ECUACIONES DE CAMPO GRAVITACIONAL POR PRINCIPIOS VARIACIONALES

RICARDO GARCIA SALCEDO

Contents

1 CALCULO TENSORIAL. **5**

2 LOS PRINCIPIOS DE MINIMO. **9**
2.1 LOS PRINCIPIOS DE MINIMO EN LA ANTIGÜEDAD 9
2.2 EL PRINCIPIO DE HERÓN. 9
2.3 EL PRINCIPIO DE FERMAT. 9
2.4 EL PRINCIPIO DE MINIMA ACCION DE MAUPERTIUS. 9
2.5 EL PRINCIPIO DE HAMILTON. 10
2.6 JUSTIFICACION "HEURISTICA" DEL PRINCIPIO DE HAMILTON. . 10

3 CALCULO VARIACIONAL. **15**
3.1 CALCULO VARIACIONAL EN INTEGRALES SENCILLAS. 15
3.1.1 PROPIEDADES DE TRANSFORMACION DE LA INTEGRAL FUNDAMENTAL. 18
3.1.2 SENTIDO FISICO DE LOS PRINCIPIOS VARIACIONALES Y LA LAGRANGIANA 19
3.1.3 ALGUNOS TEOREMAS DE CONSERVACION 20
3.2 CALCULO VARIACIONAL EN INTEGRALES MULTIPLES 22
3.2.1 PROPIEDADES DE TRANSFORMACION DE LA INTEGRAL FUNDAMENTAL MULTIPLE. 24
3.3 FORMULACION DE HAMILTON 26

4 PRINCIPIOS DE LA TEORIA DE LA RELATIVIDAD **29**
4.1 TRANSFORMACION DE GALILEO 29
4.2 PRINCIPIOS DE RELATIVIDAD 31
4.3 ESPACIO-TIEMPO 32
4.4 TRANSFOMACION DE LORENTZ 35
4.4.1 TRANSFORMACION DE VELOCIDADES 41
4.5 CONSECUENCIAS DE LA TRANSFORMACION DE LORENTZ 41
4.5.1 CONTRACCION DE LA LONGITUD 41
4.5.2 DILATACION DEL TIEMPO 42
4.6 DINAMICA RELATIVISTA 42

5 FORMULACION COVARIANTE DE LA ELECTRODINAMICA **47**
5.1 CAMPO ELECTRICO Y POTENCIAL ELECTROSTATICO 47
5.2 LEY DE GAUSS 48
5.3 FUERZA DE LORENTZ 49
5.4 LEY DE GAUSS (MAGNETICA) 49
5.5 LEY DE INDUCCION DE FARADAY 50
5.6 LEY CIRCUITAL DE AMPERE 51

5.7 REPRESENTACION DIFERENCIAL DE LAS ECUACIONES DE MAXWELL 52
5.8 FORMULACION COVARIANTE DE LA ELECTRODINAMICA 54
5.9 TENSOR DE ENERGIA MOMENTO ELECTROMAGNETICO 59

6 INTRODUCCION A LA TEORIA GENERAL DE LA RELATIVIDAD 61
6.1 CAMPO GRAVITACIONAL 61
6.2 EL PRINCIPIO DE EQUIVALENCIA 62
6.2.1 MOTIVACION 62
6.2.2 PRINCIPIO DE EQUIVALENCIA. 63
6.3 GEODESICAS EN EL ESPACIO-TIEMPO 64
6.4 RELACION ENTRE LA CONECCION AFIN Y SIMBOLOS DE CHRISTOFFEL 66
6.5 DERIVACION COVARIANTE. 68
6.6 TENSORES DE RIEMANN-CHRISTOFFEL 71
6.7 IDENTIDADES DE BIANCHI 73
6.8 TENSOR DE RICCI 74
6.9 LEY DE LA GRAVITACION DE EINSTEIN 75
6.10 MODIFICACION DE LA ECUACION DE EINSTEIN POR PRESENCIA DE MATERIA 76

7 PRINCIPIOS VARIACIONALES DE LOS CAMPOS 79
7.1 CONSTRUCCION DE FUNCIONALES DE ACCION INVARIANTES. . 79
7.2 EL CAMPO ELECTROMAGNETICO 79
7.2.1 EL PRIMER PAR DE ECUACIONES DE MAXWELL. 82
7.2.2 EL SEGUNDO PAR DE ECUACIONES DE MAXWELL 84
7.3 LA ACCION PARA EL CAMPO GRAVITACIONAL 85
7.4 LA ACCION PARA LA MATERIA 89
7.5 ACCION COSMOLOGICA 91
7.6 ACCION ELECTROMAGNETICA 92
7.7 ACCION ELECTROMAGNETICA Y GRAVITACIONAL 93

8 CONCLUSIONES 95

9 REFERENCIAS 97

Chapter 1
CALCULO TENSORIAL.

Uno de los objetivos principales de este trabajo es el de construir ecuaciones de la física que sean invariantes bajo cualquier transformación de coordenadas con el auxilio del método variacional, en especial las Transformaciones de Lorentz. Para llegar a tal objetivo, en esta sección describiremos una clase de objetos cuyas propiedades de transformación sean particularmente simples, es decir que permanezcan invariantes bajo transformaciones generales de coordenadas, de tal manera que a partir de ellos podamos construir dichas ecuaciones.

Usaremos transformaciones que sean lineales y con ello hablaremos de la necesidad de invariancia bajo estas transformaciones. Supongamos que un conjunto de coordenadas que cubren a todo el espacio-tiempo son (x^0, x^1, x^2, x^3) o en forma compacta x^μ con $\mu = 0, 1, 2, 3^*$ y que otro conjunto distinto x'^ν también lo hace. Las x'^ν deberán ser funciones de las x^μ

$$x'^\nu = x'^\nu (x^\mu) \tag{1.1}$$

y también las x^μ son funciones de las x'^ν

$$x^\mu = x^\mu (x'^\nu) \tag{1.2}$$

es decir, que la transformación deberá tener inversa, por tanto dicha transformación tiene jacobiano distinto de cero.

Antes de continuar mencionaremos que a menos que se especifique lo contrario haremos uso de la **convención de suma** de Einstein, es decir el tener índices repetidos implica suma

$$\sum_{\mu=1}^{n} \frac{\partial L}{\partial q^\mu} dq^\mu \equiv \frac{\partial L}{\partial q^\mu} dq^\mu \tag{1.3}$$

Si calculamos las diferenciales de las x^μ y x'^ν por medio de la regla de la cadena tenemos que

$$dx^\mu = \frac{\partial x^\mu}{\partial x'^\nu} dx'^\nu \tag{1.4}$$

y

$$dx'^\nu = \frac{\partial x'^\nu}{\partial x^\mu} dx^\mu \tag{1.5}$$

respectivamente.

Con lo anterior, empezaremos por mencionar que la más simple de todas las reglas de transformación es la de los **escalares**, los cuales no cambian bajo transformaciones generales de coordenadas.

---

*El uso de índices griegos (α, β, γ, μ, ν, etc.) indicará que ellos van de 0 a 4. Mientras que los índices latinos (i, j, k, etc.) tomarán los valores de 1, 2, y 3.

Otra regla de transformación es la de un **tensor o cuadrivector contravariante**, V^μ, el cual se transforma como

$$V'^\nu = \frac{\partial x'^\nu}{\partial x^\mu} V^\mu \tag{1.6}$$

Un ejemplo de un tensor contravariante es la diferenciación parcial, ya sea la dada por la ecuación (1.4) o la ecuación (1.5).

Una transformación muy relacionada con la anterior es la de un **tensor o cuadrivector covariante**, U_μ el cual se transforma como

$$U'_\nu = \frac{\partial x^\mu}{\partial x'^\nu} U_\mu \tag{1.7}$$

Por ejemplo, si ϕ es una función escalar, entonces $\dfrac{\partial \phi}{\partial x^\mu} \equiv \phi_{,\mu}$ es un tensor covariante, porque en un sistema de coordenadas transformado el **gradiente** es

$$\frac{\partial \phi}{\partial x'^\nu} = \frac{\partial x^\mu}{\partial x'^\nu} \frac{\partial \phi}{\partial x^\mu} \tag{1.8}$$

o

$$\phi'_{,\nu} \equiv \frac{\partial x^\mu}{\partial x'^\nu} \phi_{,\mu}$$

Los tensores pueden tener uno o más índices arriba o abajo. El **rango (orden) del tensor** está dado por el número de índices, por ejemplo

$$T'^{\mu\nu}_{\rho} = \frac{\partial x'^\mu}{\partial x^\alpha} \frac{\partial x'^\nu}{\partial x^\beta} \frac{\partial x^\gamma}{\partial x'^\rho} T^{\alpha\beta}_{\gamma} \tag{1.9}$$

que es un **tensor mixto,** ya que tiene tanto índices arriba como índices abajo, es de rango dos contravariante y uno covariante.

Estamos ahora en posibilidad de construir ecuaciones invariantes. Cualquier ecuación será invariante bajo transformaciones de coordenadas si ésta enuncia la igualdad de dos tensores del mismo orden tanto covariante como contravariante. Es decir, si $A^{\mu\lambda}_{\nu}$ y $B^{\mu\lambda}_{\nu}$ son dos tensores con la regla de transformación (1.9) y si en un sistema de coordenadas x^μ, $A^{\mu\lambda}_{\nu} = B^{\mu\lambda}_{\nu}$, entonces en un sistema de coordenadas x'^μ, es obvio que

$$A'^{\mu\lambda}_{\nu} = B'^{\mu\lambda}_{\nu}$$

.

Para poder lograr construir ecuaciones invariantes bajo transformaciones de coordenadas debemos considerar algunas operaciones algebraicas básicas entre tensores:

1. **Combinaciones Lineales**. Una combinación lineal de tensores con los mismos índices arriba y abajo es un tensor con estos mismos índices. Por ejemplo, sean A^μ_ν y B^μ_ν tensores mixtos y sea
$$T^\mu_\nu = aA^\mu_\nu + bB^\mu_\nu$$
donde a y b son escalares, entonces T^μ_ν es un tensor ya que
$$T'^\mu_\nu = aA'^\mu_\nu + bB'^\mu_\nu = a\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}A^\rho_\sigma + b\frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}B^\rho_\sigma$$
$$= \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}\left(aA^\rho_\sigma + bB^\rho_\sigma\right) = \frac{\partial x'^\mu}{\partial x^\rho}\frac{\partial x^\sigma}{\partial x'^\nu}T^\rho_\sigma$$

2. **Productos Directos**. El producto de las componentes de dos tensores da un tensor cuyos índices contravariantes y covariantes consisten de todos los índices de arriba y abajo de los dos tensores originales. Por ejemplo, si A^{μ}_{ν} y B^{ρ} son tensores y
$$T^{\mu\rho}_{\nu} = A^{\mu}_{\nu} B^{\rho}$$
entonces $T^{\mu\rho}_{\nu}$ es un tensor, esto es
$$\begin{aligned} T'^{\mu\rho}_{\nu} = A'^{\mu}_{\nu} B'^{\rho} &= \frac{\partial x'^{\mu}}{\partial x^{\kappa}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} A^{\kappa}_{\sigma} \frac{\partial x'^{\rho}}{\partial x^{\lambda}} B^{\lambda} \\ &= \frac{\partial x'^{\mu}}{\partial x^{\kappa}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x'^{\rho}}{\partial x^{\lambda}} A^{\kappa}_{\sigma} B^{\lambda} = \frac{\partial x'^{\mu}}{\partial x^{\kappa}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x'^{\rho}}{\partial x^{\lambda}} T^{\kappa\lambda}_{\sigma} \end{aligned}$$

3. **Contracción**. Tomando un índice superior e inferior igual y sumándolo sobre sus valores obtenemos un nuevo tensor sin estos dos índices. Sea $T^{\mu\rho\nu}_{\sigma}$ un tensor y
$$T^{\mu\nu} = T^{\mu\rho\nu}_{\rho}$$
entonces $T^{\mu\nu}$ es un tensor, dado que
$$\begin{aligned} T'^{\mu\nu} = T'^{\mu\rho\nu}_{\rho} &= \frac{\partial x'^{\mu}}{\partial x^{\kappa}} \frac{\partial x^{\sigma}}{\partial x'^{\rho}} \frac{\partial x'^{\nu}}{\partial x^{\lambda}} \frac{\partial x'^{\rho}}{\partial x^{\tau}} T^{\kappa\lambda\tau}_{\sigma} \\ &= \frac{\partial x'^{\mu}}{\partial x^{\kappa}} \frac{\partial x^{\sigma}}{\partial x^{\tau}} \frac{\partial x'^{\nu}}{\partial x^{\lambda}} T^{\kappa\lambda\tau}_{\lambda} = \frac{\partial x'^{\mu}}{\partial x^{\kappa}} \frac{\partial x'^{\nu}}{\partial x^{\lambda}} T^{\kappa\lambda} \end{aligned}$$

Estas tres operaciones pueden ser combinadas de varias maneras. Una combinación particularmente importante es el subir o bajar índices. Si tomamos el producto directo de un tensor contravariante o mixto T con otro tensor $g_{\mu\nu}$ y contraemos el índice μ con un índice contravariante de T, obtenemos un nuevo tensor en el cual este índice contravariante es remplazado por el índice covariante ν. Es decir
$$S^{\rho}_{\nu\sigma} = g_{\mu\nu} T^{\mu\rho}_{\sigma}$$
entonces por las operaciones (2) y (3), $S^{\rho}_{\nu\sigma}$ será un tensor. Similarmente tenemos que
$$R^{\nu\rho}_{\sigma} = g^{\mu\nu} S^{\rho}_{\mu\sigma}$$

Notemos que si bajamos un índice y luego lo volvemos a subir entonces nos da nuevamente el tensor original. Por ejemplo,
$$R^{\nu\rho}_{\sigma} = g^{\mu\nu} S^{\rho}_{\mu\sigma} = g^{\mu\nu} g_{\mu\lambda} T^{\lambda\rho}_{\sigma} = \delta^{\nu}_{\lambda} T^{\lambda\rho}_{\sigma} = T^{\nu\rho}_{\sigma} \tag{1.10}$$

Por último, vamos **a definir $\epsilon^{\alpha\beta\gamma\delta}$ como el símbolo totalmente antisimétrico o tensor de Levi-Civita** que tiene la propiedad
$$\epsilon^{\alpha\beta\gamma\delta} = \begin{cases} 1 & \text{si } \alpha\beta\gamma\delta \text{ es una permutación par de } 0123 \\ -1 & \text{si } \alpha\beta\gamma\delta \text{ es una permutación impar de } 0123 \\ 0 & \text{si algún índice se repite} \end{cases} \tag{1.11}$$
y $\epsilon^{0123} = 1$.

Si definimos una matriz M^{α}_{β} de 4×4, en que α y β representan renglones y columnas, entonces un resultado básico del álgebra lineal es que el determinante de M^{α}_{β} está dado por
$$\det(M^{\alpha}_{\beta}) = \epsilon^{\alpha\beta\gamma\delta} M^{0}_{\alpha} M^{1}_{\beta} M^{2}_{\gamma} M^{3}_{\delta}$$

Figure 1

o lo que es equivalente

$$\det(M^{\alpha}_{\beta})\epsilon^{\lambda\mu\nu\xi} = \epsilon^{\alpha\beta\gamma\delta}M^{\lambda}_{\alpha}M^{\mu}_{\beta}M^{\nu}_{\gamma}M^{\xi}_{\delta}$$

donde en las ecuaciones anteriores se ha hecho uso de la convención de suma sobre índices repetidos. Regresemos al concepto de tensor. ?‘Es $\epsilon^{\alpha\beta\gamma\delta}$ un tensor?. Si lo fuera , se transformaría como

$$\epsilon^{\alpha\beta\gamma\delta} \rightarrow \epsilon'^{\alpha\beta\gamma\delta} = \frac{\partial x'^{\alpha}}{\partial x^{\lambda}}\frac{\partial x'^{\beta}}{\partial x^{\mu}}\frac{\partial x'^{\gamma}}{\partial x^{\nu}}\frac{\partial x'^{\delta}}{\partial x^{\xi}}\epsilon^{\lambda\mu\nu\xi}$$

pero $\det(\frac{\partial x'}{\partial x})\epsilon^{\alpha\beta\gamma\delta} = \frac{\partial x'^{\alpha}}{\partial x^{\lambda}}\frac{\partial x'^{\beta}}{\partial x^{\mu}}\frac{\partial x'^{\gamma}}{\partial x^{\nu}}\frac{\partial x'^{\delta}}{\partial x^{\xi}}\epsilon^{\lambda\mu\nu\xi}$, y además el $\det(\frac{\partial x'}{\partial x}) = \pm 1$. Así que $\epsilon^{\alpha\beta\gamma\delta}$ se transforma como un tensor de rango 4, excepto por un cambio de signo si $\det(\frac{\partial x'}{\partial x}) = -1$.

Por definición $\epsilon^{\alpha\beta\gamma\delta}$ es un **pseudotensor**: cambia de signo frente a inversiones del tiempo o de espacio, pero salvo ese signo se tranforma como un tensor de rango 4.

Los pseudotensores de cualquier orden, en particular, los pseudoescalares, se comportan como tensores en todas las transformaciones de coordenadas, salvo respecto a aquellas que no puedan reducirse a rotaciones, es decir, quedan excluidas las reflexiones (esto es, los cambios de signo de coordenadas que no pueden reducirse a giros).

Si $A^{\alpha\beta}$ es un tensor antisimétrico, el tensor $A^{\alpha\beta}$ y el pseudotensor

$$A^{*\alpha\beta} = \frac{1}{2}\epsilon^{\alpha\beta\gamma\delta}A_{\gamma\delta} \tag{1.12}$$

se califican de **duales** entre sí. Análogamente, $\epsilon^{\alpha\beta\gamma\delta}A_{\delta}$ es un pseudotensor antisimétrico de orden tres, dual del tensor A^{α}. El producto $A^{\alpha\beta}A^{*}_{\alpha\beta}$ de dos tensores duales, es evidentemente, un pseudoescalar.

Los elementos básicos que se ha revisado nos permiten construir ecuaciones invariantes, para mayor información sobre tensores se pueden consultar, por ejemplo, [14] o [6].

Con lo anterior podemos empezar a buscar ecuaciones invariantes bajo transformaciones lineales. Posteriormente hablaremos de algunos ejemplos más físicos de tensores.

Chapter 2
LOS PRINCIPIOS DE MINIMO.

En esta parte del trabajo, expondremos las ideas básicas de donde partieron los principios variacionales que posteriormente utilizaremos para desarrollar una de las teorías físicas de la actualidad como lo es la teoría de la relatividad general de Einstein.

2.1 LOS PRINCIPIOS DE MINIMO EN LA ANTIGÜEDAD

La idea sobre la que los principios variacionales se sustentan es que en la Naturaleza algunas cantidades de un proceso físico resultan mínimas. Ya desde los primeros filósofos y científicos , los problemas de la naturaleza siempre se trataron de reducir a leyes y principios de mínimo o de lo más simple. Aristóteles (384-322 a.C) siguiendo esta idea decía que cualquier movimiento tenía que ser rectilíneo, circular o una combinación de ambos, ya que estos eran los movimientos más simples.

2.2 EL PRINCIPIO DE HERÓN.

Un intento por desarrollar un principio de mínimo fue hecho por Herón de Alejandría quien en el siglo II a.C. encontró que la ley de reflexión de la luz puede obtenerse del hecho de que cuando un rayo de luz es reflejado por un espejo, la trayectoria que toma para llegar hasta el ojo del observador es la más corta que cualquiera otra reflejada. Una construcción geométrica sencilla puede usarse para comprobar que este principio de mínimo nos puede llevar a la igualdad de los ángulos de incidencia y reflexión [12].

2.3 EL PRINCIPIO DE FERMAT.

Sin embargo, el principio de mínimo de Herón no puede suministrarnos una expresión correcta para la refracción, esto se logró hasta aproximadamente 1500 años después, en 1657 cuando Pierre de Fermat (1601-1665) enunció su propio principio, llamado principio de mínimo tiempo, que incluía tanto a la reflexión como a la refracción. Este principio enuncia que la luz se propaga de un punto a otro a lo largo de la trayectoria que le toma el mínimo tiempo, aun si tiene que desviarse de la trayectoria verdaderamente más corta para hacerlo. Uno puede demostrar que de este principio resultan tanto la ley de la reflexión como la ley de Snell, [20].

2.4 EL PRINCIPIO DE MINIMA ACCION DE MAUPERTIUS.

El astrónomo y matemático francés Maupertius (1698-1759), en 1744 enunció su famoso principio de mínima acción, le príncipe de la moindre quantité d'action*. Maupertitus basó su principio en consideraciones teológicas, consideraba que la acción es minimizada

*Este principio se mencionó por primera vez en dos artículos: uno, enviado a la Academia de Ciencias de Paris en 1744 y el otro a la Academia Prusiana dos años después.

por la sabiduría de Dios. El enunciado original del principio es el siguiente: "L'action est proportionelle au produit de la masse par la vitesse et par l'espace. Maintenant, voici ce príncipe, si sage, si digné de l' Être suprême: Lorsqu'il arrive quelque changement dans la Nature, la quantité d'action employée pour ce changement est toujours la plus petite qu'il soit possible" que brevemente dice que el movimiento dinámico tiene siempre lugar bajo una acción mínima. La definición de Maupertius para la acción fue muy obscura, ya que era diferente para cada experimento que realizó. En estudios posteriores de Leonard Euler (1707-1793) y Joseph Lagrange (1736-1813) encontraron que Maupertius definía a la acción como $S = \int_{t_o}^{t_1} v\,ds$ donde v es la velocidad y ds es el elemento de curva, posteriormente trataron hacer de este principio un teorema exacto de la dinámica† sin las consideraciones teológicas.

2.5 EL PRINCIPIO DE HAMILTON.

En trabajos publicados entre 1834 y 1835, William Rowan Hamilton (1805-1865) expuso el principio dinámico variacional sobre el cual se fundamenta toda la mecánica clásica. El principio de Hamilton puede enunciarse como:

"De todas las trayectorias posibles que puede seguir un sistema dinámico para desplazarse de un punto a otro en un intervalo de tiempo determinado, la trayectoria verdadera seguida es aquella que hace mínima la integral temporal de la energía cinética menos la energía potencial."

A partir de este principio se puede deducir el principio de mínima acción‡ ya que es más general que este último.

2.6 JUSTIFICACION "HEURISTICA" DEL PRINCIPIO DE HAMILTON.

Muchos textos de Mecánica tratan de dar un esbozo de alguna posible justificación de este principio, aquí trataremos de hacerlo de alguna forma más intuitiva o accesible§. Nos basaremos en el cápitulo que a propósito Richard P. Feynman escribió[8].

Para poder realizar esta justificación analizaremos el caso más simple, es decir, el caso de una partícula en movimiento libre. Consideremos al espacio y al tiempo como homogéneos y uniformes, y a su vez consideraremos que el espacio es isotrópico [16]. La homogeneidad del espacio y del tiempo significa que la energía cinética T de alguna partícula en movimiento no puede contener explícitamente ni el vector de posición $\mathbf{r}$ de la partícula , ni el tiempo t; es decir, T será sólo función de la velocidad $\mathbf{v}$. Puesto que el espacio es isotrópico, la energía cinética no puede depender tampoco de la dirección del vector $\mathbf{v}$; por tanto, será función solamente de su valor absoluto, es decir, es directamente proporcional al cuadrado de su velocidad, $T \sim \mathbf{v} \cdot \mathbf{v} = v^2$, más aun, la energía cinética está dada por

$$T = \frac{1}{2}mv^2 \tag{2.1}$$

donde m es la masa de la partícula.

Supongamos que deseamos viajar del punto A al punto B (consideremos el movimiento en la dimensión x solamente) de la figura 1. Es claro que para ello necesitamos un cierto intervalo de tiempo que denotaremos como τ. Existen dos formas principales para viajar de A a B, una de ellas es viajar variando la velocidad durante el trayecto, que es más complicada que la de viajar a velocidad constante.

†Para ver este teorema y como se llegó a él, ver [34] parágrafos 3 y 4.

‡Para revisar dicha deducción se puede ver, por ejemplo, [10] cap. 8-6, p.446 y ss.

§Ver por ejemplo [16] parágrafos 2, 4 y 5.

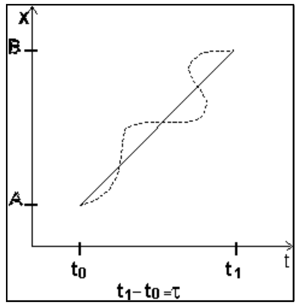


Figure 1 Formas de viajar de A a B.

La forma de calcular la velocidad promedio es tomar la distancia total recorrida y dividirla por el tiempo empleado para recorrerla.

De lo anterior podemos decir que la velocidad más simple con la que podemos ir de un lugar a otro es necesariamente la velocidad constante o velocidad promedio[¶]. Por lo tanto podemos enunciar un primer Lema:

LEMA

"La trayectoria más simple o ideal que puede seguir una partícula para viajar de un lugar a otro es aquella en la que en cada punto de ella la velocidad de la partícula es la velocidad promedio, es decir, viajar a velocidad constante".

Veamos ahora un poco de Estadística[‖]. Una **variable aleatoria continua** x está completamente definida por la **función de probabilidad** $F(x)$. Ahora introduzcamos el concepto de **densidad de probabilidad**. Si la probabilidad de la ocurrencia de x en el intervalo $(x, x + dx)$ es $dF(x)$, tenemos que $dF(x) = f\ (x)\, dx$ donde $f\ (x)$ es la densidad de probabilidad o **función de distribución**. La **condición de normalización** para una variable aleatoria continua está dada por

$$\int_{(x)} f\ (x)\, dx = 1 \tag{2.2}$$

donde la integración se realiza por todos los valores posibles de x, denotado por (x).

El **valor promedio (o medio, o esperado)** para una variable continua $v(x)$ está dado por

$$\overline{v} = \int_{(x)} v f\ (x)\, dx \tag{2.3}$$

Definimos como **desviación estardar**, σ_v , a la desviación de la variable v de su valor promedio $\overline{v}$. Una manera de medir esta desviación en términos de los valores promedio de v y v^2 es la siguiente:

$$\sigma_v^2 = \overline{v^2} - \overline{v}^2 \tag{2.4}$$

De la ecuación anterior podemos ver que

$$\overline{v^2} \geq \overline{v}^2 \tag{2.5}$$

¶Se puede comprobar fácilmente que el viajar a velocidad constante implica viajar a velocidad promedio. La herramienta para hacerlo se dará más adelante, ver ecuación 2.6.

‖Este repaso fue tomado de[30] paragrafo 2.

Como se ha considerado al tiempo uniforme entonces el **promedio temporal** de alguna variable continua, en este caso v, se obtiene a partir de (2.3) y la condición de normalización

$$\overline{v} = \frac{1}{\tau}\int_0^\tau v\, dt \tag{2.6}$$

para un cierto intervalo de tiempo τ.

Si calculamos el promedio temporal de la energía cinética encontramos usando la relación (2.1)

$$\overline{T} = \frac{1}{\tau}\int_0^\tau T dt = \frac{1}{2}m\left[\frac{1}{\tau}\int_0^\tau v^2 dt\right] = \frac{1}{2}m\overline{v^2} \tag{2.7}$$

Supongamos ahora que la partícula viaja a la velocidad promedio $\overline{v}$ entonces su energía cinética es

$$T = \frac{1}{2}m\overline{v}^2 \tag{2.8}$$

Relacionando a $\overline{v^2}$ y $\overline{v}^2$ de forma que la ec. (2.5) se satisfaga, obtenemos que

$$\overline{T} = \frac{1}{2}m\overline{v^2} \geq \frac{1}{2}m\overline{v}^2 \tag{2.9}$$

el significado de la relación (2.9) es que el promedio temporal de la energía cinética es mínimo siempre y cuando la velocidad de la partícula sea la velocidad promedio.

En conclusión, y por el Lema anterior, cuando la trayectoria que sigue una partícula es una trayectoria simple o ideal entonces el promedio temporal de la energía cinética de la partícula es mínimo. Inversamente, si el promedio temporal de la energía cinética es mínimo entonces la partícula sigue una trayectoria ideal. Este es parte del principio de Hamilton.

Definamos a la acción en un periodo de tiempo como el promedio temporal de la energía cinética

$$\frac{I}{\tau} = \overline{T} = \frac{1}{\tau}\int_0^\tau T dt$$

entonces la **acción**, I queda dada como:

$$I = \int_0^\tau T dt$$

Entonces podemos también concluir que la trayectoria que sigue la partícula es la ideal si la acción es mínima.

Hasta ahora hemos considerado una partícula libre o sin interacciones (fuerzas) externas, consideremos entonces el caso en el cual la partícula tiene interacciones externas. Cuando estas fuerzas derivan de un potencial se dice que el **sistema es conservativo** ** y por lo tanto a la partícula se le puede asociar una **energía potencial**, V, que origina dichas interacciones. De esta forma, la partícula posee ahora tanto una energía cinética, ya que está en movimiento, como una energía potencial debida a las interacciones que sufre. Esta energía potencial de alguna manera no permite que la partícula se mueva

**Un sistema conservativo se define como aquel en el que para calcular el trabajo para llevar el sistema de un punto a otro no depende de la trayectoria por donde se realice. Ver por ejemplo [20],[10], [1].

como si fuera libre, de ahí que nosotros consideraremos que la energía cinética efectiva de la partícula es su energía cinética menos su energía potencial

$$T_{ef} = T - V$$

Entonces la acción para una partícula que se encuentra dentro de una región con interacciones externas queda definida como

$$I = \int_0^\tau T_{ef} dt = \int_0^\tau (T - V)\, dt$$

A esta energía cinética efectiva generalmente se le da el nombre de **Lagrangiana** (o **densidad lagrangiana**)y se denota como L. Por lo tanto la acción para una partícula en movimiento dentro de un campo potencial queda definida como

$$I = \int_0^\tau L dt \tag{2.10}$$

Por lo general esta lagrangiana depende tanto de la posición de la partícula, debido a la energía potencial que varía con respecto a la posición (podría también depender de la velocidad y del tiempo pero entonces el sistema ya no sería conservativo), como de la velocidad de la partícula, debido a su energía cinética; la posición de la partícula depende del tiempo. De esta forma, la relación funcional de la lagrangiana queda así:

$$L = L\left(x\,(t)\,, \frac{dx}{dt}, t\right) \tag{2.11}$$

Generalmente, para definir una posición de un sistema de N partículas hay que tomar N radio vectores, es decir, $3N$ coordenadas para dicho sistema. Si ahora en nuestro problema encontramos l ecuaciones que nos restringen el problema, entonces decimos que tenemos l **ecuaciones de ligadura**[††]. Cada una de estas ecuaciones de ligadura nos permiten expresar una de las coordenadas en función de las otras y podemos decir que ahora nuestro sistema contiene $s = 3N - l$ coordenadas. En general, este número de magnitudes independientes necesarias para determinar unívocamente la posición del sistema se llama número de **grados de libertad** del sistema. Estas magnitudes no tienen que ser obligatoriamente las coordenadas cartesianas de los puntos, sino que de acuerdo con las condiciones del problema puede resultar más conveniente elegir otras coordenadas.

Se da el nombre de **coordenadas generalizadas** de un sistema de s grados de libertad a las s magnitudes cualesquiera $q_1, q_2, ..., q_s$ que caracterizan totalmente su posición, y el de **velocidades generalizadas** a sus derivadas $\dot{q}_i$. En adelante las coordenadas que utilizaremos en todo el análisis serán estas coordenadas generalizadas y sus respectivas velocidades generalizadas.

La lagrangiana encontrada anteriormente se puede expresar ahora, en términos de las coordenadas generalizada y además no tan sólo la trayectoria en una dimensión sino en varias dimensiones de la siguiente forma

$$L = L(t, q^k, \dot{q}^k) \tag{2.12}$$

donde el superíndice k indica cada uno de los grados de libertad del problema.

Antes de continuar necesitamos saber la definición de ecuación de movimiento, para ello diremos que si nosotros conocemos las posiciones y las velocidades generalizadas

[††]Para más detalles ver por ejemplo [10] de las referencias.

de un sistema en un determinado instante de tiempo t_o, es posible definir, en principio, sus posiciones, velocidades y en general su movimiento a un tiempo posterior, este es el **principio de determinancia de Newton** [1]. Desde el punto de vista matemático, al quedar definidas las q y las $\dot{q}$ en un instante de tiempo, quedan al mismo tiempo determinados los valores de las aceleraciones $\ddot{q}$ en dicho instante.

Las ecuaciones que relacionan las aceleraciones con las coordenadas y las velocidades se llaman **ecuaciones de movimiento**.

$$\ddot{q} = F\left(q, \dot{q}, t\right)$$

Por el teorema de existencia y unicidad de soluciones de ecuaciones diferenciales ordinarias, la función F y las condiciones iniciales $q\left(t_o\right)$ y $\dot{q}\left(t_o\right)$ determinan una única trayectoria de movimiento.

Chapter 3
CALCULO VARIACIONAL.

En la segunda mitad del siglo XVII. Gente como Newton, Leibniz y los hermanos Bernoulli desarrollaron una rama muy importante de la Matemática conocida hoy en día como Cálculo variacional. Con este tipo de cálculo podemos resolver problemas como el de la Braquistócrona o la catenaria*, es decir la forma que tiene una cuerda cuando está suspendida de sus extremos.

Por lo anterior estudiaremos el problema principal del Cálculo de Variaciones† de una manera condensada pero esencial.

3.1 CALCULO VARIACIONAL EN INTEGRALES SENCILLAS.

Un punto en el espacio de configuración n-dimensional lo representamos como $q^k = q^k(t)$ donde $k = 1, .., n$. Sea $q_\epsilon(t)$ una familia uniparamétrica de funciones de la forma

$$q_\epsilon(t) = q(t) + \epsilon\eta(t) \tag{3.1}$$

donde $q(t)$ y $\eta(t)$ son funciones arbitrarias. Entonces **la primera variación de I en la coordenada q en la dirección η** se define como

$$\delta I(q, \eta) = \left.\frac{dI(q + \epsilon\eta)}{d\epsilon}\right|_{\epsilon=0} \tag{3.2}$$

si la derivada existe, para I una **forma funcional** (o simplemente **funcional**), lo que en términos simples podemos considerar como una función de funciones. Esta derivada se le nombra comúnmente **derivada débil** y a la variación definida así se le conoce como **variación de Gâteaux** [25].

La definición anterior de derivada de una funcional se ha hecho con la intención de aplicar las técnicas ordinarias del cálculo diferencial para encontrar extremales. Por lo tanto si $q(t)$ es un **mínimo relativo** para la funcional I entonces

$$\delta I(q, \eta) = 0 \tag{3.3}$$

para todo η.

Ahora supongamos que $q(t) \in C_n^2[a, b]$, es decir que $q(t) = \left(q^1(t), q^2(t), ..., q^n(t)\right)$ donde cada una de las $q^k(t)$ son funciones univaluadas en el intervalo $[a, b]$ con derivadas parciales continuas hasta de segundo orden. Y además $q(a) = \boldsymbol{\alpha}$, $q(b) = \boldsymbol{\beta}$ donde $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathsf{R}^n$.

Consideremos a L como se definió en (2.12), es decir $L : \mathsf{R} \times \mathsf{R}^n \times \mathsf{R}^n \to \mathsf{R}$ es nuestra **lagrangiana**, la cual es dos veces continuamente diferenciable en cada una de sus

*Estos problemas son clásicos en el cálculo de variaciones, la solución de ellos puede encontrarse en algunos libros de Mecánica Clásica o Cálculo de Variaciones, por ejemplo ver [20].

†Para más detalles ver [1], [18], [19].

$2n+1$ argumentos. Y consideremos una funcional I llamada la **integral fundamental o integral de acción** tal y como la definimos en (2.10)

$$I(q(t)) = \int_a^b L(t, q^1(t), ..., q^n(t), \dot{q}^1(t), ..., \dot{q}^n(t))\, dt \tag{3.4}$$

o escrita en su notación vectorial como:

$$I(q(t)) = \int_a^b L(t, q(t), \dot{q}(t))\, dt \tag{3.5}$$

donde $\dot{q}(t)$ denota la derivada temporal de $q(t)$, es decir $\dot{q}(t) = \dfrac{dq(t)}{dt}$.

