + \lambda \alpha = 0$$

ya que

$$\int_a^{x_1+\alpha} F dx - \int_a^{x_1} F dx = F(x_1, y(x_1), y'(x_1))\alpha + o(|\alpha|).$$

Notemos que $\nu \neq 0$ ya que de lo contrario, la relación $\lambda\alpha = 0$ (con $\lambda \neq 0$) implicaría $\alpha = 0$, pero los incrementos en x_1 deben ser arbitrarios. Por lo tanto podemos dividir por ν , o suponer que ν vale 1. Tomando $\alpha = 0$ y usando la condición $x_1 - b = 0$, tenemos

$$\int_a^b (F_y h + F_{y'} h') dx = 0$$

para todo h en $C^1[a, b+1]$, con $h(a) = 0$, en particular para todo h en $C_0^\infty[a, b]$: por lo tanto $F_{y'}$ es C^1 , con derivada F_y , y podemos integrar por partes:

$$\int_a^b \left(F_y - \frac{d}{dx}(F_{y'}) \right) h dx + F_{y'} h \Big|_a^b + F(b, y(b), y'(b)) \alpha + \lambda \alpha = 0$$

es decir:

$$F_{y'}(b, y(b), y'(b))h(b) + (F(b, y(b), y'(b)) + \lambda)\alpha = 0$$

para todo (h, α) en $C^1[a, b+1] \times \mathbb{R}$, con $h(a) = 0$. Tomando $h(b) \neq 0$, $\alpha = 0$, obtenemos la condición “natural”

$$F_{y'}(b, y(b), y'(b)) = 0.$$

Como α es arbitrario, obtenemos también el valor del multiplicador

$$\lambda = -F(b, y(b), y'(b)).$$

Ejemplo 8.3 : Extremo móvil sobre una curva: condición de transversalidad.

El ejemplo anterior es un caso particular del siguiente problema. Encontrar los puntos críticos de $\int_a^{x_1} F(x, y, y') dx$ donde el extremo superior x_1 se mueve sobre una curva $y = \varphi(x_1)$.

Consideremos entonces el funcional

$$J(y, x_1) = \int_a^{x_1} F(x, y(x), y'(x)) dx$$

definido sobre $C^1[a, b+c] \times \mathbb{R}$, donde c es suficientemente grande para que $b+c$ sea mayor que el x_1 , candidato a punto crítico. Aquí también se extienden las funciones de manera lineal, así como F , para $x > x_1$:

$$F(x, \varphi(x_1) + c(x - x_1), c) = F(x_1, \varphi(x_1) + c(x - x_1), c) + F_x(x_1, \varphi(x_1) + c(x - x_1), c)(x - x_1).$$

Como $y(a) = A$, el espacio de trabajo será:

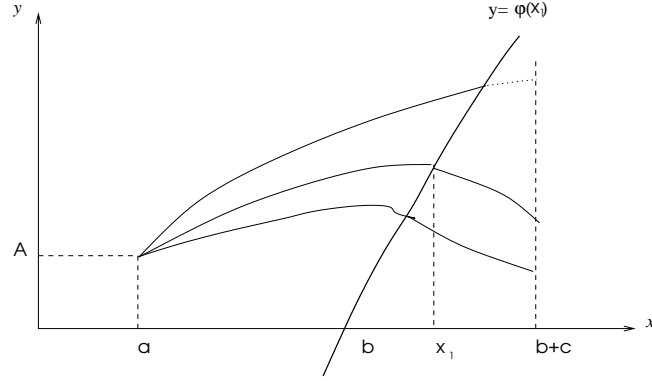


Fig. 8.2

(h, α) en $C^1[a, b+c] \times \mathbb{R}$, $h(a) = 0$, con norma

$$\|(h, \alpha)\| = |h|_{C^1[a, b+c]} + |\alpha|.$$

La condición será:

$$G(y, x_1) = y(x_1) - \varphi(x_1).$$

Entonces

$$\begin{aligned} G(y+h, x_1+\alpha) - G(y, x_1) &= y(x_1+\alpha) + h(x_1+\alpha) - \varphi(x_1+\alpha) \\ &- y(x_1) + \varphi(x_1) = \\ &= (y'(x_1) - \varphi'(x_1))\alpha + h(x_1) + o(\|(h, \alpha)\|) \end{aligned}$$

si φ es C^1 . (El término $h'(x_1)\alpha$ es de orden 2.)

Como en el ejercicio anterior, en un punto crítico tendremos

$$\nu \left(\int_a^{x_1} (F_y h + F_{y'} h') dx + F|_{x_1} \alpha \right) + \lambda \left((y'(x_1) - \varphi'(x_1))\alpha + h(x_1) \right) = 0$$

para todo (h, α) en el espacio de trabajo.

Tomando $\alpha = 0$ y h en $C_0^\infty[a, x_1]$ tendremos que $\nu F_{y'}$ es C^1 con derivada νF_y y podemos integrar por partes:

$$\nu F_{y'} h + \nu F \alpha + \lambda h + \lambda (y' - \varphi') \alpha|_{x_1} = 0.$$

Tomando $\alpha = 0$, $h(x_1) \neq 0$, y después $\alpha \neq 0$, $h(x_1) = 0$, obtenemos:

$$\begin{aligned}\nu F_{y'} + \lambda|_{x_1} &= 0 \\ \nu F + \lambda(y' - \varphi')|_{x_1} &= 0.\end{aligned}$$

De la primera ecuación, queda claro que $\nu \neq 0$, y sustituyendo λ en la segunda tenemos:

Teorema 8.1 : *En el caso de un punto crítico de $J(y)$ condicionado por el hecho que x_1 está sobre la curva $\varphi(x_1) = y$, si F y F_x son C^1 en y y en z , φ es C^1 en x , entonces el punto crítico satisface la ecuación de Euler :*

$$\frac{d}{dx}(F_{y'}) - F_y = 0 \text{ en } [a, x_1]$$

y la condición de transversalidad en x_1 :

$$F(x_1, y(x_1), y'(x_1)) - F_{y'}(x_1, y(x_1), y'(x_1))(y'(x_1) - \varphi'(x_1)) = 0.$$

Por ejemplo, si $F(x, y, z) = P(x, y)(1+z^2)^{1/2}$ con $P(x, y) \neq 0$, la condición de transversalidad es:

$$1 + y'(x_1)\varphi'(x_1) = 0$$

es decir la función crítica $y(x)$ es ortogonal a la curva $\varphi(x)$ en x_1 . Esto es el caso en los problemas de pompas de jabón.

Nota 8.1 : Otra manera de encontrar la condición de transversalidad está basada en la idea de evitar el multiplicador usando una reducción a una sola variable (es decir utilizar el teorema de la función implícita directamente). Este procedimiento, más clásico, tiene dos inconvenientes: requiere de una hipótesis de unicidad de un problema de punto crítico (lo cual no es obvio antes de ver las condiciones suficientes y es difícil de verificar) y es bastante oscuro, con la posibilidad de equivocarse. Supondremos pues que el problema de encontrar un punto crítico del funcional

$$J(y) = \int_a^{x_1} F(x, y, y') dx$$

tiene una solución única para cada x_1 vecino al candidato \bar{x}_1 y $y(x)$ vecina (en C^1) de la función crítica $\bar{y}(x)$, con las condiciones en la frontera $y(a) = A$, $y(x_1) = \varphi(x_1)$. Llamaremos a esta solución, perteneciente a una familia de soluciones indexadas por x_1 , a $y(x, x_1)$.

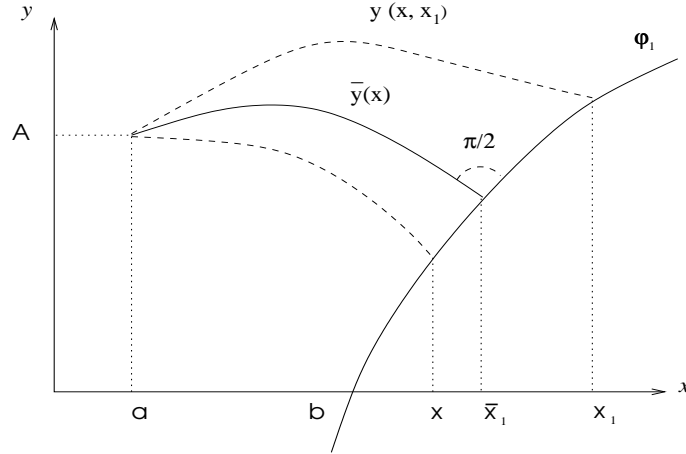


Fig. 8.3

Por lo tanto

$$\begin{aligned} y(x_1, x_1) &= \varphi(x_1) \\ y(a, x_1) &= A \\ y(x, \bar{x}_1) &= \bar{y}(x). \end{aligned}$$

Supondremos además que $y(x, x_1)$ es C^2 . Sea entonces

$$I(x_1) = \int_a^{x_1} F\left(x, y(x, x_1), \frac{\partial}{\partial x}(y(x, x_1))\right) dx.$$

Como $y(x, x_1)$ es una familia admisible de funciones para el problema, $I(x_1)$ debe tener un punto crítico en \bar{x}_1 (un extremo en caso de que $\bar{y}(x)$ sea un extremo para el problema). Por lo tanto:

$$\frac{dI}{dx_1}(\bar{x}_1) = 0 = \int_a^{\bar{x}_1} \left(F_y \frac{\partial y}{\partial x_1} + F_{y'} \frac{\partial^2 y}{\partial x_1 \partial x} \Big|_{\bar{x}_1} \right) dx + F|_{\bar{x}_1}.$$

Ahora como y es C^2 , podemos intercambiar las derivadas con respecto a x_1 y a x . Además $\bar{y}(x)$ es también un punto crítico (un extremo) para el problema

$$G(y) = \int_a^{\bar{x}_1} F(x, y, y') dx$$

con $y(a) = A$, $y(\bar{x}_1) = \varphi(\bar{x}_1)$. Por lo tanto $F_{y'}(x, \bar{y}(x), \bar{y}'(x))$ es C^1 , con derivada $F_y(x, \bar{y}(x), \bar{y}'(x))$. Podemos integrar por partes y obtener:

$$\frac{dI}{dx_1}(\bar{x}_1) = 0 = F|_{x_1} + F_{y'} \frac{\partial}{\partial x_1} (y(x, \bar{x}_1))|_{x=\bar{x}_1}.$$

Como $y(a, x_1) = A$, entonces $\frac{\partial}{\partial x_1}(y(a, x_1)) = 0$. Como $y(x_1, x_1) = \varphi(x_1)$, entonces $\varphi'(x_1) = \frac{\partial y}{\partial x}(x_1, x_1) + \frac{\partial y}{\partial x_1}(x_1, x_1)$. Por lo tanto

$$\begin{aligned} \frac{\partial}{\partial x_1} y(x, \bar{x}_1)|_{x=\bar{x}_1} &= \varphi'(\bar{x}_1) - \frac{\partial y}{\partial x}(\bar{x}_1, \bar{x}_1) \\ &= \varphi'(\bar{x}_1) - y'(\bar{x}_1). \end{aligned}$$

Obtenemos entonces la condición de transversalidad.

Ejercicio 8.1 : Si los dos extremos de integración varían sobre las curvas $y = \varphi_0(x_0)$, $y = \varphi_1(x_1)$, encontrar por el método de multiplicadores (hay una condición más: $G_0(y, x_0, x_1) = y(x_0) - \varphi_0(x_0) = 0$) y por el método de la nota, las dos condiciones de transversalidad.

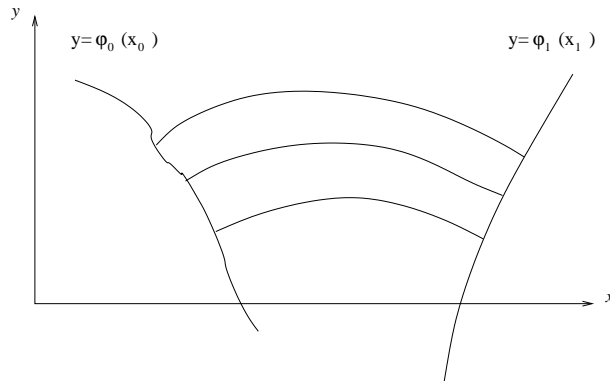


Fig. 8.4

Ejemplo 8.4 : Curvas entre dos superficies.

Supongamos que las superficies están dadas por

$$x_0 = \varphi_0(y, z), \quad x_1 = \varphi_1(y, z)$$

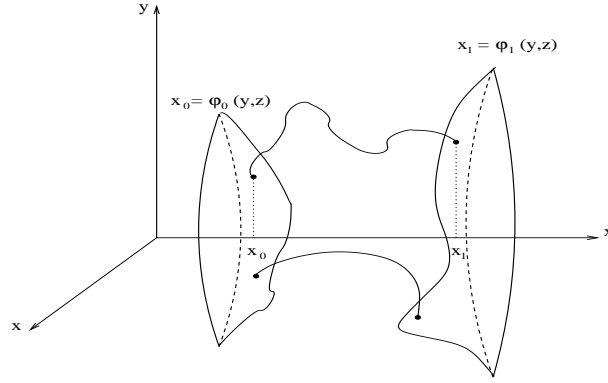


Fig. 8.5

Las curvas en el espacio son de la forma $(x, y(x), z(x))$ y se busca un punto crítico del funcional

$$J(y, z) = \int_{x_0}^{x_1} F(x, y(x), z(x), y'(x), z'(x)) dx$$

donde x_0 debe estar sobre la superficie $x_0 = \varphi(y, z)$ y x_1 sobre la otra superficie.

Extenderemos de nuevo las funciones $y(x), z(x)$ linealmente de cada lado de tal forma que podamos tomar límites fijos a y b . Si F no está ya definido en $[a, b] \times \mathbb{R}^4$, lo extendemos linealmente (esto requiere que F_x sea C^1 en a y en b). Trataremos entonces el funcional:

$$J(y, z, x_0, x_1) = \int_{x_0}^{x_1} F(x, y, z, y', z') dx$$

condicionado por:

$$G_0(y, z, x_0, x_1) = x_0 - \varphi_0(y(x_0), z(x_0)) = 0$$

$$G_1(y, z, x_0, x_1) = x_1 - \varphi_1(y(x_1), z(x_1)) = 0$$

en el espacio $C^1[a, b] \times C^1[a, b] \times \mathbb{R} \times \mathbb{R}$, con la norma $\|\cdot\|_1$ en las componentes C^1 y el valor absoluto en \mathbb{R} .

Como

$$\begin{aligned} & G_i(y + h, z + k, x_0 + \alpha, x_1 + \alpha_i) - G_i(y, z, x_0, x_1) \\ &= \alpha_i - \left[\varphi_i(y(x_i + \alpha_i) + h(x_i + \alpha_i), z(x_i + \alpha_i) + k(x_i + \alpha_i)) \right. \\ &\quad \left. - \varphi_i(y(x_i), z(x_i)) \right] \\ &= \alpha_i - \frac{\partial \varphi_i}{\partial y}(y(x_i), z(x_i))(y'(x_i)\alpha_i + h(x_i)) - \frac{\partial \varphi_i}{\partial z}(z'(x_i)\alpha_i + k(x_i)) \end{aligned}$$

$$+ o(|h|_1 + |k|_1 + |\alpha_i|)$$

ya que el término $h'(x_i)\alpha_i$ es de orden 2.

La relación

$$\nu DJ(y, z, x_0, x_1)(h, k, \alpha_0, \alpha_1) + \sum_{i=0}^1 \lambda_i DG_i(y, z, x_0, x_1)(h, k, \alpha_0, \alpha_1) = 0$$

da

$$\begin{aligned} & \nu \left[\int_{x_0}^{x_1} (F_y h + F_{y'} h' + F_z k + F_{z'} k') dx + F|_{x_1} \alpha_1 - F|_{x_0} \alpha_0 \right] + \\ & + \sum_{i=0}^1 \lambda_i (\alpha_i - \varphi_{iy}(y' \alpha_i + h) - \varphi_{iz}(z' \alpha_i + k))|_{x_i} = 0. \end{aligned}$$

Tomando $\alpha_0 = \alpha_1 = 0$, $k \equiv 0$, h en $C_0^\infty[x_0, x_1]$ con $h \equiv 0$ en $[a, x_0] \cup [x_1, b]$ y después $h \equiv 0$ y el correspondiente k , tendremos que $\nu F_{y'}$ es C^1 en (x_0, x_1) con derivada νF_y , $\nu F_{z'}$ es C^1 en (x_0, x_1) , con derivada νF_z y podemos integrar por partes la integral y obtener para ella:

$$\nu(F_{y'} h + F_{z'} k)|_{x_0}^{x_1}.$$

Como h y k son independientes, podemos tomarlos con valores 0 y 1 en x_0 y x_1 y viceversa, por lo tanto cada factor de $h(x_i)$, $k(x_i)$, α_i debe ser 0, para $i = 0$ y 1.

$$\begin{aligned} \nu F_{y'} & - \lambda_1 \varphi_{1y}|_{x_1} = 0, \quad \nu F_{y'} + \lambda_0 \varphi_{0y}|_{x_0} = 0 \\ \nu F_{z'} & - \lambda_1 \varphi_{1z}|_{x_1} = 0, \\ \nu F_{z'} & + \lambda_0 \varphi_{0z}|_{x_0} = 0 \end{aligned}$$

$$\nu F + \lambda_1(1 - \varphi_{1y}y' - \varphi_{1z}z')|_{x_1} = 0, \quad F - \lambda_0(1 - \varphi_{0y}y' - \varphi_{0z}z')|_{x_0} = 0.$$

Si $\nu = 0$ y $\lambda_0 \neq 0$, por ejemplo, vemos, de las ecuaciones en x_0 , que $\varphi_{0y} = \varphi_{0z} = 0$ lo cual contradice la tercera ecuación. Podemos por lo tanto dividir por ν y suponer que ν vale 1.

Si restamos de la tercera ecuación (en x_1) la primera multiplicada por y' , y la segunda multiplicada por z' , obtenemos:

$$\begin{aligned} F & - y'F_{y'} - z'F_{z'}|_{x_1} = -\lambda_1 \\ F & - y'F_{y'} - z'F_{z'}|_{x_0} = \lambda_0. \end{aligned}$$

Proposición 8.1 : *Tendremos entonces las condiciones de transversalidad*

$$\begin{aligned} F_{y'} + \varphi_{iy}(F - y'F_{y'} - z'F_{z'})|_{x_i} &= 0, \quad i = 0, 1 \\ F_{z'} + \varphi_{iz}(F - y'F_{y'} - z'F_{z'})|_{x_i} &= 0, \quad i = 0, 1 \end{aligned}$$

con las ecuaciones de Euler

$$\frac{d}{dx}(F_{y'}) - F_y = 0, \quad \frac{d}{dx}(F_{z'}) - F_z = 0 \quad \text{en } (x_0, x_1).$$

Notemos que la normal a la superficie $x - \varphi_i(y, z) = 0$ es $(1, -\varphi_{iy}, -\varphi_{iz})^T$ y que la tangente a la curva es $(1, y', z')^T$ de tal modo que $1 - y'\varphi_{iy} - z'\varphi_{iz}$ es el producto escalar de los dos vectores. Si $F|_{x_i} \neq 0$ la curva atraviesa a la superficie (ver la tercera ecuación).

Ejercicio 8.2 : Consideremos el funcional

$$J(y, z) = \int_0^{x_1} (y'^2 + z'^2 + 2yz) dx$$

con $y(0) = 0 = z(0)$.

1. Encontrar los puntos críticos usando las condiciones naturales en x_1 . Resolver las ecuaciones de Euler, buscando ecuaciones para $y + z$ y $y - z$.
2. Encontrar los puntos críticos usando las condiciones de transversalidad en el plano $x = x_1$.
3. Evaluando el residuo para funciones (h, k) adicionadas (pensar en senos y cosenos) probar que los puntos críticos encontrados no son ni mínimos ni máximos si $x_1 > \pi/2$. Probar que para $x_1 < \pi/2$ hay un solo mínimo y que para $x_1 = \pi/2$ hay una familia de mínimos.

Ejercicio 8.3 : Aplicar las condiciones de transversalidad de la proposición 8.1 al funcional:

$$J(y, z) = \int_{x_0}^{x_1} f(x, y, z)(1 + y'^2 + z'^2)^{1/2} dx$$

y dar una interpretación geométrica del resultado.

Ejercicio 8.4 : Resolver el problema del ejemplo 8.4 con el método de la nota 8.1, es decir, suponiendo que existe una familia de soluciones al problema con condiciones de frontera dadas parametrizada por t tal que $x_i(t) = \varphi_i(y(x, t), z(x, t))$.

Este ejercicio es para que el lector se de cuenta de lo fácil que es equivocarse con este método. En caso de dificultades puede acudir a alguno de los textos clásicos.

Ejemplo 8.5 : Problema con una discontinuidad no fija. Condición de Weierstrass II.

Veremos como usar los multiplicadores de Lagrange para obtener extremos de $J(y) = \int_a^b F(x, y, y') dx$ donde y' puede tener varias discontinuidades, $y(a) = A$, $y(b) = B$. De nuevo, como en el ejemplo 7.4, el funcional se interpreta como

$$\sum_i \int_{x_i}^{x_{i+1}} F(x, y, y') dx$$

donde (x_i, x_{i+1}) es un intervalo de continuidad de y' . Podemos suponer que el candidato a extremo y tiene una sola discontinuidad en y' en c (restringiéndonos a dos intervalos de continuidad como en el ejemplo 7.4). De todos modos no se puede trabajar en el espacio (no es lineal) de funciones con una sola discontinuidad.

De hecho tomaremos como espacio de trabajo a $X = \{h \text{ con un número finito de saltos en } h \text{ y en } h'\}$ con $h(a) = h(b) = 0$. Veremos la razón de permitir saltos en h más adelante.

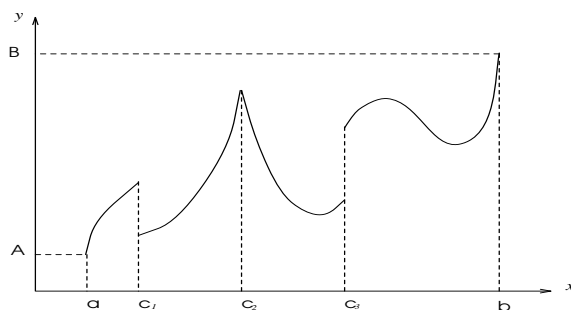


Fig. 8.6

Pondremos como norma sobre X :

$$\|h\| = \sum_j \|h\|_{C^1[I_j]}$$

donde $I_j, \dots, I_{N(h)}$ son los intervalos de continuidad de h y de h' (dependen de h). X es un espacio normado pero no es completo (límites de funciones en X podrían tener un número infinito de discontinuidades). Recordemos solamente que el teorema de multiplicadores de Lagrange es válido con un espacio normado.

Para h en X definimos dos funciones h_- y h_+ . h_- es tal que $h_-(x) = h(x)$ hasta la primera discontinuidad de h y lineal después. h_+ es tal que $h_+(x) = h(x)$ después de la última discontinuidad y lineal antes.

También definiremos y_+ y y_- para funciones C^1 por pedazos pero que no satisfacen las condiciones en a y b .

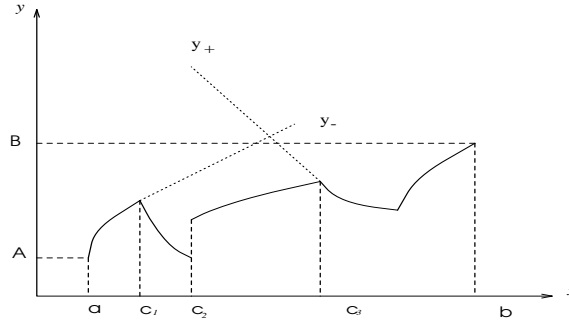


Fig. 8.7

Sea

$$J(y, c) = \int_a^c F(x, y_-, y'_-) dx + \int_c^b F(x, y_+, y'_+) dx.$$

Si condicionamos este problema por:

$$G(y, c) = y_-(c) - y_+(c) = 0$$

tendremos el funcional original con una sola discontinuidad en y' en c . Si $G(y+h, c+\alpha) = 0$, entonces tendremos una familia de curvas $(y+h)_-(x)$, para $x < c+\alpha$, $(y+h)_+(x)$, para $x \geq c+\alpha$, con una sola discontinuidad en su derivada en $c+\alpha$.

Entonces

$$\begin{aligned} J(y+h, c+\alpha) - J(y, c) &= \int_a^c (F(x, (y+h)_-, (y+h)_-') - F(x, y_-, y'_-)) dx + \\ &\int_c^{c+\alpha} (F(x, (y+h)_-, (y+h)_-') - F(c, y_-(c), y'_-(c))) dx + \\ &\int_c^b (F(x, (y+h)_+, (y+h)_+') - F(x, y_+, y'_+)) dx + \\ &\int_c^{c+\alpha} (F(c, y_+(c), y'_+(c)) - F(x, (y+h)_+, (y+h)_+')) dx + \\ &\alpha F(c, y_-(c), y'_-(c)) - \alpha F(c, y_+(c), y'_+(c)). \end{aligned}$$

Tomaremos direcciones h con una sola discontinuidad en c (en h o en h'), ya que estas serán suficientes para dar el resultado y que no está claro que J tenga derivada de Gâteaux para todo h en X .

Para esta clase de funciones uno tiene :

$$\begin{aligned} (y+h)_- &= y_- + h_- \text{ para } x \leq c \\ (y+h)_+ &= y_+ + h_+ \text{ para } x \geq c. \end{aligned}$$

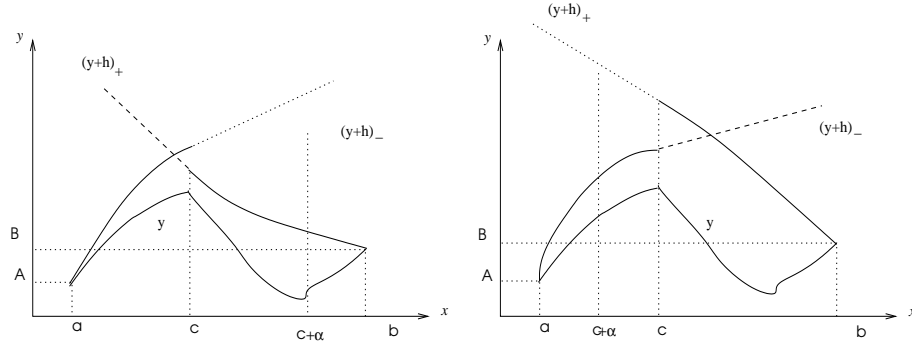


Fig. 8.8

La función $(y+h)_-$ para $x \leq c+\alpha$, $(y+h)_+$ para $x > c+\alpha$, que se usa en $J(y+h, c+\alpha)$, tendrá una discontinuidad en $c+\alpha$. En la figura 8.8, h tiene un salto en c , pero invitamos al lector a verificar que aún para h en C^∞ , se tendrá ese salto en $c+\alpha$. De ahí la necesidad de ampliar el espacio X .

El primer término es entonces:

$$\int_a^c (F_y(x, y_-, y'_-) h_- + F_{y'}(x, y_-, y'_-) h'_-) dx + o(|h_-|_1),$$

el tercero:

$$\int_c^b (F_y(x, y_+, y'_+) h_+ + F_{y'}(x, y_+, y'_+) h'_+) dx + o(|h_+|_1).$$

El segundo y el cuarto son acotados por $|\alpha|(|\alpha| + \|h\|)$ (vea la figura 8.8). Por lo tanto la derivada de Gâteaux será

$$D_{(h,\alpha)} J(y, c) = \int_a^c (F_y(-)h + F_{y'}(-)h') dx + \int_c^b (F_y(+)h + F_{y'}(+)h') dx + [F]_c \alpha$$

donde $F(-)$, $F(+)$ corresponden a y_- y a y_+ .

Del mismo modo

$$\begin{aligned} & G(y + h, c + \alpha) - G(y, c) \\ &= (y + h)_-(c + \alpha) - (y + h)_-(c) + (y + h)_-(c) - y_-(c) \\ &\quad - [(y + h)_+(c + \alpha) - (y + h)_+(c) + (y + h)_+(c) - y_+(c)] \\ &= y'_-(c)\alpha - y'_+(c)\alpha + [h]_c + h'_-(c)\alpha - h'_+(c)\alpha + \dots = [y']_c \alpha + [h]_c + o(|\alpha| + \|h\|) \end{aligned}$$

ya que $h'(c)\alpha$ es de orden cuadrático.

Por lo tanto, para esas direcciones

$$D_{(h,\alpha)}G(y, c) = [y']_c\alpha + [h]_c.$$

Por el teorema de multiplicadores (dejamos al lector ver que la demostración es la misma) :

$$\nu D_{(h,\alpha)}J(y, c) + \lambda D_{(h,\alpha)}G(y, c) = 0$$

para todo h en $C^1([a, c] \cup (c, b])$ con $h(a) = h(b) = 0$, y todo α en \mathbb{R} .

Tomando $\alpha = 0$, $[h]_c = 0$, vemos que $\nu F_{y'}(-)$ tiene derivada en (a, c) igual a $\nu F_y(-)$; y $\nu F_{y'}(+)$ tiene derivada en (c, b) igual a $\nu F_y(+)$. Podemos integrar por partes y obtener:

$$\nu(F_{y'}(-)h_- - F_{y'}(+)h_+)_c + \nu[F]_c\alpha + \lambda([y']_c\alpha + [h]_c) = 0$$

es decir:

$$[\nu F_{y'}h + \nu F\alpha + \lambda y'\alpha + \lambda h]_c = 0.$$

Tomando $\alpha = 0$, $[h]_c = 0$, $h(c) \neq 0$, tenemos $\nu[F_{y'}]_c = 0$, por lo tanto:

$$(\nu F_{y'} + \lambda)[h]_c = 0, \quad [\nu F + \lambda y']_c = 0$$

donde $[h]$ es arbitrario.

Esto implica que ν no puede ser 0, ($\nu F_{y'} + \lambda = 0$), y sustituyendo λ en la segunda ecuación

$$[F - y'F_{y'}]_c = 0, \quad [F_{y'}]_c = 0$$

las dos condiciones de Weierstrass, obtenidas en el ejemplo 7.4.

El lector se preguntará porqu  se definieron las funciones y_- y y_+ en lugar de trabajar directamente con el funcional $J(y, c)$ el cual es perfectamente definido para funciones en X y de hecho Fr chet diferenciable. La raz n es que en ese caso el funcional $G(y, c)$ no es ni siquiera continuo (uno cuantos dibujos convencer n al lector). Aqu  se ve claramente la diferencia entre un funcional en forma integral, como J , y un funcional en forma puntual (o local) como G .

Si bien se puede definir el funcional J para un espacio m s grande, el funcional G no puede ser continuo sobre un espacio de funciones discontinuas. Esto nos lleva tambi n a restringir las direcciones de las derivadas de G teaux. Si bien en este caso la prueba presentada en el ejemplo 7.4 es m s simple, esta no se aplica en el ejemplo siguiente.

Al final de la nota 8.2 daremos una  ltima manera de probar este resultado.

Ejemplo 8.6 : Reflexi n sobre una curva.

Supongamos que $y(x)$ es un punto cr tico del funcional

$$J(y) = \int_a^b F(x, y, y') dx$$

$y(a) = A$, $y(b) = B$, con la condición de que las curvas $y(x)$ deben tocar una curva dada $y = \varphi(x)$. En el punto de contacto y' puede ser discontinua.

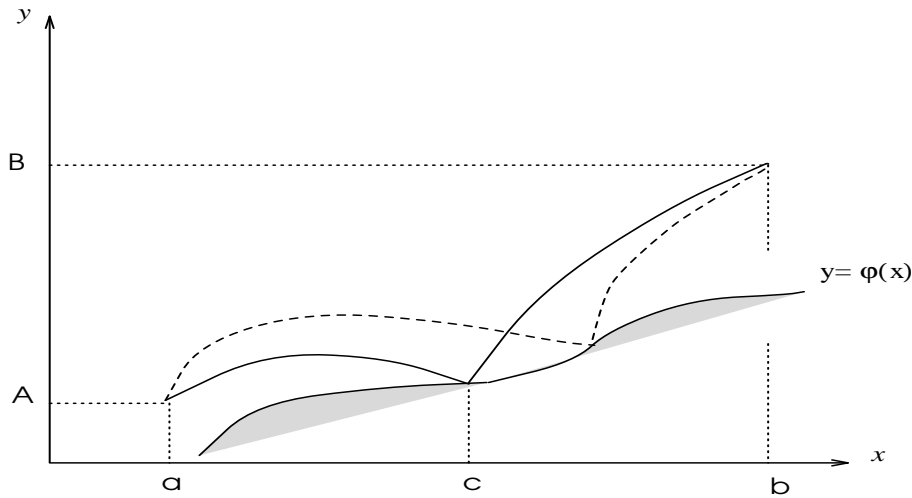


Fig. 8.9

Esto pasa con rayos de luz emitidos en (a, A) , observados en (b, B) después de una reflexión sobre el espejo $y = \varphi(x)$. Si la curva es vertical en un punto c , tendríamos el problema con una discontinuidad y la condición de continuidad de Weierstrass.

Sea y el candidato a extremo, con una discontinuidad en y' en c . Sea

$$B = \{h \text{ en } C^1[a, c] \cup C^1[c, b], h(a) = h(c) = h(b) = 0\}$$

con norma: $\|h\| = |h|_{C^1[a, c]} + |h|_{C^1[c, b]}$. B es un espacio de Banach y y es un punto crítico de J sobre B , ya que podemos tomar curvas $y(x)$ tocando la curva $y = \varphi(x)$ en $(c, \varphi(c))$.

La derivada de Fréchet de J es:

$$DJ(y)h = \int_a^c (F_y h + F'_y h') dx + \int_c^b (F_y h + F'_y h') dx$$

tomando h en $C_0^\infty[a, c]$ o en $C_0^\infty[c, b]$, tenemos que $F_{y'}$ es C^1 con derivada F_y en (a, c) y en (c, b) .

Tomemos ahora el espacio X , del ejemplo 8.5, de las funciones C^1 por partes con discontinuidades en h y en h' . Definimos el mismo funcional $J(y, c)$ (que tiene derivadas de Gâteaux para h en $C^1([a, c] \cup (c, b])$). Daremos ahora dos condiciones:

$$G_+(y, c) = y_+(c) - \varphi(c) = 0$$

$$G_-(y, c) = y_-(c) - \varphi(c) = 0.$$

Tenemos entonces

$$\begin{aligned} G_+(y + h, (c + \alpha)) - G_+(y, c) \\ = (y + h)_+(c + \alpha) - (y + h)_+(c) + (y + h)_+(c) - y_+(c) - (\varphi(c + \alpha) - \varphi(c)) \\ = (y + h)'_+(c)\alpha + h_+(c) - \varphi'(c)\alpha + o(|\alpha| + |h|_{C^1[c, b]}) \end{aligned}$$

para h, y , con una única singularidad en c . Por lo tanto, usando integración por partes:

$$\nu ([F_{y'}h] + [F]\alpha) + \mu_+ ((y'_+ - \varphi')\alpha + h_+) + \mu_- ((y'_- - \varphi')\alpha + h_-) = 0.$$

Tomando $\alpha = 0$, $h_+(c) = 0$, $h_-(c) \neq 0$: $\nu F_{y'}(-) + \mu_- = 0$.

Tomando $\alpha = 0$, $h_+(c) \neq 0$, $h_-(c) = 0$: $-\nu F_{y'}(+) + \mu_+ = 0$.

Tomando $\alpha \neq 0$, y $h = 0$: $\nu[F] + \mu_+(y'_+ - \varphi') + \mu_-(y'_- - \varphi') = 0$.

De las dos primeras ecuaciones vemos que $\nu \neq 0$: podemos ponerlo igual a 1. Substituyendo μ_+ , μ_- en la tercera:

$$[F] + F_{y'}(+)(y'_+ - \varphi') - F_{y'}(-)(y'_- - \varphi') = 0$$

es decir:

$$[F - (y' - \varphi')F_{y'}]_c = 0.$$

Notemos que aquí $F_{y'}$ no es continua en c (sumando las ecuaciones obtenidas para $\alpha = 0$ tenemos $[F_{y'}]_c = -\frac{(\mu_- + \mu_+)}{\nu}$).

Por ejemplo si $F = (1 + y'^2)^{1/2}v^{-1}(x, y)$, como en el caso de un rayo de luz en el plano (x, y) , obtenemos:

$$\left[(1 + y'\varphi')(1 + y'^2)^{-1/2} \right]_c = 0$$

$(v(x, y)$ es continuo). En términos de ángulos:

$$\begin{aligned} \cos \theta_1 + \sen \theta_1 \tan \psi &= \cos \theta_2 + \mathbf{sen} \theta_2 \tan \psi \\ \text{o sea : } \cos(\theta_1 - \psi) &= \cos(\theta_2 - \psi). \end{aligned}$$

Con soluciones $\theta_1 = \theta_2$: el rayo atraviesa el espejo, o:

$$\theta_2 - \psi = \psi - \theta_1$$

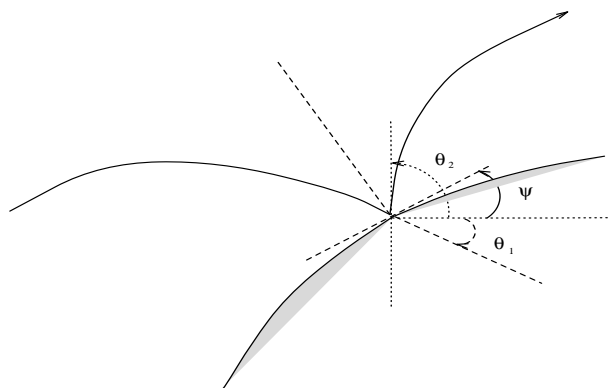


Fig. 8.10

es decir el ángulo del rayo incidente con la tangente al espejo es igual al ángulo de la tangente al espejo con el rayo reflejado.

Nota 8.2 : Daremos otra prueba de este resultado, usando el siguiente truco. Sea $y(x)$ una curva admisible con una discontinuidad en $(c + \alpha, \varphi(c + \alpha))$

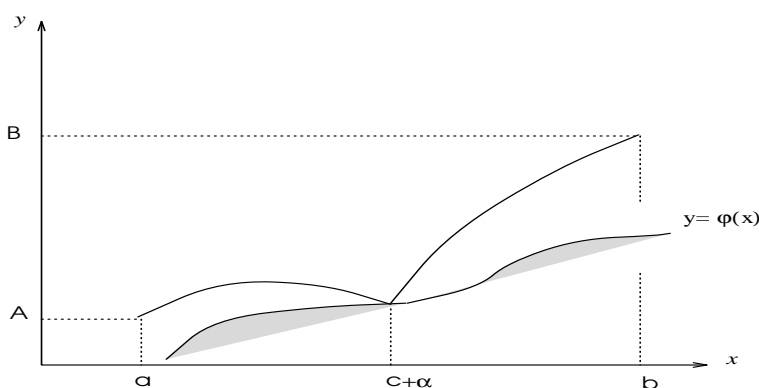


Fig. 8.11

Podemos parametrizar esta curva por $[0, 1]$, con $(x(t), y(t))$ de tal forma que $x(1/2) = c + \alpha$. Notemos que con este cambio de parametrización $y(t)$ puede tener una discontinuidad sólo si $t = 1/2$, pero no aseguramos que la curva toque la curva $y = \varphi(x)$ en $t = 1/2$.

Pondremos como condición :

$$G(x, y) = y(1/2) - \varphi(x(1/2)).$$

El funcional se vuelve entonces :

$$\begin{aligned}
J(x, y) &= \int_0^{1/2} \\
F(x(t), & \quad y(t), y'(t)/x'(t))x'(t)dt \\
&+ \int_{1/2}^1 F(x, y, y'/x')x' dt.
\end{aligned}$$

Tomemos como espacio de trabajo al conjunto $\{h(t), k(t) \text{ en } C^1[0, 1/2] \cup C^1[1/2, 1], \text{ con } h(0) = h(1) = 0, k(0) = k(1) = 0, h \text{ y } k \text{ continuas en } 1/2\}$ con la norma $\| \cdot \|_{C^1[0, 1/2]} + \| \cdot \|_{C^1[1/2, 1]}$.

Las condiciones en 0 y en 1 vienen del hecho que $x(t)$ y $y(t)$ son dadas en los extremos. Si F es C^1 es claro que J y G tienen derivadas de Fréchet en ese espacio, tendremos pues:

$$\begin{aligned}
&\nu \left(\int_0^{1/2} (Fh' + F_x x'h - F_z(y'/x')h' + F_y x'k + F_z k')dt \right. \\
&+ \int_{1/2}^1 (Fh' + F_x x'h - F_z(y'/x')h' + F_y x'k + F_z k')dt) \\
&+ \lambda(k(1/2) - \varphi'(x(1/2))h(1/2)) = 0
\end{aligned}$$

para todo (h, k) en el espacio.

Tomando h, k alternativamente 0 y en $C_0^\infty(0, 1/2)$ o en $C_0^\infty(1/2, 1)$, tenemos que $F - F_z y'/x'$ es C^1 con derivada $F_x x'$ y que F_z es C^1 con derivada $F_y x'$. (Estas dos ecuaciones de Euler se reducen de hecho a la ecuación de Euler para $y(x)$, como lo veremos en la sección de problemas paramétricos.) Lo que nos interesa aquí es que podemos integrar por partes:

$$\nu[(F - F_z(y'/x'))h(1/2) + F_z k(1/2)]_{1/2} + \lambda(k - \varphi'h)|_{1/2} = 0.$$

Escogiendo adecuadamente $h(1/2)$ y $k(1/2)$, tenemos:

$$\nu[F - F_z y'/x']_{1/2} - \lambda\varphi'(x(1/2)) = 0$$

$$\nu[F_z]_{1/2} + \lambda = 0.$$

De la segunda ecuación, $\nu \neq 0$ y recordando que $\frac{dy}{dx} = y'/x'$ obtenemos,

$$[F - (\frac{dy}{dx} - \varphi')F_z]_c = 0.$$

Notemos que en el caso del problema con discontinuidades, del Ejemplo 8.5, podemos usar el mismo método, pero sin condición. Los mismos resultados (con $\lambda = 0$) son válidos, obtenemos entonces:

$$[F - F_z y' / x']_c = 0$$

$$[F_z]_c = 0,$$

es decir, (con $\frac{dy}{dx} = y' / x'$) las dos condiciones de Weierstrass.

Ejercicio 8.5 : Se puede también tomar h y k discontinuas en $1/2$. En ese caso hay dos condiciones más:

$$x_-(1/2) - x_+(1/2) = 0, \quad y_-(1/2) - y_+(1/2) = 0.$$

Usando 3 multiplicadores, llegar al mismo resultado.

Ejercicio 8.6 : Tratar el problema de difracción por una curva de las mismas maneras (como en el ejemplo 8.6, como en la nota 8.2 y como en el ejercicio anterior).

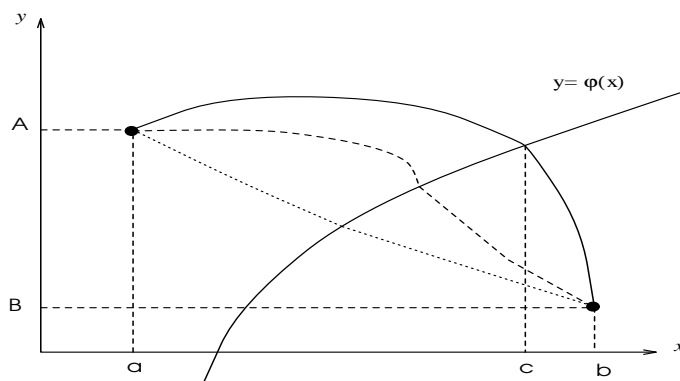


Fig. 8.12

Estudiar en particular el caso $F(x, y, z) = (1+z^2)^{1/2} v(x, y)^{-1}$ con su interpretación geométrica.

Ejercicio 8.7 : Condiciones de continuidad para segundas derivadas.

1. Para problemas con $\int_a^b F(x, y, y', y'') dx$, con discontinuidades no fijas, usar el método de multiplicadores, como en el ejemplo 8.5, para encontrar las tres condiciones de continuidad del ejercicio 7.3.
2. Usar el método de la nota 8.2 para obtener el mismo resultado.
3. Usar cualquiera de estos dos métodos si además hay una condición:

- α) Que la curva tenga su discontinuidad en y'' sobre una curva dada $y = \varphi(x)$.
 β) Que la curva tenga una discontinuidad en y' o en y'' sobre una curva dada $y = \varphi(x)$.

4.9 Problemas isoperimétricos

Una gran fuente de ejemplos donde se usan los multiplicadores de Lagrange proviene de los problemas de Geometría y de Mecánica. A todos ellos se les agrupa bajo el nombre de problemas isoperimétricos y se dividen en tres clases según la condición que se impone al problema.

La primera categoría es la de los problemas condicionados por un funcional en forma integral. Esto es el caso del problema de la curva de longitud dada que encierra una área máxima. La segunda categoría engloba a los problemas condicionados por un *operador no lineal* que depende de (x, y) . Estos se llaman restricciones holonómicas en Mecánica y corresponden a posición de cuerpos con ataduras rígidas, o por ejemplo, a encontrar una geodésica sobre una superficie dada (la superficie *no* es un funcional).

La última categoría reúne a los problemas donde las condiciones son *operadores no lineales* que dependen de las derivadas. A éstas se les llama restricciones no holonómicas o ataduras elásticas o viscosas.

La diferencia más importante desde el punto de vista matemático es que las dos últimas categorías son totalmente diferentes: las condiciones son a valores en otro espacio y no en los reales. Queda por lo tanto claro que no se pueden aplicar los mismos argumentos.

Veremos ejemplos de cada una de estas categorías.

4.9.1 Problema isoperimétrico clásico

Si el funcional $J(y) = \int_a^b F(x, y, y') dx$ con $y(a) = A$, $y(b) = B$ está condicionado por $G(y) = \int_a^b G(x, y, y') dx - K = 0$ donde F y G son C^1 , entonces tenemos:

$$\nu \int_a^b (F_y h + F_{y'} h') dx + \lambda \int_a^b (G_y h + G_{y'} h') dx = 0,$$

es decir:

$$\int_a^b ((\nu F_y + \lambda G_y)h + (\nu F_{y'} + \lambda G_{y'})h') dx = 0.$$

Teorema 9.1 : *En este caso $\nu F_{y'} + \lambda G_{y'}$ es C^1 con derivada*

$$\nu F_y + \lambda G_y.$$

Notemos que la generalización a más funciones, más derivadas, más variables y más condiciones no presenta ninguna dificultad. Dejamos al lector hacer estas generalizaciones como ejercicio.

Ejemplo 9.1 : Catenaria.

Si tenemos una cadena de longitud L , sujeta entre dos puntos (a, A) y (b, B) , la posición de reposo será tal que la energía potencial debida al peso sea mínima.

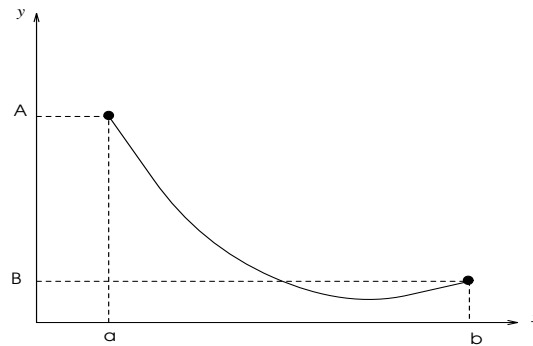


Fig. 9.1

La energía potencial es $J(y) = \int_0^L \rho g y(s) ds$ y la longitud $L = \int_a^b (1 + y'^2)^{1/2} dx$. Aquí $J(y) = \int_a^b \rho g y (1 + y'^2)^{1/2} dx$. Por lo tanto,

$$(\nu \rho g y y' (1 + y'^2)^{-1/2} + \lambda y' (1 + y'^2)^{-1/2})' = \nu \rho g (1 + y'^2)^{1/2}.$$

Si $\nu = 0$ tendremos : $y' (1 + y'^2)^{1/2} = \text{constante}$, es decir $y' = \text{constante}$, la solución es una recta pasando por los dos puntos, pero será la solución al problema sólo si L es la distancia entre los dos puntos. En cualquier otro caso $\nu \neq 0$. La ecuación es entonces:

$$(z z' (1 + z'^2)^{-1/2})' = (1 + z'^2)^{1/2}$$

con $z = y + \lambda/(\nu \rho g)$, $z' = y'$.

Ahora bien, esta ecuación ya fue estudiada en el problema de la superficie de revolución: Ejemplos 1.5 y 2.2. La solución es un coseno hiperbólico (o dos o ninguno según los parámetros).

Ejemplo 9.2 : Área encerrada por una curva de longitud dada.

Aquí $J(x, y) = \int_0^{2\pi} (x y' - x' y) dt / 2$ con la condición $I(x, y) = \int_0^{2\pi} (x'^2 + y'^2)^{1/2} dt - L = 0$ sobre el espacio de funciones 2π -periódicas, ya que la curva debe ser cerrada.

Tendremos entonces dos ecuaciones de Euler:

$$\begin{aligned}(\nu F_{x'} + \lambda G_{x'})' &= \nu F_x + \lambda G_x \\ (\nu F_{y'} + \lambda G_{y'})' &= \nu F_y + \lambda G_y,\end{aligned}$$

es decir:

$$\begin{aligned}\nu y' &= \left(\frac{\lambda x'}{(x'^2 + y'^2)^{1/2}} \right)' \\ -\nu x' &= \left(\frac{\lambda y'}{(x'^2 + y'^2)^{1/2}} \right)'. \end{aligned}$$

Integrando obtenemos :

$$\begin{aligned}\nu y &= \frac{\lambda x'}{(x'^2 + y'^2)^{1/2}} + c_1 \\ -\nu x &= \frac{\lambda y'}{(x'^2 + y'^2)^{1/2}} + c_2.\end{aligned}$$

Luego, si $\nu \neq 0$

$$(x - x_0)^2 + (y - y_0)^2 = \left(\frac{\lambda}{\nu} \right)^2,$$

con $x_0 = -c_2/\nu$, $y_0 = c_1/\nu$. Por lo tanto la curva está sobre el círculo de centro (x_0, y_0) y radio $|\lambda/\nu|$. Como la curva debe ser cerrada, la curva debe cubrir todo el círculo y, siendo simple, debe cubrirlo una sola vez. Por lo tanto $|\lambda/\nu| = L/(2\pi)$ y $J(x, y) = L^2/(4\pi)$. Claramente el centro es arbitrario ya que una traslación de ejes no cambia ni la longitud ni el área. Falta todavía probar que esta solución corresponde a un máximo.

Por otro lado, si $\nu = 0$, $x' = k_1$, $y' = k_2$, por la constricción :

$$k_1^2 + k_2^2 = L^2/(4\pi).$$

De aquí que $x = k_1 t + x_0$, $y = k_2 t + y_0$; lo cual corresponde a una recta y por lo tanto a una solución inadmisibles.

Ejercicio 9.1 : Estudiar el problema del funcional:

$$J(y) = \int_a^b y \, dx, \quad y(a) = y(b) = 0 \quad y(x) \geq 0,$$

condicionado por $G(y) = \int_a^b (1 + y'^2)^{1/2} \, dx - L = 0$.

