

Grupo A. Temas generales

Grupo A.1 Geodesia y Geofísica

- Tema 1. El campo de la gravedad terrestre. Sus componentes. Fuerza y potencial gravitatorios. Potencial gravitatorio de una Tierra con simetría esférica. Propiedades del potencial gravitatorio. Aceleración centrífuga, potencial centrífugo. Aceleración y potencial de la gravedad.
- Tema 2. Aceleración y potencial de las mareas. Mareas terrestres. Modelos para el cálculo del potencial teórico de las mareas terrestres para una tierra rígida. Cálculo de los coeficientes de marea. Medida de las mareas terrestres.
- Tema 3. Medidas absolutas y relativas de la gravedad. Métodos pendulares y de caída libre. Observación sobre móviles. Determinación de las segundas derivadas del potencial de la gravedad. Medida del gradiente de la gravedad. Calibración e intercomparación de instrumentos. Correcciones a las observaciones: mareas, movimiento del polo, carga oceánica y carga atmosférica.
- Tema 4. Sistemas geodésicos de referencia. Sistema cartesiano espacial, movimiento del Polo. Sistema de coordenadas en el campo de la gravedad terrestre. Sistemas astronómicos general y local, transformaciones entre ambos.
- Tema 5. El geoide como superficie de referencia para las altitudes. Nivel medio del mar. Altitudes sobre el nivel del mar. Definiciones, objeto de su determinación, precisiones. Altitudes dinámicas, normales y ortométricas. Nivelación geométrica y trigonométrica.
- Tema 6. Sistemas elipsoidales de referencia. Parámetros del elipsoide. Latitud geodésica, geocéntrica y reducida. Curvatura del elipsoide.
- Tema 7. La esfera celeste. El movimiento diurno. Sistemas de coordenadas en Astronomía: horizontales, horarias, ecuatoriales absolutas, eclípticas. Transformaciones. Movimiento aparente del Sol. Teoría de las anomalías. La ecuación del tiempo. El tiempo: tiempo rotacional, tiempo de efemérides, tiempo atómico. Correcciones astronómicas: movimiento propio, precesión, nutación, paralaje, aberración, refracción atmosférica.
- Tema 8. Métodos de transformación entre Sistemas Geodésicos de Referencia Clásicos y Geocéntricos. Transformación de cinco parámetros. Transformación de siete parámetros. Ecuaciones de regresión. Métodos basados en la eliminación de la distorsión de la red.
- Tema 9. Sistemas de Referencia Celestes. Sistemas de Referencia Geocéntricos. ITRS, ETRS, ETRS89. El IERS. Marcos. Transformación de parámetros entre Sistemas Geocéntricos Terrestres.
- Tema 10. Métodos de precisión para el levantamiento de un punto Laplace mediante procedimientos ópticos: Determinación del acimut astronómico por el método de series a la polar, procedimiento y precisión. Determinación del ángulo de latitud astronómica por el método de Sterneck, procedimiento y precisión. Determinación del ángulo de longitud astronómica por el método de Mayer, procedimiento y precisión. Reducción de los datos astronómicos al polo medio de rotación.
- Tema 11. Redes geodésicas: objeto y definiciones. Precisión. Triangulaciones clásicas: Longitud de los lados, utilización de las mismas. Medida de ángulos y distancias en Geodesia: Instrumentación, métodos de observación acimutal. Errores y compensación de una estación. Reducciones de las medidas. Calibración y contrastación de instrumentos.
- Tema 12. Sistemas de posicionamiento y navegación: GPS, EGNOS, Galileo, GLONASS. Sistemas de corrección diferencial y de aumentación. Estaciones virtuales GPS.
- Tema 13. Estructura interna de la Tierra. Corteza y manto superior. Manto inferior y núcleo. Densidad y parámetros elásticos. Propiedades anelásticas. Ecuación de estado y de composición.
- Tema 14. Distribución espacial de terremotos. Características de terremotos en márgenes convergentes, divergentes y transcurrentes. El ciclo sísmico: modelos de recurrencia. Distribución de magnitudes. Modelos temporales de recurrencia.
- Tema 15. Caracterización de terremotos. Identificación de fases sísmicas en un sismograma. Localización hipocentral. Intensidad sísmica. Escala EMS-98. Definiciones de magnitud.
- Tema 16. Instrumentación sísmica. Teoría del sismómetro mecánico. Sismómetro electromagnético. Sismómetro de banda ancha. Acelerómetro. Funciones de respuesta y de transferencia. Determinación de amplitudes del suelo a través de sismogramas digitales.
- Tema 17. Movimientos sísmicos fuertes. Acelerogramas. Características de un acelerograma en el tiempo y en la frecuencia. Estimación empírica de la aceleración máxima en un punto. Espectro de respuesta y de diseño.
- Tema 18. Peligrosidad y riesgo sísmico. Conceptos. Características de los métodos determinista y probabilista. Periodo de retorno. Normativa de construcción sismorresistente en España.
- Tema 19. Tsunamis. Generación, propagación e inundación. Magnitud e intensidad del Tsunami. Características de los terremotos productores del Tsunami. Sistemas de alerta de Tsunami.
- Tema 20. Volcanismo. Materiales volcánicos. Mecánica de los fenómenos eruptivos. Proyección de piroclastos. Extrusión y dinámica de domos y coladas. Mapas de peligrosidad.
- Tema 21. Campo magnético terrestre. Componentes y división según su origen. Campo magnético de un dipolo. Dipolo terrestre. Variación secular. Origen del campo magnético interno.
- Tema 22. Campo magnético externo. Variaciones del campo magnético externo. Composición de la ionosfera. Estructura de la magnetosfera. Anillos de radiaciones y Auroras.
- Tema 23. Observaciones del campo magnético. Medidas absolutas y relativas. Métodos clásicos y modernos de medidas del campo magnético. Observación desde satélites.
- Tema 24. Radioactividad de la Tierra. Elementos radioactivos. Leyes de la desintegración radioactiva. Principios de geocronología. Edad de la Tierra. Evolución térmica de la Tierra.

Tema 1. El campo de la gravedad terrestre. Sus componentes. Fuerza y potencial gravitatorios. Potencial gravitatorio de una Tierra con simetría esférica. Propiedades del potencial gravitatorio. Aceleración centrífuga, potencial centrífugo. Aceleración y potencial de la gravedad.

1.1. El campo de la gravedad terrestre

De las cuatro interacciones básicas que ocurren en la naturaleza (fuerza gravitatoria, fuerza electromagnética, fuerza nuclear fuerte y fuerza nuclear débil), la fuerza gravitatoria es la más débil de todas ellas, siendo despreciable en las interacciones entre partículas elementales (moléculas, átomos, núcleos, etc.), o entre objetos del tamaño de las personas pero siendo de gran importancia en las interacciones entre cuerpos muy grandes tales como satélites, planetas, estrellas, etc.

Por campo de la gravedad terrestre no sólo se entiende a la fuerza gravitatoria ejercida por la Tierra, sino a la suma de todas aquellas fuerzas que actúan sobre un cuerpo en reposo sobre la superficie terrestre.

La astronomía no es la única ciencia que se encarga del estudio del campo de la gravedad de la Tierra. Uno de los principales objetivos de la geodesia es la determinación de este campo de gravedad ya que, entre otras razones, las magnitudes medidas en geodesia tienen como sistema de referencia fundamental, el campo de la gravedad terrestre. De hecho, el geoide es una superficie de nivel del campo de gravedad. Además, la distribución de los valores de gravedad en superficie junto con otras medidas geodésicas permite determinar la forma de la superficie terrestre.

La gravimetría también tiene como principal objetivo el estudio del campo de gravedad terrestre debido a que las medidas de gravedad en superficie dan información sobre la estructura y las características del interior de la Tierra y la variación temporal del campo de gravedad revela fenómenos como las oscilaciones de los polos terrestres, la redistribución de la masa de la Tierra, etc.

1.2. Sus componentes

Las fuerzas que actúan sobre un cuerpo en reposo que se encuentra en un punto fijo sobre la superficie terrestre son:

- la fuerza gravitatoria terrestre
- la fuerza centrífuga debida a la rotación de la Tierra
- la fuerza de atracción de otros cuerpos celestes como son el Sol y la Luna
- la fuerza de atracción de la atmósfera terrestre
- Otras

La fuerza resultante de todas ellas se denomina fuerza de la gravedad y se caracteriza por ser función de la posición y del tiempo. Para estudios no muy precisos, sólo se tienen en cuenta las dos primeras fuerzas mencionadas ya que la influencia de las demás fuerzas sobre el valor de la gravedad es muy pequeña y se suelen despreciar (como haremos en este tema).

Esta fuerza de la gravedad $\vec{F}$, se obtiene de multiplicar la aceleración de la gravedad $\vec{g}$ por la masa del cuerpo m :

$$\vec{F} = m\vec{g}$$

Al vector $\vec{g}$ se le denomina campo de la gravedad terrestre y se define como la fuerza por unidad de masa ejercida por la Tierra sobre cualquier objeto.

Para la representación del campo de gravedad, se suele considerar a la Tierra como un sólido rígido que gira con velocidad de rotación ω , constante sobre un eje invariable. En un sistema de coordenadas geocéntrico, el origen se sitúa en el centro de gravedad de la Tierra y se hace coincidir el eje z con el eje medio de rotación. El eje x está contenido en el plano del meridiano de Greenwich y el eje y se elige de tal manera que esté en el plano del ecuador y sea perpendicular a los ejes x y z .

En un punto de la superficie terrestre, el campo de la gravedad $\vec{g}$ vendrá dado por la suma del campo gravitatorio $\vec{g}_m$ más la aceleración centrífuga $\vec{a}_c$:

$$\vec{g} = \vec{g}_m + \vec{a}_c$$

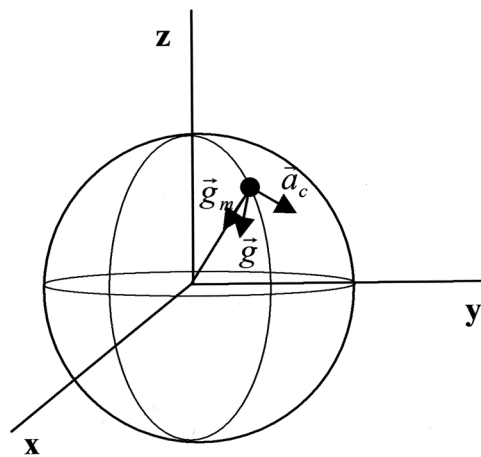


Figura 1. Campo de gravedad en una tierra esférica

La gravedad es función de la posición. Diferencias de gravedad con la latitud o la altitud pueden llegar a ser del orden de $5 \cdot 10^{-3}g$ (siendo g un valor medio de gravedad). Los efectos de las mareas afectan en torno a los $3 \cdot 10^{-7}g$ y los desplazamientos de las masas terrestres influyen entre $10^{-8}g$ y $10^{-9}g$.

La unidad de la gravedad en el Sistema Internacional es m/s^2 . Sin embargo, se suele utilizar la unidad gravimétrica para medidas de desviaciones de la gravedad o para errores de medida:
 $1u.g. = \mu m/s^2 = 10^{-6} m/s^2$

En el sistema cgs la unidad de la gravedad es el Gal (por Galileo): $1gal = 1cm/s^2$

Y para desviaciones de la gravedad o errores de medida se utiliza el mgal: $1mgal = 10^{-5} m/s^2$

La unidad del gradiente de gravedad es s^{-2} en el Sistema Internacional o E (Eötvös) en el sistema cgs: $1E = 10^{-9}s^{-2}$.

Y la unidad del potencial es m^2/s^2 , dimensión de trabajo por unidad de masa.

1.3. Fuerza y potencial gravitatorios

En 1687 Newton publicó la ley de la gravitación en su “Philosophiae Naturalis Principia Mathematica”. Esta ley postula que existe una fuerza de atracción entre cada par de objetos que es proporcional al producto de las masas de los objetos e inversamente proporcional al cuadrado de la distancia que los separa. Esta fuerza se denomina fuerza gravitatoria y está dirigida a lo largo de la línea que conecta los dos objetos. Veámosla con más detalle.

Si tenemos dos objetos en un sistema de coordenadas rectangulares, uno con masa m_1 que se encuentra en la posición $\vec{r}_1$ y otro con masa m_2 en la posición $\vec{r}_2$, la fuerza ejercida por m_1

sobre m_2 $\vec{F}_{1,2}$ es:

$$\vec{F}_{1,2} = -G \frac{m_1 m_2}{r^2} \hat{r}_{1,2}$$

G es la constante de gravitación universal, cuyo valor es $6,67 \cdot 10^{-11} \text{ Nm}^2/\text{kg}^2$.

$\hat{r}_{1,2}$ es el vector unitario del vector definido desde la masa m_1 a la masa m_2 : $\hat{r}_{1,2} = \frac{\vec{r}_{1,2}}{|\vec{r}_{1,2}|}$

La fuerza gravitatoria $\vec{F}_{1,2}$ está situada en la recta que une m_1 y m_2 , y dirigida hacia m_1 . De modo que $\vec{F}$ y $\vec{r}$ apuntan en direcciones opuestas.

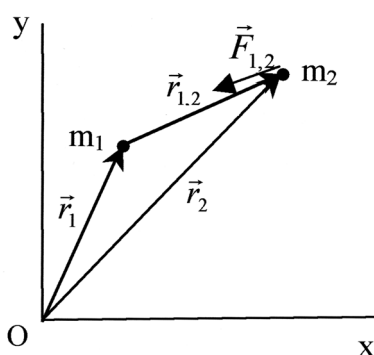


Figura 2. Fuerza gravitacional actuando entre dos objetos

En este caso, m_1 es la masa atrayente y m_2 la masa atraída.

De acuerdo con la tercera ley de Newton, la fuerza $\vec{F}_{2,1}$ (ejercida por m_2 sobre m_1) es el valor negativo de $\vec{F}_{1,2}$:

$$\vec{F}_{2,1} = -\vec{F}_{1,2}$$

El módulo de la fuerza gravitatoria es por tanto:

$$F = G \frac{m_1 m_2}{r^2}$$

A la fuerza gravitatoria por unidad de masa ejercida por la Tierra sobre cualquier objeto de masa

m se denomina campo gravitatorio terrestre $\vec{g}_m$:

$$\vec{g}_m = \frac{\vec{F}}{m}$$

$\vec{g}_m$ es la denominada aceleración gravitacional o gravitación.

El campo gravitatorio en un punto debido a un conjunto de masas puntuales es igual a la suma vectorial de los campos debidos a las masas individuales en dicho punto:

$$\vec{g}_m = \sum_i \vec{g}_{mi}$$

Y el campo gravitatorio en un punto debido a un cuerpo continuo se calcula considerando el campo $d\vec{g}_m$ debido a un pequeño elemento de masa dm del cuerpo, suponiendo que se trata de una masa puntual y se integra sobre el cuerpo entero: $\vec{g}_m = \int d\vec{g}_m$

Si en un sistema de coordenadas rectangulares designamos las coordenadas de la masa atrayente M por (x_0, y_0, z_0) y las coordenadas del punto atraído, de masa m igual a la unidad, por (x, y, z) , las componentes de la fuerza gravitatoria $\vec{F}$ serán:

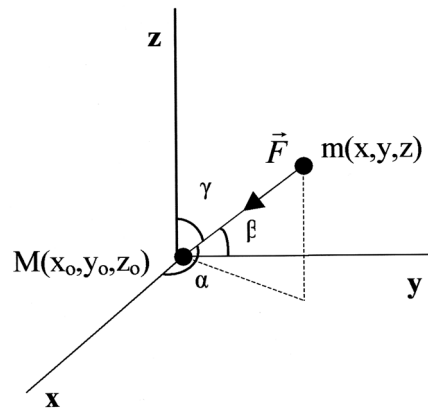


Figura 3. Fuerza de atracción entre dos objetos en un sistema de coordenadas rectangulares

$$\left. \begin{aligned} F_x &= -F \cos \alpha = -\frac{GM}{r^2} \frac{(x-x_0)}{r} = -\frac{GM}{r^3} (x-x_0) \\ F_y &= -F \cos \beta = -\frac{GM}{r^2} \frac{(y-y_0)}{r} = -\frac{GM}{r^3} (y-y_0) \\ F_z &= -F \cos \gamma = -\frac{GM}{r^2} \frac{(z-z_0)}{r} = -\frac{GM}{r^3} (z-z_0) \end{aligned} \right\} \vec{F} = (F_x, F_y, F_z)$$

siendo $r = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$ la distancia entre el elemento de masa M y el punto atraído.

Si se introduce una función escalar de la forma $V = \frac{GM}{r}$, se verifica que las componentes de

la fuerza gravitatoria $\vec{F}$ vienen dadas por: $F_x = \frac{\partial V}{\partial x}$ $F_y = \frac{\partial V}{\partial y}$ $F_z = \frac{\partial V}{\partial z}$

Por tanto: $\vec{F} = (F_x, F_y, F_z) = \nabla V$

Gracias a que $\nabla \times \vec{F} = 0$, el vector fuerza es el gradiente de la función escalar V . Y V se denomina potencial gravitatorio e indica el trabajo requerido para mover una unidad de masa dentro del campo gravitatorio.

Este resultado es de gran importancia ya que las tres componentes del vector $\vec{F}$ se pueden reemplazar por una única función escalar V . Esto hace que en casos complicados, donde se considera la atracción de muchos puntos materiales o de cuerpos sólidos, la función V es simplemente la suma de las contribuciones de las respectivas partículas.

En un sistema con varios puntos materiales $m_1, m_2, m_3 \dots$ el potencial gravitatorio es la suma de

las contribuciones individuales:
$$V = G \sum_{i=1}^n \frac{m_i}{r}$$

Si los puntos materiales están distribuidos continuamente sobre un volumen v de densidad

$\rho = \frac{dm}{dv}$, donde dv es un elemento de volumen y dm un elemento de masa, el potencial

gravitatorio vendrá dado por:

$$V = G \iiint_v \frac{dm}{r} = G \iiint_v \frac{\rho}{r} dv = G \iiint_v \frac{\rho(x_0, y_0, z_0)}{\sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}} dx_0 dy_0 dz_0$$

1.4. Potencial gravitatorio de una Tierra con simetría esférica

En la determinación del potencial gravitatorio terrestre se considera, en primera aproximación, la Tierra de forma esférica con una estructura de densidad centralmente simétrica.

Para su cálculo, utilizaremos coordenadas esféricas (r, θ, λ) , siendo r la distancia geocéntrica con origen en el centro de masas de la Tierra. El ángulo θ o colatitud geocéntrica, se mide en sentido horario desde el eje de rotación y su complemento ϕ o latitud geocéntrica ($\theta = 90^\circ - \phi$), se mide desde el plano ecuatorial al radio vector $\vec{r}$, positivo en el hemisferio norte y negativo en el hemisferio sur. La longitud λ se mide positiva hacia el este, tomando como origen el meridiano de Greenwich.

Y ello con la orientación acostumbrada de este sistema con respecto al sistema global x, y, z (el eje $\theta = 0$ coincide con el eje z situado a lo largo del eje de rotación y el eje $\lambda = 0$ coincide con el eje x).

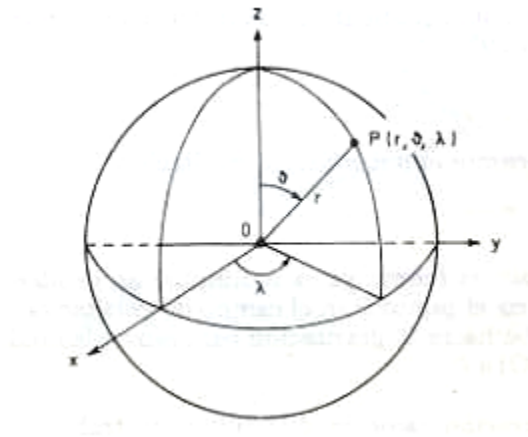


Figura 4. Coordenadas esféricas. Fuente: Torge W. Geodesia

El radio vector en coordenadas esféricas es:

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin \theta \cos \lambda \\ r \sin \theta \sin \lambda \\ r \cos \theta \end{pmatrix}$$

Por simplicidad, supondremos la masa atraída m , igual a la unidad.

En primer lugar calcularemos el potencial de una corteza esférica homogénea para, a partir de ella, obtener el potencial de la tierra esférica.

El potencial de una corteza esférica homogénea de densidad ρ y grosor infinitesimal dr' viene

dado por:

$$V' = G \iiint_V \frac{\rho}{r} dv = G \rho \iiint_V \frac{r'^2 \sin \theta'}{r} d\theta' d\lambda' dr'$$

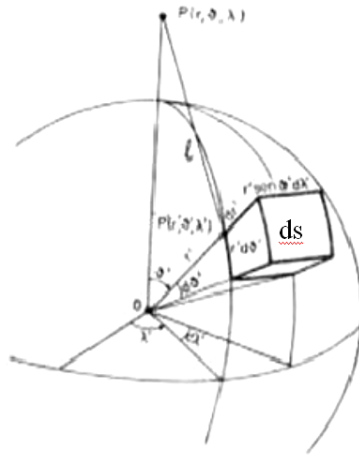


Figura 5. Corteza esférica. Fuente: Torge W. Geodesia

Ya que el elemento de volumen de la corteza esférica es: $dv = r'^2 \sin \theta' d\theta' d\lambda' dr'$

Y el elemento diferencial de masa de la corteza esférica será: $dm' = 4\pi\rho r'^2 dr'$

Distingamos dos casos según el punto atraído P, se encuentre fuera de la Tierra o dentro de ella.

1.4.1. P exterior a la Tierra

Si el punto atraído está fuera de la corteza esférica, el potencial vendrá dado por:

$$V'_{ext} = 4\pi G\rho \frac{r'^2}{r} dr' = G \frac{dm'}{r}$$

Si consideramos que la Tierra está compuesta por un conjunto continuo de cascarones esféricos homogéneos y concéntricos, el potencial gravitatorio para una Tierra esférica vendrá dado por:

$$V_{ext} = G \iiint_T \frac{dm}{r} = \frac{GM}{r}$$

Donde la integral se extiende a toda la Tierra y M es la masa de la Tierra.

Este resultado equivale al potencial de toda la masa de la Tierra concentrada en el centro de masas.

1.4.2. P interior a la Tierra

El potencial gravitatorio de un punto interior a la corteza esférica será:

$$V'_{\text{int}} = 4\pi G \rho \frac{r'^2}{r'} dr' = G \frac{dm'}{r'}$$

Al ser el potencial constante, para puntos del interior de la corteza esférica el campo gravitatorio es cero.

Para calcular el potencial de la Tierra esférica de radio R para un punto situado en su interior a una distancia r del centro de masas, consideraremos que la Tierra está formada por cascarones esféricos continuos, por lo que el valor del potencial será la contribución del potencial para la tierra esférica de radio r , como si el punto se encontrase en su superficie, y la contribución de la corteza esférica de grosor $R - r$:

$$V_i = 4\pi G \int_0^r \rho \frac{r'^2}{r} dr' + 4\pi G \int_r^R \rho \frac{r'^2}{r'} dr' = \frac{4\pi G \rho}{r} \int_0^r r'^2 dr' + 4\pi G \int_r^R r' dr' = 2\pi G \rho \left(R^2 - \frac{r^2}{3} \right)$$

1.5. Propiedades del potencial gravitatorio

Volvemos a distinguir los dos casos anteriores.

1.5.1. P exterior a la Tierra

Como hemos visto, el potencial gravitatorio para todo punto exterior a la Tierra viene dado por

la expresión: $V_{\text{ext}} = G \iiint_T \frac{dm}{r}$ Con $r = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$.

Así, V_{ext} es una función continua para $r > 0$, que se anula en el infinito como $1/r$:

$$\lim_{r \rightarrow \infty} V_{\text{ext}} = 0$$

Para distancias r muy grandes el cuerpo actúa, aproximadamente, como un punto material, y el potencial viene dado por: $V_{ext} = \frac{GM}{r}$ lo que equivale al potencial de toda la masa M de la Tierra concentrada en el centro de masas.

Veamos la continuidad de las derivadas del potencial gravitatorio en este caso:

La primera derivada del potencial, es decir, la componente de la fuerza, para la componente x , vendrá dada por:

$$\frac{\partial V_{ext}}{\partial x} = G \iiint_T \frac{d}{dx} \left(\frac{1}{r} \right) dm = -G \iiint_T \frac{(x - x_0)}{r^3} dm \quad \text{Función continua para } r > 0.$$

La segunda derivada, para la componente x será:

$$\frac{\partial^2 V_{ext}}{\partial x^2} = \frac{\partial}{\partial x} \left(-G \iiint_T \frac{(x - x_0)}{r^3} dm \right) = -G \iiint_T \frac{\partial}{\partial x} \left(\frac{(x - x_0)}{r^3} \right) dm = -G \iiint_T \left(\frac{1}{r^3} - \frac{3(x - x_0)^2}{r^5} \right) dm$$

También es una función continua para $r > 0$

Análogamente se opera con las otras dos componentes.

De modo que el potencial gravitatorio para un punto P exterior a la Tierra, sus primeras y segundas derivadas son valores finitos y funciones continuas en todo el espacio exterior, anulándose en el infinito.

Aplicando el operador laplaciano a V_{ext} :

$$\Delta V_{ext} = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -G \iiint_v \left(\frac{1}{r^3} - \frac{3(x - x_0)^2}{r^5} \right) dm - G \iiint_v \left(\frac{1}{r^3} - \frac{3(y - y_0)^2}{r^5} \right) dm - G \iiint_v \left(\frac{1}{r^3} - \frac{3(z - z_0)^2}{r^5} \right) dm = 0$$

$$\Delta V_{ext} = 0 \quad \text{Ecuación de Laplace.}$$

V_{ext} cumple la ecuación de Laplace. Las soluciones de la ecuación de Laplace son funciones armónicas o también llamadas funciones potenciales. Y toda función armónica es analítica

dentro de la región donde satisface dicha ecuación. Esto es, es continua y tiene derivadas continuas de cualquier orden, como hemos comprobado.

1.5.2. P interior a la Tierra

Para calcular el potencial gravitatorio de un punto que se encuentra en el interior de la Tierra hay que tener en cuenta que ahora el caso $r = 0$ es posible, por lo que es posible la discontinuidad debida a $1/r$.

Para calcular el potencial gravitatorio, consideramos al punto P a una distancia q del centro de una esfera E que lo envuelve, la cual tiene densidad homogénea ρ , radio p y está centrada en P_0 .

Figura 6. Potencial gravitatorio dentro de la Tierra. Fuente: Torge W. Geodesia

El potencial gravitatorio será la suma de las contribuciones debidas a las masas que están en el interior de la esfera y de las masas que se encuentran en su exterior:

$$V_i = G \iiint_{T-E} \frac{dm}{r} + 2\pi G \rho \left(p^2 - \frac{q^2}{3} \right)$$

V_i es una función continua para $r > 0$. En el caso en que $p \rightarrow 0, q \rightarrow 0$, se obtiene la expresión del potencial para un punto exterior a la Tierra.

Estudiemos la primera y segunda derivada del potencial para la componente x (se haría de igual modo para las otras componentes):

$$\frac{\partial V_{\text{int}}}{\partial x} = -G \iiint_{T-E} \frac{(x-x_0)}{r^3} dm - \frac{4}{3} \pi G \rho (x-x_0) \quad \text{Función continua para } r > 0$$

$$\frac{\partial^2 V_{\text{int}}}{\partial x^2} = -G \iiint_{T-E} \left(\frac{1}{r^3} - \frac{3(x-x_0)^2}{r^5} \right) dm - \frac{4}{3} \pi G \rho$$

Las segundas derivadas muestran discontinuidades para cambios súbitos de densidad.

El laplaciano será:

$$\Delta V_{\text{int}} = \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = -4\pi G \rho$$

$$\Delta V_{\text{int}} = -4\pi G \rho \quad \text{Ecuación de Poisson}$$

En el interior de la Tierra se verifica la ecuación de Poisson. El potencial gravitatorio y sus primeras derivadas son funciones de valor único, finitas y continuas para $r > 0$, mientras que las segundas derivadas muestran discontinuidades para cambios súbitos de densidad.

(La superficie más importante de discontinuidades de densidades es la superficie Tierra-atmósfera con un salto de densidad de aproximadamente $0,0013 \text{ g/cm}^3$ para la atmósfera con unos $2,7 \text{ g/cm}^3$ aproximadamente para la corteza terrestre).

Por consiguiente, el potencial gravitatorio es una función armónica fuera de las masas atrayentes pero no dentro de las mismas, donde se satisface la ecuación de Poisson.

1.6. Aceleración centrífuga, potencial centrífugo

La aceleración centrífuga aparece solamente en los sistemas rotacionales no inerciales. Está causada por el movimiento de rotación de la Tierra sobre su propio eje.

Para calcularla se considera, en primera aproximación, que la Tierra gira con una velocidad angular ω constante alrededor de un eje fijo. Tomaremos el sistema de coordenadas rectangulares donde el eje z coincide el eje medio de rotación.

La aceleración centrífuga que actúa sobre un punto situado en la superficie terrestre a una distancia d del eje de rotación es:

$$\vec{a}_c = \omega^2 \vec{d} = (\omega^2 x, \omega^2 y, 0)$$

Siendo la distancia del punto donde calculamos la fuerza al eje de rotación:

$$|\vec{d}| = \sqrt{x^2 + y^2}$$

La aceleración centrífuga $\vec{a}_c$ tiene la dirección del vector $\vec{d}$ y por lo tanto es perpendicular al eje de rotación y dirigida hacia el exterior de la Tierra.

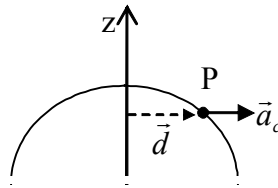


Figura 7. Aceleración centrífuga

La velocidad angular de rotación de la Tierra: $\omega = 7,292 \cdot 10^{-5} \text{ rad/s}$.

La fuerza centrífuga es el producto de la masa del punto donde estamos evaluando la fuerza por la aceleración centrífuga:

$$\vec{f} = m\vec{a}_c$$

Al igual que para el caso gravitatorio y gracias a que $\nabla \times \vec{f} = 0$, se puede considerar una función escalar que defina, a través de sus derivadas, la función vectorial $\vec{f}$ que está actuando sobre una masa unidad. A esta función escalar se la denomina potencial centrífugo y viene dado por la expresión:

$$\Phi = \frac{1}{2} \omega^2 d^2 = \frac{1}{2} \omega^2 (x^2 + y^2) \quad \vec{f}_c = \nabla \Phi = \left(\frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y}, \frac{\partial \Phi}{\partial z} \right)$$

El potencial es una función continua y verifica que: $\lim_{d \rightarrow 0} \Phi = 0$

Las derivadas del potencial centrífugo son también funciones continuas. Para la componente x :

$$\frac{\partial \Phi}{\partial x} = \omega^2 x \quad \frac{\partial^2 \Phi}{\partial x^2} = \omega^2$$

El laplaciano del potencial es:
$$\Delta\Phi = \frac{\partial^2\Phi}{\partial x^2} + \frac{\partial^2\Phi}{\partial y^2} + \frac{\partial^2\Phi}{\partial z^2} = \omega^2 + \omega^2 + 0 = 2\omega^2 \neq 0$$

Al contrario que V , esta función no es armónica.

El valor del potencial centrífugo en el ecuador es de $1,1 \cdot 10^5 \text{ m}^2/\text{s}^2$ mientras que la aceleración centrífuga vale $0,03 \text{ m/s}^2$. En los polos, tanto la fuerza centrífuga como el potencial centrífugo son nulos.

Por último, destacar que si el cuerpo se encontrase en movimiento, además de la fuerza centrífuga actuaría la fuerza de Coriolis, la cual es proporcional a la velocidad del cuerpo con respecto a la Tierra, de modo que es nula para todos aquellos cuerpos que estén en reposo sobre la superficie terrestre.

1.7. Aceleración y potencial de la gravedad

La fuerza de la gravedad es la resultante de la suma de la fuerza gravitatoria y de la fuerza centrífuga que en coordenadas esféricas se expresan como:

- fuerza gravitatoria, $\vec{F}$:
$$\vec{F} = (F_r, F_\theta, F_\lambda) = \left(-\frac{GM}{r^2}, 0, 0\right)$$
- fuerza centrífuga, $\vec{f}$:
$$\vec{f} = (f_r, f_\theta, f_\lambda) = (\omega^2 r \sin^2 \theta, \omega^2 r \sin \theta \cos \theta, 0)$$

La fuerza de la gravedad por unidad de masa se denomina aceleración de la gravedad $\vec{g}$. Su dirección se conoce como dirección de la plomada y su magnitud, $g = |\vec{g}|$, se denomina intensidad de la gravedad o simplemente gravedad.

La fuerza de la gravedad se podrá obtener también del gradiente de un potencial, el llamado potencial gravítico o potencial de la gravedad W , suma del potencial gravitatorio V y del potencial centrífugo Φ :

$$W = V + \Phi = G \iiint_T \frac{\rho}{r} dv + \frac{1}{2} \omega^2 (x^2 + y^2) = \frac{GM}{r} + \frac{1}{2} \omega^2 r^2 \sin^2 \theta$$

Sus derivadas cumplen:

$$\left. \begin{aligned} g_x &= \frac{\partial W}{\partial x} = -G \iiint_v \frac{x-x_0}{r^3} \rho dv + \omega^2 x \\ g_y &= \frac{\partial W}{\partial y} = -G \iiint_v \frac{y-y_0}{r^3} \rho dv + \omega^2 y \\ g_z &= \frac{\partial W}{\partial z} = -G \iiint_v \frac{z-z_0}{r^3} \rho dv \end{aligned} \right\} \vec{g} = \nabla W = \left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}, \frac{\partial W}{\partial z} \right) = (g_x, g_y, g_z)$$

El potencial W y sus primeras derivadas son de valor único, finitas y continuas en todo el espacio, como consecuencia de las características de los potenciales V y Φ , con excepción del caso en que $\vec{r} \rightarrow \infty$, ya que entonces $\Phi \rightarrow \infty$, y $\vec{g} = 0$ (la dirección de la plomada no es única). Las segundas derivadas poseen discontinuidades a variaciones de densidad irregulares.

El laplaciano de W , cumple que en un punto:

- exterior a la Tierra: $\Delta W = 2\omega^2$ Ecuación de Laplace generalizada
- interior a la Tierra: $\Delta W = -4\pi G\rho + 2\omega^2$ Ecuación de Poisson generalizada

Debido a la forma de la Tierra que no es totalmente esférica sino que está achatada por los polos y ensanchada por el ecuador, el valor de la gravedad en el ecuador es de $9,78 \text{ m/s}^2$ mientras que en los polos toma un valor de $9,83 \text{ m/s}^2$.

Bibliografía

- [1] Heiskanen W.A. y Moritz H. Geodesia física. Trad. Sevilla de Lerma, J. Instituto Geográfico Nacional; Instituto de Astronomía y Geodesia, 1985.
- [2] Torge, W. Geodesia. Editorial Diana, 1983.
- [3] Torge, W. Gravimetry. Editorial Walter de Gruyter 1989.

Tema 2. Aceleración y potencial de las mareas. Mareas terrestres. Modelos para el cálculo del potencial teórico de las mareas terrestres para una tierra rígida. Cálculo de los coeficientes de marea. Medida de las mareas terrestres

2.1. Aceleración y potencial de las mareas

El campo de la gravedad terrestre varía con el tiempo por múltiples causas: procesos tectónicos, transferencias de masa en el manto terrestre, actividad volcánica, movimientos de fallas, etc. Pero sin ninguna duda, la principal variación temporal del campo de gravedad se debe a las mareas que producen variaciones en el campo de la gravedad del orden de $10^{-7}g$ (siendo g un valor medio de gravedad).

Las mareas están ocasionadas por la atracción que sobre la Tierra ejercen los cuerpos celestes, principalmente la Luna por su cercanía y el Sol por su gran masa, y al efecto inercial del sistema astro-Tierra en torno a su baricentro. Las mareas se aprecian de forma más notable en los océanos aunque también afectan a la parte sólida de la Tierra, Por ello, el estudio de las mareas permite conocer mejor las características elásticas de la Tierra.

En una primera aproximación, se considera a la Tierra como un sólido rígido y a la Luna y al Sol como puntos masa. Los cálculos se realizan separadamente para el sistema Tierra-Luna y el sistema Tierra-Sol y posteriormente se suman. Los resultados obtenidos con esta aproximación se denominan mareas teóricas.

Se toma un sistema de referencia fijo en la Tierra, considerada esta de forma esférica. El origen se sitúa en el centro de masas terrestre y no se considera la parte de la rotación terrestre. El sistema gira en torno al centro de gravedad de la Tierra y el astro, el baricentro.

Calculemos el efecto de la marea sobre una masa unidad que se encuentra situada en un punto P de la superficie terrestre.

Haremos los cálculos para el efecto producido por un astro cualquiera y luego los resultados los aplicaremos al Sol y a la Luna.

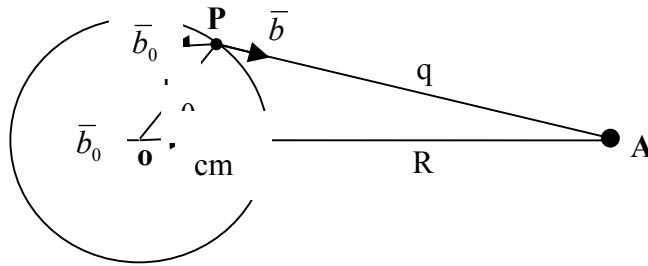


Figura 1. Aceleración de la marea lunar

En esta figura, R es la distancia geocéntrica del centro del astro A , al centro de la Tierra. Suponemos que la distancia R se encuentra situada sobre el eje x . La distancia entre el punto P y el centro del astro se expresa por q . La distancia cenital geocéntrica θ es el ángulo entre el radio vector $\vec{r}$ del centro de la Tierra al punto P y el eje x . Expresaremos la masa del astro por M . Y el centro de masas del sistema, cm , se encuentra situado dentro del radio de la Tierra.

En el punto P actuará:

- La fuerza de atracción gravitatoria $\vec{b}$ del astro causante de la marea:

$$\vec{b} = \frac{GM}{q^2} \hat{u}_q \quad \text{siendo } \hat{u}_q \text{ el vector unitario en la dirección de } \vec{q}.$$

- La fuerza centrífuga $\vec{b}_0$:

Para calcularla se tiene que en cuenta que hemos considerado una Tierra rígida, de modo que la fuerza está actuando, en todos los puntos rígidamente conectados, de igual manera y por tanto, tendrá un valor constante. Para obtener este valor, se calcula esta fuerza en el centro de masas de la Tierra, ya que es aquí donde se compensa la fuerza centrífuga con la fuerza gravitatoria del astro al encontrarse el sistema en equilibrio.

$$\vec{b}_0 = -\frac{GM}{R^2} \hat{u}_R \quad \text{fuerza inercial de la rotación del sistema en torno al baricentro}$$

La suma de estas dos fuerzas recibe el nombre de fuerza de marea:

$$\vec{b}_t = \vec{b} + \vec{b}_0 = \frac{GM}{q^2} \hat{u}_q - \frac{GM}{R^2} \hat{u}_R$$

Esta fuerza puede expresarse como el gradiente de un potencial:

$$\vec{b}_i = \nabla \psi \quad b_{ii} = \frac{\partial}{\partial x_i} \left(\frac{GM}{q} - \frac{GM}{R^2} x \right) = \frac{\partial}{\partial x_i} \left(\frac{GM}{q} - \frac{GM}{R^2} r \cos \theta \right)$$

$$\psi = GM \left(\frac{1}{q} - \frac{r \cos \theta}{R^2} + cte \right)$$

En el potencial aparece una constante. Para determinarla se tiene en cuenta que ψ se debe

anular en el centro de la Tierra: $\psi = 0$ para $r = 0$ ($q = R$) $cte = -\frac{1}{R}$

El potencial de las mareas es:
$$\psi = GM \left(\frac{1}{q} - \frac{r \cos \theta}{R^2} - \frac{1}{R} \right)$$

La distancia q se puede expresar como:
$$q = \sqrt{r^2 + R^2 - 2rR \cos \theta}$$

$$\frac{1}{q} = \frac{1}{R} \left[\left(\frac{r}{R} \right)^2 + 1 - 2 \left(\frac{r}{R} \right) \cos \theta \right]^{-1/2} = \frac{1}{R} \sum_{n=0}^{\infty} \left(\frac{r}{R} \right)^n P_n(\cos \theta)$$

Donde se han utilizado los polinomios de Legendre:
$$(1 - 2zx + x^2)^{-1/2} = \sum_{n=0}^{\infty} x^n P_n(z)$$

Con
$$P_n(z) = \frac{1}{2^n n!} \frac{d^n}{dz^n} (z^2 - 1)^n$$

De modo que el potencial de las mareas se puede escribir como:

$$\psi = GM \left[\frac{1}{R} \sum_{n=0}^{\infty} \left(\frac{r}{R} \right)^n P_n(\cos \theta) - \frac{r \cos \theta}{R^2} - \frac{1}{R} \right]$$

Tomando todos los términos hasta $n = 2$, el potencial de las mareas queda de la siguiente manera:

$$\psi = \frac{1}{2} GM \frac{r^2}{R^3} (3 \cos^2 \theta - 1)$$

Si introducimos C , distancia media del centro de la Tierra al astro y $\overline{G}(r) = \frac{3}{4}GM \frac{r^2}{C^3}$, la constante de Doodson (que depende de la latitud del punto donde se observa la marea), el potencial de la marea producida por un astro A de masa M es:

$$\psi = \overline{G}(r) \frac{2}{3} \left(\frac{C}{R} \right)^3 (3 \cos^2 \theta - 1) = \overline{G}(r) \left(\frac{C}{R} \right)^3 \left(\cos 2\theta + \frac{1}{3} \right)$$

El potencial de la marea lunisolar, marea producida tanto por la Luna, L, como por el Sol, S, es:

$$\psi_{LS} = \psi_L + \psi_S = \overline{G}_L \frac{2}{3} \left(\frac{C_L}{R_L} \right)^3 (3 \cos^2 \theta - 1) + \overline{G}_S \frac{2}{3} \left(\frac{C_S}{R_S} \right)^3 (3 \cos^2 \theta' - 1)$$

Para un punto en la superficie terrestre:

$$\overline{G}_L(R_T) = 2,628 m^2/s^2 \quad \overline{G}_S(R_T) = 1,207 m^2/s^2$$

De aquí que la proporción entre las mareas solares y las mareas lunares sea: $\overline{G}_S = 0,46 \overline{G}_L$

El efecto producido por las mareas solares es sólo un 46% del producido por las mareas lunares. Esto es debido a que el campo gravitatorio de un objeto esférico varía con la distancia según $1/r^2$. La fuerza ejercida por la Luna es más intensa en los puntos de la Tierra más próximos a ella que en los puntos más alejados. Y aunque el Sol ejerce sobre los océanos una fuerza mucho mayor que la ejercida por la Luna, la diferencia de fuerzas ejercida por la Luna entre el océano más próximo y el más alejado es mucho mayor que la correspondiente fuerza diferencial ejercida por el Sol. Y esta fuerza diferencial es la responsable de las mareas.

2.2. Mareas terrestres

Hasta aquí se han considerado las mareas teóricas donde se ha supuesto que la Tierra es un cuerpo rígido. Sin embargo, la Tierra no es totalmente rígida y las fuerzas de las mareas producen, además del movimiento del mar, deformaciones de la Tierra sólida. A estas deformaciones se las denomina mareas terrestres.

Si la Tierra fuese un cuerpo totalmente elástico, se deformaría debido a las fuerzas de las mareas y recobraría inmediatamente su forma original. Estas deformaciones serían proporcionales a la fuerza de marea. Sin embargo, el material de la Tierra no es perfectamente elástico con lo que

esta tarda en adaptarse en cada momento a la fuerza de mareas. Debido al retraso en volver a su posición de equilibrio, el material contenido entre la superficie original y la de la Tierra deformada produce una fuerza cuyo momento tiende a retardar el giro de la Tierra, fricción de la marea lunar, pero a su vez, también tiende a acelerar el movimiento de revolución de la Luna alrededor de la Tierra. De modo que el momento angular total del sistema se conserva.

2.3. Modelos para el cálculo del potencial teórico de las mareas terrestres para una tierra rígida

El potencial de las mareas terrestres es difícil de calcular debido a que no se conocen perfectamente las propiedades elásticas de la Tierra. Pero se puede calcular de forma aproximada como veremos a continuación. Para ello, tenemos en cuenta los distintos tipos de altura de la marea que nos podemos encontrar:

- Altura de la marea sobre la Tierra rígida, $\bar{\xi}$
- Altura de la marea sobre la Tierra deformada, ξ^*
- Altura de la marea terrestre, ξ

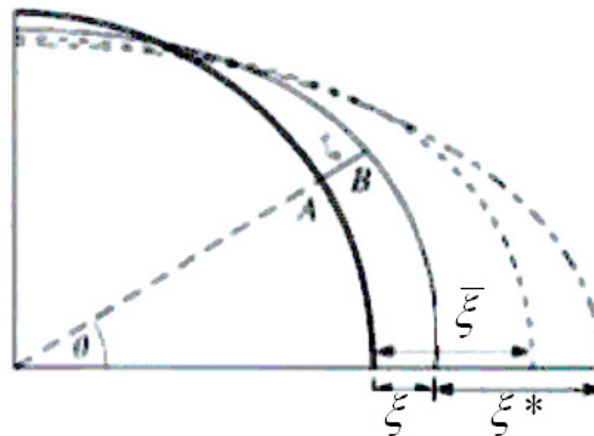


Figura 2. Alturas de la marea estática. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

Antes de calcular el potencial en un punto de la superficie, consideramos al astro causante de la marea estacionario con respecto a la Tierra rígida, cubierta por una capa líquida uniforme de densidad muy pequeña, y resolveremos el problema desde el punto de vista estático.

Así, la capa líquida tomará la forma de una superficie equipotencial que no varía con tiempo, suma del potencial gravitatorio de la Tierra V y del potencial de la marea ψ producida por el astro:

$$W = V + \psi = cte$$

Si V_0 es el potencial de la superficie de la Tierra rígida, el potencial en un punto A en el que actúa la fuerza de marea será:

$$V_A = V_0 + \bar{\xi} \frac{\partial V}{\partial r} = V_0 - \bar{\xi} g$$

Ya que: $\frac{\partial V}{\partial r} = -\frac{GM}{r^2} = -g$

Y si tomamos $V + \psi = V_0 = cte$ obtenemos la altura de la marea estática de equilibrio para una Tierra rígida:

$$\bar{\xi} = \frac{\psi}{g}$$

Como la posición del astro es estacionaria, $\bar{\xi}$ es función del ángulo θ . Si la Tierra girase alrededor del astro y sobre sí misma, el ángulo θ sería función del tiempo y por tanto, $\bar{\xi}$ también variaría con el tiempo.

Para una Tierra deformable, el potencial sobre un punto B sobre la superficie de la Tierra deformada será función de la altura de la marea terrestre ξ :

$$W = V + \frac{\partial V}{\partial r} \xi + \psi + V^*$$

Donde V^* es el potencial debido a la masa contenida entre la superficie original y la de la Tierra deformada.

Al no conocerse perfectamente la elasticidad de la Tierra, para determinar el potencial y la altura de la marea terrestre, se utilizan unos coeficientes que relacionan la deformación de la Tierra con el potencial de la marea para la Tierra rígida. Estos coeficientes son los denominados números h y k de Love y número l de Shida.

Los números de Love h y k , son funciones de r (distancia al centro de la Tierra), siendo constantes en la superficie. Son adimensionales y dependen del grado de expansión de los armónicos esféricos.

- El número h representa la proporcionalidad entre la altura de la marea terrestre ξ y la altura de la marea de equilibrio de la Tierra rígida $\bar{\xi}$: $\xi = h\bar{\xi} = h\frac{\psi}{g}$
- El número k representa la proporcionalidad entre el potencial de la marea para una Tierra rígida y el potencial de la masa contenida entre la superficie original y la de la Tierra deforma debida a la marea terrestre: $V^* = k\psi$

De modo que en el punto B que se encuentra en la superficie de la Tierra deformada, el potencial se puede escribir como:

$$W = V + \frac{\partial V}{\partial r}\xi + \psi + V^* = V - g\xi + \psi + V^* = V - h\psi + \psi + k\psi = V + (1 + k - h)\psi$$

Y el potencial debido a la marea de la Tierra deformable es: $\psi_D = (1 + k - h)\psi = \gamma\psi$

Y la altura de la marea sobre la Tierra deformable ξ^* es: $\xi^* = \gamma\bar{\xi} = (1 + k - h)\bar{\xi}$

$\gamma = (1 + k - h) = \frac{\xi^*}{\bar{\xi}}$ es llamado factor de disminución. Es el cociente entre la altura de la marea de la Tierra deformable y la altura teórica de la Tierra rígida.

Diferenciando ψ se obtiene la aceleración de la marea para una Tierra rígida. La componente

radial (positiva hacia fuera) es: $b_r = \frac{\partial \psi}{\partial r} = \frac{2}{r}\psi$

Lo que indica que la gravedad de la Tierra está afectada por la componente radial de la aceleración de la marea, que es positiva hacia fuera causando una disminución de la gravedad de la Tierra:

En una Tierra deformable: $b_{rD} = \frac{\partial \psi_D}{\partial r} = (1 - \frac{3}{2}k + h)b_r$

La componente tangencial en una Tierra deformable será:

$$b_{\theta D} = -\frac{1}{r} \frac{\partial \psi_D}{\partial \theta} = (1 + k - h)b_{\theta}$$

La desviación de la vertical $\bar{i}$, debida al efecto de la marea en una Tierra rígida vendrá dada

por: $\bar{i} \cong tg \bar{i} = \frac{1}{gr} \frac{\partial \psi}{\partial \theta}$

En una Tierra deformable: $i = \frac{1+k}{gr} \frac{\partial \psi}{\partial \theta}$

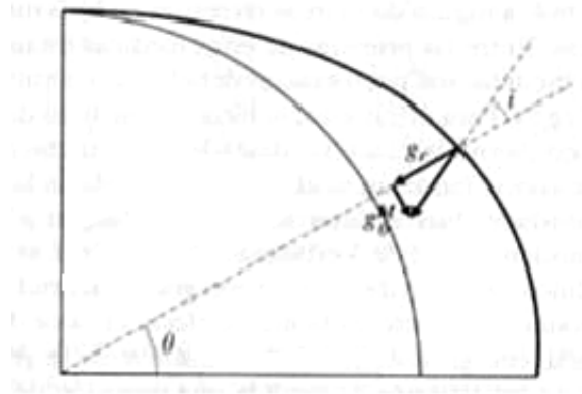


Figura 3. Desviación de la vertical debida a la marea. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

La desviación de la vertical debida a las deformaciones horizontales del suelo en las mareas

terrestres, i' viene dada por: $i' = \frac{l}{gr} \frac{\partial \psi}{\partial \theta}$

Siendo l el número de Shida: $l = \frac{i'}{i}$

De modo que la desviación total de la vertical debida a la marea terrestre es:

$$i - i' = (1 + k - l) \frac{1}{gr} \frac{\partial \psi}{\partial \theta}$$

2.4. Cálculo de los coeficientes de marea

El potencial de la marea $\psi = \bar{G}(r) \left(\frac{C}{R} \right)^3 \left(\cos 2\theta + \frac{1}{3} \right)$ es función de la posición del astro en su órbita con respecto a la Tierra. Para estudiar su variación con el tiempo, expresaremos el ángulo

cenital θ en función de las coordenadas geocéntricas (ϕ, λ) del punto donde lo estamos calculando.

Para ello se tiene en cuenta la esfera celeste y se resuelve el triángulo esférico con vértices en el Polo Norte Celeste, la Luna y el punto P (proyección de la vertical de un punto de la superficie terrestre de coordenadas (ϕ, λ) sobre la esfera celeste). Así, el ángulo θ se puede expresar en función de la latitud geocéntrica ϕ del punto atraído P, la declinación δ_L y el ángulo horario τ de la Luna.

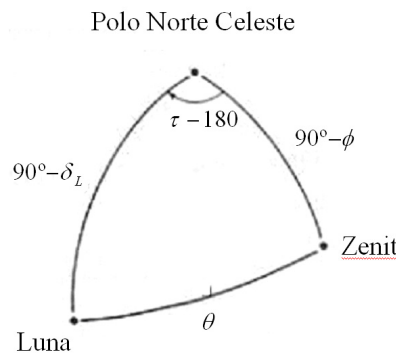


Figura 4. Triángulo astronómico

$$\cos \theta = \sin \delta_L \sin \phi - \cos \delta_L \cos \phi \cos \tau$$

Sustituyendo en el potencial de la marea, operando y teniendo en cuenta las relaciones trigonométricas, se llega a que el potencial de marea de la Luna es:

$$\psi_L = \bar{G}(r) \frac{2}{3} \left(\frac{C_L}{R_L} \right)^3 \left[3(\sin \delta_L \sin \phi - \cos \delta_L \cos \phi \cos \tau)^2 - 1 \right]$$

Esta función se puede separar en tres términos de modo que:

$$\psi = \psi_0^L + \psi_1^L + \psi_2^L \begin{cases} \psi_0^L = 3A \left(\frac{1}{3} - \sin^2 \delta_L \right) \left(\frac{1}{3} - \sin^2 \phi \right) \\ \psi_1^L = -A \sin 2\delta_L \sin 2\phi \cos \tau \\ \psi_2^L = A \cos^2 \delta_L \cos^2 \phi \cos 2\tau \end{cases}$$

Siendo $A = G \left(\frac{C}{R} \right)^3$

Estos tres términos tienen distintas características, ya que las cantidades R , δ_L y τ varían con distintos periodos (el periodo de τ es de un día lunar medio, 24h 50m 47s):

ψ_0^L : son las mareas de largo periodo, de 14 días para la Luna. Este término no depende de la rotación de la Tierra y consta de una parte no periódica que ocasiona una deformación permanente de las superficies de nivel (estas son disminuidas en los polos en aproximadamente 0,26 m y elevadas en el ecuador unos 0,13m, ya que hay una disminución permanente de la gravedad en el ecuador de aproximadamente $0.3 \mu\text{m/s}^2$ y un aumento en los polos de unos $0.61 \mu\text{m/s}^2$).

ψ_1^L : término que oscila con periodos diurnos.

ψ_2^L : es el componente semidiurno. Su periodo cambia cada 12 horas. Es mínimo en los polos y máximo en el ecuador.

De todos ellos, el término más importante en las mareas es el semidiurno (ya que δ_L es pequeño y $\sin 2\delta_L < \cos^2 \delta_L$)

De la misma manera, se obtienen estas ecuaciones para el efecto de marea producido por el Sol. Sólo hay que tener en cuenta que se debe sustituir τ por t , el tiempo solar, cuya periodicidad es de un día solar medio (24h). En este caso, el periodo de ψ_0^S es de 0.5 años.

2.5. Medida de las mareas terrestres

Como se ha visto, las mareas terrestres afectan tanto al módulo de la gravedad $\vec{g}$ como a las desviaciones de la vertical.

Las variaciones en el módulo de $\vec{g}$ se suelen medir con gravímetros, las fluctuaciones en la dirección de la línea de la plomada se determinan con clinómetros y las deformaciones de la corteza se miden con extensómetros. Todos ellos tienen que alcanzar una gran precisión para

poder estudiar adecuadamente el efecto de las mareas terrestres. Por ello, muchos de los aparatos suelen estar conectados a un registro continuo para detectar las variaciones temporales obteniéndose precisiones por encima de 10^{-8} m/s^2 :

- Gravímetros registradores: pueden ser gravímetros de campo unidos con una unidad de grabación o gravímetros de mareas especiales (como es el Lacoste-Romberg, Geodynamics, etc). En todos ellos, los cambios en la posición de la masa se convierten a una tensión eléctrica y se amplifica y se graba junto con el tiempo. Con ellos se puede obtener una precisión relativa de 10^{-3} .
- Gravímetros superconductores: Utilizan los efectos de la superconductividad para mantener una bolita en equilibrio mediante un campo magnético. Los desplazamientos verticales, causados por variaciones de gravedad, son detectados y compensados a través de un sistema de retroalimentación, tomando con ello una medición. En estos gravímetros, las perturbaciones microsísmicas y los efectos de deriva son muy pequeñas. La precisión que se obtiene es superior a 10^{-9} m/s^2 .
- Clinómetros: miden inclinaciones. Son péndulos horizontales y verticales y las fluctuaciones en la dirección de la línea de la plomada con respecto a la superficie terrestre se determinan de sus dos componentes mutuamente perpendiculares (NS, EW).
 - Péndulos horizontales: consisten en dos hilos verticales que sostienen una barra horizontal con una masa adherida. Debido a la pequeña inclinación del eje de rotación con respecto a la dirección de la vertical, una fuerza horizontal ocasiona una deflexión angular fuertemente amplificada, teniendo con ello una medida.
 - Péndulos verticales: estos péndulos están suspendidos de tal manera que pueden oscilar libremente y las deflexiones se perciben por dos detectores colocados en ángulos rectos entre sí, que después de amplificadas, se graban por medios analógicos y digitales.
- Extensómetros: miden deformaciones. Para determinar las seis componentes del tensor de deformación, los extensómetros deben orientarse en diferentes direcciones espaciales. También miden directamente los valores del número de Love h y del número de Shida l y en especial el cociente l/h . Hoy en día, pueden ser barras de superinvar con longitudes de entre 10 y 20 m o interferómetros láser que permiten longitudes de hasta de 1km. En cada caso, se tiene un extremo fijo a la roca y otro libre donde se miden las deformaciones de la corteza terrestre.

Una vez recogida la señal por estos aparatos, hay que tener en cuenta que a la señal de la marea terrestre se le superpondrán otros efectos como son:

- Errores instrumentales sistemáticos: incertidumbres asociadas con la determinación de la calibración, efectos directos de la presión atmosférica y la temperatura, etc. Muchos de ellos se pueden determinar y eliminar, en la medida de lo posible, mediante análisis de regresión.
- La deriva del instrumento: influye sobre todo en el análisis de las mareas de largos periodos.
- Los efectos causados por las mareas oceánicas. Son difíciles de evaluar ya que no se conocen suficientemente los parámetros elásticos de la corteza terrestre.
- Influencias locales de los lugares donde se realizan las medidas.
- Efectos secundarios que surgen de las variaciones de la presión atmosférica, la temperatura, las radiaciones solares, procesos tectónicos, etc.
- Etc.

Los gravímetros proporcionan buenos resultados cuando se colocan en compartimentos de temperatura controlada. Los clinómetros deben protegerse de las influencias de la superficie mediante una capa de roca suficientemente gruesa. Así, los péndulos horizontales se instalan en túneles y los péndulos verticales en agujeros perforados.

Actualmente existe en todo el mundo un gran número de estaciones de mareas terrestres. Los valores obtenidos muestran que la aproximación de una Tierra rígida no es adecuada. Por lo que utilizan modelos de Tierra más realistas que consideran una Tierra elíptica en rotación, con océanos y atmósfera y con unos parámetros elásticos en su interior de acuerdo con los últimos resultados de las investigaciones sismológicas.

Los resultados obtenidos de las observaciones de las mareas terrestres tienen un gran número de aplicaciones: reducir medidas geodésicas, medidas gravimétricas u otras medidas precisas como pueden ser el posicionamiento por satélite o los métodos radiointerferométricos; para establecer modelos geofísicos como pueden ser modelos regionales de corteza-manto o para la verificación de modelos globales y regionales de mareas oceánicas, entre otras.

Bibliografía

- [1] Torge, W. Geodesia. Editorial Diana, 1983.
- [2] Torge, W. Gravimetry. Editorial Walter de Gruyter 1989.
- [3] Udías, A. y Mezcuá, J. Fundamentos de Geofísica. Editorial Alianza, 1996.

Tema 3. Medidas absolutas y relativas de la gravedad. Métodos pendulares y de caída libre. Observación sobre móviles. Determinación de las segundas derivadas del potencial de la gravedad. Medida del gradiente de la gravedad. Calibración e intercomparación de gravímetros. Correcciones a las observaciones: mareas, movimiento del polo, carga oceánica y carga atmosférica.

3.1. Medidas absolutas y relativas de la gravedad

En las medidas absolutas se determina el valor total de la aceleración de la gravedad en el punto de observación (módulo de la componente vertical de la aceleración de la gravedad). Las primeras determinaciones de la longitud del péndulo fueron realizadas por Richer en La Cayenne en 1672. También Maupertuis 1736-37 y Bouguer 1735-44 (junto con Jorge Juan y Antonio de Ulloa) efectuaron medidas pendulares en los viajes para la determinación del achatamiento de la Tierra. Los marinos españoles Malaspina y Bustamante (1789) también realizaron experiencias con el péndulo invariable. En España la primera determinación de la gravedad que tenemos conocimiento fue realizada por Gabriel Ciscar en 1800 en Madrid con péndulos de lenteja. En el IGN Barraquer entre 1877 y 1883 mide la gravedad por el método absoluto pendular mediante el péndulo de Bessel. Actualmente el método de caída libre es el método absoluto más exacto y preciso de determinación de la componente vertical de la gravedad. En las medidas relativas se pueden distinguir dos casos:

- Espaciales: Se determina la diferencia o incremento de la aceleración de la gravedad del sensor del gravímetro relativo en un punto con respecto a la posición del sensor en otro punto inicial o de referencia, en el cual habitualmente se conoce el valor absoluto de la fuerza de la gravedad.
- Temporales: En las medidas relativas temporales se determinan los valores de la fuerza de la gravedad en instantes de tiempo distintos, como ocurre en el caso de la medida de las mareas terrestres (variaciones de largo período).

Las medidas relativas se realizan por Sterneck en 1881 y se extienden por el mundo para una mayor densificación de los valores de la gravedad hasta el desarrollo de los gravímetros relativos, hacia 1940. En España se realizan medidas desde 1901 hasta 1950. Desde la década de 1940 hasta la actualidad las medidas relativas se realizan mediante los gravímetros de muelle de cuarzo y de metal. También desde 1970, el gravímetro superconductor sirve para realizar

medidas relativas temporales, constituyendo el gravímetro relativo más preciso, aunque también el más complejo. Finalmente, los métodos de medida de aceleraciones de la gravedad se pueden dividir en terrestres, marinos, submarinos, aéreos y por satélite. Un resumen general de las precisiones y exactitudes se observa en la tabla 1.

SENSOR	Exactitud (mGal)	Exactitud (ms^{-2})	RESOLUCIÓN N	Comentarios
Gravimetría estática				
gravímetros relativos	$1-2 \cdot 10^{-2}$	$1-2 \cdot 10^{-7}$	valor en un punto	limitaciones por accesibilidad
gravímetros absolutos	$1-2 \cdot 10^{-3}$	$1-2 \cdot 10^{-8}$	valor en un punto	limitaciones por ruido ambiental
gravímetros superconductores	$1-2 \cdot 10^{-5}$	$1-2 \cdot 10^{-9}$	valor en un punto	limitado por ruido del entorno ambier
Gravimetría de plataforma móvil				
gravímetros aéreos	2-4 mGal	$2-4 \cdot 10^{-5}$	10-20 km	requiere amortiguamiento del sensor
gradiómetros aéreos	$5-10 \cdot 10^{-9} \text{ s}^{-2}$		1-10 km	no operacional
gravímetros marinos	0.2-0.4 mGal	$0.2-0.4 \cdot 10^{-5}$	< 1km	extensión de área y densificación cost
gradiómetros marinos	$5-10 \cdot 10^{-9} \text{ s}^{-2}$		< 1km	utilizada por la US Navy
sistemas inerciales	1-4 mGal	$1-4 \cdot 10^{-5}$	< 1km	deriva del giróscopo es el problema p
Métodos por satélite				
Seguimiento terrestre de satéli	0.3-0.5 m		500 km	
Seguimiento satélite a satélite	1-4 mGal		100-200 km	
Gradiometría de satélite	1-2mGal		50-100 km	
Altimetría de satélite	<0.15 m		5-30 km	

Tabla 1. Métodos de medida de la gravedad.

3.2. Métodos pendulares y caída libre

3.2.1. El péndulo simple o matemático

La ecuación del péndulo simple o matemático de longitud l es:

$$m \frac{d^2 \alpha}{dt^2} = -mg \sin \alpha$$

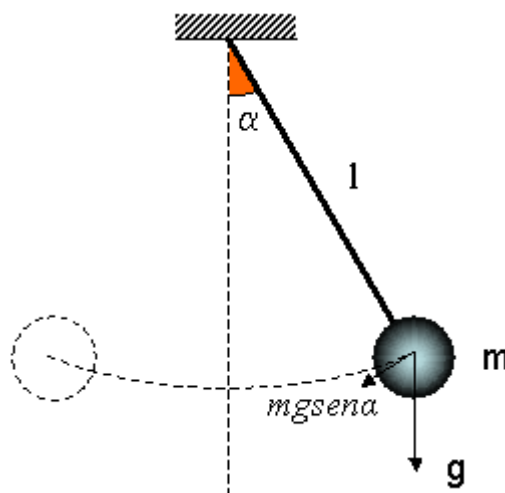


Figura 1. Péndulo simple.

y su solución, en ausencia de rozamiento, para $\text{sen } \alpha = \alpha$: $\alpha = \alpha_0 \text{sen}(\omega t + \varphi_0)$

En general un péndulo simple o matemático tiene un periodo en función del ángulo α :

$$T = 2\pi \sqrt{\frac{l}{g}} \left[1 + (1/2)^2 \text{sen}^2 \frac{\alpha}{2} + (1 \cdot 3 / 2 \cdot 4)^2 \text{sen}^4 \frac{\alpha}{2} + \dots + (2n-1 / 2n)^2 \text{sen}^{2n} \frac{\alpha}{2} + \dots \right]$$

Y en función de la longitud del péndulo y de la máxima altura b de amplitud máxima:

$$T = 2\pi \sqrt{\frac{l}{g}} \cdot \left[1 + (1/2)^2 \frac{b}{2l} + \dots + \frac{2n-1}{2n} \left(\frac{b}{2l} \right)^n + \dots \right]$$

En aproximación para oscilaciones infinitesimales $\alpha < 5^\circ$, $\alpha < 1^\circ$, $\alpha < 30$ minutos de arco, para $b=0$, respectivamente:

$$T = 2\pi \sqrt{\frac{l}{g}} \left[1 + (1/2)^2 \text{sen}^2 \frac{\alpha}{2} \right] \approx 2\pi \sqrt{\frac{l}{g}} \left[1 + \frac{\alpha}{16} \right] \approx 2\pi \sqrt{\frac{l}{g}}$$

Teniendo en cuenta que para oscilaciones pequeñas $b = l - l \cos \alpha = l(1 - \cos \alpha) = 2l \text{sen}^2(\alpha/2)$.
Despejando en cada caso de oscilación la gravedad g obtendremos la fórmula de obtención del valor de la gravedad.

3.2.2. El péndulo compuesto o físico

La ecuación del péndulo compuesto o físico es la misma que para el péndulo simple, pero substituyendo la masa por el momento de inercia del sólido rígido I , y la distancia del centro de gravedad al centro de rotación h :

$$I \frac{d^2 \alpha}{dt^2} = -mg h \text{sen} \alpha$$

y su solución, en ausencia de rozamiento, es para ángulos pequeños ($\text{sen} \alpha = \alpha$):

$$\alpha = \alpha_0 \text{sen}(\omega t + \varphi_0)$$

$$\text{con periodo: } T = 2\pi \sqrt{\frac{I}{mgh}} \cdot \left[1 + (1/2)^2 \frac{b}{2h} + \dots + \frac{2n-1}{2n} \left(\frac{b}{2h} \right)^n + \dots \right]$$

En aproximación para oscilaciones infinitesimales $\alpha < 5^\circ$, $\alpha < 1^\circ$, $\alpha < 30$ minutos de arco, para $b=0$, respectivamente:

$$T = 2\pi \sqrt{\frac{I}{mgh}} \left[1 + (1/2)^2 \text{sen}^2 \frac{\alpha}{2} \right] \approx 2\pi \sqrt{\frac{I}{mgh}} \left[1 + \frac{\alpha}{16} \right] \approx 2\pi \sqrt{\frac{I}{mgh}} = 2\pi \sqrt{\frac{l_e}{g}}$$

donde l_e , la longitud equivalente del péndulo matemático, es función de I el momento de inercia, h la distancia del centro de gravedad al centro de suspensión del péndulo y m la masa del péndulo.

El método de Kater (péndulo reversible) consistía en medir dos periodos con ejes en dos puntos simétricos respecto del centro de gravedad del péndulo físico, determinado de esta manera la longitud equivalente o longitud del péndulo simple equivalente.

$$l = \frac{I_c}{mh} = \frac{I}{mh} + h$$

El método de Bessel consistía en medir dos periodos similares de las oscilaciones, así como las distancias entre los ejes de suspensión, ejes de oscilación y centro de gravedad. De esta manera se obtenía el periodo de oscilación del lugar:

$$T^2 = \frac{\lambda_1 T_1^2 - \lambda_2 T_2^2}{\lambda_1 - \lambda_2}$$

Las medidas consistían en los pasos siguientes:

- Distancia entre los ejes de suspensión λ (entre los filos de los cuchillos), en partes de la longitud de la regla colocada verticalmente en el aparato.
- Duración de la oscilación isócrona circular (T_1 , T_2)
- Situación del centro de gravedad (λ_1 , λ_2) del péndulo respecto de los filos de los cuchillos.

Otras medidas y correcciones a realizar en los métodos pendulares eran las constantes termométricas, amplitud finita de la oscilación, acortamiento de la regla por su propio peso, longitud absoluta de la regla, marcha del reloj, movimiento oscilatorio o balance del sostén, flexión del péndulo, deformación de los filos de los cuchillos por el peso del péndulo, reducción al vacío, y reducción al nivel del mar para comparar medidas en distintos lugares.

3.2.3. Método pendular relativo

Teniendo en cuenta que $\Delta T = T_2 - T_1$, y la relación de periodos distintos del péndulo invariable en dos lugares distintos se obtiene:

$$g_2 = g_1 \frac{T_1^2}{T_2^2} = g_1 \frac{1}{\left(1 + \frac{\Delta T_1}{T_2}\right)^2} = g_1 \left[1 - 2 \frac{\Delta T}{T_1} + 3 \frac{\Delta T^2}{T_1^2} - 4 \frac{\Delta T^3}{T_1^3} + \dots \right]$$

$$\Delta g = g_2 - g_1 = \left[-2 \frac{\Delta T}{T_1} + 3 \frac{\Delta T^2}{T_1^2} - 4 \frac{\Delta T^3}{T_1^3} + \dots \right]$$

Lo cual nos indica que conociendo el valor de la gravedad en un lugar y la diferencia de periodos entre dos lugares podemos deducir el valor de la gravedad en el segundo emplazamiento. El Instituto Geográfico Nacional realizó medidas relativas pendulares entre 1903 y 1942, completando una red de estaciones gravimétricas por toda España, Portugal y

Marruecos (alrededor de 210), que permitieron el cálculo de los primeros mapas de España de anomalías de la gravedad de aire libre y de Bouguer.

3.2.4. Gravímetros relativos

Un gravímetro relativo es un instrumento, generalmente de muelle de cuarzo o de aleación de metal, que permite la medida relativa espacial y temporal de la componente vertical de la gravedad. El gravímetro Worden fue desarrollado por Sam Worden al final de la década de 1940. Basado en un sistema elástico de muelle de cuarzo, posee tres muelles para conseguir la longitud cero y una masa de unos 5 miligramos. El primero que permitió un gran número de observaciones relativas de calidad fue el gravímetro Lacoste-Romberg modelo G. A partir de una idea de 1934 de Lucien Lacoste, con patente en 1942 y 1945, el gravímetro de muelle de metal astatizado de longitud cero, que soporta una masa, registra elongaciones que son proporcionales a las diferencias de gravedad. Mediante un sistema electrónico capacitor conectado a un voltímetro externo, en los primeros tiempos analógico y posteriormente digital, permite mejorar sustancialmente las medidas mecánicas del dial óptico llegando a incertidumbres de algunos μGal . Las incertidumbres dependen de su correcta tabla de calibración y factores de calibración. El método empleado es un método de cero en el que la posición de equilibrio define la posición horizontal del haz en la dirección de la masa que pende del muelle y que es objeto de atracción (figura 2).

La ecuación de equilibrio resultado del equilibrio de fuerza de gravedad y la restauradora del muelle de constante k es:

$$mga \cos \alpha - k(s - s_0)b \sin \theta = 0$$

siendo la sensibilidad en función del ángulo $\left(\frac{\Delta \alpha}{\Delta g} \right) = \frac{\tan \alpha}{g}$

y la sensibilidad en función de la elongación $\left(\frac{\Delta s}{\Delta g} \right) = \frac{m}{k} \frac{a}{bd} \frac{s^2}{s_0}$

$$\text{entonces } g = \frac{k}{m} \frac{db}{a} \frac{s - s_0}{s}$$

El rango de medida llega a ser mundial (7000 mGal) al variar la distancia d entre el eje de giro y el punto de suspensión. Las derivas en un gravímetro recién construido son de 1 mGal al mes, bajando a 0,5 mGal al mes transcurrido cierto tiempo. Derivas estáticas de 15 $\mu\text{Gal/día}$ y menores son habituales en instrumentos envejecidos.

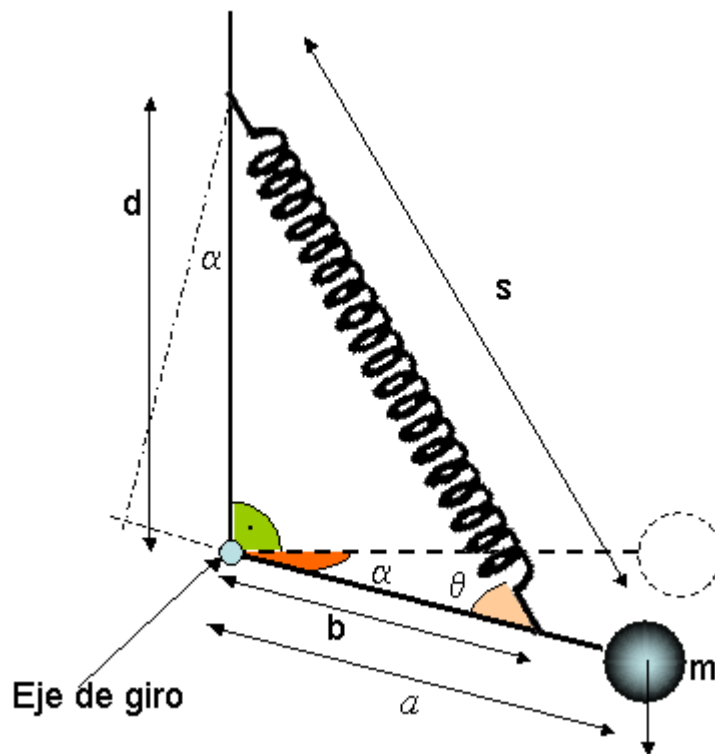


Figura 2. Esquema de un gravímetro relativo de muelle.

Dentro de los gravímetros relativos cabe citar el mencionado gravímetro superconductor. El fundamento es la levitación en condiciones de superconductividad en un campo permanente magnético de una pequeña esfera de 25 mm de diámetro, inicialmente de aluminio recubierta de plomo, actualmente de niobio. La superconductividad se consigue situando el dispositivo dentro de un vaso Dewar con vacío y condiciones criogénicas de temperatura (4 K). Un cambio en la gravedad induce un movimiento vertical en la esfera; para mantenerse en el nivel “cero” necesita de un voltaje de realimentación que será proporcional al cambio de gravedad. Para convertir voltajes en gravedad se precisa de un gravímetro absoluto que los calibre, con precisiones en el factor de calibración mejores que 0,1 %. Son instrumentos que se instalan y se mantienen permanentemente en observatorios o laboratorios.

3.2.5. Método absoluto de caída libre

La ecuación diferencial que define a una partícula en caída libre en ausencia de rozamiento, en primera aproximación lineal, considerando un campo no homogéneo y un gradiente constante γ será:

$$\ddot{z} = g_o + \gamma(z - z_o)$$

$$\left\{ \begin{array}{l} \ddot{z} - \gamma z = g_o - \gamma z_o \\ z(0) = z_o; \dot{z}(0) = v_o; \ddot{z}(0) = g_o; \end{array} \right\}$$

donde g_o es un valor inicial de la aceleración de la gravedad, la posición inicial y la velocidad inicial del grave en caída libre.

La solución n-ésima en el gradiente es:

$$z(t) = \sum_{n=0}^{\infty} a_n t^n = z_o + v_o \sum_{n=0}^{\infty} \frac{\gamma^n t^{2n+1}}{(2n+1)!} + g_o \sum_{n=0}^{\infty} \frac{\gamma^n t^{2n+2}}{(2n+2)!}$$

En general la solución anterior se puede expresar como un polinomio:

$$z(t) = \sum_{n=0}^{\infty} a_n t^n = a_o + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4$$

Truncando hasta el primer orden en el gradiente, la solución de la ecuación es:

$$z(t) = z_o + v_o t + \frac{g_o}{2} t^2 + v_o \frac{\gamma}{6} t^3 + g_o \frac{\gamma}{24} t^4$$

Por tanto si consideramos solamente los términos cuadráticos, y diferenciamos $ds = \partial s / \partial g \, dg + \partial s / \partial t \, dt$. Si queremos errores relativos no mayores de 10^{-9} , la distancia de 0.5 m y tiempos de caída de 0.3 s nos llevan a requerir exactitudes de ± 0.5 nm y ± 0.2 ns. Por tanto, solamente empleando métodos interferométricos o atómicos, con patrones de frecuencia de relojes atómicos y longitudes de onda de láser, es posible alcanzar las exactitudes mencionadas.

Se realiza en cada caída un ajuste mínimo cuadrático con un gradiente vertical de la gravedad conocido a priori γ_o . En forma matricial para cada caída $n=1,2, \dots, n$:

$$\begin{bmatrix} z_1 \\ z_o \\ z_o \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} 1 & (t_1 + \frac{\gamma}{6} t_1^3) & (t_1^2 + \frac{\gamma}{24} t_1^4) \\ 1 & (t_2 + \frac{\gamma}{6} t_2^3) & (t_2^2 + \frac{\gamma}{24} t_2^4) \\ 1 & (t_3 + \frac{\gamma}{6} t_3^3) & (t_3^2 + \frac{\gamma}{24} t_3^4) \\ \vdots & \vdots & \vdots \\ 1 & (t_n + \frac{\gamma}{6} t_n^3) & (t_n^2 + \frac{\gamma}{24} t_n^4) \end{bmatrix} \begin{bmatrix} z_o \\ v_o \\ g_o \end{bmatrix}$$

Comenzando con un sistema interferométrico de luz blanca en el año 1962, los predecesores del gravímetro absoluto de caída FG5 fueron seis gravímetros de la serie JILAG construidos en 1985, con diversas instituciones apoyando el proyecto (NIST, DMA, NOAA, GSC, FGI, Universidades de Hannover y de Viena). El principio de este tipo de gravímetros es reproducir la aceleración libre de un cuerpo grave en caída en el campo gravitatorio, al mismo tiempo que

se miden los pares distancia-tiempo (t_i , x_i) mediante un contador de intervalos de tiempo y un interferómetro láser.

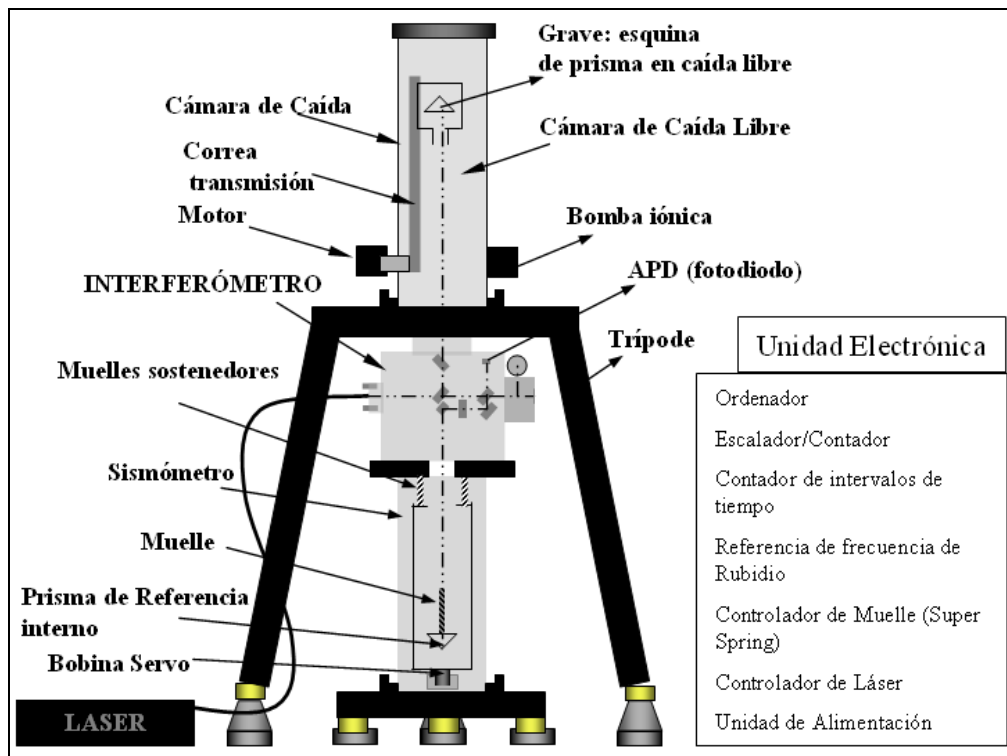


Figura 3. Esquema de los elementos principales que componen el gravímetro absoluto FG5.

El láser de He-Ne estabilizado en frecuencia por una célula de Iodo, juega un papel de patrón físico primario y (en principio) no requeriría calibración. Dentro de la cámara de caída libre, donde es necesario realizar el vacío por debajo del nivel de 10^{-4} Pa, se precipita una esquina de prisma cúbico (espejo retroreflector), que genera franjas de interferencias láser dentro de un interferómetro del tipo Mach-Zender (en modelos anteriores del tipo Michelson). Puesto que la masa de unos 200 g cae durante unos 20 cm (unos 600 nms^{-2}), hay que promediar de alguna manera el valor final a la altura efectiva de cálculo de la gravedad. El conjunto de fuentes de error en cada una de las variables que influyen en la medida absoluta de la gravedad por el método de caída libre se cifra en una incertidumbre instrumental de $1.1 \mu\text{Gal}$ (tabla 2). El ruido microsísmico, incluyendo el generado por el ser humano, afecta a la distancia al prisma de referencia, con periodos de ruido entre 100 s y 0,01 s. Un sismómetro de largo periodo de 60 s, cuya masa inercial soporta el prisma fijo de referencia, reduce el ruido considerablemente. Es una mejora considerable respecto al gravímetro de la serie JILA. La rotación del prisma móvil es causa de errores en la longitud del camino óptico medido. Se sospecha que esta rotación se debe a la flexión de la cámara de caída libre al liberar la masa. La rotación del prisma móvil no debe superar los 0,03 rad/s. El grave que cae está construido de tal manera que su centro de

gravedad y centro óptico coinciden dentro de los límites de $2.5 \cdot 10^{-5}$ m más una cierta incertidumbre debida a la imperfección en la determinación de esta distancia. Las tasas de rotación medidas son del orden de 10 mrad/s. El error final en el valor de la gravedad es proporcional al error en la medida de la distancia entre el centro de gravedad y el centro óptico del prisma.

Fuente de error	Incertidumbre	Comentarios
Presión residual	0,1	Depende de presión
Temperatura diferencial	0,1	Depende de temperatura y presión.
Grad. campo magnético	0,1	Difícil de estimar
Electrostática	0,1	Difícil de estimar
Atracción del instrumento	0,1	Desviación fija en diseño
Verticalidad	0,1	Siempre negativa
Rotación prisma	0,3	Se degrada con el tiempo
Longitud onda laser	0,1	Laser con celda de yodo
Efecto de Coriolis	0,4	Depende latitud y orientación instrumento
Inclinación y retroceso	0,1	Depende de la estación
Modulación de “air-gap”	0,6	Depende estacionamiento
Cambio de fase por electrónica	0,6	
Patrón de frecuencia	0,2	
Bordes de lentes	0,3	
Límite de difracción	0,2	Depende del láser
Incertidumbre total	1,1	μGal (11 nms^{-2})

Tabla 2. Incertidumbres principales que componen la incertidumbre instrumental del gravímetro absoluto FG5.

3.3. Observación sobre móviles

Existe la necesidad de medir la gravedad en todo el globo debido a que el 70% de la superficie terrestre está cubierta por agua. Por otra parte, también en otros medios como en el interior de los océanos (submarinos), el aire o incluso los satélites artificiales se hace útil la medición para un conocimiento completo del campo de gravedad. La gravimetría aérea comprende la gravimetría medida en los aviones y en los helicópteros. Las mediciones en submarino por razones militares se realizaron ya desde 1940 entre profundidades de 200 y 1000 m. En general las medidas se hacen integrando a un intervalo de tiempo que depende de las aceleraciones de perturbación características de cada plataforma móvil. Hacia 1929, Vening-Meinesz diseñó el doble péndulo y en 1939 el triple péndulo para la medición de la componente vertical, eliminando las aceleraciones horizontales no deseadas.

Para cada péndulo hemos de aplicar la ecuación diferencial, en ausencia de aceleraciones horizontales:

$$\frac{d^2\theta_i}{dt^2} + \frac{g}{l}\theta_i = 0$$

Y con aceleraciones:

$$\frac{d^2\theta_i}{dt^2} + \frac{g}{l}\theta_i = -\frac{l}{l} \frac{d^2 y}{dt^2}$$

Las aceleraciones horizontales se anulan al restar las ecuaciones correspondientes a los péndulos θ_1 y θ_2 :

$$(\ddot{\theta}_1 - \ddot{\theta}_2) + \frac{g}{l}(\theta_1 - \theta_2) = 0$$

La integración sobre un periodo de tiempo permite determinar el valor de la gravedad eliminando las aceleraciones horizontales.

En los métodos modernos, la ecuación general de la oscilación forzada de un gravímetro en una plataforma móvil es:

$$\ddot{z} + 2\delta\dot{z} + \omega_o^2 z = g + a_z - f$$

en donde, a_z es la aceleración perturbadora vertical (en dirección de la línea de la plomada), δ es el factor de amortiguamiento, ω_o la frecuencia propia del resorte y f la contrafuerza del resorte. Los gravímetros o acelerómetros se suelen situar en el metacentro del barco o en, general, en el centro de gravedad de móvil, para que las aceleraciones perturbadoras sean mínimas.

Respecto de las observaciones en tierra firme existen tres diferencias fundamentales:

- Nivelación de una base móvil.
- Existencias de aceleraciones de perturbación.
- Aparición de aceleraciones de Coriolis.

La nivelación de la base móvil se puede realizar a través de una suspensión tipo Cardan o nivelando la plataforma estabilizada con giróscopo. En el caso de la suspensión Cardan, el sistema se alinea automáticamente con la resultante de la gravedad $g(t)$, la magnitud medida es:

$$g^*(t) = [g^2(t) + h^2(t)]^{1/2}$$

en donde la composición de las aceleraciones horizontales es $h(t) = (a_x + a_y)^{1/2}$. Este sistema quedó obsoleto a mediados de los años 60 del pasado siglo. Los sistemas de nivelación estabilizados por giróscopo o GPS pueden dar alineaciones con la vertical del orden de $\pm 1'$ de arco o mejores. Si el ángulo de inclinación es $v(t)$ y la aceleración de perturbación horizontal $h(t)$, la gravedad medida se puede expresar como:

$$g^* = g \cos v(t) + h(t) \operatorname{sen} v(t)$$

$$h(t) = h_o \operatorname{sen} \omega t$$

$$v(t) = v_o \operatorname{sen}(\omega t + \varepsilon)$$

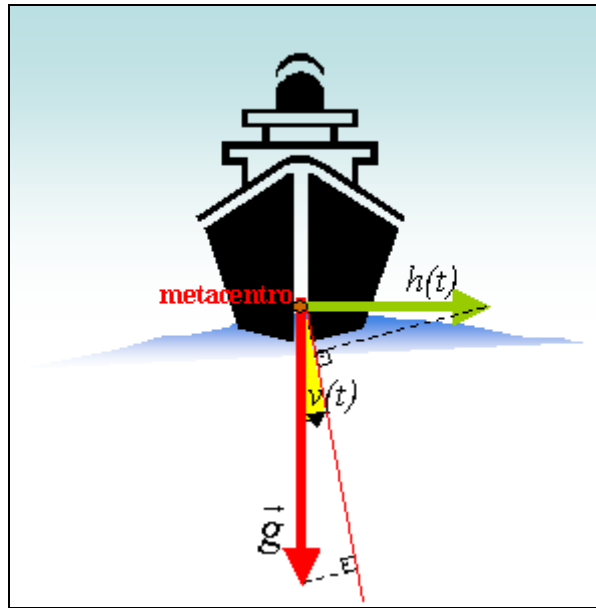


Figura 4. Esquema del error de nivelación debido a la aceleración horizontal $h(t)$ e inclinación residual $v(t)$ en una plataforma móvil.

Para ángulos pequeños el error de desnivelación se puede expresar:

$$\delta g = g - g^* = g \frac{v^2}{2} - h(t)v(t)$$

en donde el primer término es el error estático de inclinación y el segundo es el error dinámico.

Realizando una integración a un período T, la corrección queda:

$$\delta g = g - g^* = g \frac{v_o^2}{4} - \frac{h_o v_o}{2} \cos \varepsilon$$

Si las aceleraciones de perturbación horizontales y verticales tienen el mismo período se pueden producir efectos de acoplamiento, que también habrá que corregir.

Debido al movimiento de rotación de la Tierra se inducen aceleraciones que afectan a la medida de la componente vertical de la aceleración de la gravedad en plataformas móviles. La aceleración de Coriolis tiene valores máximos en el movimiento según los paralelos y mínimos si el movimiento se realiza según los meridianos. Al aumentar la latitud la corrección disminuye. La aceleración de Coriolis debida al movimiento del móvil con velocidad v en dirección α ha de ser corregida, y puede ser evaluada con:

1. La expresión para la aproximación esférica:

$$\delta g = 2\omega v \cos \varphi \sin \alpha + \frac{v^2}{r} \approx 4.0v \cos \varphi \sin \alpha + 0.0012v^2 \quad (10^{-5} \text{ ms}^{-2})$$

2. Aproximación elipsoidal (utilizada en aviones):

$$\delta g = \left(1 + \frac{h}{a}\right) \left(2\omega v \cos \varphi \sin \alpha + \frac{v^2}{r}\right) - f \frac{v^2}{a} (1 - \cos^2 \varphi (3 - 2\sin^2 \alpha))$$

Los gravímetros marinos (como el LCR) tienen un rango de unos 12000 mGal y una deriva de 1 mGal/mes. Actualmente los errores de orientación y navegación han disminuido enormemente

por la utilización del posicionamiento GPS. Las precisiones que se obtienen mediante estos métodos oscilan entre 1 y 0.1 mGal. Las incertidumbres que tienen los distintos métodos de medición de la gravedad sobre plataformas móviles se pueden resumir en (tabla 1):

- Gravímetros sobre el mar $\pm 0.5 \cdot 10^{-5} \text{ ms}^{-2}$ con resoluciones del orden de 1-2 km.
- Gravímetros submarinos alcanza entre ± 1 y $\pm 3 \cdot 10^{-5} \text{ ms}^{-2}$.
- Helicópteros ± 1 y $\pm 3 \cdot 10^{-5} \text{ ms}^{-2}$.
- Gravímetros aerosuspendidos: entre ± 2 y $\pm 4 \cdot 10^{-5} \text{ ms}^{-2}$.
- Expediciones lunares (Apolo 11, 12, 14, 17): en la zona de alunizaje con $\pm 18 \cdot 10^{-5} \text{ ms}^{-2}$.

3.4. Determinación de las segundas derivadas del potencial de la gravedad

El gradiente del potencial total de la gravedad es el vector gravedad:

$$\vec{g} = \vec{f}_c + \vec{f}_N = \overrightarrow{\text{grad}W} = \nabla W = \left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}, \frac{\partial W}{\partial z} \right) = (W_x, W_y, W_z)$$

Si diferenciamos nuevamente, obtenemos las segundas derivadas del potencial de la gravedad, el tensor gradiente de la gravedad, también llamado tensor de Eötvös:

$$\overrightarrow{\text{grad} \vec{g}} = \text{grad}(\overrightarrow{\text{grad}W}) = \nabla(\nabla W) = \Delta W = \begin{pmatrix} W_{xx} & W_{xy} & W_{xz} \\ W_{yx} & W_{yy} & W_{yz} \\ W_{zx} & W_{zy} & W_{zz} \end{pmatrix}$$

Puesto que el campo de la gravedad es irrotacional, podemos deducir que el tensor gradiente de la gravedad solamente tiene cinco componentes independientes, ya que $W_{xy}=W_{yx}$, $W_{xz}=W_{zx}$, $W_{yz}=W_{zy}$.

$$\Delta W(x, y, z) = W_{xx} + W_{yy} + W_{zz} = -4\pi G\rho + 2\omega^2$$

Aunque en la realidad difieren bastante de los valores teóricos, sobre el elipsoide del sistema de referencia geodésico GRS80 los valores teóricos de las componentes del tensor gradiente de la gravedad son, teniendo en cuenta que $U_{xy}=U_{zy}=0$, y siendo la componente más utilizada U_{zz} el gradiente normal de la gravedad:

$$U_{xx} \approx U_{yy} = -1540 \text{ ns}^{-2}$$

$$U_{yy} - U_{xx} = -10.4 \cos^2 \varphi \text{ ns}^{-2}$$

$$U_{zz} = \frac{\partial \gamma}{\partial z} = 3086 \text{ ns}^{-2}$$

$$U_{zx} = \frac{\partial \gamma}{\partial x} = 8.1 \sin 2\varphi \text{ ns}^{-2}$$

Un gradiómetro es un sensor que puede medir el cambio de aceleración de la gravedad en el espacio. Los primeros gradiómetros fueron diseñados por Eötvös a principios de siglo, a partir de la balanza de torsión de Cavendish. El IGN adquirió una balanza de torsión de Eötvös-

Schweydar en 1923. La unidad de medida en la medida de gradientes es el Eötvös (E), equivalente a 10^{-9} s^{-2} .

El principio de la balanza de torsión es la medida del ángulo de giro $\theta - \theta_0$ en la dirección α con dos masas m_1 y m_2 en los extremos de la balanza, con la condición de equilibrio:

$$\tau(\theta - \theta_0) = \frac{ml^2}{2}((W_{yy} - W_{xx})\sin 2\alpha + 2W_{xy}\cos 2\alpha) - \frac{mlh}{2}(W_{zx}\sin \alpha - W_{xx}\cos \alpha)$$

La precisiones alcanzadas son de ± 1 Eötvös en tierra firme en terreno llano.

Los gradiómetros inerciales miden la gravedad mediante acelerómetros en dos puntos tan próximos que podemos suponer que la variación de la gravedad entre ellos es lineal. Existen dos tipos: los traslacionales y los giratorios (o rotatorios).

En el giratorio, la ecuación de las fuerzas f_i medidas por los pares de acelerómetros (en cada una de las direcciones) separados una distancia l de una plataforma circular que rota con velocidad ω y que está estabilizada es:

$$\Delta f = (f_1 + f_2) - (f_3 + f_4) = l(W_{yy} - W_{xx})\sin 2\omega t + 2lW_{xy}\cos 2\omega t$$

$$\Delta f_1 = (f_1 + f_2)' - (f_3 + f_4)' = l(W_{zz} - W_{xx})\sin 2\omega t + 2lW_{xz}\cos 2\omega t$$

$$\Delta f_2 = (f_1 + f_2)'' - (f_3 + f_4)'' = l(W_{zz} - W_{yy})\sin 2\omega t + 2lW_{yz}\cos 2\omega t$$

Por último, el gradiómetro absoluto de caída libre es un instrumento con dos cámaras de caída libre similar a la del gravímetro absoluto, y dos masas test que caen en el vacío al mismo tiempo. La precisión en cinco minutos de medidas es de 5 E y la exactitud es de 12 E.

3.5. Medida del gradiente de la gravedad

La medición del gradiente vertical de la gravedad en tierra es necesario para introducirlo en la fórmula del cálculo de la medida absoluta de la gravedad. La componente $g_z = \partial g / \partial z$ corresponde a la segunda derivada del potencial total W con respecto a la componente z dos veces, W_{zz} .

$$W_{zz} = \frac{\partial^2 W}{\partial z^2} = \frac{\partial g}{\partial z} = -(W_{xx} + W_{yy}) - 4\pi G\rho + 2\omega^2$$

La observación del gradiente vertical de la gravedad en tierra se realiza mediante gravímetros relativos de muelle y plataformas o trípodes a distintas alturas. Para obtener resultados hemos de aplicar las correcciones del párrafo 7, además de considerar las correcciones de deriva estática y dinámica de cada uno de los gravímetros. La corrección de deriva dinámica es muy pequeña, pues el desplazamiento del gravímetro en el laboratorio (en la vertical) es muy cuidadoso y no está sujeto a grandes movimientos. Con gravímetros sin sistema de realimentación son necesarias unas 50 reiteraciones, mientras que son 10 reiteraciones las que requieren los

gravímetros con sistema de realimentación. En el cálculo del gradiente vertical asumimos que el campo de gravedad sobre las estaciones de observación se puede aproximar por una función polinómica de segundo grado en función de la altura sobre la marca de la estación, aunque se podrían utilizar polinomios de mayor orden:

$$g(h) = c_0 + c_1 h + c_2 h^2$$

donde los coeficientes c_0 , c_1 y c_2 se obtienen para cada punto por mínimos cuadrados. El gradiente en cada punto de la vertical será la derivada de esa función, por tanto, la función lineal:

$$\gamma(h) = c_1 + 2 c_2 h$$

3.6. Calibraciones e intercomparaciones

Los tres elementos que necesitan de algún tipo de control o calibración en los gravímetros absolutos son: el reloj de rubidio, el láser, y el barómetro. Como frecuencia se utiliza un reloj atómico de rubidio, de estabilidad del orden de 10^{-10} . El reloj de rubidio ha de calibrarse contra patrones más estables, como el reloj de cesio y el máser de hidrógeno. Las calibraciones de la frecuencia del Láser de He-Ne se realizan contra el patrones similares que se encuentran en condiciones más estables, por ejemplo en el CEM (Centro Español de Metrología) o del BIPM (Bureau Internacional de Pesas y Medidas de Sèvres). El barómetro se debe calibrar en el CEM (Centro Español de Metrología) y en las intercomparaciones. Las calibraciones se realizan con el máser de hidrógeno del Observatorio Astronómico de Yebes y con el rubidio disciplinado a un patrón de cesio del Real Observatorio de la Armada de San Fernando (Cádiz), que es el laboratorio de tiempo oficial asociado al CEM.

Las primeras comparaciones con métodos pendulares se realizaron en Potsdam entre 1909 y 1971. Las comparaciones oficiales entre gravímetros absolutos datan del año 1976 en Sèvres, aunque ya en 1968 Faller y Hammond comenzaron a practicarlas. Desde 1981 la IAG reconoce la necesidad de comparaciones periódicas de gravímetros absolutos para detectar posibles errores sistemáticos y definir el nivel de exactitud de la metodología. Los resultados finales están reducidos al un punto de referencia y a una altura de referencia de 90 cm.

3.7. Correcciones a aplicar a las observaciones

Para obtener los valores finales en las mediciones relativas y absolutas, las observaciones originales de la gravedad se corrigen por mareas terrestres, movimiento del polo, atracción gravitatoria y carga procedente de la carga oceánica de las mareas oceánicas (efecto indirecto oceánico), cambios de gravedad debidos a las variaciones de presión atmosférica.

3.7.1. Corrección por mareas terrestres

Las mareas son un fenómeno planetario en su origen y terrestre en sus consecuencias. La Tierra en su conjunto y todos los cuerpos sólidos, líquidos y gaseosos que la componen son deformados y tensionados. Por tanto, todo tipo de medida precisa se encuentra afectada de este fenómeno. La elipticidad, la inelasticidad y las heterogeneidades laterales de la estructura terrestre influyen en las deformaciones finales y han de tenerse en cuenta a la hora de elaborar modelos. El efecto máximo pico a pico de las mareas terrestres en una estación gravimétrica absoluta situada es de unos 300 μGal , y las exactitudes con las que actualmente los modelos predicen las mareas terrestres se encuentran muy por debajo de 1 μGal .

La corrección por marea se puede calcular por diversos métodos:

- A partir de un potencial de la Tierra rígida con un factor gravimétrico de 1,16 y fase cero. Dependiendo del potencial utilizado se pueden obtener mayores o menores incertidumbres.
- Cálculo de un potencial de Tierra rígida con los factores de amplitud de un modelo de Tierra elásticos (modelo 1066 de Gilbert y Dziewonski, Modelo de Gutenberg-Bullen, modelo PREM, modelo DDW, etc.) y corrigiendo de los efectos producidos por la marea oceánica en amplitud y fase para las ondas principales.
- Cálculo de un potencial de Tierra rígida con los factores de amplitud de un modelo de Tierra elásticos y medida experimental de mareas en la estación y utilizar los factores gravimétricos experimentales medidos en esa estación.

El potencial gravitatorio de marea W en un punto de coordenadas esféricas (r, φ, λ) , si la masa de la Tierra es M , G la constante de gravitación universal, r el radio, φ la latitud y λ la longitud, y d la distancia al astro perturbador, se puede expresar como:

$$W(r, \varphi, \lambda) = GM \sum_{n=2}^{\infty} \frac{r^n}{d^{n+1}} P_n(\cos z)$$

siendo las funciones $P_n(\cos z)$ los polinomios de Legendre de orden n , y $\cos z$ del triángulo esférico de la astronomía de posición es:

$$\cos z = \sin \varphi \sin \delta + \cos \varphi \cos \delta \cos H(r, \varphi, \lambda)$$

en donde (δ, H) son las coordenadas ecuatoriales horarias del astro perturbador y (φ, λ) las coordenadas astronómicas del punto.

Para el término de orden 2 tendremos:

$$W_2(r, \varphi, \lambda) = \frac{GM}{2} \frac{r^2}{d^3} (3 \cos^2 \varphi - 1)$$

El factor gravimétrico en una Tierra esférica sin rotación se define como la relación de la amplitud de la marea observada Δg (por un gravímetro) y la amplitud de la marea sobre una Tierra perfectamente rígida, es decir, es la constante de proporcionalidad adimensional entre ambas, y se escribe (h y k son los números de Love):

$$\delta_2 = 1 + h_2 - \frac{3}{2} k_2$$

$$\Delta g = \delta_2 \frac{\partial W_2}{\partial r}$$

En 1981, Wahr demostró que el factor gravimétrico δ depende en general de la latitud, al resolver las ecuaciones del movimiento para una Tierra elástica, elíptica en rotación y autogravitante. El factor δ_2 se puede obtener según la expresión:

$$\delta_2 = 1.160 - 0.005 \left[\frac{\sqrt{3}}{2} (7 \sin^2 \varphi - 1) \right]$$

En general, la expresión que define el factor gravimétrico δ_n es (Melchior, 1983):

$$\delta_n = 1 + \frac{2}{n} h_n - \frac{n+1}{n} k_n$$

La exactitud del potencial de Cartwright-Tayler-Eden (1971, 1973) es de 0.24 μGal y para el de Tamura es de 0.06 μGal . Las principales ondas de marea se pueden distinguir o separar con observaciones suficientemente largas en el tiempo. Cuanto más largas sean las observaciones se pueden separar un conjunto mayor de constituyentes u ondas de marea, siendo necesarias observaciones continuas de más de un año, generalmente realizadas con un gravímetro relativo (de muelle o superconductor).

3.7.2. Corrección por movimiento del polo

La corrección por el movimiento del polo, que es un efecto de largo período, requiere de las coordenadas instantáneas del polo obtenidas de IERS por diversas técnicas (VLBI, SLR y GPS principalmente). Para el factor de amplitud se pueden introducir los valores 1.0 a 1.2, aunque las recomendaciones son de 1.164.

El vector rotación de la Tierra $\vec{\Omega}$ se puede descomponer en una parte constante $\vec{\omega}$ y otra parte variable $\delta\vec{\omega}$:

$$\vec{\Omega} = \vec{\omega}_0 + \delta\vec{\omega} = [\omega_1, \omega_2, \omega_3]$$

$$\vec{\omega}_0 = [0, 0, \Omega]$$

$$\delta\vec{\omega} = \Omega[m_1, m_2, m_3]$$

$$\vec{\Omega} = \Omega[m_1\vec{i} + m_2\vec{j} + (1 + m_3)\vec{k}]$$

La perturbación de primer orden del potencial en función de la colatitud y la longitud (θ, λ) se obtiene al tomar solo los dos términos en m_1 y m_2 :

$$V(r, \theta, \lambda) = -\frac{\Omega^2 r^2}{2} \sin 2\theta (m_1 \cos \lambda + m_2 \sin \lambda)$$

La variación de la gravedad en la superficie por el movimiento del polo se obtiene por la expresión en donde δ es el factor gravimétrico, combinación de números de Love, con $\delta = 1.16$:

$$\delta g = \delta \frac{2}{a} V(a, \theta, \lambda)$$

La corrección por el movimiento del polo tiene en cuenta los cambios diarios en la aceleración centrífuga debidos a la diferencia de distancia entre el eje de rotación y la estación donde se realiza la medida de la gravedad. Para cada estación se calcula utilizando las posiciones del polo que están más cerca del momento de observación. Por tanto la fórmula a aplicar, recomendada por la IAG es:

$$\delta g = -1.164 \cdot 10^8 \omega^2 a \sin \varphi \cos \varphi (x \cos \lambda - y \sin \lambda)$$

donde δg es la corrección de movimiento del polo (μGal); ω es la velocidad angular de rotación de la Tierra (rad/s), a es el radio ecuatorial (semi-eje mayor) del elipsoide de referencia (m); φ y λ son la latitud y longitud geodésicas, respectivamente, de la estación (rad) y x, y son las coordenadas del polo.

El efecto máximo que el movimiento del polo puede inducir en el desplazamiento radial puede llegar a alrededor de 25 mm y en desplazamientos horizontales a 7 mm. El efecto máximo que el movimiento del polo puede inducir en los valores de gravedad puede llegar a 15 μGal pico a pico en varios años. Este efecto es fuertemente dependiente de la latitud, y a 45° de latitud tiene su efecto mayor.

3.7.3. Carga oceánica

La Tierra considerada como una esfera o un elipsoide, se encuentra cubierta por una capa de agua en superficie. La superficie de océanos y mares presenta ondulaciones solamente debidas a las mareas que alcanzan amplitudes de un metro en océano abierto y dos metros en aguas costeras y que se producen con una frecuencia de dos veces al día. La Tierra es deformada, no solamente por las mareas terrestres, sino también por el peso de las mareas oceánicas, hecho dependiente del lugar del continente en que se encuentra el punto donde queremos calcular el efecto. En la Península Ibérica hay estudios que obtienen factores gravimétricos observados y

comparan el comportamiento de las distintas ondas de marea para la Península Ibérica. En la aceleración de la gravedad las variaciones máximas entre el máximo y el mínimo se cuantifican aproximadamente $\pm 11 \mu\text{Gal}$ para la Península Ibérica, hecho que se puede acentuar en las estaciones cercanas a la costa, especialmente en el Atlántico y el Cantábrico.

Las causas del efecto indirecto de las mareas oceánicas son tres: deformación por presión de esa masa sobre la corteza (carga por peso), una atracción directa de las masas oceánicas desplazadas, y una redistribución de las masas de la corteza. Los efectos se manifiestan en una estación de gravedad absoluta en Tierra en:

- Desplazamientos horizontales y verticales.
- Variaciones en la aceleración de la gravedad.
- Variaciones en inclinaciones, es decir, en la desviación de la vertical (“tilt”).
- Variaciones en esfuerzos corticales (“strain”).

Cualquier tipo de carga puede calcularse mediante el formalismo de Farrell (1973). La carga oceánica producida es la convolución entre las funciones de Green para el modelo de Tierra PREM y el modelo de mareas oceánicas a través de la integral:

$$L(\vec{r}) = \rho_w \iint_{\text{oceános}} G(\vec{r} - \vec{r}') * H(\vec{r}') dA'$$

En donde r es la posición del punto en donde queremos calcular la carga oceánica, ρ es la densidad media del agua del mar, H es el modelo de marea oceánica, que representa la altitud de la marea en el elemento de superficie dA' situado en un lugar de coordenadas geográficas r' . La función de Green G representa la respuesta de un determinado modelo de Tierra sometida a una carga puntual unidad. La integral ha de evaluarse para cada onda de marea separadamente. El resultado es un conjunto de parámetros de amplitud y fase denominados factores gravimétricos para cada frecuencia en la corrección por carga oceánica, que representan el efecto de carga oceánica sobre un determinado punto de la superficie de la Tierra.

3.7.4. Corrección por variación en la presión atmosférica y carga

La presión atmosférica perturba también la gravedad. Este efecto una señal de largo período que oscila entre horas, días y también tiene variaciones con periodo estacional. Se ha demostrado que la señal gravimétrica y la presión local atmosférica están correladas con un factor de admitancia de alrededor de $-0,30 \mu\text{Gal/hPa}$.

Habitualmente, la gravedad observada se normaliza a una presión nominal en cada lugar aplicando una corrección basada en la presión atmosférica observada durante las medidas. La

corrección barométrica local C_p (μGal) se aplica a cada caída a través de la fórmula (Resolución de la Asociación Internacional de Geodesia num. 9, 1983):

$$C_p = A (P_o - P_n)$$

donde A es el factor barométrico de admitancia. Este valor se encuentra usualmente entre -0.30 y $-0.42 \mu\text{Gal/hPa}$. El valor habitual es -0.30 , como se recomienda en la resolución anteriormente reseñada; P_o (hPa) es la presión atmosférica observada, y P_n (hPa) es la presión nominal en el emplazamiento. La presión nominal es calculada de acuerdo con:

$$P_n = 1013.25 (1 - 0.0065 h_m / 288.15)^{5.2559}$$

donde h_m es la altitud media sobre el nivel del mar en metros.

Es necesaria una distribución global de barómetros con mayor densidad cerca de la estación gravimétrica para conseguir una corrección mejor por este efecto. Se han observado desplazamientos superficiales de 20 a 60 nms^{-2} (2 a $6 \mu\text{Gal}$) en gravedad, y entre 6 y 20 mm en desplazamientos debidos ambos a efectos atmosféricos en la banda de periodos entre días y meses.

La carga atmosférica se puede calcular mediante el formalismo de Farrell (1973) a través de la integral:

$$L(\vec{r}) = \rho_w \iint_{\text{tierra}} G(\vec{r} - \vec{r}') * H(\vec{r}') dA'$$

A diferencia de la carga oceánica, $H(r)$ es la función de marea atmosférica. Este formalismo, como se ha mencionado en la corrección por carga oceánica, considera los tres efectos de carga por marea: deformación por presión de esa masa sobre la corteza (carga por peso), una atracción directa de las masas atmosféricas desplazadas, y una redistribución de las masas de la corteza. Es pues una corrección más completa que la del factor de admitancia. De nuevo, la función de Green G representa la respuesta de un determinado modelo de Tierra sometida a una carga puntual unidad.

Bibliografía

- [1] Torge, W. Gravimetry. Walter de Gruyter. Berlín, 1989
- [2] Melchior, P. The tides of the Planet Earth. Pergamon Press. 1978.
- [3] Udías, A. y Mézcua, J. Fundamentos de Geofísica. Editorial Alianza, 1996.

Tema 4. Sistemas geodésicos de referencia. Sistema cartesiano espacial, movimiento del Polo. Sistema de coordenadas en el campo de la gravedad terrestre. Sistemas astronómicos general y local, transformaciones entre ambos.

4.1. Sistemas Geodésicos de Referencia

Un Sistema de Referencia consiste en un conjunto de modelos necesarios para la descripción cuantitativa de posiciones y movimientos de cuerpos celestes, incluida la Tierra, (sistemas celestes), o de cuerpos sobre la Tierra (sistemas terrestres). Deben definirse: origen, escala, orientación y plano principal. Un marco asociado a un Sistema de Referencia es una serie de puntos fiduciaros que, junto con sus coordenadas, constituyen la realización práctica del sistema de referencia.

4.2. Sistema cartesiano espacial

El Sistema Cartesiano Espacial es un sistema fijo a la Tierra y convencional, tiene su origen en el centro de masas terrestre generalmente incluyendo atmósfera y océanos, su eje Z que coincide con un eje de rotación medio, en los ITRS corresponde al PMR (Polo Medio de Rotación), su plano XZ corresponde a un meridiano medio de Greenwich y su plano XY corresponde a un ecuador medio, siendo dextrógiro.

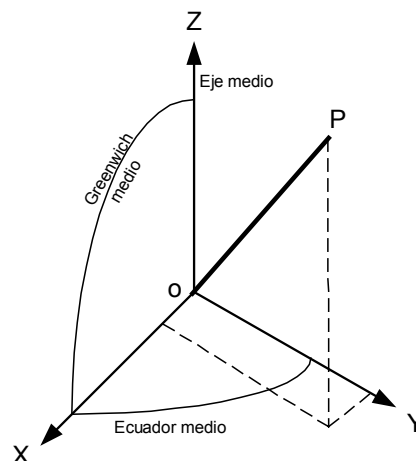


Figura 1.

4.3. Movimiento del polo

Las mediciones extraterrestres juegan una gran importancia en la Geodesia, en tanto que la gran mayoría de las observaciones proceden o bien de satélites artificiales, o bien de estrellas, es importante entender algunos conceptos sobre el movimiento de la Tierra [4]:

1. Se mueve solidariamente en la galaxia respecto a otras galaxias.
2. Se mueve con el sistema solar dentro de nuestra galaxia.
3. Se traslada en su movimiento anual alrededor del sol.
4. Rota sobre su eje instantáneo de rotación.

De los dos primeros movimientos se ocupa la astronomía, por lo general en geodesia, observaremos a objetos dentro de nuestra galaxia. Por tanto, nos centraremos en los movimientos de carácter anual y diurno. El movimiento anual puede ser adecuadamente explicado usando la mecánica celeste, esto es, asumiendo que la Tierra y otros cuerpos celestes son cuerpos puntuales con su masa concentrada en el centro de gravedad. Para explicar el fenómeno diurno, con sus efectos de precesión y nutación el modelo de Tierra tiene que ser refinado teniendo en cuenta la reología (comportamiento ante esfuerzos), atmósfera, océanos...

Movimiento anual de la Tierra

En la descripción del movimiento anual, las dimensiones de la Tierra y otros cuerpos celestes pueden ser consideradas despreciables con respecto a las del sistema solar. Bajo estas condiciones se cumplen las leyes de Kepler:

- Los planetas describen órbitas elípticas con el sol en uno de sus focos
- La velocidad superficial es constante
- El cociente $\text{periodo}^2 / \text{semi\_eje\_mayor}^3$ es constante

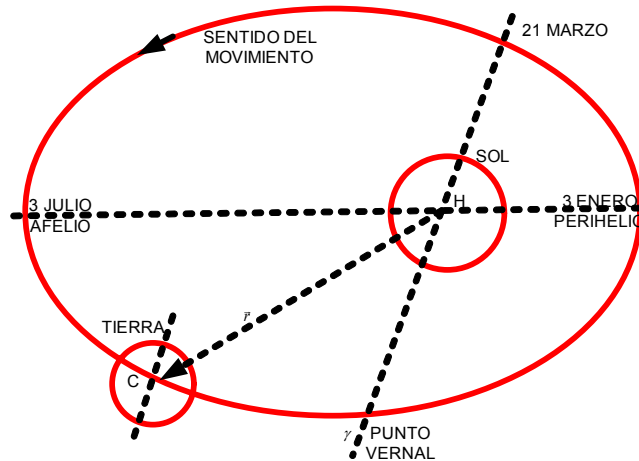


Figura 2.

Debido a la segunda ley, ya sabemos que la Tierra se mueve más rápidamente en el perihelio, en realidad, la luna y los planetas perturban la órbita provocando que no sea elíptica ni plana. Para muchos propósitos prácticos estas perturbaciones son inapreciables, además, la totalidad de la elipse orbital se mueve respecto a las estrellas que la rodean dentro de la galaxia, este movimiento es extraordinariamente lento.

Rotación, precesión y nutación

Para describir la rotación de la Tierra no se pueden despreciar sus dimensiones, el siguiente modelo dinámico más simple es aquel que toma la Tierra como un cuerpo rígido que viaja alrededor del sol y gira a través de un eje que pasa por su cuerpo. En mecánica se denomina giróscopo a este cuerpo. A la manifestación diaria del efecto giroscópico, la rotación alrededor del eje polar de la Tierra es el movimiento diurno y a la Tierra le lleva 366.2564 rotaciones respecto de la estrellas, lo que se llama día sidéreo, 365.2564 respecto del Sol, lo que llamamos día solar medio, para completar una revolución alrededor del Sol.

En el siguiente grado de exactitud el eje de rotación, denominado instantáneo, no coincide con el eje principal que corresponde al máximo momento de inercia que pasa por el centro de masas de la Tierra. Cuando existe un momento externo ejercido sobre el giróscopo, el eje de giro describe un cono circular en cuyo centro está el centro de masas del giróscopo, este movimiento es conocido como precesión. Para la Tierra los cuerpos celestes son los que proporcionan este momento.

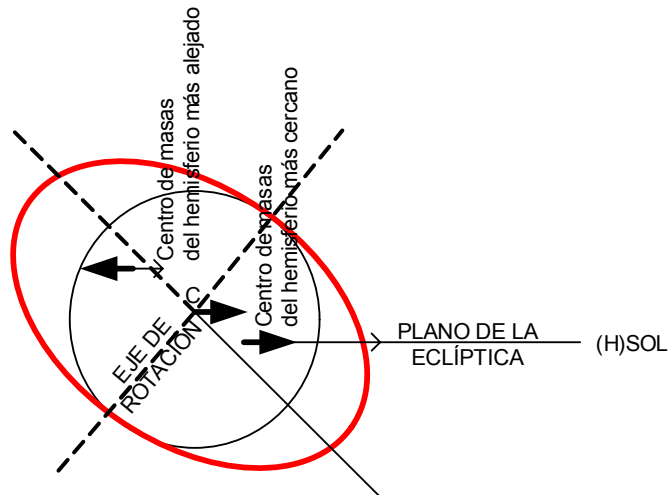


Figura 3.

El hemisferio más cercano al sol es atraído más que el alejado, para obtener el momento con respecto a C, punto de referencia para describir la precesión, debemos de restar la fuerza que actúa en C de ambas fuerzas hemisféricas. Si no existiera abultamiento ecuatorial el momento desaparecería debido a que el par de fuerzas descansan sobre el plano C-H. En nuestro caso, el eje de rotación terrestre no está fijo en el espacio y se mueve describiendo un cono perpendicular al plano de la Eclíptica. El ciclo de precesión es de 26000 años, un año platónico. El movimiento del punto vernal es de unos 50" por año en el sentido de las agujas del reloj.

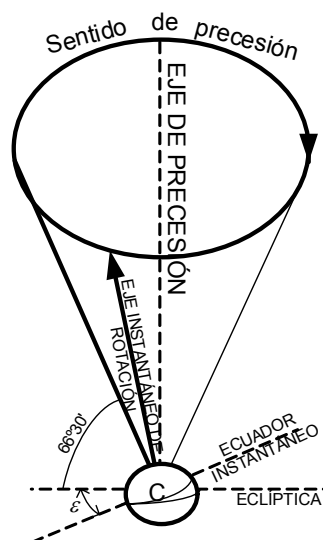


Figura 4.

La presencia de la luna hace que el estudio de la cinemática de la Tierra sea todavía más complicado.

El primer hecho importante es que la luna órbita en un plano inclinado $5^{\circ}11'$ respecto a la Eclíptica, la intersección de estos dos planos es la línea nodal, que rota cada 18.6 años, esto introduce un momento adicional.

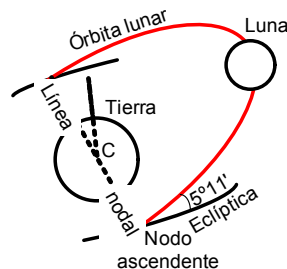


Figura 5.

Por tanto, la luna además de perturbar la órbita anual, también perturba la precesión. Esta perturbación da como resultado otro movimiento del eje de rotación llamado nutación forzada o simplemente nutación. Este cono de nutación es mucho más estrecho que el de precesión, su ángulo es de $18^{\circ}42'$ (47° para la precesión). La nutación es mucho más rápida que la precesión completándose un ciclo en 18.6 años en vez de 26000.

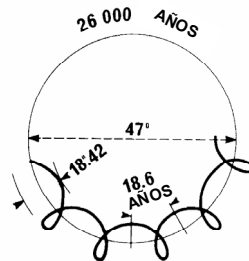


Figura 6.

Evidentemente la nutación también afecta al punto vernal, la descripción matemática del momento es sumamente tediosa debido a que el momento lunisolar que causa estos movimientos es función de las posiciones del sol y de la luna y ésta varía constantemente. El movimiento compuesto, es decir, precesión lunisolar+nutación tiene periodicidades bien definidas.