Apliquemos ahora (3.3), para ello sea $q_\epsilon(t)$ las variaciones de las funciones como en (3.1) que hacen mínima a la funcional,

$$q_\epsilon(t) = q(t) + \epsilon\eta\,(t) \tag{3.6}$$

con $\eta\,(t) = \left(\eta^1\,(t)\,, \eta^2\,(t)\,, ..., \eta^n\,(t)\right) \in C_n^2\,[a,b]$ y $\eta\,(a) = \eta\,(b) = \mathbf{0}$. Observemos que (3.6) es la notación corta de

$$q_\epsilon^k(t) = q^k(t) + \epsilon\eta^k\,(t) \qquad\qquad k = 1, 2, ..., n \tag{3.7}$$

Continuando, necesitamos calcular $\delta I\,(q\,(t)\,, \eta\,(t))$ y hacerla igual a cero. Por (3.6) tenemos que

$$\delta I\,(q\,(t)\,, \eta\,(t)) = \left[\frac{d}{d\epsilon}\int_a^b L\,(t, q(t) + \epsilon\eta\,(t)\,, \dot{q}(t) + \epsilon\dot{\eta}\,(t))\,dt\right]_{\epsilon=0}$$

Por la regularidad tomada sobre L y dado que ϵ no depende de t entonces la derivada $\dfrac{d}{d\epsilon}$ puede ser introducida dentro de la integral y aplicando la regla de la cadena obtenemos que

$$\delta I\,(q\,(t)\,, \eta\,(t)) = \left[\int_a^b \left(\frac{\partial L}{\partial q_\epsilon^k}\frac{\partial q_\epsilon^k}{\partial \epsilon} + \frac{\partial L}{\partial \dot{q}_\epsilon^k}\frac{\partial \dot{q}_\epsilon^k}{\partial \epsilon}\right) dt\right]_{\epsilon=0}$$

donde se aplica en convenio de suma. Si tomamos de (3.7) que $\dfrac{\partial q_\epsilon^k}{\partial \epsilon} = \eta^k(t)$ y $\dfrac{\partial \dot{q}_\epsilon^k}{\partial \epsilon} = \dot{\eta}^k(t)$ y evaluamos en $\epsilon = 0$, entonces

$$\delta I\,(q\,(t)\,, \eta\,(t)) = \int_a^b \left(\frac{\partial L}{\partial q^k}\eta^k(t) + \frac{\partial L}{\partial \dot{q}^k}\dot{\eta}^k(t)\right) dt$$

Si el segundo término lo integramos por partes obtenemos

$$\delta I = \int_a^b \left[\frac{\partial L}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right)\right]\eta^k(t)dt + \left.\frac{\partial L}{\partial \dot{q}^k}\eta^k(t)\right|_a^b$$

Dado que $\eta^k(a) = \eta^k(b) = 0$ para todo $k = 1, 2, ..., n$ entonces

$$\delta I = \int_a^b \left[\frac{\partial L}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k} \right) \right] \eta^k(t)dt \qquad (3.8)$$

la cual es la **primera variación de** I y esta variación se despreciará si $q(t)$ es un mínimo.

Antes de continuar es conveniente mencionar un lema debido a Dubois-Reymond, que es llamado **Lema fundamental del cálculo de variaciones**[25], que podemos plantear de la siguiente manera; sea $f(t)$ una función de valor real continua definida en $a \leq t \leq b$ y supongamos que

$$\int_a^b f(t)\,\phi(t)dt = 0 \qquad (3.9)$$

para toda $\phi \in C^2[a, b]$ que satisface que $\phi(a) = \phi(b) = 0$. Entonces

$$f(t) \equiv 0 \qquad t \in [a, b]$$

La demostración de este lema se realiza por contradicción. Tomemos que existe un punto $t_0 \in (a, b)$ para la cual $f(t_0) > 0$. Por la continuidad de f, existe un intervalo (t_1, t_2) contenido en (a, b) alrededor de t_0 sobre el cual f es estrictamente positiva. Ahora sea

$$\phi(t) = \begin{cases} (t - t_1)^3 (t_2 - t)^3 & \forall t \in (t_1, t_2) \\ 0 & \forall t \notin (t_1, t_2) \end{cases}$$

donde $\phi \in C^2[a, b]$ y

$$\int_a^b f(t)\,\phi(t)dt = \int_{t_1}^{t_2} f(t)\,\phi(t)dt > 0$$

es una contradicción a (3.9). ■

Volvamos a (3.8) y dado que en esta ecuación hicimos la consideración de que $\eta(a) = \eta(b) = \mathbf{0}$, en particular podemos ver que $\eta(t) = \left(0, ..., \eta^i(t), 0, ..., 0\right)$ para algún i fijo, entonces esto implica que

$$\int_a^b \left[\frac{\partial L}{\partial q^i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i} \right) \right] \eta^i(t)dt = 0 \qquad \text{(no hay suma sobre } i\text{)}$$

para todo $\eta^i(t)$ de clase $C^2[a, b]$ que se anula en a y b. Las condiciones de regularidad de L implican que la expresión $\dfrac{\partial L}{\partial q^i} - \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}^i} \right)$ es continua en $[a, b]$ y por tanto, por (3.9) se sigue que $\dfrac{\partial L}{\partial q^i} - \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}^i} \right) = 0$. Pero i es arbitrario, entonces se sigue que una condición necesaria para que la función $q(t)$ suministre un mínimo relativo para la funcional I definida por (3.4) es tal que sus componentes $q^1(t), q^2(t), ..., q^n(t)$ satisfacen las n ecuaciones

$$E_k = \frac{\partial L}{\partial q^k} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k} \right) = 0 \qquad k = 1, 2, ..., n \qquad (3.10)$$

para todo $a \leq t \leq b$.

Las ecuaciones (3.10) forman un sistema de n ecuaciones diferenciales ordinarias de segundo orden, las cuales se conocen como **Ecuaciones de movimiento de Euler-Lagrange**.

Cabe hacer notar que, análogamente al cálculo, no podemos estar seguros de que en realidad $q(t)$ sea un mínimo local, esto se logra haciendo la **segunda variación de Gâteaux**, pero esta investigación no es relevante en el estudio de la Física por lo que podemos remitir su estudio a libros de Cálculo de Variaciones [25].

3.1.1 PROPIEDADES DE TRANSFORMACION DE LA INTEGRAL FUNDAMENTAL.

Ahora veremos algunas propiedades de la integral fundamental bajo una transformación de coordenadas. Esta transformación tiene la siguiente forma

$$q'^k = q'^k \left(q^h\right)$$

como en (1.1) y no afecta al parámetro t. Supongamos que esta transformación tiene inversa como en (1.2)

$$q^h = q^h \left(q'^k\right)$$

Las derivadas de la trasformación (1.2) constituyen las componentes de un tensor contravariante, es decir

$$\dot{q}^h = \frac{\partial q^h}{\partial q'^k}\,\dot{q}'^k = \dot{q}^h\left(q'^k, \dot{q}'^k\right) \tag{3.11}$$

Además, se puede ver de (3.11) que

$$\frac{\partial \dot{q}^h}{\partial \dot{q}'^k} = \frac{\partial q^h}{\partial q'^k} \tag{3.12}$$

junto con

$$\frac{\partial \dot{q}^h}{\partial q'^k} = \frac{\partial^2 q^h}{\partial q'^k \partial q'^l}\,\dot{q}'^l \tag{3.13}$$

El hecho de que la integral fundamental (3.4) sea invariante bajo (1.1) es equivalente a decir que la lagrangiana L sea un escalar relativo a (1.1). Esta condición es muy importante. Si denotamos que la transformación de L bajo (1.1) es $\bar{L}$, nuestro requerimiento de invariancia es equivalente a decir que

$$\bar{L}\left(t, q'^k, \dot{q}'^k\right) = L\left(t, q^h\left(q'^k\right), \dot{q}^h\left(q'^k, \dot{q}'^k\right)\right) \tag{3.14}$$

en donde se han usado las relaciones (1.2) y (3.11). Esto puede ser válido para todos los valores $\left(q'^k, \dot{q}'^k\right)$.

Ahora vamos a diferenciar con respecto a $\dot{q}'^k$ la ecuación (3.14) y usamos (3.12) para obtener

$$\frac{\partial \bar{L}}{\partial \dot{q}'^k} = \frac{\partial L}{\partial \dot{q}^h}\frac{\partial \dot{q}^h}{\partial \dot{q}'^k} = \frac{\partial L}{\partial \dot{q}^h}\frac{\partial q^h}{\partial q'^k} \tag{3.15}$$

esta ecuación hace evidente que las derivadas $\dfrac{\partial L}{\partial \dot{q}^h}$ constituyen las componentes de un tensor covariante. Sin embargo, si nosotros diferenciamos (3.14) con respecto a q'^k y aplicamos (3.13), encontramos que

$$\frac{\partial \bar{L}}{\partial q'^k} = \frac{\partial L}{\partial q^h}\frac{\partial q^h}{\partial q'^k} + \frac{\partial L}{\partial \dot{q}^h}\frac{\partial \dot{q}^h}{\partial q'^k} = \frac{\partial L}{\partial q^h}\frac{\partial q^h}{\partial q'^k} + \frac{\partial L}{\partial \dot{q}^h}\frac{\partial^2 q^h}{\partial q'^k \partial q'^l}\,\dot{q}'^l \tag{3.16}$$

la cual implica que $\dfrac{\partial L}{\partial q^h}$ no sean las componentes de un tensor.

Sabemos que las derivadas $\dfrac{\partial \phi}{\partial q^h}$ de una función escalar $\phi\left(q^h\right)$ representa un tensor covariante, llamado el gradiente de ϕ; sin embargo, en el caso del escalar $L\left(t, q^h, \dot{q}^h\right)$ esto no resulta, dada la dependencia de L sobre las variables $\dot{q}^h$.

Por lo tanto es necesario construir un tensor covariante que represente el gradiente generalizado de L. Para ello, eliminaremos el segundo término de (3.16), esto lo conseguiremos diferenciando (3.15) con respecto a t, lo cual nos dará

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial \bar{L}}{\partial \dot{q}'^k}\right) &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^h}\right)\frac{\partial q^h}{\partial q'^k} + \frac{\partial L}{\partial \dot{q}^h}\frac{d}{dt}\left(\frac{\partial q^h}{\partial q'^k}\right) \\ &= \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^h}\right)\frac{\partial q^h}{\partial q'^k} + \frac{\partial L}{\partial \dot{q}^h}\frac{\partial^2 q^h}{\partial q'^k \partial q'^l}\,\dot{q}'^l \qquad (3.17)\end{aligned}$$

Cuando a esta última ecuación le restamos (3.16) obtenemos que

$$\frac{d}{dt}\left(\frac{\partial \bar{L}}{\partial \dot{q}'^k}\right) - \frac{\partial \bar{L}}{\partial q'^k} = \frac{\partial q^h}{\partial q'^k}\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^h}\right) - \frac{\partial L}{\partial q^h}\right] \qquad (3.18)$$

de la cual es evidente que las cantidades $E_h = \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}^h}\right) - \dfrac{\partial L}{\partial q^h}$ que coinciden con las ecuaciones de Euler-Lagrange (3.10) constituyen las componentes de un tensor covariante de orden uno. Este tensor puede ser interpretado como el **gradiente generalizado** de la lagrangiana L. Con lo anterior se ha comprobado que las ecuaciones de Euler-Lagrange se comportan como un tensor siempre que la lagrangiana sea un escalar relativo a la transformación de coordenadas.

Esto es muy importante ya que las ecuaciones de Euler-Lagrange son las ecuaciones de movimiento de un sistema físico el cual está descrito por la lagrangiana, el que éstas ecuaciones de movimiento se transformen como un tensor nos garantiza que serán las mismas para cualquier sistema de coordenadas.

3.1.2 SENTIDO FISICO DE LOS PRINCIPIOS VARIACIONALES Y LA LAGRANGIANA

Veamos ahora las implicaciones físicas de las ecuaciones de Euler-Lagrange. Como ya se ha mencionado, el principio de Hamilton nos dice que un sistema físico evoluciona por la trayectoria o trayectorias (si el sistema tiene varios grados de libertad) que hace mínima la integral de acción, entonces el tratamiento matemático anterior nos sirve para encontrar dichas trayectorias.

Existen ventajas para una formulación variacional de los sistemas físicos, una de ellas es que puede hacerse sin hacer referencia al sistema en particular de coordenadas generalizadas, por tanto podemos decir que es invariante respecto a la elección de las coordenadas del sistema. Otra es que nosotros podemos añadir a la lagrangiana una derivada total con respecto del tiempo de una función cualquiera de las posiciones y del tiempo sin que afecte el comportamiento variacional de la integral de acción.

Otra ventaja muy importante es que este tratamiento puede amplicarse fácilmente a otros sistemas no precisamente mecánicos, es decir podemos considerar por ejemplo el campo electromagnético, gravitacional, etc.

Por último, otra ventaja es que se pueden describir dos o más sistemas físicos distintos por lagrangianos de la misma forma, esto significa que el investigar un sistema físico podemos obtener resultados, también, en el otro sistema y viceversa.

La lagrangiana y el principio de Hamilton (que en adelante llamaremos simplemente de mínima acción, por conveniencia) forman, juntos, una manera de implicar ecuaciones de movimiento. Casi en todos los campos de la física se pueden utilizar principios variacionales para expresar "ecuaciones de movimiento", ya sean las ecuaciones de Newton, las ecuaciones de Maxwell, las ecuaciones de Schrödinger o las ecuaciones de Einstein, por mencionar sólo algunas. Por tanto se utiliza un principio variacional como base de la formulación, entonces podemos hablar de cierta analogía estructural en todos estos campos. Cuando los resultados experimentales muestran la necesidad de alterar el contenido físico de la teoría de un campo, este grado de analogía ha indicado, muchas veces, como pueden efectuarse alteraciones semejantes en otros campos.

Así, los experimentos de principios de siglo mostraron que era necesaria la cuantización de la radiación electromagnética y las partículas elementales. Pero esta cuantización ya se había dado anteriormente para sistemas mecánicos, por lo que describiendo al campo electromagnético mediante una lagrangiana y el correspondiante principio variacional de Hamilton, fue posible desarrollar los métodos de cuantización en la radiación y dar lugar a la Electrodinámica Cuántica desarrollada por R.P. Feynmann, Schwinger y Tomonaga [3],[26].

3.1.3 ALGUNOS TEOREMAS DE CONSERVACION

En ciertos casos especiales es posible encontrar soluciones parciales a las ecuaciones de Euler-Lagrange (3.10) en forma de primeras integrales, es decir que $f\left(t, q^1, q^2, ..., q^n\right)$ es constante sobre la extremal, es decir $f\left(t, q^1, q^2, ..., q^n\right) = cte.$

Analicemos algunos casos, pero antes consideremos sistemas sometidos a fuerzas que deriven de un potencial que sólo depende de las posición, entonces

$$\begin{aligned}\frac{\partial L}{\partial \dot{q}^k} &\equiv \frac{\partial T}{\partial \dot{q}^k} - \frac{\partial V}{\partial \dot{q}^k} = \frac{\partial T}{\partial \dot{q}^k} \\ &= \frac{\partial}{\partial \dot{q}^k}\left(\frac{1}{2} m_k \dot{q}_k^2\right) = p_k\end{aligned}$$

es decir

$$\frac{\partial L}{\partial \dot{q}^k} = p_k \tag{3.19}$$

A p_k se le llama **cantidad de movimiento generalizada canónica o conjugada** o simplemente **momento canónico o conjugado**. Esta no tiene precisamente unidades de cantidad de movimiento.

Cuando la lagrangiana de un sistema no contenga una coordenada q^k (aun cuando pueda contener su velocidad generalizada correspondiante $\dot{q}^k$) diremos que la **coordenada** es **cíclica o ignorable**. Entonces la ecuación de Euler-Lagrange correspondiente a esa coordenada

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) - \frac{\partial L}{\partial q^k} = 0$$

se reducirá a

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) = 0$$

donde sustituyendo la definición del momento canónico (3.46) tenemos que

$$\frac{dp_k}{dt} = 0 \tag{3.20}$$

entonces $p_k =$ constante. Por lo que podemos enunciar un primer teorema de conservación:

La cantidad de movimiento generalizada canónica a una coordenada cíclica se conserva.

Existen dos clases de momento, cuando la variación se hace translacionalmente, diremos que se conserva el momento lineal y cuando la variación es rotacional, se conservará el momento angular.

Otro teorema importante es el siguiente. Cuando el tiempo es homogéneo, es decir, cuando estamos en un sistema de referencia inercial (es decir que viaja a velocidad constante respecto otro‡) o cuando el sistema se encuentra en un campo de fuerzas uniforme tenemos que

$$\frac{\partial L}{\partial t} = 0$$

es decir que L no depende explícitamente del tiempo, entonces

$$\frac{dL}{dt} = \frac{\partial L}{\partial q^k}\frac{dq^k}{dt} + \frac{\partial L}{\partial \dot{q}^k}\frac{d\dot{q}^k}{dt}$$

donde se aplica la convención de suma. Si sustituimos las ecuaciones de Euler-Lagrange tenemos que la ecuación anterior queda como

$$\frac{dL}{dt} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right)\dot{q}^k + \frac{\partial L}{\partial \dot{q}^k}\ddot{q}^k = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\dot{q}^k\right)$$

o lo que es lo mismo

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\dot{q}^k - L\right) = 0$$

si

$$\frac{\partial L}{\partial \dot{q}^k}\dot{q}^k - L = h$$

entonces diremos que h se conserva. Si la energía potencial sólo depende de las coordenadas $V(q^k)$, entonces

$$h = \frac{\partial T}{\partial \dot{q}^k}\dot{q}^k - L \tag{3.21}$$

Sabemos que cuando T se expresa en coordenadas rectangulares, podemos escribirla, de manera general como

$$T = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=i}^{3} m_i \dot{x}_{ij}^2$$

Ahora queremos determinar como depende T de las coordenadas y velocidades generalizadas tales que

$$x_{ij} = x_{ij}\left(q^j, t\right)$$

Haciendo los cálculos§ llegamos a que

$$T = a_{ij}\dot{q}^i\dot{q}^j$$

donde a_{ij} son ciertas constantes, entonces

$$\frac{\partial T}{\partial \dot{q}^k} = a_{ik}\dot{q}^i + a_{kj}\dot{q}^j$$

‡En un capítulo posterior hablaremos con más detalle de los sistemas de referencia inerciales.

§Estos pueden encontrarse en [20] página 241.

Si ahora la multiplicamos por $\dot{q}^k$ tenemos que

$$\dot{q}^k \frac{\partial T}{\partial \dot{q}^k} = a_{ik}\dot{q}^i\dot{q}^k + a_{kj}\dot{q}^j\dot{q}^k$$
$$= 2a_{ik}\dot{q}^i\dot{q}^k = 2T$$

Por lo tanto (3.21) queda como

$$h = 2T - (T - V) = T + V = \text{constante}$$

A h generalmente, para sistemas conservativos, se le conoce como la **energía del sistema** (E).

3.2 CALCULO VARIACIONAL EN INTEGRALES MULTIPLES

Ahora veremos como son las Ecuaciones de Euler-Lagrange para el problema de integrales múltiples, es decir que tiene m parámetros t^α donde $\alpha = 1, ..., m$. Este tipo de problema es también muy importante dentro de la Física, especialmente en la llamada Teoría de los Campos. Para encontrar las ecuaciones de Euler-Lagrange para integrales múltiples nos basaremos en el procedimiento seguido para las integrales simples.

Fijemos ahora una región de integración D que denota un rectángulo cerrado o cilindro en R^m.Un punto en D se denotará por $t = \left(t^1, ..., t^m\right)$ y sea $dt = dt^1...dt^m$. Denotaremos a la frontera de D como ∂D. Para desarrollos posteriores necesitaremos del **teorema de la divergencia** el cual enuncia que:

Sea $D \subseteq \mathsf{R}^m$ y sea $f(t) = f\left(t^1, ..., t^m\right)$ continua en D con frontera, con primeras derivadas parciales continuas en D. Entonces

$$\int_D \frac{\partial f}{\partial t^\alpha} dt = \int_{\partial D} f(t) \cos\left(\mathsf{n}, t^\alpha\right) d\sigma \tag{3.22}$$

donde n denota el vector unitario normal hacia afuera de la frontera de D, $d\sigma$ es un elemento de superficie de ∂D y $\cos\left(\mathsf{n}, t^\alpha\right)$ denota el coseno del ángulo entre n y la $t^\alpha -$ *ésima* coordenada. ■

Supongamos que $q(t) \in C_n^2(D)$, es decir que $q(t) = \left(q^1(t), q^2(t), ..., q^n(t)\right)$ donde cada una de las $q^k(t)$ son funciones sobre D con derivadas parciales continuas hasta de segundo orden tal que $q(t) = \phi(t)$ con $t \in \partial D$ y ϕ dada sobre ∂D.

Definimos nuestro problema variacional de la siguiente manera:

Sea $L = L\left(t^1, ..., t^m, q^1, ..., q^n, y^1, ..., y^{mn}\right)$ una función de valor real definida sobre R^{m+n+mn} la cual es dos veces diferenciable en cada uno de sus argumentos. Denotaremos a las primeras y segundas derivadas parciales de las componentes $q(t)$ por

$$\dot{q}_\alpha^k = \frac{\partial q^k}{\partial t^\alpha}, \qquad \ddot{x}_{\alpha\beta}^k = \frac{\partial^2 q^k}{\partial t^\alpha \partial t^\beta} \tag{3.23}$$

Entonces la integral fundamental se define como

$$I(q(t)) = \int_D L\left(t^1, ..., t^m, q^1(t), ..., q^n(t), \dot{q}_1^1, ..., \dot{q}_1^n, \dot{q}_m^1, ..., \dot{q}_m^n\right) dt \tag{3.24}$$

con dt como se definió anteriormente, o la integral en forma vectorial se escribirá como

$$I(q(t)) = \int_D L\left(t, q(t), \frac{\partial q(t)}{\partial t}\right) dt \tag{3.25}$$

donde $\dfrac{\partial q(t)}{\partial t}$ denota el conjunto de todas las primeras derivadas parciales $\dfrac{\partial q^k}{\partial t^\alpha}$, con $k = 1, ..., n$ y $\alpha = 1, ..., m$.

Siguiendo la aproximación general para integrales sencillas sea $q(t)$ tal que haga que (3.25) sea mínima. Entonces sea

$$q(t, \varepsilon) = q(t) + \varepsilon\eta(t) \tag{3.26}$$

donde $\eta(t) = \left(\eta^1(t), ..., \eta^n(t)\right) \in C_n^2$, $\eta(t) = \mathbf{0}$ donde $t \in \partial D$ y $q(t, \varepsilon) = q^1(t, \varepsilon), ..., q^n(t, \varepsilon)$. Calculemos la primera variación de I aplicando la variación de Gâteaux (3.3)

$$\delta I = \left[\frac{d}{d\varepsilon}\int\limits_D L\left(t, q(t) + \varepsilon\eta(t), \frac{\partial q(t)}{\partial t} + \varepsilon\frac{\partial\eta(t)}{\partial t}\right) dt\right]_{\varepsilon=0} = 0 \tag{3.27}$$

para todo $\eta(t)$ como se definió anteriormente. Tanto L como $\dfrac{\partial L}{\partial\varepsilon}$ son funciones continuas de ε y t, la derivada en (3.27), por tanto, puede entrar dentro de la integral y entonces por regla de la cadena y evaluando en $\epsilon = 0$ tenemos que

$$\delta I = \int\limits_D \left(\frac{\partial L}{\partial q^k}\eta^k(t) + \frac{\partial L}{\partial\dot{q}_\alpha^k}\dot{\eta}_\alpha^k(t)\right) dt = 0 \tag{3.28}$$

Ahora, haciendo uso de la identidad

$$\frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial\dot{q}_\alpha^k}\eta^k\right) = \frac{\partial L}{\partial\dot{q}_\alpha^k}\dot{\eta}_\alpha^k + \frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial\dot{q}_\alpha^k}\right)\eta^k \tag{3.29}$$

la ecuación (3.28) se puede reescribir como

$$\int\limits_D \left[\frac{\partial L}{\partial q^k} - \frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial\dot{q}_\alpha^k}\right)\right]\eta^k(t)dt + \int\limits_D \frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial\dot{q}_\alpha^k}\eta^k(t)\right) dt = 0 \tag{3.30}$$

Dado que las funciones $\dfrac{\partial L}{\partial\dot{q}_\alpha^k}\eta^k(t)$ son continuamente diferenciables sobre D, entonces satisfacen las condiciones del teorema de la divergencia (3.22) y al aplicarlo a la segunda integral obtenemos

$$\int\limits_D \left[\frac{\partial L}{\partial q^k} - \frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial\dot{q}_\alpha^k}\right)\right]\eta^k(t)dt + \int\limits_{\partial D} \frac{\partial L}{\partial\dot{q}_\alpha^k}\eta^k(t)\cos\left(\eta^k(t), t^\alpha\right) d\sigma = 0 \tag{3.31}$$

La integral sobre la frontera de D se desprecia dado que $\eta^k(t) = 0$ sobre ella y por lo tanto obtenemos que

$$\int\limits_D \left[\frac{\partial L}{\partial q^k} - \frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial\dot{q}_\alpha^k}\right)\right]\eta^k(t)dt = 0 \tag{3.32}$$

Mencionaremos un lema análogo al de Dubois-Reymond (fundamental del cálculo de variaciones). Si $f(t)$ es continua y de valor real sobre D y si

$$\int_D f(t)h(t)dt = 0$$

para toda $h(t) \in C^2(D)$ la cual se desprecia sobre ∂D, entonces

$$f(t) \equiv 0$$

para todo $t \in D$. ■

Si

$$E_k(t) = \frac{\partial L}{\partial q^k} - \frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial \dot{q}_\alpha^k}\right)$$

entonces haciendo uso del lema anterior del que se puede ver la demostración en [18] y que es análogo al (3.9), la ecuación (3.32) queda como

$$\int_D E_k(t)\,\eta^k(t)dt = 0 \tag{3.33}$$

para todo $\eta^k(t)$ con $\eta^k(t) = 0$ para todo $t \in \partial D$. Sea $\eta^k(t) = 0$ para $k \neq i$ donde i es un índice fijo entre 1 y n. Entonces

$$\int_D E_i(t)\,\eta^i(t)dt = 0$$

para todos los posibles $\eta^i(t) \in C^2(D)$ con $\eta^i(t) = 0$ sobre ∂D. Entonces $E_i(t)$ es continua sobre D y por el Lema (3.9) tenemos que $E_i(t) = 0$ sobre D. Haciendo $i = 1, ..., n$, obtenemos que si $q(t)$ como lo definimos anteriormente es un mínimo local de la funcional I definida por (3.25), entonces las componentes $q^k(t)$ de $q(t)$ satisfacen las n ecuaciones

$$E_k(t) = \frac{\partial L}{\partial q^k} - \frac{\partial}{\partial t^\alpha}\left(\frac{\partial L}{\partial \dot{q}_\alpha^k}\right) = 0 \tag{3.34}$$

sobre D. Estas son las **Ecuaciones de Euler-Lagrange para m parámetros**.

3.2.1 PROPIEDADES DE TRANSFORMACION DE LA INTEGRAL FUNDAMENTAL MULTIPLE.

Ya definimos que un punto en nuestro espacio de configuración de las integrales múltiples se puede denotar como $q(t) = \left(q^1(t), q^2(t), ..., q^n(t)\right)$ donde $t = \left(t^1, ..., t^m\right)$, es decir

$$q^h = q^h(t^\alpha)$$

donde supusimos que las funciones $q^h(t^\alpha)$ son de clase C^2 y por tanto podemos construir las derivadas como

$$\dot{q}_\alpha^h = \frac{\partial q^h}{\partial t^\alpha}$$

Ahora, bajo una transformación de coordenadas de clase C^2 del tipo (1.1) con un jacobiano no nulo, es decir, que existe la transformación inversa; las derivadas de esta transformación respecto de t^α serán

$$\dot{q}_\alpha'^j = \frac{\partial q'^j}{\partial t^\alpha} = \frac{\partial q'^j}{\partial q^h}\frac{\partial q^h}{\partial t^\alpha} = \frac{\partial q'^j}{\partial q^h}\dot{q}_\alpha^h \tag{3.35}$$

lo cual hace evidente que $\dot{q}_\alpha^h$ representa las componentes de un tensor contravariante relativo a la transformación (1.1). Sin embargo, bajo una transformación paramétrica de clase C^2 del tipo

$$t'^\alpha = t'^\alpha\left(t^\beta\right) \tag{3.36}$$

también con jacobiano no nulo, tenemos que

$$\frac{\partial q^j}{\partial t'^\beta} = \frac{\partial q^j}{\partial t^\alpha}\frac{\partial t^\alpha}{\partial t'^\beta} = \frac{\partial t^\alpha}{\partial t'^\beta}\dot{q}_\alpha^j \tag{3.37}$$

de lo que se sigue que las derivadas $\dot{q}_\alpha^j$ son las componentes de un tensor covariante relativo a la transformación (3.36).

Supongamos que L es una escalar o invariante bajo la transformación de coordenadas (1.1), entonces tenemos que

$$\bar{L}\left(t^\alpha, q'^j, \dot{q}'^j_\alpha\right) = L\left(t^\beta, q^k, \dot{q}^k_\beta\right) \tag{3.38}$$

de tal manera que

$$\frac{\partial L}{\partial \dot{q}^k_\beta} = \frac{\partial \bar{L}}{\partial \dot{q}'^j_\alpha}\frac{\partial \dot{q}'^j_\alpha}{\partial \dot{q}^k_\beta} \tag{3.39}$$

y

$$\frac{\partial L}{\partial q^k} = \frac{\partial \bar{L}}{\partial q'^j}\frac{\partial q'^j}{\partial q^k} + \frac{\partial \bar{L}}{\partial \dot{q}'^j_\alpha}\frac{\partial \dot{q}'^j_\alpha}{\partial q^k} \tag{3.40}$$

Ahora de (3.35) se sigue directamente que

$$\frac{\partial \dot{q}'^j_\alpha}{\partial \dot{q}^k_\beta} = \frac{\partial q'^j}{\partial q^h}\frac{\partial \dot{q}'^h_\alpha}{\partial \dot{q}^k_\beta} = \frac{\partial q'^j}{\partial q^h}\delta^h_k\delta^\alpha_\beta = \frac{\partial q'^j}{\partial q^k}\delta^\alpha_\beta \tag{3.41}$$

junto con

$$\frac{\partial \dot{q}'^j_\alpha}{\partial q^k} = \frac{\partial^2 q'^j}{\partial q^h \partial q^k}\dot{q}^h_\alpha \tag{3.42}$$

Cuando sustituimos (3.41) en (3.39) obtenemos

$$\frac{\partial L}{\partial \dot{q}^k_\beta} = \frac{\partial \bar{L}}{\partial \dot{q}'^j_\alpha}\frac{\partial q'^j}{\partial q^k}\delta^\alpha_\beta = \frac{\partial q'^j}{\partial q^k}\frac{\partial \bar{L}}{\partial \dot{q}'^j_\beta} \tag{3.43}$$

donde podemos ver que estas cantidades representan las componentes de un tensor covariante. Sin embargo, si sustituimos (3.42) en (3.40) encontramos que

$$\frac{\partial L}{\partial q^k} = \frac{\partial \bar{L}}{\partial q'^j}\frac{\partial q'^j}{\partial q^k} + \frac{\partial \bar{L}}{\partial \dot{q}'^j_\alpha}\frac{\partial^2 q'^j}{\partial q^h \partial q^k}\dot{q}^h_\alpha \tag{3.44}$$

donde podemos ver que $\dfrac{\partial L}{\partial q^k}$ no son las componentes de un tensor. Si diferenciamos (3.43) respecto de t^β (donde no hay suma sobre β) tenemos que

$$\frac{d}{dt^\beta}\left(\frac{\partial L}{\partial \dot{q}^k_\beta}\right) = \frac{\partial q'^j}{\partial q^k}\frac{d}{dt^\beta}\left(\frac{\partial \bar{L}}{\partial \dot{q}'^j_\beta}\right) + \frac{\partial \bar{L}}{\partial \dot{q}'^j_\beta}\frac{\partial^2 q'^j}{\partial q^h \partial q^k}\dot{q}^h_\beta$$

Si restamos de esta última ecuación la ecuación (3.44) resulta que

$$\frac{d}{dt^\beta}\left(\frac{\partial L}{\partial \dot{q}^k_\beta}\right) - \frac{\partial L}{\partial q^k} = \frac{\partial q'^j}{\partial q^k}\left[\frac{d}{dt^\beta}\left(\frac{\partial \bar{L}}{\partial \dot{q}'^j_\beta}\right) - \frac{\partial \bar{L}}{\partial q'^j}\right] \tag{3.45}$$

de la cual deducimos que las cantidades definidas por (3.34) constituyen las componentes de un tensor covariante.

A partir de estas ecuaciones de movimiento de Euler-Lagrange (3.10) y (3.34) se pueden construir casi todas las teorías físicas de la actualidad.

3.3 FORMULACION DE HAMILTON

Esta formulación no tiene nada de nuevo y no es superior a la técnica de Lagrange, que hemos estudiado hasta ahora, lo único que haremos es obtener otro método, que es más potente para trabajar con los principios físicos ya establecidos con anterioridad.

Una de las utilidades de la formulación Hamiltoniana consiste en proporcionar un marco para extensiones teóricas en muchos campos de la Física. Dentro de la mecánica clásica es base de la teoría de Hamilton-Jacobi y los métodos de perturbaciones [10],[24]. Dentro de la mecánica estadística y la mecánica cuántica constituye gran parte del lenguaje usado ahí.

De aquí en adelante consideraremos a los sistemas mecánicos de manera que sus fuerzas deriven de un potencial que sólo dependa de las posiciones.

Hemos visto anteriormente que el momento canónico conjugado lo definimos como

$$\frac{\partial L}{\partial \dot{q}^k} = p_k \tag{3.46}$$

En adelante llamaremos a las cantidades (q, p) como variables canónicas del sistema.

Si observamos como un problema estrictamente matemático, lo que haremos para pasar de la formulación de Lagrange a la de Hamilton es hacer un cambio de variables, es decir pasaremos de las variables $\left(q^k, \dot{q}^k, t\right)$ a las variables $\left(q^k, p_k, t\right)$ donde p_k se relaciona con q^k y $\dot{q}^k$ como en (3.46). El método para conmutar las variables de esta forma lo proporciona la **Transformación de Legendre**, planeada para este tipo de cambio de variables.

Consideremos una función $f(x^1, ..., x^n)$ de clase, al menos, C^2 tal que su diferencial sea

$$df = \frac{\partial f}{\partial x^k} dx^k \tag{3.47}$$

donde $k = 1, 2, ..., n$ y se tome en cuanta la convención de suma (1.3). Definimos a p_k como

$$p_k = \frac{\partial f}{\partial x^k} \qquad k = 1, ..., n \tag{3.48}$$

Ahora deseamos cambiar las primeras r variables $\left(x^1, x^2, ..., x^r\right)$ por las r variables $(p_1, p_2, ...p_r)$ de manera que las diferenciales se expresen en función de las primeras r diferenciales dp_k. Sea g una función de las variables p_k tal que

$$g = \sum_{k=1}^{r} x^k p_k - f \tag{3.49}$$

entonces su diferencial será

$$dg = \sum_{k=1}^{r} x^k dp_k + \sum_{k=1}^{r} p_k dx^k - df$$

y de (3.47) tenemos que

$$dg = \sum_{k=1}^{r} x^k dp_k - \sum_{k=r+1}^{n} p_k dx^k$$

que es exactamente la que buscabamos. Podemos ver que

$$x^k = \frac{\partial g}{\partial p_k} \qquad k = 1, ..., n$$

$$p_k = -\frac{\partial g}{\partial x^k} \qquad k = 1, ..., n$$

son de hecho las inversas de (3.48).

Regresemos a la mecánica clásica, para cambiar de las coordenadas $\left(q^k, \dot{q}^k, t\right)$ a $\left(q^k, p_k, t\right)$ tomaremos a nuestra función, en analogía con (3.49) como

$$H\left(q^k, p_k, t\right) = q^k p_k - L\left(q^k, \dot{q}^k, t\right) \tag{3.50}$$

A esta función se le llama **Hamiltoniana**. La diferencial de la hamiltoniana será

$$dH = \frac{\partial H}{\partial q^k} dq^k + \frac{\partial H}{\partial p_k} dp_k + \frac{\partial H}{\partial t} dt \tag{3.51}$$

pero según (3.50) tenemos que

$$\frac{\partial H}{\partial q^k} = -\frac{\partial L}{\partial q^k}$$

y

$$\frac{\partial H}{\partial p_k} = \dot{q}^k + p_k \frac{\partial \dot{q}^k}{\partial p_k} - \frac{\partial L}{\partial \dot{q}^k} \frac{\partial \dot{q}^k}{\partial p_k}$$

entonces

$$dH = -\frac{\partial L}{\partial q^k} dq^k + \dot{q}^k dp_k + p_k d\dot{q}^k - \frac{\partial L}{\partial \dot{q}^k} d\dot{q}^k - \frac{\partial L}{\partial t} dt$$

Por la ecuación (3.46) tenemos que las ecuaciones de Euler-Lagrange son

$$\dot{p}_k = \frac{\partial L}{\partial q^k} \tag{3.52}$$

entonces introduciendo tanto (3.46) como (3.52) en la última diferencial de H tenemos que los términos tercero y cuarto se anulan y sólo nos queda que

$$dH = -\dot{p}_k dq^k + \dot{q}^k dp_k - \frac{\partial L}{\partial t} dt$$

Si comparamos esta ecuación con la ecuación (3.51) tenemos que

$$\begin{aligned} \dot{q}^k &= \frac{\partial H}{\partial p_k} \\ \dot{p}_k &= -\frac{\partial H}{\partial q^k} \\ \frac{\partial L}{\partial t} &= \frac{\partial H}{\partial t} \end{aligned} \tag{3.53}$$

Estas son las llamadas **Ecuaciones canónicas de Hamilton** y constituyen un sistema de $2n$ ecuaciones diferenciales de primer orden que sustituyen a las ecuaciones de Euler-Lagrange. Notemos que estas ecuaciones tienen un aspecto simétrico.