Encontrar la solución y probar que ésta es única si $b - a \leq L \leq \pi(b - a)/2$. Probar que no hay solución si alguna de estas desigualdades no se cumple. Interpretación geométrica.

Ejercicio 9.2 : Encontrar la solución al problema de hallar un punto crítico del funcional

$$J(y) = \int_a^b (y^2 + y'^2 + y''^2) dx$$

con condiciones de frontera naturales y condicionado por:

$$G(y) = \int_a^b y dx - (b - a) = 0.$$

Se encontrará una ecuación de 4o. orden, no homogénea. Probar que la ecuación homogénea tiene una única solución, multiplicando la ecuación por y e integrando por partes. Probar que la solución obtenida corresponde a un mínimo (estudiar el residuo).

4.9.2 Condiciones holonómicas

Consideremos el funcional:

$$J(y, z) = \int_a^b F(x, y, z, y', z') dx$$

con $y(x)$, $z(x)$ dados en a y en b y condicionadas por $G(x, y, z) = 0$.

Ahora bien $G(x, y(x), z(x))$ es una aplicación de $C^1[a, b] \times C^1[a, b]$ en $C^1[a, b]$, si G es C^1 . Es decir no es un funcional, ya que a un par de funciones da una función. Sin embargo, ya que tenemos una herramienta (los multiplicadores) que funciona para problemas condicionados (con *funcionales*), ¿cómo podemos convertir este problema en un problema donde sólo hay funcionales? La respuesta es usar otro espacio y definir el funcional:

$$G(y, z, u) = \int_a^b G(x, y(x), z(x))u(x) dx = 0$$

para u en $C_0^1[a, b]$. Claramente si G es C^1 el funcional $G(y, z, u)$ tiene derivada de Fréchet.

Además si para algún $(y(x), z(x))$ tenemos $G(x, y(x), z(x)) = 0$, entonces $G(y, z, u) = 0$ para todo u en $C_0^1[a, b]$ e inversamente, por el lema III.2, si $G(y, z, u) = 0$ para todo u en $C_0^1[a, b]$, entonces $G(x, y(x), z(x)) = 0$. Hemos pues encontrado una manera más complicada de expresar la condición. Por lo tanto tenemos:

$$\nu DJ(y, z, u)(h, k, l) + \lambda DG(y, z, u)(h, k, l) = 0$$

para todo (h, k, l) en $C^1[a, b] \times C^1[a, b] \times C^1[a, b]$ con valor 0 en a y en b .

$$\nu \int_a^b (F_y h + F_{y'} h' + F_z k + F_{z'} k') dx + \lambda \int_a^b ((G_y h + G_z k)u + Gl) dx = 0.$$

Por lo tanto tenemos tres ecuaciones de Euler:

$$\nu F_y + \lambda G_y u = \nu (F_{y'})'$$

$$\begin{aligned}\nu F_z + \lambda G_z u &= \nu (F_{z'})' \\ \lambda G &= 0.\end{aligned}$$

Si $\lambda = 0$ (entonces $\nu \neq 0$) tendremos un punto crítico libre de J . Si $\lambda \neq 0$, tendremos la condición $G(x, y(x), z(x)) = 0$ y, poniendo a $\lambda(x) = -\lambda u(x)$:

Teorema 9.2 : *Si $\int_a^b F(x, y, y', z, z') dx$ está condicionado por $G(x, y, z) = 0$, entonces existen una constante ν y una función $\lambda(x)$, con $|\nu| + |\lambda(x)| \neq 0$ tales que:*

$$\begin{aligned}\nu((F_{y'})' - F_y) + \lambda(x)G_y &= 0 \\ \nu((F_{z'})' - F_z) + \lambda(x)G_z &= 0.\end{aligned}$$

Ejemplo 9.3 : Geodésicas sobre una superficie.

Sea $F(t, x, y, z) = (x'^2 + y'^2 + z'^2)^{1/2}$, dando la longitud de una curva parametrizada por t , condicionada por el hecho que la curva debe estar sobre la superficie $G(x, y, z) = 0$, y pasando por dos puntos dados. Tendremos entonces:

$$\begin{aligned}\nu(x'(x'^2 + y'^2 + z'^2)^{-1/2})' + \lambda(t)G_x &= 0 \\ \nu(y'(x'^2 + y'^2 + z'^2)^{-1/2})' + \lambda(t)G_y &= 0 \\ \nu(z'(x'^2 + y'^2 + z'^2)^{-1/2})' + \lambda(t)G_z &= 0.\end{aligned}$$

Si s es la longitud de arco de la curva, $ds/dt = (x'^2 + y'^2 + z'^2)^{1/2}$, $\dot{x} = dx/ds = x'(x'^2 + y'^2 + z'^2)^{-1/2}$, y las ecuaciones se vuelven:

$$\nu \ddot{x}_i = -\lambda(t)(x'^2 + y'^2 + z'^2)^{-1/2} G_{x_i} \quad \text{para } x_i = x, y, z,$$

es decir la normal principal a la curva es paralela a ∇G , o sea a la normal a la superficie. Recuperamos así el resultado obtenido en el ejemplo 3.3.

Nota 9.1 : La manera clásica de obtener este resultado es usar el teorema de la función implícita de la manera siguiente:

A partir de $G(x, y, z) = 0$, podemos encontrar $z = z(x, y)$ si $G_z \neq 0$ y $G(a, A, \alpha) = 0$ (por ejemplo), en una vecindad de (a, A, α) de manera única. Ahora bien, si en lugar de (y, z) ponemos $(y(x), z(x))$ obtendremos $z = z(x, y(x))$, con $z(a) = z(a, A) = \alpha$ por la unicidad. Recordamos que entonces

$$\frac{dz}{dx} = z_x + z_y y' = -(G_x + G_y y')/G_z.$$

El funcional se vuelve ahora dependiente solamente de y :

$$I(y) = J(y, z(\cdot, y(\cdot))) = \int_a^b F(x, y, y', z(x, y(x)), z_x + z_y y') dx$$

con ecuación de Euler (aquí se supone que y y z son C^2 , así como G) :

$$F_y + F_z z_y + F_{z'}(z_{xy} + z_{yy} y') = (F_{y'} + F_{z'} z_y)'.$$

Por lo tanto:

$$F_y - (F_{y'})' + (F_z - (F_{z'})') z_y = 0.$$

Si definimos $\lambda(x) = -(F_y - (F_{y'})')/G_z$ tenemos:

$$F_y - (F_{y'})' + \lambda(x) G_y = 0$$

$$F_z - (F_{z'})' + \lambda(x) G_z = 0.$$

Notemos que quizás esta “prueba” sea más directa pero requiere de más hipótesis y sobre todo no se generaliza fácilmente al caso de problemas no holonómicos, como lo veremos a continuación.

4.9.3 condiciones no holonómicas

Consideramos de nuevo el funcional

$$J(y, z) = \int_a^b F(x, y, y', z, z') dx$$

con y, z dadas en a y b y condicionadas por

$$G(x, y, y', z, z') = 0.$$

Notemos primero que si $G_{z'} \neq 0$ podemos expresar $z' = f(x, y, y', z)$ y para $y(x)$ dada, tendríamos una ecuación diferencial para z con dos condiciones: $z(a) = \alpha$, $z(b) = \beta$, es decir, sin solución en general.

Si el lector quiere pasar por encima de esta dificultad, lo dejamos tratar de seguir el razonamiento de la nota 9.1. Por nuestra parte, seguiremos los pasos de la prueba del teorema 9.2.

Sea $G(y, z, u) = \int_a^b G(x, y(x), y'(x), z(x), z'(x))u(x) dx$ para u en $C_0^1[a, b]$. Como antes $G(y, z, u) = 0$ para todo u si y sólo si se cumple la condición no holonómica. Tendremos entonces:

$$\begin{aligned}\nu F_y + \lambda u G_y &= (\nu F_{y'} + \lambda u G_{y'})' \\ \nu F_z + \lambda u G_z &= (\nu F_{z'} + \lambda u G_{z'})' \\ \lambda G &= 0.\end{aligned}$$

Si $\lambda = 0$, tendremos un punto crítico libre de J . Si $\lambda \neq 0$, tendremos la condición $G = 0$ y poniendo $\lambda u(x) = \lambda(x)$:

Teorema 9.3 : *Si $\int_a^b F(x, y, y', z, z') dx$ está condicionado por $G(x, y, y', z, z') = 0$, existen (para un punto crítico) $\lambda(x)$, ν con $|\nu| + |\lambda(x)| \neq 0$ tales que:*

$$\begin{aligned}\nu F_y + \lambda(x) G_y - (\nu F_{y'} + \lambda(x) G_{y'})' &= 0 \\ \nu F_z + \lambda(x) G_z - (\nu F_{z'} + \lambda(x) G_{z'})' &= 0.\end{aligned}$$

Ejemplo 9.4 : Consideremos el funcional

$$\int_a^b F(x, y, z) dx$$

condicionado por $z - y' = 0$, (es decir el funcional clásico $\int_a^b F(x, y, y') dx$).

Tendremos:

$$\begin{aligned}\nu F_y + \lambda'(x) &= 0 \\ \nu F_z + \lambda(x) &= 0.\end{aligned}$$

De la segunda ecuación vemos que $\nu \neq 0$ y sustituyendo en la primera tenemos: $F_y - (F_{y'})' = 0$.

Ejemplo 9.5 : Problema isoperimétrico.

En el ejemplo 2.3, vimos que el problema de encontrar la curva encerrando una área máxima para una longitud dada, podía descomponerse en dos pedazos (ver la figura 2.4) y resultaba en encontrar la curva pasando por $(a, 0)$, $(b, 0)$ de longitud L_1 y con área:

$$\int_0^{L_1} yx' ds = \int_0^{L_1} y(1 - y'^2)^{1/2} ds = \int_a^b y dx$$

donde se usó la relación $x'^2 + y'^2 = 1$ y la condición $x' \geq 0$.

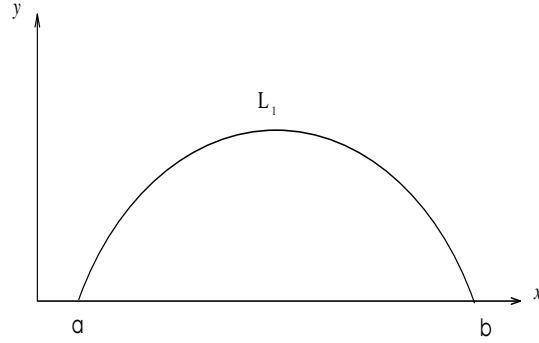


Fig. 9.2

Consideremos el funcional:

$$A(x, y) = \int_0^{L_1} yx' ds$$

con las condiciones de frontera $x(0) = a$, $x(L_1) = b$, $y(0) = y(L_1) = 0$ y la condición no holonómica:

$$G(x, y, x', y') = x'^2 + y'^2 - 1 = 0.$$

Tendremos entonces:

$$\nu x' - (2\lambda y')' = 0$$

$$(\nu y + 2\lambda x')' = 0.$$

Multiplicando la primera ecuación por $-\lambda y'$, la segunda por $\lambda x'$ y sumando:

$$(\lambda^2(x'^2 + y'^2))' = 0.$$

Como $x'^2 + y'^2 = 1$, λ es una constante.

Como en el ejemplo 9.2, $\nu \neq 0$, $\lambda \neq 0$ ($\nu = 0$ correspondería al segmento (a, b)). Obtenemos entonces

$$\begin{aligned}\nu(x - x_0) &= 2\lambda y' \\ \nu(y - y_0) &= -2\lambda x' .\end{aligned}$$

Tomando $k = \nu/(2\lambda)$ y $X = x - x_0$, $Y = y - y_0$, tenemos

$$\begin{aligned}X'' + k^2 X &= 0 \\ Y'' + k^2 Y &= 0 ,\end{aligned}$$

cuyas soluciones son :

$$\begin{aligned}X &= a \cos(k\theta + \varphi) \\ Y &= b \operatorname{sen}(k\theta + \psi) .\end{aligned}$$

De la relación $kX = Y'$, obtenemos $b = a$, $\varphi = \psi$. Es decir,

$$\begin{aligned}x(s) &= c_1 + c \cos(ks + \varphi), \\ y(s) &= c_2 + c \operatorname{sen}(ks + \varphi) .\end{aligned}$$

Ejercicio 9.3 : Calcular las constantes c_1, c_2, k, φ, c a partir de las condiciones en a y en b y a partir de la condición $x'^2 + y'^2 = 1$. Probar que la solución es única (si $b - a \leq L_1$) y es un arco de círculo que pasa por los dos puntos. ¿Por qué en este caso hay una sola condición ($b - a \leq L_1$) cuando en el ejercicio 9.1 está también la condición $L_1 \leq \pi(b - a)/2$?

Capítulo 5

Cambios de variables y Teoría de Hamilton-Jacobi

En este Capítulo veremos como se comportan las ecuaciones de Euler bajo cambios de variables y como, a partir de estos cambios de variables podemos derivar integrales primeras. Esto nos llevará además a examinar más a fondo el principio de mínima acción, los fundamentos de la mecánica de varias partículas y, a partir de la ecuación de Hamilton-Jacobi, como asentar las bases para el estudio de las condiciones suficientes y la teoría de campos.

5.1 Ecuaciones en forma paramétrica

Cuando le pedimos a un niño que nos dibuje una curva que una a dos puntos en el plano, el resultado será en general de la forma siguiente:

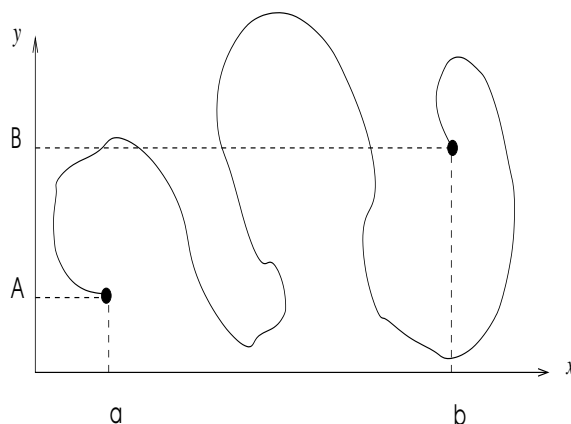


Fig. 1.1

Ahora bien, en estas notas hemos hablado de muchas curvas, casi todas de la forma $y = f(x)$. Seguramente que la reina Dido no tenía en mente solamente curvas parametrizadas por x cuando resolvió su problema.

Así pues cuando uno trabaja un problema geométrico, como el de las curvas, hay que tener presente el hecho que tal curva debe parametrizarse como $(x(t), y(t))$ y no solo como $y(x)$. (Ver el ejercicio IV-9.3) En este caso tendremos que considerar

$$I(x, y) = \int_0^T F(t, x(t), y(t), x'(t), y'(t)) dt$$

con dos ecuaciones de Euler:

$$\begin{aligned} \text{Euler}_x(F) &\equiv F_x - (F_{x'})' \\ \text{Euler}_y(F) &= F_y - (F_{y'})'. \end{aligned}$$

Ahora bien, el funcional es un funcional que depende de la curva: \tilde{I} (curva), no de su parametrización:

$$\begin{aligned} \tilde{I}(\text{curva}) &= \int_{t_0}^{t_1} F(t, x(t), y(t), x'(t), y'(t)) dt \\ &= \int_{\tau_0}^{\tau_1} F(\tau, x(\tau), y(\tau), x'(\tau), y'(\tau)) d\tau \end{aligned}$$

si τ es otro parámetro tal que $x(\tau_i) = x(t_i)$, $y(\tau_i) = y(t_i)$ para $i = 0, 1$. Si definimos el cambio de parámetros $\tau = \tau(t)$, con $\tau' > 0$ (para preservar la orientación de la curva), y denotamos por $\dot{x} = dx/d\tau$, $x' = dx/dt$, $\dot{y} = dy/d\tau$, $y' = dy/dt$, tenemos

$$\begin{aligned} \tilde{I}(\text{curva}) &= \int_{\tau_0}^{\tau_1} F(t(\tau), x, y, \dot{x}/\dot{t}, \dot{y}/\dot{t}) \dot{t} d\tau \\ &= \int_{\tau_0}^{\tau_1} F(\tau, x, y, \dot{x}, \dot{y}) d\tau \end{aligned}$$

para todo τ_0, τ_1 .

Debemos tener:

$$F(t(\tau), x, y, \dot{x}/\dot{t}, \dot{y}/\dot{t}) \dot{t} = F(\tau, x, y, \dot{x}, \dot{y})$$

para todo τ, x, y .

Por lo tanto F no debe depender explícitamente de τ y debe ser homogéneo de grado 1 en (\dot{x}, \dot{y}) .

Por la identidad de Euler:

$$\dot{x}F_{\dot{x}} + \dot{y}F_{\dot{y}} = F$$

(esto es un caso particular del teorema de Noether).

Además, como se vio en la nota IV.3.1, las ecuaciones de Euler no son independientes:

$$\dot{x} \text{ Euler}_x(F) + \dot{y} \text{ Euler}_y(F) = 0.$$

Para más variables, por ejemplo con superficies en lugar de curvas, el argumento que acabamos de exponer no se generaliza pero es claro que el hecho que una superficie sea un punto crítico de un funcional no debe depender de la parametrización escogida. Veremos esto en la siguiente sección.

Nota 1.1: Si, en el funcional $J(y) = \int_a^b F(x, y, y') dx$, parametrizamos $y(x)$ por $(x(t), y(t))$ tendremos, con $\dot{x} = dx/dt > 0$, (para preservar la orientación de la curva),

$$J(y) = \int_0^1 F(x(t), y(t), \dot{y}(t)/\dot{x}(t)) \dot{x}(t) dt = \tilde{J}(x, y)$$

Las ecuaciones de Euler para (x, y) , punto crítico, serán:

$$\begin{aligned} F_x \dot{x} - \frac{d}{dt}(F - F_{y'} \dot{y}/\dot{x}) &= 0 \\ F_y \dot{x} - \frac{d}{dt}(F_{y'}) &= 0. \end{aligned}$$

Es decir:

$$\begin{aligned} F_x - \frac{d}{dx}(F - y' F_{y'}) &= 0 \\ F_y - \frac{d}{dx}(F_{y'}) &= 0 \end{aligned}$$

recuperamos de esta manera las dos ecuaciones de Euler. Esto no debe sorprender al lector ya que para obtener la segunda ecuación de Euler se usó un caso particular de este argumento.

Nota: Hemos probado que la integral primera es consecuencia de tener un punto crítico.

5.2 Invariancia de las ecuaciones de Euler

En problemas de geometría (como en la sección anterior) o de física, los extremos de un funcional no dependen del sistema de coordenadas escogido. Veremos que en efecto los puntos críticos, es decir las soluciones de las ecuaciones de Euler, son independientes de los cambios de coordenadas.

Caso 2.1: Cambio de una variable independiente

Consideremos el funcional

$$J(y, z) = \int_a^b F(x, y, z, y', z') dx.$$

Si cambiamos la variable de x a $\tau(x)$, tendremos

$$\begin{aligned} J(y, z) &= \int_{\alpha}^{\beta} F(x(\tau), y, z, \dot{y}/\dot{x}, \dot{z}/\dot{x}) \dot{x} d\tau \\ &= \int_{\alpha}^{\beta} H(\tau, y, z, \dot{y}, \dot{z}) d\tau \end{aligned}$$

donde \cdot denota las derivadas con respecto a τ , $\alpha = \tau(a)$, $\beta = \tau(b)$, y' denota las derivadas con respecto a x .

Podemos escribir, si F es C^1 :

$$\begin{aligned} J(y + h, z + k) - J(y, z) &= \int_a^b (F_y h + F_{y'} h' + F_z k + F_{z'} k') dx + \dots \\ &= \int_{\alpha}^{\beta} (H_y h + H_{\dot{y}} \dot{h} + H_z k + H_{\dot{z}} \dot{k}) d\tau + \dots \end{aligned}$$

Por la unicidad de la derivada de Fréchet (Proposición II.3.1) los dos términos lineales son iguales. Cambiando de variable en la primera integral:

$$\int_{\alpha}^{\beta} ((F_y \dot{x} - H_y) h + (F_{y'} - H_{\dot{y}}) \dot{h} + (F_z \dot{x} - H_z) k + (F_{z'} - H_{\dot{z}}) \dot{k}) d\tau = 0$$

para toda (h, k) en el espacio de trabajo, en particular en $C_0^{\infty}[\alpha, \beta]$.

Por el Teorema III.1:

$$\begin{aligned} \frac{d}{d\tau}(F_{y'} - H_{\dot{y}}) &= F_y \dot{x} - H_y \\ \frac{d}{d\tau}(F_{z'} - H_{\dot{z}}) &= F_z \dot{x} - H_z. \end{aligned}$$

Que podemos escribir:

$$F_y \dot{x} - \frac{d}{d\tau}(F_{y'}) = (F_y - \frac{d}{dx}(F_{y'})) \dot{x} = H_y - \frac{d}{d\tau}(H_{\dot{y}}).$$

Es decir,

$$\begin{aligned} \text{Euler}_y(F)(x) \frac{dx}{d\tau} &= \text{Euler}_y(F \frac{dx}{d\tau})(\tau) \\ \text{Euler}_z(F)(x) \frac{dx}{d\tau} &= \text{Euler}_z(F \frac{dx}{d\tau})(\tau) \end{aligned}$$

Por lo tanto los puntos críticos no dependen de la parametrización.

Quizas el lector pensará que esto es mucho trabajo para algo obvio: en efecto, como $H = F\dot{x}$, tenemos $H_y = F_y\dot{x}$, $H_{\dot{y}} = F_{y'}\dot{x}/\dot{x} = F_{y'}$, y las dos primeras relaciones son solamente dos maneras de escribir $0 = 0$! Sin embargo, para más variables, el cálculo directo resulta mucho más trabajoso.

Caso 2.2: Cambio de varias variables independientes

Sea

$$J(u) = \int_{\Omega} \dots \int F(x_1, \dots, x_n, u, u_{x_1}, \dots, u_{x_n}) dx_1, \dots, dx_n$$

un funcional definido para $u(x_1, \dots, x_n)$ con (x_1, \dots, x_n) en Ω , un dominio acotado de \mathbb{R}^n . Supondremos que F y u son C^2 para poder usar el teorema de la divergencia. Si hacemos el cambio de variables

$$(x_1, \dots, x_n) \rightarrow (\xi_1, \dots, \xi_n), \Omega \rightarrow \Omega', X \rightarrow \xi$$

entonces:

$$\begin{aligned} J(u) &= \int_{\Omega'} \dots \int F(X(\xi), u, \sum_1^n u_{\xi_j} \partial \xi_j / \partial x_i) \left| \frac{D(x_1, \dots, x_n)}{D(\xi_1, \dots, \xi_n)} \right| d\xi \\ &= \int_{\Omega'} \dots \int H d\xi. \end{aligned}$$

Entonces

$$\begin{aligned} J(u+h) - J(u) &= \int_{\Omega} \dots \int (F_u h + \sum F_{u_{x_i}} h_{x_i}) dX + \dots \\ &= \int_{\Omega'} \dots \int (H_u h + \sum H_{u_{\xi_i}} h_{\xi_i}) d\xi + \dots \end{aligned}$$

Podemos identificar las partes lineales y usar el teorema de la divergencia en cada una, suponiendo $h = 0$ en $\partial\Omega$ (y por lo tanto en $\partial\Omega'$, (por ejemplo $h \in C_0^\infty(\Omega)$)).

$$\begin{aligned} &\int_{\Omega} \dots \int \left(F_u - \sum \frac{\partial}{\partial x_i} (F_{u_{x_i}}) \right) h dX = \int_{\Omega'} \dots \int \left(H_u - \sum \frac{\partial}{\partial \xi_i} (H_{u_{\xi_i}}) \right) h d\xi \\ &= \int_{\Omega'} \dots \int \left(F_u - \sum \frac{\partial}{\partial x_i} (F_{u_{x_i}}) \right) h \left| \frac{D(x_1, \dots, x_n)}{D(\xi_1, \dots, \xi_n)} \right| d\xi \end{aligned}$$

(cambiando de variables), para todo h en $C_0^\infty(\Omega')$. Por el lema III.8:

$$\left(F_u - \sum \frac{\partial}{\partial x_i} (F_{u_{x_i}}) \right) \left| \frac{D(x_1, \dots, x_n)}{D(\xi_1, \dots, \xi_n)} \right| = H_u - \sum \frac{\partial}{\partial \xi_i} (H_{u_{\xi_i}}).$$

Teorema 2.1 : *Bajo el cambio de variables $(x_1, \dots, x_n) \rightarrow (\xi_1, \dots, \xi_n)$ tenemos:*

$$Euler_u(F)(x_1, \dots, x_n) \frac{D(x_1, \dots, x_n)}{D(\xi_1, \dots, \xi_n)} = Euler_u \left(F \frac{D(x_1, \dots, x_n)}{D(\xi_1, \dots, \xi_n)} \right) (\xi_1, \dots, \xi_n).$$

En particular los puntos críticos no dependen de las variables.

Notemos que el Jacobiano cambia de signo en ambos lados y, por lo tanto, podemos quitar el valor absoluto.

Ejercicio 2.1 : Probar esta relación directamente con la regla de la cadena.

Ejemplo 2.1 : Superficie de area mínima. Problema de Plateau.

El area de una superficie parametrizada por (x_1, x_2) está dada por:

$$A(u_1, u_2, u_3) = \int_{\Omega} \int \|N(x_1, x_2)\| dx_1 dx_2$$

donde N es la normal a la superficie dada por

$$M(x_1, x_2) = (u_1(x_1, x_2), u_2(x_1, x_2), u_3(x_1, x_2))^T, \text{ en } \mathbb{R}^3.$$

$$N = M_{x_1} \times M_{x_2}.$$

Por lo tanto: $\|N\| = \|M_{x_1}\| \|M_{x_2}\| |\sin \theta|$, θ el ángulo entre los dos vectores tangentes M_{x_1} y M_{x_2} .

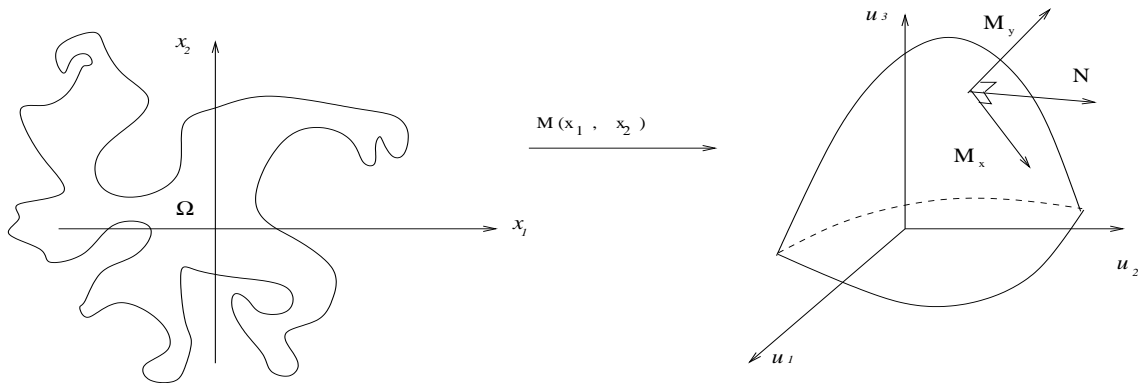


Fig. 2.1

Por lo tanto:

$$\begin{aligned}\|N\| &= \|M_{x_1}\| \|M_{x_2}\| (1 - \cos^2 \theta)^{\frac{1}{2}} \\ &= \left(\|M_{x_1}\|^2 \|M_{x_2}\|^2 - (M_{x_1} \cdot M_{x_2})^2 \right)^{\frac{1}{2}}.\end{aligned}$$

Si llamamos a $E = \|M_{x_1}\|^2$, $F = M_{x_1} \cdot M_{x_2}$, $G = \|M_{x_2}\|^2$ (ver el ejemplo IV.3.3), tenemos

$$\|N\| = (EG - F^2)^{\frac{1}{2}} = (\det A)^{\frac{1}{2}}$$

donde A es la matriz:

$$\begin{pmatrix} E & F \\ F & G \end{pmatrix}.$$

Notemos que $\|N\|$ es homogéneo de grado 1 en M_{x_1} y también en M_{x_2} , por lo tanto, tendremos las identidades de Euler:

$$\sum u_{ix_1} \|N\|_{u_{ix_1}} = \sum u_{ix_2} \|N\|_{u_{ix_2}} = \|N\|$$

(este es otro caso particular del teorema de Noether).

Las ecuaciones de Euler serán:

$$\text{Euler}_{u_i}(F)(x_1, x_2) = \frac{\partial}{\partial x_1} \left((\det A)^{\frac{1}{2}}_{u_{ix_1}} \right) + \frac{\partial}{\partial x_2} \left((\det A)^{\frac{1}{2}}_{u_{ix_2}} \right) = 0$$

estas ecuaciones se llaman de Laplace-Beltrami y tienen las propiedades siguientes:

Si hacemos el cambio de variables $(x_1, x_2) \rightarrow (\xi_1, \xi_2)$ la ecuación de Euler se transforma según el teorema 2.1. Ahora bien

$$\begin{aligned}M_{\xi_1} &= M_{x_1} \frac{\partial x_1}{\partial \xi_1} + M_{x_2} \frac{\partial x_2}{\partial \xi_1} \\ M_{\xi_2} &= M_{x_1} \frac{\partial x_1}{\partial \xi_2} + M_{x_2} \frac{\partial x_2}{\partial \xi_2}.\end{aligned}$$

Si \tilde{A} es la matriz correspondiente a (ξ_1, ξ_2) , dejamos al lector, como ejercicio, verificar que:

$$\tilde{A} = J^T A J, \quad \text{donde } J = \begin{pmatrix} \frac{\partial x_1}{\partial \xi_1} & \frac{\partial x_1}{\partial \xi_2} \\ \frac{\partial x_2}{\partial \xi_1} & \frac{\partial x_2}{\partial \xi_2} \end{pmatrix} \text{ es la matriz jacobiana.}$$

Por lo tanto

$$(\det \tilde{A})^{\frac{1}{2}} = (\det A)^{\frac{1}{2}} \left| \frac{D(x_1, x_2)}{D(\xi_1, \xi_2)} \right|$$

el teorema 2.1 dice solamente:

$$\text{Euler}_{u_i} \left((\det A)^{\frac{1}{2}} \right) (x_1, x_2) \left| \frac{D(x_1, x_2)}{D(\xi_1, \xi_2)} \right| = \text{Euler}_{u_i} \left((\det \tilde{A})^{\frac{1}{2}} \right) (\xi_1, \xi_2)$$

y para el problema de área mínima, que la ecuación de Laplace-Beltrami se satisface independientemente de las coordenadas escogidas.

En particular dado un punto M_0 , se pueden escoger las variables (ξ_1, ξ_2) tales que sean las coordenadas en la base dada por los vectores propios de la matriz A , entonces \tilde{A} será diagonal. Para que $\xi_i = \xi_i(E, F, G)$ pueda ser invertible, este procedimiento solo es válido en una vecindad de M_0 . Podemos además, mediante un cambio de escala (ya que E, F, G son derivadas con respecto a ξ_1, ξ_2) tener $\tilde{E} = \tilde{G}$ en esa vecindad. Entonces $(\det \tilde{A})^{\frac{1}{2}} = (\tilde{E}\tilde{G})^{\frac{1}{2}} = (\tilde{E} + \tilde{G})/2$.

En las coordenadas (ξ_1, ξ_2) , las ecuaciones son $\Delta u_i = 0$. Esto quiere decir que en la métrica dada por A , las funciones u_i son armónicas. A partir de aquí hay una muy bonita conexión entre superficies de área mínima y superficies de Riemann. El lector interesado podrá consultar algún libro avanzado de geometría diferencial.

Ejemplo 2.2 : El laplaciano en coordenadas polares.

Sabemos que la ecuación de Laplace Δu es la ecuación de Euler de $F = -(u_x^2 + u_y^2)/2$.

Si hacemos el cambio de variables a coordenadas polares:

$$\begin{aligned} x &= r \cos \theta, y = r \sin \theta, \text{ entonces : } \frac{D(r, y)}{D(r, \theta)} = r, \\ u_r &= u_x \cos \theta + u_y \sin \theta \\ u_\theta &= ru_x \sin \theta + ru_y \cos \theta \\ u_r^2 + u_\theta^2/r^2 &= u_x^2 + u_y^2. \end{aligned}$$

Por el teorema 2.1:

$$r\Delta u = \frac{\partial}{\partial r}(ru_r) + \frac{\partial}{\partial \theta}(ru_\theta/r^2).$$

es decir

$$\Delta u = u_{rr} + u_r/r + u_{\theta\theta}/r^2.$$

Ejercicio 2.2 : Encontrar Δu , en tres dimensiones, en coordenadas.

1. Cilíndricas
2. Esféricas

con este método y directamente.

Caso 2.3: Transformación del espacio, regla de la cadena

Si para el funcional

$$J(y) = \int_a^b F(x, y(x), y'(x)) dx$$

cambiamos, no solamente la variable independiente x , pero también la función y , es decir $(x, y) \rightarrow (u, v)$:

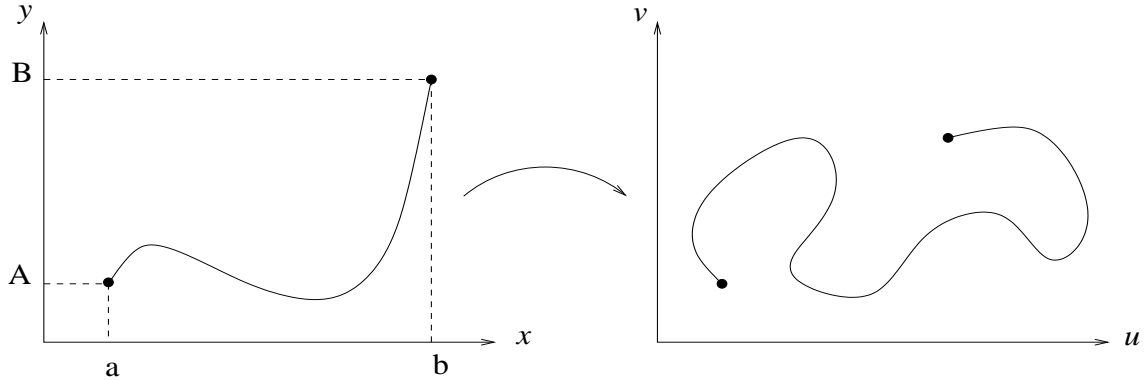


Fig. 2.2

Entonces una curva $y = y(x)$ da una curva en el espacio (u, v) pero no necesariamente de la forma $v = v(u)$, como es el caso cuando solo se transforma a x .

Para evitar el papel especial de la variable x como parámetro, pasaremos primero a un parámetro independiente t , por ejemplo en $[0, 1]$. La curva $(x(t), y(t))$ dará entonces una curva $(u(t), v(t)) \equiv (u(x(t), y(t)), v(x(t), y(t)))$ e inversamente.

$$\begin{aligned} J(x, y) &= \int_0^1 F(x(t), y(t), \dot{y}/\dot{x}) \dot{x} dt \\ &= \int_0^1 F(x(u(t), v(t)), y(u(t), v(t)), (y_u \dot{u} + y_v \dot{v}) / (x_u \dot{u} + x_v \dot{v})) (x_u \dot{u} + x_v \dot{v}) dt \\ &\equiv \int_0^1 G(u, v, \dot{u}, \dot{v}) dt = J_1(u, v). \end{aligned}$$

Es decir tenemos la composición de operadores:

$$\begin{aligned} C^1[0, 1] \times C^1[0, 1] &\xrightarrow{I} C^1[0, 1] \times C^1[0, 1] \xrightarrow{J} \mathbb{R} \\ (u(t), v(t)) &\longrightarrow (x(u(t), v(t)), y(u(t), v(t))) \longrightarrow J(x, y) \end{aligned}$$

donde suponemos que la transformación $(u, v) \longrightarrow (x, y)$ es C^1 .

Tenemos pues:

$$J_1(u, v) = J(I(u, v)) = J \circ I(u, v)$$

y nos interesa su derivada de Fréchet.

Teorema 2.2 : (Regla de la cadena). Sean X, Y, Z tres espacios normados, I, J dos operadores (no lineales en general) con derivadas de Fréchet

$$X \xrightarrow{I} Y \xrightarrow{J} Z.$$

Entonces $J \circ I$ tiene una derivada de Fréchet con

$$D(J \circ I)(\varphi)h = DJ(I(\varphi))DI(\varphi)h,$$

para toda h en X , donde las derivadas de Fréchet son operadores lineales

Prueba: Recordaremos que

$$J(y + k) - J(y) = DJ(y)k + o(\|k\|_Y)$$

donde $DJ(y) : Y \rightarrow Z$ es un operador lineal continuo. Además $I(\varphi + h) - I(\varphi) = DI(\varphi)h + o(\|h\|_X)$ donde $DI(\varphi) : X \rightarrow Y$ es un operador lineal continuo.

Entonces

$$\begin{aligned} & J(I(\varphi + h)) - J(I(\varphi)) \\ &= DJ(I(\varphi))(I(\varphi + h) - I(\varphi)) + o(\|I(\varphi + h) - I(\varphi)\|_Y) \\ &= DJ(I(\varphi))DI(\varphi)h + DJ(I(\varphi))o(\|h\|_X) + \\ &+ o(\|DI(\varphi)h\|_Y + \|o(\|h\|_X)\|_Y). \end{aligned}$$

Como los operadores y residuos son continuos:

$$J(I(\varphi + h)) - J(I(\varphi)) = DJ(I(\varphi))DI(\varphi)h + o(\|h\|_X).$$

Por la unicidad de la derivada de Fréchet (Proposición II.3.1), el resultado queda demostrado.

Q.E.D.

Aquí

$$\begin{aligned}
DJ \circ I(u, v)(h, k) &= \int_0^1 ((G_u - (G_{\dot{u}})')h + G_v - (G_{\dot{v}})')k) dt \\
&= \int_0^1 \begin{pmatrix} \text{Euler}_u(G)(t) \\ \text{Euler}_v(G)(t) \end{pmatrix}^T \begin{pmatrix} h \\ k \end{pmatrix} dt.
\end{aligned}$$

(Supusimos que G, u, v son C^2 para poder integrar por partes para h y k en $C_0^\infty[0, 1]$. Esto es inecesario. Pedimos al lector que use el mismo argumento que en el Caso 2.1 para llegar al resultado).

Del mismo modo

$$DJ(x, y)(\xi, \eta) = \int_0^1 \begin{pmatrix} \text{Euler}_x(F\dot{x})(t) \\ \text{Euler}_y(F\dot{x})(t) \end{pmatrix}^T \begin{pmatrix} \xi \\ \eta \end{pmatrix} dt.$$

$$\text{Adem\'as } DI(u, v)(h, k) = \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

ya que $x(u+h, v+k) - x(u, v) = x_u h + x_v k + o(|h| + |k|)$ y una expresi3n an3loga para la y .

Tenemos entonces:

Teorema 2.3 : *Con la transformaci3n $(x, y) \rightarrow (u, v)$, las ecuaciones de Euler se transforman:*

$$\begin{aligned}
&\begin{pmatrix} \text{Euler}_u[F(x(u, v), y(u, v), (y_u \dot{u} + y_v \dot{v})/(x_u \dot{u} + x_v \dot{v}))] (x_u \dot{u} + x_v \dot{v})(t) \\ \text{Euler}_v[F(x(u, v), y(u, v), (y_u \dot{u} + y_v \dot{v})/(x_u \dot{u} + x_v \dot{v}))] (x_u \dot{u} + x_v \dot{v})(t) \end{pmatrix} = \\
&\begin{pmatrix} x_u & y_u \\ x_v & y_v \end{pmatrix} \begin{pmatrix} \text{Euler}_x[F(x, y, \dot{y}/\dot{x})\dot{x}](t) \\ \text{Euler}_y[F(x, y, \dot{y}/\dot{x})\dot{x}](t) \end{pmatrix}
\end{aligned}$$

Recordemos que (Nota 1.1):

$$\text{Euler}_x(F\dot{x})(t) = \dot{x} \quad (\text{Segunda ecuaci3n de Euler}),$$

$$\text{Euler}_y(F\dot{x})(t) = \text{Euler}_y(F)(x)(\dot{x}), \quad (\text{Teorema 2.1}).$$

En particular si el funcional $J(y)$ tiene un punto cr3tico (por lo tanto las dos ecuaciones de Euler son 0), bajo la transformaci3n, el funcional J_1 tendr3 un punto cr3tico e inversamente.

Notemos que si ponemos $H = F\dot{x}$, entonces para un cambio de variables en las funciones $(x(t), y(t))$ tenemos la forma m3s compacta:

$$\begin{pmatrix} \text{Euler}_u(H)(t) \\ \text{Euler}_v(H)(t) \end{pmatrix} = \begin{pmatrix} x_u & y_u \\ x_v & y_v \end{pmatrix} \begin{pmatrix} \text{Euler}_x(H)(t) \\ \text{Euler}_y(H)(t) \end{pmatrix}$$

reflejando solamente el hecho que estamos usando la regla de la cadena para calcular la derivada de Fr3chet de $J_1(u, v) = J \circ I(x, y)$.

Ejercicio 2.3 : Encontrar la transformación de las ecuaciones de Euler, cuando se cambia $(x, y, z) \rightarrow (u, v, w)$, para el funcional: $J(y, z) = \int_a^b F(x, y(x), y'(x), z(x), z'(x)) dx$.

Ejercicio 2.4 : Encontrar la transformación de las ecuaciones de Euler, cuando se cambia $(x, y, z) \rightarrow (u, v, w)$ para el funcional:

$$J(z) = \int_{\Omega} \int F(x, y, z(x, y), z_x, z_y) dx dy.$$

Hacer primero una transformación $(x, y) \rightarrow (\xi, \eta)$, para obtener un funcional

$$J(x, y, z) = \int_{\Omega'} \int H(x, y, z, x_{\xi}, y_{\xi}, z_{\xi}, x_{\eta}, y_{\eta}, z_{\eta}) d\xi d\eta.$$

Usar después la regla de la cadena.

Ejemplo 2.3 : Condiciones de transversalidad.

Veremos como obtener las condiciones de transversalidad cuando un extremo de la curva se puede mover sobre una curva dada $y = \varphi(x)$, como en el ejemplo IV.8.3.

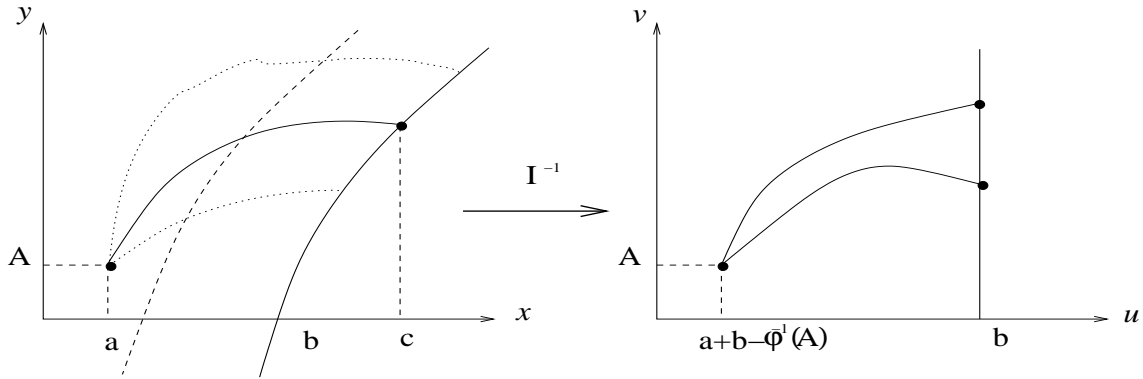


Fig. 2.3

Si φ es monótona y sobre, para cada (x, y) hay una curva $y = \varphi(x + b - u)$ que pasa por ese punto.

Sea entonces la transformación:

$$\left\{ \begin{array}{l} u = x + b - \varphi^{-1}(y) \\ v = y \end{array} \right\} \Longleftrightarrow \left\{ \begin{array}{l} x = u - b + \varphi^{-1}(v) \\ y = v \end{array} \right\}.$$

La curva $y = \varphi(x)$ corresponde a la recta $u = b$, y los puntos críticos están ligados por la fórmula del Teorema 2.3.

Ahora bien para $J_1(u, v) = J \circ I(u, v)$, los incrementos (h, k) son tales que $h(0) = k(0) = 0$ (ligados en (a, A)), $h(1) = 0$ (ligado en b), $k(1)$ es libre.

De la regla de la cadena obtenemos:

$$\begin{aligned} \int_0^1 \begin{pmatrix} \text{Euler}_u H(t) \\ \text{Euler}_v H(t) \end{pmatrix}^T \begin{pmatrix} h \\ k \end{pmatrix} + \begin{pmatrix} H_{\dot{u}} \\ H_{\dot{v}} \end{pmatrix}^T \begin{pmatrix} h \\ k \end{pmatrix} \Big|_0^1 = \\ \int_0^1 \begin{pmatrix} D \left(\text{Euler}_x H(t) \right) \\ D \left(\text{Euler}_y H(t) \right) \end{pmatrix}^T \begin{pmatrix} h \\ k \end{pmatrix} + \begin{pmatrix} D \left(H_{\dot{x}} \right) \\ D \left(H_{\dot{y}} \right) \end{pmatrix}^T \begin{pmatrix} h \\ k \end{pmatrix} \Big|_0^1 \end{aligned}$$

donde

$$D = \begin{pmatrix} x_u & y_u \\ x_v & y_v \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \varphi'^{-1} & 1 \end{pmatrix}.$$

Para (h, k) en $C_0^\infty[0, 1]$ obtenemos de nuevo la invariancia de las ecuaciones de Euler, por lo tanto:

$$\begin{pmatrix} H_{\dot{u}} \\ H_{\dot{v}} \end{pmatrix}^T \begin{pmatrix} h \\ k \end{pmatrix} \Big|_0^1 = \begin{pmatrix} D \left(H_{\dot{x}} \right) \\ D \left(H_{\dot{y}} \right) \end{pmatrix}^T \begin{pmatrix} h \\ k \end{pmatrix} \Big|_0^1.$$

Como solo $k(1)$ puede ser diferente de 0:

$$H_{\dot{v}} \Big|_{t=1} = \varphi'^{-1} H_{\dot{x}} + H_{\dot{y}} \Big|_{t=1} = 0,$$

ya que, en un punto crítico, la derivada de Fréchet debe ser 0.

Como

$$H = F(x, y, \dot{y}/\dot{x})\dot{x}, \quad H_{\dot{x}} = F - F_{y'}\dot{y}/\dot{x}, \quad H_{\dot{y}} = F_{y'},$$

se obtiene

$$F - (y' - \varphi')F_{y'}|_c = 0.$$

Ejercicio 2.5 : Tratar al caso de la difracción por una curva (ver el ejercicio IV.8.6) de esta manera.

5.3 Coordenadas Lagrangianas

Los principios variacionales en la Mecánica han permitido el desarrollo de toda esta parte de la Física de una manera mucho más eficiente que si se hubiera tratado de mantener únicamente un punto de vista microscópico al igualar fuerzas y aceleraciones, según la ley de Newton, en particular por la dificultad de conceptualizar fácilmente esas nociones para

sistemas complicados. Además hubo siempre la idea filosófica que Dios, o la Naturaleza, trabaja eficientemente y que, por lo tanto, algo debe ser minimizado para que cada cosa en el universo funcione (ver por ejemplo el principio de Fermat). Los primeros éxitos del Cálculo de Variaciones, llevaron a los Bernoulli, Euler y Lagrange (entre otros) a buscar funcionales cuya minimización llevara a las leyes (y a los resultados experimentales) ya conocidos. Este principio de mínima acción de Lagrange fue transformado por Hamilton (hacia 1835) en principio de acción estacionaria (es decir buscar puntos críticos) y pulido por Jacobi en 1848.

Si, por ejemplo, en el caso de la ley de Newton para N partículas con coordenadas $X_i(t)$, en \mathbb{R}^3 para cada i , buscamos una función $L(t, X_i, \dot{X}_i)$ tal que las ecuaciones de Euler del funcional

$$I(X_1, \dots, X_N) = \int_{t_0}^{t_1} L(t, X_1, \dots, \dot{X}_N) dt$$

sean las ecuaciones de Newton:

$$m_i \ddot{X}_i = F_i(t, X_1, \dots, X_N)$$

debemos tener:

$$(L_{\dot{X}_i})' = L_{X_i} \quad (\text{son tres ecuaciones para cada } i).$$

Si identificamos cada término (Nota: esto no es una manera de resolver ecuaciones! sólo una justificación):

$$m_i \dot{X}_i = L_{\dot{X}_i} \quad , \quad F_i = L_{X_i}$$

tendremos

$$L = \sum m_i |\dot{X}_i|^2 / 2 - U(t, X_1, \dots, X_n)$$

(integrando la primera ecuación),

$$F_i = -U_{X_i}.$$

Hemos llegado entonces a definir la energía cinética del sistema:

$T = \sum m_i |\dot{X}_i|^2 / 2$ y a la energía potencial U tal que $F_i = -U_{X_i}$; las ecuaciones de Newton son las ecuaciones de Euler para el funcional

$$\int_{t_0}^{t_1} (T - U) dt.$$

Ahora bien, si esto funciona para un sistema cerrado de N partículas, en general hay muchas restricciones físicas o geométricas que se deben satisfacer. Por ejemplo, en un sólido, las partículas son ligadas las unas a las otras. Para estos casos podemos pensar en el movimiento de los centros de gravedad, o más generalmente **suponer** que el sistema está descrito por

n variables linealmente independientes q_1, \dots, q_n , llamadas las **coordenadas Lagrangianas**. Evidentemente debe de haber una relación entre la posición de las partículas X y las coordenadas Lagrangianas Q , es decir $X = G(Q)$.

Cuando las partículas se mueven debemos tener:

$$\dot{X} = G_Q \dot{Q}$$

G_Q la matriz jacobiana de la transformación, y la energía cinética se vuelve:

$$\sum m_i |\dot{X}_i|^2 / 2 = \sum a_{ij}(Q) \dot{q}_i \dot{q}_j / 2$$

donde las $a_{ij}(Q)$ provienen de G_Q .

$$L = T - U = \sum a_{ij}(Q) \dot{q}_i \dot{q}_j / 2 - U(Q).$$

Por el teorema 2.3 (sólo hay transformación de X , no en t) los puntos críticos de L son independientes de la transformación, tenemos entonces el **principio de Hamilton**:

Las trayectorias de las partículas en las coordenadas Lagrangianas son los puntos críticos del funcional de acción:

$$I(q_1, \dots, q_n) = \int_{t_0}^{t_1} L(t, Q, \dot{Q}) dt$$

con $q_i(t_0), q_i(t_1)$ dados.

Las ecuaciones del movimiento son entonces:

$$\frac{d}{dt}(L_{\dot{q}_i}) = L_{q_i}, \quad i = 1, \dots, n, \quad q_i(t_0), q_i(t_1) \text{ dados.}$$

Notemos que si T y U no dependen de t y U no depende de \dot{q}_i (sistema conservativo) entonces la energía total $T + U$ se conserva. (ver el ejemplo IV.3.1.).

5.4 Sistema canónico y transformada de Legendre

Consideremos ahora el funcional

$$I(q_1, \dots, q_n) = \int_{t_0}^{t_1} L(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) dt$$

donde L es cualquier función C^1 (Lagrangiana en Mecánica) con condiciones $q_i(t_0) = A_i, q_i(t_1) = B_i$.