- Los dos periodos mencionados 18.6 y 26000 años son sólo los principales cuya contribución tiene las amplitudes más grandes.
- Otros periodos que contribuyen apreciablemente son:
 - a) Semianual de amplitudes 0.15"
 - b) Quincenal de amplitudes 0.1"

La posición y la orientación en un momento dado del eje terrestre es el resultado de todos los movimientos descritos. Todos ellos tienen un resultado directo en las observaciones satelitarias y astronómicas. Esto es parte de la mecánica celeste. El movimiento del eje instantáneo de rotación con respecto a la Tierra es otro problema. En este caso el movimiento concierne a la Tierra de modo más interno.

Nutación libre

Dinámicamente el movimiento conocido como nutación libre de la Tierra o "bamboleo" (wobble) es la nutación libre de momento. Adoptemos el sistema de coordenadas geocéntrico cartesiano relacionado con los principales momentos de inercia.

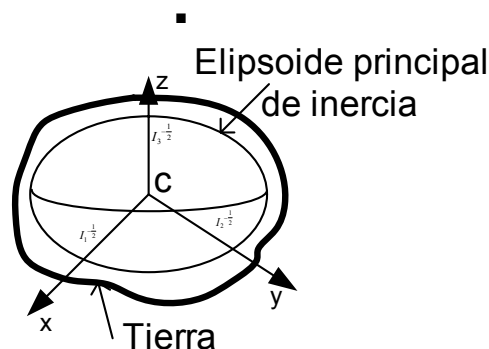


Figura 7.

Los ejes de este sistema de referencia están dados por los vectores propios del tensor de inercia J . En otras palabras es un sistema dextrógiro xyz con origen en el centro de masas de la Tierra y ejes coincidentes con el tensor principal de inercia formando el elipsoide de inercia (no confundir con el elipsoide geodésico)

Para un cuerpo rígido este sistema de coordenadas está rígidamente unido al cuerpo, en caso contrario el sistema está definido instantáneamente por la distribución de las masas dentro del cuerpo.

Si denomino I_1, I_2, I_3 a los principales momentos de inercia respecto a x, y, z respectivamente, la ecuación de Euler para la nutación libre es $J\dot{\vec{w}} + \vec{w} \times J\vec{w} = 0$, donde $\vec{w}_1, \vec{w}_2, \vec{w}_3$ son las componentes del vector velocidad angular y $\dot{\vec{w}}_1, \dot{\vec{w}}_2, \dot{\vec{w}}_3$ son las componentes del vector aceleración angular, de esta manera, la ecuación de Euler puede quedar como sigue:

$$\begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} \dot{w}_1 \\ \dot{w}_2 \\ \dot{w}_3 \end{pmatrix} + \begin{pmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{pmatrix} \begin{pmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = 0 \quad (1)$$

Donde I_1 e I_2 no difieren significativamente el uno del otro, pero si difieren de I_3 , desarrollando la expresión anterior tengo las tres ecuaciones diferenciales que definen la nutación libre:

$$\dot{w}_1 + \left(\frac{I_3 - I_1}{I_1} \right) w_2 w_3 = 0; \quad \dot{w}_2 + \left(\frac{I_3 - I_1}{I_1} \right) w_1 w_3 = 0; \quad \dot{w}_3 = 0 \quad (2)$$

Se observa que $w_3 = \text{cte} = \mu$, si diferenciamos las ecuaciones respecto del tiempo las dos ecuaciones se transforman en ecuaciones diferenciales de segundo grado con variables separadas.

$$\dot{w}_1 + \left(\frac{I_3 - I_1}{I_1} \right)^2 \mu^2 w_1 = 0; \quad \dot{w}_2 + \left(\frac{I_3 - I_1}{I_1} \right)^2 \mu^2 w_2 = 0 \quad (3)$$

Resolviéndolas queda

$$\begin{aligned} w_1(\tau) &= \beta \cos\left(\left(\frac{I_3 - I_1}{I_1}\right) \mu \tau + \psi\right) \\ w_2(\tau) &= \beta \sin\left(\left(\frac{I_3 - I_1}{I_1}\right) \mu \tau + \psi\right) \end{aligned} \quad (4)$$

Con ψ β constantes de integración, si observamos la figura inferior vemos que el polo instantáneo de rotación describe un cono circular alrededor del eje principal de inercia de sentido contrario al de las agujas del reloj mirándolo desde el polo. (amplitud, fase y periodo)

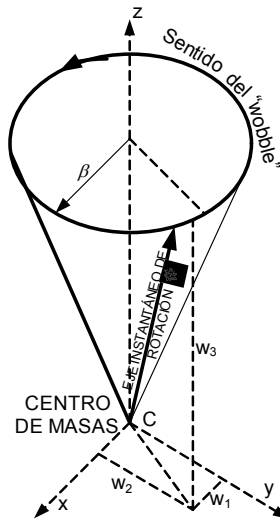


Figura 8.

La solución de la ecuación de Euler no muestra información sobre las constantes que han de ser definidas por observación. , la observación si indica que w_1 y w_2 son muy pequeños comparados con w_3 , en otras palabras, el eje instantáneo de rotación se desvía un ángulo muy pequeño. Por tanto $w_3 = \omega$, prácticamente igual a la velocidad angular de rotación de la Tierra., consecuentemente $w_3 = \mu = \omega = 2\pi / 1 \text{ día sidereo}$, con periodo $P = 2\pi \frac{I_1}{(I_3 - I_1)\mu}$, lo que

$$H^{-1} = \frac{I_1}{I_3 - I_1}$$

define el aplanamiento dinámico como correspondiente a 305 días sidéreos. Este valor es el denominado periodo de Euler. A final de siglo se descubrió que el periodo realmente era un 40% mayor que el de Euler (Newcomb), ello es debido a la falta de rigidez de la Tierra, esta falta de rigidez eleva el periodo a 435 días y es el llamado periodo de Chandler, Debido a que la Tierra es un sólido deformable en vez de rígido, se hace necesario tener en cuenta la fricción interna y por tanto la disipación de energía, a medida que un sistema disipa energía se produce un amortiguamiento del movimiento, en nuestro caso, al menos teóricamente, debemos de esperar una amortiguación de β exponencialmente en el tiempo. La descripción de la nutación para un cuerpo deformable está dada por la ecuación de Liouville

4.4. Movimiento del polo observado y variación en la velocidad de rotación de la Tierra

Para determinar los parámetros desconocidos del "bamboleo" (wobble) se inició el programa del servicio internacional de la latitud, pues tradicionalmente la variación del movimiento del polo ha sido observada por la variación de latitud de observatorios, se comenzó a observar el fenómeno a partir de 1899 en la latitud $39^{\circ}08'$ (Mizusawa, Kitab, Caloforte, Gaithersburs y Ukiah) y las dos agencias que las controlaban era el IPMS (International Pole Motion Service) y el BIH. Los EEUU introdujeron una nueva, el DPMS (Dal Pole Motion Service) monitorizándolo el polo a través de la red TRANSIT. En la actualidad hay una gran abundancia de datos comparativamente homogéneos revelando que el movimiento es relativamente más complicado de lo que se pensó en un principio.

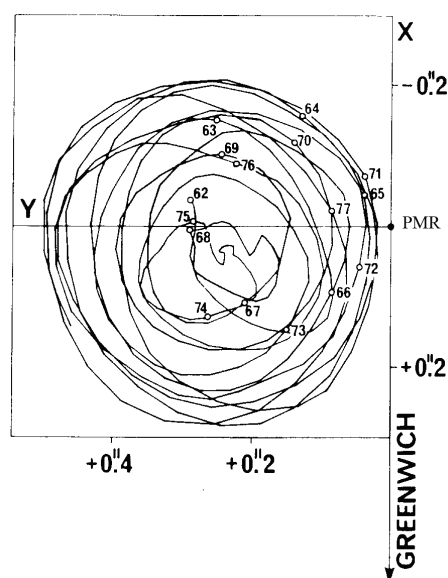


Figura 9.

En la figura anterior se observa el movimiento del polo respecto del Polo Medio de Rotación, antiguamente denominado CIO (polo medio 1900-1905), a primera vista no revela un amortiguamiento, la explicación más verosímil es que detrás del amortiguamiento existe un mecanismo que ocasionalmente o continuamente excita el "wobble". La hipótesis de que la excitación está generada por terremotos tectónicos no es muy realista (Jeffreys 1970 et al.) aunque no se tienen datos totalmente precisos y convincentes del mecanismo de excitación. Es difícil obtener un valor de β debido a los fenómenos de amortiguamiento y excitación. Varios análisis muestran que está en torno a $0.1''$ y $0.2''$, lo que corresponde a un movimiento del polo de 3 a 6 metros. Varios análisis también revelan la existencia de dos componentes significativas en el movimiento del polo, la componente estacional y secular. El mecanismo que controla estos fenómenos tampoco está muy bien conocido.

Las variaciones estacionales, de periodo anual proporcionan una considerable variación al periodo de Chandler en amplitud. Según algunos autores la amplitud es de : 0.04"-0.12" ó 1-4 metros. La variación está íntimamente relacionada con la variación de la temperatura, presión, carga de nieve, precipitación. Otros autores piensan que la variación estacional es simplemente espuria en cuanto a su aparición y no revela sino ruidos locales a la estación de observación provocados por los efectos estacionales experimentados por éstas. Quizá de todos los periodos residuales el menos entendido es la variación secular, se manifiesta como una deriva del polo de unos 0.002" a 0.003" por año y la explicación está probablemente ligada a la tectónica de placas.

En cuanto a la rotación terrestre, la Tierra efectúa aproximadamente 366.3564 rotaciones (días sidéreos) en completar un año sidéreo, la velocidad de rotación se ha considerado tradicionalmente constante prácticamente hasta el empleo de relojes atómicos (1930), se demuestra que existe una variación en la velocidad de rotación de la Tierra, en la actualidad se reconocen 3 tipos de fluctuación en la velocidad de giro: Secular, periódica e irregular:

- La continua o secular provoca una disminución de la rotación (fricción de marea) e incrementa la longitud del día en 2 milisegundos al año.
- Las estacionales con periodos de 1 y 1/2 años y otras variaciones periódicas (ej. 1 mes) llevan a variaciones de varios milisegundos, en parte son debidas a efectos mareales y de viento
- Las variaciones irregulares son las más interesantes, pueden ser de hasta 10 milisegundos por día. El mecanismo no está muy claro y tampoco algunos autores aceptan su existencia.

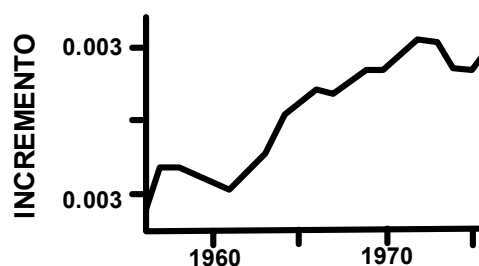


Figura 10.

4.5. Rotación de la Tierra. El Polo Celeste Intermedio (CIP)

(Extractadote [2])

Cuando se buscan resultados con muy altas precisiones, el estudio del movimiento de rotación terrestre es considerablemente complejo. Ello es debido a que tanto la posición espacial del eje como la velocidad angular de rotación son variables que dependen de numerosos factores, gran parte de ellos relacionados con la disposición y las distintas densidades de las masas del planeta. Sin embargo, hoy en día el mencionado estudio es de capital importancia para múltiples aspectos propios de la Geodesia y la Astronomía, por lo que se hace necesario atender a nuevos métodos y técnicas capaces de rendir las precisiones requeridas. Tal es el propósito del modelo IAU2000 en el tema que nos ocupa.

La descripción de la rotación terrestre en cada época consiste en determinar el movimiento del sistema de referencia que gira con la revolución diurna del planeta (ITRS), respecto del sistema de referencia de orientación fija en el espacio (GCRS). Para ello las resoluciones IAU2000 definen un eje intermedio situado muy próximo al eje instantáneo de rotación, y desglosan el movimiento a estudiar en dos: un movimiento del eje intermedio respecto del GCRS conocido como movimiento conjunto de precesión – nutación, y un movimiento del eje intermedio respecto del ITRS, que se denomina movimiento del polo. El primero se estudia considerando los parámetros de desplazamiento celeste del polo en longitud y oblicuidad, los cuales describen conjuntamente el efecto de precesión y de nutación de acuerdo con el modelo IAU2000. Adicionalmente se tiene en cuenta una función temporal (s) de muy pequeño valor ligada con las definiciones de las nuevas referencias.

En cuanto al segundo movimiento, su estudio se realiza mediante los valores de dos pequeños ángulos análogos a los de empleo tradicional (Apdo. 6.8), y considerando también una función temporal (s') de magnitud muy pequeña motivada por la definición de las nuevas referencias. En su conjunto, los mencionados valores paramétricos utilizados en el estudio del movimiento de precesión – nutación y del movimiento del polo se conocen como Parámetros de Orientación de la Tierra (EOP), y como veremos después servirán como argumentos para el intercambio de coordenadas entre el ITRS y el GCRS.

4.6. Forma observacional

Por medio de observaciones hechas desde la Tierra a objetos localizados en el espacio: Estrellas, satélites artificiales, la Luna y quasars. Todos ellos proporcionan direcciones de referencia para medir la rotación de la Tierra. Durante décadas se han fotografiado las estrellas para determinar el movimiento de rotación del polo. Actualmente el uso de láseres y radiotelescopios proporciona métodos de medida más precisos. El IERS (International Earth Rotation Service) fue creado en 1988 por la Unión Internacional de Geodesia y Geofísica (IUGG). Reemplazó a la Sección de Rotación de la Tierra del BIH (Bureau International de l'Heure) y al Servicio

Internacional del Movimiento del Polo (IPMS). Actualmente el IERS es la organización internacional responsable de la coordinación de todas las observaciones del movimiento del polo, nutación y tiempo astronómico. La oficina central se encuentra localizada en el observatorio de París.

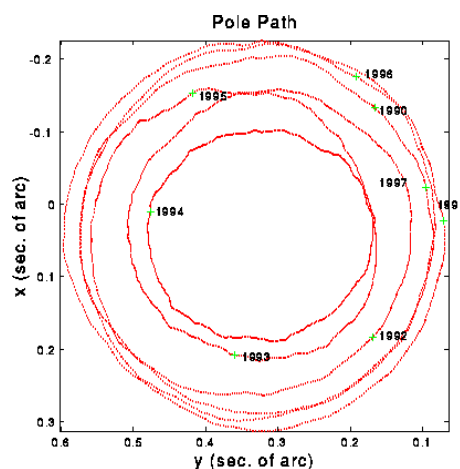
Existen dos puntos de vista:

- a) Mediante cambios en el eje de rotación en el espacio respecto al sistema absoluto de la esfera celeste, referido al plano de la Eclíptica y a su polo. Movimientos de precesión de los equinoccios, nutación y cambio secular de la oblicuidad de la Eclíptica.
- b) Cambios del eje de rotación con respecto al sistema convencional. Este movimiento puede ser periódico, la llamada nutación libre (Euler), que en realidad comprende precesión, nutación y deriva secular del polo. También puede ser no periódico, por ejemplo por la inelasticidad de la Tierra.

4.7. Descripción de las coordenadas del polo. Movimiento del CIP en el ITRS

Los ángulos que caracterizan la dirección del polo rotacional dentro de la Tierra es lo que se llaman coordenadas polares x e y . Estas coordenadas refieren el polo instantáneo a en un marco de referencia adoptado por los observatorios encargados de determinar estos parámetros. La coordenada x está medida a lo largo del meridiano de Greenwich y la coordenada y medida a lo largo del meridiano 90W. Estas dos coordenadas determinan sobre un plano las direcciones del movimiento del polo.

El meridiano, el ecuador, las coordenadas o el acimut astronómicos de una estación están siempre referidos al sistema de referencia instantáneo y en consecuencia están sometidos a cambios en el tiempo.



El movimiento del CIP en el ITRS, que tiene en cuenta las oscilaciones con períodos menores de dos días, es obtenido por el IERS mediante observaciones VLBI y resulta descrito a través de los correspondientes cosenos directores del CIP en cada época.

4.8. Boletín IERS

En el Boletín IERS las coordenadas x e y del polo son las coordenadas del Polo Celeste de Efemérides respecto al polo convencional IERS (IERP IERS Reference Pole). El CEP difiere del polo instantáneo de rotación en aproximadamente $0.01''$. Estos desplazamientos del polo celeste fueron descubiertos gracias a las observaciones VLBI y LLR poniendo de manifiesto los modelos de precesión y nutación de la IAU (International Astronomy Union). Las correcciones del polo celeste se denominan $d\psi$ y $d\xi$, y proporcionan los desplazamientos en longitud y oblicuidad del polo celeste respecto a la posición definida por los modelos convencionales IAU, un ejemplo de este boletín se encuentra en la tabla inferior:

MJD	$x(\text{arcsec})$	$y(\text{arcsec})$	UT1-UTC(sec)
51080	0.1177	0.4545	-0.15193
51081	0.1193	0.4528	-0.15295
51082	0.1209	0.4511	-0.15389
51083	0.1223	0.4495	-0.15475
51084	0.1237	0.448	-0.15558

Tabla 1. Ejemplo de coordenadas del polo

Donde MJD es el día Juliano modificado, que puede obtenerse con el siguiente algoritmo

Forma de obtención: M=Mes: Y=año, D=día

Si $M \leq 2$ entonces $m=M+12$, en caso contrario $m=M$

Si $M \leq 2$ entonces $y=Y-1$, caso contrario $y=Y$

$JD = \text{ENTERO}(365.25*y) + \text{ENTERO}(30.6001*(m+1)) + D + 1720981.5$

$MJD = JD - 2400000.5$

4.9. Fórmulas clásicas de reducción de observaciones

Sean u', v', w' los cosenos directores de la vertical local. Sean Λ', ϕ' la longitud y latitud astronómica observada.

$$u' = \cos \Phi' \cos \Lambda'$$

$$v' = \cos \Phi' \sin \Lambda'$$

$$w' = \sin \Phi'$$

Para pasar de estas componentes del coseno director a las convencionales, tendremos que aplicar dos rotaciones: de eje X y valor -yp, y de eje Y y valor -xp. Las matrices quedan:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos yp & -\sin yp \\ 0 & \sin yp & \cos yp \end{pmatrix} \begin{pmatrix} \cos xp & 0 & \sin xp \\ 0 & 1 & 0 \\ -\sin xp & 0 & \cos xp \end{pmatrix} \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix}$$

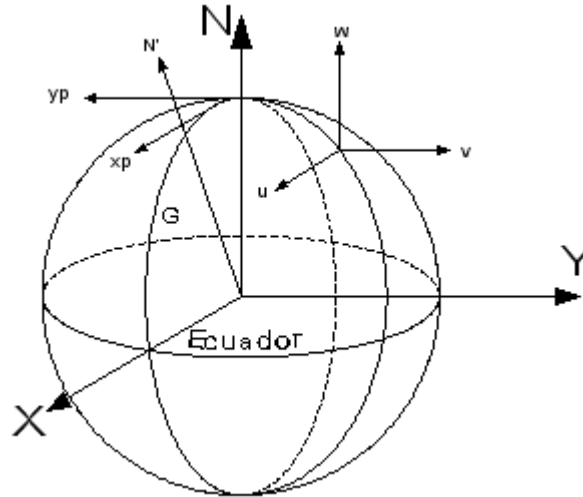


Figura 12.

Aproximando el seno al arco, el coseno a la unidad y despreciando dobles productos tengo:

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} 1 & 0 & xp \\ 0 & 1 & -yp \\ -xp & yp & 1 \end{pmatrix} \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix}$$

Para obtener finalmente las coordenadas corregidas emplearemos.

$$\Lambda = \arctag\left(\frac{v}{u}\right)$$

$$\Phi = \arctag \frac{w}{\sqrt{u^2 + v^2}}$$

También podemos emplear las fórmulas reducidas y obtener directamente la corrección.

$$\Phi = \Phi' - xp \cdot \cos \Lambda' + yp \cdot \sin \Lambda'$$

$$\Lambda = \Lambda' - (xp \cdot \sin \Lambda' + yp \cdot \cos \Lambda') \tan \Phi'$$

Corrección de acimut

De la misma forma el acimut astronómico determinado en un instante debemos referirlo al polo convencional. La fórmula es:

$$A_g = A' \cdot (xp \cdot \sin \Lambda + yp \cdot \cos \Lambda) / \cos \Phi'$$

4.10. Sistema de coordenadas en el campo de la gravedad terrestre. Sistemas astronómicos general y local, transformaciones entre ambos

Sistema astronómico global

Es posible definir en el campo de la gravedad terrestre un sistema de coordenadas denominado generalmente "natural". Este sistema es al que se refieren las observaciones hechas con equipos clásicos. Las coordenadas naturales son Λ , Φ y W . Longitud, latitud astronómica y potencial. Λ como el ángulo medido en el plano meridiano entre el meridiano origen y el meridiano que pasa a través del punto. Φ como el ángulo medido en el plano meridiano entre el plano ecuatorial y la dirección de la línea de la plomada en el punto P, W permite fijar el punto P en la superficie de nivel. De las tres cantidades, la última no puede determinarse por observación directa, sino que se calcula a partir de observaciones de gravedad y nivelación. Un punto definido en el SAG (Sistema Astronómico Global) está definido por la intersección de las superficies coordenadas no ortogonales ($\Phi = \text{cte}$, $\Lambda = \text{cte}$, $W = \text{cte}$). El meridiano astronómico, el paralelo astronómico y la isocénital ($W = \text{cte}$) son curvas alabeadas [3].

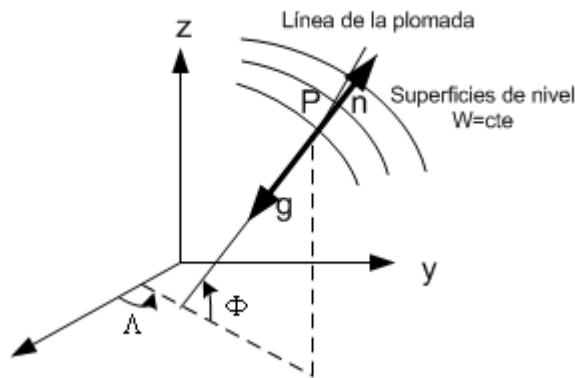


Figura 13.

Con $\vec{g} = \text{grad}W = -g\vec{n} = -g \begin{pmatrix} \cos \Phi \cos \Lambda \\ \cos \Phi \sin \Lambda \\ \sin \Phi \end{pmatrix}$, evidentemente la expresión inversa es obvia

$$\Phi = \arctan \frac{-W_z}{\sqrt{W_x^2 + W_y^2}} \text{ y } \Lambda = \arctan \frac{W_y}{W_x}$$

Sistema astronómico local (SAL)

El sistema astronómico local se denomina habitualmente: NEU (North East Up) ,NEH (North East Height)

, NEZ (North East Zenith), xyz (con x=norte, y=este, z=cenital. No dextrógiro), xyz (con x=este, y=norte, z=cenital. Dextrógiro) o Triedro Laplaciano y es el sistema tradicionalmente usado en topografía, el meridiano local del observador contiene al vector gravedad del observador y al paralelo al eje medio PMR, El SAL es generalmente instantáneo y deberá ser reducido por movimiento del polo al polo convencional.

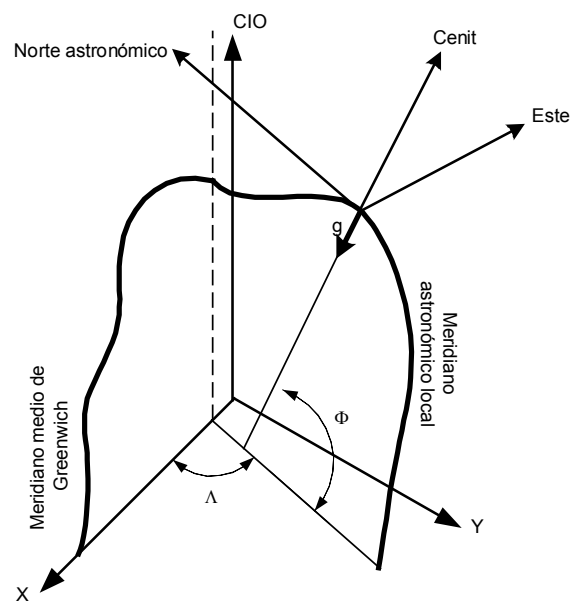


Figura 14.

Sistema geodésico local (SGL), neu

Corresponde a la “versión geodésica” del anterior donde se consideran los elementos geodésicos equivalentes a los astronómicos anteriores

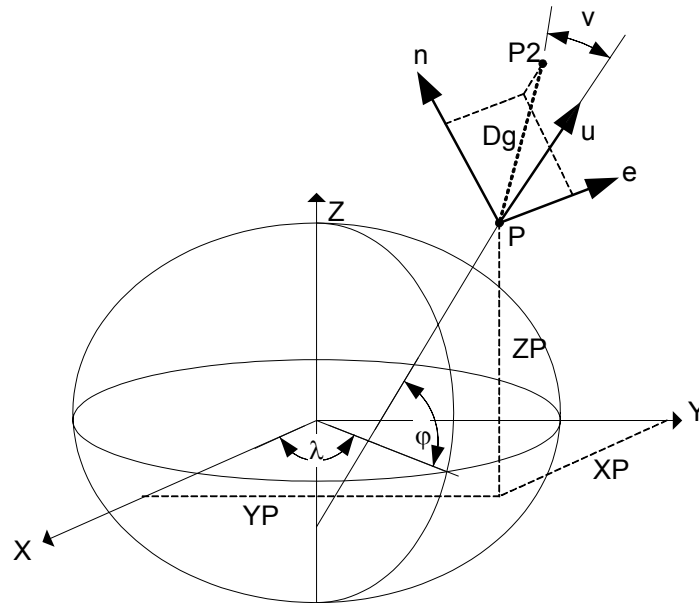


Figura 15.

Paso de un sistema local a uno global

Pongamos como ejemplo el astronómico, siendo idéntica la matriz final al geodésico.

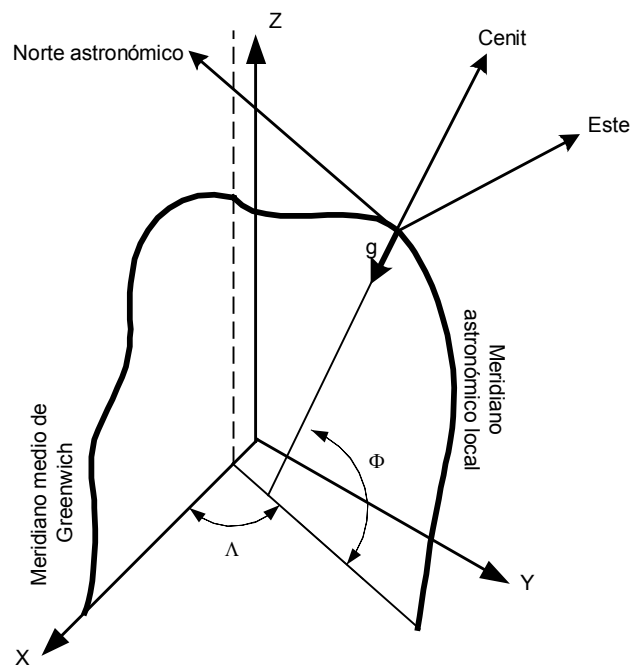


Figura 16.

Construcción de las matrices de rotación [1]:

$$\begin{aligned}
 j &= i(\text{módulo } 3) + 1 \\
 k &= j(\text{módulo } 3) + 1 \\
 r_{jj} &= r_{kk} = \cos \alpha \\
 r_{ij} &= r_{ji} = r_{ik} = r_{ki} = 0
 \end{aligned}$$

$$\begin{aligned}
 r_{jk} &= \sin \alpha \\
 r_{kj} &= -\sin \alpha
 \end{aligned}$$

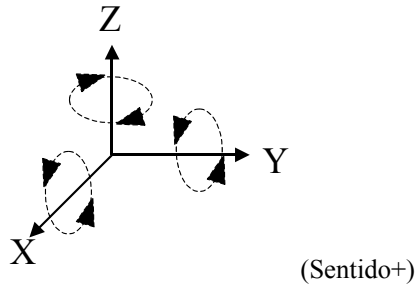


Figura 17.

Primera rotación, $-(90 - \Phi)$ sobre $x(E)$

$$R1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin \Phi & -\cos \Phi \\ 0 & \cos \Phi & \sin \Phi \end{pmatrix}$$

Segunda rotación

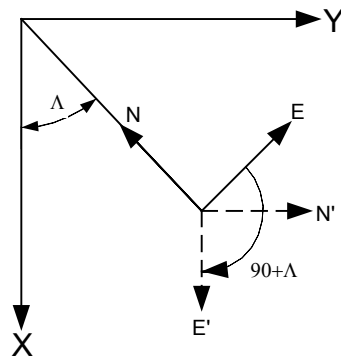


Figura 18.

$-(90 + \Lambda)$ sobre el eje $z(U)$ formando la matriz de rotación $R2$, a saber:

$$R2 = \begin{pmatrix} -\sin \Lambda & -\cos \Lambda & 0 \\ \cos \Lambda & -\sin \Lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Quedando por fin:

$$\begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} = R2 R1 \begin{pmatrix} E \\ N \\ U \end{pmatrix} = \begin{pmatrix} -\sin \Lambda & -\sin \Phi \cos \Lambda & \cos \Phi \cos \Lambda \\ \cos \Lambda & -\sin \Phi \sin \Lambda & \cos \Phi \sin \Lambda \\ 0 & \cos \Phi & \sin \Phi \end{pmatrix} \begin{pmatrix} E \\ N \\ U \end{pmatrix}$$

Paso del sistema global al local

La transformación inversa es inmediata. $(90+\Lambda)$ sobre el eje Z formando la matriz de rotación R1, a saber:

$$R1 = \begin{pmatrix} -\operatorname{sen}\Lambda & \cos\Lambda & 0 \\ -\cos\Lambda & -\operatorname{sen}\Lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

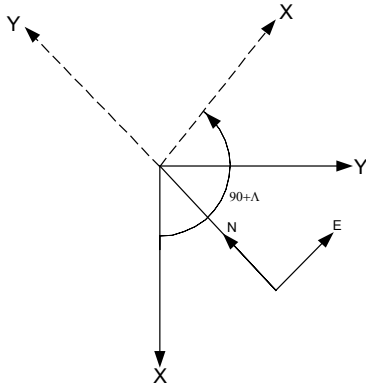


Figura 19.

Segunda rotación, $(90-\Phi)$ sobre X

$$R2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \operatorname{sen}\Phi & \cos\Phi \\ 0 & -\cos\Phi & \operatorname{sen}\Phi \end{pmatrix}$$

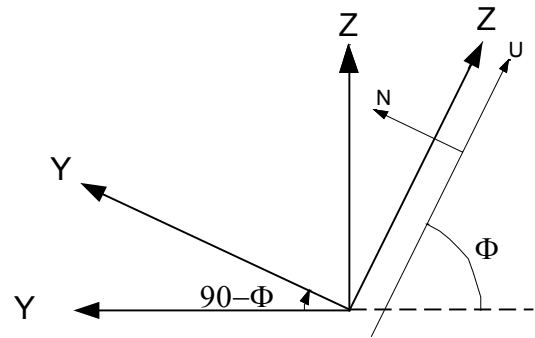


Figura 20.

El resultado, como cabía esperar, es la traspuesta de la matriz de rotación anterior.

$$\begin{pmatrix} E \\ N \\ U \end{pmatrix} = R2R1 \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix} = \begin{pmatrix} -\operatorname{sen}\Lambda & \cos\Lambda & 0 \\ -\operatorname{sen}\Phi \cos\Lambda & -\operatorname{sen}\Phi \operatorname{sen}\Lambda & \cos\Phi \\ \cos\Phi \cos\Lambda & \cos\Phi \operatorname{sen}\Lambda & \operatorname{sen}\Phi \end{pmatrix} \begin{pmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{pmatrix}$$

Las coordenadas de un punto B en el SCG definido por las coordenadas locales en el SAL de A serán:

$$\begin{pmatrix} X_B \\ Y_B \\ Z_B \end{pmatrix} = \begin{pmatrix} -\operatorname{sen}\Lambda & -\operatorname{sen}\Phi \cos\Lambda & \cos\Phi \cos\Lambda \\ \cos\Lambda & -\operatorname{sen}\Phi \operatorname{sen}\Lambda & \cos\Phi \operatorname{sen}\Lambda \\ 0 & \cos\Phi & \operatorname{sen}\Phi \end{pmatrix} \begin{pmatrix} E_B \\ N_B \\ U_B \end{pmatrix} + \begin{pmatrix} X_A \\ Y_A \\ Z_A \end{pmatrix}$$

o si nos referimos al sistema geodésico

$$\begin{pmatrix} X_B \\ Y_B \\ Z_B \end{pmatrix} = \begin{pmatrix} -\operatorname{sen} \lambda & -\operatorname{sen} \varphi \cos \lambda & \cos \varphi \cos \lambda \\ \cos \lambda & -\operatorname{sen} \varphi \operatorname{sen} \lambda & \cos \varphi \operatorname{sen} \lambda \\ 0 & \cos \varphi & \operatorname{sen} \varphi \end{pmatrix} \begin{pmatrix} E_B \\ N_B \\ U_B \end{pmatrix} + \begin{pmatrix} X_A \\ Y_A \\ Z_A \end{pmatrix}$$

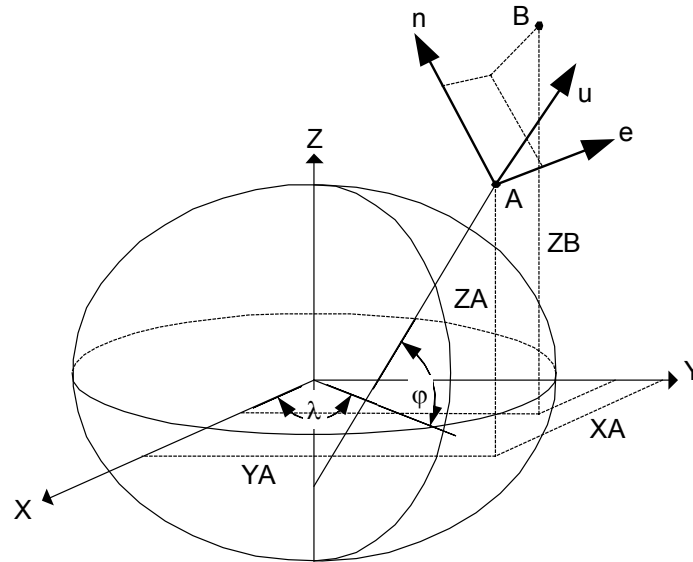


Figura 21.

Bibliografía

- [1] A. Kleusberg y P. J. G. Teunissen, *GPS for geodesy*, Springer, Berlin ; New York, 1998.
- [2] J. Mena, *Geodesia superior (en prensa)*, Instituto Geográfico Nacional, Madrid, 2008.
- [3] W. Torge, *Geodesy*, De Gruyter, Berlin ; New York, 2001.
- [4] P. Vaníček y E. J. Krakiwsky, *Geodesy : the concepts*, North- Holland, Amsterdam, 1987.

Tema 5. El geoide como superficie de referencia para las altitudes. Nivel medio del mar. Altitudes sobre el nivel del mar. Definiciones, objeto de su determinación, precisiones. Altitudes dinámicas y ortométricas. Nivelación geométrica y trigonométrica.

5.1. El geoide como superficie de referencia para las altitudes

Una superficie equipotencial es el lugar geométrico de los puntos de un campo en los cuales el potencial de campo es constante.

Cuando el campo de estudio es el gravitatorio, podemos definir el *Geoide* como la *superficie equipotencial del campo gravitatorio terrestre que, en promedio, coincide con el nivel medio de los océanos en mar abierto*. Para obtenerlo, se tienen en cuenta parámetros tales como la superficie media del mar, el oleaje, las mareas, las corrientes y la rotación terrestre.

Etimológicamente Geoide significa “forma de la Tierra”. Y la forma geométrica (superficie matemática) mas parecida al Geoide (superficie física) es el elipsoide de revolución. De hecho, el Geoide sería un elipsoide de revolución si:

- La densidad fuera uniforme.
- No existiera el relieve.
- La rotación fuese constante.
- No existieran campos gravitatorios exteriores.

Se presentan variaciones de hasta 100m en la separación elipsoide-geoide y variaciones del paralelismo entre las dos superficies que van de algunos segundos a un minuto en zonas montañosas.

La forma del geoide puede determinarse mediante:

- *Medidas gravimétricas*: midiendo la magnitud de la intensidad de la gravedad en numerosos puntos de la superficie terrestre. Estas mediciones de la gravedad terrestre tienen que ser corregidas para eliminar las anomalías locales debido a variaciones de la densidad.
- *Mediciones astronómicas*: Se basan en medir la vertical del lugar y ver sus variaciones. Esta variación se relaciona con su forma.

- *Medición de las deformaciones producidas en la órbita de los satélites:* causadas porque la Tierra no es homogénea.

5.2. Nivel medio del mar

La *superficie del nivel del mar*, ordinariamente empleada como *superficie altimétrica de referencia*, tiene una concreta entidad física y es normalmente empleada como origen altimétrico de cualquier sistema.

Para establecer el geoide como una superficie de referencia para alturas, se registra el nivel del mar usando mareógrafos y se promedia para eliminar las variaciones periódicas del nivel del mar con el tiempo. Pero el nivel medio de los mares no es fácil de determinar. Se conocen las principales ondas armónicas que definen este nivel, y alguna son de largo periodo. Según la teoría vigente de mareas, para obtener el nivel correcto harían falta un mínimo de 18 años y 8 meses de registro. Aún en este caso, hay otras muchas causas que influyen falseando los resultados, principalmente meteorológicas, y que no se pueden controlar. Por todo esto la definición del nivel medio toma un carácter puramente convencional, correspondiente a una época y un lugar dado. En consecuencia, en dos puntos distintos, el nivel cero no tiene por qué coincidir, en el sentido de que el desnivel observable entre ellos puede ser diferente de cero.

El *nivel medio del mar (NMM)* obtenido representa una aproximación al geoide. Sin embargo no se puede hablar estrictamente de que sea una superficie de nivel, pues el NMM puede tener discrepancias de ± 1 m. a ± 2 m. respecto de ella.

Con lo mareógrafos queda determinada el N.M.M en las zonas de costa. Pero es la altimetría de satélites la que proporciona información sobre los mares abiertos y se refiere a la superficie de mar instantánea. La altura de la superficie del océano sobre el geoide representa la *topografía de la superficie del mar*. Habría que distinguir entre la topografía de la superficie *instantánea* y la topografía de la superficie del mar *cuasiestacionaria* la cual resulta después de tener en cuenta las variaciones dependientes del tiempo. Estas variaciones incluyen:

- Mareas oceánicas que pueden desviarse de los valores teóricos debido a profundidades desiguales y al obstáculo que representan los continentes a su movimiento.
- Las de naturaleza meteorológica (presión atmosférica, vientos...)
- Las de naturaleza oceanográfica (corrientes, diferencias de densidad, salinidad, presión...)
- Aquellas debidas a la cantidad de agua.

5.3. Altitudes sobre el nivel del mar. Definiciones, objeto de su determinación, precisiones. Altitudes dinámicas y ortométricas

La altitud sobre el nivel del mar de los puntos sobre la superficie terrestre nos determina una de las coordenadas que delimitan la figura de la Tierra.

Antes hemos definido el Geoide como una superficie de nivel. Cualquier otra superficie equipotencial será aquella en la que el trabajo que hemos de realizar para elevar la unidad de masa desde la superficie del geoide (superficie de altitud cero) hasta ella, es el mismo. Este concepto quedará claro al definir la cota geopotencial

Dimensionalmente, trabajo es igual a fuerza por espacio, y fuerza es igual a masa por aceleración. Por tanto, el trabajo realizado por las fuerzas del campo (siempre que la extensión del mismo sea lo bastante limitado como para poder considerar en ella g constante) será;

$$W = F \cdot h = m \cdot g \cdot h$$

Considerando h_A y h_B el desnivel correspondiente entre dos superficies equipotenciales W_1 y W_2 en dos verticales distintas (las correspondientes a los puntos A_1 y B_1), de la definición anterior, deducimos:

$$g_A \cdot h_A = g_B \cdot h_B$$

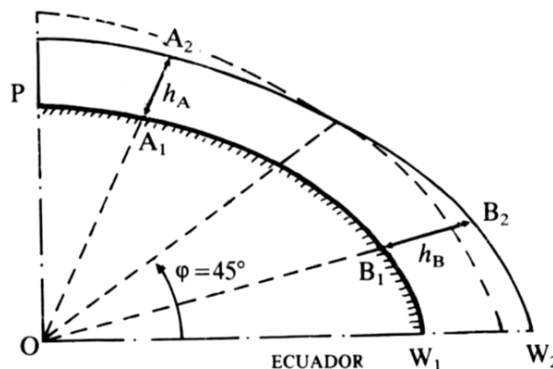


Figura 1. Fuente: "Geodesia y Cartografía Matemática"

Por otro lado, y según la Unión Geodésica y Geofísica Internacional (U.G.G.I.) el valor de la gravedad normal en un determinado punto de latitud φ será:

$$\gamma = 978,03185 \cdot (1 - 0,005278895 \cdot \text{sen}^2 \varphi + 0,000023462 \cdot \text{sen}^4 \varphi) \text{ (Gal)}$$

De donde se deduce que g es variable con la latitud. Como además se verifica que $g \cdot h = \text{cte}$, también h variará, luego $h_A \neq h_B$ y por tanto, queda demostrado que las superficies equipotenciales no son paralelas.

Además, si $g_A > g_B$, se demuestra que $h_A < h_B$ y por tanto las superficies equipotenciales se acercan del Ecuador al polo.

Por otro lado, la gravedad que actúa sobre la unidad de masa en un punto sobre la superficie terrestre se define como la resultante entre el vector fuerza de atracción de la Tierra y el vector fuerza centrífuga. En los polos no hay fuerza centrífuga, ya que por ellos pasa el eje de giro terrestre, y por tanto, en dichos puntos esta fuerza es nula. Por otra parte, atendiendo a la forma de la Tierra, se deduce que, su radio polar es mínimo. Por ambas razones, la gravedad es máxima en el polo y mínima en el ecuador.

Aclarados estos conceptos, veamos ahora cómo se obtienen las diferencias de nivel entre dos puntos de la superficie terrestre, y cuáles son los *diferentes sistemas de altitudes empleados*.

A causa del no paralelismo de las superficies de nivel, el esquema ideal de la nivelación se vuelve relativamente difícil. La dificultad estriba, en que cada observación requiere la consideración de las diferencias respecto de la elipsoidicidad de la Tierra y de la distribución no uniforme de su densidad interna.

Para obtener el desnivel debe introducirse una corrección por el no paralelismo entre la superficie de nivel y la superficie de referencia. Esta corrección representará la influencia del ángulo formado entre las tangentes a estas superficies según la línea de nivelación y se puede determinar mediante observaciones gravimétricas realizadas a lo largo del tiempo.

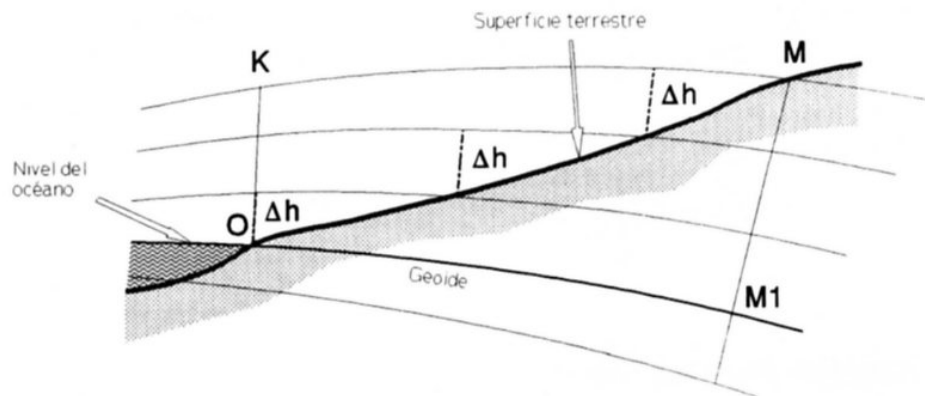


Figura 2. Fuente: "Topografía y Cartografía"

En la figura está representado el perfil de la superficie terrestre, y el nivel del océano (origen altimétrico) que pasa por el punto O. Por nivelación se pretende determinar la altura del punto M. Los segmentos Δh son los desniveles que podemos observar en las sucesivas estaciones de la nivelación OM.

El desnivel obtenible entre los dos puntos depende del camino por donde se realice el itinerario de la nivelación. Supongamos que de O a M la nivelación se realiza por dos caminos:

- El de O a K siguiendo por la superficie de nivel hasta M.
- El de O, a lo largo de la superficie del geoide hacia el punto M_1 , al punto M.

Vemos que en el primer camino la magnitud a obtener está determinada por el segmento OK, y en el segundo, por el segmento MM_1 . Debido al no paralelismo de las superficies de nivel, OK y M_1M son diferentes. Eligiendo otros caminos para el itinerario altimétrico obtendremos otros valores diferentes a los dos anteriores.

Esta dependencia del valor de la altura del punto respecto del camino de la nivelación genera una indeterminación que no es admisible en trabajos rigurosos de nivelación en territorios de extensión significativa, por lo que en grandes redes es necesario solucionarla, aunque en los más habituales trabajos la repercusión sea tolerable o irrelevante. Por tanto, el problema planteado es la determinación unívoca de las altitudes, cualquiera que sea el camino de la nivelación, debido al no paralelismo de las superficies de nivel del campo normal de la Tierra.

Cota geopotencial

La dificultad que presenta el que la suma de los desniveles no sea una función uniforme puede ser salvada si renunciamos al concepto geométrico de altitud, sustituyendo el desnivel por la diferencia de potencial gravitatorio.

Si además de diferencias de nivel tomamos valores de la gravedad, se podrá calcular a partir de un origen cualquiera el potencial gravimétrico W_M de otro punto:

$$W_M = W_O - \int_O^M g \cdot dh$$

con un número suficiente de medidas, no solo de dh , sino también de la gravedad g .

En general tomaremos como origen de altitudes $W_O = 0$ y llamaremos cota geopotencial de M a la expresión:

$$C_M = \sum_O^M g \cdot dh$$

Por tanto, la cota geopotencial es el trabajo que hay que realizar para transportar una masa de unidad de O a M. Las cotas geopotenciales son constantes sobre una superficie equipotencial. Vienen dadas en unidades geopotenciales (1 u.g.p. = 1 Kgal · m)

Estas cotas presentan el inconveniente de no tener dimensiones lineales y de que la diferencia entre los valores que se obtienen para dos puntos situados en la misma vertical no está directamente relacionada con la distancia entre ellos. Es decir, ni son altitudes, ni tienen sentido geométrico. Sus valores son aproximadamente un 2% menores ($C = 0,98 \cdot h$) que las altitudes sobre el geoide (h , expresada en metros).

El número geopotencial C es, en cierto sentido, el resultado más directo de nivelación y tiene gran importancia científica.

Partiendo de la fórmula:

$$W_M = \sum_O^M g \cdot dh$$

Como se trata de una función de punto, vamos a sustituirla por:

$$H_M = \frac{\sum_O^M g \cdot dh}{\gamma_{med}}$$

Donde γ_{med} = valor medio de la gravedad. Variable en función del tipo de altitud estudiada.

Este valor de la gravedad sólo depende de la situación del punto M, por lo que:

- H_M es una función uniforme.
- H_M tiene un valor métrico poco diferente de $\sum dh$

Podemos pasar fácilmente de H_M a C_M , y viceversa. Según sea el denominador, o valor para la gravedad, obtendremos los distintos sistemas de altitudes:

- Altitudes aproximadas
- Altitudes ortométricas
- Altitudes normales
- Altitudes dinámicas

Altitud aproximada

Estas altitudes se obtienen si no se considera el efecto del campo de la gravedad de la Tierra, es decir, sin realizar observación gravimétrica alguna ni aplicar correcciones.

En las mediciones así realizadas se asume como fuente de error la influencia del campo real de gravedad de la Tierra. El cálculo de las altitudes aproximadas en un paso previo a la obtención de las altitudes en otros sistemas.

Salvo grandes líneas o trabajos singulares, todas las nivelaciones adoptan este sistema.

Altitudes ortométricas

Se llaman altitudes ortométricas a la distancia desde los puntos sobre la superficie del geoide hasta la superficie terrestre, medidas a lo largo de las líneas verticales que pasan por esos puntos.

Las altitudes ortométricas no dependen del camino de nivelación. Además, estas altitudes pueden tener diferentes valores para puntos que están ubicados en una misma superficie de

nivel, puesto que las distancias desde el geoide hasta la superficie de nivel de dicho punto no son necesariamente constantes y dependerán de la fuerza de la gravedad.

Los puntos que tienen la misma altitud ortométrica, equidistan de la superficie de nivel de cota cero, por lo que es conocido como la “altitud natural” sobre el N.M.M.

Sin embargo, estas altitudes adolecen de un gran inconveniente: no pueden ser calculadas exactamente, ya que la magnitud de la gravedad dependen de una forma compleja de la anomalía gravimétrica local.

$$H = \frac{C}{g} = \frac{\sum_O^M g \cdot dh}{\bar{g}} \Leftrightarrow \bar{g} = \frac{1}{H} \int_0^H g dH$$

Para el cálculo de la gravedad media $\bar{g}$ a lo largo de la línea de la plomada, se necesita conocer los valores reales de gravedad entre el geoide y la superficie terrestre. Como no es posible medir directamente la gravedad dentro de la Tierra, se deben formular hipótesis respecto a la distribución de su masa. Son las denominadas reducciones de la gravedad.

Precisiones: Un error en la determinación de $\bar{g}$ de 100mgal produce un error de 10 cm en una altitud de 1000 m. Un error de 0,1 g/cm³ en la densidad produce un error de 4 mm en una altitud de 1000m.

La nivelación es una de las medidas geodésicas más precisas. Es posible una desviación típica de ± 0.1 mm por kilómetro de distancia. Aumenta con la raíz cuadrada de la distancia. Si los errores de medida, interpolación, etc., de la gravedad son despreciables, las diferencias en el número geopotencial C pueden determinarse con una precisión de 0.1 gal·m por Km de distancia. Esto corresponde a 0,1 mm en la altitud medida. Para alcanzar esto la distancia entre las estaciones gravimétricas de 2 á 3 Km en zonas llanas, de 1 a 2 Km en alturas moderadas y de 0,3 a 1,5 Km en áreas montañosas.

Altitudes normales

Se trata de un concepto similar al de altitud ortométrica con la diferencia de usar como superficie de referencia el elipsoide (que sería la superficie equipotencial correspondiente si no hubiera anomalías) en vez del geoide.

$$H^N = \frac{C}{\gamma} = \frac{\sum_O^M g \cdot dh}{\bar{\gamma}} \Leftrightarrow \bar{\gamma} = \frac{1}{H^N} \int_0^{H^N} \gamma \cdot dH^N$$

Ahora $\bar{\gamma}$ es la gravedad normal, medido a lo largo de la línea de la plomada perpendicular al elipsoide.

Las alturas normales originan otro sistema de altura sin necesidad de formular hipótesis. La superficie de referencia para las alturas normales es el *cuasigeoide* que está cerca del geoide y que no representa una superficie de nivel.

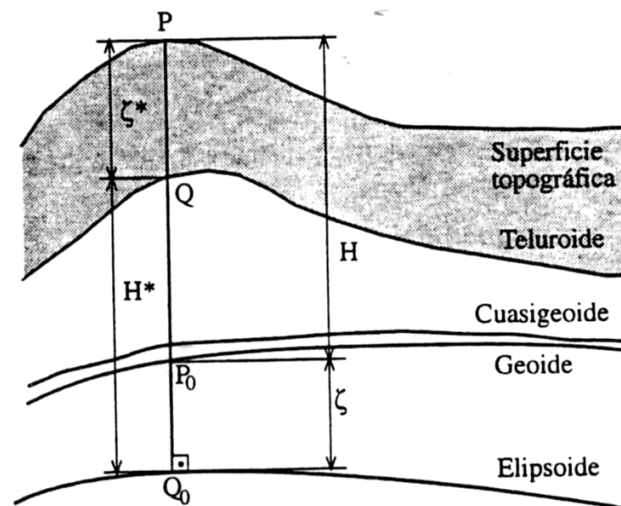


Figura 3. Fuente: "Geodesia"

Un punto Q se determina por H^* (H^N) como colocado en la línea de la plomada normal o con una buena aproximación sobre la normal del elipsoide que pasa por el punto de superficie, P . En el campo de gravedad normal, Q debería tener la misma diferencia de potencial C con respecto al elipsoide de nivel $U=U_0$ que la que el punto P tiene con respecto al geoide en el campo de gravedad de la tierra:

$$C=U_0-U_Q=W_0-W_P$$

Si se quiere que los potenciales del elipsoide y del geoide sean iguales ($U_0=W_0$) entonces tenemos que $U_Q=W_P$. La superficie en la cual para cada punto $U_Q=W_P$ se la denomina *Teluroide*. La altura normal H^* de un punto P es equivalente a la altura del correspondiente punto del teluroide, Q , sobre el elipsoide. El teluroide no es una superficie de nivel del campo de la gravedad normal.

Altitudes dinámicas

Se obtienen cuando para la gravedad se toma un valor fijo, gravedad normal para una latitud estándar arbitraria, normalmente 45° , cuyo valor en el elipsoide internacional es de 980,6294 gales.

$$H^{DIN} = \frac{\sum_{o}^M g \cdot dh}{\gamma_{45^\circ}}$$

Los puntos de una misma superficie de nivel tienen la misma altitud dinámica y al dividir el número geopotencial por una constante, la convierte en una longitud (dimensiones de altitud), pero no tiene significado geométrico.

5.4. Nivelación geométrica y trigonométrica

La nivelación geométrica se caracteriza porque las visuales son siempre horizontales. Los instrumentos empleados son los *niveles o equialtímetros*. La nivelación puede dividirse en:

Simples: Cuando el desnivel a medir se obtiene con una sola determinación.

Compuestas: Cuando se emplean varias determinaciones.

La nivelación trigonométrica permite obtener desniveles entre puntos con observaciones de distancias cenitales de cualquier inclinación. Emplea los taquímetros y las estaciones totales.

5.4.1. Métodos de Nivelación geométrica simple

a) *Método del punto medio.*

Sean A y B dos puntos cuyo desnivel se quiere determinar. Este método consiste en estacionar el nivel (E) entre A y B, de tal forma que la distancia existente a ambos puntos sea la misma, es decir $EA = EB$. El desnivel será:

$$z_A^B = m_A - m_B = m_{Espalda} - m_{Frente}$$

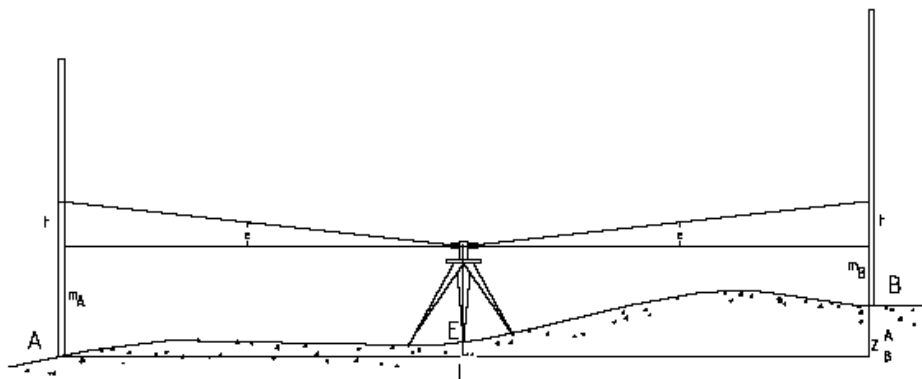


Figura 4. Fuente: "Métodos topográficos y oficina técnica"

Si el instrumento tiene un error residual de corrección, las visuales no serán exactamente horizontales, pero su influencia en las lecturas a las dos miras será igual, debido a la equidistancia del nivel respecto de los puntos y al hacer la diferencia de lecturas, esta influencia se elimina.

También debido a la igualdad de distancias, este método elimina la corrección por esfericidad y refracción, por ser igual la que corresponde a cada una de las miras.

En la práctica se demuestra que el límite de distancias para conseguir lecturas en las que se asegure el mm de precisión, es de 80 a 100 m. Esto representa distancias de 160 a 200 m, entre los puntos cuyo desnivel se desea obtener.

b) Método del punto extremo.

Sean A y B los dos puntos cuyo desnivel queremos determinar. Para ello, se estaciona el nivel en el punto A, a una altura sobre el suelo i_A y se visa a la mira situada en B, efectuándose la lectura m_B . El desnivel será:

$$z_A^B = i_A - m_B$$

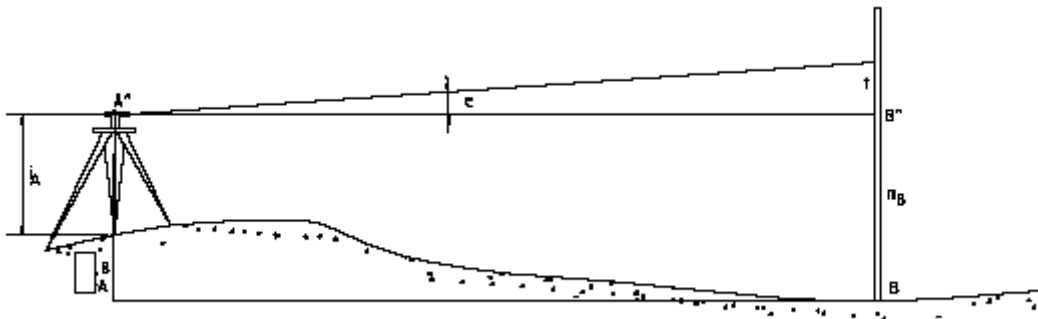


Figura 5. Fuente: “Métodos topográficos y oficina técnica”

Analizando la expresión observamos que la precisión del método es inferior a la que se obtiene con el método del punto medio. La altura del instrumento no puede medirse con precisión del milímetro. Esto supone una precisión del orden del cm. o del medio centímetro.

Por otra parte, en este método no se elimina el error del instrumento y tampoco se elimina el error de esfericidad y refracción.

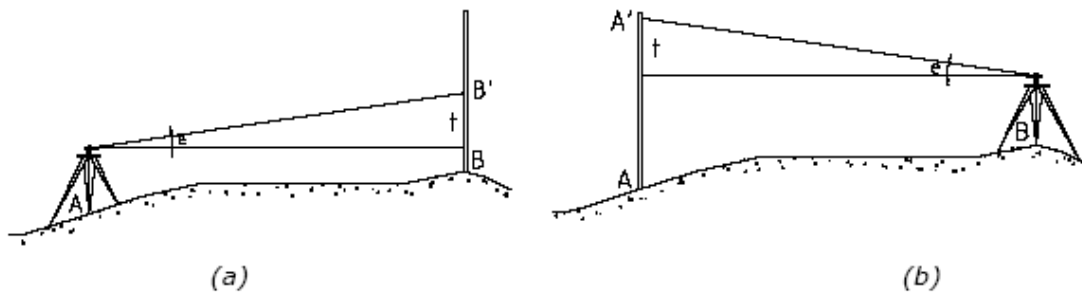
c) Método de estaciones reciprocas.

Se emplea para eliminar los efectos del error residual (e) y los efectos de la esfericidad y la refracción. Con ello se mejora también la precisión.

Se determina primeramente el desnivel z_A^B por el método del punto extremo. A continuación se intercambian las posiciones entre el nivel y la mira. El desnivel será:

$$z_A^B = \frac{i_A - i_B}{2} + \frac{m_A - m_B}{2}$$

En esta expresión no aparece el término relativo al error residual que produce errores en las lecturas, pero se siguen midiendo magnitudes con diferente precisión (i y m)



d) Método de estaciones equidistantes.

Se coloca el instrumento entre las miras A y B, pero no equidistante de estas y se realizan las lecturas. Después se sitúa el aparato en E', de modo que E'B sea igual que la distancia EA del paso anterior, y se vuelve a leer sobre las miras.

El desnivel será:

$$z_A^B = \frac{m_A - m_B}{2} + \frac{m'_A - m'_B}{2}$$

Los resultados obtenidos son más precisos que los obtenidos empleando estaciones recíprocas, ya que sólo intervienen las alturas de las miras y se eliminan además los efectos de esfericidad y refracción.

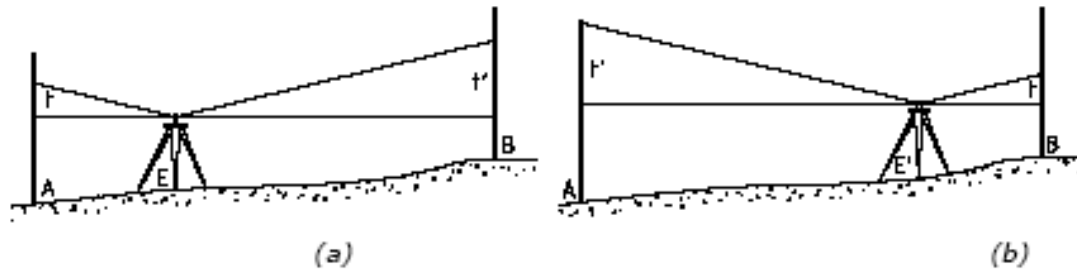


Figura 6. Fuente: "Métodos topográficos y oficina técnica"

e) Método de estaciones exteriores

Se elimina la condición de equidistancia del aparato a las miras del método anterior, colocando el instrumento exterior a las miras.

El desnivel será:

$$z_A = \frac{m_A - m_B}{2} + \frac{m'_A - m'_B}{2}$$

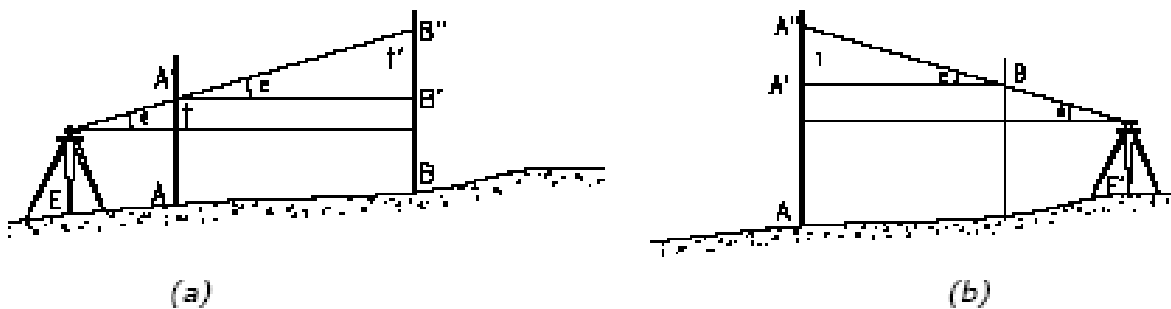


Figura 7. Fuente: "Métodos topográficos y oficina técnica"

5.4.2. Métodos de nivelación compuesta

a) Nivelación sencilla.

Se emplea cuando los puntos, cuyo desnivel se quiere determinar, están muy separados o la diferencia de nivel es mayor de la que puede medirse de una sola vez.

El procedimiento de observación es como en el método del punto medio, pero una vez realizadas las lecturas a las miras A y B, la mira situada en A, avanza hasta colocarse por delante de la B, mientras que la mira en B gira sobre su vertical, de manera que se vuelve a repetir el esquema de observación del punto medio, colocando el instrumento equidistante entre ambas.

El desnivel entre los puntos A y E se calcula:

$$z_A^E = (m_e - m_f) + (m'_e - m'_f) + \dots + (m^n_e - m^n_f) = \sum m_e - \sum m_f$$

Para tener comprobación de la bondad de las operaciones realizadas hay que empezar y terminar en puntos de altitud conocida. De esta manera se puede comprobar que el desnivel obtenido entre esos puntos z_A^E no coincide con el teórico, conocido $z_A^E = Z_E - Z_A$

A la diferencia entre estas dos cantidades se le denomina *error de cierre altimétrico*:

$$e_c = z_A^E - Z_A^E$$

Para que resulte admisible deberá ser menor que una tolerancia T dada.

$$e_c \leq e_k \sqrt{K} = T$$

donde K es la longitud de la nivelación en kilómetros y e_k el error kilométrico. Este error se determina a partir de las características del nivel (errores de puntería e_p y verticalidad e_v) y de la distancia máxima admitida entre el nivel y la mira (L_N).

$$e_k = \frac{\sqrt{e_p^2 + e_v^2}}{636620} \cdot L_N \cdot \frac{1000}{L_N}$$

$$e_p = \frac{cte^{cc}}{A} \cdot \left(1 + \frac{4 \cdot A}{100}\right)$$

Siendo; cte retículo = 150 normal, cte = 60 bisección y cte = 30 cuña.

A = Aumentos del antejo.

$$e_v = \frac{S^{cc}}{3} \text{ (Burbuja normal)} \quad S = \text{Sensibilidad del nivel}$$

$$e_v = \frac{S^{cc}}{20} \text{ (Burbuja partida)}$$

Cuando se cumple esta condición se procede a compensar el error de cierre. Para ello se pueden emplear dos criterios: repartir el error *proporcionalmente a los desniveles* parciales o a las *longitudes de los tramos* (normalmente se emplea este).

Con los desniveles compensados se calculan las altitudes de los puntos intermedios. Para ello basta sumar a la altitud del punto inicial el desnivel del primer tramo y se obtiene la del segundo punto, una vez obtenida esta, se le suma el desnivel del segundo tramo y se obtiene la altitud del

tercer punto, y así sucesivamente hasta llegar al último punto, para el que deberá encontrarse una altitud igual a la conocida. Esta operación se llama *corrida de altitudes*.

Este modo de trabajar denominado nivelación sencilla sólo tiene comprobación de los resultados al calcular el error de cierre y si este no resulta tolerable hay que repetir todo el trabajo.

b) Nivelación doble.

Para evitar repetir los trabajos cuando las medidas no son tolerables y al mismo tiempo mejorar la precisión, se efectúan las medidas por duplicado .

Se divide el conjunto de la línea en *trozos o anillos* de unos 2Km de longitud, de modo que los extremos de cada anillo sean puntos materializados en el terreno y estables. Se efectúa la nivelación en un sentido empleando el método del punto medio. Una vez concluida la nivelación de *ida*, se realiza la de *vuelta*, en sentido contrario.

Comprobando que todos los cierres de los anillos son tolerables, se suman independientemente los desniveles de ida y los de vuelta, obteniéndose dos valores del desnivel total de la línea, cuya diferencia tiene que ser también tolerable. Si alguno de los cierres de los anillos no resulta tolerable se repite sólo ese anillo en lugar de repetir toda la línea. Ahora;

$$e_c \leq e_k \sqrt{K} \sqrt{2} = T$$

Comprobada la bondad de los cierres de los anillos y de la línea, se compensa la nivelación, repartiendo los errores proporcionalmente a los desniveles parciales o a la longitud de los tramos.

5.4.3. Nivelación trigonométrica o por pendientes

a) Nivelación trigonométrica simple.

Para determinar el desnivel del punto B respecto del A se estaciona un instrumento en el punto A y se visa a B, midiendo la distancia cenital y la distancia geométrica (D_g) entre ambos puntos. La expresión será:

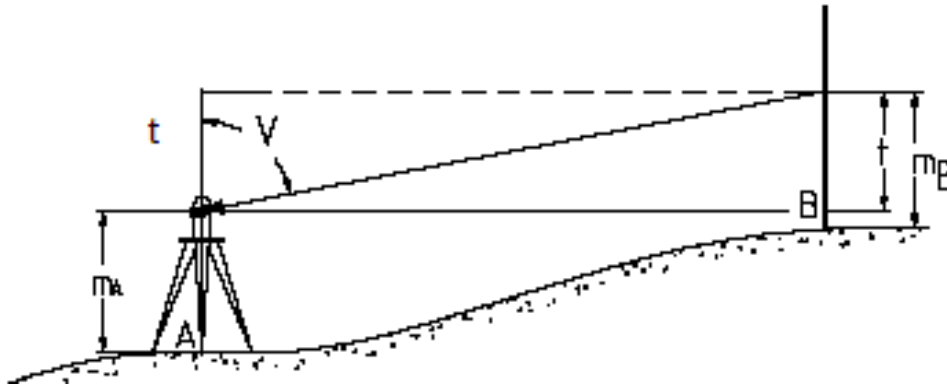


Figura 8. Fuente: "Métodos topográficos y oficina técnica"

$$z_A^B = D_g \cdot \cos V + i_A - m_B = t_A^B + i_A - m_B$$

Cuando las distancias son pequeñas, no es necesario tener en cuenta las correcciones por esfericidad y refracción. Sin embargo, en grandes distancia sí que es necesario, y la anterior expresión quedaría de la forma;

$$z = t + i - m + (c_e - c_r) = t + i - m + (0,5 - K) \frac{D_g^2}{R} = t + i - m + 0,42 \frac{D_g^2}{R}$$

Siendo K el coeficiente de refracción, y R el radio terrestre.

El *error máximo en el desnivel* será la componente cuadrática de los errores en t , en i y en m .

$$e_z \leq \sqrt{e_t^2 + e_i^2 + e_m^2}$$

- El error en t , es función de los errores en las medidas de distancias cenital e_{Tv} y geométrica e_D .

- El error total en la medida cenital $e_{Tv} = \sqrt{e_v^2 + e_p^2 + e_l^2}$ siendo

Error de verticalidad: $e_v \leq \frac{s''}{3}$ s sensibilidad.

Error de puntería: $e_p \leq \frac{150^{cc}}{A} (1 + \frac{4A}{100})$ A aumentos del anteojo.

Error de lectura: $e_l \leq \frac{2}{3} a^{cc}$ a apreciación del instrumento

- El error en la medida de la distancia, dependerá fundamentalmente de las características del instrumento utilizado, y de la distancia medida.
- El error en i es el cometido en la determinación de la altura del instrumento:

$$1mm < e_i < 5mm$$

- El error por falta de verticalidad de la mira, afecta directamente en la medida del término m .

$$e_m < m_B \frac{\beta''}{r''} \operatorname{tg}(\alpha + \beta)$$

donde α es la inclinación de la visual, β es la inclinación de la mira y m_B es la altura de la mira o prisma.

b) Nivelación trigonométrica por estaciones recíprocas.

En este caso se determina el desnivel estacionando primero en el punto A y se visa a B, a continuación, se estaciona el instrumento en B y se visa al punto A, calculándose los desniveles directo y recíproco.

Teóricamente ambos valores deberían ser iguales y de signos contrarios. La discrepancia entre ellos da una idea de la precisión alcanzada y en caso de ser tolerable, se toma como desnivel definitivo la semidiferencia:

$$z'_A = \frac{z_A^B - z_B^A}{2} \qquad T \leq e_z \sqrt{2}$$

c) Nivelación trigonométrica compuesta por estaciones recíprocas.

Si para obtener el desnivel entre dos puntos A y E, hay que hacer varias determinaciones, se empieza situando el instrumento en A y se hace observaciones sobre la mira en B. A continuación se coloca el instrumento en B y se visa a A y a C. Después desde C se realizan medidas sobre B y D y así hasta estacionar en E, desde donde se observa a D. Se podría haber hecho una nivelación simple de cada tramo, pero observando al punto anterior y al siguiente se obtiene una comprobación de los desniveles.

Los cálculos se desarrollan deduciendo los desniveles directo y recíproco de los tramos, y si las diferencias son tolerables, se toma la semidiferencia como valor para el tramo. Para obtener el desnivel entre A y E se suman estos desniveles promediados.

$$z_A^E = z_A^B + z_B^C + z_C^D + z_D^E$$

Para tener comprobación de los resultados, las líneas de nivelación comienzan y acaban en puntos de altitud conocida, por lo que se podrá calcular: $z_A^E = Z_E - Z_A$

Y el error de cierre altimétrico será:

$$e = z_A^E - z_A^E$$

que para que resulte admisible, debe ser menor que una determinada tolerancia.

d) Cálculo del desnivel por visuales recíprocas y simultáneas:

También puede calcularse el desnivel entre dos puntos A y B estacionando dos instrumentos en estos puntos y realizando observaciones verticales recíprocas y simultáneas, por lo que los efectos angulares de refracción en las dos visuales serán iguales, y quedarán eliminados también los de esfericidad. La expresión será:

$$z_A^B = D \cdot \operatorname{tg} \frac{1}{2}(V' - V)$$

Las distancias cenitales V y V', son las que correspondería observar sin alturas de instrumento y mira. Es decir, que *las distancias cenitales que han sido observadas deberán ser reducidas al terreno.*

La simultaneidad en las observaciones puede sustituirse operando en condiciones atmosféricas semejantes.

Bibliografía

- [1] Balbuena Durán, J. L. y Vara Jaén, M^a. D., “Instrumentación y metodología empleadas en las técnicas altimétricas”, Topografía y Cartografía. Revista del Colegio Oficial de Ingenieros Técnico en Topografía, nº 74. Mayo-Junio 1996.
- [2] Martín Asín, F., “Geodesia y Cartografía Matemática”, Editorial Paraninfo, pág. 135 a 167.
- [3] Ojeda, J. L., “Métodos topográficos y oficina técnica”, pág. 205 a 237
- [4] Hweiskanen, W. A. y Moritz, H., “Geodesia Física”, pág. 160 a 178.
- [5] Calero, E. “Apuntes de Geodesia de la E.P.E.S.”

Tema 6. Sistemas elipsoidales de referencia. Parámetros del elipsoide.

Latitud geodésica, geocéntrica y reducida. Curvatura del elipsoide.

6.1. Sistemas elipsoidales de referencia. Parámetros del elipsoide

La superficie de la tierra se puede aproximar por un elipsoide biaxial de revolución achatado por los polos. Este elipsoide de revolución se crea por rotación de un plano meridiano alrededor de su eje menor hasta completar 360°.

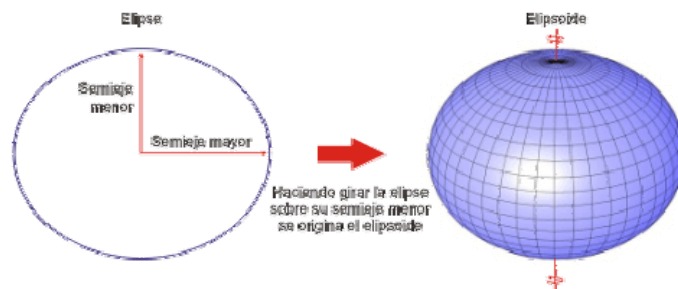


Figura 1. Elipsoide de revolución. (Fuente: gabrielortiz.com)

El elipsoide se define por dos parámetros geométricos: semieje mayor a y semieje menor b , aunque generalmente el semieje menor se suele sustituir por cantidades más pequeñas que son más adecuadas para desarrollos en serie como es el aplastamiento o achatamiento del elipsoide α , la excentricidad lineal ε o la primera y segunda excentricidades e , e' , respectivamente:

$$(1) \quad \alpha = \frac{a-b}{a}, \varepsilon = \sqrt{a^2 - b^2}, e^2 = \frac{\varepsilon}{a^2}, e'^2 = \frac{\varepsilon}{b^2}$$

En la elipse meridiana podemos observar los parámetros que la definen:

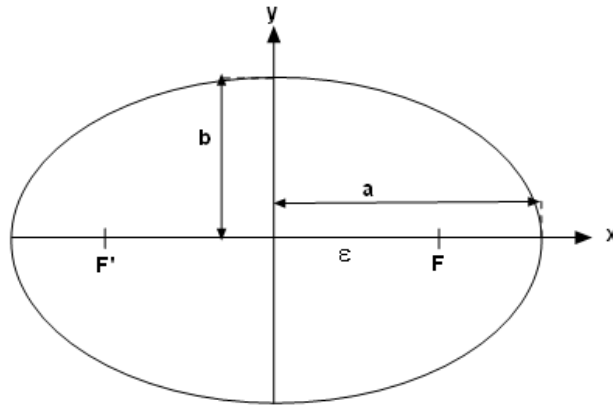


Figura 2. Elipse meridiana

Es importante también conocer las relaciones que existen entre los diferentes parámetros:

$$(2) \quad \frac{b}{a} = 1 - \alpha = \sqrt{1 - e^2} = \frac{1}{\sqrt{1 + e'^2}}$$

Para definir el sistema de coordenadas en el elipsoide, se introduce un sistema de coordenadas cartesiano espacial x, y, z. El origen de este sistema es O (ver Figura 3) y el eje z, coincide con el eje menor del elipsoide.

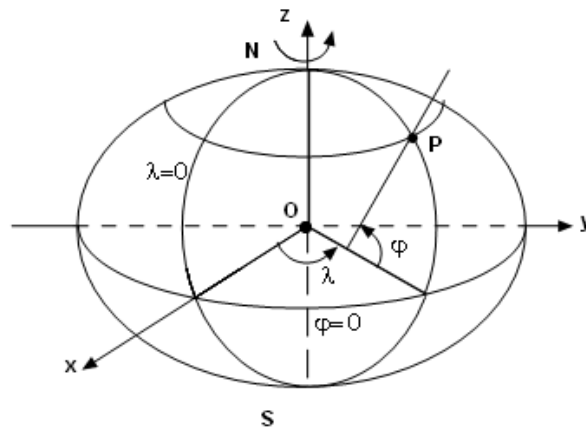


Figura 3. Sistema de coordenadas en el elipsoide de revolución.

La ecuación de la superficie del elipsoide de revolución en el sistema cartesiano es la siguiente:

$$(3) \quad \frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2} - 1 = 0$$

6.2. Latitud geodésica, geocéntrica y reducida

El sistema de coordenadas geográficas elipsóidicas se define por la latitud geográfica φ y la longitud geográfica λ (también, latitud y longitud geodésicas). La latitud φ es el ángulo medido en el plano meridiano entre el plano ecuatorial del elipsoide (plano x,y) y la normal elipsoidal en P (un punto sobre la superficie real terrestre). La latitud es positiva hacia el Norte y negativa hacia el Sur. La longitud λ es el ángulo medido en el plano ecuatorial entre el meridiano cero (eje x) y el plano del meridiano de P y el eje z. La longitud es positiva hacia el Este y negativa hacia el Oeste.

Además de la latitud geográfica o geodésica es necesario definir la latitud geocéntrica y la latitud reducida y sus relaciones. Observando la siguiente figura podemos ver cada una de estas latitudes:

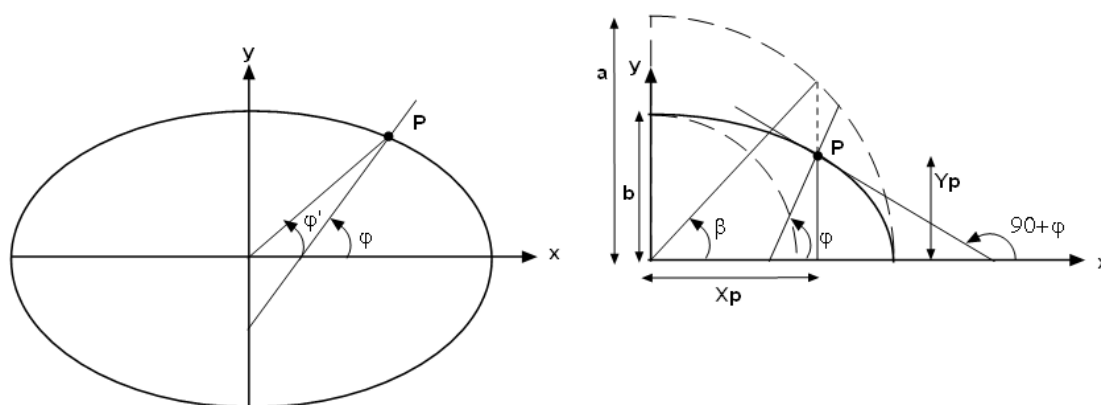


Figura 4. (a) Latitud geográfica φ y geocéntrica φ' . (b) Latitud geográfica φ y reducida β .

La pendiente de la tangente de una curva $y=f(x)$ es dy/dr . Esta tangente forma un ángulo de $90+\varphi$ con el eje r, entonces:

$$(4) \quad \frac{dy}{dx} = \operatorname{tg}(90 + \varphi) = -\operatorname{ctg}(\varphi)$$

Si derivamos en la ecuación de la elipse meridiana $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, tenemos:

$$(5) \quad \frac{2x}{a^2} dx + \frac{2y}{b^2} dy = 0 \xrightarrow{\text{despejando}} \frac{dy}{dx} = -\frac{x}{y} \frac{b^2}{a^2}$$

sustituyendo en la expresión (4), nos queda:

$$(6) \quad \operatorname{tg} \varphi = \frac{a^2}{b^2} \frac{y}{x}$$

Como la latitud geocéntrica verifica que $\tan \varphi' = \frac{y}{x}$, entonces llegamos a la relación entre la latitud geodésica y la geocéntrica:

$$(7) \quad \operatorname{tg} \varphi = \frac{a^2}{b^2} \operatorname{tg} \varphi'$$

Fijándonos de nuevo en la Figura 4.(b), se deduce que $x = a \cos \beta$; $y = b \operatorname{sen} \beta$, por lo tanto:

$$(8) \quad \frac{b}{a} \operatorname{tg} \beta = \frac{y}{x}$$

sustituyendo este valor en la expresión (6), llegamos a la expresión que relaciona la latitud geodésica con la reducida:

$$(9) \quad \operatorname{tg} \varphi = \frac{a}{b} \operatorname{tg} \beta$$

6.3. Curvatura del elipsoide

Los meridianos y los paralelos son líneas de curvatura del elipsoide de rotación. Para definir la curvatura del elipsoide tenemos que conocer los siguientes elementos de la geometría del elipsoide:

Planos normales: son planos que contienen al elipsoide que pasa por un punto considerado.

Secciones normales: son las curvas intersección de los planos normales con la superficie del elipsoide.

Secciones normales principales: en cada punto del elipsoide existen dos secciones normales mutuamente perpendiculares que tienen curvatura máxima y mínima respectivamente.

a.- Sección meridiana (ρ):

$$(10) \quad \rho = \frac{a(1 - e^2)}{(1 - e^2 \operatorname{sen}^2 \varphi)^{3/2}}$$

b.- Sección del primer vertical (N):

$$(11) \quad N = \frac{a}{(1 - e^2 \operatorname{sen}^2 \varphi)^{1/2}}$$

Con estos elementos definidos en la elipse meridiana, una de las secciones es el propio meridiano y la otra es la producida al elipsoide por el plano que conteniendo a la normal en el punto es perpendicular a dicho plano meridiano. De esta forma, todas las secciones que obtendríamos producidas por el haz de planos cuyo eje fuera normal a la superficie en el punto considerado tienen radios de curvatura con valores comprendidos entre N y ρ .

Tanto N como ρ , llamados radios de curvatura principales, dependen únicamente de la latitud φ , por lo que todos los puntos en un mismo paralelo tendrán los mismo radios de curvatura principales.

En este sentido, el valor de los radios de curvatura principales son iguales en los polos con

latitud $\varphi = 90^\circ$ y en el ecuador, con $\varphi = 0^\circ$ son: $N_0 = a$ y $\rho = \frac{b^2}{a}$.

La curvatura de una sección normal arbitraria sobre el elipsoide de revolución tendiendo en cuenta una dirección α , es calculada por ([2] Pág. 47):

$$(12) \frac{1}{R_\alpha} = \frac{\cos^2 \alpha}{\rho} + \frac{\sin^2 \alpha}{N}$$

siendo R_α , el radio de curvatura. El acimut geodésico α es el ángulo medido en el plano horizontal entre el plano de la elipse meridiana por el punto estación y el plano vertical definido por la normal al punto estación y por el punto visado.

El radio medio de curvatura R se define como el límite de la media aritmética de los radios de curvatura de las secciones normales cuando el número de éstos tiende a infinito¹:

$$(13) R_i = \frac{\sum_{\alpha=0}^{2\pi-\Delta\alpha} \frac{\rho N}{N \cos^2 \alpha + \rho \sin^2 \alpha}}{\frac{2\pi}{\Delta\alpha}} = \frac{\sum_{\alpha=0}^{2\pi-\Delta\alpha} \frac{\rho N \Delta\alpha}{N \cos^2 \alpha + \rho \sin^2 \alpha}}{2\pi}$$

$$(14) R = \lim_{\Delta\alpha \rightarrow 0} R_i = \frac{4}{2\pi} \int_0^{\frac{\pi}{2}} \frac{\rho N}{N \cos^2 \alpha + \rho \sin^2 \alpha} d\alpha = \sqrt{\rho N}$$

Este resultado tiene un significado geométrico y es conocido como radio medio de curvatura y su inversa, como curvatura de la curva en un punto, en este caso curvatura del elipsoide en un punto.

Una vez conocidos todos los elementos de la geometría del elipsoide podemos estudiar el sistema de coordenadas elipsoidal espacial. Además de las coordenadas latitud y longitud geográficas ya introducidas, para puntos sobre la superficie física de la Tierra, debemos considerar la altitud h sobre el elipsoide. Las coordenadas elipsoidicas espaciales φ , λ , h se denominan geodésicas.

¹ [3] Zakatov, pág. 44

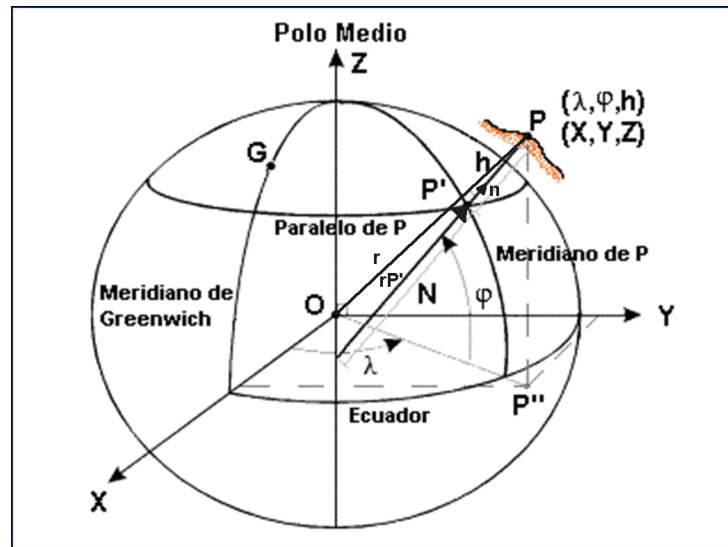


Figura 5. Sistema de coordenadas elipsoidico espacial. (Fuente: gabrielortiz.com)

Las coordenadas del punto P' se calculan por:

$$(15) \ rP' = \begin{pmatrix} x_{P'} \\ y_{P'} \\ z_{P'} \end{pmatrix} = N \begin{pmatrix} \cos \varphi \cdot \cos \lambda \\ \cos \varphi \cdot \text{sen} \lambda \\ (1 - e^2) \cdot \text{sen} \varphi \end{pmatrix}$$

siendo N el radio de curvatura del primer vertical.

Para el punto P tenemos:

$$(16) \ r = r_p + hn, \ n = \begin{pmatrix} \cos \varphi \cdot \cos \lambda \\ \cos \varphi \cdot \text{sen} \lambda \\ \text{sen} \varphi \end{pmatrix}$$

Y finalmente llegamos a la expresión que relaciona el sistema de coordenadas cartesiano con el sistema de coordenadas elipsoidico:

$$(17) \ r = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (N + h) \cos \varphi \cdot \cos \lambda \\ (N + h) \cos \varphi \cdot \text{sen} \lambda \\ ((1 - e^2)N + h) \cdot \text{sen} \varphi \end{pmatrix}, \text{ nombramos cada ecuación: } \begin{pmatrix} I \\ II \\ III \end{pmatrix}$$

En el caso de tener las cartesianas y necesitar las elipsoidicas, se resuelve con un proceso iterativo en un sistema de ecuaciones que converge rápidamente ya que $h \ll N$:

$$(18) \ h = \frac{\sqrt{x^2 + y^2}}{\cos \varphi} - N \quad \text{Resultado de I+II}$$

$${}^{(19)} \varphi = \operatorname{arctg} \left\{ \frac{z}{\sqrt{x^2 + y^2}} \left(1 - e^2 \frac{N}{N + h} \right)^{-1} \right\} \quad \text{Res. de: (I+II)/III}$$

$${}^{(20)} \lambda = \operatorname{arctg} \frac{y}{x} \quad \text{Resultado de: II/I}$$

Bibliografía

- [1] Martín Asín, F. *Geodesia y cartografía matemática*. 3ª Edición, Editorial Paraninfo. Madrid 1990
- [2] Torge, W. *Geodesy*. 2nd Edition. Walter de Gruyter, Berlin – New York, 1991. Pags. 44-49.
- [3] Zakatov, P.S. *Curso de Geodesia Superior*. Editorial Rubiños-1860. Madrid, 1997.

Tema 7. La esfera celeste. El movimiento diurno. Sistemas de coordenadas en Astronomía: horizontales, horarias, ecuatoriales absolutas, eclípticas. Transformaciones. Movimiento aparente del Sol. Teoría de las anomalías. La ecuación del tiempo. El tiempo: tiempo rotacional, tiempo de efemérides, tiempo atómico. Correcciones astronómicas: movimiento propio, precesión, nutación, paralaje, aberración, refracción atmosférica.

Seguidamente se exponen los principales conceptos de la Astronomía geodésica clásica, atendiendo a lo desarrollado en la referencia Mena (2008) de la cual se han extraído los distintos apartados que siguen.

7. 1. La esfera celeste. El movimiento diurno

Se entiende por esfera celeste una esfera de radio infinito cuyo centro coincide con el de la Tierra, en la cual se suponen proyectados todos los astros según líneas radiales. El radio de la esfera celeste se toma igual a la unidad, reduciéndose la Tierra al centro T de la esfera. Prolongando el eje terrestre de rotación y el ecuador, se obtienen respectivamente los *polos celestes* P y P' y el *ecuador celeste* QQ' . Los planos paralelos al ecuador celeste se denominan *paralelos celestes*, y los planos que contienen a la línea de los polos, *meridianos celestes*.

Consideremos un punto de la Tierra de latitud Φ y el meridiano celeste que pasa por el mismo (plano de la figura 1, izquierda. Ref. [11]). La vertical física en el lugar de estación corta a la esfera celeste en los puntos C y C' , denominados respectivamente *cenit astronómico* y *nadir astronómico*. Todo plano que contiene a la vertical se denomina *vertical astronómico*. Entre los infinitos verticales de cada lugar de estación son especialmente importantes el meridiano y el perpendicular al mismo, conocido como *primer vertical*. El plano trazado perpendicularmente a la vertical por el centro de la esfera celeste se llama *horizonte astronómico* del lugar. En este plano se encuentra la dirección Norte – Sur (línea NS) denominada *meridiana*. Como consecuencia de estas definiciones, el arco QC sobre el plano meridiano es la *latitud astronómica* del punto, y el arco PC , complemento del anterior, la *colatitud astronómica*. La línea de polos PP' divide al meridiano del lugar en dos semiplanos: el *meridiano superior*, que contiene al cenit y al punto Sur, y el *meridiano inferior o antimeridiano*, que pasa por el nadir y por el punto Norte.

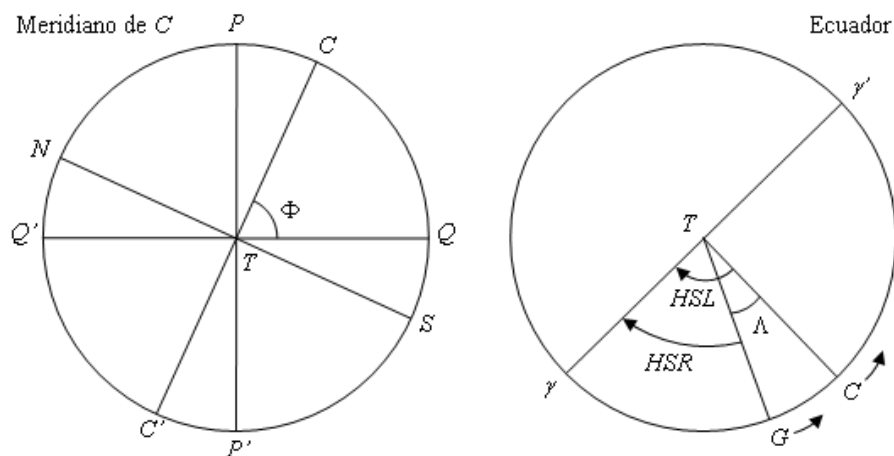


Figura 1.

Atendiendo ahora al plano del ecuador celeste (plano de la figura 1, derecha. Ref. [11]), su intersección con el plano definido por la órbita de traslación de la Tierra alrededor del Sol, denominado plano de la *eclíptica*, determina la recta llamada *línea de los equinoccios* ($\gamma\gamma'$). De los dos equinoccios, que son puntos relativamente “fijos” en el espacio si comparamos su movimiento con la rotación terrestre, el referente al inicio de la primavera en el hemisferio norte y del otoño en el hemisferio sur se denomina *punto vernal o punto Aries* (γ).

El punto vernal y los polos celestes determinan el meridiano celeste que se toma como referencia para definir la *Hora Sidérea Local* (HSL) de cualquier meridiano. De este modo, para cada lugar C de la Tierra, dicho ángulo es el formado en el plano del ecuador por el meridiano del lugar (TC) con el meridiano de referencia ($T\gamma$). Particularmente, la hora sidérea local del meridiano de Greenwich (TG), origen de longitudes, se denomina Hora Sidérea Greenwich (HSG), o bien *Hora Sidérea Reducida* (HSR). La hora sidérea se mide en sentido horario de 0 a 24 horas, estando determinada la longitud astronómica Λ de cualquier meridiano por la igualdad:

(1)

$$\boxed{\Lambda = HSL - HSR}$$

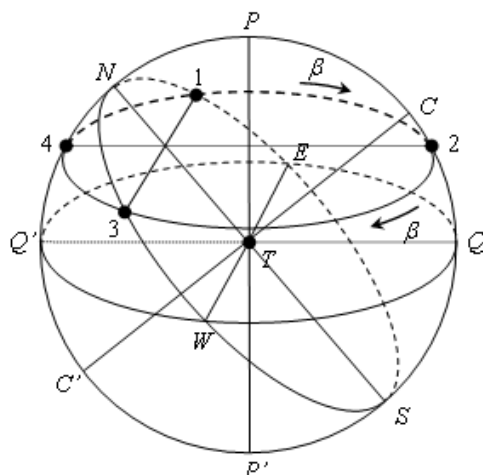


Figura 2.

Consideremos una estrella β cualquiera que suponemos “fija” en la esfera celeste, y sea C un lugar cualquiera de la Tierra. Para un observador situado en esta latitud, la rotación terrestre hace que aparentemente el astro describa en sentido Este–Oeste un paralelo celeste completo en cada giro de 360° del planeta. Como se muestra en la figura 2 (Ref. [11]), en el caso general la trayectoria tiene un *orto* (1) y un *ocaso* (3), coincidiendo dichos puntos con los cortes del paralelo del astro con el horizonte del lugar. Entre estas dos posiciones, la estrella recorre la parte visible de su ruta al Este (E) del observador, hasta alcanzar el meridiano (2). Este instante se corresponde con la *culminación o paso superior* del astro. A partir de aquí la estrella describe la zona visible de su ruta al Oeste (W) del observador, hasta llegar al ocaso. Finalmente, entre el ocaso y el nuevo orto la estrella se encuentra bajo el horizonte del lugar, y por tanto oculta, cruzando nuevamente el meridiano del lugar (4) en el llamado *paso inferior*.

Las estrellas cuyo paralelo está situado íntegramente por encima del horizonte del lugar de observación, y que por tanto tienen su paso inferior en el arco NP , son visibles de forma permanente desde dicha latitud. Estas estrellas se conocen como *circumpolares* del lugar. No debe confundirse este concepto con el de estrellas *cenitales* o *circuncenitales* de un lugar, las cuales son aquellas que culminan sobre el cenit de un lugar o muy próximas al mismo.

7.2. Sistemas de coordenadas en Astronomía: horizontales, horarias, ecuatoriales absolutas, eclípticas. Transformaciones.

Como la posición de una estrella respecto del punto vernal es teóricamente independiente de la rotación terrestre y por lo tanto constante, se establece el sistema de *coordenadas ecuatoriales* o *coordenadas absolutas*, con objeto de definir la posición de un astro sobre la esfera celeste. Ta-

les coordenadas se representan en la figura 3 (Ref. [11]) y son dos magnitudes angulares denominadas *ascensión recta* (α) y *declinación* (δ). La primera se mide sobre el ecuador celeste, y queda definida por el diedro que determinan el meridiano que contiene al punto vernal y el que pasa por la estrella. Varía de 0 a 24 horas en sentido antihorario a partir del meridiano de referencia. La segunda coordenada se define como el ángulo medido en el meridiano celeste de la estrella, formado por el ecuador y el radio que pasa por la estrella. La declinación se expresa de 0 a 90 grados sexagesimales, norte o sur, o bien positiva y negativa respectivamente.

Si en lugar del ecuador y los polos celestes, consideramos la eclíptica y los polos de este plano, análogamente a las coordenadas ecuatoriales se definen las *coordenadas eclípticas*. Dichas coordenadas se expresan ambas en grados sexagesimales y son: la *longitud eclíptica* (μ), definida tomando como referencia el punto vernal de forma análoga a la ascensión recta, y la *latitud eclíptica* (l), semejante a la declinación. Al igual que el sistema ecuatorial, las coordenadas eclípticas de un astro son independientes de la rotación terrestre.

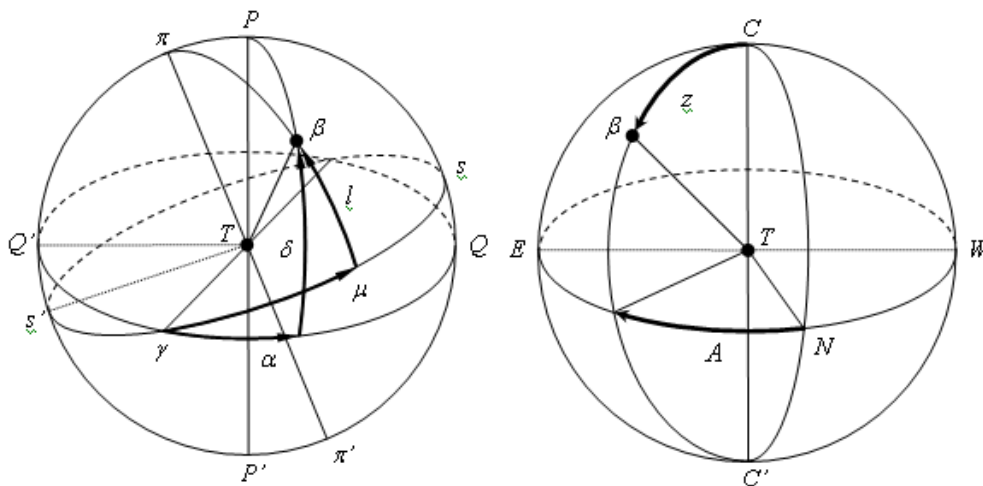


Figura 3. a y b

Por otra parte, para establecer una dirección radial en la esfera celeste con relación al horizonte y al cenit astronómicos de un lugar de la Tierra, se utilizan las denominadas *coordenadas horizontales*. Éstas son el *acimut astronómico* (A) y la *distancia cenital* (z), las cuales podemos materializar con un aparato de observación (figura 3b. Ref. [11]). Para una estrella sobre la esfera celeste, el acimut es el ángulo diedro que forma el meridiano del lugar con el plano vertical que pasa por la estrella. Este ángulo se mide en grados sexagesimales a partir del norte y en sentido horario. La segunda coordenada también se mide en grados; en este caso tomando como origen la dirección del cenit y como plano de observación el vertical que pasa por el astro.

Como los datos de posición de un astro en el sistema horizontal varían continuamente, las coordenadas en dicho sistema deben ir acompañadas de un parámetro adicional: la hora para la cual fueron determinadas. Como tal puede utilizarse la hora sidérea local, o bien la *Hora de Tiempo Universal* (HUT) correspondiente que se trata más adelante.

Al ser el cenital la distancia angular del cenit a la estrella y la declinación la distancia angular de la estrella al ecuador, si se observa un astro de declinación conocida en el momento en que se encuentra en el meridiano del lugar, entonces el ángulo de latitud correspondiente se obtiene mediante la siguiente expresión elemental:

(2)

$$\Phi = \delta \pm z$$

Expresión donde, según la figura 4 (Ref. [11]), el signo positivo es para culminación sur y el negativo para el caso de culminación norte.

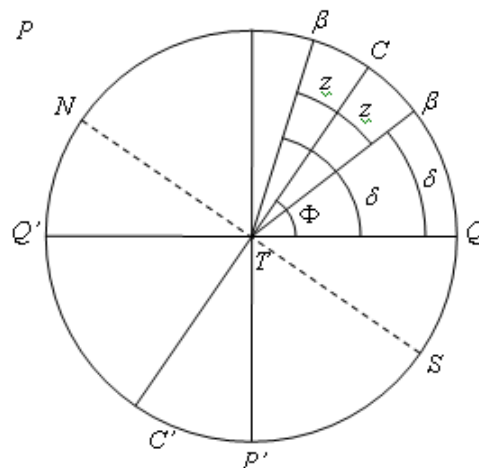


Figura 4.

El enlace entre el sistema horizontal y el sistema ecuatorial se realiza a través del denominado *ángulo horario t*. Este ángulo se define en cada instante para un punto β sobre la esfera celeste tomando como referencia un lugar de observación C , y se entiende como el diedro formado por los meridianos celestes de β y C . Se mide pues en el ecuador y en sentido horario desde el meridiano del lugar. Según se observa en la figura 5 (Ref. [11]), la suma del ángulo horario y la ascensión recta es, en todo instante y para cualquier astro, igual a la hora sidérea local del meridiano de C :

(3)

$$HSL = \alpha + t$$

Como consecuencia de esta relación el valor de la hora sidérea local permite obtener los valores (t, δ) correspondientes de un astro para un instante y lugar, a partir de sus coordenadas ecuatoriales. Esta pareja de variables definen el *sistema de coordenadas horarias*.

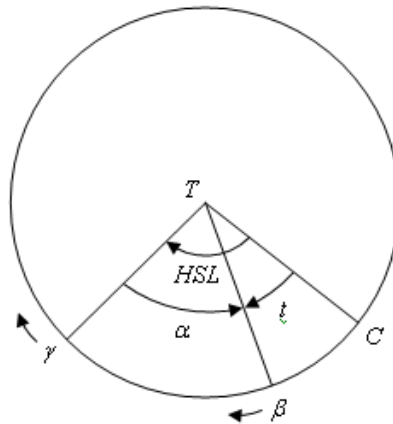


Figura 5.

La relación existente entre las coordenadas ecuatoriales y las horarias es inmediata. Asimismo, tampoco es compleja la ligadura entre coordenadas horarias y horizontales. Ésta se establece mediante el *triángulo polar*, el cual queda definido para un astro β y lugar C por los arcos de círculo máximo que sobre la esfera celeste unen el astro, el cenit del lugar y el polo celeste en cada instante. Como se aprecia en la figura 6 (Ref. [11]), los lados de este triángulo son: $P\beta$ el complemento de la declinación del astro, PC la colatitud astronómica del lugar, y $C\beta$ la distancia cenital. Por su parte los ángulos son los siguientes: P el ángulo horario, C un valor igual a 360° menos el acimut norte, y β una magnitud denominada *ángulo paraláctico* del astro. Evidentemente, la resolución del triángulo polar se simplifica en ciertos casos particulares. Entre estos casos, los principales son: (1) que el ángulo en C sea recto, lo que tiene lugar cuando el astro se encuentra en el primer vertical del lugar; (2) que el lado $C\beta$ sea recto, en cuyo caso el astro se halla sobre el horizonte en el correspondiente orto u ocaso; y (3) que el ángulo paraláctico sea recto, lo que implica que, de tratarse de una estrella, ésta se encuentra en su *máxima digresión* observada desde el punto de estación. No obstante, en el caso general basta conocer tres de los elementos del triángulo polar para determinarlo completamente. De este modo, y con vistas a la transformación en ambos sentidos entre los sistemas de coordenadas horizontales y ecuatoriales, será necesario disponer, aparte de las coordenadas del astro en uno de los sistemas, de la hora sidérea local y del ángulo de latitud del lugar; o sea, un total de cuatro datos debido al movimiento continuo del sistema horizontal.

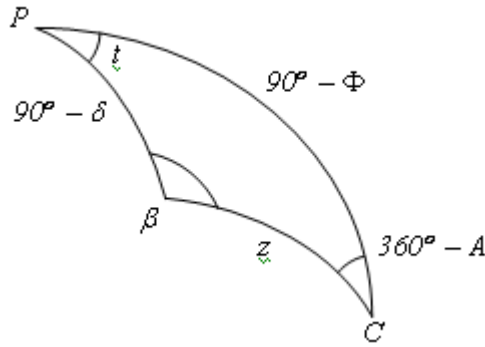


Figura 6.

La transformación descrita puede realizarse en forma vectorial si se definen previamente los correspondientes triedros directos con origen en el centro de la Tierra O . Según se representa en la figura 7 (Ref. [11]), en el sistema ecuatorial (EC) dicho triedro se establece situando el eje OX en la dirección del punto vernal y el eje OZ en la del polo celeste. En el sistema horario (HR), el eje OZ es coincidente con el anterior, trazándose el eje OX en el meridiano del lugar correspondiente. En ambos casos los ejes OY quedan subordinados al carácter directo de cada triedro. Con esta disposición, y teniendo en cuenta los operadores de giro:

(4)

$$G_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \text{sen} \alpha \\ 0 & -\text{sen} \alpha & \cos \alpha \end{pmatrix}, G_y(\alpha) = \begin{pmatrix} \cos \alpha & 0 & -\text{sen} \alpha \\ 0 & 1 & 0 \\ \text{sen} \alpha & 0 & \cos \alpha \end{pmatrix}, G_z(\alpha) = \begin{pmatrix} \cos \alpha & \text{sen} \alpha & 0 \\ -\text{sen} \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

donde se toma positivo el sentido antihorario de la rotación visto desde el extremo positivo del eje de giro, el paso del sistema ecuatorial al horario responde a la siguiente expresión matricial:

(5)

$$V_{EC} = (\cos \delta \cos \alpha, \cos \delta \text{sen} \alpha, \text{sen} \delta) \leftrightarrow V_{EC} (\cos \delta \cos t, -\cos \delta \text{sen} t, \text{sen} \delta)$$

$$V_{HR} = G_z(HSL)V_{EC}$$

Es decir, basta un giro positivo del ángulo HSL respecto al eje común OZ , para superponer en un instante dado el triedro ecuatorial (fijo) sobre el horario (móvil).

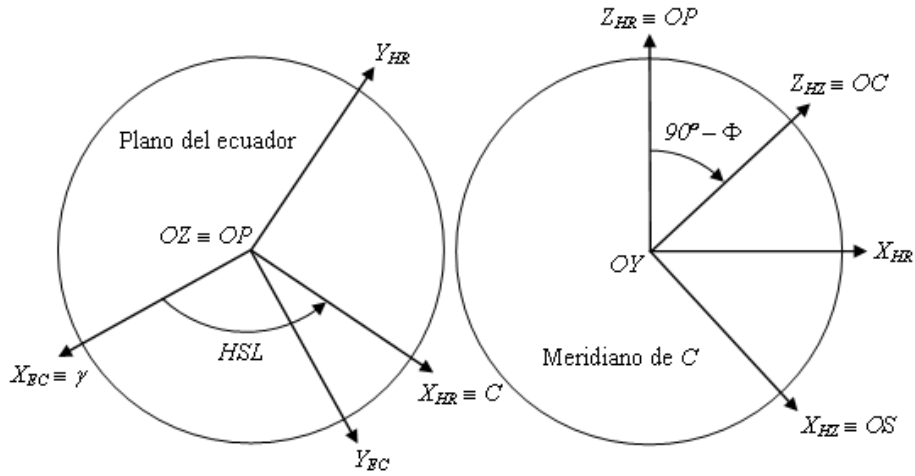


Figura 7.

Situándonos ahora en el plano meridiano del lugar C , definimos el triedro horizontal (HZ) mediante el eje OZ dirigido al cenit y el eje OX en la dirección Sur. De este modo el eje OY queda superpuesto en dirección y sentido con el correspondiente del sistema horario. Con ello el paso del sistema horario al horizontal resulta determinado por la expresión matricial:

(6)

$$\mathbf{V}_{HR} = (\cos \delta \cos t, \cos \delta \sin t, \sin \delta) \leftrightarrow \mathbf{V}_{HZ} (-\sin t \cos A, \sin t \sin A, \cos z)$$

$$\mathbf{V}_{HZ} = G_y(90^\circ - \Phi) \mathbf{V}_{HR}$$

De las dos relaciones anteriores resultan estas otras que ligan directamente el sistema ecuatorial con el de coordenadas horizontales de un lugar para un instante concreto:

(7)

$$\mathbf{V}_{HZ} = G_y(90^\circ - \Phi) G_z(HSL) \mathbf{V}_{EC} \Leftrightarrow \mathbf{V}_{EC} = G_z(-HSL) G_y(\Phi - 90^\circ) \mathbf{V}_{HZ}$$

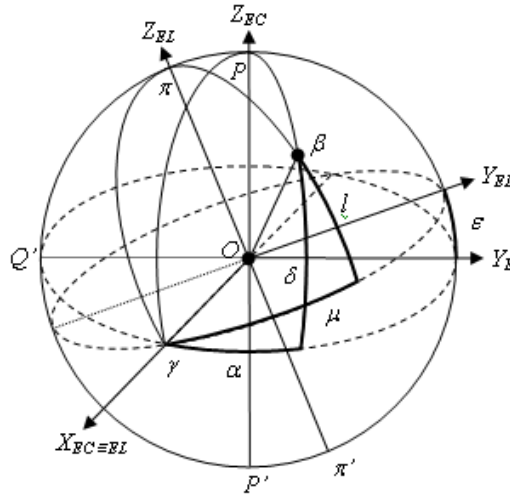


Figura 8.

La transformación entre coordenadas ecuatoriales y eclípticas se encuentra recordando que el plano del ecuador de la Tierra y la eclíptica se cortan en la línea de los equinoccios; y que en esta línea, situado sobre la esfera celeste, se encuentra el punto vernal (γ). Por lo tanto estos dos planos forman un ángulo diedro (ε), que se conoce como *inclinación u oblicuidad de la eclíptica*, y cuyo valor aproximado es de $23^\circ 26.5'$. Atendiendo al plano y polos de la eclíptica, se define el triedro directo (EL) correspondiente a las coordenadas eclípticas (μ, l), de manera que el eje OX se dirige al punto vernal y el eje OZ hacia el polo π tal y como se representa en la figura 8 (Ref. [11]). Con dicha referencia, al considerar a la vez el sistema ecuatorial (EC), resulta que los ejes OX de ambos triedros son coincidentes y, por tanto, también los planos OYZ . Como consecuencia, para obtener en coordenadas eclípticas los cosenos directores de la posición de un astro a partir de los respectivos ecuatoriales, bastará simplemente con realizar un giro positivo del ángulo de inclinación ε con respecto al eje común OX . La ecuación matricial de la transformación en ambos sentidos es pues la siguiente:

(8)

$$\mathbf{V}_{EC} = (\cos \delta \cos \alpha, \cos \delta \sin \alpha, \sin \delta) \leftrightarrow \mathbf{V}_{EL} (-\cos l / \cos \mu, \cos l / \sin \mu, \sin l)$$

$$\mathbf{V}_{EL} = G_x(\varepsilon) \mathbf{V}_{EC} \Leftrightarrow \mathbf{V}_{EC} = G_x(-\varepsilon) \mathbf{V}_{EL}$$

7.3. Movimiento aparente del Sol

7.3.1. Conceptos elementales

Considerado como una estrella más, la posición del centro del Sol en una época cualquiera puede definirse mediante sus coordenadas ecuatoriales (α , δ). Sin embargo, la proximidad del astro a la Tierra motiva que el movimiento de traslación de nuestro planeta a su alrededor haya de tenerse muy en cuenta a la hora de establecer las correspondientes posiciones relativas entre los dos cuerpos. Como ya se dijo, el movimiento de traslación terrestre se produce sobre la eclíptica; dibujando entonces este plano junto al ecuador, ambos con sus polos respectivos, si prescindimos de la rotación diurna y consideramos el movimiento aparente que visto desde la Tierra afecta al Sol a lo largo del año, resulta que, proyectado sobre la esfera celeste, el Sol (S) describe una vuelta completa a la Tierra en dicho tiempo. Ello provoca que las coordenadas ecuatoriales del astro no sean constantes, sino que varíen continuamente correspondiéndose en cada época con la posición de la Tierra en su órbita de traslación. De este modo, y según se representa en la figura 9 (Ref. [11]), la ascensión recta del Sol se incrementa cada año en 24^h y la declinación oscila en el intervalo $[-\varepsilon, +\varepsilon]$. En este movimiento relativo las posiciones del Sol definidas por las coordenadas $(0^h, 0^\circ)$ y $(12^h, 0^\circ)$ constituyen respectivamente los *equinoccios* de primavera (γ) y otoño (γ') para el hemisferio norte, y las posiciones $(6^h, +\varepsilon)$ y $(18^h, -\varepsilon)$ definen los *solsticios* de verano (s) e invierno (s') en el citado hemisferio. Asimismo, si atendemos a las coordenadas eclípticas del Sol, vemos que el ángulo de latitud eclíptica del mismo es constantemente igual a cero, y también que los equinoccios tienen lugar para los valores de longitud eclíptica (μ) 0° y 180° , en tanto que los solsticios ocurren cuando esta coordenada vale 90° ó 270° . Estas cuatro posiciones del Sol también suelen identificarse dos a dos diciendo que el Sol ocupa respectivamente el *coluro de los equinoccios*, o bien el *coluro de los solsticios*, entendiéndose por *coluro* todo plano que contiene a los polos π y π' de la eclíptica (*).

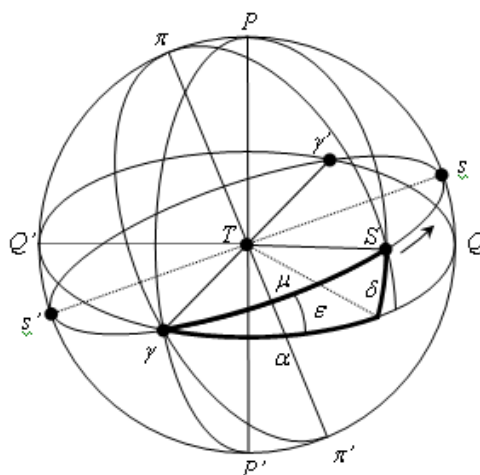


Figura 9.

(*) Para un astro cualquiera el coluro que pasa por el mismo es pues el plano que determina su longitud eclíptica y donde se mide su correspondiente ángulo de latitud eclíptica.

En los párrafos anteriores se ha supuesto que las coordenadas del Sol son constantes a lo largo de un mismo día. Sin embargo, en rigor esto no es así puesto que los incrementos en ascensión recta y declinación no se producen a saltos, sino de forma continua. Por consiguiente, sólo de modo grosero puede afirmarse que la trayectoria diurna aparente del Sol sea un paralelo celeste. En realidad, dicha trayectoria es helicoidal, por lo que existirá un desplazamiento relativo entre Sol y el punto vernal o cualquier estrella fija, en sus correspondientes rutas de movimiento diurno. Para estudiar este desplazamiento observemos que el incremento continuo en la ascensión recta del Sol supone un avance cuyo sentido es contrario al del movimiento aparente causado por la rotación terrestre. Resulta pues que el movimiento angular diurno relativo del Sol es más lento que el del punto vernal o cualquier estrella fija. Esto motiva que se produzcan antes dos pasos consecutivos del punto vernal por un mismo meridiano, que dos pasos consecutivos del Sol; o lo que es igual, que el giro de rotación de la Tierra necesario para observar el Sol dos veces seguidas en el mismo meridiano, sea mayor de 360° . En términos de tiempo, este retraso del Sol sobre los astros y referencias lejanas se conoce como *aceleración de las fijas* y, según veremos después, su valor medio es de casi 4 minutos diarios.

Por otra parte, ha de tenerse en cuenta que la variación anual de la ascensión recta del Sol, en definitiva se debe a la variación correspondiente de la longitud eclíptica, y esta variación, si bien es continua, no se produce de forma uniforme puesto que en primera aproximación el movimiento de traslación terrestre verifica las leyes de Kepler. Según éstas, el planeta describe una órbita elíptica en uno de cuyos focos se encuentra el Sol (*primera ley*), y lo hace a una velocidad angular tal que barre áreas iguales en tiempos iguales (*segunda ley o ley de las áreas*). Esto implica que la máxima velocidad angular se produce en la zona del punto de la órbita más próximo al Sol (*perihelio*), y la mínima en la zona correspondiente al punto más lejano (*afelio*). Dichos puntos se denominan *ápsides* de la órbita y son precisamente los extremos del eje mayor de la elipse de traslación de la Tierra. Se tiene pues que también la velocidad lineal de traslación de nuestro planeta, cifrada en unos $100\,000\text{ km/h}$, es variable según la época y día del año, con lo cual no existe uniformidad en la variación del incremento de ascensión recta del Sol en cada época. Como consecuencia, en el movimiento diurno aparente la ralentización del Sol respecto del punto vernal y las estrellas fijas adquiere valores distintos en función del día del año, por lo que si medimos con rigor el tiempo transcurrido entre dos pasos consecutivos del Sol por un mismo meridiano, comprobamos que este intervalo varía ligeramente de unos días a otros, yendo en aumento o disminución según la época del año en que se efectúan las observaciones.

7.3.2. Teoría de las anomalías

Para determinar la posición del Sol en su órbita aparente, durante muchos años se ha utilizado la denominada teoría de las *anomalías* (figura 10. Ref. [11]). En ella se supone que la Tierra (T) está situada en uno de los focos de la elipse, y que es el Sol (S) el cuerpo que orbita a su alrededor haciéndolo en correspondencia con el movimiento real del planeta. En este marco, los puntos donde la distancia del Sol a la Tierra es mínima y máxima se denominan respectivamente *perigeo* (P) y *apogeo* (A), y el ángulo polar que respecto del eje mayor fija la posición del Sol en cada instante se conoce como *anomalía verdadera* (θ). Este ángulo es la variable que figura en la ecuación polar de la elipse, donde r es la distancia radial, a el semieje mayor, y e la primera excentricidad:

(9)

$$r = \frac{a(1 - e^2)}{1 + e \cos \theta}$$

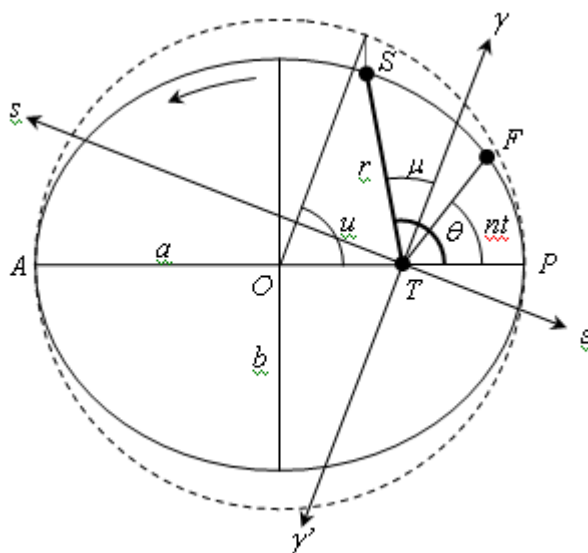


Figura 10.

La anomalía verdadera está ligada en todo instante con el ángulo u que en el centro de la elipse forman el eje polar y el radio correspondiente a la proyección del Sol sobre la circunferencia directriz mayor de la curva. Dicho ángulo se denomina *anomalía excéntrica*, y su relación con el ángulo polar se establece mediante la siguiente igualdad cuya demostración figura en Mena (2008):

(10)

$$\operatorname{tg} \frac{\theta}{2} = \sqrt{\frac{1+e}{1-e}} \operatorname{tg} \frac{u}{2}$$

Por otra parte, la posición instantánea del Sol verdadero S en su órbita puede relacionarse con la determinada por un sol ficticio F que recorre la elipse con una velocidad angular constante n , pero de forma que coincide con el verdadero en los ápsides. Esta relación se establece a través de la anomalía excéntrica mediante el ángulo polar de F llamado *anomalía media* (nt); se trata de la *ecuación de Kepler*, y su demostración también puede consultarse en Mena (2008):

(11)

$$nt = u - e \operatorname{sen} u$$

Evidentemente, conocido el valor de una anomalía cualquiera, las fórmulas anteriores permiten determinar la distancia entre el Sol y la Tierra en cada instante. Sin embargo, más frecuente que el empleo de anomalías es el uso de las longitudes eclípticas, lo que tiene como ventaja la posibilidad de situar la elipse sobre el plano de la eclíptica en su posición correcta. Como sabemos, las longitudes eclípticas se miden desde la dirección del punto vernal en sentido antihorario; pero como ni la línea de los equinoccios ni la de los solsticios coinciden con la línea de ápsides de la elipse, para relacionar la anomalía verdadera con la longitud eclíptica del Sol se hace intervenir como parámetro la longitud μ_P del perigeo de la órbita. De esta manera las dos variables de posición del Sol resultan ligadas por la ecuación:

(12)

$$\theta = \mu - \mu_P$$

Y análogamente, si denominamos μ_m a la longitud eclíptica del sol ficticio F :

(13)

$$nt = \mu_m - \mu_P$$

La diferencia (12) permite escribir la ecuación polar en la forma tradicional:

(14)

$$r = \frac{a(1-e^2)}{1+e \cos(\mu - \mu_P)}$$

Esta expresión proporciona la distancia del Sol a la Tierra en función de la longitud eclíptica de aquél, o bien la citada longitud en función de la distancia. Sin embargo, previamente será necesario determinar los valores de las constantes para que la órbita de traslación anua esté por completo definida en forma y tamaño (a, e) y en su posición sobre la eclíptica (μ_p). Sobre la forma de evaluar estos parámetros se remite al lector a la referencia SGE (1972).

