Para Mercurio las medidas astronómicas indican $u_a = 1.43 \cdot 10^{-11} m$ y $u_p = 2.19 \cdot 10^{-11} m$ y el r_0 correspondiente al Sol (la masa que genera la métrica de Schwarzschild) es, como habíamos mencionado, $2.96 \cdot 10^3 m$. Calculando numéricamente la integral (con ayuda de un ordenador), se obtiene $\Delta = 5.04 \cdot 10^{-7}$. Teniendo en cuenta que Mercurio tarda 0.24 años en dar una vuelta alrededor del Sol, cada siglo habrá dado 416.67 vueltas y la variación del ángulo se multiplicará por este número, siendo

$$\text{Variación secular} = 2.10 \cdot 10^{-4} rad = 43.32''$$

lo que coincide con gran precisión con la cantidad observada experimentalmente. En realidad no es necesario utilizar el ordenador para estimar Δ, con aproximaciones de Taylor sencillas se puede deducir

$$\Delta \approx 3\pi r_0 (u_a + u_p)/2.$$

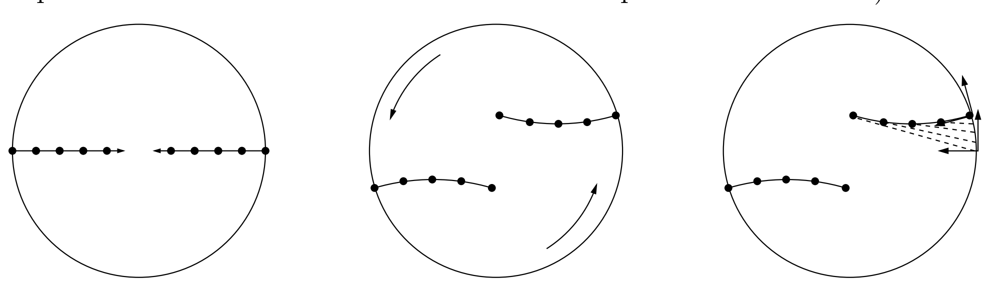
En particular el efecto está en relación inversa con la distancia al Sol.

3.3. Geodésicas y conexiones

Para estudiar la variación de un vector no sólo debemos tener en cuenta la variación de sus componentes sino la del propio sistema de coordenadas. Consideremos por ejemplo en $\mathbb{R}^2$ una base $B = \{\vec{e}_1, \vec{e}_2\}$ que va rotando con el tiempo, entonces la derivada del vector $\vec{v} = \big(a(t), b(t)\big) = a(t)\vec{e}_1 + b(t)\vec{e}_2$ no es $\big(a'(t), b'(t)\big)$ sino $a'(t)\vec{e}_1 + b'(t)\vec{e}_2 + a(t)\vec{e}_1{}' + b(t)\vec{e}_2{}'$. Como B es una base, existen Γ_j^k con $j, k \in \{1, 2\}$ tales que $\vec{e}_j{}'(t) = \Gamma_j^1 \vec{e}_1(t) + \Gamma_j^2 \vec{e}_2(t)$ y la derivada en la base B es $\big(a', b'\big) + \big(a\Gamma_1^1 + b\Gamma_2^1, a\Gamma_1^2 + b\Gamma_2^2\big)$.

Con esta definición la derivada es *absoluta* en el sentido de que se transforma tensorialmente (como un vector, el vector aceleración). En mecánica, estos términos que se añaden a la derivada de las componentes en sistemas que rotan da lugar a la llamada *aceleración de Coriolis*. Una de las demostraciones más claras de su existencia es considerar en una plataforma circular dos chorros de agua que apuntan al centro desde puntos diametralmente opuestos. Al hacer girar la plataforma, los chorros se desvían dejando de converger en el centro.

Para un observador en el borde, la velocidad del chorro sigue la dirección del eje radial y es constante, por ello la desviación es sorprendente desde su punto de vista (lo mismo que lo es la aceleración de Coriolis desde la superficie de la Tierra).



Un observador exterior nota que hay una velocidad tangencial inicial imprimida por el giro y la trayectoria de cada gota es la recta que debería ser (no exactamente radial por efecto de la velocidad tangencial) pero cada una parte de puntos diferentes por la variación del sistema de coordenadas.