Chapter 4
PRINCIPIOS DE LA TEORIA DE LA RELATIVIDAD

Aquí daremos algunas ideas a cerca de como tuvo que cambiarse el pensamiento de que el tiempo era absoluto por el de un tiempo relativo. Todo ello nos conduce hacía un mundo distinto al que podemos percibir cotidianamente.

Las ideas básicas en las que se basó A. Einstein en su famosa teoría de la Relatividad Especial, publicada en 1905 no como una teoría nueva sino más bien como una explicación a ciertos fenómenos que se observaban dentro de la electrodinámica clásica y que no podían ser totalmente explicados simplemente con la mecánica clásica, son las que desarrollaremos a lo largo de este capítulo.

Obviamente es, creo yo, difícil describir nuevamente todos los fenómenos que se estudian dentro de la mecánica clásica con este nuevo enfoque, dada la extensión que debería que tener este capítulo; sin embargo, dentro de las referencias de este trabajo existe una gran variedad de libros que pueden ser consultados para obtener mayor precisión en algunos de los conceptos tratados aquí o algunos otros que no trataremos.

4.1 TRANSFORMACION DE GALILEO

Para describir cualquier proceso físico necesitamos poder definir la posición de éste en un sistema de coordenadas y además donde podamos medir el tiempo, al cual entenderemos como un **sistema de referencia**.

Si consideramos una partícula que viaje a velocidad constante y además tomamos un sistema de referencia para describir a dicha partícula en el que permanezca en reposo, estaremos considerando que dicho sistema de referencia lleva una velocidad constante y por lo tanto lo llamaremos **inercial**.

Si tenemos dos sistemas de referencia que se mueven con movimiento rectilíneo y uniforme uno con respecto de otro y si además uno de ellos es inercial, el otro también lo será.

La experiencia nos ha mostrado que todas las leyes de la naturaleza son las mismas en todos los sistemas de referencia inerciales, es decir que las ecuaciones que expresan dichas leyes son invariantes respecto a transformaciones de coordenadas y de tiempo, las cuales hacen pasar de un sistema inercial a otro. Al razonamiento anterior lo conocemos como principio de relatividad. Bondi expresa esto de la siguiente manera al comienzo de un artículo en el que pasa revista a la relatividad*:

El propósito de cualquier teoría física es describir de una manera concisa una gran variedad de fenómenos. En muchos casos esto necesitará, como parte de la teoría, una prescripción para aplicar la teoría a sistemas que se encuentren en estados de movimiento diferentes. Una prescripción de este tipo, es decir, una especie de código de traducción, consistirá generalmente en un sistema matemático de leyes de transformación.

*La transcripción posterior es de la referencia [[9]] en donde se menciona la siguiente referencia: Bondi, H., Rept. Progr. Phys. 22, 97-120 (1959)

Pertenece a la naturaleza de las leyes de transformación el cambiar la mayor parte de las cantidades y el dejar invariantes algunas de ellas. Estas últimas son las denominadas invariantes de la transformación y sirven para definir su carácter. Una afirmación en el terreno de la física de cuáles son estas invariantes se denomina principio de relatividad, y las ecuaciones fundamentales de una teoría definen generalmente el principio de relatividad que le es aplicable.

Galileo es a quien debemos la primera discusión clara de tales cuestiones. En su libro en el que aboga por la idea de Copernico acerca del sistema solar en contra de la de Ptolomeo[†], Galileo decía que el hecho de que se dejara caer libremente un objeto y este describiera una trayectoria vertical no podía implicar que la Tierra se encontrara en reposo. Una analogía que se hacia es la de una piedra que es dejada caer desde lo más alto del mástil de un barco, ésta caerá al pie del mástil sin importan si el barco se encuentra en reposo o en movimiento rectilíneo y uniforme. Entonces se habla de un principio de relatividad, puesto que contiene la afirmación de que una cierta ley de la naturaleza (la ley de la caída libre de los cuerpos) es la misma en todos los sistemas de referencia que difieran tan sólo en una velocidad constante.

Supongamos ahora que tenemos dos sistemas de referencia, el barco en movimiento horizontal con velocidad constante y otro fijo, por ejemplo en tierra. Un observador situado en el barco observará que la trayectoria que describe la piedra es una línea vertical, mientras que uno que este colocado en el sistema fijo observará que la piedra describe una trayectoria parabólica. Si analizamos estos movimientos mediante la dinámica de Newton descubriremos que, aunque los movimientos son diferentes, las conclusiones que pueden obtenerse acerca del valor de la aceleración, y por consiguiente de la fuerza que causa el movimiento, son iguales.

La relación entre las medidas tomadas, tanto en un sistema de referencia inercial como en otro, esta expresada mediante un conjunto de ecuaciones de transformación, que en la mecánica Newtoniana son conocidas como Ecuaciones de **Transformación de Galileo**, estas son:

$$\begin{aligned} x' &= x - vt \\ y' &= y \\ z' &= z \\ t' &= t \end{aligned} \tag{4.1}$$

donde v es la velocidad que un sistema de referencia inercial lleva con respecto del otro. Además haciendo algunos cálculos podemos encontrar que

$$\begin{aligned} u'_x &= u_x - v \\ a'_x &= a_x \end{aligned} \tag{4.2}$$

donde u_x y a_x son la velocidad del objeto en el sistema de referencia en reposo y su aceleración respectivamente.

Las ecuaciones de la **transformación inversa de Galileo** son

$$\begin{aligned} x &= x' + vt \\ y &= y' \\ z &= z' \\ t &= t' \end{aligned} \tag{4.3}$$

de donde podemos obtener

$$\begin{aligned} u_x &= u'_x + v \\ a_x &= a'_x \end{aligned} \tag{4.4}$$

[†]Galileo, **Dialogos acerca de los dos sistemas principales del Universo -el de Ptolomeo y el de Copérnico**, Imprenta de la Universidad de California, Berkeley, 1953.

Las tres primeras ecuaciones (4.1) bastan por sí solas para definir la relación entre dos sistemas inerciales en movimiento relativo. No se hubiese considerado la cuarta, puesto que el concepto de tiempo universal es inherente al fundamento de la mecanica Newtoniana: "El tiempo absoluto, verdadero y matemático, por sí mismo y debido a su propia naturaleza transcurre sin relación alguna con el exterior[‡]"

Dadas estas ecuaciones, la relación de las medidas de la velocidad en sistemas inerciales diferentes queda definida por completo y la medida de una aceleración es la misma para todos esos sistemas. En la mecánica newtoniana la aceleración es invariante.

4.2 PRINCIPIOS DE RELATIVIDAD

La interacción de partículas materiales se describe en la mecánica newtoniana mediante una energía potencial de interacción que es función de las coordenadas de las partículas que interactuan. De lo anterior podemos ver que describir interacciones supone una propagación instantánea de las mismas. En efecto, un cambio en la posición de cualquiera de las partículas repercute inmediatamente sobre las otras partículas. Sin embargo, muestra la experiencia que no existen interacciones instantáneas en la naturaleza. Esto quiere decir que si en una de las partículas ocurre un cambio, éste influirá sobre alguna otra partícula después de transcurrido cierto intervalo de tiempo. Dividiendo la distancia entre las dos partículas por dicho intervalo de tiempo obtenemos la **velocidad de propagación de interacción**.

Dado que la velocidad de propagación de interacción en todos los sistemas debe ser la misma (principio de relatividad), entonces esta velocidad es una constante universal. Esta velocidad constante, es la velocidad de propagación de la luz en el vacío, se designa como c y su valor, determinado por medio de muchos experimentos (ver por ejemplo [9], capítulo 2), es igual a

$$c = 2.99793X10^{10}cm/seg \tag{4.5}$$

Se puede ver que el hecho de que esta velocidad sea tan grande implica que la mecanica clásica o newtoniana considere que la velocidad de propagación sea infinita, es decir que las interacciones sean instantáneas.

Al conjunto: Principio de relatividad y el de la finitud de la velocidad de propagación de las interacciones, se le llama **Principio de Relatividad de Einstein** y fue formulado por el mismo Albert Einstein en 1905. Al principio de relatividad con el hecho de que la velocidad de propagación de las interacciones sea infinita se le conoce como **Principio de Relatividad de Galileo**.

En mecánica clásica, el concepto de distancia es relativo, es decir, las relaciones espaciales entre sucesos diferentes dependen del sistema de referencia en que se describen. El tiempo, como ya dijimos, es absoluto, es decir que las propiedades del tiempo no dependen del sistema de referencia.

La idea de un tiempo absoluto está en contradicción con la relatividad especial de Einstein. Para convencernos mencionaremos que de la mecánica clásica vale la ley de composición de velocidades, la cual dice que la velocidad de un movimiento compuesto es igual a la suma (vectorial) de las velocidades que componen dicho movimiento. Y dado que esta ley es universal, también debería aplicarse a la propagación de interacciones. Entonces la velocidad de esta propagación sería diferente en distintos sistemas inerciales, en contradicción con el principio de relatividad.

Mediciones realizadas, primeramente, por Michelson en 1881 y otros posteriormente confirmaron el principio de relatividad de Einstein, es decir, se pone de manifiesto que la velocidad de la luz es independiente de la dirección de propagación de ésta.

[‡]Newton, **Principia**, Libro I, "Scholium after Definitions"

Por lo tanto, el principio de relatividad conduce a que el tiempo no es absoluto. Los intervalos de tiempo son distintos en distintos sistemas de referencia. En particular, sucesos que son simultáneos en un sistema de referencia, no serán simultáneos en otro.

4.3 ESPACIO-TIEMPO

Un **suceso** se define por el lugar en el que ocurre y el instante en el que ocurre. Así, si algún suceso ocurre en una cierta partícula, este se definirá mediante las tres coordenadas de la partícula y por el instante en que ocurre.

Vamos a considerar un espacio ficticio cuadridimensional, cuyos ejes serán las tres coordenadas espaciales y el tiempo, comúnmente se le llama a este espacio ficticio **espacio-tiempo**. Los puntos en dicho espacio les llamaremos **puntos universo**. Un punto universo lo denotaremos por

$$x^\mu = (x^0, x^1, x^2, x^3) \text{ con } \mu = 0, 1, 2, 3$$

y donde para el sistema cartesiano tenemos que $x^0 = ct$, $x^1 = x$, $x^2 = y$ y $x^3 = z$. A cada partícula le corresponde una cierta trayectoria en este espacio-tiempo, a la que llamaremos **línea de universo,** donde cada punto de esta trayectoria representa la posición que ocupa la partícula en todo instante.

Mostremos una propiedad importante en la relatividad especial. Para ello consideremos dos sistemas de referencia K y K' que se mueven con movimiento relativo uniforme. En estos sistemas, los ejes X y X' coinciden, mientras que los ejes Y y Z son paralelos a Y' y Z', el tiempo correspondiente a los dos sistemas son t y t' respectivamente y además que la velocidad de la luz es un invariante.

Supongamos que un primer suceso consiste en que una señal, que viaja a la velocidad de la luz, de un punto x^1, y^1, z^1 al tiempo t^1 y un segundo suceso consiste en que dicha señal llega a x^2, y^2, z^2 en el instante t^2 en el sistema K. La distancia recorrida por la señal, por un lado, es $c(t^2 - t^1)$ y por el otro $\left[\left(x^2 - x^1\right)^2 + \left(y^2 - y^1\right)^2 + \left(z^2 - z^1\right)^2\right]^{\frac{1}{2}}$, por lo que podemos escribir la siguiente relación entre las coordenadas de ambos sucesos en el sistema K como

$$c^2(t^2 - t^1)^2 - \left(x^2 - x^1\right)^2 - \left(y^2 - y^1\right)^2 - \left(z^2 - z^1\right)^2 = 0 \tag{4.6}$$

Los mismos sucesos se pueden observar desde el sistema K' donde x'_1, y'_1, z'_1, t'_1 y x'_2, y'_2, z'_2, t'_2 son las coordenadas de los sucesos respectivamente. Entonces, como (4.6), tenemos que

$$c^2(t'^2 - t'^1)^2 - \left(x'^2 - x'^1\right)^2 - \left(y'^2 - y'^1\right)^2 - \left(z'^2 - z'^1\right)^2 = 0 \tag{4.7}$$

A la cantidad

$$s_{12} = \left[c^2(t^2 - t^1)^2 - \left(x^2 - x^1\right)^2 - \left(y^2 - y^1\right)^2 - \left(z^2 - z^1\right)^2\right]^{\frac{1}{2}} \tag{4.8}$$

se le llama **intervalo** entre dos sucesos.

Del principio de invariancia de las leyes físicas se sigue, por consiguiente que si el intervalo entre dos sucesos es cero en un sistema de referencia, es también cero en cualquier otro.

Si los dos sucesos son infinitamente próximos entre sí, para el intervalo ds entre ellos se tiene

$$ds^2 = c^2\,(dt)^2 - (dx)^2 - (dy)^2 - (dz)^2 \tag{4.9}$$

Dado que consideramos a la velocidad de la luz como un invariante entonces

$$ds^2 = ds'^2$$

A este espacio-tiempo, como el que estamos trabajando, se le conoce como **espacio de Minkowski**. Al conjunto de coeficientes de la parte derecha de la ecuación (4.9), es decir $(+1, -1, -1, -1)$ se le conoce como **signatura del espacio** y en algunos textos se nombra como la suma de ellos, en este caso sería -2. La ecuación (4.9) puede escribirse en notación más compacta como

$$ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu \tag{4.10}$$

en donde hay convención de suma. Las componentes $\eta_{\mu\nu}$ forman una matriz cuadrada

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{4.11}$$

y $\eta_{\mu\nu}$ es llamado **tensor**[§] **métrico para el espacio de Minkowski**.

Ahora, el **tensor métrico para cualquier espacio** se define como

$$g_{\mu\nu} = \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\mu} \frac{\partial \xi^\beta}{\partial x^\nu} \tag{4.12}$$

donde ξ^α es un sistema de coordenadas para el espacio de Minkowski, $\eta_{\alpha\beta}$ es como el definido en (4.11) y x^μ son las coordenadas en cualquier otro espacio. En algún otro espacio diferente x'^μ el tensor métrico es

$$\begin{aligned} g'_{\mu\nu} &= \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x'^\mu} \frac{\partial \xi^\beta}{\partial x'^\nu} \\ &= \eta_{\alpha\beta} \frac{\partial \xi^\alpha}{\partial x^\rho} \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial \xi^\beta}{\partial x^\sigma} \frac{\partial x^\sigma}{\partial x'^\nu} \end{aligned}$$

y por tanto

$$g'_{\mu\nu} = g_{\rho\sigma} \frac{\partial x^\rho}{\partial x'^\mu} \frac{\partial x^\sigma}{\partial x'^\nu} \tag{4.13}$$

Lo que demuestra que $g_{\mu\nu}$ es un tensor covariante de orden dos. Su inverso es un tensor contravariante definido por $g^{\lambda\mu}$ entonces

$$g_{\mu\nu} g^{\lambda\mu} = \delta^\lambda_\nu \tag{4.14}$$

donde δ^λ_ν es la **delta de Kronecker** definida como

$$\delta^\lambda_\nu \equiv \begin{cases} 1 & \text{si } \lambda = \mu \\ 0 & \text{si } \lambda \neq \mu \end{cases}$$

podemos ver que

$$\begin{aligned} g'_{\lambda\nu} g'^{\mu\nu} &= g'_{\lambda\mu} \frac{\partial x'^\nu}{\partial x^\rho} \frac{\partial x'^\mu}{\partial x^\sigma} g^{\rho\sigma} \\ &= \frac{\partial x'^\nu}{\partial x^\rho} \frac{\partial x'^\mu}{\partial x^\sigma} g^{\rho\sigma} \frac{\partial x^\sigma}{\partial x'^\mu} \frac{\partial x^\eta}{\partial x'^\lambda} g_{\sigma\eta} \\ &= \frac{\partial x'^\nu}{\partial x^\rho} g^{\rho\sigma} \frac{\partial x^\eta}{\partial x'^\lambda} g_{\sigma\eta} \\ &= \frac{\partial x'^\nu}{\partial x^\sigma} \frac{\partial x^\sigma}{\partial x'^\lambda} = \delta'^\nu_\lambda \end{aligned}$$

[§] Aquí sólo lo definiremos así, aunque se puede demostrar que efectivamente es un tensor.

o, que también se puede demostrar que

$$g_{\lambda\nu}g^{\mu\nu} = \delta^{\nu}_{\lambda}$$

de aquí que

$$g'^{\lambda\mu}g'_{\lambda\mu}\frac{\partial x'^{\nu}}{\partial x^{\rho}}\frac{\partial x'^{\mu}}{\partial x^{\sigma}}g^{\rho\sigma} = g'^{\lambda\mu}\delta'^{\nu}_{\lambda}$$

por lo tanto

$$\frac{\partial x'^{\nu}}{\partial x^{\rho}}\frac{\partial x'^{\mu}}{\partial x^{\sigma}}g^{\rho\sigma} = g'^{\nu\mu} \tag{4.15}$$

como se requería para un tensor contravariante. Finalmente, la delta de Kronecker δ^{λ}_{ν} es un tensor mixto ya que

$$\delta^{\mu}_{\nu}\frac{\partial x'^{\rho}}{\partial x^{\mu}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}} = \frac{\partial x'^{\rho}}{\partial x^{\mu}}\frac{\partial x^{\nu}}{\partial x'^{\sigma}} = \delta^{\rho}_{\sigma} \tag{4.16}$$

El tensor métrico es útil para subir o bajas índices como se hizo en la ecuación (1.10).

Mencionaremos que el tensor obtenido por subir un índice con el tensor métrico $g^{\mu\nu}$ o bajar un índice con el tensor $g_{\mu\nu}$, es precisamente el tensor de Kronecker (4.14).

También, el subir los dos índices de $g_{\mu\nu}$ da como resultado su tensor inverso

$$g^{\mu\lambda}g^{\nu\sigma}g_{\lambda\sigma} = g^{\mu\lambda}\delta^{\nu}_{\lambda} = g^{\mu\nu}$$

y el bajar ambos índices de $g^{\mu\nu}$ nos da el tensor métrico $g_{\mu\nu}$.

Veamos algunos ejemplos de algo que no son tensores, pero casi. Un ejemplo muy importante de un no tensor es el determinante del tensor métrico

$$g \equiv -\det\ g_{\mu\nu} \tag{4.17}$$

La regla de transformación para el tensor métrico puede ser arreglada como ecuación matricial

$$g'_{\lambda\mu} = \frac{\partial x^{\rho}}{\partial x'^{\lambda}}g_{\rho\sigma}\frac{\partial x^{\sigma}}{\partial x'^{\mu}}$$

La primera matriz tiene sus líneas especificadas por λ y sus columnas especificadas por ρ, la segunda matriz tiene sus líneas especificadas por ρ y sus columnas por σ y por último, la tercera matriz tiene sus líneas especificadas por σ y sus columnas por μ. Así, este producto es igual a la matriz $g'_{\lambda\mu}$ con líneas λ y columnas μ. Ahora tomamos el determinante de la ecuación anterior

$$g' = \left|\frac{\partial x}{\partial x'}\right|^{2} g \tag{4.18}$$

donde $\left|\frac{\partial x}{\partial x'}\right|$ es el **jacobiano** de la transformación $x' \to x$, esto es el determinante de la matriz $\frac{\partial x^{\rho}}{\partial x'^{\lambda}}$. Una cantidad como g, la cual se transforma como un escalar excepto por factores extras del jacobiano, es llamada **densidad escalar**, y similarmente una cantidad que se transforme como un tensor excepto por factores extras del jacobiano es llamado **tensor de densidad**. El número de factores del determinante $\left|\frac{\partial x}{\partial x'}\right|$ es llamado **peso de la densidad** y se denota como W, por ejemplo se ve de (4.18) que g es una densidad de peso -2, porque

$$\left|\frac{\partial x}{\partial x'}\right| = \left|\frac{\partial x'}{\partial x}\right|^{-1}$$

como se observa si tomamos el determinante de la ecuación

$$\frac{\partial x^{\mu}}{\partial x'^{\sigma}}\frac{\partial x'^{\sigma}}{\partial x^{\nu}} = \delta^{\mu}_{\nu}$$

La ecuación (4.18) también puede escribirse como

$$\sqrt{g'} = J\sqrt{g}$$

Cualquier tensor de densidad con peso W puede ser expresado como un tensor ordinario con un factor $g^{-\frac{W}{2}}$. Por ejemplo, un tensor de densidad Ψ'^{μ}_{ν} de peso W tiene la regla de transformación

$$\Psi'^{\rho}_{\sigma} = \left|\frac{\partial x'}{\partial x}\right|^{W} \frac{\partial x^{\mu}}{\partial x'^{\sigma}}\frac{\partial x'^{\rho}}{\partial x^{\nu}}\Psi^{\mu}_{\nu}$$

y usando (4.18) vemos que

$$g'^{\frac{W}{2}}\Psi'^{\rho}_{\sigma} = g^{\frac{W}{2}}\frac{\partial x^{\mu}}{\partial x'^{\sigma}}\frac{\partial x'^{\rho}}{\partial x^{\nu}}\Psi^{\mu}_{\nu} \tag{4.19}$$

La importancia del tensor de densidad se ve reflejada en el teorema fundamental del cálculo integral, [21] ó [32], que bajo una transformación de coordenadas $x \to x'$, el elemento de volumen d^4x viene como

$$d'^4x = \left|\frac{\partial x'}{\partial x}\right| d^4x$$

Entonces el producto de d^4x con un tensor de densidad de peso -1 se transforma como un tensor ordinario de acuerdo a (4.19). En particular

$$\sqrt{g}d^4x \tag{4.20}$$

es el **elemento de volumen invariante**.

Suponga un campo escalar S tal que $S = S'$. Entonces

$$\int S'\sqrt{g'}d^4x' = \int S'\sqrt{g}Jd^4x = \int S\sqrt{g}d^4x$$

si la región de integración para las x es la región para las x', entonces

$$\int S\sqrt{g}d^4x = \text{ invariante} \tag{4.21}$$

Este invariante es muy importante ya que a partir de él formularemos un principio variacional con el cual obtendremos las ecuaciones del campo gravitacional.

4.4 TRANSFOMACION DE LORENTZ

La transformación de Galileo nos dan una invariancia en las ecuaciones de la Mecánica clásica o Newtoniana, pero esta no es toda la física, existen fenómenos en los cuales la transformación de Galileo no llevan a la invariancia de las ecuaciones que definen a estos fenómenos¶. Por ejemplo, consideremos una partícula que interacciona con otra, la

¶Por ejemplo, la teoría electromgnética que está regida por las ecuaciones de Maxwell, no son invariantes bajo la transformación de Galileo, es fácil ver esto. Supongamos que tenemos una partícula cargada que se mueve respecto de un sistema de referencia fijo, en este sistema, el movimiento de la carga genera un campo magnético (como se verá más adelante), mientras que en un sistema de referencia que este montado sobre la partícula, ésta permanecerá fija, por lo que no generará ningún campo magnético.

velocidad de propagación de interacción, como ya vimos, es la velocidad de la luz (4.5) que además es la máxima, todo esto en un sistema de referencia. Ahora supongamos que existe otro sistema de referencia que viaja con respecto al anterior a una velocidad v, entonces, según la trasformación de Galileo, la velocidad de propagación de interacción visto desde el sistema en movimiento sería mayor o menor que la velocidad de la luz dependiendo de la dirección del movimiento del segundo sistema de referencia, lo que contradice al segundo postulado del principio de relatividad de Einstein.

Por lo anterior es necesario proponer unas nuevas ecuaciones de transformación que hagan que la velocidad de la luz sea constante y además sea la máxima y que también se puedan reducir a la transformación de Galileo cuando la velocidad del segundo sistema de referencia sea muy pequeña comparada con la velocidad de la luz. No olvidemos que estamos trabajando en un espacio de Minkowski.

Debemos encontrar la forma de una transformación de coordenadas que deje invariante el valor s_{12}. Consideremos un suceso en el origen del espacio-tiempo, entonces la distancia[‖] (al cuadrado) entre el suceso del origen y cualquier otro suceso con coordenadas (ct, x, y, z) es, por definición, ecuación (4.8):

$$s^2 = c^2t^2 - x^2 - y^2 - z^2 \tag{4.22}$$

Queremos encontrar una transformación de un sistema de referencia cartesiano de coordenadas (ct, x, y, z) a otro sistema de coordenadas (ct', x', y', z') tal que

$$s^2 = s'^2$$

donde

$$s'^2 = c^2t'^2 - x'^2 - y'^2 - z'^2$$

En la geometría clásica existe una transformación que deja invariante a la distancia

$$l^2 = x^2 + y^2 + z^2$$

(en coordenadas cartesianas) entre el origen y el punto (x, y, z), esta transformación es una rotación de ejes. En general, una rotación arbitraria se puede descomponer en una combinación de rotaciones en los planos xy, yz, y zx. Por ejemplo, la rotación por un ángulo θ en el plano xy está dada por la transformación

$$\begin{aligned} x \rightarrow x' &= x\cos\theta + y\sin\theta \\ y \rightarrow y' &= -x\sin\theta + y\cos\theta \\ z \rightarrow z' &= z \end{aligned} \tag{4.23}$$

y deja invariante la distancia $x^2 + y^2$, que es igual a $x'^2 + y'^2$.

Del mismo modo, podemos hablar de una "rotación" en el espacio-tiempo; esta será una combinación de rotaciones en los "planos" xy, yz, zx, ctx, cty y ctz. Los tres primeros son rotaciones comunes como (4.23) y no representan nada nuevo. En cambio, el equivalente a una rotación en el "plano" ctx debe ser una transformación de coordenadas deje invariante a

$$s^2 = c^2t^2 - x^2 \tag{4.24}$$

siguiendo la analogía con la transformación (4.23), hagamos

$$\begin{aligned} ct' &= ct\cosh\psi + x\sinh\psi \\ x' &= ct\sinh\psi + x\cosh\psi \\ y' &= y \qquad\qquad z' = z \end{aligned} \tag{4.25}$$

[‖]Generalmente se le denomina seudodistancia.

y es fácil comprobar que la condición (4.24) se satisface, ya que

$$\cosh^2\psi - \sinh^2\psi = 1 \tag{4.26}$$

Identifiquemos ahora a ψ. El origen del sistema K' tiene coordenada $x' = 0$, así de acuerdo con (4.25), el origen K' satisface la condición

$$\tanh\psi = -\frac{x}{ct}$$

Pero, visto desde K, $\frac{x}{t}$ es precisamente la velocidad v del sistema K'. Entonces

$$\tanh\psi = -\frac{v}{c} = -\beta$$

y sabiendo que

$$1 - \tanh^2\psi = \frac{1}{\cosh^2\psi} \tag{4.27}$$

entonces

$$\cosh\psi = \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

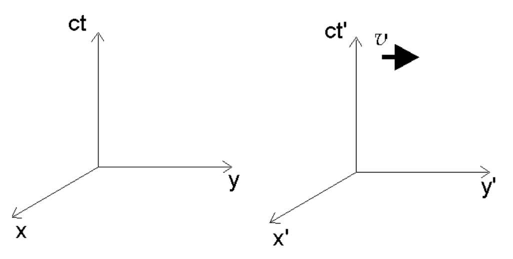
y sustituyendo en la identidad trigonométrica (4.26) obtenemos que

$$\sinh\psi = \pm\beta\gamma$$

Por lo tanto la transformación (4.25) queda como

$$\begin{aligned} ct' &= \gamma ct \pm \beta\gamma x \\ x' &= \pm\beta\gamma ct + \gamma x \\ y' &= y \qquad\qquad z' = z \end{aligned} \tag{4.28}$$

El signo se determina conociendo la dirección del movimiento, es decir si el sistema de referencia inercial en reposo es el sistema no primado, como es el caso, entonces tenemos la siguiente situación



Sistemas de referencia inerciales con $v = cte$.

entonces el signo de (4.28) es negativo, de tal manera que

$$\begin{aligned} ct' &= \gamma ct - \beta\gamma x \\ x' &= -\beta\gamma ct + \gamma x \\ y' &= y \qquad\qquad z' = z \end{aligned} \tag{4.29}$$

Estas ecuaciones fueron obtenidas por primera vez por el físico holandés H.A. Lorentz y por lo tanto son llamadas la **Transformación de Lorentz****. Con ellas

**Una forma alternativa de encontrar las transformaciones (4.28) y en particular (4.29) es la siguiente:

demostró que las ecuaciones básicas del electromagnetismo son invariantes bajo esta transformación, las cuales no lo son bajo la transformación de Galileo, ver [2].

Propondremos la nueva transformación de la forma

$$x' = \gamma(x - vt) \tag{4.30}$$

donde γ es un factor de proporcionalidad que no depende ni de x ni de t, pero que podría depender de v. Hay tres razones principales por las que se escogió de esta manera la nueva transformación:

1. Es lineal en x y x', así que en un evento simple en el sistema de referencia en reposo, equivale a un evento simple en el sistema de referencia inercial.
2. Es simple.
3. Puede reducirse a la transformación de Galileo, la cual es correcta para la mecánica clásica.

Como las ecuaciones de la física deben tener la misma forma para ambos sistemas de referencia, tanto el inercial como el que está en reposo, entonces podemos escribir a x en términos de x' y t' como

$$x = \gamma(x' + vt') \tag{4.31}$$

donde el signo de la velocidad ha cambiado debido, precisamente, al sistema de referencia. El factor γ puede ser el mismo para ambos sistemas de referencia. Como estamos considerando que el sistema inercial sólo se mueve en una dirección, en este caso x, como se consideró para la transformación de Galileo, entonces las coordenadas y y z quedan iguales, por lo tanto

$$\begin{aligned} y &= y' \\ z &= z' \end{aligned} \tag{4.32}$$

Sin embargo las coordenadas temporales t y t' no son iguales como es que esperamos. Esto se puede ver si sustituimos el valor de x' de (4.30) en (4.31), es decir

$$x = \gamma^2(x - vt) + \gamma v t'$$

de donde

$$t' = \gamma t + \left(\frac{1-\gamma^2}{\gamma v}\right) x \tag{4.33}$$

Las ecuaciones (4.30), (4.32) y (4.33) constituyen la transformación de coordenadas que satisfacen el primer postulado del principio de relatividad.

Ahora el segundo postulado, la velocidad de la luz constante, nos permitirá encontrar el valor de γ. Supongamos que al tiempo $t = 0$ los orígenes de ambos sistemas de referencia coinciden y por lo tanto también $t' = 0$. Supongamos que se enciende una flama en el origen común de ambos sistemas cuando $t = t' = 0$ y que los observadores de ambos sistemas pueden medir la velocidad de expansión de la luz de ella. Ambos encuentran la misma velocidad c lo que significa que en el sistema de referencia K, $x = ct$ mientras que en el sistema de referencia K', $x' = ct'$. Si en esta última relación sustituimos las ecuaciones (4.30) y (4.33) tenemos que

$$\gamma(x - vt) = c\left[\gamma t + \left(\frac{1-\gamma^2}{\gamma v}\right) x\right]$$

entonces

$$x = ct\left[\frac{1 + \frac{v}{c}}{1 - \left(\frac{1-\gamma^2}{\gamma^2 v}\right) c}\right]$$

así lo que está entre los paréntesis cuadrados debe ser igual a uno para que se cumpla la relación en el sistema de referencia K, por lo tanto haciendo algunos pasos algebráicos obtenemos que

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{4.34}$$

Por lo tanto la transformación buscada es la siguiente

$$\begin{gathered} t' = \frac{t - \frac{vx}{c^2}}{\sqrt{1 - \frac{v^2}{c^2}}} \qquad x' = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}(x - vt) \\ y' = y \\ z' = z \end{gathered} \tag{4.35}$$

Para encontrar la Transformación Inversa de Lorentz es semejante a la transformación de Lorentz dado que las leyes de la física deben ser semejantes en cualquier sistema de referencia inercial, lo cual es el primer postulado del principio de relatividad, salvo el signo de la velocidad.

Para encontrar la **Transformación Inversa de Lorentz** despejamos las coordenadas no primadas de la transformación de Lorentz (4.29), haciendo esta cálculo obtenemos que

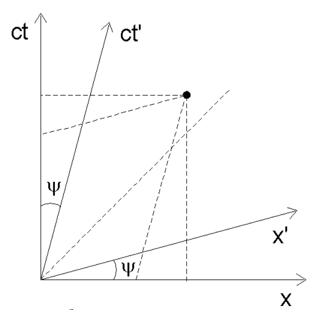
$$\begin{aligned} ct &= \gamma ct' + \beta\gamma x' \\ x &= \beta\gamma ct' + \gamma x' \\ y = y' & \qquad\qquad z = z' \end{aligned} \tag{4.37}$$

que es precisamente lo que esperabamos.

Tanto en las ecuaciones (4.29) como en las ecuaciones (4.37) se ve claramente que cuando $c \to \infty$ se pasa a la mecánica clásica, las ecuaciones de las transformaciones de Lorentz e inversa de Lorentz se convierten en las transformaciones de Galileo (4.1) y su inversa.

Si en las ecuaciones (4.29) y (4.37) $v > c$, las coordenadas x, x', t y t' toman valores imaginarios, cosa que esta de acuerdo con el hecho de que sea imposible la existencia de velocidades mayores que las de la luz.

Geométricamente, la transformación de Lorentz corresponde a "cerrar" los ejes (a diferencia de la rotación común). Sin embargo, hay que tener cuidado con la medición de distancias en un diagrama de espacio-tiempo, ya que la seudodistancia no corresponde a la distancia aparente entre dos sucesos. A diferencia de las rotaciones comunes, la transformación de Lorentz preserva la seudodistancia, no la distancia que se ve en una figura como la siguiente



Una rotación en el espacio-tiempo equivale a cerrar los ejes.

Es evidente de la figura anterior, entre otras cosas, que dos sucesos simultáneos (que ocurren en los tiempos $t^1 = t^2 = t$) en el sistema K no lo son en el sistema K' (ocurren en los tiempos $t^1 \neq t^2$). Así, en la teoría de la relatividad, simultaneidad es un concepto relativo. Este hecho tiene consecuencias muy importantes, como veremos más adelante.

Estas son las ecuaciones de la **Transformación de Lorentz**. La **Transformación inversa de Lorentz**, es fácil verlo, se obtienen intercambiando las cantidades primadas por las no primadas y además cambiando el signo de la velocidad de tal manera que las ecuaciones que la forman son:

$$t = \frac{t' + \frac{vx'}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}} \qquad x = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}(x' + vt') \qquad y = y' \qquad z = z' \tag{4.36}$$

Anteriormente vimos que un punto universo lo definimos por las coordenadas $x^\mu = (x^0, x^1, x^2, x^3)$ donde $x^0 = ct$ y las últimas tres son x, y y z respectivamente. Entonces la transformación de Lorentz (4.29) se puede escribir en forma matricial como

$$\begin{pmatrix} ct' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

o en una forma más compacta

$$x'^\mu = L^\mu_\nu x^\nu$$

donde

$$(L)^\mu_\nu = \frac{\partial x'^\mu}{\partial x^\nu} = \begin{pmatrix} \gamma & -\beta\gamma & 0 & 0 \\ -\beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4.38}$$

es la matriz asociada a la transformación de Lorentz. Es claro que la matriz transpuesta de Lorentz es la misma matriz, es decir

$$L^\mu_\nu = L^\nu_\mu$$

Ahora la matriz asociada a la transformación inversa de Lorentz la denominaremos por L^{-1} y la escribiremos como, ecuación (4.37)

$$\mathsf{L}^{-1} = \begin{pmatrix} \gamma & \beta\gamma & 0 & 0 \\ \beta\gamma & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \eta^\sigma_\mu L^\rho_\sigma \eta^\nu_\rho \tag{4.39}$$

Notemos que

$$\frac{\partial x'^\mu}{\partial x^\nu} = L^\mu_\nu, \qquad \frac{\partial x^\nu}{\partial x'^\mu} = \left(\mathsf{L}^{-1}\right)^\nu_\mu = \left(\boldsymbol{\eta}\mathsf{L}^T\boldsymbol{\eta}\right)^\nu_\mu = \eta^\sigma_\mu L^\rho_\sigma \eta^\nu_\rho \equiv L^\nu_\mu$$

donde $\eta_{\mu\nu} = \eta^\mu_\nu$ es la métrica de nuestro espacio como se definió en (4.11). Así, en nuestra notación tensorial podemos ver que

$$L^\mu_\rho L^\sigma_\mu = \delta^\sigma_\rho \tag{4.40}$$

Si realizamos el producto de $\mathsf{L}\mathsf{L}^{-1}$ obtenemos que

$$\mathsf{L}\left(\boldsymbol{\eta}\mathsf{L}^T\boldsymbol{\eta}\right) = \mathsf{I} = \mathsf{L}\mathsf{L}^{-1} \tag{4.41}$$

El determinante de la matriz L esta dado de la siguiente manera

$$1 = \det(\mathsf{I}) = \det\left(\mathsf{L}\left(\boldsymbol{\eta}\mathsf{L}^T\boldsymbol{\eta}\right)\right) = \det{}^2(\mathsf{L}) \tag{4.42}$$

por lo que

$$\det(\mathsf{L}) = \pm 1 \tag{4.43}$$

Las transformaciones de Lorentz tienen una propiedad importante y es que forma un grupo, llamado el grupo de Lorentz, que por supuesto no estudiaremos aquí, por lo que a los interesados los remitiremos, por ejemplo, a Joseph C. Várilly, **Teoría de Grupos en Cuantización**, CINVESTAV, 1992.