7. 3.3. El sol medio y la ecuación del tiempo

Siendo que la longitud del perigeo es una magnitud que puede considerarse constante en primera aproximación, la igualdad (13) muestra que la longitud eclíptica de F , al contrario de lo que sucede con la del Sol, varía de manera uniforme a lo largo del año. Pues bien, si restamos miembro a miembro las dos ecuaciones (12) y (13), se obtiene una función que expresa la diferencia existente en cada instante entre las posiciones del Sol verdadero y del sol ficticio, tanto en términos de anomalías como de longitudes eclípticas. Dicha función, que se conoce como *ecuación del centro*, se define así:

(15)

$$E_c = \theta - nt = \mu - \mu_m$$

Seguidamente introduciremos una nueva función cuyo objeto es relacionar entre sí la longitud eclíptica del Sol con su ascensión recta en cada instante. Esta función se denomina *reducción al ecuador* y está determinada por la diferencia:

(16)

$$R_e = \mu - \alpha$$

Su valor puede ser calculado directamente a partir de la longitud del Sol y del ángulo de inclinación de la eclíptica mediante la expresión siguiente (SGE, 1972):

(17)

$$R_e = \arctan \left(\frac{\tan^2(\varepsilon/2) \sin 2\mu}{1 - \tan^2(\varepsilon/2) \cos 2\mu} \right)$$

La correspondencia entre el Sol verdadero y ficticio se prolonga al objeto de establecerla sobre el ecuador, es decir, entre ascensiones rectas. Para ello, el sol ficticio se sustituye por otro que recorre el ecuador con movimiento uniforme, de forma que su ascensión recta es en todo momento igual que la longitud eclíptica de F . Este sol se conoce como *sol medio* (M) y su movimiento se define así:

(18)

$$\alpha_m = \mu_m$$

De este modo, la relación entre las posiciones del Sol verdadero y medio quedará determinada en todo instante por la diferencia entre sus correspondientes ascensiones rectas. Este parámetro se llama *ecuación del tiempo* y resulta precisamente como la diferencia entre la ecuación del centro y la reducción al ecuador:

(19)

$$E_t = E_c - R_e = \alpha - \alpha_m$$

El sol medio constituye la referencia básica de tiempo por la que regimos nuestra vida diaria (*tiempo medio*). Dicha referencia está pues estrechamente ligada al movimiento del Sol pero con la ventaja de girar de modo uniforme, prescindiendo así de las oscilaciones en la velocidad aparente del Sol características del movimiento kepleriano. Por consiguiente, la ecuación del tiempo puede definirse como la cantidad numérica que debe añadirse a la posición de la referencia de tiempo verdadero, observable físicamente, para obtener la posición correspondiente de la referencia de tiempo uniforme mediante el cual desarrollamos nuestra actividad cotidiana. Esta magnitud oscila a lo largo del año entre +14 y -17 minutos de tiempo (Mena, 2008).

7. 3.4. Período de revolución anua

A lo largo de los párrafos anteriores nos hemos referido con frecuencia al término “año”. Sin embargo este concepto tiene distintas e importantes acepciones según las referencias que se tomen para su determinación. Entre las diversas posibilidades, la duración del período de la traslación terrestre alrededor del Sol puede evaluarse principalmente de tres formas distintas. La primera consiste en atender al denominado *año sidéreo*, el cual se define como el tiempo necesario para que la Tierra efectúe una revolución completa de 360° en su órbita. Tomando como unidad de medida el día medio, del que hablaremos después, el período del año en cuestión resulta en primera aproximación considerando la *tercera ley de Kepler*; según ésta, el cuadrado del período orbital es proporcional al cubo del semieje mayor de la elipse de traslación, interviniendo en la razón de proporcionalidad la constante G de atracción universal, la masa M del Sol y la masa m de la Tierra. No obstante, si el valor inicialmente obtenido se corrige por otras causas adicionales que influyen en el equilibrio del planeta en el sistema solar, la duración del año sidéreo resulta ser aproximadamente la siguiente:

(20)

$$T_s^2 = \frac{4\pi^2}{G(M+m)} a^3 \quad (\text{factores de corrección}) \Rightarrow T_s = 365,25636^{dm}$$

El segundo término importante es el *año trópico o año astronómico*, que constituye la base del calendario. Si consideramos el movimiento aparente del Sol, como tal se entiende el tiempo transcurrido entre dos pasos consecutivos del Sol por el punto vernal; o lo que es igual, el tiempo necesario para que el Sol incremente su longitud eclíptica en 360°. Esta definición difiere de la del año sidéreo en que ahora se tiene en cuenta el retroceso medio que anualmente experimenta el punto vernal debido al movimiento de precesión de la Tierra, con lo cual el encuentro del Sol con el equinoccio de primavera se produce un poco antes de completarse una revolución exacta. Como se verá después, dicho retroceso tiene un valor aproximado de unos 50.2'' por año; cantidad que varía ligeramente de unos años a otros. En consecuencia, un valor grosero de la duración del año trópico puede obtenerse a partir del período sidéreo mediante la proporción siguiente, si bien el resultado que se presenta está apoyado en observaciones y por tanto es más preciso:

(21)

$$\frac{T_s}{360^\circ} = \frac{T_t}{360^\circ - 50,2''} \quad (\text{factores de corrección}) \Rightarrow T_t = 365,24219^{dm}$$

El tercer término usado tradicionalmente en la definición del año es el llamado *año anomalístico*. En este caso se evalúa el tiempo necesario para que el Sol, en su movimiento aparente, efectúe dos pasos consecutivos por el perigeo de la órbita. En este caso, y puesto que la línea de ápices tiende a juntarse con la de los equinoccios a razón de unos 11.7'' por año, la duración del año anomalístico es algo mayor que la del año sidéreo; o sea, que el período anomalístico requiere que el Sol sobrepase en una pequeña cantidad la revolución exacta. Para su evaluación grosera podemos emplear una proporción semejante a la anterior, la cual se incluye seguidamente junto al valor más preciso que se ofrece como resultado:

(22)

$$\frac{T_s}{360^\circ} = \frac{T_a}{360^\circ + 11,7''} \quad (\text{factores de corrección}) \Rightarrow T_a = 365,25964^{dm}$$

7. 4. El tiempo

7. 4.1. Tiempo rotacional

En su concepción más simple, por medida del tiempo se entiende la determinación del ángulo recorrido por una determinada referencia en movimiento circular uniforme a partir de una dirección radial origen.

* El más elemental de los distintos tipos de tiempo rotacional es el *Tiempo Sidéreo* (TS). En esta concepción se toma como referencia el punto vernal (o cualquier estrella fija) en su movimiento aparente sobre el ecuador (o sobre su paralelo celeste) con respecto a un meridiano cualquiera. Decimos entonces que un *día sidéreo* es el tiempo transcurrido entre dos pasos consecutivos del punto vernal por un mismo meridiano, considerado éste como una de las dos partes en que queda dividido por la línea de los polos. De este modo, el día sidéreo, que se corresponde evidentemente con una rotación completa de la Tierra, se fracciona en 24 *horas sidéreas*, 24 x 60 *minutos sidéreos*, 24 x 3600 *segundos sidéreos*, etc. Resulta entonces que, para cada lugar concreto del planeta, la *Hora Sidérea Local* correspondiente a cada instante está determinada por el ángulo HSL ya definido en el apartado 1.

* Otro tipo de tiempo rotacional es el *Tiempo solar Medio* (TM). Ahora la referencia de movimiento es el *sol medio* (*M*) introducido en la teoría de las anomalías. Recordemos que esta referencia es un sol ficticio que se mueve uniformemente sobre el ecuador, pero en correspondencia con el movimiento aparente del sol verdadero sobre la eclíptica. Se pretende así dar uniformidad a la medición del tiempo, y a la vez mantener un ajuste con el Sol que luce y por el cual regimos nuestras actividades cotidianas. Mediante este tiempo se definen, de forma análoga al caso del tiempo sidéreo, los conceptos correspondientes de *día medio* y *hora media* (HM) para un meridiano cualquiera (*C*). No obstante, como el comienzo del día medio en un lugar se acuerda establecerlo en el instante en que el sol medio cruza el antimeridiano correspondiente, la hora media se define con 12^h angulares de diferencia, es decir, a partir del meridiano inferior del lugar. Con esta disposición, en el caso particular del meridiano de Greenwich (*G*) el tiempo medio se denomina *Tiempo Universal* (UT), y la hora, única para todos los puntos de la Tierra, *Hora de Tiempo Universal* (HUT). Evidentemente, al considerar negativas las longitudes Oeste y positivas las Este, la hora media de un meridiano cualquiera cuya longitud es conocida, se obtiene a partir de la hora universal mediante la expresión (figura 11. Ref. [11]):

(23)

$$HM = HUT + \Lambda$$

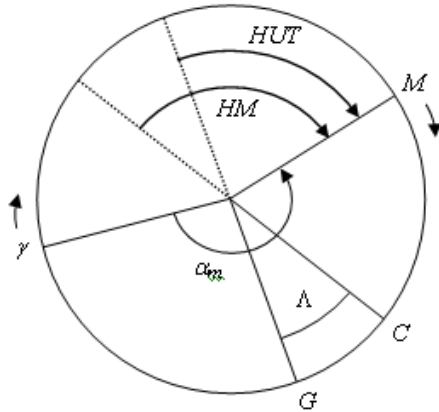


Figura 11.

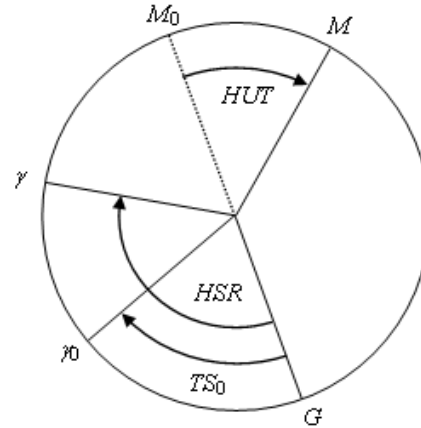


Figura 12.

Para obtener la relación de conversión entre intervalos de tiempo medio y sidéreo basta con tener en cuenta que la duración del año trópico es aproximadamente igual a 365.2422 días medios. Fijando entonces un meridiano cualquiera como origen y suponiendo inicialmente que el punto vernal y el sol medio son coincidentes en él, al cabo de un año, cuando ambas referencias vuelvan a coincidir, el sol habrá dado 365.2422 vueltas con respecto a este meridiano, mientras que el punto vernal habrá rotado una vez más debido al incremento continuo de la ascensión recta del sol (aceleración de las fijas), o sea, 366.2422 veces. Quiere esto decir que la duración de 365.2422 días medios equivale a la de 366.2422 días sidéreos, con lo cual la transformación lineal entre un intervalo medio (IM) y el correspondiente intervalo sidéreo (IS) será:

(24)

$$\frac{IM}{IS} = \frac{365.2422}{366.2422}$$

En cuanto a la relación entre la hora de tiempo universal (hora media del meridiano de Greenwich) y la correspondiente hora de tiempo sidéreo (figura 12. Ref. [11]), ésta se establece a partir del parámetro denominado *Tiempo Sidéreo a 0 horas UT* (TS_0), el cual determina la hora sidérea local en Greenwich en el instante en que el sol medio se encuentra en el meridiano inferior de este lugar.

Dicho valor puede calcularse en base a la *fecha juliana (FJ)* del día en cuestión; fecha que se corresponde con el número de días transcurridos desde un instante origen. Por lo tanto, conocida la magnitud TS_0 para un día medio, la *Hora Sidérea Greenwich (HSG / HSR)* a cualquier hora de tiempo universal (HUT) resulta haciendo uso de la expresión (24):

(25)

$$HSG = HSR = TS_0 + \frac{366.2422}{365.2422} HUT$$

A partir de aquí puede calcularse la hora sidérea local que en el mismo instante corresponde a un meridiano de longitud dada:

(26)

$$HSL = HSR + \Lambda$$

Las dos últimas expresiones permiten obtener la hora universal a la cual un astro β cuya ascensión recta α es conocida, cruza el meridiano de un lugar de longitud Λ . Para ello basta tener en cuenta que en el instante en el cual β culmina en el meridiano en cuestión, la hora sidérea local será igual a α (ángulo horario cero) y, consecuentemente, la hora sidérea reducida tiene por valor $\alpha - \Lambda$. Despejando entonces de (25) la incógnita HUT, a la que denominamos aquí *Hora de Paso Meridiano (HPM)*, encontramos enseguida la siguiente igualdad:

(27)

$$\left. \begin{array}{l} HSL = \alpha \\ HSR = \alpha - \Lambda \end{array} \right\} \Rightarrow HPM_{\beta} = (\alpha - \Lambda - TS_0) \frac{365.2422}{366.2422}$$

* Un tercer tipo de tiempo rotacional es el *Tiempo solar Verdadero (TV)*, es decir, el regido por el movimiento aparente del Sol real en su recorrido por la eclíptica unido al movimiento diurno de rotación terrestre. El *día verdadero* es pues el tiempo transcurrido entre dos pasos consecutivos del Sol por un mismo meridiano; definiéndose la *hora verdadera (HV)* de un lugar como el ángulo que forma el meridiano inferior del sitio en cuestión con el meridiano que contiene al Sol en un instante cualquiera.

Para establecer la correspondencia entre el tiempo medio y el verdadero se atiende a la siguiente expresión (Mena, 2008):

$$24^{hv} = 24^{hm} + \Delta E_t$$

Donde el segundo término del segundo miembro es la diferencia entre el valor de la ecuación del tiempo medido a las 0 horas UT del día $D + 1$ menos el observado previamente a las 0 horas UT del día D . Por consiguiente, la razón mediante la cual se relacionan los intervalos de tiempo verdadero (IV) y medio (IM) para una fecha concreta resulta ser:

(28)

$$\frac{IV}{IM} = \frac{24}{24 + \Delta E_t}$$

Por otra parte, la relación entre la *Hora Verdadera Greenwich* (HVG) y la hora de tiempo universal para cualquier instante se define mediante el valor de la ecuación del tiempo, considerado con su signo, en dicho instante:

(29)

$$HVG = HUT - E_t$$

O también:

(30)

$$HVG + \alpha = HSR + 12^h$$

Desde luego, y sea cual sea la forma de calcular HVG, la hora verdadera (HV) de un meridiano de longitud conocida está determinada para el mismo instante por la siguiente expresión:

(31)

$$HV = HVG + \Lambda$$

7. 4.2. Tiempo de efemérides

Según hemos visto, el tiempo rotacional medio, y en particular el tiempo universal, es teóricamente uniforme. Sin embargo, el movimiento del polo, así como las oscilaciones periódicas debidas a la precesión y nutación de la Tierra, generan imprecisiones en la posición de los distintos meridianos, incluyendo la del meridiano origen. Tales imprecisiones hacen que el tiempo universal, medido en la forma descrita más arriba, no sea todo lo uniforme que exigen determinadas aplicaciones y usos. De hecho, la falta de uniformidad del tiempo universal, y en definitiva, las irregularidades en el movimiento de rotación de la Tierra, se pusieron de manifiesto al observar que las ecuaciones de movimiento de la Mecánica Celeste no respondían con toda precisión a las observaciones cuando como variable tiempo se introducía el tiempo medio. Como consecuencia de ello, a finales del siglo XIX y principios del XX se pensó en tomar como referencia de tiempo para confeccionar las efemérides de los astros, no ya la rotación diaria del sol medio, sino la traslación anua del mismo, lo que dio lugar al denominado *Tiempo de Efemérides*

(TE). Esta escala de tiempo, al estar pues basada en el movimiento de traslación de la Tierra, elimina las irregularidades de la rotación; y como además la referencia es el sol medio en su movimiento anual prácticamente uniforme sobre el ecuador, resulta que el tiempo de efemérides es mucho más estable que el universal, permitiendo disponer de unas efemérides de los astros consonantes con la realidad de las mediciones.

El origen del tiempo de efemérides resultó establecido al comienzo del año 1900 por Simon Newcomb apoyándose en el estudio del movimiento del sol ficticio F (figura 11) cuya longitud eclíptica es constantemente igual a la ascensión recta del sol medio. Para ello Newcomb determina previamente el denominado *año de efemérides*, el cual define como la duración del año trópico de 1900, y a partir de aquí establece la definición del *segundo de efemérides* en la forma siguiente (Mena, 2008):

(32)

$$1 \text{ segundo de efemérides} = \frac{\text{Año trópico de 1900}}{365,2421988 \cdot 86400} = \frac{\text{Año trópico de 1900}}{31556925,975}$$

Incluso contando con la estabilidad del TE, el segundo de efemérides no es adoptado por la comunidad internacional como unidad fundamental hasta 1956, fecha en que el tiempo puede medirse ya mediante osciladores de cuarzo. A partir de entonces comienzan a desarrollarse los anuarios con argumento t en tiempo de efemérides, hasta que algunos años más tarde, a raíz de la aparición del tiempo atómico, la escala de efemérides es sustituida por otras más precisas que trataremos enseguida. No obstante, esta sustitución es más una prolongación que un cambio radical, puesto que la mencionada estabilidad del tiempo de efemérides permitió que muchos de los cálculos relativos a los procesos de carácter dinámico realizados entonces, fuesen únicamente actualizados con las posteriores escalas de tiempo.

Cuando se considera la vida cotidiana la escala de tiempo de efemérides ya no resulta práctica, pues al ser independiente de la rotación terrestre los días de efemérides se van desfasando progresivamente de los naturales. Esta es la razón por la cual el tiempo de efemérides no ha sustituido al tiempo universal, manteniéndose ambos para distintas aplicaciones específicas. Además, dado que la conversión entre UT y TE siempre es posible mediante observación, el tiempo de efemérides sirvió también para medir con precisión las faltas de uniformidad de la rotación de la Tierra, es decir, para estudiar con detalle la variabilidad del tiempo universal.

La conversión entre el tiempo de efemérides y el universal exigió realizar observaciones experimentales con objeto de determinar el incremento que en cada momento separaba ambas escalas según la diferencia:

(33)

$$TE = UT + \Delta T \quad \Leftrightarrow \quad \Delta T = TE - UT$$

Este incremento ΔT representa la falta de estabilidad de la rotación terrestre con respecto al tiempo de efemérides, Su valor en el año 1900 era de -3.79 segundos mientras que a primeros del año 2006, rondaba ya los 65. Por las razones que veremos después conviene tener presente que el 1 de enero de 1958, fecha de origen de la escala de Tiempo Atómico Internacional (TAI), el parámetro ΔT tenía un valor de 32.184 segundos. Por otra parte, es importante considerar que la precisión con la cual se determina la diferencia $TE - UT$, está influenciada directamente por la aproximación que el tiempo universal mantiene en cada momento con la verdadera posición de la Tierra en su movimiento de rotación. Para especificar dicha aproximación, tradicionalmente se han utilizado las escalas UT1 y UT2, las cuales ofrecen el tiempo universal pero considerando un número distinto de factores en la determinación del mismo.

Lo dicho en los párrafos anteriores muestra la progresiva separación del tiempo universal con la recta estable determinada por el tiempo de efemérides; pero como hemos visto, este desfase tan solo es de algo más de un minuto después de más de 100 años. Por ello, cuando se trata de convertir intervalos de tiempo relativamente pequeños, con frecuencia éstos suelen tomarse iguales; es decir, podemos asumir la siguiente relación aproximada entre un intervalo de tiempo de efemérides (IE) y el intervalo correspondiente de tiempo medio (IM):

(34)

$$IE \cong IM$$

Tal y como se ha definido, podemos decir que el tiempo de efemérides resulta como consecuencia de igualar la posición calculada de un astro a través de su ecuación de movimiento, con la posición obtenida mediante observación. Sin embargo, cuando se sustituye el astro en cuestión por otro cuya ecuación ha sido deducida con la misma precisión del primero, se comprueba que no existe una absoluta coincidencia entre las posiciones observada y calculada si el argumento utilizado es el tiempo de efemérides definido por el astro inicial. Esto implica que *cada astro genera su propio tiempo de efemérides*, por lo que para disponer de un único tiempo completamente uniforme de cara a las aplicaciones dinámicas, sería necesario contar con un exacto y total conocimiento de las constantes de la Astronomía, las cuales, por otra parte y a excepción de la velocidad de la luz, no son rigurosamente constantes. Por este motivo, a partir del descubrimiento de los relojes atómicos se abandonó la idea de apoyar la escala de tiempo estable en el movimiento de los astros, dirigiéndose todas las atenciones hacia las oscilaciones sumamente regulares y uniformes de ciertos átomos. No obstante, la nueva definición de la unidad de se-

gundo mantuvo el mismo valor que había sido utilizado con el tiempo de efemérides, prolongándose la escala TE con una mucho más estable que en un principio se denominó *Tiempo Dinámico Terrestre* (TDT), la cual permitiría dar continuidad al conjunto de observaciones y cálculos realizados años antes con el tiempo de efemérides (Mena, 2008).

7. 4.3. Tiempo atómico

Poco después de finalizar la segunda guerra mundial los científicos comenzaron el desarrollo del primer reloj atómico de la Historia. Inicialmente, y adoptando como unidad de tiempo atómico el segundo de efemérides, el día 1 de enero de 1958, a las 0 horas UT2, se genera la escala de *Tiempo Atómico Internacional* (TAI); escala que se estableció con carácter mundial a partir del promedio de las oscilaciones de numerosos relojes atómicos ya existentes en diversos países. Como en el instante origen del TAI la diferencia TE – UT2 era de 32.184 segundos, la estabilidad de ambas escalas hace que esta cantidad sea el incremento constante que desde entonces permite la conversión de una en otra (figura 13. Ref. [11]):

(35)

$$TE - TAI = 32.184 \text{ seg}$$

A la vez que el TAI se pone en marcha una escala nueva de tiempo universal cuyo objeto es reemplazar a UT2 y que será denominada como *Tiempo Universal Coordinado* (UTC). Estas siglas sustituirán progresivamente a las correspondientes del *Tiempo del Meridiano de Greenwich* (GMT) para designar el tiempo universal. No obstante, como UTC es rigurosamente estable y por tanto independiente de la rotación de la Tierra, la escala se desfaza de forma continua del movimiento diurno, irregular por naturaleza pero base del concepto de tiempo medio que gobierna nuestra vida diaria. Por ello, en estos primeros años se opta por realizar constantes ajustes en la frecuencia de oscilación de los relojes UTC a fin de casar la estabilidad del tiempo atómico con la rotación terrestre.

Llegamos así al año 1967, fecha en que se adopta como referencia para la definición del *segundo internacional* las vibraciones del átomo de cesio. En concreto, un segundo atómico es 9.192.631.770 veces el período de la radiación correspondiente al salto entre los dos niveles de la estructura hiperfina del estado fundamental de los átomos del nucleido Cesio 133.

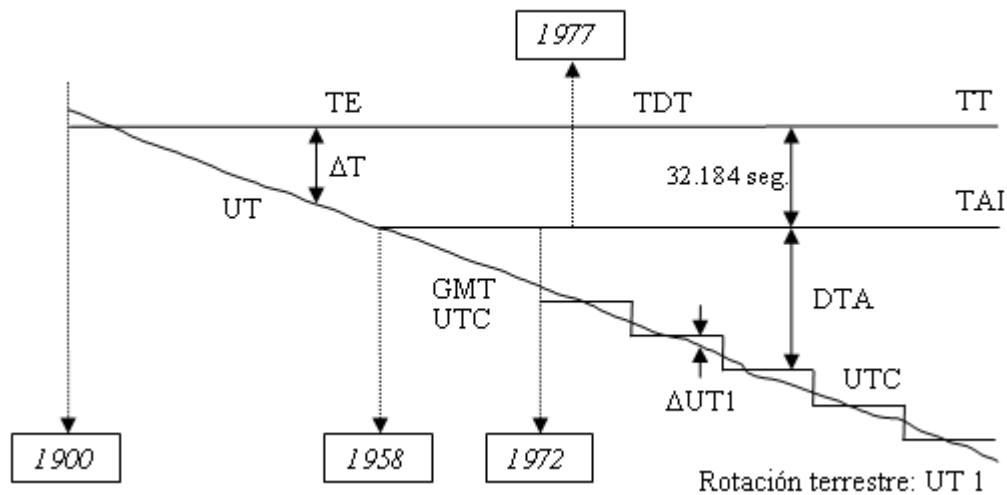


Figura 13.

El método para mantener el carácter de tiempo medio aprovechando a la vez la estabilidad del tiempo atómico, consistió en insertar cada cierto tiempo un salto al UTC cuyo valor es exactamente de un segundo, lo que mantiene dicha escala ajustada en todo instante con la rotación terrestre dentro de este margen. Como consecuencia de ello UTC se representa a modo de una escalera cuyos peldaños son paralelos al TAI permanentemente, dado que ambas escalas de tiempo atienden a la misma unidad patrón. El parámetro que determina el número exacto de segundos que se han ido aplicando al UTC se conoce como *Diferencia con el Tiempo Atómico* (DTA), y expresa la separación que el tiempo universal mantiene con el TAI en la época correspondiente:

(36)

$$DTA = TAI - UTC$$

A fecha de 2007, el último segundo intercalar fue introducido a las $23^h 59^m 59^s$ UTC del 31 de diciembre de 2005, quedando entonces actualizado el DTA al valor +33 segundos.

No obstante, en muchas aplicaciones se requiere un conocimiento más preciso de la posición de un meridiano (variable ligada a la rotación terrestre) que la que puede determinarse con UTC. Para ello existe la escala UT1, la cual representa con exactitud la posición del meridiano de referencia en base a observaciones y otros datos adicionales al UTC. El incremento existente entre ambas escalas se establece en todo instante mediante el parámetro *Diferencia con el Tiempo Universal 1* ($\Delta UT1$, o bien DUT1 cuando se aproxima a la décima de segundo), posibilitando así el acceso a UT1 a partir de UTC:

(37)

$$\boxed{UT1 = UTC + \Delta UT1}$$

Este parámetro, siempre menor de un segundo, es suministrado con toda la precisión que se quiera por los organismos correspondientes. Uno de ellos, sito en España y que también nos proporciona el DTA, es el Real Observatorio de la Armada (San Fernando de Cádiz).

Por su parte, el tiempo de efemérides cede su lugar al tiempo atómico a las 0 horas TAI del 1 de enero de 1977, momento en el cual es sustituido por el *Tiempo Dinámico Terrestre* (TDT). Sin embargo, al objeto de mantener en vigor los múltiples datos existentes referenciados al tiempo de efemérides, la nueva escala se creó como una continuación de aquél. Esta rigurosa escala, diseñada para su uso en los procesos dinámicos no relativistas, posteriormente cambió su denominación por la de *Tiempo Terrestre* (TT), escala que se mantiene en la actualidad junto con otra aún más refinada como es el *Tiempo Coordinada Geocéntrico* (TCG). Vemos así que la mencionada continuidad con el tiempo de efemérides es la razón por la que se escogió precisamente la actual definición del segundo atómico en lugar de cualquier otra, y también el motivo de la diferencia constante:

(38)

$$TT - TAI = 32.184 \text{ seg}$$

Esta expresión permite el acceso al tiempo terrestre desde UTC de acuerdo con la relación:

(39)

$$\boxed{TT = UTC + DTA + 32.184 \text{ seg}}$$

Además de las escalas mencionadas se han generado otras para atender a distintas aplicaciones. Algunas de ellas son específicas de determinados sistemas, y sólo se diferencian del TAI en el incremento constante que las separa; tal es la escala de referencia GPS *Time* que estudiamos en Geodesia espacial. Sin embargo, otras van más allá, profundizando en el análisis de los efectos relativistas debidos a la gravedad y a la velocidad del movimiento del planeta. En este marco se encuentra la escala de *Tiempo Dinámico Baricéntrico* (TDB) y la más reciente de *Tiempo Coordinada Baricéntrico* (TCB), que consisten básicamente en el transporte del tiempo terrestre al baricentro del sistema solar (Mena, 2008).

7. 4.4. Época juliana

Se dice que un determinado instante está definido en cómputo juliano cuando se expresa mediante el número de días y fracción de éstos que han transcurrido desde un cierto momento ori-

gen hasta el instante en cuestión. Dicho origen, contabilizado hacia atrás con el actual calendario y considerando la existencia de un año cero conforme al cómputo astronómico, queda fijado en el día 1 de enero del año – 4712 a las 12 horas de tiempo universal, siendo este momento el correspondiente a la fecha juliana 0. El *día juliano*, por consiguiente, comienza y acaba al mediodía del día medio, y tiene pues una duración de 24 horas. Como unidad de gran interés figura el *siglo juliano*, que ya hemos utilizado, y que abarca un período de 36525 días julianos.

Propuesta precisamente en el año 1582 por el religioso y erudito francés Joseph Justus Scaliger, en honor a su padre fallecido Julius Caesar Scaliger, esta manera de contar el tiempo encuentra su principal aplicación en los procesos astronómicos toda vez que facilita notablemente el operar con intervalos temporales ya sean éstos de gran o pequeña duración. Entre las fechas julianas que conviene tener en cuenta destacamos las siguientes:

- FJ = 0000000.0 01 / 01 / 4712 a.C. 12h UT
- FJ = 1721058.0 01 / 01 / 0000 a.C. 12h UT
- FJ = 1721424.0 01 / 01 / 0001 d.C. 12h UT
- FJ = 2299160.0 04 / 10 / 1582 d.C. 12h UT
- FJ = 2299161.0 15 / 10 / 1582 d.C. 12h UT
- FJ = 2415020.5 01 / 01 / 1900 d.C. 00h UT
- FJ = 2451545.0 01 / 01 / 2000 d.C. 12h UT

En este contexto, el intervalo en días julianos transcurrido entre una fecha de referencia FJ_0 y el instante determinado por la hora universal HUT de un cierto día, que ha sido definido por su fecha juliana FJ a las 0 horas UT, está dado pues por la sencilla igualdad:

(40)

$$t = FJ + \frac{HUT}{24} - FJ_0$$

No obstante, por los años cincuenta se promovió la idea de cambiar el origen de la fecha juliana, situándolo a las 0 horas UT del 17 de noviembre de 1858 con objeto de aminorar las magnitudes numéricas. De este modo nació la denominada *fecha juliana modificada (FJM)*, que se obtiene a partir de la anterior mediante la diferencia: $FJM = FJ - 2400000,5$. Información adicional encontrará el lector en Mena (2008).

7. 5. Correcciones astronómicas

7. 5.1. Factores de corrección y referencia de cálculo

Con objeto de obtener precisión en las observaciones de los astros, es necesario aplicar una serie de correcciones a las coordenadas contenidas en los catálogos, a fin de calcular las posiciones donde realmente serán observados en un instante y lugar determinados. Estas correcciones, cuyas expresiones analíticas se deducen atendiendo a las matrices y sentidos de giro expuestos en (4), son debidas principalmente a las siguientes causas:

- El movimiento propio del astro.
- El movimiento de precesión terrestre.
- El movimiento de nutación terrestre.
- La paralaje anua ocasionada por la traslación de la Tierra.
- El fenómeno de aberración de la luz.
- La refracción atmosférica.

Para la determinación sucesiva de todas estas correcciones se parte de los datos iniciales relacionados en los distintos catálogos, donde figuran para cada estrella sus coordenadas ecuatoriales respecto de las posiciones del ecuador y equinoccio en una época determinada, junto con una serie de parámetros adicionales necesarios para los distintos cálculos. Uno de estos catálogos, ampliamente utilizado, es el FK5 (*Fifth Fundamental Catalogue*), el cual ha sido posteriormente reemplazado por el FK6 a la vez que se ha presentado una nueva teoría sobre los sistemas de referencia y las correcciones astronómicas (Sevilla, 2003; Mena, 2008).

Consideremos pues el catálogo FK5 donde el sistema de referencia es el baricentro O del sistema solar junto con el ecuador y equinoccio medios referidos a las 12 horas de tiempo terrestre del 1 de enero de 2000. La primera tarea consiste en el cálculo del intervalo de tiempo terrestre que separa la época de referencia (época E_0) del instante de observación (época E). Para ello basta con tener en cuenta: la fecha juliana de referencia ($FJR = 2451545.0$), la fecha juliana de la época E (FJ), y la hora de tiempo terrestre (HTT) correspondiente a dicho instante, la cual se obtiene a partir de la hora UTC mediante (39). Con estos datos, y expresado en siglos julianos puesto que así lo requieren todas las fórmulas de corrección que siguen, el intervalo de cálculo resulta mediante la ecuación:

(41)

$$t = \frac{1}{36525} \left(\frac{HTT}{24} + FJ - FJR \right)$$

7. 5.2. Corrección por movimiento propio

La primera corrección que se aplica a las coordenadas ecuatoriales (α_0, δ_0) de una estrella contenida en el FK5, es la correspondiente al pequeño movimiento propio del astro. Considerando la esfera celeste en la época E_0 (figura 14. Ref. [11]), a causa de este movimiento la estrella pasa de su posición inicial β_0 reflejada en el catálogo, a una posición futura β que está relacionada con la anterior por el ángulo Ψ_0 y por una determinada velocidad angular media μ_m función del intervalo de tiempo (41) relativo a la época actual E .

Puesto que el movimiento de la estrella se supone uniformemente acelerado, como velocidad angular media μ_m para todo el recorrido se toma la instantánea en el momento $t/2$. Para ello se calcula primeramente la velocidad angular μ correspondiente a la parte uniforme del movimiento, mediante la siguiente ecuación donde μ_α y μ_δ son los incrementos seculares en ascensión recta y declinación tabulados en el catálogo:

(42)

$$\mu = \sqrt{\mu_\delta^2 + \mu_\alpha^2 \cos^2 \delta_0}$$

A partir del valor de μ , y considerando el ángulo Π de paralaje y la velocidad radial de la estrella, ambos datos contenidos en el catálogo, así como el intervalo de tiempo (41), se obtiene la velocidad angular buscada a través de la siguiente igualdad donde K es una constante para el acuerdo de las unidades (Mena, 2008):

(43)

$$\mu_m = \mu (1 - K \Pi v_r t)$$

Asimismo, el seno y coseno del ángulo Ψ_0 que indica la dirección del movimiento, se calculan mediante las expresiones:

(44)

$$\cos \Psi_0 = \frac{\mu_\delta}{\mu}, \quad \sin \Psi_0 = \frac{\mu_\alpha \cos \delta_0}{\mu}$$

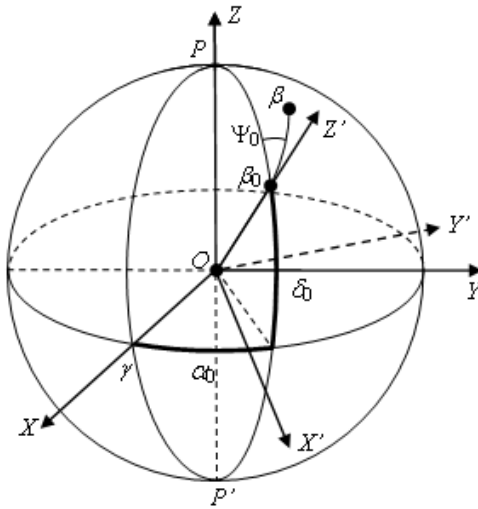


Figura 14.

Una vez determinada la velocidad angular media, se obtiene la distancia esférica ($\mu_m t$) que recorre la estrella entre las épocas E_0 y E . Con este valor, junto a las funciones trigonométricas del ángulo de dirección (Ψ_0), se buscan las componentes del vector de posición $\mathbf{V}'$ de β en el triedro directo con origen en el punto O cuyos ejes OZ' y OX' están situados en el meridiano de β_0 , de forma que OZ' pasa por β_0 y OX' se dirige 90° hacia el Sur. Con tal disposición, y según la figura, las coordenadas de β resultan ser:

(45)

$$\mathbf{V}' = \begin{pmatrix} -\text{sen}(\mu_m t) \cos \Psi_0 \\ \text{sen}(\mu_m t) \text{sen} \Psi_0 \\ \cos(\mu_m t) \end{pmatrix}$$

Finalmente se transforma este vector al triedro ecuatorial $OXYZ$ utilizado como referencia en la época del catálogo. Para ello basta con aplicar los dos giros siguientes que superponen el triedro auxiliar sobre el ecuatorial. Se obtienen así los cosenos directores que proporcionan la posición angular de β corregida de movimiento propio:

(46)

$$\mathbf{V}_{mp} = G_z(-\alpha_0) G_y(\delta_0 - 90^\circ) \mathbf{V}'$$

7. 5.3. Corrección por precesión

La siguiente corrección está motivada por el movimiento de *precesión* terrestre (figura 15. Ref. [11]). Como es sabido el eje de rotación de la Tierra no mantiene su dirección en el espacio, sino que describe un cono con vértice en el centro de gravedad terrestre cuyo período es de aproximadamente 26000 años. Este movimiento, estudiado ya con cierto rigor desde principios del siglo pasado, no es desdeñable toda vez que el ángulo en el vértice del cono ronda el valor $23^{\circ},5$. Su principal efecto es que, al variar la línea de los polos celestes y como consecuencia la posición del ecuador, se produce un movimiento retrógrado del punto vernal en el plano de la eclíptica, lo que motiva una modificación en las coordenadas ecuatoriales de los astros. Este efecto se denomina *precesión de los equinoccios* y tiene un valor aproximado de $50''.3$ / año.

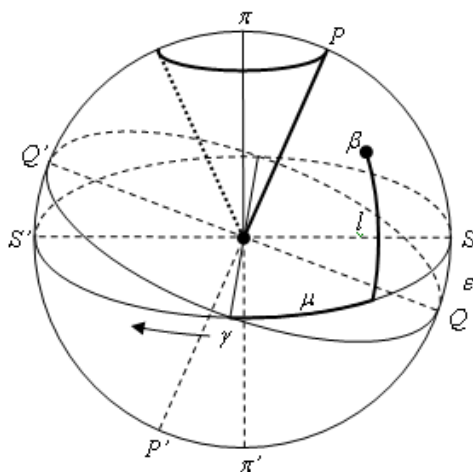


Figura 15.

Además del anterior, la precesión terrestre causa como segundo efecto una *variación en la oblicuidad de la eclíptica*; es decir, influye permanentemente en el valor del ángulo diedro ε formado por la eclíptica y el ecuador. El promedio de tal variación es mucho menor que el correspondiente a la precesión de los equinoccios, siendo por término medio igual a $47''$ / siglo. En la actualidad este efecto, cuyo período se estima en 41000 años, tiende hacia la superposición de los dos planos citados, lo que implica un signo negativo del mencionado incremento secular.

El fenómeno de la precesión general está ocasionado por la atracción que sobre la Tierra ejercen la Luna, el Sol y los restantes planetas del sistema solar, por lo que se trata de un problema que entra dentro de la Mecánica Celeste y que se estudia bajo la denominación de *movimiento o efecto giroscópico*. Básicamente, consiste en el análisis del par de fuerzas de perturbación que actúa sobre la Tierra debido a la atracción gravitatoria de las masas cercanas al planeta, y que se

manifiesta mediante una aceleración complementaria perpendicular al plano determinado por cada fuerza de atracción y el eje del mundo. Como consecuencia de ello, el estudio de la precesión suele dividirse en dos términos: la *precesión lunisolar* y la *precesión planetaria*. El primero es con diferencia el más importante, motivando la práctica totalidad de los valores correspondientes al retroceso de los equinoccios y la oblicuidad de la eclíptica. En particular, es la Luna el astro que produce los mayores efectos dada su proximidad a la Tierra; de hecho, es causante de unos 34'' / año en el movimiento retrógrado de los equinoccios, mientras que al Sol le corresponden unos 16'' / año y a los planetas sólo unos 12'' / siglo.

En lo relativo a la corrección, ésta se aplica sobre el vector obtenido en (46) atendiendo a las posiciones del ecuador y del equinoccio en las épocas E_0 y E . En este contexto, siendo OM la recta intersección de las dos posiciones del ecuador (figura 16. Ref. [11]), y considerando los respectivos triedros ecuatoriales en ambas épocas, la superposición del sistema E_0 sobre el actual E se realiza mediante los ángulos $\gamma_0 M$, θ y $\gamma_m M$ representados en la figura. El primero es la ascensión recta de M en la época E_0 , el segundo la inclinación de un ecuador respecto del otro, y el tercero la ascensión recta de M en la época E . Estos tres ángulos pueden ser obtenidos a partir del intervalo de tiempo (41) mediante los denominados *parámetros precesionales*, los cuales han sido calculados durante años por las fórmulas publicadas en Lieske et al. (1977). Seguidamente se incluyen estos parámetros y su relación con los mencionados ángulos de giro, si bien en la actualidad estas fórmulas están sustituidas por otras más precisas (Mena, 2008):

(47)

$$\begin{aligned} \eta &= 2306''.2181t + 0''.30188t^2 + 0''.017998t^3 \\ z &= 2306''.2181t + 1''.09468t^2 + 0''.018203t^3 \\ \theta &= 2004''.3109t - 0''.42665t^2 - 0''.041833t^3 \end{aligned} \quad \begin{cases} \gamma_0 M = 90^\circ - \eta \\ \gamma_m M = 90^\circ + z \end{cases}$$

Una vez encontrados estos valores, la superposición de los sistemas coordenados requiere realizar tres giros. Inicialmente un giro respecto de OZ en sentido positivo y cuyo argumento es $\gamma_0 M$. A continuación una rotación respecto de OX de ángulo θ , también en sentido positivo. Finalmente, una nueva rotación respecto de OZ de argumento $\gamma_m M$ y realizada en sentido negativo. Como resultado del movimiento descrito, el vector unitario y corregido por la precesión terrestre que proporciona las denominadas *coordenadas medias* para la época E , se obtiene a través de la relación matricial:

(48)

$$V_m = G_z(-90^\circ - z)G_x(\Theta)G_z(90^\circ - \eta)V'$$

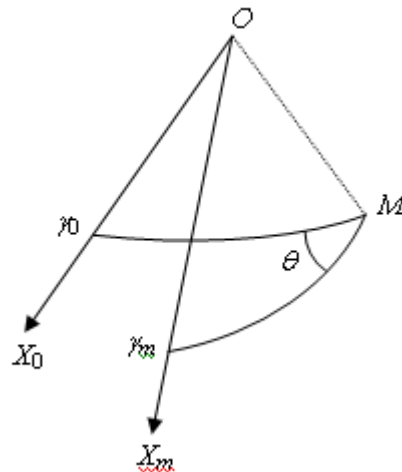


Figura 16.

7. 5.4. Corrección por nutación

Aparte del cono de precesión descrito por el movimiento del *polo medio* terrestre, la atracción gravitatoria de la Luna principalmente, aunque también la del Sol, provocan un movimiento superpuesto al anterior que llamamos *nutación* (figura 17. Ref. [11]). Éste se estudia con relación al de precesión diciendo que en cada instante el *polo verdadero* de la Tierra tiende a describir una elipse cuyo centro es la posición del polo medio. El movimiento de nutación consiste pues en oscilaciones del polo respecto de la trayectoria establecida por la precesión.

El período de la nutación es aproximadamente igual a 18.6 años; tiempo que coincide con el de una revolución completa del *nodo lunar* ^(*) en la órbita del satélite. En cuanto al tamaño de las elipses que representan la vibración del polo verdadero, se trata de figuras muy pequeñas comparadas con el cono de precesión, ya que sus semiejes tienen por valores angulares $9''.2$ y $6''.8$. No obstante los efectos de la nutación sobre las coordenadas ecuatoriales de los astros son importantes, puesto que hacen variar la posición del ecuador con respecto de la eclíptica. Estos efectos consisten en la modificación de la posición del punto vernal en la eclíptica (*nutación en longitud*), unida al incremento del ángulo de inclinación ε de dicho plano respecto de su valor medio correspondiente al movimiento de precesión (*nutación en oblicuidad*). Veamos cómo obtener estos valores.

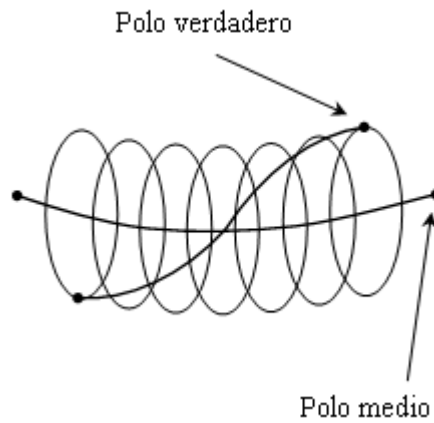


Figura 17.

Como se indica en la figura 19 (Ref. [11]), es la dirección paralela al semieje mayor de la elipse instantánea, la causa de la nutación en oblicuidad $\Delta\varepsilon$. Por ello, para obtener el valor verdadero (ε_v) del ángulo de inclinación de la eclíptica en cada momento, habremos de sumar al valor medio (ε_m) el término $\Delta\varepsilon$:

(49)

$$\varepsilon_v = \varepsilon_m + \Delta\varepsilon$$

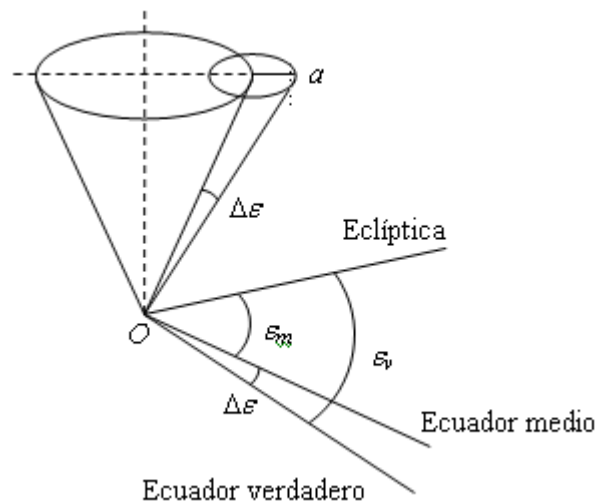


Figura 18.

(*) Punto intersección de la órbita de la Luna con el plano de la eclíptica, en general considerando la rama ascendente de dicha órbita, es decir, la transición de la Luna a valores positivos de su latitud eclíptica.

Seguidamente veamos la manera de calcular las dos componentes de la nutación para el intervalo (41) de tiempo terrestre. Las fórmulas correspondientes recomendadas por la IAU en 1980 consistieron en los siguientes sumatorios de 106 términos, denominados *armónicos de nutación*, los cuales se encuentran actualizados hoy en día (Mena, 2008):

(52)

$$\Delta\psi'' = \sum_{i=1}^{106} (A_i + A'_i t) 10^{-4} \sin(a_i l + b_i l' + c_i F + d_i D + k_i \Omega)$$

$$\Delta\varepsilon'' = \sum_{i=1}^{106} (B_i + B'_i t) 10^{-4} \cos(a_i l + b_i l' + c_i F + d_i D + k_i \Omega)$$

Aquí los factores l , l' , F , D y Ω son los *parámetros nutacionales*, que se definen atendiendo a las longitudes eclípticas de diversos puntos referenciadas al equinoccio medio de la época E . Dichos puntos son: las posiciones del Sol (S) y la Luna (L), el perigeo de la órbita del Sol (PS) y el perigeo de la órbita de la Luna (PL), y el nodo ascendente (N) de la órbita lunar. En este marco, los parámetros nutacionales se corresponden con los ángulos siguientes:

(53)

$$l = \mu_L - \mu_{PL}, \quad l' = \mu_S - \mu_{PS}, \quad F = \mu_L - \mu_N, \quad D = \mu_L - \mu_S, \quad \Omega = \mu_N$$

Sus valores resultan directamente a través de las expresiones publicadas en Seidelmann (1982); son las que siguen, asimismo sustituidas en la actualidad por otras más precisas (Mena, 2008):

(54)

$$\begin{aligned} l &= 134^\circ.96298139 + 477198^\circ.86739806 t + 0^\circ.00869722 t^2 + 0^\circ.0000178 t^3 \\ l' &= 357^\circ.52772333 + 35999^\circ.05034 t - 0^\circ.00016028 t^2 - 0^\circ.0000033 t^3 \\ F &= 93^\circ.27191028 + 483202^\circ.01753806 t - 0^\circ.00368250 t^2 + 0^\circ.0000031 t^3 \\ D &= 297^\circ.85036306 + 445267^\circ.11148 t - 0^\circ.00191417 t^2 + 0^\circ.0000053 t^3 \\ \Omega &= 125^\circ.04452222 - 1934^\circ.13626083 t + 0^\circ.00207083 t^2 + 0^\circ.0000022 t^3 \end{aligned}$$

Respecto a los coeficientes que aparecen en (52), éstos se encuentran relacionados en un fichero que el lector puede obtener acudiendo a organismos especializados.

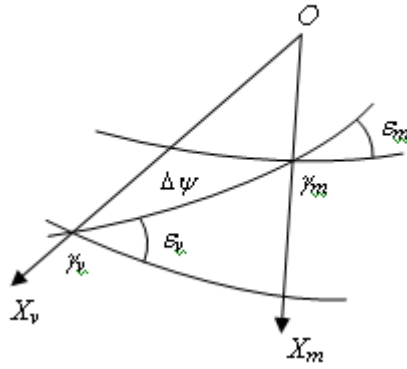


Figura 20.

Una vez obtenidas las componentes de la nutación, así como los valores medio y verdadero del ángulo de inclinación de la eclíptica, la corrección correspondiente sobre las coordenadas medias del astro se realiza superponiendo el triedro ecuatorial medio sobre el verdadero. Ello se logra mediante los tres giros siguientes (figura 20. Ref. [11]). Primeramente un giro respecto de OX , positivo y de argumento igual al ángulo de inclinación medio. En segundo lugar un giro negativo respecto de OZ cuyo valor es la nutación en longitud. Finalmente un giro respecto de OX que tiene por amplitud el ángulo de inclinación verdadero y se ejecuta en sentido negativo. Tras el citado movimiento de ejes, resulta la siguiente relación matricial que rinde las denominadas *coordenadas ecuatoriales verdaderas*:

(55)

$$V_v = G_x(-\varepsilon_v)G_z(-\Delta\Psi)G_x(\varepsilon_m)V_m$$

7. 5.5. Tiempo sidéreo medio y tiempo sidéreo verdadero a 0^h UT

Como se ha indicado anteriormente, el valor de la hora sidérea Greenwich en el instante inicial de cada día de tiempo universal y referido al ecuador y equinoccio medios de la fecha, puede ser obtenido directamente mediante el cálculo. Dicho valor, que denominamos Tiempo Sidéreo Medio (TSM) a 0^h UT, se encuentra a través de la fórmula de Aoki (Aoki et al., 1982) cuya expresión clásica es la que sigue, si bien hoy en día se encuentra actualizada (Mena, 2008):

(56)

$$TSM \ 0^h \ UT = 6^h \ 41^m \ 50^s .54841 + 8640184^s .812866 \ t_0 + 0^s .093104 \ t_0^2 - 6^s .2 \cdot 10^{-6} \ t_0^3$$

Fórmula cuyo argumento es el siguiente intervalo expresado en siglos julianos:

(57)

$$t_0 = \frac{\text{Dia Juliano} - 2451545}{36525}$$

Aunque el valor TSM tiene precisión suficiente para numerosas aplicaciones, en muchas otras es necesario corregirlo a fin de que resulte referido al ecuador y equinoccio verdaderos de la fecha. Para ello basta con añadirle el término correspondiente a la ecuación de los equinoccios, calculado asimismo para el instante t_0 . Este término puede obtenerse por la fórmula elemental (51), o mejor, utilizando la siguiente expresión más precisa, que también ha sido actualizada hace poco tiempo (Mena, 2008):

(58)

$$e_{-e} = \frac{1}{15} (\Delta\psi \cos \varepsilon_v + 0''.00264 \sin \Omega + 0''.000063 \sin 2\Omega)$$

Aquí el factor 1/15 constituye la transformación a segundos de tiempo, calculándose las distintas variables para el momento (57) a través de las relaciones dadas en la corrección de nutación.

En estas condiciones el tiempo sidéreo en Greenwich, corregido también por el movimiento de nutación, se denomina *Tiempo Sidéreo verdadero o Aparente* (TSA) a 0^h UT, y tiene por valor el siguiente (ver fórmula actual en Mena, 2008):

(59)

$$\boxed{TSA\ 0^h\ UT = TSM\ 0^h\ UT + e_{-e}}$$

A raíz de este parámetro y teniendo en cuenta (25) y (26), se obtiene la *Hora Sidérea Local verdadera o Aparente* (HSLA) para un meridiano cualquiera de longitud Λ y en cualquier instante de tiempo universal (HUT):

(60)

$$HSLA = \Lambda + TSA\ 0^h\ UT + HUT \frac{366.2422}{365.2422}$$

Expresión donde como tiempo universal debe tomarse la escala UT1.

7.5.6. Corrección de paralaje

Puesto que el origen utilizado en el FK5 es el baricentro O del sistema solar, a continuación se procede a trasladar el vector de coordenadas verdaderas al centro de gravedad de la Tierra (T). Esta corrección se conoce como paralaje anual, y en su aplicación interviene el parámetro Π incluido en el catálogo y el vector $\mathbf{r}$ de posición de T con respecto de O en la época E . La corrección por paralaje diurna no se considera debido al pequeño valor del radio terrestre comparado con la distancia a la estrella.

Como se observa en la figura 21 (Ref. [11]), los vectores de posición del astro visto desde los puntos O y T están relacionados en cada instante mediante la expresión:

$$\mathbf{OT} + \mathbf{T}\beta = \mathbf{O}\beta \Leftrightarrow \mathbf{r} + \mathbf{X} = \mathbf{V}$$

Por lo tanto, dividiendo por la distancia $O\beta$ y despejando el unitario en la dirección $\mathbf{X}$, resulta:

$$\frac{\mathbf{X}}{O\beta} = \frac{\mathbf{V}}{O\beta} - \frac{\mathbf{r}}{O\beta} \Leftrightarrow \frac{\mathbf{X}}{O\beta} = \frac{\mathbf{V}}{O\beta} - \frac{\mathbf{r}}{r} \cdot \frac{r}{O\beta}$$

De los vectores del segundo miembro, el primero es el unitario determinado por las coordenadas verdaderas (55). En cuanto al segundo, podemos escribir:

$$\frac{r}{O\beta} = \sin \Pi \cong \Pi \text{ rad.}$$

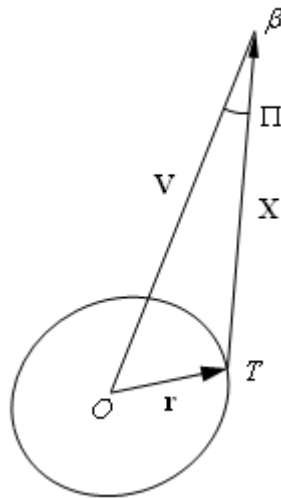


Figura 21.

Se tiene entonces:

$$\frac{\mathbf{X}}{O\beta} = \mathbf{V}_v - \Pi \frac{\mathbf{r}}{r}$$

Y puesto que este vector no es unitario, lo normalizamos; resulta así el correspondiente vector corregido de la paralaje anua:

(61)

$$\mathbf{V}_{par} = \frac{1}{\left| \mathbf{V}_v - \Pi \frac{\mathbf{r}}{r} \right|} \left(\mathbf{V}_v - \Pi \frac{\mathbf{r}}{r} \right)$$

Para obtener el vector $\mathbf{r}$ de posición de la Tierra que interviene en esta expresión, puede interpolarse en las efemérides correspondientes publicadas periódicamente por determinados organismos, como por ejemplo el Real Observatorio de la Armada de San Fernando (Cádiz). Dicha interpolación se realiza con la fecha juliana y la hora de tiempo terrestre de la época E .

7. 5.7. Corrección por aberración de la luz

El fenómeno de aberración de la luz consiste en el desplazamiento aparente de la posición de un objeto, debido a la relación existente entre la velocidad del movimiento relativo observador – objeto y la velocidad de la luz. Descubierto por Bradley en 1728, este fenómeno es causa de la siguiente corrección en las coordenadas del astro, puesto que la velocidad de la Tierra, sobre todo la de traslación, si bien no es comparable a la de la luz, ronda los cien mil kilómetros por hora y por tanto es relativamente grande.

Suponiendo el caso más general, aquél en que tienen movimiento tanto el observador como el objeto observado, la denominada aberración planetaria se estudia mediante el ángulo formado por las direcciones que unen el objeto con el observador en dos instantes E_0 y E . Este fenómeno se descompone en dos: la aberración por tiempo luz (A_c), que se debe al movimiento exclusivo del objeto, y la aberración estelar (A_e), motivada por el movimiento único del observador.

La aberración por tiempo luz implica que la posición observada del objeto cuando la luz llega al observador, supuestamente fijo, es distinta de la posición que realmente tiene el objeto en ese momento a causa del tiempo que tarda la luz en realizar el recorrido. No obstante, la corrección correspondiente no es importante cuando el objeto es una estrella cuyo movimiento propio es

relativamente pequeño. De hecho este valor puede ser despreciado en el caso que nos incumbe, aunque ha de ser tenido en consideración en otras aplicaciones.

Por su parte el estudio de la aberración estelar supone que el objeto permanece fijo mientras que el observador está movimiento. Dicho movimiento ocasiona que al llegar la luz al observador, éste vea el objeto en una posición más adelantada conforme a la dirección y sentido de su movimiento (*ley general de la aberración*). En consecuencia será necesario introducir las correcciones pertinentes a las coordenadas del objeto a fin de obtener las denominadas *coordenadas aparentes*; o sea, las coordenadas del punto donde el astro será observado. Para ello, y puesto que el movimiento de un observador sobre la Tierra es la composición de varios, el fenómeno de aberración estelar se descompone a su vez en tres; a saber: la *aberración secular*, producida por el movimiento del sistema solar en la galaxia; la *aberración anua*, que es debida a la traslación terrestre; y la *aberración diurna*, causada por el movimiento de rotación.

* La aberración secular se encuentra incluida en las coordenadas que figuran en el FK5, por lo que no ha de ser tomada en cuenta nuevamente. En cuanto a la aberración anua, deberemos calcular y normalizar el vector **X** representado en la figura 22 (Ref. [11]) para obtener la corrección correspondiente a la posición β' aparente del astro. Consideremos los vectores **X**, **V** y **d**:

$$\mathbf{X} = \mathbf{V} + \mathbf{d}$$

Y dividiendo por la distancia $T\beta$, resulta:

$$\frac{\mathbf{X}}{T\beta} = \frac{\mathbf{V}}{T\beta} + \frac{\mathbf{d}}{T\beta}$$

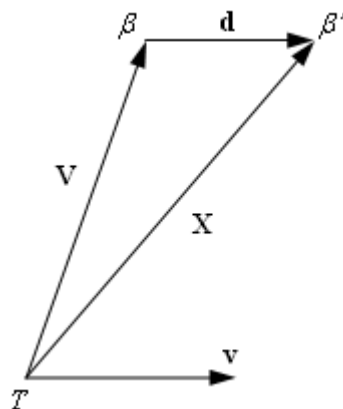


Figura 22.

En esta expresión el primer sumando del segundo miembro es el unitario (61) resultante de la corrección por paralaje anua. El segundo es un vector que tiene la misma dirección que la velocidad instantánea $\mathbf{v}$ de traslación terrestre, y cuyo módulo está relacionado con la distancia d que en dicha dirección separa las posiciones aparente y real del astro. Esta distancia está determinada por la ley general de la aberración; y según dicha ley se obtiene mediante el producto del *factor de aberración anua* (que es la razón entre la velocidad v de traslación del observador y la velocidad c de la luz) por la distancia del astro al observador. Por lo tanto, vectorialmente se tiene:

$$\mathbf{d} = \frac{\mathbf{v}}{c} \cdot T\beta$$

Como consecuencia, sustituyendo ésta en la expresión anterior nos queda:

(62)

$$\frac{\mathbf{X}}{T\beta} = \mathbf{V}_{par} + \frac{1}{c} \mathbf{v}$$

Encontramos así que el vector unitario mediante el cual se definen las coordenadas ecuatoriales del astro corregidas por aberración anua, resulta por la igualdad:

(63)

$$\mathbf{V}_{aa} = \frac{1}{\left| \mathbf{V}_{par} - \Pi \frac{1}{c} \mathbf{v} \right|} \left(\mathbf{V}_{par} + \frac{1}{c} \mathbf{v} \right)$$

Para encontrar las componentes del vector $\mathbf{v}$ de la velocidad lineal de la Tierra en su órbita en la fecha juliana y hora de tiempo terrestre de cada instante, puede optarse por interpolar en las Efemérides Astronómicas del Observatorio de San Fernando.

* Sólo nos resta ya aplicar la corrección por aberración diurna para que las coordenadas del astro queden dispuestas para su transformación al sistema horizontal. Esta corrección es debida a que el observador está situado en la superficie de la Tierra y no en su centro de gravedad, por lo que se encuentra sometido a una velocidad lineal $\mathbf{v}'$ que sólo será nula en los polos terrestres. Aunque algunos autores abogan porque esta corrección no es necesaria en razón a su pequeñez, la incluimos aquí toda vez que el proceso requerido para su determinación es similar al seguido en el caso de la aberración anua. De hecho, la expresión (62) es completamente válida cuando se considera en lugar de $\mathbf{V}_{par}$ el unitario resultante de (63), y en lugar de $\mathbf{v}$ la velocidad lineal $\mathbf{v}'$ motivada por la rotación. Por lo tanto, en este caso se tiene:

(64)

$$\frac{\mathbf{X}}{T\beta} = \mathbf{V}_{aa} + \frac{1}{c} \mathbf{v}'$$

Evidentemente la velocidad lineal del observador depende ahora de la velocidad angular ω' de rotación terrestre y de la distancia del observador al eje de giro. Tomando como aproximación de la forma de la Tierra la correspondiente al elipsoide WGS – 84, y suponiendo que el punto de observación se encuentra a una altitud elipsoidal h en la latitud geodésica φ , el módulo de la velocidad lineal se determina entonces por la siguiente expresión donde N es el respectivo valor de la normal principal:

$$v' = \omega'(N + h) \cos \varphi$$

Además, si consideramos el sistema de coordenadas ecuatoriales con origen en el centro de gravedad terrestre, al estar el vector $\mathbf{v}'$ situado en el plano perpendicular al eje de rotación que pasa por el punto de observación, resulta que dicho vector no tiene componente z , por lo que sólo habremos de obtener las componentes x e y . Estos dos valores se definen fácilmente atendiendo al ángulo formado en cada instante por el meridiano del observador con respecto al meridiano que contiene al punto vernal; es decir, considerando la hora sidérea local. Y puesto que el punto vernal ha sido corregido a su posición verdadera por las correcciones anteriores, como tal hora tomamos la hora sidérea local verdadera o aparente (60). Por lo tanto, la expresión que define la velocidad lineal del observador resulta ser:

$$v' = \omega'(N + h) \cos \varphi \begin{pmatrix} -\sin HSLA \\ \cos HSLA \\ 0 \end{pmatrix}$$

Introduzcamos aquí el unitario $\mathbf{u}'$ y la función auxiliar $A(\varphi)$, definidos como sigue:

(65)

$$\mathbf{u}' = \begin{pmatrix} -\sin HSLA \\ \cos HSLA \\ 0 \end{pmatrix}, \quad A(\varphi) = (N + h) \cos \varphi$$

Nos queda entonces:

$$\mathbf{v}' = \omega' A(\varphi) \mathbf{u}'$$

Dispuesto $\mathbf{v}'$, el vector (64) en estudio se escribe entonces así:

$$\frac{\mathbf{X}}{T\beta} = \mathbf{V}_{aa} + \frac{\omega'}{c} A(\varphi) \mathbf{u}'$$

Con ello, el vector unitario definitivo que proporciona las denominadas *coordenadas ecuatoriales aparentes* del astro, resulta determinado por la expresión:

(66)

$$\mathbf{V}_{aa} = \frac{1}{\left| \mathbf{V}_{aa} + \frac{\omega'}{c} A(\varphi) \mathbf{u}' \right|} \left(\mathbf{V}_{aa} + \frac{\omega'}{c} A(\varphi) \mathbf{u}' \right)$$

7. 5.8. Corrección por la refracción atmosférica

Calculadas las coordenadas aparentes del astro, la observación del mismo requiere transformarlas al sistema horizontal del punto de estación. Para ello aplicaremos sobre el vector (66) la primera de las expresiones (7) donde intervienen como argumentos el ángulo de latitud astronómica y la hora sidérea local. Desde luego, para estos dos parámetros pueden considerarse valores más o menos precisos según la finalidad de la medición. Sin embargo, una corrección de la que no se puede prescindir en la observación es la correspondiente a la refracción atmosférica, lo que está motivado por la importante influencia de este fenómeno sobre los ángulos cenitales.

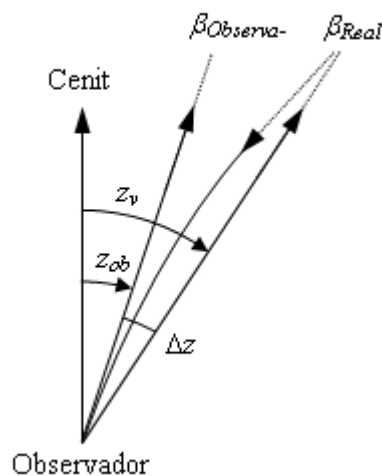


Figura 23.

El estudio de la refracción atmosférica es un tema complejo por su variabilidad con el tiempo y el lugar. Y aunque existen diversos modelos atmosféricos que, tomando el índice de refracción como variable, se apoyan principalmente en la subdivisión en distintas capas (Mena, 2008), el recorrido real del rayo de luz a través de la atmósfera no es fácil de obtener. De hecho, la desviación total que sufre la trayectoria de un rayo depende de la función que rige el índice de refracción a cada altitud; función que sólo puede determinarse de forma muy aproximada y que además únicamente es válida para un corto período de tiempo en la zona donde se mide. No obstante, se sabe que los principales factores que influyen en la refracción atmosférica son la presión y la temperatura; así como también se conoce que la refracción motiva una disminución del ángulo cenital bajo el cual se observa el astro (z_{ob}) respecto del valor real correspondiente (z_v), creciendo rápidamente la diferencia con el valor del ángulo. Escribimos por tanto (figura 23. Ref. [11]):

(67)

$$z_v = z_{ob} + \Delta z$$

Dada pues la dificultad de obtener con toda precisión el incremento a introducir en el ángulo cenital, en las observaciones astronómicas se opta por utilizar alguna de las muchas expresiones existentes que, basadas en determinadas hipótesis experimentales, son funciones del valor de z y de la presión y temperatura del momento. Una de ellas, muy sencilla de aplicar y que proporciona resultados bastante aceptables, es la *fórmula de Bradley*:

(68)

$$\Delta z'' = 57''.8 \operatorname{tg}(z - 230'' \operatorname{tg} z) \frac{278(1 - 0.00016 T)}{753.5(268 + T)} b$$

Aquí, T es la temperatura ambiente expresada en grados centígrados y b la presión definida en milímetros de mercurio. Otro modelo recomendable, que incrementa ligeramente la precisión de éste pero a cambio de un método de cálculo bastante más costoso, es el que incluyen las Efemérides Astronómicas del Observatorio de San Fernando. Tanto en un caso como en el otro, se pone claramente de manifiesto cómo el valor de la corrección aumenta considerablemente a medida que lo hace el ángulo cenital, incrementándose con ello los errores. Por esta razón, en las mediciones de cierta precisión es de la máxima importancia que los ángulos cenitales sean lo más pequeños posible, para lo cual los astros a observar deben ser adecuadamente elegidos.

7. 5.9. Otras correcciones a la posición aparente de un astro

Anteriormente hemos citado que ciertas correcciones que en rigor afectan a las observaciones, pueden ser despreciadas en razón a su orden de magnitud. Tal es el caso de la corrección por paralaje diurna así como también podría serlo el de la aberración diurna. Sin embargo, cuando se utilizan procedimientos electrónicos de medición, determinados fenómenos que no afectan prácticamente en el posicionamiento clásico debido al error inherente al mismo, comienzan a tomar importancia. Uno de estos fenómenos es el conocido como *deflexión de la luz* que comentaremos ahora muy brevemente.

En la teoría de la relatividad general se indica que al pasar un rayo de luz por las proximidades de un objeto de masa relevante, sufre una desviación causada por la atracción gravitatoria que ejerce la masa sobre él. Este fenómeno, del que actualmente existen numerosas evidencias procedentes de la observación de cuásares y galaxias distantes, se midió por primera vez con ocasión de un eclipse de Sol. Para ello se eligió una estrella cuya posición aparente estaba muy próxima al borde del Sol, estableciéndose su situación respecto de otras estrellas más alejadas. Reiterando las mediciones en una época posterior al eclipse, y durante la noche al objeto de eliminar la influencia del Sol, la estrella seleccionada aparecía desplazada de su posición inicial con relación a las que se habían tomado como referencia. Pues bien, el valor angular del citado desplazamiento es precisamente la distorsión por la deflexión de la luz; distorsión que en la primera observación había permitido detectar una estrella oculta tras el Sol. Este experimento, que ha sido realizado numerosas veces, demuestra que la discrepancia entre los valores medios obtenidos y los calculados por la teoría de la relatividad es de tan solo dos o tres centésimas de segundo. En concreto, para la masa del Sol el ángulo de deflexión es inversamente proporcional al cuadrado de la velocidad de la luz y a la mínima distancia entre el rayo y el centro solar (Mena, 2008); ello confirma que, salvo para visuales próximas al Sol, el ángulo en cuestión es muy pequeño y por tanto obviarse cuando las observaciones se realizan en el arco nocturno.

Bibliografía

- [1] Aoki, S., Guinot, B., Kaplan, G.H., Kinoshita H., et al. 1982. The New Definition of Universal Time. *Astronomy and Astrophysics*, 105, pp. 359 – 361.
- [2] Bakulin, P.I., Kononovich, E.V., Moroz, V.I., 1983. *Curso de Astronomía General*. Ed. Mir. Moscú.
- [3] Bate, R., Mueller, D., White, J., 1971. *Fundamentals of Astrodynamics*. Dover P, N.Y.
- [4] Bradt, H., 2004. *Astronomy methods*. Cambridge University Press. U.K.
- [5] Brouwer, D., Clemence, G.M., 1971. *Methods of Celestial Mechanics*. Academic Press. New York.
- [6] Danjon, A., 1982. *Astronomie Générale*. Senac. Paris.
- [7] Eichhorn, H., 1975. *Astronomy of Star Positions*. Frederick Ungar P. in. New York.
- [8] Ferrer Torio, R., Piña Patón, B., 1988. *Astronomía de posición*. Servicio de publicaciones de la ETS de Ingenieros de caminos, canales y puertos. Universidad de Cantabria.
- [9] Lieske, J.H., Lederle, T., Fricke, W., Morando, B., 1977. Expression for the Precession Quantities Based upon the IAU (1976) System of Astronomical Constants. *Astron. and Astrophysics*, Vol. 58, no. 1-2, pp. 1-16.
- [10] Mena, J.B., 2004. Cálculo de las correcciones astronómicas para el levantamiento astrogeodésico. VII Jornadas de Trabajo en Mecánica Celeste, 28-30 de Junio de 2004. Real Observatorio de la Armada, S. Fernando (Cádiz).
- [11] Mena, J.B., 2008. *Geodesia Superior*. Texto en dos tomos. Centro Nacional de Información Geográfica.
- [12] Orús, J.J., Catalá, M.A., 1987. *Apuntes de Astronomía*. Departamento de Física de la Atmósfera, Astronomía y Astrofísica. Universidad de Barcelona.
- [13] Roy, A., 1978. *Orbital Motion*. Adam Hilger. Bristol.
- [14] Seidelmann, P.K., 1982. 1980. IAU Nutation: The Final Report of the IAU Working Group on Nutation. *Celestial Mechanics*, Vol. 27, May 1982, pp. 79-106.
- [15] Sevilla, M.J., 1989. *Mecánica Celeste Clásica*. Instituto de Astronomía y Geodesia. Universidad Complutense – C. S. I. C. Madrid.
- [16] Sevilla, M.J., 2003. *Actualización de cálculos astronómicos*. Instituto de Astronomía y Geodesia. Universidad Complutense – C. S. I. C. Madrid.
- [17] SGE, 1972. *Astronomía Geodésica*. Servicio Geográfico del Ejército. Escuela de Geodesia. Madrid.
- [18] Zakatov, P.S., 1981. *Curso de Geodesia Superior*. Editorial Mir. Moscú.

Tema 8. Métodos de transformación entre Sistemas Geodésicos de Referencia Clásicos y Geocéntricos. Transformación de cinco parámetros. Transformación de siete parámetros. Ecuaciones de regresión. Métodos basados en la eliminación de la distorsión de la red.

8.1. Métodos de transformación entre Sistemas Geodésicos de Referencia Clásicos y Geocéntricos

Los datums clásicos definidos por 8 parámetros: 6 que definen la posición en el espacio de un elipsoide de referencia y 2 para la forma/tamaño del mismo, fueron concebidos para ajustarse a una porción determinada de la Tierra. Su posición relativa a la figura de la misma se hizo lo más tangente posible al territorio que se deseaba abarcar: Datum Madrid centrado en el observatorio del Retiro para proporcionar la base geodésica a España, Datum ED50 para Europa centrado en Potsdam. Es claro que si deseamos abarcar la totalidad de la figura terrestre no se puede hacer tangencia en ninguna parte de la superficie de la misma, el objetivo es centrarlo de la manera más precisa posible en su geocentro, más concretamente en el centro de masas considerando Tierra, océanos y atmósfera [1].

Por ejemplo, el datum ED50 fue concebido para dotar a toda Europa de una homogeneidad geodésica y en consecuencia cartográfica, en realidad no fue adoptado por todos los países. Hoy en día la propia globalización de los Sistemas Globales de Navegación por Satélite (GNSS), la necesidad de unificar la cartografía mundial, usos en defensa o aeronáutica hacen necesario el uso de Sistemas Geodésicos de Referencia (SGR) globales. La transformación entre un sistema clásico y uno geocéntrico no es sencilla [2].

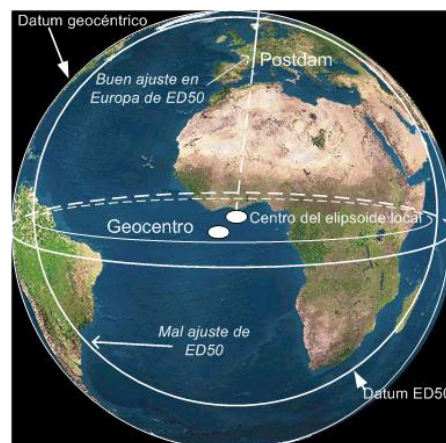


Figura 1. Datum local y geocéntrico

Cualquier levantamiento geodésico o topográfico, y la cartografía a que pueda dar lugar, están necesariamente georeferenciados utilizando un SGR. El avance tecnológico con las precisiones alcanzables hoy día ha permitido definir y usar en la práctica sistemas “globales” y considerar en éstos, además de las coordenadas habituales (latitud, longitud y altitudes elipsoidal y ortométrica), su época de referencia (ITRSyy), lo que equivale a introducir una cuarta coordenada, el tiempo y/o las distintas épocas en las que se expresan los parámetros, coordenadas y el datum. Sin embargo, en este tema no se tiene en cuenta la dependencia temporal, considerando que los datos corresponden a coordenadas de puntos bajo la hipótesis de “Tierra rígida”. Es decir, se cuenta con un conjunto de puntos “único”, pero en ellos se dispone de coordenadas al menos en un SGR (SGR1) y, en algunos, también las correspondientes a un segundo sistema (SGR2) [3]. Actualmente, debido a la proliferación de equipos GPS, sistema que permite ahorros importantes en tiempo y dinero, es muy común tratar en todo trabajo con dos sistemas, el primero de ellos un sistema geocéntrico convencional para fijar ambigüedades y compensar vectores y el segundo uno clásico donde se expresan los resultados finales [4]. Conviene recordar que se trata de puntos ubicados en una parte de la corteza perteneciente a una sola placa tectónica que rota en milisegundos de arco por año. No es menos cierto que las coordenadas en ambos sistemas son dos “aproximaciones” distintas de una “única realidad”, considerada indeformable, que es el conjunto de puntos en el terreno. Cada conjunto de coordenadas adolece de un cierto “grado de aproximación”, y es función de un sin número de causas, accidentales o sistemáticas, entre las que se pueden citar las siguientes [3]:

- o la precisión de las observaciones primitivas (las medidas) que dieron lugar finalmente al conjunto de coordenadas;
- o el grado de exactitud de las correcciones aplicadas a las observaciones para adecuarlas a la realidad física o a la superficie de referencia para los cálculos y el ajuste final (p. ej. “reducción” de distancias o ángulos);
- o la herramienta (programas) utilizada para el ajuste de red y forma en que se ha aplicado (p. ej. “puntos fijos”; ¿se conoce el orden de magnitud de la exactitud de sus coordenadas y el de la precisión relativa entre ellos?);
- o la adecuación al modelo geométrico de representación: bidimensional (latitud y longitud, o abscisa y ordenada en una “proyección cartográfica” plana; en ambos casos con la elección unívoca de elipsoide-datum), o bien, tridimensional (cartesianas X, Y, Z , o lo que es igual φ, λ, h referidas al elipsoide asociado con el sistema). No confundir este último caso con el anterior cuando a éste se le añade la altitud, en cuyo caso se puede hablar de coordenadas “dos más una” donde las primeras son relativas al elipsoide y la tercera lo es al geoide ($\approx$ nivel medio del mar en ...).

En consecuencia, es preciso tener presentes algunas reglas o conceptos esenciales, partiendo de la base de que se pretende encontrar un “modelo matemático” (parámetros), que satisfaga la necesidad de transformar las coordenadas de puntos de un primer sistema de referencia geodésico (SRG1) en las que le corresponden en un segundo sistema (SRG2). Siempre se considera la transformación entre un Sistema Geocéntrico, por ejemplo ETRS89, y un Sistema clásico, por ejemplo ED50.

Es útil recordar algunas ideas esenciales:

- o La transformación, en primera instancia, debe responder al tipo de “transformación de semejanza” por su propia naturaleza [5].
- o El tipo de acuerdo de las coordenadas transformadas con las conocidas en el segundo sistema responde a la imprecisión que tengan las del conjunto menos preciso. Las coordenadas no mejoran con la transformación.
- o Se debe elegir como conjunto primero aquel que se estime más preciso y con mayor coherencia interna.
- o Los parámetros obtenidos (el modelo) serán diferentes si la cantidad de puntos comunes considerados es diferente, o, incluso considerando el mismo número de vértices, si se elige otro conjunto de puntos comunes.
- o La elección de puntos comunes debe procurar, en cuanto sea posible, cubrir la zona holgadamente y con uniformidad. Si se dispone de mayor densidad en una parte de ella, los puntos ignorados en la obtención del modelo pueden ser utilizados como test para estimar la bondad del ajuste, junto con los residuos del ajuste por mínimos cuadrados.
- o La aplicación de los parámetros de transformación debe hacerse a todos los puntos, los comunes y los no comunes, y es muy conveniente que figuren junto con la relación final de coordenadas transformadas; así es posible reconstruir el conjunto primitivo posteriormente. Aún más, si en el futuro se dispusiera de mejor conjunto con coordenadas en ambos sistemas - más puntos y mejor distribuidos -, se podrían calcular mejores parámetros.

Se pueden encontrar más detalles pueden en Dalda y González.-Matesanz [3].

8.2. Transformación de cinco parámetros o Molodensky

El planteamiento teórico correcto de los sistemas geodésicos locales implica que el eje de rotación del elipsoide es paralelo al geocéntrico, que el plano origen de longitudes sea también paralelo al del sistema geocéntrico, y por último, que los acimutes Laplace mantienen en el

sistema local la reorientación de la red. En resumen, el centro del elipsoide del sistema local se encuentra desplazado respecto al centro de masas, origen del sistema ECEF, y las coordenadas de éste $(\Delta X_0, \Delta Y_0, \Delta Z_0)$ junto con las diferencias de semieje y aplastamiento $(\Delta a, \Delta f)$ constituyen los cinco parámetros que dan lugar a las correspondientes fórmulas de transformación.

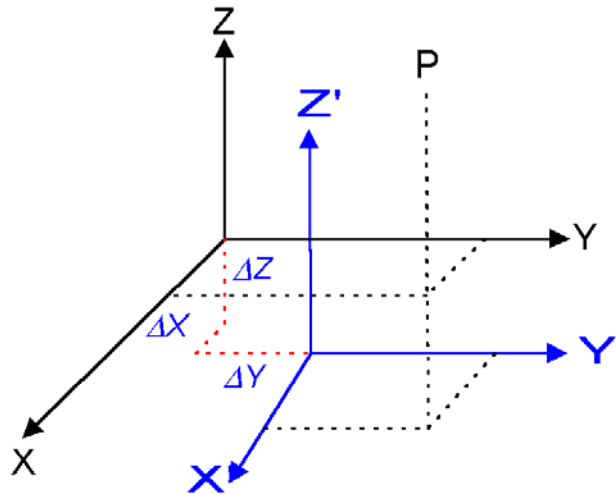


Figura 2. Transformación entre sistemas paralelos

Bajo esta consideración, y después de algunas simplificaciones se obtiene [6]:

$$\begin{aligned} a\Delta\varphi &= \sin\varphi\cos\lambda\Delta X_0 + \sin\varphi\sin\lambda\Delta Y_0 - \cos\varphi\Delta Z_0 - 2a\sin\varphi\cos\varphi\Delta f, \\ a\cos\varphi\Delta\lambda &= \sin\lambda\Delta X_0 - \cos\lambda\Delta Y_0, \end{aligned} \quad (1)$$

y

$$\Delta h = -\cos\varphi\cos\lambda\Delta X_0 - \cos\varphi\sin\lambda\Delta Y_0 - \sin\varphi\Delta Z_0 + \Delta a - a\sin^2\varphi\Delta f.$$

Siendo

$$\begin{aligned} \Delta X_0 &= X_{local} - X_W, \Delta Y_0 = Y_{local} - Y_W, \Delta Z_0 = Z_{local} - Z_W, \\ \Delta a &= a_{local} - a_W, \Delta f = f_{local} - f_W, \\ \Delta\varphi &= \varphi_{local} - \varphi_W, \Delta\lambda = \lambda_{local} - \lambda_W. \end{aligned} \quad (2)$$

Se puede encontrar otras dos ternas de fórmulas que responden con esta a las tres formas de fijar el datum local:

Posición del centro del elipsoide: $(\Delta X_0, \Delta Y_0, \Delta Z_0)$

Coordenadas del punto origen (Datum): $(\varphi_1, \lambda_1, h_1)$

Separación geoide-elipsoide en origen: (ξ_1, η_1, N_1)
 (deflexión, ondulación)

Una forma más completa de las fórmulas anteriores son las conocidas como “fórmulas estándar de Molodensky” [7]:

$$\begin{aligned}
 (R_M + h)\Delta\varphi &= \sin\varphi \cos\lambda\Delta X_0 + \sin\varphi \sin\lambda\Delta Y_0 - \cos\varphi\Delta Z_0 \dots \\
 \dots\dots\dots - (R_M a/b + R_N b/a) \sin\varphi \cos\varphi\Delta f - (R_N e^2 \sin\varphi \cos\varphi)\Delta a, \\
 (R_N + h)\cos\varphi\Delta\lambda &= \sin\lambda\Delta X_0 - \cos\lambda\Delta Y_0, \\
 \Delta h &= -\cos\varphi \cos\lambda\Delta X_0 - \cos\varphi \sin\lambda\Delta Y_0 - \sin\varphi\Delta Z_0 + a/R_N\Delta a - R_N b/a \sin^2\varphi\Delta f.
 \end{aligned}
 \tag{3}$$

Este modelo, con mayor o menor simplificación, es el utilizado comúnmente en receptores y software GPS, aunque no el único. A veces se pueden apreciar diferencias en los resultados que pueden ser achacables a los valores adoptados para los 3 parámetros de traslación $(\Delta X_0, \Delta Y_0, \Delta Z_0)$ ya que los otros dos $(\Delta a, \Delta f)$ dependen de los elipsoides asociados en cada sistema, y obviamente son bien conocidos.

8.3. Transformación de siete parámetros entre sistemas casi paralelos

Responde en la práctica a que, si se considera que las coordenadas en ambos sistemas se refieren a una realidad física “única” (los puntos del terreno), la función de transformación debe ser una razón de semejanza 3D, por lo que al llevar el poliedro (puntos) en el sistema local sobre el poliedro en el global los ejes de los triedros de referencia presentan una ligera falta de paralelismo o lo que es igual, una rotación infinitesimal del sistema local. La demostración de la conformidad de esta transformación se puede ver en Grafarend et al.[8].

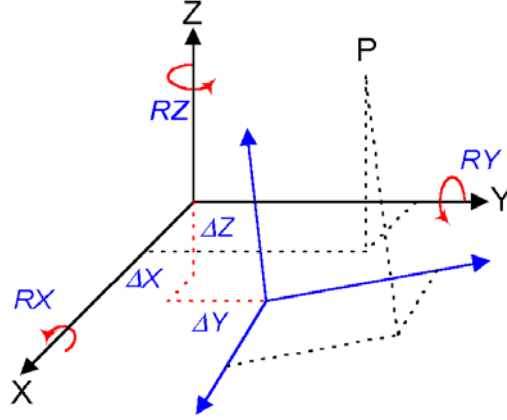


Figura 3. Transformación entre sistemas casi paralelos

Para abordarlo hay que disponer de coordenadas cartesianas 3D en cada uno de los sistemas para los vértices de un “poliedro” en el espacio (figura 3), de manera siguiente:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_{W/L} = \begin{bmatrix} (N+h) \cos \varphi \cos \lambda \\ (N+h) \cos \varphi \sin \lambda \\ (N+h) \sin \varphi - e^2 N \sin \varphi \end{bmatrix}_{W/L} = (N+h)u_{W/L} - \begin{bmatrix} 0 \\ 0 \\ e^2 N \sin \varphi \end{bmatrix}_{W/L} \quad (4)$$

Si se sustituye h por N_{WL} (ondulación) se está considerando el punto sobre el elipsoide donde corta la normal desde P , en vez del propio punto P , y al transformar ese punto se obtiene la ondulación del geode en el segundo sistema de referencia.

La forma de plantear la resolución de la relación de semejanza (tres componentes de la translación, un factor de escala, y tres componentes de la rotación, en total 7 parámetros) da lugar a dos modelos 3D conocidos como el de Bursa-Wolf y el de Molodensky-Badekas. Teniendo presente que las tres componentes de giro se pueden considerar infinitesimales, sus cosenos se igualan a 1 y los senos al ángulo, con lo que la formulación de la matriz se simplifica

Modelo de Bursa-Wolf

Consideramos la expresión siguiente:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_L = \begin{bmatrix} \Delta X_0 \\ \Delta Y_0 \\ \Delta Z_0 \end{bmatrix} + (1+\mu) \begin{bmatrix} 1 & -\Omega_z & \Omega_y \\ \Omega_z & 1 & -\Omega_x \\ -\Omega_y & \Omega_x & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_W, \quad (5)$$

que también se puede expresar de esta otra manera, más adecuada para representar el sistema de ecuaciones para compensar por mínimos cuadrados:

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ \vdots \\ X_i \end{bmatrix}_L = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ \vdots \\ X_i \end{bmatrix}_W + \begin{bmatrix} 1 & 0 & 0 & X_1 & 0 & -Z_1 & Y_1 \\ 0 & 1 & 0 & Y_1 & Z_1 & 0 & -X_1 \\ 0 & 0 & 1 & Z_1 & -Y_1 & X_1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & X_i & 0 & -Z_i & Y_i \end{bmatrix}_W \begin{bmatrix} \Delta X_0 \\ \Delta Y_0 \\ \Delta Z_0 \\ \mu \\ \Omega_X \\ \Omega_Y \\ \Omega_Z \end{bmatrix} \quad i=\{1 \dots n\}, \quad (6)$$

conformaría la expresión de Bursa-Wolf para la transformación de siete parámetros.

Modelo de Molodensky-Badekas

En esencia, se basa en lo mismo que el anterior método, establecer una relación de semejanza 3D, pero es más indicado para determinar los 7P correspondientes a una zona limitada dado que en caso contrario la matriz sería singular (p. ej. la Península Ibérica). El problema se puede abordar en un doble paso; primero obteniendo los parámetros de escala y las tres rotaciones (translación cero), segundo aplicando estos a los baricentros calcular los tres de translación, de esta manera se pueden utilizar estos 7P en la expresión de Bursa-Wolf. Siendo $x_1 = X_1 - \bar{X}$, $y_1 = Y_1 - \bar{Y}$ y $z_1 = Z_1 - \bar{Z}$ se tiene:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ \vdots \end{bmatrix}_L = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ \vdots \end{bmatrix}_W + \begin{bmatrix} x_1 & 0 & -z_1 & y_1 \\ y_1 & z_1 & 0 & -x_1 \\ z_1 & -y_1 & x_1 & 0 \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}_W \begin{bmatrix} \mu \\ \Omega_X \\ \Omega_Y \\ \Omega_Z \end{bmatrix}, \quad (7)$$

hallados estos 4 parámetros los 3 restantes vienen definidos por:

$$\begin{bmatrix} \Delta X_0 \\ \Delta Y_0 \\ \Delta Z_0 \end{bmatrix} = \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix}_L - \begin{bmatrix} \bar{X} \\ \bar{Y} \\ \bar{Z} \end{bmatrix}_W - \begin{bmatrix} \bar{X} & 0 & -\bar{Z} & \bar{Y} \\ \bar{Y} & \bar{Z} & 0 & -\bar{X} \\ \bar{Z} & -\bar{Y} & \bar{X} & 0 \end{bmatrix}_W \begin{bmatrix} \mu \\ \Omega_X \\ \Omega_Y \\ \Omega_Z \end{bmatrix}. \quad (8)$$

El primer modelo citado es muy adecuado para transformación entre sistemas globales, pues se puede contar con puntos comunes repartidos por todo el Globo, mientras que su utilización en un conjunto regional los parámetros hallados presentan una alta correlación entre ellos, en cuyo caso es más aconsejable el segundo.

Es importante resaltar una vez más que en ambos casos surge la dificultad de disponer de coordenadas geodésicas completas en el sistema local puesto que en mayoría de los casos no se dispone de altitudes elipsoidales en el sistema local sino únicamente de ortométricas. Además en el caso español la inmensa mayoría de la red (ROI) procede de compensación de desniveles trigonométricos. Por ello es necesario aplicar un proceso iterativo partiendo de un conocimiento

aproximado del valor de la ondulación del geoide en el sistema local y aplicando los parámetros provisionales a los puntos del geoide (ondulación de IBERGEO95) se calcula un nuevo valor de ella en el local y se procede nuevamente a buscar el ajuste, ver [2] para más información. No considerar coordenadas espaciales en ambos sistemas duplica el error en altura [3].

8.4. Ecuaciones de regresión

En cualquier sistema en el que exista una variable cuantitativa que cambie y se necesite examinar los efectos que otras variables ejerzan sobre ésta u otras, el método de regresión múltiple merece especial atención. De hecho, una de las primeras aproximaciones a la resolución del problema de transformación entre SGR clásicos y geocéntricos fue mediante métodos de regresión [9]. Puede de hecho existir una relación simple funcional entre variables; en la mayoría de los procesos físicos esto es una excepción en lugar de una regla. A menudo existe una relación funcional que es demasiado complicada de describir mediante términos simples [10]. En este caso, se pretende encontrar una aproximación a esta relación funcional mediante una función matemática simple, como por ejemplo un polinomio, que contenga las variables apropiadas y que aproxime la función verdadera sobre un rango limitado de variables involucradas.

Se pueden distinguir dos tipos de variables: variables predictoras y variables respuesta. Por variable predictora se entiende aquella que puede tomar un valor determinado (por ejemplo un valor de temperatura), por variable respuesta se entiende aquella afectada del cambio que ocurre en las variables predictoras, es decir, se genera una transmisión hacia la variable o variables respuesta. Si somos capaces de descubrir una relación simple o una dependencia de la variable respuesta utilizando uno o varios predictores se habrá tenido éxito en nuestro proceso de regresión. Estas variables tienen distintos nombres en la literatura [11]: *Variables predictoras o explicativas: variables de entrada, entrada, variables x, regresores o variables independientes. Variables respuesta: variables de salida, salidas, variables y o variables dependientes*

En regresión generalmente se utiliza el método de mínimos cuadrados tanto para examinar los datos como para comprobar la dependencia y las relaciones que pueden existir entre variables. El modelo general de regresión es el siguiente: *Variable respuesta = función asociada a un modelo + errores aleatorios*

La función modelo generalmente será conocida o especificada de tal manera que las variables predictoras, así como los parámetros a ser estimados a partir de los datos, serán los elementos a determinar. La distribución de errores aleatorios se asume generalmente como una distribución

normal de media cero y los errores son independientes, esto es, no están correlados. Todas estas suposiciones generalmente se evalúan después de ajustar el modelo.

El modelo general de regresión

El modelo general de regresión es la extensión para k variables explicativas del modelo simple, en general, una variable respuesta y depende de muchas otras variables $x_1, \dots, x_n$ aunque alguna de estas pueden ser no observables o incluso desconocidas. El modelo de regresión intenta medir el efecto de las más importantes, el resto se representa mediante una variable aleatoria que se denomina perturbación del modelo. El modelo de regresión múltiple queda como

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + R_i \quad (9)$$

La solución de este modelo se realiza mediante el procedimiento habitual de mínimos cuadrados, $Ax=R$ donde:

$$A = \begin{pmatrix} 1 & x_{11} & \dots & x_{k1} \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_{1n} & \dots & x_{kn} \end{pmatrix}, \quad x = \begin{pmatrix} \beta_0 \\ \vdots \\ \vdots \\ \vdots \\ \beta_k \end{pmatrix}, \quad K = \begin{pmatrix} y_1 \\ \vdots \\ \vdots \\ \vdots \\ y_n \end{pmatrix}, \quad R = \begin{pmatrix} r_1 \\ \vdots \\ \vdots \\ \vdots \\ r_n \end{pmatrix}, \quad (10)$$

con

$$\hat{x} = \begin{pmatrix} \hat{\beta}_0 \\ \vdots \\ \vdots \\ \vdots \\ \hat{\beta}_k \end{pmatrix} = (A^T A)^{-1} A^T K, \quad (11)$$

los valores más probables estimados con $(A^T A)^{-1}$ no singular. Siendo la varianza a posteriori del observable de peso unidad la siguiente:

$$\hat{\sigma}^2 = \frac{RR^T}{Rango(A)}, \quad (12)$$

la estimación de la matriz varianza covarianza a posteriori de los parámetros [12] se realiza de la manera usual:

$$\Sigma(\hat{\beta}) = \hat{\sigma}^2 (A^T A)^{-1}, \quad (13)$$

y llamando q_{ij} a los elementos de la matriz $(A^T A)^{-1}$ se deduce [10] que

$$\hat{\beta}_j \sim N(\beta_j, \hat{\sigma} \sqrt{q_{jj}}). \quad (14)$$

Lo que indica que un determinado parámetro sigue una distribución normal de media y desviación típica conocidas, lo que nos lleva inmediatamente a definir los intervalos de confianza y los contrastes necesarios para los parámetros. Los intervalos de confianza para los coeficientes individuales se obtienen mediante:

$$\hat{\beta}_i \pm t(\alpha/2, \text{Rango}(A)) \hat{\sigma}_{\beta_i}, \quad (15)$$

con

$$\hat{\sigma}_{\beta_i} = \hat{\sigma} \sqrt{q_{ii}}. \quad (16)$$

Siendo t la distribución de Student con el número de grados de libertad determinado por el rango de la matriz de diseño A y con un nivel de confianza α . El contraste $E[\hat{\beta}_i] = \beta_i^*$ se reduce entonces al tener el estadístico $t = \frac{\hat{\beta}_i - \beta_i^*}{\hat{\sigma}_{\beta_i}}$ y compararlo con la distribución t .

El contraste más habitual que vamos a realizar es el de $\beta_i^* = 0$ con lo que el estadístico sigue la siguiente expresión

$$t = \frac{\hat{\beta}_i}{\hat{\sigma}_{\beta_i}}, \quad (17)$$

quedando las hipótesis de contraste para comprobar si un parámetro es cero o no, de la siguiente manera: $H_0 : \hat{\beta}_i = 0$, $H_1 : \hat{\beta}_i \neq 0$

Si se verifica la hipótesis nula, esto es, el estadístico es considerablemente grande y, si falla el test a dos colas, debería considerarse la eliminación de dicho parámetro en la regresión.

La variabilidad de la respuesta puede descomponerse de la forma siguiente [10]:

$$\sum (K_i - \bar{K}_i)^2 = \sum (\hat{K}_i - \bar{K}_i)^2 + R^T R, \quad (18)$$

ecuación que expresa la variación total (VT) como suma de la varianza explicada por los regresores (VE) y la varianza no explicada por los regresores o residual (VNE), es decir: VT=VE+VNE, en nuestro caso las dos varianzas quedarían como

$$VE = (Ax - m\bar{K})^T (Ax - m\bar{K}), \quad (19)$$

y

$$VNE = (A\hat{x} - K)^T (A\hat{x} - K), \quad (20)$$

donde m es el número de filas de la matriz de términos independientes K con $\bar{K}$ y $\hat{K}$ valores medios y estimados respectivamente de la matriz de términos independientes. El contraste de regresión comprueba que la varianza explicada es significativamente mayor que la no explicada y se define por el estadístico:

$$F = \frac{VE}{kVNE}, \quad (21)$$

que sigue una distribución F con k , $n-k-1$ grados de libertad y la hipótesis nula sería la de que todos los coeficientes son nulos. Siendo k el número de parámetros menos uno. Los grados de libertad de cada una de las varianzas se muestran en la tabla siguiente:

	Expresión	Grados de libertad
Varianza explicada por los regresores (VE)	$(Ax - m\bar{K})^T (Ax - m\bar{K})$	k
Varianza residual (VNE)	$(A\hat{x} - K)^T (A\hat{x} - K)$	$n-k-1$
Varianza total	$(K - \bar{K})^T (K - \bar{K})$	$n-1$

Tabla 1. Contrastes y descomposición de la varianza

Coeficiente de correlación y método de eliminación progresiva

Un parámetro esencial utilizado en regresión múltiple es el coeficiente de determinación o coeficiente de correlación múltiple, permite construir una medida descriptiva del ajuste global del modelo y consiste simplemente en el coeficiente entre la varianza explicada por la regresión y la varianza total

$$R^2 = \frac{(K - \bar{K})^T (K - \bar{K})}{(Ax - m\bar{K})^T (Ax - m\bar{K})}. \quad (22)$$

Las propiedades del coeficiente de correlación son las siguientes [10, 11]:

- o Su valor absoluto siempre es menor o igual que uno. Si fuera exactamente igual a uno existiría una relación funcional exacta entre las variables respuestas y las explicativas.
- o Coincide con el coeficiente de correlación lineal en el modelo simple de regresión.
- o El inverso porcentual sería la varianza no explicada, $100(1-R)^2$

El coeficiente de determinación se utiliza generalmente para comparar la eficacia de distintas regresiones, aunque no es justificado este hecho del todo dado que por una parte este coeficiente aumenta significativamente cuando se introducen nuevas variables en el modelo a pesar de que su efecto no sea significativo, por otra parte, este coeficiente depende enormemente de la formulación del modelo y qué variable dependiente se haya escogido; este coeficiente tiene sentido únicamente cuando el cociente entre el número de variables y el tamaño muestral no sea excesivamente alto. Para evitar que el coeficiente de determinación aumente significativamente al introducir nuevas variables se utiliza generalmente lo que se denomina el " coeficiente de determinación corregido por grados de libertad " cuya expresión es la siguiente:

$$\bar{R}^2 = 1 - \frac{\text{VNE}}{\sigma_K^2}, \quad (23)$$

siendo σ_K^2 la varianza del término independiente.

Con la mayoría de las herramientas de regresión múltiple descritas hasta ahora, es lógico pensar que un polinomio pueda absorber de manera más eficaz las heterogeneidades de una red clásica, como ED50. En el caso de utilizar variables reales se plantean estos dos polinomios:

$$\begin{aligned} \varphi' &= C_1 + C_2\lambda + C_3\varphi + C_4\lambda^2 + C_5\lambda\varphi + C_6\varphi^2 + C_7\lambda^3 \dots, \\ \lambda' &= C'_1 + C'_2\lambda + C'_3\varphi + C'_4\lambda^2 + C'_5\lambda\varphi + C'_6\varphi^2 + C'_7\lambda^3 \dots, \end{aligned} \quad (24)$$

o bien, se considera uno solo en variable compleja, lo que equivale a imponer a los dos anteriores la condición de conformidad

$$\lambda' + i\psi' = \sum_{j=0}^n (K_j + iK'_j)(\lambda + i\psi)^j, \quad (25)$$

donde ψ es la latitud isométrica, es necesario que el sistema de coordenadas sea conforme, podrían ser coordenadas reducidas a una proyección cartográfica o un sistema curvilíneo como el de la latitud isométrica.

La elección del polinomio óptimo es uno de los problemas más estudiados en cualquier bibliografía sobre regresión. Los métodos más comunes son: “eliminación progresiva”, “introducción progresiva” y “regresión paso a paso” [13]. El primero de ellos se basa en la regresión de la variable respuesta sobre todas aquellas que le pueden influir, reduciendo el modelo término a término mediante el estadístico t ; el segundo de los métodos es básicamente inverso al anterior; por último, en el tercero (paso a paso) se van incluyendo las variables más correladas con la respuesta, evaluando de nuevo las ya introducidas y eliminando en su caso alguna de ellas. El método de regresión “paso a paso” fue utilizado por Appelbaum (1982) para el cálculo de los polinomios de la NIMA (TR8350.2). Todos los métodos presentan ventajas e inconvenientes. En este capítulo se ha decidido el método de “eliminación progresiva” porque, si bien presenta como inconveniente disponer de gran capacidad de cálculo y en ocasiones conduzca a problemas de multicolinealidad (singularidad de la matriz de diseño), como contrapartida resulta excelente para evitar la exclusión de variables significativas [10]. Como advertencia general a cualquier método de regresión “cualquier juego de polinomios nunca deben ser empleados fuera de la zona límite de los datos”.

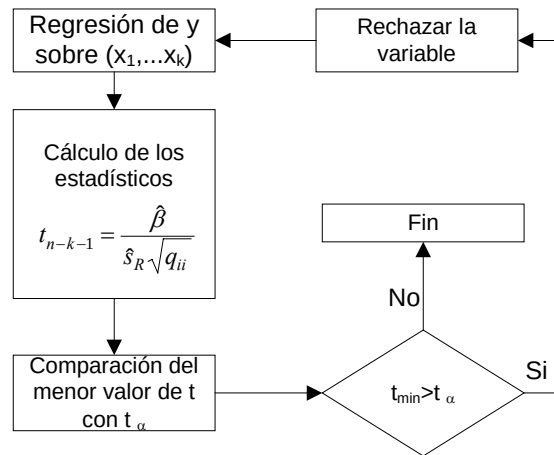


Figura 4. Eliminación progresiva

En la aplicación del método se deben tener en cuenta la mayoría de las herramientas de regresión para la elección de variables, entre ellas:

- o Coeficiente de determinación: R^2 explica el ajuste global del modelo y corresponde al cociente entre la varianza explicada por la regresión y la varianza total.
- o Coeficiente de determinación corregido: evita el efecto de que R^2 aumente ficticiamente cuando el número de variables introducidas crece.
- o Contrastes individuales de los coeficientes. $H_0: \beta=0$, permite la inclusión de variables, o su exclusión en el avance del proceso.
- o Tabla ADEVA. Analiza la varianza explicada y no explicada por el modelo de regresión.
- o Estadístico F. Evalúa los coeficientes globalmente.
- o Multicolinealidad. Verifica la correlación entre las variables explicativas, el número de condición de la matriz.
- o Estudio de los residuos. Explica problemas de falta de linealidad del modelo, variabilidad de la varianza (heterocedasticidad) y adecuación del modelo de regresión.

8.5. Métodos basados en la eliminación de la distorsión de la red

En sí mismo, el cambio de Sistema Geodésico de Referencia debería poder llevarse acabo mediante traslación, rotación y cambio de escala, es decir, mediante 7 parámetros, cuya transformación es conforme [8], sin embargo, la materialización de la realidad terreno en cada caso provoca ligeros cambios de forma a lo largo de la red, fundamentalmente provocado por el uso de métodos de medida distintos, criterios de ajuste diferentes, errores de medida etc; por tanto, este hecho no puede ser modelado mediante una simple transformación conforme [14].

Uno de los procedimientos ya ensayados por otros países consiste en los siguientes pasos esenciales:

- o Calcular la mejor transformación conforme posible entre los dos datums, esto eliminaría la diferencia entre los mismos debida únicamente al cambio de Sistema Geodésico de Referencia pero no tendría en cuenta el cambio de forma.
- o Obtener las diferencias entre el valor en el datum de llegada y el calculado a partir de esta transformación conforme. Este residuo representa la distorsión de la red.
- o Modelar la distorsión.
- o Obtener la transformación conjunta “conformidad+modelo de distorsión”

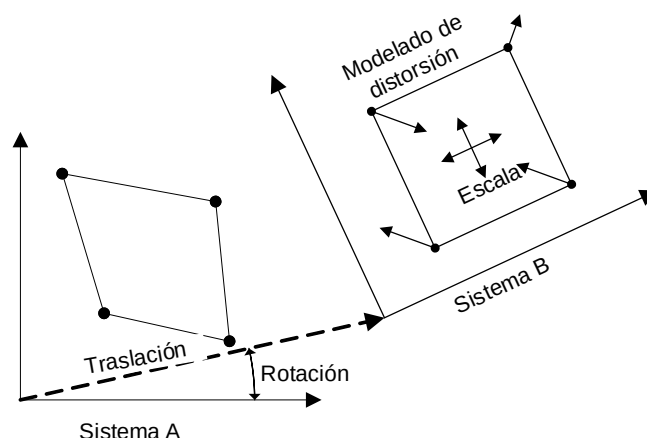


Figura 5. Fundamento del método de modelado de distorsión

Por distorsión se entiende la parte residual de la transformación teórica entre dos Sistemas Geodésicos de Referencia, entendiendo dicha parte contenida en el residuo proporcionado por una transformación de 7 parámetros de semejanza. Común a los métodos de modelado de distorsión son las tres fases siguientes: primero, el cálculo de la mejor transformación posible que asuma el cambio teórico de Sistema de Referencia; segundo, el análisis de los residuos de esta transformación para obtener una idea de los valores que han de ser absorbidos por el método; y por último, depurar los puntos anómalos.

La posibilidad de modelar la distorsión se basa en la existencia de un patrón regular de comportamiento ya que ésta no presenta aleatoriedad sino una cierta correlación espacial. La forma de modelar esta distorsión ha sido llevada a cabo por distintos autores mediante cuatro técnicas: Superficies de Mínima Curvatura (SMC), Colocación Mínimo Cuadrática (CMC), Regresión Múltiple (MRE) y Rubber-Sheeting (RS).

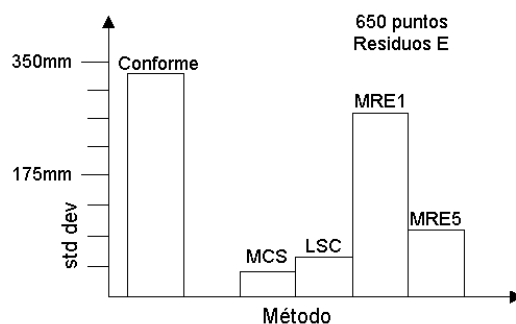


Figura 6. Métodos de modelado de distorsión en la red, resultados para la red australiana de Victoria [5]

En la Figura 6 se pone de manifiesto la comparación de ambos métodos para la red australiana de la ciudad de Victoria. Con un criterio puramente numérico, y en el caso australiano, el mejor de los métodos es el de mínima curvatura que ha servido de modelo para la transformación de NAD27 a NAD83 en EEUU. La Colocación Mínimo Cuadrática tiene similar bondad, con la ventaja de permitir estimar en cada punto el “error esperado” que se está cometiendo al aplicar esta transformación. Además, resulta más versátil si la densidad de puntos es irregular, permitiendo el recálculo de la malla sin necesidad de introducir todos los puntos de nuevo. Este método de colocación (CMC) ha sido el elegido en países como Australia o Canadá. Como se observa en la Figura 6 los métodos de regresión múltiple de grado 1 (MRE1) y grado 5 (MRE5) presentan los peores ajustes. En el caso español se cuenta con una distribución de puntos extraordinariamente regular y una dimensión territorial muy inferior a las de Australia, Canadá o EEUU.

Rubber-Sheeting

El proceso consiste en encontrar una transformación entre las dos figuras de la red en distintos datums de manera que sean topológicamente equivalentes. La forma más adecuada de identificar las figuras de ambas redes es mediante una triangulación de Delaunay y determinar un homomorfismo lineal por partes. La forma más común consiste en obtener los parámetros de una transformación afín para cada triángulo y derivar de esta descomposición del espacio una rejilla. Este método ha sido el elegido por Gran Bretaña y Suiza.

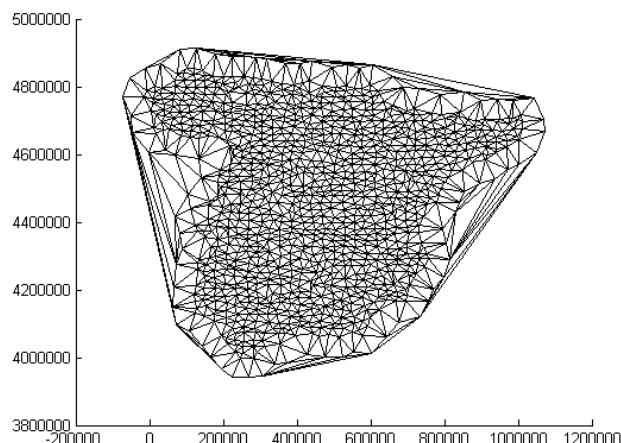


Figura 7. Triangulación Delaunay de REGENTE

En la figura 7 se puede observar la triangulación de Delaunay de REGENTE a la que se ha añadido “puntos virtuales” cuyo vector ED50-ETRS89 se ha obtenido por transformación lineal de los más cercanos. El propósito de estos “puntos virtuales” es evitar triángulos demasiado alargados cerca de la línea de la costa. Para cada triángulo así formado se ha calculado una transformación del tipo de la ecuación siguiente, no existiendo residuo alguno en esta transformación.

$$\begin{bmatrix} x_{ETRS89} \\ y_{ETRS89} \end{bmatrix} = \begin{bmatrix} \mu_x \cos \alpha_x & \mu_y \cos \alpha_y \\ \mu_x \sin \alpha_x & \mu_y \sin \alpha_y \end{bmatrix} \begin{bmatrix} x_{ED50} \\ y_{ED50} \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \quad (26)$$

Superficies de mínima curvatura

El método de mínima curvatura fue inicialmente desarrollado por Briggs [27], se basa en minimizar la curvatura total sobre una rejilla creada a partir de una distribución discreta de puntos. La idea proviene de la ingeniería mecánica y la teoría de la elasticidad. Consiste en considerar una lámina en equilibrio donde las fuerzas actúan perpendiculares a ella y no existen ni fuerzas de cizalla ni tensiones en la misma. Las fuerzas actuantes sobre esta lámina corresponderían a los puntos de la red, en concreto a la distorsión de la red en esos puntos. Formando la ecuación biarmónica que define el comportamiento de la placa plana cuya función de Airy proporciona las tensiones, en equilibrio en el contorno con las fuerzas exteriores actuantes. En el problema de la placa, suponiendo $p(x,y)$ la carga por unidad de superficie en

cada punto y $D = \frac{Eh^3}{12}$ la rigidez de la placa, la deformación de esta, w , debe satisfacer la ecuación de Lagrange o ecuación de las placas [28]

$$\frac{\partial^4 w}{\partial x^4} + \frac{\partial^4 w}{\partial y^4} + 2 \frac{\partial^4 w}{\partial x^2 \partial y^2} = -\frac{p(x,y)}{D}, \quad (27)$$

En nuestro caso particular, se conocen las deformaciones, que constituyen la distorsión en ϕ y λ de cada uno de los vértices de una hipotética red geodésica. El objetivo es determinar numéricamente la superficie, en la figura siguiente se muestra la distorsión obtenida entre ED50 y ETRS89 para REGENTE.

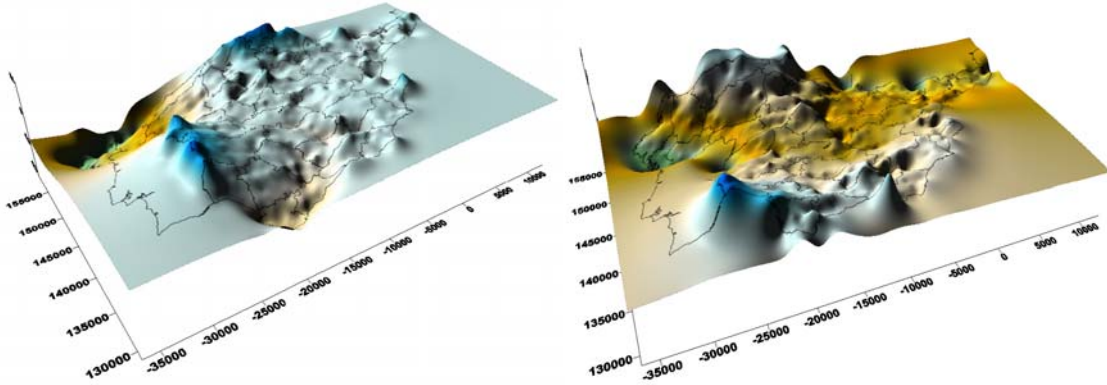


Figura 8. Superficie de Mínima Curvatura para REGENTE, (longitud a la izquierda y latitud a la derecha) [2]

El método empleado en la transformación entre NAD27 y NAD83 en Estados Unidos se basa en la descomposición de la expresión (27) por medio de diferencias finitas [29]. Este método fue inicialmente desarrollado para interpolar valores de gravedad donde su valor varía suavemente en relación al espacio [27]. Supongamos una malla regular (**¡Error! No se encuentra el origen de la referencia.**) de puntos sobre la cual se va a ir trasladando una roseta de iteración [30] que modificará la deformación de la misma, la expresión siguiente muestra el caso general donde partiendo de los valores de alrededor de un punto $A_{i+2,j}$, $A_{i,j+2}$, ..., $A_{i,j+1}$, $A_{i,j-1}$ se obtiene el valor de la deformación en el punto $A_{i,j}$. Este método se denomina de diferencias finitas. Existen otros métodos que proporcionan resultados similares, por ejemplo el de splines en tensión, aunque los resultados fuera de la zona de datos son bastante peores.

$$A_{i,j} = -[A_{i+2,j} + A_{i,j+2} + A_{i-2,j} + A_{i,j-2} \dots + 2(A_{i+1,j+1} + A_{i-1,j+1} + A_{i+1,j-1} + A_{i-1,j-1}) \dots - 8(A_{i-1,j} + A_{i-1,j} + A_{i,j+1} + A_{i,j-1})] / 20 \quad (28)$$

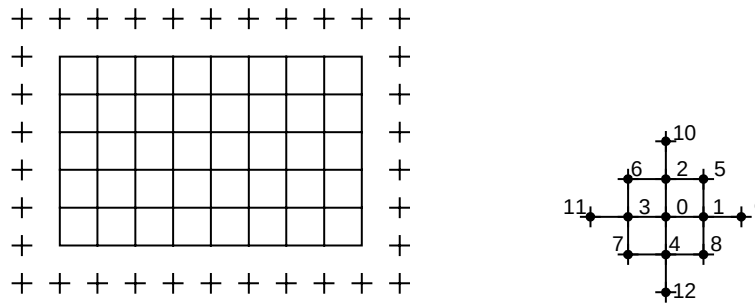


Figura 9. Malla regular a priori y roseta de iteración a la derecha

Colocación mínimo cuadrática

La colocación mínimo cuadrática (CMC) se puede definir como un método de estimación generalizado que combina ajuste, filtro y predicción. Se considera, aparte de los parámetros, dos conjuntos de variables aleatorias relacionadas a partir de sus matrices varianza-covarianza. Estas dos variables aleatorias se denominan ruido (los errores de medida) y señal. Tradicionalmente los métodos de colocación se han empleado en la interpolación de anomalías gravimétricas, en realidad, pueden utilizarse en cualquier problema clásico de mínimos cuadrados expandiendo las posibilidades de estos métodos mediante las técnicas de filtrado y predicción. Se considera la expresión habitual en mmcc expresada de la siguiente forma

$$K = Ax + s' + n, \quad (29)$$

De esta forma, se divide el vector de residuos R en una cantidad aleatoria s' que denominaremos de ahora en adelante señal, y otra, también aleatoria, denominada ruido n . Esta separación del vector residuos en dos es el denominado modelo de “Colocación Mínimo Cuadrática” incluyendo parámetros sistemáticos. Por lo tanto las medidas K dependen de una parte sistemática Ax y dos partes aleatorias s' y n .

Sea la curva mostrada en la Figura 12 como una línea continua, la cual se ha obtenido mediante una serie de valores a partir de observaciones discretas. Dichas observaciones estarán afectadas por errores de observación o ruido, denotado por n . La curva interpolada estará compuesta de una parte sistemática Ax , y una parte aleatoria s' , ambas a tener en cuenta en la teoría de la Colocación Mínimo Cuadrática.

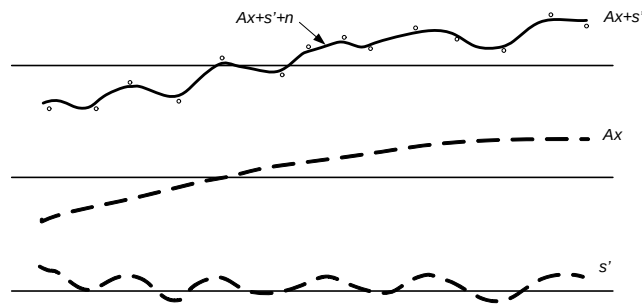


Figura 10. Ejemplo de valores de señal en punto [31]

El método de colocación consiste en la determinación de los parámetros x mediante ajuste, el filtrado y eliminación del ruido así como la predicción de los valores de la señal s en otros puntos distintos a los medidos. De aquí se deriva la definición fundamental que subyace en el método de colocación, se puede decir que el método de Colocación Mínimo Cuadrática combina ajuste, filtrado y predicción.

La teoría de predicción mediante Colocación muestra que se deben construir dos matrices covarianza. La primera de ellas entre los puntos conocidos y aquel que se desee estimar (C_{sx} , de señal desconocida) y la segunda entre los puntos conocidos, denominada $\bar{C}$.

Por señal se entienden las diferencias en latitud y longitud entre ETRS89 y ED50 menos la contribución de los 7 parámetros determinados en IV.2.4.