Si deseamos derivar campos de vectores en una variedad, la dependencia será en el punto más que en un parámetro t y debemos especificar en la dirección en la que movernos. Si el campo es $V = V^k\partial_k$ y elegimos la dirección ∂_i, entonces un razonamiento como el anterior sugiere que esta "derivada direccional absoluta" debiera tener como componente k-ésima

$$\frac{\partial V^k}{\partial x^i} + \Gamma^k_{ij}V^j$$

para ciertos Γ^k_{ij} que indican la variación del sistema de coordenadas $\phi = (x^1, x^2, \ldots, x^n)$. La elección de unos Γ^k_{ij} que den resultados coherentes en cartas compatibles es lo que da lugar a una conexión afín: una manera absoluta (tensorial) de derivar en una variedad. Sin embargo una definición a través de estas funciones sería demasiado fea (aunque posible [Spi79, p.233]) y se prefiere habitualmente otra más sintética y operativa pero menos transparente.

Definición: Una *conexión afín* en una variedad M es una aplicación que asigna a cada par de campos de vectores X, Y un tercero, denotado mediante $\nabla_X Y$, tal que se verifican las siguientes propiedades

$$\begin{array}{ll} a) & \nabla_{fX+gY}Z = f\nabla_X Z + g\nabla_Y Z \\ b) & \nabla_X(Y+Z) = \nabla_X Y + g\nabla_X Z \\ c) & \nabla_X(fY) = f\nabla_X Y + X(f)Y \end{array}$$

donde f y g son funciones arbitrarias $M \longrightarrow \mathbb{R}$.

Tomando f muy concentrada alrededor de un punto y $g = 0$, se deduce de las propiedades anteriores que $\nabla_X Y$ en cada punto p sólo depende de $X\big|_p$ y del comportamiento local de Y alrededor de p. La terminología "conexión" proviene de T. Levi-Civita que en un trabajo pionero introdujo un concepto de paralelismo generalizado en variedades que permitía *conectar* vectores paralelos de espacios tangentes cercanos.

Ahora reconciliamos esta definición con los comentarios previos.

Lema 3.3.1 *Sea una conexión afín ∇ y sean X y V campos de vectores que se expresan como $X = X^i\partial_i$ y $V = V^i\partial_i$ con una carta dada $\big(\mathcal{U}, \phi = (x^1, \ldots, x^n)\big)$, entonces*

$$\nabla_X V = \Big(\frac{\partial V^k}{\partial x^i} + \Gamma^k_{ij}V^j\Big)X^i\partial_k$$

donde Γ^k_{ij} es la componente k-ésima de $\nabla_{\partial_i}\partial_j$.

Demostración: Usando las propiedades de la definición de conexión afín, $\nabla_X V = X^i \nabla_{\partial_i}\big(V^j \partial_j\big) = X^i\big(V^j \nabla_{\partial_i}\partial_j + \frac{\partial V^j}{\partial x^i}\partial_j\big)$ y basta renombrar la variable muda j como k en el último sumando. $\Box$

Cuando uno tiene que hacer cálculos explícitos con conexiones, por ejemplo en relatividad, la notación se vuelve un poco tediosa y existe una especie de taquigrafía bastante extendida que hace más pintorescos los textos de cálculo tensorial.

Se escribe $f_{,i}$ en lugar de $\frac{\partial f}{\partial x^i}$ cuando no se quiere hacer explícito el sistema de coordenadas y $V^k_{;i}$ en lugar de $V^k_{,i} + \Gamma^k_{ij} V^j$.

Definición: Dado un campo de vectores $V = V^k \partial_k$ en una variedad M, se llama *derivada covariante* de V al tensor de tipo $(1,1)$ que tiene componentes $V^k_{;i}$. Si $c : I \longrightarrow M$ es una curva parametrizada, se llama *derivada covariante* de V a lo largo de c, y se denota $\frac{DV}{dt}$, al vector $\nabla_u V$ en $c(t)$ con u el vector tangente a c en $c(t)$.

A veces a la derivada covariante a lo largo de una curva o a la derivada covariante en general, se le llama *derivada absoluta*.

Por la definición libre de coordenadas de ∇ y el Lema 3.3.1, realmente las $V^k_{;i}$ se transforman como componentes de un tensor bajo cambios de carta. Esto obliga a que los Γ^k_{ij} tengan que cumplir una regla de transformación no tensorial (véase la Prop. 4.5 de [Aub01] y su prueba).