4.4.1 TRANSFORMACION DE VELOCIDADES

Vamos a obtener una regla para comparar las velocidades tanto en un sistema K y un sistema K' que se mueve a una velocidad constante respecto a K.

La velocidad de un punto en K' medido desde el origen del sistema K tiene componentes

$$V_x = \frac{dx}{dt}, \qquad V_y = \frac{dy}{dt}, \qquad V_z = \frac{dz}{dt}$$

donde x, y, z y t vienen dadas por la transformación inversa de Lorentz (4.36).

Similarmente, las componentes de la velocidad del punto medida desde el origen de K' son

$$V'_x = \frac{dx'}{dt'}, \qquad V'_y = \frac{dy'}{dt'}, \qquad V'_z = \frac{dz'}{dt'}$$

las cuales, a partir de la trasformación de Lorentz (4.35) obtenemos que

$$V'_x = \frac{dx'}{dt'} = \frac{V_x - v}{1 - \dfrac{vV_x}{c^2}} \tag{4.44}$$

De manera análoga llegamos a que

$$V'_y = \frac{dy'}{dt'} = \frac{V_y}{\gamma\left(1 - \dfrac{vV_x}{c^2}\right)} \qquad V'_z = \frac{dz'}{dt'} = \frac{V_z}{\gamma\left(1 - \dfrac{vV_x}{c^2}\right)} \tag{4.45}$$

donde v es la velocidad de un sistema de referencia con respecto del otro que está en reposo.

Las ecuaciones (4.44) y (4.45) dan la ley de transformación de Lorentz para las velocidades. Nuevamente se reduce a las ecuaciones de ley de velocidades para la transformación de Galileo (dos primeras ecuaciones 4.2) cuando $v \ll c$. De acuerdo con esta transformación de velocidades, c es la máxima.

4.5 CONSECUENCIAS DE LA TRANSFORMACION DE LORENTZ

4.5.1 CONTRACCION DE LA LONGITUD

La longitud de un objeto puede definirse como la distancia entre sus extremos. Sin embargo, si el objeto cuya longitud se mide, se encuentra en movimiento relativo respecto a un observador, las posiciones de sus extremos deben ser medidas simultáneamente. Consideremos una barra en reposo relativo a K' y paralela al eje X'. Designando sus extremos por a y b, su longitud medida desde K' es $L' = x'_a - x'_b$. La simultaneidad no es necesaria para K' debido a que él ve la barra en reposo. Sin embargo, el observador en K, quien ve la barra en movimiento, debe medir las coordenadas x_a y x_b de los extremos al mismo tiempo t, obteniendo $L = x_a - x_b$. Aplicando la transformación de Lorentz tenemos que

$$x'_a = \gamma\,(x_a - \beta ct)$$
$$x'_b = \gamma\,(x_b - \beta ct)$$

Ahora restando la primera relación de la segunda tenemos que

$$L' = x'_a - x'_b = \gamma\,(x_a - x_b)$$

o lo que es lo mismo

$$L' = \gamma L$$

Escribiendo ahora la longitud en el sistema K (en reposo) en términos del sistema K' (en movimiento relativo) tenemos que

$$L = \sqrt{1 - \frac{v^2}{c^2}} L' \tag{4.46}$$

Puesto que el factor $\sqrt{1 - \frac{v^2}{c^2}}$ es menor que uno, tenemos una situación en la cual L es menor que L'; esto es, el observador del sistema en reposo, quien ve el objeto en movimiento, mide una longitud menor que el observador en K', quien ve el objeto en reposo.

4.5.2 DILATACION DEL TIEMPO

Un intervalo de tiempo puede definirse como el tiempo que transcurre entre dos sucesos, medido por un observador. Consideremos dos sucesos que ocurren en el mismo lugar x' con respecto a un observador en K'. El intervalo entre dos sucesos es $T' = t'_a - t'_b$. Para un observador en K con respecto a quien K' se está moviendo con velocidad constante v en la dirección positiva de la X, el intervalo es $T = t_a - t_b$. Para encontrar la relación entre los tiempos en los cuales ocurren los dos sucesos, registrados por ambos observadores, usamos nuevamente la transformación de Lorentz para el tiempo y entonces

$$\begin{aligned} t_a &= \gamma\left(ct'_a + \beta x'\right) \\ t_b &= \gamma\left(ct'_b + \beta x'\right) \end{aligned}$$

Ahora restando t'_a de t'_b tenemos que

$$t_a - t_b = \gamma\left(t'_a - t'_b\right)$$

o lo que es lo mismo

$$T = \gamma T' \tag{4.47}$$

Ahora T' es el intervalo de tiempo medido por un observador K', en reposo con respecto al punto en el cual tienen lugar los sucesos, y T es el intervalo de tiempo medido por un observador en K, relativo al cual el punto está en movimiento cuando los sucesos ocurren. Esto es, el observador en K ve que los sucesos ocurren en dos posiciones distintas del espacio. Puesto que el factor γ es menor que uno, la ecuación (4.47) indica que T es mayor que T'. Por consiguiente los procesos parecen tomar más tiempo cuando ocurren en un cuerpo en movimiento relativo a un observador que cuando el cuerpo está en reposo relativo al observador.

4.6 DINAMICA RELATIVISTA

Consideremos dos sucesos que, en un sistema K, ocurren en el mismo lugar, pero en tiempos t_1 y t_2 distintos. La distancia (4.8) entre estos dos sucesos será

$$s_{12}^2 = c^2\left(t_2 - t_1\right)^2$$

O sea, $\frac{\sqrt{s_{12}^2}}{c}$ es el tiempo medido por un reloj que está fijo en K. Si ahora definimos el **tiempo propio** entre dos sucesos como

$$\tau_{12} = \frac{\sqrt{s_{12}^2}}{c} \tag{4.48}$$

o en forma diferencial

$$d\tau = \frac{\sqrt{ds^2}}{c}$$

vemos que el tiempo así definido es invariante a una transformación de Lorentz (o simplemente invariante de Lorentz) y corresponde al tiempo medido dt en el sistema K en el que el reloj está parado.

Ahora, una partícula que se mueve de algún modo describe en el espacio de Minkowski, una línea de universo. Esta línea, como toda curva, puede describirse por medio de ecuaciones paramétricas

$$x^\mu = x^\mu(\tau)$$

donde τ es un parámetro que varía de manera continua.

Es conveniente escoger como parámetro al tiempo propio a lo largo de la línea de universo, así

$$d\tau = \frac{1}{c}\sqrt{ds^2} = \frac{1}{c}\sqrt{c^2dt^2 - |d\mathsf{x}|^2} = \sqrt{1 - \frac{|\mathsf{v}|^2}{c^2}}dt$$

que es

$$d\tau = \gamma^{-1}dt \tag{4.49}$$

Definamos ahora

$$u^\mu = \frac{dx^\mu}{cd\tau}$$

u^μ es el vector tangente a la línea de universo y es un cuadrivector, llamado **cuadrivector de velocidad**, en efecto, ya que dx^μ es un cuadrivector y $d\tau$ es un escalar; por lo tanto, frente a una transformación de Lorentz, u^μ se transforma como un tensor contravariante.

La magnitud de u^μ es

$$\eta_{\mu\nu}u^\mu u^\nu = \frac{1}{c^2}\eta_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} = \frac{ds^2}{c^2d\tau^2} = 1 \tag{4.50}$$

así que u^μ es un tensor o cuadrivector unitario.

Veamos cual es la relación entre los componentes de u^μ y la velocidad v. Tenemos que

$$u^0 = \frac{dx^0}{cd\tau} = \tfrac{dt}{d\tau} = \gamma$$
$$u^1 = \frac{dx^1}{cd\tau} = \tfrac{dx}{c\gamma^{-1}dt} = \gamma\frac{V_x}{c}$$

de donde

$$u^\mu = \left(\gamma, \gamma\frac{\mathsf{v}}{c}\right) \tag{4.51}$$

o en sus componentes covariantes

$$\eta_{\mu\nu}u^\mu = u_\nu = \left(\gamma, -\gamma\frac{\mathsf{v}}{c}\right) \tag{4.52}$$

Se puede comprobar a partir de (4.51) y (4.52) que

$$u^\mu u_\mu = 1$$

Consideremos el caso de una partícula libre, es decir, una partícula que se mueve sin influencia de las fuerzas externas. Este caso es casi trivial en la mecánica clásica. La lagrangiana de la partícula libre es como en la ecuación (2.1)

$$T = \frac{1}{2}mv^2$$

donde m es la masa de la partícula y v es su velocidad. Su acción correspondiente de acuerdo a (3.4) es

$$I = \int_{t_1}^{t_2} L \, dt$$

Sin embargo, para el caso relativista esta definición ya no es tan trivial dado que ahora no podemos tomar al tiempo como el invariante sobre el cual se va a integrar ya que, como hemos visto, este cambia bajo una transformación de Lorentz, por lo que tomaremos como el parámetro de integración al tiempo propio ya que este si es un invariante, como se puede ver en (4.48).

Lo anterior implica que la acción debe tener la forma

$$I = \alpha \int_{t_1}^{t_2} d\tau \tag{4.53}$$

donde α es cierta constante. Esta forma de la acción garantiza que sea la misma en cualquier sistema de coordenadas.

Si recordamos de (4.49)

$$d\tau = \sqrt{1 - \frac{v^2}{c^2}} dt$$

tenemos que la lagrangiana relativista será

$$L = \alpha \sqrt{1 - \frac{v^2}{c^2}}$$

que en el límite no relativista se reduce, haciendo el desarrollo en serie de $\sqrt{1-x}$, a

$$L \approx \alpha \left(1 - \frac{v^2}{2c^2}\right)$$

despreciando lo términos de orden mayor que dos en $\frac{v}{c}$.

Si ahora ponemos

$$\alpha = -mc^2$$

entonces

$$L \approx -mc^2 + \frac{1}{2} m v^2$$

que es exactamente la lagrangiana clásica, excepto por el término constante $-mc^2$. Pero esta constante es irrelevante, pues esta no entra en las ecuaciones de movimiento de Euler-Lagrange (3.10), justamente por ser constante.

La lagrangiana

$$L = -mc^2 \gamma^{-1} \tag{4.54}$$

es la única que garantiza que que la acción sea invariante de Lorentz. A partir de esta lagrangiana relativista, podemos calcular el momento generalizado, ecuación (3.46) de la partícula, el cual es

$$p_i = \frac{\partial L}{\partial v_i} = \gamma m v_i \tag{4.55}$$

o sea, el momento relativista es

$$\mathsf{p} = m\gamma \mathsf{v} \tag{4.56}$$

que difiere del clásico por el factor γ.

Del mismo modo, el hamiltoniano está dado como en (3.50) y en el caso relativista resulta ser

$$H = p \cdot v - L = \sqrt{m^2c^4 + p^2c^2}$$

que es igual a la energía. Esta energía puede escribirse en función de la velocidad, utilizando la ecuación (4.55), y resulta ser

$$E = mc^2\gamma \tag{4.57}$$

que es una cantidad conservada.

La ecuación (4.57) implica que, incluso si la partícula esta en reposo, posee una energía

$$E_0 = mc^2 \tag{4.58}$$

que es la famosa ecuación de Einstein que relaciona la masa con la energía.

De acuerdo con la ecuación (4.57), $m\gamma$ puede interpretarse como la masa en reposo por un factor γ. También se dice que la masa de un cuerpo aumenta cuando se mueve, así

$$m = m_0\gamma \tag{4.59}$$

Volvamos nuevamente a la ecuación (4.56), p es un vector en el espacio tridimensional, pero en la teoría de la relatividad es más conveniente trabajar con cuadrivectores o tensores. Si definimos el cuadrivector momento como

$$p^\mu = mcu^\mu \tag{4.60}$$

donde u^μ es el cuadrivector de velocidad, tenemos, según la ecuación (4.56) que

$$p^\mu = (mc\gamma, m\gamma\mathsf{v})$$

Las componentes espaciales de este cuadrivector son, precisamente, las componentes del momento p, según la ecuación (4.56). Además, resulta que la componente temporal es la masa en movimiento multiplicada por $c\gamma$, o sea, según la ec. de Einstein (4.57), la energía total deividida por c. Tenemos así que

$$p^\mu = (\frac{E}{c}, \mathsf{p})$$

En la teoría de la relatividad, la energía y el momento forman un cuadrivector. Nótese que la magnitud de p^μ es

$$p^\mu p_\mu = m^2c^2$$

que es un escalar.

De la ecuación (4.57) se ve que la energía necesaria para que una partícula alcance la velocidad de la luz es infinita. Esto implica que, según la teoría de la relatividad, ninguna partícula puede viajar a la velocidad de la luz. La única excepción la constituyen las patículas cuya masa en reposo es nula: el fotón y, posiblemente, el neutrino. La velocidad de tales partículas es siempre c.

Chapter 5
FORMULACION COVARIANTE DE LA ELECTRODINAMICA

En el presente capítulo se pretende llegar a una formulación covariante de estas ecuaciones de Maxwell, es decir que sean invariantes bajo transformación de Lorentz.

Para mayor profundidad de la electrodinámica se pueden revisar [23], [13].

5.1 CAMPO ELECTRICO Y POTENCIAL ELECTROSTATICO

Las ecuaciones de Maxwell no son ecuaciones que hayan salido totalmente de la teoría sino más bien éstas han sido probadas y deducidas de experimentos realizados por varios científicos del siglo pasado tales como Michael Faraday (1791-1867), Charles Coulomb, Karl F.Gauss (1777-1855), Andrè Marie Ampère (1775-1836), etc. Empezaremos introduciendo el concepto de atracción entre dos cargas eléctricas.

Coulomb determinó experimentalmente que la fuerza de atracción o repulsión entre dos cargas está dada por

$$\mathsf{F}_E = \frac{q_1 q_2}{4\pi\epsilon}\frac{\mathsf{r}}{r^3} \tag{5.1}$$

donde $\mathsf{r} = \mathsf{r}_1 - \mathsf{r}_2$ y $r = |\mathsf{r}|$ y r_1 es el radio vector de la carga q_1 y r_2 es el radio vector de la carga q_2 en un sistema de referencia cartesiano. Donde la constante ϵ se conoce como **permitividad eléctrica** del medio. Para el caso especial del vacío, la permitividad del espacio libre está dada por

$$\epsilon_o = 8.8542 \times 10^{-12} C^2 N^{-1} m^{-2}$$

La permitividad de un material se puede expresar en términos de ϵ_o como $\epsilon = K_e \epsilon_o$ donde K_e,constante dieléctrica, es una cantidad adimensional y es la misma para todos los sistemas de unidades.

Sabemos, por experimentos, que cargas situadas en el vacío, experimentan cierta interacción mutua. Una posible explicación para esta interacción podría ser que cada carga emite o absorbe un flujo de partículas indetectables. Otra posible explicación, y la más aceptada, es que cada carga está rodeada de algo llamado **campo eléctrico**. Entonces supongamos que cada carga interacciona directamente con el campo eléctrico en el que está sumergido. Si una carga q experimenta una fuerza F_E, el campo eléctrico en la posición de la carga está definido por

$$\mathsf{F}_E = q\mathsf{E} \tag{5.2}$$

Es fácil ver que el campo eléctrico debido a una carga a una distancia r de ella puede verse como el gradiente de cierta función escalar φ evaluada en dicho punto r, ver [23], sección 2.3, es decir,

$$\mathsf{E}(\mathsf{r}) = -\boldsymbol{\nabla}\varphi \tag{5.3}$$

φ es llamado generalmente **potencial escalar electrostático**. En general el potencial electrostático debido a una carga puntual q_1 es

$$\varphi(\mathsf{r}) = \frac{q_1}{4\pi\epsilon}\frac{1}{\mathsf{r}} \tag{5.4}$$

5.2 LEY DE GAUSS

De acuerdo a las ecuaciones (5.1) y (5.2) tenemos que el campo eléctrico debido a la carga q_1 que experimenta la carga q_2 es

$$\mathsf{E} = \frac{\mathsf{F}_E}{q_2} = \frac{q_1}{4\pi\epsilon}\frac{\mathsf{r}}{r^3} \tag{5.5}$$

Ahora tomemos la integral de la componente normal del campo eléctrico sobre una superficie cerrada (Fig. 1) que encierra tanto al origen como a la carga q_1

$$\oint \mathsf{E}\cdot\mathsf{n}\,dS = \frac{q_1}{4\pi\epsilon}\oint \frac{\mathsf{r}}{r^3}\cdot\mathsf{n}\,dS \tag{5.6}$$

Veamos que valor tiene la integral del lado derecho de la ecuación anterior. Supongamos que siempre es posible encontrar una superficie esférica de radio fijo r que contiene a la superficie cerrada que hemos considerado hasta ahora. La integral sobre toda la superficie de la esfera de la proyección de dS en un plano perpendicular a r, $\left(\frac{\mathsf{r}}{r}\cdot\mathsf{n}\,dS\right)$, es igual al área de la esfera, es decir

$$\oint \frac{\mathsf{r}}{r}\cdot\mathsf{n}\,dS = 4\pi r^2$$

entonces si nosotros dividimos por r^2 ambos lados y dado que consideramos que la esfera tiene un radio fijo, este puede entrar a la integral de tal manera que

$$\oint \frac{\mathsf{r}}{r^3}\cdot\mathsf{n}\,dS = 4\pi \tag{5.7}$$

El integrando del lado izquierdo de la ecuación anterior es llamado comúnmente el **ángulo sólido** y es denotado por $d\Omega$. Entonces combinando las ecuaciones (5.6) y (5.7) obtenemos que

$$\oint \mathsf{E}\cdot\mathsf{n}\,dS = \frac{q_1}{\epsilon} \tag{5.8}$$

Supongamos ahora que no sólo se encuentra q_1 en el interior de la superficie cerrada sino que hay un número finito de cargas, entonces (5.8) la podemos ver como

$$\oint \mathsf{E}\cdot\mathsf{n}\,dS = \frac{1}{\epsilon}\sum_{i=1}^{N} q_i$$

También podemos generalizar esta relación para una distribución de cargas continua caracterizada por una densidad de carga, entonces (5.8) queda definida como

$$\Phi_E = \oint \mathsf{E}\cdot\,d\mathsf{S} = \frac{1}{\epsilon}\int_V \rho\,dV \tag{5.9}$$

donde ρ depende de la posición dentro de la superficie cerrada.

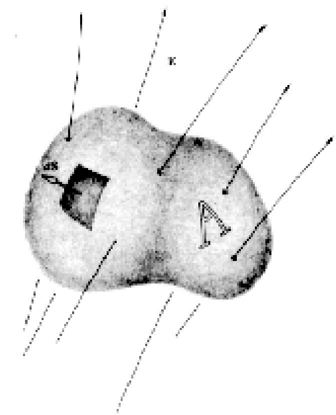


Figure 1 Campo E a través del área cerrada A.

A Φ_E se le conoce como **flujo eléctrico**. La ecuación (5.9) se conoce como la **Ley de Gauss para el campo eléctrico**.

Una interpretación para el flujo eléctrico es que de cierta manera mide la diferencia entre las líneas de campo eléctrico que entran a la superficie y las que salen ella, como se muestra en la figura 1. Como sabemos las cargas pueden ser consideradas como "fuentes" o "sumideros" de líneas de campo y la parte derecha de la ecuación (5.9) sería que tantas líneas de campo se perdieron o se crearon dependiendo de la distribución de carga continua.

5.3 FUERZA DE LORENTZ

La ecuación (5.1) nos da la fuerza de interacción entre dos cargas fijas, ahora si una carga se mueve respecto a la otra con una velocidad v, ésta experimenta una fuerza F_M la cual es proporcional a esta velocidad. Entonces definimos a la llamada **Inducción Magnética** B, como

$$\mathsf{F}_M = q\mathsf{v} \times \mathsf{B}$$

generalmente a B se le conoce también como el **campo magnético**. Si ambas fuerzas ocurren simultáneamente, entonces se dice que la carga se encuentra en un región donde se encuentran ambos campos y la carga experimenta una fuerza

$$\mathsf{F} = q\mathsf{E}+q\mathsf{v} \times \mathsf{B} \tag{5.10}$$

llamada **Fuerza de Lorentz**.

5.4 LEY DE GAUSS (MAGNETICA)

Hasta ahora, experimentalmente no se ha encontrado una contraparte magnética de la carga eléctrica, es decir, no se conocen de manera aislada los monopolos magnéticos, estos monopolos (polo norte y polo sur) sólo se han observado juntos para formar los imanes.

A diferencia del campo eléctrico, la inducción magnética B no diverge o converge hacia algún tipo de carga magnética. Estos campos magnéticos se pueden describir en función de una distribución de corriente, por tanto podemos considerar un magneto elemental como si fuera una corriente circular donde las líneas de B son continuas y cerradas. Cualquier superficie cerrada en una región de campo magnético podría tener entonces un número igual de líneas de campo entrando y saliendo. Esto se produce debido a la falta

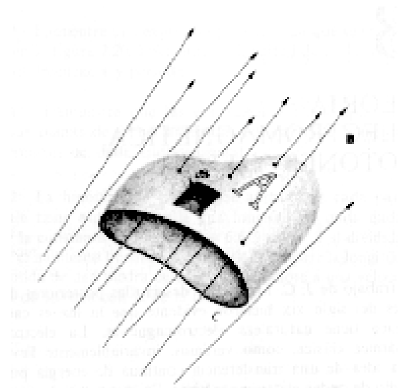


Figure 2 Campo B a través de un área abierta A.

de monopolos dentro del volumen encerrado por la superficie. Entonces el **flujo de inducción magnética**, Φ_B, a través de dicha superficie es cero, entonces el equivalente magnético de la ley de Gauss es

$$\Phi_B = \oint \mathsf{B} \cdot d\mathsf{S} = 0 \tag{5.11}$$

5.5 LEY DE INDUCCION DE FARADAY

Hasta aquí hemos encontrados dos ecuaciones que involucran al campo eléctrico y a la inducción magnética por separado, (5.9) y (5.11). Ahora trataremos de encontrar expresiones que nos indiquen como llegar a expresar un campo en términos del otro, es decir, ecuaciones de interacción entre el campo eléctrico y la inducción magnética.

La primera de estas ecuaciones de interacción fue encontrada experimentalmente por Michael Faraday, esta nos dice que el flujo magnético variable en el tiempo, pasando a través de un circuito conductor cerrado, da como resultado la generación de una corriente alrededor del circuito. El flujo de la inducción magnética (o **densidad de flujo magnético**) Φ_B a través de una área abierta A limitada por el circuito conductor, ver figura 2, esta dado por

$$\Phi_B = \int_A \mathsf{B} \cdot d\mathsf{S}. \tag{5.12}$$

Si el flujo a través de un circuito cerrado C es constante, de modo que $\frac{d\Phi_B}{dt} = 0$, entonces puede observarse que no existe corriente en el circuito. Sin embargo, Faraday encontró que si el flujo no es constante a través de C, de modo que $\frac{d\Phi_B}{dt} \neq 0$, entonces existe una corriente provocada en C. Se dice que esta corriente fue "inducida" por el cambio de flujo, por lo que el valor numérico de la corriente depende de la resistencia del circuito, por lo que es más conveniente expresar el resultado cuantitativo del experimento en función de la **fuerza electromotríz** inducida, es decir, el trabajo realizado por unidad de carga. De ahí que fuerza electromotríz inducida o $\mathcal{E}_{ind}$ producida alrededor del circuito es

$$\mathcal{E}_{ind} = -\frac{d\Phi_B}{dt} \tag{5.13}$$

Puesto que nuestro interés son los campos eléctrico y magnético y dado que la $\mathcal{E}_{ind}$ existe solamente como resultado de la presencia de un campo eléctrico inducido no

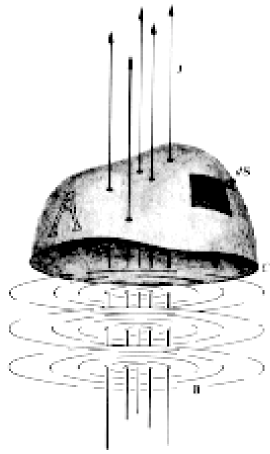


Figure 3 Densidad de corriente a través de un área abierta A.

conservativo a lo largo del alambre,

$$\mathcal{E}_{ind} = \oint_C \mathsf{E}_{ind} \cdot d\mathsf{l} \tag{5.14}$$

alrededor de la curva C, entonces si relacionamos las ecuaciones (5.12), (5.13) y (5.14) obtenemos que

$$\oint_C \mathsf{E}_{ind} \cdot d\mathsf{l} = -\frac{d}{dt}\int_A \mathsf{B} \cdot d\mathsf{S}$$

Decimos entonces que la fuerza electromotriz inducida es proporcional a la derivada total del flujo respecto al tiempo. Podemos variar el flujo variando la inducción magnética o variando la forma, orientación o posición del circuito. Diremos en este caso que lo que varía es la inducción magnética y por lo tanto

$$\oint_C \mathsf{E}_{ind} \cdot d\mathsf{l} = -\int_A \frac{\partial \mathsf{B}}{\partial t} \cdot d\mathsf{S} \tag{5.15}$$

Esta ecuación nos indica que el campo magnético variable en el tiempo tendrá un campo eléctrico inducido asociado con él.

5.6 LEY CIRCUITAL DE AMPERE

Otra de las ecuaciones de interacción entre los campos eléctrico y magnético es la que, experimentalmente, encontró Andrè Marie Ampère (1775-1836). Se conoce como ley circuital y relaciona una integral de línea de B tangente a una curva cerrada C, con la corriente total i que pasa dentro de los confines de C

$$\oint_C \mathsf{B} \cdot d\mathsf{l} = \mu \int_A \mathsf{J} \cdot d\mathsf{S} = \mu\, i \tag{5.16}$$

donde la superficie A es abierta y está limitada por C y J es la corriente por unidad de área, como se muestra en la figura 3.

A la cantidad μ se le llama **permeabilidad del medio**. Para el caso del vacío $\mu = \mu_o$ que se define como

$$\mu_o = 4\pi \times 10^{-7} N s^2 C^{-2}$$

como en el caso de la permitividad, $\mu = K_m \mu_o$ donde K_m es la permeabilidad relativa adimensional.

Las cargas móviles no son la única fuente de campo magnético. Esto se evidencia por el hecho de que mientras se está cargando o descargando un capacitor, uno puede medir un campo B en la región entre sus placas. Este campo es indistinguible del que rodea los alambres, aún cuando ninguna corriente en realidad atraviesa el capacitor. Observamos, sin embargo, que si A es el área de cada placa y Q la carga en ella, el campo eléctrico se expresa como *

$$E = \frac{Q}{\epsilon A}$$

Cuando la carga varía, el campo eléctrico cambia y

$$\epsilon \frac{\partial \mathsf{E}}{\partial t} = \frac{i}{A}$$

lo cual es efectivamente una densidad de corriente. J.C. Maxwell supuso la existencia de tal mecanismo al que llamó **densidad de corriente de desplazamiento**† definida por

$$\mathsf{J}_D = \epsilon \frac{\partial \mathsf{E}}{\partial t} \tag{5.17}$$

La reformulación de la ley de Ampère, queda como

$$\oint_C \mathsf{B} \cdot d\mathsf{l} = \boldsymbol{\mu} \int_A \left(\mathsf{J} + \epsilon \frac{\partial \mathsf{E}}{\partial t} \right) \cdot d\mathsf{S} \tag{5.18}$$

y fue una de las contribuciones más grandes de Maxwell. Aclara que cuando $\mathsf{J} = 0$, un campo E variable en el tiempo estará acompañado por un campo B.

Hasta aquí hemos presentado las ecuaciones de Maxwell en forma integral, ahora con ayuda de los famosos teoremas integrales de la Divergencia y de Stokes veremos su representación diferencial que es la más conocida y de donde partiremos para encontrar la formulación covariante de la Electrodinámica.

5.7 REPRESENTACION DIFERENCIAL DE LAS ECUACIONES DE MAXWELL

Las ecuaciones (5.9) , (5.11), (5.15) y (5.18) pueden escribirse en representación diferencial si a estas últimas ecuaciones se les aplican dos teoremas integrales del cálculo vectorial, estos son el teorema de Stokes y el teorema de la divergencia o de Gauss. A continuación los mencionaremos, su demostración puede verse en libros como [21].

TEOREMA DE LA DIVERGENCIA O DE GAUSS

Sea Ω una región del espacio. Denotemos por $\partial\Omega$ la superficie cerrada orientada que acota Ω. Sea F un campo vectorial suave definido en Ω. Entonces

$$\int_\Omega \boldsymbol{\nabla} \cdot \mathsf{F} = \int_{\partial\Omega} \mathsf{F} \tag{5.19}$$

*Esta expresión puede ser revisada en [23], [13], etc.

†Se puede ver en A.M. Bork, Am. J. Phys. **31**, 854 (1963).

o alternamente

$$\int_{\Omega} div\mathsf{F}\; dV = \int_{\partial\Omega} (\mathsf{F}\cdot\mathsf{n})\; dS \tag{5.20}$$

■

TEOREMA DE STOKES

Sea S una superficie orientada definida por una función de C^2, $z = f(x,y)$, $(x,y) \in D$ donde D es el área proyectada por la superficie en el plano xy, y sea F un campo vectorial C^1 en S. Entonces si ∂S denota la curva frontera orientada de S, tal que la orientación obedece a la regla de la mano derecha, tenemos

$$\int_{S} \nabla\times\mathsf{F}\; dS = \int_{\partial S} \mathsf{F}\cdot d\mathsf{l} \tag{5.21}$$

Recordemos que $\int_{\partial\Omega} \mathsf{F}\cdot d\mathsf{l}$ es la integral a lo largo de ∂S de la componente tangencial de F, mientras que $\int_{S} \mathsf{G}$ es la integral sobre S de $\mathsf{G}\cdot\mathsf{n}$, la componente normal de G. ■

Consideremos primeramente la ley de Gauss eléctrica, ecuación (5.9) y al miembro izquierdo de esta ecuación le aplicamos el teorema de la divergencia, ecuación (5.20), de tal manera que

$$\int_{V} div\mathsf{E}\; dV = \frac{1}{\epsilon}\int_{V} \rho\, dV$$

dado que estamos considerando un volumen de integración arbitrario y además el mismo en ambos lados de la ecuación anterior, tenemos que

$$\int_{V} \left(div\mathsf{E} - \frac{1}{\epsilon}\rho \right) dV = 0$$

y así, para que la integral anterior sea cero, lo que está entre paréntesis debe ser cero de tal manera que

$$\nabla\cdot\mathsf{E} = \frac{\rho}{\epsilon} \tag{5.22}$$

Un procedimiento análogo para la ley de Gauss magnética nos da como resultado que

$$\nabla\cdot\mathsf{B} = 0 \tag{5.23}$$

Ahora consideremos la ley de Inducción de Faraday, ecuación (5.15) y al miembro izquierdo de esta ecuación le aplicamos el Teorema de Stokes (5.21) de tal manera que

$$\int_{S} \nabla\times\mathsf{E}\cdot d\mathsf{S} = -\int_{S} \frac{\partial\mathsf{B}}{\partial t}\cdot d\mathsf{S}$$

dado que estamos hablando de la misma superficie de integración y además arbitraria tenemos que

$$\int_{S} \left(\nabla\times\mathsf{E} + \frac{\partial\mathsf{B}}{\partial t} \right)\cdot d\mathsf{S} = 0$$

para que sea cero la integral anterior, los términos entre paréntesis deben ser cero, entonces

$$\nabla\times\mathsf{E} = -\frac{\partial\mathsf{B}}{\partial t} \tag{5.24}$$

Si realizamos la misma operación en el caso de la ley circuital de Ampère obtenemos que

$$\nabla \times \mathsf{B} = \mu \left(\mathsf{J} + \epsilon \frac{\partial \mathsf{E}}{\partial t} \right) \tag{5.25}$$

Así las **ecuaciones de Maxwell** en su representación diferencial estan dadas por las ecuaciones (5.22), (5.23), (5.24) y (5.25).

Otra ecuación importante dentro de la Teoría Electromagnética es la que resulta de tomar la divergencia de la ecuación (5.25) e intercambiar la derivada temporal con el operador nabla en el segundo término del lado derecho de la ecuación resultante y sustituir la ecuación (5.22) y, además observar que el lado izquierdo de la misma ecuación resultante es igual a cero, para obtener que

$$\nabla \cdot \mathsf{J} + \frac{\partial \rho}{\partial t} = 0 \tag{5.26}$$

esta importante ecuación nos da la **conservación de la carga eléctrica** y se llama **ecuación de continuidad**. En general J esta dada por $\mathsf{J} = \rho \mathsf{v}$

5.8 FORMULACION COVARIANTE DE LA ELECTRODINAMICA

Primeramente, es fácil ver que las ecuaciones de Maxwell en su representación diferencial predicen la propagación de ondas electromagnéticas[‡] a través de un medio lineal, con la consideración de que la densidad de carga ρ es cero y que la única densidad de corriente J surge de la respuesta pasiva de un medio óhmico al campo eléctrico de la onda, es decir, J surge de la ley de Ohm , $\mathsf{J} = g\mathsf{E}$ donde g es una constante para el vacío, de aquí surgen dos ecuaciones de onda para los campos, estas son:

$$\nabla^2 \mathsf{B} - \epsilon_0 \mu_0 \frac{\partial^2 \mathsf{B}}{\partial t^2} = 0 \tag{5.27}$$

y

$$\nabla^2 \mathsf{E} - \epsilon_0 \mu_0 \frac{\partial^2 \mathsf{E}}{\partial t^2} = 0 \tag{5.28}$$

donde el producto

$$\epsilon_0 \mu_0 = \frac{1}{c^2} \tag{5.29}$$

donde c es la velocidad de la luz.

Veamos que existen ciertos potenciales que pueden generar tanto a E como a B. Como la inducción magnética tiene divergencia cero ecuación (5.23), se puede representar como

$$\mathsf{B} = \nabla \times \mathsf{A} \tag{5.30}$$

donde A es el **potencial vectorial**, que es arbitrario.

Sustituyendo B de la ecuación anterior en la ley de Faraday (5.24) obtenemos:

$$\nabla \times \mathsf{E} + \frac{\partial}{\partial t} \nabla \times \mathsf{A} = 0$$

suponiendo suficiente continuidad de los campos para intercambiar las derivaciones espacial y temporal, esto puede escribirse como

[‡]Ver [23] cap. 16.

$$\nabla \times (\mathsf{E} + \frac{\partial}{\partial t}\mathsf{A}) = 0$$

Como para todo campo escalar φ ocurre que $\nabla \times (\nabla\varphi) = 0$ entonces

$$\mathsf{E} = -\nabla\varphi - \frac{\partial}{\partial t}\mathsf{A} \tag{5.31}$$

donde φ es un **potencial escalar** arbitrario. Los potenciales escalar y vectorial satisfacen ecuaciones de onda semejantes a las de los campos, ecuaciones (5.27) y (5.28). Consideraremos que estamos en el vacío, por lo tanto vamos a usar la relación (5.29). La ecuación de onda para A se deduce sustituyendo en la ecuación (5.25) las expresiones de B y E dadas en (5.30) y (5.31)

$$\nabla \times (\nabla \times \mathsf{A}) = \mu_o \mathsf{J} + \frac{1}{c^2}\frac{\partial}{\partial t}\left[-\nabla\varphi - \frac{\partial}{\partial t}\mathsf{A}\right]$$

Si la identidad vectorial $\nabla\times(\nabla \times \mathsf{F}) = \nabla\nabla\cdot\mathsf{F} - \nabla^2\mathsf{F}$ para algún campo vectorial F la aplicamos en la ecuación anterior obtenemos que

$$\nabla\left(\nabla \cdot \mathsf{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t}\right) - \nabla^2\mathsf{A} + \frac{1}{c^2}\frac{\partial^2\mathsf{A}}{\partial t^2} = \mu_0\mathsf{J}$$

que será semejante a las ecuaciones de onda para los campos siempre y cuando se satisfaga que

$$\nabla \cdot \mathsf{A} + \frac{1}{c^2}\frac{\partial\varphi}{\partial t} = 0 \tag{5.32}$$

que es llamada la **norma de Lorentz**.

Entonces la ecuación de onda para el potencial vectorial A será

$$\nabla^2\mathsf{A} - \frac{1}{c^2}\frac{\partial^2\mathsf{A}}{\partial t^2} = -\mu_0\mathsf{J} \tag{5.33}$$

Ahora si sustituimos la ecuación (5.31) en (5.22), se tiene

$$-\nabla \cdot \nabla\varphi - \nabla \cdot \frac{\partial}{\partial t}\mathsf{A} = \frac{\rho}{\epsilon_0}$$

si intercambiamos el orden de la divergencia y de la derivada con respecto al tiempo que se aplica a A y utilizando la norma de Lorentz (5.32), se tiene

$$\nabla^2\varphi - \frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} = -\frac{\rho}{\epsilon_0} \tag{5.34}$$

Por lo tanto, imponiendo la norma de Lorentz, tanto el potencial escalar como el vectorial están obligados a satisfacer ecuaciones de onda inhomogéneas de forma semejante.

Las ecuaciones de Maxwell se escriben en función de las derivadas respecto al tiempo y al espacio de los campos E y B. En el sistema tridimensional común, el tiempo interviene en las ecuaciones como un escalar, pero las tres derivadas espaciales entran en ciertas combinaciones simétricas (divergencia y rotacional). La ley de Gauss eléctrica, la podemos escribir como

$$\frac{\partial E^i}{\partial x^i} = \frac{\rho}{\epsilon_o}$$

que es la divergencia y la ecuación del rotacional (es decir, la ley de Ampère) como

$$\frac{\partial B^i}{\partial x^j} - \frac{\partial B^j}{\partial x^i} = \mu_o J^k + \mu_o \epsilon_o \frac{\partial E^k}{\partial t}$$

la cual, en realidad nos representa tres ecuaciones.