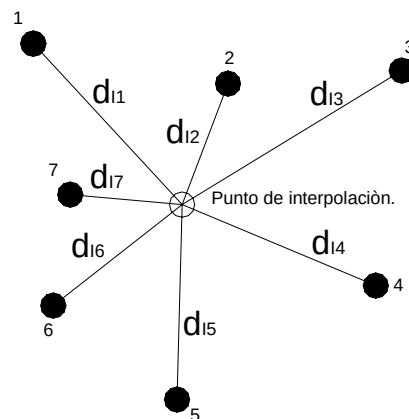


Figura 11. Función covarianza

A partir de la función covarianza experimental determinada se obtiene el vector covarianza para los puntos muestrales

$$CI = (C(d_1) \ C(d_2) \ C(d_3) \ C(d_4) \ C(d_5) \ C(d_6) \ C(d_7)). \quad (30)$$

Aplicando la misma función, a continuación se determina la matriz covarianza entre los puntos muestrales

$$CD = \begin{pmatrix} C(d_{11}) & C(d_{12}) & C(d_{13}) & C(d_{14}) & C(d_{15}) & C(d_{16}) & C(d_{17}) \\ C(d_{21}) & C(d_{22}) & C(d_{23}) & C(d_{24}) & C(d_{25}) & C(d_{26}) & C(d_{27}) \\ C(d_{31}) & C(d_{32}) & C(d_{33}) & C(d_{34}) & C(d_{35}) & C(d_{36}) & C(d_{37}) \\ C(d_{41}) & C(d_{42}) & C(d_{43}) & C(d_{44}) & C(d_{45}) & C(d_{46}) & C(d_{47}) \\ C(d_{51}) & C(d_{52}) & C(d_{53}) & C(d_{54}) & C(d_{55}) & C(d_{56}) & C(d_{57}) \\ C(d_{61}) & C(d_{62}) & C(d_{63}) & C(d_{64}) & C(d_{65}) & C(d_{66}) & C(d_{67}) \\ C(d_{71}) & C(d_{72}) & C(d_{73}) & C(d_{74}) & C(d_{75}) & C(d_{76}) & C(d_{77}) \end{pmatrix} \cdot R = \begin{pmatrix} \Delta\lambda_1 \\ \Delta\lambda_2 \\ \Delta\lambda_3 \\ \Delta\lambda_4 \\ \Delta\lambda_5 \\ \Delta\lambda_6 \\ \Delta\lambda_7 \end{pmatrix} \quad (31)$$

Por último, se construye el vector de distorsiones, considerado como el residuo R de la transformación teórica entre los dos Sistemas Geodésicos de Referencia. La ecuación básica del método de interpolación medio-cuadrática para predecir el valor de la distorsión del punto de interpolación mediante la expresión

$$R \equiv \Delta\lambda = \Delta\varphi = C_I C_D^{-1} R. \quad (32)$$

Estimación de la función covarianza para el método de colocación

Para determinar las matrices covarianzas anteriores hay que evaluar la variabilidad espacial, tanto el semivariograma [32] como la función covarianza, miden ésta sobre una variable aleatoria. Generalmente es más sencillo determinar el semivariograma y luego la covarianza.

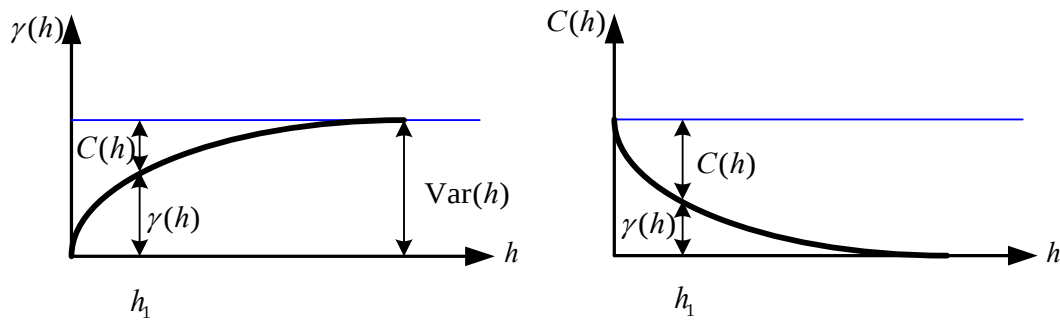


Figura 12. Figura 1 Relación entre la función covarianza y el semivariograma

Generalmente, el semivariograma es una función monótona decreciente [33], se defina como $\gamma(h)$, donde h es la distancia a x , y su expresión es

$$\gamma(h) = \frac{1}{2} E[(f(x+h) - f(x))^2], \quad (33)$$

donde $E[f(x)]$ es la esperanza matemática de $f(x)$.

El semivariograma se define como el momento de segundo orden de la función aleatoria $E[(f(x+h) - f(x))]^2$ [34]. La expresión (33) da lugar al estimador denominado semivariograma muestral o experimental [35]. Esta función se estimará más adelante.

Se puede adoptar un estimador $\gamma^*(h)$ del semivariograma como:

$$\gamma(h)^* = \frac{1}{2} \frac{1}{M(h)} E[(f(x+h) - f(x))^2], \quad (34)$$

donde $M(h)$ es el número total de parejas sobre las que se estima la función experimental. El estimador $\gamma(h)$ presenta un inconveniente, en el caso de que los datos $f(x)$ estén distribuidos de forma irregular, existirán valores de h para los que $M(h)$ será un valor muy pequeño; por lo cual la varianza de $\gamma(h)$ será muy grande. La media muestral de tamaño M tiene una varianza inversamente proporcional a M , es deseable aumentar el número de parejas $M(h)$ empleadas en la estimación de $\gamma(h)$. A tal efecto, se definen una serie de intervalos de distancia y consideramos de forma conjunta todas aquellas parejas de puntos cuyas distancias se encuentren dentro del mismo intervalo. De esta forma, se aumenta el número de parejas a costa de discretizar el semivariograma, ya que se obtiene un solo valor del semivariograma muestral para cada intervalo $\gamma(h)$ [36]. En el caso que nos ocupa, también debe definirse un ángulo de tolerancia alrededor de la dirección definida por el vector h , ver la siguiente Figura.

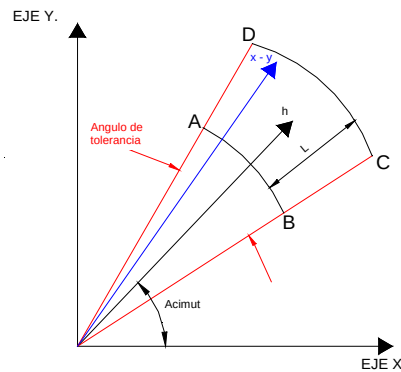


Figura 13. Representación esquemática de los intervalos y de la tolerancia para el cálculo del semivariograma muestral en dos dimensiones.

El vector separación h viene definido por su acimut y módulo de h . A lo largo de la dirección definida por h se definen una serie de intervalos de longitud y se considera una zona de tolerancia

definida entre las dos direcciones con azimut (*azimut $\pm$ ángulo de tolerancia*). De esta forma cada intervalo define ahora un cuadrilátero curvilíneo *ABCD*, para el que el semivariograma muestral se calcula tomando todas aquellas parejas de puntos cuyo vector de separación ($x - y$) cae dentro de dicho cuadrilátero. Puesto que $\gamma(h)$ es función de la distancia h y la dirección, conviene calcular $\gamma^*(h)$ en varias direcciones para comprobar si el semivariograma es isótropo o no, es decir, si depende del acimut. Con anterioridad se ha apuntado que el semivariograma y la función covarianza tienen una relación directa y ambas muestran la variabilidad espacial de los datos. Por definición se tiene que la función covarianza es [37]

$$C(h) = \text{Var}(h) - \gamma^*(h). \quad (35)$$

Una vez que se ha estimado la función covarianza, es posible ajustarla a una función convencional, por ejemplo la de Reilly, correspondiente a la expresión siguiente y evaluar de forma completa las matrices covarianza para estimar la distorsión en puntos desconocidos.

$$C(h) = C_0 \left(1 - \frac{1}{2} \frac{s^2}{d^2} \right) e^{\left(-0.5 \frac{s^2}{d^2} \right)}. \quad (36)$$

Bibliografía

- [1] McCarthy DD, Petit G. IERS Conventions (2003). IERS Technical Note No. 32. In: Centre IC, editor, 2003.
- [2] González.-Matesanz J. Aportaciones al estudio de los modelos de distorsión para el cambio de datum entre ED50 y ETRS89. Departamento de matemáticas. España: Universidad de Alcalá, 2007.
- [3] Dalda A, González.-Matesanz J. Parámetros de transformación entre sistemas de referencia. Jornadas Técnicas sobre Topografía, Cartografía, Fotogrametría, Geodesia y Teledetección. 2001. p.189.
- [4] Hofmann-Wellenhof B, Collins J, Lichtenegger H. Global Positioning System: theory and practice. New York: Springer-Verlag, 2001.
- [5] Collier P, Leahy FJ, Argeanu VS. Transition to the geocentric datum of Australia. Consultants report to the office of surveyor general. Victoria: University of Melbourne, 1996. p.91.
- [6] Heiskanen WA, Moritz H. Physical geodesy. San Francisco,: W. H. Freeman, 1967.
- [7] National Imagery and Mapping Agency. Department of Defense World Geodetic System 1984. Its definition and relationships with Local Geodetic Systems: NIMA, 2000.
- [8] Grafarend E, Knickmeyer EH, Schaffrin B. Geodätische Datumtransformationen. Z Vermess 1982;107:15.
- [9] Appelbaum LT. Geodetic datum transformation by multiple regression equations. Proceedings of the 3rd International Symposium on Satellite Doppler. Las Cruces, New Mexico: New Mexico State University, 1982. p.207.
- [10] Peña Sánchez de Rivera D. Estadística : modelos y métodos. vol.2, Modelos lineales y series temporales. Madrid: Alianza, 1999.
- [11] Draper NR, Smith H. Applied regression analysis. New York: Wiley, 1998.
- [12] Leick A. GPS satellite surveying. New York: Wiley, 1995.
- [13] Peña D. Estadística: modelos y métodos. vol.2, Modelos lineales y series temporales. Madrid: Alianza, 1999.
- [14] Collier P, Argeanu VS, Leahy FJ. Distorsion modelling and the transition to GDA94. The Australian surveyor 1998;43:29.
- [15] Burrough PA, McDonnell R, Burrough PA. Principles of geographical information systems. Oxford ; New York: Oxford University Press, 1998.
- [16] White MS, Griffin P. Piecewise Linear rubber-sheet map transformation. The American Cartographer 1985;12:123.
- [17] Saalfeld A. A fast rubber-sheeting transformacion using simplicial coordinates. The American Cartographer 1985;12:169.

- [18] Landau P, Schwartz E. Subset Warping: Rubber Sheetting with Cuts. CVGIP: Graphical Models and Image Processing 1994;56:247.
- [19] Sack JR, Urrutia J. Handbook of computational geometry. Amsterdam ; New York: Elsevier, 2000.
- [20] Greaves M, Cruddace P. The adoption of ETRS89 as the National Mapping System for GB, via a Permanent GPS Network and Definitive Transformation. EUREF Symposium. Dubrovnik, 2001.
- [21] Ordnance Survey. A Guide to coordinate systems in Great Britain. In: Ordnance Survey S, UK., editor: <http://www.gps.gov.uk/guidecontents.asp>, 1999.
- [22] Ordnance Survey. OSTN97 and OSGM91 User Guide. In: Ordnance Survey S, UK., editor: <http://www.gps.gov.uk/guidecontents.asp>, 1999.
- [23] Swisstopo. Formulas and constants for the calculation of the Swiss conformal cylindrical projection and for the transformation between coordinate systems. Federal Office of Topography. Federal Directorate of Cadastral Surveying, 2001.
- [24] Ollikainen M, Mäkinen J, Jokela J, Chen R. National report of Finland. EUREF 2004. Bratislava, 2004.
- [25] Jäger R, Schneid S. GPS-Höhenbestimmung mittels Digitaler Finite-Elemente Höhenbezugsfläche (DFHBF) - das Online-Konzept für DGPS-Positionierungsdienste. XI. Internationalen Geodätischen Woche Obergurgl, vol. 19. Österreich: Institut für Geodäsie der Innsbruck, 2001. p.195.
- [26] Yang QH, Snyder JP, Tobler WR. Map projection transformation : principles and applications. London ; New York: Taylor & Francis, 2000.
- [27] Briggs IC. Machine contouring using minimum curvature. Geophysics 1974:34.
- [28] Love AEH. A treatise on the mathematical theory of elasticity. New York: Dover Publications, 1944.
- [29] Dewhurst WT. NADCON : the application of minimum curvature-derived surfaces in the transformation of positional data from the North American Datum of 1927 to the North American Datum of 1983. Rockville, MD: U.S. Dept. of Commerce National Oceanic and Atmospheric Administration National Ocean Service Charting and Geodetic Services, 1990.
- [30] García de Arango A. Elasticidad teorica: Instituto Geográfico Nacional 1974.
- [31] Moritz H. Least-squares collocation. München,: Bayerische Akademie der Wissenschaften, Beck in Kommission, 1973.
- [32] Olea RA. Geostatistics for engineers and earth scientists. Boston: Kluwer Academic, 1999.
- [33] Chiles J-P, Delfiner P. Geostatistics: modeling spatial uncertainty. New York: Wiley, 1999.

- [34] Isaaks EH, Srivastava RM. Applied geostatistics. New York: Oxford University Press, 1989.
- [35] Clark I. Practical geostatistics. London: Applied Science Publishers, 1979.
- [36] Goovaerts P. Geostatistical incorporation of spatial coordinates into supervised classification of hyperspectral data. Journal of Geographical Systems 2002;4:99.
- [37] Samper Calvete FJ, Carrera Ramírez J, Centro Internacional de Métodos Numéricos en Ingeniería. Geoestadística : aplicaciones a la hidrogeología subterránea. Barcelona: Centro Internacional de Métodos Numéricos en Ingeniería, 1996.

Tema 9. Sistemas de Referencia Celestes. Sistemas de Referencia Geocéntricos. ITRS, ETRS, ETRS89. El IERS. Marcos. Transformaciones entre Sistemas Geocéntricos Terrestres

9.1. Introducción

Un Sistema de Referencia consiste en un conjunto de modelos necesarios para la descripción cuantitativa de posiciones y movimientos de cuerpos celestes, incluida la Tierra, (sistemas celestes), o de cuerpos sobre la Tierra (sistemas terrestres). Deben definirse: origen, escala, orientación y plano principal. Un marco asociado a un Sistema de Referencia es una serie de puntos fiduciaros que, junto con sus coordenadas, constituyen la realización práctica del sistema de referencia.

Ejemplos:

- catálogo de posiciones precisas (y movimientos si cabe) de estrellas u objetos extragalácticos vistos desde el baricentro del sistema solar en una época determinada (se toma J2000.0).
- coordenadas de un conjunto de estaciones terrestres, y sus variaciones en el tiempo, con relación a un sistema coordinado tridimensional.

9.2. Sistemas de Referencia Celestes

Un Sistema de Referencia ("Reference System") [1] es una estructura geométrica para referir las coordenadas de puntos del espacio, queda definido por la situación del origen, las direcciones de los ejes, la escala, los algoritmos necesarios para sus transformaciones espaciales y temporales y las constantes utilizadas en las definiciones y correcciones. Un Marco de Referencia ("Reference Frame") es la materialización de un Sistema de Referencia, es decir, el conjunto de elementos que determinan de forma práctica un sistema de referencia y está constituido por las coordenadas de los puntos de definición, las técnicas aplicadas en las observaciones o medidas y los métodos de cálculo aplicados para la obtención de los parámetros.

En astronomía, geodesia y geodinámica son necesarios dos sistemas de referencia fundamentales (y sus correspondientes marcos). Uno que se considera fijo en el espacio y otro que se considera fijo en la Tierra. Al primero se le llama Sistema de Referencia Celeste (CRS) y al segundo Sistema de Referencia Terrestre (TRS). A los marcos correspondientes se les llama Marco de Referencia Celeste (CRF) y Marco de Referencia Terrestre (TRF). Esta nomenclatura,

ha sido adoptada por la Asociación Internacional de Geodesia (AIG) y por la Unión Astronómica Internacional y será utilizada en lo que sigue.

En estos sistemas de referencia se estudian y determinan las posiciones y velocidades de puntos del espacio y de la Tierra respectivamente. El CRS se considera como un sistema casiinercial y en él se estudian los movimientos de los cuerpos celestes y en el geocéntrico el de los satélites artificiales de uso geodésico, como los de la constelación NAVSTAR del GPS. El TRS se utilizará para fijar las posiciones, coordenadas, de cualquier punto de la Tierra y para estudiar y determinar las velocidades de estos puntos que, al considerar la Tierra como un cuerpo deformable, estarán en continuo movimiento, aunque eso si, de pequeña magnitud. Será de gran importancia establecer una formulación correcta que nos relacione ambos sistemas de referencia. La rotación de la Tierra viene dada por la orientación del TRS en el CRS.

Sistemas de Referencia Celestes Baricéntrico y Geocéntrico

La Resolución A4 de la XXI Asamblea General de la UAI de 1991 definió dos sistemas celestes de referencia con coordenadas espacio-temporales en el contexto de la relatividad general, uno para el sistema solar (ahora llamado Sistema de Referencia Celeste Baricéntrico, BCRS), con origen en el baricentro del sistema solar y otro para la Tierra (ahora llamado Sistema de Referencia Celeste Geocéntrico, GCRS), con origen en el centro de masas de la Tierra. La propiedad más importante de estos sistemas es que las coordenadas espaciales de puntos a ellos referidas no sufren rotaciones globales con respecto a objetos alejados del universo, es decir las direcciones de los ejes se mantienen fijas respecto a ellos. El Comité Conjunto BIPM/IAU. sobre Relatividad en Sistemas de Referencia espaciotemporales y Metrología dejó claro que debido a las altas precisiones alcanzadas en la medida del tiempo por relojes atómicos era necesaria una ampliación de las Recomendaciones de la UAI de 1991.

La Resolución BI.3 de la XXIV Asamblea General de la UAI de 2000 dice que se extiendan las definiciones de los sistemas básicos fijos en el espacio BCRS y GCRS en el contexto de la relatividad general por medio del uso de coordenadas armónicas en ambos; se dan las formas propuestas de los tensores métricos y las fórmulas para la transformación de coordenadas post-Newtoniana entre BCRS y GCRS, transformación de Lorentz generalizada que contiene la aceleración del geocentro y potencial gravitatorio. Todo esto es necesario por la gran precisión alcanzada por las observaciones. El clásico sistema celeste casi inercial Newtoniano se sustituye por dos: el BCRS y el GCRS. El BCRS debe usarse como sistema de referencia básico para astrometría y efemérides en el sistema solar. Como es un sistema baricéntrico las estrellas no están sujetas a paralajes y sus coordenadas sólo varían por movimiento propio. El tiempo junto a las coordenadas sólo se usa para dar la época de la observación, pero no hay tiempo o época asociada al propio sistema BCRS porque está definido cinemáticamente. Puede ser asociado con

la esfera celeste de los cuásares (objetos remotos). Las direcciones baricéntricas de las radiofuentes extragalácticas no muestran rotación global respecto de estos objetos, se les supone con movimientos propios despreciables. Ahora bien, el movimiento del baricentro del sistema solar en su órbita alrededor del centro de la galaxia es no lineal, por tanto existe una precesión geodésica galáctica pero este no está incluida en la definición del BCRS

El GCRS es un sistema local en movimiento acelerado (el del geocentro) alrededor del baricentro del sistema solar pero manteniendo fijas las direcciones de los ejes con respecto a radiofuentes extragalácticas, su orientación sigue la condición cinemática de ausencia de rotación global de las direcciones geocéntricas de las radiofuentes; en él no puede describirse la esfera celeste. A este sistema GCRS se refieren las observaciones hechas desde la Tierra y a él deben referirse los procesos dinámicos en la Tierra y su entorno así como los parámetros de rotación de la Tierra (EOP), pues es en este sistema geocéntrico no rotante en el que se mide la rotación de la Tierra y no depende del movimiento de la Tierra. En particular en este sistema deben describirse los movimientos de los satélites artificiales y de la Luna. Debe indicarse que si se efectúan correcciones de precesión-nutación y se usa el modelo IAU2000 deben quitarse los términos de precesión y nutación geodésica que este modelo lleva incluidos. Este GCRS reemplaza los conceptos clásicos tales como eclíptica, equinoccio, etc.

Sistema de Referencia Celeste Internacional ICRS

El Sistema de Referencia Celeste Internacional (ICRS) se define como un sistema de referencia celeste ecuatorial con origen en el baricentro del sistema solar y con los ejes en direcciones fijas respecto a radiofuentes extragalácticas principalmente cuásares, está basado en una definición cinemática. No es otra cosa que el BCRS. Va acompañado de las teorías de precesión-nutación y rotación de la Tierra con sus correspondientes constantes, coeficientes y amplitudes. También se necesita la velocidad de la luz ($c = 299792458 \text{ ms}^{-1}$) y el sistema y la escala de tiempo. Queda materializado por el Marco de Referencia Celeste Internacional (ICRF). En este sistema las coordenadas esféricas de un objeto celeste se denominan simplemente ascensión recta y declinación sin calificativos. Se trata pues de un sistema de referencia casi inercial ecuatorial en el que se establecen las teorías del movimiento en el entorno del sistema solar y en el que se tienen las coordenadas de las estrellas y objetos espaciales. Siguiendo la Resolución A4 de la UAI de 1991 podemos decir de forma esquemática que el ICRS tiene su origen 0 en el baricentro del sistema solar, sus ejes (OX,OY,OZ) en direcciones fijas respecto a quasares. Su plano principal lo más próximo posible al ecuador medio J2000.0, (eje OZ perpendicular a este plano dirigido hacia el polo CO), la dirección del eje OX (origen de coordenadas) lo más

próxima posible al equinoccio dinámico medio J2000.0 (hacia el punto Σ_0 en el ecuador de CO) y el eje OY formando triedro directo con los anteriores.

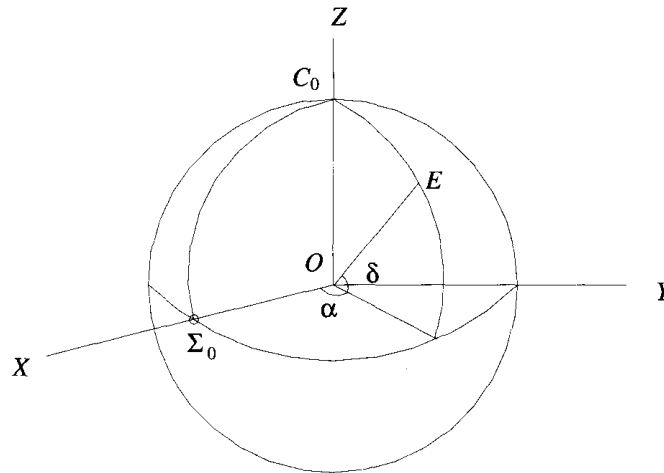


Figura 1.

El ICRS debe mantener fijas las direcciones de los ejes a lo largo del tiempo. El Servicio Internacional de Rotación de la Tierra y Sistemas de Referencia (IERS) viene determinando esta referencia cada año desde 1989 a 1995. El número de radiofuentes que definen la referencia ha ido creciendo desde 23 en 1989 a 212 en 1995, estas determinaciones llegan a converger con una precisión de 0.01 milisegundos de arco (mas). La versión propuesta por el IERS en 1995 es la que se toma como ICRS; según la Resolución B2 de la Asamblea General UAI de 1997. Oficialmente reemplaza al FK5 a partir del 1 de enero de 1998, y se ha determinado que es consistente con él con una precisión de ± 50 milisegundos de arco (mas), la precisión del FK5. Para fijar el origen de ascensiones rectas del ICRS lo más próximo posible al equinoccio dinámico J2000.0, el Σ_0 del eje X del sistema del IERS fue definido adoptando la media de las ascensiones rectas de 23 radiofuentes que se calcularon fijando la ascensión recta del cuasar 3C273B (Algol en Perseo) en su valor del catálogo FK5, es decir en 12h 29m 6.6997s en J2000.0

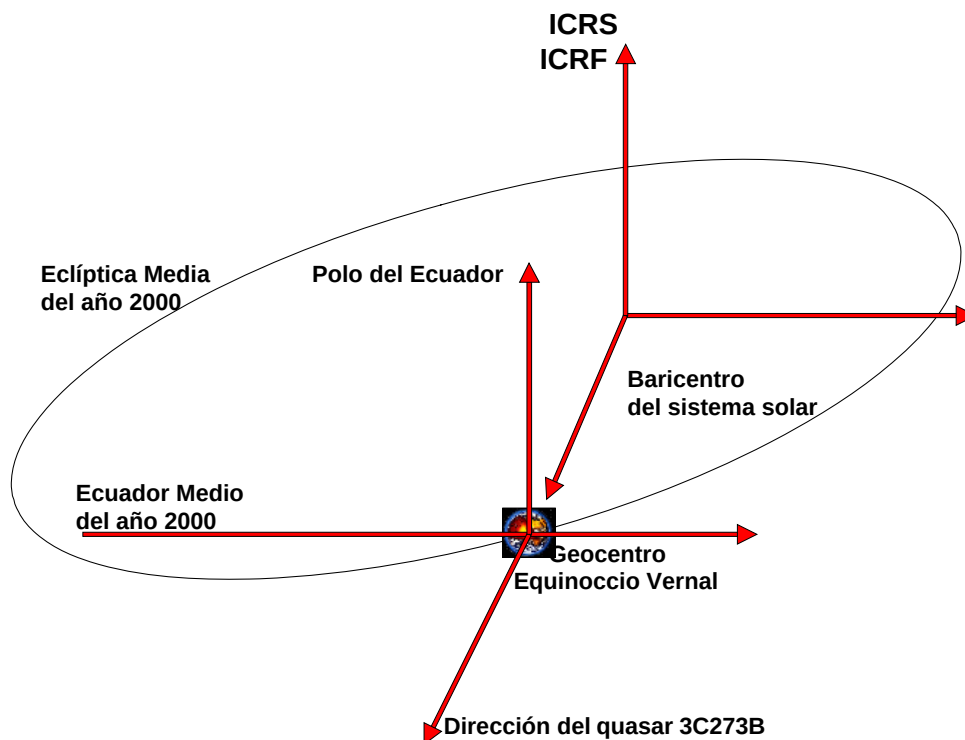


Figura 2.

Los datos de observación revelan que el polo Co del ICRS está desplazado respecto al polo celeste medio C2000.0 en J2000.0 en +16.617 mas en la dirección 12h y +6.819 mas en la dirección 18h (con la teoría de nutación MBH2000.0). El análisis de datos láser a la Luna (LLR) indica que el origen Σ_O , dirección del eje X del ICRS, está desviado del equinoccio medio inercial $\gamma_{2000.0}$ en J2000.0 en -55.4 mas en el plano de referencia del ICRS (correspondiente a -4.6 en el ecuador medio J2000.0) y esto está de acuerdo con las exigencias de la resolución UAI de 2000. La desviación entre el origen del ICRS y el equinoccio medio dinámico de las efemérides DE200 del Jet Propulsion Laboratory (JPL) es de 78 mas. El equinoccio de las nuevas efemérides planetarias DE405 y lunares LE405, recomendadas por el IERS, ha sido orientado hacia el ICRS con una precisión de 1 mas. También se ha determinado que el equinoccio del FK5 está desviado -22.9 mas respecto del lo del ICRS.

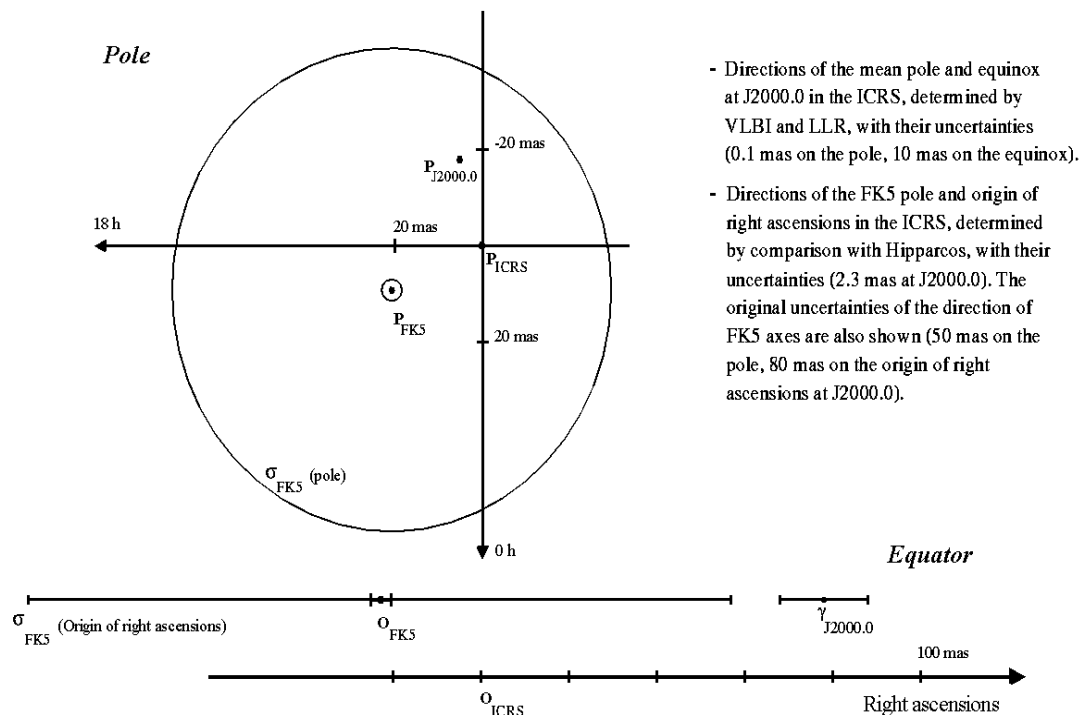


Figura 3. Desplazamiento de los polos respecto al ICRS.

http://www.iers.org/iers/earth/icrs/axes_icrs.gif

En definitiva, el Sistema de Referencia Celeste Internacional es un sistema fijo, cinemático, independiente de la época, independiente de la dinámica del sistema solar, no depende ni de la rotación del polo de la Tierra ni de la del polo de la eclíptica, los orígenes de coordenadas son estables y no precesionan a lo largo del ecuador en movimiento, es pues independiente el modelo de precesión-nutación adoptado.

9.3. Sistemas de Referencia Geocéntricos

Este sistema de referencia se asocia con un triedro que rota con la Tierra en su movimiento diurno, y respecto del cual las posiciones de los puntos de la superficie terrestre sólo se ven modificadas por los fenómenos geofísicos. Su origen O se sitúa en el centro de masas del planeta considerando océanos y atmósfera, y el eje Oz se dirige hacia el polo R_0 , punto situado en el eje medio de la figura terrestre (eje medio de Tisserand de la Tierra), que inicialmente se hace coincidir con el Polo Medio de Rotación (PMR).

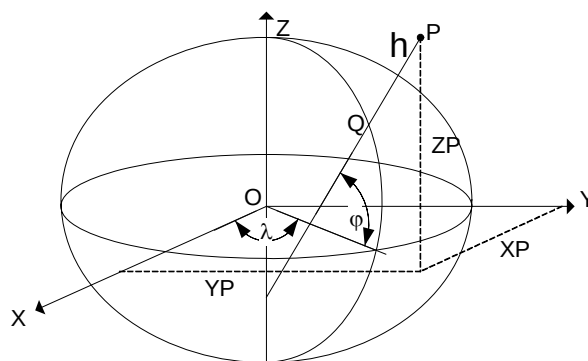


Figura 4. Relación entre el Sistema Geocéntrico y el Geodésico

9.4. ITRS (Internacional Terrestrial Reference System)

Un Sistema de Referencia Terrestre (TRS) es un sistema de referencia espacial que rota de manera solidaria con la Tierra en su movimiento diurno en el espacio. En este sistema, las posiciones de los puntos están ancladas a la superficie sólida de la Tierra, estas posiciones pueden tener pequeñas variaciones en el tiempo debidas a efectos geofísicos, como por ejemplo deformaciones tectónicas o cargas de marea. Un marco de referencia terrestre (TRF) es un conjunto de puntos físicos en los que se ha determinado de forma muy precisa sus coordenadas en un TRS específico que puede ser cartesiano, geodésico o proyectado; se dice que un TRF es una realización de un TRS.

Un sistema de referencia terrestre ideal se define como un marco tridimensional de referencia cercano a la Tierra que rota de manera solidaria con ella. Los sistemas geocéntricos se escogen de manera que su origen sea lo más cercano posible al centro de masas de la Tierra, y su orientación sea ecuatorial, es decir, que su eje Z sea paralelo al polo. Sobre esta hipótesis, el modelo general de transformación entre un conjunto de coordenadas cartesianas de cualquier punto cercano a la Tierra de un TRS a otro se determinará por medio de una transformación de semejanza de siete parámetros [2].

El IERS (International Earth Rotation Service) es el órgano encargado de definir, realizar y promover el International Terrestrial Reference System (ITRS) según se definió en la resolución 2 de la IUGG en Viena en 1991. El IERS proporciona la definición del CTRS (Conventional Terrestrial Reference System) sobre la base del ITRS. Las siguientes definiciones se aplican al CTRS:

- Se define a partir de un sistema geocéntrico sin rotación que permita ser considerado como un sistema quasi-cartesiano
- Es idéntico al definido por las resoluciones de la IAU (International Astronomy Union) y denominado Sistema de Referencia Geocéntrico (GRS)
- El sistema de tiempo será el Tiempo Geocéntrico Coordinado (TCG)
- El origen del sistema será el geocentro de la Tierra incluyendo atmósfera y océanos
- El sistema no tiene ninguna rotación residual con respecto a los movimientos horizontales de la corteza terrestre.

De esta manera, el ITRS con las condiciones anteriores cumplirá:

- Es geocéntrico, el centro de masas está definido por toda la Tierra incluyendo océanos y atmósfera
- La unidad de longitud es el metro y esta escala es consistente con el TGC mediante las adecuadas modelizaciones relativistas
- Su orientación fue proporcionada de forma inicial por el BIH (Bureau International de l'Heure) en la época 1984.0
- Su evolución en el tiempo está garantizada mediante el uso de la condición de no rotación con respecto a cualquier movimiento tectónico sobre la superficie de la Tierra

Las realizaciones de los ITRS las genera el “IERS ITRS Product Center” (ITRS-PC) bajo el nombre de Marcos de Referencia Internacionales (ITRF). El procedimiento consiste en combinar todas las soluciones individuales calculadas en los centros de análisis del IERS utilizando observaciones procedentes de geodesia espacial como SLR, VLBI, LLR, GPS y DORIS. Estas soluciones individuales contienen las posiciones y las velocidades de las estaciones junto con la matriz varianza-covarianza de la solución en formato SINEX (Solution, Software/Technique, INdependent Exchange Format). El modelo que se utiliza en la combinación se basa principalmente en fórmulas de transformación de semejanza espacial de 7 parámetros y hace uso de los “local ties” o excentricidades entre los elementos activos de varias técnicas (por ejemplo, entre la antena GPS y la VLBI). Estas excentricidades se introducen en el cálculo como una observación más con sus varianzas. Actualmente, las soluciones están disponibles casi de forma anual; se denotan como ITRF_{yy}, donde “yy” indica que la solución integra todos los datos disponibles hasta el año yy-1 [3].

El IERS comienza a generar productos en 1964, el primero de ellos se denominó BTS84 y fue determinado mediante las coordenadas de las estaciones de sistemas como VLBI, LLR, SLR y

Doppler/Transit [4]. El sistema BTS84 (Bureau International de l'Heure Terrestrial System 1984) fue realizado en el marco de las actividades del BIH como coordinador del proyecto internacional MERIT (Monitoring of Earth Rotation And Inter-comparison of Techniques), al cual sucedieron otras realizaciones siendo la última la denominada BTS87 cuando en 1988 fue creado el IERS por la Asociación Internacional de Geodesia y Geofísica y la Unión Astronómica Internacional (IAU). Hasta el momento existen 11 versiones de ITRF, comenzando con ITRF88 y acabando en la última realización denominada ITRF2005, esta última solución mejora ampliamente la combinación de distintas soluciones procedentes de las distintas técnicas geodésicas empleadas.

Una explicación detallada de los distintos sistemas, sus orígenes y alineaciones se puede encontrar en McCarthy y Petit [2]. ITRF2000 ha intentado desde sus comienzos ser un estándar para cualquier tipo de georeferenciación en las Ciencias de la Tierra, además de las estaciones fiduciarias observadas mediante técnicas VLBI, LLR, SLR, GPS y DORIS, ITRF2000 ha sido densificado mediante redes regionales en Alaska, la región antártica, Asia, Europa, América y el Pacífico, es de esperar que el reciente ITRF2005 tenga aún más éxito que su predecesor ITRF2000.

9.5. ETRS89 (European Terrestrial Reference System 1989)

EUREF (European REference Frame) es el nombre de la Subcomisión para Europa correspondiente a la comisión X de la IAG (Asociación Internacional de Geodesia), creada en 1987 como sucesora de RETRIG. El trabajo de esta comisión está orientado a las redes de control vertical, horizontal, nacionales o continentales, o globales mediante técnicas espaciales así como sus conexiones y evoluciones. Esta comisión X tiene dos subdivisiones: la subcomisión para grandes áreas geográficas y los grupos de trabajo para determinados aspectos técnicos [5].

EUREF nace como consecuencia de la utilidad de las técnicas geodésicas espaciales, en particular el GPS, para el mantenimiento de los marcos geodésicos de referencia globales y continentales, es por ello que su objetivo principal es dotar de un marco de referencia continental a Europa. La Comisión X tuvo su origen en la Asamblea General de la IUGG (International Union of Geodesy and Geophysics) en Vancouver en 1987 para continuar las labores de las Subcomisiones RETrig y REUN creadas por la IAG a mediados del siglo XX, que fueron responsables entre otras del “European Datum” basado en técnicas terrestres y el ajuste de las redes europeas de nivelación. Los objetivos son los siguientes [6]:

- Realizar un marco de referencia geocéntrico para cualquier proyecto geodésico o geodinámico sobre la placa europea
- Ser una referencia precisa muy cercana a WGS84 para ser usada no solo en geodesia sino en cualquier tipo de navegación en Europa
- Ser una referencia continental para conjuntos de datos cartográficos

Ni ED50 ni ED87 podían cumplir estos requerimientos, especialmente en cuanto a la posición y orientación tridimensional. Por otra parte, WGS84 tampoco garantizaba la precisión requerida, debido principalmente a que en sus inicios era derivado de observaciones Doppler.

El marco de referencia establecido y mantenido por EUREF es la infraestructura geodésica para la precisa georeferenciación en términos de posición tridimensional y posicionamiento dependiente del tiempo, geodinámica, navegación precisa y geoinformación. EUREF colabora estrechamente con EuroGeographics (la unión de CERCO y MEGRIN), así como la NMCA (Agencias Cartográficas y Catastrales Nacionales).

La mayoría de las actividades de EUREF han sido el establecimiento y el mantenimiento de ETRS89 (European Terrestrial Reference System) y EVRS (European Vertical Reference System). ETRS89 proporciona un posicionamiento tridimensional con precisiones de milímetro [5] mientras que EVRS tiene la misma funcionalidad para las alturas. Estos dos sistemas constituyen la base de la georeferenciación en Europa y han sido recomendadas a la Unión Europea, Eurocontrol y Eurogeographics; más información se puede encontrar en Agria [7].

Los cambios en la estructura del IAG acontecidos en la Asamblea General 2003 en Sapporo, provocaron que EUREF fue integrada en la Subcomisión 1.3 denominada “Regional Reference Frames” de la Comisión 1 “Reference Frames” bajo el nombre de “Reference Frame Subcommittee for Europe”.

El instrumento clave para el mantenimiento de ETRS89 es la Red de Estaciones Permanentes (EPN) que cubre toda Europa [6]. Las estaciones EPN proporcionan datos GPS y GLONASS para garantizar la estabilidad a largo plazo de ETRS89.

De la misma forma, EPN contribuye al Servicio Internacional GNSS (IGS) para la realización del “International Terrestrial Reference System” (ITRS), para monitorizar las deformaciones de

naturaleza tectónica en Europa, cambio climático y el desarrollo de los estándares, así como proporcionar los medios operacionales para la distribución de sus productos a través de Internet.

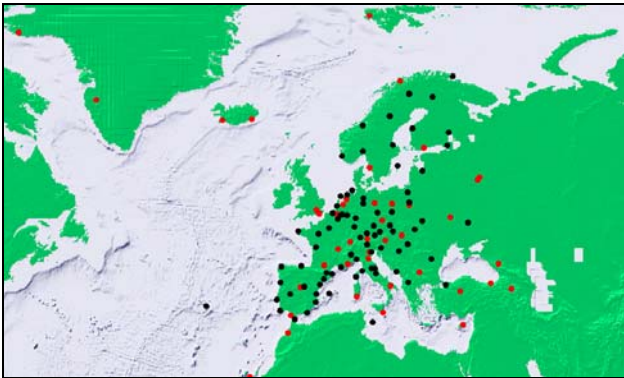


Figura 5. Red EPN

El “Spatial Reference Workshop” en París [8] recomendó a la Comisión Europea la adopción de ETRS89 como el datum geodésico que expresara y almacenara las posiciones en forma de coordenadas elipsoidales o geodésicas sobre el elipsoide GRS80 [9] así como el uso de tres proyecciones cartográficas para su representación plana (ver **¡Error! No se encuentra el origen de la referencia.**). Estas proyecciones son la Universal Transversa de Mercator, Lambert Azimutal Equiárea y Lambert Cónica Conforme. Este mismo “Workshop” recomendaba el uso de la Red Vertical Europea Unificada (EUVN) y la Red Europea Unificada de nivelación (EULN) como datum vertical. Estas recomendaciones venían seguidas de otras en las que se manifestaba la necesidad de que las Agencias Cartográficas Nacionales proporcionaran transformaciones entre los sistemas locales y ETRS89 con precisiones cercanas a los dos metros e instaba al grupo de trabajo técnico de EUREF a proporcionar dichos datos en Noviembre de 2000 para EUROSTAT, la oficina encargada de proporcionar información estadística a la Comisión Europea.

ETRS89	Sistema de Referencia Elipsoidal,	Para el almacenamiento de posiciones geográficas, en general latitud, longitud y altura elipsoidal basadas en SGR80
ETRS-TMzn	ETRS Transversa Mercator huso (zona) ‘zn’	Para mapas topográficos a escalas menores o iguales a 1:500.000, similar a UTM para el hemisferio norte
ETRS-LCC	ETRS Cónica Conforme de Lambert	Para mapas topográficos de escala mayor a 1:500.000 Proyección de Lambert con 2 paralelos automecoicos (estándar)
ETRS-LAEA	ETRS89 Azimutal Equivalente de Lambert	Para fines estadísticos y aquellas aplicaciones que requieran una proyección equivalente

Tabla 1. Proyecciones y Sistema Geodésico de Referencia recomendados para Europa [8]

ETRS89 está reconocido como el datum más apropiado para Europa por la comunidad científica. Está definido con una precisión de 1cm y es consistente con ITRS. El IERS (International Earth Rotation Service) es el encargado de realizar y proporcionar el International Terrestrial Reference System (ITRS) según se definió en la resolución nº 2 de la UIGG en Viena en 1991, posterior a la resolución EUREF90 en Florencia donde se dictó la recomendación siguiente:

The IAG Subcommision for the European Reference Frame (EUREF) , following its Resolution 1 adopted in Firenze meeting in 1990, recommends that the terrestrial reference system to be adopted by EUREF will be coincident with ITRS at the epoch 1989.0 and fixed to the stable part of the Eurasian Plate. It will be named European Terrestrial Reference System 89 (ETRS89).

La Subcomisión de la AIG para el Marco de Referencia Europeo (EUREF) de acuerdo a la Resolución I adoptada en la reunión de Florencia en 1990, recomienda que el sistema de referencia terrestre adoptado por EUREF será coincidente con ITRS en la época 1989.0 y será fijo a la parte estable de la placa euroasiática. Se denominará Sistema de Referencia Terrestre Europeo 89 (ETRS89)

La figura 6 muestra el campo de velocidades a partir de los resultados del IERS, en ella se aprecia que la placa continental europea se está moviendo a razón de unos 3cm por año en dirección noreste con la excepción de Grecia y Turquía. Este movimiento justifica que con el objetivo de tener un conjunto de coordenadas estables para Europa, el Sistema de Referencia deba estar anclado a la placa europea. Este SGR anclado a la placa europea se denomina ETRS89 por coincidir con ITRS en 1989. Desde 1989 las coordenadas se han fijado a la placa europea a partir de las realizaciones posteriores de otros ITRS, la facilidad de relación entre ETRS89 y los ITRS se debe a que la deriva en esta placa es conocida con suficiente precisión para ello, de esta forma las transformaciones entre estos SGR geocéntricos se lleva a cabo al menos con 1cm de precisión.

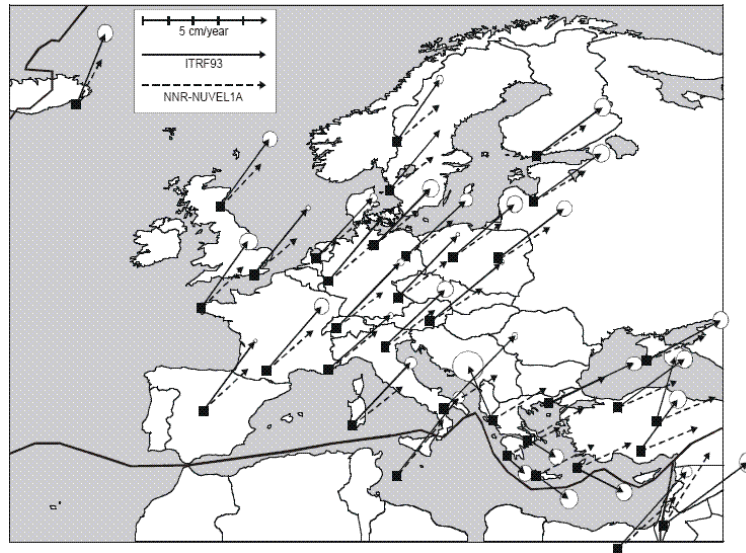


Figura 6. Velocidades para Europa en ITRF93 y con NUVEL1A

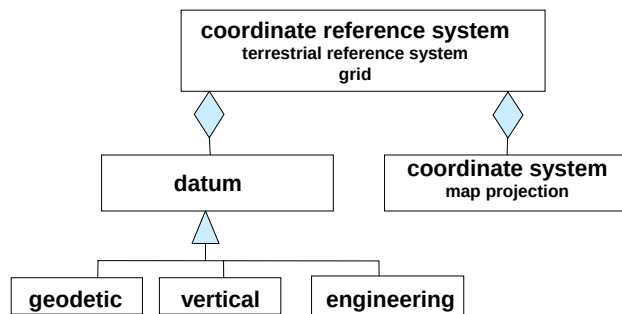


Figura 7. Descripción UML (Unified Modelling Language) de un CRS [10]

En la figura 7 tiene la descripción mediante lenguaje UML [11] de un Sistema de Referencia por Coordenadas, acorde a la ISO19111. Siguiendo la notación UML un CRS estaría compuesto de un datum y un Sistema de Coordenadas bajo una proyección determinada (relación de composición). Por su parte, los sistemas geodésico, vertical y otros, como por ejemplo un sistema usado en ingeniería, serían casos particulares de un datum (relación de herencia). Las relaciones entre ETRS, EVRS y los casos particulares de ETRS89 y datums verticales se encuentran descritos en la figura 8.

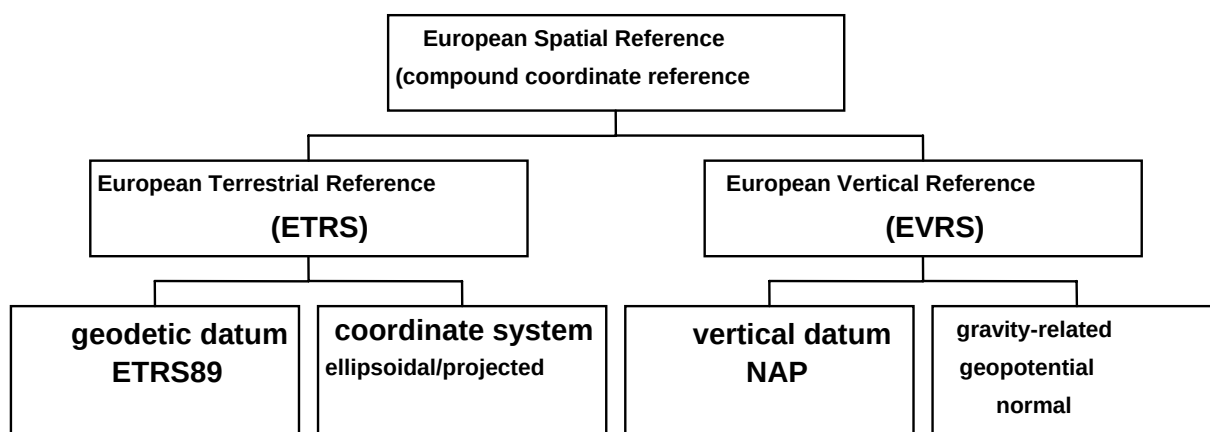


Figura 8. Componentes del Sistema de Referencia Europeo [10]

Acorde a la ISO19111 la descripción formal de los sistemas elipsoidales y geocéntricos de ETRS89 se pueden ver en el Anexo VI (en soporte informático).

La Subcomisión de EUREF en 1995 decidió coordinar todas las actividades de EPN para el mantenimiento del ETRF mediante un proceso distribuido. El esquema de proceso [12] consta de 15 centros de análisis encargados de analizar una parte de la red EPN, uno de ellos, el BKG (Bundesamt für Kartographie und Geodäsie) es el encargado de combinar todas las soluciones individuales en una única solución, esta solución es enviada al IGS. Están disponibles soluciones de geometría libre desde 1996 con valores de error medio cuadrático de unos 2mm para las componentes horizontales y 6mm para las verticales. Entre estas soluciones se garantiza una calidad similar en cada una de ellas y en la repetibilidad semanal de las mismas. A partir de 1997 se comenzó a enviar estas soluciones al IERS para formar parte de los sucesivos ITRS, de hecho en la resolución número 1 del Symposium de EUREF 1990 en Florencia fue cuando se decidió que ETRS89 sería anclado a la parte estable de Europa y sería coincidente con ITRS89. En realidad ETRS89 es un subconjunto de ITRS89 [6], no ha sido obligatorio para la Comisión Europea hasta 2003 [13].

En mayo de 1995 se formó el grupo de trabajo de EPN bajo una guía de proceso [14] de las estaciones de la red EPN con el objeto de aportar al IGS soluciones regionales de proceso [15] que contribuyeran a los futuros ITRS, aceptándose oficialmente esta propuesta en mayo de 1996.

Una de las ventajas del proceso distribuido es el enriquecimiento por parte de los técnicos de distintos países que tienen un LAC, “Local Analysis Center”. Los resultados de estos LAC son distribuidos en formato SINEX (Solution-Software/technique INdependent EXchange Format)

[15] y son combinados mediante el software Bernesse [16] utilizando la técnica de ajuste secuencial mínimo cuadrática. Debido a que la técnica de combinación tiene en cuenta dos veces el punto que haya sido procesado por dos LAC se realizan dos combinaciones: utilizando una solución de red libre y utilizando las estaciones fiduciarias. En la solución libre se define en si mismo el datum especificando 6 constreñimientos (3 para las traslaciones y 3 para las rotaciones) para algunas estaciones. El objetivo de esta solución es identificar problemas en la combinación de la segunda solución. La solución fiduciaria es la oficial de EUREF para la definición del marco.

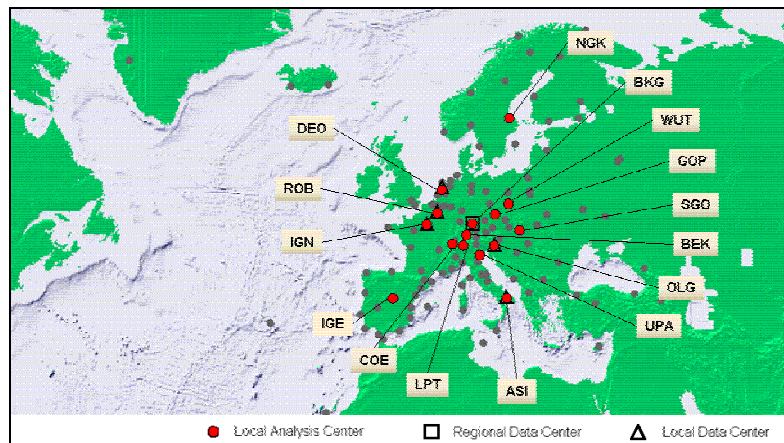


Figura 9. Centros de análisis y de datos de EPN [10]

En la realización de ITRS de 1997 el IERS empezó a tener en cuenta la solución de EUREF para su inclusión en los sucesivos ITRF. En la actualidad ETRS89 está totalmente consolidado como uno de los Sistemas Geodésicos de Referencias más precisos que existen. De cada solución ITRSy se puede derivar un marco ETRS89 mediante transformaciones simples [17] que permite conservar la calidad de la última solución ITRS y ser consistente con ETRS89.

9.6. Transformaciones entre Sistemas Geocéntricos Terrestres

Transformación rigurosa

En un sentido clásico, un datum geodésico es la superficie referencia, generalmente un elipsoide de revolución de un tamaño-forma convencionalmente escogido, con origen, orientación y escala definidas por un marco terrestre [18]. Una vez que el elipsoide es seleccionado las coordenadas de un punto en el espacio pueden ser dadas en forma cartesiana o curvilínea (latitud, longitud y altura).

Además, en sentido clásico se presupone que las coordenadas de los puntos no cambian excepto por algún motivo tectónico local (subsistencia, deslizamientos, erupciones volcánicas...). Por tanto, los marcos de un datum geodésico están ligados a una placa tectónica de manera que solidariamente las coordenadas se muevan respecto de ésta sin cambios respecto a la rotación de las placas [19].

El concepto moderno es el contrario, el marco está fijo al manto terrestre mientras que las placas son rotadas a su posición original en la época t_D (la época del datum) [20].

Esto es llevado a cabo aplicando el mismo “tipo” de corrección a cada punto. La magnitud esta corrección es determinada a través de la matriz de velocidades angulares asociada a la placa continental donde los puntos estén localizados.

En esencia, todos los puntos son movidos hacia atrás a su localización en la época t_D en el marco ITRF_{yy} que por su definición es fijo. La placa y los puntos sobre esta están congelados en el espacio en la época en la que el datum fue definido. Todas las coordenadas en una misma época t son traídas hacia atrás a la época del datum.

Para realizar cualquier transformación entre sistemas terrestres geocéntricos se debe considerar que los dos marcos estén en la misma época. Se requiere esta precaución para tener en cuenta el desplazamiento posible entre estaciones debido al movimiento de placas.

Puntos en el sistema y época ITRF_{yy}, época t_0 (donde yy denota la solución elegida) siempre tienen asociado un campo de velocidad para cada punto, siendo sus velocidades lineales V_x , V_y , V_z).

La transformación rigurosa de coordenadas entre un ITRF_{yy} época t_0 a ITRF_{zz} época t viene dado por la expresión [18]

$$X_{ITRFZZ} = T_X + (1 + s)(R + I)(X_{ITRFyy} + V_{X\ ITRFyy}(t - t_0)). \quad (1)$$

Donde:

X_{ITRFZZ}	Terna de coordenadas finales
T_X	Traslación entre los dos sistemas
s	Factor de escala entre los dos sistemas

R	Matriz de rotación entre los dos sistemas
I	Matriz identidad
X_{ITRFyy}	Coordenadas iniciales
$V_{X\ ITRFyy}$	Campo de velocidades
$(t - t_0)$	Distancia temporal entre sistemas

Aproximación rigurosa, transformación de catorce parámetros

El objeto de la transformación es definir el paso entre $ITRF00(t_D)$ donde t_D denota la época asociada al datum $ITRFyy(t_D)$. En este caso particular, la ecuación para transformar las coordenadas en el marco $ITRF00$ al marco $ITRFyy$ bajo la condición de que el marco $ITRF00$ no está cambiando con el tiempo y que las coordenadas de las estaciones están fijas en el espacio (no hay velocidades), entonces, se puede describir la transformación de tres dimensiones de Helmert según la expresión (1) o de forma equivalente:

$$\begin{Bmatrix} x' \\ y' \\ z' \end{Bmatrix}_B = \begin{Bmatrix} T_x \\ T_y \\ T_z \end{Bmatrix} + (1+s) \begin{bmatrix} 1 & \varepsilon_z & -\varepsilon_y \\ -\varepsilon_z & 1 & \varepsilon_x \\ \varepsilon_y & -\varepsilon_x & 1 \end{bmatrix} \begin{Bmatrix} x \\ y \\ z \end{Bmatrix}_A, \quad (2)$$

para el caso que nos ocupa:

$$\{x(t_D)\}_{ITRFyy} = \{T_x\} + (1+s)[\delta R]\{x(t_D)\}_{ITRF00}, \quad (3)$$

donde la matriz de rotación diferencial denominada $[\delta R]$ es dada de forma explícita por:

$$[\delta R] = [I] + [\underline{\varepsilon}]^t = [I] + \begin{bmatrix} 0 & \varepsilon_z & -\varepsilon_y \\ -\varepsilon_z & 0 & \varepsilon_x \\ \varepsilon_y & -\varepsilon_x & 0 \end{bmatrix}, \quad (4)$$

y cuyo sentido de rotación es antihorario de ejes. Esta transformación tiene sentido $ITRF00 \triangleright ITRFyy$ en un instante t_D común a las dos series de coordenadas.

Asumamos ahora que las coordenadas en el marco $ITF00$ se está moviendo a una cierta velocidad respecto al marco, el cual permanece fijo, es decir, se conocen las velocidades y las coordenadas de una determinada época t . En este caso la transformación de Helmert quedaría

$$\{x(t_D)\}_{ITRFyy} = \{T_x\} + (1+s)[\delta R]\{x(t) + (t_D - t)\{v_x\}\}_{ITRF00}, \quad (5)$$

finalmente, si asumimos los parámetros de Helmert cambian respecto al tiempo y que éstos son proporcionados en una determinada época t_k generalmente distinta de la época t_D . Los parámetros finales quedarían como sigue:

$$\{T_x\} \equiv \{T_x(t_D)\} = \{T_x(t_k)\} + (t_D - t_k)\{\dot{T}_x\}, \quad (6)$$

$$[\underline{\varepsilon}]^t \equiv [\underline{\varepsilon}(t_D)]^t = [\underline{\varepsilon}(t_k)]^t + (t_D - t_k)[\underline{\dot{\varepsilon}}]^t, \quad (7)$$

$$s \equiv s(t_D) = s(t_k) + (t_D - t_k)\dot{s}, \quad (8)$$

donde el punto indica la derivada respecto del tiempo, por ejemplo:

$$[\underline{\dot{\varepsilon}}]^t = \frac{\partial [\underline{\varepsilon}]^t}{\partial t} = \begin{bmatrix} 0 & \dot{\varepsilon}_z & -\dot{\varepsilon}_y \\ -\dot{\varepsilon}_z & 0 & \dot{\varepsilon}_x \\ \dot{\varepsilon}_y & -\dot{\varepsilon}_x & 0 \end{bmatrix}, \quad (9)$$

si sustituimos (4), (5) y (6) en (3) quedaría la fórmula final de transformación:

$$\begin{aligned} \{x(t_D)\}_{ITRF_{yy}} &= \{T_x(t_k)\} + (t_D - t_k)\{\dot{T}_x\} + \\ &+ \left[(1 + s(t_k)) [\partial \mathbf{R}] + (t_D - t_k) \times \left((1 + s(t_k)) [\underline{\varepsilon}]^t + s[\partial \mathbf{R}] \right) + (t_D - t_k)^2 s[\underline{\dot{\varepsilon}}]^t \right] \\ &\times \left(\{x(t)\}_{ITRF_{00}} + (t_D - t_k)\{v_x(t)\}_{ITRF_{00}} \right) \end{aligned} \quad (10)$$

el sentido de las rotaciones es anti horario, sin embargo si se aplican rotaciones de cuerpo (rotación de vectores) como en el caso de la tectónica de placas la matriz antisimétrica de rotación tiene los signos opuestos. En este caso particular los ejes del marco permanecen fijos mientras que los vectores de posición (coordenadas) son rotados en sentido anti horario. Esta rotación es denominada "rotación de Euler" o "rotación activa" [21]

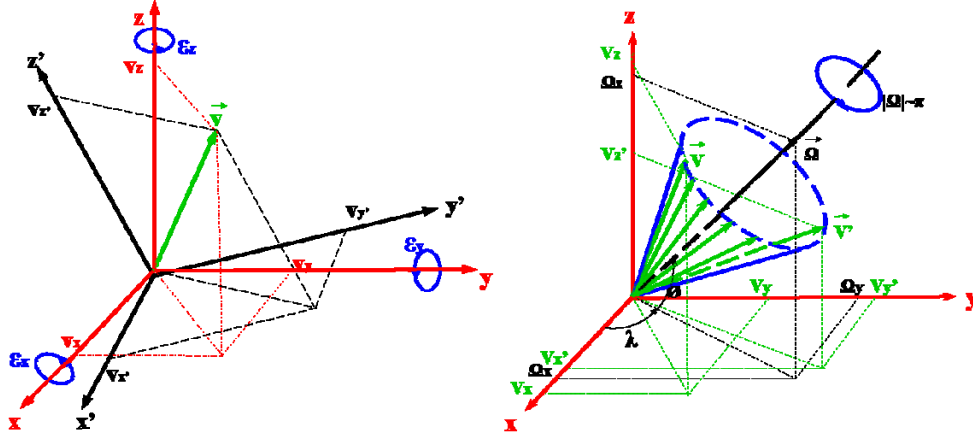


Figura 10. Rotación activa y rotación de vectores (derecha)

$$\begin{bmatrix} 0 & \varepsilon_z & -\varepsilon_y \\ -\varepsilon_z & 0 & \varepsilon_x \\ \varepsilon_y & -\varepsilon_x & 0 \end{bmatrix} = [\underline{\varepsilon}]^t \left\{ \begin{array}{l} \text{rotaciones de ejes} \\ \text{contrarias a las agujas del reloj} \\ \text{positivas} \end{array} \right. \quad (11)$$

$$\begin{bmatrix} 0 & -\Omega_z & \Omega_y \\ \Omega_z & 0 & -\Omega_x \\ -\Omega_y & \Omega_x & 0 \end{bmatrix} = [\underline{\dot{\Omega}}] \begin{cases} \text{rotaciones de puntos} \\ \text{contrarias a las agujas del reloj} \\ \text{positivas} \end{cases} \quad (12)$$

donde la velocidad de un determinado punto dadas las componentes de la velocidad angular de la placa tectónica según el modelo REVEL [22] y NNR-NUVEL-1A [23] serían:

Placa		Norte América	Eurasia	Australia	Sudamérica
Velocidad angular	Modelo	mas/año	mas/año	mas/año	mas/año
$\dot{\Omega}_x$	NNR-NUVEL-1A	0.0532	-0.2023	1.6169	-0.2141
	REVEL	0.1358	-0.1030	1.4534	-0.2454
$\dot{\Omega}_y$	NNR-NUVEL-1A	-0.7423	-0.4940	1.0569	-0.3125
	REVEL	-0.7036	-0.4763	1.1463	-0.2422
$\dot{\Omega}_z$	NNR-NUVEL-1A	-0.0316	0.6503	1.2957	-0.1794
	REVEL	-0.0299	0.7882	1.2893	-0.1669

Tabla 2. Velocidades angulares de las placas más importantes

El conjunto de épocas para los Sistemas Convencionales más frecuentes son

NAD83 (CORS96)	ETRS89	GDA94	SIRGAS
ITRF00->NAD83	ITRF00->ITRF89	ITRF00->ITRF92	ITRF00->ITRF94
$t_D = 2002.00$	$t_D = 1989.00$	$t_D = 1994.00$	$t_D = 1995.40$

Tabla 3. Épocas de referencia de los Sistemas Convencionales más frecuentes

Y los parámetros de transformación corresponderían a la tabla siguiente

ITRF00->		NAD83	ITRF89	ITRF92	ITRF94
		$t_k=1997.0$	$t_k=1988.0$	$t_k=1988.0$	$t_k=1997.0$
T_X	cm	99.56	2.97	1.47	0.67
T_Y	cm	-190.13	4.75	1.35	0.61
T_Z	cm	-52.15	-7.39	-1.39	-1.85
$\mathcal{E}_X$	mas	25.915	0.00	0.00	0.00
$\mathcal{E}_Y$	mas	9.426	0.00	0.00	0.00

ITRF00->		NAD83	ITRF89	ITRF92	ITRF94
ε_z	mas	11.599	0.18	0.18	0.00
s	ppb	0.62	5.85	0.75	1.55
$\dot{T}_x$	cm/año	0.07	0.00	0.00	0.00
$\dot{T}_y$	cm/año	-0.07	-0.06	-0.06	-0.06
$\dot{T}_z$	cm/año	0.05	-0.14	-0.14	-0.14
$\dot{\varepsilon}_x$	mas/año	0.013	0.00	0.00	0.00
$\dot{\varepsilon}_y$	mas/año	-0.015	0.00	0.00	0.00
$\dot{\varepsilon}_z$	mas/año	-0.020	-0.02	-0.02	-0.02
$\dot{s}$	ppb/año	-0.18	0.01	0.01	0.01

Tabla 4. 14 parámetros de transformación entre ITRF00 y distintos sistemas [18]

Aproximación de Boucher y Altamimi

El objetivo de este apartado es mostrar cómo se calcula una campaña GPS observada en un determinado momento t_c y la obtención de coordenadas en un sistema comparable al Europeo ETRS89. El IGS (Servicio internacional de GPS) distribuye las efemérides precisas de la constelación GPS en el sistema ITRF más reciente. Cada solución contiene una combinación de observaciones de distintas técnicas espaciales: DORIS, GPS, SLR y VLBI. Las efemérides precisas del IGS, por ejemplo, se expresaron en ITRF96 a partir de las 0 horas UTC del 1 de marzo de 1998, el marco en el que se expresan varía dependiendo de la realización más cercana. Por tanto, las coordenadas de los puntos observados en un instante determinado se obtienen de acuerdo con el marco que definan las efemérides en ese momento.

Es evidente, por tanto, que se debe de utilizar la última realización de ITRF disponible antes de la campaña si empleamos las efemérides IGS. Por tanto, las estaciones usadas como referencia durante la campaña GPS deben de ser fijadas (o fuertemente constreñidas) a los valores siguientes

$$X_{yy(t_c)} = X_{yy(t_0)} + V_{yy}(t_c - t_0). \quad (13)$$

El siguiente paso es convertirlos en ETRS 89. La expresión que permite realizar esta transformación es

$$X_{ETRS89(t_c)} = X_{YY(t_c)} + T_{yy} + V_{Ryy} [X_{YY(t_c)} (t_c - 1989.0)]. \quad (14)$$

Para ello se utilizan los valores de la Tabla 5 y Tabla 6. [17].

yy		T_X (cm)	T_Y (cm)	T_Z (cm)
1989		0	0	0
1990	A	1.9	2.8	-2.3
	B	2.6 ± 0.7	2.5 ± 0.7	-2.6 ± 0.7
1991	A	2.1	2.5	-3.7
	B	2.3 ± 0.7	2.1 ± 0.7	-3.1 ± 0.7
1992	A	3.8	4.0	-3.7
	B	4.3 ± 0.8	3.4 ± 0.8	-3.2 ± 0.8
1993	A	1.9	5.3	-2.1
	B	1.0 ± 0.5	5.9 ± 0.5	-1.4 ± 0.5
1994	A	4.1	4.1	-4.9
	B	2.9 ± 0.4	4.3 ± 0.4	-3.6 ± 0.4
1996	A	4.1	4.1	-4.9
	B	3.9 ± 0.4	4.1 ± 0.4	-3.9 ± 0.4
1997	A	4.1	4.1	-4.9
	B	3.4 ± 0.4	4.4 ± 0.4	-4.3 ± 0.4
2000	A	5.4	5.1	-4.8
	B	4.2 ± 0.4	5.1 ± 0.4	-4.6 ± 0.4
2005	A	5.6	4.8	-3.7
	B	3.6 ± 0.4	4.2 ± 0.4	-4.1 ± 0.4

Tabla 5. Parámetros de traslación de distintos ITRS a ETRS89 [17]

Yy	R_X (0.001"/año)	R_Y (0.001"/año)	R_Z (0.001"/año)
1989	0.11	0.57	-0.71
1990	0.11	0.57	-0.71
1991	0.21	0.52	-0.68
1992	0.21	0.52	-0.68
1993	0.32	0.78	-0.67
1994	0.20	0.50	-0.65
1996	0.20	0.50	-0.65
1997	0.20	0.50	-0.65
2000	0.081 ± 0.021	0.490 ± 0.008	-0.792 ± 0.026
2005	0.054 ± 0.009	0.518 ± 0.006	-0.781 ± 0.011

Tabla 6. Velocidades de rotación de distintos ITRS a ETRS89 [17]

Por último, dado que cualquier transformación siempre se realiza en la misma época, quedaría pasar las coordenadas a 1989. Se utiliza la expresión: $XE(89.0)=XE(tc)+VE(1989.0-tc)$, donde VE es la estimación de la velocidad de la estación en ETRS. Para la parte estable de Europa se puede considerar $VE=0$.

Bibliografía

- [1] Dalda A, González-Matesanz J, Cano M, Valdes M, Sánchez JA, editors. VII Curso GPS en geodesia y cartografía. Colombia, 2007.
- [2] McCarthy DD, Petit G. IERS Conventions (2003). IERS Technical Note No. 32. In: Centre IC, editor, 2003.
- [3] Ray J. Report of the Working Group on ITRF Datum. IERS Gazette: IERS Gazette, 1999.
- [4] Boucher C, Altamimi Z. Towards an improved realization of the BIH terrestrial frame. Columbus, Ohio, EEUU: OSU/DGS, 1985.
- [5] EUREF. Terms of Reference. EUREF2004 Symposium. Bratislava, 2004.
- [6] Adam J, Augath W, Boucher C, Bruyninx C, Dunkley P, Gubler E, Gurtner W, Hornik H, Marel H, SCHLÜTER W, Seeger H, Vermeer M, Zielinski JB. The European Reference System Coming of Age. Geodesy Beyond 2000 - The Challenges of the First Decade. IAG Symposia, vol. 121, 1999.
- [7] Agria J. EUREFTerms of reference. EUREF 2004. Bratislava, Slovakia, 2004.
- [8] Megrim. Conclusions and recomendations. Spatial Reference Workshop. Marne-la-Vallée, Francia, 1999.
- [9] Moritz H. Geodetic Reference System 1980. Bulletin Geodesique 1984;58:388.
- [10] Ihde J. The European Datum ETRS89 and Related Map Projections. 1st European Workshop on Reference Grids JRC Milan: European Commision, 2004.
- [11] Booch G, Rumbaugh J, Jacobson I. El lenguaje unificado de modelado. Madrid: Addison-Wesley, 2000.
- [12] Brunner FK. Advances in positioning and reference frames IAG Scientific Assembly : Rio de Janeiro, Brazil, September 3-9, 1997. Berlin: Springer, 1998.
- [13] European Commision. Mandatory use of Projection Systems within the Commission 2003.
- [14] Gurtner W. Guidelines for a Permanent EUREF GPS Network. In: Veröffentlichungen der Bayerischen Kommission für die Internationale Erdmessung, editor. EUREF 1995, vol. 26. Helsinki, 1995.
- [15] Blewitt G, Bock Y, Kouba J. Constructing the IGS polyhedron by distributed processing. In: Zumbege JF, Liu R, editors. IGS Workshop 1994. Jet Propulsion Laboratory, Pasadena, California, USA, 1994.

- [16] Brockmann E. Combination of solutions for geodetic and geodynamic applications of the global positioning system (GPS) / Elmar Brockmann. Zürich: Institut für Geodäsie und Photogrammetrie, Eidg. Technische Hochschule Zürich, 1997.
- [17] Boucher C, Altamimi Z. Memo: Specifications for reference frame fixing in the analysis of a EUREF GPS campaign. vol. 75: EUREF IAG Subcommission, 2001. p.7.
- [18] Soler T, Marshall J. A note on frame transformations with applications to geodetic datums. GPS solutions 2003;7:23.
- [19] Steblov GM. Contemporary space geodesy and reference frames. Doklady Earth Sciences 2004;399:1139.
- [20] González.-Matesanz J. Aportaciones al estudio de los modelos de distorsión para el cambio de datum entre ED50 y ETRS89. Departamento de matemáticas. España: Universidad de Alcalá, 2007.
- [21] Soler T. A compendium of transformations formulas useful in GPS Work. Journal of Geodesy 1998;72:482.
- [22] Sella GF, Dixon TH, Mao AL. REVEL: A model for Recent plate velocities from space geodesy. Journal of Geophysical Research-Solid Earth 2002;107.
- [23] Argus DF, Gordon RG. No-Net-Rotation Model of Current Plate Velocities Incorporating Plate Motion Model Nuvel-1. Geophysical Research Letters 1991;18:2039.

Tema 10. Métodos de precisión para el levantamiento de un punto Laplace mediante procedimientos ópticos: Determinación del acimut astronómico por el método de series a la polar, procedimiento y precisión. Determinación del ángulo de latitud astronómica por el método de Sterneck, procedimiento y precisión. Determinación del ángulo de longitud astronómica por el método de Mayer, procedimiento y precisión. Reducción de los datos astronómicos al polo medio de rotación

10.1. Introducción

Entendiendo por punto Laplace aquél vértice de la red geodésica donde se determinan las coordenadas astronómicas y el acimut del mismo tipo de un lado de la red, estudiaremos a continuación tres de los principales procedimientos clásicos de precisión que han sido utilizados durante años con tal objeto. El primero atiende a la observación astronómica del acimut, y los dos siguientes a la medición del ángulo de latitud y del ángulo de longitud respectivamente. Otros detalles en relación con lo que sigue se encuentran en Caturla (1978) y particularmente en Mena (2008), referencia esta última de la que se ha extraído el presente resumen.

10.2. Determinación del acimut astronómico por el método de series a la polar

La primera tarea a realizar para el levantamiento astronómico de un punto consiste en obtener el acimut de una dirección con la mayor precisión posible. Esta labor ha de ser la primera porque el dato de acimut facilitará notablemente la determinación posterior de las coordenadas astronómicas. Uno de los procedimientos diseñados para tal fin es el *método de las series a la estrella polar* que pasamos a describir (ver también Sevilla y Chueca, 1977, y Sevilla, 1979).

* Como se observa en la figura 1 (Ref. [10]), el cálculo del acimut a una referencia lejana R puede obtenerse a partir del acimut conocido de un astro β en un determinado instante, a través de la siguiente expresión:

$$(1) \quad A_R = A_\beta + L_R - L_\beta$$

Igualdad donde interviene el acimut del astro y la diferencia entre lecturas acimutales del aparato a la referencia y al astro. En esta relación resulta evidente que los problemas principales estriban en el cálculo del acimut del astro, así como en la determinación precisa de la lectura al mismo, dado su continuo movimiento aparente relativo al observador.

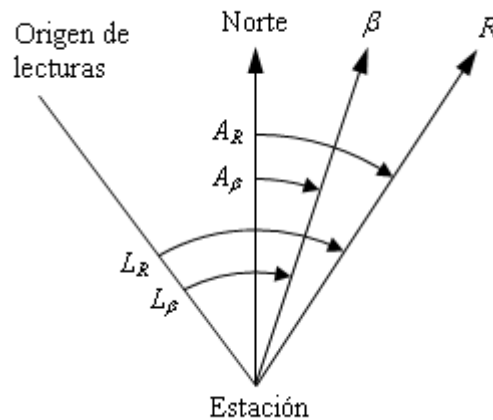


Figura 1.

La lectura al astro puede conseguirse de forma adecuada mediante un dispositivo que registre la hora en la cual el astro cruza el centro del retículo del aparato. Manteniendo inmóvil el anteojo en dicho instante, se procede a tomar la lectura horizontal. En cuanto al cálculo del acimut, éste se realiza para el día en cuestión y la hora registrada, atendiendo al triángulo polar de la figura 2 (Ref. [10]). Según esta disposición, las fórmulas de la Trigonometría esférica proporcionan la siguiente igualdad que define el acimut del astro como función de la latitud, la declinación y el ángulo horario:

(2)

$$\tan A_{\beta} = \frac{\text{sent}}{\text{sen}\Phi \cos t - \cos \Phi \tan \delta} \Leftrightarrow A_{\beta} = \arctan\left(\frac{\text{sent}}{\text{sen}\Phi \cos t - \cos \Phi \tan \delta}\right)$$

Para encontrar las mejores condiciones de aplicación de esta fórmula, diferenciamos la primera igualdad tomando Φ y t como variables en el segundo miembro ya que la declinación la supondremos carente de error en las efemérides. Tras despejar el diferencial de acimut, se tiene:

$$dA_{\beta} = \frac{\cos^2 A_{\beta}}{(\text{sen}\Phi \cos t - \cos \Phi \tan \delta)^2} \left[-\text{sent}(\cos \Phi \cos t + \text{sen}\Phi \tan \delta) d\Phi + (\text{sen}\Phi - \cos \Phi \cos t \tan \delta) dt \right]$$

Y sustituyendo el factor que figura en el denominador por su valor en función de la tangente del acimut, resulta:

$$dA_{\beta} = \text{sen}^2 A_{\beta} \left(-\frac{\cos \Phi \cos t + \text{sen} \Phi \tan \delta}{\text{sen} t} d\Phi + \frac{\text{sen} \Phi - \cos \Phi \cos t \tan \delta}{\text{sen}^2 t} dt \right)$$

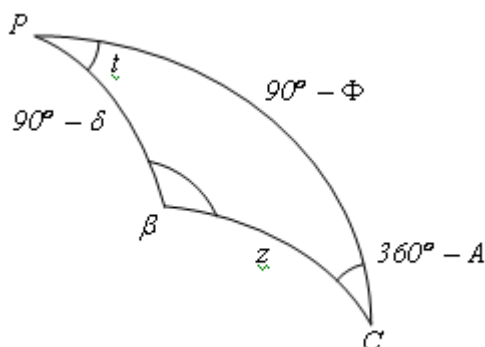


Figura 2.

Esta expresión nos indica la conveniencia de elegir un astro cuyo acimut sea lo más cercano a cero posible; es decir, que se encuentre en las proximidades del meridiano. Pero consideremos además las siguientes relaciones trigonométricas en el triángulo polar:

$$\cos \Phi \cos t \cos \delta + \text{sen} \Phi \text{sen} \delta = \cos z$$

$$\cos \delta \text{sen} \Phi - \cos \Phi \cos t \text{sen} \delta = \text{sen} z \cos \beta$$

$$\text{sen} t \cos \delta = -\text{sen} A \text{sen} z$$

Tras sustituirlas en el diferencial de acimut, se tiene:

(3)

$$dA_{\beta} = \text{sen} A_{\beta} \cotg z d\Phi + \frac{\cos \delta \cos \beta}{\text{sen} z} dt$$

Encontramos pues que también interesan valores grandes en la declinación del astro. Esta consecuencia y la obtenida inicialmente conducen a elegir una estrella muy próxima al polo celeste, ya que de este modo, a la vez que la declinación tiene un valor cercano a 90°, los acimutes observados desde una latitud media son muy pequeños en todo instante. Asimismo, la

expresión deducida muestra que lo ideal será realizar las mediciones cuando la estrella se halla en la zona de su máxima digresión ($\beta = 90^\circ$); y también aconseja considerar ángulos cenitales no excesivamente pequeños.

Por lo dicho, la estrella más idónea para observar en el hemisferio norte terrestre es $\alpha - Ursae Minoris$, en tanto que en el hemisferio sur este lugar lo ocupa $\sigma - Octantis$, más difícil de observar que la primera en razón a su mayor magnitud, pero también muy próxima al polo celeste. Por sus elevadas declinaciones, estas estrellas tienen velocidades lineales pequeñas en sus movimientos aparentes, siendo circumpolares para la gran mayoría de las latitudes de su respectivo hemisferio.

En cuanto al margen de latitudes donde el método es aplicable, se estima que no deberían ser inferiores a los 15° puesto que entonces la refracción toma un valor excesivo interfiriendo en la visual al astro, aunque el ángulo cenital en sí no intervenga directamente en el cálculo del acimut. Asimismo también es aconsejable que no superen los 55° a fin de que los lados del triángulo polar no tengan valores angulares demasiado pequeños o el acimut se aparte en demasía del meridiano.

Finalmente, en lo relativo a la referencia terrestre cuyo acimut se pretende levantar, lo ideal es que ésta se encuentre muy próxima a la dirección del meridiano con objeto de poder observarla mediante el hilo móvil del retículo del aparato, o sea sin necesidad de actuar sobre el movimiento horizontal (método micrométrico). Por otra parte, si en la igualdad (3) se considera la citada referencia en lugar de la estrella, entonces también resulta conveniente que la altitud del punto visado sea similar a la del vértice estación para que el ángulo cenital sea grande. Téngase en cuenta que, si bien la refracción será alta en este caso, la distancia a la referencia no es comparable con la distancia a la estrella, con lo cual el posicionamiento horizontal de la visual no se verá prácticamente afectado.

* En el marco del levantamiento astronómico clásico, para llevar a cabo el proceso con la precisión necesaria es imprescindible contar con un equipo de observación de calidad. Este equipo consta generalmente de un teodolito capaz de apreciar valores angulares a la décima del segundo sexagesimal (T-4) o bien al medio segundo (DKM3-A), más un dispositivo detector de la señal horaria que proporcione ésta con precisión. Dicha señal horaria se toma en tiempo universal, siendo suficiente en este procedimiento con el uso de la escala UTC y la aproximación del segundo. Asimismo resulta necesario un barómetro y un termómetro para efectuar posteriores correcciones.

Dispuesto el citado equipo, el método se divide en series independientes de observaciones. En cada serie se requiere conocer inicialmente los valores aproximados de longitud y latitud del lugar, así como también el parámetro TSM o TSA a 0^h UT (Mena, 2008) del día en que se realiza la medición, siendo indistinto usar uno u otro por las razones que veremos enseguida.

Considerando una serie aislada, el procedimiento consiste en realizar 8 observaciones sobre la referencia y otras tantas sobre la estrella polar. En el caso de la referencia, estas observaciones comprenden 8 lecturas acimutales y 4 lecturas cenitales; mientras que en el caso de la polar, además se añaden 8 tomas de la señal horaria, una por cada medida del ángulo acimutal. Las citadas observaciones se realizan alternando periódicamente los círculos directo (ocular a la derecha) e inverso (ocular a la izquierda) del aparato, y se distribuyen en la serie de forma que, comenzando y terminando por la referencia, todas las observaciones a la polar se efectúen en bloque en la zona central de la serie según el esquema que se expone a continuación (las X indican las correspondientes lecturas de datos):

Num.	Círculo	PV	L.Az.	L.Cen.	O/I	O/D	HUTC
(Toma de datos de presión y temperatura)							
1	CD	R	X	X	X	X	
2	CD	R	X				
3	CI	R	X				
4	CI	R	X	X	X	X	
5	CI	Polar	X	X	X	X	X
6	CI	Polar	X				X
7	CI	Polar	X				X
8	CI	Polar	X	X	X	X	X
9	CD	Polar	X	X	X	X	X
10	CD	Polar	X				X
11	CD	Polar	X				X
12	CD	Polar	X	X	X	X	X
13	CD	R	X	X	X	X	
14	CD	R	X				
15	CI	R	X				
16	CI	R	X	X	X	X	
(Toma de datos de presión y temperatura)							

Tabla 1.

Además de las lecturas mencionadas, si el aparato tiene un nivel caballero suspendido del eje de muñones, entonces las medidas cenitales se acompañan con las lecturas a los extremos de la burbuja en las dos posiciones del nivel, que se corresponden respectivamente con la situación del cero de su graduación a la izquierda y la derecha de la visual (0 / I, 0 / D). Estas lecturas serán utilizadas después para introducir la corrección relativa a la inclinación del eje secundario del aparato.

Para obtener el valor de acimut resultante de la serie, se comienza calculando el promedio R de todas las lecturas acimutales a la referencia así como el valor medio de las horas registradas:

(4)

$$R = \frac{1}{8} \sum L_{Ri}, \quad HUTC_m = \frac{1}{8} \sum HUTC_i$$

Acudiendo a las efemérides con el valor medio de la hora universal, se buscan seguidamente las coordenadas α y δ de la estrella polar para dicho instante; coordenadas que deberán estar actualizadas y corregidas por las diversas causas que se tratan en la Astronomía geodésica (Mena, 2008). A partir de aquí, para cada $HUTC_i$ se calcula el correspondiente valor de la hora sidérea local y después el respectivo ángulo horario. Introduciendo ahora este ángulo en (2) se encuentra el acimut bajo el cual se observó la polar en el instante $HUTC_i$. Este valor, llevado a la expresión (1) proporciona el correspondiente acimut parcial a la referencia. Esquemáticamente, el proceso es así:

$$\begin{aligned} HSL_i &= \Lambda + TSA + HUTC_i \frac{366.2422}{365.2422} \Rightarrow t_i = HSL_i - \alpha \Rightarrow A_{\beta_i} \Rightarrow \\ &\Rightarrow A_{Ri} = R + A_{\beta_i} - L_{\beta_i} \end{aligned}$$

Con los ocho acimutes parciales se obtiene entonces el promedio de la serie:

(5)

$$A_R = \frac{1}{8} \sum_i A_{Ri}$$

* El acimut resultante ha de corregirse a continuación; principalmente a causa de la inclinación del eje secundario del aparato, aunque también suele introducirse una corrección adicional debida a la aberración diurna. La primera se calcula sobre la base de las lecturas del nivel caballero utilizando la siguiente expresión:

(6)

$$C_{in} = F_R \cotg z_R - F_{\beta} \cotg z_{\beta}$$

Aquí los ángulos cenitales son los respectivos promedios a la referencia y a la polar corregidos por refracción, para lo cual, tanto al inicio como al final de la serie se habrán medido los valores de presión y temperatura utilizándose las medias aritméticas de estas medidas para entrar en la fórmula de Bradley (Mena, 2008):

(7)

$$\Delta z'' = 57''.8 \operatorname{tg}(z - 230'' \operatorname{tg} z) \frac{278(1 - 0.00016 T)}{753.5(268 + T)} b$$

En cuanto a los otros términos, éstos dependen del valor μ_c que, expresado en segundos, se corresponde con una subdivisión del nivel caballero; y también intervienen aquí los sumatorios de todas las lecturas del nivel realizadas con el cero a la derecha por un lado y con el cero a la izquierda por el otro, de acuerdo con las relaciones:

$$F_R = \frac{\mu_c}{16} \left(\sum 0/D - \sum 0/I \right)_{\text{referencia}}, \quad F_\beta = \frac{\mu_c}{16} \left(\sum 0/D - \sum 0/I \right)_{\text{estrella}}$$

Por su parte, la corrección por aberración diurna se establece en un único valor estándar, el cual resulta consecuente con los márgenes de latitud recomendados para el método. Dicho valor se estima en:

(8)

$$C_{ab} = \pm 0''.32$$

Éste se toma con signo positivo o negativo según que los ocho acimutes calculados a la polar vayan en aumento o disminución respectivamente.

Como consecuencia de estas correcciones, el acimut final resultante de la serie y referido al polo instantáneo de rotación, el cual se sustituye hoy en día por el CIP (Mena, 2008), resulta dado por la igualdad:

(9)

$$A_{PIR}(\text{una serie}) = A_R + C_{in} + C_{ab}$$

* Del proceso descrito se desprenden las exigencias en precisión de los distintos parámetros que intervienen en el cálculo. Así por ejemplo, carece de interés obtener las marcas horarias de forma más exacta, o exigir el empleo del TSA en lugar del TSM, toda vez que la

hora sidérea local resulta mediante un valor aproximado de la longitud, siendo además que el ángulo horario aparece en la fórmula (2) como argumento de funciones trigonométricas. No obstante, el resultado (9) de una sola serie no alcanza la suficiente precisión como para ser utilizado en la red geodésica. De hecho, el cálculo del error medio teórico que afecta a este valor, suponiendo condiciones de observación adecuadas y utilizando valores medios para las variables de que depende, se estima en aproximadamente 1.2". Por lo tanto, para rebajar este número hasta un margen consecuente con la precisión exigible, como pueden ser las dos décimas de segundo sexagesimal, la teoría de errores indica que será necesario establecer como acimut resultante el promedio de 36 series realizadas en las mismas condiciones. Ello se demuestra teóricamente considerando el error medio del valor más probable como función del error medio de una observación aislada y del número n de mediciones (Sevilla, 1993; Mena, 1997):

$$E_{mc} = \frac{e_{mc}}{\sqrt{n}} \quad \Rightarrow \quad n = \frac{e_{mc}^2}{E_{mc}^2} = \frac{1.2^2}{0.2^2} = 36$$

Como curiosidad nótese que para alcanzar un error en el resultado final de tan sólo una décima de segundo, sería necesario realizar 144 series. Por esta razón, en general podemos escribir que aproximadamente y de forma estándar se verifica:

(10)

$$A_{PIR}(\text{método}) = \frac{1}{36} \sum_i A_{PIR}(\text{serie } i), \quad E_{mc} \cong \pm 0.2''$$

10.3. Determinación del ángulo de latitud astronómica por el método de Sterneck. Procedimiento y precisión.

* Uno de los procedimientos más utilizados para el cálculo de la latitud en los levantamientos astronómicos de precisión ha sido durante muchos años el denominado método de Sterneck. Su fundamento es sumamente simple, ya que basta considerar que si la declinación de un astro es conocida, al observarlo en el instante de su culminación el ángulo de latitud se obtiene mediante la siguiente expresión elemental (Mena, 2008):

(11)

$$\Phi = \delta \pm z$$

Recordemos que el signo positivo se corresponde con el caso de culminación sur, y el negativo con la culminación norte. Por consiguiente, en el método de Sterneck teóricamente sólo es necesaria la medida del ángulo cenital en el momento de la culminación de un astro para obtener la latitud del lugar.

Como se deduce de la igualdad anterior, la precisión del resultado está íntimamente ligada con las precisiones de los dos sumandos del segundo miembro. Esto exige, por un lado que la declinación de la estrella esté corregida según todos los factores conocidos lo más exactamente posible, y por otro que los ángulos cenitales sean medidos justamente en el meridiano del lugar, y además con la máxima minuciosidad. La primera exigencia no es difícil de satisfacer aplicando a las coordenadas ecuatoriales del astro la serie completa de correcciones que se estudian en la Astronomía geodésica, desde el movimiento propio hasta la aberración diurna, utilizando la hora de tiempo terrestre correspondiente al instante de la culminación (Mena, 2008). Esta hora bastará determinarla con la aproximación del segundo.

Sin embargo la medición exacta del ángulo cenital es bastante más problemática de lograr; sobre todo si como es habitual en el método clásico se utilizan aparatos ópticos de observación, por muy fieles que éstos sean. La razón de ello estriba en diversas causas de difícil evaluación. Una consiste en la puesta rigurosa en estación del aparato; lo que implica la necesidad de una perfecta nivelación para establecer la vertical astronómica (causa inevitable de errores), unida al establecimiento posterior de la orientación al meridiano a través de un dato de acimut (asimismo afectado de error). Otra causa adicional consiste en la evaluación de la corrección de refracción, imprescindible cuando se miden ángulos cenitales y que según dijimos es imposible de conocer exactamente. Aparte de estas causas, existen muchas otras que están asimismo presentes tales como la precisión del observador en las lecturas y punterías, los errores residuales en las posiciones relativas de los ejes del aparato, la progresiva degeneración de la nivelación inicial, etc.

Considerando pues los principales factores que acarrear errores en el resultado, y partiendo de la base de que se cuenta con un observador experimentado y un equipo óptico y de medida de tiempo de calidad, el método de Sterneck propone: primeramente establecer el aparato en el meridiano utilizando el valor de acimut obtenido recientemente por series a la polar; y en segundo lugar observar un número suficientemente elevado de estrellas, clasificándolas por parejas de culminación norte y sur de forma que sus cenitales se correspondan lo mejor posible al objeto de compensar en gran parte el efecto de la refracción atmosférica. Además de ello, cada observación ha de ir acompañada de todas aquellas correcciones adicionales que puedan ser evaluadas, buscando así para cada estrella un ángulo cenital lo más depurado posible con el

cual establecer la ecuación correspondiente según (11). Una vez hayan sido planteadas todas las ecuaciones, la resolución del sistema por mínimos cuadrados proporcionará los valores de las incógnitas que nos permiten obtener el ángulo de latitud.

* En el método de Sterneck las observaciones se distribuyen en 6 series, cada una de las cuales está compuesta por unas 20 parejas de estrellas. Estas parejas serán seleccionadas de un listado a confeccionar inicialmente, denominado *puntero de observación*, donde figuren para un amplio conjunto de estrellas del catálogo los siguientes datos: su número de identificación, la magnitud, la hora de tiempo universal aproximada de paso meridiano, los cenitales aparentes esperados calculados en cada uno de los círculos del aparato, y el sentido norte o sur de la culminación.

El número y la magnitud son valores que se toman directamente del catálogo. En cuanto a la hora universal de cruce por el meridiano del lugar, ésta se obtiene por la siguiente expresión (Mena, 2008) en la cual tomaremos como referencia el tiempo sidéreo aparente (TSA) a 0 horas UT. Dicha fórmula resulta ser:

(12)

$$HPM_{\beta} = \left(\alpha - \Lambda - TSA \ 0^h \ UT \right) \frac{365.2422}{366.2422}$$

Evidentemente, la hora obtenida será tanto más precisa cuanto más lo sean los valores de la ascensión recta y la longitud astronómica del lugar. Y aunque ahora no sea necesario, resulta conveniente corregir el valor de α para trasladarlo de la época del catálogo hasta la fecha del día de observación. De este modo el error de la hora resultante sólo dependerá de la aproximación con que se conozca el ángulo de longitud. Por su parte, el cenital aparente es el ángulo bajo el cual será esperado el paso meridiano de la estrella a la hora anteriormente obtenida. Dicho ángulo se calcula a partir de una latitud aproximada del punto de estación y de la declinación de la estrella que al igual que α deberá estar actualizada, despejándolo de (11). Además, y puesto que la refracción es importante, se calculará el efecto de ésta por la fórmula de Bradley (7), restándose entonces del valor de z obtenido al objeto de conseguir el correspondiente valor bajo el cual teóricamente hemos de observar (Mena, 2008). Calculado el ángulo cenital, se incluirán en el puntero los dos valores del mismo correspondientes a los círculos del aparato. Finalmente, el sentido de la culminación resulta de forma inmediata como consecuencia de comparar la declinación con el valor aproximado de latitud.

Dispuesto así el puntero de estrellas para una serie determinada, se procede entonces a seleccionar un total de unos 40 astros, los cuales se distribuyen en parejas de modo que las

culminaciones de cada par sean lo más simétricas posible respecto del cenit, procurando además que en todos los casos los ángulos cenitales se mantengan por debajo de los 30°. Bajo este criterio, es deseable que en cada serie la suma algebraica de todos los ángulos cenitales ronde los cero grados, no debiendo de superarse en ningún caso el margen de los 5° ó 10°.

Antes de comenzar las observaciones reviste gran importancia determinar la posición del origen de las lecturas cenitales del aparato (error de eclímetro), así como ajustar el reloj a la hora UTC con la precisión suficiente de un segundo. Si como es habitual en los equipos de precisión el eclímetro tiene acceso manual, se obtendrá el valor z_0 correspondiente a la posición del origen de cenitales mediante 10 pares de lecturas en ambos círculos a una referencia bien definida, y calculando después el promedio de todos los valores resultantes. Para cada par de lecturas, y suponiendo graduación sexagesimal, dicho valor responde a la expresión:

(13)

$$z_0 = \frac{360^\circ - (L_{CD} - L_{CI})}{2}$$

En el caso de que este número sea excesivo, se procede a corregirlo manualmente y se reitera su medición a fin de obtener el valor residual.

Elegidas las 20 parejas, calculado el parámetro z_0 y ajustada la central horaria, la observación de la serie es un procedimiento muy sencillo. Para ello, una vez situado el aparato en estación y orientado al meridiano, el movimiento acimutal no se vuelve a tocar durante el proceso, manipulándose exclusivamente el movimiento cenital. De este modo las estrellas de culminación norte serán observadas en un círculo del aparato, que supondremos es el directo (CD), y las de culminación sur en el otro (CI). La medición sobre cada estrella se realiza entonces introduciendo previamente al momento de la culminación el ángulo cenital que figura en el puntero. Cuando el astro aparece en el anteojo, lo que se producirá sobre la hora estimada si todos los cálculos han sido correctos, se ajusta el movimiento cenital al objeto de hacerla pasar rigurosamente por el centro del retículo. En dicho instante se toma la señal horaria y la correspondiente lectura cenital. Estos datos se acompañarán con lecturas de presión y temperatura necesarias para la corrección posterior de refracción. Asimismo, como en los aparatos de eclímetro no automático existe un nivel asociado al mismo, se tomarán las lecturas de éste antes y después del paso meridiano de la estrella con objeto de corregir adicionalmente el ángulo cenital. Además de los datos anteriores, que son de toma obligada para cada estrella, cada dos o tres observaciones deberán ser leídos también los extremos de la burbuja del nivel caballero con el cero al este y el cero al oeste, a fin de calcular posteriormente, de forma análoga

a lo hecho en el método de acimut, la corrección del ángulo cenital debida a la inclinación del eje secundario del aparato.

El proceso descrito para una serie se realiza en la misma forma hasta completar seis series por lo menos. Ello proporciona el número suficiente de ecuaciones de observación (unas 120), como para esperar un resultado aceptable en el valor del ángulo de latitud. Veamos seguidamente cómo preparar los datos y cómo establecer después las mencionadas ecuaciones.

* Correcciones a la declinación. Respecto a la declinación de cada estrella, ésta será actualizada desde la época del catálogo al instante de su culminación, utilizando para ello la fecha y la hora de tiempo terrestre (TT) obtenida a partir de la señal UTC tomada para el paso meridiano. Dicha hora se calcula con rigor mediante la expresión (Mena, 2008):

(14)

$$TT = UTC + DTA + 32.184 \text{ seg}$$

No obstante, si en la preparación del puntero se corrigieron las coordenadas ecuatoriales convenientemente, no será imprescindible volver a realizarlo ahora, ya que la modificación de estas coordenadas en un pequeño intervalo de tiempo no es apreciable en relación con la precisión del método, sobre todo empleando medios ópticos.

* Correcciones al ángulo cenital. La relativa permisividad citada para el cálculo de la declinación corregida desaparece en el tratamiento de los ángulos cenitales, dado que en la fidelidad de los mismos reside la precisión del resultado final para el ángulo de latitud. Así, partiendo del ángulo leído, la primera operación será convertirlo al círculo directo de no estarlo ya, introduciendo a la vez la corrección de eclímetro z_0 . Se obtiene entonces un primer valor z_1 que resulta de la forma siguiente:

$$z_1 = z_{CD} + z_0 \quad \text{o bien} \quad z_1 = 360^\circ - z_{CI} - z_0$$

A continuación se procede a calcular la corrección por nivel de eclímetro C_{ne} en base a las cuatro lecturas realizadas de los extremos izquierdo y derecho de la burbuja de este nivel (dos lecturas inmediatamente antes de medir el cenital y dos inmediatamente después). Esta corrección se obtiene en función del valor angular μ_e de una división del nivel correspondiente, mediante la igualdad:

(15)

$$C_{ne} = \frac{\mu_e}{4} (\sum L_I - \sum L_D)$$

El valor así calculado se añade entonces al ángulo z_1 , con signo positivo o negativo según que éste proceda del círculo directo o inverso respectivamente; es decir, el cenital observado a la estrella será:

(16)

$$z_{ob} = z_{CD} + z_0 + C_{ne} \quad \text{o bien} \quad z_{ob} = 360^\circ - z_{CI} - z_0 - C_{ne}$$

Dispuesto el valor observado, se introducen ahora cuatro correcciones más. La principal es la corrección por refracción atmosférica (C_R) que determinamos mediante la fórmula de Bradley (7) y siempre es positiva. La segunda es la corrección por inclinación del eje secundario (C_{in}). Este valor se deduce a partir del triángulo polar considerando una inclinación I que es función del número n de lecturas realizadas con el nivel caballero en sus dos posiciones, así como también del parámetro μ_c correspondiente a una división del citado nivel. Expresada en segundos, esta corrección responde a la fórmula:

(17)

$$I'' = \frac{\mu_c}{2n} \left(\sum 0/E - \sum 0/W \right), C_{in} = -\frac{I^2}{2} \frac{\cos(\Phi - \delta)}{\cos \delta} \operatorname{sen} \Phi \frac{\pi}{648000}$$

El tercer parámetro a tener en cuenta corresponde a la corrección por desvío acimutal del aparato respecto del meridiano (C_a). Dicha desviación depende del error que afecta al acimut mediante el cual se establece la orientación del meridiano, pero principalmente está causada por el inevitable error de lectura y puntería propio de la pareja observador – aparato. Cuando este error es desconocido como frecuentemente sucede, entonces suele aceptarse por término medio un valor a de falta de orientación estimado en $5''$. Con ello, la fórmula correspondiente que expresa el valor de la corrección en segundos, deducida asimismo del triángulo polar, es esta:

(18)

$$C_a = -\frac{a^2}{2} \cdot \frac{\operatorname{sen}|\Phi - \delta|}{\cos \delta} \cos \Phi \cdot \frac{\pi}{648000}$$

Finalmente, también de las relaciones que ligan los distintos elementos del triángulo polar se desprende la igualdad que proporciona el valor en segundos de la denominada corrección por colimación. Este término está causado por la existencia de un cierto error c en la dirección del eje óptico del antejo respecto del eje geométrico del mismo. Para un aparato suficientemente

contrastado, la respectiva corrección sólo se introduce cuando las dos correcciones anteriores son apreciables. En tal caso, y dado c en segundos de arco, su valor se calcula así:

(19)

$$C_c = -\frac{c^2}{2} \cdot \operatorname{tg} \delta \cdot \frac{\pi}{648000}$$

Aplicando las indicadas correcciones, el ángulo cenital verdadero correspondiente a la estrella en cuestión queda definitivamente determinado en la forma siguiente:

(20)

$$z = z_{ob} + C_R + C_{in} + C_a + C_c$$

* Corregidos para cada estrella la declinación y el ángulo cenital, y observadas todas las series, se procede a la construcción del sistema de ecuaciones tomando por incógnitas el incremento de latitud $\Delta\Phi$ sobre el valor aproximado conocido Φ_a , más un incremento en ángulo cenital Δz que se supone igual en todo el proceso y que está motivado por las múltiples causas no evaluadas. De este modo, y teniendo en cuenta (11), cada ecuación adopta una de las dos formas que siguen, dependiendo de si la estrella culminó al norte o al sur respectivamente:

(21)

$$\Delta\Phi + \Delta z = \delta_i - z_i - \Phi_a \quad (N), \quad \Delta\Phi - \Delta z = \delta_i + z_i - \Phi_a \quad (S)$$

Estas ecuaciones pueden ponderarse atendiendo a diversas circunstancias; una muy socorrida consiste en asignar peso 1 a las observaciones que finalmente han sido emparejadas, y peso 0.5 a las que por diversos motivos no lo fueron.

Respecto de la resolución del sistema, ésta debe hacerse mediante mínimos cuadrados a través de un método robusto, es decir, aplicando ponderación de residuos y reiterando los cálculos hasta alcanzar una solución estabilizada (Peña, 1999; Björck, 1996). No obstante, puede prescindirse del carácter robusto en el cálculo si previamente a la introducción de cada ecuación en el sistema nos aseguramos que ésta no lleva asociado algún error grosero, lo que se realiza comparando la suma algebraica de la declinación y el cenital con el valor aproximado de latitud. En cualquier caso, una vez resuelto el sistema el resultado de latitud astronómica referido al polo instantáneo resulta mediante la expresión:

(22)

$$\Phi_{PIR} = \Phi_a + \Delta\Phi$$

En cuanto a la precisión de este valor, se evalúa en función del error medio estimado para el cenital de una observación aislada y del número de ecuaciones del sistema. Considerando un aparato óptico de precisión y un observador experimentado, si para el primer parámetro se toman 2'' y 120 para el segundo, entonces podemos cifrar el error del resultado en torno a la siguiente magnitud:

(23)

$$E_{mc} = \frac{e_z}{\sqrt{n}} = \frac{2''}{\sqrt{120}} \cong 0.2''$$

Esto supone un margen de incertidumbre sobre la superficie terrestre alrededor de los 6 metros.

10.4. Determinación del ángulo de longitud astronómica por el método de Mayer. Procedimiento y precisión.

* Sin duda es el levantamiento de la longitud lo que determina la precisión final en las coordenadas astronómicas del vértice. Esto es así porque, si bien existen distintos métodos para calcular la longitud, todos ellos están basados en la diferencia horaria entre el meridiano del lugar y el meridiano de referencia. Se trata pues de un problema de cálculo de tiempo, y como tal se encuentra sujeto a la precisión con la que éste se mida. Téngase en cuenta que un segundo de tiempo expresado en distancia sobre la superficie terrestre, supone aproximadamente 450 metros. Por tanto, para alcanzar resultados aceptables en las determinaciones astronómicas con fines geodésicos, dicho margen debería rebajarse al menos hasta los 6 metros consecuentemente con el error del levantamiento en latitud. Ello supone trabajar con la milésima de segundo al objeto de que los errores inherentes al proceso de observación aseguren la centésima en el resultado, es decir, teóricamente unos 4.5 metros.

Por la Astronomía geodésica sabemos que el ángulo de longitud se obtiene mediante la diferencia entre la hora sidérea del lugar (HSL) y la hora sidérea que en dicho instante corresponde al meridiano de referencia (HSR). En consecuencia, si observamos una estrella de ascensión recta conocida en el momento de su paso meridiano, considerando que en ese instante el ángulo horario es nulo ($t = 0$), la hora sidérea local será entonces igual a la ascensión recta, y con ello la longitud del punto de estación estará dada por la igualdad:

(24)

$$\Lambda = \alpha - HSR$$

En esta ecuación, el valor de α puede ser calculado con gran precisión mediante las correcciones astronómicas clásicas, actualizando así este parámetro desde la época del catálogo hasta la fecha y hora de tiempo terrestre correspondiente al paso meridiano. El problema real está pues en la determinación exacta de la hora sidérea que en dicho instante es en el meridiano de Greenwich. Este valor responde a la siguiente, pero tomando como tiempo sidéreo a las 0 horas UT el aparente del día de la observación:

(25)

$$HSR = TSA \ 0^h \ UT + HUT \frac{366.2422}{365.2422}$$

Aquí la hora de tiempo universal (HUT) ha de ser precisamente la del instante justo en que la estrella considerada culmina en el lugar de estación. Se plantean de este modo dos importantes problemas. El primero consiste en el exacto establecimiento del aparato de medición en el meridiano del lugar. Y el segundo concierne a la toma de la hora universal de paso meridiano con una precisión de milésima de segundo; hora que además deberá obtenerse en la escala UT1 al objeto de asegurar su máximo ajuste con la rotación terrestre.

Para resolver estos problemas, o al menos minimizar su influencia, el método de Mayer preconiza: por un lado estudiar la importancia y su corrección de las distintas causas que intervienen en la orientación física del aparato de observación sobre el meridiano; y por otro, obtener la hora de tiempo universal de paso meridiano a través de la media aritmética de múltiples tomas correspondientes a posiciones simétricas de la estrella con respecto al meridiano.

Atendiendo al primero de los problemas citados, resulta evidente que si el aparato de observación no se encuentra perfectamente orientado sobre el meridiano, entonces existirá un valor residual Δt del ángulo horario que motivará posteriores errores en la hora resultante. Este valor ha de tenerse muy en cuenta en los métodos de precisión, por lo que en la práctica, en lugar de la expresión (24) se utiliza esta otra:

(26)

$$\Lambda = \alpha + \Delta t - HSR$$

Para evaluar el incremento en ángulo horario se recurre al triángulo polar. Tras un análisis del mismo, la denominada fórmula de Mayer (SGE, 1972) muestra que dicho ángulo residual puede definirse como función de tres factores principales: el desvío acimutal a del aparato, la inclinación b de su eje secundario, y el error de colimación c del anteojo. Dicha fórmula es la siguiente, que escribimos también de manera abreviada:

(27)

$$\Delta t = -\frac{\sin(\Phi - \delta)}{\cos \delta} a - \frac{\cos(\Phi - \delta)}{\cos \delta} b - \frac{1}{\cos \delta} c \Leftrightarrow \Delta t = Aa + Bb + Cc$$

De los tres factores a , b y c , el primero es desconocido, el segundo puede medirse, y el tercero puede ser eliminado mediante un procedimiento adecuado de observación. Por consiguiente, asumiendo tales circunstancias, la expresión (26) adopta la forma siguiente:

(28)

$$\Lambda = \alpha + Aa + Bb - HSR$$

Además, como el factor a es una incógnita siempre presente por muy fielmente que se intente orientar el aparato, lo deseable para que dicho término tenga la menor influencia posible sería que su coeficiente A esté muy próximo a cero. Por esta razón en el método de Mayer se aboga por observar estrellas cuyas declinaciones sean muy parecidas a la latitud del lugar. Esta afirmación constituye la primera condición teórica del método de Mayer.

Por otra parte, si consideramos el triángulo polar junto con el desvío angular m del eje secundario del aparato respecto del primer vertical, se obtiene la siguiente relación trigonométrica entre la latitud, la declinación, y el ángulo horario (SGE, 1972):

$$\operatorname{tg} \Phi \operatorname{sen}(m + t) = \operatorname{tg} \delta \operatorname{sen} m$$

Pues bien, diferenciando esta relación, dividiendo el resultado entre la primera y despejando el error en t , se obtiene:

(29)

$$dt = \frac{2}{\cot(m + t)} \left(\frac{d\delta}{\operatorname{sen} 2\delta} - \frac{d\Phi}{\operatorname{sen} 2\Phi} \right)$$

De esta igualdad se deduce que para minimizar el valor de la función, tanto la declinación de cada estrella como el ángulo de latitud del lugar deberían tomar valores próximos a los 45°. Esta consecuencia es la segunda condición teórica del método de Mayer. Asimismo, la fórmula

indica que la precisión del resultado de longitud se incrementa cuando dicha coordenada se determina mediante observaciones en el meridiano, es decir, cuando m y t son cero.

* Considerando el proceso experimental de medición de la hora universal de paso meridiano, deberá disponerse de un teodolito de alta precisión que incluya en el retículo un hilo móvil independiente del movimiento acimutal del aparato (micrómetro), conectado directamente con una central horaria capaz de apreciar la milésima de segundo sincronizada con UTC. Dicha central ha de registrar de algún modo preciso las horas correspondientes a las distintas posiciones del hilo móvil en su recorrido. Evidentemente, si las horas registradas tanto antes como después del cruce meridiano de la estrella se corresponden dos a dos con posiciones simétricas del astro respecto del meridiano, el promedio de todos los valores medios será una buena aproximación de la hora UTC de paso meridiano (HPM). Posteriormente esta hora será convertida a la escala UT1 e introducida como argumento en (25).

La observación de la longitud por el método de Mayer se subdivide en 10 series parciales compuestas por 8 estrellas cada una. En cada una de tales series las estrellas se eligen de un puntero construido en la misma forma indicada para el método de Sterneck, pero realizando esta elección conforme a los siguientes requerimientos derivados de lo dicho en los párrafos anteriores. En primer lugar la fórmula de Mayer aconseja eliminar el posible error c de colimación del aparato mediante el propio procedimiento de observación. Para ello cada estrella será observada en ambos círculos acimutales del teodolito, seleccionándose además por cada serie 4 estrellas de culminación sur y otras tantas de culminación norte.

En cuanto a los valores de la declinación, y dada la imposibilidad de que todos ellos sean iguales que la latitud, se opta por elegir las estrellas de modo que el coeficiente A que afecta al desvío acimutal a en la formula (28) no tenga un peso excesivo en el futuro sistema de ecuaciones. Para ello se acotan las declinaciones forzando que el citado coeficiente no supere los márgenes que siguen:

$$-0.6 < -\frac{\text{sen}(\Phi - \delta)}{\cos \delta} < +0.6 \Rightarrow \arctan\left(\frac{\text{sen}\Phi + 0.6}{\cos \Phi}\right) > \delta > \arctan\left(\frac{\text{sen}\Phi - 0.6}{\cos \Phi}\right)$$

Además debe procurarse que en cada serie la suma algebraica de todos los valores de estos coeficientes sea de módulo inferior a la unidad:

$$\sum_1^8 A_i \in (-1,1)$$

Como condición adicional, las declinaciones tampoco deberán sobrepasar los 60°; ello está motivado tanto por el valor ideal de los 45° como por la necesidad de que la velocidad aparente de la estrella durante su persecución con el hilo móvil del retículo no sea demasiado lenta. Asimismo, y también por causa del seguimiento, los cenitales habrán de mantenerse siempre por debajo de los 45°.

Una vez seleccionadas las estrellas a observar, sincronizada exactamente la central horaria con UTC, y situado el aparato en estación con orientación precisa al meridiano, se introduce el ángulo cenital correspondiente a la primera estrella de la serie y se sitúa el hilo móvil del micrómetro en el extremo del retículo por el cual se espera la entrada de ésta a la hora aproximada indicada en el puntero (posición p_1 de la figura 3. Ref. [10]). Cuando aparece la estrella y se superpone sobre el hilo citado, se procede a seguirla de forma suave y continua hasta que el hilo móvil alcanza alguna de las marcas fijas más próximas al hilo fijo central del retículo (meridiano), pero sin llegar a sobrepasarlo (posición p_2). Llegados a este punto, se cambia el círculo del aparato y se vuelve a orientar sobre el meridiano, con lo cual el hilo móvil aparecerá ahora en la posición q_2 . Introducido el cenital correspondiente para visualizar la estrella en el anteojo, ésta no deberá haber rebasado la citada posición q_2 . Siendo así, cuando la estrella llegue aquí se vuelve a perseguir en su movimiento; persecución que finalizará cuando el hilo móvil alcance la posición q_1 equivalente a la inicial p_1 . Puesto que la central horaria registra el movimiento del hilo móvil (normalmente por pulsos), se habrán obtenido tantas marcas horarias en el recorrido $p_1 \rightarrow p_2$ como en el recorrido $q_2 \rightarrow q_1$, correspondiéndose unas con las otras según las distintas posiciones simétricas de la estrella con respecto al meridiano.

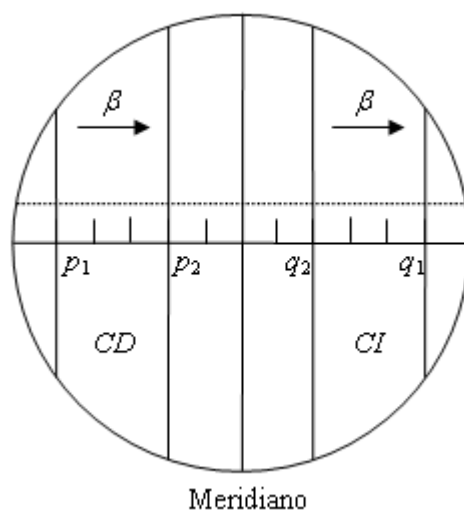


Figura 3.

Con objeto de aplicar después la corrección por inclinación del eje secundario, la observación de cada estrella se acompaña de las correspondientes lecturas del nivel caballero con el cero al este y al oeste del meridiano. Dichas lecturas se realizan tanto inmediatamente antes de que aparezca la estrella en el retículo, como inmediatamente después de la observación de ésta. El estadillo de observación de cada estrella consiste pues en la relación de las cuatro lecturas del mencionado nivel, más el listado de marcas de tiempo registradas por la central horaria.

* Una vez observadas todas las series, y antes de construir el sistema de ecuaciones que proporcionará el resultado de longitud, es necesario calcular para cada estrella los valores de los distintos términos y coeficientes ponderables que intervienen en la expresión (28). Para ello, primeramente se calcula la hora UTC de culminación como la media de los n promedios de marcas de tiempo correspondientes a las n parejas de posiciones simétricas del astro respecto al meridiano. Esta hora se convierte entonces a la escala UT1 mediante la relación (Mena, 2008):

(30)

$$UT1 = UTC + \Delta UT1$$

Y el resultado se introduce como argumento en (25) para obtener la respectiva hora sidérea reducida. Considerando la estrella i -ésima, esquemáticamente se tiene:

$$HUTC_i = \frac{1}{n} \sum HPM_j \quad \Rightarrow \quad HUT1_i \quad \Rightarrow \quad HSR_i$$

A partir de la hora UTC media de la culminación se calcula también la hora de tiempo terrestre (HTT) correspondiente a dicho instante mediante (14), procediéndose a corregir por todas las causas conocidas el valor de ascensión recta que figura en el catálogo:

$$HUTC_i \Rightarrow HTT_i \Rightarrow \alpha_i$$

Análogamente al caso de la declinación en el método de Sterneck, y a la vista de los márgenes de precisión del procedimiento, la actualización de esta coordenada no será necesaria ahora si ya fue realizada para un instante cercano en la confección del puntero.

A continuación se obtiene la corrección por inclinación del eje secundario con las lecturas observadas del nivel caballero; lo que realizamos a través de la siguiente expresión donde interviene el valor angular de una subdivisión del nivel:

(31)

$$b_i = \frac{\mu_c}{4} (\sum 0/E - \sum 0/W)$$

Y finalmente, con la declinación actualizada se calculan los coeficientes A y B mediante sus respectivas definiciones extraídas de la fórmula de Mayer:

(32)

$$A_i = -\frac{\text{sen}(\Phi - \delta_i)}{\cos \delta_i}, \quad B_i = -\frac{\cos(\Phi - \delta_i)}{\cos \delta_i}$$

Tomando como incógnita principal el incremento de longitud $\Delta\Lambda$ sobre el valor aproximado conocido Λ_a (éste último expresado con su signo), y como incógnita secundaria el desvío acimutal a , se genera por cada estrella observada la ecuación siguiente en la cual todos los términos deben figurar en unidades de tiempo:

(33)

$$\Delta\Lambda - A_i a = \alpha_i - HSR_i + B_i b_i - \Lambda_a$$

El sistema así formado se resuelve entonces de la misma manera que se expuso en el método de Sterneck; es decir, mediante mínimos cuadrados, ya sea aplicando un método robusto o bien asegurando previamente la carencia de errores groseros en todas las ecuaciones a través de la igualdad (24). En consecuencia, el valor de longitud astronómica referido al polo instantáneo resulta por la suma:

(34)

$$\Lambda_{PIR} = \Lambda_a + \Delta\Lambda$$

En cuanto al error que se espera en este resultado, es función del error en el cálculo de la ascensión recta y del error en la determinación de la hora sidérea reducida. El primero se supone inexistente toda vez que las coordenadas ecuatoriales han sido corregidas con la suficiente precisión. Sin embargo, el segundo depende de numerosos factores como el sincronismo de la central horaria, el desvío acimutal del meridiano, la precisión óptica del aparato de observación, el espesor de los contactos del micrómetro y el tiempo muerto que se pierde en la transmisión de los pulsos al reloj, la experiencia del observador, etc... Muchas de estas causas son insignificantes y otras son corregidas por el proceso de observación y cálculo. No obstante, en la precisión de la hora sidérea reducida existe un factor que impera notablemente sobre todos los demás y que no es fácil de corregir: se trata de la fidelidad con la cual se realiza el seguimiento de la estrella y que influye directamente en la hora UTC observada de paso meridiano. Este término es determinante en la precisión del resultado final, suponiendo que no existen errores groseros en los restantes factores. De hecho, aunque la central horaria esté perfectamente sincronizada a la milésima de segundo y el meridiano esté bien definido mediante un aparato preciso de observación, un mal seguimiento puede ocasionar incluso errores de décimas de segundo en la hora UTC; errores que se reflejan con toda su magnitud en la longitud resultante. Como consecuencia de ello, el cálculo numérico del error final mediante el error teórico del ajuste del reloj y el número de observaciones, no es indicativo, por lo que el error buscado suele estimarse concretamente para cada pareja observador – aparato mediante procedimientos experimentales. Como norma general, en los levantamientos clásicos de precisión realizados por profesionales, suele aceptarse por término medio un valor aproximado en torno a las 2 centésimas de segundo de tiempo; lo que supone un margen en distancia entre 5 y 10 metros:

(35)

$$E_{mc} \cong 0.02 \text{ seg}$$

10.5. Reducción de los datos astronómicos al polo medio de rotación

Los tres datos del levantamiento astronómico en un vértice están obtenidos con referencia al eje y polos instantáneos de rotación terrestre, por lo que los resultados sólo son válidos durante un cierto intervalo de tiempo tras la observación. Para evitar esta circunstancia, proporcionando continuidad en el tiempo al levantamiento astronómico, los valores calculados, tanto en acimut como en longitud y latitud, han de ser referidos a una posición media del eje de rotación; posición que está determinada por el centro de gravedad terrestre y por el polo medio de rotación (PMR). Actualmente, este eje es precisamente el correspondiente del sistema terrestre internacional ITRS.

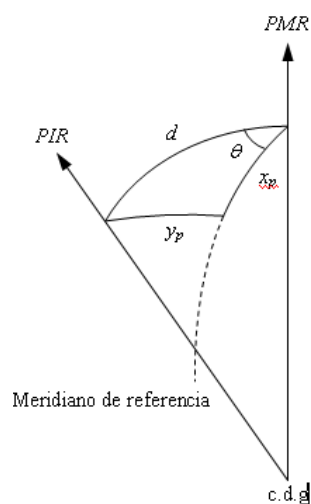


Figura 4.

La posición instantánea del polo (PIR), actualmente el Polo Celeste Intermedio (CIP), respecto al PMR, se precisa mediante los parámetros x_p e y_p representados en la figura 4 (Ref. [10]). Estos ángulos se miden con vértice en el centro de gravedad terrestre, respectivamente en el plano meridiano origen (Greenwich) y en el meridiano ortogonal al mismo. Dado su carácter infinitesimal, consideramos estos valores expresados en radianes.