Si escribimos $x^i(t)$ y $V^k(t)$ para indicar $(x^i \circ c)(t)$ y $V^k\big(x^1(t), \dots, x^n(t)\big)$, entonces aplicando la regla de la cadena al Lema 3.3.1

$$\frac{DV}{dt} = \Big(\frac{dV^k(t)}{dt} + \Gamma^k_{ij}\big(c(t)\big) V^j(t) \frac{dx^i(t)}{dt}\Big)\partial_k\big|_{c(t)}. \tag{3.11}$$

En particular para determinar DV/dt basta con que V esté definido en la imagen de c, no hace falta tener relamente un campo vectorial en toda la variedad. Si $DV/dt = 0$ se dice que $V(t)$ es el *transporte paralelo* de $V(0)$ a $V(t)$ a lo largo de c. Éste es el concepto generalizado de paralelismo al que antes nos hemos referido.

De todas las conexiones afines posibles, hay sólo una que es natural en una variedad semiriemanniana. Implícitamente, sin repetirlo cada vez, consideraremos que es esta conexión afín la que usamos a partir de ahora.

Teorema 3.3.2 *Sea M una variedad semiriemanniana con métrica $G = g_{ij} dx^i \otimes dx^j$. Existe una única conexión ∇, llamada* conexión de Levi-Civita*, tal que*

$$\nabla_{\partial_i}\partial_j = \nabla_{\partial_j}\partial_i \qquad y \qquad \frac{d}{dt} G(V, W) = G\big(\frac{DV}{dt}, W\big) + G\big(V, \frac{DW}{dt}\big)$$

donde $V = V(t)$ y $W = W(t)$ son campos de vectores arbitrarios definidos sobre una curva. Además los Γ_{ij}^k dando la k-ésima componente de $\nabla_{\partial_i}\partial_j$, llamados símbolos de Christoffel, *responden a la fórmula*

$$\Gamma_{ij}^k = \frac{1}{2} g^{mk}(g_{mi,j} + g_{jm,i} - g_{ij,m})$$

donde (g^{ij}) es la matriz inversa de (g_{ij}).

Demostración: Supongamos primero la existencia de ∇. De $\Gamma_{ij}^k = \Gamma_{ji}^k$ y las propiedades de una conexión, se deduce (véase la definición de $[X,Y]$ en el capítulo anterior)

$$\nabla_X Y - \nabla_Y X = [X,Y]. \tag{3.12}$$

Por otro lado, considerando curvas integrales de un campo vectorial X, la hipótesis sobre la dervada de un producto escalar equivale a

$$X\big(G(Y,Z)\big) = G\big(\nabla_X Y, Z\big) + G\big(Y, \nabla_X Z\big). \tag{3.13}$$

Con unos cálculos tediosos pero simples, utilizando (3.12) y (3.13) permutando las variables, se sigue [dC92, p.55]

$$\begin{aligned} 2G\big(Z,\nabla_Y X\big) \;&=\; X\big(G(Y,Z)\big) + Y\big(G(Z,X)\big) - Z\big(G(X,Y)\big) \\ &\quad -G\big([X,Z],Y\big) - G\big([Y,Z],X\big) - G\big([X,Y],Z\big). \end{aligned} \tag{3.14}$$

Tomando $Y = \partial_i$, $X = \partial_j$, $Z = \partial_k$, esta igualdad afirma

$$g_{kl}\Gamma_{ij}^l = \frac{1}{2}\big(g_{ik,j} + g_{kj,i} - g_{ji,k}\big).$$

Usando la simetría $g_{ij} = g_{ji}$ y multiplicando por la matriz inversa de la métrica, se obtiene la fórmula esperada.

Para la existencia, notemos primero que por dualidad (las componentes de una métrica forman una matriz no singular), la función $Z \mapsto G(Z,X)$ determina el vector X, por tanto (3.14) define $\nabla_Y X$. Además, gracias a la relación $[X, fY] = f[X,Y] + X(f)Y$, no es difícil ver que realmente ∇ tiene las propiedades de una conexión afín.