Las ecuaciones de Euler son entonces:

$$(L_{\dot{q}_i})' - L_{q_i} = 0, \quad i = 1, \dots, n$$

es decir, un sistema de n ecuaciones de segundo orden, **sistema Hamiltoniano**, en general muy difícil de resolver.

Si ponemos $z_i = \dot{q}_i$ tendremos:

$$\begin{aligned}\dot{q}_i &= z_i & q_i(t_0) = A_i, q_i(t_1) = B_i, \\ (L_{z_i})' &= L_{q_i}\end{aligned}$$

un sistema de $2n$ ecuaciones de primer orden.

Hay, sin embargo, una manera más simétrica de llegar a un sistema de $2n$ ecuaciones de primer orden: Sea

$$p_i \equiv L_{z_i}(t, q_1, \dots, q_n, z_1, \dots, z_n), \quad i = 1, \dots, n,$$

(aquí $q_1, \dots, q_n, z_1, \dots, z_n$ son variables, no funciones, independientes).

Podemos invertir este sistema, es decir expresar z_1, \dots, z_n en función de $(t, q_1, \dots, q_n, p_1, \dots, p_n)$ si usamos el teorema de la función implícita. Para esto es necesario que:

$$\det(L_{z_i z_j})_{ij} \neq 0.$$

Mantendremos esta hipótesis en todo el resto de estas notas.

Denotaremos por

$$\tilde{L}(t, q_1, \dots, q_n, p_1, \dots, p_n) =$$

$$L(t, q_1, \dots, q_n, z_1(t, q_1, \dots, q_n, p_1, \dots, p_n), \dots, z_n(t, q_1, \dots, q_n, p_1, \dots, p_n))$$

y definiremos al **Hamiltoniano** H por:

$$H(t, q_1, \dots, q_n, p_1, \dots, p_n) =$$

$$\sum_1^n p_i z_i(t, q_1, \dots, q_n, p_1, \dots, p_n) - \tilde{L}(t, q_1, \dots, q_n, p_1, \dots, p_n).$$

(Daremos más adelante su interpretación física).

Tenemos entonces:

$$\begin{aligned}H_t &= \sum_1^n p_i z_{it} - \tilde{L}_t = \sum_1^n p_i z_{it} - \sum_1^n L_{z_i} z_{it} - L_t \\ H_t &= -L_t, \quad \text{por la definición de } p_i. \\ H_{q_j} &= \sum_1^n p_i z_{iq_j} - \tilde{L}_{q_j} = \sum_1^n p_i z_{iq_j} - \sum_1^n L_{z_i} z_{iq_j} - L_{q_j}\end{aligned}$$

$$\begin{aligned} H_{q_j} &= -L_{q_j} \quad \text{por la definición de } p_i \\ H_{p_j} &= z_j + \sum_1^n p_i z_{ip_j} - \sum_1^n L_{z_i} z_{ip_j} = z_j \end{aligned}$$

Además en un punto crítico:

$$L_{q_i} = (L_{z_i})' = \dot{p}_i, \quad z_i = \dot{q}_i$$

Tenemos entonces:

Teorema 4.1 : Si la matriz $L_{z_i z_j}$ es invertible, definiendo los **momentos** $p_i = L_{z_i}$ y el **Hamiltoniano** $H = \sum_1^n p_i z_i - \tilde{L}$, las ecuaciones de Euler se transforman en el sistema canónico:

$$\dot{q}_i = H_{p_i}$$

$$\dot{p}_i = -H_{q_i}$$

$$q_i(t_0) = A_i, q_i(t_1) = B_i,$$

$2n$ ecuaciones de primer orden.

$$\text{Además } H_t = -L_t, \quad H_{q_i} = -L_{q_i}, \quad H_{p_i} = z_i.$$

Notemos que no solamente si (q_i, z_i) son soluciones de las ecuaciones $(L_{z_i})' = L_{q_i}$, $\dot{q}_i = z_i$ entonces p_i y q_i son soluciones de las ecuaciones canónicas, sino también al revés.

Notemos que, según el ejercicio IV.3.2, los puntos críticos son siempre C^2 .

Nota 4.1: Energía del sistema

Si definimos a la energía del sistema como

$$E(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = \sum_1^n L_{z_i} \dot{q}_i - L$$

(si $L = \sum a_{ij}(Q) \dot{q}_i \dot{q}_j / 2 - U(t, q_1, \dots, q_n)$ entonces

$$E = \sum a_{ij}(Q) \dot{q}_i \dot{q}_j - L = 2T - (T - U) = T + U,$$

entonces por la segunda ecuación de Euler tenemos (ejercicio IV.3.1):

$$(L - \sum \dot{q}_i L_{\dot{q}_i})' = L_t,$$

por lo tanto $\frac{dE}{dt}(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = -L_t = H_t$. Además, ya que $L_{z_i} = p_i$, $\dot{q}_i = z_i$:

$$E(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) = H(t, q_1, \dots, q_n, p_1, \dots, p_n).$$

En particular si L no depende explícitamente de t : $L_t = -H_t = 0 : E = H$ se conserva sobre las trayectorias.

Más generalmente, sea $\varphi(t, q_1, \dots, q_n, p_1, \dots, p_n)$ cualquier función. Si $q_i(t)$, $p_i(t)$ son soluciones de las ecuaciones canónicas, tendremos:

$$\frac{d\varphi}{dt}(t, q_i(t), p_i(t)) = \varphi_t + \sum \varphi_{q_i} H_{p_i} - \sum \varphi_{p_i} H_{q_i} = \varphi_t + [\varphi, H],$$

donde $[\varphi, H]$ denota al “bracket” de Poisson.

Tendremos entonces:

Proposición 4.1. Condición de Frobenius. φ es una integral primera (es decir $\frac{d\varphi}{dt} = 0$), con φ independiente de t , si y solo si $[\varphi, H] = 0$. En particular esto pasa para $\varphi = H$.

Nota 4.2: Consideremos la **acción generalizada**

$$\begin{aligned} A(q_1, \dots, q_n, p_1, \dots, p_n) &= \int_{t_0}^{t_1} \tilde{L}(t, q_i, p_i) dt \\ &= \int_{t_0}^{t_1} \left(\sum_1^n p_i \dot{q}_i - H(t, q_1, \dots, q_n, p_1, \dots, p_n) \right) dt, \end{aligned}$$

con $q_i(t_0) = A_i$, $q_i(t_1) = B_i$.

Las ecuaciones de Euler para A son:

$$\dot{p}_i + H_{q_i} = 0, \quad \dot{q}_i - H_{p_i} = 0,$$

es decir las ecuaciones canónicas. Notemos que no hay condiciones para p_i en t_0 y en t_1 (las condiciones naturales $\tilde{L}_{\dot{p}_i} = 0$ siempre se cumplen). Además si $H_t = 0$, entonces $H = E =$ constante. Podemos considerar solamente (cambiando a un parámetro τ común a todas las curvas):

$$\int_0^1 \sum p_i q'_i d\tau \text{ condicionado por } H(P, Q) = E.$$

Es decir un funcional definido sobre una superficie.

Ejemplo 4.1 : Sistemas de N partículas.

Retomemos el sistema de la sección anterior con:

$$\begin{aligned} L = T - U &= \sum a_{ij}(Q) \dot{q}_i \dot{q}_j / 2 - U(t, Q) \\ &= \sum m_i |\dot{X}_i|^2 / 2 - U(t, X) \end{aligned}$$

en coordenadas cartesianas. Entonces la matriz $L_{z_i z_j}$ es positiva definida, ya que lo es en las coordenadas cartesianas.

En este caso

$$H = E = \sum_i p_i z_i - \tilde{L},$$

con $z_i = \dot{q}_i$, $p_i = L_{z_i} = \sum_j a_{ij}(Q) z_j$.

Aquí es fácil expresar z_j en función de Q y p_i : $z_i = \sum a^{ij}(Q) p_j$, donde $(a^{ij})_{ij} = (a_{ij})_{ij}^{-1}$, (en coordenadas cartesianas $P_i = m_i \dot{X}_i$, es el momento), por lo tanto

$$H = 2T - (T - U) = T + U.$$

Si ponemos $A = (a_{ij})$, entonces $T = Z^T A Z / 2 = P^T A^{-1} P / 2$ ya que $Z = A^{-1} P$ y A es simétrica. Las ecuaciones canónicas son:

$$\begin{aligned} \dot{q}_i &= \sum_j a^{ij}(Q) p_j \\ \dot{p}_i &= -U_{q_i} + P^T A_{q_i}^{-1} P / 2. \end{aligned}$$

Nota 4.3: Interpretación geométrica: Transformada de Legendre

Fijando t, q_1, \dots, q_n , sea $f(Z) = L(t, q_1, \dots, q_n, z_1, \dots, z_n)$.

La superficie $U = f(Z)$ tiene un plano tangente en $(Z_0, U_0 = f(Z_0))$ con ecuación

$$U - f(Z_0) = \nabla f(Z_0) \cdot (Z - Z_0).$$

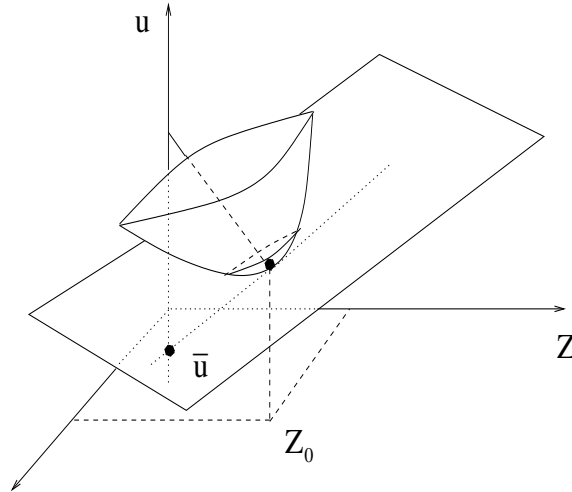


Fig. 4.1

Este plano tangente intersecta el eje U (en $Z = 0$) en el punto

$$\bar{U} = -\nabla f(Z_0) \cdot Z_0 + f(Z_0).$$

Sea $p_i = f_{z_i} = L_{z_i}$, $P = \nabla f(Z_0)$

$$H = -U + \sum p_i z_i = -L + P \cdot Z_0 = -\bar{U}.$$

Definiremos a la **transformada de Legendre**:

$$(Z_0, U) \xrightarrow{\mathcal{L}} (P, H),$$

como una transformación de \mathbb{R}^{n+1} en \mathbb{R}^{n+1} , tal que si $U = f(Z)$,

$$(Z, f(Z)) \xrightarrow{\mathcal{L}} (\nabla f(Z), H),$$

en donde

$$H = \nabla f(Z) \cdot Z - f(Z) = - \text{intersección del plano tangente con el eje } U.$$

Como la transformación $P = \nabla f(Z)$ es invertible si la matriz jacobiana $(f_{z_i z_j})_{ij}$ es invertible (teorema de la función inversa), escribiendo $Z = g(P)$ tenemos:

$$H = P \cdot Z - f(Z) = P \cdot g(P) - f(g(P)) \equiv H(P)$$

es decir, bajo la transformación de Legendre, la superficie $U = f(Z)$, se transforma en la superficie $H = H(P)$.

Si repetimos la transformación:

$$(Z, U) \xrightarrow{\mathcal{L}} (P, H) \xrightarrow{\mathcal{L}} (X, K)$$

con $X = \nabla H(P)$, $K = \nabla H(P) \cdot P - H(P)$, tenemos $x_i = \frac{\partial H}{\partial p_i} = z_i + \sum p_j \frac{\partial z_j}{\partial p_i} - \sum \frac{\partial f}{\partial z_j} \frac{\partial z_j}{\partial p_i}$, pero, como $p_j = \frac{\partial f}{\partial z_j}$, tenemos $x_i = z_i$.

$$K = \sum x_i p_i - H = \sum z_i p_i - H = U$$

y finalmente $(H_{p_i p_j}) = \left(\frac{\partial z_i}{\partial p_j}\right) = (U_{z_i z_j})^{-1}$, es decir P puede escribirse en función de X .

Tenemos entonces:

Proposición 4.2: *La transformada de Legendre es una involución, es decir $\mathcal{L}^2 = Id$.*

Notemos que si desarrollamos a segundo orden la función

$$f(Z) = f(Z_0) + \nabla f(Z_0) \cdot (Z - Z_0) + (Z - Z_0)^T \text{Hess}(f(Z_0))(Z - Z_0)/2 + \dots$$

si en un punto crítico de f ($\nabla f(Z_0) = 0$), el Hessiano es positivo definido, tendremos un mínimo local de $f(Z)$.

Como $H(P_0) = -f(Z_0)$, ya que $\nabla f(Z_0) = 0$, y $\nabla H(P_0) = Z_0$, por la involución, y $\text{Hess}(H(P_0)) = (\text{Hess } f(Z_0))^{-1}$, entonces si $Z_0 = 0$, la superficie $H = H(P)$ tendrá también un mínimo. Este tipo de argumentos es la base de los principios variacionales duales.

Nota 4.4: Ya hemos visto directamente que las ecuaciones de Euler para $\int_{t_0}^{t_1} L(t, q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n) dt$ tienen las mismas soluciones que las ecuaciones canónicas. Veremos de nuevo este hecho, sin usar cálculos, a partir de la transformación de Legendre. En la Nota 4.2 vimos que las ecuaciones canónicas son las ecuaciones de Euler de la acción generalizada:

$$A(q_1, \dots, q_n, p_1, \dots, p_n) = \int_{t_0}^{t_1} (P \cdot \dot{Q} - H(t, P, Q)) dt.$$

Si consideramos la transformación (de las variables dependientes)

$$(Q, P) \xrightarrow{\text{Id} \times \mathcal{L}^{-1}} (Q, Z),$$

por el teorema 2.3, debemos tener **los mismos puntos críticos**, cuando expresamos el integrando en (Q, Z) , con respecto a Q, Z :

Como $H = P \cdot Z - L = H(t, Q, P)$, y $P = \nabla_Z L$, tenemos

$$A(Q, P) = \int_{t_0}^{t_1} (\nabla_Z L \cdot (\dot{Q} - Z) + L(t, Q, Z)) dt = \tilde{A}(Q, Z)$$

y las ecuaciones de Euler para \tilde{A} son:

$$\begin{aligned} (L_{z_i})' &= \left(\sum_j L_{q_i z_j} (\dot{q}_j - z_j) + L_{q_i} \right) = 0 \\ &= \left(\sum_j L_{z_i z_j} (\dot{q}_j - z_j) - L_{z_i} + L_{z_i} \right) = 0. \end{aligned}$$

Como $L_{z_i z_j}$ es invertible, tenemos:

$$\dot{q}_j = z_j, \quad (L_{z_i})' - L_{q_i} = 0$$

las cuales son las ecuaciones de Euler originales.

Nota 4.5: Ecuaciones en forma paramétrica

Consideremos de nuevo la acción generalizada:

$$A(Q, P) = \int_{t_0}^{t_1} (P \cdot \dot{Q} - H(t, Q, P)) dt$$

con $Q(t_0) = A$, $Q(t_1) = B$.

Podemos cambiar de variable, de t a τ , de tal forma que para $\tau(t)$, $\tau(t_0) = 0$, $\tau(t_1) = 1$.

$$\begin{aligned}
\text{Como } dt &= t' d\tau, \\
\dot{Q} &= Q'/t' \\
A(Q, P) &= \int_0^1 (P \cdot Q' - H(t, Q, P)t') d\tau.
\end{aligned}$$

Si ponemos $q_0 = t$ y p_0 , el momento conjugado de q_0 , como $p_0 = -H(t, Q, P)$ tendremos:

$$A(\tilde{Q}, \tilde{P}) = \int_0^1 \tilde{P}(\tau) \cdot \dot{\tilde{Q}}(\tau) d\tau$$

donde

$$\begin{aligned}
\tilde{P} &= (p_0, p_1, \dots, p_n) && \text{con } p_0 = -H(\tilde{Q}, P) \\
\tilde{Q} &= (q_0, q_1, \dots, q_n) && \text{con } q_0 = t.
\end{aligned}$$

Sabemos que para $A(\tilde{Q}, P)$ (sin p_0) las ecuaciones de Euler son las ecuaciones canónicas (Nota 4.2 y Caso 2.1). Además si trabajamos con $A(\tilde{Q}, \tilde{P})$ (con p_0) condicionado por

$$G(\tilde{Q}, \tilde{P}) = H(\tilde{Q}, P) + p_0 = 0 \quad (\text{condición holonómica}),$$

las ecuaciones de Euler para este funcional condicionado son (Teorema IV.9.2):

$$\begin{aligned}
\nu \dot{q}_i + \lambda(\tau) H_{p_i} &= 0 && i = 1, \dots, n \\
\nu \dot{p}_i - \lambda(\tau) H_{q_i} &= 0 && i = 1, \dots, n \\
\nu \dot{q}_0 + \lambda(\tau) &= 0 \\
\nu \dot{p}_0 - \lambda(\tau) H_{q_0} &= 0
\end{aligned}$$

De la tercera ecuación vemos que $\nu \neq 0$ y podemos suponer $\nu = 1$. Si ponemos:

$$t = - \int_0^\tau \lambda(\xi) d\xi$$

tendremos: $\frac{dq_i}{dt} = H_{p_i}$, $\frac{dp_i}{dt} = -H_{q_i}$, $q_0 = t$, $\frac{dp_0}{dt} = -H_{q_0}$, es decir las ecuaciones canónicas.

Nota 4.6: Transformaciones canónicas

Una de las ventajas de pasar a la formulación de acción generalizada, con las ecuaciones canónicas, es que tenemos una integral primera del movimiento si H no depende explícitamente del tiempo: es decir, la energía (H) se conserva sobre las soluciones de las ecuaciones.

Ahora bien, como estas ecuaciones son difíciles de resolver, uno trata de encontrar más integrales primeras, como en la proposición 4.1, a partir de transformaciones de las variables

dependientes (Q, P) (como las ecuaciones de Euler preservan los puntos críticos, tendremos integrales primeras a partir de las mismas ecuaciones!).

Sea

$$A(Q, P) = \int_{t_0}^{t_1} (P \cdot \dot{Q} - H(t, Q, P)) dt.$$

Si hacemos la transformación:

$$(Q, P) \longrightarrow (S, R),$$

entonces sea $\tilde{H}(t, S, R) = H(t, Q(S, R), P(S, R))$

$$\begin{aligned} P \cdot \dot{Q} &= P(S, R) \cdot \left(\sum_i \frac{\partial Q}{\partial r_i} \dot{r}_i + \frac{\partial Q}{\partial s_i} \dot{s}_i \right) \\ \text{y } A(Q, P) &= \tilde{A}(S, R). \end{aligned}$$

Definición 4.1. Diremos que la transformación

$$(Q, P) \longrightarrow (S, R)$$

es **canónica**, si las ecuaciones de Euler para $\tilde{A}(S, R)$ son canónicas, es decir:

$$\dot{s}_i = \tilde{H}_{r_i}, \quad \dot{r}_i = -\tilde{H}_{s_i}.$$

Podríamos por ejemplo pedir que

$$P \cdot \dot{Q} - H = R \cdot \dot{S} - \tilde{H}$$

pero esto es demasiado fuerte, ya que es suficiente que:

$$P \cdot \dot{Q} - H = R \cdot \dot{S} - \tilde{H} + \frac{d}{dt} G(t, Q(t), P(t), S(t), R(t))$$

para cualquier función (diferenciable) $G(t, Q, P, S, R)$, ya que al integrar esa diferencial exacta, tendremos $G(t, Q, P, S, R)|_{t_0}^{t_1}$ lo cual será una constante si P, Q, R, S son dados en los extremos.

Evidentemente, dada una transformación $(Q, P) \longrightarrow (S, R)$ entonces todos los términos son conocidos y tendremos una transformación canónica por casualidad. La idea es pues, a partir de una parte de la transformación, por ejemplo S o R , encontrar una función G , de solamente esas variables, que permita por una parte completar la transformación y, por otra parte, garantizar que sea canónica.

Por ejemplo, con Q y S , tomemos como función G , **función generadora de la transformación canónica**, una función $G(t, Q, S)$ tal que:

$$\frac{\partial G}{\partial t} = \tilde{H} - H, \quad p_i = \frac{\partial G}{\partial q_i}$$

(esto define parte de G si los otros términos son conocidos)

$$r_i = -\frac{\partial G}{\partial s_i}$$

(esto define el resto de la transformación).

Entonces

$$\frac{dG}{dt} = \frac{\partial G}{\partial t} + \sum \frac{\partial G}{\partial q_i} \dot{q}_i + \sum \frac{\partial G}{\partial s_i} \dot{s}_i = (\tilde{H} - H) + P \cdot \dot{Q} - R \cdot \dot{S}.$$

En particular si se escoge G independiente de t , para obtener $\tilde{H}(S, R)$ bastará sustituir Q, P en $H(Q, P)$.

Si por ejemplo, por razones físicas, es más conveniente pensar en una función $\Phi(t, Q, R)$, sea entonces

$$G(t, Q, R, S) = \Phi(t, Q, R) - R \cdot S,$$

$$\begin{aligned} \frac{d}{dt}(G) &= \Phi_t + \sum (\Phi_{q_i} \dot{q}_i + \Phi_{r_i} \dot{r}_i - r_i \dot{s}_i - s_i \dot{r}_i) \\ &= \tilde{H} - H + P \cdot \dot{Q} - R \cdot \dot{S} \end{aligned}$$

si tomamos $\Phi_t = \tilde{H} - H$, $p_i = \Phi_{q_i}$ (esto define parte de Φ) y podemos tomar $s_i = \Phi_{r_i}$.

Sin embargo, una de las maneras más elegantes de generar integrales primeras es a través de las simetrías del problema, como lo veremos en la siguiente sección. (Recordemos que para poder resolver completamente las ecuaciones canónicas necesitamos $2n$ integrales primeras, ya que así reducimos el sistema de $2n$ ecuaciones de primer orden a un sistema algebraico).

Nota 4.7: La Ecuación de Hamilton-Jacobi.

Uno de los propósitos de las transformadas canónicas es el de poder integrar fácilmente las ecuaciones asociadas. Siguiendo a Jacobi, las ecuaciones canónicas más fáciles de integrar son aquellas para las cuales $\tilde{H}(t, S, R) = 0$.

En efecto, entonces:

$$\begin{aligned} \dot{s}_i &= \tilde{H}_{r_i} = 0 \\ \dot{r}_i &= -\tilde{H}_{s_i} = 0 \end{aligned}$$

las soluciones son $s_i = \text{constante}$, $r_i = \text{constante}$.

Para esto tendremos que buscar una función generatriz de la transformación canónica $(Q, P) \longrightarrow (S, R)$ de la forma $G(t, Q, S)$ tal que

$$\begin{aligned}\frac{\partial G}{\partial t} &= -H(t, Q, P) \quad (\tilde{H} = 0) \\ p_i &= \frac{\partial G}{\partial q_i}\end{aligned}$$

y definir r_i a través de $r_i = -\frac{\partial G}{\partial s_i}$.

Como $r_i = \text{constante}$, podemos buscar una función generatriz que no dependa de S , o una familia de funciones parametrizadas por n constantes s_1, \dots, s_n : $G(t, Q)$ tal que:

$$\frac{\partial G}{\partial t} + H(t, Q, \frac{\partial G}{\partial q_1}, \dots, \frac{\partial G}{\partial q_n}) = 0.$$

Esta es la **ecuación de Hamilton-Jacobi**.

Es una ecuación en derivadas parciales de primer orden para $G(t, Q)$. Si pudiésemos resolverla, con n constantes independientes, obtendríamos la transformación $(P, Q) \longrightarrow (S, R)$ (a partir de $r_i = -\frac{\partial G}{\partial s_i}$).

Esta derivación de la ecuación de Hamilton-Jacobi es un poco mágica. Además no queda muy claro si perdimos información en el proceso (es decir, si podemos regresar a la solución de las ecuaciones canónicas o mejor aún, a la solución de $q_i(t)$) o no. En la siguiente sección veremos como probar esta equivalencia a partir del estudio de la ecuación de Hamilton-Jacobi. Veremos en particular que si, conceptualmente, la ecuación de Hamilton-Jacobi es más elegante que las ecuaciones de Euler, el resolverla es exactamente resolver las ecuaciones de Euler. Es decir, que en la práctica no se avanzó nada!

5.5 Teorema de Noether y leyes de conservación.

Como ya lo hemos notado, varios sistemas físicos son independientes de las coordenadas o del origen del tiempo. Por ejemplo, la acción total de un sistema, donde la Lagrangiana no depende del tiempo, acumulada entre t_0 y t_1 debe ser la misma que entre $t_0 + T$ y $t_1 + T$. Esto debe tener consecuencias físicas, en este caso que la energía se conserva. Puede haber otras cantidades o transformaciones que dejen invariante a la acción, ¿Cuáles son los resultados? En general consideremos el funcional

$$J(y_1, \dots, y_n) = \int_{t_0}^{t_1} F(t, y_1, \dots, y_n, y'_1, \dots, y'_n) dt.$$

Supongamos que hay una familia de aplicaciones, parametrizada por ϵ de $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R} \times \mathbb{R}^n$

$$(t, y_i, z_i) \rightarrow (\tau_\epsilon, Y_\epsilon) = (\tau_\epsilon(t, y_i, z_i), Y_{j\epsilon}(t, y_i, z_i))$$

tal que $(\tau_\epsilon, Y_\epsilon)$ sean C^2 (también en ϵ) y

$$\tau_0 = t, \quad Y_0 = y.$$

Esto implica que para ϵ pequeño, podemos invertir parcialmente en la forma

$$t = t(\tau_\epsilon, Y_\epsilon, z), \quad y = y(\tau_\epsilon, Y_\epsilon, z).$$

Si tenemos $y_i = y_i(t)$, $z_i = y'_i(t)$, entonces $\tau_\epsilon = \tau_\epsilon(t, y_i(t), y'_i(t)) = \tau_\epsilon(t)$, con $\tau_0 = t$, podemos expresar, para ϵ pequeño, $t = t(\tau_\epsilon)$, $Y_\epsilon(t, y_i(t), y'_i(t)) = Y_\epsilon(t) = Y_\epsilon(t(\tau_\epsilon))$ a través de la transformación de $t = t(\tau_\epsilon)$.

Definición 5.1: Diremos que el funcional es **invariante** bajo la familia de aplicaciones si:

$$\begin{aligned} J_\epsilon(Y_{1\epsilon}, \dots, Y_{n\epsilon}) &= \int_{\tau_{0\epsilon}}^{\tau_{1\epsilon}} F(\tau_\epsilon, Y_{i\epsilon}(t(\tau_\epsilon)), \frac{dY_{i\epsilon}(t(\tau_\epsilon))}{d\tau_\epsilon}) d\tau_\epsilon \\ &= J(y_1, \dots, y_n) \quad \text{para todo } y_1, \dots, y_n, t_0, t_1. \end{aligned}$$

Cuidado: No se trata de un cambio de variables, sino de una substitución de t por τ_ϵ (incluyendo en los argumentos de las funciones y en las derivadas) y de y_i por $Y_{i\epsilon}$.

Ejemplo 5.1 : Si consideramos la familia $\tau_\epsilon = t + \epsilon$, $Y_{i\epsilon} = y_i$ entonces $t = \tau_\epsilon - \epsilon$, $y_i(t) = y_i(\tau_\epsilon - \epsilon) = Y_{i\epsilon}(t(\tau_\epsilon))$.

$$J_\epsilon = \int_{t_0+\epsilon}^{t_1+\epsilon} F(\tau_\epsilon, y_i(\tau_\epsilon - \epsilon), y'_i(\tau_\epsilon - \epsilon)) d\tau_\epsilon.$$

Al hacer el cambio de variables $t = \tau_\epsilon - \epsilon$:

$$J_\epsilon = \int_{t_0}^{t_1} F(t + \epsilon, y_i(t), y'_i(t)) dt.$$

es invariante si F no depende de t .

En general si hacemos el cambio de variables $\tau_\epsilon = \tau_\epsilon(t)$ tendremos $d\tau_\epsilon = \frac{d\tau_\epsilon}{dt} dt$ con $\tau_\epsilon = \tau_\epsilon(t, y_i(t), y'_i(t))$

$$\frac{dY_{i\epsilon}(t(\tau_\epsilon))}{d\tau_\epsilon} = \frac{dY_{i\epsilon}(t, y_i(t), y'_i(t))}{dt} \frac{dt}{d\tau_\epsilon}.$$

Por lo tanto:

$$J_\epsilon(Y_{1\epsilon}, \dots, Y_{n\epsilon}) = \int_{t_0}^{t_1} F\left(\tau_\epsilon(t), Y_{i\epsilon}(t, y_i(t), y'_i(t)), \frac{dY_{i\epsilon}}{dt} \frac{dt}{d\tau_\epsilon}\right) \frac{d\tau_\epsilon}{dt} dt.$$

Teorema 5.1 : (Noether 1918). Si J_ϵ es invariante, entonces en un punto crítico de J_0 , uno tiene la integral primera:

$$\sum_1^n F_{z_i} \frac{\partial Y_{i\epsilon}}{\partial \epsilon} \Big|_{\epsilon=0} + \left(F - \sum_1^n y'_i F_{z_i} \right) \frac{\partial \tau_\epsilon}{\partial \epsilon} \Big|_{\epsilon=0} = \text{constante}$$

Prueba: Como $J_\epsilon = J_0$, entonces

$$\begin{aligned} & \left. \frac{dJ_\epsilon}{d\epsilon} \right|_{\epsilon=0} = 0 \\ &= \int_{t_0}^{t_1} \left[\left(F_t \frac{\partial \tau_\epsilon}{\partial \epsilon} + \sum F_{y_i} \frac{\partial Y_{i\epsilon}}{\partial \epsilon} + \sum F_{z_i} \frac{\partial}{\partial \epsilon} \left(\frac{dY_{i\epsilon}}{dt} \left(\frac{d\tau_\epsilon}{dt} \right)^{-1} \right) \right) \frac{d\tau_\epsilon}{dt} + F \frac{\partial}{\partial \epsilon} \left(\frac{d\tau_\epsilon}{dt} \right) \right]_{\epsilon=0} dt. \end{aligned}$$

Ahora, como $\tau_0 = t$, $\left. \frac{d\tau_\epsilon}{dt} \right|_{\epsilon=0} = 1$, $\left. \frac{dY_{i\epsilon}}{dt} \right|_{\epsilon=0} = \frac{dy_i}{dt} = y'_i$,
tenemos entonces:

$$0 = \int_{t_0}^{t_1} \left[F_t \frac{\partial \tau_\epsilon}{\partial \epsilon} + \sum F_{y_i} \frac{\partial Y_{i\epsilon}}{\partial \epsilon} + \sum F_{z_i} \left(\frac{d}{dt} \left(\frac{\partial Y_{i\epsilon}}{\partial \epsilon} \right) - \frac{d}{dt} \left(\frac{\partial \tau_\epsilon}{\partial \epsilon} \right) y'_i \right) + F \frac{d}{dt} \left(\frac{\partial \tau_\epsilon}{\partial \epsilon} \right) \right] dt$$

donde hemos usado el hecho que τ_ϵ , $Y_{i\epsilon}$ son C^2 .

Notando que para $\epsilon = 0$, los argumentos de F_{z_i} son (t, y_i, y'_i) y, por ser y_i un punto crítico, podemos integrar por partes, ya que F_{z_i} y $F - \sum y'_i F_{z_i}$ son C^1 (primera y segunda ecuaciones de Euler) con derivadas F_{y_i} y F_t respectivamente:

$$\sum F_{z_i} \frac{\partial Y_{i\epsilon}}{\partial \epsilon} \Big|_{\epsilon=0} + (F - \sum y'_i F_{z_i}) \frac{\partial \tau_\epsilon}{\partial \epsilon} \Big|_{\epsilon=0} \Big|_{t_0}^{t_1} = 0.$$

Como esto debe ser cierto para todo t_0 y t_1 obtenemos el resultado.

Q.E.D.

En el Ejemplo 5.1, tenemos que $F - \sum y'_i F_{z_i} = \text{constante}$, lo cual es la integral primera dada por la segunda ecuación de Euler. Por la nota 4.1 tenemos conservación de la energía total.

Ejercicio 5.1 : Teorema de Noether en dos dimensiones.

Sea la familia de transformaciones, de \mathbb{R}^5 en \mathbb{R}^3 :

$$\begin{aligned} x(\epsilon) &= \varphi_1(x, y, u, v, w, \epsilon) & \text{con } x(0) &= x, & \varphi_1 &\text{ es } C^2 \\ y(\epsilon) &= \varphi_2(x, y, u, v, w, \epsilon) & \text{con } y(0) &= y, & \varphi_2 &\text{ es } C^2 \\ u(\epsilon) &= \varphi(x, y, u, v, w, \epsilon) & \text{con } u(0) &= u, & \varphi &\text{ es } C^2, \end{aligned}$$

y sea $I(u) = \int \int_\Omega F(x, y, u, u_x, u_y) dx dy$

Para cada función $u(x, y)$, $v = u_x(x, y)$, $w = u_y(x, y)$, tendremos $x(\epsilon) = x(x, y, \epsilon)$, $y(\epsilon) = y(x, y, \epsilon)$, $u(\epsilon) = u(x, y, \epsilon)$. Sea entonces

$$I_\epsilon(u) = \int_{\Omega_\epsilon} \int F \left(x(\epsilon), y(\epsilon), u(\epsilon), \frac{\partial u(\epsilon)}{\partial x(\epsilon)}, \frac{\partial u(\epsilon)}{\partial y(\epsilon)} \right) dx(\epsilon) dy(\epsilon),$$

donde, si (x, y) pertenece a Ω , $(x(\epsilon), y(\epsilon))$ pertenece a Ω_ϵ .

Como $(x(0), y(0)) = (x, y)$, para ϵ pequeño podemos hacer el cambio de variables $(x(\epsilon), y(\epsilon)) \rightarrow (x, y)$, con

$$I_\epsilon(u) = \int_{\Omega} \int F \left(x(\epsilon), y(\epsilon), u(\epsilon), \frac{\partial u(\epsilon)}{\partial x} \frac{\partial x}{\partial x(\epsilon)} + \frac{\partial u(\epsilon)}{\partial y} \frac{\partial y}{\partial x(\epsilon)}, \frac{\partial u(\epsilon)}{\partial x} \frac{\partial x}{\partial y(\epsilon)} + \frac{\partial u(\epsilon)}{\partial y} \frac{\partial y}{\partial y(\epsilon)} \right) \left| \frac{D(x(\epsilon), y(\epsilon))}{D(x, y)} \right| dx dy.$$

Diremos que I es invariante bajo la familia de transformaciones si $I_\epsilon(u) = I(u)$ para todo ϵ pequeño y todo Ω . Probar que $\frac{dI_\epsilon}{d\epsilon}$ vale:

$$\int_{\Omega} \int V \text{ Euler}_u(F)(x, y) dx dy + \int_{\partial\Omega} \begin{pmatrix} F_{u_x} V + F \varphi_{1\epsilon} \\ F_{u_y} V + F \varphi_{2\epsilon} \end{pmatrix} \cdot \vec{n} ds$$

donde

$$\text{Euler}_u(F)(x, y) = F_u - \frac{\partial}{\partial x}(F_{u_x}) - \frac{\partial}{\partial y}(F_{u_y}),$$

$$V = \varphi_\epsilon - u_x \varphi_{1\epsilon} - u_y \varphi_{2\epsilon},$$

$$(\varphi_{i\epsilon} = \frac{\partial \varphi_i}{\partial \epsilon}).$$

Concluir que si I_ϵ es invariante en un punto crítico de I entonces existe una función Φ tal que:

$$\begin{aligned} F_{u_x} V + F \varphi_{1\epsilon} &= \Phi_y \\ F_{u_y} V + F \varphi_{2\epsilon} &= -\Phi_x. \end{aligned}$$

Ejercicio 5.2 : Si en el problema de encontrar las condiciones de transversalidad, hacemos la siguiente transformación:

$$\begin{aligned} \tau_\epsilon(x, y, z) &= x + \epsilon h(x) \\ Y_\epsilon(x, y, z) &= y + \varphi(x + \epsilon h(x)) - \varphi(x) \end{aligned}$$

para h en $C^1[a, \infty[$ con $h(a) = 0$,
y si aplicamos el teorema de Noether, a pesar de que

$$I(y) = \int_a^{x_1} F(x, y, y') dx$$

no es invariante bajo esta transformación; se encuentra:

$$F_{y'} \varphi'(x_1) h(x_1) + (F - y' F_{y'}) h(x_1) = 0$$

(en $x = a$ esto vale 0). Como $h(x_1)$ es arbitrario, tendríamos

$$F - (y' - \varphi') F_{y'} \Big|_{x_1} = 0.$$

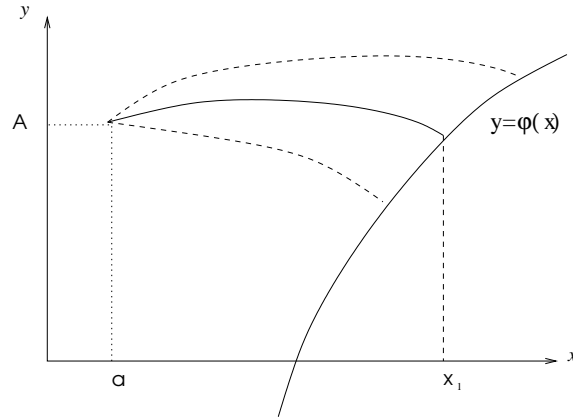


Fig. 5.1

Explicar por qué el resultado es cierto a pesar que I_ϵ no sea invariante, examinando con cuidado la demostración del teorema de Noether.

Ejemplo 5.2 : Conservación de momento.

Ya vimos, en el ejemplo 5.1, que si F no depende de t , entonces $F - \sum y'_i F_{y'_i} = -E$ se conserva (ver la Nota 4.1). Supongamos ahora que $F(t, y_1, \dots, y_n, z_1, \dots, z_n)$ no depende explícitamente de y_1 , por ejemplo.

Si hacemos la transformación:

$$\tau_\epsilon = t, Y_{1\epsilon} = y_1 + \epsilon, Y_{i\epsilon} = y_i \quad i \geq 2,$$

tendremos

$$F(\tau_\epsilon, Y_{1\epsilon}, Y_{2\epsilon}, \dots, Y_{n\epsilon}, \frac{dY_{1\epsilon}}{d\tau_\epsilon}, \frac{dY_{2\epsilon}}{d\tau_\epsilon}, \dots) d\tau_\epsilon = F(t, y_2, \dots, y_n, y'_1, \dots, y'_n) dt$$

y por lo tanto $J_\epsilon(Y_{1\epsilon}, \dots, Y_{n\epsilon}) = J(y_1, \dots, y_n)$.

Por el teorema de Noether:

$$F_{z_1} = \text{constante sobre las soluciones de las ecuaciones de Euler.}$$

Por ejemplo si $F = \frac{1}{2} \sum m_i |\dot{X}_i|^2 - U(t, X_1, \dots, X_n)$, para n partículas X_1, \dots, X_n , X_i en \mathbb{R}^3 . Tendremos que si U no depende explícitamente de X_{i1} , la primera componente de X_i , entonces $p_{i1} = F_{z_{i1}}$ es constante: el momento de cada partícula con respecto al eje x_1 , es constante. Si al reemplazar X_{i1} por $X_{i1} + \epsilon$ para cada i , U no cambia, entonces la suma de los momentos con respecto al eje x_1 es constante.

Ejemplo 5.3 : Conservación del momento angular.

Supongamos ahora que

$$L = \frac{1}{2} \sum m_i |\dot{X}_i|^2 - U(t, X_1, \dots, X_n)$$

es tal que para una familia de rotaciones $A(\epsilon)$ en \mathbb{R}^3 , con $A(0) = I$, tenemos

$$U(t, A(\epsilon)X_1, \dots, A(\epsilon)X_n) = U(t, X_1, \dots, X_n).$$

(Por ejemplo, si la energía potencial depende sólo de la distancia a un eje, podemos tomar:

$$A(\epsilon) = \begin{pmatrix} \cos \epsilon & -\sin \epsilon & 0 \\ \sin \epsilon & \cos \epsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{si el eje es el eje } x_3.$$

Notando que con la transformación:

$$\tau_\epsilon = t, \quad X_i(\epsilon) = A(\epsilon)X_i$$

entonces

$$\left| \frac{dX_i(\epsilon)}{d\tau_\epsilon} \right|^2 = |\dot{X}_i(\epsilon)|^2 = \dot{X}_i^T A(\epsilon)^T A(\epsilon) \dot{X}_i = |\dot{X}_i|^2$$

ya que $A(\epsilon)$ es una rotación. Por lo tanto la acción es invariante y, por el teorema de Noether:

$$\sum_i P_i \cdot A'(0)X_i = \text{constante},$$

es decir, la suma de los momentos angulares con respecto al eje de rotación definido por $A'(0)$ es constante. Notemos que como $A^T(\epsilon)A(\epsilon) = I$ y $A(0) = I$, entonces $A'^T(0) + A'(0) = 0$, es decir $A'(0)$ es antisimétrica (con determinante 0, ya que $A'(0)$ es una matriz en \mathbb{R}^3), tiene un vector propio con valor propio 0. Este vector propio define al eje de rotación).

Ahora bien si

$$A'(0) = \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

entonces $A'(0)X = X \times \begin{pmatrix} c \\ -b \\ a \end{pmatrix} = X \times K$ (verificarlo).

$X = K$ es el eje de rotación,

y $P \cdot A'(0)X = (P, X, K) = (P \times X) \cdot K$, donde usamos la propiedad cíclica del producto mixto.

Por lo tanto $(\sum P_i \times X_i) \cdot K = \text{constante}$, es decir, el momento angular del sistema con respecto al eje K es constante.

En el caso de una rotación con respecto al eje x_3 ,

$$A'(0) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad P_i \cdot A'_i(0)X_i = (P_i \times X_i) \cdot (0, 0, -1)^T = P_{i2}X_{i1} - P_{i1}X_{i2}.$$

Ejercicio 5.3 : Problema de fuerza central. Supongamos que $L = \|P\|^2/(2m) + k/r$ para una partícula en el espacio, a una distancia r del origen. $P^T = (p_1, p_2, p_3)$ es el momento.

1. A partir de rotaciones con respecto a cada eje, probar que sobre la trayectoria $X \times \dot{X} = \text{constante}$.
2. Derivando esta relación, probar que la torsión de la curva es 0, es decir, que la trayectoria es plana: Primera Ley de Kepler.
3. Poniéndose en el plano de la trayectoria y en polares, derivar la Segunda Ley de Kepler.
4. A partir de la conservación de energía, encontrar la trayectoria. Discutir los distintos casos según el valor de la energía.

5.6 La ecuación de Hamilton-Jacobi

Es esta sección veremos como, en principio, podemos reducir el sistema de $2n$ ecuaciones canónicas a una sola ecuación de primer orden (desafortunadamente en derivadas parciales. ¡no hay milagros!). Ya vimos, en la Nota 4.7, como derivar la ecuación de Hamilton-Jacobi a partir de las transformadas canónicas. En esta sección la haremos en forma más geométrica, a través de las funciones geodésicas y la teoría de campos.

5.6.1 Función distancia geodésica

Consideremos el funcional:

$$J(y_1, \dots, y_n) = \int_{t_0}^T F(t, y_1(t), \dots, y_n(t), y'_1(t), \dots, y'_n(t)) dt$$

y supongamos que el problema

$$(F_{y'_i})' - F_{y_i} = 0, \quad y_i(t_0) = A_i, \quad y_i(T) = Y_i$$

tiene una única solución $y_i(t, T, Y_1, \dots, Y_n)$ tal que y_i y y'_i sean C^1 en (T, Y_1, \dots, Y_n) .

Aquí la hipótesis fuerte es la de existencia y unicidad de la solución. Si F es C^2 y $\det F_{z_i z_j} \neq 0$ entonces, como en la proposición IV.1.1, y_i es C^2 y es diferenciable en los datos.

Definición 6.1: *La función distancia geodésica es la cantidad:*

$$S_0(T, Y_1, \dots, Y_n) = \int_{t_0}^T F(t, y_1(t, T, Y_1, \dots, Y_n), \dots, y'_n(t, T, Y_1, \dots, Y_n)) dt.$$

Para (T, Y_1, \dots, Y_n) fijo, $S_0(T, Y_1, \dots, Y_n)$ es el valor del funcional en su punto crítico. La curva $y(t, T, Y_1, \dots, Y_n)$ es la curva “geodésica” entre (t_0, A_1, \dots, A_n) y (T, Y_1, \dots, Y_n) .

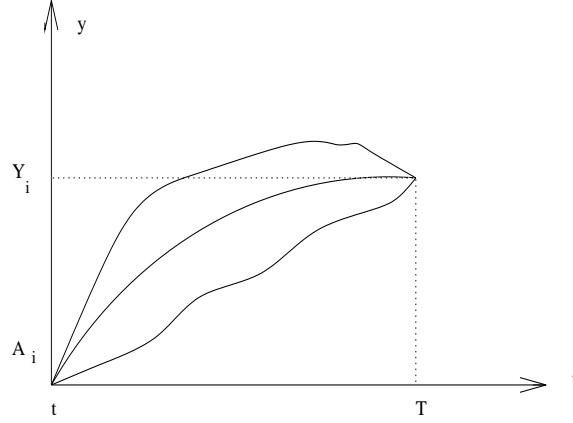


Fig. 6.1

Notemos que si pedimos la existencia de una única curva geodésica, cuando (t_0, A_1, \dots, A_n) son arbitrarias, entonces para (T, Y_1, \dots, Y_n) y $(\bar{T}, \bar{Y}_1, \dots, \bar{Y}_n)$ hay una sola curva geodésica que los une.

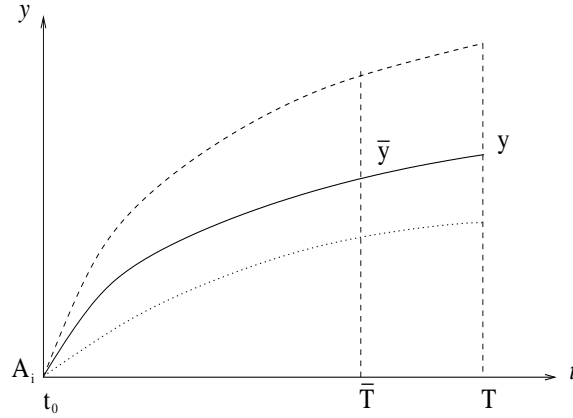


Fig. 6.2

Ahora, la función S_0 es diferenciable: de hecho:

$$\frac{\partial S_0}{\partial Y_i} = \int_{t_0}^T \left(\sum_j F_{y_j} \frac{\partial y_j}{\partial Y_i} + F_{z_j} \frac{\partial^2 y_j}{\partial Y_i \partial t} \right) dt$$

como y'_j es C^1 en los datos y $F_{z_j}(t, y_1(t, T, Y_1, \dots, Y_n), \dots, y'_n(t, T, \dots, Y_n))$ es C^1 con derivadas F_{y_j} (por ser punto crítico), podemos integrar por partes y obtener:

$$\frac{\partial S_0}{\partial Y_i} = \sum_j F_{z_j} \frac{\partial y_j}{\partial Y_i} \Big|_{t_0}^T.$$

Pero $y_j(t_0, T, Y_1, \dots, Y_n) = A_j$ es independiente de Y_i y $y_j(T, T, Y_1, \dots, Y_n) = Y_j$, por lo tanto:

$$\frac{\partial S_0}{\partial Y_i} = F_{z_i}(T, Y_1, \dots, Y_n, y'_1(T, T, Y_1, \dots, Y_n), \dots, y'_n(T, T, Y_1, \dots, Y_n)) \equiv p_i.$$

Del mismo modo:

$$\frac{\partial S_0}{\partial T} = F \Big|_T + \int_{t_0}^T \sum_j \left(F_{y_j} \frac{\partial y_j}{\partial T} + F_{z_j} \frac{\partial^2 y_j}{\partial T \partial t} \right) dt.$$

Integrando de nuevo por partes, y usando la ecuación de Euler:

$$\frac{\partial S_0}{\partial T} = F \Big|_T + \sum_j F_{z_j} \frac{\partial y_j}{\partial T} \Big|_{t_0}^T.$$

Como $y_j(T, T, Y_1, \dots, Y_n) = Y_j$, tenemos:

$$\frac{dy_j}{dT} = 0 = \frac{\partial y_j}{\partial t} + \frac{\partial y_j}{\partial T},$$

por lo tanto:

$$\frac{\partial S_0}{\partial T} = F(T, Y_1, \dots, Y_n, y'_1(T, T, \dots, Y_n) \dots) - \sum p_i y'_i(T, T, Y_1, \dots, Y_n).$$

Recordando que si $(F_{z_i z_j})_{ij}$ es invertible sobre la trayectoria, la transformación de Legendre está bien definida (Nota 4.3), podemos expresar z_i en función de p_i , y tendremos

$$\frac{\partial S_0}{\partial T} = -H(T, Y_1, \dots, Y_n, p_1, \dots, p_n).$$

Usando el hecho que $p_i = \frac{\partial S_0}{\partial Y_i}$, tenemos:

Proposición 6.1: Si $(F_{z_i z_j})_{ij}$ es invertible y la función distancia geodésica está definida, entonces $S_0(T, Y_1, \dots, Y_n)$ satisface a la ecuación de Hamilton-Jacobi:

$$\frac{\partial S_0}{\partial T} + H(T, Y_1, \dots, Y_n, \frac{\partial S_0}{\partial Y_1}, \dots, \frac{\partial S_0}{\partial Y_n}) = 0.$$

Recordemos que si $(F_{z_i z_j})_{ij}$ es invertible, entonces las ecuaciones de Euler son equivalentes a las ecuaciones canónicas (Nota 4.4):

$$\dot{y}_i = H_{p_i}, \quad \dot{p}_i = -H_{y_i}$$

con $y_i(t_0) = A_i$, $y_i(T) = Y_i$.

Notemos que si $T = t_0$, entonces la función distancia geodésica está definida sólo si $Y_i = A_i$ con valor $S(t_0, A_1, \dots, A_n) = 0$. Es decir, $S(t_0, Y_1, \dots, Y_n)$ no está definida en general. Para poder “resolver” la ecuación de Hamilton-Jacobi es necesario tener unos datos iniciales no singulares. Para poder obtenerlos daremos una hipótesis más fuerte.

5.6.2 Condiciones de transversalidad

Consideremos la superficie siguiente, en forma paramétrica:

$$\begin{cases} t = t_0(s_1, \dots, s_n) \\ y_i = A_i(s_1, \dots, s_n) \end{cases} \equiv \Sigma(s_1, \dots, s_n)$$

y el funcional $J(y_1, \dots, y_n) = \int_{t_0}^T F(t, y_1, \dots, y_n, y'_1, \dots, y'_n) dt$.