Si denominamos con Λ , Φ y A_{ob} respectivamente a la longitud, latitud y acimut astronómicos obtenidos por los métodos descritos, y con λ_A , φ_A y A_A respectivamente a los correspondientes valores referidos al PMR, la transformación de unos en otros se realiza a través de las fórmulas siguientes cuyas deducciones las encontrará el lector en Mena (2008):

$$\begin{aligned}\lambda_A &= \Lambda - x_p \operatorname{sen} \Lambda + y_p \cos \Lambda \operatorname{tg} \Phi \\ (36) \varphi_A &= \Phi - x_p \cos \Lambda + y_p \operatorname{sen} \Lambda \\ A_A &= A_{ob} - x_p \operatorname{sen} \lambda_A + y_p \cos \lambda_A \sec \phi\end{aligned}$$

Debe tenerse muy en cuenta que los valores de las coordenadas y del acimut astronómicos reducidos al polo medio no son valores geodésicos, ya que no se trata de magnitudes referidas a la recta normal del elipsoide sino a la vertical física. De hecho, para su comparación con los correspondientes datos geodésicos, dichos valores astronómicos han de ser corregidos también por el ángulo de desviación de la vertical. Así, en el caso concreto del acimut, que es el factor principal a considerar, a partir del resultado corregido se obtiene el denominado *acimut geodésico de Laplace*, el cual es un valor procedente de la observación astronómica si bien se encuentra reducido a las referencias geodésicas con objeto de realizar la mencionada comparación dentro de un mismo marco (ver Mena, 2008, tema 8).

Bibliografía

- [1] Björck, A., 1996. Numerical methods for least squares problems. Siam. Philadelphia.
- [2] Bolshakov, V., Gaidayev, P., 1989. Teoría de la elaboración matemática de las mediciones geodésicas. Editorial Mir. Moscú.
- [3] Bradt, H., 2004. Astronomy methods. Cambridge University Press. U.K.
- [4] Caturla, J.L., 1978. Determinación de Puntos Laplace. Dirección General del Instituto Geográfico Nacional. Talleres del I. G. N., Madrid.
- [5] Danjon, A., 1982. Astronomie Générale. Senac. Paris.
- [6] Ferrer Torio, R., Piña Patón, B., 1988. Astronomía de posición. Servicio de publicaciones de la ETS de Ingenieros de caminos, canales y puertos. Universidad de Cantabria.
- [7] Mena, J.B., 1995. Observación astrogeodésica en la Antártida. Resultado y conclusiones. Informe interno del Servicio Geográfico del Ejército. Madrid.
- [8] Mena, J.B., 1997. Teoría del tratamiento matemático de las mediciones experimentales. Servicio Geográfico del Ejército. Escuela de Geodesia. Madrid.
- [9] Mena, J.B., 2004. Cálculo de las correcciones astronómicas para el levantamiento astrogeodésico. VII Jornadas de Trabajo en Mecánica Celeste, 28-30 de Junio de 2004. Real Observatorio de la Armada, S. Fernando (Cádiz).
- [10] Mena, J.B., 2008. Geodesia Superior. Texto en dos tomos. En fase de publicación por el Instituto Geográfico Nacional. Madrid.
- [11] Peña, D., 1999. Estadística. Alianza Universidad Textos. Editorial Alianza. Madrid.
- [12] Saastomainen, J., 1972-1973. Contributions to the theory of atmospheric refraction.

Bulletin Geodesique, 105, pp. 279-298; 106: pp. 383-397; 107, pp. 13-34.

[13] Sevilla, M.J., Chueca, R., 1977. Determinación de acimutes por observación de la polar. Método micrométrico. Universidad Complutense de Madrid, Facultad de Ciencias. Publicación núm. 95.

[14] Sevilla, M.J., 1979. Determinación del acimut de una referencia por observación de la estrella polar. Programa de cálculo automático. VRANIA, 1979, núm. 291 y 292.

[15] Sevilla, M.J., 1993. Teoría de errores de observación. Física de la Tierra, núm. 5, pp. 133-166. Editorial Complutense. Madrid.

[16] SGE, 1972. Astronomía Geodésica. Servicio Geográfico del Ejército. Escuela de Geodesia. Madrid.

[17] Zakatov, P.S., 1981. Curso de Geodesia Superior. Editorial Mir. Moscú.

Tema 11. Redes geodésicas: objeto y definiciones. Precisión. Triangulaciones clásicas: Longitud de los lados, utilización de las mismas. Medida de ángulos y distancias en Geodesia: Instrumentación, métodos de observación acimutal. Errores y compensación de una estación. Reducciones de las medidas. Calibración y contrastación de instrumentos.

11.1. Redes Geodésicas

Cuando hablamos de redes geodésicas nos referimos a las que tenían como fin principal determinar las coordenadas geodésicas latitud y longitud, ó actualmente las que determinan las coordenadas cartesianas tridimensionales globales (X,Y,Z), de un conjunto de puntos. En otros temas se exponen las redes que tienen como fin la determinación en sus puntos de algún tipo de altitud con significado físico, como sería la altitud ortométrica, y que son llamadas redes de nivelación de alta precisión. Aunque, ciertamente son también otro tipo de redes geodésicas diferente.

Por lo tanto, para la materialización de puntos cuyas coordenadas sean conocidas con la mayor precisión posible en un territorio dado, las llamadas redes geodésicas, han sido varios los métodos utilizados a lo largo del tiempo. Antes de los años cincuenta del siglo XX, las redes geodésicas consistían en triangulaciones angulares controladas por un número de bases de las que se conocía su distancia y un conjunto adicional de azimutes astronómicos. La introducción de la medida de distancias electromagnética (MDE) hizo que la triangulación fuera poco a poco suprimida por las poligonales y la trilateración con MDE [1]. La constatación de la utilidad geodésica de la observación doppler de satélites, en sistemas como el TRANSIT estadounidense o el TSIKADA soviético, hizo que renaciera la llamada geodesia espacial. El anterior precedente en la geodesia espacial, aparte de la pura geodesia astronómica, fue la observación simultánea a globos aerostáticos. La posterior aparición de los sistemas de navegación por satélite (GNSS) provocó el auge de la geodesia espacial como fuente principal de las observaciones con las que actualmente se realizan las redes geodésicas. El GPS estadounidense, en concreto, ha sido el sistema más utilizado hasta el momento. Otros sistemas como el ruso GLONASS, con graves problemas de despliegue, ó el futuro europeo Galileo, también con problemas económicos, presumiblemente serán utilizados en un futuro no muy lejano tanto como el GPS. Otros tipos de técnicas, con sus redes asociadas, también son utilizados en geodesia hoy en día. Entre ellos, tenemos por un lado la observación por distanciometría a

satélites o a la Luna con láser (SLR, LLR), y por otro lado la interferometría de muy larga base (VLBI), que observa a los lejanos cuasares. En cualquier caso, ambos merecerían por su complejidad un estudio aparte, aunque buena parte de sus principios de funcionamiento los encontramos en la MDE. Otro tanto se puede decir de los sistemas GNSS, que son más ampliamente explicados en otros temas. Por lo tanto, en este tema, nos centraremos más ampliamente en los instrumentos utilizados en los llamados métodos de geodesia clásica: la medición angular con teodolitos y la medición de distancias por MDE con distanciómetros. No olvidaremos mencionar la antigua metodología de medición de bases por medio de reglas ó invar.

Todavía hoy en día son muchas las redes geodésicas existentes que se observaron por triangulación clásica. En algunos casos las antiguas observaciones son todavía útiles y pueden ser reutilizadas en nuevos cálculos; de tal forma, que se ajustan conjuntamente con nuevas observaciones con MDE ó GNSS. Este es, por ejemplo, el caso de la Red de Orden Inferior española (ROI), de unos 10000 puntos. Con las antiguas observaciones de vueltas de horizonte angulares y nuevas con GPS ha sido recalculada aprovechando la existencia de los puntos de la Red Geodésica Nacional por Técnicas Espaciales (REGENTE). La red REGENTE tiene aproximadamente 1000 puntos con sus coordenadas obtenidas con GPS [17].

11.1.1. Objeto y definiciones

Los objetivos de las redes geodésicas clásicas eran básicamente los siguientes [1]:

- Constituir la red principal sobre la cual se puedan apoyar observaciones menos precisas, por lo cual pueden formar el marco para mapas topográficos ó catastrales.
- Combinando observaciones de latitud y longitud y de gravedad ayudar en la determinación de la forma y tamaño de La Tierra, las variaciones de densidad dentro de ella y la forma de las superficies equipotenciales fuera de ella.

De los dos objetivos, el primero de ellos es una necesidad práctica. En grandes territorios es necesario un adecuado marco geodésico para que se pueda realizar una cartografía correcta. Tradicionalmente ha sido difícil justificar el cuantioso gasto que supone la geodesia para aspectos puramente científicos, pero su necesidad para apoyo de la topografía y cartografía ha hecho este gasto justificable. La red geodésica principal de un país, siempre que cumpliera unos determinados estándares de observación y precisión, era la llamada *red primaria ó de primer orden*. Éstas tienen unos lados normalmente entre 30 y 50 Km entre vértices [2]. La red que se apoyaba en ésta, mediante triangulación ó con poligonales y con el fin de la determinación de los puntos necesarios para una topografía detallada, es la llamada *red terciaria ó de tercer orden* con lados entre 4 y 8 Km [2]. En algunos países existe una *red de segundo orden* que enlazaba

las dos primeras, con lados entre 15 y 20 Km [1]. A partir del tercer orden, y con lados más pequeños, están las llamadas redes de cuarto y quinto orden.

El uso de observaciones GNSS ó astronómicas (puntos Laplace) de las redes resultaban más precisas que las anteriores de primer orden y provocó el nacimiento de las *redes de orden cero*. Dentro de éstas también se pueden distinguir distintas clases, por ejemplo para Europa el VIII grupo de trabajo de la comisión CERCO decidió distinguir tres clases: A, B y C.

- Las de clase A tienen una precisión de más de 1 cm. en sus coordenadas y tienen observación continua y por lo tanto observación del vector velocidad del punto en cuestión. Ejemplos de estas redes serían la red del Internacional GNSS Service (IGS) [3] de carácter mundial, la red permanente europea de EUREF (EPN) [4] o la red de estaciones permanentes del IGN (ERGPS)[5].
- Las de clase B tienen una precisión de más de 1 cm. en sus coordenadas, pero no se conoce su vector velocidad. En la Península Ibérica un ejemplo de esta red sería la red IBERIA95.
- Las de clase C con una precisión de más de 5 cm. en sus coordenadas y no se conoce su vector velocidad, en España existe tenemos la red REGENTE [17].

Hoy en día, la existencia de redes geodésicas con observación GNSS de carácter continuo ha multiplicado los objetivos posibles en este nuevo tipo de redes geodésicas. Entre ellos podemos señalar: la monitorización continua en prevención de desastres naturales (terremotos, erupciones volcánicas), la observación meteorológica del contenido de vapor de agua de la atmósfera, la emisión de correcciones diferenciales para posicionamiento preciso en tiempo real.

11.1.2. Precisión

Generalmente las redes geodésicas son hechas con tanta precisión como es razonablemente posible. Para hacer mapas topográficos y levantamientos topográficos un error de escala de unas 100 a 200 partes por millón (ppm) puede ser admisible. Por lo tanto, es lógico que la red geodésica necesaria para conseguir esta precisión en todo el terreno sea mucho más precisa. Teniendo en cuenta esto último, es razonable conseguir una precisión final entre 10 y 20 ppm para los lados de los triángulos o poligonales de la red. [1]

En cuanto a los objetivos científicos, el primero de ellos es medir la diferencia entre la figura del nivel del mar prolongada bajo los continentes (geoide) de una forma esferoidal perfecta (elipsoide). Estas diferencias probablemente no son mayores de unos 60 m. ó 10 ppm, como máximo, respecto del radio de La Tierra. Para delinearlas necesitaremos una precisión de 2 ppm en levantamientos de extensión continental, lo que necesita unos errores accidentales o aleatorios de unas 10 a 20 ppm en cada lado de poligonal o triangulación. Son las mismas precisiones exigidas por los requerimientos cartográficos. Hoy en día es un objetivo que se puede lograr sobradamente.

En triangulación de primer orden el error medio triangular ha estado generalmente entre 0,3"-1,0". Una fórmula aproximada, aceptada para el error estándar deseable (error medio cuadrático) en la posición relativa de dos puntos geodésicos separados una distancia D, es [6]

$$\sigma = 100000 \sqrt{\frac{S}{30}}$$

Las actuales redes geodésicas disfrutaban de una grandísima precisión, que llega a ser del orden milimétrico además de pertenecer a sistemas de referencia globales, y no locales ó regionales como se definían en la época de las triangulaciones. Las mayores precisiones se consiguen sobre todo en las redes VLBI, las más precisas con precisiones bajo el milímetro, pero también en las redes GNSS con precisiones en todo caso por debajo del centímetro.

11.2. Triangulaciones clásicas

Las diferentes formas en que se han realizado las triangulaciones clásicas han sido dos principalmente como una red continua o como un sistema de cadenas. Las ventajas de una red eran: que necesitaba menos operadores de triangulación (topógrafos) y que permitía que el ajuste fuera llevado a cabo de modo satisfactorio ya que daba un resultado más preciso. La ventaja de un sistema de cadenas era que, excepto en áreas muy pequeñas, era menos costoso y más sencillo de ajustar. También era más rápido de observar y calcular, ya que los resultados ajustados podían ser obtenidos antes de que la mayoría de los trabajos topográficos hubieran sido realizados [1]. Las ventajas apuntadas en el sistema de cadenas condujeron a los sistemas de poligonales con MDE.

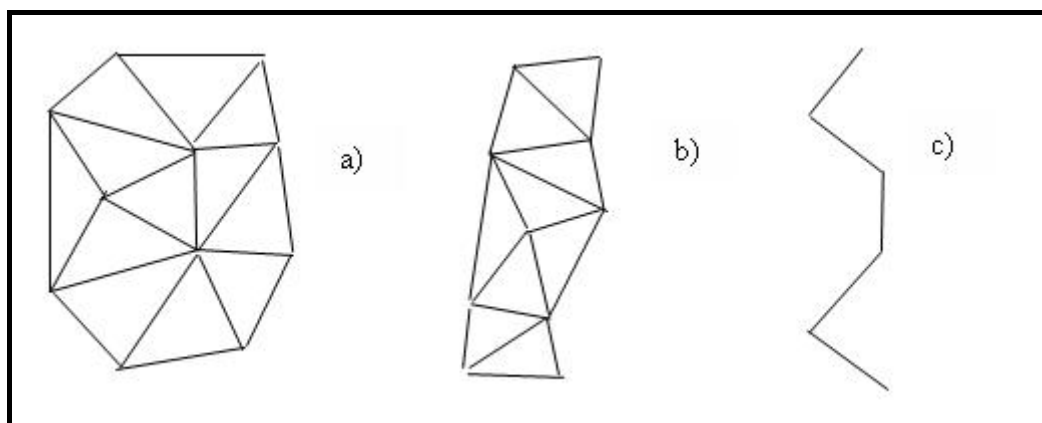


Figura 1. Tipos de redes geodésicas clásicas, a)red, b)cadena, c)poligonal

El punto débil de toda triangulación radica en que tiende a acumular errores de escala y azimuth, ya que cada lado obtiene su escala y azimuth del lado precedente. Por lo tanto, aunque la determinación de una única base (distancia) sería suficiente matemáticamente, así como una única determinación de azimuth. Es necesario controlar escala y azimuth mediante la

fueron añadiendo más observaciones astronómicas. La base fundamental fue la de Madrideo (Toledo) a la que se añadieron 5 más distribuidas por la periferia de la Península. Las primeras bases eran medidas con reglas, hasta la aparición del invar. Hoy en día parece una red de geometría dudosa, pero hay que tener en cuenta que en su día fueron figuras razonables y bien estudiadas dada la dificultad de encontrar sitios apropiados para la medida con reglas ó bases invar y el coste que suponía aumentar su extensión. [7]. Ochenta años después, fruto de la cooperación internacional en este campo y gracias a la aparición de los ordenadores, se procedió al ajuste conjunto de las observaciones geodésicas de primer orden aportadas por los países europeos. El primer ajuste fue realizado por el estadounidense “Army Map Service” [8] y, posteriormente, la subcomisión europea RETRIG de la Asociación Internacional de Geodesia (AIG) tomó el relevo y continuó sus trabajos. Esta solución de la red de primer orden sería adoptada, años más tarde, como marco en el que se ajustaron las densificaciones de la red geodésica española (parte continental), que es conocida como Red de Orden Inferior (ROI, Datum ED50 o Potsdam). Tanto este datum ED50 como el Datum Madrid se consideran *datums locales*, y, en consecuencia, las coordenadas geodésicas de la red referidas a ellos responden, también, a Sistemas Geodésicos de Referencia (SGR) locales.

La MDE hizo fácil medir un número considerable de lados largos de triángulos. En muchos países las nuevas redes observadas hicieron uso de este hecho, por ejemplo, Australia cubrió gran parte de su territorio con redes de trilateración. Las antiguas redes fueron en muchos sitios reobservadas, por ejemplo, España lo hizo sobre el 30% de su primer orden, además se observaron nuevos puntos Laplace para mantener cierta homogeneidad entre la precisión de azimut y escala, y se cubrió el territorio con nuevo tercer orden (ROI).

11.2.2. Cadenas de triangulación

Una cadena de triángulos podía consistir en diferentes tipos de figuras, como simples triángulos encadenados, cuadriláteros adosados, cuadriláteros y/o pentágonos y/o hexágonos centrados ó figuras aun más complejas, respectivamente a), b), c) y d) en la siguiente figura. Incluso cadenas que eran composición de todas las anteriores fueron finalmente construidas. El tipo adoptado era decidido generalmente al realizar el reconocimiento y construcción de la red sobre el terreno.

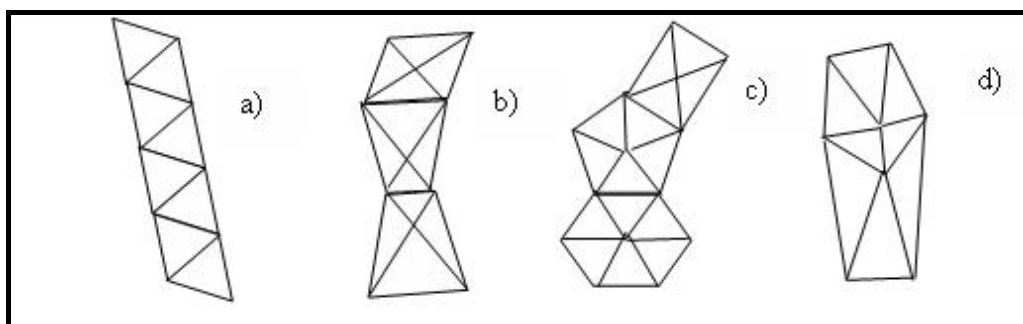


Figura 3. Tipos de cadenas de triángulos.

11.2.3. Redes continuas

Una red de triangulación debe constituir al menos una red continua de triángulos simples. Los tres ángulos de todos los triángulos deben, en general, ser observados. Adicionalmente, se pueden observar otras líneas además de las estrictamente necesarias. Realmente desde cada punto se observan todos los vértices visibles desde él. Más líneas dan más fuerza al posterior ajuste. Una sola línea que cruce muchas de las demás mejorará el ajuste final considerablemente. La figura ideal de los triángulos es la de triángulos equiláteros. Es ineludible la variación de las figuras sobre la ideal, sobre todo entre terreno muy llano o muy montañoso, pero se deben evitar cambios bruscos en la geometría, tanto como sea posible.

11.2.4. Longitud de las líneas de triangulación

Una buena distancia en terreno montañoso son 50 Km, aunque se han llegado a observar líneas de hasta 270 Km (enlaces hispano-argelino y de Baleares)[6]. En llanuras se podía llegar a tener lados medios de hasta 15 Km si no era posible utilizar torres de observación [1].

Las líneas largas se hacían para progresar rápidamente, pero daban problemas si había malas condiciones atmosféricas, además, la comunicación entre puntos se hacía más insegura. Las líneas también se hacían largas para conseguir máxima precisión, pero eso sólo ocurría si la luz sólo se curvaba en el plano vertical. Si esto no es así, los errores que se producían podían llegar a ser demasiado grandes. Por ello, los triángulos demasiado grandes no tienen por que ser los más precisos. Los lados largos también aumentan la imprecisión en el transporte de la escala. Además, las líneas más cortas hacen las cosas más fáciles para la densificación con redes de tercer orden, pero este problema era eludible con la construcción de pequeñas cadenas de triángulos que se unían a la red primaria.

El decrecimiento de la densidad del aire con la altura causa que la luz se curve en el plano vertical. La refracción lateral ó curvatura en el plano horizontal se puede dar cuando las condiciones no sean las mismas a los dos lados de una línea, o cuando la visual pasa cerca del

terreno o de una ladera , y por lo tanto, se deben evitar. Líneas bajo estas condiciones suponen un serio riesgo. En terrenos montañosos este efecto se puede eludir fácilmente, pero en terreno llano no es tan sencillo y sus efectos sólo se minimizan construyendo o usando grandes torres o acortando los lados. La MDE ha constituido un sustituto ideal para este tipo de terrenos. Líneas con mar en un extremo y tierra en otro, o con desierto en un extremo y vegetación en el otro pueden dar lugar a refracción lateral, incluso con visuales despejadas. Por lo tanto, siempre que sea posible, las estaciones deben estar en los puntos más altos de montañas, y preferiblemente en la más alta de un grupo dado. En todo caso es menester estudiar la necesidad de aumentar el número de bases y de puntos Laplace.

11.3. Redes Geodésicas con MDE

La introducción de la medida de distancias electromagnética supuso una revolución en la forma de diseñar, ejecutar y calcular las redes geodésicas de primer orden. Así, nació la posibilidad de realizar poligonales e incluso redes continuas de trilateración con MDE, pero esta se miden todos los lados y se añaden diagonales ya que la figura elemental es el cuadrilátero con sus dos diagonales. Lo más usual era la combinación de medidas de ángulos y distancias con MDE, ya sea en cadenas ó redes continuas [1].

11.3.1. Poligonales con MDE

Las poligonales con MDE se convirtieron en el método recomendado de observación de redes de primer orden en grandes países, sustituyendo a las triangulaciones. Las ventajas que presentaba comparadas con la triangulación eran:

- La facilidad con que una simple línea de estaciones fácilmente accesibles podía ser seleccionada, comparada con la selección de, por ejemplo, una adecuada cadena de triángulos.
- La facilidad de organizar observaciones solamente a dos estaciones a la vez.
- Facilidad de comunicación, ya que todas las estaciones están en la misma línea de comunicación. En vez de observar a seis ó más estaciones a la vez, como sucedía en las redes continuas.
- En una poligonal no hay acumulación de error de escala, ya que se miden todos los lados.
- Es más fácil de ajustar que una cadena y mucho más que una red continúa.
- Dada la velocidad con que se completa, necesita menos publicación intermedia de información.

La distancia entre nodos de una red de poligonales, para un territorio dado, dependía principalmente de: la posibilidad de un ajuste simultáneo, la economía para las posteriores redes

de segundo o tercer orden y sobretodo de la geografía de la región. La densificación -redes de segundo y tercer orden- también se puede hacer con poligonales. Al final, la red de poligonales se construye a base de ‘cuadriláteros’, aunque también se puede formar por ‘triángulos’.

La longitud de lado en las poligonales depende del terreno. En terreno llano pueden tener una media de 10-20 Km, usando torres de observación móviles de 8 a 15 m. En terreno montañoso era conveniente que el valor medio no superara los 40 Km. Los puntos Laplace debían de darse con tanta homogeneidad como se tenía de escala, con el fin de obtener la misma precisión para escala y azimut en la red. Lo recomendado era intentar un punto Laplace en cada nodo de la red poligonal. Si alguno fallaba, principalmente por motivos metereológicos, no era decisivo. En países con malas condiciones climatológicas era posible dar un punto Laplace cada 5 ó 6 vértices. Pero se debía tener especial cuidado en la medición de cada ángulo de la poligonal. Los puntos Laplace, con medición adicional de latitud astronómica, eran útiles por que proporcionaban perfiles del geoide necesarios para la reducción al elipsoide. [1]

11.3.2. Trilateración

Si en una cadena de triángulos medimos cada uno de los lados con MDE, ésta podrá ser calculada geoméricamente, aunque será menos económico que una poligonal. Será susceptible de acumular errores azimutales, por lo que el control con puntos Laplace deberá ser muy denso. Una cadena de trilateración que consiste en simples triángulos no tiene un chequeo interno, como el error de cierre de cada triángulo en la triangulación. Estas cadenas de simple trilateración, sin ninguna observación angular, no eran aconsejadas para redes de primer orden. Sin embargo, una red continua de trilateración puede ser lo suficientemente rigurosa si esta asociada con un adecuado numero de puntos Laplace para su control, y llega a ser la solución adecuada para territorios pequeños o medios. Por ejemplo, en España se observo la región gallega por trilateración incluyendo diagonales. En grandes áreas, podían existir problemas de capacidad de cálculo del ajuste, al igual que ocurría con las redes continuas con medición angular. Trilateraciones puras sin control Laplace fueron utilizadas para segundos órdenes en algunos países, se usaban distanciómetros de microondas en épocas de buena visibilidad. Una figura robustamente medida sería una cadena en la que todos los ángulos y distancias fueran medidos, si esto se acompañaba de un denso control Laplace, tendríamos una primera posibilidad de obtener un red de un orden superior al primero, obtendríamos al menos una red de orden cero del tipo C. Pero antes de que se pusieran realmente en práctica este tipo de redes aparecieron sistemas como los Shoran, la observación a las constelaciones de satélites por doppler como fue el TRANSIT americano o el TSIKADA ruso, o la distanciometría con

medición de fase a las actuales constelaciones GPS, GLONASS y la futura europea Galileo, los llamados GNSS.

11.4. Redes geodésicas actuales GNSS

Actualmente, las redes geodésicas principales han sido observadas ó reobservadas con GNSS, ó más concretamente GPS. Este nuevo tipo de instrumentación se aplicó inmediatamente a su aparición a las redes geodésicas de muchos países, entre ellos España. Lógicamente, las características de las redes cambian al cambiar el tipo de instrumentación con que se observan. Para empezar, los puntos de una red geodésica GNSS ya no tienen porque ser visibles entre si. Por otro lado, las distancias entre puntos no están tan limitada, con GNSS se pueden observar puntos a miles de kilómetros entre si.

Un aspecto principal de la medición precisa de redes geodésicas con GNSS es el empleo de la técnica diferencial. Lo que permite calcular los vectores entre pares de puntos con muchísima precisión (del orden de hasta 0,1 ppm), ya que el posicionamiento absoluto o global no proporciona esa precisión. Para mantenerla globalmente, se aprovechan las estaciones permanentes de redes superiores, públicas, de las cuales se conoce con gran exactitud su posición en un determinado sistema de referencia geocéntrico (tridimensional). Es importante reseñar respecto a los sistemas de referencia locales o regionales -usados hace unas décadas- que han sido sustituidos por sistemas de referencia globales, como el ITRS, con los que es posible enlazar mediante observaciones GNSS. Aun así, los sistemas de referencia locales o regionales siguen siendo útiles para fines cartográficos, pero esta vez la relación matemática con los sistemas globales es conocida con precisión suficiente. Éste, es el caso del sistema ETRS89 europeo, adoptado en 2007 como sistema geodésico de referencia oficial para España [13].

Las redes de este tipo, que afectan a nuestro país, son la red del *Intenational GNSS Service* (IGS) y la red permanente europea de EUREF (EPN). La red del IGS comprende unas 340 estaciones permanentes GNSS repartidas por todo el mundo [3], sus puntos forman el marco de referencia físico para cada solución concreta del sistema ITRS -los llamados ITRFxx-, donde xx es el año de esa solución concreta [9]. En geodesia de precisión, para la consecución de puntos en el sistema ETRS89, exactamente en un marco de referencia concreto ETRFxx, es necesario, primero calcular la campaña en ITRFxx y luego hacer una transformación a ETRFxx [14]. Por ello, las coordenadas, vectores velocidad y observaciones de las estaciones permanentes del IGS forman la red primaria de donde parten las coordenadas del resto de los puntos. En España, actualmente, hay 5 estaciones permanentes GNSS que pertenecen, a su vez, a la red IGS. La red EPN de EUREF [4] es la densificación en Europa de la red del IGS, hay unas 200 estaciones y,

de ellas, unas 30 están en suelo español actualmente. Para el cálculo de la EPN se parte de las coordenadas de estaciones IGS que fueron incluidas en el cálculo del adecuado ITRFxx. Así, hay estación que pertenece a IGS, a EPN y ERGPS simultáneamente

Cuando se realizaron en España las primeras redes de orden cero observadas con GNSS, ya existían las redes geodésicas GNSS continentales y mundiales EPN e IGS, aunque menos densas que hoy en día. Después de algunos estudios previos en varias campañas, se decidió la materialización de una red que cubriera toda la Península Ibérica. Ésta sería la estructura primaria para posteriores redes más densas. Se llamó IBERIA95 y consta de 39 estaciones, 27 de ellas españolas [10]. Se observaron 12 horas durante 5 días consecutivos en 1995, por lo que es una red de orden cero clase B. Se realizó una campaña homóloga para las Islas Baleares que se llamó Balear98 [11]. A partir de estos puntos, se pudo dar coordenadas a una red más densa para toda España y que se llamo REGENTE, con unos 1100 puntos. En ambas redes se aprovechó la monumentación de las redes ya existentes, mejorándolas con centrados forzosos. Se observaron 6 horas en cada vértice, lo que produjo una red de orden cero de clase C. Por último, se han aprovechado las antiguas observaciones angulares con vueltas de horizonte de la ROI y las observaciones GPS de algunas comunidades autónomas para ajustarlas a las coordenadas de los vértices REGENTE, obteniendo coordenadas en el sistema ETRS89 de los aproximadamente 11000 vértices de la ROI.

11.5. Medida de ángulos y distancias en geodesia

Dada su relevancia histórica e incluso actualidad en algunos estudios geodésicos concretos, aún hoy en día, vamos a tratar a continuación en profundidad la medida de ángulos y distancias en geodesia. Estudios geodésicos donde se utiliza actualmente este tipo de instrumentación y técnica medida, son: las redes de microgeodesia para control de deformaciones y las redes de unificación, en un mismo sistema de referencia, de distintos instrumentos de observación geodésicas en un observatorio, como son antenas VLBI, SLR, GNSS y otros.

La medida de ángulos se realiza con teodolitos y la de distancias mediante distanciómetros MDE. Veremos para cada uno de estos instrumentos: sus principios de funcionamiento, métodos de observación, errores, compensaciones, reducciones al elipsoide además de las calibraciones y contrastaciones necesarias para su correcto funcionamiento.

11.5.1. Medida de ángulos

Teodolitos

El teodolito es el instrumento de medición de ángulos, tanto horizontales como verticales. Estructuralmente éste consta de los siguientes ejes:

- Un eje vertical que se consigue su verticalidad con niveles de burbuja. Entonces permite medir en el círculo ó limbo horizontal los ángulos horizontales.
- Un eje horizontal ó de muñones, perpendicular al anterior, que, asociado con el nivel de eclímetro o compensador, permite medir los ángulos verticales o cenitales.
- El eje de colimación, perpendicular al anterior y contenido en el mismo plano que el eje vertical, es el eje con el que se hace la puntería con un sistema óptico telescópico.

A lo largo de la historia de la geodesia han sido varios los tipos de teodolitos usados. Desde principios del siglo XIX hasta mediados del XX, los teodolitos usados eran los llamados teodolitos de micrómetro. Un ejemplo son los de tipo Repsold ó Pistor, usados en la observación de la primera red de primer orden española [7]. Posteriormente, aparecieron los llamados ‘de coincidencia’, que han sido los más ampliamente utilizados en las redes aún existentes. Estos permiten poner en coincidencia lados opuestos de los círculos, mediante un micrómetro que puede ser leído mientras se realiza la visual. Estos teodolitos generalmente han tenido círculos ó limbos de 140 mm de diámetro, como el Wild T3 ó el Kern DKM3, con una precisión de 0,1". El Wild T4, con la misma precisión pero más apropiado para observaciones astronómicas, tiene un círculo de 250 mm de diámetro. Los teodolitos podían constar de otros elementos fundamentales. El nonio, que es la parte dividida que se compara con las divisiones del limbo. El microscopio, para ver mecanismos tan pequeños. Y el micrómetro, que nos permite, a partir de giros, obtener la lectura. Éste, en el caso de limbo de vidrio puede ser de tambor o de estima. La precisión de sus medidas radica sobre todo en el micrómetro. Las observaciones realizadas con los antiguos teodolitos de micrómetro no tienen por que ser necesariamente peores que las realizadas con los modernos, ya que el largísimo tiempo de observación que necesitaban compensaba en muchos casos errores como el de la refracción lateral. Finalmente, aparecieron los teodolitos electrónicos. En ellos, la lectura no se realiza de forma óptica con un microscopio, sino que están constituido por un mecanismo electrónico que realiza la lectura (medida electromagnética de ángulos). Elemento fundamental es el transductor, que es el sensor que convierte la medida sobre el limbo en ángulo. Éstos pueden ser de tipo escobilla, ópticos ó magnéticos, según este grabado el limbo. En principio, los teodolitos electrónicos tienen círculos mucho más pequeños que los anteriores. Después, para pasar esa medida eléctrica a un número binario, se necesitan los llamados codificadores angulares, que suelen ser de tipo óptico o dinámico.

Los teodolitos usados en geodesia se deben estacionar sobre un vértice estable, como es un pilar de hormigón con sistema de centrado forzoso. En la inmensa mayoría de los casos se deben evitar estacionamientos sobre trípode.

Errores y ajustes al teodolito

Requieren, entre otros, los siguientes ajustes previos a su utilización, para evitar que sean fuentes de errores:

- Perpendicularidad entre ejes:
 - Colimación azimutal. Debe asegurarse la máxima perpendicularidad entre el eje de colimación y el horizontal.
 - Perpendicularidad entre el eje horizontal y el vertical.
- Ajuste de las burbujas, para evitar el error de verticalidad.
- Excentricidad de círculos.
- Graduación de círculos.
- Ajuste de la plomada óptica, para estacionamientos sobre trípode o en torres de observación.
- Verticalidad del retículo.
- Funcionamiento correcto del micrómetro.
- Determinación del valor de una división de cada burbuja, en especial si se realizan observaciones astronómicas.

Estos ajustes son más o menos permanentes y no necesitan ser controlados en cada estación sino en revisiones y calibraciones. Los ajustes que hay que hacer en cada estación son:

- Examinar si existen uniones perdidas o elementos deteriorados.
- Centrado sobre la marca de estación.
- Nivelación para conseguir que el eje vertical sea tangente a la línea de la plomada.
- Clara visión del retículo, enfocando su visión del retículo.
- Clara visión del objeto a observar y ausencia de paralajes, enfocando la visión del objeto.
- Clara visión del microscopio de lectura de círculos.

Señales

Otro aspecto importante era el lugar donde se iba a realizar la visual. Las señales a las que se observaba eran de distinto tipo:

- *Heliotropos*, que son un conjunto de espejos que consiguen reflejar la luz adecuada al origen de la visual, donde se encuentra el observador con el teodolito. Son necesarios en observaciones diurnas a grandes distancias.
- *Lámparas eléctricas*, consiguen emitir un fino haz de luz en la dirección del observador. Imprescindibles en observaciones nocturnas.
- *Señales opacas*, son las que requerían menor atención, pero no se ven a grandes distancias ni con poca visibilidad.

Métodos de observación

En una triangulación los ángulos horizontales pueden ser observados de distintas maneras:

- *Vueltas de horizonte*. Se comienza con una estación seleccionada o con una referencia, el teodolito es movido horizontalmente en una dirección visando a cada uno de los vértices que la componen, hasta cerrar en el primero de ellos ó referencia; se repite en círculo inverso. Normalmente, en un vértice se observan varias vueltas de horizonte.
- *El método de pares a la referencia*. Se elige una marca como referencia que sea de visibilidad óptima, o se utiliza un colimador. Entonces, se mide cada ángulo referencia y vértice independientemente. La referencia no pertenece a la vuelta.
- *El método de los ángulos*, los ángulos entre estaciones adyacentes en una vuelta de horizonte son medidos independientemente. La suma de los ángulos de la vuelta requiere una compensación al no sumar 360° .
- *Método de Schreiber*, se miden todos los ángulos entre cada dos direcciones de la vuelta. Requiere gran trabajo, pero permite programar las observaciones de 1^{er} orden evitando la suspensión de observaciones, y conseguir un peso equivalente para las diferentes direcciones como con los dos anteriores.

El método de las vueltas de horizonte es apropiado para distancias corta y si las condiciones lo permiten sin interrupciones,. Los métodos de pares a la referencia ó de Schreiber se aconsejan para distancias largas que requieren numerosas series de medida. El método de ángulos sólo es aconsejado cuando haya que realizar mediciones adicionales a un ángulo determinado para asegurar más precisión.

La regla de Bessel es de uso obligado en todos los métodos expuestos, consiste en cabecear la alidada y lee en la parte opuesta del círculo ó limbo, es decir, medir en círculo directo ó derecha (CD) y círculo inverso ó izquierda (CI). Corrige diversos errores como el de excentricidad.

Otro método, encaminado a evitar errores de graduación del limbo, es utilizar diferentes zonas del mismo. Están relacionados con diferentes tipos de teodolitos: los repetitivos, que permitían la repetición angular, y los reiterativos, que permitían nueva medida angular. Los primeros utilizan el movimiento general -limbo horizontal incluido- de tal forma que, si queremos medir el ángulo entre dos puntos, leeremos la primera dirección con 0° y la segunda no se lee, sino que llevaremos la posición del limbo con esa lectura a la puntería del primero, así tantas veces como sea necesario. Al final, dividiremos la lectura final en el primero entre el número de veces que hayamos desplazado el limbo. Obtendremos una lectura del ángulo entre las dos direcciones con precisión. Los reiterativos no nos permiten ese desplazamiento con precisión. Con ellos, para evitar el error de graduación, lo que hacemos es medir en diferentes zonas del limbo. Dividiremos medio limbo (ya que con Bessel vamos a leer en zonas opuestas) por tantas veces como queramos reiterar. Colocaremos el limbo aproximadamente en esos valores cada vez que comencemos lo que se llama una serie.

Es importante, al final de cada vuelta, cotejar como ha cerrado la vuelta de horizonte. Normalmente, se realiza el cierre en la primera estación observada. Si el cierre es anormalmente grande habría de repetirse la vuelta completa.

En poligonales, generalmente, solo hay dos estaciones a las que observar. En todo caso se recomienda utilizar el método de las vueltas, siempre cerrando las vueltas. Las series en este caso se deben hacer tanto en una dirección (sentido horario) como la opuesta (sentido antihorario).

Número de medidas de cada ángulo

Una *puntería* corresponde a una visual o dirección y su lectura angular en el círculo. Una medida de ángulo corresponde a la diferencia de *lecturas* a ambas visuales. Tales lecturas pueden provenir de varias punterías y no sólo de una. *Lecturas correspondientes* son, por la regla de Bessel, un par de lecturas en CD y CI. Una *serie completa* comprende todas las lecturas hechas –CD y CI incluidas- con el círculo graduado en una posición.

Respecto a las punterías hay que señalar que con algunos teodolitos hay que intentar hacerlas en pares. Un par sería suficiente, pero en malas condiciones se pueden necesitar más. Como regla general, si las lecturas difieren más de $2''$ otro par más sería necesario. Si el rango de las cuatro punterías supera $4''$, un tercer par sería necesario y así se seguiría. Aunque, quizás, sería mejor esperar a mejores condiciones para repetir las punterías.

Dos ó tres lecturas por puntería y dirección son necesarias con los mejores teodolitos. Otra alternativa es hacer 2 ó 3 lecturas del círculo directo en una serie y 2 ó 3 en círculo inverso en la siguiente serie.

En relación con las series, de 10 a 16 series eran observadas, generalmente, en triangulación de primer orden, con 1, 2 ó 3 lecturas correspondientes en cada una ó tantas series con círculo directo como con inverso en cada serie. En las series por reiteración, una marca de referencia ó una estación es escogida como el punto respecto al que las series son definidas. Los grados en círculo directo en este punto son espaciados regularmente en 180° , y los minutos y segundos son espaciados en el rango del micrómetro para minimizar el error propio del micrómetro.

En España el primer orden se observaba por el método de pares a la referencia con 18 series, reiteradas a 10° , y en cada serie 4 lecturas correspondientes a la referencia y al punto visado.

En las poligonales, donde las estaciones Laplace podrían ser más frecuentes, la precisión necesaria de los ángulos de la poligonal puede ser menor que la que se usó para triangulaciones. El tiempo que se tarda en hacer una serie es mucho menor, por lo que se hacía el mismo número de series. Por otro lado, si los puntos Laplace están separados tanto como en triangulación (unos 200 Km), la precisión en azimut de una cadena triangulación de primer orden sería igualada si varios días de observación se gastaran en cada ángulo de la poligonal. Ya que desde este punto de vista, una cadena de triangulación es como 3 ó 4 poligonales. En la red de EEUU los ángulos de las poligonales de orden cero fueron medidos con 16 series en cada una de dos noches, con una tercera noche si estas diferían más de $1''$.

Periodo del día

Con buenos instrumentos, usados de forma adecuada, la peor de las fuentes de error se creía que era la refracción de la luz fuera del plano vertical. El salto que experimentan las lecturas en los días muy calurosos es compensado aumentando el número de punterías. Pero la refracción puede inclinarse más a una cara que a la otra, la repetición por lo tanto aporta poco. Esto es posible que ocurra cuando la visual pasa cerca de una ladera de una montaña ó de terreno con inclinaciones, aunque plano en general. La refracción será entonces mayor a partir del mediodía, cuando el aire cerca del terreno esta sobrecalentado. Éste será más pequeño una hora ó dos tras el amanecer e igual por la tarde. Y por la noche, cuando el aire está más fresco que la tierra, el signo del efecto puede cambiarse. Lo ideal sería por tanto dividir aproximadamente igual entre los cuatro periodos del día citados, omitiendo el periodo después del mediodía si el salto es exagerado. En líneas cortas ó territorio con poca diferencia de temperaturas, estas reglas pueden

ser pasadas por alto. Por la noche, las luces se ven bien y estables pero podemos estar viéndolas separándose lentamente y dar peores resultados que luces aparentemente malas al mediodía. En España y otros países las observaciones se han hecho en su mayoría nocturnas, dando buenos resultados. Pero en climas tropicales secos lo dicho anteriormente sería obligado. Si el método de pares a la referencia o de los ángulos es el usado, las punterías a cada estación deben ser distribuidas por todo el periodo de trabajo, y no concentrarse en uno ó pocos periodos. El método de las vueltas de horizonte asegura automáticamente lo necesario. Salvo excepciones, se puede decir que la precisión con que un ángulo es medido depende de la cantidad de tiempo que se usa en observarlo.

Compensación de una estación

Si los ángulos son todos medidos independientemente alrededor de una vuelta, lo más seguro es que no sumarán exactamente 360° . El error de cierre tendrá que ser distribuido entre ellos, habrá dos opciones, igual o inversamente proporcional a sus pesos. Normalmente, se puede considerar que todos los ángulos tienen el mismo peso, pero se pueden considerar otros aspectos como son el número de series. Aún así, los ángulos raramente se miden independientemente.

Si la observación se hace en vueltas de horizonte, e incluso vueltas con puntos no vistos en todas las vueltas existen, y además la lectura aceptada a la referencia es la media de la primera y la última, no habrá error de cierre. Pero si la lectura de cierre ha sido considerada separadamente, habrá que distribuir el error de cierre en toda la vuelta.

Si el método de pares a la referencia es usado con una sola referencia, no habrá cierre que ajustar. Si dos o más referencias son utilizadas entonces habrá que ajustar, ya sea por sentido común o por mínimos cuadrados

Otro caso se da cuando los ángulos han sido medidos en diferentes combinaciones o con diferentes instrumentos e incluso en años distintos. Se pueden establecer las ecuaciones de condición que relacionan estas observaciones y resolverlo por mínimos cuadrados. Si el ajuste se hace por variación de coordenadas, se puede considerar cada ángulo ó dirección como una observación separada.

Fuentes de error en la medición de ángulos

Las principales fuentes de error de ángulos horizontales son:

- Refracción lateral

- El centrado del teodolito y de la señal a observar (lámpara, heliotropo, etc.)
- No perpendicularidad de ejes o uniones perdidas.
- Desnivelación, afectando más a visuales inclinadas.
- Inestabilidad en monumentación en torre portable.

La refracción lateral es el error menos controlable y más serio. Ni el error de graduación del círculo, ni los errores accidentales de puntería ó de lectura llegan a ser decisivos si se sigue un buen programa de observación con más de 10 series.

Reducción de ángulos

La reducción, en general, consiste en el conjunto de problemas debidos al traslado de las magnitudes indirectamente medidas en la superficie de la Tierra a sus correspondientes valores sobre la superficie de referencia, habitualmente el elipsoide. La reducción es necesaria para poder realizar un aprovechamiento conjunto matemático de los datos, a partir de las propiedades y dependencias geométricas existentes entre los elementos de la superficie del elipsoide. La reducción no debe afectar negativamente a la precisión de las magnitudes medidas. Para ello, sería necesario el conocimiento, con la debida exactitud, de las relaciones geométricas entre la superficie de la Tierra y la del elipsoide en un punto dado. Es decir, conocer magnitudes como son las diferencias de altura, la desviación de la vertical, las anomalías de la gravedad, etc..., lo que muchas veces representa un serio problema. El método más adecuado es el llamado ‘método de proyección’ en que las magnitudes medidas directamente se reducen, matemáticamente con exactitud, de la superficie de la Tierra a la superficie del elipsoide. Las reducciones se realizan, por tanto, por el traslado de las magnitudes medidas directamente a sus proyecciones. Se calculan mediante fórmulas que expresan las correcciones señaladas en función de las magnitudes. Estas fórmulas determinan las posiciones mutuas de la superficie terrestre y del elipsoide de referencia, es decir, las alturas elipsoidales y las desviaciones de las líneas verticales. Éste es un método de translación rigurosa de las magnitudes geodésicas medidas a sus proyecciones sobre la superficie del elipsoide de referencia. Se conserva la ubicación recíproca de los puntos de la superficie terrestre y hacen posible la elaboración rigurosa, independientemente de lo extensa que sea nuestra red geodésica. Para usar este método es indispensable establecer previamente las dimensiones del elipsoide de referencia y su orientación respecto al cuerpo de la Tierra. [12].

Las correcciones a considerar en caso de la observación angular son [15]:

- *Por la desviación de la vertical.* El ángulo medido directamente sobre la superficie terrestre es determinado por un ángulo diedro (entre dos planos), cuya arista es la línea que coincide con el eje vertical del instrumento, es decir con la ‘línea de la plomada’. El

ángulo en el punto correspondiente en la superficie del elipsoide se mide como un ángulo diedro, cuyas caras son dos planos normales y su arista es la normal a la superficie del elipsoide. El ángulo existente entre la línea de la plomada y la normal al elipsoide es decir la ‘desviación de la vertical’ plantea esta reducción. Para casos de ángulos de desviación de la vertical usuales este valor es despreciable.

- *De la sección normal a la línea geodésica.* Cuando se visa de un punto a otro se esta cortando figuradamente a la superficie del elipsoide por un plano formado por la normal del punto de observación y el punto observado; esta curva intersección es la llamada ‘sección normal’ y existe la recíproca, que en general no coincide con la primera. Por otro lado, la línea más corta que une dos puntos en la superficie del elipsoide es la ‘línea geodésica’. Esta línea sólo coincidiría con la sección normal si los dos puntos estuviesen en el mismo meridiano. El valor diferencia entre la sección normal, que es lo realmente observado, y la línea geodésica es igual a la tercera parte del ángulo existente entre las secciones normales recíprocas de los dos puntos implicados, de A a B y de B a A. En este caso, este valor si que tiene que ser considerado.
- *Por altura del punto observado.* Surge de la consideración geométrica de la altura que tiene el punto observado sobre el elipsoide de referencia. El plano de observación observado no coincide con el que deberíamos observar, si se enfocara al mismo punto proyectado por su normal a la superficie del elipsoide. Como regla general se debe considerar en mediciones de primer orden cuando las diferencias de altura sean grandes.
- *Por altura sobre el nivel del mar.* Las líneas de la plomada no son líneas rectas sino curvas, es decir la desviación de la vertical será distinta dependiendo de la altitud sobre el geoide de un punto dado. Pero esta corrección, por lo general, la podemos considerar despreciable.

Ángulos verticales

Se requieren observaciones algo menos precisas que para los ángulos horizontales, ya que la refracción atmosférica hace que buscar grandes precisiones sea imposible en cualquier caso. Las redes geodésicas para control vertical se observan con nivelación geométrica y por eso, las redes geodésicas para control de altitudes no han coincidido tradicionalmente con las de control horizontal. Aun así, no quedaba más remedio que dar algún tipo de altitud a los vértices geodésicos de la forma más precisa posible, dentro de lo que la refracción permita.

Un método utilizado para llevar altura a los vértices fue a través de cadenas, con al menos una línea de visuales verticales recíprocas. La omisión de algunos ángulos verticales, especialmente en largas diagonales, no era irreparable, ya que con buenos lados medidos a buenas horas del día

se podían deducir alturas a partir de cierre de triángulos. En las poligonales, sin embargo, es necesaria la medida de todos los ángulos verticales entre estaciones, ya que en cualquier caso se necesitan las inclinaciones de las líneas para su reducción al horizonte.

El método de observación consiste en lecturas correspondientes en pares de círculo directo e inverso, para eliminar el error por desajuste de burbujas ó de verticalidad. Tres pares deberían ser suficientes. Si la estación es ocupada más de un día, al menos dos pares deben ser observados cada día.

Se aconseja la observación desde el mediodía local hasta 3 horas más, ya que aunque el efecto de la refracción no es menor sí es más uniforme de lugar a lugar y día a día. Ángulos recíprocos observados simultáneamente son mejores que en días distintos, pero este era un objetivo difícil en triangulación. Las mediciones nocturnas son mejores en cualquier caso que las hechas cerca del amanecer o atardecer. La bondad del resultado dependerá de las condiciones meteorológicas entre la línea. Mediciones de presión y temperatura eran útiles para posteriores correcciones.

Calibración de teodolitos

Son el conjunto de operaciones de confirmación y reajuste que tienen como fin el perfecto funcionamiento del teodolito. En general, las podemos dividir en dos grandes grupos: por un lado las calibraciones que introducen algún tipo de incertidumbre por el uso del sistema óptico del teodolito y por otro las que este sistema no interviene.

En las calibraciones en las que interviene el sistema óptico del teodolito, existen varias tipos. Por un lado, el uso de los colimadores multiángulos, los colimadores aseguran la colimación o la visual en una dirección dada. Es normal colocar los colimadores a 0° , 15° , 37° y 90° . Para asegurar que los colimadores se encuentran en esta dirección exactamente existen dos opciones diferentes: la primera sería que existiera un teodolito de orden superior que colocase los colimadores en las posiciones dadas, la otra sería el no saber en principio cual es la posición correcta de los colimadores, pero sabiendo que siempre es la misma, comprobarla en distintas posiciones del limbo. Otra forma de calibración sería el uso de la llamada ‘mesa generadora de ángulos’ y tener una referencia externa. Los ángulos medidos por el teodolito y por la referencia externa son comparados. En el caso de ángulos verticales se suelen utilizar miras de nivelación de alta precisión perfectamente calibradas, que se usan como patrón para comprobar las mediciones de ángulos verticales.

Las calibraciones de teodolitos, en las que no interviene el sistema óptico, también presentan varios tipos. En primer lugar se encuentra la interferometría angular. En este caso se emite un rayo láser, que es variado con precisión a través de un reflector angular, de tal forma que se puede conocer con precisión cual es su giro. Otro tipo de calibración utiliza el llamado polígono óptico, con una construcción que asegura el valor de los ángulos, se coloca sobre el teodolito y se comparan las lecturas de este con los valores del polígono. Por otro lado, está el llamado autocolimador, con un sistema especular que se coloca sobre el teodolito de tal forma que se espera la llegada del rayo proveniente del colimador con unas determinadas características. [16]

11.5.2. Medida de distancias

Para la medida de distancias con precisión, como exige la geodesia, han sido varios los instrumentos utilizados a lo largo del tiempo. En la primera red geodésica española fueron las reglas metálicas ó bimetálicas las pioneras en la medición de bases de unos 4 m de longitud. En España, se midió con este instrumento la base de Madrideo de 14657m con una precisión de 3mm. Posteriormente, aparecieron los hilos invar, que mejoraban la precisión y hacían el trabajo más rápido. Se basaban en unos hilos de una aleación de hierro-níquel, con un bajísimo coeficiente de dilatación, que mediante el uso de poleas y pesos en los extremos, además de microscopios para la lectura, conseguía gran precisión en tramos de hasta 50m. Finalmente, apareció la medida de distancias electromagnética (MDE) que desarrollaremos en este capítulo. Aunque, hoy en día, los GNSS han desbancado a todas estas metodologías. La MDE se sigue aún aplicando hoy en día en trabajos específicos. Los más modernos distanciómetros pueden llegar a superar en precisión a los GNSS en distancias medias ó pequeñas.

La MDE se desarrolló a la vez que los experimentos para la determinación de la velocidad de la luz en el vacío. Actualmente, este valor se conoce con suficiente precisión en los algoritmos de la MDE, tiene un valor de 299792.5 Km/s. Antes de ver los principios de los distanciómetros normalmente usados en la medición de redes geodésicas, es importante no pasar por alto la MDE más precisa de todas, la interferometría. Aunque realmente no tenga una aplicación directa en medición de bases geodésicas.

Interferometría

En la interferometría, se miden incrementos de distancias y no distancias absolutas. Se utiliza una fuente de luz monocromática, generalmente un láser Helio-Neón de $\lambda=632.8$ nm (rojo). Un interferómetro simple podía consistir en un haz láser que incide sobre un prisma partidor, que divide el rayo en dos perpendiculares entre sí. Uno va a un prisma de reflexión total, inmóvil y

de referencia, que lo devuelve al emisor donde llega a un fotodetector. El otro haz va al prisma móvil que se puede mover en la dirección de llegada del haz. Éste, finalmente, también lo devuelve al fotodetector del emisor. En la llegada al fotodetector se comprueba la distinta fase con que llegan los dos rayos. Si están en misma fase habrá un máximo, si son contrarias se anularán y habrá un nulo. Lo que significa, que se pueden medir a partir de una posición dada, los desplazamientos del prisma móvil con precisión mejor que $\lambda/4$. En geodesia se utiliza el interferómetro en diferentes campos, como son: las calibraciones de distinto instrumental, es la tecnología en la que se basan los más modernos instrumentos de medición de gravedad absoluta y en la medición de incrementos de distancias a cuasares en VLBI.

Distanciometría

Se basa en la medición absoluta de una distancia, que tiene por lo tanto un origen y un fin determinado. Por un lado, la ecuación de una distancia es $D=c.t$, donde 'c' es la velocidad de la luz y 't' el tiempo en desplazarse por la trayectoria del rayo. Por otro lado, en función de la longitud de onda λ , la fase ϕ , y el número entero de longitud de onda N la distancia es,

$$D = N\lambda + \frac{\phi}{2\pi} \lambda.$$

Casi todos los medidores de distancia por medio de ondas funcionan por el método de comparación de fase. Una onda portadora, que parte de un extremo, va hasta el otro, donde es reflejada volviendo al origen. La portadora es modulada, y esa onda modulada es la que se usa en la medición. La onda modulada establece en el espacio una especie de cinta métrica cuyas divisiones son longitudes de onda λ . El recorrido de la onda modulada es el doble que la distancia a medir D. Si se usan dos longitudes de onda próximas λ_1 y λ_2 , tendremos

$$D = N\lambda_1 + \frac{\phi_1}{2\pi} \lambda_1 \quad D = N\lambda_2 + \frac{\phi_2}{2\pi} \lambda_2$$

Mientras que el valor de N sea el mismo, el sistema es resoluble ya que se conocen λ_1 y λ_2 , se miden ϕ_1 y ϕ_2 y se calculan D y N. Finalmente, utilizando dos frecuencias moduladoras, el problema se reduce a medir en el instrumento las diferentes fases ϕ que representan la diferencia de fase entre la salida y el retorno para cada frecuencia moduladora.

Pero existe un problema, hay una distancia llamada distancia límite D_{lim} a partir de la cual los observables ϕ_1 y ϕ_2 se repiten cada número entero de distancias límites. Su expresión es

$$D_{\text{lim}} = \frac{\lambda_1 \lambda_2}{2(\lambda_1 - \lambda_2)}$$

Por lo que interesa que λ_1 y λ_2 , sean lo más próximas posibles para que D_{lim} sea lo más grande posible. En realidad el problema se resuelve utilizando 3 o 4 frecuencias moduladoras, de forma que las distancias límites de las combinaciones entre ellas vayan creciendo exponencialmente.

Otro sistema distinto de MDE absoluta es el de ajuste de frecuencia que es el sistema que utilizan los distanciómetros de más alta precisión como el ME-5000. La distancia se calcula a partir de valores de frecuencia que mide el propio instrumento. Esos valores se miden cuando coinciden las fases de salida y de llegada, por lo tanto hay coincidencia de fase. Esta coincidencia de fase existe cuando en el recorrido hay un número entero de longitudes de onda. Dada una distancia, se pueden dar diversas ocasiones como las planteadas, tantas como ajustes de la frecuencia se lleguen a realizar. El número entero de longitudes de onda M se determina indirectamente mediante el ajuste de la frecuencia de modulación en intervalos predefinidos. El proceso de ‘ajuste de frecuencia’ se realiza en dos tiempos. En el primero, se busca una coincidencia de fase en la onda que vuelve hasta localizarla. Es un rastreo y constituye un ajuste aproximado. Después, utilizando el valor determinado, se hacen múltiples mediciones en frecuencia ligeramente mayores y menores que la correspondiente a la coincidencia de fase. La media ponderada de estos valores ofrece el valor final de la frecuencia ajustada, a esa coincidencia de fase concreta.

Las fuentes de la frecuencia son los llamados osciladores. Un oscilador es un dispositivo electrónico que convierte energía de corriente continua en corriente alterna y puede engendrar y mantener una oscilación de una frecuencia determinada por las constantes físicas del sistema oscilante. Los usados en los distanciómetros son los denominados osciladores piezoeléctricos. La piezoelectricidad es una propiedad característica de algunos cristales (como el cuarzo) que al ser sometidos a deformaciones mecánicas modifican la distribución de sus cargas eléctricas. A la inversa, si sometemos a un cristal piezoeléctrico a un campo eléctrico, éste se deforma y late bajo su acción. La lámina de cuarzo, que se usa en los osciladores de cuarzo, se corta en caras planas y paralelas perpendiculares al eje óptico de cristal. A partir de una frecuencia estabilizada se pueden obtener otras muchas múltiplos o submúltiplos, mediante los adecuados circuitos electrónicos de adición o sustracción. Otro tipo de osciladores son los llamados osciladores atómicos, que son muchos más precisos pero que generalmente no son usados en distanciometría. Se usan en geodesia como fuente de frecuencia externa de receptores GNSS y en VLBI. Existen diversos tipos como de rubidio, cesio ó másér de hidrógeno

En cuanto a la fuente de la portadora, se emplean dos tipos fundamentalmente de distanciómetros. Los que emplean longitudes de onda de la región de las microondas (Telurómetros) y las de la región del visible e infrarrojo (Geodímetros). Para la MDE de

microondas se usan como portadoras frecuencias de 2-15 GHz ($15-2 \text{ cm de } \lambda$). Se pueden generar con circuitos transistorizados o con un tubo electrónico del tipo 'Klystron reflex'. La portadora es emitida por una antena que puede ser: parabólica con dipolo en el punto focal, parabólica con montaje Cassegrain, de tipo concha o de tipo bocina. Las ventajas del uso de estas portadoras es que son poco sensibles a interferencias y ofrecen un largo alcance, por el contrario presentan problemas de desvanecimiento de la señal y reflexiones múltiples en superficies lisas como el agua. Su precisión siempre es inferior a los geodímetros. Para la MDE en el óptico se usan como portadoras frecuencias de 400-1200 nm. La MDE en el óptico se basa en fuentes que utilizan el fenómeno de la electroluminiscencia. Los hay de tres tipos: tubo de descarga luminiscente, diodos electroluminiscentes (LED) y tubos láser.

Para la devolución de la señal se emplean sistemas distintos en los telurómetros que en los geodímetros. En el caso del telurómetro, en las microondas, el elemento que devuelve la señal es activo y se llama 'transponder'. En el caso de las portadoras ópticas, de los geodímetros se utilizan prismas de reflexión total ('corner'), elemento pasivo que devuelve un rayo paralelo.

La recepción del retorno es diferente en telurómetros y geodímetros, aunque siempre pretenden separar de la portadora la señal modulada para comparar su fase. En las microondas la demodulación se basa en un proceso heterodino, que consiste en conseguir mediante combinaciones lineales de diferentes frecuencias otras diferentes manteniendo la modulación. En el caso de los geodímetros, la recepción la hace un sistema óptico que consigue la colimación del haz de llegada con un fotodetector. Los fotodetectores pueden ser fotodiodos ó fotomultiplicadores.

En cuanto a la comparación de fase, ésta es la etapa crítica, ya que excepto en los distanciómetros de ajuste de frecuencia, se reduce a hallar el desfase entre la moduladora a su salida o señal de referencia y la de vuelta o retorno. Se puede hacer por medición directa o por búsqueda de nulo. El método directo consistía, ya que hoy esta en desuso, en comparar ambas señales directamente. Usaba técnicas magnéticas o con tubo de rayos catódicos, y el observable era la observación visual de una marca en un círculo graduado. El método de búsqueda de nulo ofrece una mayor resolución. Consiste en realizar una superposición de ambas señales de tipo interferométrico haciendo un desfase controlable y mensurable y función de la fase de una de las señales, que suele ser la de retorno. Existen tres métodos: el del camino óptico ajustable, el goniómetro inductivo y el de medición digital. El primero usa directamente medidas lineales deslizándolo el fotodetector hasta conseguir interferométricamente el nulo. El desplazamiento en unidades lineales es el observable. En el caso del goniómetro inductivo se utilizan fenómenos eléctricos como la inducción, de tal forma que el observable es el giro entre dos elementos

mecánicos. Los digitales son los usados actualmente, el observable es un tiempo que se obtiene por métodos electrónicos. Se usa un contador de pulsos del oscilador principal. A través de la diferencia de tiempos de inicio de la fase de referencia y la de retorno se obtiene el desfase.

Fuentes de error en la MDE

Las fuentes de error accidentales existentes en los distanciómetros las podemos agrupar en dos: las que son proporcionales a la distancia medida y las que no lo son. Desde otro punto de vista, algunos de ellas proceden de los mecanismos del instrumento por lo que son de origen interno, y otras tienen un origen externo al instrumento.

- Proporcionales a la distancia.
 - Índice de refracción. (origen externo)
 - Frecuencia de modulación. (origen interno)
- No proporcionales a la distancia.
 - Constante del equipo. (origen interno)
 - Error cíclico. (origen interno)
 - Centrado mecánico. (origen externo)

Las proporcionales a la distancia generan un factor de escala. En el ajuste de observaciones de una red no afectan a la precisión interna, por lo que puede no ser detectado. Las no proporcionales, sin embargo, presentan valores absolutos.

Los errores por el índice de refracción provienen de la dependencia de la velocidad de propagación de la portadora de las características del medio. El índice de refracción 'n' es una forma de expresar cual es la relación entre la velocidad de la luz en el vacío y la velocidad de la luz en un medio concreto ($n=c/c_m$). El distanciómetro, normalmente, aplicará un determinado índice de refracción en función de unas características del medio estándar, que por lo general no coincidirá con la del medio de observación. Además, la afectación de pequeñas variaciones en el índice de refracción en el cálculo de la distancia es proporcional a ésta.

En cuanto a la frecuencia de modulación propia del oscilador interno del instrumento, ésta viene afectada por un envejecimiento propio que la hace variar, adicionalmente puede sufrir desajustes accidentales. En este caso, también la afectación de pequeñas variaciones en la frecuencia principal en el cálculo de la distancia es proporcional a ésta.

La constante del equipo, también llamada error de cero, se compone de la constante del equipo y la del prisma reflector, aunque se corrigen conjuntamente. En el instrumento, el lugar desde donde parte el rayo, el cero instrumental, es posible que no este exactamente en el centro

mecánico del instrumento de acuerdo al cual se haya estacionado éste. En el prisma reflector, el material que lo forma no tiene el mismo índice de refracción que el medio, esto afectará con un tiempo de recorrido distinto al supuesto. Afectará de forma constante a las medidas.

El error cíclico o de linealidad se repite en la distancia en cada semilongitud de onda. Proviene de la mala medición de la fase y generalmente es despreciable.

Los errores de centrado provienen de la limitación física de estacionamiento del distanciómetro y de colocación del prisma. No es posible colocarlos exactamente en el centro de las marcas que se pretenden medir. Por lo tanto, existe una cierta incertidumbre.

Cuando se expresa, a priori, el error propio de un distanciómetro, éste se expresa de la forma.

$$\sigma_D = \pm \sqrt{a^2 + (bD)^2}$$

En ‘a’ están representados los errores no proporcionales y en ‘b’ los proporcionales. Muchas veces sin ser rigurosamente exacto se da el error esperable como $\sigma_D = \pm(a + b(\text{ppm}))$ expresando ‘b’ en partes por millón.

Reducción de las distancias

Como ya se expuso en el caso de la reducción de ángulos, la reducción consiste en el conjunto de problemas debidos al traslado de las magnitudes indirectamente medidas en la superficie de la Tierra a sus correspondientes valores sobre la superficie de referencia, habitualmente el elipsoide [15].

En el caso de las distancias, las reducciones a aplicar son las siguientes:

- *Corrección meteorológica.* Ésta no es, ciertamente, una reducción, pero es una operación que hay que afrontar antes de realizar las reducciones. Ya explicamos en las fuentes de error que, habitualmente, el distanciómetro usa un índice de refracción estándar para unas supuestas condiciones meteorológicas. Este índice no coincidirá exactamente con el que deberíamos asignar en el caso de una observación dada. Así, la distancia correcta será, siendo ‘n’ el índice de refracción en el momento de la observación,

$$D = D_{\text{medido}} + c_{\text{meteo.}} = D_{\text{medido}} + D_{\text{medido}}(n_{\text{estandar}} - n)$$

Para calcular n existen distintas fórmulas, según se trate de geodímetros ó de telurómetros. Estas fórmulas calculan el valor del índice de refracción del medio en el momento de la observación en función de la presión, temperatura seca y húmeda, presión del vapor de agua y longitud de onda exacta que utiliza el distanciómetro. En el

caso de los geodímetros se usaba normalmente la fórmula de Barrel y Sears y con telurómetros la de Essen y Froome.

- *Corrección al horizonte.* La visual se realizó en puntos con distintas alturas. Hay que hallar cual sería la distancia horizontal entre dos puntos de misma altura, y que ésta sea la media de los dos puntos. En geodesia se aplica conjuntamente al corregir al elipsoide.
- *Corrección a la superficie del elipsoide.* Ahora, se puede bajar la distancia horizontal anterior a las proyecciones de los dos puntos con altura cero. Las fórmulas existentes utilizan el radio de curvatura en la dirección dada a partir de la fórmula de Euler. Normalmente, esta corrección se realiza conjuntamente con la anterior en un solo paso de acuerdo a la siguiente fórmula exacta.

$$D_{\text{cuerda}} = \sqrt{\frac{D^2 - \Delta h^2}{\left(1 + \frac{h_A}{R_\alpha}\right)\left(1 + \frac{h_B}{R_\alpha}\right)}} \quad \text{cuerda entre puntos sobre elipsoide}$$

Como D está medida desde una altura de instrumento i sobre h_A hasta un punto con altura m sobre la marca en B, se puede introducir, directamente, en la fórmula anterior (h_A+i) en lugar de h_A y (h_B+i) en lugar de h_B. Se hace notar que las altitudes h_A y h_B son altitudes elipsóidicas, pero si los valores introducidos corresponden a alturas ortométricas la distancia obtenida estaría referida al geoide.

- *Corrección de la cuerda al arco.* La distancia obtenida en este momento es una distancia recta en el espacio, pero nos interesa la distancia curva entre los dos puntos sobre la superficie del elipsoide. Una corrección posible, aproximada, sería.

$$S = 2 R_\alpha \arcsin \frac{D_{\text{cuerda}}}{2 R_\alpha}$$

Calibración de distanciómetros

Para la calibración o contrastación de distanciómetros se usan distintas técnicas, según estemos buscando un error proporcional a la distancia ó un error constante ó cíclico [2].

- *Determinación de la constante del equipo o error de cero.* Se basa en la medición sobre tres hitos perfectamente alineados, se utilizan sistemas que minimizan los errores de centrado mecánico. Se mide la distancia total y las intermedias, de las múltiples medidas resultantes para cada distancia se deduce este error. Se realiza un procedimiento algo distinto para el caso de los telurómetros.
- *Determinación del error cíclico.* Se utilizan bancadas de precisión en laboratorio y un interferómetro con mediciones continuas, alejando ó acercando uno sobre otro.
- *Determinación del error de escala.* En este caso, hay dos opciones: por un lado la comprobación en gabinete de la ‘frecuencia intermedia’, y por otro la comprobación en

diferentes distancias comprendidas en el rango del aparato, medidas también con medios más precisos, pero que confirmen dicho error.

Bibliografía

- [1] Bomford, G., (1975), Geodesy, Imprenta de la Universidad de Oxford. ISBN 0-19-851919-2
- [2] Ferrer, R., Piña, B., Núñez-García, A., Valbuena, J. L., Martínez, I., (2007), Distanciometría Electrónica de Alta Precisión. Servicio de Publicaciones de la Universidad de Cantabria. ISBN 84-86928-52-4
- [3] IGS (2007). Página oficial del IGS. <http://igscb.jpl.nasa.gov>.
- [4] EPN (2007). Página oficial de la EPN de EUREF . <http://epncb.oma.be>.
- [5] ERGPS (2007). Servicio de datos e información de la ERGPS del IGN. <ftp://ftp.geodesia.ign.es>.
- [6] Asplund, L., (1964) Report on specifications for fundamental networks. Trav. Ass. int. Géod.
- [7] Paladini, A., (1969) La Red Geodésica Española. Boletín de Información nº5 del Servicio Geográfico del Ejército. 1er Trimestre de 1969..
- [8] Principal Triangulation – Spain – European Datum - Internacional Ellipsoid – 1950. AMS.
- [9] IERS (2007). Página oficial del Servicio Internacional de Rotación de la Tierra. <http://www.iers.org>.
- [10] Rodríguez, E., Caturla, J. L., (1999) IBERIA95 Red Peninsular de Orden Cero. IGN. ISBN 84-95172-02X
- [11] Rodríguez, E., Hernández, F., Caturla, J. L., Barrado, J., Los proyectos Balear 98 y REGENTE Baleares. IGN y SGE.
- [12] Zakatov, P. S. (1981), Curso de Geodesia Superior.) Editorial MIR.
- [13] Ministerio de la Presidencia. Real Decreto 1071/2007 por el que se regula el sistema geodésico de referencia oficial en España. BOE del 29/08/07.
- [14] Boucher, C., Altamini, Z., Memo: Specifications for reference frame fixing in the analysis of a EUREF GPS campaign. <http://lareg.ensg.ign.fr/EUREF/memo2007.pdf>.
- [15] Martín Asín, F., “Geodesia y Cartografía Matemática”, Editorial Paraninfo.
- [16] Centro Español Metrología (2004) Procedimiento DI-034 para la calibración de teodolitos. ISBN 706-04-004-4.
- [17] Quirós et al (2003): “National Report of Spain: REGENTE project finished”. Mitteilungen des Bundesamtes für Kartographie und Geodäsie. Band 29. EUREF Publication No.12. Report on the Symposium of the IAG Subcommission for Europe (EUREF). ISBN 3-89888-873-8 pp 292.

Tema 12. Sistemas de posicionamiento y navegación: GPS, EGNOS, Galileo, GLONASS. Sistemas de corrección diferencial y de aumentación. Estaciones virtuales GPS.

12.1. Introducción

En los últimos años, ha sido sorprendente el auge de los llamados sistemas de navegación y posicionamiento, en inglés Global Navigation Satellite Systems (GNSS), más veterano es el estadounidense Global Positioning System (GPS) que desde finales del siglo pasado ya era de uso generalizado. El sistema ruso, antes soviético, fue el segundo en aparecer, su nombre Global'naya Navigatsionnaya Sputnikkovaya Sistema (GLONASS), aunque ambicioso en principio aun no ha llegado a superar al Sistema GPS en difusión. La causa es sin duda la falta de estabilidad del Sistema, no se han cumplido los plazos en la puesta y mantenimiento en órbita de los satélites. Por último Galileo, el sistema europeo, que todavía no esta desplegado. Actualmente tan solo un satélite de pruebas esta en órbita. Aunque con serios problemas en la gestión del proyecto, se promete un despliegue total en un corto periodo de tiempo.

Tan pronto se extendió el uso del GPS nació la necesidad de aumentar su precisión para navegación. Es decir, conseguir en tiempo real o casi real la mayor precisión posible en las coordenadas. Para ello nacieron los sistemas de corrección diferencial ó de aumentación como un subsistema complementario a los GNSS. Estos consisten en el apoyo en estaciones terrestres de referencia con coordenadas conocidas con gran precisión, para la emisión de correcciones fruto de la comparación de lo observado con lo ya conocido. Si la emisión de estas correcciones se emiten desde tierra ya sea por radio, radio enlace ó Internet tendremos los llamados Sistemas de Aumentación Terrestre (GBAS), pero si estas correcciones son emitidas por satélites tendremos los Sistemas de Aumentación Espacial (SBAS). En los SBAS se usan satélites geoestacionarios, los Estados Unidos mantienen la constelación WAAS, Europa EGNOS, Japón MSAT e India GAGAN. GLONASS y Galileo ya nacieron proyectados para el uso de sistemas de aumentación, en el caso de Galileo incluso del tipo SBAS con los satélites EGNOS que son parte integrante del sistema desde un principio.

Las técnicas diferenciales GNSS, donde una estación con coordenadas conocidas es usada como estación de referencia, permiten conseguir precisiones submétricas e incluso de orden centimétrico. Si usamos observaciones de código la técnica se llama GPS diferencial (DGPS). Mucha más precisión se puede obtener a partir de observaciones de fase es la técnica llamada

Real Time Kinematics (RTK). Ambas técnicas degradan su precisión en función de la distancia a la estación de referencia.

Si el fin es proporcionar un servicio de posicionamiento diferencial para una gran área, varias estaciones de referencia tienen que ser utilizadas. A diferencia de las dos técnicas anteriores, estas estaciones no tienen que ser consideradas de forma independiente con cada estación cubriendo solamente una parte de la superficie. Las observaciones de una red de estaciones tienen que ser premodeladas con el fin de generar modelos de corrección regionales. Basándonos en los parámetros del modelo y en la posición aproximada del usuario podremos predecir correcciones individuales de una estación de referencia virtual, *Virtual Reference Station* (VRS), que nos permita el posicionamiento diferencial.

12.1. Sistemas de posicionamiento y navegación

Son cada uno de los sistemas que a partir de la recepción de señales emitidas por satélites nos permiten conocer nuestra posición con una precisión razonable. Estos son, el GPS, GLONASS y Galileo, aunque este último todavía está en fase de desarrollo. Además, está en proyecto un sistema chino que se llamará BEIDOU.

12.2.1. Sistema GPS

El GPS es un sistema de radio navegación por satélite que provee a usuarios de coordenadas precisas de posicionamiento tridimensional e información sobre navegación y tiempo. Se empezó a desarrollar en 1973 como mejora del antiguo sistema americano TRANSIT que usaba medidas Doppler. Aunque diseñado para uso militar en principio, pasó a ser un sistema de uso civil con ciertas restricciones, siendo utilizado para fines geodésicos prácticamente desde 1983. La configuración final del sistema fue alcanzada en 1994, con 24 satélites utilizables. Desde entonces se ha convertido en el principal sistema (o al menos el más versátil) para posicionamiento de precisión.

Para el estudio del sistema, lo dividiremos en sus tres principales segmentos o sistemas, el espacial, el de control y el de usuario. Son muchas las publicaciones que desarrollan el GPS con más detenimiento, entre ellos [6], [7], [8] y [9].

El Segmento Espacial.

Está constituido por la constelación de satélites NAVSTAR, cuyas siglas significan sistema de navegación con cronometraje y medición de distancia, *NAVigation System with Timing and Ranging*. La constelación está pensada para dar cobertura a cualquier hora del día y en cualquier parte del mundo. El segmento espacial proporciona cobertura con 4 a 8 satélites por encima de cualquier horizonte de cualquier lugar de la Tierra. Si la máscara de elevación en la observación se reduce a 10° , se pueden llegar a observar hasta 10 satélites. Si la máscara se reduce a 0° , se puede ver la mitad de la constelación: 12 satélites.

Las características orbitales son las siguientes:

- 6 órbitas casi circulares a 20180 km de altitud, nombradas A, B, C, D, E, F, con 55° de inclinación.
- 4 satélites por órbita, 24 en total.
- Periodo de 12 horas de tiempo sidéreo. Por lo tanto La configuración se repite 4 minutos antes cada día solar.
- Visibilidad de cada satélite: 5 horas.
- Existen hasta 4 satélites desactivados y disponibles como reserva “spares”.

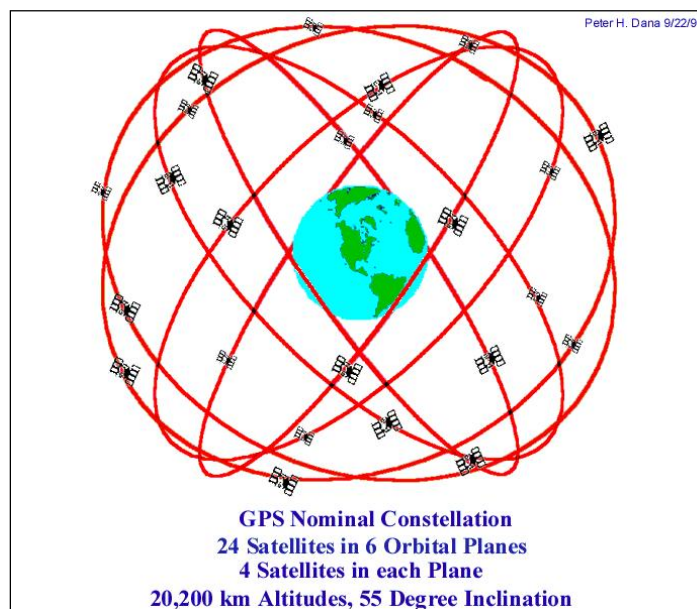


Figura 1. Segmento espacial

Los satélites de diferente generación se agrupan en los llamados "bloques", existiendo por el momento los bloques I, II, IIA, IIR, IIF (y III). Cada generación aumenta la vida útil del satélite, el tiempo posible de operación sin contacto con el segmento de control y la estabilidad de los osciladores a bordo.

- Bloque I.- Los satélites de este bloque fueron pioneros, aportaron experiencia para el resto de los bloques. Actualmente no queda ninguno en funcionamiento. Fue experimental, estos satélites fueron lanzados desde la Base de Vandenburg en California, entre los años 1978 y 1985.
- Bloque II.- Para reducir costes se decidió poner en órbita los satélites del bloque II mediante vehículos transbordadores espaciales. Se ponían en órbita tres satélites en cada viaje. El desastre del transbordador Challenger en 1986 retrasó el programa. Actualmente se están lanzando con el cohete MLV Delta2. Los satélites de este bloque se reparten en cuatro generaciones sucesivas: Bloque II, IIA, IIR los lanzados actualmente, IIR-M con el nuevo código L2C y los inmediatos II-F con la nueva portadora L5.
- Bloque III.- Está en proyecto, puede que se sustituyan sus satélites por el modelo II-R, evolución de los del Bloque II. El primer satélite estaba previsto para el 2009 y toda la constelación para el 2030. Se prevé la incorporación de dos nuevas portadoras para uso civil. Posiblemente irán provistos de relojes con osciladores más precisos de hidrógeno, tendrán mejor capacidad antiinterferencia para las portadoras L1 y L2.

La puesta en órbita se ha realizado principalmente con dos tipos de cohetes los Atlas F y Delta 2. El lanzamiento se realiza utilizando distintas orbitas en cada fase. La primera y segunda etapa sitúan al satélite en “*órbita de aparcamiento*”, una órbita elíptica con apogeo y perigeo a 870 y 180 km. de altitud respectivamente, quedando la tercera etapa del cohete y su carga útil (satélite). La tercera etapa se sitúa en “*órbita de transferencia*” a 20.180 Km, con fuerte excentricidad. Los cohetes de maniobra del satélite retocan hasta establecer la órbita de servicio definitiva. Desde el control en tierra evidentemente se puede manipular el funcionamiento, permitiendo corregir la órbita o incluso cambiar la posición dentro de la propia órbita.

Para la identificación de los satélites se pueden usar varios sistemas:

- El número NAVSTAR (SVN), que es el número de orden de lanzamiento del satélite en la constelación.
- Por el número de órbita, se usan letras de la ‘A’ a la ‘F’ y la posición que ocupa en la órbita con números.
- Por el número de catálogo dentro de todos los satélites de la NASA.
- Por la identificación internacional constituida por el año de lanzamiento, el número de lanzamiento en el año y una letra según el tipo.
- Por el número IRON. Número aleatorio asignado por la Junta de Defensa Aérea Norteamericana (NORAD) de Estados Unidos y Canadá.

- Pero la forma generalizada de identificarlos es por su PRN o ruido pseudo aleatorio, que es característico de cada satélite NAVSTAR, y es la forma como los distingue el receptor.

LAUNCH ORDER	PRN	SVN	LAUNCH DATE	FREQ STD	US PLANE	LAUNCH ORDER	PRN	SVN	LAUNCH DATE	FREQ STD	US PLANE
II-1		14	14 FEB 1989			IIA-23	4	34	26 OCT 1993	Rb	D4
II-2		13	10 JUN 1989			IIA-24	6	36	10 MAR 1994	Rb	C1
II-3		16	18 AUG 1989			IIA-25	3	33	28 MAR 1996	Rb	C2
II-4		19	21 OCT 1989			IIA-26	10	40	16 JUL 1996	Cs	E3
II-5		17	11 DEC 1989			IIA-27	30	30	12 SEP 1996	Cs	B2
II-6		18	24 JAN 1990			IIA-28	8	38	6 NOV 1997	Cs	A3
II-7		20	26 MAR 1990			***IIR-1		42	17 JAN 1997		
II-8		21	2 AUG 1990			IIR-2	13	43	23 JUL 1997	Rb	F3
II-9		15	1 OCT 1990			IIR-3	11	46	7 OCT 1999	Rb	D2
IIA-10	32	23	26 NOV 1990	Rb	E6	IIR-4	20	51	11 MAY 2000	Rb	E1
IIA-11	24	24	4 JUL 1991	Cs	D6	IIR-5	28	44	16 JUL 2000	Rb	B3
IIA-12	25	25	23 FEB 1992	Rb	A5	IIR-6	14	41	10 NOV 2000	Rb	F1
*IIA-13		28	10 APR 1992			IIR-7	18	54	30 JAN 2001	Rb	E4
IIA-14	26	26	7 JUL 1992	Rb	F2	IIR-8	16	56	29 JAN 2003	Rb	B1
IIA-15	27	27	9 SEP 1992	Cs	A4	IIR-9	21	45	31 MAR 2003	Rb	D3
IIA-16	1	32	22 NOV 1992	Cs	F6	IIR-10	22	47	21 DEC 2003	Rb	E2
*IIA-17	29	29	18 DEC 1992	Rb	F5	IIR-11	19	59	20 MAR 2004	Rb	C3
*IIA-18		22	3 FEB 1993			IIR-12	23	60	23 JUN 2004	Rb	F4
*IIA-19		31	30 MAR 1993			IIR-13	2	61	6 NOV 2004	Rb	D1
IIA-20	7	37	13 MAY 1993	Rb	C5	IIR-14M	17	53	26 SEP 2005	Rb	C4
IIA-21	9	39	26 JUN 1993	Rb	A1	IIR-15M	31	52	25 SEP 2006	Rb	A2
IIA-22	5	35	30 AUG 1993	Rb	B4	IIR-16M	12	58	17 NOV 2006	Rb	B5

Figura 2. Satélites de los bloques II/IIA/IIR/IIR-M en Septiembre de 2007

Los últimos satélites son mucho más fiables que los de la primera generación. Tienen sistemas alternativos computables desde tierra. Un satélite queda fuera de servicio por avería o envejecimiento de los paneles solares, falta de capacidad de los acumuladores, averías no computables en los sistemas eléctricos o agotamiento del combustible de maniobra y recuperación de órbita.

Los relojes son la fuente de la frecuencia, patrón de generación de las frecuencias emisoras L1 y L2. Los tipos de reloj pueden ser: Oscilador de cristal de cuarzo, de rubidio o de cesio y más de hidrógeno. Por bloques, las características de los relojes han sido las siguientes. De los satélites del Bloque I, 4 llevaban oscilador de cuarzo, 3 de rubidio y 3 de cesio. Los del Bloque II incorporan osciladores atómicos, de cesio o de cesio y rubidio. Normalmente tienen dos patrones de frecuencia cesio y dos patrones de frecuencia rubidio.

Los del bloque III, posiblemente llevarán osciladores de más de hidrógeno.

Tipo de reloj	Frecuencia de oscilación (GHz)	Estabilidad por día ($\Delta f/f$)	Tiempo para perder un segundo
Oscilador de Cristal de Cuarzo	0.05	10^{-9}	30 años
Rubidio	6834682613	10^{-12}	30000 años
Cesio	9192631770	10^{-13}	300000 años
Más de Hidrógeno	1420405751	10^{-15}	30000000 años

Figura 3. Precisión de Osciladores usados en GPS.

Otra característica importante del GPS es la utilización de su propio sistema de tiempo, el Tiempo GPS (TGPS). Mantiene una desviación constante de 19 segundos Sistema Internacional con el Tiempo Atómico Internacional (TAI). Coincidió con Tiempo Universal Coordinado (UTC) en la época inicial de TGPS, enero 6,0 del año 1980. Su unidad es el segundo del Sistema Internacional. Por lo que es una escala uniforme y continua. La diferencia entre el GPST y UTC no es constante. Se relacionan a través de los “segundos intercalares”.

El control de tiempos en GPS es el principal factor de imprecisión. Por otra parte, el GPS es una herramienta muy válida para la transmisión de tiempos.

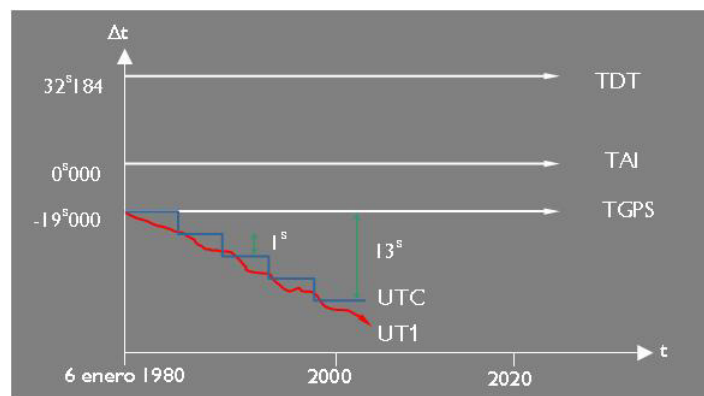


Figura 4. Relación de escalas de tiempo con tiempo GPS

A partir del Tiempo GPS se definen las llamadas unidades derivadas del tiempo GPS. Estas son, el día de Año (DOY): número del día dentro del año calculado en el intervalo 1-365, 1-366 en bisiestos, la semana GPS (GPS week): semanas transcurridas desde el 6 de enero de 1980 a 0h de UTC, el día de semana que es la numeración de días dentro de cada semana en la que corresponde el 0 para el domingo y el 6 para el sábado

Día convencional	21 de junio de 2003.
Día de Año (DOY):	172
Semana GPS (GPS week):	1223
Día de semana GPS	6 (sábado)

Tabla 1. Ejemplo de denominación de día GPS.

El diseño de la estructura de la señal GPS, está condicionado por los objetivos que persigue el propio sistema de navegación y que distinguen al propio sistema de los anteriores sistemas de posicionamiento. Algunos de los objetivos más importantes son:

- Alta precisión en posición en tiempo real
- Navegación en tiempo real para usuarios con altas velocidades

- Cobertura mundial
- Tolerancia a las interferencias intencionadas y no intencionadas
- Conseguir una posición inicial en un tiempo razonable

A excepción de la cobertura mundial, que depende del despliegue orbital de los satélites, los otros puntos tienen relación con:

- Elección de la banda de radio frecuencias
- Utilización de dos portadoras
- Modulación de las portadoras por medio de códigos
- Señales de amplio espectro
- Diseño de las antenas

Debemos tener en cuenta que para obtener posiciones instantáneas en tiempo real, hay que realizar medidas simultáneamente de varios satélites (mínimo 4), para lo cual cada satélite va a tener asignado un código único (PRN) con el fin de identificar sus señales y además vamos a conocer la posición de cada uno de los satélites en tiempo real por medio del mensaje de navegación.

Por otro lado, la obtención de velocidad (con precisión a nivel cm/s) requiere medidas Doppler sobre portadoras de longitud de onda centimétrica. La señal GPS va a consistir en ondas electromagnéticas pertenecientes a la banda L del espectro debido principalmente a que los efectos del retardo ionosférico son sustancialmente menores que en otras bandas y es más fácil la asignación del canal de comunicación (uso de parte del espectro electromagnético).

Cada satélite transmite señales en dos frecuencias, siendo estas las señales de navegación (códigos) y los datos de navegación y sistema (mensaje). Los códigos son modulados sobre la frecuencia portadora en forma de secuencias llamadas "ruido pseudoaleatorio" o "pseudo random noise" (PRN). Las dos frecuencias portadoras en la banda L (valores usados están en la banda L de radiofrecuencias que abarcan desde 1 Ghz hasta 2 Ghz) son derivadas de una frecuencia fundamental de 10,23 MHz:

- L1: $154 \times 10,23 \text{ MHz} = 1575,42 \text{ MHz}$ ($\lambda = 19,05 \text{ cm}$)
- L2: $120 \times 10,23 \text{ MHz} = 1227,60 \text{ MHz}$ ($\lambda = 24,45 \text{ cm}$)

La señal L1 contiene dos códigos: el código de alta precisión P y el código menos preciso C/A (course acquisition). La señal L2 sólo contiene el código P. Actualmente, también un código C es transmitido por la banda L2 en los satélites del bloque IIR(M). En el futuro cercano también se emitirá una tercera portadora L5 de 1176,45 MHz. Más información podemos encontrar en [5].

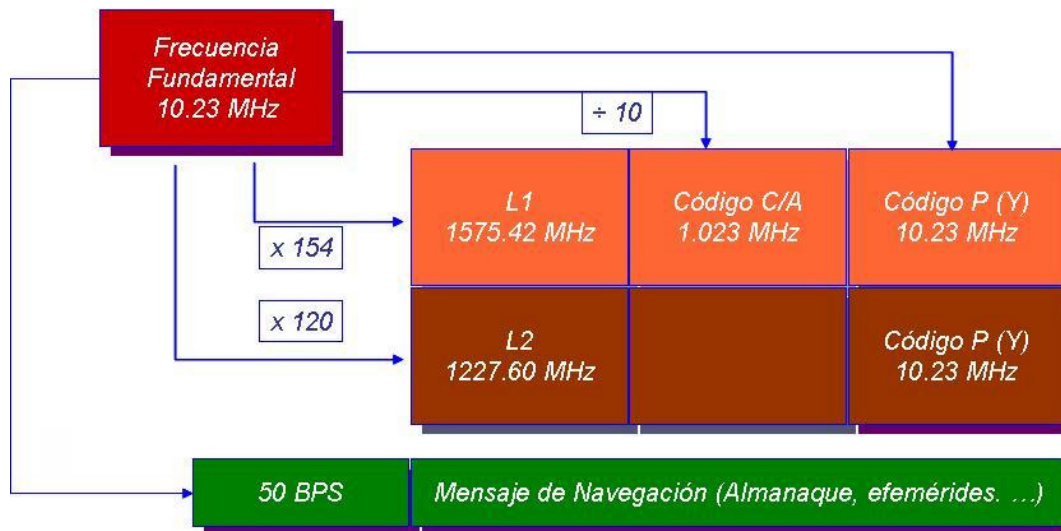


Figura 5. Señales transmitidas por el GPS

El Segmento de Control

Es el que monitoriza el sistema GPS. Tiene la función de realizar el seguimiento continuo de los satélites, calcular su posición precisa desde puntos de situación conocida, la transmisión de datos y la supervisión necesaria para el control diario de todos los sistemas de satélites.

En un principio había cinco puntos de control ó emplazamientos de las instalaciones del segmento de control. El principal en Colorado Springs y otras cuatro secundarias distribuidas homogéneamente en longitud a lo largo del perímetro terrestre. Posteriormente se Cabo Cañaveral en Florida y más tarde 6 puntos de control pertenecientes al *Nacional Imageary and Mapping Agency* (NIMA).

Todas ellas son **Estaciones Monitoras**, y su función es recibir continuamente las señales de los satélites que estén sobre el horizonte en receptores GPS. Estos datos son enviados a La **Estación Maestra** donde se establece con gran precisión las órbitas de los satélites, los estados de relojes. Todos estos datos son transmitidos por las llamadas **Antenas Terrestres** y se almacenan en la memoria del correspondiente satélite, formando parte esencial del mensaje del satélite para su radiodifusión. La sincronización del tiempo de los satélites es una de las misiones más importantes del segmento de control. Por ello la estación principal de control está conectada directamente con el tiempo estándar del Observatorio Naval de los Estados Unidos (USNO).

Los datos que son calculados y posteriormente extrapolados al usuario son básicamente:

- Efemérides Transmitidas (posiciones de los satélites) o parámetros orbitales de cada satélite.
- Datos Ionosféricos-troposféricos
- Estados de los relojes de los satélites

Una vez determinados los parámetros anteriores, la misión es emitirlos a los satélites para que éstos puedan transmitirlos a los usuarios. También prevén y realizan maniobras, lanzamientos, reprogramaciones, desactivaciones y otros.

El Segmento de Usuario

Lo constituyen todos los receptores GPS y sus programas de procesos de datos. La oferta de modelos va en continuo aumento. Su finalidad es igualmente variada y sus características son acordes con ella: navegación, topografía y geodesia y/o transferencia de hora. A la hora de elegir un equipo habrá que tener en cuenta una serie de consideraciones como el tipo de observables que es capaz de registrar, la capacidad de memoria, el peso, tamaño, consumo etc.

Aun así, presentan determinados principios comunes y evidentemente todos ellos contienen los elementos suficientes para recibir la señal y procesarla. Los componentes imprescindibles son:

- Antena con preamplificación
- Sección de Radio Frecuencia
- Microprocesador
- Oscilador
- Fuente de Alimentación
- Dispositivo de control
- Dispositivo de almacenamiento.
- Los elementos accesorios pueden ser, el interfaz de usuario, memoria, dispositivos de puesta en estación, puertos de comunicaciones, etc.

La antena y circuito de antena convierten la radiación electromagnética procedente de los satélites en señal eléctrica y la amplifica para que pueda ser tratada en la etapa de radiofrecuencia. Debe ser omnidireccional. Diseñadas para las portadoras L1 ó L1 y L2. El tipo más frecuente es el realizado con cerámica y denominado microstrip. La Sección de Radiofrecuencia, también conocida con el nombre de “canales”, recoge la señal eléctrica de los

satélites que llega desde la antena y la sigue en todo momento sin interrupciones por uno o varios canales, además realiza tratamientos previos con la señal como mezclas ó filtros.

El Microprocesador controla todo el sistema. Permite el trabajo en tiempo real (navegación) utilizando pseudodistancias, calculando continuamente la posición. El Oscilador es el encargado de generar la frecuencia patrón para control del reloj y réplica de códigos y longitudes de onda GPS. El más usual es de cuarzo, pero también puede ser una frecuencia externa al receptor: Cesio, Rubidio, Máser de Hidrógeno, que mejore la precisión del oscilador del receptor GPS. La Fuente de alimentación es el elemento encargado de aportar la energía eléctrica para los demás componentes del receptor, ya sea batería recargable ó directamente a la red eléctrica.

El dispositivo de control es el medio de comunicación con el receptor, a través de él se asignan los parámetros de trabajo, se realizan diagnósticos, y más funciones. El dispositivo de almacenamiento graba los datos observados, ya sean coordenadas ó directamente observaciones.

12.2.2. Sistema GLONASS

GLONASS es el correspondiente ruso del estadounidense GPS. Es un sistema de radio navegación por satélite que provee a usuarios de coordenadas precisas de posicionamiento tridimensional e información sobre navegación y tiempo. También tiene un origen eminentemente militar y nacieron para sustituir al anterior sistema Tsikada que se basaba en mediciones Doppler. En 1982 se lanzo el primer satélite, y aunque en 1997 había hasta 25 satélites orbitando, la situación empeoró hasta que en 2002 solo quedaban unos 10. Actualmente se esta en proceso de completar la constelación, pero muchos de ellos presentan problemas frecuentemente y son desactivados. Información detallada podemos encontrar en [1] y [4].

Para el estudio del sistema, lo dividiremos en sus tres principales segmentos o sistemas, el espacial, el de control y el de usuario.

El Segmento Espacial

La constelación de satélites GLONASS completa consistiría de 24 satélites orbitando en tres planos con 64,8° de inclinación cada uno, e igualmente separados en 120° de longitud. En cada plano orbitan 8 satélites, uno de ellos se encontraría es espera por si hay algún incidente con los demás. Orbitan a 19100 Km de altitud con un periodo aproximado de 11h15m.

Los satélites GLONASS se pueden clasificar por generaciones ó bloques según su antigüedad.



Figura 6. Satélite GLONASS

Bloque I.- Es la formada por los prototipos previos. Tres réplicas y 18 prototipos fueron lanzados entre 1982 y 1985. Su vida Los satélites de este bloque fueron pioneros, aportaron experiencia para el resto de los bloques. Actualmente no queda ninguno en funcionamiento.

Bloque II.- Formada por vehículos estabilizados de tres ejes, con un peso de unos 1250 Kg y con un modesto sistema de propulsión para maniobrar. Eran lanzados tres a la vez desde el Cosmódromo de Baikonur (Kazajistán) usando cohetes tipo Proton-K Blok-DM-2 o Proton-K Briz-M. Hay tres tipos, los del Bloque IIa, Bloque IIb y Bloque IIv. Los del Bloque IIa mejoraban los estándares de frecuencia y tiempo sobre prototipos y la estabilidad de la frecuencia, fueron lanzados en 1985-86. El Bloque IIb tenía una vida útil de 2 años. Fueron lanzados 12 en 1987, aunque la mitad fracasaron en su lanzamiento. Los del Bloque IIv, fueron usados entre 1988 a 2000, se llegaron a lanzar unos 25. Se diseñaron con una vida útil de tres años aunque alguno casi la dobló.

Bloque III.- La actual generación de satélites llamada también GLONASS-M. Se lanzaron a partir del 2001. Tienen una vida útil de unos 7 años y un peso de unos 1480 Kg. Usan relojes de Cesio. Hasta Octubre del 2007 han sido 11 los satélites lanzados usando el mismo tipo de cohetes.

Bloque IV.- La generación inmediata, son diseñados con una vida útil de hasta 12 años, un peso de aproximadamente 750 Kg. y una portadora adicional en la banda L. Se espera su primer lanzamiento en 2008. Dado su poco tamaño y peso su lanzamiento es más sencillo, permiten lanzamientos desde otros cosmódromos con naves como la Soyuz o lanzar hasta 6 a la vez con los mismos cohetes tipo que los usados en el Bloque II y desde Baikonur.

Los satélites GLONASS, se denominan *Uragan* en ruso que en español significa huracán. Pero para distinguirlos entre ellos se puede usar:

- Número para satélites operacionales consecutivo.
- Número COSMOS. Para todos los lanzamientos rusos.
- Posición orbital que se basa en dígitos, 1,2 ó 3 para la órbita, y otro número de 1 a 8 para la posición dentro de cada órbita.
- Canal de la frecuencia, que en cada satélite es diferente, y es la forma como les distingue el receptor. Cada satélite emite con una frecuencia diferente a diferencia de GPS.

Los satélites del sistema GLONASS están equipados con relojes de cesio, estos son corregidos dos veces al día, lo que permite una precisión de 15 nanosegundos en la sincronización de tiempos de los satélites respecto al Sistema de Tiempos GLONASS.

El Sistema de Tiempos GLONASS (GLONASST) es generado en la Central de Sincronización de Tiempos de Moscú, cuyos relojes de hidrógeno tiene una variación diaria inferior a 5×10^{-14} . El sistema GLONASS proporciona referencias de tiempo en el Sistema de Tiempos GLONASS, mantenido en Moscú y en el sistema UTC (CIS), mantenido en el Centro Meteorológico Principal del Servicio Ruso de Tiempos y Frecuencias de Mendeleev (VNIIFTRI). El uso de UTC en lugar del TAI (Tiempo Atómico Internacional) se debe a que el TAI es un sistema de tiempo continuo que no tiene en cuenta la disminución de la velocidad de la tierra respecto al Sol, que en la actualidad implica 1 segundo al año aproximadamente, lo que significaría un problema para la sincronización con el día solar. El tiempo GPS (GPST) no es incrementado un segundo cada año, por lo que la diferencia de tiempos GLONASS y GPS no es igual cada año.

El sistema GLONASS emplea radioseñales transmitidas de forma continua por satélites. Cada satélite de la constelación transmite dos tipos de señal, L1 de precisión estandar (SP) y L2 de alta precisión (HP). Emplea el sistema de múltiple acceso por división en frecuencia. Esto significa que cada satélite transmite señales en su propia frecuencia, lo que permite su identificación.

Las frecuencias de las portadoras de cada satélite vienen dadas por la expresión:

$$f = \left(178.0 + \frac{K}{16} \right) \cdot Z \quad (\text{MHz})$$

El valor 'K' es un entero entre -7 y +12, Z=9 para L1 y Z=7 para L2, con lo que la separación entre frecuencias contiguas es de 0.5625 MHz para L1 y 0.4375 MHz para L2. Los valores empleados para K han sido :

- Hasta 1998 : entre 0 y 12
- De 1998 hasta 2005 : entre -7 y 12
- Desde 2005 : entre -7 y 4

Si bien estos valores pueden ser modificadas en circunstancias excepcionales. Existen satélites que transmiten a la misma frecuencia, pero están situados en las antípodas, por lo que no pueden ser “vistos” simultáneamente por un usuario.

El receptor recibe señales de navegación de al menos cuatro satélites y mide sus pseudodistancias y velocidades. Simultáneamente selecciona y procesa el mensaje de navegación incluido en la señal de navegación. El microprocesador del receptor procesa toda la información recibida y calcula las tres coordenadas de posición, las tres componentes del vector velocidad y el tiempo.

El Segmento de Control

El sector encargado del control y la comunicación con los satélites consta de dos tipos de instalaciones, el Sistema Central de Control, situado en Moscú, y las Estaciones de Telemetría y Seguimiento, que toman datos de los satélites, se encuentran en San Petersburgo, Ternopol, Eniseisk y Konsomol-na-Amure. Todas las instalaciones se encuentran en el territorio de la antigua Unión Soviética.

El Segmento Usuario

El funcionamiento y características de un receptor GLONASS son muy similares a las de un receptor GPS. La recepción de satélites con frecuencias diferentes para cada uno hace que los sistemas de Antena y Radiofrecuencia sean más complejos y por lo general encarezcan los receptores.

También existen los receptores conjuntos GPS+GLONASS, que aunque más sofisticados consiguen una importante ventaja. La existencia de más satélites posibles por observar hace que sea más difícil perder la medición incluso en lugares con muy poco horizonte visible (Ej. una calle estrecha con edificios altos).

12.2.3. Sistema Galileo

Es el último de los sistemas globales de navegación por satélite. Ha sido desarrollado por la Unión Europea, con el objeto de evitar la dependencia con los sistemas GPS y GLONASS, y al contrario que estos dos será de uso civil. Además de prestar servicios de posicionamiento, podrá trabajar con los sistemas GPS y GLONASS. El usuario podrá calcular su posición con un receptor que utilizará satélites de distintas constelaciones. En los primeros planes Galileo estaría desplegado en 2008, pero este plazo ya no se va a cumplir, y las expectativas más optimistas prevén comercializar sus primeros servicios el 2011, pero parece que será más probable el 2014. Problemas entre los países, las agencias y las empresas participantes son la causa de los retrasos en el proyecto. Las últimas previsiones establecen para la implementación del sistema las fases mostradas en la tabla siguiente.

Definición	2000-2003
Desarrollo y validación en órbita	2004-2008
Despliegue	2008-2010
Explotación comercial	desde 2010-2015

Tabla 2. Fases de desarrollo de Galileo

El primer satélite se lanzó el 28 de diciembre de 2005, su nombre es *Galileo in orbit validation element* (Giove-A) desde el cosmódromo de Baikonur (Kazajistán), como casi todos los satélites GLONASS. El segundo de los satélites de prueba, el Giove-B debería haberse lanzado en abril de 2006, pero su lanzamiento fue retrasado hasta mediados o finales de 2007 por fallos en su computador de a bordo.

En el sistema Galileo, se incluyen su propio un sistema de aumentación espacial (EGNOS), entro en funcionamiento en abril de 2004 ha entrado en funcionamiento. Actualmente sólo trabaja como un sistema de aumentación al GPS para mejorar la precisión de las localizaciones. Pero sobre EGNOS se hablará en el capítulo correspondiente a los Sistemas de Aumentación Espacial.

Una de las principales características de Galileo es que garantizará la disponibilidad continua del servicio, excepto en circunstancias extremas, e informará a los usuarios en segundos en caso del fallo de un satélite. Esto lo hace conveniente para aplicaciones donde la seguridad es crucial, tal como las aplicaciones ferroviarias, la conducción de automóviles o el control del tráfico aéreo. El uso combinado de Galileo y otros sistemas GNSS ofrecerá un gran nivel de prestaciones para todas las comunidades de usuarios de todo el mundo.

Galileo realizará una importante contribución a la reducción de problemas como interferencias, fallos de los satélites, denegación o degradación de la señal al proveer en forma independiente la transmisión de señales suplementarias de radionavegación en diferentes bandas de frecuencia con el fin de mejorar la fiabilidad y evitar la vulnerabilidad de la señal.

En total, utilizará 10 radiofrecuencias, de la siguiente manera:

- 4 frecuencias en el rango de 1164-1215 MHz (E5A-E5B)
- 3 frecuencias en el rango de 1260-1300 MHz (E6),
- 3 frecuencias en el rango de 1559-1591 MHz (L1).

Analizaremos ahora sus tres principales segmentos o sistemas, el espacial, el de control y el de usuario. Información adicional y actualizada se encuentra en [3] y también en [7].

El Segmento Espacial

El sistema Galileo estará formado por una constelación de 30 satélites distribuidos en tres planos inclinados con un ángulo de 56° desde el ecuador. Orbitarán a 23.616 Km. de altitud. Cada plano contendrá diez satélites. Cada uno tardará 14 horas de tiempo sidéreo en completar la órbita. Cada plano tiene un satélite de reserva activo, capaz de reemplazar a cualquier satélite de ese plano. Los paneles solares rotarán y mirarán continuamente al Sol.

El Segmento de Control

Tendrá como misión:

- controlar la constelación.
- sincronización de los osciladores atómicos del satélite.
- determinación de las efemérides de los satélites.
- determinar y difundir, a través de los propios satélites los datos de integridad, es decir los avisos sobre la calidad del funcionamiento del sistema dado un intervalo de tiempo.
- proporcionar las interfaces necesarias tanto con los servicios comerciales como con los de búsqueda y salvamento.

Estará formado por un segmento de terreno principal y otro complementario. El principal estará formado por una red de estaciones de sincronización y determinación orbital y las estaciones remotas encargadas de llevar a cabo la comunicación y seguimiento de los satélites para el envío de comandos y recepción de telemetría. Habrá dos Centros de Control principales de Galileo que estarán en Europa. El segmento de terreno complementario será el encargado de proveer la integridad del sistema, constará de su propia red de seguimiento de la integridad (IMS) y un centro de control de integridad.

La red mundial de estaciones de comunicaciones de enlace será de tipo ascendente. Cada instalación de este tipo contendrá equipos de telemetría, telecomunicaciones, seguimiento de satélites y de transmisión de la información de misión que además controlarán la calidad de la señal. La información obtenida de estas estaciones y otras de seguimiento de satélites se transmite por la red de comunicaciones a los dos centros de control terrestres.

Aunque el sistema será de cobertura mundial y la integridad se distribuirá también a escala mundial. Galileo se ha diseñado de forma que permite la introducción de componentes regionales que podrán personalizar su propia integridad.

Las estaciones que prestan servicios regionales difundirán los datos de integridad regionales usando los canales de enlace ascendente autorizados provistos por el sistema. Se garantizará que los usuarios siempre reciban datos de integridad a través de dos satélites con un ángulo mínimo de elevación de 25°.

También contará con una componente local. Las estaciones que prestan servicios locales mejorarán las prestaciones mencionadas anteriormente con distribución de datos locales por medio de radioenlaces terrestres o redes de comunicación existentes a fin de aumentar la precisión o la integridad alrededor de aeropuertos, puertos cabeza de líneas ferroviarias y en áreas urbanas. Los componentes locales también se desplegarán para ampliar los servicios de radionavegación a los usuarios situados dentro de edificios.

El Segmento Usuario

Las características técnicas de los receptores Galileo, de forma general podemos decir que son muy similares a los de GPS. Una aportación novedosa de Galileo respecto a los demás GNSS será la existencia de distintos servicios. Ya que Galileo está concebido para usuarios distintos ó multimodales. A fin de responder a las diferentes necesidades de distintos usuarios, el sistema proveerá cinco tipos de servicios.

1. Servicio abierto, *Open Service* (OS)

Esta orientado a aplicaciones para el público en general. Proveerá señales para proporcionar información precisa de tiempo y posicionamiento en forma gratuita. Cualquier usuario equipado con un receptor podrá acceder a este servicio, sin necesidad de ninguna autorización. La precisión de posición y la disponibilidad serán superiores a las de GPS incluso con el Bloque IIF. Permitirá a los usuarios que posean receptores de uso corriente determinar su posición con un margen de error de unos pocos metros. Se estima que la mayoría de los receptores utilizarán señales conjuntas de Galileo y GPS e incluso GLONASS, lo que ofrecerá a los usuarios una notable mejora en la prestación de servicios en áreas urbanas.

Usará las frecuencias serán E5A, E5B, L1.

2. Servicio para aplicaciones críticas, *Safety of Life* (SoL)

Para aplicaciones de transporte donde la vida humana se podría poner en peligro si la prestación de los servicios del sistema de radionavegación se viera degradada sin notificación en tiempo real. Este servicio proporcionará la misma precisión en posicionamiento y en información precisa de tiempo que el servicio abierto. La diferencia principal es el alto nivel de integridad de cobertura mundial para las aplicaciones donde la seguridad es crítica, como por ejemplo la navegación aérea y las aplicaciones ferroviarias. Este servicio aumentará la seguridad, especialmente donde no hay servicios tradicionales de infraestructura terrestre. Su alcance mundial aumentará la eficiencia de las empresas que operan a escala mundial como aerolíneas y compañías marítimas transoceánicas. El servicio estará asegurado financieramente y sus prestaciones se obtendrán mediante el uso de receptores certificados de doble frecuencia.

Se ocuparán as frecuencias E5A, E5B, L1.

3. Servicio Comercial, *Commercial Service* (CS)

Trabaja para aplicaciones de mercado que requieren un nivel superior de prestaciones que las que ofrece el servicio abierto. Brindará servicios de valor añadido a cambio del pago de un canon. El servicio comercial agrega dos señales a las señales de acceso abierto. Este par de señales está protegido mediante cifrado comercial el cual será gestionado por los prestadores de servicios. El acceso será controlado a nivel de receptor con claves de protección de acceso. Ejemplos de servicios típicos de valor añadido incluyen difusión de datos, garantías de servicio, servicios de información precisa de tiempo provisión de modelos de ionosféricos y señales locales de corrección diferencial para determinar proporcionar gran precisión.

La frecuencia será la E6.

4. Servicio público regulado, *Public Regulated Service* (PRS)

Será el servicio de acceso controlado para aplicaciones gubernamentales. El servicio PRS será utilizado por usuarios tales como la policía ó aduanas. Las Instituciones civiles serán las que limiten el acceso al servicio PRS cifrado cuyo ingreso por región o grupo de usuarios cumplirá las políticas de seguridad aplicables en toda Europa. Deberá estar operativo en todo momento y en cualquier circunstancia, especialmente en períodos de crisis o cuando otros servicios puedan estar interferidos intencionadamente. El PRS es un servicio independiente, en forma tal que otros servicios pueden ser denegados sin que esto afecte a la disponibilidad del servicio PRS. Una característica que destaca a al servicio PRS es la robustez de su señal, lo cual lo protege contra los efectos de las interferencias intencionadas y de los intentos de emisión intencionada de una señal modificada. Las frecuencias serán E6 y L1.

5. Servicio de búsqueda y salvamento, *Search and Rescue Service* (SAR)

Este servicio proporcionará importantes mejoras al sistema de Búsqueda y Salvamento (SAR) ya existente. Ya que permitirá la recepción casi en tiempo real de mensajes de socorro transmitidos desde cualquier punto de la Tierra (el tiempo medio de espera es actualmente de una hora). Además también dispondrá de localización precisa de alertas (pocos metros, en lugar de los 5 Km. actualmente especificados). Por otro lado también será capaz de la detección por múltiples satélites para evitar el bloqueo en condiciones de poca visibilidad de los satélites. Y en definitiva dispondrá de una mayor disponibilidad del segmento espacial (30 satélites en órbita terrestre media que se añaden a los cuatro satélites en órbita terrestre baja y los tres satélites geoestacionarios del actual sistema).

12.2.4. Utilización conjunta de sistemas GNSS

La utilización conjunta de los sistemas GNSS permite, aunque hasta ahora solo ha sido ensayado con GPS y GLONASS, el aumento del número posible de señales de satélites captadas por un receptor dado. Lo que en definitiva se traduce en un aumento de la precisión y sobre todo de la integridad. Existen inconvenientes a esta utilización conjunta, en primer lugar desde el punto de vista del receptor, éste tendrá que ser capaz de captar más frecuencias, tendrá que mantener más canales de seguimiento, su procesador tendrá que ser capaz de transformar a un Sistema de Referencia Geodésico común las observaciones. El Sistema propio de GPS es el WGS84 mientras que el de GLONASS es el PZ-90. Además se usan sistemas de tiempo distinto GPST y GLONASST. En definitiva se encarecen los receptores.

En cuanto a Galileo, este nace desde un principio con un afán de complementariedad tanto con GLONASS como con GPS. A este respecto son varios los compromisos políticos y técnicos adquiridos en los últimos años que asegurarán la interoperabilidad entre los tres sistemas. No solo con Galileo, sino entre los mismos GPS y GLONASS, antaño sistemas enemigos.

12.3. Sistemas de corrección diferencial y aumentación (BAS)

Los sistemas de corrección diferencial y de aumentación, *Based Augmentation Systems* (BAS) nacieron como un subsistema complementario a los GNSS para aumentar sus principales rasgos de calidad, estos son:

- La *exactitud* ó diferencia entre la posición estimada y la real, es mensurable a través de los errores ó incertidumbres posicionales.
- La *integridad* ó confianza sobre la información total proporcionada, donde se establecen posibles alertas para la no utilización del servicio si este no cumple determinados umbrales.
- La *continuidad*, en la medida en que sigue en funcionamiento el sistema sin interrupciones no programadas.
- La *disponibilidad*, que es la parte del tiempo durante la cual el sistema presenta simultáneamente la exactitud, integridad y continuidad requeridas.

Consisten en apoyarse en estaciones terrestres, con coordenadas conocidas con gran precisión, para la emisión de correcciones fruto de la comparación de lo observado con lo ya conocido. Si la emisión de estas correcciones se emiten desde tierra ya sea por radio, radio enlace ó Internet tendremos los llamados Sistemas de Aumentación Terrestre *Ground Based Augmentation System* (GBAS), pero si estas correcciones son emitidas por satélites tendremos los Sistemas de Aumentación Espacial *Space Based Augmentation System* (SBAS). En los SBAS se usan satélites geoestacionarios, los Estados Unidos mantienen la constelación WAAS, Europa EGNOS, Japón MSAT e India GAGAN. GLONASS y Galileo nacieron proyectados para el uso conjunto con sistemas de aumentación, en el caso de Galileo incluso del tipo SBAS con los satélites EGNOS que son parte integrante del sistema desde un principio.

Es necesario mencionar la existencia de un caso especial de sistemas de aumentación, los *Aircraft Based Augmentation System* (ABAS). Entre los sistemas que otorgan esta aumentación a los receptores GPS están los sistemas de Receptor con Supervisión Autónoma de la Integridad (RAIM) y la función de Detección de Fallos y Exclusión (FDE). Los ABAS proporcionan la

integridad requerida para utilizar el GPS como medio único suplementario y principal de navegación durante la salida, en ruta, la llegada y para aproximaciones de precisión y de no-precisión.

12.3.1. Principios básicos sobre la corrección diferencial

Todos los sistemas de aumentación se basan en los principios básicos de la corrección diferencial de observaciones GNSS. Si la emisión de la corrección se realiza desde la superficie terrestre tendremos GBAS, si es desde el espacio SBAS. Con las técnicas diferenciales GNSS, donde una o varias estación con coordenadas conocidas es usada como estación de referencia, se pueden conseguir precisiones submétricas e incluso de orden centimétrico. Si usamos observaciones de código de una sola frecuencia como observable fundamental la técnica se llama GPS diferencial (DGPS). Mucha más precisión se puede obtener a partir de observaciones de fase de dos frecuencias fijando sus ambigüedades a través de las ecuaciones de dobles diferencias a sus valores enteros correctos. Esta técnica es la llamada *Real Time Kinematics* (RTK). Ambas técnicas degradan su precisión en función de la distancia ya que los modelos de corrección por troposfera, ionosfera y errores orbitales presentan malos resultados con la distancia. Los sistemas DGPS con precisión hasta el metro están limitados a distancias entre receptores de unos pocos cientos de kilómetros. Una precisión centimétrica con RTK esta limitada a unos 10 o 20 Km para poder resolver las ambigüedades de fase de la onda de una forma fiable.

Con el fin de proporcionar un servicio de posicionamiento diferencial para una gran área y conseguir precisiones del orden del centímetro varias estaciones de referencia tienen que ser utilizadas. A diferencia de las dos técnicas anteriores, estas estaciones no tienen que ser consideradas de forma independiente con cada estación cubriendo solamente una parte de la superficie. Las observaciones de una red de estaciones tienen que ser premodeladas con el fin de generar modelos de corrección dependientes de la distancia. Es lo que se llama un servicio de *RTK de red* ó GPS diferencial de gran área ó *wide-area DGPS* (WADGPS).

En todos los casos, el procesamiento de datos en una red RTK consiste en tres pasos diferentes. En el primero se fijan ambigüedades de los de las estaciones de la red de referencia. El cálculo de estas ambigüedades difiere del RTK clásico ya que las coordenadas de las estaciones de referencia son conocidas con gran precisión. En un segundo paso los coeficientes del modelo de corrección son calculados. Y en tercer lugar se escoge una única estación de referencia donde se crean las observaciones perfectamente corregidas (normalmente la estación de referencia más

cercana), y basándose en las diferencias de coordenadas horizontales entre la posición aproximada del usuario y la estación de referencia escogida y los modelos de corrección., las observaciones de la estación de referencia son llevadas a la posición del usuario. Cómo se define en [13].

Basándonos en los parámetros del modelo y en la posición aproximada del usuario se definen correcciones individuales de una estación de referencia virtual, *Virtual Reference Station* (VRS) que no existe realmente, que nos permita el posicionamiento diferencial a ella cómo RTK simple. Estas observaciones del VRS cercano a nuestra posición son calculadas y emitidas al usuario por la red RTK a partir de nuestras coordenadas aproximadas. Existen otras soluciones distintas a VRS como el método de *Parámetros de Corrección de Área* más conocido con el acrónimo de alemán (FKP). Es el propio receptor del usuario el que internamente se calcula sus propias observaciones internas VRS a partir del modelo de correcciones y la posición de la estación de referencia dada.

Para algunos sistemas se crearon nuevos formatos de datos para ser capaces de transmitir los parámetros del modelo a los receptores de los usuarios, los cuales predicen correcciones para sus posiciones aproximadas, esto ocurre por ejemplo en EGNOS, WAAS y algunos servicios terrestres, un ejemplo es el formato estandarizado desde 1985, de acuerdo a la propuesta de la *Radio Technical Commission for Maritime Services*, Comité 104. El estándar es conocido como formato RTCM [11]. Otros sistemas utilizan los formatos de datos originales creados para sistemas de estaciones de referencia únicas que proporcionan correcciones VRS, el clásico RTK simple.