Sin embargo, en un capítulo anterior hemos observado que las transformaciones de Lorentz mezcla las coordenadas del espacio y del tiempo y puede considerarse como una rotación en el espacio $x^0x^1x^2x^3$. Por tanto, x^0, x^1, x^2 y x^3 deberán entrar en las ecuaciones de Maxwell en forma simétrica. Debemos, de hecho, poder escribir las ecuaciones de Maxwell en función de divergencias y rotacionales en cuatro dimensiones. Una formulación de las ecuaciones de la teoría electromagnética que trata las coordenadas espaciales y temporal de manera equivalente se llama **formulación covariante de las ecuaciones de Maxwell**. Además, al utilizar las transformaciones de Lorentz con el fin de hacer esta formulación, estamos obligado a que las ecuaciones de Maxwell a ser invariantes de Lorentz y por tanto habremos encontrado ecuaciones que tengan la misma forma en los distintos sistemas de referencia inerciales.

Empezaremos con la ecuación de continuidad (5.26)

$$\boldsymbol{\nabla} \cdot \mathsf{J} + \frac{\partial \rho}{\partial t} = 0$$

Vamos a demostrar ahora que la densidad de carga ρ y la densidad de corriente J forman un cuadrivector. Para ello, consideremos el elemento de volumen dV con una carga dq en su interior. Si ρ es la densidad de carga, entonces

$$dq = \rho dV \tag{5.35}$$

La carga dq es invariante de Lorentz, pero ρ y dV no lo son por separado (recuerde la contracción de Lorentz). Así, tenemos que ρdV es un escalar. Consideremos al cuadrivector $\rho dV dx^\alpha$, entonces

$$\rho c dV dx^\alpha = \rho dV c dt \frac{dx^\alpha}{dt} \tag{5.36}$$

y por otra parte

$$c dV dt = dx dy dz c dt$$

que es el elemento de volumen en el espacio de cuatro diemsiones que es un escalar. Así, resulta que el miembro derecho de la ecuación (5.36) es un cuadrivector, por lo que $\rho \dfrac{dx^\alpha}{dt}$ también es un cuadrivector; sus componentes son

$$(c\rho, \rho \mathsf{v}) = (c\rho, \mathsf{J})$$

y se tiene que

$$J^\mu = (c\rho, \mathsf{J}) \tag{5.37}$$

Por lo que podemos escribir la ecuación de continuidad en forma covariante como

$$\frac{\partial J^\nu}{\partial x^\nu} = 0 \tag{5.38}$$

Ahora, el potencial vectorial A y el potencial escalar φ satisfacen las ecuaciones de onda inhomogéneas (5.33) y (5.34) dadas por

$$\begin{aligned} \boldsymbol{\nabla}^2 \mathsf{A} - \frac{1}{c^2}\frac{\partial^2 \mathsf{A}}{\partial t^2} &= -\mu \mathsf{J} \\ \boldsymbol{\nabla}^2 \varphi - \frac{1}{c^2}\frac{\partial^2 \varphi}{\partial t^2} &= -\frac{\rho}{\epsilon} \end{aligned} \tag{5.39}$$

Como J y ρ son las componentes de un cuadrivector, las ecuaciones (5.39) deben representar las cuatro componentes de una ecuación de cuadrivectores y, A y φ se deben combinar para formar un cuadrivector. Si definimos al **cudripotencial o potencial universal** $\mathcal{A}^\mu$ como

$$\mathcal{A}^\mu = \left(\frac{\varphi}{c}, A^1, A^2, A^3\right) \tag{5.40}$$

entonces la ecuación (5.39) puede escribirse como

$$\mathsf{2}^2\mathcal{A}^\mu = \mu_o J^\mu$$

donde $\mathsf{2}^2 \equiv \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2$ es el operador laplaciano cuadridimensional o el **operador de D'Alambert.** Esta última ecuación tambien puede escribir como

$$\eta^{\mu\nu}\frac{\partial^2\mathcal{A}^\sigma}{\partial x^\mu \partial x^\nu} = \frac{\partial^2\mathcal{A}^\sigma}{\partial x^\nu \partial x^\nu} = \mu_o J^\sigma \tag{5.41}$$

La norma de Lorentz toma la forma

$$\frac{\partial\mathcal{A}^\nu}{\partial x^\nu} = 0 \tag{5.42}$$

Estamos ahora en posición de examinar las componentes del campo electromagnético. Estas pueden obtenerse de las ecuaciones tridimensionales usuales, ecuaciones (5.30) y (5.31)

$$\mathsf{B} = \boldsymbol{\nabla} \times \mathsf{A}$$
$$\mathsf{E} = -\boldsymbol{\nabla}\varphi - \frac{\partial \mathsf{A}}{\partial t}$$

Pero A y $\frac{\varphi}{c}$ forman un cuadrivector, de modo que la última de las ecuaciones anteriores puede escribirse (en sus componentes) como

$$\frac{E_1}{c} = \frac{\partial\mathcal{A}_1}{\partial x^0} - \frac{\partial\mathcal{A}_0}{\partial x^1} \tag{5.43}$$

recordando que $x^0 = ct$, que $\mathcal{A}_\alpha = (\frac{\varphi}{c}, -\mathsf{A})$ ya que estamos en un espacio-tiempo de Minkowski.

Por tanto, B y $\frac{\mathsf{E}}{c}$ juntos forman el rotacional cuadridimensional de $\mathcal{A}$. La operación de rotacional aplicada a un vector produce realmente un tensor antisimétrico. Esto es evidente de la forma de la ecuación (5.43), puesto que es claro que debe generarse una cantidad de dos índices. Definimos el tensor de campo electromagnético $F_{\mu\nu}$ por la expresión

$$F_{\mu\nu} = \frac{\partial\mathcal{A}_\nu}{\partial x^\mu} - \frac{\partial\mathcal{A}_\mu}{\partial x^\nu} \tag{5.44}$$

donde el primer índice indica renglón y el segundo columna.

Aquí

$$\begin{aligned} F_{00} = F_{11} = F_{22} = F_{33} &= 0 \\ -F_{10} = F_{01} &= \tfrac{E_1}{c} \\ -F_{20} = F_{02} &= \tfrac{E_2}{c} \\ -F_{30} = F_{03} &= \tfrac{E_3}{c} \\ -F_{12} = F_{21} &= B_3 \\ -F_{23} = F_{32} &= B_1 \\ -F_{31} = F_{13} &= B_2 \end{aligned} \tag{5.45}$$

o en forma matricial

$$\mathsf{F} = (F)_{\mu\nu} = \begin{pmatrix} 0 & \frac{E_1}{c} & \frac{E_2}{c} & \frac{E_3}{c} \\ -\frac{E_1}{c} & 0 & -B_3 & B_2 \\ -\frac{E_2}{c} & B_3 & 0 & -B_1 \\ -\frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \tag{5.46}$$

También tenemos que las componentes contravariantes del tensor $F_{\mu\nu}$ dadas por

$$F^{\alpha\beta} = \eta^{\alpha\mu}\eta^{\beta\nu}F_{\mu\nu}$$

son explicitamente

$$F^{\alpha\beta} = \begin{pmatrix} 0 & -\frac{E_1}{c} & -\frac{E_2}{c} & -\frac{E_3}{c} \\ \frac{E_1}{c} & 0 & -B_3 & B_2 \\ \frac{E_2}{c} & B_3 & 0 & -B_1 \\ \frac{E_3}{c} & -B_2 & B_1 & 0 \end{pmatrix} \tag{5.47}$$

Ahora, definimos el tensor dual de $F_{\mu\nu}$ de acuerdo a (1.12) como

$$F^*_{\mu\nu} = \frac{1}{2}\epsilon_{\alpha\beta\mu\nu}F^{\alpha\beta} \tag{5.48}$$

y es fácil averiguar que

$$F^*_{\mu\nu} = \begin{pmatrix} 0 & -B_1 & -B_2 & -B_3 \\ B_1 & 0 & -\frac{E_3}{c} & \frac{E_2}{c} \\ B_2 & \frac{E_3}{c} & 0 & -\frac{E_1}{c} \\ B_3 & -\frac{E_2}{c} & \frac{E_1}{c} & 0 \end{pmatrix}$$

Supongamos que tomamos ahora la divergencia del tensor de campo electromagnético, debido a su forma, ecuación (5.44), obtenemos que

$$\frac{\partial F_{\mu\nu}}{\partial x^\nu} = \frac{\partial}{\partial x^\mu}\left(\frac{\partial \mathcal{A}_\nu}{\partial x^\nu}\right) - \frac{\partial^2 \mathcal{A}_\mu}{\partial x^\nu \partial x^\nu}$$

Teniendo en cuenta las ecuaciones (5.41) y (5.42), ésta se convierte en

$$\frac{\partial F_{\mu\nu}}{\partial x^\nu} = -\mu_0 J_\mu \tag{5.49}$$

Esta es una ecuación cuadrivectorial que representa una formulación covariante de dos ecuaciones de Maxwell para el vacío (5.22) y (5.25).

Además es fácil comprobar, a partir de la ecuación (5.44), que se cumple la identidad

$$\frac{\partial F_{\mu\nu}}{\partial x^\sigma} + \frac{\partial F_{\nu\sigma}}{\partial x^\mu} + \frac{\partial F_{\sigma\mu}}{\partial x^\nu} = 0 \tag{5.50}$$

donde μ, ν y σ son todos distintos y representan tres cualesquiera de los subíndices 0,1,2,3. La ecuación (5.50) representa las otras dos ecuaciones de Maxwell que faltaban, esto se hace evidente cuando sustituimos los valores (5.45) en esta misma ecuación.

El primer miembro de la ecuación (5.50) es un tensor de tercer orden, antisimétrico respecto de sus tres índices. Las únicas componentes que no son idénticamente nulas son aquellas para las que μ, ν, y σ son distintas, como ya dijimos. En total hay, por lo tanto cuatro ecuaciones distintas que, como es fácil demostrar sustituyendo (5.44), coinciden con las ecuaciones (5.23) y (5.24). Nótese que la ecuación (5.50) es precisamente la ecuación

$$\frac{\partial F^*_{\mu\nu}}{\partial x^\mu} = 0 \tag{5.51}$$

que muestra explícitamente la existencia de sólo cuatro ecuaciones independientes.

De las ecuaciones (5.49) y (5.51) se ve inmediatamente que las ecuaciones de Maxwell son invariantes de Lorentz, ya que en un nuevo sistema de referencia se tiene que

$$\begin{aligned} &\tfrac{\partial}{\partial x'^{\mu}} = \left(L^{-1}\right)^{\nu}_{\mu} \tfrac{\partial}{\partial x^{\nu}} \\ &F'_{\alpha\beta} = L^{\mu}_{\alpha} L^{\nu}_{\beta} F_{\mu\nu} \\ &J'_{\alpha} = L^{\mu}_{\alpha} J_{\mu} \end{aligned}$$

y las ecuaciones de Maxwell toman la forma

$$\frac{\partial F'_{\mu\nu}}{\partial x'^{\nu}} = -\mu_0 J'_{\mu} \qquad , \qquad \frac{\partial F'^{*}_{\mu\nu}}{\partial x'^{\mu}} = 0$$

que nos dice que, efectivamente, no cambian de forma.

A las ecuaciones, como las de Maxwell, que no cambian de forma frente a transformaciones de Lorentz se les llaman ecuaciones covariantes.

La ecuación (5.10) para la fuerza de Lorentz que actúa sobre una carga q se puede expresar en forma manifiestamente covariante como

$$\frac{dp^{\mu}}{d\tau} = qF^{\mu\nu} u_{\nu} \tag{5.52}$$

Con lo anterior podemos enunciar nuevamente el principio de Einstein para la relatividad especial como: **"En sistemas inerciales, las leyes de la física son covariantes o invariantes frente a la transformación de Lorentz"**, podemos ver que los dos postulados enunciados para este principio de relatividad visto en el capítulo anterior estan incluidos implícitamente es el anterior.

Este principio es fundamental, ya que, en la práctica, es una guía para describir los fenómenos físicos por medio de las ecuaciones matemáticas.

5.9 TENSOR DE ENERGIA MOMENTO ELECTROMAGNETICO

Un **tensor de energía momento**, [28], es un tensor de orden 2 simétrico que satisface la siguiente relación

$$\frac{\partial T^{\mu\nu}}{\partial x^{\nu}} = 0 \tag{5.53}$$

y cuyas componentes se pueden interpretar como

$$\begin{aligned} &T^{00} \rightarrow \text{ densidad de energía} \\ &T^{0i} \rightarrow \text{ densidad de flujo de energía y momento} \\ &T^{ij} \rightarrow \text{ tensor de tensiones} \end{aligned} \tag{5.54}$$

Después de haber visto la definición y significado físico de un tensor de energía momento para un medio continuio, vamos a extender ese mismo concepto al campo electromagnético.

Consideremos primero un campo electromagnético en el vacío (es decir una región donde no haya cargas o corrientes). En vacío las ecuaciones de Maxwell se ven de la siguiente manera

$$\frac{\partial F^{\alpha\beta}}{\partial x^{\alpha}} = 0 \qquad , \qquad \frac{\partial F^{*\alpha\beta}}{\partial x^{\alpha}} = 0 \tag{5.55}$$

La segunda de estas ecuaciones puede escribirse como

$$\frac{\partial F_{\beta\gamma}}{\partial x^{\alpha}} + \frac{\partial F_{\alpha\beta}}{\partial x^{\gamma}} + \frac{\partial F_{\gamma\alpha}}{\partial x^{\beta}} = 0 \tag{5.56}$$

Notar que los índices α, β y γ aparecen en orden cíclico.

Ahora, calculemos la derivada del escalar $F^{\lambda\mu}F_{\lambda\mu}$, así

$$\frac{\partial\left(F^{\lambda\mu}F_{\lambda\mu}\right)}{\partial x^{\beta}} = 2F^{\lambda\mu}\frac{\partial F_{\lambda\mu}}{\partial x^{\beta}}$$

lo cual utilizando la ecuación (5.56), podemos escribir de la forma

$$\frac{\partial}{\partial x^{\beta}}\left(F^{\lambda\mu}F_{\lambda\mu}\right) = -2F^{\lambda\mu}\left(\frac{\partial F_{\beta\lambda}}{\partial x^{\mu}} + \frac{\partial F_{\mu\beta}}{\partial x^{\lambda}}\right) \tag{5.57}$$

Si multiplicamos por $g^{\alpha\beta}$ y contraemos sobre los índices β tenemos que

$$g^{\alpha\beta}\frac{\partial}{\partial x^{\beta}}\left(F^{\lambda\mu}F_{\lambda\mu}\right) = -2F^{\lambda\mu}\left(\frac{\partial F^{\alpha}_{\lambda}}{\partial x^{\mu}} + \frac{\partial F^{\alpha}_{\mu}}{\partial x^{\lambda}}\right) \tag{5.58}$$

Ahora, dado que el tensor de campo electromagnético es antisimétrico, reordenando los índices mudos, los dos términos entre paréntesis de (5.58) son iguales y por tanto, con un cambio de índices mudos $(\mu \to \beta)$, obtenemos que

$$g^{\alpha\beta}\frac{\partial}{\partial x^{\beta}}\left(F^{\lambda\mu}F_{\lambda\mu}\right) + 4F^{\lambda\beta}\frac{\partial F^{\alpha}_{\lambda}}{\partial x^{\beta}} = 0 \tag{5.59}$$

Si definimos el tensor

$$T^{\alpha\beta} = -F^{\alpha}_{\mu}F^{\beta\mu} + \frac{1}{4}g^{\alpha\beta}\left(F^{\lambda\mu}F_{\lambda\mu}\right) \tag{5.60}$$

que es simétrico (basta ver que $F^{\alpha}_{\mu}F^{\beta\mu} = F^{\alpha\mu}F^{\beta}_{\mu}$). La divergencia de este tensor es

$$\frac{\partial T^{\alpha\beta}}{\partial x^{\beta}} = \frac{1}{4}g^{\alpha\beta}\frac{\partial}{\partial x^{\beta}}\left(F^{\lambda\mu}F_{\lambda\mu}\right) - F^{\alpha}_{\mu}\frac{\partial F^{\beta\mu}}{\partial x^{\beta}} - \frac{\partial F^{\alpha}_{\mu}}{\partial x^{\beta}}F^{\beta\mu}$$

El segundo término del lado derecho es nulo por la primera ecuación de Maxwell (5.55), mientras que el primero y tercero son justamente la ecuación (5.59) con un cambio apropiado de índices mudos.

En resumen, como

$$\frac{\partial T^{\alpha\beta}}{\partial x^{\beta}} = 0 \tag{5.61}$$

para $T^{\alpha\beta}$ definido por la ecuación (5.60), entonces $T^{\alpha\beta}$ es el **tensor de energía momento para el campo electromagnético**. De la definición (5.60) se puede ver fácilmente que se cumple con las relaciones (5.54).

Chapter 6
INTRODUCCION A LA TEORIA GENERAL DE LA RELATIVIDAD

A lo largo de este capítulo examinaremos la llamada Teoría de la relatividad General publicada en 1916 por A. Einstein y en la que propone que la materia en el Universo es la que provoca la deformación del espacio, es decir que se habla de una teoría geométrica.

El objetivo de este capítulo es dar las ideas básicas de la relatividad general, considero que este es un buen comienzo para alguien que quiera conocer más acerca de esta teoría para lograr esto, es necesario tener más conocimientos de matemáticas como Geometría o Topología Diferencial*.

Sin embargo, este podría ser un buen comienzo y no por ello no tratar de enterder lo que se expondrá aquí.

6.1 CAMPO GRAVITACIONAL

Siempre ha existido la pregunta de por qué caen los cuerpos hacia la tierra, es decir, al levantar un cuerpo y posteriormente soltarlo, siempre ha de caer al suelo. "...Para Aristóteles, la gravitación era una tendencia de los cuerpos terrestres a ocupar su ≪lugar natural≫, que es el centro de la Tierra. Es evidente que dentro de este esquema, la gravitación es un fenómeno puramente terrestre y que no puede haber atracción gravitacional entre la Tierra y un cuerpo celeste...."[27]. Así es como se interpretaba la gravitación en tiempo de los Griegos, no fue sino hasta Newton cuando la gravitación pudo ser aplicada tanto a los fenómenos terrestres como a los objetos celestes. Este es uno de los temas principales de los Principia Matemática de Newton. Sin embargo, actualmente la física moderna formula la respuesta a la pregunta anterior de un modo diferente.

A raíz de un estudio más detenido de los fenómenos electromagnéticos, se ha llegado a la conclusión de que no existen acciones inmediatas a distancia (como ya discutimos en un capítulo anterior). Cuando un imán atrae, por ejemplo, un trozo de hierro, no cabe contestarse con argüir que el imán actúa directamente sobre el hierro a través del espacio intermedio vacío. Es preciso imaginar que el imán origina siempre en el espacio circundante un "algo" físico real que se denomina campo magnético. Este campo magnético actúa a su vez sobre el trozo de hierro, de modo que este tiende a moverse hacia el imán. No expondremos la justificación de este concepto intermedio, baste indicar que los fenómenos electromagnéticos admiten una interpretación teórica mucho más satisfactoria con el concepto de campo que sin él. De manera análoga consideraremos a la gravitación.

La acción de la Tierra sobre cualquier cuerpo se produce de modo indirecto. La tierra origina en torno suyo un **campo gravitatorio**, este campo actúa sobre el cuerpo y origina su movimiento de caida. De acuerdo con la experiencia, la intensidad de la acción gravitatoria sobre un cuerpo dado disminuye a medida que nos alejamos de la Tierra y esta disminución sigue una ley perfectamente determinada, para ver una discusión más detallada en [27]..

*Un primer texto que podría leerse es [31]. Uno más matemático sería: **Differential Topology**, Brocker, Springer Verlag.

El campo gravitacional exhibe una propiedad sumamente curiosa, que no aperece en el caso del campo eléctrico y magnético, que tiene una importancia fundamental para lo que vamos a exponer posteriormente. Los cuerpos que se mueven bajo la sola acción del campo gravitatorio experimentan una aceleración que no depende lo más mínimo ni del material ni del estado físico del cuerpo. Galileo observó, y con razón, que en ausencia de aire (vacío) un trozo de madera y un trozo de plomo, por ejemplo, caen exactamente del mismo modo al dejarlos en libertad con la misma velocidad o bien sin velocidad alguna. Esta ley, que se cumple con extrema exactitud, se puede formular de otra manera basándonos en la siguiente consideración:

De acuerdo con la ley de Newton $F = m_i a$ donde m_i es la masa inercial que es una caraterística del campo acelerado y además constante. Si la fuerza aceleradora es la gravedad, tenemos, por otro lado que $F = m_g g$ donde la masa gravitacional m_g, es también una constante caraterística del cuerpo y g es la intensidad del campo gravitatorio. De ambas relaciones se sigue que

$$a = \frac{m_g}{m_i} g$$

Para que ahora, tal y como nos lo muestra la experiencia, la aceleración siempre sea la misma, indepedientemente de la naturaleza y estado del cuerpo, para un campo gravitatorio dado, es necesario que la relación entre la masa inercial y la masa gravitacional sea la misma, también para cualquier cuerpo. Entonces $\frac{m_g}{m_i} = 1$, es decir las masas gravitacional e inercial de un cuerpo son iguales.

Esta igualdad entre ambas masas fue mostrada por Galileo, Huygens, Newton, Bessel y Eötvös, ver referencia 7 del capitulo 1 de [32].

6.2 EL PRINCIPIO DE EQUIVALENCIA

6.2.1 MOTIVACION

Supongamos que nos encontramaos en una región del Universo suficientemente alejada de toda masa, es decir fuera de toda influencia de campo gravitacional y en ella colocamos un cajón. En la parte superior del cajón, colocamos un gancho al cual le sujetaremos una cuerda. Una primera situación es que si colocamos a un observador dentro de este cajón con la debida instrumentación, tanto ésta como el propio observador flotarán dentro del cajón debido a la falta de campo gravitacional. Vayamos ahora a una segunda situación independiente de la primera, en ésta, pensemos que alguien que nos es indiferente tira de la cuerda hacia arriba con una aceleración constante, esta aceleración es transmitida al cajón y a su vez , por reacción, es transmitida a todos los instrumentos y al observador mismo que se encuantran dentro de dicho cajón. Entonces el obsevador podrá permanecer erguido dado que experimenta cierta tendencia a caer al piso como si estuviera en la Tierra. Ahora el observador deja caer un cuerpo y observa que tiende a irse al piso, este experimento lo realiza varias veces y con distintos cuerpos y observa que todos caen más o menos con la misma aceleración. Por lo tanto llegará a la conclusión de que él se encuentra en un campo gravitacional más o menos constante. Por un momento quedará desconcertado por el hecho de que el cajón no caiga en el campo gravitatorio; sin embargo, descubre el gancho y la cuerda tensa en la parte superior del cajón, entonces pensará que el cajón cuelga inmóvil en el campo gravitatorio.

Supongamos ahora que el observador dentro del cajón coloca una cuerda en la parte superior del mismo y coloca en el otro extremo de la cuerda un objeto. El observador diría que el objeto suspendido experiementa una fuerza hacia abajo que se ve neutralizada por la tensión de la cuerda, mientras que un observador fuera del cajón dirá que la cuerda transmite el movimiento del cajón al cuerpo, es decir que quién esté tirando del cajón también esta tirando del objeto. Existen, pues, varias razones para extender el principio de relatividad a cuerpos de referencia acelerados uno con respecto del otro, con lo cual hemos obtenido un potente argumento en favor de un postulado generalizado de la relatividad. Este ejemplo, además, demuestra que nuestra ampliación al principio de relatividad confiere un carácter necesario a la ley de la igualdad entre la masa inercial y la gravitacional.

6.2.2 PRINCIPIO DE EQUIVALENCIA.

Supongamos que tenemos un sistema inercial K. Algunas masas que se encuentran sufiencientemente alejadas entre sí y de otras masa, están , por tanto, desprovistas de aceleración alguna respecto a K. Sea también un sistema K' que tiene una aceleración uniforme respecto a K. Repecto a K' todas las masas tienen aceleraciones iguales y parallelas entre sí, es decir, como si con respecto al sistema K, todas las masas se encontraran en un campo gravitacional y K' no tuviera aceleración.

A la suposición de la completa equivalencia física de los sistemas de coordenadas K y K' le llamaremos **Principio de Equivalencia de la Gravitación e Inercia**.

La vulnerabilidad del principio de equivalencia está en el hecho de que requiere un razonamiento que es un círculo vicioso: una masa se mueve sin aceleración si está suficientemente alejada de otros cuerpos; pero sólo sabemos que está suficientemente alejado de otros cuerpos cuando se mueve sin aceleración.

Si consideramos, en general, a todo el Universo, podemos ver que existen regiones en donde podemos suponer sistemas de referencia inerciales. Podemos, por ejemplo, considerar establecido, dentro de un alto grado de presición el principio de inercia para el espacio en el que se halla comprendido el sistema solar siempre que despreciemos los efectos del Sol, los planetas y otros cuerpos presentes en él. Dicho esto con más exactitud, el principio de equivalencia puede enunciarse como:

En cualquier punto del espacio-tiempo, en un campo gravitacional arbitrario, es posible escoger un sistema de coordenadas localmente inercial, tal que dentro de una región suficientemente pequeña del punto en cuestión, las leyes de la naturaleza toman la misma forma como en un sistema de coordenadas cartesianas sin aceleración en ausencia de gravitación.

Lo que nos indica la parte final del enunciado es que en dicha región suficientemente pequeña son válidas las leyes de la teoría de la Relatividad Especial desarrolladas anteriormente. A tales regiones se les llamó en un principio **regiones Galileanas**.

El lector pudo notar cierta coincidencia entre el principio de equivalencia y el axioma que tomó Gauss como la base de la geometría no-euclidiana. El principio de equivalencia dice que en cualquier punto del espacio-tiempo uno puede definir un sistema de coordenadas localmente inercial en el cual se satisfagan las leyes de la relatividad especial. Mientras que Gauss supuso que en cualquier punto en una superficie curva se puede definir un sistema de coordenadas localmente cartesiano en el cual las distancias obedecen el teorema de Pitágoras [32].

Con esta profunda analogía, esperamos que las leyes de la gravitación tengan una fuerte relación con la geometría Riemanniana. En particular, la suposición de Gauss implica que todas las propiedades internas de la superficie curva pueden ser descritas en términos de las derivadas $\dfrac{\partial \xi^\alpha}{\partial x^\mu}$ de la función $\xi^\alpha(x)$ que define la transformación $x \to \xi$ de algún sistema de coordenadas x^μ que cubren la superficie al sistema cartesiano localmente inercial ξ^α, mientras que el principio de equivalencia nos dice que todos los efectos del campo gravitacional pueden ser descritos en términos de las derivadas $\dfrac{\partial \xi^\alpha}{\partial x^\mu}$ de la función $\xi^\alpha(x)$ que define la transformación de las coordenadas de "laboratorio" x^μ al sistema de coordenadas localmente inercial ξ^α.

Como vimos anteriormente, ecuación (4.12), las funciones relevantes de estas derivadas son las cantidades $g_{\mu\nu}$.

6.3 GEODESICAS EN EL ESPACIO-TIEMPO

Si estamos en un espacio-tiempo, entonces tenemos que cualquier posición la escribimos como $\left(x^0, x^1, x^2, x^3\right)$ donde x^0 es la coordenada temporal ($x^0 = ct$) y las restantes son las coordenadas espaciales. Definimos una curva como una función de un parámetro, digamos τ de manera que

$$x^\mu = x^\mu\left(\tau\right) \tag{6.1}$$

donde x^μ son funciones de clase C^2. Además supondremos que a distintos valores del parámetro corresponden puntos distintos de la curva o a lo más que son cerradas las curvas.

Ahora consideramos un arco de curva que une a dos puntos A y B con (6.1). Nos propondremos encontrar las ecuaciones diferenciales cuya solución nos proporcione la curva de longitud menor que une a los puntos A y B.

Consideremos el cuadrado del elemento de arco (o elemento de longitud) dado por

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu$$

donde $g_{\mu\nu}$ es la métrica del espacio y además $g_{\mu\nu} = g_{\mu\nu}\left(x^0, x^1, x^2, x^3\right)$ y $g_{\mu\nu} = g_{\nu\mu}$ pero x^μ depende como (6.1) entonces

$$dx^\mu = \frac{dx^\mu}{d\tau} d\tau$$

y por lo tanto

$$ds^2 = g_{\mu\nu} \frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}\left(d\tau\right)^2 \qquad \Rightarrow \qquad ds = \sqrt{g_{\mu\nu} q^\mu q^\nu}\, d\tau \tag{6.2}$$

donde $q^\mu = \dfrac{dx^\mu}{d\tau}$ y $q^\nu = \dfrac{dx^\nu}{d\tau}$. Nuestro objetivo es minimizar la longitud, por lo tanto el principio variacional debe considerar a la acción, [15]:

$$I = \int_a^b ds = \int_a^b \sqrt{g_{\mu\nu} q^\mu q^\nu}\, d\tau$$

Las condiciones para que la distancia sea una extremal, en este caso mínima, nos la dan las ecuaciones de movimiento de Euler-Lagrange que vimos anteriormente (3.10), estas son:

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial q^\alpha}\right) - \frac{\partial L}{\partial x^\alpha} = 0 \tag{6.3}$$

entonces la acción que tenemos nos indica que la lagrangiana es $L = \sqrt{g_{\mu\nu} q^\mu q^\nu}$. Entonces tenemos que

$$\frac{\partial L}{\partial x^\alpha} = \frac{1}{2}\frac{g_{\mu\nu\,,\,\alpha} q^\mu q^\nu}{\sqrt{g_{\mu\nu} q^\mu q^\nu}} = \frac{1}{2}\frac{g_{\mu\nu\,,\,\alpha} q^\mu q^\nu}{\dfrac{ds}{d\tau}}$$

y

$$\frac{\partial L}{\partial q^\alpha} = \frac{1}{2}\frac{g_{\mu\nu}\,\left(q^\mu \delta^\nu_\alpha q^\nu \delta^\mu_\alpha\right)}{\sqrt{g_{\mu\nu} q^\mu q^\nu}} = \frac{g_{\alpha\nu}\, q^\nu}{\dfrac{ds}{d\tau}}$$

donde $g_{\mu\nu\,,\,\alpha} = \dfrac{\partial g_{\mu\nu}}{\partial x^{\alpha}}$ entonces

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial q^{\alpha}}\right) = \frac{\dfrac{ds}{d\tau}\left(\dfrac{\partial g_{\alpha\nu}}{\partial x^{\gamma}} q^{\gamma} q^{\nu} + g_{\alpha\nu}\dfrac{\partial q^{\gamma}}{\partial \tau}\right) - \dfrac{\partial^2 s}{\partial \tau^2} g_{\alpha\nu}\, q^{\nu}}{\left(\dfrac{ds}{d\tau}\right)^2}$$

Ahora sustituimos estas cantidades en (6.3) tenemos

$$\frac{\dfrac{ds}{d\tau}\left(\dfrac{\partial g_{\alpha\nu}}{\partial x^{\gamma}}\dfrac{dx^{\gamma}}{d\tau}\dfrac{dx^{\nu}}{d\tau} + g_{\alpha\nu}\dfrac{d^2 x^{\nu}}{d\tau^2}\right) - \dfrac{\partial^2 s}{\partial \tau^2}\, g_{\alpha\nu}\dfrac{dx^{\nu}}{d\tau}}{\left(\dfrac{ds}{d\tau}\right)^2} - \frac{1}{2}\frac{g_{\mu\nu\,,\,\alpha} q^{\mu} q^{\nu}}{\dfrac{ds}{d\tau}} = 0$$

entonces

$$\frac{\partial g_{\alpha\nu}}{\partial x^{\gamma}}\frac{dx^{\gamma}}{d\tau}\frac{dx^{\nu}}{d\tau} - \frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x^{\alpha}}\frac{dx^{\mu}}{d\tau}\frac{dx^{\nu}}{d\tau} + g_{\alpha\nu}\frac{d^2 x^{\nu}}{d\tau^2} - \frac{g_{\alpha\nu}\dfrac{dx^{\nu}}{d\tau}\dfrac{\partial^2 s}{\partial \tau^2}}{\dfrac{ds}{d\tau}} = 0$$

o lo que es lo mismo

$$\frac{1}{2}g^{\alpha\mu}\left[\frac{\partial g_{\alpha\nu}}{\partial x^{\gamma}} + \frac{\partial g_{\alpha\gamma}}{\partial x^{\nu}} - \frac{\partial g_{\gamma\nu}}{\partial x^{\alpha}}\right]\frac{dx^{\gamma}}{d\tau}\frac{dx^{\nu}}{d\tau} + \frac{d^2 x^{\mu}}{d\tau^2} - \frac{\dfrac{dx^{\mu}}{d\tau}\dfrac{\partial^2 s}{\partial \tau^2}}{\dfrac{ds}{d\tau}} = 0$$

si definimos

$$\left\{\begin{matrix}\mu\\ \gamma\nu\end{matrix}\right\} = \frac{1}{2}g^{\alpha\mu}\left[\frac{\partial g_{\alpha\nu}}{\partial x^{\gamma}} + \frac{\partial g_{\alpha\gamma}}{\partial x^{\nu}} - \frac{\partial g_{\gamma\nu}}{\partial x^{\alpha}}\right] \tag{6.4}$$

tenemos que

$$\left\{\begin{matrix}\mu\\ \gamma\nu\end{matrix}\right\}\frac{dx^{\gamma}}{d\tau}\frac{dx^{\nu}}{d\tau} + \frac{d^2 x^{\mu}}{d\tau^2} - \frac{\dfrac{dx^{\mu}}{d\tau}\dfrac{\partial^2 s}{\partial \tau^2}}{\dfrac{ds}{d\tau}} = 0 \tag{6.5}$$

Así el símbolo $\left\{\begin{matrix}\mu\\ \gamma\nu\end{matrix}\right\}$ se le conoce como **Símbolo de Christoffel de segunda especie**, a éste lo estudiaremos más adelante.

La ecuación (6.5) es la ecuación de las geodésicas para un espacio con métrica $g_{\mu\nu}$. Cuando el parámetro τ es la longitud de arco, el último término de la ecuación (6.5) se anula y por tanto la ecuación de la geodésica queda como

$$\left\{\begin{matrix}\mu\\ \gamma\nu\end{matrix}\right\}\frac{dx^{\gamma}}{d\tau}\frac{dx^{\nu}}{d\tau} + \frac{d^2 x^{\mu}}{d\tau^2} = 0 \tag{6.6}$$

Si $\left\{\begin{matrix}\mu\\ \gamma\nu\end{matrix}\right\} = 0$ la ecuación resultante es

$$\frac{d^2 x^{\mu}}{ds^2} = 0$$

que integrando obtenemos

$$x^{\mu} = k_1 s + k_2$$

que son ecuaciones de líneas rectas dependientes del parámetro s. Por lo tanto estaremos hablando de un espacio plano. Si observamos la definición de los símbolos de Christoffel podemos decir que: En cualquier espacio plano las cantidades $g_{\mu\nu}$ son cantidades constantes.

Y como conclusión a esto diremos que cuando las componentes $g_{\mu\nu}$ no son constantes entonces estamos hablando de espacios curvos.