Si buscamos los puntos críticos de J para curvas $y(t)$ tales que $y_i(T) = Y_i$, $y_i(t_0)$ esté sobre la superficie:

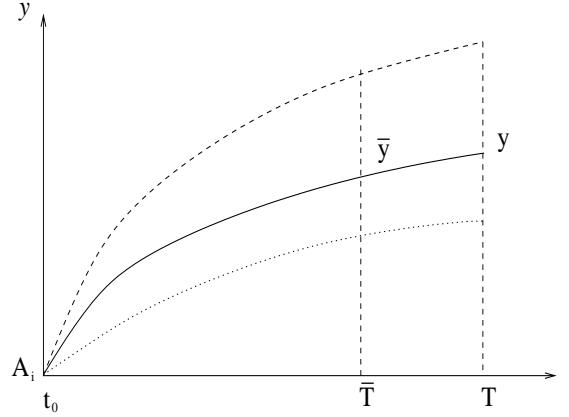


Fig. 6.3

Como lo vimos en el ejemplo IV.8.4, podemos tratar este problema como el equivalente a encontrar los puntos críticos de $J(y_1, \dots, y_n, t_0, s_1, \dots, s_n)$ condicionados por la definición de Σ . Tendremos por lo tanto $n + 2$ multiplicadores, $\nu, \lambda_0, \lambda_1, \dots, \lambda_n$ tales que:

$$\begin{aligned} & \nu D J(y_1, \dots, t_0, s_1, \dots, s_n)(h_1, \dots, h_n, \alpha_0, \dots, \alpha_n) \\ & + \sum_0^n \lambda_i D G_i(y_1, \dots, t_0, s_1, \dots, s_n)(h_1, \dots, h_n, \alpha_0, \dots, \alpha_n) = 0 \end{aligned}$$

para h_i en C^1 tales que $h_i(T) = 0$, (ya que $y_i(T) = Y_i$ es fijo), donde:

$$\begin{aligned} G_0 &= t - t_0(s_1, \dots, s_n) \\ G_i &= y_i(t_0) - A_i(s_1, \dots, s_n). \end{aligned}$$

Por lo tanto:

$$\begin{aligned} & \nu \left(\int_{t_0}^T \sum_j (F_{y_j} h_j + F_{z_j} h'_j) dt - F \Big|_{t_0} \alpha_0 \right) + \lambda_0 (\alpha_0 - \sum_j t_{s_j} \alpha_j) \\ & + \sum_1^n \lambda_i \left(h_i(t_0) + y'_i(t_0) \alpha_0 - \sum_j A_{i s_j} \alpha_j \right) = 0. \end{aligned}$$

Tomando $\alpha_0, \alpha_j = 0$, h_j en $C_0^\infty[t_0, T]$, tendremos que las ecuaciones de Euler son válidas: $(F_{z_j})' = F_{y_j}$ y podemos integrar por partes, usando $h_j(T) = 0$.

Como $h_{j_n}(t_0), \alpha_0, \dots, \alpha_n$ son arbitrarios:

$$\begin{aligned}\lambda_0 + \sum_1^n \lambda_i y_i'(t_0) &= \nu F \Big|_{t_0} \\ \lambda_i &= \nu F_{z_i} \Big|_{t_0} \\ \lambda_0 t_{s_j} + \sum_i \lambda_i A_{is_j} &= 0 \quad \text{para } j = 1, \dots, n\end{aligned}$$

De las $n + 2$ primeras ecuaciones vemos que ν no puede ser 0. Substituyendo $\lambda_0, \lambda_1, \dots, \lambda_n$ en las últimas, tendremos:

$$(F - \sum_1^n y_i' F_{z_i}) t_{s_j} + \sum_1^n A_{is_j} F_{z_i} = 0 \quad \text{para } j = 1, \dots, n.$$

Lema 6.1: Sea A la matriz definida por $A_{ij} = A_{js_i} = \partial A_j / \partial s_i$, entonces los puntos críticos de J condicionados por $y_i(T) = Y_i$, $t = t_0(s_1, \dots, s_n)$, $y_i = A_i(s_1, \dots, s_n)$ satisfacen las ecuaciones de Euler: $(F_{z_i})' = F_{y_i}$, y las condiciones de transversalidad en Σ :

$$AP = H \nabla t_0$$

$$\text{donde } P^T = (p_1, \dots, p_n) = (F_{z_1}, \dots, F_{z_n})$$

$$H = -(F - \sum_1^n y_i' F_{z_i}) = y' \cdot P - F.$$

En particular si $A_i(s_1, \dots, s_n) = s_i$, entonces A es la matriz identidad y tendremos

$$H \frac{\partial t_0}{\partial s_i} = p_i = F_{z_i}(t_0(s_1, \dots, s_n), y_i(t_0), y_i'(t_0)).$$

Hipótesis de función geodésica.

Supondremos que dado Σ , existe para (T, Y_1, \dots, Y_n) una única solución a este problema, con las condiciones de transversalidad, $y(t, T, Y_1, \dots, Y_n)$. Definimos entonces la función distancia geodésica:

$$S(T, Y_1, \dots, Y_n) = \int_{t_0}^T F(t, y_1(t, T, Y_1, \dots, Y_n), \dots, y_n'(t, T, Y_1, \dots, Y_n)) dt.$$

Notemos que t_0 depende de (T, Y_1, \dots, Y_n) : de hecho para cada (T, Y_1, \dots, Y_n) , hay un único punto $(t_0, y_1(t_0), \dots, y_n(t_0))$ sobre la superficie inicial Σ (la unicidad del punto de salida viene de la unicidad de $y(t, T, Y_1, \dots, Y_n)$), y por lo tanto tenemos una aplicación: $(T, Y_1, \dots, Y_n) \longrightarrow (s_1, \dots, s_n)$ definida por la correspondencia entre datos de llegada y los datos de salida sobre Σ .

Hipótesis de continuidad: Supondremos que $s_i(T, Y_1, \dots, Y_n)$ son C^1 . Así mismo supondremos que $y_i(t, T, Y_1, \dots, Y_n)$, $y_i'(t, T, Y_1, \dots, Y_n)$ son C^1 .

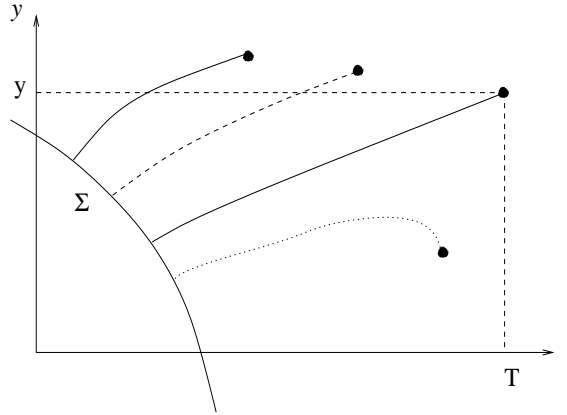


Fig. 6.4

Notemos que la segunda hipótesis implica la primera (sobre s_i) si la superficie es regular, (Por el Teorema de la Función implícita).

Podemos ahora estudiar la función:

$$S(T, Y_1, \dots, Y_n) = \int_{t_0(T, Y_1, \dots, Y_n)}^T F(t, y_1(t, T, Y_1, \dots, Y_n), \dots, y'_n(t, T, Y_1, \dots, Y_n)) dt.$$

Tenemos:

$$\frac{\partial S}{\partial Y_i} = \int_{t_0}^T \sum_j (F_{y_j} \frac{\partial y_j}{\partial Y_i} + F_{z_j} \frac{\partial^2 y_j}{\partial t \partial Y_i}) dt - \frac{\partial t_0}{\partial Y_i} F.$$

Como $F'_{z_j} = F_{y_j}$, podemos integrar por partes:

$$\frac{\partial S}{\partial Y_i} = \sum_j F_{z_j} \frac{\partial y_j}{\partial Y_i} \Big|_{t_0}^T - \frac{\partial t_0}{\partial Y_i} F.$$

Sabemos que $y_j(T, T, Y_1, \dots, Y_n) = Y_j$ y

$$y_j(t_0(T, Y_1, \dots, Y_n), T, Y_1, \dots, Y_n) = A_j(s_1, \dots, s_n) = A_j(T, Y_1, \dots, Y_n).$$

Por lo tanto:

$$\sum_j p_j \frac{\partial y_j}{\partial Y_i} \Big|_{t_0} = \sum_j p_j \left(\frac{\partial A_j}{\partial Y_i} - \frac{\partial y_j}{\partial t_0} \frac{\partial t_0}{\partial Y_i} \right).$$

(recordemos que $\frac{\partial y_j}{\partial Y_i}(t, T, Y_1, \dots, Y_n)$ es la parcial con respecto a Y_i cuando t es fijo).

$$\frac{\partial S}{\partial Y_i} = F_{z_i} \Big|_T - \sum_j p_j \frac{\partial A_j}{\partial Y_i} \Big|_{t_0} + \left(\sum_j (y'_j p_j) - F \right) \frac{\partial t_0}{\partial Y_i}.$$

Como

$$\frac{\partial A_j}{\partial Y_i} = \sum_k \frac{\partial A_j}{\partial s_k} \frac{\partial s_k}{\partial Y_i}, \quad \frac{\partial t_0}{\partial Y_i} = \sum_k \frac{\partial t_0}{\partial s_k} \frac{\partial s_k}{\partial Y_i}$$

entonces

$$H \frac{\partial t_0}{\partial Y_i} - \sum_j p_j \frac{\partial A_j}{\partial Y_i} = \sum_k \left(\frac{\partial t_0}{\partial s_k} H - \sum_j \frac{\partial A_j}{\partial s_k} p_j \right) \frac{\partial s_k}{\partial Y_i} = 0.$$

ya que esta es justamente la condición de transversalidad.

Por lo tanto:

$$\frac{\partial S}{\partial Y_i} = p_i(T, Y_1, \dots, Y_n, y'_1(T, T, Y_1, \dots, Y_n), \dots).$$

Del mismo modo:

$$\begin{aligned} \frac{\partial S}{\partial T} &= \int_{t_0}^T \sum_j \left(F_{y_j} \frac{\partial y_j}{\partial T} + F_{z_j} \frac{\partial^2 y_j}{\partial t \partial T} \right) dt + F \Big|_T - F \frac{\partial t_0}{\partial T} \\ &= \sum_j F_{z_j} \frac{\partial y_j}{\partial T} \Big|_{t_0}^T + F \Big|_T - F \frac{\partial t_0}{\partial T}. \end{aligned}$$

Como antes tenemos

$$\frac{\partial y_j}{\partial T} \Big|_T = -y'_j(T, T, Y_1, \dots, Y_n) \quad , \quad (y_j(T, T, \dots) = Y_j)$$

$\frac{\partial y_j}{\partial T} \Big|_{t_0} = \frac{\partial A_j}{\partial T} - y'_j \frac{\partial t_0}{\partial T}$ (como antes), por lo tanto:

$$\frac{\partial S}{\partial T} = -H(T, Y_1, \dots, Y_n, p_1, \dots, p_n) + \left(\sum_j (y'_j p_j) - F \right) \frac{\partial t_0}{\partial T} - \sum_j p_j \frac{\partial A_j}{\partial T}.$$

Como $\frac{\partial A_j}{\partial T} = \sum_k \frac{\partial A_j}{\partial s_k} \frac{\partial s_k}{\partial T}$, $\frac{\partial t_0}{\partial T} = \sum_k \frac{\partial t_0}{\partial s_k} \frac{\partial s_k}{\partial T}$, entonces

$$\left(\sum_j (y'_j p_j) - F \right) \frac{\partial t_0}{\partial T} - \sum_j p_j \frac{\partial A_j}{\partial T} = \sum_k \left(H \frac{\partial t_0}{\partial s_k} - \sum_j p_j \frac{\partial A_j}{\partial s_k} \right) \frac{\partial s_k}{\partial T} = 0.$$

obtenemos

$$\frac{\partial S}{\partial T} + H = 0$$

en (T, Y_1, \dots, Y_n) , donde los argumentos de F_{z_i} son $(T, Y_1, \dots, Y_n, y'_1(T, T, Y_1, \dots, Y_n) \dots)$. Podemos expresar $y'_i(T, T, Y_1, \dots, Y_n)$ en función de p_i si la matrix $(F_{z_i z_j})_{ij}$ es invertible:

Teorema 6.1 : Si $(F_{z_i z_j})$ es invertible, si las hipótesis de la función geodésica y de continuidad son satisfechas, entonces $S(T, Y_1, \dots, Y_n)$ satisface la ecuación de Hamilton-Jacobi:

$$\frac{\partial S}{\partial T} + H\left(T, Y_1, \dots, Y_n, \frac{\partial S}{\partial Y_1}, \dots, \frac{\partial S}{\partial Y_n}\right) = 0.$$

Notemos que $S\Big|_{\Sigma} = 0$, $T\Big|_{\Sigma} = t_0(s_1, \dots, s_n)$, $Y_i\Big|_{\Sigma} = A_i(s_1, \dots, s_n)$.

Evidentemente lo que nos interesa es poder resolver esta ecuación, o por lo menos decir como hacerlo. Para esto veremos como resolver una ecuación en derivadas parciales de primer orden:

$$G\left(T, Y_1, \dots, Y_n, S, \frac{\partial S}{\partial T}, \frac{\partial S}{\partial Y_1}, \dots, \frac{\partial S}{\partial Y_n}\right) = 0$$

con datos sobre una superficie inicial $T\Big|_{\Sigma}, Y_i\Big|_{\Sigma}, S\Big|_{\Sigma}$ dados.

5.6.3 Soluciones de ecuaciones en derivadas parciales de primer orden

Sea pues $G(T, Y_1, \dots, Y_n, S, p_0, p_1, \dots, p_n) = 0$, con $p_0 = \frac{\partial S}{\partial T}$, $p_i = \frac{\partial S}{\partial Y_i}$, una ecuación de primer orden, donde S está dada sobre una superficie inicial Σ definida por: $T = t_0(s_1, \dots, s_n)$, $Y_i = A_i(s_1, \dots, s_n)$.

A) Supongamos primero que tenemos una solución $S(T, Y_1, \dots, Y_n)$ (C^2) a este problema, con $S\Big|_{\Sigma} = S(s_1, \dots, s_n)$.

$S = S(T, Y_1, \dots, Y_n)$ representa entonces una superficie en \mathbb{R}^{n+2} que pasa por una superficie inicial dada por $(\Sigma, S\Big|_{\Sigma})$.

Además la normal a la superficie $S - S(T, Y_1, \dots, Y_n) = 0$ será:

$$N = (S_T, S_{Y_1}, \dots, S_{Y_n}, -1) = (\tilde{P}, -1) \quad \text{con } \tilde{P} = (p_0, p_1, \dots, p_n).$$

En particular esta normal debe ser ortogonal a todos los vectores tangentes a $M = (\Sigma, S\Big|_{\Sigma})$,

es decir a $\frac{\partial}{\partial s_i}(t_0, A_1, \dots, A_n, S)$:

$$p_0 t_{0s_i} + \sum_j^n p_j A_{js_i} = S_{s_i}(s_1, \dots, s_n) \quad , i = 1, \dots, n.$$

Evidentemente tenemos también en esa superficie:

$$G(t_0, A_1, \dots, A_n, S, p_0, p_1, \dots, p_n) = 0.$$

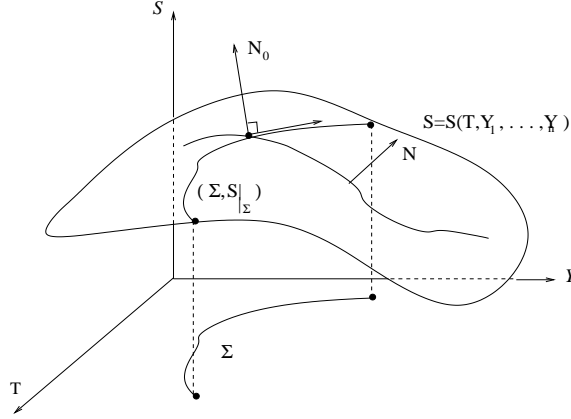


Fig. 6.5

Considerando estas $n + 1$ ecuaciones como funciones de $(s_1, \dots, s_n, p_0, \dots, p_n)$, podemos expresar (p_0, \dots, p_n) (sobre Σ) como funciones de (s_1, \dots, s_n) si

$$\begin{pmatrix} t_{0s_1} & \dots & t_{0s_n} & G_{p_0} \\ A_{1s_1} & & A_{1s_n} & G_{p_1} \\ & \vdots & & \vdots \\ A_{ns_1} & \dots & A_{ns_n} & G_{p_n} \end{pmatrix} \equiv K(s_1, \dots, s_n),$$

es invertible (condición necesaria para poder usar el teorema de la función implícita).

Supondremos que esto es el caso. Tenemos entonces $\tilde{P} = \tilde{P}(s_1, \dots, s_n)$ sobre Σ .

Notemos que como $G(T, Y_1, \dots, Y_n, S(T, Y_1, \dots, Y_n), \frac{\partial S}{\partial T}, \frac{\partial S}{\partial Y_1}(T, Y_1, \dots, Y_n), \dots) = 0$, podemos diferenciar esta relación con respecto a T, Y_1, \dots, Y_n :

$$G_T + G_S p_0 + \sum_{i=0}^n G_{p_i} p_{0Y_i} = 0, \quad \text{usando } p_{iT} = \frac{\partial^2 S}{\partial Y_i \partial T} = p_{0Y_i}, \quad (S \text{ es } C^2)$$

$$G_{Y_j} + G_S p_j + \sum_{i=0}^n G_{p_i} p_{jY_i} = 0, \quad \text{usando } p_{iY_j} = p_{jY_i}, \quad (S \text{ es } C^2)$$

donde aquí pusimos $Y_0 = T$.

Como $p_j = \frac{\partial S}{\partial Y_j}(T, Y_1, \dots, Y_n) = p_j(T, Y_1, \dots, Y_n)$ podemos también pensar en una superficie: $p_j(T, Y_1, \dots, Y_n) - p_j = 0$, con normal $(p_{jY_0}, p_{jY_1}, \dots, p_{jY_n}, -1)$, pasando por la curva inicial $(t_0(s_1, \dots, s_n), A_1(s_1, \dots, s_n), \dots, A_n(s_1, \dots, s_n), p_j(s_1, \dots, s_n))$.

Además p_j satisface la ecuación diferencial lineal:

$$G_{Y_j} + G_S p_j + \sum_{i=0}^n G_{p_i} p_{jY_i} = 0$$

donde los argumentos en G_{Y_j}, G_S, G_{p_i} son (T, Y_1, \dots, Y_n) . Podemos interpretar la ecuación como el producto escalar de la normal con el vector $(G_{p_0}, \dots, G_{p_n}, -(G_{Y_j} + G_S p_j))$.

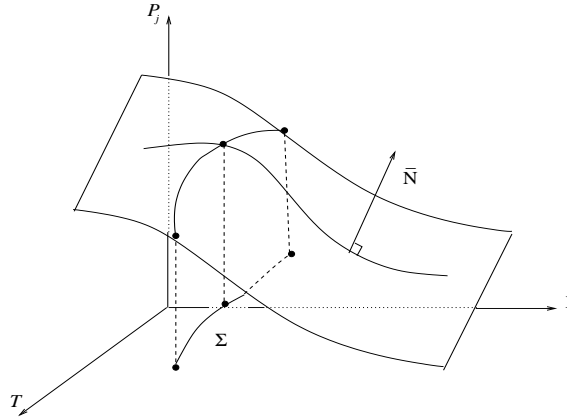


Fig. 6.6

Si fijamos (s_1, \dots, s_n) , podemos encontrar una curva $(T(\eta), Y_1(\eta), \dots, Y_n(\eta), p_j(\eta))$ sobre la superficie $p_j - p_j(T, Y_1, \dots, Y_n) = 0$, que tenga como vector tangente:

$$(T'(\eta), Y_1'(\eta), \dots, Y_n'(\eta), p_j'(\eta)) = (G_{p_0}, G_{p_1}, \dots, G_{p_n}, -(G_{Y_j} + G_{s_j} p_j))$$

y que pase, en $\eta = 0$, por $(t_0(s_1, \dots, s_n), A_1(s_1, \dots, s_n), \dots, p_j(s_1, \dots, s_n))$ si resolvemos el sistema de ecuaciones:

$$\begin{aligned} T'(\eta) &= G_{p_0}(T, Y_1, \dots, Y_n) & T(0) &= t_0 \\ Y_i'(\eta) &= G_{p_i}(T, Y_1, \dots, Y_n) & Y_i(0) &= A_i \quad i = 1, \dots, n \\ p_j'(\eta) &= -(G_{Y_j} + G_{s_j} p_j) & p_j(0) &= p_j(s_1, \dots, s_n). \end{aligned}$$

Este sistema de $(n+1)$ ecuaciones de primer orden tiene una solución única, para η pequeña. Además

$$-p_j(\eta) + p_j(T(\eta), Y_1(\eta), \dots, Y_n(\eta)) \text{ tiene derivada}$$

$$-p_j'(\eta) + \sum_{i=0}^n p_{jY_i} Y_i'(\eta) = G_{Y_j} + G_{s_j} p_j + \sum_{i=0}^n G_{p_i} p_{jY_i} = 0$$

(usando las ecuaciones para $p_j'(\eta)$ y $Y_i'(\eta)$). Por lo tanto $p_j(\eta) - p_j$ es una constante que vale 0 en $\eta = 0$: la única curva que obtuvimos se encuentra efectivamente sobre la superficie $p_j = p_j(T, Y_1, \dots, Y_n)$.

Busquemos ahora una curva $(T(\eta), Y(\eta), S(\eta))$ sobre la superficie $S = S(T, Y_1, \dots, Y_n)$ y un campo de vectores $\tilde{P}(\eta)$ tal que $(\tilde{P}(\eta), -1)$ sea normal a la superficie, y tal que para $\eta = 0$, tengamos: $T(0) = t_0(s_1, \dots, s_n)$, $Y_j(0) = A_j(s_1, \dots, s_n)$, $S(0) = S(s_1, \dots, s_n)$, $\tilde{P}(0) = \tilde{P}(s_1, \dots, s_n)$, es decir el mismo problema que hemos resuelto para cada superficie $p_j = p_j(T, Y_1, \dots, Y_n)$, pero esta vez al mismo tiempo para cada una de estas superficies y para la superficie $S = S(T, Y_1, \dots, Y_n)$.

Tendremos que resolver el sistema:

$$\begin{aligned} T'(\eta) &= G_{p_0}(T, Y_1, \dots, Y_n, S, p_0, p_1, \dots, p_n) \\ Y_j'(\eta) &= G_{p_j}(T, Y_1, \dots, Y_n, S, p_0, p_1, \dots, p_n) \\ p_j'(\eta) &= -(G_{Y_j} + G_{S p_j})(T, Y_1, \dots, Y_n, S, p_0, \dots, p_n) \\ S'(\eta) &= \sum_{i=0}^n p_i G_{p_i}(T, Y_1, \dots, Y_n, S, p_0, \dots, p_n) \end{aligned}$$

con datos iniciales: $T(0) = t_0(s_1, \dots, s_n)$, $Y_j(0) = A_j(s_1, \dots, s_n)$, $p_j(0) = p_j(s_1, \dots, s_n)$, $S(0) = S(s_1, \dots, s_n)$.

(La última ecuación viene de la regla de la cadena: si

$$S(\eta) = S(T(\eta), Y_1(\eta), \dots, Y_n(\eta)) \text{ entonces:}$$

$$S'(\eta) = \sum_{i=0}^n S_{Y_i} Y_i'(\eta) = \sum p_i Y_i'(\eta) \text{ (por definición de } p_i \text{)).}$$

Otra vez, si G es C^2 , este sistema tiene una única solución, para η pequeña. Para esa solución: $S(\eta) - S(T(\eta), Y_1(\eta), \dots, Y_n(\eta))$ tiene derivada $S'(\eta) - \sum_0^n p_i(T, Y_0, \dots, Y_n) Y_i'(\eta) = 0$.

Como en $\eta = 0$, $S(\eta) - S(t_0, A_1, \dots, A_n) = 0$, la curva está sobre la superficie. Además $p_i(\eta) = \frac{\partial S}{\partial Y_i}(T, Y_1, \dots, Y_n)$, $i = 0, \dots, n$ como ya lo vimos anteriormente para la curva sobre la superficie $p_j = p_j(T, \dots, Y_n)$ (por la unicidad de la solución a cada subsistema, $(T(\eta), Y_1(\eta), \dots, Y_n(\eta), p_j(\eta))$ es la curva que ya obtuvimos).

En resumen:

Si la matriz K es invertible, para cada (s_1, \dots, s_n) obtenemos una única solución, sobre la superficie $S = S(T, Y_1, \dots, Y_n)$, al sistema anterior.

B) Usaremos esta unicidad y existencia de la solución al sistema para generar la superficie S :

Aún si no sabemos todavía de la existencia de la superficie S , el sistema:

$$\begin{aligned} Y_j'(\eta) &= G_{p_j}(Y_0, \dots, Y_n, S, p_0, \dots, p_n) \quad , j = 0, \dots, n \\ p_j'(\eta) &= -(G_{Y_j} + G_{S p_j}) \quad , j = 0, \dots, n \\ S'(\eta) &= \sum_0^n p_i G_{p_i} \text{ con datos iniciales :} \\ Y_0(0) &= T(0) = t_0(s_1, \dots, s_n) \\ Y_j(0) &= A_j(s_1, \dots, s_n) \quad j = 1, \dots, n \\ p_j(0) &= p_j(s_1, \dots, s_n) \\ S(0) &= S(s_1, \dots, s_n) \end{aligned}$$

tiene una única solución, para η pequeña.

Notemos además que la función:

$$G(\eta) = G(Y_0(\eta), \dots, Y_n(\eta), S(\eta), p_0(\eta), \dots, p_n(\eta))$$

tiene como derivada

$$G'(\eta) = \sum_0^n G_{Y_i} G_{p_i} + G_S \left(\sum_0^n p_i G_{p_i} \right) + \sum_0^n G_{p_i} (-G_{Y_i} - G_S p_i)$$

es decir $G(\eta)$ es constante sobre las trayectorias. Pero recordando que $p_i(s_1, \dots, s_n)$ fue obtenido usando la relación

$$G(t_0, A_1, \dots, A_n, S, p_0, \dots, p_n) = 0$$

por lo tanto $G(0) = 0$. Tenemos entonces que $G(\eta) = 0$.

Ahora bien, si hacemos variar (s_1, \dots, s_n) obtendremos una familia de curvas parametrizada por (s_1, \dots, s_n) :

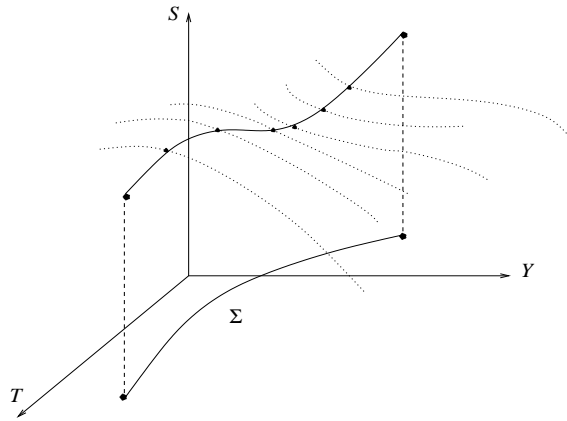


Fig. 6.7

Por la unicidad de las soluciones, estas curvas no se intersectan, y, por el teorema de dependencia continua con respecto a los datos iniciales, las soluciones son tan diferenciables con respecto a s_1, \dots, s_n como los datos y las derivadas de G . Es decir, para G en C^3 , t_0, A_j, p_j, S en C^2 , las soluciones son C^2 en (s_1, \dots, s_n) .

(Notemos que si K es invertible, $p_i(s_1, \dots, s_n)$ es C^1 si $t_0, A_j, S(s_1, \dots, s_n)$ son C^2 y G es C^1). Supondremos que esta regularidad es cierta: G es C^3 , t_0, A_j, S son C^2 , entonces (Y_j, S) son C^2 y p_j son C^1 .

Sean entonces $T(\eta, s_1, \dots, s_n)$, $Y_j(\eta, s_1, \dots, s_n)$, $S(\eta, s_1, \dots, s_n)$ esas familias de soluciones. Veremos que de hecho forman una superficie $S(T, Y_1, \dots, Y_n)$ que satisface a la ecuación diferencial.

Para esto veremos primero que para η pequeña podemos invertir la transformación: $(\eta, s_1, \dots, s_n) \rightarrow (T, Y_1, \dots, Y_n)(\eta, s_1, \dots, s_n)$: el jacobiano de la transformación es, en $\eta = 0$:

$$\begin{pmatrix} T_{s_1} & \dots & T_{s_n} & T'(\eta) \\ Y_{1s_1} & \dots & Y_{1s_n} & Y'_1(\eta) \\ \vdots & & \vdots & \vdots \\ Y_{ns_1} & \dots & Y_{ns_n} & Y'_n(\eta) \end{pmatrix}_{\eta=0} = K(\eta) \Big|_{\eta=0}$$

donde se usaron las ecuaciones para $T'(\eta)$, $Y'_j(\eta)$ y los datos iniciales. Como K es invertible, entonces $K(\eta)$ lo es también para η pequeña, hay una transformación inversa:

$$(\eta, s_1, \dots, s_n) = (\eta, s_1, \dots, s_n)(T, Y_1, \dots, Y_n).$$

Notemos que cada columna de $K(\eta)$ es un vector tangente a la superficie dada en forma paramétrica como $(T(\eta, s_1, \dots, s_n), Y_j(\eta, s_1, \dots, s_n))$. El hecho que $K(\eta)$ sea invertible dice que estos vectores tangentes son linealmente independientes: tenemos efectivamente una superficie.

Obtenemos entonces $S(T, Y_1, \dots, Y_n)$, $p_j(T, Y_1, \dots, Y_n)$, funciones C^2 y C^1 respectivamente.

Como sabemos que $G(Y_0, \dots, Y_n, S, p_0, \dots, p_n) = 0$, ya que es cierto para cada trayectoria, solo falta probar que $p_j = \frac{\partial S}{\partial Y_j}$, $j = 0, \dots, n$.

Notemos primero que la superficie $S(T, Y_1, \dots, Y_n)$ tiene normal $(S_{Y_0}, \dots, S_{Y_n}, -1)$ y pasa por la curva $(\sum, S \Big|_{\sum})$, por lo tanto la normal debe ser ortogonal a $\frac{\partial}{\partial s_i}(t_0, A_1, \dots, A_n, S)$, es decir:

$$\sum_{i=0}^n \frac{\partial S}{\partial Y_i} \frac{\partial Y_i}{\partial s_j} = S_{s_j} \quad j = 1, \dots, n.$$

Además

$$S'(\eta) = \sum_{i=0}^n \frac{\partial S}{\partial Y_i} Y'_i(\eta) = \sum_{i=0}^n \frac{\partial S}{\partial Y_i} G_{p_i}.$$

Estas relaciones pueden escribirse como:

$$K^T \frac{\partial S}{\partial Y} = \frac{\partial S}{\partial s}$$

(es únicamente la regla de la cadena para la transformación).

Ahora bien, las n primeras relaciones son también satisfechas por \tilde{P} en $\eta = 0$, ya que así se definieron los datos iniciales para \tilde{P} . La última relación también es satisfecha por \tilde{P} ya que $S'(\eta) = \sum_0^n p_i G_{p_i}$.

Por lo tanto $K^T \left(\frac{\partial S}{\partial Y} - \tilde{P} \right) = 0$.

Como K es invertible, tenemos $\frac{\partial S}{\partial Y} = \tilde{P} \Big|_{\eta=0}$.

Notemos que de las dos formas de escribir $S'(\eta)$ obtenemos:

$$\sum G_{p_i} \left(\frac{\partial S}{\partial Y_i} - p_i \right) = 0, \quad (\text{para todo } T, Y_1, \dots, Y_n).$$

Diferenciando esta relación con respecto a Y_j :

$$(*) \sum_{i=0}^n G_{p_i} \left(\frac{\partial^2 S}{\partial Y_i \partial Y_j} - p_{iY_j} \right) = - \sum_{i=0}^n \frac{\partial G_{p_i}}{\partial Y_j} \left(\frac{\partial S}{\partial Y_i} - p_i \right).$$

Ahora bien, sabemos que $G(Y_0, \dots, Y_n, S, p_1, \dots, p_n) = 0$.

Diferenciando con respecto a Y_j :

$$G_{Y_j} + G_S S_{Y_j} + \sum_{i=0}^n G_{p_i} p_{iY_j} = 0.$$

Además

$$\begin{aligned} p'_j(\eta) &= -G_{Y_j} - G_S p_j = \frac{d}{d\eta} p_j(Y_0(\eta), \dots, Y_n(\eta)) \\ &= \sum_{i=0}^n p_{jY_i} G_{p_i}, \quad (\text{con } Y'_i = G_{p_i}). \end{aligned}$$

De estas dos últimas relaciones, tenemos:

$$\sum_{i=0}^n G_{p_i} p_{iY_j} = \sum_{i=0}^n G_{p_i} p_{jY_i} - G_S (S_{Y_j} - p_j).$$

Substituyendo en (*):

$$\sum_{i=0}^n G_{p_i} (S_{Y_i Y_j} - p_{jY_i}) = -G_S (S_{Y_j} - p_j) - \sum_{i=0}^n \frac{\partial G_{p_i}}{\partial Y_j} \left(\frac{\partial S}{\partial Y_i} - p_i \right) = (S'_{Y_j} - p'_j)(\eta),$$

(ya que $S'_{Y_j} = \frac{d}{d\eta} (S_{Y_j}(Y_0, \dots, Y_n)) = \sum_{i=0}^n S_{Y_i Y_j} Y'_i(\eta)$ y S es C^2).

Por lo tanto $S_{Y_j} - p_j$ satisface a un sistema de ecuaciones lineales con coeficientes continuos en η (de hecho C^1) y condiciones iniciales homogéneas: en ese caso la solución existe y es única: $S_{Y_j} = p_j$.

La solución que encontramos satisface entonces $G(Y_0, \dots, Y_n, S, S_{Y_0}, \dots, S_{Y_n}) = 0$, en una vecindad de $(\sum, S \Big|_{\sum})$.

Notemos finalmente que esta solución es única: en efecto si hubiera otra solución, por la parte (A) llegaríamos a una familia de curvas **sobre la superficie** solución del sistema de ecuaciones, con los mismos datos. Como este sistema tiene una solución única, llegamos a que las dos superficies coinciden.

En resumen:

Teorema 6.2 : Si G es C^3 , $t_0(s_1, \dots, s_n)$, $A_i(s_1, \dots, s_n)$, $S(s_1, \dots, s_n)$ son C^2 , si

$$K = \begin{pmatrix} t_{0s_1} & \dots & t_{0s_n} & G_{p_0} \\ A_{1s_1} & \dots & A_{1s_n} & G_{p_1} \\ \vdots & & \vdots & \vdots \\ A_{ns_1} & \dots & A_{ns_n} & G_{p_n} \end{pmatrix}$$

es invertible, entonces el problema:

$$G(T, Y_1, \dots, Y_n, S, S_{Y_1}, \dots, S_{Y_n}) = 0$$

$S = S(s_1, \dots, s_n)$ sobre $T = t_0(s_1, \dots, s_n)$, $Y_i = A_i(s_1, \dots, s_n)$

tiene una única solución en la vecindad de la “curva inicial” y está dada por la solución al sistema:

$$\begin{aligned} \frac{dY_i}{d\eta} &= G_{p_i}, \quad Y_0(0) = t_0 \text{ y } Y_i(0) = A_i \\ \frac{dp_i}{d\eta} &= -G_{Y_i} - G_S p_i, \quad p_i(0) = p_i(s_1, \dots, s_n) \\ \frac{dS}{d\eta} &= \sum p_i G_{p_i}, \quad S(0) = S(s_1, \dots, s_n) \end{aligned}$$

donde $p_i(s_1, \dots, s_n)$ estan dados por la única solución a:

$$G(t_0, A_1, \dots, A_n, S, p_0, \dots, p_n) = 0$$

$$\sum_{i=0}^n p_i \frac{\partial A_i}{\partial s_j} = S_{s_j}, \quad j = i, \dots, n.$$

5.6.4 Aplicación a la ecuación de Hamilton-Jacobi

En este caso $G = S_T + H(T, Y_1, \dots, Y_n, S_{Y_1}, \dots, S_{Y_n}) = 0$ con $S \Big|_{\Sigma} = 0, T \Big|_{\Sigma} = t_0, Y_i \Big|_{\Sigma} = A_i$.

La matrix K será invertible si

$$\tilde{K} \equiv \begin{pmatrix} 1 & H_{p_1} & \dots & H_{p_n} \\ t_{0s_1} & & \cdot & \\ \vdots & & A & \\ t_{0s_n} & & & \end{pmatrix}$$

lo es (hemos intercambiado los renglones de K^T), donde $A = (\partial A_j / \partial s_i)_{ij}$ ya fue usada en el lema 6.1. Supondremos que A es invertible, es decir, que Σ puede ser parametrizada por A_1, \dots, A_n , como $t = t_0(A_1, \dots, A_n)$. Podemos escribir

$$\tilde{K} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & A & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & H_{p_1} & \dots & H_{p_n} \\ t_{0A_1} & & & \\ & I & & \\ t_{0A_n} & & & \end{pmatrix}$$

con $t_{0A_i} = \sum_j t_{0s_j} \partial s_j / \partial A_i$ de tal forma que

$$\sum_j A_{ij} t_{0A_j} = \sum_{jk} \frac{\partial A_j}{\partial s_i} \frac{\partial s_k}{\partial A_j} t_{0s_k} = \sum_{jk} A_{ij} A_{jk}^{-1} t_{0s_k} = \sum_k (AA^{-1})_{ik} t_{0s_k} = t_{0s_i}.$$

Entonces $\det K = \det \tilde{K} = \det A \det K_n$, donde K_n es la segunda matriz. Desarrollando el determinante de K_n con respecto al último renglón:

$$\det K_n = \det K_{n-1} + (-1)^n \det \begin{pmatrix} H_{p_1} & \dots & H_{p_n} \\ 1 & \ddots & \vdots \\ 0 & & \\ 0 & \dots & 1 & 0 \end{pmatrix} t_{0A_n}$$

$= \det K_{n-1} - H_{p_n} t_{0A_n}$ (desarrollarlo con respecto a la última columna). Por inducción: $\det K_n = 1 - \sum_1^n H_{p_i} t_{0A_i}$ y $\det K = \det A (1 - \sum_1^n H_{p_i} t_{0A_i})$.

Por lo tanto K es invertible si $1 - \sum_1^n H_{p_i} t_{0A_i} \neq 0$ (y $\det A \neq 0$).

En esas condiciones podemos encontrar las condiciones iniciales para p_0, \dots, p_n , a partir de las relaciones:

$$\begin{aligned} p_0 + H &= 0 \\ p_0 t_{0s_j} + \sum_i p_i A_{is_j} &= S_{s_j} = 0. \end{aligned}$$

Es decir (aquí es fácil de resolver): $AP = -p_0 \nabla_s t_0 = H \nabla_s t_0$.

Notemos que esta es justamente la condición de transversalidad (Lema 6.1), (recordemos que $P = (p_1, \dots, p_n)$, $\tilde{P} = (p_0, P)$). Por lo tanto:

$$P = HA^{-1} \nabla_s t_0 = H \left(\frac{\partial s_i}{\partial A_j} \right)_{ij} \nabla_s t_0 = H \nabla_A t_0.$$

$P|_{\Sigma}$ queda entonces determinado en función de p_0 y de los datos t_0, A_i .

$$P = -p_0 \nabla_A t_0.$$

Para poder determinar $p_0|_{\Sigma}$, usaremos la relación:

$$0 = p_0 + H(t_0(s_1, \dots, s_n), A_1(s_1, \dots, s_n), \dots, A_n(s_1, \dots, s_n), 0, -p_0 t_{0s_1}, \dots, -p_0 t_{0s_n}).$$

Para poder resolver esta ecuación en (p_0, s_1, \dots, s_n) para p_0 , podemos usar el teorema de la función implícita; para esto necesitamos que la derivada con respecto a p_0 sea diferente de 0: es decir $1 - \sum H_{p_i} t_{0A_i} \neq 0$ (lo que ya habíamos hecho globalmente).

Notemos que Σ , expresado como $t_0 = t_0(A_1, \dots, A_n)$ tiene como normal:

$(1, -\nabla_A t_0) = (p_0, P)/p_0 = \tilde{P}/p_0$ y que la superficie $S(T, Y_1, \dots, Y_n)$ tiene como normal $(\tilde{P}, -1)$.

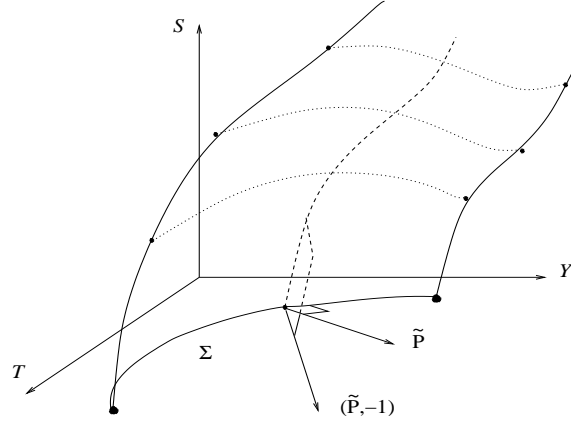


Fig. 6.8

El sistema de ecuaciones (equivalente a la ecuación en derivadas parciales) es entonces:

$$\begin{aligned}
 \frac{dT}{d\eta} &= 1 & T(0) &= t_0(s_1, \dots, s_n) \\
 \frac{dY_j}{d\eta} &= H_{p_j} & Y_j(0) &= A_j(s_1, \dots, s_n) \\
 \frac{dS}{d\eta} &= \sum p_i G_{p_i} = p_0 + \sum p_i H_{p_i} = -H + \sum \frac{dS}{d\eta} = F & S(0) &= 0 \\
 \frac{dp_0}{d\eta} &= -(G_T + p_0 G_S) = -H_T & p_0(0) &= p_0(s_1, \dots, s_n) \\
 \frac{dp_j}{d\eta} &= -(G_{Y_j} + p_j G_S) = -H_{Y_j} & p_j(0) &= p_j(s_1, \dots, s_n).
 \end{aligned}$$

Podemos integrar la primera y la tercera ecuación:

$$\begin{aligned}
 T &= t_0 + \eta \\
 S(T) &= \int_{t_0}^T F(t, Y_1(t), \dots, Y_n(t), Y_1'(t), \dots, Y_n'(t)) dt
 \end{aligned}$$

donde se usó $H_{p_i} = Y_i'(t)$. La cuarta ecuación es únicamente la derivada de la energía sobre la curva $(t, Y_1(t), \dots, Y_n(t))$.

Teorema 6.3 : Si $F_{z_i z_j}$ es invertible, si sobre la superficie inicial $\Sigma : t = t_0(s_1, \dots, s_n)$, $y_i = A_i(s_1, \dots, s_n)$, con F en C^3 , t_0 , A_i en C^2 , si además $1 - \sum H_{p_i} t_{A_i} \neq 0$ y la matriz $\left(\frac{\partial A_j}{\partial s_i}\right)_{ij}$ es invertible, entonces:

1) La ecuación de Hamilton-Jacobi:

$$S_T + H(T, Y_1, \dots, Y_n, S_{Y_1}, \dots, S_{Y_n}) = 0$$

con los datos iniciales sobre Σ

$$T = t_0(s_1, \dots, s_n), Y_i = A_i(s_1, \dots, s_n) \quad , S = 0$$

tiene una única solución en una vecindad de la superficie inicial.

2) Esta solución está dada por la solución al sistema canónico:

$$\begin{aligned} Y'_j(t) &= H_{p_j} & Y'_j(t_0) &= A_j \\ p'_j(t) &= -H_{Y_j} & p_j(t_0) &\text{ dado por } P(t_0) = -p_0 \nabla_A t_0 \\ p'_0(t) &= -H_T & p_0(t_0) &\text{ dado por } p_0 + H|_{t_0} = 0. \end{aligned}$$

3) La función $S(T) = \int_{t_0}^T F(t, Y_1(t), \dots, Y_n(t), Y'_1(t), \dots, Y'_n(t)) dt$ es la función distancia geodésica, es decir, bajo estas hipótesis la ecuación de Hamilton-Jacobi es equivalente al sistema canónico.

(Notemos que la hipótesis de invertibilidad de $F_{z_i z_j}$ es necesaria para tener la equivalencia entre las ecuaciones de Euler y las ecuaciones canónicas). Para cada (T, Y_1, \dots, Y_n) en una vecindad de Σ , tenemos un único (η, s_1, \dots, s_n) que le corresponde (ya que la transformación $(T, Y_1, \dots, Y_n) \rightarrow (\eta, s_1, \dots, s_n)$ es invertible ($K(\eta)$ es invertible)). Fijando (s_1, \dots, s_n) obtenemos, al variar η , una curva $(T(\eta), Y_1(\eta), \dots, Y_n(\eta))$ solución de las ecuaciones de Euler (equivalentes a las ecuaciones canónicas) y transversal a Σ en $(t_0(s_1, \dots, s_n), A_1(s_1, \dots, s_n), \dots, A_n(s_1, \dots, s_n))$.

Esta curva es la única de la familia que pasa por (T, Y_1, \dots, Y_n) . Esta es justamente la hipótesis de la función distancia geodésica. La hipótesis de continuidad es clara por las hipótesis de diferenciabilidad impuestas.

La familia de soluciones $Y_j(\eta)$, soluciones de las ecuaciones canónicas y transversales a la superficie inicial Σ , se llama un **campo de extremales**.

Ejemplo 6.1 : Principio de Huyghens. Si $S = \int_{x_0}^x (1 + y'^2 + z'^2)^{\frac{1}{2}} v^{-1}(x, y, z) dx$, es decir el funcional que determina el tiempo de recorrido de un rayo de luz, entonces $p_i = F_{z_i} = z_i/v(1 + z_1^2 + z_2^2)^{\frac{1}{2}}$, $H = p_1 z_1 + p_2 z_2 - F = -v^{-1}(1 + z_1^2 + z_2^2)^{-\frac{1}{2}}$.

Pero como $p_1^2 + p_2^2 = v^{-2} - v^{-2}(1 + z_1^2 + z_2^2)^{-1}$, $H = -(v^{-2} - p_1^2 - p_2^2)^{\frac{1}{2}}$ tendremos $\frac{dy}{dx} = H_{p_1} = -p_1/H$, $\frac{dz}{dx} = H_{p_2} = -p_2/H$.

La condición inicial: \tilde{P}/p_0 es normal a Σ es entonces: (p_0, p_1, p_2) es normal a Σ . Pero como $p_0 = -H$ (ecuación de Hamilton-Jacobi), la normal a Σ es $(1, \frac{dy}{dx}, \frac{dz}{dx})$: el rayo de luz es ortogonal a la superficie inicial: este es el principio de Huyghens.

Nota 6.1: La condición de invertibilidad de la matriz $(\frac{\partial A_i}{\partial s_j})_{ij}$ implica que podemos parametrizar la curva inicial por y_{10}, \dots, y_{n0} , como $t = t_0(y_{10}, \dots, y_{n0})$. La condición $1 - \sum H_{p_i} t_{y_{i0}} \neq 0$ siempre se puede cumplir al variar un poco la superficie inicial. Entonces, para cada (T, Y_1, \dots, Y_n) vecino a Σ , tendremos una única solución a las ecuaciones canónicas pasando por ese punto (ya que S es una función de (T, Y_1, \dots, Y_n) , transversal a Σ , donde \tilde{P} es un campo de vectores sobre la curva, $Y_j(T)$, normal a Σ).

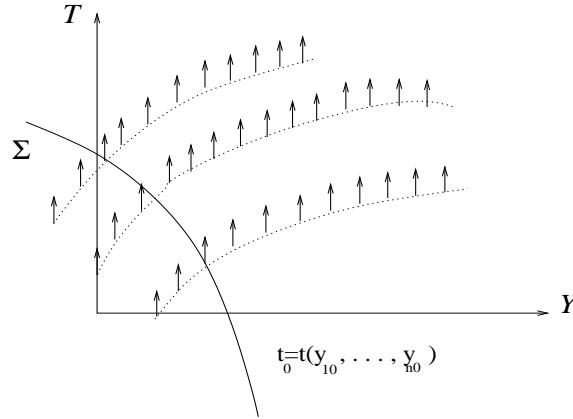


Fig. 6.9

Notemos sin embargo, que **no** hemos resuelto el problema inicial, es decir, una solución a las ecuaciones canónicas con condiciones de frontera (y no iniciales) $y_i(t_0) = y_{i0}$, $y_i(T) = Y_i$ (ni siquiera en una vecindad de Σ) y que la solución que hemos obtenido solo fue encontrada en una vecindad de Σ . El primer problema (el de encontrar una solución a la ecuación de Hamilton-Jacobi cerca de un punto inicial) puede resolverse fácilmente como lo veremos a continuación. El segundo problema (existencia y unicidad del problema de frontera para puntos de frontera arbitrarios y de manera consistente) no tiene siempre una solución y está ligado a las condiciones suficientes del próximo capítulo.

5.6.5 Campos centrales

Como lo acabamos de hacer notar, la introducción de una superficie inicial fue un artificio para poder usar los resultados de existencia y unicidad para ecuaciones en derivadas parciales de primer orden. El planteamiento inicial de la función distancia geodésica, fue a partir de un problema donde las condiciones en la frontera (en t_0 y en T) eran fijas y dadas solo para $y_i(t_0) = A_i$, $y_i(T) = Y_i$, y no para un problema con frontera en una superficie Σ :

Sin embargo, como lo vimos en la Nota 6.1, tenemos mucha libertad en la elección de la superficie Σ : sólo debe ser parametrizada en la forma $t_0(y_{10}, \dots, y_{n0})$ y satisfacer la condición $1 - \sum H_{p_i} t_{y_{i0}} \neq 0$.

Por el punto (t_0, A_1, \dots, A_n) podemos hacer pasar muchas de estas superficies, haciendo variar la “pendiente” en ese punto.

Es fácil convencerse que, en el punto (t_0, A_1, \dots, A_n) , dado cualquier vector $(\alpha_1, \dots, \alpha_n)$, con $1 - \sum H_{p_i}(t_0, A_1, \dots, A_n, -p_0\alpha_1, \dots, -p_0\alpha_n)\alpha_i \neq 0$ (donde p_0 es tal que $p_0 + H(t_0, A_1, \dots, A_n, -p_0\alpha_1, \dots, -p_0\alpha_n) = 0$), podemos encontrar una superficie Σ pasando por

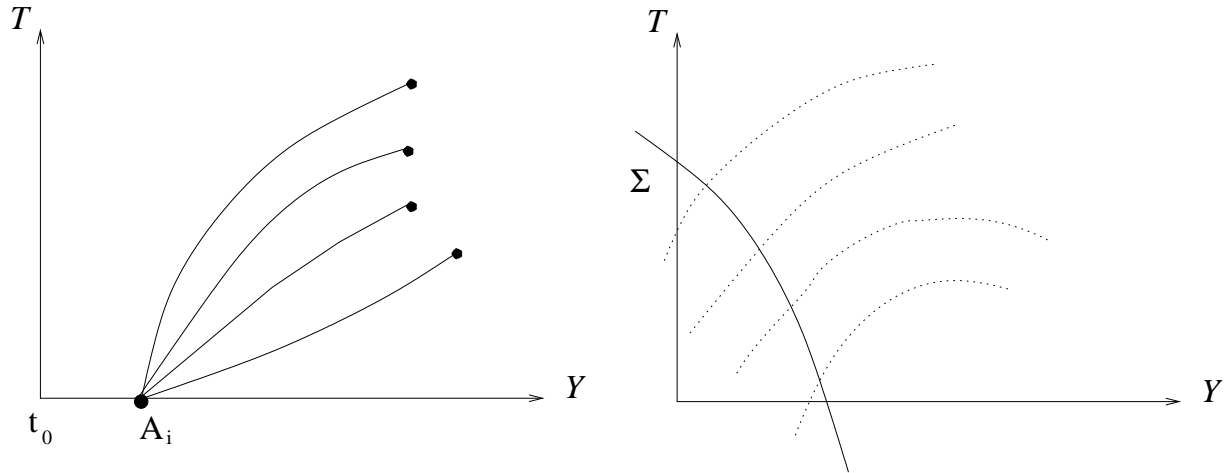


Fig. 6.10

(t_0, A_1, \dots, A_n) y con normal en ese punto el vector $(1, -\alpha_1, \dots, -\alpha_n)$. Tenemos pues la libertad de escoger (prácticamente arbitrariamente) la superficie inicial como parte de una familia de superficies $t = t_0(y_{10}, \dots, y_{n0})$ parametrizada por los parámetros s_1, \dots, s_n . Para cada una de ellas, es decir para cada (s_1, \dots, s_n) , obtendremos una superficie $S(T, Y_1, \dots, Y_n, s_1, \dots, s_n)$, solución de la ecuación de Hamilton-Jacobi, en una vecindad del punto inicial (t_0, A_1, \dots, A_n) .