Con el RTK de red, los errores diferenciales causados por la refracción ionosférica y troposférica y errores de las órbitas de los satélites son estimados con bastante precisión basándose en observaciones de fase de onda de una red local o regional de estaciones de referencia. Los parámetros del modelo de corrección son determinados para permitir la predicción de los errores diferenciales para la base línea entre una estación de referencia maestra y la posición del usuario. Aplicando estas correcciones al código y la fase de la onda de la estación de referencia maestra, medidas de una VRS son generadas para el posicionamiento RTK del receptor móvil.

En principio, no existían modelos estandarizados para la emisión de parámetros de modelos de corrección. Además, incluso no había un acuerdo sobre cual debe ser la parametrización de los modelos de corrección. Por consiguiente, el formato más común de RTK de red ha sido, usar medidas VRS para ser transmitidas a los usuarios. Las ventajas son obvias usando la técnica

VRS, formatos de datos estandarizados y receptores usuales son capaces de trabajar en modo de RTK en red.

Aun así el formato de RTK de red (VRS), el más utilizado, también tiene desventajas. En primer lugar se necesita comunicación de doble camino ya que los usuarios tienen que transmitir sus posiciones aproximadas al centro de procesamiento, el cual manda de vuelta al usuario las observaciones o correcciones de una VRS individual. En segundo lugar, el software de procesamiento de base que reside en receptor móvil no es capaz de distinguir que procesa datos de estación virtual. Considera los datos originados desde una sola estación de referencia localizada muy cerca de la posición del móvil, así que en algunos casos puede no aplicar las decisiones correctas para el procesamiento de las base. En cualquier caso, así como no existía un formato estandarizado de RTK de red, transferir observaciones VRS ha sido el método preferido de proporcionar información de red preprocesada a los usuarios.

Estrategias de procesamiento como el FKP, ó las usa EGNOS ó WAAS serán más usadas en el futuro cercano en detrimento de la estrategia pura VRS.

12.3.2. Sistemas de aumentación terrestre (GBAS)

Son sistema de corrección y aumentación de señales de GNSS a través de una red de receptores terrestres transmitiendo en las bandas VHF y UHF ó la red telefónica. Como el resto de sistemas de aumentación GNSS (SBAS) se componen comúnmente de una o varias estaciones terrestres, en las que se conoce su posición precisa, y que reciben los datos de cada GNSS. Una vez corregida la señal transmite la información a los usuarios finales. Los GBAS se diferencian de los Sistema de Aumentación Basado en Satélites en que no dependen de satélites geoestacionarios debido a que el GBAS no está diseñado para dar servicio sobre amplias regiones geográficas.

En concreto en España existen varios GBAS que diseminan correcciones en diferentes modalidades y medios de transmisión. En primer lugar las de escala nacional, existen emisiones por radio de correcciones de código suministradas por Puertos del Estado en toda la costa española. Su correspondiente en el resto del territorio nacional con emisión de correcciones de código es el llamado proyecto RECORD del IGN, que emite a través de Radio Nacional España. Todos ellos transmiten en formato RTCM. Estos proyectos han quedado, en los últimos tiempos, a la sombra de los proyectos de emisión de correcciones de medidas de fase y código a través de Internet, mediante la conexión en tiempo real a un servidor que suministra correcciones en formato RTCM de diferentes estaciones permanentes de referencia. El IGN ha

participado activamente en el proyecto europeo EUREF-IP, que ha dado como fruto la instalación de un único *caster* ó emisor de las correcciones de las estaciones de referencia del IGN (ERGPS), y que realiza la transmisión de las correcciones a través de Internet de acuerdo al formato RTCM por transmisión en Internet bajo protocolo NTRIP. Además son varias las iniciativas autonómicas y incluso privadas que actualmente funcionan en España, algunas gratuitas y otras de pago, como por ejemplo la red pionera catalana CATNET.

Una mención a parte merecen los sistemas locales de aumentación como subgrupo de los GBAS. Estos restringen su área de funcionamiento a zonas pequeñas como aeropuertos. Se utilizan pseudolites que son falsos satélites en tierra que emiten señales a partir de las cuales se pueden determinar las distancias a estos. Además de la existencia de al menos una estación GNSS.

12.3.3. Sistemas de aumentación espacial (SBAS)

Son sistema de corrección de señales de los GNSS que se transmiten al receptores GNSS del usuario desde satélites. En los SBAS se usan satélites geoestacionarios, los Estados Unidos mantienen la constelación WAAS, Europa EGNOS, Japón MSAT e India GAGAN. También existen empresas privadas que mantienen sus propias constelaciones SBAS, como por ejemplo Omnistar. GLONASS y Galileo nacieron proyectados para el uso conjunto con sistemas de aumentación, en el caso de Galileo incluso del tipo SBAS con los satélites EGNOS que son parte integrante del sistema desde un principio. Los satélites deben emitir información de exactitud, integridad, continuidad y disponibilidad de los sistemas GNSS. A continuación se amplian las características de EGNOS y WAAS.

EGNOS

Para el ámbito de la Unión Europea se dispone a principios de 2006 del sistema *European Geostationary Navigation Overlay Service* (EGNOS) que transmite correcciones diferenciales e ionosféricas, así como otros servicios de aumentación e integridad a través de satélites geoestacionarios. El sistema ha sido desarrollado por las principales agencias aeronáuticas de Europa bajo la dirección de la ESA (Agencia Espacial Europea). La información oficial podemos encontrarla en [2].

El sistema está compuesto por tres segmentos, el espacial, el terrestre y el usuario (receptores EGNOS). El segmento espacial de EGNOS está compuesto por tres satélites geoestacionarios

con cobertura global terrestre: dos satélites Inmarsat-3 (AOR-E e IOR), así como el satélite ESA Artemis. Los usuarios EGNOS pueden rastrear dos satélites geoestacionarios por lo menos. El sistema terrestre o de control está compuesto por 34 Estaciones de Referencia y de Supervisión de Integridad (RIMS), desplegadas para supervisar los satélites de las constelaciones GNSS. Cada satélite tiene que ser supervisado por múltiples RIMS antes de que se genere las correcciones y los mensajes de integridad. Existen cuatro Centros de Control de Misión (MCC), que procesan los datos de las RIMS para generar las correcciones y mensajes de integridad para cada satélite. Solo uno de estos MCC está activo y operacional, los otros MCC permanecen en reserva. Las Estaciones Terrestres de Navegación *Navigation Land Earth Stations* (NLES) transmiten los mensajes de corrección e integridad desde el MCC a los satélites geoestacionarios, encargados de radiar finalmente la señal SBAS a los usuarios finales. El sistema desplegará dos NLES (una principal y una de reserva), y una tercera NLES con fines de pruebas y validación.

El sistema EGNOS proporciona las siguientes funciones, que constituyen las aumentaciones requeridas para complementar las prestaciones de la constelación GPS/GLONASS y en el futuro Galileo.

- Telemetría, *GEO Ranging* (R-GEO): Transmisión de señales GPS desde tres satélites geoestacionarios (INMARSAT III AOR-E, INMARSAT III IOR y el ARTEMIS). Al aumentar el número de satélites de navegación, aumentará la disponibilidad, continuidad y precisión del servicio.
- Integridad, *GNSS Integrity Channel* (GIC): Distribución de información de integridad. Esto aumentará la integridad del servicio de seguridad GPS/GLONASS/EGNOS de navegación hasta el nivel requerido para la aviación civil (no precisión).
- Precisión, *Wide Area Differential* (WAD): Distribución de correcciones diferenciales. Esto incrementará la exactitud del servicio GPS/GLONASS/EGNOS de navegación y las prestaciones en general hasta alcanzar el nivel exigido para aproximaciones de categoría I.

Con el desarrollo de estas funciones, se pueden identificar cinco niveles de servicio en la utilización de los GNSS con EGNOS:

1. NIVEL 1: Nivel preoperacional. Consistirá en la transmisión de señales similares al GPS desde los satélites GEO (telemetría). Con el aumento del número de satélites disponibles se incrementará la disponibilidad. El Nivel 1 se corresponde con el mínimo suministrado por el sistema EGNOS y se garantizará en todo el área de cobertura geoestacionaria. (área GEO).

2. NIVEL 2: Suministro del servicio de telemetría y de integridad (mediante la emisión de información de integridad). El aumento de integridad permitirá que el servicio cumpla los requerimientos de la aviación civil hasta aproximaciones de no precisión donde se reciban 2 señales geoestacionarias.
3. NIVEL 3: Suministro del servicio de telemetría, integridad y precisión (mediante la emisión de correcciones diferenciales). Supondrá un aumento del nivel de precisión que permitirá cumplir los requerimientos de la aviación civil, en principio hasta aproximaciones de precisión donde se reciban 2 señales geoestacionarias.
4. NIVEL 4: Añadiendo las redundancias necesarias en los segmentos espacial y terrestre, el nivel 2 evolucionará hacia el nivel 4, mejorando las prestaciones conseguidas, de tal forma que se cumplirán los requerimientos de la aviación civil hasta aproximaciones de no precisión en todo el área de cobertura.
5. NIVEL 5: Asimismo, añadiendo las redundancias necesarias, el nivel 3 evolucionará hacia el nivel 5, obteniendo una precisión tal que se cumplan los requerimientos de la aviación civil hasta aproximaciones de precisión en todo el área de cobertura .

WAAS

Es el SBAS desarrollado por Estados Unidos. Está ideado como un complemento para la red GPS permitiendo una precisión en la posición menor de dos metros.). La información oficial podemos encontrarla en [10].

El sistema consta de 24 estaciones de vigilancia, llamadas estaciones de referencia de área amplia *Wide-area Reference Stations* (WRS), 2 estaciones maestras *Wide-area Master Stations* (WMS), 6 antenas para la retransmisión de datos *Ground Earth Stations* (GES) y 3 satélites geoestacionarios. Las 24 estaciones de vigilancia están instaladas con aproximadamente 800 kilómetros de separación a lo largo de los Estados Unidos, Alaska, Hawai y Puerto Rico. Los satélites geoestacionarios transmiten la señal de corrección en la banda de frecuencia L1 del GPS, pero usan un código pseudo-aleatorio (PRC) diferente al de los satélites GPS.

Bibliografía

- [1] Dach, R., Hugentobler, U., Fridez, P., Meindl, M., (2007), The Bernese GPS Software Version 5.0, Astronomical Institute, Universidad de Berna.
- [2] EGNOS (2007), Página oficial EGNOS en la ESA <http://www.esa.int/esaNA/egnos.html>
- [3] Galileo (2007), Página oficial Galileo en la UE, http://ec.europa.eu/dgs/energy_transport/galileo/index_en.html
- [4] GLONASS-ICD (2002), GLONASS Interface Control Document, Version 5.0, Coordination Scientific Information Center, Moscú, Rusia.
- [5] GPS-ICD (1993), GPS Interface Control Document, Revision C (ICD-GPS-200C, US Departement of Defense (DoD), <http://www.navcen.uscg.gov/pubs/gps/icd200/>
- [6] Hofmann-Wellenhof, B., H. Lichtenegger, and J. Collins (1992), GPS: Theory and Practice, Springer, ISBN 3-211-82364-6.
- [7] Leick, A. (1995), GPS Satellite Surveying, Wiley, ISBN 0-471-30626-6.
- [8] Nuñez-García, A., Valbuena, J.L., Velasco, J., (1992) GPS La nueva era de la Topografía. ISBN 84-87510-31-0
- [9] Teunissen, P. J.G., and A. Kleusberg (eds.) (1998), GPS for Geodesy, Springer, ISBN 3-540-63661-7.
- [10] WAAS (2007), Página oficial WAAS en la FAA.
http://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/techops/navservices/gnss/waas/
- [11] RTCM Radio Technical Commission for Marine Services (2001). RTCM Recommended Standards for Differential GNSS (Global Navigation Satellite Systems) Service, Version. 2.3. <http://www.rtcn.org>
- [12] Dai, L., Han, S., Wang, J., Rizos, C. (2001): A Study on GPS/GLONASS Multiple Reference Station Techniques for Precise Real-Time Carrier Phase-Based Positioning. ION GPS 2001, 392-403.
- [13] Wanninger, L. (2006), Introduction to Network RTK. International Association of Geodesy. Commission 4: Positioning & Applications Sub-Commission 4.5: Next Generation RTK Working Group 4.5.1: Network RTK. <http://www.network-rtk.info/>

Tema 13. Estructura interna de la Tierra. Corteza y manto superior. Manto inferior y núcleo. Densidad y parámetros elásticos. Propiedades anelásticas. Ecuación de estado y de composición.

13.1. Estructura interna de la Tierra

Uno de los aspectos más importantes de la sismología lo constituye el estudio de la estructura interna de la Tierra. Por una parte, el conocimiento de sus características elásticas es imprescindible para estudiar la fuente sísmica, y por otro lado la mayoría de lo que conocemos acerca de su composición, estratificación, dinámica, estado físico y temperatura está basado en observaciones sismológicas.

Ya los primeros estudios sísmicos sobre la estructura de la Tierra dieron importantes resultados como la demostración de la existencia del núcleo y de la naturaleza fluida de su parte exterior. Sin embargo, la resolución de detalles de segundo orden es crítica para conocer los procesos dinámicos que ocurren en la Tierra, puesto que grandes diferencias de temperatura o de composición química de los materiales en profundidad producen solo pequeñas variaciones en la propagación de las ondas sísmicas. Por lo tanto hay una necesidad de determinar la estructura de la Tierra con gran precisión para así poder entender su composición y los procesos dinámicos que ocurren en su interior.

Para estos estudios se utilizan todo tipos de ondas sísmicas. Desde las oscilaciones libres del planeta hasta ondas internas de altas frecuencias. A partir de un gran número de curvas dromocrónicas de ondas internas, frecuencias propias de las oscilaciones libres y medidas de ondas superficiales y de la atenuación sísmica se construyó el modelo PREM (Preliminary Reference Earth Model) (*Figura 1*) el cual incluye parámetros como la velocidad de las ondas P y S, el coeficiente de atenuación de cizalla (Q_μ), el módulo de compresibilidad adiabático (K_S), la rigidez (μ), la presión (P) y la gravedad (g). En la *Figura 2* se muestran varios modelos de la velocidad de las ondas P y S en función de la profundidad.

En general, para conocer la estructura de la Tierra, se procede a disponer los sismogramas en función de la distancia a la fuente, lo que permite identificar la llegada de las mismas fases en los distintos sismogramas. A partir de estas observaciones se crean las curvas de tiempos de llegada frente a la distancia epicentral o *curvas dromocrónicas* para cada fase sísmica, con las cuales se puede determinar la estructura interna de la Tierra por medio de los métodos de inversión.

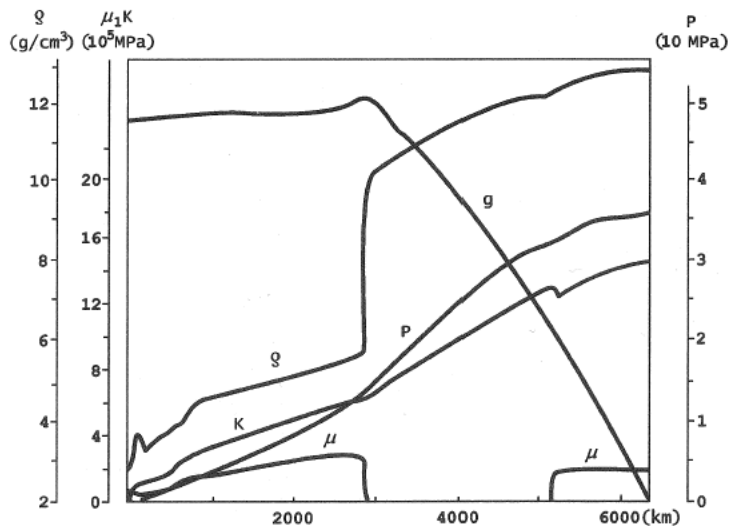


Figura 1. Distribución de la gravedad (g), densidad (ρ), presión (P), módulo de compresibilidad (K) y módulo de rigidez (μ), según el modelo PREM (Udías, 1999).

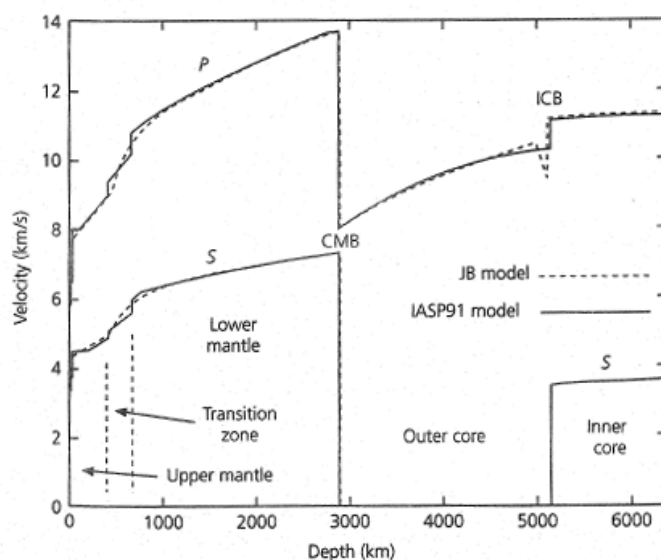


Figura 2. Comparación del modelo de Jeffreys-Bullen (Jeffreys and Bullen, 1940) y el modelo IASP91 (Kennet y Engdahl, 1991). (Stein y Wysession, 2003)

Cualquier estratificación en profundidad da lugar a reflexiones y conversiones de ondas P y S, por lo que el ajuste de las llegadas de ondas refractadas y difractadas sirve para determinar las profundidades y variaciones de velocidad de las discontinuidades internas.

Las primeras ondas utilizadas para estos estudios son aquellas producidas por terremotos, y aún hoy constituyen la mayor fuente de información disponible del interior de la Tierra. Sin embargo, para el estudio de la estructura litosférica y de las capas más superficiales, también se utilizan ondas sísmicas generadas artificialmente por medio de explosiones y otros métodos. Estos estudios a partir de ondas generadas artificialmente, dan lugar a lo que se conoce

normalmente como *perfiles sísmicos* y se basan principalmente en técnicas de dos tipos: las técnicas de refracción y reflexión de gran ángulo y las de reflexión vertical. Ambos utilizan tiempos de llegada y amplitudes, pero los primeros cubren mayores distancias que los segundos. Actualmente también se utilizan una serie de métodos englobados dentro del grupo de la *tomografía sísmica*. Estos métodos, cuyo fundamento se ha tomado de otros campos como la medicina, consisten en la observación de un gran número de rayos sísmicos que cruzan una misma región en diferentes direcciones. La aplicación de los mismos permite establecer modelos tridimensionales de la estructura de la Tierra.

A continuación describimos brevemente cada una de las partes principales de la Tierra: corteza, manto y núcleo, junto con los métodos que nos han permitido su conocimiento. Más adelante expondremos como caracterizar la densidad y algunos parámetros elásticos de la Tierra en función de la profundidad a partir del conocimiento de la distribución de velocidades de las ondas internas y finalmente introduciremos otras características del interior de la Tierra como son sus propiedades inelásticas y su composición.

13.2. Corteza y manto superior

La parte más superficial de la Tierra se denomina *corteza*. Desde el punto de vista sismológico, las primeras observaciones que pusieron de manifiesto el límite inferior de la corteza fueron las de Mohorovičić en 1909 quien observó evidencias que mostraban una discontinuidad situada a unos 30 km y que se denomina *Moho*, la cual separa las rocas de la corteza de las del manto. Una segunda discontinuidad menos profunda que el Moho fue propuesta por Conrad en 1923 (*discontinuidad de Conrad*).

El grosor de la corteza no es homogéneo. En las zonas continentales estables, esta tiene un grosor de unos 30 km, bajo los océanos de 8 a 15 km y en las regiones con altas cordilleras puede llegar a tener 60 km. En cualquier caso, representa una capa muy fina en comparación con el radio de la Tierra (~ 6370 km).

Por lo tanto la superficie del Moho tiene una topografía compleja que recuerda a una imagen inversa de la topografía de la superficie libre.

Los modelos más sencillos de corteza continental suponen dos capas conocidas como *granítica* y *basáltica* o corteza superior e inferior, cubiertas por una capa sedimentaria. Los valores aproximados de la velocidad de la onda P son 2.5-5 km/s en la capa sedimentaria 5.7-6.3 km/s en la capa granítica, 6.6-7.3 km/s en la capa basáltica y 7.8-8.3 en el manto superior.

En general, la corteza superior está formada por un material más rígido y frágil y la corteza inferior por uno más dúctil y posiblemente con mayor concentración de fluidos.

En las regiones oceánicas, la corteza sólo está formada por corteza inferior con velocidades de la onda P de 6.6-7 km/s cubierta por una fina capa sedimentaria. En mares poco profundos y zonas costeras, las características son intermedias entre las cortezas oceánicas y las continentales.

En los sismogramas se distinguen fácilmente ondas que viajan por la corteza superior (P_g y S_g), ondas reflejadas y refractadas en el Moho (P_M y P_n) y en algunas ocasiones fases refractadas en la discontinuidad de Conrad (P^* y S^*).

Frecuentemente se han observado capas de baja velocidad en la corteza. Su presencia produce retrasos en los rayos reflejados y refractados en discontinuidades más profundas. En general, los modelos de corteza con capas de velocidad constante reproducen muchas de las características de las curvas de recorrido observadas, y estos resultados mejoran si introducimos gradientes de velocidad en las capas.

El método más utilizado para conocer las características globales de la corteza es el de los perfiles de refracción aunque los métodos de reflexión casi vertical con ondas de alta frecuencia nos revelan que existe un mayor número de discontinuidades y que éstas son muy complejas. En concreto se aprecia que la corteza inferior presenta una estructura más laminada que la corteza superior, dando lugar a múltiples reflexiones y que el Moho es una zona más o menos difusa donde se producen transformaciones químicas y variaciones de la reología.

Para poder construir modelos tridimensionales de la corteza es necesario combinar los datos a partir de los perfiles de refracción y de reflexión junto con los de tomografía sísmica.

Desde un punto de vista sismológico, el material situado bajo la corteza hasta una profundidad de unos 700 km constituye el *manto superior*. Las características más importantes de esta capa en cuanto a la distribución de velocidades de las ondas sísmicas son: velocidades que aumentan rápidamente con la profundidad desde 7.8-8.3 km/s (ondas P) o 3.8-4.1 km/s (ondas S) hasta alcanzar 10.7 km/s (ondas P) y 5.9 km/s (ondas S). En general la parte superior del manto es muy heterogénea y su estructura guarda relación con la tectónica de placas.

La parte más superficial del manto, inmediatamente después del Moho, es una región con velocidades sísmicas muy altas. Esta capa se superpone a una *capa de baja velocidad* y se conoce frecuentemente como “lid”¹. El grosor del lid varía con las condiciones tectónicas, de modo que en general éste aumenta con el tiempo transcurrido desde el último evento termotectónico. Bajo corteza oceánica muy joven el lid puede estar ausente, sin embargo se observa comúnmente bajo regiones oceánicas antiguas. Asimismo en los cratones esta capa es más gruesa que en zonas continentales activas.

Aproximadamente entre 60 y 220 km está comprendida la capa de baja velocidad, más pronunciada para las ondas S. La profundidad y características de esta capa varían de una región

a otra, pudiendo llegar a desaparecer. En general la capa de baja velocidad se sitúa a poca profundidad en las dorsales oceánicas y zonas tectónicamente activas, y su profundidad aumenta en las regiones continentales antiguas, donde suele estar menos definida. Esta capa es la responsable de la *zona de sombra*² observada en torno a los 20° de distancia.

Bajo la zona de baja velocidad se producen dos aumentos bruscos de la velocidad en torno a unos 410 y 670 km³, que probablemente se deben a transformaciones en la fase de los minerales que componen el manto debido al aumento de presión. Estas discontinuidades dan lugar a reflexiones y a conversiones de las ondas sísmicas que nos revelan su existencia y nos permiten calcular su profundidad y contraste de velocidad.

Desde la discontinuidad a 410 km hasta unos 700 km está lo que se conoce como *zona de transición* del manto superior al inferior.

Desde un punto de vista tectónico y reológico se distingue entre la *litosfera* y la *astenosfera*. La litosfera es la capa más superficial, y contiene a la corteza y una parte del manto superior y está caracterizada por su rigidez y por formar las unidades de la tectónica de placas. Debajo de esta capa se sitúa la astenosfera, asociada a una zona de fusión parcial, menos rígida, con fuerte atenuación sísmica y características plásticas, permitiendo el movimiento de las unidades litosféricas. Desde el punto de vista sísmico la litosfera se extiende hasta los 60 km bajo los océanos y 120 km bajo los continentes aproximadamente. La astenosfera se identifica con la capa de baja velocidad y se extiende en promedio hasta unos 200 km. La división térmica de estas dos estructuras suele situarse en torno a la isoterma de 1500 °C.

El límite inferior del manto superior se sitúa a unos 700 km y coincide con la profundidad máxima a la que existen terremotos y aproximadamente con la isoterma de 2000 °C.

Dentro del manto superior se encuentran regiones con anomalías de velocidad del orden de 5%-10%, respecto a modelos con simetría esférica. Las anomalías positivas se asocian a materiales más fríos en las zonas de subducción, por el contrario las anomalías negativas se asocian con materiales más calientes bajo zonas de extensión.

13.3. Manto inferior y núcleo

El manto inferior se extiende desde una profundidad de 700 km hasta la discontinuidad manto-núcleo (CMB –Core Mantle Boundary–) a 2900 km. Las características sísmicas del manto inferior son muy uniformes con un aumento gradual de las velocidades de las ondas P (de 10.7 a 13.7 km/s) y de las ondas S (desde 5.9 km/s a 7.3 km/s), del mismo modo que la densidad. El

¹ Del inglés: tapa, cobertera.

² Rango de distancias epicentrales en el cual no llega una determinada fase.

³ Algunos modelos indican otra discontinuidad a 520 km.

aumento lento de la velocidad y presión en comparación con el existente en el manto superior implica que la composición y fase de los materiales en el manto inferior no varía significativamente, y que este aumento se debe principalmente al peso del material situado por encima. Las trayectorias de las ondas sísmicas por tanto, corresponden a una distribución normal de velocidades. No obstante hay evidencias de una discontinuidad no global a una profundidad de 900 km, así como la presencia de una capa de unos 250 km de espesor en la base del manto inferior donde en general se reduce el gradiente de velocidad⁴, denominada región D''.

Las propiedades promedio del manto inferior han sido determinadas por medio de la inversión clásica de curvas de tiempos de llegada y por el análisis de los modos normales. De esta forma se crean modelos con simetría esférica con gran interés para la localización de terremotos y para su utilización como modelos de referencia en estudios de tomografía sísmica.

La región D'' es una región compleja con heterogeneidades en la velocidad y se ha asociado a los procesos convectivos que se dan en el núcleo externo y que producen un flujo de calor del núcleo al manto. De este modo es probable que en esta región exista un fuerte flujo lateral y vertical de calor. Por otro lado la cercanía a la discontinuidad CMB, donde se produce un gran contraste de densidad, hace posible la existencia también de heterogeneidades químicas en D''. Algunos estudios de tomografía sísmica realizados en esta región permiten correlacionar zonas de subducción en la litosfera con las zonas con velocidades mayores en D'', de modo análogo, las zonas de baja velocidad en D'' estarían normalmente bajo puntos calientes ("hotspots") en la superficie, lo que daría lugar a una estructura convectiva en el manto. No obstante, no existe una caracterización completa y consensuada de esta región.

Mientras en las profundidades desde 700 a 2500 km los estudios de tomografía sísmica revelan variaciones laterales de la velocidad menores de $\pm 1\%$, en la región D'' la heterogeneidad es mayor, obteniendo $\pm 2\%$ para las ondas P y $\pm 3\%$ para las ondas S.

Las ondas P y S directas que se propagan en el manto inferior llegan en el rango de distancias desde 35° a 105° . El gran contraste de velocidades en el CMB (la onda P pasa de 13.7 km/s a 8.1 km/s y la onda S de 7.3 km/s a 0) produce reflexiones en esta discontinuidad que llegan a distancias de hasta 105° (PcP, ScS, PcS y ScP).

Además, dentro del manto inferior encontramos ondas reflejadas una o más veces en la superficie libre (PP, PPP, SS, PS, SP, PSS, etc.) que llegan con posterioridad a las directas y para los terremotos profundos ondas que viajan hacia arriba reflejándose en la superficie libre para luego viajar por el manto (pP).

⁴ El menor gradiente de velocidad puede esperarse dado que D'' se encuentra pegada al núcleo externo, más caliente que el manto. Sin embargo, modelos detallados de velocidad, muestran que la velocidad aumenta bruscamente, aunque de manera discontinua en la parte superior de esta capa de bajo gradiente.

La primera prueba de la existencia del núcleo a través del análisis de tiempos de llegada fue presentada por Oldham en 1906, quien dedujo que la velocidad era menor que en el manto (aproximadamente la velocidad de las ondas P decrece bruscamente desde 13.7 km/h hasta 8.2 km/h y la velocidad de las ondas S desde 7.3 km/h a cero). El núcleo comprende el 31.5% de la masa de la Tierra, aunque su volumen sólo representa el 16% del volumen de la Tierra, por lo que su densidad es tal que sólo es plausible que esté compuesto principalmente por hierro.

La existencia de un núcleo interno fue propuesta por primera vez por Lehmann en 1936. De este modo se sabe que el núcleo está formado por dos regiones, el núcleo externo (K), de unos 3500 km de radio y el núcleo interno (I) de unos 1200 km de radio.

Las ondas S no se propagan en el núcleo externo, lo que indica que este material es lo suficientemente fluido para no permitir la existencia de esfuerzos de cizalla. Pero no fue hasta principios de la década de 1970 cuando se demostró que el núcleo interno es sólido, debido a que su rigidez afecta a las oscilaciones libres de la Tierra.

El núcleo externo es por lo tanto una capa de baja velocidad que produce una zona de sombra y a continuación una duplicación en las llegadas en dos ramas: una normal y otra posterior. Esta zona de sombra (sin considerar el núcleo interno) se produce desde los 105° a los 143°. Los rayos que penetran en el núcleo externo se denominan PKP, SKS, etc. A una distancia de 143° comienza la primera rama normal (PKP₁) y termina la rama retrógrada (PKP₂).

La presencia del núcleo interno produce ondas reflejadas en la discontinuidad núcleo externo-interno (PKiKP) y ondas transmitidas en el núcleo interno (PKIKP). Los rangos de las distancias epicentrales de estas llegadas son aproximadamente: PKP₁ (de 143° a 156°), PKP₂ (de 143° a 180°), PKiKP (de 120° a 156°) y PKIKP (de 120° a 180°).

Como ya hemos comentado, las ondas S no se propagan en el núcleo externo por lo que se transforman en ondas P (K) al atravesarlo, aumentando su velocidad, por lo tanto para estas ondas no existe zona de sombra como la generada por una capa de baja velocidad sino que se aprecian los lazos que se producen en las dromocronas debido a aumentos bruscos de la velocidad con la profundidad.

Muchos de los esfuerzos dedicados al estudio del núcleo están motivados porque éste se comporta como una geodinamo, responsable del origen y mantenimiento del campo magnético de la Tierra. Esta geodinamo comprende un flujo turbulento en el núcleo externo condicionado por la rotación de la Tierra, la presencia del núcleo interno y un flujo de calor variable hacia el manto.

El material del núcleo parece ser muy homogéneo. En el núcleo externo la velocidad promedio está bien definida por medio de las curvas de tiempos de llegada de las fases PKP y SKS y esta aumenta suavemente con la profundidad. Para determinar el contraste en la discontinuidad núcleo externo-interno se utilizan las fases reflejadas y refractadas en esta interfaz, donde se pasa del estado fluido al sólido. Las velocidades de las ondas P y S en el núcleo interno son

constantes en término medio, aunque el estudio de ondas internas y modos normales revela heterogeneidades y anisotropía en su estructura. Estas características permiten estudiar una posible rotación diferencial del núcleo interno respecto al manto, como predicen los modelos para la geodinamo.

13.4. Densidad y parámetros elásticos

Para conocer la estructura de la Tierra y su composición es también necesario conocer como varían la densidad y los parámetros elásticos con la profundidad. Dado que las velocidades de las ondas internas α y β , son función de los coeficientes elásticos y de la densidad, es posible conocer la distribución de estos parámetros si conocemos de manera independiente la distribución de α y β en el interior de la Tierra.

En primer lugar, uno de los coeficientes elásticos es el coeficiente de Poisson, σ , que puede expresarse en función de los coeficientes de Lamé, λ y μ :

$$(1) \quad \sigma = \frac{\lambda}{2(\lambda + \mu)}$$

como las velocidades de las ondas internas son también función de λ y μ :

$$(2) \quad \alpha^2 = \frac{\lambda + 2\mu}{\rho}$$

$$(3) \quad \beta^2 = \frac{\mu}{\rho}$$

puede comprobarse que podemos expresar el cociente β/α en función de σ como:

$$(4) \quad \frac{\beta}{\alpha} = \left(\frac{2\sigma - 1}{2(\sigma + 1)} \right)^{1/2}$$

y por lo tanto el coeficiente de Poisson, σ , viene dado por:

$$(5) \quad \sigma = \frac{\alpha^2 - 2\beta^2}{2(\alpha^2 + \beta^2)}$$

Ecuación que indica la manera de conocer la distribución del coeficiente de Poisson a partir de la de α y β .

Por otro lado, a partir de las ecuaciones (2) y (3) y de la ecuación del módulo de compresibilidad en función de los coeficientes de Lamé:

$$(6) \quad K = \lambda + \frac{2}{3}\mu$$

podemos expresar, el módulo de rigidez, μ , el coeficiente de Lamé, λ , y el módulo de compresibilidad como:

$$(7) \quad \mu = \rho\beta^2$$

$$(8) \quad \lambda = \rho(\alpha^2 - 2\beta^2)$$

$$(9) \quad K = (\alpha^2 - 2\beta^2)$$

En estas ecuaciones aparece la densidad, ρ , por lo que para calcular estos valores en función de la profundidad necesitaremos conocer la variación de la densidad con la profundidad.

Esta variación puede calcularse bajo ciertas condiciones. Si suponemos que el material es homogéneo, existe equilibrio hidrostático y las condiciones son adiabáticas, la variación de la presión con el radio dentro de la Tierra viene dada por:

$$(10) \quad \frac{dP}{dr} = -g\rho$$

Donde g es la atracción gravitacional de la masa, m , dentro de una esfera con radio r . Para una Tierra esférica este valor en función del radio es:

$$(11) \quad g = -\frac{Gm}{r^2}$$

La variación de la densidad con el radio para una situación de equilibrio hidrostático es:

$$(12) \quad \frac{d\rho}{dr} = \frac{d\rho}{dP} \frac{dP}{dr}$$

Y por definición, K , en un material homogéneo de volumen constante y en un proceso adiabático, es:

$$(13) \quad K = -\frac{dP}{dV/V}$$

Como $\rho = \frac{m}{V}$ entonces:

$$(14) \quad \frac{d\rho}{dV} = -\frac{m}{V^2} = -\frac{\rho}{V}$$

Y a partir de las ecuaciones (13) y (14) tenemos que:

$$(15) \quad \frac{1}{K} = \frac{1}{\rho} \frac{d\rho}{dP}$$

Por lo tanto, utilizando las sustituyendo en la ecuación (12) las ecuaciones (10), (11), (15) y (9) obtenemos:

$$(16) \quad \frac{d\rho}{dr} = \frac{Gm\rho}{r^2 \left(\alpha^2 - \frac{4}{3}\beta^2 \right)}$$

Expresión conocida como ecuación de Adams-Williamson. Con la cual puede determinarse el gradiente de densidad con el radio r a partir de los valores de α , β y la masa de la Tierra hasta un radio r . Esta ecuación puede aplicarse si no hay discontinuidades en los valores de ρ y por lo tanto en los valores de α y β .

Para determinar la densidad en función del radio se comienza en la superficie de la Tierra y se aplica la ecuación sucesivamente a pequeñas capas, teniendo en cuenta que m es la masa

comprendida en el radio r que estamos considerando. La estructura de densidad obtenida por medio de la ecuación de Adams-Williamson se denomina modelo de auto-compresión, puesto que se supone que la densidad en cada punto se debe al peso del material por encima de él y no a un cambio de sustancia con distinta densidad. Dado que los materiales que componen la Tierra son distintos, suele fijarse una densidad para la parte superior del manto e ir calculando este parámetro en cada punto del manto por medio de la ecuación (16), hasta llegar al núcleo donde, debido a que el material cambia, se fija de nuevo una densidad inicial y se procede de forma análoga. En una segunda aproximación es necesario considerar los cambios de fase producidos en las discontinuidades del manto superior, los cuales producen también variaciones en la densidad.

Por otro lado, el análisis del momento de inercia de la Tierra contribuye a conocer mejor como se distribuye la masa en el interior de la misma.

Una vez que se conoce $\rho(r)$ calcular $\mu(r)$, $\lambda(r)$ y $K(r)$ por medio de las ecuaciones (7), (8) y (9).

En la *Figura 1* pueden observarse las distribuciones con la profundidad de μ , K , ρ , P y g . En ella podemos observar que la densidad aumenta gradualmente con la profundidad desde 3.4 g/cm^3 en el manto, hasta que experimenta un aumento brusco desde 5.6 g/cm^3 en la base del manto hasta $10\text{-}13 \text{ g/cm}^3$ en el núcleo. Debido a su gran densidad, la mayor contribución a la gravedad es por la masa del núcleo. La presión aumenta gradualmente en el manto y en el núcleo. El módulo de compresibilidad aumenta lentamente en el manto y más rápidamente en el núcleo y el módulo de rigidez, μ , aumenta muy poco en el manto, se anula en el núcleo externo, y tiene un pequeño valor en el núcleo interno.

13.5. Propiedades anelásticas

Si tenemos en cuenta que la Tierra no se comporta de una manera perfectamente elástica obtenemos una atenuación de la amplitud de las ondas denominada *inelástica* o *atenuación intrínseca*, en la que parte de la energía sísmica se transforma en calor. Esta atenuación suele representarse por medio del coeficiente de *disipación específica* o *factor de calidad*, cuyo inverso es la *fricción interna* y representa la fracción de energía disipada durante un periodo:

$${}^{(17)} Q^{-1} = -\frac{1}{2\pi} \left(\frac{\Delta E}{E} \right)_T$$

Por lo tanto para un material perfectamente elástico Q^{-1} es cero.

La atenuación intrínseca de la amplitud de una onda P de frecuencia ω al propagarse una distancia x puede expresarse como:

$${}^{(18)} A(x) = A_0 e^{-\gamma x} = A_0 e^{-\frac{\omega x}{2Q_\alpha}}$$

Donde γ es el coeficiente de atenuación inelástica y Q_α la disipación específica para las ondas P.

El valor de Q_α puede calcularse tomando el logaritmo de cociente de las amplitudes A_1 y A_2 del espectro de ondas P registradas a distinta distancia. Si la diferencia de distancia es Δx tenemos:

$${}_{(19)} \ln \left(\frac{A_2(\omega)}{A_1(\omega)} \right) = \ln C - \frac{\Delta x}{2\alpha Q_\alpha} \omega$$

Donde C depende de la *dispersión geométrica*. Es decir, de la disminución de la amplitud del desplazamiento debido al ensanchamiento del frente de ondas al aumentar la distancia x ⁵.

Realizando medidas para distintas frecuencias, podemos calcular mediante una regresión lineal el valor de Q_α , suponiendo que Q no varía con la frecuencia, lo cual no es cierto para frecuencias altas.

El valor de Q es distinto para las ondas P y S, en general Q_α es el doble que Q_β .

En realidad al calcular Q se obtiene un valor medio representativo de la región por la cual ha viajado la onda sísmica.

De forma análoga a como se obtiene la distribución de velocidad, densidad y parámetros elásticos con la profundidad, también se puede obtener la distribución de Q con la profundidad. El conocimiento de la atenuación en el interior de la Tierra es muy útil para estimar las variaciones de temperatura en su interior. Mientras la velocidad sísmica depende linealmente de la temperatura, la atenuación lo hace exponencialmente, y por lo tanto es más sensible a ésta. De este modo a partir de estudios combinados de velocidad y atenuación, obtenemos una valiosa información sobre las variaciones de la temperatura, las cuales revelan importantes procesos geofísicos.

En la corteza la atenuación es elevada y por lo tanto los valores de Q son relativamente bajos. Estos valores varían regionalmente, lo cual tiene importantes implicaciones en el cálculo de la peligrosidad sísmica. Bajo la corteza el valor de Q aumenta progresivamente, salvo en la astenosfera donde la atenuación es alta. Bajo la astenosfera, en el manto, Q aumenta gradualmente con la profundidad. En el núcleo externo Q_β es nulo y su valor aumenta de nuevo en el núcleo interno. Por el contrario, la atenuación de las ondas P en el núcleo interno es despreciable aunque en el núcleo interno es significativa. En la *Figura 3* se muestra un modelo de Q_β en función de la profundidad.

⁵ Fenómeno que se da en cualquier medio, sea o no perfectamente elástico y que depende de la geometría del frente de ondas. Otro fenómeno a tener en cuenta en la atenuación total es el "scattering" o esparcimiento, el cual se debe a la influencia de heterogeneidades en el medio cuyo tamaño es comparable a la longitud de onda considerada.

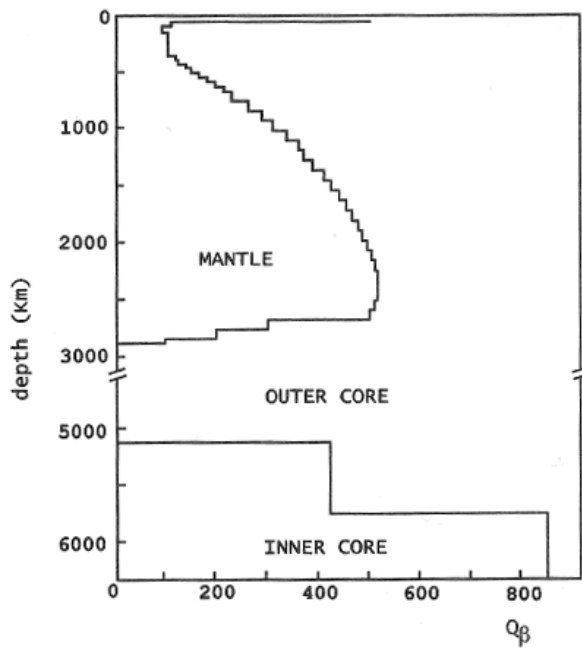


Figura 3. Q_β en el interior de la Tierra según el modelo SL8 (Anderson y Hart, 1978). (Udías, 1999).

13.6. Ecuación de estado y de composición

Las observaciones sismológicas proporcionan una fuente de información muy valiosa sobre el interior de la Tierra. A partir de ellas es posible deducir la distribución de densidad y de los parámetros elásticos e inelásticos de su interior. La distribución de estos parámetros muestra como hemos visto la existencia de tres regiones fundamentales (corteza, manto y núcleo) y dentro de ellas las subregiones del manto superior e inferior y núcleo externo e interno. También se ha deducido a partir de estas medidas, que en los 700 km más superficiales, corteza y manto superior, existen fuertes heterogeneidades laterales, mientras que en el resto, las propiedades varían principalmente con el radio.

Por medio de un perfil de densidad podemos calcular un perfil de presión, y a continuación usar el resultado de experimentos que nos muestren qué fases minerales existen a una determinada presión.

Las condiciones del material en el interior de la Tierra dependen principalmente de su presión y temperatura y se pueden representar a través de una *ecuación de estado*, que relaciona la presión y la densidad para una temperatura dada. Una de las primeras ecuaciones de estado propuestas para el material del manto es la de Birch:

$${}^{(20)}P = \frac{3}{2} K_0 \left[\left(\frac{\rho}{\rho_0} \right)^{1/3} - \left(\frac{\rho}{\rho_0} \right)^{5/3} \right]$$

Donde P es la presión, ρ la densidad y K el módulo de compresibilidad. El subíndice 0 indica condiciones a presión atmosférica.

El paso de las condiciones físicas a la composición mineralógica en el interior de la Tierra presenta todavía más incógnitas. Los datos sismológicos ponen de manifiesto discontinuidades claras en la base de la corteza y del manto, que parecen responder a cambios bruscos de composición. Dentro del manto superior, las discontinuidades situadas a profundidades de 200, 410, y 670 km parecen estar también bien establecidas, aunque pueden corresponder a cambios en la composición menos drásticos. El material de la corteza puede estudiarse directamente por sus afloramientos en la superficie y mediante sondeos, pero aún no se ha conseguido perforar hasta el manto debido a sus elevadas presiones y temperaturas.

En líneas generales se puede decir que la corteza está formada, aparte de por una capa sedimentaria donde abundan las rocas calizas y areniscas, por rocas graníticas, basálticas y peridotíticas, las cuales corresponden a rocas ácidas, básicas o ultrabásicas debido a su mayor o menor presencia de SiO_2 (70%, 50% y 45%). Otros minerales que abundan en la corteza aparte del SiO_2 son el Al_2O_3 , Fe_nO_m , MgO , CaO ,...

El manto terrestre está formado principalmente por SiO_2 , MgO y CaO . En la parte superficial, donde la presión es menor, estos compuestos se combinan para formar series de silicatos como el piroxeno y el olivino. A mayor profundidad estos minerales pasan por cambios de fase como el de olivino a espinela propuesto para la discontinuidad de 410 km y de espinela a perovskita y óxidos de magnesio en la discontinuidad de 670 km.

La información sobre la constitución del Núcleo es aún más escasa e indirecta, aunque todas las observaciones, como la existencia del campo magnético terrestre y la composición de los meteoritos apuntan a una composición de hierro con una pequeña proporción de otros elementos, como el Si, Ni, S y K que sirven para explicar la densidad y módulo de compresibilidad del núcleo, más bajos que los correspondientes al hierro puro a la presión y temperatura supuesta. La presencia de potasio-40 (K^{40}) radiactivo en el núcleo podría proporcionar parte de la energía necesaria para las corrientes de convección, que dan origen al campo magnético terrestre.

Bibliografía

- [1] Lay T. y Wallace, T. C. Modern Global Seismology. Academic Press, 1995. Pág. 251-307.
- [2] Stein, S y M. Wysession. An introduction to Seismology, Earthquakes and Earth Structure. Blackwell Publishing 2003. Pág 162-212.
- [3] Udías, A. y Mezcuá, J. Fundamentos de Geofísica. Editorial Alianza, 1996. pág . 167-181.
- [4] Udías, A. Principles of Seismology. Cambridge University Press, 1999. Pág. 153-182.

Tema 14. Distribución espacial de terremotos. Características de terremotos de márgenes convergentes, divergentes y transcurrentes. El ciclo sísmico: modelos de recurrencia. Distribución de magnitudes. Modelos temporales de recurrencia.

14.1 Distribución espacial de terremotos

La sismicidad ha sido considerada como la distribución geográfica de los terremotos y sus efectos destructores. La base para su estudio es, por tanto, la recopilación de datos sobre fecha, localización, daños producidos, y todos aquellos parámetros que sirvan para su identificación que se recogen en los llamados catálogos sísmicos. Los trabajos pioneros de catálogos se deben a J.B. Montessus de Ballore entre 1890 y 1923; los de Gutenberg y Richter que culminan con la publicación de su obra “Seismicity of the Earth” en 1954 y los más recientes de S. J. Duda en 1963 y para toda Europa los de V. Karnik y Vaniek en 1968. Los catálogos mundiales recientes son los publicados por el International Seismological Center (Reino Unido) y el National Earthquake Information Center (EE.UU.)

En el caso de la península Ibérica cabe citarse los catálogos de J.M. Sanchez Navarro-Neumann, J. Rodriguez Galbis y E. Fontseré publicados entre los años 1920 y 1940. Mas modernamente los catálogos de Rey Pastor, V. Inglada, J.M. Munuera, J. Mezcua y Martinez Solares y en la actualidad el Catálogo del Instituto Geográfico Nacional como compendio y actualización de los anteriores son los que pueden citarse como de referencia.

Una de las características más importantes de la distribución de terremotos a partir de la información de los catálogos es que éstos están concentrados en franjas relativamente estrechas, que forman una red que divide a la superficie de la Tierra en una serie de áreas cuyo interior es prácticamente asísmico. De estas franjas o alineaciones de terremotos, unas ocupan los márgenes de los continentes y otras están situadas en el interior de los océanos, coincidiendo con las grandes cordilleras submarinas. No en todos los márgenes de los continentes se producen terremotos, lo que permite separarlos en activos y pasivos; por ejemplo, el margen occidental de América es activo, mientras que el oriental no lo es. No todos los terremotos se dan a la misma profundidad, por lo que éstos se suelen clasificar en superficiales ($h < 60$ km), intermedios ($60 \text{ km} < h < 300 \text{ km}$) y profundos ($h > 300 \text{ km}$), siendo la profundidad mayor a la que se han observado aproximadamente 700 km. Así como unas zonas de la Tierra son activas y otras no, en unas zonas activas se dan terremotos profundos y en otras sólo superficiales. Respecto a la profundidad de los terremotos, los que se dan en las cordilleras oceánicas son

superficiales, mientras que los terremotos profundos se dan preferentemente en los márgenes activos de continentes, en especial los que rodean al océano Pacífico.

La distribución geográfica global de los terremotos puede resumirse esquemáticamente en las siguientes zonas activas: La primera, el cinturón circum-Pacífico, que rodea este océano, formado en su parte oriental por los terremotos de Alaska, California y Nevada, México, América Central y cordillera Andina, desde Colombia hasta el sur de Chile. En esta zona se han producido grandes terremotos que han ocasionado frecuentes catástrofes a lo largo de los años.

En América del Sur, los terremotos son superficiales en las zonas costeras y más profundos hacia el interior del continente. En su parte occidental, la costa del Pacífico está orlada de una serie de arcos de islas, desde el más septentrional de las islas Aleutianas hasta el de Fiji y Honda en el sur, en los que se dan terremotos superficiales y profundos de gran magnitud. La segunda región sísmica es la llamada «Mediterráneo-Himalaya», que se extiende desde las islas Azores hasta la costa del continente asiático. A esta región pertenecen los terremotos del sur de España, norte de África, Italia, Grecia, Turquía, Persia, Himalaya, norte de la India y China. La tercera la constituye el complejo sistema de cordilleras submarinas que se extienden en los océanos. Entre éstas hay que destacar la centro-Atlántica, que divide de norte a sur este océano en dos partes, las del Índico y las del Pacífico frente a la costa oriental de América del Sur. Los terremotos de estas zonas son superficiales y de magnitud media

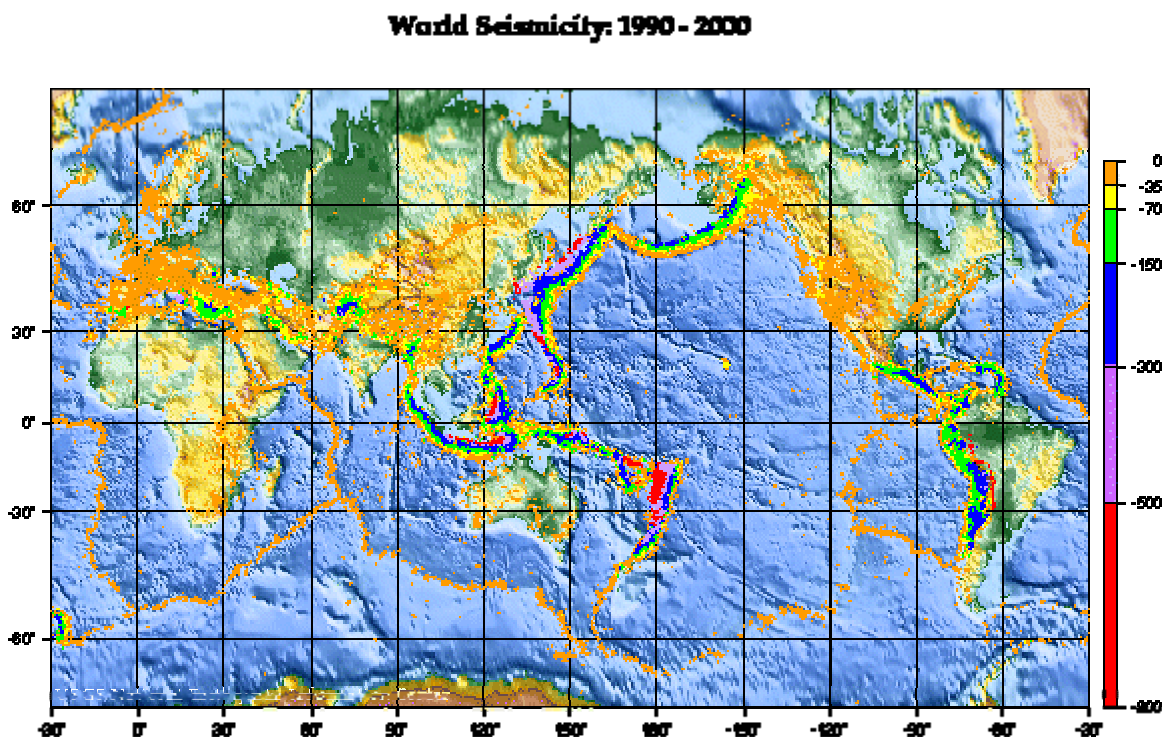


Figura 1. Sismicidad mundial por profundidades USGS.

La distribución de terremotos en Europa se extiende de oeste a este desde las islas Azores hasta el Cáucaso. La zona más activa es la del arco Helénico, seguida por la península Itálica y el arco de Sicilia-Calabria, y el sur de la península Ibérica y el norte de África. La mayoría de los terremotos son superficiales con focos de mayor profundidad (intermedios y profundos) en los arcos Helénico, Sicilia-Calabria, Cárpatos y Bético-Rif. La sismicidad de América del Sur se concentra en su margen occidental (costa del Pacífico) desde el norte de México al sur de Chile. Toda esta zona es de gran actividad con gran frecuencia de terremotos de elevada magnitud. La profundidad aumenta hacia el interior del continente llegando a los 600 km en el interior de Perú y Bolivia. Otra zona activa está situada en el Caribe desde la península de Yucatán al arco de las Pequeñas Antillas.

El estudio de la distribución de terremotos ha sido uno de los factores más importantes en el establecimiento de la teoría tectónica de placas, que interpreta esta distribución de acuerdo con un modelo de geotectónica global. La teoría de la tectónica de placas interpreta la distribución de terremotos de las figuras anteriores en la forma siguiente: la superficie de la litosfera terrestre está dividida en placas, siendo siete las más importantes, cuyos bordes coinciden con las zonas activas de terremotos. Estas placas no coinciden, como se puede ver en la figura con los continentes, sino que en una sola placa se agrupa parte continental y parte oceánica. Los bordes de las placas se pueden dividir en tres tipos. En el primero, las placas se separan unas de otras, produciendo *zonas de extensión o divergencia*. En este tipo de borde, los terremotos son superficiales y generalmente están asociados a las cordilleras oceánicas. En el segundo, las placas colisionan unas con otras, formando *zonas de convergencia o subducción* en las que la litosfera de un lado se introduce por debajo de la del otro. Estos márgenes producen en unos casos zonas de arcos, como la que se extiende desde las Aleutianas hasta Fiji-Tonga, en el Pacífico, y, en otros, grandes cordilleras, como los Andes, Himalaya y Alpes. Los márgenes del Pacífico y el que existe en algunas partes del contacto entre Eurasia y África son de este tipo. En el tercero, una placa se desliza horizontalmente con respecto a la otra, dando lugar a *zonas de deslizamiento* con grandes fallas de desgarre, como la de San Andreas, en California, y Anatolia, en Turquía.

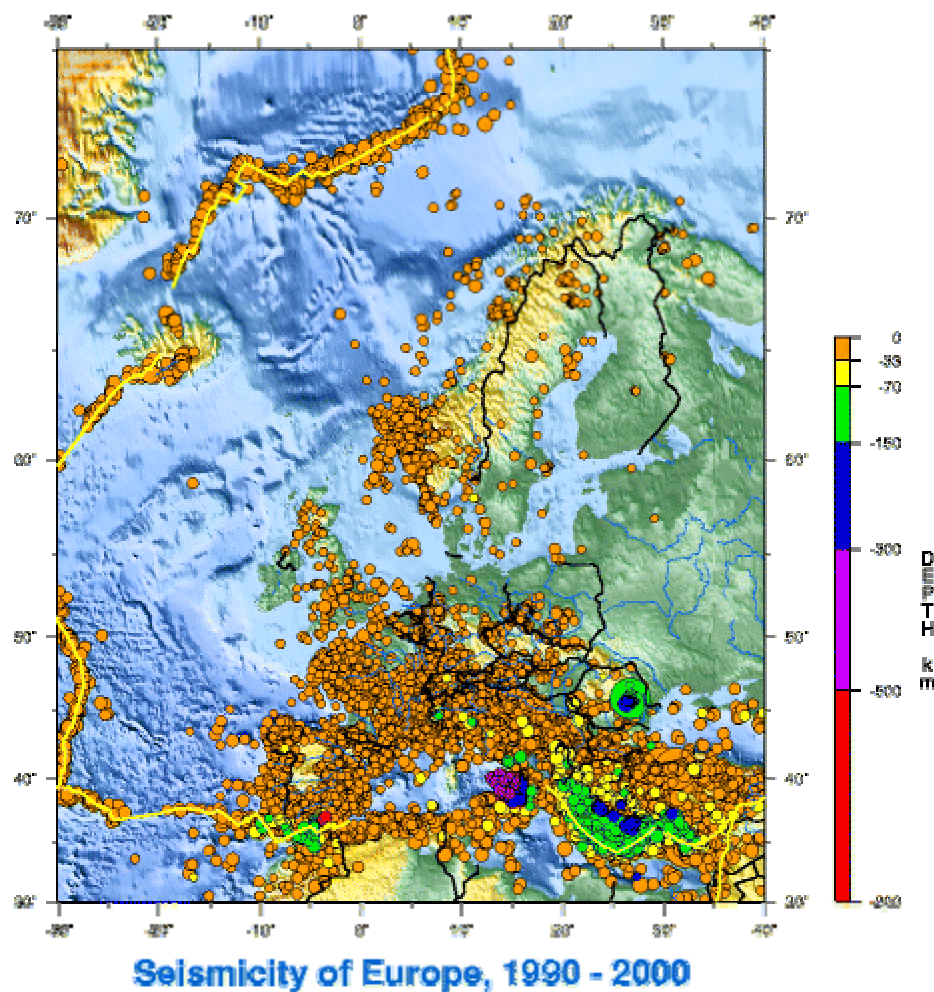


Figura 2. Sismicidad de Europa USGS

Desde el punto de vista de la sismicidad, las zonas de extensión tienen actividad superficial y de moderada magnitud. En las de convergencia, los terremotos son tanto superficiales como profundos, y pueden alcanzar magnitudes muy altas. En las fallas de desgarre, los terremotos son superficiales y su magnitud puede ser también muy grande. Aunque hemos dicho que los terremotos se dan preferentemente en los bordes de las placas, con menor frecuencia se dan también en algunas zonas de su interior. Esta circunstancia ha llevado a dividir los terremotos en interplacas e intraplacas. A veces, los terremotos intraplacas pueden ser de gran magnitud, como los acaecidos en el interior de China y en el centro y costa oriental de América del Norte.

Así como la distribución geográfica de los terremotos marca la localización de los márgenes de las placas, el estudio de su mecanismo nos indica a qué tipo de margen pertenecen. En general, se puede afirmar que en los márgenes de extensión, los terremotos están producidos por fallas normales con tensiones horizontales y perpendiculares a la dirección del margen. En los márgenes de convergencia, los terremotos superficiales corresponden a fallas inversas, con ejes de presión horizontal y normal al margen. Finalmente, en las fallas de desgarre, tanto las

presiones como las tensiones son horizontales, lo que da origen a un movimiento de deslizamiento horizontal.

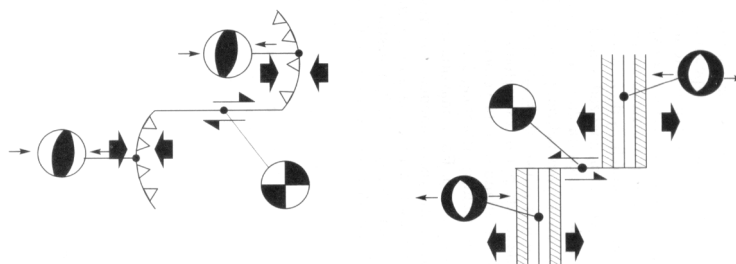


Figura 3. Mecanismos de terremotos según el tipo de margen en los que se produce. Izquierda de subducción y derecha de extensión. En ambos casos los frentes están unidos por una falla de transformación. Fuente: Udías y Mezcuá, 1997.

14.2 Características de terremotos en márgenes convergentes, divergentes y transcurrentes

Los márgenes convergentes o de subducción son aquellos en los que las placas convergen unas contra otras. Este movimiento obliga a una de ellas a introducirse por debajo de la otra, resultando que la litosfera se consume o destruye. Cuando una de las dos placas es de naturaleza continental, la placa oceánica es la que se introduce por debajo de la continental debido a la baja densidad de esta última, que opone una gran resistencia a penetrar en el manto de mayor densidad. De esta forma, la litosfera continental se ha conservado prácticamente constante, mientras la oceánica se crea y se destruye. En los márgenes en que se intenta destruir litosfera continental se produce un cambio en el sentido del movimiento y se hunde la placa oceánica opuesta, o se produce un cambio en las características del margen de las placas. En general, el frente de la placa buzante tiene una cierta curvatura penetrando desde la parte convexa. En muchos casos, el frente del margen de subducción se halla a cierta distancia de la costa continental, formando un arco de islas y existiendo entre dicho arco y el continente una cuenca marina (back-arc basin). Esta cuenca se forma a partir de un centro de extensión situado detrás (parte cóncava) del arco de islas, en el que se genera corteza oceánica de la misma manera que en una dorsal. Un ejemplo de esta situación es el mar de Japón. En otros casos, como en la costa occidental de América del Sur, la zona de subducción está directamente adosada a la costa y la placa oceánica se introduce con un ángulo pequeño bajo la litosfera continental. Esta situación se produce por un movimiento del continente hacia el frente de subducción que impide la formación de una cuenca marginal.

Uno de los indicios más importantes de la existencia de estas placas de material litosférico introducidas en el manto es la distribución de focos sísmicos en profundidad. Éstos forman alineaciones desde la superficie hasta unos 700 km de profundidad con un ángulo con la

horizontal que varía en inclinación, en muchos casos del orden de 45° , y que se denominan zonas de Benioff-Wadati.

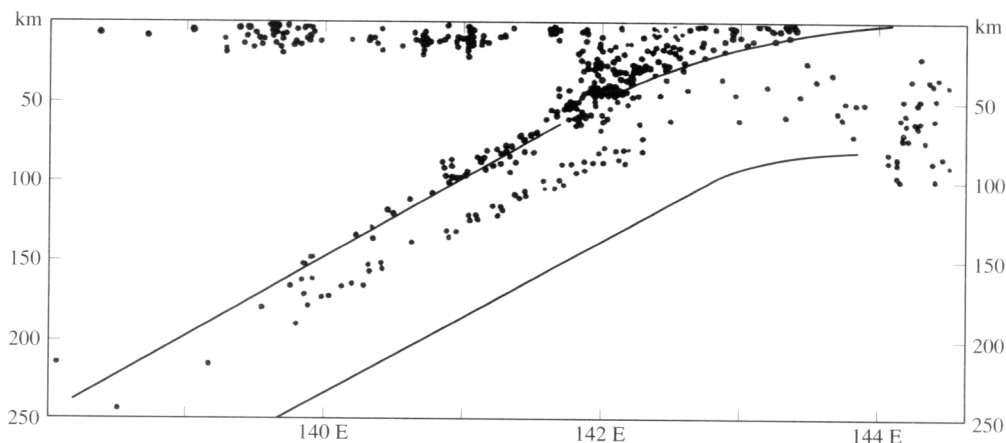


Figura 4. Distribución de terremotos en una placa en subducción. Udías y Mezcuá, 1997.

El espesor de la zona sísmica está limitado en general a la parte superior de la placa litosférica. El hecho de que su profundidad no pase de los 700 km indica que a esta profundidad la placa litosférica hundida en el manto ha perdido su rigidez y probablemente ha quedado asimilada al material del manto. El mecanismo de los terremotos en la superficie es de fallas inversas, con la parte oceánica desplazándose bajo la continental y presiones horizontales y perpendiculares al frente del arco. En la zona donde la placa se dobla se producen en la superficie superior fallas de tensión, mientras que en la inferior éstas son de compresión. El interior de la capa buzante a profundidades medias está sometido a tensiones a lo largo de la placa, mientras que en la parte más profunda lo está a compresiones, debido a la resistencia que opone el material del manto a la penetración de la placa.

En los márgenes de extensión o divergencia cuyo caso típico es el de una dorsal oceánica, como puede ser la de la cordillera Centro-Atlántica. Topográficamente, la dorsal está formada por una cadena montañosa de origen volcánico, en cuyo centro se suele dar una depresión o valle de rift, aunque no siempre sea así. El grosor de los sedimentos marinos aumenta con la distancia al eje de la dorsal, así como su edad. Esto indica que la zona de extensión actúa como centro a partir del cual se va generando la nueva litosfera oceánica. Al separarse las dos placas, el material fundido del manto surge en forma de lava a la superficie y se enfría incorporándose a la corteza. Al continuar separándose las placas, este material va ocupando el espacio abierto, creándose nueva corteza oceánica a partir del eje de la dorsal. Éstas constituyen así una importante proporción de las zonas volcánicas. Cuando en ellas el volcanismo es muy intenso, las cimas de la cordillera afloran sobre la superficie del mar formando islas.

Los datos sísmicos muestran que la distribución de epicentros está alineada en una franja estrecha que sigue el eje de la dorsal con terremotos de magnitud moderada ($M < 6,5$) y

profundidad superficial ($h < 30$ km). Estas alineaciones marcan, con asombrosa exactitud, la situación de las zonas de extensión a lo largo de la superficie de los océanos. El mecanismo de los terremotos es predominantemente de fallas normales, correspondiendo a esfuerzos tensionales horizontales y perpendiculares al eje de las dorsales. Los planos de falla buzan en la dirección de la extensión de la litosfera hacia los dos lados de la dorsal. En muchas partes, la dorsal está interrumpida por una falla perpendicular a su eje debida a una cierta diferencia relativa en la velocidad de extensión. Estas fallas reciben el nombre de fallas de transformación o transformadas

Por último los márgenes de transformación o fractura se conocen desde antiguo por la existencia de grandes fallas de movimiento predominantemente horizontal, cuyo ejemplo más notable es la falla de San Andreas, en California. La explicación del movimiento en estas fallas no se hizo clara hasta el trabajo de J. T. Wilson, de 1965, en el que explicó su función y les dio el nombre de fallas de transformación (transform faults). Una de las características que más impresionó a Wilson fue que en estas fallas el desplazamiento termina súbitamente a los dos extremos de la falla. La explicación ofrecida es que las fallas conectan zonas de extensión y subducción entre sí o unas con otras. Estas fallas son necesarias para explicar el movimiento de las placas, que no sería posible sin la existencia de este tipo de margen. En todos los tipos, el movimiento horizontal se transforma en los extremos, bien en movimiento de expansión o de subducción, lo que explica su nombre. Existen ejemplos muy claros: como la falla de San Andreas, que conecta dos zonas de extensión, y la del Caribe, que une las zonas de subducción del arco de las Antillas y de la costa de México.

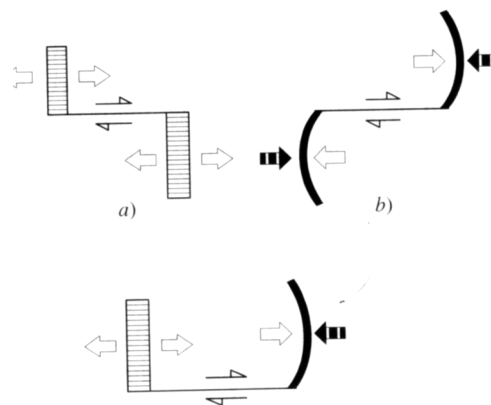


Figura 5. Fallas de transformación conectando: dos zonas de extensión, dos de subducción y una de extensión con una de subducción. Fuente: Udías y Mezcuá, 1997.

Los terremotos en estas fallas llegan a tener magnitudes muy grandes ($M > 8$), baste recordar el de San Francisco, ocurrido en la falla de San Andreas, en 1906, en el que la ruptura se extendió a más de 300 km, o los ocurridos a lo largo de la falla Azores-Gibraltar. Su mecanismo es de fallas casi verticales de movimiento horizontal, es decir, de desgarre o salto en dirección. El

sentido del movimiento, determinado en los mecanismos de estos terremotos, coincide con el que corresponde al producido por el mecanismo de transformación, y no al de fallas transcurrentes que hubieran desplazado las dorsales. La presencia y sentido del movimiento de estas fallas en las dorsales oceánicas se puede apreciar también en los desplazamientos de las franjas de las anomalías magnéticas.

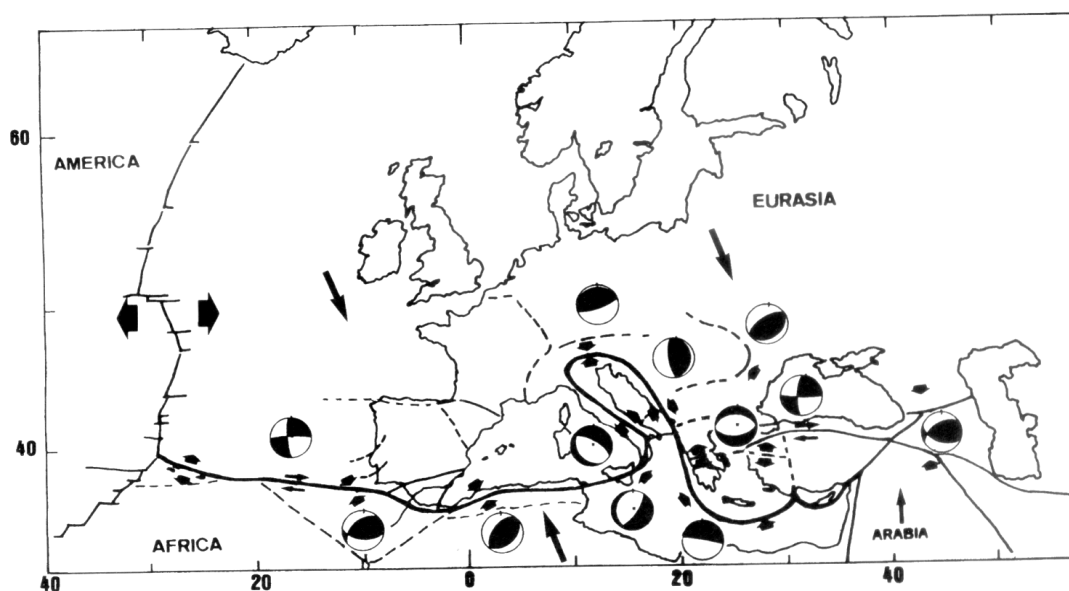


Figura 6. Mecanismos focales característicos del contacto Eurasia-Africa. Udias, 1999.

Por último, podemos observar como la distribución de mecanismos en el contacto Eurasia-Africa, desde el W son: zona de desgarre entre Azores y Gibraltar, de convergencia en el Norte de Africa y el arco de Calabria y de extensión en los Apeninos además de en la dorsal centro-oceánica.

14.3 El ciclo sísmico: modelos de recurrencia

La observación de los efectos sobre el terreno del famoso terremoto de San Francisco de 1906 realizado por H. F. Reid le permitió formular la teoría del Rebote Elástico para explicar la ocurrencia de un terremoto. Esta teoría por su simplicidad sigue siendo en la actualidad perfectamente valida aunque es conocido que únicamente explica algunos aspectos de la producción de terremotos.

En la Figura se muestran las distintas etapas en el tiempo que se producen cuando los esfuerzos aplicados en dos sentidos distintos a una cierta separación horizontal producen una deformación en el medio que se visualiza mediante el curvado de unas líneas ideales presentes en dicho medio elástico. Cuando se alcanza un máximo de deformación se produce la rotura coincidente con la ocurrencia de un terremoto. Se produce en este caso una deformación que al venir acompañada del terremoto recibe el nombre de deformación cosísmica.

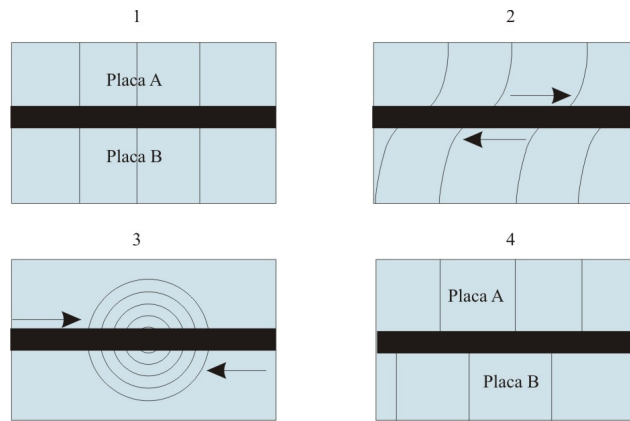


Figura 7. Modelo de Rebote elástico de Reid.

El primer modelo es la aproximación al modelo de Reid y consiste en que cada terremoto se produce al alcanzar un cierto nivel de esfuerzo crítico y se produce una caída de esfuerzos constante, hasta alcanzar otro nivel mínimo de esfuerzos. Este modelo es el denominado *Terremoto Característico* y puede considerarse como difícilmente realizable en la naturaleza. Existe un segundo modelo, en el que cada terremoto se produce igualmente a un cierto esfuerzo crítico, pero la caída de esfuerzos es en cada caso variable lo que conduce a un deslizamiento inducido en la falla también variable. Este modelo se denomina *Tiempo Predecible* porque el tiempo del siguiente terremoto puede predecirse y no la magnitud asociada. Por último, el modelo de *Deslizamiento Predecible* en el que ocurre lo contrario del modelo anterior, es decir los terremotos se producen a distintos niveles de tensiones críticas, pero el proceso de relajación posterior de las tensiones se produce hasta una tensión ambiental idéntica. En este modelo puede predecirse el deslizamiento que se va a producir, es decir la magnitud del terremoto asociado pero no el tiempo de ocurrencia.

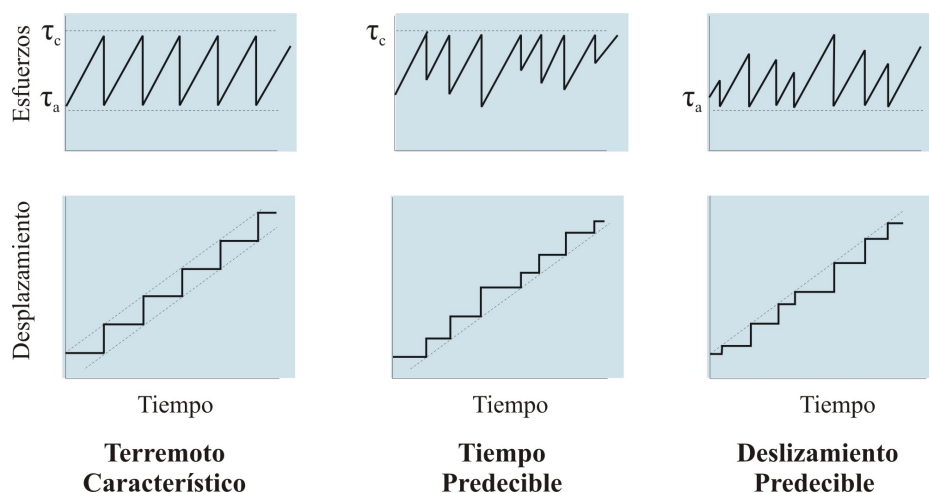


Figura 8.

Donde τ_a y τ_c son el esfuerzo mínimo y máximo respectivamente del proceso.

14.4 Distribución de magnitudes

Considerando el modelo de relajación de esfuerzos en la generación de los terremotos, es lógico pensar que para una cierta región y durante un periodo determinado de tiempo, el número de terremotos pequeños es mucho mayor que el de grandes.

Aunque se desconoce exactamente la forma en que se libera la energía elástica almacenada en una región y que da origen a terremotos de distinta magnitud, sin embargo, se ha comprobado experimentalmente que la magnitud es una variable aleatoria, cuya función acumulativa de distribución es:

$$F(M) = 1 - e^{-BM}; M > 0$$

Esta distribución fue propuesta en primer lugar por Omori, en 1889, y más tarde por Gutenberg y Richter, en 1944, en la forma:

$$\log_{10} N(M) = a - bM$$

Donde N es el número de terremotos con magnitud mayor que M; la constante a, el logaritmo del número de terremotos de magnitud mayor que cero, y b, la proporción de terremotos de una cierta magnitud. La determinación de b se hace calculando la pendiente de la recta resultante del ajuste lineal del logaritmo del número de terremotos con la magnitud. Las desviaciones con respecto a la recta en esta relación se suelen asignar a la falta de completitud en la serie de terremotos, es decir, no todos los terremotos de un cierto rango de magnitudes han sido incluidos.

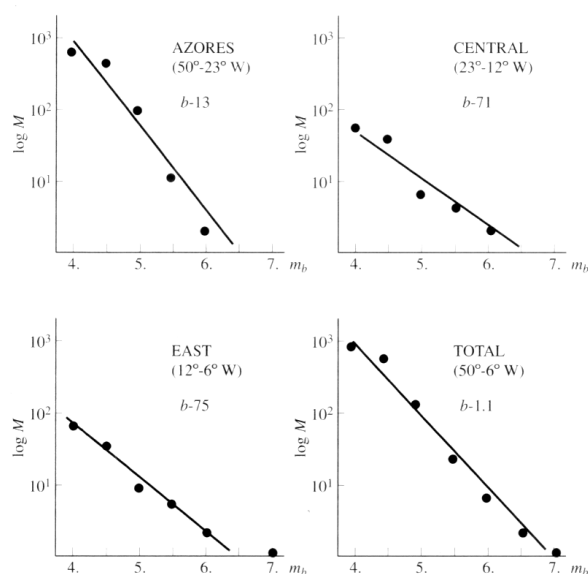


Figura 9. Distribución de frecuencia en función de la magnitud para distintas zonas del contacto Azores-Gibraltar junto con los ajustes correspondientes. Fuente: Udías y Mezcua, 1997

La constante b oscila entre 0,6 y 1,5, siendo su valor más común muy cercano a la unidad. Este valor se ha relacionado con las características físicas de cada región, de forma que un valor alto de b implica que predomina el número de terremotos de pequeña magnitud, y, por tanto, la región tiene poca resistencia, y uno bajo que predominan los terremotos de magnitud mayor, indicando una mayor resistencia en el material. La variación con el tiempo del valor de b para una región concreta se ha utilizado como índice a tener en cuenta en el problema de la predicción de terremotos. Sin embargo, esta variación puede ser producida por la falta de homogeneidad en los datos y no representar un efecto de la variación en las condiciones del material, lo que disminuye su utilidad como predictor.

Otro punto de vista para estudiar la ocurrencia de los terremotos es considerar su naturaleza fractal de acuerdo con la teoría de Mandelbrot. La primera aplicación de esta teoría a la ocurrencia de los terremotos se debe a Andrews y Aki hacia 1980. La relación entre el número y tamaño de los terremotos se puede expresar en una ley de potencia y por lo tanto considerar el fenómeno como fractal. De esta forma el coeficiente b se puede relacionar con la dimensión fractal. Como b vale aproximadamente la unidad, la dimensión fractal es dos, lo que indica que los terremotos ocupan una superficie y está de acuerdo con la observación de que los terremotos están generados por fallas. Basándose en la naturaleza fractal de los terremotos algunos autores han propuesto un mecanismo de la generación de terremotos en términos de procesos de criticalidad autoorganizada. Esto quiere decir que el sistema de la corteza terrestre bajo la acción de los esfuerzos tectónicos evoluciona por sí misma hasta un estado crítico en el que los terremotos ocurren distribuidos en tamaños según una ley de potencia. Algunos autores también apuntan que el proceso puede ser prácticamente caótico, lo que limitaría en efecto la posibilidad de su predicción.

Los valores de b de una serie de terremotos de una determinada falla o zona sismogénica pueden también servir para determinar el estado elástico del medio en el que se produce la fractura. La mayoría de los terremotos suelen suceder en forma de grupos estrechamente relacionados en el tiempo y el espacio. De estos grupos, los que están formados por la ocurrencia de un terremoto de magnitud mayor que el resto reciben el nombre de series de réplicas y premonitores, dependiendo de si ocurren después o antes que el terremoto principal. En experimentos realizados en el laboratorio se ha comprobado que las réplicas no son resultado pasivo de la ocurrencia del terremoto principal, sino una continuación activa de la liberación de energía provocada por él. Cuando en una serie de terremotos no se distingue ningún sismo principal, se denomina un enjambre de terremotos.

Una clasificación bastante extendida de estas series o grupos de terremotos es la dada por Mogi (1963), que distingue tres tipos de series:

Tipo I. No se producen premonitores. El sismo principal es el primero seguido de una serie de terremotos de menor magnitud. Estas series se dan en materiales muy homogéneos, en los que los esfuerzos acumulados no se liberan hasta que exceden la resistencia de las rocas, produciéndose el sismo principal. A éste le sigue una serie de réplicas por reajuste de los esfuerzos en el volumen que rodea la fractura que ha producido el sismo principal. Estas series están caracterizadas por un valor bajo de b entre 0.4 y 0.6 aproximadamente.

Tipo II. Se dan premonitores cuyo número aumenta rápidamente, antes del principal. La ocurrencia de estos premonitores se explica por la presencia de heterogeneidades en los materiales. Los esfuerzos acumulados se van relajando en forma de pequeños terremotos, debido a la diferencia en la resistencia de los materiales, antes de que se produzca el terremoto principal. Estas series tienen un valor de b intermedio entre 0.5 y 0.9 aproximadamente.

Tipo III. La secuencia de terremotos sucede en la forma de un enjambre en el que no se distingue un terremoto principal. Este tipo de series corresponde a zonas de materiales muy heterogéneos y distribuciones de esfuerzos no uniformes y muy concentrados. Estas series tienen un valor de b intermedio entre 0.8 y 1.5 aproximadamente.

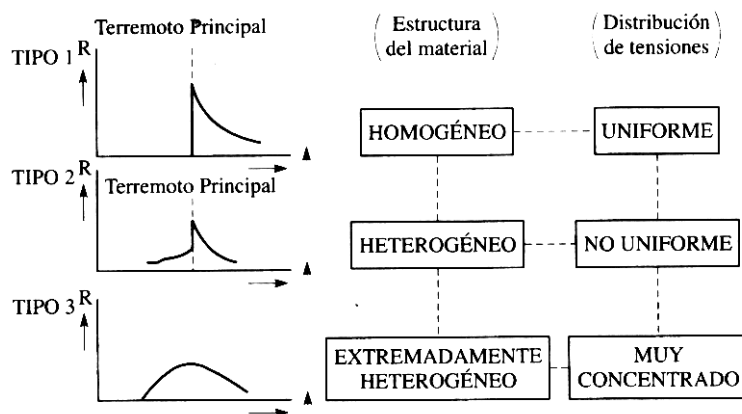


Figura 10. Distribución de premonitores y réplicas según Mogi. Fuente: Udías y Mezcua, 1997.

En las series de terremotos es de interés estudiar su distribución en el espacio, en el tiempo y la de sus magnitudes. Las réplicas se suelen encontrar espacialmente agrupadas, cubriendo la extensión de la zona total de fractura, lo que permite calcular su área a partir de la zona de réplicas. En la distribución de las réplicas a lo largo del área de fractura se aprecian a veces migraciones de los terremotos, de una parte a otra, y concentraciones en sus extremos. Del estudio de la serie de réplicas se deduce que su número y la duración de la actividad de la serie, hasta que alcanza un cierto nivel de magnitud, tienden a ser más larga según aumentamos el

valor de la magnitud del terremoto principal. En este tipo de procesos, el efecto de cada suceso puede superponerse de forma acumulativa a los efectos de los sucesos precedentes. El número acumulativo de réplicas en una serie, para un instante t , viene dado por:

$$N(t) = \frac{p}{(t + K)} N(t)$$

siendo p y K dos constantes. Esta relación fue establecida empíricamente, estudiando series de réplicas, por Omori, en 1884. Esta ley, sin embargo, no puede definir completamente el proceso, porque al estimar el número de terremotos en un intervalo, se ignora cuál es la forma de disminuir las magnitudes en él. Físicamente sería más conveniente, para representar la actividad, una expresión más claramente relacionada con la liberación de energía.

Debido a que existe un límite en la detección de las réplicas, de acuerdo con su magnitud, se puede establecer que los terremotos grandes ($M > 7$) superficiales tienen series de réplicas que pueden durar varios meses, mientras que terremotos de pequeña magnitud ($M < 5$) tienen réplicas concentradas en un corto tiempo. En el terremoto de Alaska de 1964 ($M = 8,4$), la serie de réplicas se extendió por más de un año y medio en un área de 360.000 km², de las cuales, más de 2.000 fueron localizadas.

La distribución de las magnitudes en una serie de réplicas se puede estudiar a partir de la relación vista anteriormente. Para estas series, el margen de variación de b (0,4-1,8) es mayor que el obtenido para la sismicidad durante periodos largos de tiempo en zonas activas.

En general, las series de réplicas en las que predomina el terremoto principal tienen valores altos de b y los enjambres con varios terremotos principales de casi igual magnitud valores bajos de b .

14.5 Modelos temporales de recurrencia

En el estudio de la sismicidad de una región, tan importante como la distribución espacial de los terremotos lo es la de su ocurrencia en el tiempo. Esta última puede considerarse como un proceso de sucesos puntuales, resultado de la relajación de esfuerzos que actúan de forma continua sobre una región. El análisis de las regularidades de las series temporales de terremotos, desde el punto de vista estadístico, es la forma más adecuada de estudio en este caso. Para estudiar estas series es necesario precisar los límites de la región a que se extienden. No es lo mismo estudiar la distribución de grandes terremotos a nivel global, que terremotos pequeños en una región de tan sólo algunas decenas de kilómetros cuadrados.

Desde el punto de vista estadístico, el modelo más sencillo para estudiar la ocurrencia de terremotos en el tiempo es el de la distribución de Poisson. Esta distribución supone que los terremotos son sucesos independientes, es decir, que la ocurrencia de uno no influye en la del

siguiente. De acuerdo con esta distribución, si la media de los terremotos sucedidos en un intervalo t es λ , la probabilidad de que sucedan n terremotos en dicho intervalo es:

$$P(n) = \frac{\lambda^n e^{-\lambda}}{n!}$$

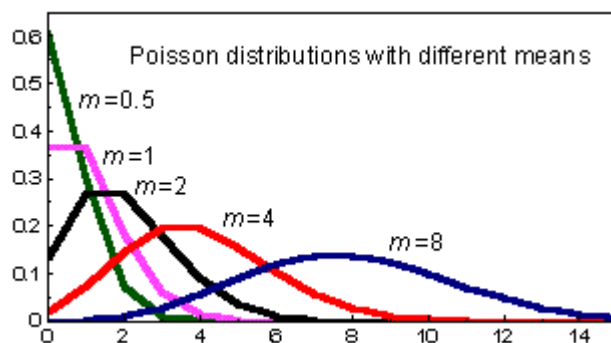


Figura 11. Distintas distribuciones de Poisson de una serie de terremotos con distintos valores de la media $\lambda=0.5$, 1, 2, 4 y 8.

De forma equivalente, si la ocurrencia de terremotos sigue una ley de Poisson, los intervalos de tiempo entre dos terremotos consecutivos δt , sigue una distribución exponencial, de forma que la probabilidad de dos terremotos separados para un intervalo δt es:

$$P(\delta t) = \lambda e^{-\lambda \delta t}$$

Esta representación de la ocurrencia de los terremotos, como un proceso de Poisson, tiene muchas limitaciones, pues físicamente es muy difícil concebir que un terremoto no influya en la ocurrencia del siguiente, si se dan en la misma zona y separados por un tiempo pequeño. Este fenómeno lleva a un agrupamiento de la ocurrencia de los terremotos, que se traduce en la definición de enjambres de terremotos, premonitores y series de réplicas que estudiaremos más adelante. Sólo si eliminamos estos sucesos y nos limitamos a terremotos grandes a nivel global, su ocurrencia puede aproximarse por un proceso de Poisson. Si consideramos la ocurrencia de terremotos de pequeña magnitud en una región relativamente pequeña, la interrelación entre ellos es grande y los agrupamientos muy predominantes. Para el estudio de las series temporales de estos terremotos, se ha propuesto una variedad de modelos estocásticos.

Bibliografía

- [1] A. Udías y J. Mezcua. (19979. Fundamentos de Geofísica. Ed. Alianza Universidad. Cap.12 y 22.

Tema 15. Caracterización de terremotos. Identificación de fases sísmicas en un sismograma. Localización hipocentral. Intensidad sísmica. Escala EMS-98. Definiciones de magnitud

15.1. Caracterización de terremotos. Identificación de fases sísmicas en un sismograma

En un sentido amplio, entendemos por *sismograma*, el registro gráfico del movimiento del suelo en un determinado lugar debido a un terremoto. La apariencia de un sismograma refleja el efecto combinado de la fuente sísmica, el camino por el que se han propagado las ondas sísmicas, las características del sismómetro y el ruido ambiental del lugar de registro.

En líneas generales, podemos decir que los procesos desencadenados en el foco producen una liberación de energía que se propaga en forma de ondas elásticas por toda la Tierra, produciendo así un movimiento en la superficie que es registrado en forma de sismogramas. Las ondas elásticas suelen clasificarse en dos grupos: ondas internas y superficiales.

Las ondas internas (P y S) siguen trayectorias en general curvas y son reflejadas, refractadas y convertidas de P a S o viceversa en las diferentes discontinuidades que encuentran a su paso. La velocidad de las ondas P es mayor que la de las S.

Las ondas superficiales circunscriben su perturbación a la corteza y manto superior y pueden ser de dos tipos: ondas Rayleigh (LR) y Love (LQ). La velocidad de las ondas LQ es mayor que la de las ondas LR y la de ambas menor que la de las ondas internas.

En el foco se libera energía que se propaga en forma de ondas internas y dependiendo de la posición del foco y del medio por el cual viajan las ondas sísmicas se irán formando unos u otros tipos de ondas, cada una de ellas con velocidades de propagación diferentes, que llegarán a la estación de registro en distintos tiempos. Cada una de estas llegadas recibe el nombre de *fase sísmica*.

Un aspecto fundamental en la interpretación de los sismogramas consiste en identificar en éste cada una de las llegadas o fases sísmicas, siendo su correcta identificación el primer paso para resolver el problema inverso en la sismología, es decir, conocer las características del foco y la estructura de la Tierra a partir de las observaciones en la superficie de la misma.

Desafortunadamente, una buena identificación de las fases sísmicas no es una tarea fácil y requiere experiencia. Debido a las reflexiones, refracciones y conversiones en multitud de discontinuidades más o menos fuertes en el interior de la Tierra, existe una gran cantidad de fases débiles que enturbian a las principales. Además, todo el *ruido sísmico* se suma a la señal

sísmica y todas las fases menos la primera llegada están oscurecidas por la coda de las fases precedentes, producida por la interacción de las ondas con pequeñas heterogeneidades de la Tierra¹. Así pues el sismograma se ve prácticamente como un continuo de ondas y no como un conjunto de llegadas aisladas.

Por todo esto es imprescindible tener un buen conocimiento de las características de cada una de las fases sísmicas (velocidad de propagación, polaridad, periodos predominantes,...) para poder identificarlas ayudándonos de curvas dromocrónicas para cada fase (llegadas teóricas), rotación del sismograma, filtrado del mismo en distintas bandas de frecuencia, etc.

A partir de ahora caracterizaremos los eventos sísmicos en distintas categorías, basándonos en su distancia epicentral y de acuerdo a su profundidad focal. La razón por la cual se hace esta clasificación, cuyos límites no están bien definidos, es que los sismogramas de cada una de estas categorías presentan características relativamente homogéneas.

En la caracterización de los terremotos respecto a la distancia epicentral, existen numerosas clasificaciones. Una de ellas considera los sismos a distancias menores de 1° como *eventos locales* y a distancias entre 1° y 10° como *eventos regionales*. Los sismogramas de ambos grupos están dominados por la energía sísmica que se propaga a través de la corteza y/o en la discontinuidad del Moho. Por otro lado, aquí consideraremos como *eventos telesísmicos* o *telesismos* los terremotos registrados a distancias epicentrales mayores de 10° aproximadamente, aunque muchos autores prefieren hablar de telesismos para referirse a terremotos registrados a más de 20° o incluso 30°. Los sismogramas registrados a distancias comprendidas entre 10° y 30° corresponden principalmente a ondas propagadas en el manto superior y por lo tanto suelen ser muy complejos debido a la interacción de estas ondas con las discontinuidades de esta capa. A distancias epicentrales entre 30° y 105°, se registran ondas propagadas por el manto inferior y corresponden a sismogramas relativamente simples. Para distancias mayores, las ondas se propagan por el núcleo o están difractadas en la discontinuidad manto-núcleo (CMB) y los sismogramas vuelven a ser más complejos.

Atendiendo a la profundidad, podemos clasificar los terremotos en *superficiales* (profundidades focales menores de 60 km), *intermedios* (profundidades entre 60 km y 300 km) y *profundos* (profundidades mayores de 300 km).

¹ Fenómeno conocido como dispersión anómala o esparcimiento (o scattering en la literatura inglesa).

15.1.1. Terremotos locales y regionales ($0^\circ < \Delta < 10^\circ$)

En la *Figura 1* puede observarse un modelo simplificado de la corteza con algunas de las ondas que en ella se propagan. Nos referimos aquí a terremotos cuyo foco se sitúa en la corteza. En general los periodos de estas ondas son pequeños² y las fases predominantes son:

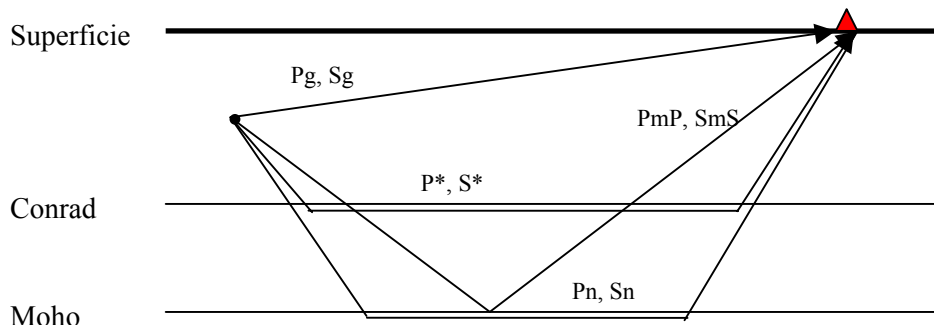


Figura 1. Modelo simplificado de la corteza con sus fases principales.

- *Pg* y *Sg*: Ondas P y S directas que viajan por la capa granítica. En modelos más simples de corteza se consideran aquí las ondas P y S directas que se propagan por toda la corteza. Estas son las primeras llegadas P y S en los eventos locales. En realidad, las ondas englobadas bajo este nombre en los sismos regionales corresponden a ondas canalizadas o refractadas en discontinuidades de la corteza superior, por lo que no aparecen en largas trayectorias oceánicas donde no existe la capa granítica ni en terremotos originados en la corteza inferior.
- *Pn* y *Sn*: Ondas difractadas en el Moho (donde viajan con la velocidad correspondiente al manto superior). Aparecen a partir de la distancia epicentral crítica de refracción (en torno a unos 100 Km en la corteza continental). A partir de distancias epicentrales de unos 150 o 200 km (dependiendo de la estructura de la corteza) estas son las primeras llegadas en el sismograma. A partir de unos 13° normalmente las amplitudes de estas fases son muy pequeñas, dejando de ser identificables.
- *P** y *S**: (También llamadas *Pb* y *Sb*). Ondas difractadas en la discontinuidad de Conrad. Aparecen también a partir de su distancia crítica y son las primeras fases en distancias entre 100 y 150 km aproximadamente. A distancias mayores se sitúan entre las fases *Pn* y *Pg* o *Sn* y *Sg* y son muy difíciles de identificar.
- *PmP* y *SmS*: Ondas reflejadas en el Moho. Constituyen fuertes llegadas a partir de los 80-120 km.

² En terremotos pequeños y moderados.

- *Lg*: Onda canalizada en toda la corteza. Forma llegadas de gran amplitud con movimiento predominantemente transversal (se observa mejor en las componentes horizontales) y periodos entre 1 y 6 s. Su velocidad de propagación es aproximadamente la velocidad media de propagación de las ondas S en la corteza superior (~ 3.5 km/s) y se observa a continuación de la llegada de la fase Sg. Esta fase se observa únicamente en trayectorias continentales³. Está formada a partir de la interferencia de ondas reflejadas de forma supercrítica en el Moho, por lo que aparece a distancias epicentrales mayores de 1° .
- *Rg*: Ondas de tipo Rayleigh generadas por eventos muy superficiales (terremotos, explosiones,...) registrados a distancias menores de unos 600 km. Se propagan con una velocidad promedio de unos 3 km/s y se registran mejor en la componente vertical mostrando una clara dispersión.
- *T*: Fase observada en sismómetros situados en el fondo del océano (OBS), islas o zonas costeras. Se debe a la propagación de ondas longitudinales en el canal de baja velocidad (~ 1.5 Km/s) situado a una profundidad de unos 800-1300 m dentro de los océanos⁴. Esta fase llega después de las fases P y S en forma de paquetes de onda de larga duración y aproximadamente monocromáticos de alta frecuencia (> 1 Hz).

15.1.2. Telesismos ($\Delta > 10^\circ$)

En la *Figura 2* podemos observar las principales fases existentes en los telesismos. A continuación describimos algunas de estas fases.

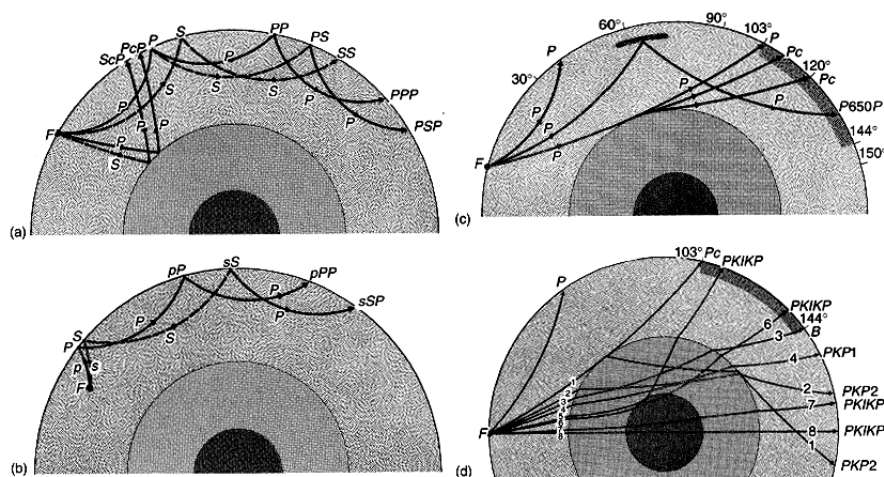


Figura 2. Modelo simplificado de la Tierra con sus fases principales. a) Foco superficial, b) Foco profundo, c) ondas difractadas en la discontinuidad manto-núcleo, d) Ondas que se propagan en el interior del núcleo de la Tierra (Kulháněk, 2002).

³ Si la línea epicentro-estación cubre un tramo de más de 2° por el océano, desaparecen las ondas Lg.

⁴ SOFAR (Sound Fixing and Ranging).

- *P* y *S*: Ondas *P* y *S* directas, que viajan por el manto sin que se produzca ninguna reflexión o refracción. Son las primeras llegadas en el rango de distancias epicentrales entre 10° y 103° aproximadamente. La fase *P* es normalmente más fuerte en la componente vertical y la *S* en las componentes horizontales.
- *PP*, *PS*, *SP*, *SS*, *PPP*, *PSP*, ...: Ondas reflejadas una o más veces en la superficie libre, donde pueden o no convertirse de *P* a *S* o viceversa. Las fases *PS* y *SP* aparecen sólo a distancias mayores de 40°. A distancias mayores de unos 100° las fases *PP* y *SS* son normalmente las mayores ondas internas registradas.
- *PcP*, *ScS*, *PcS* y *ScP*: Ondas reflejadas en la discontinuidad manto-núcleo, donde pueden convertirse. Las mayores fases de este tipo se registran a distancias de unos 40° o menores.
- *pP*, *sS*, *pS* y *sP*: (También llamadas *fases de profundidad*). Ondas que ascienden desde el foco (situado en el manto) hacia la superficie libre, donde se reflejan hacia el manto y vuelven a llegar a la superficie sin reflejarse o refractarse en discontinuidades. Preceden inmediatamente a las fases directas correspondientes. Estas fases, principalmente la *pP*, son utilizadas para conocer la profundidad hipocentral a partir de la diferencia de los tiempos de llegada entre la fase de profundidad y la directa ($t_{pP}-t_P$).
- *P_{diff}* y *S_{diff}* (o *P_{dif}* y *S_{dif}* o *Pc* y *Sc*): Ondas difractadas en la discontinuidad manto-núcleo. Llegan débilmente a partir de distancias de unos 103° y en toda la *zona de sombra*, situada entre 103° y 140°⁵. Normalmente las llegadas son emergentes y de grandes periodos.
- *PKP*, *SKS*, *PKS* y *SKP*: (Las fases *PKP* y *SKS* también se denominan *P'* y *S'*). Ondas internas refractadas hacia el interior del núcleo externo⁶ donde viajan siempre como onda *P* y refractadas de nuevo hacia el manto. Aparecen a partir de distancias de unos 144° (*PKP*) o unos 60° (*SKS*).⁷
- *PKIKP*, *PKJKP*, *PKIKS*...: Ondas refractadas hacia el núcleo externo y núcleo interno y a continuación hacia el núcleo externo y manto. Si viajan en el núcleo interno como onda *P*, se expresan con la letra *I*, y si lo hacen como onda *S*, con la letra *J* (estas últimas muy difíciles de observar debido a su baja amplitud). Las primeras pueden confundirse con las fases del tipo *PKP*, por lo que suelen designarse también como *PKP*, *PKS*, etc.
- *PKiKP*: Ondas reflejadas en la superficie de discontinuidad núcleo externo-interno.

⁵ Otros autores sitúan la zona de sombra entre 105° y 143°.

⁶ El tramo de paso por el núcleo externo se denomina *K* por la palabra alemana *Kernwellen* para designar al núcleo.

⁷ En el núcleo externo la velocidad de las ondas *P* decrece respecto a la del manto (*zona de baja velocidad*). Debido a esto, los rayos que salen ligeramente más inclinados del foco que los que producen *P_{diff}* llegan a distancias epicentrales mayores de 180°. Conforme van saliendo más inclinados del foco la distancia epicentral de llegada de esta fase va siendo menor hasta llegar a unos 144° (rama *PKP₁*). Si todavía el rayo sale más inclinado, la distancia epicentral de llegada vuelve a aumentar hasta unos 156° (rama *PKP₂*). Para llegadas epicentrales de unos 144° las dos ramas *PKP* interfieren constructivamente, dando lugar a una *cáustica* donde se concentra la energía. En realidad pueden observarse tres ramas en las curvas dromocronas *PKP* para ciertas distancias epicentrales. En la fase *SKS* no hay zona de baja velocidad, puesto que la velocidad de las ondas *P* en el núcleo externo es mayor que la de las ondas *S* en el manto.

- *PKKP*: Onda reflejada desde el interior del núcleo externo en la discontinuidad CMB.
- *LR* y *LQ*: Ondas superficiales Rayleigh y Love. Excepto a distancias epicentrales muy pequeñas (locales) las ondas superficiales contienen la mayor parte de la energía radiada por terremotos superficiales y algunos intermedios. En un sismograma se observa primero la llegada de la fase LQ (en las componentes horizontales) y a continuación la llegada de la fase LR (sobre todo en la componente vertical). Ambos tipos de fases contienen ondas dispersadas, por lo que en los sismogramas se observan estas fases como *trenes de ondas*.

En general todos los *terremotos intermedios y profundos* pueden considerarse como los lejanos, aunque en ellos no se observan ondas superficiales. Un rasgo distintivo de estos terremotos es que en la componente vertical predomina claramente la fase P sobre la fase S.

15.2. Localización hipocentral

Se denomina *hipocentro* al punto del interior de la Tierra donde se origina y comienza el proceso de fracturación cuya propagación da lugar a un terremoto. Las coordenadas del hipocentro se dan generalmente por medio de su latitud, longitud y profundidad desde la superficie de la Tierra. Se denomina *epicentro* a la proyección en superficie del hipocentro. Un cuarto parámetro de la localización hipocentral es el tiempo de inicio, t_0 .

El problema de la localización hipocentral es uno de los más antiguos de la sismología, no obstante no se ha resuelto completamente. A continuación veremos algunos de los diferentes métodos utilizados.

Las primeras determinaciones de la localización del foco sísmico están basadas en el *método macrosísmico*, el cual utiliza la información sobre los daños producidos por un terremoto. Estos métodos sitúan el epicentro en el centro de la zona de mayor daño y éste se denomina *epicentro macrosísmico*.

Con el desarrollo y mayor utilización de los sismómetros la localización hipocentral comienza a realizarse a partir de los datos registrados en los sismogramas, dando lugar a los métodos instrumentales, siendo los métodos gráficos los primeros en utilizarse.

Los *métodos gráficos* están basados en un proceso iterativo. Consideramos N estaciones ($N > 4$) donde se han medido llegadas de las fases P y S. Utilizando curvas dromocronas y la medida $\delta t = t^S - t^P$ en una estación, obtenemos la distancia epicentral, Δ , y el tiempo de viaje de la onda P, t^P , para esa estación. Por lo tanto podemos deducir el tiempo origen del terremoto, t_0 , por medio de $t_0 = t^P - t'^P$. Utilizando este valor de t_0 , calculamos el tiempo de viaje de la onda P para todas las demás estaciones, mediante $t_i^P = t_i^P - t_0$, y a partir de este tiempo, t_i^P , con las curvas dromocronas calculamos Δ_i para cada estación. Usando Δ_i como radio, trazamos

circunferencias con centro en cada estación i , de modo que la región epicentral vendrá dada por el área de intersección de todas las circunferencias. Si el epicentro se encuentra a cierta profundidad, deberemos utilizar curvas dromocronas para distintas profundidades hasta obtener el mejor ajuste.

Más precisos que los métodos gráficos son los métodos numéricos. Aunque su fundamento teórico fue desarrollado a principios del siglo XX, su aplicación no se generaliza hasta la aparición de los ordenadores hacia 1960.

La base de estos métodos consiste en la linealización de las ecuaciones que relacionan los tiempos de llegada de las fases sísmicas con la localización del terremoto y la estación.

En general, para cierta estructura de velocidades en el interior de la Tierra, el tiempo de llegada de una determinada fase a una estación vendrá dado por:

$$(1) \quad t_i = t_i(\phi_i, \lambda_i; \phi_0, \lambda_0, h_0, t_0)$$

Siendo (ϕ_i, λ_i) las coordenadas de la estación i , (ϕ_0, λ_0, h_0) las coordenadas del hipocentro y t_0 el tiempo origen del terremoto.

Si conocemos una solución aproximada de la localización del hipocentro $(\phi_0^0, \lambda_0^0, h_0^0, t_0^0)$, podemos linealizar la ecuación (1) por medio de un desarrollo de Taylor centrado en la solución inicial:

$$(2) \quad t_i = t_i^0 + \delta t_i + \left. \frac{\partial t_i}{\partial \phi} \right|_{\phi_0^0} \delta \phi + \left. \frac{\partial t_i}{\partial \lambda} \right|_{\lambda_0^0} \delta \lambda + \left. \frac{\partial t_i}{\partial h} \right|_{h_0^0} \delta h$$

Donde las derivadas están evaluadas en la solución inicial y $t_i^0 = t_i(\phi_0^0, \lambda_0^0, h_0^0, t_0^0)$. Definiendo los residuos, r_i , como las diferencias entre el tiempo observado y el calculado $r_i^0 = t_i - t_i^0$, podemos escribir la ecuación (2) en forma matricial como:

$$(3) \quad r_i^0 = A_{ij} \delta x_j^0$$

Donde i toma valores desde 1 hasta el número de estaciones, N , y j desde 1 hasta 4, que es el número de incógnitas. La matriz A está formada por las derivadas parciales de las funciones de los tiempos de llegada en cada estación evaluadas en la solución inicial.

Debido a la simplificación del modelo y a los errores en las observaciones, el problema sólo tiene solución si está sobredeterminado, es decir que tenemos más ecuaciones que incógnitas, resolviéndose la ecuación (3) por el método de *mínimos cuadrados*.

La solución de la ecuación (3) nos proporciona la variación de la solución inicial $\delta x_j^0 = (\delta \phi^0, \delta \lambda^0, \delta h^0, \delta t^0)$ para que nuestro modelo se ajuste mejor a los datos observados. De este modo construimos la siguiente solución de prueba como: $(\phi_0^1, \lambda_0^1, h_0^1, t_0^1) = (\phi_0^0, \lambda_0^0, h_0^0, t_0^0) + (\delta \phi^0, \delta \lambda^0, \delta h^0, \delta t^0)$ y repetimos el proceso de forma iterativa, hasta que los residuos, r_i , son menores que un valor dado.

A partir del estudio de la matriz A , puede deducirse que el parámetro con mayor error es la profundidad focal, y que este error sólo disminuirá con la incorporación de observaciones próximas al hipocentro.

En general los parámetros hipocentrales obtenidos para cada terremoto tienen un error asociado debido a diferencias entre el modelo de Tierra utilizado y la realidad, equivocaciones en la lectura del tiempo de llegada de cada fase, mala disposición geográfica de las estaciones para un epicentro dado, etc. Debido a esto, siempre que se realiza un cálculo de la localización hipocentral de un terremoto se incluyen datos que representan la incertidumbre de este resultado.

Algunos de los datos indicativos de la precisión que podemos esperar en la localización son:

- R_{min} : Distancia desde el epicentro a la estación más cercana a éste utilizada en el cálculo.
- Gap : Máximo ángulo acimutal, tomando el epicentro como vértice, entre dos estaciones (ver *Figura 3*).
- $Nsta$: Número de estaciones utilizadas en el cálculo.
- Nph : Número de fases utilizadas en el cálculo.

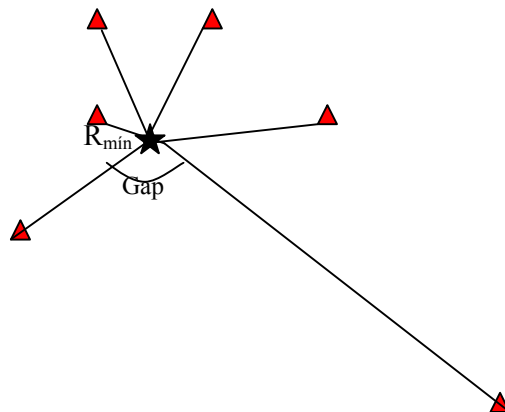


Figura 3. Disposición epicentro estaciones.

Normalmente, para un terremoto cortical necesitaremos que R_{min} sea al menos del orden de la profundidad focal para que ésta pueda ser determinada. En caso contrario suele ser necesario presuponer una profundidad fija en el proceso de cálculo. Por otro lado la localización es mejor cuanto menor sea el Gap , es decir, cuando el epicentro está rodeado de estaciones de forma uniforme. Esto a veces es difícil de conseguir, por ejemplo para terremotos pequeños y moderados situados en la costa atlántica de la Península Ibérica o en el mismo océano.

Los parámetros más utilizados para representar la incertidumbre de la localización se deducen a partir de la matriz de covarianzas obtenida en el método de mínimos cuadrados y de la diferencia entre tiempos observados y teóricos. Estos son:

- *Rms*: Raíz cuadrada de la suma de los cuadrados de las diferencias entre los tiempos de llegada observados y los teóricos para cada una de las fases partido por el número de grados de libertad.
- *dz*: Error en la profundidad focal (en km) deducido a partir de la varianza de este parámetro.
- *dt*: Error en el tiempo origen (en s) deducido a partir de la varianza de este parámetro.
- *Elipse de error*: Elipse centrada en el epicentro obtenido donde, con cierta probabilidad (normalmente el 90 %) se sitúa el epicentro real. Esta elipse se deduce a partir de la proyección de la matriz de covarianza en el plano horizontal. Se representa por medio de la longitud en km de sus semiejes mayor y menor y el azimut de su semieje mayor.

Como ejemplo mostraremos la localización hipocentral que realizó el IGN para dos terremotos ocurridos en 2007 uno al suroeste del Cabo de San Vicente ($M_w=6.1$) y otro al noroeste de Arenales de San Gregorio, Ciudad Real ($M_w=4.7$) y la localización que realizó el U. S. Geological Survey para un terremoto próximo a la costa de Perú ($M_w=8$) también en 2007.

Tiempo (UTC)		Latitud	Longitud	Profundidad	Localización	M_w
12/02/2007 10:35:25		35.9565 °	-10.4095 °	64 km	SW Cabo S. Vicente	6.1
Nsta	Nph	$R_{\min}$	Gap	Rms	dt	dz
74	108	1.95 °	165 °	0.79 s	1.26 s	31 km

Semiejes elipse de error (90%): 12.3 km; 8.7 km; (Azimut: 52°)

Tiempo (UTC)		Latitud	Longitud	Profundidad	Localización	M_w
12/08/2007 7:47:06		39.3534 °	-2.9903 °	5 km	Arenales de S. Gregorio	4.7
Nsta	Nph	$R_{\min}$	Gap	Rms	dt	dz
56	146	0.81 °	33 °	0.85 s	0.28 s	3 km

Semiejes elipse de error (90%): 2.6 km; 2.0 km; (Azimut: 110 °)

Tiempo (UTC)		Latitud	Longitud	Profundidad	Localización	M_w
15/08/2007 23:40:57		-13.354 °	-76.509 °	39 km (fija)	Cerca costa Perú	8.0
Nsta	Nph	$R_{\min}$	Gap	Rms	dt	dz
275	275	155 km	29 °	0.84 s	--	--

Incertidumbre localización epicentral: ± 5.2 km.

Como podemos observar aún en casos con buena cobertura de estaciones las incertidumbres en las localizaciones hipocentrales para redes regionales y globales son del orden de varios kilómetros.

15.3. Intensidad sísmica. Escala EMS-98

La *intensidad sísmica* o *intensidad macrosísmica* es una clasificación de la severidad de la sacudida del suelo producida por un terremoto, basada en los efectos observados en un área limitada. Es por lo tanto un parámetro descriptivo de la fuerza de un terremoto en un determinado lugar. A un mismo terremoto se le asocian un conjunto de intensidades, cada una de ellas correspondiente al lugar donde se ha evaluado este parámetro.

Aunque la intensidad se aplica directamente para expresar como se ha sentido un terremoto en un cierto lugar, puede usarse también para designar el tamaño de un terremoto, a partir del valor $I_{\text{máx}}$ o I_0 ; donde $I_{\text{máx}}$ es la intensidad máxima observada e I_0 , la intensidad que se supone en el epicentro.

A lo largo de la historia han aparecido numerosas escalas de intensidad en las que se define cada grado por la ocurrencia de una serie de fenómenos, los cuales exigen la observación de cómo la población ha percibido el sismo, de daños sobre distintos tipos de construcciones, etc.

Las primeras escalas de intensidad sísmica se deben al trabajo de Rossi y Forel en 1883 que proponen la escala Rossi-Forel de diez grados. Ésta fue revisada por Mercalli, ampliada a doce grados por Cancani y definida de modo completo por Sieberg, adoptando el nombre de Mercalli-Cancani-Sieberg (MCS). La escala MCS es la base de las distintas escalas “Mercalli Modificada (MM)” y la escala MSK⁸ y sus posteriores revisiones. En Europa, sin embargo, la escala utilizada normalmente es la *Escala Macrosísmica Europea (EMS-98)*, la cual está basada en la experiencia adquirida con las escalas anteriormente citadas. No obstante la mayoría de estas escalas tienen doce grados y son equivalentes, aunque difieren en su grado de complejidad. Otra escala importante es la japonesa JMA (Japanese Meteorological Agency). Aunque inicialmente constaba de siete grados, una reciente modificación la convierte en la práctica en una escala de diez grados (JMA 96).

Pese a que cuando se empezó a extender la localización de epicentros y el cálculo de magnitudes por medio de medidas instrumentales decreció el interés por las observaciones macrosísmicas, desde mediados de los 70 estas volvieron a cobrar importancia, ya que el cálculo de intensidades es esencial para la revisión de la sismicidad histórica y de gran importancia en los estudios de peligrosidad sísmica, vulnerabilidad y riesgo sísmico.

⁸ MSK: Medvedev, Sponheuer y Karnik.

Dada su relevancia, a continuación describimos brevemente los aspectos fundamentales de la escala EMS-98.

Como hemos mencionado anteriormente esta escala consta de doce grados (del I al XII). En la *Tabla 1* se muestra un pequeño resumen del significado de cada grado de esta escala.

Intensidad EMS	Resumen
I No sentido	No sentido
II Apenas sentido	Sentido sólo por muy poca gente en situación de reposo en casa.
III Débil	Sentido en el interior de las casas por poca gente. Las personas en reposo sienten un ligero balanceo o temblor.
IV Ampliamente Observado	Sentido en el interior por mucha gente, fuera por muy pocos. Algunas personas se despiertan. Tintineo de ventanas y vajillas.
V Fuerte	Sentido en el interior por la mayoría y en el exterior por pocos. Muchas personas se despiertan. Pocas personas se asustan. Los edificios tiemblan. Los objetos colgados oscilan considerablemente. Pequeños objetos se mueven. Puertas y ventanas se balancean o se cierran.
VI Levemente dañino	Muchas personas salen asustadas de casa. Algunos objetos se caen. Muchas casas sufren pequeños daños no estructurales como grietas o caída de pequeños trozos de yeso o enlucido.
VII Dañino	La mayoría de la gente está asustada y corre fuera de casa. Los muebles se levantan y una gran cantidad de objetos cae de las estanterías. Muchos edificios ordinarios bien contruidos sufren daños moderados: grietas en las paredes, caída de yeso, colapso parcial de la chimenea. Las construcciones antiguas tienen grandes grietas y caída de tabiques.
VIII Gravemente Dañino	A mucha gente le resulta difícil permanecer en pie. Muchas casas tienen grandes grietas en los muros. En unos pocos edificios ordinarios bien contruidos se dañan seriamente los muros mientras que las estructuras antiguas y débiles pueden colapsar.
IX Destructor	Pánico generalizado. Muchas construcciones débiles colapsan. Incluso edificios contruidos adecuadamente presentan daños importantes: serios colapsos de muros y colapso parcial de la estructura.
X Muy destructor	Muchos edificios correctamente contruidos colapsan.
XI Devastador	Muchos edificios bien contruidos colapsan, incluso algunos con diseño sismorresistente.
XII Completamente devastador	Casi todos los edificios son destruidos.

Tabla 1. Breve resumen de la escala EMS-98

La diferencia principal entre la escala EMS-98 y otras escalas de intensidad es el mayor detalle en el que se describen algunos aspectos de la misma. En particular podemos citar lo relativo a los tipos de edificación, el grado de daño y las cantidades.

Un aspecto fundamental dentro de esta escala es el que alude a la *vulnerabilidad*. Para la asignación de un grado de intensidad determinado a partir de la observación de los daños en una construcción, se tiene en cuenta la estructura donde se ha observado el daño desde el punto de vista de su mayor o menor vulnerabilidad sísmica. Así un cierto daño estructural en un edificio

de mampostería corresponde a una intensidad menor que si el mismo daño se observa en un edificio de hormigón armado. En concreto se hace una distinción entre seis tipos de vulnerabilidad: A, B, C, D, E y F y cinco tipos de daños, desde los despreciables (1) a la destrucción total (5).

Respecto al modo en que se cuantifican las observaciones se opta por no asignar un porcentaje o rango de porcentajes estricto a los términos “pocos”, “muchos” y “la mayoría” (“few”, “many”, “most”) que aparecen en las definiciones de los distintos grados de intensidad. Sino que deliberadamente se representan estas cantidades por medio de un gráfico (*Figura 4*). La resolución de posibles ambigüedades debidas al solapamiento de estos términos se deja en manos del usuario, el cual elegirá la opción más consistente con el resto de los datos.