6.4 RELACION ENTRE LA CONECCION AFIN Y SIMBOLOS DE CHRISTOFFEL

Dentro de las leyes físicas, aparece una cantidad importante la llamada **conección afín**, la cual se define como

$$\Gamma^{\lambda}_{\mu\nu} \equiv \frac{\partial x^{\lambda}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\mu} \partial x^{\nu}} \tag{6.7}$$

donde ξ^{α} son las coordenadas de un sistema localmente inercial. Es importante mostrar que la conección afín no es en realidad un tensor, para ello pasaremos de un sistema de coordenadas x^{μ} a otro sistema de coordenadas x'^{μ}, es decir

$$\Gamma'^{\lambda}_{\mu\nu} \equiv \frac{\partial x'^{\lambda}}{\partial \xi^{\alpha}} \frac{\partial^2 \xi^{\alpha}}{\partial x'^{\mu} \partial x'^{\nu}} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \frac{\partial}{\partial x'^{\mu}} \left(\frac{\partial \xi^{\alpha}}{\partial x^{\sigma}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \right)$$

$$= \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial \xi^{\alpha}} \left[\frac{\partial x^{\sigma}}{\partial x'^{\nu}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial^2 \xi^{\alpha}}{\partial x^{\tau} \partial x^{\sigma}} + \frac{\partial^2 x^{\sigma}}{\partial x'^{\mu} \partial x'^{\nu}} \frac{\partial \xi^{\alpha}}{\partial x^{\sigma}} \right]$$

e introduciendo la ecuación (6.7) nos queda como

$$\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\tau\sigma} + \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\mu} \partial x'^{\nu}} \tag{6.8}$$

Trataremos ahora de encontrar una relación entre $\Gamma^{\lambda}_{\mu\nu}$ y $g_{\mu\nu}$, para ello notemos que

$$\frac{\partial g'_{\mu\nu}}{\partial x'^{\kappa}} = \frac{\partial}{\partial x'^{\kappa}} \left(g_{\rho\sigma} \frac{\partial x^{\rho}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \right) =$$

$$= \frac{\partial g_{\rho\sigma}}{\partial x^{\tau}} \frac{\partial x^{\tau}}{\partial x'^{\kappa}} \frac{\partial x^{\rho}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} + g_{\rho\sigma} \frac{\partial^2 x^{\rho}}{\partial x'^{\mu} \partial x'^{\kappa}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} + g_{\rho\sigma} \frac{\partial^2 x^{\sigma}}{\partial x'^{\mu} \partial x'^{\kappa}} \frac{\partial x^{\rho}}{\partial x'^{\mu}}$$

de manera analoga calculamos $\dfrac{\partial g'_{\kappa\nu}}{\partial x'^{\mu}}$ y $\dfrac{\partial g'_{\kappa\mu}}{\partial x'^{\nu}}$ para que así al sumar estas últimas y restar la primera obtengamos que

$$\frac{\partial g'_{\kappa\nu}}{\partial x'^{\mu}} + \frac{\partial g'_{\kappa\mu}}{\partial x'^{\nu}} - \frac{\partial g'_{\mu\nu}}{\partial x'^{\kappa}} = \frac{\partial x^{\tau}}{\partial x'^{\kappa}} \frac{\partial x^{\rho}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \left(\frac{\partial g_{\sigma\tau}}{\partial x^{\rho}} + \frac{\partial g_{\rho\tau}}{\partial x^{\sigma}} - \frac{\partial g_{\rho\sigma}}{\partial x^{\tau}} \right) + 2 g_{\rho\sigma} \frac{\partial^2 x^{\rho}}{\partial x'^{\mu} \partial x'^{\kappa}} \frac{\partial x^{\sigma}}{\partial x'^{\kappa}}$$

que podemos escribir como

$$\left\{ \begin{matrix} \lambda \\ \mu\nu \end{matrix} \right\}' = \frac{\partial x^{\lambda}}{\partial x'^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \left\{ \begin{matrix} \rho \\ \tau\sigma \end{matrix} \right\} + \frac{\partial x^{\lambda}}{\partial x'^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\mu} \partial x'^{\kappa}} \tag{6.9}$$

donde $\left\{ \begin{matrix} \lambda \\ \mu\nu \end{matrix} \right\}$ es como lo definimos en (6.4), es decir

$$\left\{ \begin{matrix} \lambda \\ \mu\nu \end{matrix} \right\} = \frac{1}{2} g^{\lambda\kappa} \left[\frac{\partial g_{\kappa\nu}}{\partial x^{\mu}} + \frac{\partial g_{\kappa\mu}}{\partial x^{\nu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\kappa}} \right] \tag{6.10}$$

Si nosotros ahora restamos (6.9) de (6.8) tenemos que

$$\left[\Gamma^{\lambda}_{\mu\nu} - \left\{ \begin{matrix} \lambda \\ \mu\nu \end{matrix} \right\}\right]' = \frac{\partial x^{\lambda}}{\partial x'^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \left[\Gamma^{\rho}_{\tau\sigma} - \left\{ \begin{matrix} \rho \\ \tau\sigma \end{matrix} \right\}\right]$$

lo que ahora si se comporta como un tensor.

El principio de equivalencia nos dice que existe un sistema de coordenadas especial ξ_X en cual, en un punto dado X, los efectos de la gravitación son despreciables. En este sistema no puede haber fuerza gravitacional sobre las partículas, así que $\Gamma^{\lambda}_{\mu\nu}$ se desprecia, así que las primeras derivadas de $g_{\mu\nu}$ se desperecian. Entonces $\Gamma^{\rho}_{\tau\sigma} - \left\{ \begin{matrix} \rho \\ \tau\sigma \end{matrix} \right\}$ se desprecian en un sistema de coordenadas localmente inercial y dado que este es un tensor tenemos que entonces se desprecia en todos los sistemas de coordenadas, esto es

$$\Gamma^{\rho}_{\tau\sigma} = \left\{ \begin{matrix} \rho \\ \tau\sigma \end{matrix} \right\} \tag{6.11}$$

Si ahora diferenciamos

$$\delta^{\lambda}_{\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\rho}}{\partial x'^{\nu}}$$

con respecto a x'^{μ}, encontramos que

$$\frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial^2 x^{\rho}}{\partial x'^{\nu} \partial x'^{\mu}} = -\frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x^{\sigma}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \tag{6.12}$$

y ahora la sustituimos en la ecuación (6.8) de tal manera que ahora nos resulta que

$$\Gamma'^{\lambda}_{\mu\nu} = \frac{\partial x'^{\lambda}}{\partial x^{\rho}} \frac{\partial x^{\tau}}{\partial x'^{\mu}} \frac{\partial x^{\sigma}}{\partial x'^{\nu}} \Gamma^{\rho}_{\tau\sigma} - \frac{\partial x^{\rho}}{\partial x'^{\nu}} \frac{\partial^2 x'^{\lambda}}{\partial x^{\rho} \partial x^{\sigma}} \frac{\partial x^{\sigma}}{\partial x'^{\mu}} \tag{6.13}$$

Esta relación es la que ahora también nos permite realizar la transformación $x' \to x$.

Entonces podemos escribir la ecuación de las geodésicas como

$$\Gamma^{\mu}_{\gamma\nu} \frac{dx^{\gamma}}{d\tau} \frac{dx^{\nu}}{d\tau} + \frac{d^2 x^{\mu}}{d\tau^2} = 0 \tag{6.14}$$

Podemos también definir al **Símbolo de Christoffel de primera especie** como

$$\Gamma_{\mu\nu\sigma} = \frac{1}{2} \left(g_{\mu\nu,\sigma} + g_{\mu\sigma,\nu} - g_{\nu\sigma,\mu} \right) \tag{6.15}$$

Este es simétrico entre los dos últimos índices. Una consecuencia simple de (6.15) es

$$\Gamma_{\mu\nu\sigma} + \Gamma_{\nu\mu\sigma} = g_{\mu\nu,\sigma} \tag{6.16}$$

El símbolo de Cristoffel de segunda especie definido en (6.10) y (6.11) se puede obtener a partir del de primera especie

$$\Gamma^{\mu}_{\nu\rho} = g^{\mu\sigma} \Gamma_{\sigma\nu\rho} \tag{6.17}$$

El tratar la conección afín no tiene otro fin más que el de mostrar que ésta se puede reducirse al símbolo de Christoffel y además como lenguaje para alguien que desee conocer más acerca de la teoría de la relatividad general.

6.5 DERIVACION COVARIANTE.

En general, cuando derivamos un tensor, el resultado no es otro tensor, por ejemplo, consideremos en tensor contravariante V^μ, el cual se transforma como

$$V'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} V^\nu$$

si esta la diferenciamos con respacto a x'^λ tenemos

$$\frac{\partial V'^\mu}{\partial x'^\lambda} = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\lambda}\frac{\partial V^\nu}{\partial x^\rho} + \frac{\partial^2 x'^\mu}{\partial x^\rho \partial x^\nu}\frac{\partial x^\rho}{\partial x'^\lambda} V^\nu \tag{6.18}$$

Podemos ver que el primer término del lado derecho de la ecuación anterior es justamente el que esperabamos para que la derivada fuera tensor, pero el segundo término es el que destruye este comportamiento como tensor de la derivada. La ecuación (6.18) nos servirá para construir un tensor. Usando la ecuación (6.13), podemos ver que

$$\Gamma'^\mu_{\lambda\kappa} V'^\kappa = \left[\frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\lambda}\frac{\partial x^\sigma}{\partial x'^\kappa}\Gamma^\nu_{\rho\sigma} - \frac{\partial x^\rho}{\partial x'^\lambda}\frac{\partial^2 x'^\mu}{\partial x^\rho \partial x^\sigma}\frac{\partial x^\sigma}{\partial x'^\kappa}\right]\frac{\partial x'^\kappa}{\partial x^\eta} V^\eta$$

o lo que es lo mismo

$$\Gamma'^\mu_{\lambda\kappa} V'^\kappa = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\lambda}\Gamma^\nu_{\rho\sigma} V^\sigma - \frac{\partial^2 x'^\mu}{\partial x^\rho \partial x^\sigma}\frac{\partial x^\rho}{\partial x'^\lambda} V^\sigma \tag{6.19}$$

Si sumamos las ecuaciones (6.18) y (6.19), encontramos que los términos inhomogéneos se cancelan y por lo tanto

$$\frac{\partial V'^\mu}{\partial x'^\lambda} + \Gamma'^\mu_{\lambda\kappa} V'^\kappa = \frac{\partial x'^\mu}{\partial x^\nu}\frac{\partial x^\rho}{\partial x'^\lambda}\left(\frac{\partial V^\nu}{\partial x^\rho} + \Gamma^\nu_{\rho\sigma} V^\sigma\right) \tag{6.20}$$

Así, definimos la derivada covariante de un tensor contravariante como

$$V^\mu_{;\lambda} \equiv \frac{\partial V^\mu}{\partial x^\lambda} + \Gamma^\mu_{\lambda\kappa} V^\kappa = V^\mu_{,\lambda} + \Gamma^\mu_{\lambda\kappa} V^\kappa \tag{6.21}$$

y la ecuación (6.20) nos dice que $V^\mu_{;\lambda}$ es un tensor.

También podemos definir la derivada covariante de un tensor covariante V_μ, ésta es

$$V_{\mu;\nu} \equiv \frac{\partial V_\mu}{\partial x^\nu} - \Gamma^\lambda_{\mu\nu} V_\lambda = V_{\mu,\nu} - \Gamma^\lambda_{\mu\nu} V_\lambda \tag{6.22}$$

que es fácil demostrar que es un tensor haciendo un razonamiento análogo al de la derivación covariante de un tensor contravariante.

La combinación de la derivación covariante con las operaciones algebráicas para los tensores definidas anteriormante dan resultados similares a aquellos para la derivación ordinaria. En particular:

1. La derivada covariante de una combinación lineal de tensores (con coeficientes constantes) es la misma combinación lineal de las derivadas covariantes. Por ejemplo si α y β son constantes, entonces

$$(\alpha A^\mu_\nu + \beta B^\mu_\nu)_{;\rho} = \alpha A^\mu_{\nu;\rho} + \beta B^\mu_{\nu;\rho} \tag{6.23}$$

2. La derivada covariante de un producto directo de tensores obedece la **Regla de Leibniz**. Por ejemplo

$$\left(A^\mu_\nu B^\lambda\right)_{;\rho} = A^\mu_{\nu;\rho} B^\lambda + A^\mu_\nu B^\lambda_{;\rho} \tag{6.24}$$

3. La derivada covariante de la contracción de un tensor es la contracción de la derivada covariante. Por ejemplo
$$T^{\mu\lambda}_{\lambda;\rho} = \frac{\partial T^{\mu\lambda}_{\lambda}}{\partial x^{\rho}} + \Gamma^{\mu}_{\rho\nu} T^{\nu\lambda}_{\lambda} \tag{6.25}$$

En general, en cualquier sistema de coordenadas (x^{μ}), la derivada covariante de un tensor $T^{\mu_1,\mu_2,\ldots,\mu_n}_{\nu_1,\nu_2,\ldots,\nu_m}$ es el tensor $T^{\mu_1,\mu_2,\ldots,\mu_n}_{\nu_1,\nu_2,\ldots,\nu_m\,;\rho}$ dado por

$$T^{\mu_1,\mu_2,\ldots,\mu_n}_{\nu_1,\nu_2,\ldots,\nu_m\,;\rho} \equiv \frac{\partial T^{\mu_1,\mu_2,\ldots,\mu_n}_{\nu_1,\nu_2,\ldots,\nu_m}}{\partial x^{\rho}} + \Gamma^{\mu_1}_{\sigma\rho} T^{\sigma,\mu_2,\ldots,\mu_n}_{\nu_1,\nu_2,\ldots,\nu_m} + \Gamma^{\mu_2}_{\sigma\rho} T^{\mu_1,\sigma,\ldots,\mu_n}_{\nu_1,\nu_2,\ldots,\nu_m} + \cdots + \Gamma^{\mu_n}_{\sigma\rho} T^{\mu_1,\mu_2,\ldots,\sigma}_{\nu_1,\nu_2,\ldots,\nu_m} +$$
$$-\Gamma^{\sigma}_{\nu_1\rho} T^{\mu_1,\mu_2,\ldots,\mu_n}_{\sigma,\nu_2,\ldots,\nu_m} - \Gamma^{\sigma}_{\nu_2\rho} T^{\mu_1,\mu_2,\ldots,\mu_n}_{\nu_1,\sigma,\ldots,\nu_m} - \cdots - \Gamma^{\sigma}_{\nu_m\rho} T^{\mu_1,\mu_2,\ldots,\mu_n}_{\nu_1,\nu_2,\ldots,\sigma} \tag{6.26}$$

Una conclusión importante es que la derivada covariante de un tensor de orden n, es un tensor de orden $n+1$.

Definida la derivada covariante ofrece tres casos especiales en los que aparece el gradiente, rotacional y divergencia ordinarios.

El caso más simple, el **gradiente covariante**, es la derivada covariante de un campo escalar, es decir
$$T_{;\mu} \equiv \frac{\partial T}{\partial x^{\mu}} \tag{6.27}$$

Volvamos a la definición de derivada covariante para un tensor covariante (6.22). Como $\Gamma^{\gamma}_{\nu\rho}$ es simétrico (es fácil demostrar a partir de la definición) en ν y ρ, el **rotacional covariante** es, precisamente, el rotacional ordinario

$$T_{\rho;\nu} - T_{\nu;\rho} \equiv \frac{\partial T_{\rho}}{\partial x^{\nu}} - \frac{\partial T_{\nu}}{\partial x^{\rho}} \tag{6.28}$$

El último caso especial es un poco más complicado, la divergencia. Para ello tomemos la derivada covariante de un tensor contravariante como en la ec. (6.21) de manera que
$$T^{\mu}_{;\mu} \equiv \frac{\partial T^{\mu}}{\partial x^{\mu}} + T^{\lambda}\Gamma^{\mu}_{\mu\lambda} \tag{6.29}$$
veamos que $\Gamma^{\mu}_{\mu\nu}$ tiene el siguiente valor

$$\Gamma^{\mu}_{\lambda\mu} = \Gamma^{\mu}_{\mu\lambda} = \frac{1}{2} g^{\mu\rho} \left\{ \frac{\partial g_{\rho\mu}}{\partial x^{\lambda}} + \frac{\partial g_{\rho\lambda}}{\partial x^{\mu}} - \frac{\partial g_{\mu\lambda}}{\partial x^{\rho}} \right\} = \frac{1}{2} g^{\mu\rho} \frac{\partial g_{\rho\mu}}{\partial x^{\lambda}} \tag{6.30}$$

Podemos evaluar esto fácilmente si sabemos que para una matriz arbitraria M tenemos que
$$tr\left[M^{-1}(x) \frac{\partial}{\partial x^{\lambda}} M(x) \right] = \frac{\partial}{\partial x^{\lambda}} \ln\det\, M(x) \tag{6.31}$$

donde det denota el determinante y tr la traza de la matriz, es decir, la suma de los elementos de la diagonal. Para probar (6.31), consideremos la variación en $\ln\det M$, propiamente a la variación de δx^{μ} en x^{μ} :

$$\begin{aligned}
\delta \ln\det\, M &\equiv \ln\det\left(M + \delta M\right) - \ln\det(M) \\
&= \ln \tfrac{\det(M+\delta M)}{\det(M)} \\
&= \ln\det M^{-1}\left(M + \delta M\right) \\
&= \ln\det\left(1 + M^{-1}\delta M\right) \\
&\to \ln(1 + trM^{-1}\delta M) \\
&\to trM^{-1}\delta M
\end{aligned}$$

Tomando los coeficientes de δx^μ en ambos lados resulta la ecuación (6.31). Aplicando (6.31) al caso en que M sea la matriz $g_{\mu\nu}$, encontramos de (6.30) que

$$\begin{aligned}\Gamma^\mu_{\mu\lambda} &= \frac{1}{2} g^{\mu\rho} \frac{\partial g_{\rho\mu}}{\partial x^\lambda} = \frac{1}{2} \frac{\partial \ln(g_{\rho\mu})}{\partial x^\lambda} \\ &= \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^\lambda} = \Gamma^\mu_{\lambda\mu}\end{aligned} \tag{6.32}$$

Sustituyendo en la ecuación (6.29), encontramos que

$$T^\mu_{;\mu} \equiv \frac{\partial T^\mu}{\partial x^\mu} + T^\lambda \frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}}{\partial x^\lambda}$$

o lo que es lo mismo, cambiando los índices mudos

$$T^\mu_{;\mu} = \frac{1}{\sqrt{g}} \frac{\partial \left(\sqrt{g} T^\mu\right)}{\partial x^\mu} \tag{6.33}$$

que es precisamente la **divergencia covariante**.

Una consecuencia inmediata es una forma covariante del teorema de Gauss: Si T^μ se anula en el infinito entonces

$$\int d^4x \sqrt{g} T^\mu_{;\mu} = 0 \tag{6.34}$$

Note que aquí aparece el factor $\sqrt{g}$ que hace a $d^4x\sqrt{g}$ invariante, como ya lo habíamos visto antes (4.20).

La **ecuación de D'Alambert** es aquella que se obtiene igualando a cero el operador de D'Alambert, definido en el capítulo 5, es decir $\mathbf{2}^2 V = 0$. En su forma covariante la podemos escribir como

$$g^{\mu\nu} V_{;\mu;\nu} = 0$$

Por la ecuación (6.22) y sabiendo que $Y_{;\sigma} = Y_{,\sigma}$ para el caso del escalar Y, la ecuación de D'Alambert nos queda como

$$g^{\mu\nu} \left(V_{,\mu\nu} - \Gamma^\alpha_{\mu\nu} V_{,\alpha}\right) = 0 \tag{6.35}$$

Si ahora nosotros usamos ejes rectilineos en el espacio plano, cada una de las cuatro coordenadas x^λ satisface $\mathbf{2}^2 x^\lambda = 0$. Si ahora sustituimos x^λ en vez de V en (6.35), el resultado no es un tensor ya que x^λ no es un escalar como V, así que esto sólo se puede lograr en ciertos sistemas de coordenadas.

Si sustituimos x^λ por V, entonces para $V_{,\alpha}$ nosotros sustituiremos $x^\lambda_{,\alpha} = g^\lambda_\alpha$. La ecuación (6.35) se convierte ahora, con algunos pasos algebraicos, en

$$g^{\mu\nu} \Gamma^\lambda_{\mu\nu} = 0 \tag{6.36}$$

Las coordenadas que satisfacen esta condición son llamadas **coordenadas armónicas**. Ellas proporcionan una aproximación sumamente cercana a las coordenadas rectilíneas que podemos tener en un espacio curvo.

Sabemos que

$$g^{\alpha\mu}_{,\sigma} g_{\mu\nu} + g^{\alpha\mu} g_{\mu\nu,\sigma} = \left(g^{\alpha\mu} g_{\mu\nu}\right)_{,\sigma} = g^\alpha_{\nu,\sigma} = 0$$

Multiplicando por $g^{\beta\nu}$, obtenemos que

$$g^{\alpha\beta}_{,\sigma} = -g^{\alpha\mu} g^{\beta\nu} g_{\mu\nu,\sigma} \tag{6.37}$$

Ahora sustituyendo la ecuación (6.16) en la anterior, tenemos que

$$\begin{aligned} g^{\alpha\beta}_{,\sigma} &= -g^{\alpha\mu} g^{\beta\nu} \left(\Gamma_{\mu\nu\sigma} + \Gamma_{\nu\mu\sigma}\right) \\ &= -g^{\beta\nu}\Gamma^{\alpha}_{\nu\sigma} - g^{\alpha\mu}\Gamma^{\beta}_{\mu\sigma} \end{aligned}$$

Si intercambiamos los índices $\alpha \leftrightarrow \mu$ y $\beta \leftrightarrow \nu$ obtenemos

$$g^{\mu\nu}_{,\sigma} = -g^{\nu\beta}\Gamma^{\mu}_{\beta\sigma} - g^{\mu\alpha}\Gamma^{\nu}_{\alpha\sigma} \tag{6.38}$$

Así, con la ayuda de la ecuación (6.32), es decir

$$\Gamma^{\beta}_{\sigma\beta}\sqrt{g} = \left(\sqrt{g}\right)_{,\sigma} \tag{6.39}$$

la ecuación (6.38) queda como

$$\left(g^{\mu\nu}\sqrt{g}\right)_{,\sigma} = \left(-g^{\nu\beta}\Gamma^{\mu}_{\beta\sigma} - g^{\mu\alpha}\Gamma^{\nu}_{\alpha\sigma} + g^{\mu\nu}\Gamma^{\beta}_{\sigma\beta}\right)\sqrt{g} \tag{6.40}$$

Ahora haciendo $\sigma = \nu$ tenemos que

$$\left(g^{\mu\nu}\sqrt{g}\right)_{,\nu} = -\sqrt{g}g^{\nu\beta}\Gamma^{\mu}_{\beta\nu} \tag{6.41}$$

Entonces una forma alternativa para la condición de coordenadas armónicas es

$$\left(g^{\mu\nu}\sqrt{g}\right)_{,\nu} = 0 \tag{6.42}$$

6.6 TENSORES DE RIEMANN-CHRISTOFFEL

Veamos ahora como surge cierto tensor que nos interesará más adelante, para ello nos haremos la siguiente pregunta: ?‘Será lo mismo diferenciar covarientemente cierto tensor Λ_α primero respecto a x^β y luego respecto a x^σ que si intercambiamos el orden de esta diferenciación?. La respuesta parece ser que no, esto se hace evidente observando que en la derivada covariante, como ya hemos visto, aparece el símbolo de Christiffel. Pero tratemos de demostrarlo calculando la siguente expresión $\Lambda_{\alpha;\beta;\sigma} - \Lambda_{\alpha;\sigma;\beta}$. Calculemos el primer término usando la ecuación (6.26):

$$\Lambda_{\alpha;\sigma;\beta} = \frac{\partial \Lambda_{\alpha;\sigma}}{\partial x^\beta} - \Gamma^{\mu}_{\alpha\beta}\Lambda_{\mu;\sigma} - \Gamma^{\mu}_{\sigma\beta}\Lambda_{\alpha;\mu} \tag{6.43}$$

pero por la ecuación (6.22) tenemos que

$$\Lambda_{\alpha;\sigma} = \frac{\partial \Lambda_\alpha}{\partial x^\sigma} - \Gamma^{\nu}_{\alpha\sigma}\Lambda_\nu$$

sustituimos esta última ecuación en la ecuación (6.43)

$$\Lambda_{\alpha;\sigma;\beta} = \frac{\partial}{\partial x^\beta}\left(\frac{\partial \Lambda_\alpha}{\partial x^\sigma} - \Gamma^{\nu}_{\alpha\sigma}\Lambda_\nu\right) - \Gamma^{\mu}_{\alpha\beta}\left(\frac{\partial \Lambda_\mu}{\partial x^\sigma} - \Gamma^{\nu}_{\mu\sigma}\Lambda_\nu\right) - \Gamma^{\mu}_{\sigma\beta}\left(\frac{\partial \Lambda_\alpha}{\partial x^\mu} - \Gamma^{\nu}_{\alpha\mu}\Lambda_\nu\right)$$

y tenemos que

$$\begin{aligned} \Lambda_{\alpha;\sigma;\beta} = &\frac{\partial^2 \Lambda_\alpha}{\partial x^\sigma \partial x^\beta} - \Gamma^{\nu}_{\alpha\sigma}\frac{\partial \Lambda_\nu}{\partial x^\beta} - \frac{\partial \Gamma^{\nu}_{\alpha\sigma}}{\partial x^\beta}\Lambda_\nu + \\ &-\Gamma^{\mu}_{\alpha\beta}\frac{\partial \Lambda_\mu}{\partial x^\sigma} + \Gamma^{\mu}_{\alpha\beta}\Gamma^{\nu}_{\mu\sigma}\Lambda_\nu - \Gamma^{\mu}_{\sigma\beta}\frac{\partial \Lambda_\alpha}{\partial x^\mu} + \Gamma^{\mu}_{\sigma\beta}\Gamma^{\nu}_{\alpha\mu}\Lambda_\nu \end{aligned} \tag{6.44}$$

intercambiemos ahora los índices σ y β para obtener el otro orden de derivación

$$\Lambda_{\alpha;\beta;\sigma} = \frac{\partial^2 \Lambda_\alpha}{\partial x^\sigma \partial x^\beta} - \Gamma^\nu_{\alpha\beta}\frac{\partial \Lambda_\nu}{\partial x^\sigma} - \frac{\partial \Gamma^\nu_{\alpha\beta}}{\partial x^\sigma}\Lambda_\nu - \Gamma^\mu_{\alpha\sigma}\frac{\partial \Lambda_\mu}{\partial x^\beta} + \Gamma^\mu_{\alpha\sigma}\Gamma^\nu_{\mu\beta}\Lambda_\nu - \Gamma^\mu_{\beta\sigma}\frac{\partial \Lambda_\alpha}{\partial x^\mu} + \Gamma^\mu_{\beta\sigma}\Gamma^\nu_{\alpha\mu}\Lambda_\nu$$

como ya habiamos mencionado, restemos esta última ecuación de la ecuación (6.44) y obtendremos

$$\Lambda_{\alpha;\sigma;\beta} - \Lambda_{\alpha;\beta;\sigma} = \frac{\partial \Gamma^\nu_{\alpha\beta}}{\partial x^\sigma}\Lambda_\nu - \frac{\partial \Gamma^\nu_{\alpha\sigma}}{\partial x^\beta}\Lambda_\nu + \Gamma^\mu_{\alpha\beta}\Gamma^\nu_{\mu\sigma}\Lambda_\nu - \Gamma^\mu_{\alpha\sigma}\Gamma^\nu_{\mu\beta}\Lambda_\nu$$

o lo que es lo mismo

$$\Lambda_{\alpha;\beta;\sigma} - \Lambda_{\alpha;\sigma;\beta} = \left(\frac{\partial \Gamma^\nu_{\alpha\beta}}{\partial x^\sigma} - \frac{\partial \Gamma^\nu_{\alpha\sigma}}{\partial x^\beta} + \Gamma^\mu_{\alpha\beta}\Gamma^\nu_{\mu\sigma} - \Gamma^\mu_{\alpha\sigma}\Gamma^\nu_{\mu\beta}\right)\Lambda_\nu$$

si definimos

$$R^\nu_{\alpha\beta\sigma} = \frac{\partial \Gamma^\nu_{\alpha\beta}}{\partial x^\sigma} - \frac{\partial \Gamma^\nu_{\alpha\sigma}}{\partial x^\beta} + \Gamma^\mu_{\alpha\beta}\Gamma^\nu_{\mu\sigma} - \Gamma^\mu_{\alpha\sigma}\Gamma^\nu_{\mu\beta} \tag{6.45}$$

entonces

$$\Lambda_{\alpha;\beta;\sigma} - \Lambda_{\alpha;\sigma;\beta} = R^\nu_{\alpha\beta\sigma}\Lambda_\nu \tag{6.46}$$

A $R^\nu_{\alpha\beta\sigma}$ se le conoce como **tensor de Riemann** (**o Riemann-Chistoffel**) de segunda especie. El **tensor de Riemann de primera especie** puede obtenerse bajando el índice contravariante de la siguiente manera

$$R_{\mu\alpha\beta\sigma} = g_{\nu\mu}R^\nu_{\alpha\beta\sigma} \tag{6.47}$$

Entonces podemos decir que la doble derivación covariante no dependerá del orden de derivación siempre y cuando el tensor de Riemann sea idénticamente cero.

Se puede demostrar a partir de (6.15) y (6.16) que

$$R_{\mu\nu\rho\sigma} = \frac{1}{2}\left(g_{\mu\sigma,\nu,\rho} - g_{\nu\sigma,\mu,\rho} - g_{\mu\rho,\nu,\sigma} + g_{\nu\rho,\mu,\sigma}\right) + \Gamma_{\beta\mu\sigma}\Gamma^\beta_{\nu\rho} - \Gamma_{\beta\mu\rho}\Gamma^\beta_{\nu\sigma} \tag{6.48}$$

Hasta aquí no hemos demostrado que efectivamente $R_{\mu\alpha\sigma\beta}$ sea un tensor, pero es posible demostrarlo.

Algunas de las propiedades de los tensores de Riemann-Cristoffel son las siguientes:

1) Antisimetría cambiando el orden ya sea en la primera pareja de índices o en la segunda, es decir

$$R^\beta_{\nu\rho\sigma} = -R^\beta_{\nu\sigma\rho} \tag{6.49}$$

$$\begin{aligned} R_{\mu\alpha\beta\sigma} &= -R_{\alpha\mu\beta\sigma} \\ R_{\mu\alpha\beta\sigma} &= -R_{\mu\alpha\sigma\beta} \end{aligned} \tag{6.50}$$

2) Simetría al intercambiar la primera pareja de índices por la segunda pareja, es decir

$$R_{\mu\alpha\beta\sigma} = R_{\beta\sigma\mu\alpha} = R_{\sigma\beta\alpha\mu} \tag{6.51}$$

3) Cumple con la ciclicidad, es decir

$$R^\beta_{\nu\rho\sigma} + R^\beta_{\rho\sigma\nu} + R^\beta_{\sigma\nu\rho} = 0 \tag{6.52}$$

$$R_{\mu\alpha\sigma\beta} + R_{\mu\beta\alpha\sigma} + R_{\mu\sigma\beta\alpha} = 0. \tag{6.53}$$

Las dos primeras propiedades son fáciles de demostrar y sólo demostraremos la ecuación (6.53). Para ello calculemos el primer término usando las ecuaciones (6.45) y (6.47):

$$R_{\rho\alpha\beta\sigma} = g_{\nu\rho}\left(\frac{\partial\Gamma^{\nu}_{\alpha\beta}}{\partial x^{\sigma}} - \frac{\partial\Gamma^{\nu}_{\alpha\sigma}}{\partial x^{\beta}} + \Gamma^{\mu}_{\alpha\beta}\Gamma^{\nu}_{\mu\sigma} - \Gamma^{\mu}_{\alpha\sigma}\Gamma^{\nu}_{\mu\beta}\right)$$

y

$$R_{\rho\alpha\beta\sigma} = \frac{\partial\Gamma_{\rho\alpha\beta}}{\partial x^{\sigma}} - \frac{\partial\Gamma_{\rho\alpha\sigma}}{\partial x^{\beta}} + \Gamma^{\mu}_{\alpha\beta}\Gamma_{\rho\mu\sigma} - \Gamma^{\mu}_{\alpha\sigma}\Gamma_{\rho\mu\beta}$$

para los otros dos intercambiemos los índices y obtenemos

$$R_{\rho\sigma\alpha\beta} = \frac{\partial\Gamma_{\rho\beta\sigma}}{\partial x^{\alpha}} - \frac{\partial\Gamma_{\rho\beta\alpha}}{\partial x^{\sigma}} + \Gamma^{\mu}_{\beta\sigma}\Gamma_{\rho\mu\alpha} - \Gamma^{\mu}_{\beta\alpha}\Gamma_{\rho\mu\sigma}$$

y

$$R_{\rho\beta\sigma\alpha} = \frac{\partial\Gamma_{\rho\sigma\alpha}}{\partial x^{\beta}} - \frac{\partial\Gamma_{\rho\sigma\beta}}{\partial x^{\alpha}} + \Gamma^{\mu}_{\sigma\alpha}\Gamma_{\rho\mu\beta} - \Gamma^{\mu}_{\sigma\beta}\Gamma_{\rho\mu\alpha}$$

Ahora sumemos de acuerdo con (6.53) y teniendo en cuenta las propiedades de los símbolos de Christoffel obtenemos que efectivamente

$$R_{\rho\alpha\sigma\beta} + R_{\rho\beta\alpha\sigma} + R_{\rho\sigma\beta\alpha} = 0$$

El resultado de las simetrías vistas en las dos primeras propiedades de los tensores de Riemman-Christoffel es que, de las 256 componentes de $R_{\rho\beta\sigma\alpha}$, sólo 20 son independientes.

Si el espacio es plano, podemos escoger un sistema de coordenadas que sea rectilíneo, y entonces, las $g_{\mu\nu}$ son constantes. El tensor $R_{\mu\nu\sigma\alpha}$ entonces se anula. Inversamente, si $R_{\mu\nu\sigma\alpha}$ se anula, uno puede probar que el espacio es plano.

6.7 IDENTIDADES DE BIANCHI

Consideremops ahora un tensor de orden dos el cual es el producto exterior de dos tensores de orden uno, $A_{\mu}B_{\tau}$. Calculemos su segunda derivada covariante

$$(A_{\mu}B_{\tau})_{;\rho;\sigma} = A_{\mu}B_{\tau;\rho;\sigma} + A_{\mu;\sigma}B_{\tau;\rho} + A_{\mu;\rho}B_{\tau;\sigma} + A_{\mu;\rho;\sigma}B_{\tau} \tag{6.54}$$

ahora intercambiemos los índices ρ y σ para obtener

$$(A_{\mu}B_{\tau})_{;\sigma;\rho} = A_{\mu}B_{\tau;\sigma;\rho} + A_{\mu;\rho}B_{\tau;\sigma} + A_{\mu;\sigma}B_{\tau;\rho} + A_{\mu;\sigma;\rho}B_{\tau} \tag{6.55}$$

restemos la ecuación (6.55) de (6.54)

$$(A_{\mu}B_{\tau})_{;\rho;\sigma} - (A_{\mu}B_{\tau})_{;\sigma;\rho} = A_{\mu}\left(B_{\tau;\rho;\sigma} - B_{\tau;\sigma;\rho}\right) + \left(A_{\mu;\rho;\sigma} - A_{\mu;\sigma;\rho}\right)B_{\tau}$$

y apliquemos la ecuación (6.46), así

$$(A_{\mu}B_{\tau})_{;\rho;\sigma} - (A_{\mu}B_{\tau})_{;\sigma;\rho} = A_{\mu}B_{\nu}R^{\nu}_{\tau\rho\sigma} + A_{\nu}R^{\nu}_{\mu\rho\sigma}B_{\tau} \tag{6.56}$$

Un tensor general $T_{\mu\tau}$ es expresable como una suma de términos como $A_{\mu}B_{\tau}$, así este puede satisfacer (6.56), es decir

$$T_{\mu\tau;\rho;\sigma} - T_{\mu\tau;\sigma;\rho} = T_{\mu\nu}R^{\nu}_{\tau\rho\sigma} + T_{\nu\tau}R^{\nu}_{\mu\rho\sigma} \tag{6.57}$$

Ahora tomemos que $T_{\mu\tau}$ sea la derivada covariante de un tensor A_{μ}, es decir $T_{\mu\tau} = A_{\mu;\tau}$. Por la ecuación (6.57) tenemos que

$$A_{\mu;\tau;\rho;\sigma} - A_{\mu;\tau;\sigma;\rho} = A_{\mu;\nu}R^{\nu}_{\tau\rho\sigma} + A_{\nu;\tau}R^{\nu}_{\mu\rho\sigma} \tag{6.58}$$

Encontremos dos ecuaciones semejantes a la ecuación anterior haciendo la permutación cíclica de los índices τ, ρ, y σ, estas son

$$A_{\mu;\rho;\sigma;\tau} - A_{\mu;\rho;\tau;\sigma} = A_{\mu;\nu} R^{\nu}_{\rho\sigma\tau} + A_{\nu;\rho} R^{\nu}_{\mu\sigma\tau} \tag{6.59}$$

$$A_{\mu;\sigma;\tau;\rho} - A_{\mu;\sigma;\rho;\tau} = A_{\mu;\nu} R^{\nu}_{\sigma\tau\rho} + A_{\nu;\sigma} R^{\nu}_{\mu\tau\rho} \tag{6.60}$$

Si sumamos las ecuaciones (6.58), (6.59) y (6.60) obtenemos

$$\begin{aligned} &A_{\mu;\tau;\rho;\sigma} - A_{\mu;\tau;\sigma;\rho} + A_{\mu;\rho;\sigma;\tau} - A_{\mu;\rho;\tau;\sigma} + A_{\mu;\sigma;\tau;\rho} - A_{\mu;\sigma;\rho;\tau} \\ = &A_{\mu;\nu} R^{\nu}_{\tau\rho\sigma} + A_{\nu;\tau} R^{\nu}_{\mu\rho\sigma} + A_{\mu;\nu} R^{\nu}_{\rho\sigma\tau} + A_{\nu;\rho} R^{\nu}_{\mu\sigma\tau} + A_{\mu;\nu} R^{\nu}_{\sigma\tau\rho} + A_{\nu;\sigma} R^{\nu}_{\mu\tau\rho} \end{aligned} \tag{6.61}$$

Por un lado, la parte izquierda de la ecuación anterior la podemos agrupar de la siguente manera

$$\begin{aligned} &(A_{\mu;\tau;\rho;\sigma} - A_{\mu;\rho;\tau;\sigma}) + (A_{\mu;\sigma;\tau;\rho} - A_{\mu;\tau;\sigma;\rho}) + (A_{\mu;\rho;\sigma;\tau} - A_{\mu;\sigma;\rho;\tau}) \\ &= (A_{\mu;\tau;\rho} - A_{\mu;\rho;\tau})_{;\sigma} + (A_{\mu;\sigma;\tau} - A_{\mu;\tau;\sigma})_{;\rho} + (A_{\mu;\rho;\sigma} - A_{\mu;\sigma;\rho})_{;\tau} \\ &= \left(A_{\nu} R^{\nu}_{\mu\tau\rho}\right)_{;\sigma} + \left(A_{\nu} R^{\nu}_{\mu\sigma\tau}\right)_{;\rho} + \left(A_{\nu} R^{\nu}_{\mu\rho\sigma}\right)_{;\tau} \end{aligned}$$

donde para pasar del segundo al tercer renglón se ha hecho uso de la ecuación (6.46), entonces la ecuación anterior queda, aplicando la regla de Leibniz de la derivacción covariante de un producto de tensores, como

$$\begin{aligned} &(A_{\mu;\tau;\rho;\sigma} - A_{\mu;\rho;\tau;\sigma}) + (A_{\mu;\sigma;\tau;\rho} - A_{\mu;\tau;\sigma;\rho}) + (A_{\mu;\rho;\sigma;\tau} - A_{\mu;\sigma;\rho;\tau}) \\ = &A_{\nu;\sigma} R^{\nu}_{\mu\tau\rho} + A_{\nu;\rho} R^{\nu}_{\mu\sigma\tau} + A_{\nu;\tau} R^{\nu}_{\mu\rho\sigma} + A_{\nu} \left(R^{\nu}_{\mu\tau\rho;\sigma} + R^{\nu}_{\mu\sigma\tau;\rho} + R^{\nu}_{\mu\rho\sigma;\tau}\right) \end{aligned} \tag{6.62}$$

Por otro lado, la parte derecha de la ecuación (6.61) la podemos rearreglar como

$$A_{\mu;\nu} \left(R^{\nu}_{\tau\rho\sigma} + R^{\nu}_{\rho\sigma\tau} + R^{\nu}_{\sigma\tau\rho}\right) + A_{\nu;\tau} R^{\nu}_{\mu\rho\sigma} + + A_{\nu;\rho} R^{\nu}_{\mu\sigma\tau} + A_{\nu;\sigma} R^{\nu}_{\mu\tau\rho}$$

donde la primera parte (la que está entre paréntesis se anula por la propiedad de ciclicidad, ecuación (6.52) y los tres términos restantes se anulan con los mismos tres términos que aparecen en la parte izquerda de la ecuación (6.61), como se ve en la ecuación (6.62), por lo tanto, la ecuación (6.61) va a quedar como

$$A_{\nu} \left(R^{\nu}_{\mu\tau\rho;\sigma} + R^{\nu}_{\mu\sigma\tau;\rho} + R^{\nu}_{\mu\rho\sigma;\tau}\right) = 0$$

o lo que es lo mismo

$$R^{\nu}_{\mu\tau\rho;\sigma} + R^{\nu}_{\mu\sigma\tau;\rho} + R^{\nu}_{\mu\rho\sigma;\tau} = 0 \tag{6.63}$$

El tensor de Riemann satisface estas ecuaciones diferenciales que son conocidas como **Identidades de Bianchi**.