Con el cálculo antes realizado, $G\big(\partial_k, \nabla_{\partial_i}\partial_j\big) = G\big(\partial_k, \nabla_{\partial_j}\partial_i\big)$ de donde $\nabla_{\partial_i}\partial_j = \nabla_{\partial_j}\partial_i$. Finalmente, a partir de (3.14) se obtiene (3.13) que, como hemos dicho, equivale a la hipótesis sobre la dervada de un producto escalar. $\Box$

Para una subvariedad M de $\mathbb{R}^n$ la métrica natural, *inducida por la usual* (a veces también llamada *métrica usual*) es $G(\vec{v},\vec{w}) = E\big(di(\vec{v}), di(\vec{w})\big)$ donde $i : M \hookrightarrow \mathbb{R}^n$ es la inclusión y $E = (dx^1)^2 + \cdots + (dx^n)^2$ es la métrica euclídea habitual. En esta situación la derivada covariante no es más que la proyección sobre el espacio tangente de la derivada usual. De esta forma, la primera propiedad se sigue de la igualdad de las derivadas parciales cruzadas y la segunda de la regla para derivar un producto.

En términos generales, la segunda condición pedida a la conexión de Levi-Civita es una suerte de *compatibilidad* (de hecho a veces se denomina así) entre la derivada covariante y la usual al derivar productos escalares. La primera condición es algo más difícil de interpretar geométricamente[6]. Sólo diremos que tiene que ver con la existencia de un sistema de coordenadas con el que $V^k_{;i}(p) = V^k_{,i}(p)$ en un punto p [Spi79, Ch.5] lo cual equivale a que los símbolos de Christoffel se anulen en p o a que se pueda encontrar un *sistema normal de coordenadas* en el que los vectores canónicos $\partial_i|_p$ son ortonormales, en el sentido de que $G(\partial_i|_p, \partial_j|_p) = 0$ si $i \neq j$ y $G(\partial_i|_p, \partial_j|_p) = \pm 1$.

Recordemos que habíamos definido las geodésicas como curvas estacionarias para cierto funcional de energía, con la conexión afín asociada a la métrica ahora también las podemos entender como las curvas con aceleración (derivada del vector tangente) nula, obteniéndose así un principio de inercia generalizado. De nuevo las geodésicas se muestran como una generalización de las rectas parametrizadas.

Proposición 3.3.3 *Las geodésicas en una variedad semiriemanniana M son las curvas parametrizadas $c : I \longrightarrow M$ para las que $DV/dt = 0$ donde V es el campo de vectores tangentes de c. Es decir, si $(\mathcal{U}, \phi)$ es una carta y $\phi \circ c(t) = \big(q^1(t), \dots, q^n(t)\big)$, las geodésicas son las soluciones de las ecuaciones diferenciales*

$$\frac{d^2q^k}{dt^2} + \Gamma^k_{ij}\frac{dq^i}{dt}\frac{dq^j}{dt} = 0.$$

Demostración: Todo lo que hay que hacer es comprobar que para el lagrangiano $L = g_{ij}\dot{q}^i\dot{q}^j$ las ecuaciones de Euler-Lagrange son $\ddot{q}^k + \Gamma^k_{ij}\dot{q}^i\dot{q}^j = 0$.

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^k}\right) = \frac{d}{dt}(2g_{kj}\dot{q}^j) = 2g_{kj,i}\dot{q}^i\dot{q}^j + 2g_{kj}\ddot{q}^j, \qquad \frac{\partial L}{\partial q^k} = g_{ij,k}\dot{q}^i\dot{q}^j.$$

Las ecuaciones de Euler-Lagrange se pueden escribir, por tanto, como

$$2g_{kj}\ddot{q}^j - g_{ij,k}\dot{q}^i\dot{q}^j = -2g_{kj,i}\dot{q}^i\dot{q}^j.$$

En el segundo miembro podemos renombrar arbitrariamente los índices de sumación i y j. Si los intercambiamos y sumamos las ecuaciones resultantes, se deduce

$$g_{kj}\ddot{q}^j + \frac{1}{2}(g_{kj,i} + g_{ki,j} - g_{ij,k})\dot{q}^i\dot{q}^j = 0.$$

Multiplicando por g^{lk} se obtiene la ecuación del enunciado. □

[6] Una teoría de la gravitación debida a Cartan, consideraba conexiones sin esta propiedad. Más tarde Einstein intentó unificar la gravitación y el campo electromagnético mediante conexiones con $\Gamma^k_{ij} \neq \Gamma^k_{ji}$. Tal intento, incluso en tiempos de Einstein, llegó a ser poco valorado porque no tomaba en cuenta los avances realizados en la física cuántica.