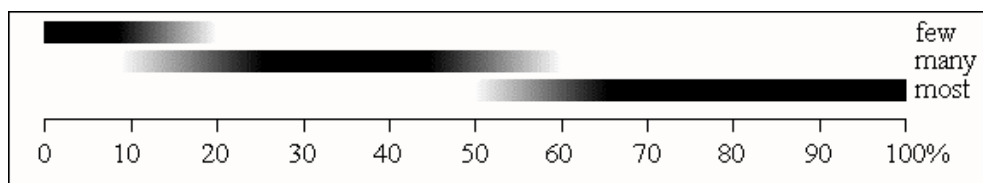


Figura 4. Definiciones de la cantidad en la EMS-98.

Algunas de las recomendaciones para su puesta en práctica son:

- El valor de intensidad siempre va asociado a un lugar lo suficientemente grande para obtener una muestra estadística, y lo suficientemente pequeño para no perder los efectos locales y que presente características homogéneas. En circunstancias generales, el lugar más pequeño no debe ser menor que un pueblo y el mayor no debe superar las dimensiones de una ciudad europea de tamaño moderado.
- Hay que entender las descripciones de cada grado como fotografías idealizadas de los efectos esperables en ese grado de intensidad y, aunque no se ajusten totalmente a la realidad optar por una u otra. Las descripciones de los grados son umbrales. De modo que si en un lugar se pasa del umbral del grado VI, pero no del VII, se le asignará un grado VI, intentando mantener el carácter de número entero. Existen casos en los que la realidad se ajusta a más de un grado de intensidad, por ejemplo al VI, VII y VIII. En ese caso se escribirá VI-VIII, lo que significa que no hay datos para decidirse por uno de estos tres valores.
- Se recomienda no considerar observaciones hechas en pisos más altos de una 5ª planta, estructuras especiales (torres, faros, puentes,...) o en lugares bajo tierra, debido a la difícil caracterización de la respuesta de estas estructuras.

En algunas ocasiones, para un terremoto dado, para el que se ha determinado la intensidad en un elevado número de puntos, es posible trazar *isosistas*, o líneas que pasan por puntos de igual intensidad y estimar I_0 .

15.4. Definiciones de magnitud

El concepto de magnitud fue introducido por Richter (1935) para proporcionar una medida instrumental objetiva del tamaño de los terremotos; frente a la intensidad sísmica, basada en percepciones hasta cierto punto subjetivas y siempre asociadas a un lugar de observación.

La magnitud, por lo tanto, trata de representar mediante un único número el tamaño de un terremoto, el cual está relacionado con la energía liberada en forma de ondas sísmicas. En la práctica es difícil, a partir de medidas sencillas en un sismograma, cumplir este objetivo, entre otras razones porque los parámetros a medir dependen del tipo de terremoto. Debido a esto se han desarrollado diferentes definiciones de magnitud, para ser usadas en un caso u otro. En la actualidad se utilizan principalmente cuatro tipos de magnitud: M_L , m_b , M_S y M_W . Aunque se ha tratado de escalar conjuntamente las diferentes magnitudes, de modo que sean consistentes en ciertos rangos, esto solo se ha logrado parcialmente.

Las primeras definiciones de magnitud se basan en las siguientes suposiciones:

- La magnitud debe ser una medida de la energía liberada en el foco en forma de ondas sísmicas, y por tanto proporcional a la velocidad del suelo, que aproximadamente viene dada por A/T , donde A es la amplitud máxima del desplazamiento del suelo y T el periodo de la onda donde se ha medido la amplitud. De modo que para una misma geometría fuente-receptor y dados dos terremotos de distinto tamaño, el terremoto más grande producirá en término medio mayores (A/T) , en el receptor.
- La atenuación de las amplitudes con la distancia, $\sigma(\Delta, h)$, es conocida en función de la profundidad del foco, h , y de la distancia epicentral, Δ .

En general, para evitar la variación de amplitudes debida al patrón de radiación de la fuente y otros efectos anómalos de cada estación, es necesario promediar las magnitudes calculadas en distintas estaciones para un mismo terremoto.

Richter definió la magnitud a partir de terremotos en California registrados a distancias menores de 600 km por sismógrafos de torsión de tipo Wood-Anderson (amplificación de 2800 y periodo natural 0.85) como:

$$(4) \quad M = \log A - \log A_0$$

Donde A es la amplitud máxima medida en mm y A_0 es una función que depende de la distancia. Esta escala fue calibrada de modo que un terremoto registrado a 100 km de distancia con una amplitud máxima de 1 mm en un sismógrafo Wood-Anderson tuviese una magnitud $M=3$ (ver Figura 5).

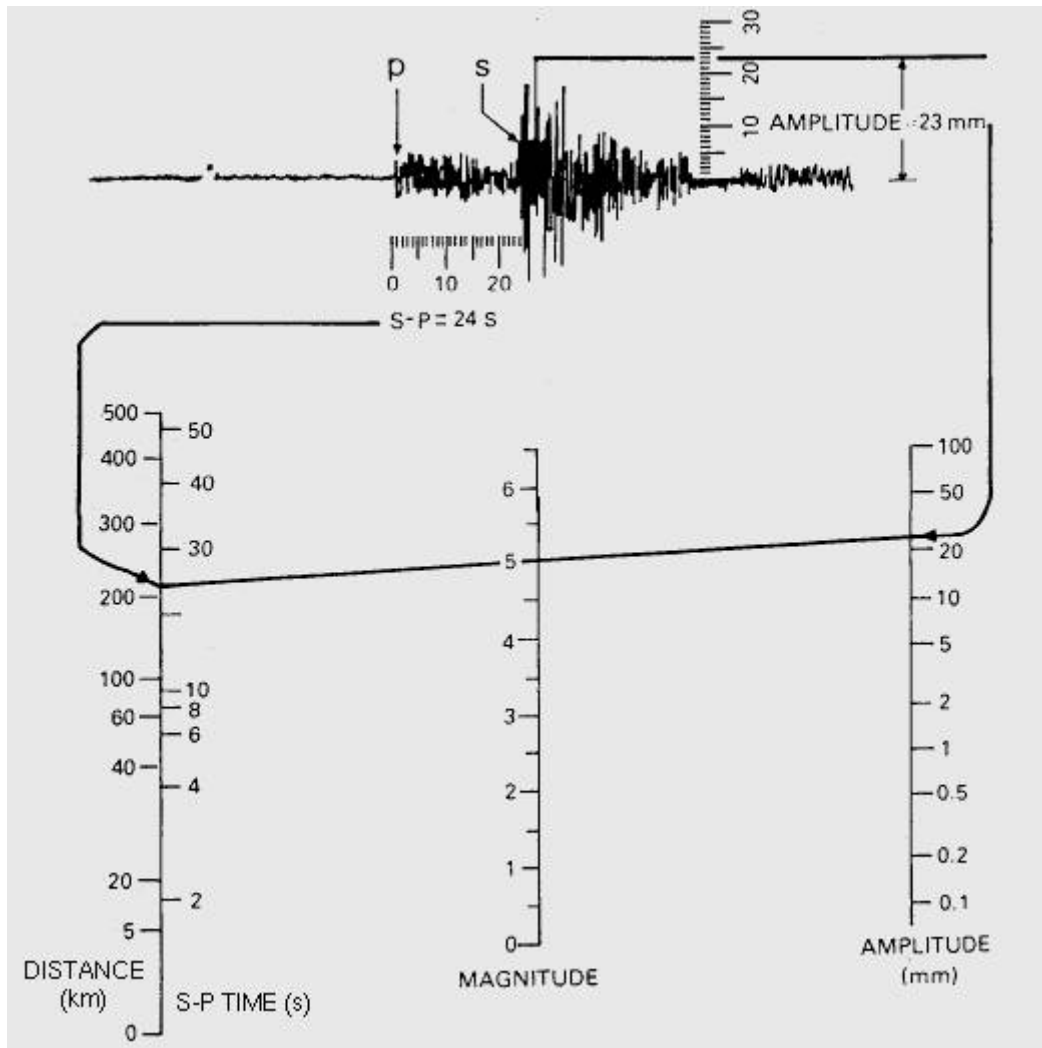


Figura 5. Determinación gráfica de la magnitud de Richter.

La magnitud Richter para terremotos locales puede generalizarse para registros obtenidos a partir de cualquier sismógrafo con periodo natural cercano a 1 s.

De este modo se define la *magnitud local* como:

$$(5) \quad M_L = \log A + 2.56 \log \Delta - 1.67$$

Donde A es la amplitud máxima del desplazamiento del suelo en micrómetros⁹ y Δ la distancia epicentral en km ($\Delta < 600$ km).

La extensión de esta definición para distancias mayores dio pie a la creación de dos escalas creadas por Gutenberg y Richter. Tenemos así:

La *magnitud de ondas internas*, m_b :

$$(6) \quad m_b = \log(A/T) + \sigma(\Delta, h)$$

Donde A es la amplitud del desplazamiento del suelo debido a las ondas internas, generalmente medida como la máxima amplitud de las ondas P en sismogramas de corto periodo ($T \sim 1$ s) en la componente vertical.

La *magnitud de ondas superficiales*, M_S :

$$(7) \quad M_S = \log(A/T) + \alpha \log \Delta + \beta$$

Donde A es la máxima amplitud en micrómetros de las ondas Rayleigh, T es el periodo (aproximadamente 20 s) y α y β son dos constantes que dependen de la atenuación de la zona.

Tanto M_S como m_b están diseñadas para ser lo más compatibles posible con M_L , es decir que obtengamos los mismos valores de magnitud en las tres escalas, pero desafortunadamente esto solo puede conseguirse en terremotos pequeños. La razón principal es que la amplitud medida en cada caso corresponde a ondas con frecuencias diferentes, y conforme un terremoto es mayor las ondas que genera contienen cada vez frecuencias menores. Debido a esto, estas *magnitudes se saturan*, es decir, que existe una cierta magnitud en cada caso, a partir de la cual aunque el terremoto sea mayor, la amplitud medida no crece (puesto que el aumento de energía está contenido en ondas con frecuencias menores que las medidas). De este modo la magnitud m_b se satura entre $m_b=5.5-6$, la magnitud M_L en torno a $M_L=6.5$ y la magnitud M_S entre $M_S=7.25-8$.

Para evitar el fenómeno descrito anteriormente Kanamori en 1977, introdujo una nueva escala de magnitud denominada *magnitud de momento*, M_W , que se define como:

$$(8) \quad M_W = \frac{2}{3} \log M_0 - 10.7$$

M_0 es el momento sísmico escalar de un terremoto en unidades de dyn·cm, fue definido por Aki en 1966 como medida del tamaño de un terremoto y se deduce a partir del espectro de bajas frecuencias, por lo que no se satura.

Además esta magnitud es consistente con M_S , dentro de su rango de validez, aunque es más difícil de calcular que las anteriores.

Para terremotos de los que no poseemos medidas instrumentales, pero sí información macrosísmica, las magnitudes pueden estimarse a partir de la intensidad epicentral I_0 , mediante relaciones empíricas. Las más frecuentes son del tipo:

$$(9) \quad M = \alpha I_0 + \beta$$

Donde α , β son constantes deducidas experimentalmente. En algunos casos también se tiene en cuenta la profundidad hipocentral, h , (si esta se conoce) y/o el área donde se ha sentido el terremoto.

⁹ Corregida del efecto instrumental.

Bibliografía

- [1] Grünthal, G. (Ed). European Macroseismic Scale 1998 (EMS-98). Cahiers du Centre Européen de Géodynamique et de Séismologie 15, 100 pp. 1998. <http://seismohazard.gfz-potsdam.de/projects/ems/index.html>
- [2] Kulháněk, O. The Structure and Interpretation of Seismograms. En Lee, W. HK., Kanamory H., Jennings P. C., Kisslinger C., (Eds). International Handbook of Earthquake and Engineering Seismology. Part A. Academic Press, 2002. pág 333-348.
- [3] Lay T. y Wallace, T. C. Modern Global Seismology. Academic Press, 1995. Pág. 200-234, 379-383.
- [4] Udías, A. Principles of Seismology. Cambridge University Press, 1999. Pág. 274-288.

Tema 16. Instrumentación sísmica. Teoría del sismómetro mecánico. Sismómetro electromagnético. Sismómetro de banda ancha. Acelerómetro. Funciones de respuesta y de transferencia. Determinación de amplitudes del suelo a través de sismogramas digitales.

16.1. Instrumentación Sísmica. Evolución Histórica

El instrumento chino, llamado “choko” hacia el siglo II para detectar la ocurrencia y también la procedencia de un terremoto, basado en la caída de las bolas de las bocas de la figura de un dragón de varias cabeza (Figura 1) que estaban en equilibrio inestable, parece ser el instrumento más antiguo conocido (Udías, 1999)

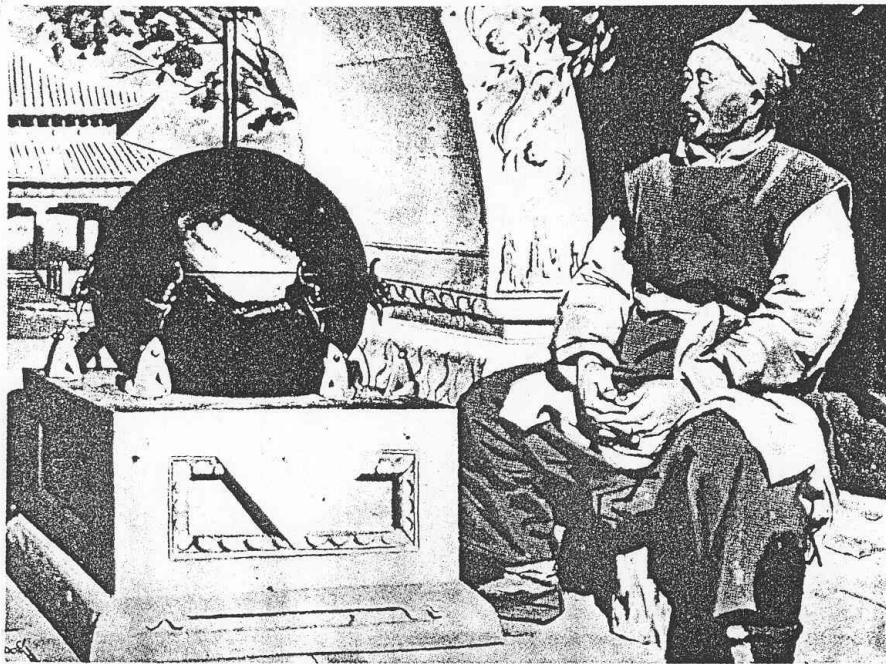


Figura 1. “Choko”

A mediados del siglo XIX, se diseñaron los primeros sismógrafos, basados en la oscilación de un péndulo ya sea vertical u horizontal, con un estilete sobre una placa de cristal ahumado. En realidad a estos aparatos que registran las señales independientemente del tiempo se les llama sismoscópios. (Figura 2).

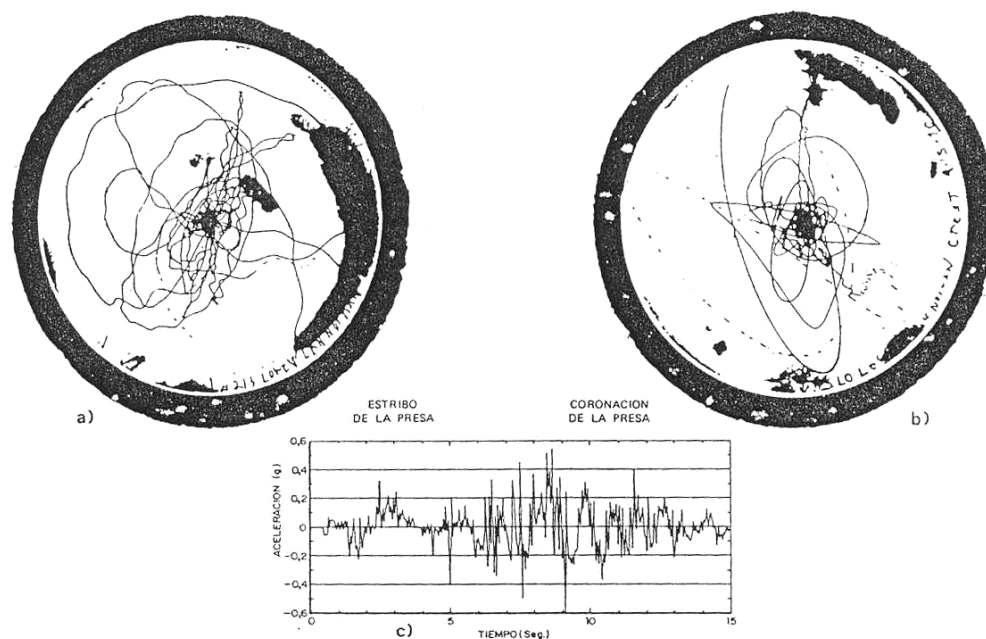


Figura 2. Sismoscopio. a) y b) Registros de sismoscopios obtenidos durante el terremoto de San Fernando, 9 de febrero de 1971 (Lower San Fernando Dam). c) Acelerograma aproximado, deducido de los registros anteriores (Scott, 1973).

A finales del siglo XIX, surgen los primeros sismógrafos de registro continuo sobre papel ahumado. Donde las amplificaciones del movimiento del suelo se consiguen mediante sistemas mecánicos y/o ópticos, y el amortiguamiento de la masa oscilante mediante elementos viscosos. Durante principios del siglo XX, se crean los sismógrafos con grandes masas (toneladas). Se pretende conseguir Sistemas Inerciales, por ejemplo los sismógrafos de Bosh-Omori, Wiechert, Mainka, etc.

En 1922, Wood y Anderson fabrican un sismómetro de reducidas dimensiones basado en la torsión de un hilo metálico. Se consiguen amplificaciones de más de 2000 veces el desplazamiento del suelo, y periodos de 0,8s. Este Sismómetro es fundamental para definir la escala de Magnitudes de Richter.

B.B. Galitzin en 1906, desarrolla su sismógrafo electromagnético, consistente en esencia en añadir a la masa oscilante un imán dentro de una bobina. La corriente que se genera en esta bobina pasa a través de un galvanómetro, cuya deflexión se registra en papel fotográfico mediante un haz de luz, y permite amplificaciones muy grandes, alrededor de 1000.

Sobre la base de esta idea surgen más tarde sismómetros con muy grandes amplificaciones y periodos más cortos, por ejemplo Benioff (reluctancia variable), Press y Ewing, etc. Utilizados en los observatorios sismológicos.

Los sensores modernos se basan en la utilización de sistemas retroalimentados y sensores de tipo capacitivo que ha permitido aumentar la sensibilidad, el rango dinámico¹ (razón entre la mayor, antes de saturarse, y la más pequeña señal que podemos medir con un instrumento) y el ancho de banda hasta límites inimaginables. Siendo además equipos muy portátiles.

En sismología la determinación exacta y precisa del tiempo es muy importante, hoy día con la evolución del sistema GPS, con una amplia cobertura en todo el mundo esto ha dejado de ser un problema.

Hasta ahora hemos hablado de los sismómetros, aparatos que hacen sus medidas correctamente en sitios muy alejados de donde ocurre el terremoto, se trata de medir la llegada de las ondas sísmicas que atraviesan cientos y miles de kilómetros, con ello se puede localizar los terremotos y estudiar el interior de la tierra.

En ingeniería sísmica se está más interesado en medir los movimientos fuertes del suelo, normalmente muy cerca del epicentro, que son los que ocasionan los daños, para ello se utilizan sensores llamados acelerómetros. Aunque como veremos el principio de funcionamiento es el mismo, es decir conseguir un *sistema inercial* de referencia.

16.2. Teoría del sismómetro mecánico

16.2.1. Descripción del sistema

Aunque este tipo de sismómetros se encuentran en desuso desde hace mucho, el principio físico en el que se basan todos los demás sismómetros posteriores es el mismo.

Vamos a considerar un sistema armónico amortiguado de un sólo grado de libertad, por ejemplo, un péndulo, o masa m suspendida de un bastidor por medio de un sistema elástico (por ejemplo un muelle) y con un sistema de amortiguamiento (como puede ser un baño de aceite), suponemos que el bastidor está solidariamente unido al suelo (Figura 3). Cuando ocurre un movimiento en el suelo, por la llegada de un terremoto, en principio la masa suspendida del muelle se pone en movimiento que en virtud de su inercia, es distinto al del bastidor, así pues la masa tiene un movimiento relativo respecto del bastidor (que a su vez está vibrando con el terremoto), o viceversa según donde coloquemos nuestro de referencia. Suponemos que la masa suspendida sólo puede vibrar en una dirección (en este caso la vertical). Como ya se ha dicho este movimiento, antiguamente se amplificaba por un sistema de palancas, y si de alguna

¹ El Rango dinámico, expresado en decibelios tiene la forma:

$$Rangodinámico(dB) = 10 \log \frac{P_{\max}}{P_{\min}} = 20 \log \frac{A_{\max}}{A_{\min}}$$

donde P es la potencia y A es la amplitud de la señal.

manera disponemos una escala graduada en el bastidor que nos permita medir estos desplazamientos relativos en función del tiempo, tendremos lo que llamamos un sismograma o acelerógrama.

16.2.2. Ecuaciones de movimiento

Sea $a(t)$; la aceleración que la llegada del terremoto produce sobre la masa, y sea $y(t)$; el desplazamiento relativo de la masa respecto del bastidor, es lo que medimos en la escala graduada unida al bastidor, sin tener en cuenta nada más, esto sería la salida del sismómetro (sismograma), téngase en cuenta que $y(t)$ no es el desplazamiento absoluto de la masa, ya que el bastidor se mueve con el terremoto.

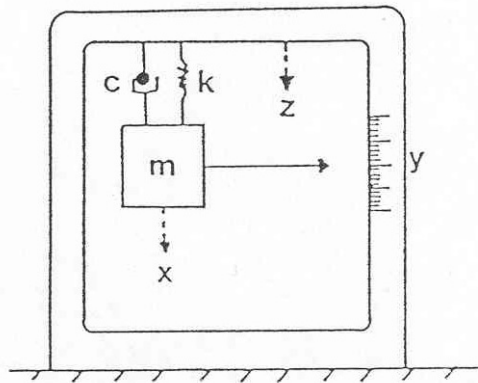


Figura 3. Esquema de un péndulo vertical con un sistema de amortiguamiento (Sistema armónico amortiguado con un sólo grado de libertad).

Las fuerzas que intervienen en el sistema y que actúan sobre la masa son:

➤ La *Fuerza elástica del Muelle*, se opone al movimiento, por la ley *de Hook*, directamente proporcional a lo que se estira o encoge el muelle, respecto a su posición de equilibrio, si consideramos movimientos relativamente pequeños. La fuerza viene dada por $F_e = -ky(t)$. k es la constante elástica de proporcionalidad del muelle.

➤ La *Fuerza de Amortiguamiento*, proporcional a la velocidad de la masa. Lo que hace es frenarla. La fuerza viene dada por $F_a = -c \dot{y}(t)$, donde c es la constante de amortiguamiento viscoso.

Hay que tener en cuenta que si llamamos $x(t)$ al desplazamiento absoluto de la masa (referido respecto a un sistema inercial) y, $z(t)$ al desplazamiento absoluto del suelo (o del bastidor), que

es lo que realmente queremos medir. Podemos escribir, que el desplazamiento de la masa respecto del bastidor:

$$y(t) = x(t) - z(t)$$

Si aplicamos la segunda ley de Newton, teniendo en cuenta las fuerzas que actúan sobre la masa:

$$\sum_i F_i = ma(t) \quad i = e, a.$$

Donde $a(t) = \ddot{x}(t) = \ddot{y}(t) + \ddot{z}(t)$

$$F_e + F_a = m \ddot{y}(t) - ky(t) - c \dot{y}(t) = m \left[\ddot{z}(t) + \ddot{y}(t) \right]$$

$$m \ddot{y}(t) + c \dot{y}(t) + ky(t) = -m \ddot{z}(t)$$

En esta última ecuación tenemos la aceleración que el terremoto produce sobre el sistema en función de los desplazamientos y sus respectivas derivadas temporales. Esta ecuación puede también escribirse como:

$$(1) \quad \ddot{y}(t) + 2h\omega_0 \dot{y}(t) + \omega_0^2 y(t) = -\ddot{z}(t)$$

Donde $\omega_0 = \sqrt{k/m}$, que es la frecuencia natural del sistema, es la frecuencia a la que vibraría la masa si no estuviese amortiguado, y $h = c/(2\sqrt{k/m})$, es el factor de amortiguamiento. $h = 1$, corresponde al amortiguamiento crítico. Para $h < 1$, la masa ante un impulso oscilará aunque sus amplitudes se amortiguen.

16.2.3. Soluciones de la Ecuación de movimiento. Funciones de Respuesta

La ecuación ec.1, puede resolverse haciendo uso de la Transformada de Fourier, a ambos miembros y aplicando sus propiedades sobre la derivada. Luego pasando al dominio del tiempo otra vez, con la transformada inversa, tenemos una solución en el dominio del tiempo en función de la aceleración del suelo:

$$y(t) = -\frac{1}{\omega_0 \sqrt{1-h^2}} \int \ddot{z}(\zeta) e^{-\omega_0 \sqrt{1-h^2} (t-\zeta)} \text{sen}[\omega_0 \sqrt{1-h^2} (t-\zeta)] d\zeta$$

Donde $\omega_a = \omega_0 \sqrt{1-h^2}$, es la frecuencia de vibración del sistema amortiguado

La ecuación de arriba es la convolución de la aceleración del suelo, $\ddot{z}(\zeta)$ con la respuesta del sistema a un impulso (oscilaciones armónicas cuya amplitud se va amortiguando).

Otra forma de estudiar las posibles soluciones de la ecuación de movimiento (1), es en el propio dominio de la frecuencia (una vez aplicada la Transformada de Fourier).

$$Y(\omega; \omega_0, h) = -\frac{\omega^2}{(\omega_0^2 - \omega^2) + 2ih\omega_0\omega} Z(\omega)$$

Donde ω , es la frecuencia angular ($\omega = 2\pi f$) a la que vibra el suelo. $Y(\omega; \omega_0, h)$ y $Z(\omega)$ son las respectivas Transformadas de Fourier de $y(t)$ y $z(t)$. El termino:

$$T(\omega; \omega_0, h) = -\frac{\omega^2}{(\omega_0^2 - \omega^2) + 2ih\omega_0\omega}$$

es la llamada Función de Transferencia del Sistema Armónico Amortiguado. Se puede demostrar que esta es la respuesta en frecuencias del sistema a una aceleración del suelo que fuese un impulso en el tiempo (Delta de Dirac). Conociendo esta función podemos averiguar la respuesta del sistema (movimiento de la masa) a cualquier tipo de aceleración del suelo (función de entrada), esto es una propiedad general de los sistemas lineales (Scherbaum, 2001). Esta función es una función compleja, que en forma polar podemos escribir como:

$$T(\omega; \omega_0, h) = M(\omega; \omega_0, h) e^{-i\phi(\omega; \omega_0, h)}$$

Donde:

a) $M(\omega; \omega_0, h)$; es el módulo de la función compleja, es la llamada *Sensibilidad en Desplazamientos* o *Magnificación Dinámica*.

$$(2) \quad M(\omega; \omega_0, h) = \frac{\omega^2}{\sqrt{[(\omega_0^2 - \omega^2)]^2 + (2h\omega_0\omega)^2}}$$

Representa el cociente entre la amplitud (módulos) del registro (sismograma) y la amplitud del desplazamiento del suelo en función de la frecuencia.

$$M = \frac{|Y|}{|Z|}$$

En esta última ecuación hemos prescindido en la notación de poner la dependencia con respecto a ω , ω_0 y h , pero es algo que debemos tener en cuenta siempre, ya que está implícita. De la ecuación (2), podemos ver, que para:

$$\omega \gg \omega_0 \Rightarrow M \rightarrow 1$$

$$\omega \rightarrow 0 \Rightarrow M \rightarrow 0$$

Cambiando el amortiguamiento podemos cambiar notablemente la forma de la curva de la Función de transferencia. Se puede demostrar que para un cierto valor amortiguamiento h ($\approx 0,707$), el sistema tiene una curva más o menos plana para frecuencias mayores que la frecuencia natural, podemos decir que el sismómetro mecánico tiene una “curva de respuesta plana en frecuencias para el desplazamiento” y, que en la frecuencia natural la amplitud del registro ha disminuido 3 dB (aprox. un 30 %) respecto de la amplitud de la zona plana. Con esto último, junto con los límites que acabamos de exponer arriba, podemos aproximar la respuesta de un sismómetro por un filtro paso alto de segundo orden, como podemos ver en la Figura 4.a.

La banda de frecuencias donde es útil trabajar con este sismógrafo es en la parte plana de la curva de amplificación, ya que podemos obtener aproximadamente la amplitud del suelo, a partir de la amplitud del registro multiplicado por una constante (no dependiente pues de la frecuencia).

$|Z| = M |Y|$ donde $M \approx \text{constante}$ para $\omega > \omega_0$.

b) $\phi(\omega; \omega_0, h)$; esta función representa el desfase entre el desplazamiento relativo de la masa sísmica y el desplazamiento del suelo.

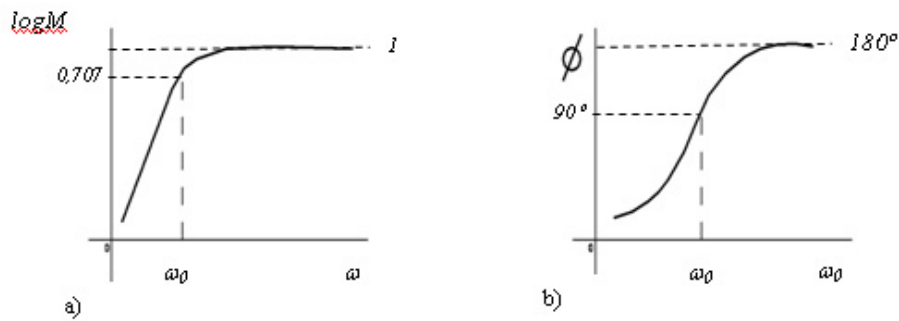


Figura 4. a) Representación aproximada de la curva de Sensibilidad en Desplazamiento en función de la frecuencia para un sistema armónico amortiguado, para un amortiguamiento $h \approx 0,707$. b) Representación aproximada de la curva de fase para el mismo sistema y el mismo amortiguamiento.

$$\phi(\omega; \omega_0, h) = \arctag \left[\frac{2h\omega_0\omega}{\omega_0^2 - \omega^2} \right] + \pi$$

La curva que se obtiene para el amortiguamiento utilizado en la construcción de sismómetros $h \approx 0,707$, la podemos ver la en la Figura 4.b. Dependiendo del rango de frecuencias de nuestra señal, hemos de tener en cuenta este desfase, para poder obtener correctamente el desplazamiento del suelo.

Físicamente cuando trabajamos en la zona plana del espectro el sismómetro se comporta como un sistema inercial, es decir para altas frecuencias la masa sísmica debido a su propia inercia permanece más o menos fija y nos permite medir los desplazamientos del bastidor (o del suelo). El desfase de π , viene del signo menos que aparece en la función de transferencia, ya que el sistema de referencia lo hemos establecido en el bastidor.

Recordar que hemos considerado desplazamientos muy pequeños que nos permita aproximar este sistema a un sistema lineal.

16.2.4. Sismómetros Horizontales

Para el registro de las componentes horizontales, Norte-Sur y Este-Oeste, si el sismómetro está orientado según los ejes geográficos, se utilizan péndulos, que pueden oscilar en un plano horizontal respecto a su posición de equilibrio, siempre considerando desplazamientos pequeños. Por ejemplo, pensemos en el más simple, en un péndulo horizontal con una pequeña inclinación, que puede oscilar como una puerta de jardín (sismómetros del tipo *gardengate*), en este caso lo que se mide es el desplazamiento angular de la masa.

16.3. Sismómetro Electromagnético

El principio físico es el mismo que para el sismómetro mecánico, es disponer de un sistema armónico amortiguado, que se comporte como un sistema inercial de forma que nos permita medir movimiento del suelo (ver Figura 5), pero ahora se disponen los componentes necesarios para que la señal de salida sea una diferencia de potencial eléctrico entre dos terminales.

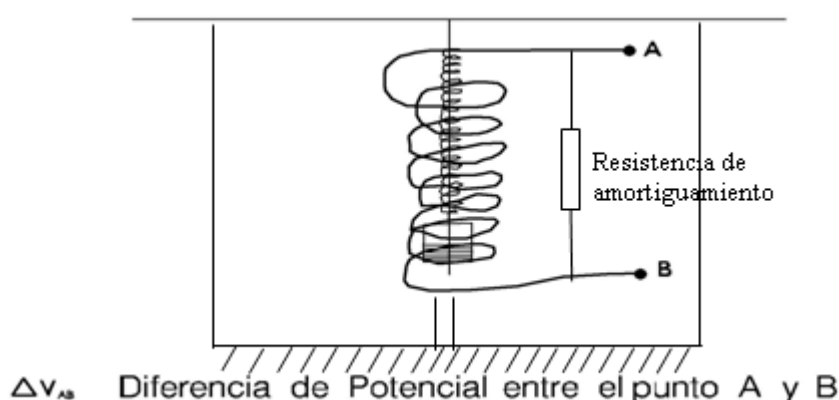


Figura 5. Representación muy esquemática de un sismómetro electromagnético, con imán que vibra dentro de una bobina.

16.3.1. Fundamentos del Transductor Electromagnético

Tradicionalmente casi todos los sismómetros han utilizado transductores electromagnéticos hasta la aparición de los sensores de banda ancha con transductores de tipo capacitivo.

Como ya se dijo al principio, los transductores electromagnéticos presentan grandes ventajas respecto a los sismómetros mecánicos, pues tienen un tamaño más reducido y además permiten

obtener el movimiento del suelo de forma más precisa, ya que la magnitud de salida es una corriente eléctrica, con las ventajas que ello supone para el tratamiento posterior de la señal.

El principio es el mismo, al sistema armónico anterior del sismómetro mecánico, se trata de unir a la masa sísmica un imán que vibra dentro de una bobina. La variación del flujo magnético, al vibrar el imán dentro de la bobina, produce en los terminales de la bobina una diferencia de potencial eléctrico $V(t)$, que es proporcional a la velocidad con que se mueve la masa vibrante respecto del bastidor, $\dot{y}(t)$ (esto que acabamos de decir se puede deducir a partir de la ley de Biot-Savart). Por eso estos sismómetros son transductores electromagnéticos, transforman la energía mecánica del movimiento del suelo en energía eléctrica. Ver Figura 5.

$$V(t) = G \dot{y}(t)$$

Hemos seguido la misma notación que en el apartado anterior.

Donde G (constante generadora), es la constante de proporcionalidad, que depende del tamaño de la bobina y de la densidad de flujo del imán.

También puede ser que la bobina sea solidaria con la masa vibrante y el campo magnético creado por el imán sea el fijo con el bastidor del sismómetro.

Una diferencia importante en estos sismómetros es que el amortiguamiento no se consigue con una sustancia viscosa, ahora es de tipo eléctrico, intercalando una resistencia entre los terminales de salida de la bobina, de tal manera que parte de la corriente eléctrica vuelve a la bobina, que produce una fuerza sobre la masa vibrante, en sentido opuesto a su movimiento, que tiende a frenarla (ley de Lenz, ver Havskov & Alguacil, 2005). Esta resistencia externa que se le suele llamar *resistencia de amortiguamiento*, se suma a la resistencia de la propia bobina. Esta resistencia de amortiguamiento se elige de forma que el amortiguamiento del sistema sea otra vez $h \approx 0,707$. La ecuación de movimiento es análoga a la del sistema mecánico, ecuación, ec.1. Hemos de tener presente que lo que registramos son ahora voltajes eléctricos $V(t)$ (que son proporcionales a la velocidad relativa de la masa vibrante). Para obtener una Curva de Respuesta con una parte plana en frecuencias (zona constante), análoga al sismómetro mecánico, la Función de Transferencia tiene que ser una relación entre la velocidad relativa de la masa sísmica (voltaje de salida) y la velocidad del suelo. Así podemos definir una *Sensibilidad en Velocidad*, S_E^V , del sensor electromagnético (o simplemente sensibilidad del sensor electromagnético):

$$(3) \quad S_E^V = \frac{|\dot{Y}|}{|\dot{Z}|} = \frac{|Y|}{|Z|} = M_e$$

En esta ecuación, en todas las funciones hay en general una dependencia con la frecuencia, como ya hemos dicho. Donde hemos utilizado la propiedad de la Transformada de Fourier, respecto de la derivada, a la que ya hemos hecho alusión antes:

$$\begin{aligned}\dot{Y} &= i\omega Y \\ \dot{Z} &= i\omega Z\end{aligned}$$

Con la ecuación (3) hemos demostrado que la sensibilidad en velocidad del sismómetro electromagnético tiene una forma a semejante a la del sismómetro mecánico, para la frecuencia. Hemos denotado por M_e , ya que las constantes del sistema que aparecen en M no son las mismas, ahora lo que aparece son constantes de tipo magnético y eléctrico.

Así pues, la sensibilidad del sismómetro electromagnético es:

$$S_E^V = \frac{\text{Voltaje de salida}}{\text{Velocidad del suelo}} \left[\text{Unidades en el Sistema Internacional de Unidades} \frac{V}{m/s} \right]$$

De lo anterior, podemos definir también una sensibilidad en aceleración, como la relación entre una determinada magnitud de salida (p.j. voltaje) y la aceleración del suelo. Es ésta la sensibilidad utilizada en los equipos llamados acelerógrafos, donde el principio físico sigue siendo igual que para el sismómetro mecánico y electromagnético, pero las constantes del sistema están elegidas de forma, que sea más difícil saturarse ante movimientos fuertes del suelo, como los que se dan cerca del epicentro del terremoto, pero que en este caso la curva de respuesta presenta una parte aproximadamente plana para la aceleración, para $\omega < \omega_0$ (Havskov & Alguacil, 2005).

16.3.2. Calibración de los Sismómetros

Normalmente todos los sismómetros disponen de una *bobina de calibración*, que realiza el proceso contrario. Se inyecta una señal eléctrica (intensidad de corriente) en esta bobina de calibración, que a su vez mueve la parte móvil del sensor (p.j. el imán), éste a su vez produce una señal (voltaje) en la bobina sensora, que utilizando los factores conversión necesarios, nos

permite volver a obtener la señal de entrada y comparar. Y esto variando la frecuencia de la señal de entrada, podemos averiguar de una forma relativa la respuesta del sensor.

16.4. Sismómetros de Banda Ancha

También llamados en la literatura inglesa como sensores “Broad Band”. Este tipo especial de sismómetros, que hoy día son muy comunes, tienen su filosofía de funcionamiento basada en los mismos principios que ya se han explicado. Los sismómetros electromagnéticos tienen un rango dinámico, sensibilidad y una banda de frecuencias muy reducida comparadas con las de este tipo de sismómetros. La mejora se debe principalmente al Transductor Capacitivo y al Principio de *Feedback*, que ahora vamos a explicar de forma somera.

16.4.1. Principio de Feedback

Sistema de “Feedback”, que podríamos traducir como Sistema Retroalimentado, o también conocido como “Force Balance” (equilibrio de fuerzas). El sistema de realimentación origina una fuerza negativa que obliga a estar en reposo la masa “vibrante”. Lo que realmente medimos es la intensidad de corriente que hay que suministrar para crear esta fuerza negativa para compensar (o equilibrar) la fuerza que el terremoto ejerce sobre la masa sísmica. Tenemos en este caso dos Bobinas la Sensora y la de Realimentación. Ya que la masa permanece casi “estacionaria”, se consigue una mejora muy importante en la *linealidad* del sistema, aunque el sismógrafo este sometido a aceleraciones más o menos fuertes.

16.4.2. Transductor Capacitivo

En un sensor de tipo capacitivo, se trata de medir la variación de capacidad que hay entre dos placas de un sistema condensador, cuando un determinado medio, entre las placas se desplaza solidario a la *masa sísmica* (Figura 6). La *variación de capacidad* es proporcional al *Desplazamiento*. Por ello este tipo de sensor es un Transductor de Desplazamiento. Pensemos en un sencillo condensador plano cuya capacidad viene dada por:

$$C = \epsilon \frac{S}{d}$$

Donde

d ; Distancia de separación entre las placas del Condensador.

S ; Superficie de las placas o armaduras.

ϵ ; Es la Constante Dieléctrica Absoluta.

En realidad, lo que se quiere medir es la *diferencia de potencial* (ΔV), que se genera entre las placas del condensador, que es inversamente proporcional a la capacidad del condensador, ésta a su vez, es inversamente proporcional a la distancia entre placas, como se acaba de ver, a si pues, un cambio en la distancia (*desplazamiento*) entre las placas, producirá **proporcionalmente** una variación de voltaje en el condensador, como habíamos dicho.

$$\Delta V = \frac{Q}{C}$$

Luego de las dos últimas ecuaciones:

$$\Delta V \propto d$$

Como se quería demostrar.

En realidad en las placas del condensador se generan dos señales de una determinada frecuencia, cuando se produce un movimiento del medio que está entre las placas, se produce una desviación de la frecuencia, que es lo que realmente se mide (Figura 7), de aquí la precisión de este transductor en medir desplazamientos.

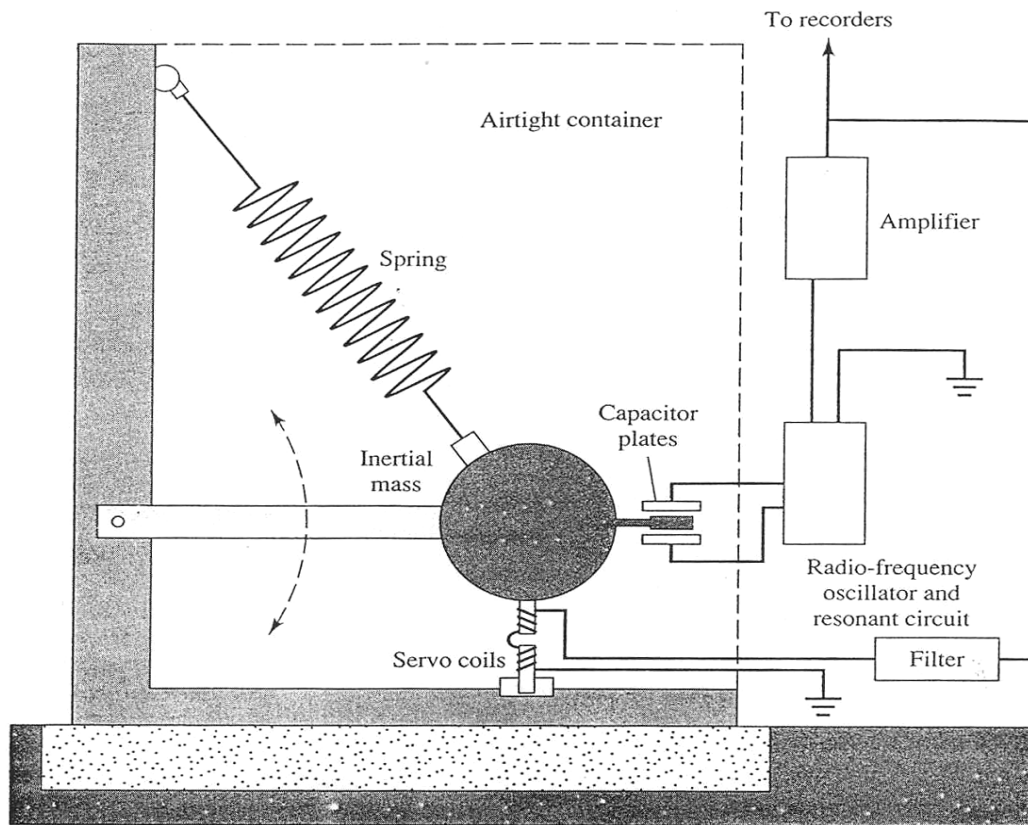


Figura 6. Esquema General de Funcionamiento de un Sensor "Broad Band". Donde solidaria a la masa sísmica, hay un Transductor Capacitivo, y un sistema de *Feedback* (de Bolt, 1993)

También se disponen de un sistema de *feedback*, para aprovechar las ventajas antes mencionadas (Figura 6). Para que la salida sea proporcional a la *Velocidad*, con una curva de respuesta como la que ya hemos visto para los sismómetros electromagnéticos, se utilizan sistemas de ecualizadores y de lazos realimentación múltiple de la señal, especialmente diseñados (ver Havskov & Alguacil, 2002).

Sin embargo, para este tipo de sensores se ha de vigilar muy de cerca los **cambios de temperatura**, por su influencia en las constantes elásticas, y de **presión atmosférica**, que afectan a la fuerza de empuje que el aire ejerce sobre las masas y por los cambios que en la inclinación de la superficie terrestre produce, por su influencia en los largos periodos.

Este tipo de sensores de banda ancha poseen frecuencias naturales del orden de los 0,01 Hz, como los sensores de la Nueva Red Sísmica Digital del I.G.N. (*CMG-3T* de *Guralp*), frente a los sensores electromagnéticos de la antigua Red Analógica de Corto Periodo, de frecuencia natural 1 Hz. Aunque hay sensores de banda ancha que pueden llegar a los 360s de periodo natural.

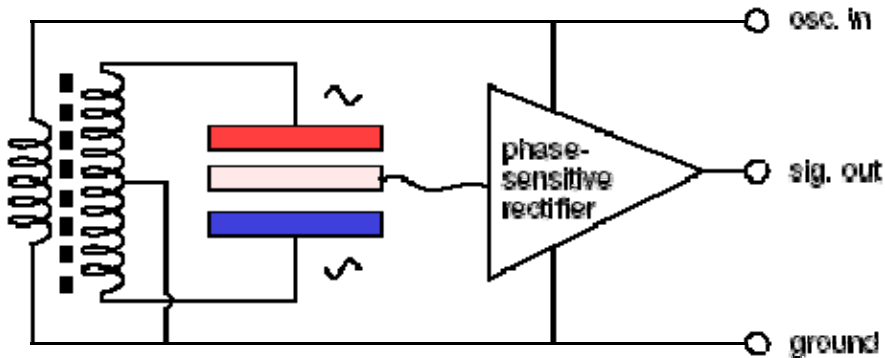


Figura 7. Transductor Capacitivo en Desplazamiento. Se mide la desviación respecto a una frecuencia de referencia cuando se mueve la armadura de en medio, que es solidaria a la *masa sísmica* (Aki & Richards, 1980)

16.5. Determinaciones de amplitudes del suelo a partir de sismogramas digitales

Los modernos sismógrafos, es decir la suma del sismómetro (sensor) y el registrador, poseen conversores analógico-digital de alta resolución (24 bit) y sistemas de amplificación electrónica de la señal, filtros activos y pasivos y procesadores digitales de la señal (DSP) que tratan la señal de forma matemática, previamente digitalizada.

Un parámetro muy importante del conversor A/D digital es su resolución, que nos dice la cantidad más pequeña que podemos medir entre dos muestras digitales (a una determinada frecuencia de muestreo). Este valor depende del número de bits del conversor y de su nivel de ruido. Se expresa en cuentas/voltio o bits/voltios (R), nos dice la cantidad de cuentas digitales que tenemos de salida para un determinado voltaje de entrada.

Para todas las demás etapas por la que pasa la señal, hemos de conocer su respuesta en función de la frecuencia, a través de sus Funciones de Transferencia, como ocurría para el propio sismómetro.

Normalmente éstas son suministradas por las casas constructoras, o se pueden y deben averiguar mediante calibraciones periódicas.

Así pues, a partir de nuestro sismograma digital, $s(t_i)$, es decir, serie de valores discretos, a intervalos constantes de tiempo, se puede obtener, la historia temporal del movimiento del

suelo, $\dot{z}(t_i)$, con una ecuación que engloba todas las etapas, como el producto de las funciones de transferencia de cada una de ellas en el dominio de la frecuencia, como:

$$(4) \quad S(\omega_i) = RF(\omega_i)T(\omega_i)\dot{Z}(\omega_i)$$

Donde.

$S(\omega_i)$; es el espectro discreto de Fourier del sismograma (en cuentas)

R ; resolución del conversor A/D, en cuentas/voltios

$F(\omega_i)$; Funciones de transferencia de todas las etapas del sismógrafo, que no es el sensor.

$T(\omega_i)$; Función de transferencia del sismómetro, en la que está incluida su sensibilidad expresada p.j. en voltios/ms⁻¹.

$\dot{Z}(\omega_i)$; es el espectro discreto de Fourier de la velocidad del suelo, siguiendo las mismas unidades de velocidad, en m/s.

Hay que tener en cuenta que el espectro de Fourier y las funciones de transferencia en el dominio de la frecuencia son funciones complejas, que tienen un módulo y una fase. Están calculadas utilizando algoritmos como la *FFT* ("Fast Fourier Transform"), a unas frecuencias determinadas, a intervalos de frecuencia regulares también, que depende de la frecuencia a que se ha muestreado la señal y de la longitud de la ventana de tiempo elegido, Es corriente expresar estas funciones como un cociente de dos polinomios (Scherbaum, 2001), a partir de sus polos (raíces del denominador, p_k) y de sus ceros (raíces del numerador, c_k). Así la ecuación (4), podemos reescribirla:

$$S(\omega_i) = R \frac{\prod_{k=1}^n (\omega_i - c_k)}{\prod_{k=1}^m (\omega_i - p_k)} \dot{Z}(\omega_i)$$

Ecuación que tiene n ceros y m polos.

Luego de la ecuación (4), podemos obtener la velocidad del suelo, a partir de un sismograma despejando el segundo miembro y aislando el término correspondiente a movimiento del suelo. Esto es lo que se llama *corrección instrumental*.

Los modernos sismógrafos digitales poseen rangos dinámicos muy grandes, de hasta 140 dB, con un solo sensor, lo cual significa que pueden registrar amplitudes en un rango de 1 a 10⁷. Esto nos permitiría registrar terremotos de forma local a distancias de aproximadamente 10

Km., con magnitudes de hasta 5 y, telesismos de Magnitud 9 a unos 30° de distancia (Udías, 1999), sin saturarse.

Bibliografía

- [1] Havskov, J.; Alguacil, G. Instrumentation in earthquake seismology. [S.l.]: Springer, 2005. 370. Varios temas.
- [2] Scherbaum, F. Of Poles and Zeros. Fundamentals of Digital Seismology. Kluwer Academic Publishers. Varios temas.
- [3] Udías, A. Principles of Seismology, 1999. Tema 21.

Figuras:

- [4] Aki, K. y Richards, P. G. Quantitative Seismology. *Freeman and Company*. New York. U. S. A., 1980. Tema 10.
- [5] Bolt, B. Earthquakes. *Freeman and Company*. New York. U. S. A., 1980. Pag. 43-50.

Tema 17. Movimientos sísmicos fuertes. Acelerogramas. Características de un acelerograma en el tiempo y en la frecuencia. Estimación empírica de la aceleración máxima en un punto. Espectro de respuesta y de diseño.

17.1 Movimientos sísmicos fuertes

Cuando las ondas producidas por terremoto se propagan por la Tierra, la inmensa mayoría de las veces no son sentidas por la población e incluso tampoco son detectadas por ningún instrumento. Sin embargo, cuando en un terremoto ya sea por su tamaño o/y por la distancia a la que es observado, las ondas que genera son percibidas por la población o producen efectos visibles decimos que estamos ante un movimiento sísmico fuerte (MSF). Este tipo de movimientos son los que tienen relevancia en Ingeniería Sísmica por su posibilidad de producir daños en las estructuras que pueda encontrar en su camino. En este sentido todos los terremotos de la época pre-instrumental que han llegado a nuestro conocimiento por informaciones de cómo fue sentido o por los efectos producidos en las estructuras son en su totalidad MSF.

Los MSF de la época histórica, al no tener instrumentos de registro han quedado descritos por sus efectos, esto es por la intensidad sísmica generalmente en forma de mapas de isosistas mientras que en la actualidad pueden además ser registrados en un instrumento. En este caso la descripción viene dada por una historia temporal de la variable que registre el instrumento que es generalmente la aceleración del suelo al paso del MSF. A este registro de un MSF se le denomina acelerograma. Por tanto la caracterización de un MSF es por la intensidad sísmica asociada a él en punto de mayor valor (Intensidad Máxima) y en algunos casos por la distribución de los valores de intensidad alrededor del epicentro donde se ha sentido y ha podido causar daños. Modernamente también se pueden caracterizar por el registro o registros en aceleración que ha podido dejar en puntos próximos al epicentro. En la actualidad muchos terremotos que son MSF tienen ambas caracterizaciones.

17.2 Acelerogramas

Son los registros de MSF obtenidos mediante instrumentos llamados acelerógrafos. Estos instrumentos tenían hasta hace pocos años registro analógico en papel fotográfico aunque modernamente todos los registros son ya digitales.

La aceleración más alta que un acelerógrafo analógico puede registrar es de 1 g ($g=980 \text{ cm/s}^2$) y el más pequeño es del orden de 0.005 g. La relación entre la aceleración más grande a la más pequeña que pueden registrar (es decir el rango dinámico) es de 200 por lo que si expresamos en escala logarítmica de decibelios tendremos que estos instrumentos analógicos tienen un rango dinámico de 50 dB. Sin embargo por experiencia vemos que los valores que se pueden dar en un terremoto dependiendo del tamaño y de la distancia a la que se observe puede llegar a ser del orden de 2g. En la actualidad los acelerógrafos son digitales y con un conversor analógico digital que puede llegar a tener de 18 a 24 bits de resolución. Para un digitalizador de 18 bits la aceleración más pequeña que puede distinguirse en un registro de estos instrumentos con un sensor con máximo registro de 2g es $\pm 2g/2^{18}=0.015 \text{ mg}$ es decir que correspondería a un rango dinámico de 108 dB. Con este tipo de instrumentos está pues asegurado la resolución pues el ruido ambiente en zonas urbanas es superior a 0.015 mg, por lo que estos instrumentos están preparados para el registro de señales por encima de este valor umbral.

El acelerograma registrado es en la actualidad en la mayoría de las redes instaladas de registro digital y por tanto este registro ha de pasar por un proceso de tratamiento para eliminar aquellas influencias perturbadoras que existen en el registro y poder finalmente obtener los parámetros necesarios para su utilización en ingeniería. Este proceso se resume en los siguientes pasos:

En el paso 1 se realiza la llamada corrección de línea base. En este proceso se trata de realizar una operación que trata de obtener un nuevo registro que tenga la propiedad de que la media total de todos sus valores sea 0. Este proceso se realiza mediante la eliminación de la línea promedio de todo el acelerograma. En el paso 2 se pretende realizar la corrección instrumental correspondiente que puede influenciar el MSF. Las correcciones de este paso se logran mediante la modelización de la respuesta del instrumento como un oscilador armónico simple de un grado de libertad al cual se puede asociar el sensor o acelerómetro. En la mayoría de los instrumentos actuales esta respuesta es plana hasta frecuencias próximas a los 10 Hz, por lo que únicamente por encima de esta frecuencia la corrección tiene sentido. Sin embargo dado que algunos acelerógrafos están colocados sobre ciertas estructuras (edificios, puentes, etc) la respuesta de la estructura también se combina con la del MSF en el rango de interés. Por supuesto esta

influencia de la estructura es necesario eliminarla. A continuación como paso 3 se efectúa la corrección de alta frecuencia necesaria para eliminar la influencia del ruido presente en la señal de alta frecuencia. Esta corrección se realiza mediante la aplicación de un filtro de Ormsby cuya curva de respuesta tiene una frecuencia de esquina de 23 Hz y una terminación a 25 Hz. A partir de este proceso y dependiendo del tipo de acelerograma, donde se haya registrado y cual va a ser el propósito para el que se va a utilizar el registro correspondiente se efectúan otros procesos como aplicación de un filtro de periodos altos (por encima de 15 s), etc.

17. 3. Características de un acelerograma en el tiempo y en la frecuencia

Un acelerograma es un registro del movimiento del suelo al paso de las ondas muy complicado pues es el resultado de muchos parámetros que van desde los correspondientes al propio fenómeno de la ruptura, pasando por lo que puede venir afectado en su viaje desde el foco hasta el instrumento y por supuesto por la influencia del instrumento. Por esta razón es muy difícil explicar el registro del MSF en su totalidad y por eso se pretende buscar aquellas características que puedan definir dicho registro y que a la vez sea relativamente fácil su conocimiento. El estudio de estas características singulares de un acelerograma se puede realizar en el dominio del tiempo o en el dominio de la frecuencia.

En el dominio del tiempo el parámetro más utilizado es la *Aceleración Máxima Horizontal* (AMH) que se corresponde con el mayor valor que se encuentra de cualquiera de las dos componentes horizontales de un acelerograma, expresado en las unidades del registro. Este parámetro es fácil de medir y además está íntimamente relacionado con las fuerzas inerciales. Este parámetro se ha relacionado con el valor de la intensidad sísmica, con el objeto de poder convertir los MSF históricos caracterizados por esta cualitativa medida de la intensidad en un valor que estuviese relacionado con el movimiento del suelo. La correlación entre estos dos parámetros es muy débil debido por una parte a la intrínseca naturaleza cualitativa de la intensidad y también con la escasa relación entre esta y el valor de la AMH. Como ejemplo de correlaciones entre estos parámetros se pueden ver en la Figura.

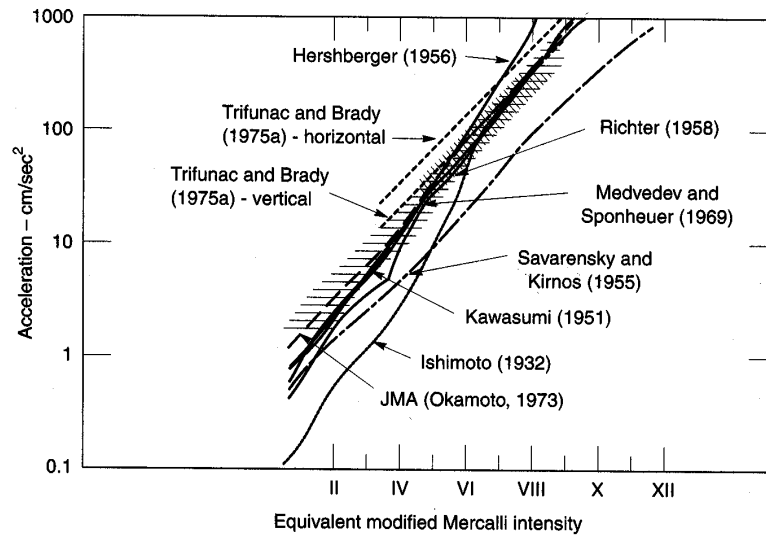


Figura 1. Algunas correlaciones Intensidad sísmica – aceleración, Kramer, 1997.

Si en el registro de aceleración realizamos una integración obtenemos el registro en velocidad del movimiento del suelo y mediante una segunda integración el registro de desplazamientos. Pues si a partir de estos registros obtenemos los valores máximos horizontales o vertical podremos obtener la *Velocidad Máxima Horizontal/Vertical* (VMH, VMV) y el *Desplazamiento Máximo Horizontal/Vertical* (DMH, DMV).

Otro parámetro que se determina en el dominio del tiempo es la *Duración*, que puede tener una gran importancia a la hora de afectar un sismo a una estructura. Existen muchas estimaciones de este parámetro.

Duración acotada. Se define como el tiempo entre la primera vez que en un acelerograma se sobrepasa un valor prefijado de aceleración y la última vez que se produce.

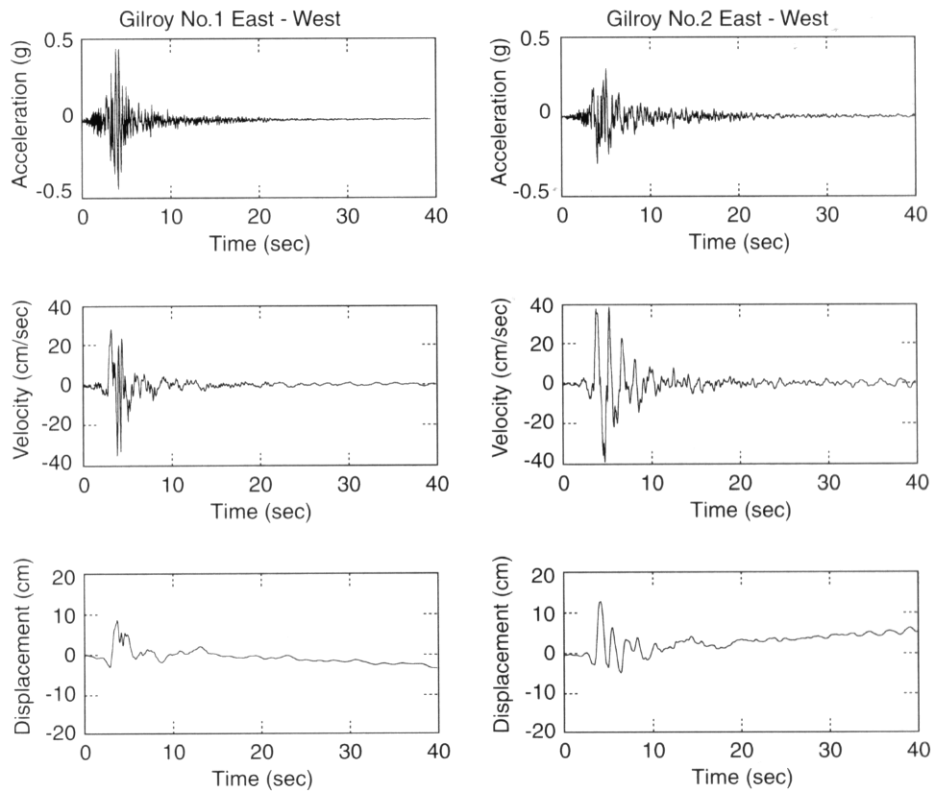


Figura 2. Acelerogramas de dos estaciones en roca (Gilroy 1) y en suelo (Gilroy 2) para la componente E-W y sus respectivas integraciones para obtener Velocidad y Desplazamiento. Kramer, 1996.

Si por el contrario el acelerograma es convertido al dominio de la frecuencia tendremos otra serie de parámetros que también pueden caracterizar al registro de aceleración. El primer parámetro que podemos considerar es el *Espectro de Fourier de la Aceleración*.

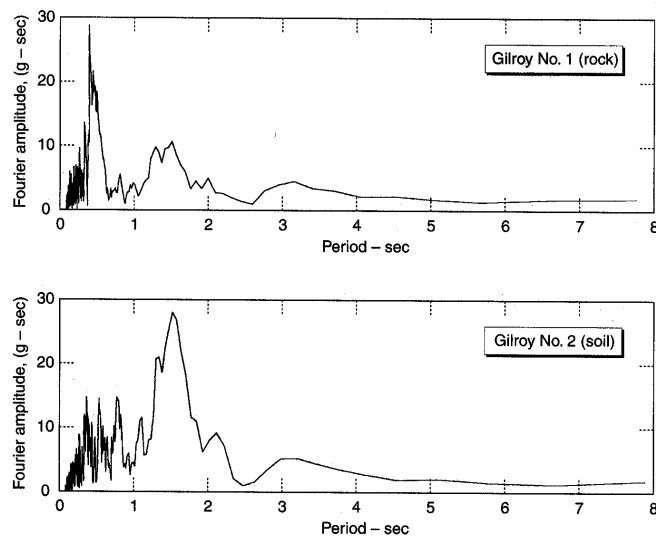


Figura 3. Espectro de Fourier de las componentes E-W de los acelerogramas en la estación Gilroy. Kramer, 1996.

El espectro expresa de qué forma la amplitud del movimiento se distribuye en relación a la frecuencia. La forma de estos espectros es diferente por tanto si el medio en el que se registra tiene una frecuencia propia de oscilación. Así en la Figura se puede observar como para un mismo terremoto registrado en dos tipos distintos de suelo y roca muestran un espectro de Fourier del MSF con sus máximos no coincidentes.

Entre los muchos parámetros que existen definidos en el dominio de la frecuencia podemos citar uno que es ampliamente utilizado por su fuerte correlación con los daños ocasionados por un sismo como es la *Intensidad de Arias* y que se define de la forma:

$$I_a = \frac{\pi}{2g} \int_0^{\infty} [a(t)^2] dt$$

Existen otros parámetros determinados en el dominio de la frecuencia como pueden ser los siguientes:

Periodo predominante. Se define como el periodo de vibración correspondiente al valor máximo del espectro de Fourier del acelerograma en un punto.

Ancho de Banda. Es el rango de frecuencias en la que el nivel promedio de la amplitud de Fourier supera un cierto valor establecido previamente, por ejemplo $1/\sqrt{2}$ el valor máximo de la amplitud de Fourier.

Frecuencia central. Es la medida de la frecuencia en la que la densidad espectral del acelerograma se concentra. Una medida de la dispersión de la función densidad espectral de potencia con relación a la frecuencia central es el llamado *Factor de Forma*.

17.4. Estimación empírica de la aceleración máxima en un punto

En el cálculo de cualquier obra de ingeniería en el que el fenómeno sísmico tenga importancia, su consideración ha de realizarse mediante algún parámetro que represente de alguna forma al sismo de que se trate. Este efecto del sismo puede ser en primer lugar el conocimiento de la historia temporal completa, es decir el conocimiento del acelerograma que tendría que haberse registrado en el lugar en el que estamos interesados de haberse producido ese sismo. Sin embargo esa estimación de un acelerograma completo es un fenómeno de difícil solución y además en la mayoría de los casos la información que proporcionaría sería muy complicada para su aplicación en el problema concreto. Por estas razones, como una primera aproximación deberíamos ser capaces de estimar algún parámetro representativo del sismo que sea más simple de calcular y además tenga un fuerte sentido para el problema de ingeniería que se pretende. Este parámetro es la aceleración sísmica máxima ya sea de entre las componentes horizontales (AMH) o de la vertical (AMV). Es decir estamos interesados en conocer cual es la

AMH de un futuro terremoto con el que se verá afectado el emplazamiento objeto de estudio. La forma de abordar este problema de una forma empírica es analizar, para casos reales ya producidos como este parámetro se comporta con la distancia y el tamaño del sismo y si es posible introduciendo las variaciones que pueden afectar a este comportamiento otros parámetros adicionales que sirven para definir el sismo como son el mecanismo focal, el tipo de terreno en el que se va a registrar, etc. Este tipo de relaciones recibe el nombre también de relaciones de atenuación. Así, la forma más general de expresar la AMH en un punto puede escribirse:

$$AMH = c_1 e^{c_2 M} R^{-c_3} e^{-c_4 r} e^{c_5 F} e^{c_6 S} e^{\varepsilon}$$

Siendo M la magnitud del terremoto, r la distancia entre la fuente y el emplazamiento y R otro parámetro dependiente de esta distancia y de la magnitud, F el parámetro que caracteriza el tipo de mecanismo, S indica el tipo de terreno en el que se registra el movimiento y ε es el término de error aleatorio con media cero y con la desviación estándar del ajuste. De forma general se suele expresar en forma logarítmica así:

$$\ln(AMH) = c_1 + c_2 M - c_3 \ln R - c_4 r + c_5 F + c_6 S + \varepsilon$$

$$\text{Siendo } R = \begin{cases} r + c_7 \exp(c_8 M) \\ \sqrt{r^2 + [c_7 + \exp(c_8 M)]^2} \end{cases}$$

El término del ajuste, representa una estimación de la desviación estándar del $\ln(AMH)$ para la magnitud y distancia de interés. Tradicionalmente este valor ha sido una constante aunque las relaciones empíricas actuales la expresan variando con la magnitud del sismo.

Estadísticamente el valor de $\sigma_{\ln AMH}$ representa una estimación de la desviación estándar del $\ln(AMH)$ para una cierta magnitud y distancia de interés. Por ejemplo, para una magnitud y distancia dadas la probabilidad de que el parámetro del movimiento del suelo Y^* (por ejemplo un cierto valor de la AMH) sea superado es precisamente $1-F(z^*)$ siendo $F(z^*)$ la función de distribución acumulada que viene dada por $z^* = (\ln Y^* - \text{media}(\ln Y)) / \sigma_{\ln Y}$.

Estas relaciones tienen un fundamento sismológico basado en el conocimiento de las leyes físicas de la propagación de ondas así como de la experimentación obtenida a partir de los datos registrados con anterioridad. Es decir la relación de la amplitud AMH con la distancia obedece por una parte a la de la amortiguación de la energía por ensanchamiento del frente de onda, que se denomina atenuación geométrica. También existe otra dependencia con relación a la distancia

r que es la correspondiente a la atenuación inelástica de la energía por efecto del alejamiento del medio de la consideración de perfectamente elástico.

Una de las variables que intervienen en el cálculo es la distancia y para este parámetro existen diferentes definiciones basadas en qué parte de la falla se considera como origen de distancias al punto de observación.

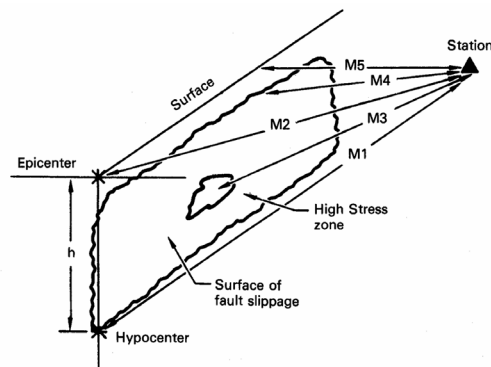


Figura 4. Diferentes definiciones de distancia utilizadas en la estimación del MSF. Reiter, 1990.

De esta forma somos capaces si para una determinada zona tectónica conocemos valores de AMH de terremotos ocurridos en la zona y disponemos del conocimiento de todos los parámetros como son la magnitud, distancia, etc. seremos capaces de deducir una expresión empírica de la forma de la descrita anteriormente que permita establecer en el futuro lo contrario; es decir que supuesto un sismo en un emplazamiento determinado de una cierta magnitud y mecanismo obtener la AMH que se registraría en un determinado emplazamiento de interés. En la Figura 5 se muestra ejemplos de curvas empíricas para diferentes zonas y ambientes tectónicos y para distintas magnitudes.

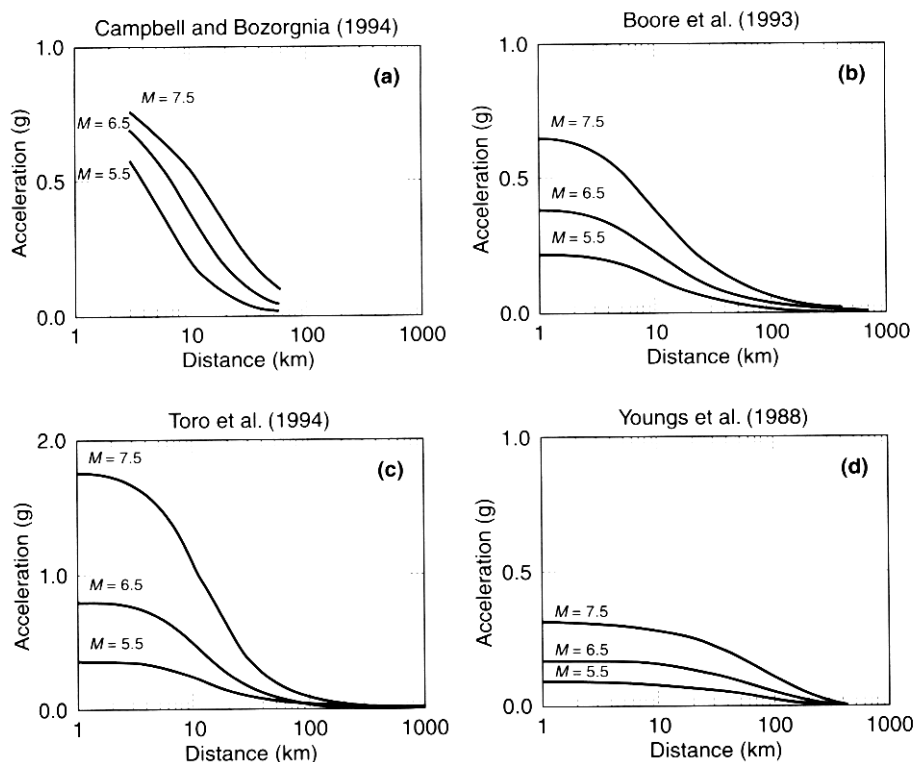


Figura 5. Correlaciones empíricas según diversos autores para diferentes situaciones tectónicas. Kramer, 1996.

17.5. Espectro de respuesta y de diseño

El espectro de respuesta es una herramienta de gran utilidad en ingeniería sísmica porque proporciona directamente la máxima respuesta, de una estructura ideal asimilable a un oscilador armónico simple de un grado de libertad ante un MSF. Además, describe el contenido en frecuencia de un MSF en una forma más utilizable en ingeniería que el espectro de Fourier de dicho movimiento, ya que se tiene en cuenta el efecto del amortiguamiento. El espectro de respuesta puede considerarse el elemento fundamental para la realización del espectro de diseño que se utiliza en todas las normativas sismorresistentes y que veremos más adelante.

En la Figura 6 se presenta un esquema de un oscilador armónico simple antes y después de incidir en su base un MSF caracterizado por una aceleración u_g

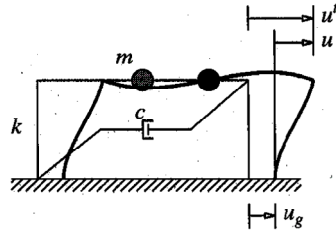


Figura 6. Respuesta de un Oscilador Armónico Simple ante una aceleración $\ddot{u}_g$

La ecuación diferencial del oscilador ante la acción del MSF es:

$$\ddot{u} + 2\xi\omega_n\dot{u} + \omega_n^2 u = -\ddot{u}_g(t)$$

Con $\omega_n = \sqrt{k/m}$ $\xi = c/2\sqrt{km}$

Si mantenemos constante el amortiguamiento y variamos la frecuencia del oscilador obtendremos para cada frecuencia una respuesta del oscilador y tomando el máximo para cada una pasaremos a representar estos máximos en función de la frecuencia en un gráfico.

Espectro de Respuesta en Desplazamiento $u_0(T_n, \xi) = \text{Max}|u(t, T_n, \xi)|$

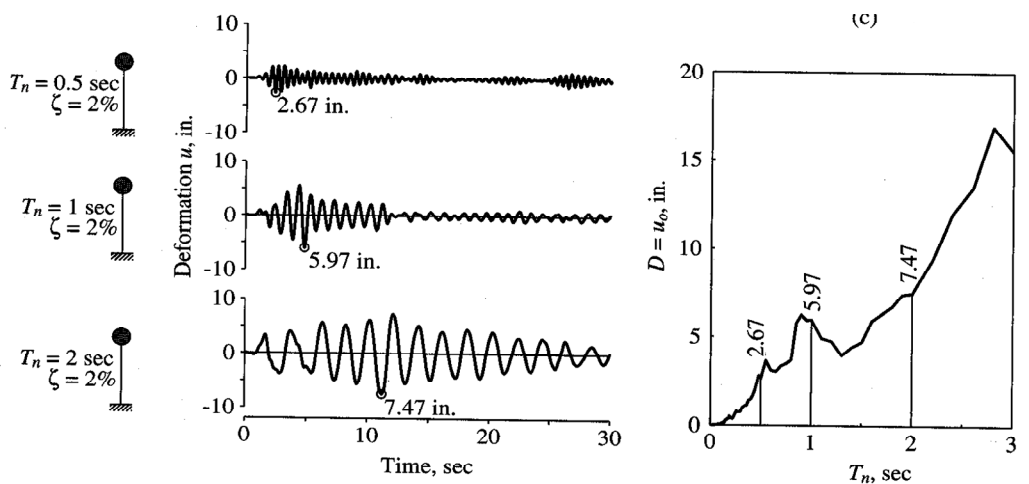


Figura 7. Respuesta de osciladores de diferentes frecuencias y la representación del Espectro de respuesta en desplazamiento, Kramer 1996.

Si calculamos la velocidad $\dot{u}(t)$ o $\ddot{u}(t)$ la aceleración y consideramos para cada frecuencia el máximo, tendremos para cada frecuencia o periodo

Espectro de Respuesta en velocidad $\dot{u}_0(T_n, \xi) = \text{Max}|\dot{u}(t, T_n, \xi)|$

Espectro de Respuesta en aceleración $u_0(T_n, \xi) = \text{Max}|u(t, T_n, \xi)|$

Podemos efectuar la siguiente simplificación: Si el registro obtenido de desplazamientos considerásemos era una función sinusoidal, lo que para amortiguamientos inferiores al 10 % del valor crítico es muy aproximado podríamos a partir de este registro obtener otro resultante de multiplicar por la frecuencia angular del oscilador. Es decir se obtiene otro espectro que al tener dimensiones de velocidad y no corresponder con el de velocidad podremos llamar de Pseudovelocidad:

Espectro de Respuesta en Pseudovelocidad $V = \omega_n u_0 = \frac{2\pi}{T_n} u_0$

Que conviene no confundir con el espectro de respuesta en velocidad $\dot{u}_0(T_n, \xi)$

Espectro de Respuesta en Pseudoaceleración $A = \omega_n^2 u_0 = \left(\frac{2\pi}{T_n}\right)^2 u_0$

Esta simplificación además permite representar de una forma simplificada los tres espectros como puede verse en la Figura 8

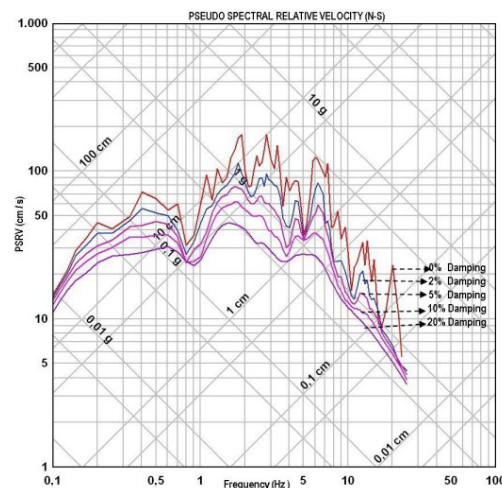


Figura 8. Representación tripartita de los tres espectros para varios amortiguamientos

Para ello hemos considerado el logaritmo de la pseudovelocidad en función del logaritmo de la frecuencia del oscilador en su representación en este gráfico que llamamos tripartito porque ese

mismo punto también se puede leer en los ejes de desplazamiento y pseudoaceleración. La razón es que el eje de pseudoaceleración tiene una pendiente +1 y el de desplazamiento de -1.

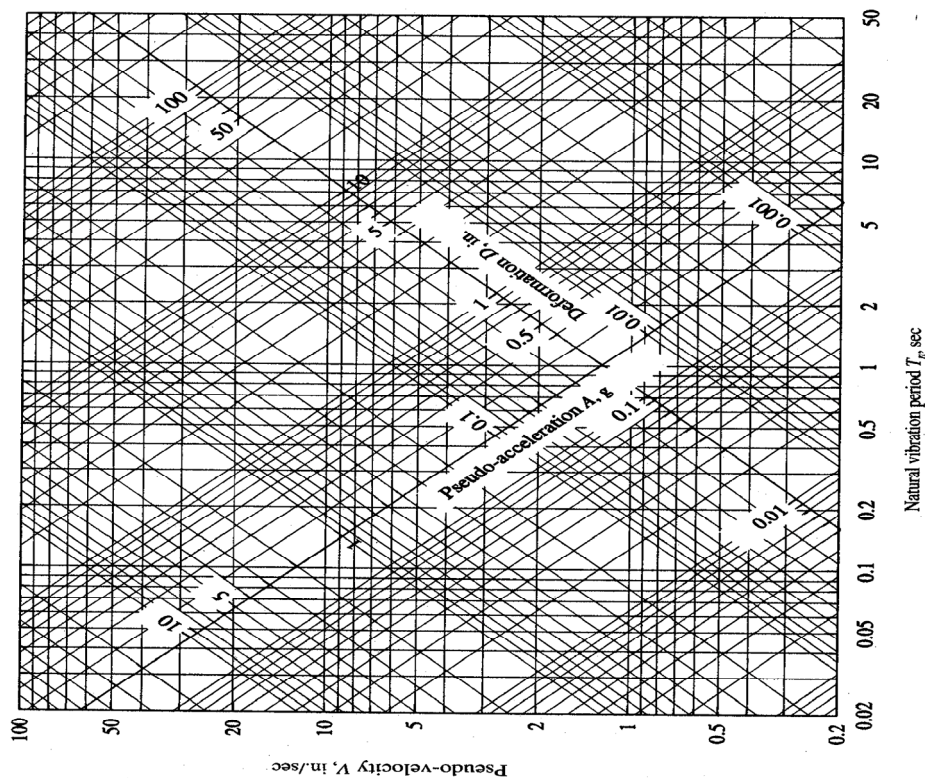
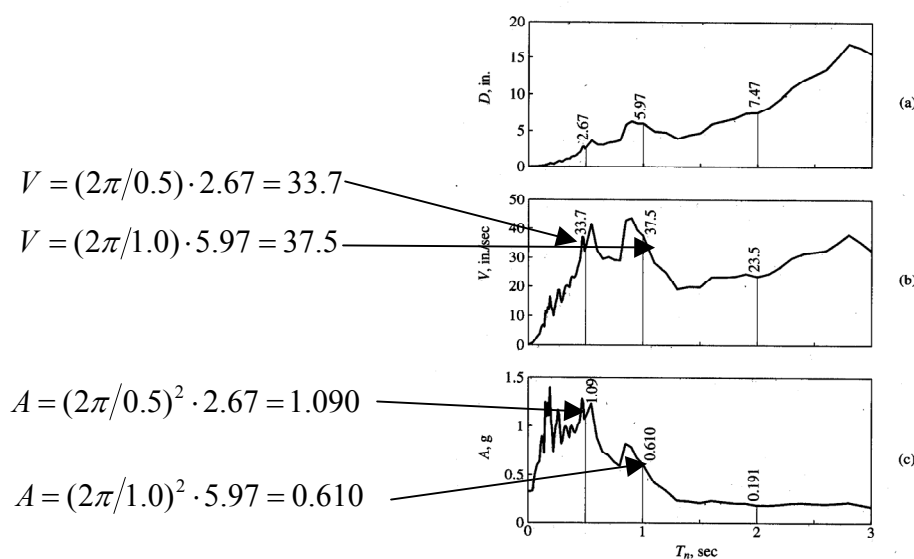


Figura 9. Representación de las variables desplazamiento(deformación), pseudovelocidad y pseudoaceleración en un gráfico tripartito.

Veamos con un ejemplo como se obtienen y se representan los tres valores para un cierto periodo o frecuencia del espectro de respuesta. Si consideramos el caso siguiente:



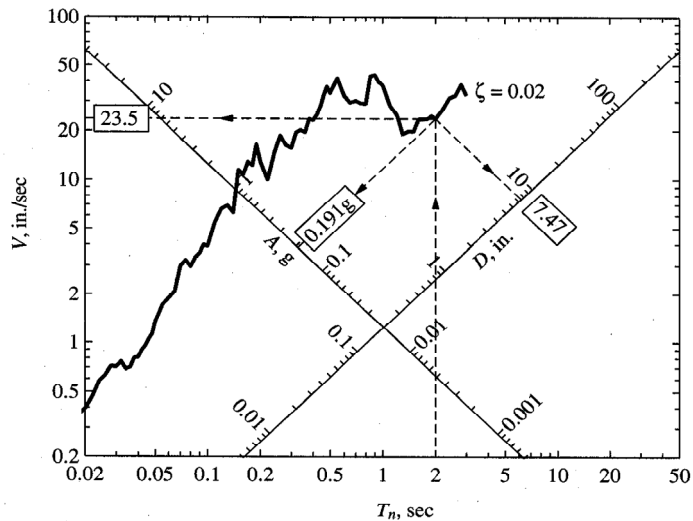


Figura 10. Pseudovelocity, pseudoacceleration and displacement in a tripartite representation of the response spectrum in tripartite graph. Chopra, 1995

Cabe por último preguntarse ¿porque es necesario calcular tres valores Desplazamiento, Pseudovelocity y Pseudoacceleration? La razón fundamental es que en el cálculo estático de acciones tenemos que la fuerza estática equivalente f_s

$$(1) \quad f_s = ku(t) = m\omega_n^2 u(t) = mA(t)$$

que como vemos depende de la pseudoacceleration. Por otra parte el valor máximo de la energía de deformación de una cierta estructura es:

$$(2) \quad E_0 = ku_0/2 = kD^2/2 = k(V/\omega_n)^2/2 = mV^2/2$$

que vemos depende de la pseudovelocity. Es decir las tres variables son de utilización en el cálculo de estructuras ante la acción sísmica por lo que conviene su representación en este diagrama tripartito.

El espectro de diseño puede considerarse como un espectro de respuesta promedio de los diferentes espectros que puedan considerarse de diferentes terremotos en un mismo emplazamiento. Por generalización, al no existir muchos espectros para un mismo lugar se consideran espectros de diferentes lugares que se considere puedan ser compatibles con el emplazamiento en cuestión. Con estos espectros obtenidos para un cierto valor de amortiguamiento (generalmente se considera el 5 %) se realiza un promedio o un espectro envolvente de todos ellos y un posterior suavizado, obteniéndose el espectro de diseño correspondiente.

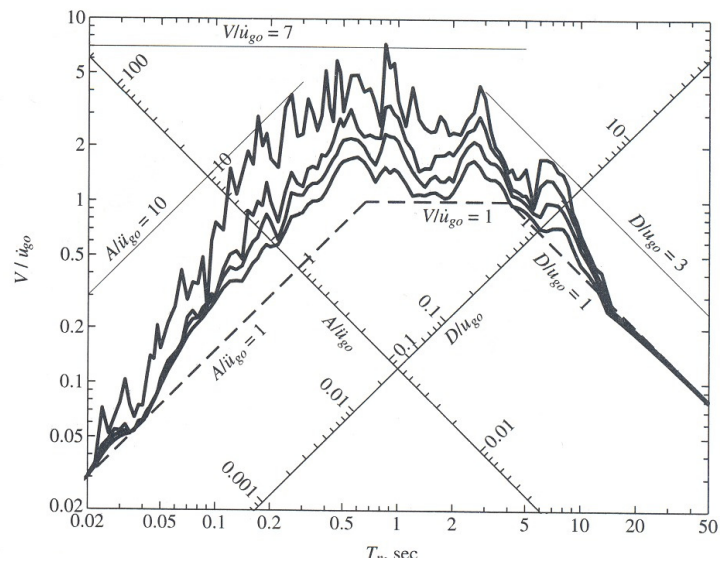


Figura 11. Espectro de respuesta normalizado de un sismo en El Centro. Chopra, 1995.

En la figura 11 se presentan los espectros normalizados a las escalas de Velocidad, $V/\dot{u}_g$, desplazamiento D/u_g y aceleración $A/\ddot{u}_g$.

Cuando esta suma se realiza para un gran número de acelerogramas se obtiene un espectro mucho más suavizado que se denomina espectro de diseño. En la Figura del espectro de diseño se puede observar que dicho espectro está dividido en cinco partes lineales con vértices con las frecuencias más significativas y que para su comparación con el espectro de respuesta de alguno de los MSF que lo conforman se representa también en dicho gráfico tripartito.

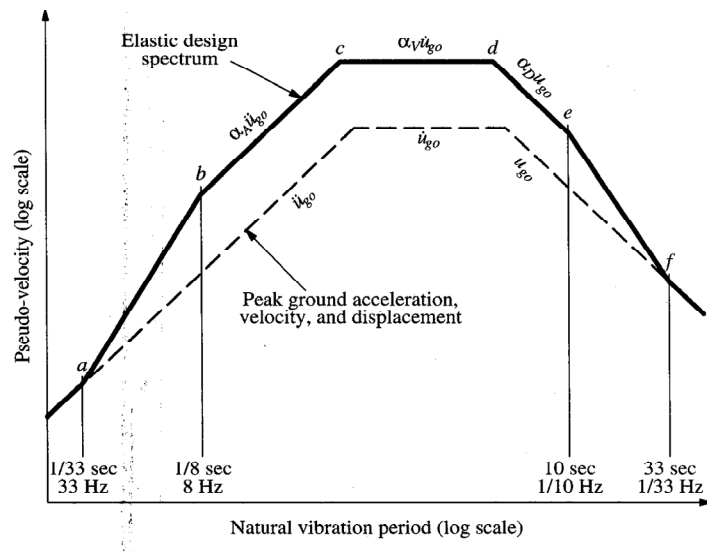


Figura 12. Espectros elásticos de respuesta (línea discontinua) y de diseño., Kramer, 1996

Las diferencias entre el espectro de respuesta y el de diseño son conceptuales y se pueden clasificar en dos grandes líneas: el espectro de respuesta es un gráfico que muestra la respuesta de un oscilador armónico simple de diferentes frecuencias y amortiguaciones a un movimiento sísmico determinado mientras que el espectro de diseño muestra un nivel de diseño ante acciones sísmicas en función de la frecuencia y el amortiguamiento de la estructura. Por otra parte para un determinado lugar el espectro de diseño es la envolvente de dos espectros de diseño diferentes que contemple otro tipo de sismos no previstos en el primer espectro de diseño.

Bibliografía

- [1] Steven L. Kramer (1996). Geotechnical earthquake engineering, Ed. Prentice Hall. Cap. 3
- [2] P. C. Jennings (2003). An Introduction to Earthquake response of structures. Part B Cap. 67. en International Handbook of Earthquake & Engineering Seismology, Editores: W. HK. Lee, H. Kanamori P. Jennings y C. Kisslinger. Ed. Academic Press.

Tema 18. Peligrosidad y riesgo sísmico. Conceptos. Características de los métodos determinista y probabilista. Periodo de retorno. Normativa de Construcción Sismorresistente en España.

18.1 Peligrosidad y riesgo sísmico. Conceptos

Estos dos términos son con frecuencia utilizados de forma indistinta aunque realmente tienen una definición que los hace diferentes. La peligrosidad sísmica describe el potencial de peligro como consecuencia de los terremotos que tiene un determinado emplazamiento y se expresa generalmente por medio de un parámetro como pueda ser la aceleración máxima horizontal, la posibilidad de liquefacción del suelo, la intensidad sísmica esperable, etc. Estos fenómenos pueden resultar en unas consecuencias adversas para la sociedad tales como la pérdida de vidas y/o la destrucción de edificios y otras estructuras de gran impacto como hospitales, etc. De esta forma, el riesgo sísmico es la probabilidad de ocurrencia de consecuencias adversas para el hombre.

El resultado de un estudio de peligrosidad sísmica puede ser la descripción de la intensidad de movimiento del terreno a causa de un determinado sismo o un mapa que muestre los niveles de movimiento del terreno en distintas partes de un territorio que tengan la misma probabilidad de poder superarse, durante un cierto tiempo. Por tanto el riesgo sísmico puede finalmente expresarse en la probabilidad de daño ocasionado por sismos y puede venir dado en su valoración económica. Por tanto para obtener el riesgo sísmico es necesario conocer previamente la peligrosidad sísmica y sus consecuencias lo que puede expresarse de forma esquemática así:

$$\text{Riesgo sísmico} = \text{Peligrosidad Sísmicas} \times \text{Vulnerabilidad} \times \text{Valor de exposición.}$$

El desastre sísmico, no es otra cosa que la realización del riesgo sísmico y existen varias formas de actuar ante él. Existen medidas de prevención que tratan de desarrollar acciones que protejan de una forma permanente ante un desastre sísmico. Es decir medidas que tratan de anular el riesgo sísmico, lo que se puede hacer únicamente o bien anulando la vulnerabilidad de las edificaciones lo que no es posible hacer de forma absoluta o bien anular el valor de exposición. Esto se puede lograr a base de no situar en la zona ningún tipo de estructura que sea susceptible de dañarse ante un terremoto. La otra medida que podemos utilizar es la de mitigación que consiste en tomar medidas para reducir el impacto de un desastre sísmico, lo que se consigue disminuyendo la vulnerabilidad. Estas medidas de mitigación se consiguen a partir de actuar

mediante la elaboración de normas constructivas que disminuyan la vulnerabilidad. Por último existen medidas de preparación que tratan de predecir, responder y tomar en consideración los efectos de un desastre. Desgraciadamente en el caso del terremoto las medidas de predicción no existen o no son fiables por lo que solamente en aquellos riesgos sísmicos en los que el fenómeno causante del daño pueda predecirse con antelación, como es el caso de tsunamis asociados a terremotos, pueden intentarse medidas de preparación.

En lo que sigue estudiaremos el peligro sísmico que forma parte de la Sismología puesto que en el estudio del riesgo, la evaluación de la vulnerabilidad de las construcciones y su estimación en términos económicos es un problema de mayor alcance e implica a la ingeniería civil, la construcción y las técnicas de seguros entre otras disciplinas.

Ante un riesgo como el sísmico u otro riesgo natural las decisiones se toman dentro de un contexto probabilístico de tal forma que el coste de reducir el riesgo este equilibrado con la probabilidad de que ese suceso se produzca. Si por ejemplo existe un cierto peligro como puede ser el sísmico con una probabilidad de ocurrencia inferior a 10^{-5} /año se considera en este caso despreciable en comparación con otros peligros que acechan al hombre por cualquier otro peligro natural. Si por el contrario la probabilidad anual de un peligro supera el valor de 10^{-2} el peligro entonces se considera muy elevado y es necesario reducirlo. Téngase en cuenta que para un país desarrollado la probabilidad de muerte para un ciudadano es del orden de 10^{-2} y por tanto cualquier peligro que lo supere debería ser reducido.

Otra consideración respecto a las medidas para reducir el peligro es que hay que tomarlas en el contexto de los demás peligros que en ese emplazamiento existen y la reducción debe ser acompañada para todos.

18.2 Características de los Métodos Determinista y Probabilista

En las primeras etapas de la determinación de la peligrosidad sísmica se desarrollo un método que se ha denominado Análisis Determinista de la Peligrosidad Sísmica (ADPS). Un ADPS se basa en el desarrollo de un escenario sísmico que permita determinar un movimiento del terreno que pueda ser peligroso para las estructuras que estén en el emplazamiento objeto de estudio. Este escenario tiene como premisa fundamental el hecho de que la sismicidad es estacionaria. Es decir que los terremotos que han ocurrido en el pasado son los que se puedan repetir en el futuro, por lo que la determinación del mayor terremoto que se pueda producir ya se ha producido con anterioridad. El proceso de un ADPS se puede esquematizar en cuatro pasos que se puede ver en la Figura 1.

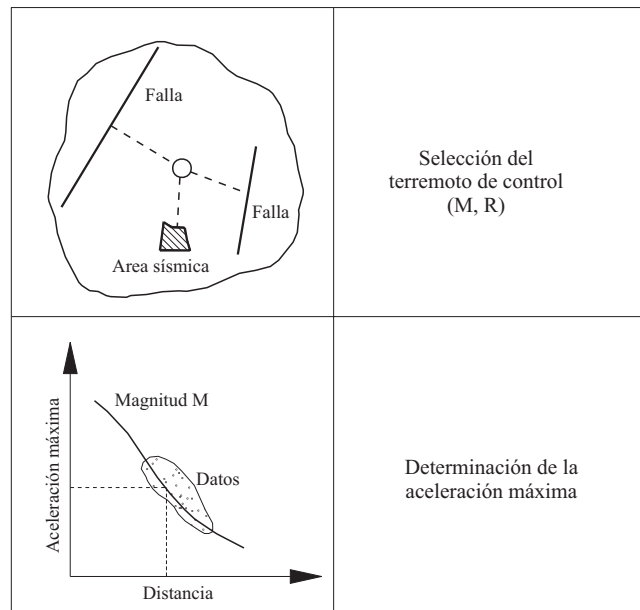


Figura 1. Esquema del proceso de un ADPS.

En el primero paso se determinan las estructuras tectónicas que es posible definir en el entorno del emplazamiento objeto de estudio como responsables de la sismicidad registrada en el pasado. A continuación se determina, para cada estructura o falla tectónica cual es el mayor terremoto que se ha producido anteriormente. Así tenemos un terremoto máximo para cada falla que se desplaza de su localización al lugar más próximo al emplazamiento. Es decir tendremos lo que se llama el *terremoto de control* de cada estructura tectónica y la distancia mínima entre el terremoto de control y el emplazamiento de posible acción.

La determinación del terremoto máximo implica en primer lugar considerar qué se entiende por este término. Existe por ejemplo el denominado *terremoto máximo posible* que define un límite en el tamaño del terremoto que una cierta estructura pueda generar, aunque sea muy improbable su ocurrencia y es el tipo de definiciones que se utilizará en el análisis probabilístico de la peligrosidad sísmica que veremos más tarde. Sin embargo en este análisis determinístico el concepto de *terremoto máximo creíble* que es el sismo que pueda producirse mediante una evaluación de aquellos procesos tectónicos razonables. Existe por último otro sismo denominado *terremoto máximo histórico* que es el que se tienen evidencias históricas o instrumentales de haberse producido en una determinada falla o zona sismógena y que sea considerado como el de mayor tamaño. De forma general se recurre para determinar el terremoto de control de una fuente a partir de un estudio detallado de ella, con determinación de los parámetros geométricos de la falla o fallas dominantes en la fuente. Esto es, determinando la longitud máxima que podría romperse en caso de sismo y a partir de esta información recurrir a ajustes empíricos determinados con terremotos actuales en los que ambos tipos de información se conozcan. Así existen algunos trabajos en este sentido como el propuesto por Wells y Coppersmith.

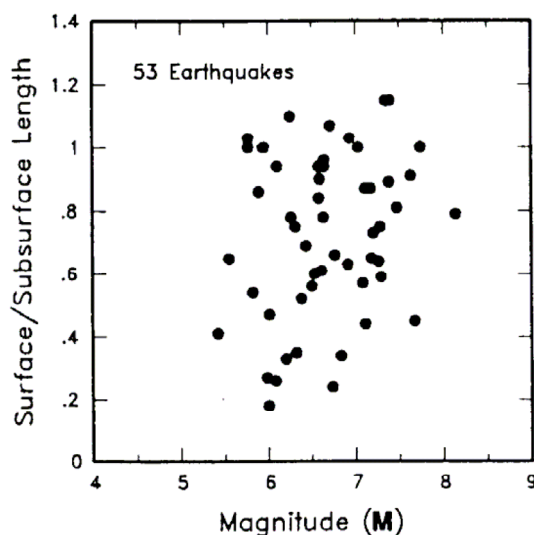


Figura 2. Relación empírica entre la longitud de rotura de falla y magnitud, Wells y Coppersmith, 1994

Con estas relaciones es posible a partir de valoraciones de posibles longitudes de rotura determinar el valor de la magnitud del sismo asociado.

A continuación, mediante relación de atenuaciones existentes para la zona trasladaremos en función de la distancia las acciones expresadas por ejemplo por la Aceleración Máxima Horizontal (AMH) o cualquier otro parámetro como puede ser la intensidad.

Por último de todas estas posibles acciones sobre el emplazamiento elegiremos la mayor de todas que consideramos como la peligrosidad sísmica que afectará al emplazamiento. Es decir consideramos el peor escenario posible.

Este método determinista tiene la ventaja de que el resultado se corresponde con la acción de un único terremoto que pudo haber ocurrido en el pasado y que podría hacerlo en el futuro. Sin embargo tiene el inconveniente de que el resultado no tiene asociado ningún parámetro que exprese la verosimilitud del resultado durante el intervalo de tiempo de interés.

De entre las acciones realizadas para realizar un ADPS es el primer paso el que más sujeto está a decisiones subjetivas, por lo que determinar lo que se denomina el potencial sísmico es un problema de alcanzar un consenso entre sismólogos, geólogos, y en general aquellos técnicos capaces de establecer como puede ser la actividad sísmica de una determinada falla tectónica.

Desde finales de los años setenta se empezó a considerar que muchos de los elementos que intervienen para estudiar los efectos de un terremoto están sujetos a incertidumbres que sería preciso considerar. Es decir incertidumbre del lugar donde se vaya a producir el próximo terremoto, incertidumbre en el tamaño asociado a él y por último incertidumbre asociada a la ley de atenuación que permite trasladar los efectos de un sismo a una cierta distancia. Esta nueva forma de determinar el peligro se denomina *Análisis Probabilista de la Peligrosidad Sísmica*

(APPS) y fue formulado por primera vez por A. C. Cornell en 1968. En la Figura se presentan las etapas fundamentales de un APPS que son:

1. Determinación de las fuentes sismogénicas, es decir de aquellas estructuras que potencialmente serán responsables de los futuros terremotos que se puedan producir. Para ello, han de determinarse la geometría de las fuentes (fallas, áreas, provincias sismotectónicas) responsables de la sismicidad de la zona de estudio. Este paso es idéntico al llevado a cabo en el ADPS. En general la distribución de terremotos en cada una de estas fuentes se considera homogénea, lo que implica que no existe información en la sismicidad pasada que permita establecer que la ocurrencia de sismos sea con probabilidad diferente en una determinada zona sismogénica. Esto se traduce en que para cada zona existe una función única de distribución de la sismicidad. Esto es lo mismo que decir que en el APPS vamos a permitir que la futura sismicidad se pueda dar en cualquier punto de cada una de las fuentes sismogénicas.

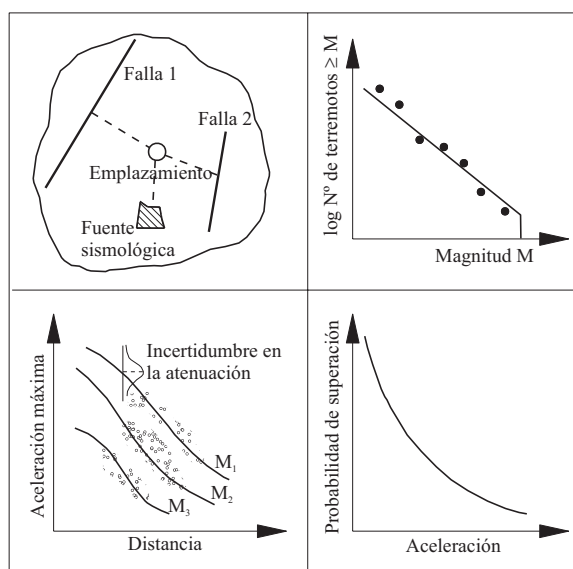


Figura 3. Esquema de un APPS.

2. Determinación de la ley de recurrencia de la sismicidad. Es preciso caracterizar para cada zona sismógena como se va a producir la distribución de la sismicidad en cuanto a tamaño se refiere. Esta caracterización se realiza mediante una relación de recurrencia que especifique la tasa anual de que sea superado un terremoto de un determinado tamaño. Además de esta tasa es necesario establecer cual va a ser el tamaño máximo (magnitud o Intensidad máxima) del terremoto que se pueda producir en dicha zona, independientemente de la ley de recurrencia determinada.

Para determinar la ley de recurrencia consideramos la ley de Gutenberg-Richter de cada zona en la que hemos establecido un valor mínimo de magnitudes y otro máximo. Esta acotación se basa en la necesidad de considerar únicamente aquellos terremotos que puedan tener interés en

ingeniería ya que valores más pequeños no van a afectar a las estructuras. Este valor mínimo se considera según la zona en la que se haga el estudio pero siempre oscila entre magnitud 3.5 a 4.5. En cuanto a la magnitud máxima es necesaria una estimación basada tanto en la actividad registrada para la zona como en las características sismotectónicas de la zona. Por consiguiente la ley de recurrencia para cada zona con ambas acotaciones será.

$$N(m) = N(m_0) \frac{e^{-\beta(m-m_0)} - e^{-\beta(m_{\max}-m_0)}}{1 - e^{-\beta(m_{\max}-m_0)}}$$

Donde $N(m_0)$ es la tasa anual de terremotos con magnitud igual o superior a la magnitud mínima m_0 , y $m_{\max}$ es el valor considerado como magnitud máxima. Esta expresión se representa así:

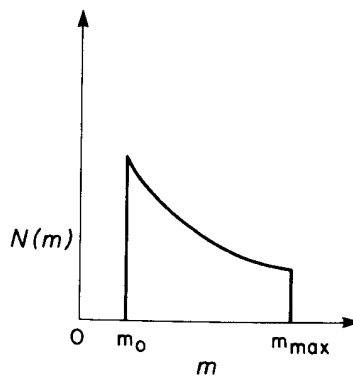


Figura 4. Distribución acotada de la ley de recurrencia de Gutenberg-Richter.

Ahora bien, como vamos a trabajar con probabilidades de ocurrencia de cada una de las magnitudes posibles en la distribución es necesario calcular la *función de distribución acumulada* de esta relación acotada de Gutenberg-Richter que nos expresa la probabilidad de que una cierta magnitud M sea menor que m dentro del intervalo definido por la acotación. Esto es:

$$F_M(m) = P[M < m | m_0 \leq m \leq m_{\max}] = \frac{1 - e^{-\beta(m-m_0)}}{1 - e^{-\beta(m_{\max}-m_0)}}$$

Esta expresión será utilizada posteriormente en el cálculo de la peligrosidad.

3. Para cada posible terremoto, que se puede producir en cualquier punto de cada zona a la que se le ha asignado una tasa y distribución de tamaño dado por la ley de recurrencia anterior para cada zona, es necesario trasladar sus efectos al emplazamiento. En otras palabras, para cada punto de la falla que se considere, tomaremos todos los posibles terremotos que la ley de recurrencia de la falla tenga asignada y trasladaremos sus efectos al punto de estudio. Para ello es necesario conocer una relación predictiva del movimiento del suelo en función de la magnitud y de la distancia al emplazamiento. En este paso existe por tanto una incertidumbre asociada a esta relación entre el movimiento del suelo y la magnitud y distancia.

Actualmente las relaciones se determinan mediante el ajuste empírico de la aceleración registrada por diferentes terremotos de magnitudes distintas y a diferentes distancias. También se puede utilizar la intensidad máxima como caracterizador del tamaño del sismo y entonces la ley empírica se obtiene a partir de las intensidades registradas a distintas distancias para terremotos de la zona.

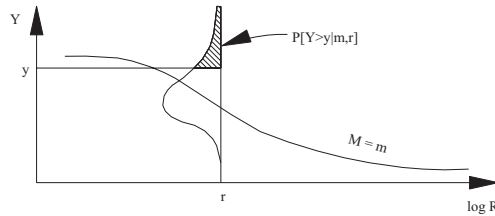


Figura 5. Para una magnitud y distancia dadas se puede obtener la probabilidad condicional de que Y sea $> y^*$

Considerada una ley de amortiguación, para una determinada magnitud y distancia, la probabilidad de que el sismo supere un cierto valor de aceleración y es:

$$P[Y > y^* | m, r] = 1 - F_Y(y^*)$$

Siendo $F_Y(y^*)$ la función de distribución acumulada de la ley de atenuación.

4. Todas las incertidumbres asociadas tanto a la posible localización del terremoto del futuro como las correspondientes al tamaño y a la relación predictiva vista en el apartado anterior se combinan para obtener la probabilidad de que el movimiento del suelo supere un cierto valor durante un intervalo de tiempo de interés considerado.

Finalmente la determinación de la peligrosidad sísmica probabilista se define como la suma de todos los efectos sobre el emplazamiento de los hipotéticos terremotos de diferentes tamaños (M) que puedan producirse en cualquier parte de cada zona sismógena siguiendo en el futuro una ley de recurrencia idéntica a la determinada para el pasado.

Este resultado, suma de todos los efectos posibles sobre el emplazamiento puede expresarse en función de un parámetro de interés en ingeniería, como es por ejemplo una aceleración del movimiento del suelo y como es lógico también del número de veces que se alcanza o supera un cierto valor de aceleración de referencia, en un determinado intervalo de tiempo.

Considerando un intervalo de tiempo de un año, la frecuencia anual de que en un emplazamiento, la aceleración calculada A supere un cierto valor de aceleración previamente fijado a^* , $N^o(A > a^*) / 1 \text{ año} = \lambda_a^*$.

$$\lambda_a^* = \sum_{i=1}^N \lambda_i \int_{m=M_{\min}}^{M_{\max}} \int_r P[A > a^* | M, R] f_{m_i}(M) f_{r_i}(R) dm dr$$

donde :

λ_i es la tasa anual de sismos superiores a una magnitud $M_{\min}$, para la fuente i.

λ_a^* es la tasa anual de superación de la aceleración a^* .

$f_{m_i}(M)$ y $f_{r_i}(R)$ son las funciones densidad de probabilidad en magnitud y distancia de cada fuente i.

$P[A > a^* | M, R]$ es la probabilidad de que un determinado sismo de magnitud M a la distancia R del emplazamiento produzca una aceleración superior a a^* mediante la relación de atenuación.

N es el número de zonas sismógenas.

Si suponemos además que el modelo temporal de producción de terremotos es un modelo probabilístico de Poisson, podremos expresar el resultado anterior en términos de probabilidad de que la aceleración calculada A supere esa aceleración previamente fijada a^* en un determinado periodo de tiempo T , denominado *tiempo de exposición, vida útil o periodo de interés* de la estructura que pretendemos proteger:

$$P(A > a^*) = 1 - e^{-\lambda_a^* T} = 1 - e^{-T/T_R}$$

donde $T_R = 1/\lambda_a^*$ es el denominado Periodo de Retorno.

En particular para valores de aproximadamente $\lambda_a^* T \leq 0.1$ se puede aproximar la expresión anterior:

$$P(A > a^*) \cong \lambda_a^* T$$

Si el cálculo descrito para las tasas anuales de superación de una determinada aceleración λ_a^* lo repetimos para distintos valores de aceleración a^* obtenemos una curva que describe cual es la probabilidad de superación de los distintos valores a^* , es decir lo que hemos denominado el peligro sísmico en un emplazamiento y que es la curva que se presenta en el cuarto paso de la Figura.

El considerar que los terremotos siguen una distribución de Poisson implica que la probabilidad de ocurrencia de un sismo no depende del tiempo que ha pasado desde el último sismo. En otras palabras, los terremotos son sucesos independientes en el tiempo.

Existen otros modelos temporales como son: el modelo no homogéneo de Poisson, modelos de Markov, etc.

18.3 Periodo de Retorno

Otra forma de expresar los resultados anteriores se obtiene si en vez de fijar una aceleración a^* fijamos un nivel de probabilidad de superación en la expresión anterior sin aproximar:

$$\lambda_{a^*} = -\frac{\ln[1 - P(A > a^*)]}{T}$$

y en función del denominado periodo de retorno, llamado así por tener dimensiones de tiempo y que no es otra cosa que la inversa de la tasa anual de superación de la aceleración λ_{a^*} correspondiente, por lo que podemos expresarlo en función de la probabilidad de superación :

$$T_R = \frac{-T}{\ln[1 - P(A > a^*)]}$$

Por tanto otra forma de expresar el peligro sísmico es también mediante este periodo de retorno T_R .

Esta expresión es igualmente válida si en vez de considerar como variable que representa el movimiento del suelo una aceleración como es la AMH lo hacemos para cada valor de la aceleración del espectro de respuesta para unos determinados valores del periodo del oscilador armónico simple y del amortiguamiento considerado. El resultado final sería lo que se denomina espectro uniforme de respuesta.

El valor a considerar del periodo de retorno o lo que es lo mismo de la probabilidad de superación dependerá del grado que estimemos necesario de protección ante el sismo y este valor debe estar en consonancia con otros riesgos naturales que tenga el emplazamiento. En otras palabras no se puede proteger de forma desigual un emplazamiento ante distintas amenazas naturales. En general los estudios de peligrosidad están enfocados a la protección de la edificación de las ciudades, por lo que es necesario establecer el tiempo de vida útil de estas construcciones. En general se considera una vida útil de 50 años, tiempo en el que se considera amortizada una edificación normal. Pues bien el nivel de protección que generalmente se estima es de una probabilidad del 10 % de que el nivel de aceleración supere al valor de diseño de la construcción durante la vida útil de esta. Estas condiciones nos conducen a un periodo de retorno de 475 años o sea aproximadamente de 500 años. Es decir que los estudios de peligrosidad de un emplazamiento en el que se va a construir una edificación, consistirán en

evaluar la aceleración que tenga un 10 % de ser superada durante la vida útil de dicha construcción o lo que es equivalente que tenga un 90 % de no superar el valor de esa aceleración.

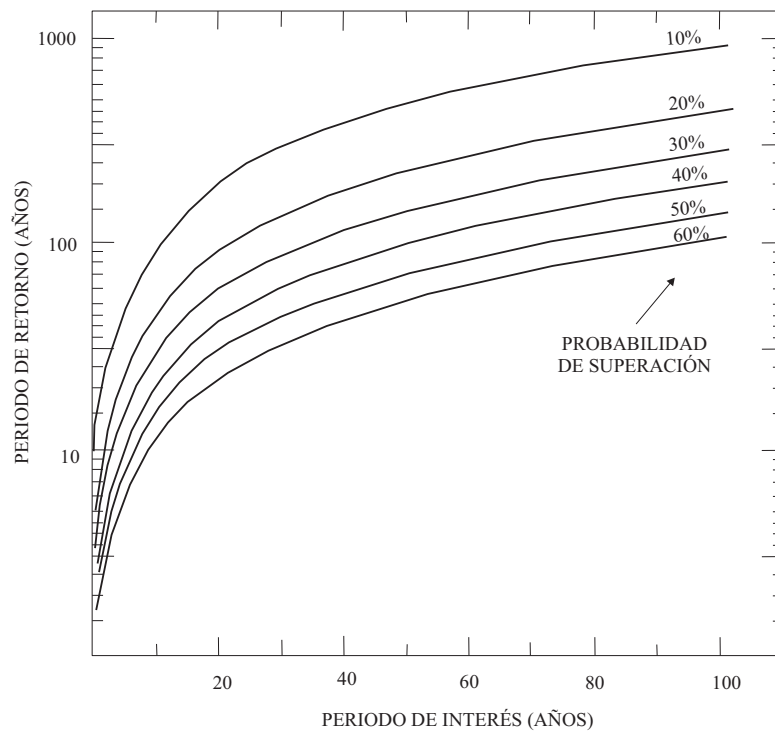


Figura 6. Relación entre el periodo de retorno, la probabilidad de superación de un cierto valor de la variable de peligrosidad y la vida útil de la estructura objeto de estudio. Mezcua, 2002.

18.4 Normativa de Construcción Sismorresistente en España

La *normativa sísmica* constituye la principal estrategia de protección ante los terremotos al incidir sobre los dos factores determinantes del *riesgo sísmico* que pueden modificarse: los elementos expuestos (*exposición*) y la *vulnerabilidad* de dicha exposición (la *peligrosidad sísmica*, no puede modificarse). En un sentido amplio, abarca tanto las normas de uso del suelo, urbanísticas y de instalaciones urbanas (de distribución de gas, por ejemplo), como las normas de cálculo y diseño sismorresistente de las construcciones.

De estas normas, las primeras inciden fundamentalmente sobre la exposición, protegiéndonos básicamente de los fallos del terreno provocados por los sismos (deslizamientos, licuefacción, etc.) y de los efectos indirectos, como el fuego o los maremotos. Las normas de cálculo y diseño sismorresistente contribuyen en cambio a la reducción de la vulnerabilidad de las

construcciones, y su aplicación debería ser subsiguiente a la de las normativas de uso del suelo y urbanística.

Ciñéndonos a nuestro país, no existe ninguna norma sísmica específica de uso del suelo, si bien, la consideración del fenómeno sísmico se ha tenido en cuenta -o al menos, citado- como un factor físico limitativo en algún que otro Plan de Ordenación del territorio. No obstante, la Norma de Construcción Sismorresistente NCSE-02 establece en su primer artículo que la consecución de los objetivos de protección está condicionada por el cumplimiento de los *preceptos limitativos de uso del suelo* que dicten las Administraciones Públicas competentes (Ayuntamientos y Comunidades Autónomas). En cambio, en España disponemos desde el año 1968 de Normas de Construcción Sismorresistente de Edificación, así como -desde el año 1962- de preceptos -artículos de otras normas- que determinan la acción sísmica a aplicar a diferentes tipos de construcciones.

En la actualidad, la *normativa vigente de carácter sismorresistente* está constituida por las normas:

NCSE-02. Norma de Construcción Sismorresistente: Partes General y de Edificación

NCSP-07. Norma de Construcción Sismorresistente: Parte de Puentes

Eurocódigo 8 (EN 1998)

a lo que habría que añadir las disposiciones relativas a las acciones sísmicas establecidas en otras normas (Instrucción y Reglamento Técnico de Seguridad de Presas, Instrucción de Hormigón Estructural, Recomendaciones de Obras Marítimas, etc.), así como la reglamentación -no española- exigida por el Consejo de Seguridad Nuclear para las Centrales Nucleares e Instalaciones radioactivas de nuestro país.

El *objetivo* de todas las normas sismorresistentes es lograr un determinado *nivel seguridad* de las construcciones caracterizado por:

- la protección de la vida humana
- el evitar el colapso ante un sismo de cálculo relativamente grande
- la reducción de los daños y el coste económico para sismos pequeños
- el mantenimiento del servicio en ciertas construcciones críticas (hospitales, por ejemplo).

Esto obliga a *clasificar las construcciones* de acuerdo con su uso y las consecuencias de su fallo, definiéndose para cada clase un *sismo de cálculo* diferente, que las normas de los países más desarrollados, como la NCSE o el Eurocódigo, establecen para *edificaciones ordinarias* en el que tiene un periodo de retorno de unos 500 años (probabilidad anual de superación 1/500). Este periodo de retorno se duplica, al menos, para aquellas construcciones *de importancia especial o crítica* por su uso o el servicio que se presta en ellas.

La *Norma de Construcción Sismorresistente de Edificación (NCSE-02)* es la norma española que proporciona los criterios básicos relativos a la sismorresistencia aplicables en el proyecto, construcción y conservación de las edificaciones y obras ubicadas en el territorio nacional. Su *ámbito de aplicación* se extiende a todas las edificaciones de nueva planta y -parcialmente- a las obras de reforma o rehabilitación, a fin de que el nivel de seguridad de los elementos afectados de éstas sea superior al que poseían en su concepción original. Además, las denominadas *prescripciones de índole general*, es decir, los *criterios de aplicación*, el *cumplimiento de la norma en las fases de proyecto, construcción y explotación* y aquellos otros artículos que definen los parámetros que determinan la *aceleración sísmica de cálculo (clasificación de las construcciones, mapa de peligrosidad sísmica, aceleración sísmica básica* e, implícitamente, *coeficiente de amplificación del terreno)*, son de aplicación supletoria a los demás tipos de construcciones, siempre que no existan disposiciones específicas de carácter sismorresistente que les afecten.

De estos parámetros, la *aceleración sísmica básica -aceleración característica* que mide la peligrosidad sísmica en el emplazamiento para un periodo de retorno de 500 años- es el parámetro cartografiado en el *mapa de peligrosidad sísmica* y define los *criterios de aplicación* de la NCSE, constituyendo además el *parámetro de referencia* para múltiples actuaciones ligadas a campos tan dispares como la Protección Civil o el aseguramiento contra terremotos. En cambio, la *aceleración sísmica de cálculo* representa la aceleración del suelo en el emplazamiento concreto de la estructura para el periodo de retorno exigido de acuerdo con el tipo o clase de construcción considerada. A partir de ella, se determinan el *terremoto de cálculo* (acelerograma o espectro de respuesta de cálculo) y las *acciones sísmicas* de la estructura.

En relación con los métodos de cálculo, la NCSE-02 fija especificaciones para el *análisis directo en el dominio del tiempo* y para el *análisis modal espectral*, si bien, sólo desarrolla un *método simplificado de cálculo* basado en la aplicación de fuerzas estáticas horizontales. Para que se pueda aplicar este método, la edificación debe satisfacer una serie de condiciones ligadas a su *geometría y simplicidad estructural*, y a la *regularidad en planta y alzado de la distribución de masas y rigideces*, condiciones que pueden conseguirse con un buen *diseño conceptual*. Sucintamente, el método de cálculo simplificado de la NCSE-02 se basa en aplicar a las edificaciones, a nivel de cada planta, fuerzas estáticas horizontales (no dinámicas como el sismo), equivalentes a las máximas fuerzas de inercia que origina el terremoto, y cuya resultante viene dada por el producto de la *masa (de cálculo) del edificio* por la *aceleración espectral*, valor este último que depende de la aceleración sísmica de cálculo y del periodo fundamental y amortiguamiento de la estructura, así como de determinadas características constructivas y de los materiales que posibilitan la disipación de energía en el rango post-elástico, es decir, de su

comportamiento dúctil. La consecución de esto último es uno de los objetivos fundamentales de todas las normas sismorresistentes actuales, ya que, además de reducir sensiblemente las fuerzas de cálculo, contribuye a disminuir notablemente el número de víctimas, al evitarse los *fallos frágiles*. Para lograr este comportamiento dúctil, la NCSE-02 dedica un capítulo entero a *reglas de diseño y prescripciones constructivas* relacionadas con la configuración estructural y los distintos elementos y materiales, cuyo cumplimiento exige explícitamente.

La *Norma de Construcción Sismorresistente: Parte de Puentes (NCSP-07)* contiene los *criterios específicos para la consideración de la acción sísmica en el proyecto de puentes de carretera y de ferrocarril*. Su ámbito de aplicación abarca aquellos puentes en los que las acciones horizontales son resistidas básicamente por los estribos o mediante la flexión de las pilas, aplicándose también -de forma parcial- a los puentes en arco o atirantados.