Nota 6.2: Hemos cambiado aquí la parametrización de cada superficie de la notación (s_1, \dots, s_n) a (y_{10}, \dots, y_{n0}) y ahora (s_1, \dots, s_n) representa la parametrización de la familia de superficies iniciales.

Lemma 6.2: Si $t_0(y_{10}, \dots, y_{n0}, s_1, \dots, s_n)$ es C^2 entonces $S(T, Y_1, \dots, Y_n, s_1, \dots, s_n)$ es C^2 y $\frac{\partial S}{\partial s_i} = \text{constante sobre los extremales } Y_j(\eta, s_1, \dots, s_n)$.

Prueba. La diferenciabilidad viene de la regularidad de las soluciones de sistemas de ecuaciones ordinarias con respecto a los parámetros (ver cualquier libro de ecuaciones ordinarias). Tenemos además:

$$\frac{d}{d\eta} \left(\frac{\partial S}{\partial s_i} \right) = \frac{d}{dT} \left(\frac{\partial S}{\partial s_i} \right) (T, Y_1(T), \dots, Y_n(T), s_1, \dots, s_n) = \frac{\partial^2 S}{\partial s_i \partial T} + \sum_j \frac{\partial^2 S}{\partial s_i \partial Y_j} H_{p_j}$$

(usando $Y_j'(\eta) = H_{p_j}$).

Pero como tenemos a la ecuación de Hamilton-Jacobi:

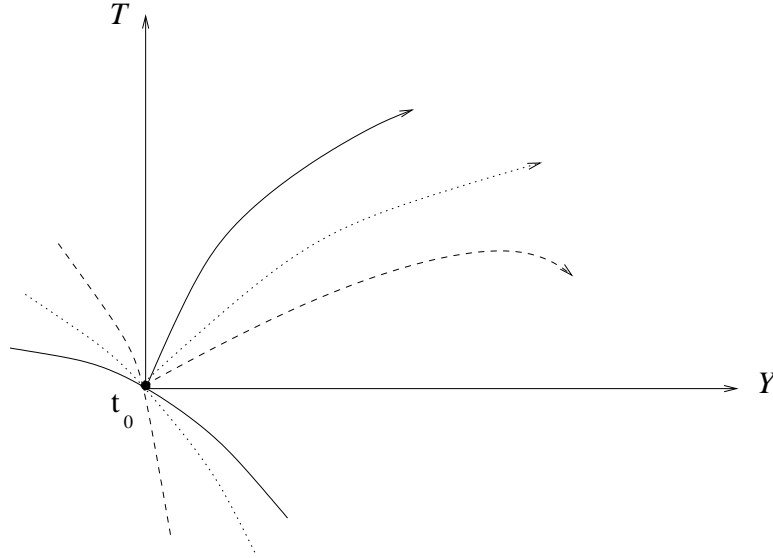


Fig. 6.11

$$\frac{\partial S}{\partial T} + H\left(T, Y_1, \dots, Y_n, \frac{\partial S}{\partial Y_1}, \dots, \frac{\partial S}{\partial Y_n}\right) = 0$$

para todo (s_1, \dots, s_n) , podemos diferenciar:

$$\frac{\partial^2 S}{\partial s_i \partial T} + \sum_j H_{p_j} \frac{\partial^2 S}{\partial s_i \partial Y_j} = 0$$

es decir

$$\frac{d}{d\eta} \left(\frac{\partial S}{\partial s_i} \right) = 0.$$

Q.E.D.

Para poder construir de manera consistente todas las trayectorias que salen de (t_0, A_1, \dots, A_n) , usaremos el siguiente resultado:

Proposición 6.2 (Jacobi): Supongamos que $\left(\frac{\partial^2 S}{\partial s_i \partial Y_j}\right)_{ij}$ sea invertible en $(T_0, Y_{10}, \dots, Y_{n0}, s_{10}, \dots, s_{n0})$. Sean $r_i = -\frac{\partial S}{\partial s_i}(T_0, Y_{10}, \dots, s_{n0})$ entonces:

1. Las ecuaciones $r_i = -\frac{\partial S}{\partial s_i}(T, Y_1, \dots, Y_n, s_1, \dots, s_n)$ tiene una única solución $Y_i(T, s_1, \dots, s_n)$ en una vecindad de ese punto.

2. Si

$$p_i(T) = \frac{\partial S}{\partial Y_i}(T, Y_1(T, s_1, \dots, s_n), \dots, s_1, \dots, s_n)$$

entonces:

$$Y'_i(T) = H_{p_i}, \quad p'_i(T) = -H_{Y_i}.$$

3. Si dejamos variar r_i, s_i , entonces la transformación $(Y, P) \rightarrow (S, R)$ es canónica, con función generatriz $S(T, Y_1, \dots, Y_n, s_1, \dots, s_n)$ y ecuaciones canónicas:

$$s'_i(T) = 0, \quad r'_i(T) = 0.$$

Prueba: Notemos primero que la condición de invertibilidad quiere decir que las normales $\tilde{P}(s_1, \dots, s_n)$ son linealmente independientes en el punto indicado. Esta condición permite usar el teorema de la función implícita para expresar Y_i en función de las otras variables (y de r_1, \dots, r_n si se permite su variación).

Derivando la relación $r_i + \frac{\partial S}{\partial s_i}(T, Y_1(T), \dots, Y_n(T)) = 0$ con respecto a T , obtenemos:

$$0 = \frac{d}{dT} \frac{\partial S}{\partial s_i} = S_{Ts_i} + \sum_j S_{s_i Y_j} Y'_j(T).$$

Pero de $S_T + H(T, Y_1, \dots, Y_n, S_{Y_1}, \dots, S_{Y_n}) = 0$, donde S depende de s_1, \dots, s_n , tenemos:

$$S_{Ts_i} + \sum_j H_{p_j} S_{s_i Y_j} = 0$$

y, por la invertibilidad: $Y'_j(T) = H_{p_j}$.

Del mismo modo:

$$\frac{dp_j}{dT} = \frac{d}{dT} \frac{\partial S}{\partial Y_j}(T, Y_1(T), \dots, Y_n(T)) = S_{TY_j} + \sum_i S_{Y_i Y_j} Y'_i(T).$$

Pero $S_{TY_j} + \sum_i S_{Y_i Y_j} H_{p_i} + H_{Y_j} = 0$.

Por lo tanto, usando la primera ecuación canónica:

$$\frac{dp_j}{dT} = -H_{Y_j}.$$

Haciendo variar s_i y r_i obtendremos $2n$ soluciones a las ecuaciones canónicas. El hecho que la transformación es canónica fue visto en la Nota 4.7.

Q.E.D.

Corolario 6.1: Si $(F_{z_i z_j})_{ij}$ es invertible y, en (t_0, A_1, \dots, A_n) , la matriz $(\partial p_i / \partial s_j)_{ij}$ o la matriz $(\partial z_i / \partial s_j)_{ij}$ es invertible, entonces, en una vecindad de (t_0, A_1, \dots, A_n) la función

distancia geodésica (ver la Definición 6.1) está bien definida. Es decir, para cada punto (T, Y_1, \dots, Y_n) vecino, existe una y una sola solución de las ecuaciones de Euler, con $y_i(T) = Y_i, y_i(t_0) = A_i$.

Prueba: Notemos primero que como $p_i = \frac{\partial S}{\partial Y_i}$, si $\left(\frac{\partial p_i}{\partial s_j}\right)_{ij}$ es invertible en (t_0, A_1, \dots, A_n) entonces las hipótesis de la Proposición 6.2 se cumplen. Por otra parte:

$$p_i = F_{z_i}, \quad \frac{\partial p_i}{\partial s_j} = \sum_k F_{z_i z_k} \frac{\partial z_k}{\partial s_j}.$$

Como $(F_{z_i z_j})_{ij}$ es invertible, la matriz $\left(\frac{\partial p_i}{\partial s_j}\right)_{ij}$ es invertible si y solo si la matriz $\left(\frac{\partial z_k}{\partial s_j}\right)_{kj}$ lo es. (Como función de s_1, \dots, s_n en (t_0, A_1, \dots, A_n) solo p_i y z_i varían).

Ahora bien, $S(s_1, \dots, s_n, t_0, A_1, \dots, A_n) = 0$ por construcción, por lo tanto, para las trayectorias que salen de (t_0, A_1, \dots, A_n) tenemos que $r_i = 0$, ya que eso valen en (t_0, A_1, \dots, A_n) y después se usa el Lema 6.2; son constantes sobre las trayectorias.

Para ese valor de r_i , por la Proposición 6.2, $Y_i(T, s_1, \dots, s_n)$ y $p_i(T, s_1, \dots, s_n)$ son soluciones de las ecuaciones canónicas, es decir, ya que $(F_{z_i z_j})_{ij}$ es invertible, $Y_i(T, s_1, \dots, s_n)$ es solución de las ecuaciones de Euler.

Fijemos (s_1, \dots, s_n) , entonces por $(T, Y_i(T, s_1, \dots, s_n))$ pasan dos soluciones de las ecuaciones de Euler: la dada por la Proposición 6.1 (con $r_i = 0$) y la trayectoria $y_i(\eta)$ transversal a la superficie $\Sigma(s_1, \dots, s_n)$ (también con $r_i = 0$). Estas dos trayectorias coinciden, ya que de lo contrario, tendríamos para el mismo (T, s_1, \dots, s_n) dos puntos $(Y_1(T), \dots, Y_n(T))$ y $(y_1(\eta), \dots, y_n(\eta))$ vecinos pero distintos con el mismo valor de r_i , contradiciendo (1) de la Proposición 6.2.

Del mismo modo el punto de intersección de esa trayectoria con $\Sigma(s_1, \dots, s_n)$ debe ser en (t_0, A_1, \dots, A_n) . Si no, a partir de (t_0, A_1, \dots, A_n) tendríamos una trayectoria vecina, con $r_i = 0$, que debe llegar, en T , en un punto diferente de (Y_1, \dots, Y_n) por la unicidad del campo de extremales asociado a $\Sigma(s_1, \dots, s_n)$. Esto contradice de nuevo (1) de la Proposición 6.2.

Resta únicamente probar que por cada punto pasa una y sola una trayectoria. Para esto bastaría tener que, a partir de la relación

$$0 = \frac{\partial S}{\partial s_i}(T, Y_1, \dots, Y_n, s_1, \dots, s_n)$$

podríamos escribir $Y_i = Y_i(T, s_1, \dots, s_n)$, para cada T y (Y_1, \dots, Y_n) es decir, que en el teorema de la función implícita sepamos que podemos tener $\left(\frac{\partial Y_i}{\partial s_j}\right)_{ij}$ invertible.

A partir de

$$\frac{\partial S}{\partial s_i}(T, Y_1(s_1, \dots, s_n), \dots, Y_n(s_1, \dots, s_n), s_1, \dots, s_n) = 0$$

tenemos

$$\frac{\partial^2 S}{\partial s_i \partial s_j} + \sum_k \frac{\partial^2 S}{\partial s_i \partial s_{Y_k}} \frac{\partial Y_k}{\partial s_j} = 0$$

entonces $\left(\frac{\partial Y_i}{\partial s_j}\right)_{ij}$ es invertible si y solo si $\left(\frac{\partial^2 S}{\partial s_i \partial s_j}\right)_{ij}$ lo es.

Esto es una consecuencia del siguiente resultado:

Lema 6.3: Sea $(F_{z_i z_j})_{ij}$ invertible. Sea $y_i(t, s_1, \dots, s_n)$ una familia de soluciones de las ecuaciones de Euler tales que: $y_i(t_0, s_1, \dots, s_n) = A_i$, $\left(\frac{\partial y'_i}{\partial s_j}(t_0, 0, \dots, 0)\right)_{ij}$ es invertible. Entonces, en una vecindad de $y_i(t, 0, \dots, 0)$, para $t \neq t_0$, $t - t_0$ pequeño, pasa uno y solo un elemento de la familia y $\left(\frac{\partial y_i}{\partial s_j}\right)_{ij}$ es invertible para $t \neq t_0$.

Prueba: Podemos escribir las ecuaciones de Euler en la forma:

$$\begin{aligned} Y'(t) &= Z(t) \\ Z'(t) &= (F_{ZZ})^{-1}(F_Y - F_{tZ} - F_{ZY}Z) \equiv G(t, Y, Z) \end{aligned}$$

ya que

$$\frac{d}{dt}(F_{z_i}) = F_{tz_i} + \sum_j F_{y_j z_i} y'_j + \sum_j F_{z_i z_j} y''_j = F_{y_i}$$

(como $(F_{z_i z_j})_{ij}$ es invertible las soluciones son C^2 : Ejercicio IV.3.2).

Aquí $Y = (y_1, \dots, y_n)^T$, $Z = (z_1, \dots, z_n)^T$, F_{ZZ} es la matriz $(F_{z_i z_j})_{ij}$.

Como estas ecuaciones son válidas para cada (s_1, \dots, s_n) , podemos diferenciarlas con respecto a $(s_1, \dots, s_n) = S$ y evaluarlas en $S = 0$:

$$Y'_S(t) = Z_S(t)$$

$$Z'_S(t) = G_Y Y_S + G_Z Z_S$$

(aquí Y_S, Z_S son matrices), con

$$\begin{aligned} Y_S(t_0) &= 0 \text{ ya que } Y(t_0, s_1, \dots, s_n) = (A_1, \dots, A_n)^T \\ Y'_S(t_0) &= Z_S(t_0) \text{ es invertible.} \end{aligned}$$

Como $G_Y(t, Y(t, 0, \dots, 0), Z(t, 0, \dots, 0))$, $G_Z(t, Y(t, 0, \dots, 0), Z(t, 0, \dots, 0))$ son funciones de t solamente, tenemos un sistema de ecuaciones lineales en Y_S y Z_S , con solución (formalmente):

$$\begin{aligned} Y_S(t) &= Y_S(t_0) + \int_{t_0}^t Z_S(\tau) d\tau \\ Z_S(t) &= Z_S(t_0) + \int_{t_0}^t (G_Y Y_S + G_Z Z_S) d\tau \end{aligned}$$

es decir:

$$Y_S(t) = (t - t_0)Z_S(t_0) + \int_{t_0}^t \int_{t_0}^{\tau} (G_Y Y_S + G_Z Z_S) d\tau_1 d\tau.$$

Notemos que $G_Y Y_S + G_Z Z_S$ no crece más rápido que una exponencial $Me^{N(t-t_0)}$, por lo tanto la doble integral es acotada por: $M(e^{N(t-t_0)} - 1 - N(t-t_0))/N^2$, es de orden $(t-t_0)^2$, para $t-t_0$ pequeño. (El sistema se puede escribir como:

$$U' = A(t)U.$$

Multiplicando cada columna de U' por las columnas correspondientes de U , obtenemos

$$\frac{d}{dt}\|U\|^2 = 2A(t)U \cdot U \leq N\|U\|^2$$

o sea

$$(e^{-N(t-t_0)}\|U\|^2)' \leq 0$$

integrando

$$e^{-N(t-t_0)}\|U(t)\|^2 \leq \|U(t_0)\|^2 = M).$$

Por lo tanto, para $t-t_0$ pequeño, $t \neq t_0$, Y_S es invertible (como $Z_s(t_0)$). Dado que Y_S es invertible, la relación:

$$Y(t, s_1, \dots, s_n) = Y$$

tiene una única solución para Y en una vecindad de $Y(t, 0, \dots, 0)$.

Q.E.D.

Nota 6.3: Dada cualquier matriz invertible C , con columnas C_1, \dots, C_n , sea $Y(t, s_1, \dots, s_n)$ la única solución (para $t-t_0$ pequeño) del problema de valores iniciales:

$$\begin{aligned} Y'(t) &= Z(t) & Y(t_0) &= (A_1, \dots, A_n)^T \\ Z'(t) &= G(t, Y, Z) & Z(t_0) &= s_1 C_1 + \dots + s_n C_n + C_0 (C_0 \text{ cualquier vector}) \end{aligned}$$

En este caso $\left(\frac{\partial Y'_i(t_0)}{\partial s_j}\right)_{ij} = C$. Podemos aplicar el Lema 6.3.

Definición 6.2: Una familia $y_j(t, s_1, \dots, s_n)$ de soluciones a las ecuaciones de Euler, con $y_j(t_0, s_1, \dots, s_n) = A_j$, $(\partial y'_j / \partial s_i)_{ij}$ invertible en $(t_0, 0, \dots, 0)$, $(\frac{\partial y_i}{\partial s_i})_{ij}$ invertible en $(t, 0, \dots, 0)$, se llama un **campo central, con foco en** (t_0, A_1, \dots, A_n) .

En ese caso el problema de valores iniciales:

$$\begin{aligned} Y'(t) &= Z(t) & Y(t_0) &= (A_1, \dots, A_n)^T \\ Z'(t) &= G(t, Y, Z) & Z(t_0) &= Y'(t_0, s_1, \dots, s_n) \end{aligned}$$

tiene una única solución para cada (s_1, \dots, s_n) , y podemos aplicar el Lema 6.3. Como además la relación

$$y_j(T, s_1, \dots, s_n) = Y_j$$

tiene una única solución (por el teorema de la función implícita) para (Y_1, \dots, Y_n) vecino a $(y_1(T, 0, \dots, 0), \dots, y_n(T, 0, \dots, 0))$ entonces, por cada (T, Y_1, \dots, Y_n) con $T \neq t_0$, tenemos uno y un solo elemento de la familia pasando por ese punto.

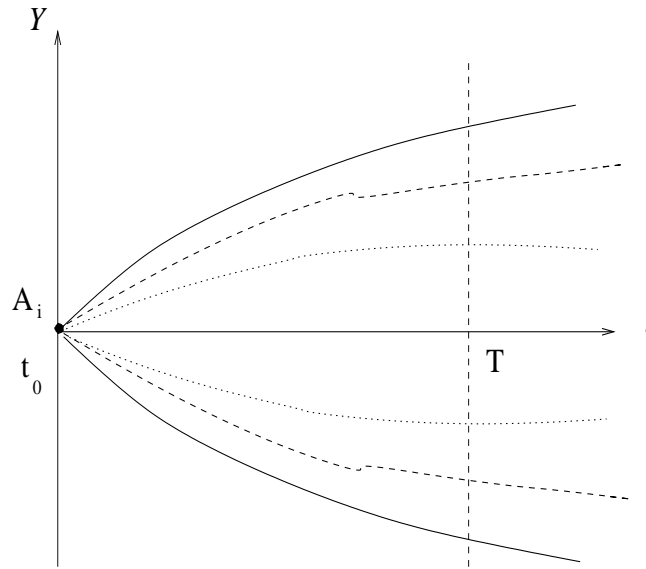


Fig. 6.12

Por la Proposición 6.1, en esa vecindad la función distancia geodésica S_0 está bien definida y satisface la ecuación de Hamilton-Jacobi. Notaremos que esto es independiente del Corolario 6.1, Proposición 6.2 y lo anterior.

Notemos que cerca de t_0 , solo se necesita la invertibilidad de $\frac{\partial y'_i}{\partial s_i}(t_0, 0, \dots, 0)$, o, como $P_S = F_{ZZ}Y_S$, de $P_S(t, 0, \dots, 0)$. Como ya se dijo al principio de esta subsección, a partir de $P_S(t_0, s_1, \dots, s_n)$ se construye $\Sigma(s_1, \dots, s_n)$ con esa normal en (t_0, A_1, \dots, A_n) . La función geodésica resultante $S(T, Y_1, \dots, Y_n, s_1, \dots, s_n)$ coincide con

$$S_0(T, Y_1, \dots, Y_n, s_1, \dots, s_n) = \int_{t_0}^T F(t, y_1(t, s_1, \dots, s_n), \dots, y'_n(t, s_1, \dots, s_n)) dt$$

para la única solución de las ecuaciones de Euler, (garantizada por el Lema 6.3), pasando por (t_0, A_1, \dots, A_n) y $(T, y_1(T, s_1, \dots, s_n), \dots, y_n(T, s_1, \dots, s_n))$: esto es debido al Corolario 6.1 y a la unicidad. Podemos entonces usar la Proposición 6.2 para generar (en teoría) la transformación canónica.

Nota 6.4: A partir de un campo central, podemos dar un campo de extremales, es decir, transversales a una superficie, de la manera siguiente:

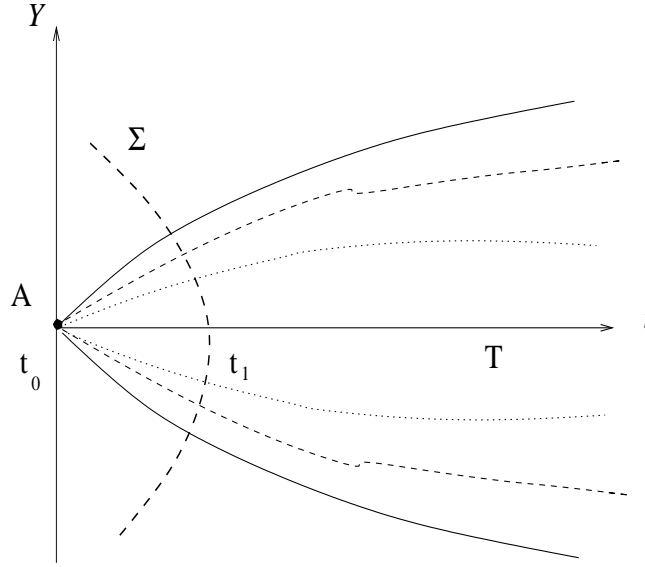


Fig. 6.13

Escojamos (t_1, y_1, \dots, y_n) sobre la curva $y_i(t, 0, \dots, 0)$ tal que

$$H(t_1, y_1(t_1, 0, \dots, 0), \dots, p_1(t_1, 0, \dots, 0), \dots, p_n(t_1, 0, \dots, 0)) \neq 0.$$

Consideremos la relación $S_0(T, Y_1, \dots, Y_n) = S_0(t_1, y_1, \dots, y_n) = \text{constante}$.

S_0 es la función distancia geodésica (a partir de t_0, A_1, \dots, A_n
con $S_{0T} = -H \quad S_{0Y_i} = p_i$.

Por el teorema de la función implícita, ya que $S_{0T}(t_1, y_1, \dots, y_n) \neq 0$, obtenemos una única superficie $T = T(Y_1, \dots, Y_n)$ satisfaciendo la relación, con:

$$\frac{\partial S_0}{\partial T} \frac{\partial T}{\partial Y_i} + \frac{\partial S_0}{\partial Y_i} = 0$$

es decir:

$$\frac{\partial T}{\partial Y_i} = p_i/H$$

o sea la superficie es transversal al campo.

Notemos que si $H(T, y_1(T, s_1, \dots, s_n), \dots, p_n(T, s_1, \dots, s_n)) \equiv 0$, entonces $S_{0T} = 0$, S_0 es constante sobre cada elemento de la familia (y vale 0 en t_0, A_1, \dots, A_n), por lo tanto $S_0 \equiv 0$, $p_i = 0$, lo cual contradice la invertibilidad de $(\partial p_i / \partial s_j)_{ij}$ en la definición del campo central.

En términos físicos hemos reemplazado el foco, o la fuente, por una fuente distribuida sobre la superficie alcanzada en el tiempo $t_1(y_1, \dots, y_n)$.

5.6.6 Campos exactos

Los campos centrales son un ejemplo de una clase más grande de campos, los exactos. Necesitaremos esta categoría de campos en el estudio de las condiciones suficientes.

Recordemos que un **campo** en \mathbb{R}^n es una aplicación de $\mathbb{R} \times \mathbb{R}^n$ a $\mathbb{R}^n : Z(T, Y)$, continuamente diferenciable. Entonces el sistema:

$$Y'(T) = Z(T, Y) \quad , Y(t_0) = Y_0$$

tiene una y una sola solución $Y(T, t_0, Y_0)$, con $Y(t_0, t_0, Y_0) = Y_0$. $Y(T, t_0, Y_0)$ es C^1 con respecto a los datos.

Definición 6.3: Diremos que el campo $Z(T, Y)$ es **exacto** para $F(T, Y, Z)$ si la diferencial:

$$(F(T, Y, Z(T, Y)) - F_Z(T, Y, Z(T, Y)) \cdot Z(T, Y)) dT + F_Z(T, Y, Z(T, Y)) \cdot dY$$

es exacta. Es decir, existe una función S tal que:

$$\begin{aligned} S_T &= F - P \cdot Z \equiv -H, \quad \text{con} \quad p_i = F_{z_i}(T, y_1, \dots, y_n, z_1, \dots, z_n) \\ S_Y &= F_Z \quad \text{es decir,} \quad S_{y_i} = F_{z_i} = p_i. \end{aligned}$$

Vemos por lo tanto que si $Z(T, Y)$ es un campo exacto entonces la función S es solución de la ecuación de Hamilton-Jacobi. Inversamente si S es solución de la ecuación de Hamilton-Jacobi, entonces el campo $Z(T, Y)$ definido al invertir la relación:

$$S_Y = F_Z(T, Y_1, \dots, Y_n, Z_1, \dots, Z_n) = S_Y(T, Y_1, \dots, Y_n)$$

para Z como función de (T, Y_1, \dots, Y_n) , (esto requiere que F_{ZZ} sea invertible), es un campo exacto ($dS = S_T dT + S_Y \cdot dY$).

Sabemos que, en un dominio simplemente conexo (sin hoyos), una condición necesaria y suficiente para que una diferencial sea exacta, es que las derivadas cruzadas sean iguales:

$$S_{TY_i} = S_{Y_iT}, \quad S_{Y_iY_j} = S_{Y_jY_i}$$

es decir

$$- \frac{\partial}{\partial Y_i} (H(T, Y_1, \dots, Y_n, p_1(T, Y_1, \dots, Y_n), \dots, p_n(T, Y_1, \dots, Y_n))) = \frac{\partial p_i}{\partial T}$$

$$\frac{\partial p_i}{\partial Y_j} = \frac{\partial p_j}{\partial Y_i}.$$

La primera relación se escribe:

$$-H_{Y_i} - \sum_j H_{p_j} \frac{\partial p_j}{\partial Y_i} = \frac{dp_i}{dT} - \sum_j \frac{\partial p_i}{\partial Y_j} \frac{dY_j}{dT}$$

como $H_{p_j} = Z_j = \frac{dY_j}{dT}$ (por definición del campo).

$$\frac{dp_i}{dT} + H_{Y_i} = - \sum_j \left(\frac{\partial p_j}{\partial Y_i} - \frac{\partial p_i}{\partial Y_j} \right) \frac{dY_j}{dT}.$$

Veremos entonces que si el campo es exacto, entonces

$$\frac{dp_i}{dT} = -H_{Y_i}$$

es decir, las ecuaciones canónicas son válidas (esto ya lo sabíamos a partir de la equivalencia de la ecuación de Hamilton-Jacobi con el sistema de ecuaciones canónicas).

Inversamente para que una familia de soluciones de las ecuaciones canónicas (que forman un campo a partir de las ecuaciones $Y'_i(T) = H_{p_i}$) sea un campo exacto, es necesario que:

$$\frac{\partial p_i}{\partial Y_j} = \frac{\partial p_j}{\partial Y_i} \quad (\text{Siempre cierto si } n = 1).$$

Por ejemplo, si la familia está dada en la forma

$$Y_i(T, s_1, \dots, s_n), \quad p_i(T, s_1, \dots, s_n), \quad \text{con } \left(\frac{\partial Y_i}{\partial s_j} \right)_{ij} \text{ invertible}$$

entonces:

$$\frac{\partial p_i}{\partial Y_j} = \sum_k \frac{\partial p_i}{\partial s_k} \frac{\partial s_k}{\partial Y_j}$$

es decir, en forma matricial

$$P_Y = P_S S_Y = P_Y^T = S_Y^T P_S^T.$$

Por lo tanto: $P_S Y_S^{-1} = Y_S^{-1T} P_S^T$ o sea:

$$Y_S^T P_S - P_S^T Y_S = 0 \quad (\text{relación de Lagrange}).$$

Notemos que un campo central es exacto, por el Corolario 6.1, ya que la función distancia geodésica está bien definida.

Hilbert introdujo la noción de campos exactos en el Cálculo de Variaciones con la siguiente idea:

$$I(\gamma) = \int_{\gamma} F(T, Y, Z(T, Y)) ds$$

es independiente del camino γ escogido y depende solo de los extremos del camino. Esto es cierto si y solo si $F ds = dS$.

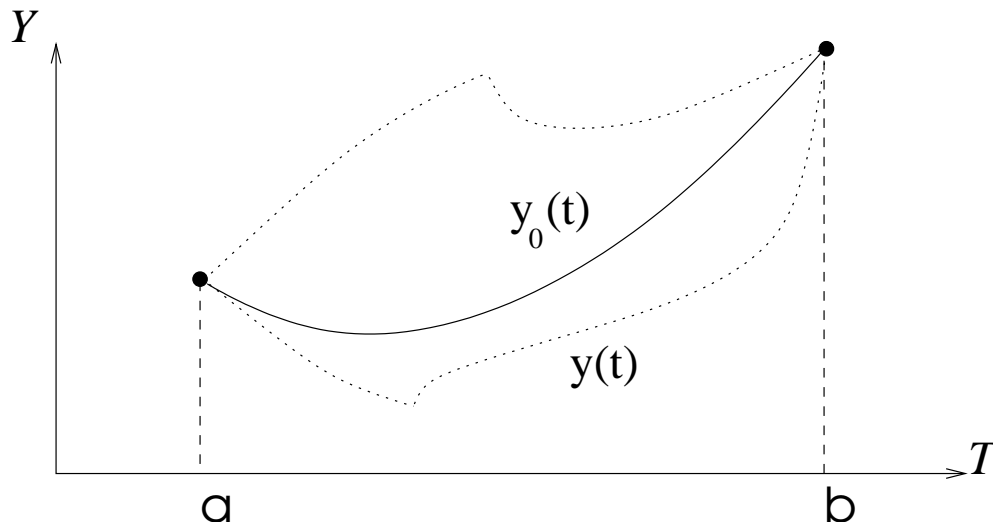


Fig. 6.14

Esta integral se llama la **integral invariante** de Hilbert.

Si γ es un camino, $Y(t)$, C^1 por pedazos, entonces:

$$\begin{aligned} I(\gamma) &= \int_{\gamma} (-H dT + P \cdot dY) = \int_a^b (-H + P \cdot Y'(T)) dT \\ &= \int_a^b F(T, Y(T), Z(T, Y(T))) + P(T, Y(T)) \cdot (Y'(T) - Z(T, Y(T))) dT \end{aligned}$$

por definición de H .

Pero si γ_0 es un camino $Y_0(T)$ que es parte de las trayectorias del campo, es decir $Y'_0(T) = Z(T, Y_0(T))$, entonces:

$$I(\gamma_0) = \int_a^b F(T, Y_0(T), Y'_0(T)) dT.$$

Si γ y γ_0 tienen las mismas extremidades, $I(\gamma_0) = I(\gamma)$ y:

$$\begin{aligned} \int_a^b F(T, Y(T), Y'(T)) dT & - \int_a^b [F(T, Y_0(T), Y'_0(T)) dT \\ & = \int_a^b [F(T, Y(T), Y'(T)) - F(T, Y(T), Z(T, Y(T)) \\ & - F_Z(T, Y(T), Z(T, Y(T)) \cdot (Y'(T) - Z(T, Y(T))] dT \end{aligned}$$

Esta es la fórmula de Hilbert, y la usaremos para estudiar los mínimos y máximos del funcional $\int_a^b F(T, Y, Y') dT$.

Nota 6.5: En todo este capítulo hemos supuesto que F_{ZZ} es invertible. Sin embargo, hay casos donde F_{ZZ} es siempre singular: Por ejemplo si F es homogéneo de grado uno en Z , entonces F_Z es homogéneo de grado 0 en Z , y la relación de Euler:

$$0 = \sum_j F_{z_i z_j} z_j$$

implica que Z es un vector propio de F_{ZZ} , con valor propio 0: en este caso F_{ZZ} es siempre singular. (Ver la Nota IV.3.1 y §V.1).

Supongamos que F no depende de T y es homogéneo de grado 1 en Z , como en §V.1, podemos cambiar de variables y escribir:

$$\int_a^b F(y_1, \dots, y_n, y'_1, \dots, y'_n) dt = \int_{y_n(a)}^{y_n(b)} F(y_1, \dots, y_{n-1}, y_n, \frac{dy_1}{dy_n}, \dots, \frac{dy_{n-1}}{dy_n}, 1) dy_n.$$

Si suponemos que

$$(F_{z_i z_j}(y_1, \dots, y_n, z_1, \dots, z_{n-1}, 1))_{i,j=0,\dots,n-1}$$

es invertible, entonces podemos aplicar todo lo visto en este capítulo.

En particular la ecuación de Hamilton-Jacobi se escribe:

$$S_{y_n} + H = 0 = S_{y_n} + \sum_1^{n-1} (p_i \frac{dy_i}{dy_n}) - F(y_1, \dots, y_n, z_1, \dots, z_{n-1}, 1)$$

con $S_{y_i} = p_i$, $i = 1, \dots, n-1$.

Pero de la identidad de Euler:

$$\sum_1^n F_{z_i} z_i = F = \sum_1^{n-1} p_i z_i + p_n z_n, \text{ con } z_i = \frac{dy_i}{dt}$$

tenemos

$$\begin{aligned}
F(y_1, \dots, y_n, y'_1, \dots, y'_n) &= \left(\sum_1^{n-1} p_i \frac{dy_i}{dy_n} + p_n \right) \frac{dy_n}{dt} \\
&= F \left(y_1, \dots, y_n, \frac{dy_1}{dy_n}, \dots, \frac{dy_{n-1}}{dy_n}, 1 \right) \frac{dy_n}{dt}
\end{aligned}$$

(por la homogeneidad de F).

La ecuación de Hamilton-Jacobi se reduce entonces a:

$$S_{y_n} = p_n, \quad S_{y_i} = p_i.$$

Lo cual es de esperarse, ya que para el sistema con y_1, \dots, y_n , tendríamos $S_T + H(T, Y_1, \dots, Y_n, p_1, \dots, p_n) = 0$, $S_{Y_i} = p_i$.

Pero $H(T, Y_1, \dots, Y_n, p_1, \dots, p_n) = \sum_1^n p_i z_i - F = \sum_1^n F_{z_i} z_i - F = 0$

(esta es la relación de Euler).

Capítulo 6

Condiciones suficientes

En este último capítulo veremos como probar que el punto crítico obtenido es un extremo. Según la norma usada en el problema, podemos usar argumentos similares a los del Cálculo sobre el signo de la segunda derivada o argumentos de convexidad global. De hecho consideremos el funcional:

$$J(y) = \int_a^b F(x, y, y') dx$$

definido para funciones C^1 en $[a, b]$ con $y(a) = A$, $y(b) = B$. Sea y_0 un punto crítico de $J(y)$, es decir una solución a la ecuación de Euler, con los requerimientos de regularidad del Capítulo IV. Para que y_0 sea un mínimo local, necesitamos probar que $J(y) - J(y_0) \geq 0$ para y vecino a y_0 .

Ahora bien, según la norma que pongamos al espacio de trabajo, tenemos varias posibilidades para definir la noción de vecindad. Una primera posibilidad sería la de usar la norma natural de C^1 , es decir $\|\cdot\|_1$: la literatura clásica llama a esta clase de extremos, extremos débiles. Otra posibilidad es la de usar en el espacio de trabajo (funciones C^1) la norma $\|\cdot\|_0$: la literatura clásica llama a esta clase de extremos, los extremos fuertes. Esta nomenclatura, sin embargo, es confusa ya que podríamos pensar que, siendo la topología C^1 más fuerte que la topología C^0 , los nombres deberían ser al revés. Por lo tanto, no la usaremos y veremos extremos en la norma C^1 o en la norma C^0 .

Ahora bien, si F es C^2 y y_0 es un punto crítico, podemos escribir $y = y_0 + h$, para todo y en el espacio de trabajo con $h(a) = h(b) = 0$ y,

$$\begin{aligned} J(y) - J(y_0) &= \int_a^b (F_y(x, y_0, y'_0)h + F_{y'}h') dx \\ &+ \frac{1}{2} \int_a^b (F_{yy}(x, y_0, y'_0)h^2 + 2F_{yy'}hh' + F_{y'y'}h'^2) dx \\ &+ \frac{1}{2} \int_a^b \left[(F_{yy}(x, y_0 + \theta h, y'_0 + \theta h') - F_{yy}(x, y_0, y'_0))h^2 \right. \\ &\quad \left. + 2(F_{yy'}(x, y_0 + \theta h, y'_0 + \theta h') - F_{yy'}(x, y_0, y'_0))hh' \right. \end{aligned}$$

$$+(F_{y'y'}(x, y_0 + \theta h, y'_0 + \theta h') - F_{y'y'}(x, y_0, y'_0))h'^2 \Big] dx$$

donde hemos usado para la función de una variable

$$f(t) = F(x, y_0 + th, y'_0 + th')$$

el desarrollo de Taylor a segundo orden:

$$f(1) - f(0) = f'(0) + f''(\theta)/2, \text{ con algún } \theta, 0 < \theta < 1.$$

Como y_0 es un punto crítico, los términos lineales (derivadas de Gâteaux) desaparecen y quedan únicamente los términos cuadráticos y un residuo, el cual es acotado (si F es C^2) por $0(|h|_1)|h|_1^2 = o(|h|_1^2)$. Por lo tanto si $|h|_1$ es pequeño, el término cuadrático dominará esta expresión y su signo dará el signo de $J(y) - J(y_0)$.

Notemos que si dividimos el intervalo $[a, b]$ en un número finito de intervalos donde y_0 y h son C^1 (de tal modo que el espacio de trabajo será el espacio de funciones continuas y C^1 por pedazos) entonces este argumento sigue válido. Esta situación se generaliza a cualquier funcional.

6.1 Extremos de funcionales

Sea $I(y)$ un funcional sobre un espacio normado X , con una derivada de Fréchet en y , para todo y en una vecindad de y_0 . Entonces para h pequeño, tal que $y_0 + h$ esté en esa vecindad, podemos usar el teorema fundamental del Cálculo y escribir:

$$I(y_0 + h) - I(y_0) = \int_0^1 DI(y_0 + th) h dt$$

ya que $\frac{d}{dt}(I(y_0 + th)) = DI(y_0 + th)h$.

Ahora bien, para cada h fijo, $DI(y_0 + th)h$ es un funcional en t , con t en X . Supongamos que $f(t) = DI(y_0 + th)h$ tiene una derivada en $t = 0$, que llamamos $D^2I(y_0)(h, h)$. Podemos entonces seguir el desarrollo de $DI(y_0 + th)h = DI(y_0)h + D^2I(y_0)(h, th) + o(th)$. Para que esto nos sea útil, es decir, para tener derivada de Fréchet, necesitamos que $D^2I(y_0)(h, th)$ sea lineal en t , (entonces la integral de t da $1/2$), y que el residuo sea de hecho de orden $o(\|h\|^2)$ uniformemente. Por lo tanto:

Definición 1.1 :

1) $I(y)$ tiene derivada segunda de Gâteaux, en y_0 , en las direcciones (h, k) , si la función de dos variables

$$g(t, \tau) = I(y_0 + \tau h + tk)$$

es C^2 en $(0, 0)$. $D_{(h,k)}^2 I(y_0) \equiv g_{\tau t}(0, 0)$

2) $I(y)$ tiene derivada segunda de Fréchet en y_0 , si

$$I(y_0 + h) = I(y_0) + DI(y_0)h + \frac{1}{2}D^2I(y_0)(h, h) + o(\|h\|^2)$$

donde $D^2I(y_0)(h, k) \equiv B(h, k)$ es un funcional bilineal simétrico, continuo, es decir:

1. $B(h, k) = B(k, h)$
2. $B(th, k) = tB(h, k)$
3. $B(h_1 + h_2, k) = B(h_1, k) + B(h_2, k)$
4. $|B(h, k)| \leq M\|h\| \|k\|$

Notemos que si $I(y)$ tiene segunda derivada de Fréchet, entonces:

$$\begin{aligned} g(t, \tau) &= g(0, 0) + DI(y_0)(\tau h + tk) + D^2I(y_0)(\tau h + tk, \tau h + tk)/2 + o(\|\tau h + tk\|^2) \\ &= g(0, 0) + g_\tau(0, 0)\tau + g_t(0, 0)t + D^2I(y_0)(h, h)\tau^2/2 + D^2I(y_0)(h, k)t\tau \\ &\quad + D^2I(y_0)(k, k)t^2/2 + o(\tau^2 + t^2). \end{aligned}$$

Por lo tanto $D_{(h,k)}^2 I(y_0) = g_{t\tau}(0, 0) = g_{\tau t}(0, 0)$.

Ejemplo 1.1 : Sea $J(y) = \int_a^b F(x, y(x), y'(x)) dx$. Si F es C^2 , entonces, por el desarrollo que hicimos en la introducción,

$$D^2J(y_0)(h, h) = \int_a^b (F_{yy}h^2 + 2F_{yy'}hh' + F_{y'y'}h'^2) dx$$

ya que el residuo es uniforme en una vecindad (en C^1) de y_0 . Así mismo

$g(t, \tau) = \int_a^b F(x, y_0 + th + \tau k, y'_0 + th' + \tau k') dx$ es C^2 , con

$$\begin{aligned} g_t(0, \tau) &= \int_a^b (F_y(x, y_0 + \tau k, y'_0 + \tau k')h + F_z(x, y_0 + \tau k, y'_0 + \tau k')h') dx \\ &= DJ(y_0 + \tau k)h \end{aligned}$$

y

$$\begin{aligned} g_{t\tau}(0, 0) &= \int_a^b (F_{yy}(x, y_0, y'_0)hk + F_{yz}(h'k + hk') + F_{zz}h'k') dx \\ &= D^2J(y_0)(h, k). \end{aligned}$$

Se ve claramente que esta expresión es bilineal, simétrica y continua. Además el residuo es acotado por $0(\|h\|_1) \int_a^b (h^2 + h'^2) dx$. Esto se generaliza sin problemas si trabajamos en el espacio de funciones C^1 por pedazos.

Ejemplo 1.2 : Si $Y = (y_1, \dots, y_n)^T$, si $F(x, Y, Y')$ es C^2 entonces $J(Y) = \int_a^b F(x, Y, Y') dx$ es tal que:

$$\begin{aligned} J(Y + H) - J(Y) &= \int_a^b (F_Y \cdot H + F_Z \cdot H') dx \\ &+ \frac{1}{2} \int_a^b (H^T F_{YY} H + H^T F_{YZ} H' + H'^T F_{ZY} H + H'^T F_{ZZ} H') dx \\ &+ o(\|H\|_1^2) \end{aligned}$$

donde F_{YY} es la matriz $(F_{y_i y_j})_{ij}$, $F_{YZ} = (F_{y_i z_j})_{ij} = F_{ZY}^T$, etc...

En efecto, si $g(t, \tau) = \int_a^b F(x, Y + tH + \tau K, Y' + tH' + \tau K') dx$, tenemos:

$$\begin{aligned} g_t(0, \tau) &= \int_a^b (F_Y(x, Y + \tau K, Y' + \tau K') \cdot H + F_Z(x, Y + \tau K, Y' + \tau K') \cdot H') dx \\ &= DJ(Y + \tau K)H \end{aligned}$$

y,

$$\begin{aligned} g_{t\tau}(0, 0) &= D^2 J(Y)(H, K) \\ &= \int_a^b \sum_{ij} (F_{y_i y_j} h_i k_j + F_{y_i z_j} h_i k'_j + F_{z_i y_j} h'_i k_j + F_{z_i z_j} h'_i k'_j) dx \\ &= \int_a^b (H^T F_{YY} K + H^T F_{YZ} K' + H'^T F_{ZY} K + H'^T F_{ZZ} K') dx. \end{aligned}$$

Notemos que $H^T F_{YZ} K' = K'^T F_{ZY} H$, ya que $F_{YZ}^T = F_{ZY}$, se ve así claramente la simetría de la segunda derivada. Esto se generaliza a funciones C^1 por pedazos.

Exactamente como en Cálculo, tenemos el siguiente resultado.

Teorema 1.1 : Sea y_0 un punto crítico de $I(y)$, funcional con segunda derivada de Fréchet en y_0 , entonces:

a) Si y_0 es un mínimo local (es decir $I(y) - I(y_0) \geq 0$, para $\|y - y_0\|$ pequeño), tenemos:

$$D^2 I(y_0)(h, h) \geq 0 \text{ para todo } h.$$

b) Si $D^2 I(y_0)(h, h) \geq \alpha \|h\|^2$, para todo h , entonces y_0 es un mínimo local.

Prueba : (a) Si en una dirección h , tenemos $D^2 I(y_0)(h, h) < 0$, con $DI(y_0)h = 0$, entonces

$$I(y_0 + th) - I(y_0) = t^2 D^2 I(y_0)(h, h)/2 + o(t^2)$$

es negativa para t suficientemente pequeño.

(b) Para h , con $\|h\|$ pequeño, tendremos que $I(y_0 + h) \geq I(y_0)$.

Q.E.D.

Notemos que si solamente tenemos como hipótesis:

$$D^2I(y_0)(h, h) > 0 \text{ para } h \neq 0$$

entonces podemos probar que para cada h existe un $t(h)$ tal que $I(y_0 + th) - I(y_0) \geq 0$, para $t \leq t(h)$. Sin embargo, en dimensión infinita, esto no es suficiente para generar una vecindad completa de y_0 (no se puede usar el argumento de compacidad del Cálculo de varias variables), a menos que se tengan estimaciones más precisas sobre $t(h)$.

Ejemplo 1.3 : Sea $J(u) = \iint_{\Omega} F(x, y, u, u_x, u_y) dx dy$, donde Ω es un dominio acotado del plano. Entonces al considerar $g(t, \tau) = J(u, tv + \tau w)$, con $v = w = 0$ sobre $\partial\Omega$, tenemos

$$g_t(0, \tau) = \iint_{\Omega} (F_u(x, y, u + \tau w, u_x + \tau w_x, u_y + \tau w_y)v + F_{u_x}v_x + F_{u_y}v_y) dx dy$$

$$\begin{aligned} g_{t\tau}(0, 0) &= D^2J(u)(v, w) \\ &= \iint_{\Omega} \left[F_{uu}vw + F_{uu_x}(vw_x + v_xw) + F_{uu_y}(vw_y + v_yw) \right. \\ &\quad \left. + F_{u_xu_x}v_xw_x + F_{u_xu_y}(v_xw_y + v_yw_x) + F_{u_yu_y}v_yw_y \right] dx dy \end{aligned}$$

Por lo tanto, si suponemos que u es C^2 y F es C^3

$$\begin{aligned} D^2J(u)(v, v) &= \iint_{\Omega} \left[\left(F_{uu} - \frac{\partial}{\partial x}(F_{uu_x}) - \frac{\partial}{\partial y}(F_{uu_y}) \right) v^2 + \right. \\ &\quad \left. + F_{u_xu_x}v_x^2 + 2F_{u_xu_y}v_xv_y + F_{u_yu_y}v_y^2 \right] dx dy \end{aligned}$$

donde usamos el teorema de la divergencia y el hecho que $v = 0$ sobre $\partial\Omega$. Si para u sabemos que $F_{uu}(x, y, u, u_x, u_y) - \frac{\partial}{\partial x}(F_{uu_x}) - \frac{\partial}{\partial y}(F_{uu_y}) \geq 0$ (por ejemplo si F no depende de u : en ese caso no se necesita que F sea C^3 o que u sea C^2), si además suponemos que la matriz

$$A \equiv \begin{pmatrix} F_{u_xu_x} & F_{u_xu_y} \\ F_{u_xu_y} & F_{u_yu_y} \end{pmatrix}$$

es definida positiva, entonces:

$$D^2J(u)(v, v) \geq \alpha \iint_{\Omega} (v_x^2 + v_y^2) dx dy.$$

Vemos que la segunda derivada no es fuertemente positiva con respecto a la norma $\|\cdot\|_1$, es decir, no satisface la hipótesis (b) del teorema 1.1, pero sí lo es con respecto a la norma $\|v\|_1^2 = \iint_{\Omega} (v_x^2 + v_y^2) dx dy$ sobre el espacio de funciones v en $C^1(\overline{\Omega})$, con $v = 0$ en $\partial\Omega$ (si

$\|v\|_1 = 0$ entonces $v = \text{constante} = 0$ en $\partial\Omega$). Se puede probar que entonces, si u es C^1 y es solución de la ecuación de Euler, tenemos un mínimo.

En efecto,

$$\begin{aligned} J(u+v) - J(u) &= D^2 J(u)(v, v)/2 \\ &+ \iint_{\Omega} (F_{uu}(x, y, u + \theta v, u_x + \theta v_x, u_y + \theta v_y) - F_{uu}(x, y, u, u_x, u_y)) v^2 dx dy / 2 \\ &+ \iint_{\Omega} ((\bar{F}_{uu_x} - F_{uu_x}) v v_x + (\bar{F}_{uu_y} - F_{uu_y}) v v_y + \nabla v^T (\bar{A} - A) \nabla v / 2) dx dy \end{aligned}$$

donde \bar{F} , \bar{A} corresponden a los argumentos $(x, y, u + \theta v, u_x + \theta v_x, u_y + \theta v_y)$. Si F es C^2 y Ω es acotado, vemos que el residuo es acotado por $O(\|v\|_1) \iint_{\Omega} (v^2 + |\nabla v|^2) dx dy$.

Ahora bien, como Ω es acotado, podemos ponerlo dentro de un rectángulo, y, como $v = 0$ en $\partial\Omega$, podemos extender v como 0 fuera de Ω y tener una función continua (ver figura 1.1)

Podemos escribir $v(x, y) - v(a, y) = \int_a^x v_x(\tau, y) d\tau$ (v_x es continuo por pedazos). Por lo tanto:

$$|v(x, y)| = \left| \int_a^x 1 \cdot v_x d\tau \right| \leq \left(\int_a^x d\tau \right)^{1/2} \left(\int_a^x v_x^2 d\tau \right)^{1/2}$$

(usando la desigualdad de Schwarz).