Estas identidades de Bianchi que son válidas para el tensor de Riemann de segunda especie, también lo son para el tensor de Riemann de primera especie, es decir

$$R_{\nu\mu\tau\rho;\sigma} + R_{\nu\mu\sigma\tau;\rho} + R_{\nu\mu\rho\sigma;\tau} = 0 \tag{6.64}$$

6.8 TENSOR DE RICCI

Por supuesto que otros tensores pueden ser formados usando al tensor métrico mismo para formar combinaciones lineales de $R^{\sigma}_{\mu\nu\rho}$. Las más comunes son las formas contraídas

$$R_{\mu\nu} = R^{\sigma}_{\mu\sigma\nu} = g^{\sigma\rho} R_{\rho\mu\sigma\nu} \tag{6.65}$$

El tensor obtenido $R_{\mu\nu}$ es llamado el **tensor de curvatura de Ricci**. Volvemos a escribir la relación de simetría (6.51) como

$$R_{\rho\mu\sigma\nu} = R_{\sigma\nu\rho\mu}$$

si la multiplicamos por $g^{\sigma\rho}$ obtenemos que

$$R_{\mu\nu} = R_{\nu\mu} \tag{6.66}$$

esto nos dice que el tensor de Ricci es simétrico. Podemos volver a contraer el tensor $R_{\mu\nu}$ como

$$g^{\mu\nu} R_{\mu\nu} = R^{\mu}_{\mu} = R$$

y el resultado es un escalar que denotaremos como R y que es llamado el **escalar de curvatura** o **curvatura total**†. Esta curvatura está definida de tal manera que sea positiva para la superficie de una esfera en tres dimensiones.

Se puede comprobar fácilmente de (6.45) que

$$R_{\mu\nu} = \Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha} - \Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} + \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} \tag{6.67}$$

La relación de Bianchi (6.63) involucra 5 subíndices. Contraigamos estas relaciones dos veces con el tensor métrico $g^{\mu\rho}$ y hagamos $\tau = \alpha$ de manera que (6.63) queda como

$$g^{\mu\rho}\left(R^{\nu}_{\mu\tau\rho;\nu} + R^{\nu}_{\mu\rho\nu;\tau} + R^{\nu}_{\mu\nu\tau;\rho}\right) = 0$$

o lo que es lo mismo

$$\left(g^{\mu\rho}R^{\nu}_{\mu\tau\rho}\right)_{;\nu} + \left(g^{\mu\rho}R^{\nu}_{\mu\rho\nu}\right)_{;\tau} + \left(g^{\mu\rho}R^{\nu}_{\mu\nu\tau}\right)_{;\rho} = 0 \tag{6.68}$$

dado que las derivadas covariantes del tensor métrico se anulan. Ahora como

$$\begin{aligned} g^{\mu\rho}R^{\nu}_{\mu\tau\rho} &= g^{\mu\rho}g^{\nu\beta}R_{\beta\mu\tau\rho} \\ &= g^{\mu\rho}g^{\nu\beta}R_{\mu\beta\rho\tau} = R^{\nu}_{\tau} \end{aligned}$$

Uno puede escribir R^{ν}_{τ} con un índice sobre el otro o también como $R_{\nu\tau}$ dado que es simétrico. La ecuación (6.68) se puede escribir entonces como

$$R^{\nu}_{\tau;\nu} + \left(g^{\mu\rho}R_{\mu\tau}\right)_{;\rho} - R_{;\tau} = 0 \tag{6.69}$$

El último término tiene signo negativo dado que en $R^{\nu}_{\mu\rho\nu}$ los índices a contraer son antisimétricos en un intercambio de ellos. La ecuación (6.69) también se puede escribir como

$$2R^{\nu}_{\tau;\nu} - R_{;\tau} = 0 \tag{6.70}$$

dado que en el segundo término de la ecuación (6.69) el índice ρ es mudo y se puede cambiar por el índice ν.

La relación (6.70) es la relacióndе Bianchi para el tensor de Ricci. Si nosotros subimos el índice τ obtenemos

$$\left(R^{\tau\nu} - \frac{1}{2}g^{\tau\nu}R\right)_{;\nu} = 0 \tag{6.71}$$

6.9 LEY DE LA GRAVITACION DE EINSTEIN

A lo largo de este capítulo, a excepción de la ecuación de las geodésicas, se han hecho sólo desarrollos matemáticos y su aplicación al espacio curvo en cualquier número de dimensiones. En donde aparece explicitamente el número de dimensiones de nuestro espacio en este formalismo es en la ecuación

$$g^{\mu}_{\mu} = \#\text{ de dimensiones}$$

†Ver referencia [5].

Einstein supuso que en el espacio vacío

$$R^{\mu\nu} = 0 \tag{6.72}$$

esto constituye su ley de gravitación, "vacío" significa que el espacio está libre de materia y no hay campos físicos excepto el campo gravitacional el cual no perturba la vacuidad caso contrario a los oros campos. Las condiciones de espacio vacío se pueden aplicar en buena aproximación para el espacio entre los planetas en el sistema solar y entonces la ecuación (6.72) se aplica ahí.

Obviamente un espacio plano satisface la ecuación (6.72), entonces las geodésicas son líneas rectas y ahí las partículas se mueven sobre éstas. Donde el espacio no es plano, la ley de Einstein impone restricciones sobre la curvatura que combinado con el hecho de que los planetas se mueven sobre las geodésicas, esto da alguna información sobre su movimiento.

Nosotros podemos ver a las componentes $g^{\mu\nu}$ como los potenciales que describen al campo gravitacional. Hay entonces 10 de ellos (debido a la simetría de $g^{\mu\nu}$), no como en la teoría Newtoniana donde sólo hay uno. Estos no sólo describen al campo gravitacional, sino que también al sistema de coordenadas. El campo gravitacional y el sistema de coordenadas están íntimamente relacionados en la teoría de Einstein y uno no puede ser descrito sin el otro.

La ecuación (6.72) es una ecuación de segundo orden en $g^{\mu\nu}$ (puede verse en la definición de $R^{\mu\nu}$ y de los símbolos de Christoffel), entonces en este caso la ecuación es no lineal.

La no linealidad significa que las ecuaciones son complicadas y de aquí su dificultad para obtener soluciones aproximadas.

6.10 MODIFICACION DE LA ECUACION DE EINSTEIN POR PRESENCIA DE MATERIA

Las **ecuaciones de Einstein en ausencia de materia** son, como ya vimos

$$R^{\mu\nu} = 0$$

o su contracción

$$R = 0$$

y también

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = 0 \tag{6.73}$$

Si comenzamos con las ecuaciones (6.73) obtenemos por contracción que

$$R - 2R = 0$$

y así podemos regresar a (6.72). Nosotros podemos, por tanto, usar a (6.72) o (6.73) de manera indistinta como las ecuaciones básicas para el espacio vacío.

En presencia de materia, estas ecuaciones son modificadas. Supongamos que (6.72) esta igualada

$$R^{\mu\nu} = X^{\mu\nu} \tag{6.74}$$

y (6.73) a

$$R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R = Y^{\mu\nu} \tag{6.75}$$

Aquí $X^{\mu\nu}$ y $Y^{\mu\nu}$ son tensores simétricos de orden dos que indican la presencia de materia. Podemos ver que (6.75) es más conveniente para trabajar porque tenemos la relación de Bianchi (6.71) la cual nos dice que

$$\left(R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right)_{;\nu} = 0$$

Por lo que (6.75) requiere que

$$Y^{\mu\nu}_{;\nu} = 0 \tag{6.76}$$

Cualquier tensor $Y^{\mu\nu}$ producido por materia puede satisfacer esta condición, de otra forma las ecuaciones (6.75) no serían consistentes.

Tratatremos de ver que el tensor $Y^{\mu\nu}$ multiplicado por alguna constante se interpreta como la densidad y el flujo de la energía y el momento (no gravitacional). $Y^{\mu 0}$ es la densidad y $Y^{\mu i}$ es el flujo.

Supongamos que tenemos una distribución de materia que viaja con velocidad continuamente variable de un punto a una vecindad de él. Si x^μ denota las coordenadas de un elemento de materia, podemos introducir el vector velocidad $v^\mu = \dfrac{dx^\mu}{ds}$ la cual seá una función continua de las x's como un campo de funciones. Esta tiene las propiedades

$$g_{\mu\nu}v^\mu v^\nu = 1 \tag{6.77}$$

como en (4.50) simplemente reemplazando el tensor métrico del espacio de Minkowski por el tensor métrico para cualquier espacio, además

$$0 = \left(g_{\mu\nu}v^\mu v^\nu\right)_{;\sigma} = g_{\mu\nu}\left(v^\mu v^\nu_{;\sigma} + v^\mu_{;\sigma}v^\nu\right) = 2g_{\mu\nu}v^\mu v^\nu_{;\sigma}$$

de tal menera que

$$v_\nu v^\mu_{;\sigma} = 0 \tag{6.78}$$

Podemos introducir un campo escalar ρ tal que el campo vectorial ρv^μ determina la densidad y el flujo de la materia así como J^μ determina la densidad y el flujo para el electromagnetismo.

Recordemos que dos de las ecuaciones de Maxwell pueden ser escritas como

$$F^{\mu\nu}_{,\nu} = \mu_0 J^\mu$$

esta ecuación pasa a ser manifiestamente covariante para la relatividad general sustituyendo la derivada por una derivada covariante, es decir

$$F^{\mu\nu}_{;\nu} = \mu_0 J^\mu$$

dada la ecuación (6.33), la cual se aplica también a cualquier campo tensorial de orden dos[‡], tenemos que

$$\left(F^{\mu\nu}\sqrt{g}\right)_{,\nu} = \mu_0 J^\mu \sqrt{g}$$

Esto nos da inmediatamente que

$$\left(J^\mu\sqrt{g}\right)_{,\mu} = \mu_0^{-1}\left(F^{\mu\nu}\sqrt{g}\right)_{,\mu,\nu} = 0 \tag{6.79}$$

y por la ecuación (6.33) y la ecuación (6.34) vemos que (6.79) nos da una ley de conservación para la electricidad.

[‡]Ver por ejemplo sección 21 de [5]

De lo anterior podemos decir que $\rho v^0 \sqrt{g}$ es la densidad y $\rho v^i \sqrt{g}$ es el flujo. En analogía con (6.79), la conservación de la materia es

$$(\rho v^\mu \sqrt{g})_{,\mu} = 0 \tag{6.80}$$

o lo que es lo mismo

$$(\rho v^\mu)_{;\mu} = 0 \tag{6.81}$$

La materia que hemos considerado tendrá una densidad de energía $\rho v^0 v^0 \sqrt{g}$ y el flujo de energía $\rho v^i v^i \sqrt{g}$, similarmente tendrá una densidad de momento $\rho v^i v^0 \sqrt{g}$ y el flujo de momento $\rho v^i v^j \sqrt{g}$. Tomemos

$$T^{\mu\nu} = \rho v^\mu v^\nu \tag{6.82}$$

Entonces $T^{\mu\nu} \sqrt{g}$ da la densidad y el flujo de energía y momento. $T^{\mu\nu}$ es llamado **tensor de energía-momento material**. Este es, por supuesto, simétrico.

Veamos si es posible introducir este tensor de tal manera que dsciba a la materia en la ecuación de Einstein (6.75). Para ver esto requeriremos de la condición de que $T^{\mu\nu}$ satisfaga la condición (6.76), es decir $T^{\mu\nu}_{;\nu} = 0$. Tenemos de la definición (6.82) que

$$T^{\mu\nu}_{;\nu} = (\rho v^\mu v^\nu)_{;\nu} = v^\mu (\rho v^\nu)_{;\nu} + (\rho v^\mu)_{;\nu} v^\nu$$

El primer término se anula dada la condición de la conservación de la masa (6.81). El segundo término se anula si la materia se mueve a lo largo de la geodésica, es decir

$$\frac{dv^\mu}{ds} = v^\mu_{,\nu} v^\nu$$

Así de (6.14), sabiendo que $\dfrac{dx^\gamma}{ds} = v^\gamma$ y $\dfrac{dx^\nu}{ds} = v^\nu$, tenemos que

$$\left(v^\mu_{,\nu} + \Gamma^\mu_{\gamma\nu} v^\gamma\right) v^\nu = 0$$

pero lo que está entre paréntesis es la derivada covariante de v^μ, es decir que la ecuación anterior queda como

$$v^\mu_{;\nu} v^\nu = 0 \tag{6.83}$$

De todo lo expuesto anteriormente, entonces podemos sustituir el tensor de energía-momento material (6.82) salvo un coeficiente constante k, en la ecuación de Einstein (6.75) y obtendremos

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = k T^{\mu\nu} \tag{6.84}$$

Para determinar el valor de la constante k se tiene que hacer una aproximación newtoniana que no haremos aquí pero que puede encontrarse en el capítulo 7 de la referencia [32] o en la sección 25 de la referencia [5], el valor de esta constante es $k = -8\pi$, entonces la **ecuación de Einstein para la presencia de una distribución de materia con un campo de velocidades** se puede escribir como

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = -8\pi T^{\mu\nu} \tag{6.85}$$

Chapter 7
PRINCIPIOS VARIACIONALES DE LOS CAMPOS

En este último capítulo encontraremos las ecuaciones de campo electromagnético (ecuaciones de Maxwell en su formulación covariante) que se obtuvieron en el capítulo 5 y las ecuaciones de campo gravitacional (ecuaciones de Einstein) encontradas en el capítulo anterior, a partir de un principio variacional, es decir, propondermos integrales de acción que sean invariantes bajo transformaciones de Lorentz para posteriormente aplicarles la ecuaciones de Euler-Lagrange encontradas en al capítulo 3 o realizar la variación de dichas integrales respecto de las variables que describen al propio campo.

7.1 CONSTRUCCION DE FUNCIONALES DE ACCION INVARIANTES.

Antes de comenzar a construir la acción para los diversos campos, mencionaremos una conclusión importante que ha surgido a lo largo de todo este trabajo: la acción debe ser necesariamente un escalar o invariante de Lorentz, pues, de lo contrario, podría tener un mínimo para un observador diferente que para otro. La condicion de que la acción sea un invariante de Lorentz impone restricciones muy fuertes.

Es claro a partir de lo anterior que toda acción que se construya sea invariante. Dada la forma que tiene la acción, es decir la de una integral, podemos ver que un elemento de volumen invariante en la teoría de la relatividad especial es el tiempo propio (4.49). Para constuir la acción que involucre los efectos gravitacionales necesitamos construir un invariante a partir de la ecuación (4.21).

A lo largo de todo el presente trabajo hemos dicho que toda integral de acción debe ser un escalar y como los escalares obedecen a la regla de adición, entonces cualquier acción puede ser escrita como suma de varias de ellas. Además recordando como calculamos la variación, ecuación (3.2), podemos ver que ésta es lineal, así, las variación de una acción que es la suma de varias, será la suma de las variaciones de cada una de las integrales de acción que la conforman.

Entonces ya estamos en posibilidad de empezar a construir la acción para el campo electromagnético y para el campo gravitacional.

7.2 EL CAMPO ELECTROMAGNETICO

Encontremos la función de acción I para el sistema constituido por el campo electromagnético y las partículas situadas en él. Esta acción ha de constar de tres partes

$$I = I_p + I_{cp} + I_c \tag{7.1}$$

donde I_p es la parte de la acción que sólo depende de las propiedades de las partículas, I_{cp} es la parte que depende de la interacción de las partículas con el campo e I_c es la parte que depende sólo de las propiedades del campo cuando no existen cargas.

Averiguemos como podemos escribir estas tres partes. La parte de la acción que corresponde a I_p es precisamente la correspondiente a las partículas libres, esto es, a las partículas en ausencia de campo. La acción para una partícula libre se puede construir de la lagrangiana dada por la ecuación (4.54):

$$I_p = -\int mc^2\gamma^{-1}dt \tag{7.2}$$

o para varias partículas

$$I_p = -\sum^N \int mc^2\gamma^{-1}dt$$

donde N es el número total de partículas en el sistema que se encuentra en el campo electromagnético.

Las propiedades de una partícula respecto de su interacción con el campo electromagnético resulta estar determinada por un sólo parámetro, la carga q de la partícula, que puede ser una cantidad positiva o negativa (o igual a cero)*. Las propiedades del campo, en cambio, están caracterizadas por el potencial vectorial $\mathcal{A}_\mu$ cuyas componentes son funciones de las coordenadas y del tiempo como ya se vió en el capítulo cinco. Una acción escalar natural que involucra ambas propiedades es de la forma

$$I_{pc} = -k_1 q\int \mathcal{A}_\mu dx^\mu \tag{7.3}$$

donde las funciones $\mathcal{A}_\mu$ se toman en los puntos de la línea de universo de la partícula y k_1 es una constante que evaluaremos de tal manera que las ecuaciones que se obtengan en esta sección tengan las mismas unidades que las encontradas en el capítulo 5. Si tenemos un sistema de N partículas en el campo electromagnético

$$I_{pc} = -\sum^N k_1 q\int \mathcal{A}_\mu dx^\mu \tag{7.4}$$

Hasta aquí, podemos hablar de una acción formada por las acciones (7.2) y (7.4), que es la de un número N de cargas en un campo electromagnético, es decir

$$I' = -\int mc^2\gamma^{-1}dt - \sum^N k_1 q\int \mathcal{A}_\mu dx^\mu \tag{7.5}$$

En vez de considerar a las cargas como puntos, por conveniencia matemática las supondremos distribuidas continuamente en el espacio. Se puede entonces introducir la densidad de carga ρ tal que ρdV es la carga total contenida en el volumen dV, la densidad ρ es, como ya vimos anteriormente, una función de las coordenadas y del tiempo. Entonces

$$\begin{aligned} dqdx^\mu &= \rho dV dx^\mu \\ &= \rho dV \tfrac{dx^\mu}{dt} dt \end{aligned}$$

y también como ya vimos, ecuación (5.37) tenemos que

$$dqdx^\mu = J^\mu dV dt \tag{7.6}$$

entonces introduzcamos esto en (7.3)

$$\begin{aligned} I_{pc} &= -k_1 \textstyle\int \mathcal{A}_\mu dx^\mu dq \\ &= -k_1 \textstyle\int \mathcal{A}_\mu J^\mu dV dt \end{aligned}$$

*Esta proposición debe considerarse, en buena medida, como resultado de datos experimentales.

o lo que es lo mismo

$$I_{pc} = -\frac{k_1}{c}\int \mathcal{A}_\mu J^\mu d\Omega \tag{7.7}$$

donde $d\Omega = dx^0 dx^1 dx^2 dx^3 = cdtdV$. Por lo tanto

$$I' = -\int mc^2\gamma^{-1}dt - \frac{k_1}{c}\int \mathcal{A}_\mu J^\mu d\Omega$$

Finalmente, I_c es la parte de la acción que sólo depende de las propiedades del campo en sí, es decir, cuando no existen cargas. Este término en la acción es necesario cuando se tratan de encontrar las ecuaciones que determinan al propio campo, es decir las ecuaciones de campo o de "movimiento".

Para establecer la forma de la acción correspondiente al campo, partiremos de la siguiente propiedad, muy importante, de los campos electromagnéticos. Muestra la experiencia que el campo electromagnético obedece al llamado **principio de superposición**. Este principio afirma que si una carga produce cierto campo y otra carga produce un segundo campo, el campo producido por las dos cargas resulta de la simple composición vectorial de los campos producidos por cada una de ellas individualmente. Esto significa que la intensidad del campo resultante en cada punto es igual a la suma vectorial de las intensidades de cada uno de dichos campos en aquel punto.

Cualquier solución de las ecuaciones del campo representan un campo que puede existir en la naturaleza. Según el principio de superposición, la suma de un número cualquiera de tales campos debe ser un campo que puede existir en la naturaleza, esto es, debe satisfacer las ecuaciones de campo. En analogía con las ecuaciones diferenciales lineales homogéneas que poseen precisamente la propiedad de que la suma de un número cualquiera de soluciones es también solución. Luego, las ecuaciones de campo deben ser ecuaciones diferenciales lineales homogéneas.

De esta discusión se sigue que el integrando correspondiente a la acción I_c debe ser una expresión cuadrática en las componentes del campo. Sólo en este caso serán lineales las ecuaciones del campo ya que estas se obtienen variando la acción, y en la variación el grado del integrando disminuye una unidad.

Los potenciales no pueden aparecer en la expresión de la acción I_c ya que no están determinados unívocamente (en I_{pc} esta falta de unicidad carecía de importancia). Por consiguiente, I_c debe ser la integral de una cierta función del tensor de campo electromagnético $F^{\mu\nu}$. El integrando de I_c no puede contener derivadas de $F^{\mu\nu}$ puesto que la lagrangiana puede contener, además de las coordenadas, tan sólo sus primeras derivadas respecto del tiempo, y el papel de las "coordenadas" (es decir de los elementos que se deben variar en el principio de mínima acción) lo representan en esta caso el potencial vectorial $\mathcal{A}_\mu$ del campo.

Si bien los campos eléctrico y magnético cambian de un sistema de referencia a otro, podemos contar con cantidades invariantes que no dependan de la velocidad del observador. Esto se logra gracias a que se puede construir un escalar y un pseudoescalar a partir del tensor $F^{\mu\nu}$, es decir, los escalares $F_{\mu\nu}F^{\mu\nu}$ y $F^*_{\mu\nu}F^{\mu\nu}$. A partir de la forma explícita de $F^{\mu\nu}$ y $F^{*\mu\nu}$ vista en el capítulo 5, se puede ver que estos tienen los valores

$$\begin{aligned} F_{\mu\nu}F^{\mu\nu} &= 2\left(\mathsf{B}^2 - \frac{\mathsf{E}^2}{c^2}\right) \\ F^*_{\mu\nu}F^{\mu\nu} &= \frac{4}{c}\left(\mathsf{E}\cdot\mathsf{B}\right) \end{aligned} \tag{7.8}$$

Lo anterior implica que $\mathsf{B}^2 - \frac{\mathsf{E}^2}{c^2}$ y $\mathsf{E}\cdot\mathsf{B}$ son invariantes de Lorentz.

Entonces para constriur la acción debemos contar con alguno de los invariantes anteriores. La la magnitud $F^*_{\mu\nu}F^{\mu\nu}$ se excluye de la acción dado que esta cantidad no es un escalar sino un seudoescalar, ver la parte final del capítulo 1.

Por lo tanto, I_c debe ser de la forma

$$I_c = -k_2 \int\int F_{\mu\nu}F^{\mu\nu} dV\, dt$$

En el integrando aparece la cantidad $F_{\mu\nu}F^{\mu\nu}$ con signo negativo, el cual tendrá un valor de $2(\frac{\mathsf{E}^2}{c^2} - \mathsf{B}^2)$. Esto es ya que el campo E contiene la derivada $\frac{\partial \mathcal{A}}{dt}$; pero es fácil ver que $\left(\frac{\partial^2 \mathcal{A}}{dt^2}\right)$ debe aparecer en la acción con signo positivo (y por consiguiente E^2 debe tener también signo positivo). En efecto, si $\left(\frac{\partial^2 \mathcal{A}}{dt^2}\right)$ apareciera en I_c con signo negativo, un cambio suficientemente rápido del potencial con el tiempo podría siempre hacer de I_c una cantidad negativa de valor absoluto arbitrariamente grande. En estas condiciones, I_c no podría tener un mínimo, en contra de lo exige el principio de mínima acción.

El valor de la constante k_2 lo escogeremos al final ya que desamos que las ecuaciones obtenidas tengan las mismas unidades como en el capítulo 5. Resumiendo, la acción para el campo tiene la forma

$$I_c = -\frac{k_2}{c} \int F_{\mu\nu}F^{\mu\nu} d\Omega \tag{7.9}$$

La acción I entonces toma la forma, a partir de las ecuaciones (7.2), (7.7) y (7.9):

$$I = -\int mc^2\gamma^{-1} dt - \frac{k_1}{c}\int \mathcal{A}_\mu J^\mu d\Omega - \frac{k_2}{c}\int F_{\mu\nu}F^{\mu\nu} d\Omega \tag{7.10}$$

7.2.1 EL PRIMER PAR DE ECUACIONES DE MAXWELL.

Una carga situada en un campo no sólo esta sujeta a la fuerza que el campo ejerce sobre ella, sino que a su vez actúa sobre el campo modificándolo. Sin embargo, si la carga q no es grande, se puede prescindir de los efectos de la carga sobre el campo. En este caso, al considerar el movimiento de la carga en un campo dado, podemos admitir que el campo en sí no depende ni de las coordenadas ni de la velocidad de la carga[†].

Se trata, entonces de hallar las ecuaciones de movimiento de una carga en un campo electromagnético dado. Estas ecuaciones se encuentran variando la acción, esto es vienen dadas por las ecuaciones de Euler-Lagrange

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \mathsf{v}}\right) = \frac{\partial L}{\partial \mathsf{r}} \tag{7.11}$$

Dado que sólo queremos encontrar las ecuaciones de movimiento para una partícula que se mueve en un campo electromagnético, la acción a considerar es I', es decir la acción del movimiento de una sola carga y su interacción con el campo (7.5)

$$I' = -\int mc^2\gamma^{-1} dt - k_1 q \int \mathcal{A}_\mu dx^\mu \tag{7.12}$$

Así, como vimos en el capítulo 5, las tres componentes espaciales del cuadrivector $\mathcal{A}_\mu$ forman un vector en tres dimensiones, ver ecuación (5.40)

$$\mathcal{A}^\mu = \left(\frac{\varphi}{c}, \mathsf{A}\right)$$

[†]De hecho existen más condiciones que debe cumplir la carga para que se pueda considerar pequeña, pero no mencionaremos aquí, para ello recomendamos la lectura de [[2]], [[13]] o [[17]].

o también
$$\mathcal{A}_\mu = \left(\frac{\varphi}{c}, -\mathsf{A}\right)$$
La integral de acción se puede escribir como
$$I' = \int \left(-mc^2\gamma^{-1} + k_1 q \mathsf{A}\cdot\mathsf{v} - k_1 q\varphi\right) dt \tag{7.13}$$
donde se ha introducido la velocidad de la partícula $\mathsf{v} = \frac{d\mathsf{r}}{dt}$ y se ha tomado como la variable de integración a t. El integrando es precisamente la lagrangiana de una carga en un campo electromagnético
$$L' = -mc^2\gamma^{-1} + qk_1\left(\mathsf{A}\cdot\mathsf{v} - \varphi\right) \tag{7.14}$$
Aplicando las ecuaciones de Euler-Lagrange a L' tenemos que
$$\mathsf{P} = \frac{\partial L'}{\partial \mathsf{v}} = m\mathsf{v}\gamma + qk_1\mathsf{A} = \mathsf{p} + qk_1\mathsf{A} \tag{7.15}$$
donde P es el momento generalizado como lo vimos en la ecuación (4.55) y (4.56). Además
$$\frac{\partial L'}{\partial \mathsf{r}} = \nabla L' = qk_1\boldsymbol{\nabla}\left(\mathsf{A}\cdot\mathsf{v}\right) - qk_1\boldsymbol{\nabla}\varphi$$
pero según la fórmula del análisis vectorial
$$\boldsymbol{\nabla}\left(\mathsf{A}\cdot\mathsf{B}\right) = (\mathsf{A}\cdot\boldsymbol{\nabla})\mathsf{B} + (\mathsf{B}\cdot\boldsymbol{\nabla})\mathsf{A} + \mathsf{B}\times(\boldsymbol{\nabla}\times\mathsf{A}) + \mathsf{A}\times(\boldsymbol{\nabla}\times\mathsf{B})$$
donde A y B son vectores arbitrarios, tenemos que
$$\boldsymbol{\nabla}\left(\mathsf{A}\cdot\mathsf{v}\right) = (\mathsf{v}\cdot\boldsymbol{\nabla})\mathsf{A} + \mathsf{v}\times(\boldsymbol{\nabla}\times\mathsf{A})$$
donde se ha recordado que la derivación es con respecto a r manteniendo constante a v. Por tanto
$$\frac{\partial L'}{\partial \mathsf{r}} = qk_1\left[(\mathsf{v}\cdot\boldsymbol{\nabla})\mathsf{A} + \mathsf{v}\times(\boldsymbol{\nabla}\times\mathsf{A})\right] - qk_1\nabla\varphi \tag{7.16}$$
Pero la diferencial total $\frac{d\mathsf{A}}{dt}dt$ consta de dos términos: el cambio $\frac{\partial\mathsf{A}}{\partial t}dt$ del potencial vectorial con el tiempo en un punto fijo del espacio y el cambio debido al paso de un punto del espacio a otro separado una distancia $d\mathsf{r}$. Este segundo término, como es sabido por el análisis vectorial, vale $(d\mathsf{r}\cdot\boldsymbol{\nabla})\,\mathsf{A}$. Por tanto
$$\frac{d\mathsf{p}}{dt} = -qk_1\frac{\partial\mathsf{A}}{\partial t} - qk_1\nabla\varphi + qk_1\left[\mathsf{v}\times(\boldsymbol{\nabla}\times\mathsf{A})\right] \tag{7.17}$$

Esta es la ecuación de movimiento de una carga en un campo electromagnético. Del lado izquierdo de la ecuación aparece la expresión $\frac{d\mathsf{p}}{dt}$ que es la fuerza que actúa sobre la carga en el campo electromagnético. Pero en el capítulo 5 vimos que esta fuerza es llamada Fuerza de Lorentz (5.10)
$$\mathsf{F} = q\mathsf{E} + q\mathsf{v}\times\mathsf{B}$$
Si comparamos ambas ecuaciones, (7.17) y (5.10), nos resultan las ecuaciones (5.31) y (5.30), esto si la constante $k_1 = 1$:
$$\mathsf{E} = -\frac{\partial\mathsf{A}}{\partial t} - \nabla\varphi \tag{7.18}$$

y

$$\mathsf{B} = \boldsymbol{\nabla} \times \mathsf{A} \tag{7.19}$$

Si tomamos el rotacional de la ecuación (7.18) obtenemos la ecuación (5.24) y tomando la divergencia de (7.19) obtenemos la ecuación (5.23) que son dos de las ecuaciones de Maxwell vistas en el capítulo 5.

7.2.2 EL SEGUNDO PAR DE ECUACIONES DE MAXWELL

Ahora para obtener las otras dos ecuaciones de Maxwell a partir de un principio variacional hemos de suponer que ya se conoce el movimiento de la carga (se supone que ya se ha variado la trayectoria) para sólo variar el campo electromagnético, es decir, el potencial vectorial. Por tanto, la variación del primer término en (7.10) es nula y en el segundo hay que conservar invariable la corriente J^μ. Por consiguiente, y dado que $k_1 = 1$ tenemos que la variación de (7.10) esta dada por

$$\delta I = \int \left[-\frac{1}{c} J^\mu \delta \mathcal{A}_\mu - \frac{k_2}{c} \left(F^{\mu\nu} \delta F_{\mu\nu} + \delta F^{\mu\nu} F_{\mu\nu} \right) \right] d\Omega = 0$$

pero $F^{\mu\nu}\delta F_{\mu\nu} = \delta F^{\mu\nu} F_{\mu\nu}$ entonces

$$\delta I = \int \left[-\frac{1}{c} J^\mu \delta \mathcal{A}_\mu - 2\frac{k_2}{c} F^{\mu\nu} \delta F_{\mu\nu} \right] d\Omega = 0$$

pero de la definición de $F_{\mu\nu}$ dada en el capítulo 5, ecuación (5.44) tenemos que

$$\delta I = \int \left[-\frac{1}{c} J^\mu \delta \mathcal{A}_\mu - 2\frac{k_2}{c} F^{\mu\nu} \frac{\partial \delta \mathcal{A}_\nu}{\partial x^\mu} + 2\frac{k_2}{c} F^{\mu\nu} \frac{\partial \delta \mathcal{A}_\mu}{\partial x^\nu} \right] d\Omega = 0$$

ahora intercambiamos los índices μ y ν en el segúndo término del integrando y, además, tomamos en cuenta que $F^{\mu\nu} = -F^{\nu\mu}$ para llegar a que

$$\delta I = \int \left[-\frac{1}{c} J^\mu \delta \mathcal{A}_\mu + 4\frac{k_2}{c} F^{\mu\nu} \frac{\partial}{\partial x^\nu} \delta \mathcal{A}_\mu \right] d\Omega = 0 \tag{7.20}$$

si hacemos un razonamiento análogo al hecho en la ecuación (3.29) tenemos que

$$\frac{\partial}{\partial x^\nu} \left(F^{\mu\nu} \delta \mathcal{A}_\mu \right) = \frac{\partial}{\partial x^\nu} F^{\mu\nu} \delta \mathcal{A}_\mu + F^{\mu\nu} \frac{\partial}{\partial x^\nu} \delta \mathcal{A}_\mu$$

y sustituyendo en (7.20) concluimos que

$$\delta I = \int \left[-\frac{1}{c} J^\mu \delta \mathcal{A}_\mu - 4\frac{k_2}{c} \frac{\partial}{\partial x^\nu} F^{\mu\nu} \delta \mathcal{A}_\mu \right] d\Omega + 4\frac{k_2}{c} \int \frac{\partial}{\partial x^\nu} \left(F^{\mu\nu} \delta \mathcal{A}_\mu \right) d\Omega = 0 \tag{7.21}$$

Ahora apliquemos el teorema de la divergencia visto en el capítulo 3, ecuación (3.22) al segundo término del lado izquierdo de la ecuación (7.21) para obtener

$$\delta I = \int \left[-\frac{1}{c} J^\mu - 4\frac{k_2}{c} \frac{\partial}{\partial x^\nu} F^{\mu\nu} \right] \delta \mathcal{A}_\mu d\Omega + 4\frac{k_2}{c} \int\limits_{\partial\Omega} F^{\mu\nu} \delta \mathcal{A}_\mu dS_\nu = 0 \tag{7.22}$$

En la segunda integral de la parte izquierda de la ecuación anterior hay que introducir los valores correspondientes a los límites de integración. Los límites para las coordenadas son iguales a infinito, donde el campo es nulo. En cambio, en los límites de la integración respecto al tiempo, es decir, para los instantes inicial y final, la variación de los potenciales es nula, puesto que, de acuerdo con el principio de mínima acción, los potenciales en estos instantes son datos fijos. El segundo término de (7.22) es, por tanto, igual a cero y nos resulta que

$$\delta I = \int \left[-\frac{1}{c} J^\mu - 4\frac{k_2}{c}\frac{\partial}{\partial x^\nu} F^{\mu\nu} \right] \delta \mathcal{A}_\mu d\Omega = 0$$

y aplicando el lema fundamental del cálculo de variaciones que dice que las variaciones $\delta \mathcal{A}_\mu$ son cualesquiera entonces los coeficientes de las $\delta \mathcal{A}_\mu$ deben ser nulos, con lo que

$$-4\frac{k_2}{c}\frac{\partial}{\partial x^\nu} F^{\mu\nu} = \frac{1}{c} J^\mu$$

Si comparamos esta ecuación con la ecuación encontrada en el capítulo 5 veremos que

$$-4k_2 = \frac{1}{\mu_0}$$

o lo que es lo mismo

$$k_2 = -\frac{1}{4\mu_0} = -\frac{\epsilon_0 c^2}{4}$$

Por lo que el segundo par de ecuaciones de Maxwell se obtienen a partir de la ecuación

$$F^{\mu\nu}_{,\nu} = -\mu_0 J^\mu$$

Por lo tanto, la acción para partículas situadas en un campo electromagnético será

$$I = -\int mc^2\gamma^{-1} dt - \int \mathcal{A}_\mu J^\mu d\Omega - \frac{\epsilon_0 c^2}{4} \int F_{\mu\nu} F^{\mu\nu} d\Omega \tag{7.23}$$

7.3 LA ACCION PARA EL CAMPO GRAVITACIONAL

Para hallar las ecuaciones de movimiento que determinan al campo gravitatorio es necesario, en primer lugar, determinar la acción I_g de dicho campo. Las ecuaciones buscadas se obtienen entonces variando la suma de las acciones del campo y de las particulas materiales. Al igual que en el campo electromagnético, la acción I_g debe expresarse en forma de una integral invariante o escalar $\int S\sqrt{g}d\Omega$, como se vió también en (4.21), extendida a todo el espacio y a todos los valores de la coordenada temporal entre dos valores dados. Para determinar este escalar partiremos de que las ecuaciones del campo gravitatorio deben contener derivadas de los "potenciales" de orden no superior al segundo (de manera análoga a como ocurre en el caso del campo electromagnético). Dado que las ecuaciones del campo se obtienen variando la acción, es necesario que el integrando S contenga derivadas de las $g_{\mu\nu}$ de orden no superior a uno; por consiguiente S debe contener solamente el tensor $g_{\mu\nu}$ y las cantidades $\Gamma^\mu_{\nu\rho}$.