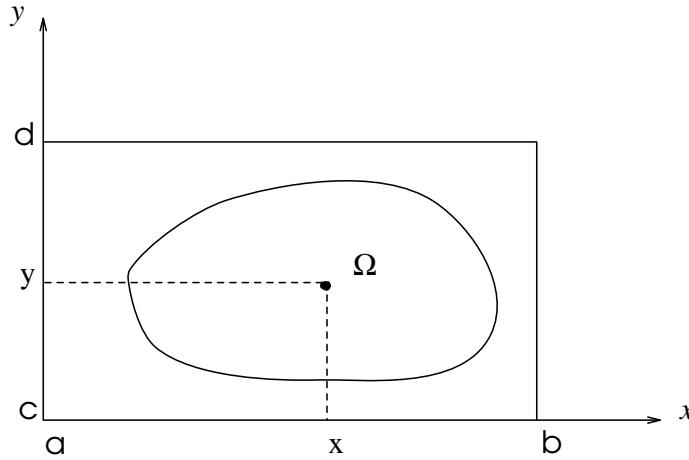


Fig. 1.1

$$\begin{aligned} \iint_{\Omega} v^2 dx dy &= \int_a^b \int_c^d v^2 dx dy \leq \int_a^b \int_c^d (x-a) \int_a^x v_x^2(\tau, y) d\tau dx dy \\ &\leq (b-a)^2 \int_c^d \int_a^b v_x^2 d\tau dy / 2 = (b-a)^2 \iint_{\Omega} v_x^2 dx dy / 2. \end{aligned}$$

Por lo tanto, existe una constante c (que no depende de v) tal que:

$$\iint_{\Omega} v^2 dx dy \leq c \iint_{\Omega} |\nabla v|^2 dx dy.$$

Esto se llama la *desigualdad de Poincaré*.

Entonces el residuo es acotado por $0(|v|_1) \int_{\Omega} \int |\nabla v|^2 dx dy$. Por lo tanto para v , tal que $|v|_1$ sea pequeño,

$$J(u+v) - J(u) \geq 0.$$

Veremos más adelante que si A tiene un valor propio negativo, entonces u **no** es un mínimo, independientemente del signo de $F_{uu} - \frac{\partial}{\partial x}(F_{uu_x}) - \frac{\partial}{\partial y}(F_{uu_y})$.

Para los problemas condicionados, (en el caso en que los gradientes de las condiciones son linealmente independientes) tenemos el siguiente resultado:

Teorema 1.2 : *Sea y_0 un punto crítico de $J(y)$ condicionado por $G_i(y) = 0$, $i = 1, \dots, n$, tal que:*

$$DJ(y_0)h + \sum_1^n \lambda_i DG_i(y_0)h = 0 \text{ para toda } h.$$

a) *Entonces si y_0 es un mínimo del funcional*

$$J(y) + \sum_1^n \lambda_i G_i(y)$$

también lo es de $J(y)$ condicionado por $G_i(y) = 0$. En particular, si G_i, J tienen segunda derivada de Fréchet en y_0 tal que

$$D^2(J + \sum_1^n \lambda_i G_i)(y_0)(h, h) \geq \alpha \|h\|^2$$

entonces J tiene un mínimo local condicionado en y_0 .

b) *Para toda curva diferenciable $y(t)$, con $y(0) = y_0$, y $G_i(y(t)) = 0$, entonces, si $J(y)$ tiene un mínimo local condicionado:*

$$D^2(J + \sum_1^n \lambda_i G_i)(y_0)(y'(0), y'(0)) \geq 0.$$

Prueba : (a) $J(y) + \sum \lambda_i G_i(y) - (J(y_0) + \sum \lambda_i G_i(y_0))$ es $J(y) - J(y_0)$ si $G_i(y) = G_i(y_0) = 0$. Poniendo $h = y - y_0$, el argumento del teorema 1.1. funciona.

(b) Si para una dirección $y'(0)$, la segunda derivada es negativa, entonces

$$\begin{aligned} J(y(t)) - J(y_0) &= D^2(J + \sum \lambda_i G_i)(y_0)(y(t) - y_0, y(t) - y_0)/2 \\ &\quad + o(\|y(t) - y_0\|^2) \end{aligned}$$

(ya que $G_i(y(t)) = 0$ y y_0 es un punto crítico de $J + \sum \lambda_i G_i$).

Para t pequeño, $y(t) - y(0) = y'(0)t + o(t)$

$$(J(y(t)) - J(y_0))/t^2 = D^2(J + \sum \lambda_i G_i)(y_0)(y'_0, y'_0)/2 + o(1)$$

sería negativo.

Q.E.D.

Nótese que, como en el caso del Cálculo en varias variables, la parte (b) del Teorema 1.2 corresponde al “*test de la segunda derivada*” restringiéndose al conjunto de direcciones tangentes (dadas por $y'(0)$) a las hipersuperficies $G_i(y) = 0$.

6.2 Condiciones suficientes para extremos en la norma C^1

Sea el funcional $J(y) = \int_a^b F(x, y, y') dx$, con $y(a) = A$, $y(b) = B$. F es C^2 . Sabemos que

$$D^2 J(y)(h, h) = \int_a^b (F_{yy} h^2 + 2F_{yz} h h' + F_{zz} h'^2) dx.$$

Sabemos que una condición necesaria para que y_0 sea un mínimo es que $D^2 J(y_0)(h, h) \geq 0$, pero queremos tener condiciones necesarias que involucren el integrando solamente.

Tenemos primero una condición necesaria para un mínimo:

Teorema 2.1 : (*Legendre*) Si y_0 es un mínimo local, para la norma C^1 , entonces $F_{zz}(x, y_0(x), y'_0(x)) \geq 0$ para x en $[a, b]$.

Prueba : Supongamos que $F_{zz} < 0$ en un punto x_0 que podemos tomar, por continuidad, en (a, b) . Entonces $F_{zz} < -c$ en un intervalo $[x_0 - \alpha, x_0 + \alpha]$ contenido en $[a, b]$.

Sea h una función con gráfica:

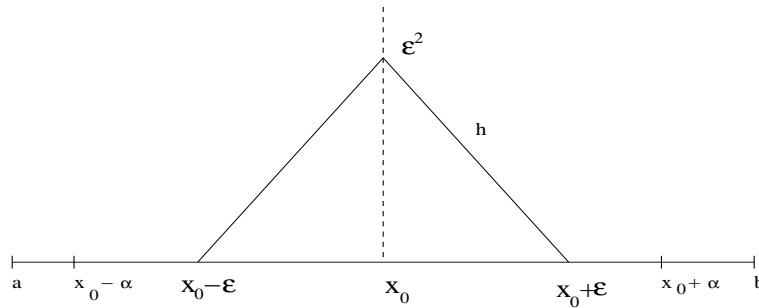


Fig. 2.1

Con $|h'| = \epsilon$, $\epsilon < \alpha$. Entonces

$$\begin{aligned} D^2 J(y_0)(h, h) &= \int_{x_0 - \epsilon}^{x_0 + \epsilon} (F_{yy} h^2 + 2F_{yz} h h' + F_{zz} h'^2) dx \\ &\leq 2M\epsilon^5 + 4N\epsilon^4 - 2c\epsilon^3 \end{aligned}$$

donde M es una cota para F_{yy} , N para F_{yz} en $[a, b]$. Por lo tanto esto será negativo si ϵ es suficientemente pequeño. Por otra parte $\int_a^b (h^2 + h'^2) dx \leq 2\epsilon^5 + 2\epsilon^3$ tiende a 0 cuando ϵ tiende a 0. $|h|_1 = \max |h| + \max |h'| \leq \epsilon + \epsilon^2$ también tiende a 0. Notemos que h es C^1 sólo por partes. Pero podemos aproximar h por \tilde{h} en C^1 , tal que $|h - \tilde{h}|_1 \leq \epsilon^3$. En ese caso:

$$D^2 J(y_0)(\tilde{h}, \tilde{h}) = D^2 J(y_0)(h, h) + D^2 J(y_0)(\tilde{h} - h, \tilde{h} + h).$$

Como $|h - \tilde{h}|_1 \leq \epsilon^3$, $|h + \tilde{h}|_1 \leq |h - \tilde{h}|_1 + 2|h|_1 \leq 3\epsilon$, para ϵ pequeño, el término $D^2 J(y_0)(\tilde{h} - h, \tilde{h} + h)$ es acotado por $M\epsilon^4$, y el resultado es válido para \tilde{h} .

Q.E.D.

Ejercicio 2.1 : Se considera el funcional

$$J(y) = \int_a^b F(x, y, y', y'') dx$$

con y y y' dados en a y en b .

a) Si F es C^2 , calcular $D^2 J(y)(h, k)$.

b) Si F es C^4 y y_0 lo es también (siempre cierto si $F_{y''y''} \neq 0$ y y_0 es un punto crítico), probar que:

$$D^2 J(y_0)(h, h) = \int_a^b (Ah^2 + Bh'^2 + Ch''^2) dx$$

con

$$A = F_{yy} - (F_{yy'})' + (F_{yy''})''$$

$$B = F_{y'y'} - (F_{y'y''})' - 2F_{yy''}$$

$$C = F_{y''y''}$$

c) Probar que $C(x) \geq 0$ es una condición necesaria para tener un mínimo. (Si $C(x_0) < 0$, tomar $h(x) = ((x - x_0)^2 - \epsilon^2)^2$ si $|x - x_0| \leq \epsilon$, 0 si $|x - x_0| > \epsilon$.)

Ejercicio 2.2 : En el ejemplo 1.3, probar que si la matriz A tiene un valor propio negativo en (x_0, y_0) entonces u no puede ser un mínimo. (Hacer un cambio de variables de tal forma que cerca de (x_0, y_0) la matriz A se vuelve diagonal. Cuidado con la diferenciabilidad. Escoger v con soporte en una vecindad de (x_0, y_0) , con gradiente esencialmente en la dirección del vector propio asociado al valor propio negativo.)

Ahora bien, es claro que $F_{zz} > 0$ no es una condición suficiente, ya que es una condición local: la “suma” de dos mínimos (en el sentido de que uno es la continuación del otro) no tiene por qué ser un mínimo; por ejemplo en el caso de las geodésicas:

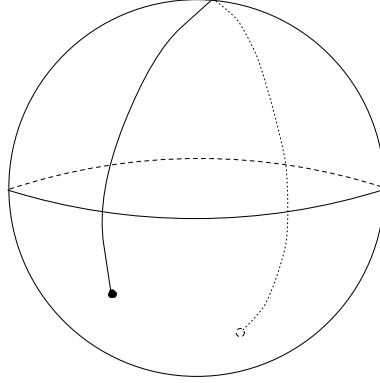


Fig. 2.2

Supondremos que en punto crítico y_0 , tenemos $F_{zz}(x, y_0(x), y'_0(x)) > 0$. Sabemos entonces que y_0 es C^2 (proposición IV.1.1). Si F es C^3 entonces podemos escribir:

$$\begin{aligned} J(y_0 + h) - J(y_0) &= \int_a^b ((F_{yy} - \frac{d}{dx}(F_{yz}))h^2 + F_{zz}h'^2) dx/2 + o(\|h\|_1) \|h\|_1^2 \\ &\equiv \int_a^b (Qh^2 + Ph'^2) dx/2 + o(\|h\|_1) \|h\|_1^2 \end{aligned}$$

donde integramos por partes el término $F_{yz}(h^2)'$, ($h = 0$ en a y b), (esto se puede hacer para funciones C^1 por pedazos) y Q y P son funciones de x , con $P > 0$ en $[a, b]$. (Ver el ejemplo 1.1, aquí $\|h\|_1^2 = \int_a^b (h^2 + h'^2) dx$.)

Sea

$$I(h) = \int_a^b (Qh^2 + Ph'^2) dx$$

un funcional cuadrático. Si $I(h) > 0$ para $h \neq 0$, probaremos que de hecho $I(h) \geq 2\alpha \int_a^b (h^2 + h'^2) dx$, para cierta α . En ese caso

$$J(y_0 + h) - J(y_0) \geq \alpha \|h\|_1^2 - o(\|h\|_1) \|h\|_1^2 \geq \alpha \|h\|_1^2/2,$$

si $\|h\|_1$ es pequeño: tenemos un mínimo.

Ahora bien $I(0) = 0$, queremos por lo tanto ver bajo que condiciones 0 es el único mínimo de $I(h)$. Para esto tenemos primero que estudiar los puntos críticos de $I(h)$. (Parece

que podríamos continuar *ad infinitum*, pero la gran diferencia es que $I(h)$ es un funcional cuadrático, cuya ecuación de Euler es lineal.)

Los puntos críticos de $I(h)$ sobre el espacio $C^1[a, b] \cap C_0[a, b]$, están dados por las soluciones a la ecuación de Euler asociada:

$$\frac{d}{dx}(Ph') - Qh = 0 \quad , \quad h(a) = h(b) = 0.$$

Si P es C^1 , como $P > 0$, h es C^2 y solución de:

$$Ph'' + P'h' - Qh = 0 \quad , \quad h(a) = h(b) = 0.$$

Es una ecuación lineal de segundo orden: se llama la *ecuación variacional* para $J(y)$.

Notemos que $h \equiv 0$ es solución.

Sabemos que el problema de valores iniciales:

$$(Ph')' - Qh = 0 \quad , \quad h(a) = 0, \quad h'(a) = \alpha$$

tiene una única solución. Si $u(x)$ es la solución para $\alpha = 1$, entonces, por la linealidad y la unicidad, la solución para α cualquiera es $\alpha u(x)$.

Definición 2.1 : \tilde{a} es un punto conjugado de a si existe una solución C^2 , no trivial al problema $(Ph')' - Qh = 0$, $h(a) = h(\tilde{a}) = 0$ ($\tilde{a} \neq a$).

(No trivial quiere decir $h \not\equiv 0$.) Por la unicidad es suficiente pedir $h'(a) \neq 0$, y podemos suponer $h'(a) = 1$, ya que si tenemos una solución con $h'(a) \neq 0$, entonces $h(x)/h'(a)$ tiene derivada 1 en a . En este libro, cada vez que se menciona la noción de puntos conjugados, se supone que la ecuación variacional es no singular, es decir que $P(x) \neq 0$ para x en $[a, b]$.

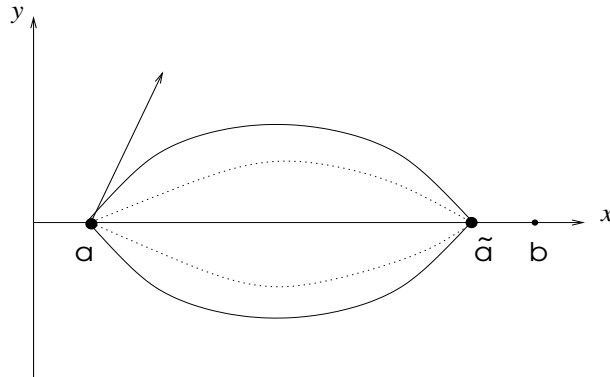


Fig. 2.3

Notemos que en el caso de funciones C^1 por pedazos, la condición de continuidad de Weierstrass (Teorema IV.7.1) es $[Ph']_c = 0$, es decir h' es continua: los puntos críticos de la ecuación variacional son C^1 y, como $P > 0$, son C^2 . Tenemos entonces:

Teorema 2.2 : *Sea $P > 0$ en $[a, b]$, P y Q continuos, entonces: $I(h) > 0$ para $h \not\equiv 0$ (con $h(a) = h(b) = 0$) si y solo si a no tiene puntos conjugados en $[a, b]$.*

Prueba : Supongamos que a no tiene puntos conjugados, entonces la solución a la ecuación

$$(Ph')' - Qh = 0, \quad h(a) = 0, \quad h'(a) = 1$$

es tal que $h(x) > 0$ en $]a, b]$.

Por continuidad de la solución con respecto a los datos iniciales, la solución $u(x)$ a $(Pu')' - Qu = 0$, $u(a) = \epsilon$, $u'(a) = 1$, con $\epsilon > 0$, también es estrictamente positiva en $[a, b]$, para ϵ pequeño: $Q = (Pu')'/u$.

$$\begin{aligned} I(h) &= \int_a^b (Qh^2 + Ph'^2) dx = \int_a^b ((Pu')'h^2/u + Ph'^2) dx \\ &= \int_a^b (Pu'(-2hh'/u + h^2u'/u^2) + Ph'^2) dx \end{aligned}$$

(integrando por partes y dado que $h(a) = h(b) = 0$).

$$I(h) = \int_a^b P(h' - hu'/u)^2 dx.$$

Por lo tanto $I(h) > 0$, al menos que $h' = hu'/u$. Pero como $h(a) = 0$, tendríamos $h'(a) = 0$ y por la unicidad al problema de valores iniciales para esa ecuación diferencial de primer orden, $h(x) \equiv 0$.

Inversamente, si b es punto conjugado de a , sea u solución de $(Pu')' - Qu = 0$ $u(a) = u(b) = 0$. Entonces

$$I(u) = \int_a^b ((Pu')'u + Pu'^2) dx = \int_a^b (-Pu'^2 + Pu'^2) dx = 0$$

integrando por partes. Esto contradice el hecho que $I(u) > 0$. Por lo tanto b no es punto conjugado de a .

Ahora, si \tilde{a} en $]a, b[$ es punto conjugado de a , consideremos la familia de funcionales parametrizados por t , dada por:

$$\begin{aligned} E(h, t) &= \int_a^b (tPh'^2 + tQh^2 + (1-t)h'^2) dx = \\ &= tI(h) + (1-t) \int_a^b h'^2 dx \end{aligned}$$

para todo t en $[0, 1]$, $E(h, t) > 0$, al menos que $h \equiv 0$. (De hecho $E(h, t) > 0$ si $t < 1$ y $I(h) \geq 0$.)

La ecuación de Euler para el funcional $E(h, t)$ es:

$$(tPh' + (1-t)h')' = tQh.$$

Sea $h(x, t)$ la solución de esta ecuación con datos iniciales

$$h(a, t) = 0, \quad h'(a, t) = 1.$$

Para $t = 0$, $h(x, 0) = x - a$, ya que en ese caso la ecuación es : $h'' = 0$. Para $t = 1$, sabemos que $h(\tilde{a}, 1) = 0$.

Consideremos el conjunto de puntos (x, t) tales que $h(x, t) = 0$: es decir x es conjugado de a para el problema $E(h, t)$.

Para cada punto (x_0, t_0) de este conjunto: $h_x(x_0, t_0) = h'(x_0, t_0) \neq 0$ (si no tendríamos $h(x_0, t_0) = h'(x_0, t_0) = 0$. Por la unicidad de la solución al problema de valores iniciales, en x , tendríamos $h(x, t_0) \equiv 0$, lo cual no es posible ya que $h'(a, t_0) = 1$). Por lo tanto, a partir de $(\tilde{a}, 1)$ tenemos una curva $x = x(t)$, con $h(x(t), t) = 0$. Hay entonces las siguientes posibilidades para esta curva.

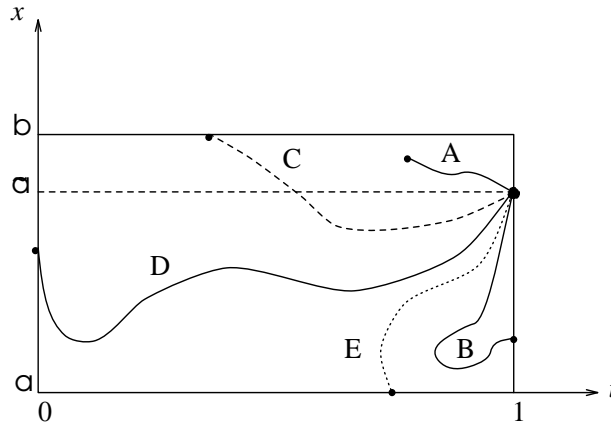


Fig. 2.4

Caso A : La curva desaparece antes de llegar a la frontera: esto no puede ser por el teorema de la función implícita: como $h_x(x, t) \neq 0$, la curva debe continuar.

Caso B : Esto es imposible ya que no es de la forma $x = x(t)$.

Caso C : Entonces b es conjugado de a para $E(h, t)$. Pero como $E(h, t) > 0$, el argumento que usamos para el caso $t = 1$ es válido y no podemos tener este caso. De hecho si $I(h) \geq 0$, tampoco podríamos tener este caso ya que el punto terminal debe tener $t < 1$, al menos que $\tilde{a} = b$.

Caso D : Para $t = 0$, el punto terminal es conjugado de a para $E(h, 0)$, es decir para la ecuación $h'' = 0$. Pero entonces $h(x) = x - a$ y no se puede cumplir que $h(x) = 0$ para x distinto de a .

Caso E : Sea (a, t_0) el punto terminal. Entonces tenemos que $x(t)$ tiende a a , cuando t tiende a t_0 . Pero $h(x(t), t) = 0 = h(a, t)$. Por el teorema del valor medio, existe \tilde{x} entre a y $x(t)$ tal que $h_x(\tilde{x}, t) = 0$, para cada t . Por lo tanto cuando t tiende a t_0 , tendríamos $h_x(a, t_0) = 0$. Pero entonces $h(x, t_0) \equiv 0$, por la unicidad del problema de valores iniciales; esto contradice $h'(a, t) = 1$.

Como hemos eliminado todos los casos, la conclusión es que \tilde{a} en $[a, b[$ no puede ser punto conjugado de a si $I(h) \geq 0$.

Q.E.D.

Con los dos teoremas anteriores tenemos:

Teorema 2.3 : Sea F en C^3 .

a) Si y_0 es un mínimo local, en C^1 o en C^1 por pedazos, para $J(y)$ con $y(a)$, $y(b)$ dados, entonces: $P \equiv F_{zz}(x, y_0(x), y'_0(x)) \geq 0$, $\frac{d}{dx}(F_z) - F_y = 0$ y, si $P > 0$ en $[a, b]$, no hay puntos conjugados de a en $[a, b]$.

b) Si y_0 es solución de $(F_z)' - F_y = 0$ con $F_{zz}(x, y_0(x), y'_0(x)) > 0$ y si no hay puntos conjugados de a en $[a, b]$, entonces y_0 es un mínimo local (para la norma $\|\cdot\|_1$) de $J(y)$.

Prueba : (a) Por el teorema 1.1, tenemos que $I(h) \geq 0$, y por el teorema 2.1, $F_{zz} \geq 0$. La última parte viene del comentario final en la prueba del teorema 2.2.

(b) Por la hipótesis que no hay puntos conjugados de a en $[a, b]$, entonces la solución a la ecuación:

$$(Ph')' = Qh, \quad h(a) = 0, \quad h'(a) = 1$$

es positiva en $]a, b]$. Podemos escoger α tan pequeño y positivo que la solución al problema:

$$((P - \alpha)h'_1)' = (Q - \alpha)h_1, \quad h_1(a) = 0, \quad h'_1(a) = 1$$

sea también positiva en $]a, b]$ y tal que $P - \alpha > 0$ en $[a, b]$. (Dejaremos al lector probar esta dependencia continua de las soluciones de ecuaciones ordinarias con respecto a los coeficientes. Aquí se puede probar directamente encontrando una ecuación no homogénea para $h - h_1$, con parte no homogénea $\alpha(h''_1 - h_1)$. Usando la fórmula de variación de constantes el lector debe probar que $|h - h_1| \leq c\alpha$.)

Por el teorema 2.2, tendremos que

$$I_\alpha(h) = \int_a^b ((P - \alpha)h'^2 + (Q - \alpha)h^2) dx \geq 0.$$

Por lo tanto $I(h) \geq \alpha \int_a^b (h^2 + h'^2) dx$.

Como ya vimos que el residuo en la expresión

$$J(y_0 + h) - J(y_0) = I(h)/2 + \text{Residuo}$$

es $O(\|h\|_1) \int_a^b (h^2 + h'^2) dx$, entonces tenemos el resultado.

Q.E.D.

Ejemplo 2.1 : En el Ejercicio III.3, vimos que el funcional

$$J(y) = \int_0^1 (1 + y'^2)^{-1} dx$$

$y(0) = 0$, $y(1) = 1$, tiene como único punto crítico a la función $y(x) = x$.

Aquí $F_{y'y'} = 2(2y'^2 - 1)(1 + y'^2)^{-3}$.

Para $y(x) = x$, $F_{y'y'} = P = \frac{1}{4}$.

Como $F_{yy} = F_{yy'} = 0$, la ecuación variacional se reduce a: $h''/2 = 0$, es decir, con los datos $h(0) = 0$, $h'(0) = 1$: $h(x) = x$. No hay puntos conjugados de 0 y $y(x) = x$ es un mínimo local para la norma $\|\cdot\|_1$.

Notemos que si F no depende explícitamente de y , entonces $Q = 0$ y la ecuación variacional se reduce a $(Ph')' = 0$, cuya solución con las condiciones iniciales $h(a) = 0$, $h'(a) = 1$, es $h(x) = \int_a^x \frac{P(a)}{P(\xi)} d\xi$, y por lo tanto $h(x) > 0$ para $x > a$, si la condición $P > 0$ es válida. Es decir, no hay puntos conjugados de a y, para $P > 0$, cualquier punto crítico del funcional, es un mínimo local en la norma $\|\cdot\|_1$. Dado que $P \geq \alpha > 0$ y $Q = 0$, se tiene que

$$I(h) \geq \alpha \int_a^b h'^2 dx \geq C \|h\|_1^2,$$

donde se usó la desigualdad de Poincaré para $h(x)$ con $h(a) = 0$. Por lo tanto, el resultado se puede ver directamente sin pasar por los puntos conjugados.

Ejemplo 2.2 : En uno de los estudios para el problema isoperimétrico (Ejercicio IV.9.1) se llegó a un funcional de la forma:

$$J(y) = \int_a^b (y - \lambda(1 + y'^2)^{1/2}) dx$$

$y(a) = y(b) = 0$, $\lambda > 0$. Por el Teorema 1.2, un máximo de $J(y)$ dará un máximo del problema condicionado original, cuya solución es un arco de círculo, si $b-a \leq L \leq \pi(b-a)/2$. (El lector puede probar, a partir de la ecuación de Euler y el requisito que $y(x) \geq 0$ que λ debe ser positivo: $y'(a) > a$). En este caso $F_{y'y'} = -\lambda(1 + y'^2)^{-3/2} < 0$, es decir la condición de Legendre para un *máximo* se cumple.

Como $F_{yy} = F_{yy'} = 0$, la ecuación variacional se reduce a:

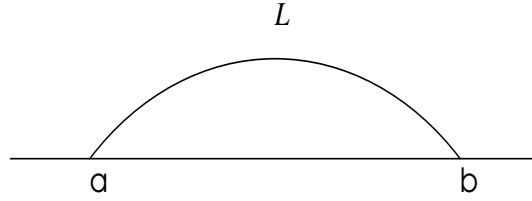


Fig. 2.5

$$(Ph')' = 0, \quad h(a) = 0, \quad h'(a) = 1.$$

Por lo tanto $h(x) = \int_a^x P(a)/P(\eta) d\eta$.

Como $P(a)/P(\eta) > 0$, h es creciente y no hay puntos conjugados: la solución encontrada es un máximo local en la norma $||\cdot||_1$. Veremos más adelante que es de hecho un máximo en la norma $||\cdot||_0$.

Ejercicio 2.3 : Sea $J(y) = \int_a^b f(x, y)(1 + y'^2)^{1/2} dx$ con $y(a) = y(b) = 0$, $f(x, y) \geq 0$.

a) Probar que la condición de Legendre siempre se cumple.

b) Si $f(x, y) \equiv 1$, encontrar el único punto crítico, probar que no hay puntos conjugados y concluir.

c) Si $f(x, y) = (1 + y^2)^{-1/2}$, probar que $y \equiv 0$ es solución de la ecuación de Euler. Probar que para esa solución hay puntos conjugados si y solo si $b \geq \pi$. Concluir.

Ejercicio 2.4 : Sea $J(y) = \int_0^a y^3 dx$ con $y(0) = 0$, $y(a) = b > 0$.

a) Encontrar el único punto crítico.

b) Probar que no hay puntos conjugados de 0. Concluir.

Ejemplo 2.3 : Geodésicas en una superficie.

Como ya lo vimos en el ejemplo IV.3.3, las geodésicas en una superficie, representada en forma paramétrica, son los mínimos del funcional:

$$I(u, v) = \int_{t_0}^{t_1} \|M'(t)\| dt = \int_{t_0}^{t_1} (Eu'^2 + 2Fu'v' + Gv'^2)^{1/2} dt$$

con $E = \|M_u\|^2$, $F = (M_u, M_v)$, $G = \|M_v\|^2$ (por lo tanto $F^2 < EG$, la igualdad sería solo si M_u y M_v son colineales y la superficie degenerada).

Este problema es con dos funciones, pero como se vio en la Nota V.6.1, $\det(\|M'\|_{u',v'}) = 0$. Por lo tanto la teoría correspondiente para este caso (tratada en la siguiente sección) no se aplica. Supondremos pues que la curva candidata a geodésica se puede parametrizar por u . En este caso:

$$I(v) = \int_{u_1}^{u_2} (E + 2Fv' + Gv'^2)^{1/2} du$$

tiene $F_{v'} = (F + Gv')\|M'\|^{-1/2}$, $F_{v'v'} = (EG - F^2)\|M'\|^{-3/2} > 0$.

Tomemos como ejemplo el caso de la esfera, con $E = \sin^2\theta$, $F = 0$, $G = 1$ y supongamos que la curva candidata a geodésica no pasa por los polos ($\theta = 0, \pi$) y que la parametrizamos por φ en $[0, 2\pi]$.

$$I(\theta) = \int_{\varphi_1}^{\varphi_2} (\sin^2\theta + \theta'^2)^{1/2} d\varphi.$$

En el ejercicio IV.3.2 se pidió resolver este problema, aquí (aunque no sea esencial para este ejemplo) pedimos al lector que compruebe que la solución a la ecuación de Euler es:

$$\theta = \text{Arctan}((\alpha \cos(\varphi - c))^{-1}), \quad \text{con } \alpha > 0.$$

y $c = \varphi_1 - \text{Arcos}(\alpha^{-1} \cot \theta_1)$ de tal modo que $\theta(\varphi_1) = \theta_1$. α es determinado por $\theta(\varphi_2) = \theta_2$, o por el dato inicial:

$$\alpha = (\theta_1'^2 + \sin^2 2\theta_1/4)^{1/2} / \sin^2 \theta_1.$$

Ahora bien $P = F_{\theta'\theta'} = \sin^2\theta(\sin^2\theta + \theta'^2)^{-3/2}$, con

$$F_{\theta\theta} = \cos 2\theta(\sin^2\theta + \theta'^2)^{-1/2} - \sin^2\theta \cos^2\theta(\sin^2\theta + \theta'^2)^{-3/2}$$

$$F_{\theta\theta'} = -\sin \theta \cos \theta \theta'(\sin^2\theta + \theta'^2)^{-3/2}.$$

De la ecuación de Euler: $(\theta'(\sin^2\theta + \theta'^2)^{-1/2})' = \sin\theta \cos\theta(\sin^2\theta + \theta'^2)^{-1/2}$ obtenemos $\theta'' = \cos\theta(2\theta'^2 + \sin^2\theta)/\sin\theta$.

Calculando

$$P' = -4\sin\theta \cos\theta \theta'(\sin^2\theta + \theta'^2)^{-3/2}.$$

Por lo tanto $Q = F_{\theta\theta} - (F_{\theta\theta'})' = (\cos 2\theta \sin^2\theta - 2\theta'^2)(\sin^2\theta + \theta'^2)^{-3/2}$.

(Dejamos como ejercicio al lector verificar estos cálculos.)

La ecuación variacional $(Ph')' = Qh$ se reduce a:

$$\begin{aligned} h'' - (4 \cot \theta \theta')h' - (\cos 2\theta - 2\theta'^2/\sin^2\theta)h &= 0 \\ &= \sin^2\theta \left[(h/\sin^2\theta)'' + (h/\sin^2\theta) \right]. \end{aligned}$$

Como $\sin^2\theta \neq 0$ y $h(\varphi_1) = 0$, la solución a esta ecuación es:

$$h(\varphi) = k \sin(\varphi - \varphi_1) \sin^2\theta.$$

El primer punto conjugado ($h(\overline{\varphi}) = 0$) es para $\overline{\varphi} = \varphi_1 + \pi$, en ese caso: $\theta(\overline{\varphi}) = \pi - \theta_1$. Este punto es el opuesto al punto inicial sobre la esfera.

Ejercicio 2.5 : Retomar este ejemplo, pero esta vez parametrizar por $\varphi(\theta)$ (es más fácil!).

Ejemplo 2.4 : Problema de frontera libre.

Sea y_0 un punto crítico del funcional:

$$J(y) = \int_a^b F(x, y, y') dx$$

con $y(b) = \tilde{B}$ y $y(a)$ libre. F es C^3 .

Por lo tanto y_0 es solución del problema:

$$(F_z)' - F_y = 0, \quad y(b) = \tilde{B}, \quad F_z(a, y(a), y'(a)) = 0$$

(Teorema IV.6.1).

Sea $X = \{h \text{ en } C^1[a, b], \text{ con } h(b) = 0 \text{ y } \|h\| = |h|_1\}$.

Como en el ejemplo 1.1, $J(y)$ tiene segunda derivada de Fréchet:

$$D^2 J(y_0)(h, k) = \int_a^b (F_{yy}hk + F_{yy'}hk' + F_{y'y}h'k + F_{y'y'}h'k') dx.$$

Por lo tanto, si y_0 es un punto crítico:

$$J(y_0 + h) - J(y_0) = \int_a^b (F_{yy}h^2 + 2F_{yz}hh' + F_{zz}h'^2) dx/2 + o(\|h\|_1^2)$$

Si integramos por partes el término $2F_{yz}hh'$, tendremos:

$$D^2 J(y_0)(h, h) = \int_a^b ((F_{yy} - \frac{d}{dx}(F_{yz}))h^2 + F_{zz}h'^2) dx - F_{yz}(a, y_0(a), y_0'(a))h^2(a).$$

Claramente la condición de Legendre $F_{zz}(x, y_0(x), y_0'(x)) \geq 0$ sigue válida, ya que podemos tomar $h(a) = 0$ para un subconjunto de X .

Si escribimos

$$D^2 J(y_0)(h, h) = I(h) = \int_a^b (Ah'^2 + 2Bhh' + Ch^2) dx$$

entonces la ecuación de Euler para $I(h)$, la ecuación variacional para $J(y)$, es

$$(Ah' + Bh)' - (Bh' + Ch) = 0 = (Ah')' + (B' - C)h$$

(suponiendo F en C^3 y $F_{zz} > 0$, entonces y_0 es C^2 y B es diferenciable).

Es decir:

$$(Ph')' - Qh = 0.$$

Con $P = F_{zz}$, $Q = F_{yy} - (F_{yz})'$.

Esto no es sorprendente ya que $C^1[a, b] \cap C_0[a, b]$ está contenido en X . Aquí solo las condiciones de frontera cambian: la condición natural para la ecuación variacional es $Ah' + Bh|_a = 0$.

$$F_{zz}(a, y_0(a), y_0'(a))h'(a) + F_{yz}(a, y_0(a), y_0'(a))h(a) = 0.$$

Supondremos $P > 0$ en $[a, b]$; la ecuación de segundo orden $(Ph')' = Qh$, $h(a) = \epsilon$, $h'(a) = -\epsilon B/A$ (por lo tanto satisfaciendo la condición natural para el problema variacional), tiene una única solución. Notemos que, debido a la linealidad, la solución con ϵ es $\epsilon u(x)$, con $u(a) = 1$, $u'(a) = -B/A$, $(Pu')' = Qu$.

Definición 2.2 : \tilde{a} en $[a, b]$ es **conjugado** de a si $u(\tilde{a}) = 0$.

Notemos que \tilde{a} no depende de ϵ (otra vez la relación $h(x) = \epsilon u(x)$).

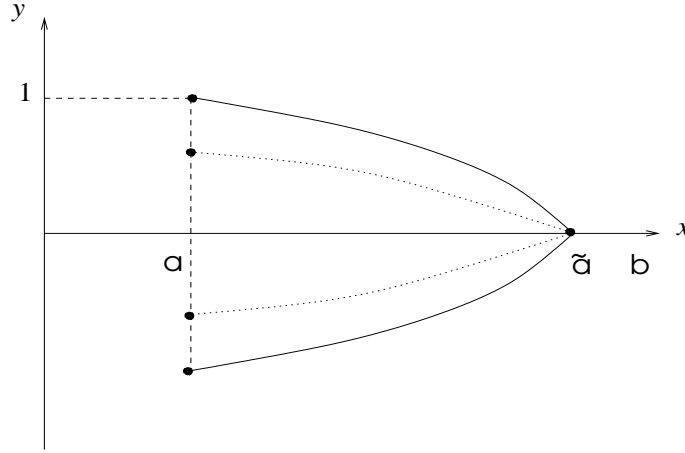


Fig. 2.6

Teorema 2.4 : Para el problema de frontera libre en C^1 o C^1 por pedazos:

- a) $D^2J(y_0)(h, h) > 0$, para h en X si y sólo si no hay puntos conjugados de a en $[a, b]$.
- b) Si y_0 es un mínimo local, entonces: $(F_{y'})' - F_y = 0$, $F_{y'}|_a = 0$, $y_0(b) = \tilde{B}$, $P \equiv F_{y'y'} \geq 0$ y si $P > 0$ en $[a, b]$ no hay puntos conjugados de a en $[a, b]$.
- c) Si y_0 es solución de $(F_{y'})' - F_y = 0$, $F_{y'}|_a = 0$, $y_0(b) = \tilde{B}$, $F_{y'y'} > 0$ y no hay puntos conjugados de a en $[a, b]$, entonces y_0 es un mínimo local en X , con la norma C^1 .

Prueba : (a) Si no hay puntos conjugados de a en $[a, b]$, entonces la solución u de $(Pu')' - Qu = 0$, $u(a) = 1$, $u'(a) = -B/A$ es siempre positiva en $[a, b]$.

$$\begin{aligned}
 D^2J(y_0)(h, h) &= \int_a^b (Ph'^2 + Qh^2) dx - F_{yz}(a)h^2(a) \\
 &= \int_a^b (Ph'^2 + (Pu')'h^2/u) dx - F_{yz}h^2|_a \\
 &= \int_a^b P(h' - u'h/u)^2 dx - (Pu'h^2/u + F_{yz}h^2)|_a
 \end{aligned}$$

$$= \int_a^b P(h' - u'h/u)^2 dx,$$

(por el valor de u y u' en a).

Tendremos $D^2J(y_0)(h, h) > 0$ al menos que $h' = u'h/u$. Como $h(b) = 0$, entonces $h'(b) = 0$, y la única solución es $h \equiv 0$.

Inversamente, si b es conjugado de a , entonces u pertenece a X , con

$$\begin{aligned} D^2J(y_0)(u, u) &= \int_a^b (Pu'^2 + Quu) dx - F_{yz}u^2(a) \\ &= \int_a^b (Pu'^2 - Pu'^2) dx - (Pu' + F_{yz}u)u|_a \end{aligned}$$

donde se usó la ecuación $Qu = (Pu')'$ y se integró por partes. Usando las condiciones iniciales para u , tenemos $D^2J(y_0)(u, u) = 0$, lo cual contradice la hipótesis.

Si consideramos ahora:

$$I(h, t) = \int_a^b (t(Ch^2 + 2Bhh' + Ah'^2) + (1-t)Ah'^2) dx, \text{ sobre } X.$$

Nótese que, en el Teorema 2.2, se hubiera podido utilizar $I(h, t)$ en lugar de $E(h, t)$.

Como $I(h, 1) > 0$, entonces $I(h, t) > 0$ para $h \not\equiv 0$, ya que para $t = 0$, si $I(h, 0) = 0$ entonces $h' \equiv 0$, $h(x) = h(b) = 0$. (Si $I(h, 1) \geq 0$ entonces $I(h, t) > 0$ para $t < 1$ y $h \not\equiv 0$.)

La ecuación variacional asociada a $I(h, t)$ es:

$$(Ph')' = tQh, \quad h(b) = 0, \quad tBh(a) + Ph'(a) = 0.$$

Esta ecuación con condiciones iniciales:

$$h(a) = 1, \quad h'(a) = -tB/P$$

tiene una única solución $h(x, t)$.

Para $t = 0$, $h(x, 0)$ es tal que $(Ph')' = 0$, es decir $Ph' = \text{constante}$, pero con $h'(a, 0) = 0$, tenemos $h' \equiv 0$, $h(x, 0) = 1$. No hay puntos conjugados de a para $t = 0$.

Ahora bien, si $h(x_0, t) = 0$ (es decir x_0 es conjugado de a para $I(h, t)$), por la unicidad de la solución, $h_x(x_0, t) \neq 0$ y podemos expresar el conjunto de puntos conjugados como un conjunto de curvas de la forma $x(t)$. Entonces si \tilde{a} es un punto conjugado de a , para $t = 1$, tenemos las siguientes posibilidades :

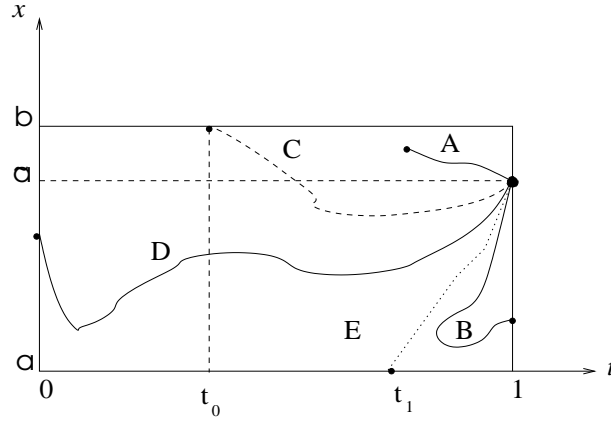


Fig. 2.7

Dejamos al lector analizar cada caso y ver que ninguno es posible (para $C : I(h, t_0) = 0$ ya que b es conjugado de a para ese funcional; para $E : h(a, t_1) = 0$, por continuidad, pero $h(a, t) = 1$).

Vemos también que si $I(h, 1) \geq 0$, entonces la única posibilidad es $\tilde{a} = b$. Esto prueba (b).

(c) Si no hay puntos conjugados de a para $I(h)$ tampoco los hay para $I(h) - \epsilon \int_a^b (h^2 + h'^2) dx$ para ϵ pequeño. El resto del argumento es entonces como antes.

Q.E.D.

6.3 Condiciones suficientes para varias funciones en la norma C^1

Consideremos ahora el funcional:

$$J(y_1, \dots, y_n) = \int_a^b F(x, y_1, \dots, y_n, y'_1, \dots, y'_n) dx$$

con $y_i(a) = A_i$, $y_i(b) = B_i$. F es C^3 .

Ya vimos en el ejemplo 1.2 que:

$$\begin{aligned} J(Y + H) - J(Y) &= \int_a^b (F_Y \cdot H + F_Z \cdot H') dx \\ &+ \frac{1}{2} \int_a^b (H^T F_{YY} H + H^T F_{YZ} H' + H'^T F_{ZY} H + H'^T F_{ZZ} H') dx \end{aligned}$$

$$+0(\|H\|_1) \int_a^b (H^T H + H'^T H') dx.$$

Aplicando el teorema 1.1 vemos que si Y es un mínimo entonces

$$D^2 J(Y)(H, H) \geq 0$$

y que si $D^2 J(Y)(H, H) \geq \alpha \|H\|_1^2$ en un punto crítico Y , entonces Y es un mínimo local, con respecto a la norma $\|\cdot\|_1$.

De manera análoga tenemos la condición necesaria:

Teorema 3.1 : (Legendre).

En un mínimo local Y_0 , la matriz simétrica $F_{ZZ}(x, Y(x), Y'(x))$ no tiene valores propios negativos.

Prueba : Ejercicio. Si F_{ZZ} tiene un valor propio negativo en x_0 , diagonalizar F_{ZZ} cerca de x_0 y tomar un incremento H localizado cerca de x_0 y en la dirección del vector propio correspondiente.

Supondremos que $P = F_{ZZ}$ es definida positiva y estudiaremos el funcional:

$$I(H) = \int_a^b (H^T C H + H^T A H' + H'^T A^T H + H'^T P H') dx$$

donde $C = F_{YY}$, $A = F_{YZ}$.

Para poder probar que $I(H) > 0$, para $H \neq 0$, necesitamos ver que $H = 0$ es el único mínimo para $I(H)$ sobre $C^1[a, b] \cap C_0[a, b]$, o en C^1 por pedazos.

Los puntos críticos de $I(H)$, tienen

$$DI(H)K = 0$$

para toda K en el espacio, es decir:

$$\int_a^b (K^T C H + K^T A H' + K'^T A^T H + K'^T P H') dx$$

(donde se usó $C = C^T$, $P = P^T$). Como P es no singular, Y es C^2 (Ejercicio IV.3.2). Supondremos que F es C^3 , podemos entonces integrar por partes:

$$\int_a^b K^T (C H + A H' - \frac{d}{dx}(A^T H) - \frac{d}{dx}(P H')) dx = 0$$

para toda K en el espacio.

La ecuación variacional es entonces:

$$(PH')' - (A - A^T)H' - (C - A^{T'})H = 0$$

con $H(a) = H(b) = 0$.

(En el caso de C^1 por pedazos, las condiciones de continuidad de Weierstrass implican que H es C^1 y, como P es definida positiva, H es C^2 .)

Definición 3.1 : Diremos que \tilde{a} es **punto conjugado** de a si la ecuación variacional tiene una solución (no idénticamente 0) $H(x)$ con $H(a) = 0$, $H(\tilde{a}) = 0$, H es C^2 .

Sean H_1, \dots, H_n , n soluciones con $H_i(a) = 0$, $H'_i(a) = (0, \dots, 1, \dots, 0)$ (1 en la i -ésima función). (Con estas condiciones, las soluciones son automáticamente linealmente independientes cerca de a , ver cualquier libro de ecuaciones diferenciales.)

Sea $B = (H_1, \dots, H_n)$ la matriz cuyas columnas son H_1, \dots, H_n . B se llama la *matriz fundamental* del sistema. Entonces toda solución a la ecuación variacional con $H(a) = 0$, $H'(a) = \alpha$, se escribe: $H(x) = B(x)\alpha$, ya que $B'(a) = I$ y por la unicidad al problema de valores iniciales.

Por lo tanto \tilde{a} es conjugado de a si y sólo si $\det B(\tilde{a}) = 0$.

Teorema 3.2 : a) Si P es definida positiva y F es C^3 , entonces: $I(H) > 0$ para $H \neq 0$, si y sólo si no hay puntos conjugados de a en $[a, b]$.

b) Si Y es un mínimo local, entonces: P no tiene valores propios negativos en $[a, b]$, $(F_{z_i})' - F_{y_i} = 0$, y no hay puntos conjugados de a en $[a, b]$. (Probaremos algo más fuerte con respecto a este punto: si $I(H) \geq 0$ entonces no hay puntos conjugados de a en $[a, b]$.)

c) Si P es definida positiva, Y es solución de las ecuaciones de Euler y no hay puntos conjugados de a en $[a, b]$, entonces Y es un mínimo local de $J(Y)$ en la norma $\|\cdot\|_1$.

(En la Nota 4.1 daremos otra demostración, basada en la teoría de campos: de este modo el lector podrá darse cuenta de que los argumentos que parecen simples en una dimensión no lo son en más dimensiones y que presentando las cosas de manera un poco distinta (en el espacio de funciones C^1 por pedazos y usando más la teoría de ecuaciones diferenciales) uno puede llegar a una demostración más elegante o bien a resultados más generales).

Prueba : (a) Supongamos que no hay puntos conjugados en $[a, b]$, es decir $\det B(x) \neq 0$ para x en $[a, b]$.

Sea U una matriz solución del problema:

$$-(PU' + A^T U)' + AU' + CU = 0$$

$$U(a) = \epsilon I, \quad U'(a) = I + \epsilon P^{-1}A.$$

Por continuidad con respecto a los datos iniciales, si ϵ es pequeño, $\det U(x)$ es vecino de $\det B(x)$ y, cerca de a , las columnas de $U(x)$ son linealmente independientes.

Ahora para probar que $I(H) > 0$ queremos escribir el integrando como un cuadrado (como en una dimensión). Para esto daremos el siguiente argumento:

Sea W una matriz simétrica, entonces:

$$\int_a^b (H^T W H)' dx = 0 = \int_a^b (H^T W' H + 2H^T W H') dx$$

ya que $H(a) = H(b) = 0$ y $W = W^T$.

Además como $P = P^T$ es definida positiva, $P = T^T \Lambda T$ donde T es ortogonal, compuesta de los vectores propios de P , y Λ es diagonal. Podemos definir entonces $P^{1/2} \equiv T^T \Lambda^{1/2} T$. Claramente $(P^{1/2})(P^{1/2}) = P$, $P^{1/2}$ es simétrica positiva.

Entonces:

$$\begin{aligned} I(H) = \int_a^b & ((P^{1/2} H')^T P^{1/2} H' + 2(P^{1/2} H')^T P^{-1/2} W H + H^T W' H \\ & + 2(P^{1/2} H')^T P^{-1/2} A^T H + H^T C H) dx \end{aligned}$$

donde agregamos 0 a $I(H)$ y donde se usó el hecho que $H^T A H' = H'^T A^T H$, siendo esta última expresión un número y por lo tanto igual a su transpuesta.

$$\begin{aligned} I(H) = \int_a^b & [(P^{1/2} H' + P^{-1/2}(W + A^T)H)^T (P^{1/2} H' + P^{-1/2}(W + A^T)H) \\ & - (P^{-1/2}(W + A^T)H)^T (P^{-1/2}(W + A^T)H) + H^T W' H + H^T C H] dx. \end{aligned}$$

Tendremos un cuadrado si:

$$H^T (W' H + C H - (W + A^T)^T P^{-1} (W + A^T) H) = 0$$

para todo H . Es decir:

$$W' + C = (W + A)P^{-1}(W + A^T)$$

(ya que $W^T = W$) o bien:

$$W' - W P^{-1} W - A P^{-1} W - W (A P^{-1})^T + C - A P^{-1} A^T = 0.$$

Esta es una ecuación diferencial de Riccati.

Para resolverla se escribe $W = VU^{-1}$ donde V y U son por determinar. Entonces:

$$W' = V'U^{-1} - VU^{-1}U'U^{-1}.$$

(Derivando $UU^{-1} = I$, se obtiene $(U^{-1})' = -U^{-1}U'U^{-1}$.)

Multiplicando la ecuación de Riccati (por la derecha) por U se obtiene: se obtiene:

$$\begin{aligned} V' - VU^{-1}U' - VU^{-1}P^{-1}V - AP^{-1}V - VU^{-1}(AP^{-1})^TU + CU - AP^{-1}A^TU \\ = V' - AP^{-1}V + (C - AP^{-1}A^T)U - VU^{-1}(U' + P^{-1}V + (AP^{-1})^TU) = 0. \end{aligned}$$

Pongamos

$$V = -P(U' + (AP^{-1})^TU) = -(PU' + A^TU).$$

Entonces:

$$\begin{aligned} V' - AP^{-1}V + (C - AP^{-1}A^T)U &= 0 \\ = -(PU')' - A^T'U - A^TU' + AU' + AP^{-1}A^TU + CU - AP^{-1}A^TU \\ &= -(PU')' - (A^T)'U + (A - A^T)U' + CU \\ &= -(PU' + A^TU)' + AU' + CU = 0. \end{aligned}$$

Es decir, U es solución de la ecuación variacional.

Entonces para poder regresarnos, necesitamos ver que la ecuación variacional tiene una solución U , invertible en $[a, b]$, tal que $W = VU^{-1} = -(PU' + A^TU)U^{-1}$ sea simétrica.

Tomemos U como al principio de la prueba; notemos que W satisface a la ecuación de Riccati:

$$W' = WP^{-1}W + ((AP^{-1})W + W(AP^{-1})^T - (C - AP^{-1}A^T))$$

(regresandose hasta la ecuación).

Como C , $AP^{-1}A^T$, P son simétricas, W^T es también solución de esta ecuación (tomando su transpuesta). Además esta ecuación tiene una solución única si se da $W(a)$.

Aquí

$$W(a) = -P(I + \epsilon P^{-1}A)I/\epsilon - A^T = -P/\epsilon - (A + A^T) = W(a)^T.$$

Por lo tanto $W = W^T$. $I(H)$ es entonces 0 sólo si $PH' = -(W + A^T)H$. Pero como $H(a) = 0$, la única solución es $H = 0$.

Notemos entonces que si no hay puntos conjugados de a en $[a, b]$, tampoco lo habrá para

$$I(h) - \epsilon \int_a^b (H^T H + H'^T H') dt$$

si ϵ es tan pequeño que $P - \epsilon I$ sigue positiva definida. Del hecho que este funcional debe ser positivo, concluimos la prueba de (c) para funciones que son C^1 por pedazos.

Para poder probar los conversos no podemos usar el argumento de continuación que dió el resultado para $n = 1$. En efecto, si $\det B(x)$ determina los puntos conjugados como sus ceros, en un punto conjugado \tilde{a} no necesariamente $(\det B(\tilde{a}))' \neq 0$, es decir que no podemos representar el conjunto de puntos conjugados, para el problema dependiente del parámetro t , como curvas. (Por ejemplo en a , $\det B(a) = 0$, $(\det B(a))' = 0$.) Por otra parte, si especificamos $H(x)$ tal que $H(\tilde{a}) = 0$, los ceros de $H(t, x)$ serían, en el mejor de los casos, n curvas y no una sola.

Aprovechamos esta dificultad para dar un resultado más general: el teorema del índice de Morse. (Daremos otra demostración en la Nota 4.1, usando la teoría de campos.)

En el espacio X de funciones C^1 por pedazos entre a y b , continuas y con $H(a) = H(b) = 0$, consideramos el funcional:

$$I(H) = \int_a^b (H^T C H + H^T A H' + H^{T'} A^T H + H^{T'} P H') dx$$

con $P(x)$ definida positiva, C simétrica, A, P, C son C^1 .

La derivada de Fréchet de $I(H)$ es dos veces la forma bilineal:

$$I(H, K) = \int_a^b (K^T C H + K^T A H' + K^{T'} A^T H + K^{T'} P H') dx.$$

Para cada c entre a y b definiremos la forma bilineal:

$$I_c(H, K) = \int_a^c (K^T C H + K^T A H' + K^{T'} A^T H + K^{T'} P H') dx$$

sobre el espacio X_c de funciones C^1 por pedazos, continuas con $H(a) = H(c) = 0$, $K(a) = K(c) = 0$. Igualmente definiremos $I_c(H) = I_c(H, H)$.

Vemos entonces que c es conjugado de a si y sólo si $I_c(H, K) = 0$ para algún H en X_c y todo K en X_c . En efecto, si esto sucede entonces $DI_c(H)K \equiv 0$, H satisface la ecuación variacional:

$$-(PH' + A^T H)' + AH' + CH = 0, \quad H(a) = H(c) = 0,$$

$[PH' + AH] = 0$ en cada discontinuidad, pero con $[H] = 0$, P definida, entonces $[H'] = 0$, H es C^1 y como P es definida positiva (y por tanto distinta de cero) H es C^2 . El inverso se prueba por integración por partes. Vemos además que $I_c(H) = 0$.

Como $H(x) = B(x)H'(a)$, donde $B(x)$ es la matriz fundamental para las soluciones con datos iniciales $H(a) = 0$, no hay más de n soluciones $H(x)$ linealmente independientes.

Definición 3.2 : La **multiplicidad** de c como punto conjugado de a es el número de soluciones linealmente independientes de la ecuación variacional con $H(a) = H(c) = 0$.

Teorema 3.3 : (Teorema del índice de Morse).

Sea X_- el subespacio de X donde $I(H) < 0$ si $H \neq 0$. Entonces:

- a) Índice de $I \equiv \dim X_-$ es finito.
- b) El número de puntos conjugados de a entre a y b es finito.
- c) El índice de I es el número de puntos conjugados de a , estrictamente entre a y b y contados cada uno con su multiplicidad.

Fin de la prueba del teorema 3.2 : (Ver también la Nota 4.1 para una prueba directa.)

Si $I(H) \geq 0$ (caso de un mínimo) entonces el índice de I es 0 y no hay puntos conjugados en $[a, b]$.

Si $I(H) > 0$ ($H \neq 0$) el único candidato a punto conjugado podría ser b , pero, si este fuese el caso, entonces $I(H) = 0$ para las soluciones de la ecuación variacional con $H(a) = H(b) = 0$ (es decir miembros de X).

Notemos que el Teorema 3.2 está planteado sobre el espacio de funciones es C^1 (y no X , espacio de funciones C^1 por pedazos). Ahora bien si $I(H) \geq 0$ para H en C^1 , por continuidad $I(H) \geq 0$ para H en X (de hecho para H en $H_1[a, b]$). Por lo tanto (b) del teorema 3.2 es válido y el único punto conjugado ser b .

Si $I(H) > 0$ en C^1 , $I(H) \geq 0$ en X , el único punto conjugado sería b , pero entonces tendríamos una solución de la ecuación variacional (en C^2) con $I(H) = 0$. Por la primera

parte $I(H) > 0$ en X también. Otra manera de ver esto es que los puntos conjugados no dependen del espacio de trabajo.

Q.E.D.

Prueba del Teorema 3.3:

Advertencia al lector: el nivel de dificultad matemática de esta demostración es ligeramente superior al del resto de estas notas. Se recomienda hacer una primera lectura superficial y después, en caso de desearlo, un estudio más profundo.

Dividiremos la prueba en varias secciones.

Lema 3.1 : *Si c y \tilde{c} son conjugados entre sí entonces existe una constante k (independiente de c y \tilde{c}) tal que $|c - \tilde{c}| \geq k$.*

Prueba : Si \tilde{c} es conjugado a c entonces la ecuación variacional con datos $H(c) = H(\tilde{c}) = 0$ tiene una solución no trivial. Si $B(x, c)$ es la matriz fundamental para la ecuación variacional con datos $B(c, c) = 0$, $B'(c, c) = I$, entonces

$$B(x, c) = (x - c)(I + 0(x - c))$$

(desarrollo de Taylor) donde $0(x - c)$ es independiente de c : ver la prueba del lema V.6.3: la cota depende únicamente de los coeficientes de la ecuación variacional.

Por lo tanto $B(x, c)$ es invertible para $0 < |x - c| \leq k$ (tal que $\|0(|x - c|)\| < 1$). Pero $\det B(\tilde{c}, c) = 0$.

Q.E.D.

Podemos entonces subdividir el intervalo $[a, b]$ en subintervalos $[c_i, c_{i+1}]$, $i = 0, \dots, N$, $c_0 = a, c_{N+1} = b$, de longitud menor que k : en cada intervalo no hay puntos conjugados.

Sea \tilde{X} el subespacio de X (por lo tanto $\tilde{H}(a) = \tilde{H}(b) = 0$) de $\tilde{H}(x)$ soluciones en cada intervalo $[c_i, c_{i+1}]$ de la ecuación variacional ($\tilde{H}(x)$ tiene discontinuidades sólo en \tilde{H}' y en c_i). Sea X' el subespacio de X de $K(x)$ con $K(c_i) = 0$, $i = 0, \dots, N$.

Lema 3.2 : *a) La dimensión de \tilde{X} es finita, e igual a Nn .*

b) $X = \tilde{X} \oplus X'$ de tal modo que $I(\tilde{H}, K) = 0$, para \tilde{H} en \tilde{X} , K en X' . $I(K, K) > 0$, si K está en X' , $K \neq 0$.

Prueba : (a) Sean \tilde{H}_1 y \tilde{H}_2 dos elementos de \tilde{X} , con $\tilde{H}_1(c_i) = \tilde{H}_2(c_i)$. Por lo tanto $\tilde{H}_1 - \tilde{H}_2$ es solución de la ecuación variacional, con datos 0 en cada c_i . Entonces c_i es conjugado de c_{i+1} , lo cual no es posible al menos que $\tilde{H}_1 \equiv \tilde{H}_2$. (Esto demuestra también que $\tilde{X} \cap X' = \{0\}$.) Por lo tanto cada elemento de X está determinado por completo por su valor en cada c_i ,

$i = 1, \dots, N$.

Sean $B_k(x)$ las matrices fundamentales con $B_k(c_k) = 0$, $B'_k(c_k) = I$, $k = 0, \dots, N$. Sean $\alpha_1, \dots, \alpha_N$, N vectores en \mathbb{R}^n .

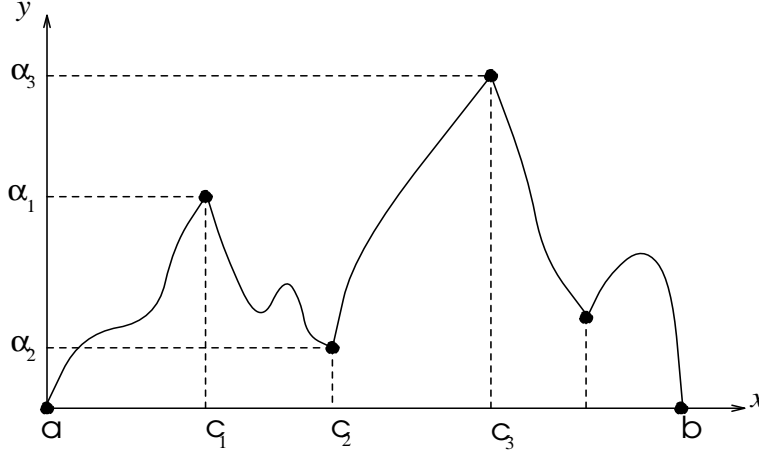


Fig. 3.1

$B_k(x)$ es invertible en $(c_k, c_{k+1}]$. Por lo tanto existe β_0 tal que $B_0(c_1)\beta_0 = \alpha_1$.

$B_0(x)\beta_0$ será la primera parte de la solución \tilde{H} . También existe β_1 tal que:

$$B_1(c_2)\beta_1 = \alpha_2 - B_0(c_2)\beta_0.$$

Por lo tanto: $\tilde{H}(x) \equiv B_0(x)\beta_0 + B_1(x)\beta_1$, en $[c_1, c_2]$, es tal que $\tilde{H}(c_1) = B_0(c_1)\beta_0 = \alpha_1$, $\tilde{H}(c_2) = \alpha_2$.

Por inducción: sea β_k tal que

$$B_k(c_{k+1})\beta_k = \alpha_{k+1} - \sum_0^{k-1} B_i(c_{k+1})\beta_i$$

y $\tilde{H}(x) \equiv \sum_0^k B_i(x)\beta_i$, en $[c_k, c_{k+1}]$.

$$\begin{aligned} \tilde{H}(c_k) &= \sum_0^{k-1} B_i(c_k)\beta_i \quad (\text{ya que } B_k(c_k) = 0) \\ &= B_{k-1}(c_k)\beta_{k-1} + \sum_0^{k-2} B_i(c_k)\beta_i = \alpha_k \quad (\text{por construcción}) \\ \tilde{H}(c_{k+1}) &= \alpha_{k+1} \end{aligned}$$

(para $k = N$ se escoge $\alpha_{N+1} = 0$). Por lo tanto la dimensión de \tilde{X} es Nn .

b) Sea H cualquier elemento de X . Sea \tilde{H} el único elemento de \tilde{X} tal que $\tilde{H}(c_i) = H(c_i)$, $i = 1, \dots, N$. Entonces $K \equiv H - \tilde{H}$ pertenece a X' . Como $X' \cap \tilde{X} = \{0\}$, tenemos que $X = \tilde{X} \oplus X'$.

Además $I(\tilde{H}, K) = \sum_{i=1}^N K^T [P\tilde{H}' + A^T \tilde{H}]_{c_i} = \sum_{i=1}^N K^T [P\tilde{H}']_{c_i}$ (integrando por partes en cada intervalo). Pero $K(c_i) = 0$.

Finalmente, en cada intervalo $[c_k, c_{k+1}]$ no hay puntos conjugados de c_k (por construcción). Por la primera parte del teorema 3.2 (ya probado)

$$I(K) = \sum_0^N I(K)|_{[c_k, c_{k+1}]} > 0, \text{ al menos que } K \equiv 0.$$

Notemos que $I(\tilde{H}) = \sum_{i=1}^N \tilde{H}^T [P\tilde{H}']_{c_i}$. Pero

$$[\tilde{H}']_{c_i} = \sum_0^i B'_j(c_i)\beta_j - \sum_0^{i-1} B'_j(c_i)\beta_j = B'_i(c_i)\beta_i = \beta_i.$$

Entonces

$$I(\tilde{H}) = \sum_{i=1}^N \alpha_i^T P \beta_i.$$

Q.E.D.

Lema 3.3 : a) *El índice de I sobre X es igual al índice de I sobre \tilde{X} y es finito.*

b) *Si b es conjugado de a en X con multiplicidad ν , también lo es en \tilde{X} con la misma multiplicidad.*

Prueba : Para cada H en X tenemos \tilde{H} en \tilde{X} , único, y definido por $\tilde{H}(c_i) = H(c_i)$. La aplicación $H \xrightarrow{P} \tilde{H}$ es claramente lineal, con $PH = \tilde{H}$ (es una proyección sobre el espacio, de dimensión Nn , \tilde{X}).

(a) Si H pertenece a X_- , entonces $I(H) = I(\tilde{H}) + I(K) < 0$, por lo tanto $I(\tilde{H}) = I(H) - I(K) < 0$ (la única posibilidad que $I(\tilde{H}) = 0$ es $K = 0$, $H = 0$). Por lo tanto: $P|_{X_-}$ es uno a uno, sobre \tilde{X}_- (ya que si \tilde{H} está en \tilde{X}_- , entonces para $H = \tilde{H}$, $I(H) < 0$). Como \tilde{X}_- es de dimensión finita, el índice es finito.

(b) Si b es conjugado de a en X , sean H_1, \dots, H_n soluciones linealmente independientes (C^2) de la ecuación variacional con $H_j(a) = H_j(b) = 0$, entonces H_j pertenece a X , $\tilde{H}_j = H_j$ (por la unicidad de las soluciones en cada intervalo): H_j está en \tilde{X} . En ese caso $\beta_1 = \dots = \beta_N = 0$ (son los saltos en la derivada) $\tilde{H}_j(x) = B_0(x)\beta_{0j}$.

Inversamente sea \tilde{H} representado por $(\alpha_1, \dots, \alpha_N)$ (con sus saltos β_0, \dots, β_N) tal que para toda \overline{H} (con $\overline{\alpha}_1, \dots, \overline{\alpha}_N, \overline{\beta}_0, \dots, \overline{\beta}_N$) en \tilde{X} tengamos:

$$I(\tilde{H}, \overline{H}) = \sum_{i=1}^N \alpha_i P \overline{\beta}_i + \overline{\alpha}_i P \beta_i = 0.$$

Como $I(\tilde{H}, K) = 0$ para todo K en X' , tendremos: $I(\tilde{H}, H) = 0$, para todo H en X . Esto implica que \tilde{H} es un punto crítico de $I(H)$; las condiciones de continuidad de Weierstrass implican entonces que: $\beta_1 = \dots = \beta_N = 0$, $\tilde{H}(x) = B_0(x)\beta_0$ y b es un punto conjugado de a , con la misma multiplicidad.

Q.E.D.

Consideremos ahora el funcional $I_c(H)$ sobre X_c . Sea $i(c)$ el índice de $I_c(H)$. Tenemos entonces:

Lema 3.4 :

- a) $i(c)$ es una función (en los naturales) no decreciente.
- b) $i(c) = 0$ para c en $]a, c_1]$.
- c) $i(c - \epsilon) = i(c)$, para $\epsilon \geq 0$ pequeño.
- d) Si c es conjugado de a con multiplicidad $\nu(c)$, entonces $i(c + \epsilon) = i(c) + \nu(c)$, para $\epsilon > 0$ pequeño. Hay un número finito de puntos conjugados de a en $[a, b]$.

Prueba : a) Si $c < c'$, entonces existe un subespacio de X_c de dimensión $i(c)$ tal que $I_c(H) < 0$, para H en ese espacio. Si para cada H en X_c , extendemos H por 0 en $[c, c']$, tendremos un elemento de $X_{c'}$ y $I_{c'}(H) = I_c(H) < 0$. Evidentemente la independencia lineal se mantiene y $i(c') \geq i(c)$.

b) Como no hay puntos conjugados de a en $]a, c_1]$, entonces, por la primera parte del teorema 3.2, $I(H) > 0$ para todo H en X_c , (extendiendo H como 0 en $[c, c_1]$). Como $i(b)$ es finito, la función monótona (en \mathbb{N}) $i(c)$ tiene solamente un número finito de saltos, en x_1, \dots, x_M , por lo tanto $i(c) = i(a) = 0$ para c en $]a, c_1]$.

c) Podemos escoger c_1, \dots, c_N entre a y b tales que no coincidan con los saltos y que no haya más de un salto en cada intervalo. Sea c un salto entre c_i y c_{i+1} : $X_c = \tilde{X}_c \oplus X'_c$, como en el lema 3.2, donde \tilde{X}_c está determinado por los valores de las soluciones a la ecuación variacional en c_1, \dots, c_i . $\tilde{X}_c \cong \mathbb{R}^\epsilon$. Para c' vecino a c podemos mantener la misma subdivisión y por lo tanto $\tilde{X}_{c'} = \tilde{X}_c$.

Ahora bien $I_c(\tilde{H}) = \sum_{j=1}^i \alpha_j^T P \beta_j$ (P valuado en cada c_j) donde, para $\alpha_1, \dots, \alpha_i$ dados, $\beta_1, \dots, \beta_{i-1}$ no dependen de c (ver la definición de β_j).

$$\beta_i = -B_i^{-1}(c) \sum_0^{i-1} B_j(c) \beta_j.$$

Por lo tanto:

$$I_c(\tilde{H}) - I_{c'}(\tilde{H}) = -\alpha_i^T P \sum_0^{i-1} (B_i^{-1}(c) B_j(c) - B_i^{-1}(c') B_j(c')) \beta_j$$

es decir, $I_c(\tilde{H})$ es continuo en c .

Entonces si $I_c(\tilde{H}) < 0$, también $I_{c'}(\tilde{H})$ para c' vecino a c : $i(c') \geq i(c)$.

Pero como $i(c') \leq i(c)$ para $c' \leq c$, tenemos $i(c) = i(c')$ para $c' \leq c$, c' vecino a c . Si c no es un salto, entonces el resultado es evidente.

d) Sea c entre c_i y c_{i+1} . Entonces $\tilde{X}_c \cong \mathbb{R}^{ni}$ es independiente de c y, por el lema 3.2, la forma cuadrática $I_c(\tilde{H})$ es definida positiva sobre un subespacio de \tilde{X}_c de dimensión

$$p(c) \equiv ni - i(c) - \nu(c).$$

Repitiendo el argumento de continuidad de (c) vemos que: $p(c') \geq p(c)$, para c' vecino a c . (No se puede repetir el argumento de (a) ya que la extensión como 0 entre c y c' no da un elemento de $\tilde{X}_{c'}$.)

Por lo tanto: $i(c') \leq i(c') + \nu(c') \leq i(c) + \nu(c)$ para c' vecino de c .

Sean ahora $H_1, \dots, H_{i(c)}, i(c)$ elementos linealmente independientes de X_c tales que la matriz $(I_c(H_i, H_j))_{ij}$ sea negativa definida.

(Por el lema 3.1 podemos escoger esos H_i en \tilde{X}_c : H_i son soluciones de la ecuación variacional en $[a, c_1] \dots, [c_i, c]$ con $H_i(a) = H_i(c) = 0$.) Sean $\tilde{H}_1, \dots, \tilde{H}_{\nu(c)}, \nu(c)$ elementos linealmente independientes de \tilde{X}_c tales que $I(\tilde{H}_i, H) = 0$ para todo H en X_c . Por lo tanto $\tilde{H}_j = B_0(x) \beta_{0j}$, $\beta_1, \dots, \beta_i = 0$ y $B_0(c) \beta_{0j} = 0$.

Ahora bien $\tilde{H}'_j(c)$ son linealmente independientes (si $\sum \lambda_j \tilde{H}'_j(c) = 0$ entonces $H(x) = \sum \lambda_j \tilde{H}_j(x)$ es solución de la ecuación variacional, vale 0 en c así como su derivada: por la unicidad sería idénticamente 0, lo cual contradice la independencia lineal: $H'(a) = 0 = \sum \lambda_j \tilde{H}'_j(a) = \sum \lambda_j \beta_{j0}$, pero los β_{j0} son linealmente independientes). También lo son

$P(c)\tilde{H}'_j(c)$, ya que $P(c)$ es invertible. Podemos completar este conjunto de vectores por vectores $\alpha_{\nu(c)+1}, \dots, \alpha_n$, obteniendo una matriz invertible cuyas $\nu(c)$ primeras columnas son $P(c)\tilde{H}'_j(c)$. La matriz inversa tendra como $\nu(c)$ primeros renglones los vectores $K_1, \dots, K_{\nu(c)}$ tales que $K_i^T P(c)\tilde{H}'_j(c) = \delta_{ij}$.

Sea c' entre c y c_{i+1} . Sea $K_j(x)$ una función (en \mathbb{R}^n) C^2 tal que:

$$K_j(x) \equiv 0 \text{ para } x \leq c_i, K_j(c') = 0, K_j(c) = K_j.$$

Extenderemos \tilde{H}_j y H_j como 0 entre c y c' . Obtenemos entonces elementos de $X_{c'}$.

$I_{c'}(H_k, H_j) = I_c(H_k, H_j)$ es negativa definida.

$$I_{c'}(H_k, \tilde{H}_j) = I_c(H_k, \tilde{H}_j) = 0 = I_{c'}(\tilde{H}_k, \tilde{H}_j).$$

$$I_{c'}(H_k, K_j) = \int_{c_i}^c (K_j^T C H_k + K_j^T A H'_k + K_j^{T'} A^T H_k + K_j^{T'} P H'_k) dx$$

(por las definiciones de H_k y K_j).

$$I_{c'}(H_k, K_j) = K_j^T P H'_k(c)$$

ya que H_k es solución del ecuación variacional en $[a, c']$ (por pedazos), $K_j(c_i) = 0$ y $H'_k(c_+) = 0$.

Del mismo modo $I_{c'}(\tilde{H}_k, K_j) = K_j^T P \tilde{H}'_k(c) = \delta_{jk}$ por construcción. $I_{c'}(K_k, K_j)$ es fijo.

Entonces, si consideramos $I_{c'}$ sobre el espacio generado por:

$$H_1, \dots, H_{i(c)}, \tilde{H}_1/\epsilon - \epsilon K_1, \dots, \tilde{H}_{\nu(c)}/\epsilon - \epsilon K_{\nu(c)}$$

tendremos la matriz

$$\begin{pmatrix} I_c(H_k, H_j) & -\epsilon K_j^T P H'_k(c) \\ -\epsilon K_k^T P H'_j(c) & -2I + \epsilon^2 I_{c'}(K_k, K_j) \end{pmatrix}$$

ya que $I_{c'}(\tilde{H}_k/\epsilon - \epsilon K_k, \tilde{H}_j/\epsilon - \epsilon K_j) = -2\delta_{jk} + \epsilon^2 I_{c'}(K_k, K_j)$.

Claramente, para ϵ pequeño, esa matriz será negativa definida. Por lo tanto los $i(c) + \nu(c)$ vectores que hemos encontrado son linealmente independientes y pertenecen a $X_{c'}$, es decir $i(c') \geq i(c) + \nu(c)$.

Tenemos entonces $i(c') = i(c) + \nu(c)$ para todo c' entre c y c_{i+1} .

Como hemos escogido la subdivisión c_1, \dots, c_N , de tal modo que no hay más de un salto por subintervalo $[c_i, c_{i+1}]$, y como cada punto conjugado de a , entre c_i y c_{i+1} , produce un salto,

tenemos que a lo más hay puntos conjugados en c_1, \dots, c_N , y a lo más uno por subintervalo: hay un número finito de puntos conjugados de a entre a y b . Finalmente si c no es conjugado de a , entonces $i(c') = i(c)$ ($\nu(c) = 0$) y no hay salto de $i(c)$.

Q.E.D.

Fin de la prueba del teorema del índice :

Ya hemos probado (a) y (b). Como hay un número finito de puntos conjugados, podemos escoger c_1, \dots, c_N , tales que ninguno de ellos sea un punto conjugado de a y que en $[c_i, c_{i+1}]$ no haya más de un salto de $i(c)$ (es decir un punto conjugado de a). Tendremos $i(b) = i(a_+) + \sum$ (saltos) y (como $i(a_+) = 0$) $i(b) = \sum \nu(c_i)$.

Q.E.D.

6.4 Puntos conjugados y campos centrales.

En esta sección veremos una manera conveniente de caracterizar los puntos conjugados (por lo menos el primero).

Recordemos que para el funcional

$$J(Y) = \int_a^b F(x, Y, Y') dx$$

con $Y(a) = A$, $Y(b) = B$, un campo central, con foco en (a, A) , es una familia $Y(x, s_1, \dots, s_n)$ de puntos críticos (con solo $Y(a) = A$) tal que las matrices

$$F_{ZZ}(x, Y(0), Y'(0))$$

$$\left(\frac{\partial y'_i}{\partial s_j}\right)_{ij}(a, 0, \dots, 0)$$

$$\left(\frac{\partial y_i}{\partial s_j}\right)_{ij}(x, 0, \dots, 0)$$

con $x > a$, son invertibles. El candidato a extremo es $Y(x, 0, \dots, 0)$. Aquí suponemos que $F_{ZZ}(x, Y(x, 0), Y'(x, 0))$ es definida positiva. (Ver la definición V.6.2.) Recordemos que toda familia de soluciones con $Y(a, s_1, \dots, s_n) = A$, $(\partial y'_i / \partial s_j)_{ij}(a, 0, \dots, 0)$ invertible, es un campo central (Lema V.6.3) para x cerca de a . También hemos visto que para x vecino a a siempre existe un campo central (solución de la ecuación de Hamilton-Jacobi) y que se puede construir explícitamente: nota V.6.3.

Teorema 4.1 : *El primer punto conjugado de a entre a y b es el primer punto donde $(\partial y_i / \partial s_j)_{ij}(x, 0, \dots, 0)$ es singular para cualquier campo central en $[a, x]$.*

Prueba : Sea $Y(x, s_1, \dots, s_n)$ el campo central, con $Y(x, 0, \dots, 0) \equiv Y_0(x)$ el candidato a extremo. Sea $Y_S(x)$ la matriz $(\partial y_i / \partial s_j)(x, 0, \dots, 0)$.

Como $\frac{d}{dx}(F_Z(x, Y(x, S), Y'(x, S))) = F_Y(x, Y(x, S), Y'(x, S))$ para todo S , podemos diferenciar estas ecuaciones con respecto a s_i y obtenemos, en $S = 0$:

$$\frac{d}{dx}(F_{ZY}Y_S + F_{ZZ}Y'_S) = F_{YY}Y_S + F_{YZ}Y'_S$$

es decir:

$$(A^T Y_S + P Y'_S)' = C Y_S + A Y'_S$$

o bien,

$$(P Y'_S)' - (A - A^T) Y'_S - (C - A^T) Y_S = 0$$

es decir cada columna de Y_S es solución de la ecuación variacional (ver la definición 3.1).

Como $Y(a, S) = A$, $Y_S(a, 0) = 0$ y Y_S es una matriz solución de la ecuación variacional, invertible para x vecino (pero distinto) de a (es una matriz fundamental), $Y_S Y_S'^{-1}(a)$ es también solución con valor I en a para su derivada. $Y_S(x) = B(x) Y'_S(a)$ (por la unicidad) donde $B(x)$ es la matriz fundamental que ya hemos usado. $Y_S(x)$ es singular si y sólo si $B(x)$ lo es, es decir en los puntos conjugados de a . Además estos puntos no dependen de la familia $Y(x, s_1, \dots, s_n)$.

Q.E.D.

Ejercicio 4.1 : En el ejemplo de las geodésicas sobre una esfera de radio 1, ejemplo 2.3, se encontró una familia de puntos críticos:

$$\theta(\varphi, \alpha) = \text{Arctan} \left((\alpha \cos(\varphi - \varphi_1 + \text{Arccos}(\alpha^{-1} \cot \theta_1)))^{-1} \right)$$

con $\alpha = (\theta_1'^2 + \sin^2 2\theta_1/4)^{1/2} \sin^2 \theta_1$.

Probar que $\frac{\partial \theta}{\partial \alpha} = g(\alpha) \sin(\varphi - \varphi_1)$ para cierta función $g(\alpha)$. Por lo tanto el primer punto donde $\frac{\partial \theta}{\partial \alpha} = 0$ es para $\varphi = \varphi_1 + \pi$. (Desarrollar el coseno de la suma).

Ejercicio 4.2 : Probar que la familia $y \pm \alpha x$ es un campo para el problema del ejemplo 2.1.

Nota 4.1 : Podemos dar otra demostración del teorema 3.2 (no del teorema del índice).

Supongamos primero que Y_0 es un mínimo local. Construyamos un campo central $Y(x, S)$, con foco en (a, A) , como en la nota V.6.3., es decir, la única solución, para x vecino a a , al sistema:

$$Y'(x) = Z(x) \quad Y(a) = A$$

$$Z'(x) = G(x, Y, Z) \quad Z(a) = S + Y'_0(a).$$

Aquí la matriz C de la nota V.6.3 es la identidad y c_0 es $y'_0(a)$. $Z_S(a) = I$.

Si c es el primer punto conjugado de a , con $c < b$, entonces $Y_S(0)$ es singular en c . Sea β un vector propio de $Y_S(c, 0)$, con $Y_S(c, 0)\beta = 0$. Sea $\tilde{H}(x) = Y_S(x, 0)\beta$, para $x \leq c$, $\tilde{H}(x) = 0$ para $x \geq c$. Entonces $\tilde{H}(x)$ pertenece a X , como $Y_S(x, 0)$ es una matriz solución de la ecuación variacional, también lo es $\tilde{H}(x)$, en $[a, c[$ y $]c, b]$. Por lo tanto $I(\tilde{H}) = \tilde{H}^T(c)[P\tilde{H}' + A^T\tilde{H}]_c = 0$.

Como $I(H) \geq 0$ en X , entonces \tilde{H} es un mínimo de I sobre X : debe satisfacer la condición de continuidad de Weierstrass $[P\tilde{H}' + A^T\tilde{H}]_c = 0$, por lo tanto $\tilde{H}'(c) = Y'_S(c, 0)\beta = 0$. Pero entonces $\tilde{H}(x) \equiv 0$, por la unicidad de la solución al problema de valores iniciales en c . Como $\tilde{H}'(a) = Y'_S(a, 0)\beta$, y $Y'_S(a, 0)$ es invertible, entonces $\beta = 0$, lo cual es una contradicción. Notemos que este argumento funciona también (sobre X) reemplazando $Y_S(x, 0)$ por $B(x)$.

Vemos también que si $I(H) > 0$, el punto b también es eliminado como punto conjugado de a (si no habría un H en X tal que $I(H) = 0$).

El inverso viene del siguiente resultado y de la aplicación de los resultados de la siguiente sección.

Teorema 4.2 : *Si P es definida positiva, F es C^3 , Y_0 es un punto crítico de $H(Y)$, entonces:*

No hay puntos conjugados de a en $[a, b]$ si y sólo si Y_0 es parte de un campo central con foco en $(a - \epsilon, Y_0(a - \epsilon))$, para todo $\epsilon > 0$, pequeño.

Prueba : Notemos primero que como Y_0 es C^2 podemos extender Y_0 como única solución de las ecuaciones de Euler con condiciones $Y_0(a)$, $Y'_0(a)$ entre $a - \epsilon$ y a .

Si Y_0 es parte de un campo central con foco en $(a - \epsilon, Y_0(a - \epsilon))$ entonces no hay puntos conjugados de $a - \epsilon$ en $[a - \epsilon, b]$. Ya que este hecho es independiente del campo, podemos suponer que el campo es solución de:

$$Y'(x) = Z(x) \quad Y(a - \epsilon) = Y_0(a - \epsilon)$$

$$Z'(x) = G(x, Y, Z) \quad Z(a - \epsilon) = S + Y'_0(a - \epsilon).$$

Claramente la solución $Y(x, S, \epsilon)$ es diferenciable en ϵ . Por lo tanto para cada $x > a$ fijo, $Y_S(x, 0, 0) = \lim Y_S(x, 0, \epsilon)$ es invertible. Tenemos un campo central saliendo de $(a, Y_0(a))$ con $Y_S(x, 0)$ invertible para todo x , $a < x \leq b$.

Inversamente, el mismo argumento prueba que $Y_S(x, 0, \epsilon)$ es invertible para $x > a + \alpha$ para todo $\epsilon \leq \epsilon(\alpha)$. Pero como los puntos conjugados no pueden estar cerca (lema 3.1), podemos tomar α tan pequeño que $|a + \alpha - (a - \epsilon)| < k$ (k dado en el lema 3.1), (por ejemplo $\alpha \leq k/2$ y tomaremos $\epsilon \leq \min(\epsilon/\alpha, k/2)$). Por lo tanto no hay puntos conjugados de $a - \epsilon$ en $[a - \epsilon, b]$ y $Y(x, S, \epsilon)$ es un campo central.

Q.E.D.

Ejemplo 4.1 : Sea $I(y) = \int_0^a (1 + y'^2)^{1/2} / y \, dy$, con $y(0) = d > 0$, $y(a) = \alpha > 0$.

La ecuación de Euler da:

$$yy'' + 1 + y'^2 = 0 = (y^2)''/2 + 1.$$

Por lo tanto, con las condiciones iniciales:

$$y^2 + x^2 = 2cx + d^2 \quad \text{ó} \quad y^2 + (x - c)^2 = c^2 + d^2.$$

(Para la solución $c = \frac{a^2 + \alpha^2 - d^2}{2a}$)

Tenemos entonces una familia de círculos, con centros en $(c, 0)$ y pasando por $(0, d)$: $y(x) = (d^2 + 2cx - x^2)^{1/2}$.

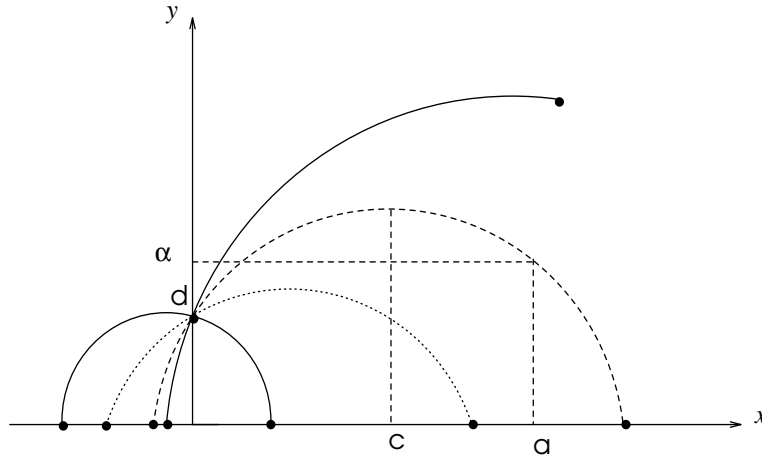


Fig. 4.1

Como $\frac{\partial y}{\partial c}(x, c) = x/y \neq 0$ para $x \neq 0$, no hay puntos conjugados de 0 en $[0, a]$ si $\alpha > 0$. Esta familia es un campo central con foco en $(0, d)$ (también se puede ver esto directamente ya que por cada punto (x, y) , $x > 0$ y $y > 0$, pasa uno y sólo un círculo).

Además, $F_{zz} = y^{-1}(1 + z^2)^{-3/2} > 0$, por lo tanto la solución es un mínimo local en la norma C^1 .

Notemos que si $d = 0$, la ecuación de Euler tiene también una solución

$$x^2 + y^2 = 2cx.$$

Pero $y' = (c - x)/y$, $1 + y'^2 = c^2/y^2$ y

$$\int_0^a \frac{(1 + y'^2)^{1/2}}{y} dx = c \int_0^a \frac{dx}{2cx - x^2} = \frac{1}{2} \log(x(2c - x)^{-1})|_0^a = -\infty$$

es decir, el funcional no está definido para el candidato a mínimo.

Ejemplo 4.2 : En el ejemplo 2.4, tratamos los problemas de frontera libre en a , llegando a la ecuación variacional:

$$(Ph')' = Qh \quad h(b) = 0, \quad Ah'(a) + Bh(a) = 0.$$

Ahora bien si tenemos una familia de soluciones de:

$$(F_{y'})' - F_y = 0, \quad F_{y'}(a, y(a, s), y'(a, s)) = 0, \quad y(b, s) = B$$

derivando la ecuación de Euler con respecto a s obtenemos que y_s es solución de la ecuación variacional (como antes). Derivando la condición de transversalidad: $F_{yz}y_s + F_{zz}y'_0 = 0$, es decir $Ay'_s + Py_s|_a = 0$. y_s es solución de la ecuación variacional con sus condiciones de frontera.

Además (ejercicio) los puntos conjugados de a , en el sentido de la definición 2.2, son exactamente los puntos donde $y_s(x, 0) = 0$.

6.5 Condiciones suficientes en la norma C^0

Veremos finalmente cómo en algunos casos se puede usar la convexidad de F en Z para probar que no solamente el punto crítico Y_0 es un mínimo en la norma C^1 (es decir para $|Y - Y_0|_1$ pequeño) pero también en la norma C^0 (es decir para Y en C^1 con $|Y - Y_0|_0$ pequeño: las derivadas de Y pueden ser muy diferentes de las de Y_0).

Supondremos que Y_0 es parte de una familia de soluciones de las ecuaciones de Euler $Y(x, s_1, \dots, s_n) = Y(x, S)$, tal que $Y_0(x) = Y(x, 0)$, y $Y_S(x, 0)$ es invertible para x en $[a, b]$ (por ejemplo si tomamos un campo central con foco en $(a - \epsilon, Y_0(a - \epsilon))$). Supondremos además que $Z = Y'(x, S)$ es un campo exacto: como Y_S es invertible entonces para cada (x, Y) vecino a $(x, Y_0(x))$, podemos invertir S como función de Y . El campo es entonces $Z = Y'(x, S(x, Y)) = Z(x, Y)$ (ver la sección V.6.6). Como supusimos el campo exacto entonces

$$I(\gamma) = \int_{\gamma} F(x, Y, Z(x, Y)) dx$$

no depende del camino entre dos puntos en \mathbb{R}^{n+1} y tenemos la fórmula de Hilbert:

$$J(Y) - J(Y_0) = \int_a^b E(x, Y, Z, Y') dx$$

donde $E(x, Y, Z, Y')$ es la función de Weierstrass y está dada por:

$$E(x, Y, Z, Y') = F(x, Y, Y') - F(x, Y, Z) - F_Z(x, Y, Z) \cdot (Y' - Z).$$

Es decir, en cada punto de $Y(x)$ se compara $Y'(x)$ con la dirección del campo $Z(x, Y(x))$ (ver figura 5.1).

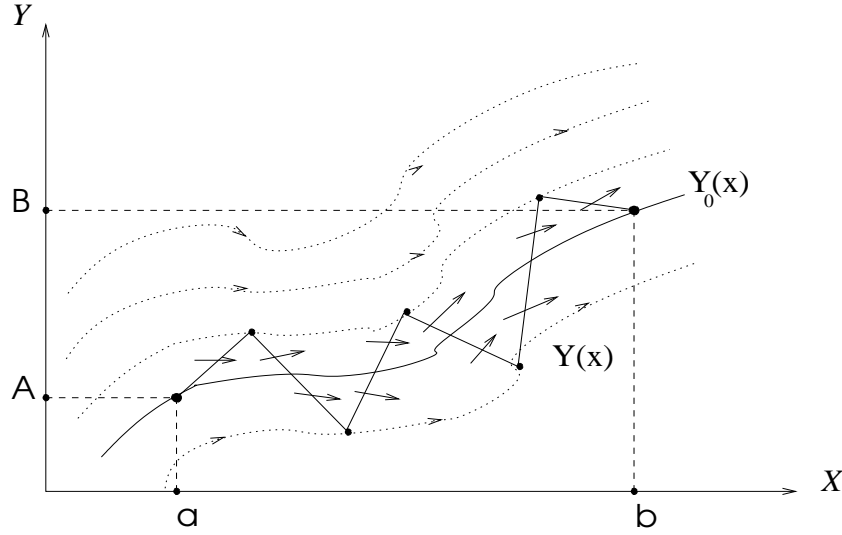


Fig. 5.1

Tenemos entonces el siguiente resultado.

Teorema 5.1 : a) Si Y_0 es una trayectoria de un campo exacto $Z(x, Y)$, tal que $E(x, Y(x), Z(x, Y(x)), W) \geq 0$ para todo W en \mathbb{R}^n entonces Y_0 es un mínimo, en la norma C^0 , para todo $Y(x)$ en el dominio del campo, $Y(x)$ es C^1 por pedazos.

b) Si $E(x, Y, Z, W) \geq 0$ con igualdad sólo si $W = Z$, entonces Y_0 es el único mínimo.

c) Si $F(x, Y, Z)$ es C^2 , la condición de (a) es $F_{ZZ}(x, Y, Z)$ es definida no negativa para todo $Y(x)$ cerca de $Y_0(x)$ y para toda Z en \mathbb{R}^n . La condición de (b) es que $F_{ZZ}(x, Y(x), Z)$ sea definida positiva.

d) Si $F_{ZZ}(x, Y_0(x), Y'_0(x))$ es positiva definida y $Y_0(x)$ es parte de un campo exacto, entonces Y_0 es un único mínimo local para la norma C^1 , en el espacio de funciones C^1 por pedazos.

Prueba : (a) es evidente.

Para (b) basta ver que si hay dos mínimos, entonces $J(Y) = J(Y_0)$ y por lo tanto $Y'(x) = Z(x, Y(x))$. Es decir, $Y(x)$ es parte de las trayectorias del campo. Pero como sólo hay una que pasa por (a, A) , $Y(x) = Y_0(x)$.

c) Si F es C^2 , entonces:

$$E(x, Y, Z, W) = \int_0^1 (W - Z)^T F_{ZZ}(x, Y, Z + t(W - Z))(W - Z)(1 - t) dt$$

donde se usó el hecho que si $f(t) = F(x, Y, Z + t(W - Z))$, entonces

$$f(1) - f(0) = \int_0^1 f'(t) dt = \int_0^1 (1 - t)f''(t) dt + f'(0)$$

(integrando por partes).

Aquí $f(1) = F(x, Y, W)$, $f(0) = F(x, Y, Z)$, $f'(0) = F_Z(x, Y, Z) \cdot (W - Z)$, $f''(t)$ es el integrando.

Entonces, si F_{ZZ} es definida no negativa, $E(x, Y, Z, W) \geq 0$, mientras que si $F_{ZZ}(x, Y, Z)$ es negativa para algún (x, Y, Z_0) (es decir tiene un vector propio U para el cual la forma cuadrática es negativa), tomando $W = Z_0 + \epsilon U$, tendremos $E(x, Y, Z_0, W) < 0$.

(d) Es obvio ya que entonces $F_{ZZ}(x, Y, Z)$ será positiva definida para (Y, Z) vecinos de $(Y_0(x), Y'_0(x))$.

Q.E.D.

Notemos en particular que si no hay puntos conjugados de a en $[a, b]$ (y F_{ZZ} es invertible), por el teorema 4.2, Y_0 es parte de un campo central con foco en $(a - \epsilon, Y_0(a - \epsilon))$, por lo tanto $E > 0$ y tenemos un único mínimo local en la norma C^1 . Esto termina la prueba del teorema 3.2 usando campos.

Ejemplo 5.1 : El máximo para el problema isoperimétrico encontrado en el ejemplo 2.2 es un máximo en la norma C^0 , y es global.

Ejemplo 5.2 : La geodésica sobre la esfera es un mínimo para la norma C^0 si (φ_1, θ_1) es unido a (φ_2, θ_2) por un camino que cubra menos de media esfera.

Ejemplo 5.3 : El mínimo encontrado en el ejemplo 4.1 es un mínimo global en la norma C^0 .

Ejercicio 5.1 : Probar un resultado análogo para el problema de frontera libre del ejemplo 4.2.

Ejercicio 5.2 : Estudiar el problema del ejercicio 2.3.

Notemos que la positividad de E es también necesaria:

Teorema 5.2 : Si $Y_0(x)$ es un mínimo local, en la norma C^0 , para el espacio de funciones C^1 por pedazos, entonces $E(x, Y_0(x), Y'_0(x), W) \geq 0$ para todo W en \mathbb{R}^n (F, F_Y, F_Z son continuos). $F_{ZZ}(x, Y_0(x), W)$ es definida no negativa si F es C^2 , para todo W en \mathbb{R}^n .

Prueba : Podemos escribir: $F(x, Y, Y') - F(x, Y_0, Y'_0)$ como

$$\begin{aligned} & F(x, Y, Y') - F(x, Y_0, Y') + F(x, Y_0, Y') - F(x, Y_0, Y'_0) \\ &= F_Y(x, Y_0, Y')(Y - Y_0) + o(|Y - Y_0|) + F_Z(x, Y_0, Y'_0)(Y' - Y'_0) + o(|Y' - Y'_0|) \end{aligned}$$

donde los residuos son uniformes en x en $[a, b]$, $Y_0(x), Y'_0(x)$ si $|Y'(x) - Y'_0(x)|, |Y(x) - Y_0(x)|$, son acotados.

Sea c entre a y b , h positivo, sea W un vector de \mathbb{R}^n . Definamos

$$Y(x) = Y_0(x) + H(x), \quad \text{con } H(x) = \begin{cases} 0 & \text{si } x \leq c \\ (x - c)W & \text{si } c \leq x \leq c + h \\ hW(b - x)/(b - c - h) & \text{si } x \geq c + h \end{cases}.$$

$H(x)$ es C^1 en pedazos, $H(b) = 0$. (Para $c = b$, se tomará $H(x)$ con un pico en $c - h$.) $|H(x)| \leq h|W|$, por lo tanto para h pequeño $J(Y) - J(Y_0) \geq 0$ si Y_0 es un mínimo en la norma C^0 . Escribiendo la integral como tres integrales de $F(x, Y, Y') - F(x, Y_0, Y'_0)$, la contribución sobre $[a, c]$ es 0. Sobre $[c, c + h]$ tenemos:

$$\begin{aligned} \int_c^{c+h} [F(x, Y_0 + (x - c)W, Y'_0 + W) - F(x, Y_0, Y'_0 + W) + F(x, Y_0, Y'_0 + W) - F(x, Y_0, Y'_0)] dx \\ = o(h) + h (F(c, Y_0(c), Y'_0(c) + W) - F(c, Y_0(c), Y'_0(c))). \end{aligned}$$

Sobre $[c + h, b]$ tenemos:

$$\int_{c+h}^b (F_Y(x, Y_0, Y') \cdot H + F_Z(x, Y_0, Y'_0) \cdot H') dx + o(h)$$

ya que $H(x)$ es lineal en x , en ese intervalo, y $H'(x) = -hW/(b - c - h)$.

Ahora, como Y_0 es un punto crítico, F_Z es C^1 por partes, con derivada $F_Y(x, Y_0, Y'_0)$, excepto en los puntos de discontinuidad de Y'_0 , donde sólo $F_Z(x, Y_0, Y'_0)$ es continuo. Podemos entonces integrar por partes, sobre cada intervalo de continuidad de Y'_0 y los términos de frontera en una discontinuidad de Y'_0 se cancelan, ya que $H(x)$ es C^1 .

Tenemos:

$$\begin{aligned} \int_{c+h}^b (F_Y(x, Y_0, Y') - F_Y(x, Y_0, Y'_0)) \cdot H dx - \\ - F_Z(c + h, Y_0(c + h), Y'_0(c + h)) \cdot H(c + h) + o(h) \\ = -F_Z(c, Y_0(c), Y'_0(c)) \cdot H(c + h) + o(h) \end{aligned}$$

por la continuidad de $F_Z(x, Y_0, Y'_0)$, $H(c + h) = hW$. Dividiendo por h ($h > 0$) y tomando el límite cuando h tiende a 0:

$$\begin{aligned} 0 \leq F(c, Y_0(c), Y'_0 + W) - F(c, Y_0(c), Y'_0(c)) - F_Z(c, Y_0(c), Y'_0(c)) \cdot W \\ = E(c, Y_0(c), Y'_0(c), Y'_0(c) + W). \end{aligned}$$

Q.E.D.

Notemos que si sabemos que F_{ZZ} es definido positivo sobre $Y_0(x)$, entonces, dado que un mínimo en la norma C^0 lo será en la norma C^1 , no hay puntos conjugados de a en $[a, b[$ (Teorema 3.2). Y_0 es parte de un campo central, con foco (a, A) , que puede tener una singularidad en b .

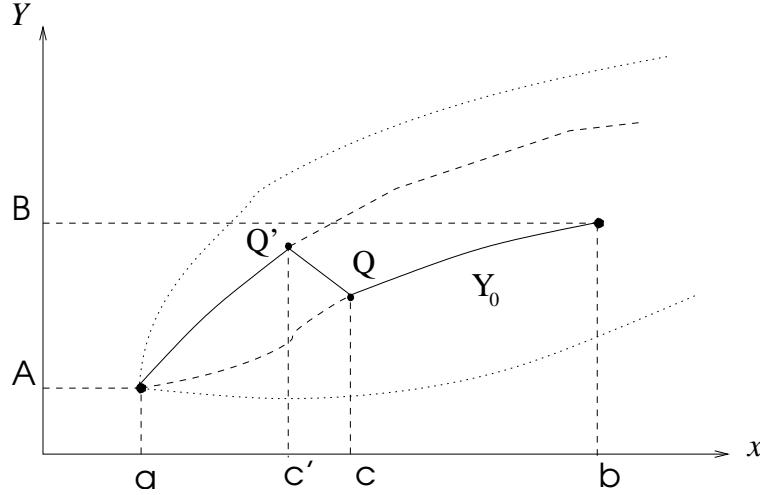


Fig. 5.2

Supongamos que $E(c, Y_0(c), Y'_0(c), W) < 0$ para algún vector W . Sea $c' < c$ tal que $E(x, Y(x), Z(x, Y(x)), W) < 0$, para $c' \leq x \leq c$, donde $Y(x)$ es parte del campo y vecino a $Y_0(x)$.

Sea $Q' = Y_0(c) + W(c' - c)$. Escojamos una curva $AQ'QB$, tal que la porción AQ' sea la trayectoria del campo $Y(x, c', Q')$ que pasa por (c', Q') , la porción $Q'Q$ sea $Y(x) = Y_0(c) + W(x - c)$, y la porción QB sea $Y_0(x)$. Entonces

$$\int_a^b E(x, Y, Z, Y') dx = \int_{c'}^c E(x, Y, Z, Y') dx < 0,$$

ya que sobre (a, c') y (c, b) , $Y'(x) = Z(x, Y(x))$: E es 0.