Sin embargo, es imposible construir un escalar a partir de las cantidades $g_{\mu\nu}$ y $\Gamma^\mu_{\nu\rho}$ únicamente. Esto se ve sin más que considerar que, mediante una elección adecuada del sistema de coordenadas, podemos siempre reducir a cero todas las cantidades $\Gamma^\mu_{\nu\rho}$ en un punto dado. Existe, sin embargo, el escalar R (la curvatura del espacio cuadridimensional) que si bien contiene, además de las $g_{\mu\nu}$ y sus primeras derivadas, las segundas derivadas de $g_{\mu\nu}$, es lineal respecto de éstas. Gracias a este carácter, la integral invariante

$$I_g = \int R\sqrt{g}d\Omega \tag{7.24}$$

se puede transformar en la integral de una expresión que no contiene derivadas segundas. Veamos esto:

De la expresión (6.67) tenemos que

$$R = g^{\mu\nu} R_{\mu\nu} = g^{\mu\nu}\left(\Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha}\right) - g^{\mu\nu}\left(\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} - \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\right)$$

o también lo podemos escribir como

$$R = R^* - L$$

donde

$$R^* = g^{\mu\nu}\left(\Gamma^{\alpha}_{\mu\alpha,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha}\right)$$

y

$$L = g^{\mu\nu}\left(\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta} - \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\right) \tag{7.25}$$

Como ya dijimos, I_g involucra segundas derivadas de $g^{\mu\nu}$, pero estas sólo están en el término R^*. Pero esto ocurre sólo linealmente, así que estas pueden ser removidas por una integración parcial. Tenemos que

$$R^*\sqrt{g} = \left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\alpha}\right)_{,\nu} - \left(\sqrt{g}g^{\mu\nu}\Gamma^{\sigma}_{\mu\nu}\right)_{,\sigma} - \left(\sqrt{g}g^{\mu\nu}\right)_{,\nu}\Gamma^{\sigma}_{\mu\sigma} - \left(\sqrt{g}g^{\mu\nu}\right)_{,\sigma}\Gamma^{\sigma}_{\mu\nu} \tag{7.26}$$

Los primeros dos términos son diferenciales exactas por lo que no contribuirán a I_g. Así que sólo nos quedan los dos últimos términos de (7.26). Con ayuda de (6.41) y (6.40) estos últimos dos términos se pueden escribir como

$$g^{\nu\beta}\Gamma^{\mu}_{\beta\nu}\Gamma^{\sigma}_{\mu\sigma}\sqrt{g} + \sqrt{g}\left(-2g^{\nu\beta}\Gamma^{\mu}_{\beta\sigma} + g^{\mu\nu}\Gamma^{\beta}_{\sigma\beta}\right)\Gamma^{\sigma}_{\mu\nu}$$

esto es justamente $2L\sqrt{g}$, como en (7.25). Así (7.24) estará dada por

$$I_g = \int L\sqrt{g}d\Omega \tag{7.27}$$

la cual involucra sólo a $g^{\mu\nu}$ y sus primeras derivadas. Esta es homogénea de grado dos en estas primeras derivadas.

Tomaremos a $L\sqrt{g}$ como la densidad lagrangiana de la acción para el campo gravitacional (esto salvo una constante que determinaremos posteriormente).

Podemos ver a las $g^{\mu\nu}$ como las coordenadas dinámicas y sus derivadas temporales como las velocidades. Además vemos que la lagrangiana es cuadrática (no homogénea) en las velocidades, como ocurre generalmente en la dinámica ordinaria.

Hagamos la variación de I_g e igualémosla a cero, tal que

$$\delta I_g = \delta\int L\sqrt{g}d\Omega = 0$$

o, dado que la variación se hace sobre las $g^{\mu\nu}$

$$\delta I_g = \int \delta\left(L\sqrt{g}\right)d\Omega = 0 \tag{7.28}$$

Así

$$\delta\left(L\sqrt{g}\right) = \delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}\right) - \delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\right) \tag{7.29}$$

Calculemos el primer término de la ecuación anterior

$$\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}\right) = \Gamma^{\alpha}_{\mu\nu}\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\beta}_{\alpha\beta}\right) + \sqrt{g}g^{\mu\nu}\Gamma^{\beta}_{\alpha\beta}\delta\Gamma^{\alpha}_{\mu\nu}$$

en el primer término del lado derecho de la ecuación anterior, podemos sustituir lo que está entre paréntesis de acuerdo a la ecuación (6.32) y el segundo término sustituirlo por

$$\sqrt{g}g^{\mu\nu}\delta\Gamma^{\alpha}_{\mu\nu} = \delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\nu}\right) - \Gamma^{\alpha}_{\mu\nu}\delta\left(\sqrt{g}g^{\mu\nu}\right) \tag{7.30}$$

de tal manera que

$$\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}\right) = \Gamma^{\alpha}_{\mu\nu}\delta\left(g^{\mu\nu}\sqrt{g}_{,\alpha}\right) + \Gamma^{\beta}_{\alpha\beta}\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\nu}\right) - \Gamma^{\beta}_{\alpha\beta}\Gamma^{\alpha}_{\mu\nu}\delta\left(\sqrt{g}g^{\mu\nu}\right)$$

el segundo término del lado derecho se sustituye de acuerdo a (6.41), por tanto

$$\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\nu}\Gamma^{\beta}_{\alpha\beta}\right) = \Gamma^{\alpha}_{\mu\nu}\delta\left(g^{\mu\nu}\sqrt{g}_{,\alpha}\right) - \Gamma^{\beta}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right)_{,\nu} - \Gamma^{\beta}_{\alpha\beta}\Gamma^{\alpha}_{\mu\nu}\delta\left(\sqrt{g}g^{\mu\nu}\right) \tag{7.31}$$

Ahora, el segundo término de (7.29) queda como

$$\begin{aligned}\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\right) &= \sqrt{g}g^{\mu\nu}\left(\Gamma^{\alpha}_{\mu\beta}\delta\Gamma^{\beta}_{\nu\alpha} + \Gamma^{\beta}_{\nu\alpha}\delta\Gamma^{\alpha}_{\mu\beta}\right) + \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\delta\sqrt{g}g^{\mu\nu}\\ &= 2\sqrt{g}g^{\mu\nu}\Gamma^{\beta}_{\nu\alpha}\delta\Gamma^{\alpha}_{\mu\beta} + \Gamma^{\beta}_{\nu\alpha}\Gamma^{\alpha}_{\mu\beta}\delta\sqrt{g}g^{\mu\nu}\\ &= \Gamma^{\beta}_{\nu\alpha}\left(2\sqrt{g}g^{\mu\nu}\delta\Gamma^{\alpha}_{\mu\beta} + \Gamma^{\alpha}_{\mu\beta}\delta\sqrt{g}g^{\mu\nu}\right)\end{aligned}$$

haciendo uso nuevamente de la relación (7.30) con los índices adecuados tenemos que el primer término del lado derecho de la última ecuación queda como

$$\begin{aligned}\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\right) &= \Gamma^{\beta}_{\nu\alpha}\left[2\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\right) - 2\Gamma^{\alpha}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right) + \Gamma^{\alpha}_{\mu\beta}\delta\sqrt{g}g^{\mu\nu}\right]\\ &= \Gamma^{\beta}_{\nu\alpha}\left[2\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\right) - \Gamma^{\alpha}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right)\right]\end{aligned}$$

y sustituyendo la ecuación (6.38), con los adecuados cambios de índices mudos, en el primer término del lado derecho de la ecuación anterior, tenemos que

$$\delta\left(\sqrt{g}g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\right) = -\Gamma^{\beta}_{\nu\alpha}\delta\left(\sqrt{g}g^{\nu\alpha}_{,\beta}\right) - \Gamma^{\beta}_{\nu\alpha}\Gamma^{\alpha}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right) \tag{7.32}$$

Ahora ya podemos calcular δI_g, sustituyendo (7.31) y (7.32) en (7.29) y cambiando algunos índices mudos,

$$\begin{aligned}\delta I_g = \int \Big\{&\Gamma^{\alpha}_{\mu\nu}\delta\left(g^{\mu\nu}\sqrt{g}_{,\alpha}\right) - \Gamma^{\beta}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right)_{,\nu} - \Gamma^{\beta}_{\alpha\beta}\Gamma^{\alpha}_{\mu\nu}\delta\left(\sqrt{g}g^{\mu\nu}\right)\\ &+\Gamma^{\beta}_{\nu\alpha}\delta\left(\sqrt{g}g^{\nu\alpha}_{,\alpha}\right) + \Gamma^{\beta}_{\nu\alpha}\Gamma^{\alpha}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right)\Big\}\, d\Omega = 0\end{aligned}$$

o lo que es lo mismo

$$\begin{aligned}\delta I_g = \int \Big\{&\Gamma^{\alpha}_{\mu\nu}\delta\left(g^{\mu\nu}\sqrt{g}\right)_{,\alpha} - \Gamma^{\beta}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right)_{,\nu}\\ &+\left[\Gamma^{\beta}_{\nu\alpha}\Gamma^{\alpha}_{\mu\beta} - \Gamma^{\beta}_{\alpha\beta}\Gamma^{\alpha}_{\mu\nu}\right]\delta\left(\sqrt{g}g^{\mu\nu}\right)\Big\}\, d\Omega = 0\end{aligned}$$

Los dos primeros términos de la integral anterior, se pueden expresar como sigue

$$\Gamma^{\alpha}_{\mu\nu}\delta\left(g^{\mu\nu}\sqrt{g}\right)_{,\alpha} = \left[\Gamma^{\alpha}_{\mu\nu}\delta\left(g^{\mu\nu}\sqrt{g}\right)\right]_{,\alpha} - \Gamma^{\alpha}_{\mu\nu,\alpha}\delta\left(g^{\mu\nu}\sqrt{g}\right)$$

y

$$\Gamma^{\beta}_{\mu\beta}\delta\left(\sqrt{g}g^{\mu\nu}\right)_{,\nu} = \left[\Gamma^{\beta}_{\mu\beta}\delta\left(g^{\mu\nu}\sqrt{g}\right)\right]_{,\nu} - \Gamma^{\beta}_{\mu\beta,\nu}\delta\left(g^{\mu\nu}\sqrt{g}\right)$$

entonces en la δI_g tenemos otros dos términos los cuales son derivadas totales, y dado que estos no contribuyen a las ecuaciones de movimiento, entonces

$$\begin{aligned}\delta I_g = \int \Big\{ & -\Gamma^{\alpha}_{\mu\nu,\alpha}\delta\left(g^{\mu\nu}\sqrt{g}\right) + \Gamma^{\beta}_{\mu\beta,\nu}\delta\left(g^{\mu\nu}\sqrt{g}\right) \\ & + \left[\Gamma^{\beta}_{\nu\alpha}\Gamma^{\alpha}_{\mu\beta} - \Gamma^{\beta}_{\alpha\beta}\Gamma^{\alpha}_{\mu\nu}\right]\delta\left(\sqrt{g}g^{\mu\nu}\right)\Big\} d\Omega = 0\end{aligned}$$

o lo que es lo mismo

$$\delta I_g = \int \left(\Gamma^{\beta}_{\mu\beta,\nu} - \Gamma^{\alpha}_{\mu\nu,\alpha} + \Gamma^{\beta}_{\nu\alpha}\Gamma^{\alpha}_{\mu\beta} - \Gamma^{\beta}_{\alpha\beta}\Gamma^{\alpha}_{\mu\nu}\right)\delta\left(\sqrt{g}g^{\mu\nu}\right) d\Omega = 0$$

Por lo tanto podemos escribir a la variación de la acción gravitacional, según la ecuación (6.67), como

$$\delta I_g = \int R_{\mu\nu}\delta\left(\sqrt{g}g^{\mu\nu}\right) d\Omega = 0 \tag{7.33}$$

Una manera de dar las ecuacones de movimiento para el campo gravitacional es diciendo que $\delta g^{\mu\nu}$ es arbitraria, entonces las cantidades $\delta\left(\sqrt{g}g^{\mu\nu}\right)$ son también arbitrarias e independientes, así lo que se tiene que anular para que la integral (7.33) sea cero es que $R_{\mu\nu} = 0$ que la Ley de Einstein, como ya lo vimos en (6.72).

Desarrollemos la variación $\delta\left(\sqrt{g}g^{\mu\nu}\right)$, así

$$\delta\left(\sqrt{g}g^{\mu\nu}\right) = \sqrt{g}\delta g^{\mu\nu} + g^{\mu\nu}\delta\sqrt{g} \tag{7.34}$$

podemos deducir, por el mismo método usado para deducir la ecuación (6.37) que

$$\delta g^{\mu\nu} = -g^{\mu\alpha}g^{\nu\beta}\delta g_{\alpha\beta} \tag{7.35}$$

y también por el mismo método con el que llegamos a la expresión (6.39) podemos obtener que

$$\delta\sqrt{g} = \frac{1}{2}\sqrt{g}g^{\alpha\beta}\delta g_{\alpha\beta} \tag{7.36}$$

Así, sustituyendo (7.35) y (7.36) en (7.34) tenemos que

$$\delta\left(\sqrt{g}g^{\mu\nu}\right) = -\left(g^{\mu\alpha}g^{\nu\beta} - \frac{1}{2}g^{\mu\nu}g^{\alpha\beta}\right)\sqrt{g}\delta g_{\alpha\beta}$$

Por tanto, podemos escribir a (7.33) como

$$\begin{aligned}\delta I_g &= -\int R_{\mu\nu}\left(g^{\mu\alpha}g^{\nu\beta} - \tfrac{1}{2}g^{\mu\nu}g^{\alpha\beta}\right)\sqrt{g}\delta g_{\alpha\beta} d\Omega \\ &= \int \left(R^{\alpha\beta} - \tfrac{1}{2}g^{\alpha\beta}R\right)\sqrt{g}\delta g_{\alpha\beta} d\Omega = 0\end{aligned}$$

y, otra vez, como las variaciones $\delta g_{\alpha\beta}$ son arbitrarias, entonces

$$R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}R = 0 \tag{7.37}$$

que es la forma alterna de escribir la Ley de Einstein para el espacio vació.

7.4 LA ACCION PARA LA MATERIA

Ahora consideraremos la acción para una distribución continua de materia. Para un flujo continuo de materia tenemos el vector velocidad u^μ (con un factor desconocido) en cada punto. Definamos un tensor contravariante de densidad p^μ, en la dirección de u^μ, el cual determina tanto la cantidad del flujo como su velocidad de acuerdo a las relaciones siguientes:

$p^0dx^1dx^2dx^3$, es la cantidad de materia dentro del elemento de volumen $dx^1dx^2dx^3$ a un cierto tiempo y

$p^1dx^0dx^2dx^3$, la cual es la cantidad de fluido a través del elemeto de superficie dx^2dx^3 durante el intervalo de tiempo dx^0. Además supondremos que la materia se conserva , así

$$p^\mu_{,\mu} = 0 \tag{7.38}$$

Supongamos que cada elemento de materia es desplazado de z^μ a $z^\mu + b^\mu$ con b^μ pequeño. Determinaremos el cambio resultante en p^μ en un punto dado x.

Tomemos primero el caso de $b^0 = 0$. El cambio en la cantidad de materia dentro de un cierto volumen tridimensional V es igual a menos la cantidad desplazada a través de la frontera de V

$$\delta \int\limits_V p^0dx^1dx^2dx^3 = -\int p^0b^r dS^r$$

con $r = 1, 2, 3$, donde dS^r denota un elemento de superficie de la frontera de V. Podemos transformar el lado derecho a una integral de volumen por medio del teorema de Gauss (5.20) y encontramos que

$$\delta \int\limits_V p^0dx^1dx^2dx^3 = -\int\limits_V \left(p^0b^r\right)_{,r} dV$$

o lo que es lo mismo

$$\delta p^0 = -\left(p^0b^r\right)_{,r} \tag{7.39}$$

Podemos generalizar este resultado al caso en que $b^0 \neq 0$. Podemos, también, hacer uso de la condición de que si b^μ es proporcional a p^μ, cada elemento de materia es desplazada a lo largo de su línea de universo y entonces no hay cambio en p^μ. La generalización de (7.39) es evidentemente

$$\delta p^0 = \left(p^rb^0 - p^0b^r\right)_{,r}$$

dado que obtenemos (7.39) cuando $b^0 = 0$ y nos resulta $\delta p^0 = 0$ cuando b^μ es proporcional a p^μ. Hay una ecuación correpondiente para las otras componentes de p^μ, esta es

$$\delta p^\mu = \left(p^\nu b^\mu - p^\mu b^\nu\right)_{,\nu} \tag{7.40}$$

Para descibir un flujo continuo de materia, las cantidades p^μ son las variables básicas a usarse en la función de acción material. Ellas pueden ser variadas de acuerdo a la ecuación (7.40), y, entonces, después de algunas integraciones parciales, podemos hacer los coeficientes de cada b^μ iguales a cero. Estas nos darán como resultado las ecuaciones de movimiento para la materia.

Como vimos en el capítulo 4, ver ecuación (4.53) , la acción para una partícula aislada de masa m es

$$I = -m\int d\tau \tag{7.41}$$

Podemos obtener la acción para una distribución de materia de (7.41) reemplazando m por $p^0 dx^1 dx^2 dx^3$ e integrando, de tal manera que

$$I_m = -\int p^0 dx^1 dx^2 dx^3 d\tau$$

Si comparamos la ecuación (7.38) de la conservación de la materia y la comparamos con la ecuación de la conservación de la materia vista en el capítulo anterior, ecuación (6.80), podemos deducir que

$$p^\mu = \rho v^\mu \sqrt{g} \tag{7.42}$$

donde ρ es un escalar que determina la densidad y v^μ es el tensor definido previamente u^μ normalizado a la longitud 1. Así

$$I_m = -\int \rho v^0 \sqrt{g} dx^1 dx^2 dx^3$$

o también puede escribirse como

$$I_m = -\int \rho \sqrt{g} d\Omega \tag{7.43}$$

donde $v^0 d\tau = dx^0$.

Esta forma para la acción material no es conveniente para aplicarle variaciones, dado que ρ y v^μ no son variables independientes. De la ecuación (7.42) podemos ver que

$$(p^\mu p_\mu)^{\frac{1}{2}} = \rho \sqrt{g}$$

Entonces la acción (7.43) se podrá escribir como

$$I_m = -\int (p^\mu p_\mu)^{\frac{1}{2}} d\Omega \tag{7.44}$$

Ahora calculemos la variación de la acción gravitacional

$$\begin{aligned}\delta I_m = -\int \delta (p^\mu p_\mu)^{\frac{1}{2}} d\Omega &= -\int \left[\tfrac{1}{2} (p^\sigma p_\sigma)^{-\frac{1}{2}} (p^\mu p^\nu \delta g_{\mu\nu} + 2 p_\mu \delta p^\mu) d\Omega\right] \\ &= -\int \tfrac{1}{2} \rho v^\mu v^\nu \sqrt{g} \delta g_{\mu\nu} d\Omega - \int v_\mu \delta p^\mu d\Omega\end{aligned}$$

Por lo tanto podemos escribir a la variación de la acción material como:

$$\delta I_m = -\int \frac{1}{2} \rho v^\mu v^\nu \delta g_{\mu\nu} \sqrt{g} d\Omega - \int v_\mu \delta p^\mu d\Omega =$$

Veamos que resulta del último término de la ecuación anterior, si sustituimos la ecuación (7.40), tenemos que

$$\begin{aligned}-\int v_\mu \delta p^\mu d\Omega &= -\int v_\mu (p^\nu b^\mu - p^\mu b^\nu)_{,\nu} d\Omega \\ &= \int (v_{\mu,\nu} - v_{\nu,\mu}) p^\nu b^\mu d\Omega \\ &= \int (v_{\mu;\nu} - v_{\nu;\mu}) \rho v^\nu b^\mu \sqrt{g} d\Omega\end{aligned}$$

Si tomamos en cuenta la ecuación (6.78) la ecuación anterior resulta ser

$$-\int v_\mu \delta p^\mu d\Omega = \int v_{\mu;\nu} \rho v^\nu b^\mu \sqrt{g} d\Omega \tag{7.45}$$

Si igualamos a cero el coeficiente de b^μ ontenemos la ecuación de la geodésica (6.83). Por lo que si la materia sigue la geodésica, la variación de la acción material queda entonces como

$$\delta I_m = -\int \frac{1}{2}\rho v^\mu v^\nu \delta g_{\mu\nu}\sqrt{g}d\Omega \tag{7.46}$$

Si tomamos a la acción I_1 como

$$I_1 = I_g + I_m$$

y la variamos respecto de la métrica obtenemos que $\delta I_1 = \delta I_g + \delta I_m$, entonces

$$\begin{aligned}\delta I_1 &= \int \left(R_{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}R\right)\sqrt{g}\delta g_{\mu\nu}d\Omega - \int \tfrac{1}{2}\rho v^\mu v^\nu \delta g_{\mu\nu}\sqrt{g}d\Omega = 0\\ &= \int \left[\left(R_{\mu\nu} - \tfrac{1}{2}g^{\mu\nu}R\right) - \tfrac{1}{2}\rho v^\mu v^\nu\right]\sqrt{g}\delta g_{\mu\nu}d\Omega = 0\end{aligned}$$

y como las variaciones $\delta g_{\mu\nu}$ son arbitrarias entonces las ecuaciones de movimiento para el campo gravitacional con presencia de material quedan como

$$\left(R_{\mu\nu} - \frac{1}{2}g^{\mu\nu}R\right) = \frac{1}{2}\rho v^\mu v^\nu$$

si vemos la ecuación (6.82) obtendrmos la ecuación (6.85) salvo un factor de -16π por lo que a la acción gravitacional tendríamos que multiplicarla por dicgo factor de tal manera que la acción gravitacional I_g queda como

$$I_g = (16\pi)^{-1}\int R\sqrt{g}d\Omega \tag{7.47}$$

7.5 ACCION COSMOLOGICA

Como lo hizo Einstein, consideraremos una generalización de sus ecuaciones de campo para el espacio vacío de tal manera que

$$R_{\mu\nu} = \lambda g_{\mu\nu} \tag{7.48}$$

donde λ es una constante. Esta es una ecuación tensorial, así que es posible admitirla como una ley de la naturaleza. Podemos considerar la observación del sistema solar con una muy buena aproximación cuando λ sea suficientemente pequeño. Dado que $R_{\mu\nu}$ contiene segundas derivadas de $g_{\mu\nu}$, λ tendrá dimensiones de (distancia)$^{-2}$. Entonces cuando λ es pequeño, esta distancia será muy grande, esta es una distancia cosmológica, del orden del radio del universo, por lo que a λ se le conoce como constante cosmológica.

Este término extra es importante para las teorías cosmológicas, pero tiene un efecto despreciable en la física de objetos cercanos. Por lo que en la teoría de campos lo tomaremos solamente como un término extra de la acción gravitacional, este término será

$$I_\lambda = G\int \sqrt{g}d\Omega \tag{7.49}$$

con una constante arbitraria G. La variación de (7.54) estará dada por

$$\delta I_\lambda = G\int \delta\sqrt{g}d\Omega$$

que por la ecuación (7.36)

$$\delta I_\lambda = G\int \frac{1}{2}g^{\mu\nu}\delta g_{\mu\nu}\sqrt{g}d\Omega \tag{7.50}$$

Así que para la acción $I_2 = I_g + I_\lambda$ tenemos la variación siguiente

$$\delta I_g = (16\pi)^{-1} \int \left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \right) \sqrt{g} \delta g_{\mu\nu} d\Omega + G \int \frac{1}{2} g^{\mu\nu} \delta g_{\mu\nu} \sqrt{g} d\Omega = 0$$

o lo que es lo mismo

$$\delta I_g = (16\pi)^{-1} \int \left[\left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \right) + G \frac{1}{2} g^{\mu\nu} \right] \sqrt{g} \delta g_{\mu\nu} d\Omega = 0$$

y dado que $\delta g_{\mu\nu}$ es arbitrario obtenemos que

$$(16\pi)^{-1} \left(R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R \right) = -G \frac{1}{2} g^{\mu\nu}$$

Así, para obtener la ecuación (7.48) tenemos que $\lambda = -8\pi G$.

7.6 ACCION ELECTROMAGNETICA

La acción construida del campo electromagnético anteriormente tenía la información a cerca del espacio sólo en el elemento de volumen, ahora, cuando la información acerca del espacio la da el tensor métrico, la acción para algún escalar esta dada como en (4.21).

La acción I_c se construye a partir del escalar $F_{\mu\nu} F^{\mu\nu}$ entonces la acción para el campo electromagnético, en general esta dado por

$$I_c = -\frac{\epsilon_0 c^2}{4} \int F_{\rho\sigma} F^{\rho\sigma} \sqrt{g} d\Omega \tag{7.51}$$

Entonces la variación de esta acción se puede hacer respecto al potencial electromagnético o con respecto a la métrica $g_{\mu\nu}$.

La variación con respecto al potencial electromagnético no da como resultado el segundo par de ecuaciones de Maxwell para el vacío.

Hagamos la variación de la acción I_c con respecto al tensor métrico manteniendo al potencial electromagnético constante, si asociamos a $\sqrt{g}$ con el tensor $F^{\rho\sigma}$, entonces $F_{\rho\sigma}$ permanecerá constante pero $F^{\rho\sigma}\sqrt{g}$ no lo será. Así que

$$\begin{aligned} \delta\left(F_{\rho\sigma} F^{\rho\sigma} \sqrt{g}\right) &= F_{\rho\sigma} F^{\rho\sigma} \delta\sqrt{g} + F_{\rho\sigma} F_{\alpha\beta} \sqrt{g} \delta\left(g^{\rho\alpha} g^{\sigma\beta}\right) \\ &= \tfrac{1}{2} F_{\rho\sigma} F^{\rho\sigma} g^{\mu\nu} \sqrt{g} \delta g_{\mu\nu} - 2 F_{\rho\sigma} F_{\alpha\beta} \sqrt{g} g^{\rho\mu} g^{\alpha\nu} g^{\sigma\beta} \delta g_{\mu\nu} \end{aligned}$$

con la ayuda de las ecuaciones (7.36) y (7.35). Lo anterioe podemos escribirlo como

$$\begin{aligned} \delta\left(F_{\rho\sigma} F^{\rho\sigma} \sqrt{g}\right) &= \left(\tfrac{1}{2} F_{\rho\sigma} F^{\rho\sigma} g^{\mu\nu} - 2 F^{\mu}_{\sigma} F^{\nu\sigma}\right) \sqrt{g} \delta g_{\mu\nu} \\ &= E^{\mu\nu} \sqrt{g} \delta g_{\mu\nu} \end{aligned} \tag{7.52}$$

donde $E^{\mu\nu}$ es el tensor de energía-momento del campo electromagnético y está definido por

$$E^{\mu\nu} = \frac{1}{4} F_{\rho\sigma} F^{\rho\sigma} g^{\mu\nu} - F^{\mu}_{\sigma} F^{\nu\sigma} \tag{7.53}$$

Este tensor ya lo vimos en el capítulo 5, en la ecuación (5.60).

Si tomamos la acción $I_3 = I_g + I_\lambda + I_c$ y la variamos respecto de $g_{\mu\nu}$ tenemos que las ecuaciones de movimiento serán

$$R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R = \lambda g^{\mu\nu} - \frac{\epsilon_0 c^2}{4} E^{\mu\nu} \tag{7.54}$$

7.7 ACCION ELECTROMAGNETICA Y GRAVITACIONAL

Existe una acción que involucra tanto al campo electromagnético como al campo gravitacional [22] la cuál está dada por

$$S = \int \sqrt{g} d\Omega \left[\frac{1}{16\pi} (R + 2\lambda) + L \right] \tag{7.55}$$

donde R es la curvatura escalar y λ es la constante cosmológica. La densidad lagrangiana L es la densidad lagrangiana para el campo electromagnético.

Si esta acción la podemos variar con respecto de $g_{\mu\nu}$ y A_μ podemos obtener las ecuaciones de campo electromagnético de Maxwell para el vacío y las ecuaciones de campo gravitacional, ecuaciones de Einstein (7.54).

Chapter 8
CONCLUSIONES

Un primer resultado de relevencia, en el presente trabajo, es haber dado una justificación heurística del principio de Hamilton, ya que regularmente se deja de lado por considerarlo demasiado elemental o por que no se considera necesario. La justificación que se presenta es importante ya que se le da un enfoque didáctico y se coloca como la base sobre la cual se construye la formulación lagrangiana.

La formulación que hemos utilizado en todo el trabajo implica que tanto la lagrangiana (o densidad lagrangiana) que describe a cierto fenómeno físico, como la acción en general, sean invariantes bajo transformaciones de Lorentz, esto con el objeto de que las ecuaciones de movimiento que surgen a partir de ellas sean las mismas tanto para un observador que esté en un cierto sistema de referencia como para otro observador en cualquier otro sistema de referencia, sean o no inerciales.

Además, dado un problema físico, lo primero que hay que hacer es identificar cuales son sus variables independientes (o dinámicas) y a partir de ellas construir todos los invariantes posibles para posteriormente determinar con cuales de ellos construir su acción, de tal manera que las ecuaciones de movimiento que resulten tengan algún significado físico.

Establecemos que de las relaciones cuantitativas experimentales que describen fenómenos, tanto eléctricos como magnéticos o ambos, llegamos a una formulación de la electrodinámica que sea invariante de Lorentz, es decir una formulación covariante. Por otro lado, a partir de principios variacionales llegamos a las mismas expresiones, lo que nos garantiza la covariancia de las ecuaciones de movimiento de la electrodinámica, ya que, tanto la densidad lagrangiana que se proponga para describir estos fenómenos como la acción, son invariantes de Lorentz.

En el presente trabajo se tiene como base el espacio-tiempo de Minkowski (en el caso de espacio plano), y más generalmente un espacio-tiempo curvo. Existen otras formulaciones donde la funcional de los campos vectoriales, que están definidos en tres dimensiones, se define por medio de una convolución temporal [29], para de esta manera considerar, tanto al espacio como al tiempo en el planteamiento de la acción correspondiente. Obviamente los resultados que se obtienen en una y otra formulación han de ser equivalentes.

Esto no es todo, es posible continuar un trabajo en base a este, es decir, hasta aquí sólo hemos dado la formulación lagrangiana de la teoría de la relatividad general, pero, ?'Porqué no hacer lo que se hace en mecánica? Es decir, una formulación hamiltoniana de la teoría de la relatividad. La construcción clásica de la teoría hamiltoniana a partir de la lagrangiana implica construir los momentos canónicos de las variables dinámicas, es decir las 10 componentes del tensor métrico, tomando la derivada parcial de la densidad lagrangiana respecto de las velocidades. Pero ?'cómo definimos tales velocidades? Para hacer eso, nos damos cuenta que es necesario privilegiar a alguna de las coordenadas como el tiempo para poder definir velocidades. Esto rompe la covariancia de las ecuaciones de Einstein, pero además está cortando el espacio-tiempo en rebanadas de $x^0 = ct$. La complicación algebraica a la que lleva tal procedimiento es enorme. podemos darnos una idea de dicha complejidad, con el hecho de que P.A.M. Dirac dedicara más de una década en tal intento sin llegar a resultados satisfactorios. Por lo que para hacer esto es necesario profundizar en las bases matemáticas y, tal vez, algún día podamos lograr esto. Pero no empezaríamos de cero, existe un trabajo [4] en que da la introducción a este formalismo.

Chapter 9

REFERENCIAS

1. V. I. Arnold, et.al., **Mathematical methods of classical mechanics**. Springer-Verlag, New York, 1980.
2. Barut A.O., **Electrodynamics and Classical Theory of Fields & particles**, Dover New York,1964.
3. Byalinicki-Birula I, **Quantum Electrodynamics**, 1a. ed. Pergamon Press, Poland, 1975.
4. Corichi, A. y Núñez, Dario, Rev. Mex. Fís. **37**(4), 1991.
5. Dirac, P.A.M, **General Theory of Relativity**, John Wiley & Sons., N.Y., 1975.
6. Eisenhart, **An introduction to Differential Geometry With use of the Absolut Calculus**, Princeton University Press, USA, 1927.
7. Einstein, A. **El significado de la relatividad**, Origen/Planeta, México, 1985.
8. Feynmann, Richard P., et.al. **Física** , Vol.II, Addison-Wesley, 1990.
9. French, A.P., **Relatividad Especial**, Reverté, España, 1991.
10. H. Goldstein, **Mecánica Clásica**. Reverté, España, 1987.
11. Carlos Graeff F., **Apuntes de Relatividad General**. CCH,1990
12. Hecht-Zajac, **Optica**. Addison Wesley, USA, 1986
13. Jackson J.D. , **Classical Electrodynamics**, 2a. ed. Wiley, New York ,1975.
14. David C. Kay, **Cálculo tensorial**, Serie Schaum, Mc. Graw-Hill, España, 1989.
15. I. R. Kenyon, **General Relativity**. Oxford, 1990.
16. L.Landau y E.M. Lifshitz, **Mecánica**. Reverté, España, 1991.
17. L.Landau y E.M. Lifshitz, **Teoría Clásica de los Campos**, Física Teórica Vol. 2, Reverté, España, 1992.
18. Logan, **Invariant Variational Principles**, 1a ed. Academic Press, 1977.
19. Lovelock, David y Rund, Hanno, **Tensors, Differential Forms, and Variational Principles**, 1a. ed. Wiley, New York, 1975.
20. Jerry B. Marion, **Dinámica Clásica de las partículas y sistemas**. Reverté, España, 1991.
21. Marsden, J.E. & Tromba, A.J., **Cálculo Vectorial**, 3ra. ed.,Addison-Wesley, 1994.
22. M. Przanowski, "Mystery of Duality Rotation", **Proc. of the Internacional Conference on Aspects of General Relativity and Mathematical Physics**, CINVESTAV, México,1993, pp. 123-146.
23. Reitz, Milford, Christy, **Foundations of Electromagnetic Theory**, 3^{th}ed. Adisson-Wesley, Mass., 1979.
24. H. Rund, **The Hamilton-Jacobi theory in the calculus of variations**. D. Van Nostrand Co., Great Britain, 1966.
25. Hans Sagan, **Introduction to tha calculus of variations**, Dover, USA, 1992.
26. Schwinger, Julian. **A path to Quantum Electrodynamics**, Physics Today. Febrero 1983 pp. 42-48.

27. Shahen, Hacyan, Rev. Mex. Fís. **28**(3), 1982.
28. Shahen Hacyan, **Relatividad especial para estudiantes de física**, F.C.E. y E.C.U., México, 1995.
29. Tao, L.N, Journal of Mathematical Physics, Vol. 7, No. 3, 1966.
30. Ya. P.Terletskii, **Statistical physics**. North Holland, 1971.
31. Kopczynski, W. & Trautman, A. **Spacetime and Gravitation**, Wiley, Poland, 1992.
32. Weinberg, Steven. **Gravitation and Cosmology**, John Wiley, 1972.
33. Robert Weinstock, **Calculus of Variations, with applications to physics & engineering**, Dover, NY,1974.
34. W.Yourgrau, S. Mandelstam, **Variational Principles in Dynamics and Quantum Theory**. Dover, New York, 1968.