Notemos que si fijamos x , entonces

$$\begin{aligned} E(x, Y(x), Z(x, Y(x)), W) &= F(x, Y(x), W) - F(x, Y(x), Z(x, Y(x))) \\ &\quad - F_Z(x, Y(x), Z(x, Y(x))) \cdot (W - Z(x, Y(x))) \\ &\equiv E(W) \end{aligned}$$

$$F(x, Y(x), W) \equiv F(W)$$

$F = F(W)$ es una superficie en \mathbb{R}^{n+1} . La ecuación del plano tangente a esa superficie en $Z(x, Y(x))$ es:

$$F = F(Z(x, Y(x))) + F_Z(Z(x, Y(x))) \cdot (W - Z(x, Y(x))) \equiv F_T(W)$$

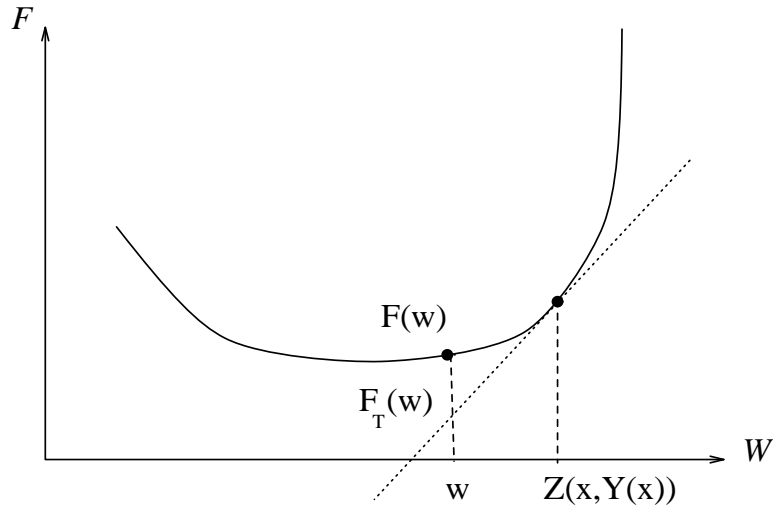


Fig. 5.3

Por lo tanto, decir que $E \geq 0$ en $(x, Y(x))$ para todo W quiere decir que $F(W) - F_T(W) \geq 0$: la superficie está por encima del plano tangente.

Ahora si $E(x, Y(x), Z, W) \geq 0$ para todo Z, W , x fijo, quiere decir que la superficie es convexa. Esto es el caso, si F es C^2 , si y sólo si $F_{ZZ}(x, Y(x), Z)$ es positiva definida, para convexidad estricta, o definida no negativa, para convexidad simple.

Si x varía, entonces tendremos una familia de superficies parametrizadas por x y una curva $W = Z(x, Y(x))$ tal que para cada x , la superficie correspondiente está por encima del plano tangente en el punto $(x, Z(x, Y(x)), F(x, Y(x), Z(x, Y(x))))$.

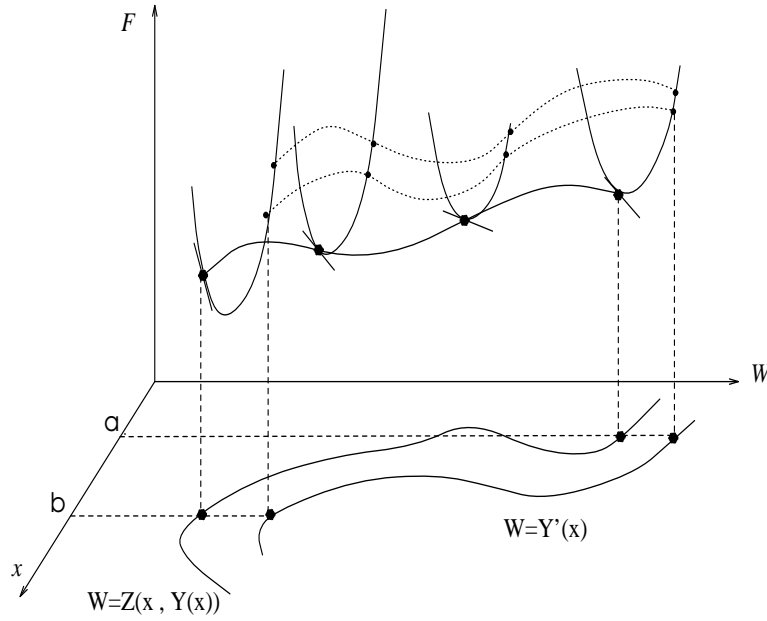


Fig. 5.4

Vemos así que podemos usar directamente la convexidad en Z de $F(x, Y, Z)$ para cada (x, Y) fijo para probar la existencia de mínimos. Esto es útil cuando la función F no es C^2 .

Ejemplo 5.4 : En el ejemplo 2.1, vimos que el funcional

$$J(y) = \int_0^1 (1 + y'^2)^{-1} dx, \quad y(0) = 0, y(1) = 1$$

tiene un único punto crítico $y(x) = x$, el cual es un mínimo local para la norma C^1 . Sin embargo $F_{y'y'} = 2(2y'^2 - 1)(1 + y'^2)^{-3}$ por lo tanto $F_{zz}(x, y, w) < 0$ si $w^2 < 1/2$: el mínimo en C^1 no es un mínimo para la norma C^0 . Notemos que esto ya fue probado directamente en el Ejercicio III.3.

Ejercicio 5.3 : Probar que el extremo encontrado en el ejercicio 2.4 no es un mínimo para la norma C^0 .

Ejemplo 5.5 : La Braquistocrona.

Para el funcional

$$J(y) = \int_0^a \frac{(1 + y'^2)^{1/2}}{y^{1/2}} dx,$$

con $y(0) = 0$, $y(a) = A$, tenemos una familia de soluciones de la ecuación de Euler:

$$x(\theta) = c(\theta - \sin \theta)$$

$$y(\theta) = c(1 - \cos \theta)$$

(ver el ejemplo IV.2.1).

Además en el primer capítulo se probó que para cada $x > 0$, $\alpha > 0$, hay una y una sola cicloide que une el origen con (x, α) .

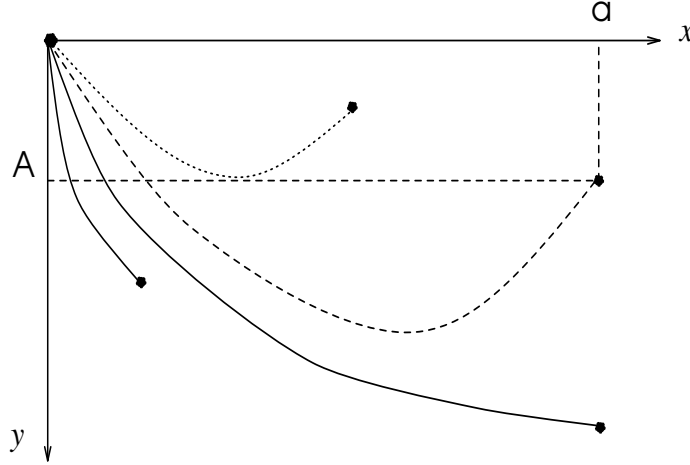


Fig. 5.5

Además como $\frac{\partial y(x,c)}{\partial c} = 1 - \cos \theta + c \operatorname{sen} \theta \frac{\partial \theta}{\partial c}$ (ya que θ es función de x y de c).

De la relación, $c(\theta(x, c) - \operatorname{sen} \theta(x, c)) - x = 0$, obtenemos, derivando con respecto a c :

$$\theta_c = -(\theta - \operatorname{sen} \theta)/c(1 - \cos \theta).$$

Por lo tanto $y_c = (2 - 2 \cos \theta - \theta \operatorname{sen} \theta)/(1 - \cos \theta)$.

Si ponemos $F(\theta) = 2 - 2 \cos \theta - \theta \operatorname{sen} \theta$, $F'(\theta) = \operatorname{sen} \theta - \theta \cos \theta$ se anula si $\theta = \tan \theta$ es decir, para θ en $]0, 2\pi[$ en un sólo punto θ_0 entre π y $3\pi/2$ (graficar $\tan \theta$ y θ), para el cual $\cos \theta_0 < 0$, $\operatorname{sen} \theta_0 < 0$. Entonces, $F(\theta_0) > 0$. Como $F(0) = F(2\pi) = 0$, tenemos un sólo extremo de $F(\theta)$ entre 0 y 2π , en θ_0 ; $F(\theta) > 0$ para θ en $]0, 2\pi[$. Vemos entonces que $y_c > 0$, para θ en $]0, 2\pi[$. (Recordemos que solamente tenemos un arco de cicloide, es decir θ está en $[0, 2\pi]$.)

El campo que obtenemos es central con foco en $(0, 0)$. Además

$$F_{zz} = y^{-1/2}(1 + z^2)^{-3/2} > 0$$

Para poder concluir que obtenemos un mínimo global, en la norma C^0 , debemos estudiar el funcional cerca de $y = 0$, ya que ahí F_{zz} es singular. Por el momento tenemos que:

$$E(x, y(x), z(x, y(x)), w) > 0$$

para $w \neq z(x, y(x))$ y $x > 0$.

Usaremos el siguiente argumento. Recordemos que la integral invariante de Hilbert:

$$I(\gamma) = \int_{\gamma} (-H dx + p dy)$$

es independiente del camino γ entre dos puntos del plano (ver V.6.6).

Aquí $H = pz - (1 + z^2)^{1/2} y^{-1/2}$, con $p = F_z = zy^{-1/2}(1 + z^2)^{-1/2}$. Por lo tanto,

$$H = -y^{-1/2}(1 + z^2)^{-1/2}, \quad 1 + z^2 = (1 - yp^2)^{-1}$$

es decir,

$$H = -(1 - yp^2)^{1/2} y^{-1/2}.$$

Ahora para cualquier curva $y(x)$ con $y(0) = 0$, $y(a) = A$, tomemos los siguientes caminos γ y $\tilde{\gamma} + \sigma$, uniendo los puntos $(\epsilon, y_0(\epsilon))$ y (a, A) (ver figura 5.6).

Tenemos $I(\gamma) = I(\tilde{\gamma}) + I(\sigma) = \int_{\epsilon}^a F(x, y_0(x), y'_0(x)) dx$. Sea $J_{\epsilon}(y) = \int_{\epsilon}^a F(x, y, y') dx$.

Tenemos

$$\begin{aligned} J_{\epsilon}(y) - J_{\epsilon}(y_0) &= J_{\epsilon}(y) - I(\gamma) = J_{\epsilon}(y) - I(\tilde{\gamma}) - I(\sigma) \\ &= \int_{\epsilon}^a E(x, y(x), z(x, y(x)), y'(x)) dx - I(\sigma) \end{aligned}$$

donde usamos la fórmula de Hilbert (ver V.6.6).

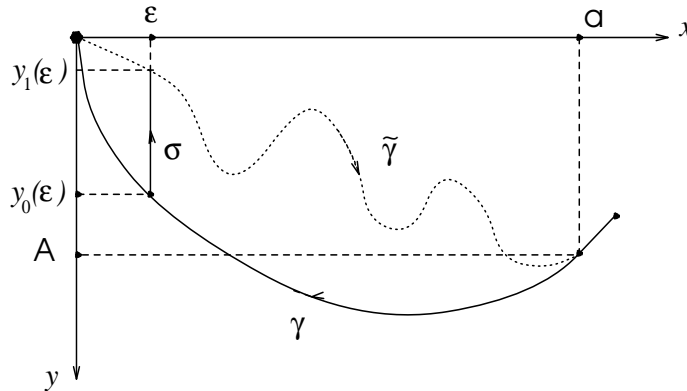


Fig. 5.6

Notemos que como $F(x, y, y') \geq 0$ y $E \geq 0$, entonces

En conclusión invitamos al lector a volver a leer la introducción y juzgar por él mismo. Además lo invitamos a completar sus conocimientos, viendo un buen libro como el de Troutman, a diversificarlos en el estudio de problemas con varias variables, en especial en el libro de Rektorys, y a profundizarlos con el estudio de desigualdades variacionales, referencias [7] y [10].

Apéndice

Notas sobre métodos directos

En este libro partimos de un funcional y vimos que un punto crítico es solución de una ecuación diferencial. Los métodos directos en Cálculo de Variaciones consisten en hacer el recorrido inverso: partir de una ecuación diferencial, de tipo elíptico lineal o no lineal, llegar a un funcional (energía por ejemplo) y analizar la ecuación diferencial a partir del estudio del funcional.

El propósito de estas notas es el dar una introducción a estos métodos muy poderosos y más modernos, en el caso más sencillo, es decir el de las ecuaciones diferenciales ordinarias (la teoría de Sturm-Liouville) y, de paso, dar otra prueba del Teorema del Índice de Morse.

Para problemas definidos sobre un intervalo $[a, b]$, los espacios más idóneos para los métodos directos son los espacios de Sobolev

$$H^k[a, b] = \{f(x) \text{ con } \|f\|_k^2 \equiv \int_a^b \sum_0^k |f^{(j)}(x)|^2 dx < \infty\}.$$

Aquí se trata de integral de Lebesgues y $H^0[a, b]$ es $L^2[a, b]$.

Estos espacios son completos y con producto escalar, es decir son espacios de Hilbert (Ver Ejemplos II.1.3 y II.1.4).

Nota: En este apéndice supondremos conocidos, o por lo menos no sorprendentes, los resultados más elementales de análisis y de análisis funcional.

Ahora bien, si u es localmente integrable y si existe v , localmente integrable, tal que

$$\int_a^b v \varphi dx = (-1)^n \int_a^b u \varphi^{(n)} dx$$

para toda φ en $C_0^\infty[a, b]$, se dice que v es **derivada débil** de u , o en el sentido de **distribuciones**. El argumento del Teorema III.1 permite probar que si u y v están en $L^2[a, b]$,

entonces v es **derivada fuerte** de u , es decir u está en $H^n[a, b]$.

Además, para cualquier función u , continuamente diferenciable, uno tiene la desigualdad, para $y > x$:

$$|u(y) - u(x)| = \left| \int_x^y u' d\tau \right| \leq |y - x|^{1/2} \left(\int_x^y u'^2 d\tau \right)^{1/2} \leq |y - x|^{1/2} \|u'\|_1.$$

Por lo tanto, si u está en $H^1[a, b]$, usando el hecho que, por construcción, C^1 es denso en H^1 , uno obtiene que

$$H^1[a, b] \subset C^{0,1/2}[a, b]$$

es decir, todo elemento de H^1 es continuo de Hölder con exponente $1/2$. En particular, una función u en $H^1[a, b]$ es continua y se puede hablar del valor de u en un punto. Además, toda sucesión acotada en H^1 será uniformemente continua y uniformemente acotada en C^0 , y por lo tanto, por el Teorema de Arzela-Ascoli, tendrá una subsucesión convergente en C^0 . Este es un caso particular del teorema de **compacidad de Rellich**, mientras que la inclusión de H^1 en $C^{0,1/2}$ es un caso particular del teorema de **encaje de Sobolev**.

Esto implica que el subespacio

$$H_0^1[a, b] = \{u \in H^1[a, b], u(a) = u(b) = 0\},$$

es un subespacio cerrado de H^1 , y por lo tanto es de Hilbert. Además, por la desigualdad anterior, tomando $x = a$ con $u(a) = 0$, tenemos la **desigualdad de Poincaré** (Ver Ejemplo VI.1.3)

$$\|u\|_0^2 \leq C \|u'\|_0^2.$$

para todo u en $H_0^1[a, b]$, lo que implica que $\|u'\|_0$ es una norma, en H_0^1 , equivalente a la norma de H^1 .

Como en el Teorema IV.1.1, si y_0 en H_0^1 es un punto crítico de $J(y) = \int_a^b F(x, y, y') dx$, es decir, si F es C^1 y

$$\int_a^b (F_y h + F_{y'} h') dx = 0,$$

para toda h en $H_0^1[a, b]$, entonces $F_{y'}$ tiene derivada en L^2 igual a F_y , en otras palabras, la ecuación de Euler es válida en L^2 . Por otra parte se puede probar, como en la Proposición IV.1.1, que si F es C^2 y $F_{y'y'}(x, y_0, y'_0) \geq \alpha > 0$, entonces y_0 está en $H^2[a, b]$: esto es un caso particular del Teorema de **regularidad de Schauder** y requiere de algunas estimaciones más delicadas.

Ejercicio A.1: Usando lo anterior, encontrar el complemento ortogonal de $H_0^1[a, b]$ en $H^1[a, b]$, es decir el conjunto de u 's en $H^1[a, b]$ tales que $\int_a^b (uh + u'h') dx = 0$, para todo h en $H_0^1[a, b]$.

Después de estos resultados preliminares podemos considerar el problema de Sturm-Liouville en el contexto del Teorema VI.3.2, es decir el estudio del operador diferencial

$$\mathcal{A}U = -(PU')' - (A^T - A)U' + (C - A'^T)U$$

para vectores $U(x)$ en \mathbb{R}^n , con $U(a) = U(b) = 0$, y tal que A , P y C sean matrices $n \times n$, con P y A de clase C^1 , $P^T = P$, $C^T = C$ y $P(x)$ es definida positiva para toda x en $[a, b]$.

$\mathcal{A}U = 0$ es la ecuación variacional del funcional

$$\begin{aligned} 2I(U) &= \int_a^b U^T \mathcal{A}U dx \\ &= \int_a^b (U'^T P U' + 2U^T A U' + U^T C U) dx, \end{aligned}$$

asociada a la forma bilineal $DI(U)V \equiv I(U, V) = \int_a^b V^T \mathcal{A}U dx$, con

$$I(U, V) = \int_a^b (V'^T P U' + V^T A U' + V'^T A^T U + V^T U) dx.$$

De esta última fórmula es inmediato ver que $I(U, V) = I(V, U)$, es decir que \mathcal{A} es un operador simétrico con respecto al producto escalar de L^2 .

Podemos entonces considerar al operador \mathcal{A} como un operador de $C^2[a, b] \cap \{U(a) = U(b) = 0\}$ en $C^0[a, b]$, o como un operador autoadjunto de $H^2[a, b] \cap H_0^1[a, b]$ en $L^2[a, b]$.

Estudiaremos el **espectro** de \mathcal{A} , es decir el conjunto de λ 's tal que $\mathcal{A}U = \lambda U$ tiene una solución no trivial con $U(a) = U(b) = 0$.

Si esto sucede, λ es un **valor propio** y U un **vector propio**.

Notemos que el conjunto de vectores propios asociados a un λ es un espacio vectorial de dimensión a lo más n (se trata de un sistema de ecuaciones diferenciales) llamado $\ker(\mathcal{A} - \lambda I)$.

Notemos también que el espectro de \mathcal{A} es acotado inferiormente: en efecto, existe una constante positiva α tal que

$$U'^T P U' \geq \alpha |U'|^2,$$

y, usando la continuidad de A y C , se tiene

$$|2U^T A U' + U^T C U| \leq k(\epsilon |U'|^2 + \epsilon^{-1} |U|^2 + |U|^2),$$

para todo $\epsilon > 0$. Escogiendo $\epsilon = \alpha/2$, se tiene que

$$\int_a^b U^T (\mathcal{A}U + MU) dx = 2I(U) + M \int_a^b |U|^2 dx > C \|U\|_1^2 > 0,$$

para U no idénticamente zero y M suficientemente grande.

Otras propiedades sencillas son las siguientes:

Lemma A.1:

1. $\ker(\mathcal{A} - \lambda I)$ y $\ker(\mathcal{A} - \mu I)$ son ortogonales (en L^2) si $\mu \neq \lambda$.
2. Si $\mathcal{A}U = f$ entonces f es ortogonal (en L^2) a $\ker \mathcal{A}$.

Prueba: Si $\mathcal{A}U = \lambda U$ y $\mathcal{A}V = \mu V$, entonces

$$\begin{aligned} I(U, V) &= \int_a^b V^T \mathcal{A}U dx = \lambda \int_a^b V^T U dx \\ I(V, U) &= \mu \int_a^b V^T U dx. \end{aligned}$$

Para $\lambda \neq \mu$, la única posibilidad es $\int_a^b V^T U dx = 0$.

Para otra parte, si $\mathcal{A}U = f$, entonces

$$\int_a^b V^T f dx = I(U, V) = I(V, U) = \int_a^b U^T \mathcal{A}V dx,$$

en particular, si V está en $\ker \mathcal{A}$, esto vale 0.

Q.E.D.

La implicación inversa de (2) es la **alternativa de Fredholm** y será probada más adelante. Antes, notemos que la forma bilineal $I(U, V)$ está bien definida y acotada para U y V en $H_0^1[a, b]$ y que sólo toma la forma $\int_a^b V^T \mathcal{A}U dx$ si U está en $H^2[a, b]$, es decir, el dominio de I es más grande que el dominio de \mathcal{A} . Por otra parte, si U en $H_0^1[a, b]$ es tal que

$$I(U, V) - \int_a^b V^T f dx = 0,$$

para todo V en $H_0^1[a, b]$, es decir si U es un punto crítico del funcional $I(U) - \int_a^b U^T f dx$, entonces, debido a la versión del teorema de regularidad que ya hemos mencionado, se tiene que U está en $H^2[a, b]$ y $\mathcal{A}U = f$ en L^2 . Para poder estudiar la existencia de U y sus propiedades como función de f , necesitamos algunos resultados de análisis funcional.

Teorema A.1: Representación de Riesz. Sea H un espacio de Hilbert y L un funcional lineal continuo sobre H . Entonces existe un único l en H tal que $L(U) = (U, l)$ para todo U en H . Además $\|l\| = \|L\|$.

Prueba. Sea $M = \{U : L(U) = 0\} = \ker L$. Queda claro que M es un subespacio lineal cerrado de L (por la continuidad). Sea M^\perp su complemento ortogonal. Si L no es idénticamente 0, entonces existe U_0 en M^\perp , con $LU_0 \neq 0$ (es un escalar). Para cualquier otro U_1 en

M^\perp , tendremos, para $\alpha = -LU_1/LU_0$, que $L(\alpha U_0 + U_1) = 0$, es decir $\alpha U_0 + U_1$ está en M y en M^\perp , por lo tanto $\alpha U_0 + U_1 = 0$, en otras palabras la dimensión de M^\perp es 1.

Ahora bien, usando la descomposición ortogonal

$$H = M \oplus M^\perp,$$

todo U en H se escribe como $U = \alpha U_0 + V$, con V en M y $\alpha = (U, U_0) / \|U_0\|^2$. Entonces $LU = \alpha LU_0 = (U, l)$, definiendo

$$l = U_0 \overline{LU_0} / \|U_0\|^2.$$

La unicidad de l viene del hecho de que si $LU = (U, l_1) = (U, l_2)$, entonces $l_1 - l_2$ es ortogonal a todo U en H y por lo tanto 0.

Finalmente, $|LU| \leq \|U\| \|l\|$, lo cual implica que $\|L\| \leq \|l\|$. Por otra parte, $|Ll| = \|l\|^2 \leq \|L\| \|l\|$, probando la igualdad de las normas.

Q.E.D.

Regresemos a $I(U)$ y tomemos M suficientemente grande para que $2I(U) + M \int_a^b |U|^2 \geq C \|U\|_1^2$.

Definamos entonces

$$\begin{aligned} (U, V)_E &= I(U, V) + M \int_a^b U^T V dx, \\ \|U\|_E^2 &= (U, U)_E. \end{aligned}$$

Por la desigualdad anterior y el hecho que $\|U\|_E^2 \leq C \|U\|_1^2$, queda claro que $\|U\|_E$ define una norma equivalente a $\|U\|_1$ sobre $H_0^1[a, b]$, con producto escalar $(U, V)_E$. Esta es la **norma de energía**.

Teorema A.2. Lax Milgram. El funcional $F(u) = \|U\|_E^2 - 2 \int_a^b f^T U dx$ tiene un único mínimo U_0 en $H_0^1[a, b]$, llamado la **solución generalizada** de $(\mathcal{A} + MI)U = f$.

Prueba. Consideremos el funcional $L(U) = \int_a^b f^T U dx$ sobre $H_0^1[a, b]$.

Claramente $L(U)$ es un funcional lineal acotado, por lo tanto continuo. Por el Teorema de Riesz, existe un único U_0 en H_0^1 tal que

$$L(U) = (U, U_0)_E = (U, f)$$

donde $(U, f) = \int_a^b U^T f dx$. Además $\|U_0\|_E \leq \|f\|$, tomando $U = U_0$. Entonces

$$\begin{aligned}
F(U) &= \|U\|_E^2 - 2(U, U_0)_E = \|U - U_0\|_E^2 - \|U_0\|_E^2 \\
&= F(U_0) + \|U - U_0\|_E^2.
\end{aligned}$$

Por lo tanto F tiene un mínimo en U_0 y este es único.

Q.E.D.

Notemos que, por el Teorema de regularidad, U_0 es una solución fuerte (en el caso de más variables este es un punto no trivial). Por otra parte, la relación $(U, U_0)_E = (U, f)$ para todo U en H_0^1 , con U_0 único, permite definir un operador $U_0 = Kf$, de L^2 en H_0^1 , el cual es claramente lineal y, como $\|Kf\|_E \leq \|f\|$, tiene norma a lo más 1. El operador K se llama el **operador de Green**, y por el Teorema de regularidad, es acotado de L^2 en $H^2 \cap H_0^1$, es decir es el operador **inverso** de $\mathcal{A} + MI$.

Lema A.2. *El operador K , como operador de L^2 en L^2 , es autoadjunto y compacto, es decir manda conjuntos acotados en compactos.*

Prueba. Sea $U_0 = Kf$ y $U = Kg$, entonces de la igualdad que define a K tenemos que

$$\begin{aligned}
(Kg, f) &= (Kg, Kf)_E = (Kg, (\mathcal{A} + MI)Kf) \\
&= ((\mathcal{A} + MI)Kg, Kf) = (g, Kf),
\end{aligned}$$

donde se usó el hecho que $\mathcal{A} + MI$ es autoadjunto y es el inverso de K . Por otra parte, si f_n es una sucesión acotada en L^2 , entonces Kf_n es una sucesión acotada en H^1 y por lo tanto en $C^{0,1/2}$. Tendremos entonces una subsucesión convergente en C^0 y, con más razón, en L^2 .

Q.E.D.

Lemma A.3: *Para $\lambda > -M$, sea $\mu = (\lambda + M)^{-1}$, entonces*

- (a) $\mathcal{A}U = \lambda U$ si y solo si $KU = \mu U$.
- (b) $(\mathcal{A} - \lambda I)U = f$ si y solo si $(\mu I - K)U = \mu Kf$.
- (c) *Rango $(K - \mu I)$ es cerrado y uno tiene la descomposición ortogonal*

$$L^2[a, b] = \ker(K - \mu I) \oplus \text{Rango}(K - \mu I).$$

Prueba. Si $(\mathcal{A} - \lambda I)U = f$ entonces $(\mathcal{A} + M)U = (\lambda + M)U + f$. Aplicando K se obtiene (b) y (a), con $f = 0$. En particular los núcleos son iguales y los rangos son isomorfos.

Para probar (c), supongamos que tenemos una sucesión en el rango convergiendo a w :

$$(K - \mu I)U_n \rightarrow w.$$

Si $\|U_n\| \leq C$, entonces por la compacidad de K , existe una subsucesión U_{n_j} con KU_{n_j} convergiendo a v . Por lo tanto U_{n_j} converge a $U = (v - w)/\mu$ y, por la continuidad de K se tiene que $(K - \mu I)U = w$, es decir el rango es cerrado. Por otra parte, si $\{U_n\}$ no es acotada podemos suponer que U_n está en el complemento ortogonal (cerrado) de $\ker(K - \mu I)$ y que $V_n = U_n / \|U_n\|$ es tal que $(K - \mu I)V_n$ converge a 0. El razonamiento anterior implica que V_{n_j} converge a V , con $\|V\| = 1$ y $(K - \mu I)V = 0$, lo cual contradice el hecho que V es ortogonal a $\ker(K - \mu I)$. Finalmente, si v es ortogonal a Rango $(K - \mu I)$, entonces

$$((K - \mu I)u, v) = (u, (K - \mu I)v) = 0,$$

para todo u en L^2 y, por lo tanto, $(K - \mu I)v = 0$.

Q.E.D.

Teorema A.3: Alternativa de Fredholm. *La ecuación $\mathcal{A}U = f$, en L^2 , tiene solución U en $H^2[a, b] \cap H_0^1[a, b]$ si y solo si f es ortogonal, en L^2 a $\ker \mathcal{A}$.*

Prueba. Por el lema anterior, $\mathcal{A}U = f$ si y solo si $(\mu I - K)U = \mu Kf$, con $\mu = M^{-1}$, y esto pasa si y solo si Kf es ortogonal, en L^2 , a $\ker(K - \mu I)$, es decir, para v en $\ker(K - \mu I)$

$$0 = (Kf, v) = (f, Ku) = \mu(f, v).$$

Q.E.D.

Para un operador acotado K , se define su norma como

$$\|K\| = \sup_{x \neq 0} \frac{\|Kx\|}{\|x\|}.$$

En el caso del operador de Green se tiene el siguiente resultado

Lema A.4. $\|K\| = \sup \frac{(Kf, f)}{\|f\|^2}$ y $\mu_1 = \|K\|$ es el más grande de los valores propios de K .

Prueba. De la relación $(Kf, f) = (Kf, Kf)_E$ se tiene que K es un operador positivo. Además, si $M = \sup \frac{(Kf, f)}{\|f\|^2}$, entonces $M \leq \|K\|$. Pero

$$\begin{aligned} |(K(f+g), f+g) - (K(f-g), f-g)| &= 4|(Kf, g)| \\ &\leq M(\|f+g\|^2 + \|f-g\|^2) = 2M(\|f\|^2 + \|g\|^2). \end{aligned}$$

Por lo tanto, si $\|f_0\| = \|g_0\| = 1$, se tiene $|(Kf_0, g_0)| \leq M$, y en general

$$|(Kf, g)| \leq M \|f\| \|g\|.$$

Tomando $g = Kf$, obtenemos

$$\|Kf\|^2 \leq M \|f\| \|Kf\|,$$

y por lo tanto $\|K\| \leq M$, probando la igualdad.

Sea entonces una sucesión f_n con $\|f_n\| = 1$ y (Kf_n, f_n) tendiendo a $\|K\|$.

Debido a la compacidad de K podemos extraer una subsucesión, que denotaremos también Kf_n , que converge a un elemento w de L^2 . Pero de la relación

$$(Kf, f) = \|Kf\|_E^2,$$

obtenemos que $\|K(f_n - f_m)\|_E^2 \leq \|K(f_n - f_m)\| \|f_n - f_m\|$, y por lo tanto Kf_n es también una sucesión de Cauchy en H_0^1 , es decir w está en H_0^1 y Kf_n converge en H_0^1 a w . El teorema de regularidad permite dar una convergencia en H^2 . Aplicando el operador $\mathcal{A} + MI$, se tiene que f_n converge a $f_1 \equiv (\mathcal{A} + MI)w$. Por lo tanto $\|f_1\| = 1$ y $Kf_1 = w$. Por lo tanto

$$(Kf_1, f_1) = \|K\| \|f_1\|^2 \leq \|Kf_1\| \|f_1\|.$$

Pero la desigualdad de Cauchy-Schwarz es una igualdad solo si los dos vectores son colineales, es decir, $Kf_1 = \mu_1 f_1$, aquí con $\mu_1 = \|K\|$. Que cualquier otro valor propio sea menor o igual a μ_1 , es inmediato.

Q.E.D.

Lema A.5. *El espectro de K consiste en un número infinito de valores propios μ_n , decreciendo a 0, y con vectores propios f_n , ortonormales en $L^2[a, b]$, que forman una base de L^2 .*

Prueba. Ya probamos la existencia del valor propio μ_1 más grande, con vector propio f_1 . Sea L_1 el complemento ortogonal de f_1 en $L^2[a, b]$. El espacio L_1 es de Hilbert y K manda L_1 en el mismo ya que

$$(Kf, f_1) = (f, Kf_1) = \mu_1 (f, f_1).$$

Por lo tanto, como antes, $\mu_2 = \|K\|_{L_1} = \sup_{L_1} \frac{(Kf, f)}{\|f\|^2}$ es un valor propio de K , con $\mu_2 \leq \mu_1$ ya que $L_1 \subset L^2[a, b]$.

Siguiendo este proceso se genera una sucesión decreciente de valores propios μ_n , de funciones propias f_n , ortonormales, y de espacios encajados L_n . Si L_∞ es el complemento ortogonal de

todos los f_n , entonces K manda L_∞ en el mismo y $Kf = 0$ para f en L_∞ (si no se generaría otro valor propio). Pero K es uno a uno y por lo tanto $L_\infty = 0$.

Por otra parte, como los espacios propios son ortogonales y de dimensión finita, este proceso incluye a todos los valores propios de K . Además, si μ_n se acumulan en $\mu > 0$, usando la compacidad de K , tenemos una subsucesión Kf_{n_j} que converge a w y f_{n_j} converge a w/μ , lo cual es imposible y $\mu = 0$.

Finalmente sea $g_N = \sum_1^N (f, f_n) f_n$, la proyección ortogonal de f sobre el espacio generado por (f_1, \dots, f_N) . Por el teorema de Pitágoras, se tiene

$$\|f\|^2 = \|f - g_N\|^2 + \sum_1^N (f, f_n)^2.$$

En particular $\|g_N\| \leq \|f\|$ (desigualdad de Bessel) y la sucesión g_N converge a un elemento g de $L^2[a, b]$, con

$$g = \sum_1^\infty (f, f_n) f_n.$$

Como $(f - g, f_k) = 0$, para todo k , uno tiene que $f - g$ está en $L_\infty = \{0\}$.

Q.E.D.

Teorema A.4. Sturm-Liouville. *El espectro de \mathcal{A} es de la forma*

$$-M < \lambda_1 \leq \lambda_2 \leq \dots, \lambda_n \rightarrow \infty,$$

con funciones propias $U_n = (M + \lambda_n)^{-1/2} f_n$, ortonormales en la norma de energía y cualquier U en $H_0^2[a, b]$ se escribe de forma única como

$$\begin{aligned} U &= \sum_1^\infty (U, U_n)_E U_n = \sum_1^\infty (U, f_n) f_n, \\ I(U) &= \sum_1^\infty \lambda_n (U, f_n)^2. \end{aligned}$$

Prueba. La primera parte viene de los Lemas A.3 y A.5, con el hecho que $(U_n, U_m)_E = (U_n, Kf_m)_E \mu_m^{-1/2} = (U_n, f_m) \mu_m^{-1/2} = (\mu_n / \mu_m)^{1/2} (f_n, f_m) = \delta_{nm}$. Por otra parte, si U es ortogonal, en la norma de energía en H_0^1 , a todas las U_n 's, entonces

$$0 = (U, U_n)_E = \mu_n^{-1/2} (U, Kf_n)_E = \mu_n^{-1/2} (U, f_n),$$

y U es 0. El argumento del Lema A.5 permite concluir que las U_n son una base de H_0^1 .

Finalmente,

$$\begin{aligned}
I(U) &= \|U\|_E^2 - M \int_a^b |U|^2 dx \\
&= \sum_1^\infty (U, U_n)_E^2 - M \sum_1^\infty (U, f_n)^2.
\end{aligned}$$

Pero $(U, U_n)_E^2 = \mu_n^{-1}(U, f_n)^2 = (\lambda_n + M)(U, f_n)^2$, dando la última fórmula.

Q.E.D.

La fórmula dando $I(U)$ permite dar una caracterización de los valores propios de \mathcal{A} independiente de la base f_n :

Teorema A.5. Principio de minimax de Courant. Si V_{k-1} es un subespacio de H_0^1 de dimensión $k-1$, denotemos por V_{k-1}^\perp al subespacio de H_0^1 ortogonal, en la norma de L^2 , a V_{k-1} . Entonces,

$$\lambda_k = \max_{V_{k-1}} \left(\min_{V_{k-1}^\perp} \frac{I(U)}{\|U\|^2} \right).$$

Prueba. Si U es ortogonal, en la norma de L^2 , a f_1, \dots, f_{k-1} , entonces

$$I(U) = \sum_k^\infty \lambda_n (U, f_n)^2 \geq \lambda_k \sum_k^\infty (U, f_n)^2 = \lambda_k \|U\|^2$$

Por lo tanto, para U ortogonal a f_1, \dots, f_{k-1} , uno tiene

$$\lambda_k \leq \frac{I(U)}{\|U\|^2}.$$

Además, tomando $U = f_k$, entonces $I(f_k) = \lambda_k$, es decir

$$\lambda_k = \min_{M_{k-1}^\perp} \frac{I(U)}{\|U\|^2},$$

donde M_{k-1} es el espacio generado por f_1, \dots, f_{k-1} .

Sea ahora $V_{k-1} \subset H_0^1$, generado por $\varphi_1, \dots, \varphi_{k-1}$, linealmente independientes.

Es fácil ver que $M_k \cap V_{k-1}^\perp$ no se reduce a 0, porque sino tendríamos que cualquier $U = \sum_1^k a_n f_n$, en M_k , y ortogonal a $\varphi_1, \dots, \varphi_{k-1}$, sería 0. Pero $(U, \varphi_j) = \sum_1^k a_n (f_n, \varphi_j)$, $j = 1, \dots, k-1$, es un sistema lineal de $k-1$ ecuaciones y de k incógnitas, por lo tanto hay una solución diferente de 0.

Para U en $M_k \cap V_{k-1}^\perp$, se tiene

$$I(U) = \sum_1^k \lambda_n a_n^2 \leq \lambda_k \|U\|^2.$$

Por lo tanto

$$\min_{V_{k-1}^\perp} \frac{I(U)}{\|U\|^2} \leq \lambda_k = \min_{M_{k-1}^\perp} \frac{I(U)}{\|U\|^2}.$$

Esta última desigualdad prueba el teorema.

Q.E.D.

A partir de los resultados anteriores, el teorema del índice se deduce fácilmente. Sean $a < c < c'$ dos puntos. Podemos estudiar el problema $\mathcal{A}U = f$, con $U(a) = U(c) = 0$, via la teoría desarrollada anteriormente. Sea $H_0^1(a, c), L^2(a, c)$ con norma $\|U\|_c$, y $I_c(U) = \int_a^c (U'^T P U' + 2U^T A U' + U^T C U) dx$ definido sobre $H_0^1(a, c)$. El subespacio de $H_0^1(a, c)$ donde $I_c(U)$ es negativo es generado por f_1, \dots, f_N tales que $\lambda_1(c) \leq \dots \leq \lambda_N(c) < 0$ y es de dimensión N . Notemos que c es un punto conjugado de a , de multiplicidad m , si y solo si $\ker \mathcal{A}$ tiene dimensión m en $H_0^1(a, c)$.

Para $c < c'$, podemos extender U en $H_0^1(a, c)$ como $U \equiv 0$ en (c, c') y obtener un elemento de $H_0^1(a, c')$, tal que $\|U\|_c = \|U\|_{c'}$ y $I_c(U) = I_{c'}(U)$. De esta forma podemos también considerar cualquier subespacio $V_{k-1}(c)$ en $H_0^1(a, c)$ como un subespacio de $H_0^1(a, c')$, y tal que

$$H_0^1(a, c) \cap V_{k-1}^\perp(c) \subset H_0^1(a, c') \cap V_{k-1}^\perp(c').$$

Por lo tanto,

$$\min_{V_{k-1}^\perp(c)} \frac{I_c(U)}{\|U\|_c^2} \geq \min_{V_{k-1}^\perp(c')} \frac{I_{c'}(U)}{\|U\|_{c'}^2},$$

y

$$\lambda_k(c) \geq \lambda_k(c').$$

Para t en $[a, c']$, sea $\tau = t - \left(\frac{c'-c}{c'-a}\right)(t-a)$ en $[a, c]$ (y $t = \tau + \left(\frac{c'-c}{c'-a}\right)(\tau-a)$). Entonces si U pertenece a $H_0^1(a, c)$, la función

$$\tilde{U}(t) = U(\tau)$$

pertenece a $H_0^1(a, c')$. Además

$$\|\tilde{U}\|_{H_0^1(a, c')}^2 = \int_a^{c'} (|\tilde{U}|^2 + |\tilde{U}'|^2) dt = \int_a^c \left(\left(\frac{c'-a}{c-a} \right) |U|^2 + \left(\frac{c-a}{c'-a} \right) |U'|^2 \right) d\tau$$

Por lo tanto el cambio de variables es un isomorfismo de $H_0^1(a, c')$ a $H_0^1(a, c)$. Se tiene además

$$\| \tilde{U} \|_{L^2(a,c')}^2 = \frac{c' - a}{c - a} \| U \|_{L^2(a,c)}^2.$$

Del mismo modo, si $\tilde{P}(t) = P(\tau)$, $\tilde{A}(t) = A(\tau)$, $\tilde{C}(t) = C(\tau)$,

$$\begin{aligned} I_{c'}(\tilde{U}) = I_c(U) &+ \int_a^c \left(\frac{c - c'}{c' - a} U'^T P U' + \frac{c' - c}{c - a} U^T C U \right) d\tau + \\ &\int_a^{c'} (\tilde{U}'^T (P - \tilde{P}) \tilde{U}' + 2\tilde{U}^T (A - \tilde{A}) \tilde{U}' + \tilde{U}^T (C - \tilde{C}) \tilde{U}) dt. \end{aligned}$$

Por lo tanto,

$$|I_{c'}(\tilde{U}) - I_c(U)| \leq k(|c - c'| + \epsilon) \| U \|_1^2,$$

donde ϵ depende de la continuidad de A y C y puede ser reemplazado por $|c - c'|$ si esas matrices son Lipschitz continuas.

Como el cociente $I_c(U) / \| U \|_{L^2(a,c)}^2$ es homogéneo de grado 0, podemos considerarlo solo para $U(\tau)$ con $\| U \|_1 = 1$, probando, ya que a todo $V_{k-1}^\perp(c)$ corresponde un $\tilde{V}_{k-1}^\perp(c')$, que

$$|\lambda_k(c') - \lambda_k(c)| \leq \alpha(|c - c'| + \epsilon)$$

es decir $\lambda_k(c)$ es continuo con respecto a c y Lipschitz continua si A y C lo son.

En conclusión, dado que $\lambda_k(c) > 0$ si c es vecino a a (por la desigualdad de Poincaré), y dada la continuidad de $\lambda_k(c)$ y el hecho que son no-crecientes, entonces, conforme c va creciendo, los valores propios van decreciendo y, cuando pasan por 0, se obtienen tantos como la multiplicidad del punto conjugado. Esto prueba el teorema del índice.

BIBLIOGRAFIA ADICIONAL

Los métodos directos son clásicos y cualquier libro de ecuaciones en derivadas parciales, que tenga un buen capítulo sobre los métodos modernos en ecuaciones elípticas, contiene estos resultados. Una buena referencia es [13]. Aquí el tratamiento para ecuaciones diferenciales ordinarias permite una exposición menos técnica y más corta.

Para un tratamiento más avanzado del Cálculo de Variaciones, el lector puede consultar las referencias [7], [10] y los libros de Vainberg.

El Cálculo de Variaciones es una herramienta muy importante en la investigación actual: Para ello, el lector puede ver:

1. Ekeland, I. *Convexity methods in Hamiltonian mechanics*. Springer 1990.
2. Mawhin, J. y Willem M. *Critical point theory and Hamiltonian systems*. Springer 1989.

3. Rabinowitz, P.H. *Minimax methods in critical point theory with applications to differential equations*. CBMS Regional Conf. **65**, A.M.S., 1986.
4. Struwe, M. *Variational Methods*, Springer 1990.

Bibliografía

- [1] G.A. Bliss. Lectures on the Calculus of Variations. University of Chicago Press; Chicago, 1946.
- [2] O. Bolza. Lectures on the Calculus of Variations. Dover Publications, Inc: New York, 1961.
- [3] C. Caratheodory. Calculus of Variations and Partial Differential Equations of the First Order: I, II. Holden-Day, Inc.: San Francisco, 1965, 1967.
- [4] R. Courant. Calculus of Variations and Supplementary Notes and Exercises. New York University Lectures Notes, 1957.
- [5] R. Courant and D. Hilbert. Methods of Mathematical Physics. I. Interscience Publishers: New York, 1953.
- [6] L.E. El'Sgol'C. Calculus of Variations. Addison-Wesley Publishing Co., Inc.: Reading, Mass., 1962.
- [7] I. Ekeland and R. Temam. Convex Analysis and Variational Problems, North-Holland: Amsterdam, 1977.
- [8] I.M. Gelfand and S.V. Fomin. Calculus of Variations. Prentice-Hall, Inc.: Englewood Cliffs, New Jersey, 1963.
- [9] A.D. Ioffe and V.M. Tihomirov. Theory of Extremal Problems. North-Holland Publishing Company: Amsterdam, 1979.
- [10] D. Kinderlehrer and G. Stampacchia. An introduction to Variational Inequalities and Their Applications. Academic Press, Inc., New York, 1980.
- [11] S.G. Mikhlin. Variational Methods in Mathematical Physics. The Macmillan Company: New York, 1964.
- [12] J. Milnor. Morse Theory. Princeton University Press. Princeton 1963.
- [13] K. Rektorys. Variational Methods in Mathematics, Science and Engineering. D. Reidel Publishing Company: Boston, 1975.
- [14] J.L. Troutman. Variational Calculus with Elementary Convexity. Undergraduate Texts in Mathematics. Springer Verlag, New York, 1983.

- [15] M.M. Vainberg. Variational Methods for the Study of Nonlinear Operators. Holden-Day, Inc.: San Francisco, 1964.
- [16] M.M. Vainberg. Variational Methods and Method of Monotone Operators in the Theory of Nonlinear Equations. John Wiley and Sons: New York, 1973.

Índice

- $C^0[a, b]$, 36
- $C^\infty[a, b]$, 36
- $C_0^\infty[a, b]$, 36
- $DJ(y_0)$, 45
- $D_h J(f_0)$, 44
- $H^1[a, b]$, 37
- $L^2[a, b]$, 37

- abierto, 40
- acción generalizada, 160
- alisador, 62
- alternativa de Fredholm, 256, 259
- area encerrada, 135
- area mínima, 89

- barra elástica, 94
- barra pesada, 101
- bracket de Poisson, 160
- braquistocrona, 15, 21, 44, 68, 70, 75, 248

- cambios de variables, 143
- campo de extremales, 190, 199
- campo exacto, 200, 242
- campos centrales, 191, 198, 238
- catenaria, 135
- Cauchy-Schwarz, 39
- cerrado, 40
- cerradura, 40
- cicloide, 23
- coercividad, 42
- compacidad de Rellich, 254
- compacto, 42
- completación, 38
- condición de Frobenius, 160
- condición de Legendre, 220
- condición de Weierstrass, 110, 125
- condición holonómica, 164
- condiciones de frontera, 71
- condiciones de transversalidad, 117, 154, 170, 176
- condiciones holonómicas, 137
- condiciones naturales, 115
- condiciones necesarias, 67
- condiciones no holonómicas, 139
- condiciones suficientes, 212, 225
- conjugado, 223
- conservación de momento, 171
- conservación del momento angular, 171
- continuidad, 40
- continuidad para segundas derivadas, 133
- convergencia uniforme, 37
- coordenadas Lagrangianas, 155
- cuerda vibrante, 89

- denso, 40
- derivada débil, 253
- derivada de Fréchet, 45
- derivada de Gâteaux, 44
- derivada fuerte, 254
- derivada segunda de Fréchet, 207
- derivada segunda de Gâteaux, 207
- derivadas, 44
- derivadas de orden más alto, 93
- desigualdad de Poincaré, 211, 254
- Dido, 15
- diferencias finitas, 24
- difracción, 133
- discontinuidad no fija, 110, 125
- discontinuidades en las segundas derivadas, 114
- discontinuidades, condición de Weierstrass, 106
- distancia, 37
- distribuciones, 253

Du Bois-Reymond, 60

ecuación de Euler, 97, 98

ecuación de Euler-Lagrange, 67

ecuación de Hamilton-Jacobi, 166, 187

ecuación de onda, 91

ecuación variacional, 215

ecuaciones canónicas, 160

ecuaciones de Cauchy-Riemann, 88

ecuaciones de primer orden, 180

ecuaciones en forma paramétrica, 143

encaje de Sobolev, 254

energía, 159

energía potencial, 156

energía cinética, 81, 156

energía potencial, 81

espacio completo, 38

espacio de Banach, 38

espacio lineal, 35

espacio normado, 37

espectro, 255

existencia, unicidad, 30

extremo móvil, 117

extremos, 48

extremos débiles, 205

extremos de funcionales, 206

extremos fuertes, 205

fórmula de Hilbert, 203, 250

forma paramétrica, 143, 163

frontera libre, 221, 242

frontera libre: condiciones naturales, 99

fuerza central, 173

función distancia geodésica, 173, 190, 195

función generadora, 165

funcional invariante, 168

funcional lineal, 43

funcionales, 16

geodésica, 16, 17, 27, 84, 138, 220

Haar, 64

Hamilton-Jacobi, 143, 173

Hamiltoniano, 158

Hilbert, 39

hipótesis de continuidad, 177

hipótesis de función geodésica, 177

identidad de Euler, 86

índice de Morse, 230, 253

integral primera, 72

invariancia, 145

invariante de Hilbert, 202, 250

Jacobi, 193

Lagrange, 55

Laplace-Beltrami, 149

laplaciano, 150

Lax Milgram, 257

Legendre, 212, 226

ley de Hooke, 90

ley de Snell, 108

leyes de conservación, 167

Lipschitz, 41

máximo local, 48

métodos directos, 253

mínima acción, 81

membrana elástica, 20, 88

Meyers-Serrin, 61

minimax, 262

momentos, 159

multiplicadores de Lagrange, 49, 115

multiplicidad, 231

norma de energía, 257

operador de Green, 258

operador lineal, 43

Plateau, 16

potencial electrostático, 108

pre-Hilbert, 39

principio de Fermat, 83

principio de Hamilton, 157

principio de Huyghens, 190

principios variacionales duales, 163

problema de Plateau, 148

problema de Robin, 106

problema isoperimétrico, 18, 25, 52, 77,
134, 140, 244

problema mixto, 102

producto escalar, 39
 propagación de luz, 82
 punto conjugado, 215, 227
 punto crítico, 68
 punto de acumulación, 40

 reflexión, 128
 regla de la cadena, 151, 152
 regularidad de Schauder, 254
 relación de Lagrange, 202
 representación de Riesz, 256
 restricciones holonómicas, 134
 restricciones no holonómicas, 134
 Riccati, 229

 segunda ecuación de Euler, 72, 99
 segundas derivadas discontinuas, 110
 semicontinuidad, 41
 semicontinuo superiormente, 41
 sistema canónico y transformada de Legendre, 157
 sistema conservativo, 81, 157
 sistema Hamiltoniano, 158
 solución de Bernoulli, 23
 solución generalizada, 257
 Stone-Weierstrass, 40
 Sturm-Liouville, 253, 261
 sucesión de Cauchy, 37
 superficie de revolución, 19, 71, 75, 100, 251

 teorema de Noether, 145, 167
 topología, 37
 topología débil, 43
 transformaciones canónicas, 164
 transformada de Legendre, 161

 unicidad de la derivada, 47

 valor propio, 115, 255
 Vandermonde, 62
 vecindad, 39
 vector propio, 255