La Norma de Puentes define los criterios de proyecto, los sismos de cálculo y las exigencias de comportamiento para los periodos de construcción y explotación. Además, caracteriza la acción sísmica (a estos efectos, añade al espectro de respuesta de la NCSE-02 una nueva rama, correspondiente a los periodos largos), desarrolla diversas metodologías de cálculo, y establece los criterios y requisitos para el dimensionamiento y las comprobaciones resistentes a llevar a cabo en los elementos estructurales y de la cimentación. Sus disposiciones son totalmente congruentes con el marco establecido en la NCSE-02 y acordes con la Parte de Puentes del Eurocódigo 8.

El *Eurocódigo 8* es la norma de la Unión Europea que complementa los nueve restantes Eurocódigos de Construcción en los aspectos relacionados con la sismorresistencia. Está dividido en seis *Partes* que abarcan -además de las *Reglas generales* y los *Puentes*- otras cuatro de las que no disponemos de normas sismorresistentes específicas españolas (*Refuerzo de edificios*; *Silos, depósitos y conducciones*; *Cimentaciones*; y *Torres, mástiles y chimeneas*). En la actualidad, todas las Partes están aprobadas por la Comisión Europea, por lo que son ya *Normas de la Unión Europea* (normas EN 1998-1 a EN-1998-6). Una vez que nuestro país realice y apruebe su traducción al castellano, serán además *Normas UNE-EN*.

Todos los Eurocódigos de construcción se complementan con *Anejos Nacionales* en los que cada país puede fijar el valor de *determinados* parámetros, coeficientes, expresiones o condiciones, con objeto de lograr el nivel de protección deseado (los Eurocódigos dan unos valores recomendados). Actualmente, se está redactando el Anejo español al Eurocódigo 8 que el Ministerio de Fomento, como responsable de la seguridad de la construcción, habrá de aprobar. Este Anejo deberá garantizar que la seguridad sismorresistente que proporciona el

Eurocódigo 8 no será inferior a la que se logra mediante la aplicación de las normas de construcción españolas.

No existe ninguna norma española de Centrales Nucleares; todas las centrales existentes en España se calcularon con las normas del país origen de la tecnología de la central. En la actualidad, el Consejo de Seguridad Nuclear exige la aplicación de la normativa americana de la *Nuclear Regulatory Commission*, (NRC), junto con la de la Agencia de la Energía Nuclear de la OCDE, (ENA), para los proyectos relacionados con las Instalaciones radioactivas, incluidos sus aspectos sismorresistentes.

Bibliografía

- [1] Steven L. Kramer (1996). *Geotechnical Earthquake Engineering*. Ed. Prentice Hall. Cap. 4.
- [2] Julio Mezcua (2002). Peligrosidad sísmica Capítulo 14. en *Riesgos Naturales*, Coord. F.Ayala-Carcedo y J. Olcina Ed. Ariel.

Tema 19. Tsunamis. Generación, propagación e inundación. Magnitud e intensidad del Tsunami. Características de los terremotos productores del Tsunami. Sistemas de alerta de Tsunami.

19.1 Tsunamis. Generación, propagación e inundación

Un tsunami o maremoto es una ola que se propaga con una cierta velocidad a través de los océanos. Cuando alcanza la costa correspondiente, la ola puede sufrir alteraciones en su altura y velocidad debido a la geometría de la plataforma continental y fenómenos asociados a la geometría de la costa.

El origen de esta ola está asociado a un desplazamiento brusco del fondo submarino producido por un terremoto o una intrusión masiva de material en una zona tal que se produce un desplazamiento brusco de una importante masa de agua.

El periodo de estas olas es muy diferente al del oleaje marino producido por el viento. Así, en un tsunami el periodo puede oscilar entre 5 minutos y 2 horas lo que equivale longitudes de onda entre 20-300 km con alturas de un máximo de 1 m en océanos abiertos con velocidades de propagación del orden de 900 km/h mientras que en el oleaje estos valores son de 6-10 s, longitudes de onda de 100-200 m, alturas de 3-5 m y velocidades de 2-30 km/h. Vemos que las características de ambos tipos de ola son muy diferentes, aparte de su origen completamente diferente. Cuando el tsunami se acerca a la costa, como veremos más adelante cambian las propiedades de propagación, ralentizándose su velocidad y aumentando de forma espectacular la altura de la ola. Así en la costa la velocidad disminuye hasta 50-350 km/h con acortamiento de la longitud de onda a valores de 1.5-3 km y las alturas de ola aumentan a valores de 3-30 m.

La palabra japonesa tsunami quiere decir “onda de puerto” es el nombre que de forma casi exclusiva se utiliza en todos los idiomas aunque en español tiene el nombre de maremoto que tiene la limitación de que su origen es por terremoto y no por otra causa como los de origen volcánico.

Existen muchos ejemplos de tsunamis de gran importancia en la historia de fenómenos naturales responsables de gran cantidad de personas afectadas. Merecen citarse el tsunami generado por el Terremoto de Lisboa de 1 de noviembre de 1755, que afectó muy extensamente a las costas ibéricas y del norte de África. También es importante el tsunami de Chile de 22 de mayo de 1960 que recorrió todo el Pacífico llegando 23 horas después de generarse a las costas

japonesas, produciendo muchas víctimas. Más recientemente el 26 de diciembre de 2004 se produjo un terremoto de magnitud 9.2 que originó un tsunami en las costas de Sumatra en Indonesia y que se transmitió por todo el Océano Indico, produciendo alrededor de 290.000 víctimas.

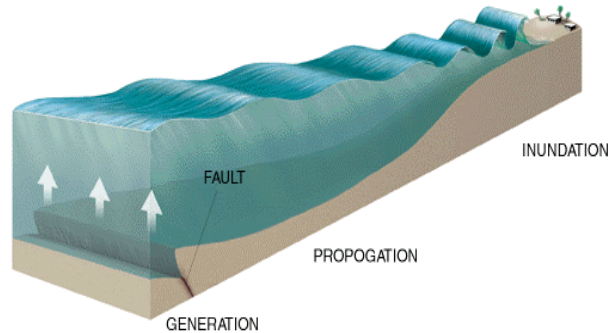


Figura 1. Fases de estudio de un Tsunami

Un tsunami se genera de forma general por un desplazamiento brusco del fondo marino a causa de un sismo. El problema de la generación de desplazamientos del suelo marino está asociada a la formación y rotura a partir de una falla del terreno por acumulación de esfuerzos generados por motivaciones tectónicas. El conocimiento de la geometría de la fractura y de las propiedades elásticas del medio permite aplicar diversas teorías elásticas para determinar el campo de desplazamientos generado en el fondo marino. Realizando estudios sobre modelos teóricos se observa que cuando se supone una falla de tipo inverso, generalmente asociada a fenómenos de subducción de una placa oceánica bajo una continental, el campo de desplazamientos en la superficie del medio es muy grande. En la Figura 1 se presenta una situación como la descrita y el campo de desplazamientos generado en el fondo del mar. Este estudio sería un estudio típico de resolución de un problema de generación de tsunamis.

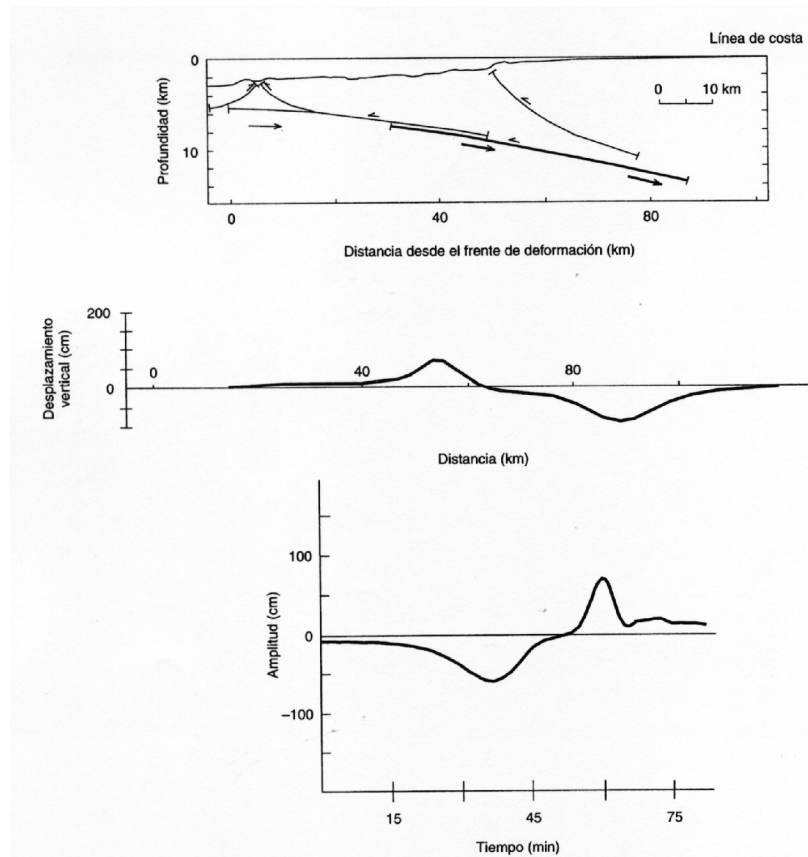
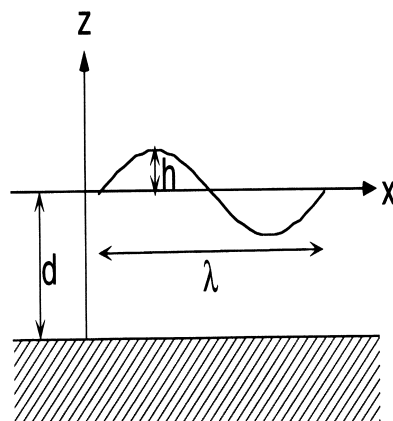


Figura 2. Falla en subducción, campo de desplazamientos y la ola generada. Mezcua, 2002

Una vez se genera el tsunami la ola se propaga por el océano siguiendo las leyes de la hidrodinámica y existen varias teorías para estudiar el problema. La más sencilla es la llamada *teoría de aguas superficiales u ondas largas*, en las que la escala horizontal del movimiento, es decir su longitud de onda es, mucho más grande que la profundidad del agua del medio por el que se propaga. En estas condiciones la ecuación de movimiento que ha de aplicarse



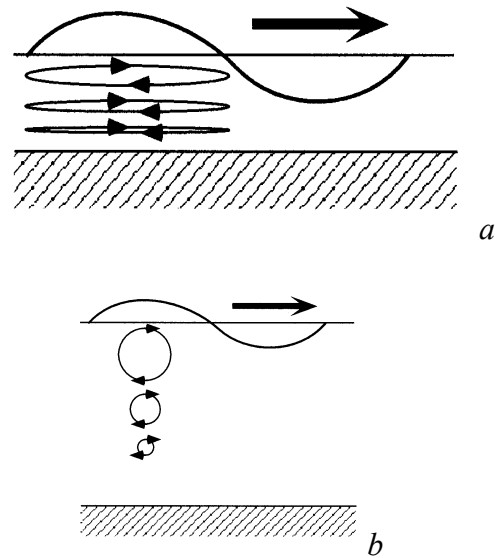


Figura 3.

es la llamada Ecuación de Eüler,

$$\frac{d\vec{V}}{dt} = -\vec{g} - \frac{1}{\rho} \nabla p$$

Donde $\vec{V}$ es el vector velocidad cuyas componentes en el plano (x,z) son (u,w), $\vec{g}$ es el vector gravedad y ρ y p la densidad del agua y la presión respectivamente. Sin embargo,

$$\frac{d\vec{V}}{dt} = \frac{\partial \vec{V}}{\partial t} + \vec{V} \nabla \vec{V}$$

Si como hemos establecido la escala horizontal del movimiento es mucho mayor que la profundidad del agua, es decir la longitud de onda $\lambda \gg d$, entonces la aceleración vertical del agua en su movimiento es despreciable frente a la acción de la gravedad. Esto se traduce en que el movimiento de la columna de agua se mueve de forma uniforme desde la superficie hasta el fondo.

Si nos fijamos únicamente en la componente horizontal del movimiento, la gravedad no actuará y la componente horizontal de la presión se sustituye por la pendiente de la superficie del agua

$$\frac{du}{dt} = -g \frac{\partial h}{\partial x}$$

Despreciando el término advectivo, la ecuación resultante será:

$$\frac{\partial u}{\partial t} = -g \frac{\partial h}{\partial x}$$

Si la amplitud de la ola $h \ll d$, la ecuación de continuidad será:

$$\frac{\partial h}{\partial t} = -\frac{\partial}{\partial x}(du)$$

Se puede por tanto igualando las dos últimas expresiones

$$\frac{\partial^2 h}{\partial t^2} = c^2 \frac{\partial^2 h}{\partial x^2}$$

Siendo $c = \sqrt{gd}$ la velocidad de propagación de la ola en este caso de ondas largas o también llamado de propagación en aguas someras o superficiales. Es preciso tener en cuenta que el concepto de agua superficial es relativo, pues dada la longitud de onda considerada, del orden de varias decenas de kilómetros, la profundidad así denominada es del orden de 4000- 5000 m. Este tipo de propagación se puede aplicar a la parte de trayectoria en océanos profundos, pero cuando se acerca la ola a la costa ya no puede aplicarse esta simplificación. En estos casos en los que $d \gg \lambda$ estamos en el caso de *propagación en aguas profundas u ondas cortas* Figura b resultando una propagación dispersiva de velocidad

$$c = \sqrt{\frac{g\lambda}{2\pi}}$$

Si la longitud de onda del tsunami es mucho menor que la escala de la heterogeneidad de velocidad, se puede estudiar la propagación utilizando la teoría de rayos. Los frentes de onda pueden localizarse como el lugar geométrico de los rayos con idéntico tiempo de propagación y que se han denominado diagramas de refracción, como puede verse en la Figura.

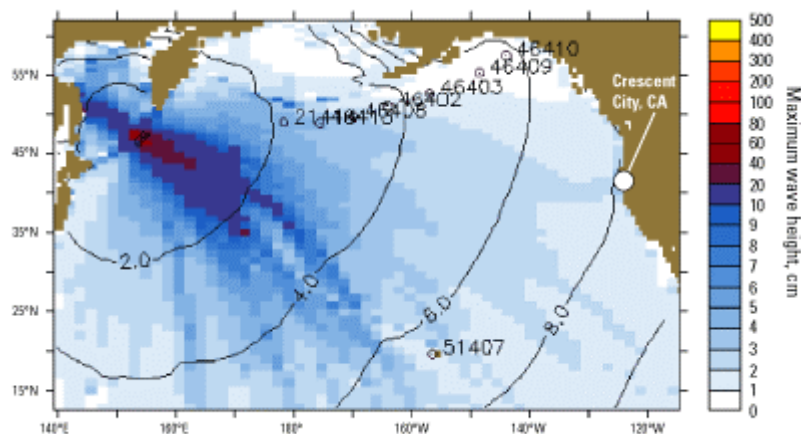


Figura 4. Diagrama de refracción para un tsunami generado en las islas Kuriles y su propagación por el Océano Pacífico. En color también se da la altura máxima de la ola del Tsunami.

Por último en la propagación es necesario conocer no solamente a que tiempo va a llegar una ola y cual ha sido su trayectoria hasta el punto de interés sino también conocer cual va a ser la altura de la en un punto de su trayectoria. Para ello existen varias aproximaciones a su

resolución. Entre las mas sencillas esta la ley de Green que se basa en la conservación de la energía a lo largo del rayo lo que se puede traducir en $b_0 d_0^{1/2} h_0^2 = b_1 d_1^{1/2} h_1^2$ siendo d la profundidad del agua y b la distancia entre dos rayos consecutivos y h la altura de la ola. El punto 0 es un punto de la trayectoria en el que conocemos toods los valores y 1 es el punto de interés. Por lo tanto en este punto 1 la altura de la ola vendrá dada por la expresión:

$$h_1 = \left(\frac{b_0}{b_1} \right)^{1/2} \left(\frac{d_0}{d_1} \right)^{1/4} h_0$$

Existen dos conceptos ligados a la altura que alcanza la ola en su choque contra la costa. Uno es lo que se denomina *altura de la ola*, que es la altura observada con un mareógrafo que alcanza el tsunami al alcanzar la costa. Esta altura es observada en un instrumento y por tanto se refiere al nivel medio del mar en ese punto. Sin embargo la ausencia de muchos instrumentos de este tipo en las costas dificultaría su observación en caso de tsunami, por lo que se considera otra altura denominada *altura de inundación* (altura run up) que se refiere a la altura que en una determinada zona alcanza la ola en su proceso de inundación del terreno. Esta altura se determina por observación de los restos que deja la ola al llegar a esta altura e iniciar el proceso de retroceso hasta el nivel de costa.

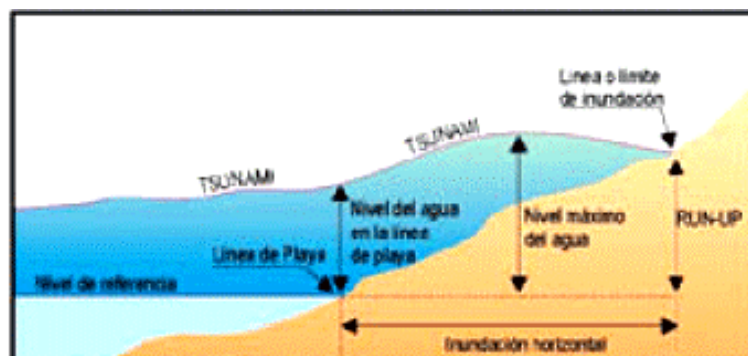


Figura 5.

19.2 Intensidad y magnitud del tsunami

La intensidad de un tsunami es un valor descriptivo de la fuerza con que un tsunami ha afectado un determinado lugar de la costa. Esta asignación se hace a partir de la comparación de efectos que ocasiona sobre las personas, las edificaciones y las embarcaciones y en general los daños que ocasiona. Es por tanto una asignación cualitativa que al igual que la intensidad en los terremotos se expresa en números romanos para así impedir la utilización algebraica de estos valores. Existen numerosas escalas de intensidad, por lo que únicamente mencionaremos aquí

una de ellas. La escala de Sieberg- Ambraseys está dividida en seis valores de intensidad y son: I para aquellos que son imperceptibles y solo detectado por mareógrafos, III: Apreciable. Se observa retirada del mar en playas muy horizontales. Daños leves en estructuras ligeras. Pequeñas embarcaciones quedan sobre la playa. O en el caso más dañino VI: Desastroso. Destrucción total o parcial de estructuras. Grandes barcos son lanzados a tierra. Árboles arrancados De esta forma la intensidad de un tsunami en un cierto emplazamiento será función de la vulnerabilidad de las construcciones y densidad de población de la zona, además del tamaño del terremoto responsable del tsunami.

La magnitud del tsunami en cambio intenta reflejar una estimación del tamaño atendiendo a la observación de una magnitud física como puede ser la altura de la ola. En este caso la asignación de un valor es observado hace de la magnitud una variable objetiva del tsunami.

Así, puede considerarse la magnitud de Inamura-Iida

$$m = \log_2 h$$

donde h es la altura máxima de inundación registrada en ese lugar, expresada en metros. En la práctica los valores que resultan oscilan entre -1 y +4, dando la impresión de estar ante una escala de intensidades de seis grados. A la vista de esta definición, Soloviev propuso una definición de intensidad de tsunami :

$$i_s = \log_2 \sqrt{2}h$$

siendo h la altura media alcanzada en esa parte de la costa, expresada en metros. Aunque se denomina a este valor intensidad es en realidad una magnitud pues se deduce de la observación de una variable como es la altura de inundación del tsunami.

Siguiendo la filosofía del cálculo de magnitudes en el caso de los terremotos, Abe definió la magnitud del tsunami:

$$M_t = a \log h + b \log \Delta + D$$

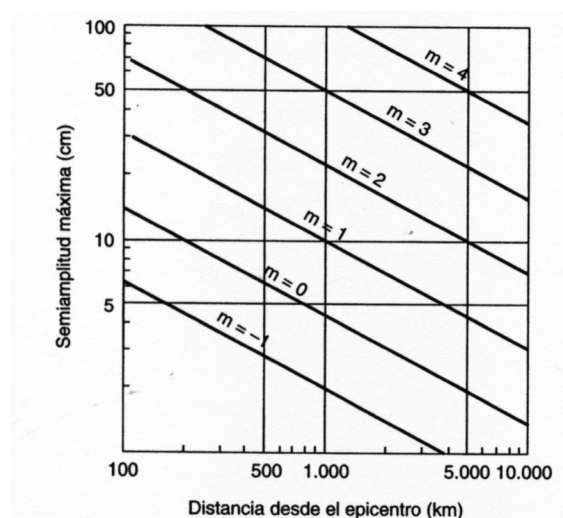


Figura 6. Estimación de la magnitud en función de la amplitud de la ola y distancia al epicentro, Mezcuca, 2002.

Donde h es la altura observada en un mareógrafo situado en la costa, medida en metros, Δ es la distancia en kilómetros entre el epicentro del sismo y la estación mareográfica correspondiente y las constantes a , b y D son el resultado de una calibración para que el resultado de magnitud del tsunami sea idéntica a la magnitud del sismo que lo generó, según la definición de magnitud momento.

Existen una gran variedad de definiciones de magnitud de tsunami fundamentalmente asociadas a diferentes áreas geográficas o bien a conceptos a medir diferentes, como puede ser la magnitud de Murty y Loomis,

$$ML=2(\log E-19)$$

Donde E es la energía potencial del tsunami expresada en ergios.

19.3 Características de los terremotos productores de tsunamis

Como se ha indicado la generación de la gran mayoría de los tsunamis se produce por deformación del suelo oceánico inducida por un terremoto. De la observación de los tsunamis más importantes, observamos los tipos de falla en los que se han producido:

- Sismo en la **falla de empuje entre-placas**. (Chile 1960) Generan los mayores terremotos tsunamigénicos. Importantes teletsunamis.
- Sismo **entre-placas, previo al frente sísmico**. Mecanismos de falla de empuje (Nicaragua 1992). Terremotos tsunami. Grandes tsunamis.
- Sismo de **falla de empuje intraplaca**. En la zona posterior del arco de isla (Flores 1992).
- Sismo **intraplaca en la zona oceanica**. Si es superficial puede ser en el promontorio pre-trinchera (Falla normal Sanriku, 1933) o a gran profundidad (Tonga, 1977).

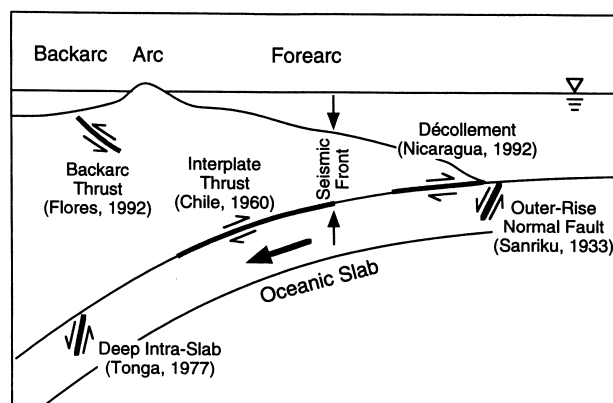


Figura 7. Localización de los distintos tipos de terremotos productores de tsunamis.

Puede observarse que los *Terremotos tsunamigénicos*, es decir aquellos terremotos con mecanismo de falla de empuje de pequeño ángulo entre-placas litosféricas que están alejados de

la trinchera de subducción, son de profundidad superficial y producen tsunamis. Estos tsunamis asociados a estos terremotos se propagan a grandes distancias por lo que el tiempo de llegada a diferentes lugares es previsible y se pueden tomar medidas de preparación. Sin embargo, existen sismos entre-placas con mecanismo de falla de empuje que se producen en las zonas previas al frente de subducción, en la zona del prisma de acrección. Estos terremotos excitan tsunamis de amplitud mucho mayor de lo que se debería esperar para la magnitud correspondiente. A este tipo de sismos se les denomina *Terremotos tsunami* y son mucho más peligrosos porque la magnitud del sismo no se corresponde con la magnitud del tsunami generado. Con menor magnitud del sismo se produce un tsunami mucho mayor. Estos terremotos se denominan también *terremotos lentos* por tener una velocidad de ruptura de la falla mucho más lenta al estar en la zona de acumulación de material sedimentario y poseer por tanto menor rigidez.

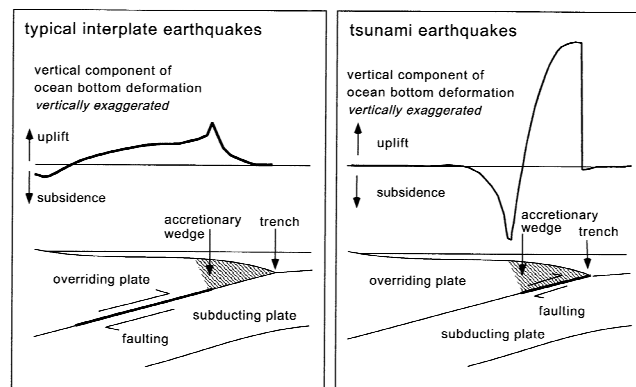


Figura 8. Desplazamiento del suelo oceánico de un terremoto entraplacas(izquierda) y un “terremoto tsunami” o “lento”

Por tanto las características de los terremotos tsunami son de falla superficial (menor de 20 km) y cercana a la trinchera previa a la subducción. La magnitud a partir de la cual se comienza a excitar tsunami es $M_w=6.5$. El mecanismo de estos sismos exhibe un gran vector deslizamiento que se traduce en una deformación vertical mucho mayor que la que le correspondería en un material más rígido. Es decir la velocidad de ruptura de la falla es mucho más lenta y por tanto el tiempo de ruptura es mucho mayor. De esta propiedad reciben el nombre de *terremoto lento*. Estos terremotos son mucho más peligrosos porque con magnitudes moderadas se consiguen tsunamis de gran altura de ola.

19.4. Sistemas de alerta de tsunami

En este peligro natural es posible actuar con medidas de preparación puesto que es posible, en algunos casos, determinar si un sismo va a tener características de terremoto tsunamigénico o tsunami y predecir cual va a ser el instante de impacto a las costas correspondientes. Es decir,

podemos utilizar la diferencia enorme entre la velocidad de las ondas sísmicas del terremoto y la velocidad del tsunami generado por él. Por tanto un sistema de alerta de tsunamis esta basado en tres acciones independientes:

- Detección del terremoto.
- Determinación de las características tsunamigénicas.
- Confirmación de la propagación del tsunami.

La primera acción es común en todas las redes sísmicas y consiste en la detección y cálculo de localización y tiempo origen de cualquier terremoto dentro del área de interés. Esta acción se realiza de forma automática en todos los centros sísmicos de los diferentes países y también en los centros regionales que tienen a su cargo la vigilancia de grandes extensiones de la tierra como es el caso del Centro Sismológico Europeo o bien más específicamente para la vigilancia de tsunamis el Centro de Alerta de Tsunamis de los países del Océano Pacífico.

A continuación es necesario determinar las características en cuanto a mecanismo focal y determinación de otros parámetros tales como duración de la función fuente del sismo, Energía del sismo, etc que permitan determinar si un sismo que ha ocurrido tiene características que le hace potencialmente productor de un tsunami. En la figura se puede ver como la determinación de la alerta, una vez analizadas sus características para el Centro de Alerta de Tsunamis de los países del Océano Pacífico, es cada vez más corta.

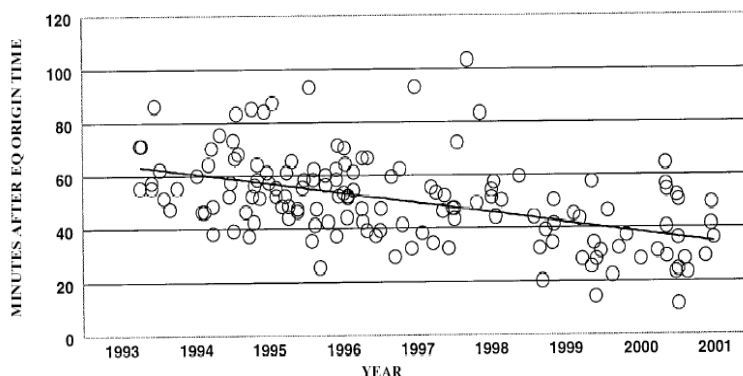


Figura 9. Tiempos de respuesta desde la hora origen del sismo según el año

Por último dada la trascendencia de una alerta de tsunami es necesario que una vez determinado que el terremoto estudiado reúne las características de ser productor de tsunami comprobar su propagación y altura de la ola.

En la Figura se presenta un dibujo esquemático de cómo se realiza la confirmación de un tsunami, mediante su observación en el Océano al paso por boyas en los que existe la información de un sensor de presión que mide el paso del tsunami.

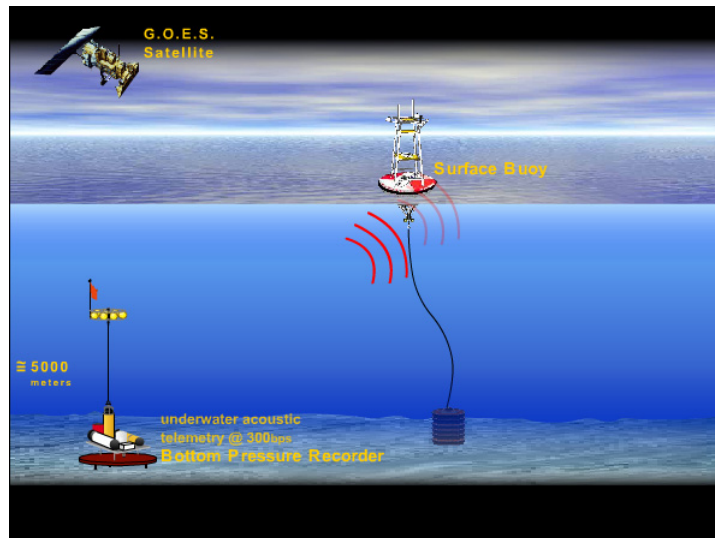


Figura 10. Sistema de seguimiento de un tsunami

Bibliografía

- [1] Julio Mezcua (2002). Riesgo de Tsunamis. Capitulo 57. en Riesgos Naturales, Coord. F.Ayala-Carcedo y J. Olcina Ed. Ariel
- [2] K. Satake.(2002). Tsunamis. En Internacional Handbook of Earthquake and Engineering Seismology. Editores. W.H.K. Lee, H. Kanamori, P. C. Jennings y C. Kissliger. Ed. Academic Press. Cap. 28.

Tema 20. Volcanismo. Materiales volcánicos. Mecánica de los fenómenos eruptivos. Proyección de piroclastos. Extrusión y dinámica de domos y coladas. Mapas de peligrosidad.

20.1 Materiales volcánicos

El estudio de los materiales emitidos por un volcán permite una clasificación simple en función del estado en que se encuentren estos materiales. Existe una parte gaseosa y una parte líquida que no obstante lleva gases disueltos:

- **Gases.**

- componentes principales en el sistema H-O-C-S.

- **Lava.** Fase líquida con gases disueltos.

- Fragmentario. Piroclastos: *Tephra*

- Masivo. *Domos y coladas*.

- Fragmentario con explosión o colapso: *colada de Piroclastos*

Los gases emanados en una erupción volcánica no responden exactamente a los correspondientes en el magma original sino que en el proceso de ascensión los mecanismos de cambio de flujo y contaminación en la parte superficial por los acuíferos hacen que estos se vean modificados. No obstante, los componentes fundamentales H-O-C-S se considera que sufren de fenómenos de oxidación dando como consecuencia tres productos que se han detectado en todos los procesos eruptivos analizados: H₂O, CO₂ y SO₂. De entre todos estos productos es con diferencia el agua el que contribuye con más proporción y además se pone en evidencia porque al estar a altas temperaturas se volatiliza en forma de vapor de agua.

También se producen, en el proceso de expulsión a la superficie reacciones de reducción formándose además de los ya existentes de H₂O, CO₂ y SO₂ los siguientes componentes CO, y CH₄. De una forma no generalizada también se encuentran otros materiales en algunos volcanes como son el ClH y FH. El primero de ellos es muy importante por el impacto que tiene en la destrucción de la capa de Ozono y su posterior efecto invernadero.

Existen otros componentes con una menor concentración como son: Ar, He, N₂ y otros que se manifiestan al enfriarse y su mezcla con el aire, condensándose a las bajas condiciones de presión y temperatura. Los sublimados que se forman constituyen aerosoles que se adosan o

depositan en las paredes de las grietas. Los compuestos más frecuentes son los cloruros y sulfatos de Na, K, Ca, Mg, Al y Fe así como óxidos de Si y Fe.

Como ejemplo de composición de gases de algunos volcanes se presenta en la tabla siguiente:

Volcano Tectonic Style Temperature	Kilauea Hot Spot 1170°C	Summit Spot 1130°C	Erta` Divergent Plate 1130°C	Ale Convergent Plate 820°C	Momotombo Plate
H ₂ O	37.1		77.2		97.1
CO ₂	48.9		11.3		1.44
SO ₂	11.8		8.34		0.50
H ₂	0.49		1.39		0.70
CO	1.51		0.44		0.01
H ₂ S	0.04		0.68		0.23
HCl	0.08		0.42		2.89
HF	---		---		0.26

Tabla 1. Contenido en gases de algunos volcanes en diferentes estilos tectónicos. Symonds et al. 1994

Los contenidos lávicos con gases disueltos al contacto con la atmósfera sufrirán un proceso de liberación de los gases disueltos, un enfriamiento y su consiguiente consolidación dando lugar a rocas volcánicas. Este material sólido puede ser de carácter fragmentario o piroclastos cuya salida del volcán puede ser por proyección aérea o por caída. Los piroclastos con tamaños desde 2 mm hasta 1 m, cuando se depositan en el terreno se denominan de forma genérica Tephra, aunque existe la tendencia a asociar a este nombre únicamente aquellos depósitos de piroclastos que se encuentra suelto y no cementado.

Los diferentes tipos y denominaciones de Tephra son:

- Ceniza.(<2 mm)
- Lapilli.(2-64 mm)
- Bloques. (>64 mm)
- Escorias.
- Bombas.
- Hilos de cristal volcánico.
- Pómez.

Las cenizas son fácilmente arrastradas por el viento y se depositan a grandes distancias del volcán. Los lapilli, son materiales de composición basáltica de color negro. Las escorias son fragmentos de lava correspondientes a magmas poco viscosos que al caer todavía calientes se sueldan entre si. La pómez son piroclastos de composición acida y se encuentra en fragmentos de todos los tamaños. Son de color claro y muy porosos lo que les permite flotar en el agua.

En cuanto a las lavas de carácter masivo tenemos que las emitidas por los volcanes basálticos son más abundantes que las correspondientes a volcanes calcoalcalinos, aunque en ambos casos el volumen emitido es inferior al de piroclastos. Como ejemplo podemos decir que los volcanes de Hawai, que son los que mas lava emiten, están en un valor medio de $0.1 \text{ km}^3/\text{año}$. La clasificación en cuanto a lavas masivas se refiere tenemos dos productos:

- Domos.
- Coladas.

Los domos son material lávico de gran viscosidad cuya principal característica es su enraizamiento en la chimenea volcanica. Según se produzca este enraizamiento por acumulación sobre la boca de sucesivos derrames lávicos se denomina domo exógeno, mientras que si la formación se desarrolla a partir de la expansión de la lava de una masa magmática que se extruye lentamente se llama endógeno.

En relación a las coladas, su recorrido en el terreno depende fundamentalmente de la viscosidad que tengan y de la pendiente topográfica del terreno alcanzándose en algún caso distancias del centenar de kilómetros. El espesor de estas depende también de la pendiente del terreno. Una clasificación permite distinguir las siguientes tipos:

- Malpais o aa
- Bloques
- Lisas o Pahoehoe.
- Cordadas.

Cuando el enfriamiento de la superficie de la colada se produce rápidamente y su flujo es todavía alto se produce un cuarteamiento con rotura de la superficie, generando una superficie de escoria. Si la viscosidad es alta esta división se lleva a cabo en grandes bloques que se distribuyen de forma caótica. Estas dos formas se denominan Malpais o aa y de bloques respectivamente. Si por el contrario hay una gran fluidez y elevada temperatura en la colada, estas adquieren una superficie lisa, de color negro que se conocen con el nombre hawaiano de pahoehoe o superficie fácil de caminar. Como una variedad de estas podemos situar las

cordadas que son generadas porque la superficie experimenta arrastres locales cuando todavía está en estado plástico.

En general la salida de estos materiales por la boca del volcán así como la altura que alcanza la columna de material fragmentario unido a la frecuencia de erupciones que se puede determinar para un volcán dado ha permitido realizar una clasificación en lo que se denomina *Índice de Explosividad* de un volcán (IEV).

IEV	Clasificación	Descripción	Altura <u>columna eruptiva</u>	Volumen material arrojado	Periodicidad	Ejemplo	Total erupciones históricas
0	<u>Erupción hawaiana</u>	no-explosiva	< 100 m	> 1000 m ³	diaria	<u>Kilauea</u>	-
1	<u>Erupción stromboliana</u>	ligera	100-1000 m	> 10,000 m ³	diaria	<u>Stromboli</u>	-
2	Erupción vulcaniana/ stromboliana	explosiva	1-5 km	> 1,000,000 m ³	semanal	<u>Galeras, 1992</u>	3477
3	<u>Erupción Vulcaniana</u> (sub-pliniana)	violenta	5-15 km	> 10,000,000 m ³	anual	<u>Nevado del Ruiz, 1985</u>	868
4	Vulcaniana (sub-pliniana)/ pliniana	cataclísmica	10-25 km	> 0.1 km ³	cada 10 años	<u>Galunggung, 1982</u>	278
5	<u>Pliniana</u>	paroxística	> 25 km	> 1 km ³	cada 100 años	<u>St. Helens, 1980</u>	84
6	Pliniana/ Ultra-Pliniana (krakatoana)	colosal	> 25 km	> 10 km ³	cada 100 años	<u>Krakatoa, 1883</u>	39
7	<u>Ultra-Pliniana</u> (krakatoana)	super-colosal	> 25 km	> 100 km ³	cada 1.000 años	<u>Tambora, 1815</u> <u>Maipo, 500.000 adC.</u>	4
8	Ultra-Pliniana (krakatoana)	mega-colosal	> 25 km	> 1000 km ³	cada 10.000 años	<u>Toba, 69.000 adC.</u>	1

Tabla 2. Índice de explosividad volcánica.

20.2 Mecánica de los fenómenos eruptivos

Una erupción generalmente comienza con una emisión de gases que en procesos previos a la llegada a la superficie han sido separados del material lávico. Al llegar estos gases a la superficie se produce una expansión muy rápida lo que genera una fuerte explosión. A veces y debido al grado de viscosidad de la lava, los gases han podido ir separándose de la fase líquida desde su salida de la cámara magmática con lo que no se produce una acumulación de estos y por tanto la presión alcanzada no es muy grande. Esto sucede con las lavas poco viscosas y por tanto tienen una erupción que viene precedida de emisión de gases sin alta presión, mientras que si la lava es muy viscosa, al quedar atrapada en los conductos de la chimenea, mantiene los gases disueltos en ella, aumentando la presión interna, por lo que una expansión rápida provoca una erupción con una gran explosividad. A continuación se produce la salida de la lava que se realiza a través de alturas sobre el volcán más bajas que por las que han salido los gases. Cuando termina la salida de material lávico, el proceso de salida de gases continúa hasta el total enfriamiento de la chimenea y de las cámaras magmáticas correspondientes. Esta actividad final de la erupción puede ir acompañada de otros fenómenos asociados como son: aparición de geysers, fuentes termales, etc.

Para estudiar el proceso mecánico del ascenso del magma, es necesario establecer una serie de hipótesis que permitan simplificar el problema. En primer lugar consideraremos una columna de material lávico formada por una fase líquida con cristales y otra gaseosa en un proceso isoterma. La concentración en volumen es:

$$C = \frac{V_{gas}}{V_{magma}} = H/k \frac{P_0 - P}{P} \rho_{magma}$$

Donde V es el volumen de ambas fases, H la constante de la ley de Henry que relaciona el peso del contenido en agua en el magma y P la presión, k relaciona la densidad del gas con la presión.

Si esta mezcla asciende por un conducto cilíndrico y en condiciones de régimen laminar con velocidad media v_{media} , el flujo de magma Q, cumplirá la siguiente expresión:

$$\frac{4}{3} v_{media} \frac{dv_{media}}{dp} = 8\pi\mu \frac{v_{media}^2}{Q} - \frac{H}{k} \frac{P_0 - P}{P}$$

En esta expresión se pone de manifiesto que el flujo de magma depende del contenido en gases y de las diferencias de presiones así como de la velocidad media del ascenso.

A continuación de este ascenso del material lávico se produce su salida por la boca o bocas que existen en la mayoría de los volcanes. Es frecuente la existencia de una boca principal y alrededor otras secundarias, aunque con frecuencia se produce una emisión en bocas situadas de forma alineada, siguiendo fracturas produciéndose en este caso un volcanismo denominado fisural.

La actividad fisural (magmatismo toleítico) produce erupciones de muy baja explosividad y con magmas muy fluidos eyectándose en grandes chorros verticales a lo largo de centenares de metros lo que produce una sensación de emisión en cortinas verticales de lava.

La boca de erupción esta relacionada con el cráter y por tanto también con la caldera. El cráter es la parte final del conducto de ascenso y su geometría se modifica de forma continua durante un proceso eruptivo. Así los cráteres de grandes dimensiones responden a la acumulación de muchos procesos eruptivos, al nivel freático de la zona y a los sucesivos colapsos de la caldera que se hayan podido producir. Por ejemplo el volcán de St Helen que tenia una cima de 2950 metros antes de la erupción de 1980, redujo en más de 500 m su cima por una explosión que produjo el cráter actual con dimensiones de 1.5x 3 km . En cuanto a la formación de la caldera se puede asociar a la emisión de productos de elevada concentración de gases cuyos procesos explosivos facilitan la rotura de de los materiales situados encima de la cámara facilitándose el colapso de estas.

Como hemos indicado existe otro fenómeno asociado al mecanismo de erupción y es el de las explosiones volcánicas. Para estudiar estos mecanismo explosivos es necesario establecer una serie de condicionantes geométricos en lo referente a las dimensiones del cráter; esto es, considerado cónico con radio R y altura Z correspondiente a una explosión de origen puntual la energía liberada en la explosión puede determinarse es:

$$E = \frac{5\pi\rho c^2}{r_0} Z^4 \left[1 + \left(\frac{R}{Z}\right)^2\right]^3$$

Donde r_0 es el radio de la fuente de explosión, c la velocidad y R/Z es el factor de forma del volcán cuya importancia en la energía liberada es, como podemos ver, muy importante.

Los tipos de erupciones explosivas han sido objeto de clasificación desde los primeros registros escritos sobre ellos. La clasificación comúnmente establecida es así:

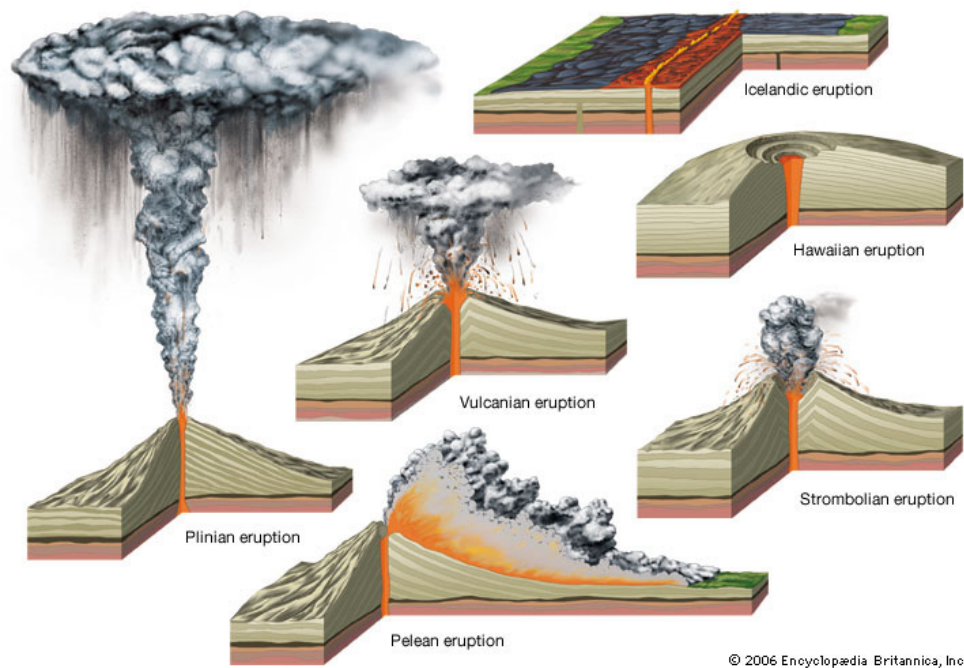


Figura 1. Tipos principales de erupciones volcánicas.

- ***Erupciones plinianas.*** Violenta explosión, continuada emisión de productos fragmentarios.
- ***Erupciones vulcanianas.*** Violenta explosión después de una etapa de tranquilidad. También ocurren en erupciones de magmas muy viscosos.
- ***Erupciones strombolianas.*** Explosiones producidas por burbujas de gas que se escapan del magma fluido.
- ***Erupciones Islandesas.*** Efusiones de lava basáltica a través de fisuras largas y paralelas que conforman mesetas de lava.
- ***Erupciones submarinas.*** Se asocian grandes depósitos de cenizas
- ***Erupciones freatomagmáticas.*** Cuando un magma fragmentado a alta temperatura alcanza el nivel freático.
- ***Erupciones anulares rasantes.*** Son superficiales, con materiales ligeros y bajo agua.
- ***Erupciones dirigidas.*** Se producen cuando existe salida de magmas muy viscosos.

20.3 Proyección y depósito de piroclastos

La emisión de cenizas en un volcán, junto con los gases es la parte más espectacular que puede darse en una erupción volcánica. Esta emisión se produce en forma de columna que en primera aproximación clasificaremos como estacionaria o de explosión. En las columnas estacionarias se

considera que la emisión de material en la boca del volcán es constante durante un cierto tiempo, mientras que en las columnas de tipo explosivo el tiempo de emisión es relativamente pequeño

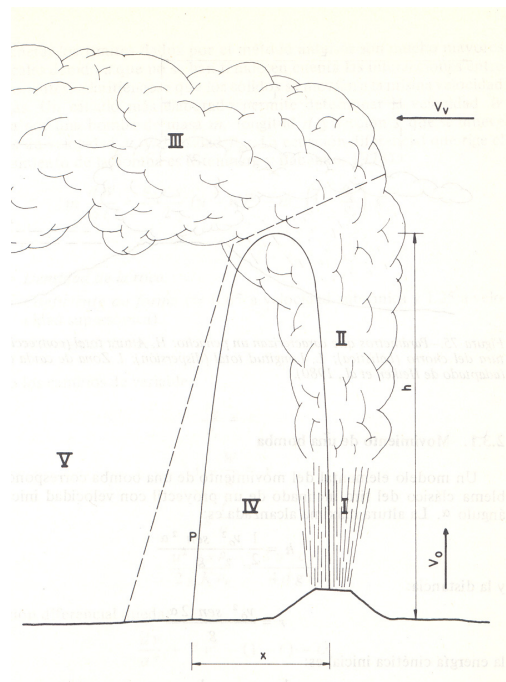


Figura 2. Zonas principales en la proyección de piroclastos. Araña y Ortiz, 1984

Para estudiar la dinámica de una columna de gas y piroclastos tomaremos una aproximación simplificada del problema. Si consideramos un ascenso adiabático, tendremos:

$$v \frac{dv}{dh} = \left(\frac{1}{q} - 1 \right) g - cv^2$$

Donde el término de rozamiento c es función de la forma de la masa que asciende y de las características de la atmósfera, v la velocidad de ascenso y q la relación entre las densidades de la columna de ascenso y la del aire que la rodea. Si se sustituye q por ρ_c / ρ y para una atmósfera típica, es decir:

$$\frac{dT}{T} = (1 - \gamma) \frac{dh}{H}$$

Que nos permite calcular el enfriamiento de la columna por dh de elevación y por tanto la velocidad de ascenso de la columna:

$$V(Z) = V(0) \left(1 - \frac{Z}{H} \right)^b$$

Donde la velocidad a cualquier altura es función de la velocidad inicial de emisión, de la altura máxima H alcanzada por la columna y b una constante adimensional.

En relación a la altura máxima alcanzada por una columna convectiva viene dada para el caso de una explosión como inicio de la columna:

$$H = 1.37 Q^{1/4}$$

Donde Q es la energía en julios.

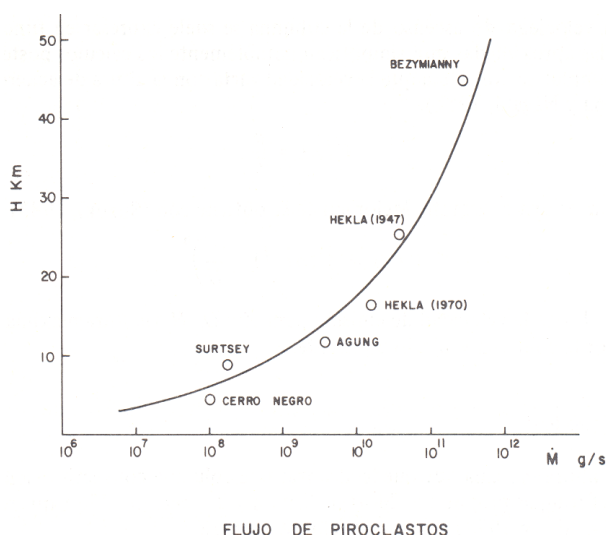


Figura 3. Altura de la columna eruptiva de diferentes volcanes Araña y Ortiz, 1984

La columna eruptiva en su ascenso está regida por la energía cinética inicial en el momento de salida de la boca. En cuanto al descenso de los materiales se produce de forma regular en los casos que la columna se alimenta de aire exterior. Sin embargo, cuando la columna, por su tamaño, forma en su interior celdas de convección interna de gran amplitud en la dimensión vertical con escasa incorporación de aire exterior, los materiales descienden con velocidades similares a la del ascenso, formando una envoltura descendente alrededor de la columna, constituyendo el modelo de colapso de una columna eruptiva. Este colapso puede ser violento cuando se produce una disminución del contenido en gases, por un descenso en la velocidad de emisión o un aumento producido por la erosión del tamaño de los conductos. Sin embargo, es de gran importancia conocer también cómo se desarrolla un flujo horizontal de piroclastos, cuyo modelo se presenta en la Figura. que en primera aproximación consideraremos con densidad constante.

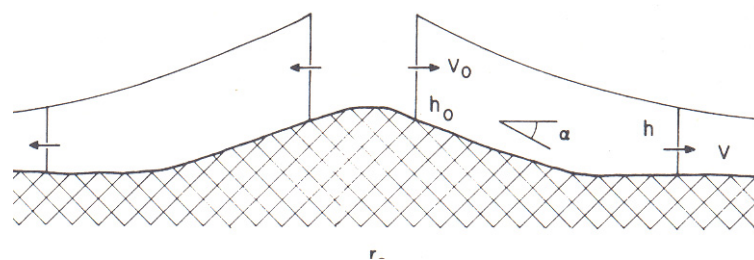


Figura 4. Modelo de desarrollo horizontal , Araña y Ortiz, 1984.

En estas condiciones la ecuación resultante es:

$$\rho v \frac{dv}{dr} = (g \sin \alpha)(\rho - \rho_0) - cv^2$$

Siendo ρ_0 la densidad del aire, ρ la densidad de la masa de piroclastos, c un coeficiente de rozamiento, v la velocidad horizontal y α la pendiente del terreno.

Cuando la erupción es de carácter explosivo el material se proyecta verticalmente, pero también horizontalmente. Estos transportes horizontales de material piroclástico dentro de una masa de gases se suele también denominar *nube ardiente*.

Las coladas piroclásticas se inician a partir de explosiones en la boca del volcán mientras que la denominadas *oleadas piroclásticas*, se originan cuando además existe un componente hidromagmático cuya importancia dependerá de la proporción de agua que intervenga en relación al magma. En las coladas piroclásticas más características, la nube inicial es muy densa por lo que el transporte es fundamentalmente gravitacional, siguiendo la topografía del terreno.

En cuanto a los depósitos que producen ambos tipos de formaciones piroclásticas es fácil distinguir ya que los originados por coladas u oleadas manifiestan claramente mediante los rasgos puramente magmáticos para las primeras y freatomagmáticos para las oleadas. Los depósitos de las coladas tienen en general una gran potencia mientras que los correspondientes a las oleadas se caracterizan por una estratificación correspondiente a las distintas explosiones generadas. Los depósitos cubren la topografía pero tienen tendencia a acumularse en los valles. Se caracterizan por presentar estructuras sedimentarias unidireccionales. Así las capas de los niveles inferiores tienen mayor carga de líticos y abundan en las proximidades del origen de la explosión mientras que en las superiores el grano es más fino y aparecen a mayores distancias.

20.4 Mapas de peligrosidad

Los peligros asociados a las erupciones volcánicas son diferentes para cada volcán debido a sus especiales características. En general los peligros se pueden enumerar de la siguiente forma:

- Flujos de lava.
- Bombas volcánicas
- Lluvia de cenizas.
- Flujos piroclásticos.
- Lahares.
- Avalanchas de rocas y tierras.
- Terremotos.
- Deformación del terreno.
- Tsunamis.
- Gases y aerosoles

Aunque siempre se ha considerado como el peligro más importante el producido por la efusión de lavas, veremos que no es este el que más daño ha producido en la mayor parte de las erupciones que han afectado al hombre. Las características de estos fenómenos son:

- **Flujos de lava.**
 - Temperaturas entre 750-1100° C
 - Enfriamiento superficial.
 - La tasa de fluencia depende de la viscosidad (pocos m/h hasta 60 km/h).
 - Espesores de 1-2 m a decenas de m.
 - Capacidad de flotación de objetos sobre ella.
 - Seguimiento de depresiones topográficas.
 - Tubos de lava la transmiten a gran distancia.
- **Proyección de bombas.**
 - área circular de 3-5 Km. alrededor de la boca.
 - proyectiles de más de 100 mm viajan a alta velocidad.
 - gran energía de impacto.
- **Caída de cenizas.**
 - Columnas eruptivas de hasta 55 Km. de altura.
 - Velocidad de viento de transporte entre 10 - 100 km/h.
 - Gran influencia de la velocidad del viento.
 - Extensión posible de más de 1000 Km.
 - Afecta la visibilidad.

- Lluvia afecta su estabilidad.
- Altamente conductiva en contacto con agua.
- **Flujos piroclásticos.**
 - La mayor parte del gas es vapor de agua y granos 60μ.
 - Velocidades de hasta 160 m/s.
 - Temperaturas de hasta 900° C.
 - Distancias de hasta 10 Km. con salto de obstáculos de hasta 1000 m.
- **Lahares.**
 - Generados en áreas de depósitos no consolidados de espesor >150 mm, sobre pendientes de terreno, con precipitación de >10 mm/h o donde los depósitos obstruyen los canales.
 - Pueden ser calientes o fríos.
 - Producen grandes áreas de sedimentación.
 - Puede ser muy erosivo dependiendo del volumen.
 - Sigue las depresiones topográficas.
 - Viajan distancias de decenas de kilómetros.
- **Avalanchas.**
 - Generadas por colapso de parte del volcán.
 - Volúmenes >10-20 millones de metros cúbicos.
 - Distancias de recorrido de 20-30 Km.
 - Pueden ser calientes o frías.
 - Velocidades de hasta 100 m/s.
 - Pueden crear nueva topografía, rellenar lagos, etc.
- **Terremotos.**
 - Foco a profundidad muy superficial.
 - Magnitud de hasta 6.5
 - Daños dependientes del tipo de suelo.
- **Deformación.**
 - Inclinaciones de decenas de metros
 - áreas de daños proximas a 15-20 Km de radio.
- **Tsunamis.**
- **Ondasónicas.**
- **Inducción de rayos.**
- **Gases y aerosoles.**

Dependiendo en cada caso cual es el peligro que se quiere determinar su distribución en la zona, se modeliza el fenómeno correspondiente obteniéndose así diferentes mapas.

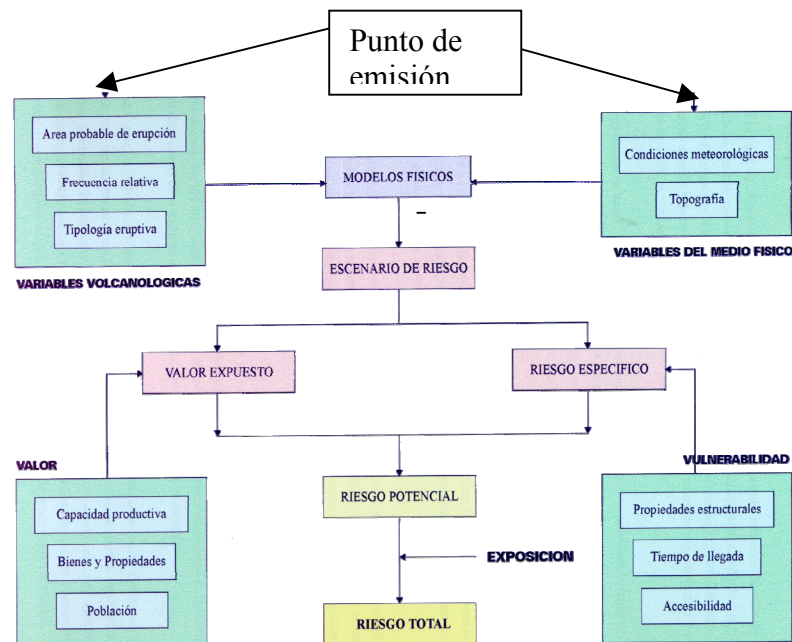


Figura 5. Modelo conceptual de datos, Gomez,1996.

Según uno de los últimos estudios realizado por Felpeto para las islas Canarias considera como peligros fundamentales la caída de cenizas y las coladas de lava. En otros lugares como por ejemplo los volcanes de Centroamérica el debido a lahares se considera de la mayor importancia. Para la caída de cenizas se considera un modelo basado en una erupción pliniana mientras que para la colada de lava utiliza dos modelos uno determinista y otro probabilista que determina la probabilidad de cada punto del área de estudio de ser cubierto por la lava. Para la caída de cenizas es necesario conocer también datos sobre la estadística de vientos dominantes en la zona de estudio.

A continuación presentaremos algunos mapas de peligrosidad de algunos de los peligros volcánicos que pueden existir determinados por diferentes autores para diferentes partes del mundo.

En primer lugar presentamos el mapa de peligro de flujos piroclásticos y lahares para el volcan Tungurahua en Ecuador realizado por Hall y otros (1999) y expresado mediante flechas las trayectorias de los flujos piroclásticos recogidas en época histórica y los valores 1, 2 y 3 representan alta peligrosidad para flujos piroclásticos en toda el área y lahares en los valles,

peligrosidad baja en flujos piroclásticos y alta peligrosidad para lahares en el valle de Pastaza respectivamente.

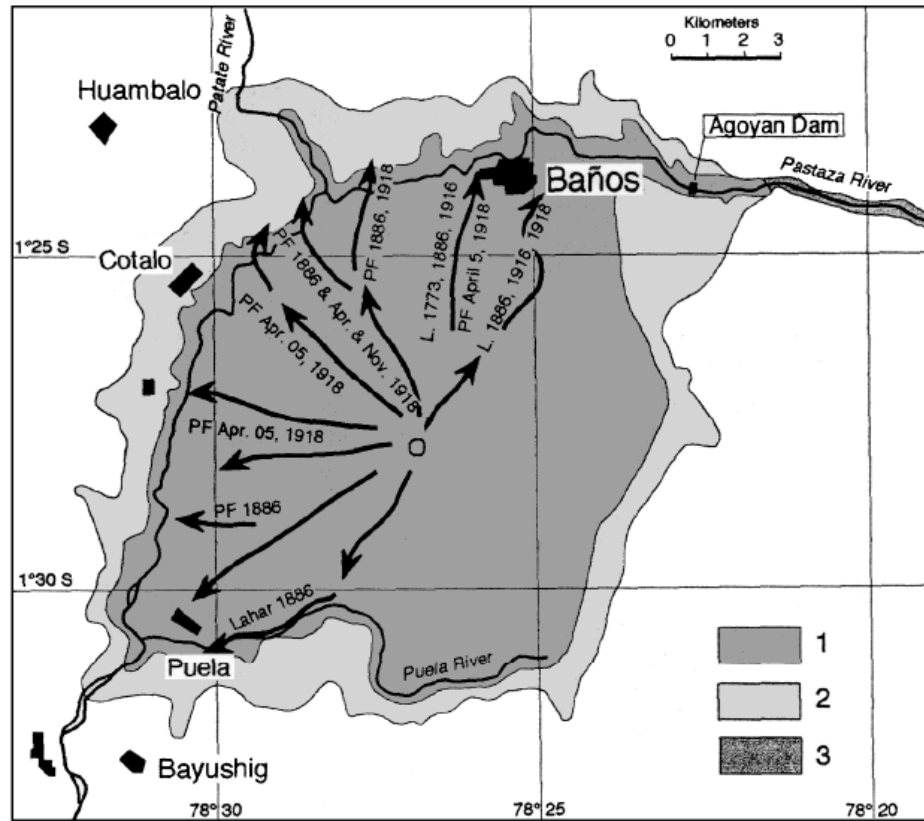


Figura 6. Mapa de peligrosidad de flujos piroclásticos y lahares en el volcán Tungurahua, Hall et al 1999.

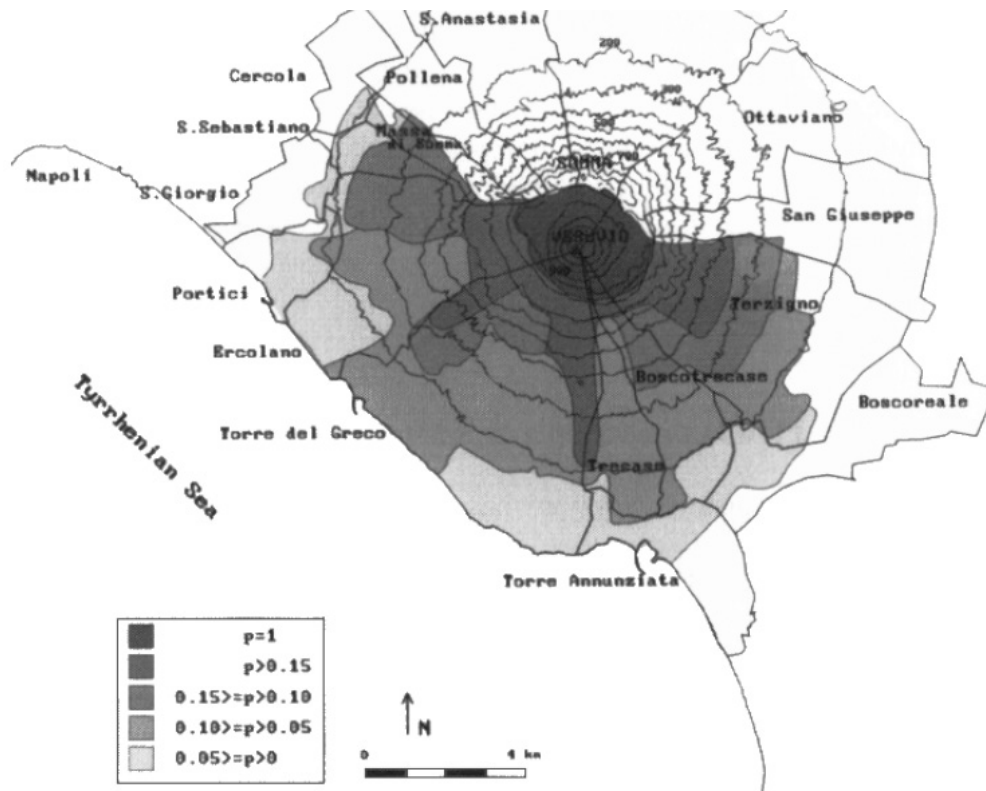


Figura 7. Mapa de peligrosidad de flujo de lava en el Vesubio, Lirer et al., 1998.

Bibliografía

- [1] V. Araña y R. Ortiz. Volcanología. (1984). Ed. Rueda. 2ª parte.
- [2] E. Ancochea y J. L. Barrera. (2002) Peligrosidad Volcanica. En Riesgos Naturales Coord. F. Ayala-Carcedo y J. Olcina. Ed. Ariel. Cap. 12.

Tema 21. Campo magnético terrestre. Componentes y división según su origen. Campo magnético de un dipolo. Dipolo terrestre. Variación secular. Origen del campo magnético interno.

21.1. Campo magnético terrestre

Una de las propiedades de la Tierra es que tiene un campo magnético asociado a ella- el campo geomagnético. Este campo varía con el tiempo y la posición de una manera complicada. El geomagnetismo estudia la forma detallada y el comportamiento del campo magnético terrestre y su relación con otros fenómenos geofísicos.

La tierra está formada por un núcleo líquido metálico con una parte interna sólida, un manto sólido y una corteza, una atmósfera aislante, y un plasma envolvente llamada la magnetosfera. Todos ellos, exceptuando la atmósfera, contribuyen o influyen en el campo geomagnético. Además muchas rocas tienen una “memoria magnética” que dura miles de millones de años. Por todo ello no es de extrañar que el estudio del campo geomagnético sea capaz de revelar una considerable cantidad de información acerca de la estructura e historia de la Tierra. Dentro de las más recientes contribuciones a la física de la Tierra a partir del geomagnetismo es la evidencia, a partir de la presencia de un campo magnético y su variación de largo periodo, del movimiento del núcleo; la teoría de la expansión del suelo oceánico “*sea floor spreading*” basado en anomalías magnéticas lineales en los océanos y reforzada por las direcciones paleomagnéticas; la complicada interacción entre la Tierra y la radiación de partículas procedentes del Sol a partir de las tormentas magnéticas. También el geomagnetismo contribuye de una forma más inmediata y práctica a la sociedad, como por ejemplo en la exploración minera, en la navegación aérea y en estudios arqueológicos.

Dado que los cuerpos no son directamente sensibles a la presencia de un campo magnético, la existencia del campo geomagnético no se conoció hasta finales del siglo XVI. Ahora bien la brújula puede haber sido ya usada en China antes del Siglo I a.c, y los imanes naturales eran conocidos por los griegos en épocas clásicas. De Magnesia, en Tesalia, viene el nombre de “magnet” (imán en inglés).

Sobre el nombre que se da a esta ciencia, hay que decir que antiguamente se utilizaba la denominación “magnetismo terrestre” siendo en la actualidad más utilizado el término “geomagnetismo” siguiendo la líneas de geodesia, geomorfología, geofísica,..

21.2. Componentes y división según su origen

21.2.1. Componentes

Todo campo magnético es una magnitud vectorial, es decir, en todo punto tiene una dirección definida y suele representarse por la letra **B** (en todo el texto se utilizará la negrita para representar magnitudes vectoriales). Por lo tanto se necesitan tres cantidades para especificarlo. A lo largo de los años se ha desarrollado una nomenclatura específica para las componentes del campo geomagnético, normalmente denominadas “elementos del campo geomagnético”. El sistema de referencia utilizado es un sistema cartesiano de coordenadas XYZ en dirección Norte Este y Nadir, o sistema geográfico, en el que el Norte es el Norte geográfico. Otros elementos utilizados frecuentemente son:

D- declinación; ángulo entre el norte geográfico y la componente horizontal del campo, contado desde el norte hacia el este, es decir el ángulo que forma el norte geográfico y el Norte magnético.

I- inclinación; el ángulo que forman la dirección del campo y la horizontal, contado hacía abajo desde la horizontal.

H- intensidad horizontal; magnitud de la intensidad horizontal (proyección horizontal del campo).

F-intensidad total; magnitud del vector campo.

Diferentes combinaciones de estos elementos son utilizados para definir el campo, la combinación particular está determinada por el tipo de observaciones realizadas. X, Y y Z se utiliza fundamentalmente para estudios teóricos, H, D y Z son los que normalmente se registran en la práctica.

Las siguientes relaciones entre los distintos elementos del campo geomagnético se deducen por simple trigonometría.

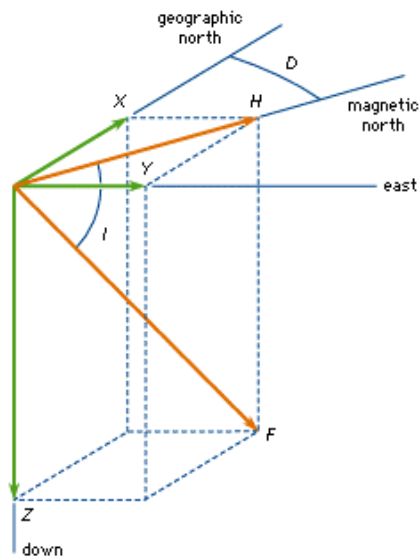


Figura 1. Componentes del campo magnético terrestre. Fuente: Enciclopedia Britanica

- Intensidad del campo

$$F = \sqrt{H^2 + Z^2} = \sqrt{X^2 + Y^2 + Z^2}$$

- Intensidad horizontal

$$H = \sqrt{X^2 + Y^2}$$

- Declinación

$$D = \tan^{-1} Y/X$$

- Inclinación

$$I = \tan^{-1} Z/H$$

21.2.2 División según su origen

El campo como un todo varía de una forma complicada tanto en el tiempo como en el espacio, pero en general la variación temporal varía suavemente con la posición y los cambios espaciales tienen poca dependencia temporal. Esto se traduce en que $\partial \mathbf{B} / \partial x$ y $\partial \mathbf{B} / \partial t$ pueden ser grandes, pero $\partial^2 \mathbf{B} / \partial x \partial t$ suele ser pequeño.

A partir de registros del campo magnético en diferentes puntos de la Tierra y en diferentes épocas, se observa por la forma de su variación, que el campo que medimos en la superficie de la tierra es la suma vectorial de diferentes campos constituyentes, cada uno de ellos generado en diferentes regiones y varían de manera distinta tanto en el espacio como en el tiempo. Un análisis profundo es capaz de separar cada uno de estos campos en dos partes: uno creado por fuentes en el interior de la Tierra y el otro debido a fuentes externas a ella. El interno a su vez está formado fundamentalmente por dos constituyentes: uno global de gran amplitud, cuya fuente debe estar situada en el interior de la tierra y otro, de amplitud mucho menor y que debe ser creado por fuentes mucho más superficiales. Conviene tratar a estos como dos campos diferentes, y se les suele llamar "campo fundamental" al de origen interno y "campo local" al de origen superficial, superpuesto a estos, está el "campo externo", producido por corrientes eléctricas que circulan por la Ionosfera y Magnetosfera, que al variar en el tiempo generan campos inducidos, que se suman al "campo anómalo o cortical", originado por rocas magnetizadas, bien inducida, bien remanente, para crear lo que hemos llamado campo local.

Un resumen de los diferentes campos constituyentes del campo geomagnético, mostrando su origen, intensidad media, morfología variación temporal y métodos de observación, aparece en la siguiente tabla:

Campo	Localización	Intensidad	Morfología	Variación temporal
Fundamental	Núcleo externo	50000 nT	Principalmente bipolar	Variación secular (1000 a)
Local	Corteza	100 nT (10^5 nT)	Muy irregular	Ninguna
Tormentas regulares	Magnetosfera	150 nT (500 nT)	Casi uniforme	4 a 10 horas
Tormentas irregulares	Ionosfera y magnetosfera	100 nT	Global. Mayor en zonas aurorales	5 a 100 minutos
Variación diaria	Ionosfera	50 nT	Global	Periódica 24, 12 y 8 horas.
Pulsaciones	Magnetosfera	Pocos nT	Cuasi-global. Mayor en zonas aurorales	Cuasi-periódica. 1-300 s.
Campos inducidos	Corteza	Aprox. 1/2 de los 4 campos anteriores	Global pero irregular	Igual que los 4 campos anteriores

21.3. Campo magnético de un dipolo

El campo magnético terrestre se puede aproximar, en un 90 por 100, por el que produce un dipolo situado en su centro, inclinado $11,5^\circ$ con respecto a la, dirección del eje de rotación. El polo “negativo” de este dipolo está en la dirección del polo Norte geomagnético.

Un campo magnético dipolar se forma por una corriente eléctrica I , que circula por una espira de área S y su momento viene dado por:

$$m = IS$$

Por analogía con el comportamiento de cargas eléctricas de distinto signo, un dipolo magnético puede considerarse como formado por dos monopolos de distinto signo e intensidad P , separados por una distancia s . Si el dipolo está en la dirección del eje X_3 , expresando el potencial escalar del campo de cada monopolo situado en los puntos $(0,0, s/2)$ y $(0,0, -s/2)$, por un desarrollo de Taylor con respecto al origen, tomando sólo el primer término se obtiene el potencial del dipolo en un punto a distancia r del origen:

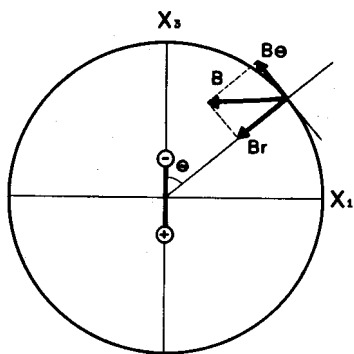
$$\Phi = \frac{m\mu_0}{4\pi} \frac{\partial}{\partial x_3} \left(\frac{1}{r} \right) = -Cm \frac{\cos \theta}{r^2}$$

donde $m = Ps$ es el momento magnético del dipolo, $C = \mu_0 / 4\pi$ y θ está medido desde el polo negativo del dipolo. En unidades SI, el momento magnético del dipolo m viene dado en amperios metro cuadrado (Am^2) (Cm puede expresarse en Tm^3 , ya que $\mu_0 = 4 \pi \times 10^{-7}$ henrios/m).

La inducción magnética B del campo del dipolo vendrá dada por el gradiente del potencial Φ . Tomando coordenadas esféricas (r, θ, λ) , si X_3 está en la dirección del eje del dipolo, las componentes de B son:

$$B_r = - \frac{\partial \Phi}{\partial r} = - 2Cm \cos \theta / r^3$$

$$B_\theta = (1/r) \frac{\partial \Phi}{\partial \theta} = - 2Cm \sin \theta / r^3$$



El ángulo θ es aquí el formado por el eje del dipolo y el radio vector r . La relación de estas componentes con X, Y, Z definidas anteriormente con relación a las direcciones geográficas es:

$$Z = -B_r$$

Figura 2. Campo magnético del dipolo

$$H = -B_\theta$$

El módulo de $\mathbf{B}$ o intensidad del campo F vendrá dado por:

$$|\mathbf{B}| = F = Cm (1 + 3 \cos^2 \theta)^{1/2} / r^3$$

La relación entre el ángulo de inclinación I en un punto y el ángulo θ viene dada por:

$$\operatorname{tg} I = Z/H = B_r / B_\theta = 2 \cotg \theta$$

Esta expresión es muy importante, pues relaciona la inclinación en un punto de la superficie con el ángulo que forma el radio en dicho punto con el eje del dipolo. El valor del campo para $\theta = \pi/2$, sobre la superficie de la Tierra ($r = a$), sólo tiene componente horizontal:

$$Z = -B_r = 0$$

$$H = -B_\theta = Cm/a^3 = B_0$$

denominándose B_0 la constante geomagnética.

Para $\theta = 0$ el campo es vertical y la componente horizontal es nula:

$$Z = -B_r = 2B_0$$

$$H = -B_\theta = 0$$

Si el eje X_3 no coincide con el del dipolo, las componentes horizontales del campo magnético X e Y pueden deducirse del potencial Φ en la forma:

$$X = -B_\theta = -\frac{1}{r} \frac{\partial \Phi}{\partial \theta}$$

$$Y = B_\lambda = -\frac{1}{r \sin \theta} \frac{\partial \Phi}{\partial \lambda}$$

Donde θ y λ son ahora las coordenadas geográficas.

Las líneas de fuerza del dipolo son simétricas con respecto al eje del mismo y a su plano normal y se extienden para valores r_0 cada vez mayores en el espacio. Aplicado al campo magnético terrestre, las líneas de fuerza representarían una imagen de dicho campo en la ausencia de influencias externas.

21.4. Dipolo terrestre

En primera aproximación el campo magnético terrestre puede aproximarse a un dipolo magnético, cuyo eje coincide con el eje de rotación y otro dipolo perpendicular a él, de modo que pueden combinarse para formar el campo de un dipolo inclinado respecto al eje de rotación. Los puntos donde dicho eje corta la superficie terrestre se denominan “polos geomagnéticos”,

Boreal	79° N ,	70° W
Austral	79° S ,	110° E

El dipolo está dirigido hacia el Sur, de tal modo en el hemisferio Norte cerca del polo Norte geográfico se ubica un polo Sur magnético y en el hemisferio Sur cerca del polo Sur geográfico se ubica un polo Norte magnético. Por convención se denomina el polo magnético ubicado cerca del polo Norte geográfico polo Norte magnético y el polo magnético situado cerca del polo Sur geográfico polo Sur magnético. El ángulo entre el eje de rotación y el del dipolo es aproximadamente 11,5°.

No hay que confundir los polos geomagnéticos con los polos magnéticos, que son los puntos en los que la media temporal del campo geomagnético es vertical.

Las coordenadas respecto al eje del dipolo se denominan coordenadas geomagnéticas definidas de la siguiente forma:

ϕ^* Latitud geomagnética, medida desde el ecuador geomagnético, definido como el plano normal al eje del dipolo. (No coincide con $I=0$)

λ^* Longitud geomagnética con origen en el meridiano geográfico que pasa por el polo geomagnético.

Por la inclinación del eje dipolar geomagnético con respecto al eje de rotación de la tierra los sistemas de coordenadas geográficas y geomagnéticas no coinciden.

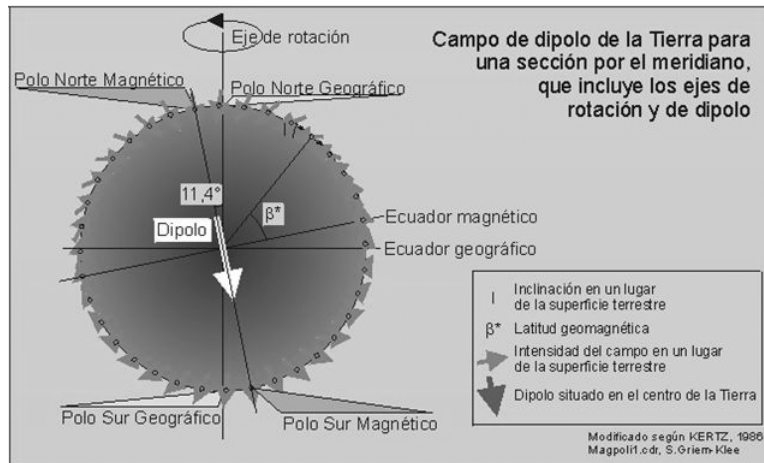


Figura 3. Campo de dipolo de la Tierra

Las coordenadas geomagnéticas para un lugar en la superficie terrestre se calculan de modo siguiente, conociendo las coordenadas geográficas del lugar en cuestión y de las intersecciones boreal o austral respectivamente.

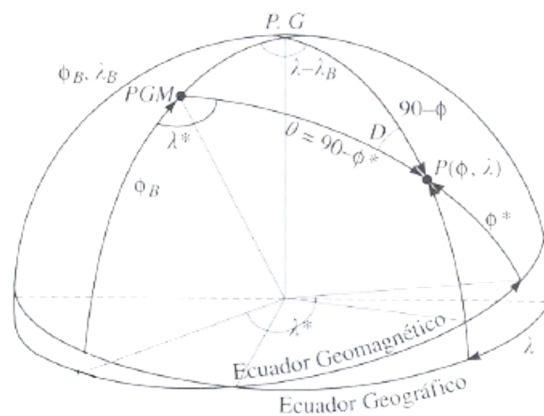


Figura 4. Coordenadas geomagnéticas

$$\begin{aligned} \sin \varphi^* &= \sin \varphi_B \sin \varphi + \cos \varphi_B \cos \varphi \cos(\lambda - \lambda_B) \\ \sin \lambda^* &= [\cos \varphi \sin(\lambda - \lambda_B)] / \cos \varphi^* \end{aligned}$$

Las componentes bipolares en función de las coordenadas geomagnéticas:

Declinación:

$$\sin D = [\cos \varphi_B \sin(\lambda - \lambda_B)] / \cos \varphi^* \quad \varphi^* = 90^\circ - \theta$$

Inclinación

$$\tan I = 2 \cotg \theta = 2 \tan \varphi^*$$

Intensidad vertical y horizontal

$$Z = B_0 \sin \varphi^* \quad H = B_0 \cos \varphi^*$$

B_0 es la constante geomagnética, aproximadamente 30951 nT.

21.5. Variación secular

Las variaciones temporales del campo magnético terrestre, de periodo tan largo que sólo se aprecian al comparar valores medios anuales durante varios años, reciben el nombre de variación secular. Gellibrand en 1634 se percató de que el valor de la declinación en Londres había ido cambiando de forma continuada desde 1580, en ese periodo había decrecido desde 11,3°E hasta 4,1° E.

Un fenómeno de la variación secular hace referencia a que la distribución del campo geomagnético se mueve lentamente hacia el Oeste. El promedio de avance es del orden de 0,18v de longitud por año. A esta velocidad, la distribución del campo daría la vuelta completa a la Tierra en unos 2.000 años. A diferencia de las tempestades magnéticas, que ocurren por causas externas, las anomalías a largo plazo y su marcha hacia el Oeste se deben a causas localizadas en el interior de la Tierra. Los cambios internos tienen lugar de modo muy lento y abarcan hasta millares de millones de años. En comparación, dos mil años es, pues, un tiempo muy corto. Este elemento constituye una de las claves fundamentales en el estudio del magnetismo terrestre.

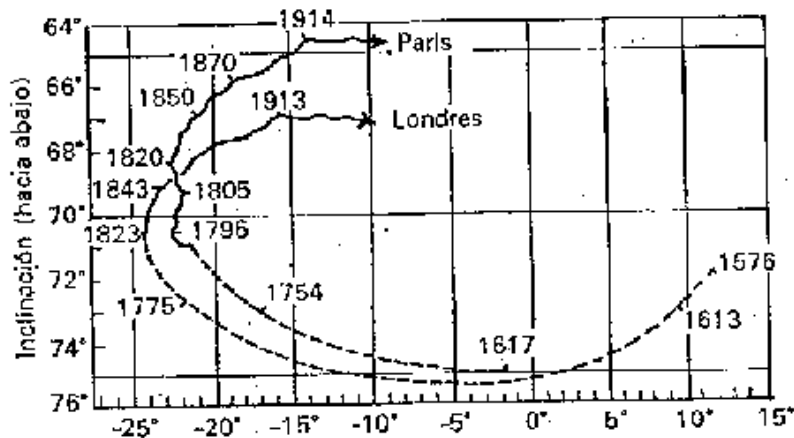


Figura 5. Variación secular

La causa de la variación secular está ciertamente relacionada con la causa del propio campo interno, pero presenta algunos aspectos propios, como la deriva hacia el oeste, que son interesantes de analizar por ellas mismas. La conductividad del núcleo es suficientemente elevada como para que el movimiento de las líneas de fuerza a través de él, deba ser lento, por tanto la variación secular debe ser más una cuestión de arreglar las líneas de fuerza que su destrucción o creación.

Una interpretación de la deriva hacia el oeste es que representa la diferencia en la velocidad angular entre el manto y el núcleo externo. Otra teoría es que en el núcleo se cree algún tipo de movimiento ondular hidromagnético y la deriva observada en la superficie este causada por una onda progresiva hacia el oeste. Tampoco se excluye otra teoría que supone grandes movimientos del núcleo.

21.6. Origen del campo magnético interno

Durante muchos años la explicación del origen de campo fundamental ha sido una de los problemas sin resolver de la geofísica, por lo que se fueron proponiendo diferentes teorías, que morían antes de estar del todo desarrolladas.

En su libro clásico “Die Magnete”, Gilbert (1600) describía la tierra como un gran imán. Una esfera uniformemente magnetizada crea un campo externo en la forma de un dipolo. La magnetización requerida para crear un campo de la intensidad del campo fundamental, sin tener en cuenta la magnetización inducida, sería de 80Am^{-1} . Algunas rocas se han encontrado con magnetizaciones tan altas como ésta, la magnetita, pero entonces el porcentaje de este mineral en la tierra debería de ser del 50%. Pero las rocas pueden estar permanentemente magnetizadas

por debajo de la temperatura de Curie y esta se alcanza a una profundidad del orden de los 30 km; por lo que para explicar el campo fundamental habría que aumentar la magnetización de las rocas hasta 5650 A m^{-1} , cifra que hizo descartar esta teoría en cuanto se supo que la mayor parte del interior de la tierra está a altas temperaturas.

Al no poder explicar el campo magnético interno mediante magnetización de rocas surgió otra teoría basada en corrientes eléctricas generadas por algún tipo de dinamo. El problema era encontrar el mecanismo que mantuviera las corrientes a lo largo de la historia.

Larmor (1857-1942) fue el primero en proponer un efecto de dinamo autoexcitada como origen del campo magnético. Basándose en la existencia de un núcleo terrestre fluido y de alta densidad, compuesto casi en su totalidad de hierro, la circulación de este material conductor en presencia de un campo magnético genera corrientes eléctricas que a su vez realimentan el campo inductor.

DINAMO AUTOINDUCIDA

Dinamo de disco homopolar

Fue propuesta por Bullard en 1955. Este modelo está constituido por un disco conductor que gira en presencia de un campo magnético B_0 , perpendicular a él. En el disco se forman corrientes eléctricas que van del eje al borde del disco. Si se coloca un hilo conductor en el borde del disco se produce una corriente I que va a lo largo de la espira hasta el eje del disco. Esta corriente, al circular por la espira del hilo, genera un campo magnético B que realimenta el campo inductor B_0 . Si la velocidad angular Ω es suficientemente grande, el campo generado B es suficiente para mantener el proceso y el inductor B_0 acaba por no ser necesario; la dinamo es autoinducida.

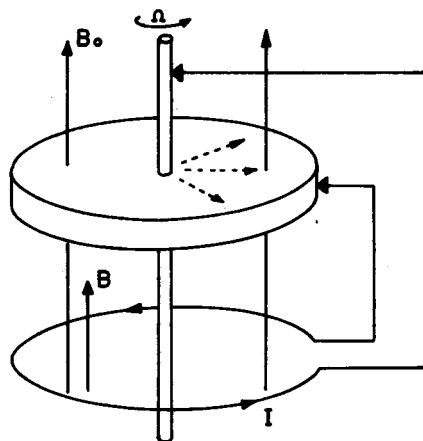


Figura 6. Modelo simplificado de la dinamo de disco homopolar

Para un modelo de dimensiones de laboratorio, con un disco de 10 cm de radio y un circuito de 10 ohmios de resistencia, la velocidad angular necesaria es del orden de 10^7 revoluciones por segundo. Como se ve, para estas dimensiones es necesaria una velocidad angular muy alta para producir el fenómeno de autoexcitación. Para la Tierra, sin embargo, debido a su tamaño, este efecto se obtendría con velocidades relativamente bajas.

Modelos de dinamos homogéneas.

Dinamo homogénea es aquella que resulta del movimiento de un fluido conductor homogéneo. Estos modelos son los más utilizados para explicar la generación del campo magnético terrestre y se basan en la interacción de movimientos de material conductor, con los correspondientes campos electromagnéticos. Dada la simetría esférica del medio, los campos pueden separarse en dos tipos, uno poloidal, con componente radial y otro toroidal, contenido en una superficie esférica.

Las ideas fundamentales del funcionamiento de la dinamo se pueden deducir del modelo simplificado de E. N. Parker. En este modelo, un cierto campo magnético inicial es modificado al entrar en interacción con un campo de velocidad de un fluido conductor. En el caso de la Tierra se distinguen dos efectos principales, en uno el campo magnético poloidal (campo del dipolo) interacciona con un campo de velocidad toroidal, para producir un campo magnético toroidal más complejo. A este proceso se le llama efecto ω . En el otro, la interacción de un campo magnético toroidal con un campo de velocidad poloidal produce de nuevo un campo magnético poloidal, proceso que se llama efecto α . Los modelos que combinan estos dos efectos se denominan $\alpha\omega$, como, por ejemplo, el del Bullard-Gellman. Este no es el único tipo de modelo posible, ya que se pueden combinar dos efectos para generar campos magnéticos toroidales y poloidales. Las dinamos resultantes se llaman de α^2 . Ejemplos de estos modelos son los presentados por P. H. Roberts, en 1971, y S. Childress, en 1978.

Para comprender de una manera cualitativa el proceso de un modelo de dinamo $\alpha\omega$, explicaremos brevemente el de Bullard y Gellman. En este modelo el núcleo fluido de la Tierra posee dos tipos de movimiento de material, uno en la dirección del radio, o campo poloidal P^V , y otro en forma anular en torno al eje de rotación o campo toroidal T^V . De momento no se especificará el mecanismo causante de estos dos tipos de flujo de material. La velocidad postulada para ambos movimientos es del mismo orden, aproximadamente 0,04 cm/s. Los modelos consideran dos campos magnéticos principales P_1^H , campo poloidal del dipolo, observado en la superficie de la Tierra y T_2^H un campo toroidal fuerte, contenido en la superficie del núcleo y causado por la interacción entre el campo P_1^H y el flujo de material T^V . La

intensidad de este campo se supone del orden de 5×10^{-2} T. El movimiento del material conductor, en presencia de campos magnéticos, produce campos eléctricos que a su vez generan otros campos magnéticos.

La secuencia simplificada de autoexcitación del campo geomagnético es la siguiente. El campo dipolar inicial P_1^H en interacción con las corrientes T^V , produce corrientes eléctricas P_2^E de carácter radial. Estas corrientes generan el fuerte campo magnético toroidal T_2^H confinado a la superficie del núcleo. Las corrientes de material P^V crean también un campo magnético toroidal, que junto con T_2^H induce en interacción con T^V y P^V campos eléctricos toroidales.

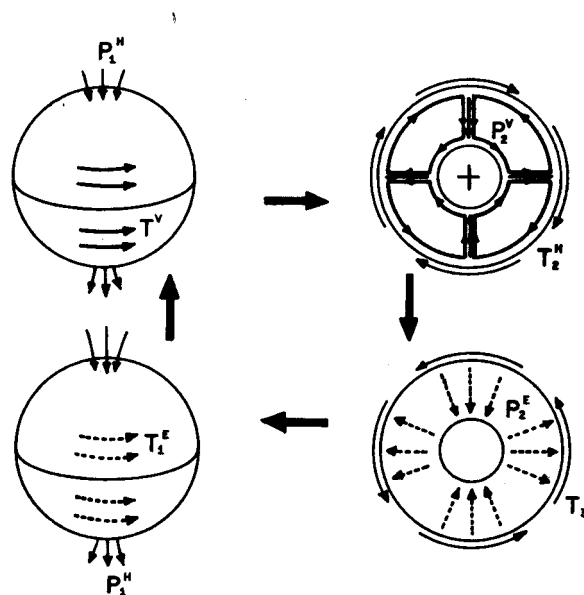


Figura 7. Esquema simplificado de la generación del campo geomagnético bipolar por el modelo de dinamo de Bullard-Gellman.

Estos a su vez generan un campo magnético poloidal que alimenta el campo inicial P^H . El proceso viene representado esquemáticamente en la figura 7, dejando sólo los cuatro pasos más importantes. Para empezar el proceso se necesita un campo magnético inicial de carácter poloidal cuya existencia se explica de diversas maneras.

El proceso básico es la realimentación del campo P_1^H a partir del campo T_2^H que a su vez es alimentado por la energía derivada del transporte de material conductor T^V y P^V . Las dos interacciones principales corresponden, por tanto, a los efectos ω y α . Las corrientes del material del núcleo pueden estar producidas por diversos mecanismos en los que está siempre la rotación de la Tierra. La influencia de la rotación en la producción de T es la causa de que el dipolo equivalente esté orientado en una dirección cercana a la del eje de giro. Para mantener una dinamo auto inducida, el proceso descrito debe ser regenerativo y, por tanto, la energía aportada por el transporte de material debe ser suficientemente grande.

ELEMENTOS DE MAGNETOHIDRODINAMICA

El problema complejo de la formulación matemática de los procesos de la dinamo es una aplicación de las ecuaciones de la magnetohidrodinámica. Fundamentalmente se trata del comportamiento de un fluido conductor en movimiento en un campo magnético. No vamos a desarrollar aquí las formulaciones completas que se han propuesto para los diversos modelos, sino solamente las ecuaciones fundamentales a partir de las cuales se pueden establecer ciertas propiedades de éstos. Si llamamos $\mathbf{E}$ al campo eléctrico; $\mathbf{B}$, la inducción magnética, y $\mathbf{v}$, la velocidad del fluido, la densidad de corriente $\mathbf{J}$ viene dada por la ley de Ohm en la forma:

$$\mathbf{J} = \sigma \mathbf{E} + \sigma \mathbf{v} \times \mathbf{B}$$

La ecuación del movimiento de un fluido incompresible y sin viscosidad, incluyendo la fuerza de Coriolis de la rotación de la Tierra (frecuencia angular Ω) es:

$$\rho \frac{d\mathbf{v}}{dt} = -\nabla p - 2\rho(\boldsymbol{\Omega} \times \mathbf{v}) + \mathbf{J} \times \mathbf{B} + \mathbf{F}$$

donde $\mathbf{J} \times \mathbf{B}$ es la fuerza de Lorentz, debida a la interacción de la corriente $\mathbf{J}$ con el campo magnético $\mathbf{B}$; $\mathbf{F}$, las fuerzas que dependen de elementos de volumen y p la presión.

A partir de las ecuaciones de Maxwell, siendo $\mathbf{D} = 0$ y tomando el rotacional en la ecuación resultante se obtiene:

$$\frac{1}{\mu} \nabla \times (\nabla \times \mathbf{B}) = \sigma (\nabla \times \mathbf{E}) + \sigma \nabla \times (\mathbf{v} \times \mathbf{B})$$

Sustituyendo el rotacional de $\mathbf{E}$ por su valor y poniendo los rotacionales en función del gradiente y la divergencia

$$\nabla \times \nabla \times \mathbf{B} = \nabla(\nabla \cdot \mathbf{B}) - \nabla^2 \mathbf{B} = -\nabla^2 \mathbf{B}$$

$$\nabla \times (\mathbf{v} \times \mathbf{B}) = (\mathbf{B} \cdot \nabla) \mathbf{v} - (\mathbf{v} \cdot \nabla) \mathbf{B}$$

obtenemos finalmente la ecuación:

$$-\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\sigma \mu} \nabla^2 \mathbf{B} + (\mathbf{B} \cdot \nabla) \mathbf{v} - (\mathbf{v} \cdot \nabla) \mathbf{B}$$

Si $\mathbf{v} = 0$; no hay movimiento en el fluido, la ecuación se reduce a:

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\sigma \mu} \nabla^2 \mathbf{B}$$

Esta es la ecuación de difusión, lo que indica que el campo magnético $\mathbf{B}$ decae exponencialmente con el tiempo en ausencia de transporte de material. En consecuencia, para mantener el campo magnético $\mathbf{B}$ se necesita la energía derivada del flujo de fluido de velocidad $\mathbf{v}$.

El término $(\mathbf{v} \cdot \nabla) \mathbf{B}$ representa la convección del campo $\mathbf{B}$ con el flujo de material $\mathbf{v}$. El término $(\mathbf{v} \cdot \nabla) \mathbf{B}$, el gradiente del flujo en la dirección del campo magnético $\mathbf{B}$. Este último término es el responsable del alargamiento de las líneas de fuerza de $\mathbf{B}$ en la dirección del flujo $\mathbf{v}$, que resulta en un aumento de la intensidad del campo magnético, a partir de la energía cinética del flujo $\mathbf{v}$. De forma cualitativa tenemos aquí una posibilidad de realimentar el campo magnético a partir de las corrientes de material.

Un parámetro importante en esta discusión es R , el número magnético de Reynold:

$$R = \mu \sigma v L$$

Donde σ es la conductividad del medio; μ , la permeabilidad magnética; v , la velocidad del fluido, y L , una longitud característica, por ejemplo, el radio del núcleo, o el espesor del núcleo externo. Este número debe ser mayor de 10 para que la dinamo tenga una acción regeneradora, lo que se logra si L es constante, aumentando σ o v . Si se toma la velocidad de la deriva hacia el oeste del campo no dipolar, como una indicación del orden de magnitud de la velocidad del material en el núcleo, teniendo en cuenta la diferencia de radio entre la superficie de la Tierra y del núcleo, ésta será del orden de 10^{-2} cm/s. Como μ en el núcleo es prácticamente igual a μ_0 , el valor de la conductividad en el núcleo debe ser del orden de 10^5 S/m.

Debido a la rotación de la Tierra, en la generación del campo magnético es necesario introducir la fuerza de Coriolis. La rotación de la Tierra causa además un acoplamiento entre los campos magnéticos, poloidales y toroidales y su inclusión complica las soluciones de la dinamo y plantea problemas a su estabilidad.

Un tipo de solución al problema de la dinamo autoinducida que incluye oscilaciones de amplitud finita en el campo magnético resultante se ha propuesto para explicar la disminución

de la magnitud del dipolo, observada en la variación secular, y los cambios de polaridad o inversiones observados en el paleomagnetismo.

Algunos autores, como Rikitake y Gellman, han tratado matemáticamente el caso de dos dinamos de disco acopladas, cuya resolución presenta oscilaciones en amplitud, con inversiones periódicas de polaridad. La aplicación a la Tierra de estos modelos supone comparar cada una de las células de convección en el núcleo con una dinamo de disco y el sistema total, como un acoplamiento de estas dinamos. El resultado es un sistema que admite oscilaciones de amplitud y cambios de polaridad, como los observados en la Tierra a lo largo de las eras geológicas.

Dinamos cinemáticas y dinámicas-Las ecuaciones básicas que se han presentado para los procesos que dan origen a la dinamo auto inducida pueden resolverse introduciendo una gran variedad de modelos. Estos pueden dividirse en dos tipos, modelos cinemáticos, que asumen una velocidad de transporte de material v sin explicar su origen, y los dinámicos, que proponen modelos que explican el mecanismo de este flujo. En todos los casos se supone que los procesos tienen lugar en la región del núcleo externo, entre 1300 y 3500 km a partir del centro de la Tierra. Esta región se considera formada por un fluido homogéneo conductor en el que se produce un flujo de velocidad v . La región del manto que rodea el núcleo está formada por material prácticamente aislante. Las soluciones que se buscan son aquellas que expliquen la existencia de un campo magnético B resultante, que observado sobre la superficie de la Tierra tiene la forma del campo geomagnético. Este campo no debe amortiguarse exponencialmente con el tiempo, sino mantenerse, o en todo caso oscilar en intensidad y reproducir las características de la variación secular. Un punto importante en el modelo Elsasser, Bullard y Gellman es la existencia de corrientes regulares toroidales en el campo de velocidad de transporte de material. Existen serias objeciones a esta hipótesis, por lo que se han propuesto otros modelos más complejos. En algunos de estos modelos se postula la existencia de corrientes de turbulencia en el núcleo, afectadas por la fuerza de Coriolis, derivada de la rotación de la Tierra, a partir de la cual adquieren una cierta naturaleza helicoidal. A este tipo de modelos pertenecen los propuestos por Parker, en 1955, y, más tarde, por Steenbeck, Krause, Lortz y otros, en 1969. Actualmente, dentro de las teorías de dinamos hidromagnéticas hay que considerar los modelos propuestos por Busse, en 1976, con flujo en forma de «rollos» de convección paralelos al eje de rotación y los que implican fenómenos de turbulencia, como los propuestos en primer lugar por Braginsky y desarrollados más tarde por Gubbins y Moffat.

La fuerza de Coriolis y la de Lorentz suponen que existe ya una velocidad v . Para explicar su origen se supone que este movimiento de material está formado por la convección térmica, debida a la existencia de un gradiente de temperatura entre el exterior y el interior del núcleo y

un flujo de calor del núcleo al manto. Para conseguir velocidades del orden de 0,01 cm/s bastaría un pequeño gradiente de no más de un 1 a un 10 por 100 y un flujo térmico resultante del orden 10^{12} W para todo el núcleo. Para mantener este flujo es necesaria una fuente térmica, que en general se supone de naturaleza radiactiva, aunque este proceso crea dificultades a la hora de establecer qué elementos lo producen. Algunos autores suponen la presencia de una proporción de isótopos de potasio en el núcleo, responsables de la generación de calor por procesos radiactivos.

Una fuente de energía no térmica propuesta por Stewartson, Roberts y Malkus, en 1963, es la derivada del movimiento de precesión de la Tierra. La velocidad de precesión depende de la elipticidad dinámica. Si la elipticidad del núcleo es diferente de la de la Tierra existirá una diferencia en la velocidad de precesión entre el núcleo y el manto. Este proceso llevaría a movimientos en la superficie del núcleo con velocidades del orden de 1 cm/s, mayores que las postuladas en los modelos de dinamos homogéneas. Otros autores han propuesto también un mecanismo de movimiento de material en el núcleo de origen gravitacional, con una continua deposición del material más pesado hacia su interior.

Desde el punto de vista energético, la energía total del campo magnético terrestre es del orden de 10^{23} J. Sin una continua realimentación, este campo se disiparía con una constante de tiempo del orden de quince mil años. Si el calor perdido por el núcleo es de 10^{12} J/s, con una eficiencia del 10 por 100, el proceso convectivo dispondría de suficiente energía para servir de motor a la dinamo. Por otro lado, un aporte de energía de este orden de magnitud a partir de procesos originados por la rotación de la Tierra, no implica ninguna dificultad. Algunos autores, como Stacey, concluyen que los procesos en el núcleo, generados por el movimiento de precesión de la Tierra, proporcionan el mecanismo básico para la dinamo geomagnética, aunque la convección térmica tome parte también en él.

TEORIAS DE LA VARIACION SECULAR

Hemos dividido el efecto de la variación secular del campo magnético terrestre en una variación que afecta al campo del dipolo, y otra que afecta al campo no dipolar. La variación del campo dipolar está asociada a los procesos de la dinamo del núcleo terrestre. Las variaciones del campo no-dipolar se pueden dividir en dos partes: una deriva hacia el oeste y una variación temporal de las características de este campo. Las primeras explicaciones de estas variaciones pusieron su origen en efectos de magnetización y sus cambios en la corteza terrestre. Después de los estudios ya mencionados de Elsasser, Bullard y Gellman se admite que estos cambios deben tener también su origen en los procesos del núcleo terrestre. La deriva al oeste del campo no-dipolar, de acuerdo con Bullard, se explica por un acoplamiento viscoso y electromagnético

entre el manto y la superficie exterior del núcleo. Debido a las corrientes de convección en el núcleo, éste gira más rápidamente en su interior que en la superficie. Se da, por tanto, en la superficie del núcleo una velocidad que es a su vez menor que la velocidad en el interior del núcleo, y también menor que la existente en el manto. Aunque esta diferencia es muy pequeña, esta zona de velocidad angular menor es suficiente para que se cree un campo toroidal, debido a la diferente conductividad entre la capa externa del núcleo y la capa interna del manto. Este mecanismo es el que se supone responsable de la deriva hacia el oeste del campo no-di polar (Fig. 8). Otros autores, como R. Hyde y S. I. Braginsky, suponen que esta deriva del campo no-dipolar es un fenómeno de naturaleza ondulatoria con periodos del orden de millones de años. Esta interpretación exige que el campo magnético toroidal confinado a la superficie del núcleo sea muy intenso. Esta explicación hace innecesaria la existencia de la capa de velocidad angular más lenta en la parte exterior del núcleo.

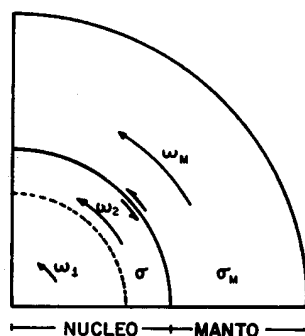


Figura 8. Explicación de la deriva hacia el oeste del campo no bipolar.

Las variaciones temporales del campo no-dipolar (variaciones de su estructura) se deben, de acuerdo con la mayoría de los autores, a las corrientes de convección en la superficie del núcleo. Estas corrientes ascendentes o descendentes se distribuyen sobre la superficie del núcleo y su variación y la interacción de estas variaciones de flujo con las líneas de fuerza del campo magnético darían origen a los cambios del campo residual.

Bibliografía

- [1] Miguel, L. de (1980): *Geomagnetismo*. I.G.N. Madrid
- [2] Parkinson, W.D. (1983): *Introduction to Geomagnetism*. Elsevier, Amsterdam.
- [3] Udías, A. y J.Mézcua(1986): *Fundamentos de Geofísica*. Alhambra Universidad. Madrid.

Tema 22. Campo magnético externo. Variaciones del campo magnético externo. Composición de la ionosfera. Estructura de la magnetosfera. Anillos de radiaciones y Auroras

22.1. Campo magnético externo

El campo magnético terrestre es principalmente de origen interno, resultado de fenómenos ocurridos en el núcleo y en la corteza terrestre por debajo del punto de Curie. Sin embargo, existe también un campo de origen externo debido a procesos localizados fundamentalmente en la ionosfera y la magnetosfera, causados principalmente por la actividad del Sol.

El campo magnético externo se caracteriza por sus variaciones en el tiempo, de periodos cortos y fácilmente observables, mientras que las variaciones temporales del campo interno sólo son apreciables en periodos largos de tiempo.

Las variaciones del campo magnético externo producen, en algunas ocasiones, campos inducidos en la corteza, el manto y los océanos, siendo, la mayor parte de estas perturbaciones magnéticas, consecuencia de la entrada en la atmósfera terrestre de partículas procedentes del Sol.

Se denomina “viento solar” a la emisión continua de materia por el Sol. Esta nube de plasma alcanza velocidades supersónicas de aproximadamente 400km/s y transporta un campo magnético de alrededor 5nT, denominado campo magnético interplanetario. Las partículas cargadas llegan a la Tierra entre 20 y 40 horas después de su emisión y al alcanzar la parte superior de la atmósfera terrestre origina, entre otros fenómenos, modificaciones en el campo magnético terrestre.

22.2. Variaciones del campo magnético externo

Las variaciones del campo magnético externo pueden ser periódicas o irregulares.

22.2.1. Variaciones periódicas

Son debidas a procesos que ocurren de manera periódica en el Sol y la Luna, como son los periodos de las órbitas de estos astros, los periodos de la rotación de la Tierra y de la rotación del Sol o los periodos asociados a las manchas solares entre otros. Se aprecian mejor en los días

tranquilos, cuando las perturbaciones de la actividad solar son pequeñas. Las más importantes son la “variación solar diurna” y la “variación lunar”.

- **Variación solar diurna**

El análisis de los magnetogramas en días tranquilos muestra una variación periódica del campo magnético terrestre, de unas pocas nanoteslas que depende del tiempo solar local. Esta variación se conoce con el nombre de variación solar tranquila o variación diaria y se representa por la letra S_q (solar quiet).

Su periodo predominante es de 1 día solar, con un máximo positivo o negativo a las 12 horas solares. Pero también tiene periodicidades de 8, 12 y 24 horas, de 13.5 y 27 días, de 6 y 12 meses y periodos semianuales, anuales y uno de 22 años asociado al ciclo de actividad de las manchas solares.

La amplitud del máximo es de aproximadamente de 100nT en la componente horizontal, H y de 20nT en la componente vertical, Z. Las variaciones en el ángulo de declinación D, oscilan de 2 a 4 minutos de arco.

Pero S_q también depende de la época del año y de la latitud.

Para analizar la variación estacional, los meses del año se agrupan en los denominados “meses de Lloyd”:

- Meses D: meses de solsticio de invierno en el Hemisferio Norte (noviembre, diciembre, enero y febrero).
- Meses E: meses de equinoccio (marzo, abril, septiembre y octubre).
- Meses J: meses de solsticio de verano en el Hemisferio Norte (mayo, junio, julio y agosto).

Las componentes H y Z presentan variaciones máximas en los meses E mientras que la componente D es máxima en los meses de verano (meses J en el hemisferio norte y meses D en el hemisferio sur)

En cuanto a la variación con la latitud magnética, se observa que la variación de:

- Z es nula en el ecuador magnético, se hace máxima para latitudes entre los 5 y los 10° y desde ahí disminuye al aumentar la latitud.
- D es nula en el ecuador magnético y máxima en altas latitudes.

- H es máxima en el ecuador magnético y mínima cerca de los 30° para ambos hemisferios magnéticos.

Las características de las componentes X , Y , Z son:

- las tres componentes muestran variación nula para todas las latitudes durante la noche.
- X tiene un máximo cerca del mediodía para bajas latitudes que desaparece al aumentar la latitud y pasa a ser un mínimo bien definido en altas latitudes.
- Y se caracteriza porque en el hemisferio norte tiene un máximo en el amanecer y un mínimo por la tarde (tanto mayores cuanto mayor es la latitud). En el hemisferio sur, la secuencia de los máximos y mínimos es la contraria.
- Z presenta un mínimo al mediodía en ambos hemisferios.

• Variación lunar

La influencia de la Luna en las variaciones del campo magnético terrestre se denomina variación lunar y se representa con la letra L . Se caracteriza por su pequeña amplitud (aproximadamente la veinteava parte de la producida por el Sol), por lo que no se aprecia directamente en los magnetogramas, sino que debe obtenerse por medio de un análisis armónico de varios meses de observación.

L está referida al tiempo lunar local y su periodo predominante es semidiurno. Se características principales son:

- H presenta un mínimo a las 0 y 12 horas y un máximo a las 6 y 18 horas. En el ecuador magnético su variación es máxima. La máxima amplitud de H sólo alcanza los 20nT.
- la componente Z es aproximadamente cero en el ecuador magnético y máxima en latitudes bajas, disminuyendo de nuevo para latitudes altas. Su variación aumenta con la latitud.
- la variación de las tres componentes, D , H y Z es máxima en los meses de verano en los dos hemisferios. Pero las amplitudes son mayores en el hemisferio Sur para los meses D y en el hemisferio norte en los meses J .

22.2.2. Variaciones no periódicas

Aparte de las variaciones diurnas suaves y regulares que se producen con mayor o menor amplitud todos los días, existe una gran diversidad de variaciones irregulares que ocurren de vez

en cuando. Se conocen con el nombre de campos de perturbación y generalmente se expresan con la letra D.

Están causadas principalmente por procesos ocurridos de manera esporádica en el Sol en los cuales se emite radiación corpuscular, produciendo entre otros efectos, las tormentas magnéticas, las subtormentas y las bahías.

• Tormentas magnéticas

Las tormentas magnéticas son perturbaciones no periódicas del campo magnético externo de la Tierra que pueden llegar a tener amplitudes de hasta 400nT. Su comienzo es prácticamente simultáneo en todas las partes de la Tierra. Su origen se debe a la interacción del viento solar con el campo magnético terrestre.

Generalmente se distinguen dos tipos de tormentas: las tormentas de comienzo brusco y las tormentas de comienzo gradual.

En la morfología de las tormentas de comienzo brusco y para observatorios de latitudes medias pueden distinguirse las siguientes fases:

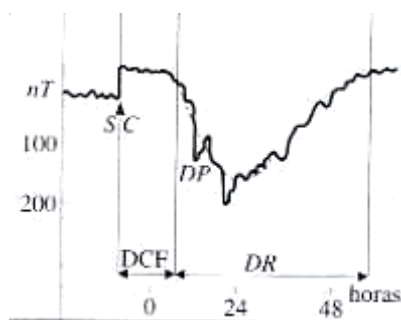


Figura 1. Tormenta magnética. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

- Fase inicial:

La tormenta comienza con una variación muy brusca, generalmente un incremento de la componente H, debido al choque de las partículas emitidas por el Sol sobre la magnetosfera. Es el campo denominada DCF (Disturbed Corpuscular flux).

A veces, este pequeño incremento inicial va seguido de un ligero descenso de H y la tormenta se cataloga como ssc (storm sudden commencement). Generalmente, después del ssc, la componente H aumenta y permanece entre los 30 y los 50nT por encima del valor previo al comienzo súbito.

En periodos de perturbación, muchos de estos cambios bruscos no llegan a constituir una tormenta. Entonces se denominan impulsos súbitos.

- Fase principal:

Comienza entre 2 y 10 horas después del ssc y se caracteriza por una disminución de la componente H de entre 100 y 200nT que dura unas 12 - 24 horas. Este cambio de H es mayor en las proximidades del ecuador y disminuye hacia los polos. Por el contrario, Z no cambia en el ecuador y aumenta ligeramente en los dos hemisferios al aumentar la latitud. Pero en cualquier caso, los cambios en la componente Z son menores que los de la componente H. El campo se denomina DR (Disturbed Ring Current) y está asociado con la corriente de anillo producida por el giro de las partículas en torno a la Tierra en órbitas zonales situadas a 5-6 radios terrestres. Sobre DR se observan a veces variaciones de corta duración que se denominan DP (Disturbed Polar current), originadas por las corrientes ionosféricas en las altas latitudes.

En esta fase es frecuente la ocurrencia simultánea de auroras.

- Fase de recuperación:

Se extiende desde el mínimo de la componente H hasta la recuperación del nivel previo a la tormenta. Dura varios días y el magnetograma adopta una forma exponencial.

De modo que la perturbación magnética D está formada por dos partes:

- Una parte que no depende del tiempo local, Dst (Disturbance storm time), que forma la tormenta en sentido estricto. Está formada por DCF, DR y DP.
- Otra parte, DS (Disturbance Solar), que sí es función del tiempo local y está relacionado con fenómenos localizados en la ionosfera.

Para nuestra latitud, en un año suele haber una tormenta con un incremento de H mayor de 250nT y unas 10 tormentas con un incremento de unos 50nT.

Las tormentas de comienzo gradual son estados de agitación más o menos intensa (pero con amplitudes inferiores a las que aparecen en las ssc) con un principio y un final difíciles de situar temporalmente. Son producidas por emisiones corpusculares ocurridas en regiones permanentes del Sol sin manchas solares, por lo que estas tormentas se suelen repetir cada 27 días y suelen ser más frecuentes en épocas de baja actividad solar.

Además, en el registro de una tormenta aparecen multitud de anomalías denominadas fluctuaciones irregulares que se extienden desde la fase inicial hasta el comienzo de la fase de recuperación. Pueden tener una duración de entre minutos y horas y su amplitud puede variar

desde muy pocos nT hasta varias decenas de ellos. Las más importantes son las subtormentas y las bahías.

- Subtormentas: Son perturbaciones de características análogas a las tormentas pero de menor amplitud y duración aunque mucho más frecuentes (al menos 1 ó 2 veces por mes).
- Bahías: son pequeñas perturbaciones con una duración de entre decenas y centenas de minutos. Se denominan así porque su forma recuerda al contorno de la línea de costa.

22.3. Composición de la ionosfera

La ionosfera es la capa de la atmósfera que se caracteriza por ser un medio eléctricamente neutro, donde las partículas de carga positiva igualan en número a las de carga negativa.

Esta zona de la atmósfera comienza en torno a los 50-60 km. de altura, presentando un máximo de densidad electrónico alrededor de 220–300 km. Su límite superior no está bien definido extendiéndose hasta los 2000–2500 km. de altura, lugar donde comienza la protonosfera.

Las principales capas que se definen en la ionosfera están caracterizadas por su densidad electrónica. Estas son:

- capa D: se extiende entre los 60 y los 90km de altura y se caracteriza por una elevada absorción de ondas electromagnéticas causada por la alta densidad de partículas neutras que contiene. Esta capa desaparece por la noche y sus principales componentes son N_2 , O_2 , CO_2 , Ar, Ne, He y en menor medida NO.
- capa E: está compuesta por O_2 , O, NO^+ , O^+ y electrones entre los 90 y los 140 km. de altura. La ionización es debida a la radiación ultravioleta y rayos X, por lo que esta capa depende estrechamente de la radiación solar, debilitándose por las noches y aumentando durante el verano
- capa F: se distinguen dos zonas dentro de esta capa:
 - F_1 : entre los 140 y los 200 km. Compuesta principalmente por N_2 , NO^+ , O^+ y electrones, siendo más importante en verano y en los mínimos del ciclo solar. Tiende a desaparecer por la noche.

- F₂: se extiende desde los 200 km. hasta los 500 km. aproximadamente y está formada por O, N, O⁺ y electrones. La concentración de electrones es mayor en invierno que en verano y muy grande en los equinoccios. No disminuye por la noche y existe un mínimo en la región ecuatorial donde la radiación es máxima por lo que no sigue una pauta diaria de acuerdo con el Sol.

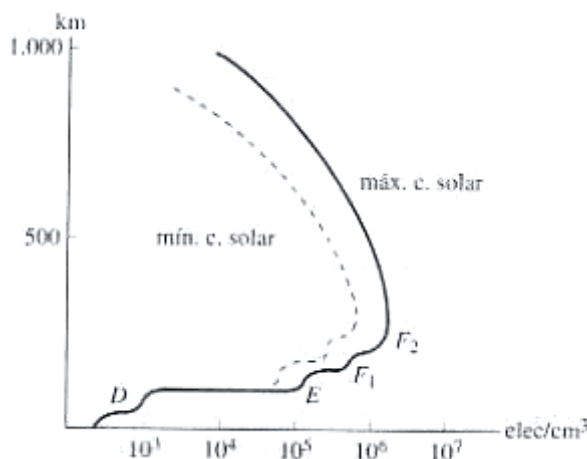


Figura 2. Distribución de densidad de iones en la ionosfera para el máximo y el mínimo de manchas solares.

Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

La teoría Chapman establece que las capas ionosféricas se forman por la ionización de las partículas de la alta atmósfera, originalmente neutras, por efecto de la radiación ultravioleta y rayos X procedentes del Sol. En este proceso, la radiación solar es absorbida, produciéndose en cada paso un ión positivo y un electrón libre, lo que resulta en un número igual de partículas con carga positiva que con carga negativa.

A una altura determinada, la energía absorbida por la ionización dependerá de la intensidad de la radiación monocromática y de la densidad de partículas. Así, la cantidad de iones producidos para una intensidad fija, aumentará con la densidad de partículas, mientras que para una densidad fija, aumentará con la intensidad de la radiación. (Durante el día, al aumentar la intensidad de la radiación aumenta la densidad de electrones, pero por la noche, la fuente de radiación es eliminada y la densidad electrónica decae).

Sin embargo, este modelo simplifica la alta atmósfera y la considera homogénea (formada por un solo elemento), distribuida en capas horizontales en reposo y eléctricamente neutra. Además supone que la radiación es monocromática, y que solo existe un mecanismo de ionización con las partículas de la atmósfera. Aún así, esta teoría es buena para describir el funcionamiento de las capas más bajas, capas D y E, pero no para describir a la capa F.

22.4. Estructura de la magnetosfera

La magnetosfera que es la zona que rodea a la Tierra y que se caracteriza porque fuera de ella, la influencia del campo magnético de la Tierra desaparece.

Su estructura viene determinada por la interacción de las partículas de baja energía del viento solar con el campo magnético terrestre (ya que las partículas de alta energía que constituyen la radiación cósmica, atraviesan el campo magnético terrestre sin ser apenas afectadas por él).

Esta interacción produce una deformación de las líneas de fuerza del campo magnético terrestre, comprimiéndolas en la dirección frontal hacia el Sol (entre 10 y 15 radios terrestres) y alargándolas en la dirección opuesta (hasta 60 radios terrestres) en la denominada cola magnetosférica. En los denominados modelos abiertos de la magnetosfera, la cola se extiende hasta distancias de 100-200 radios terrestres o incluso mayores, fundiéndose en muchos casos las líneas de fuerza del campo magnético con las del campo interplanetario.

El límite exterior de la magnetosfera viene determinada por la superficie llamada magnetopausa, fuera de la cual, el campo magnético de origen terrestre desaparece. El límite inferior no está bien definido ya que existe bastante continuidad con la ionosfera.

A la distancia de unos 4 radios terrestres en la dirección hacia el Sol se produce el frente de choque debido a la propagación supersónica de las partículas del viento solar. La región que se encuentra entre este frente de choque y la magnetopausa se denomina magnetovaina y en ella, las partículas tienen un movimiento irregular.

El interior de la magnetosfera se encuentra lleno de plasma, principalmente de origen ionosférico aunque en ocasiones se cuela plasma de origen solar. Dentro de ella, podemos destacar:

- la plasmaesfera: constituye la región donde se encuentra la mayor densidad de partículas de baja energía que rota con la Tierra. Se encuentra a una altura entre los 3 y los 6 radios terrestres, dependiendo del estado de perturbación de la magnetosfera (la cual depende del estado de perturbación del Sol). La superficie exterior que delimita esta región se denomina plasmapausa y fuera de ella, la densidad de electrones y iones decrece muy rápidamente.
- Hoja de plasma u hoja neutra: separa en dos partes a la cola magnetosférica. A un lado de la hoja las líneas del campo magnético se dirigen hacia la Tierra mientras que en el otro lado, se alejan de ella.

Además existen dos puntos neutros (A y B) donde el campo magnético es nulo. Estos puntos se encontrarían sobre los polos magnéticos si el campo magnético no fuese perturbado por el viento solar.

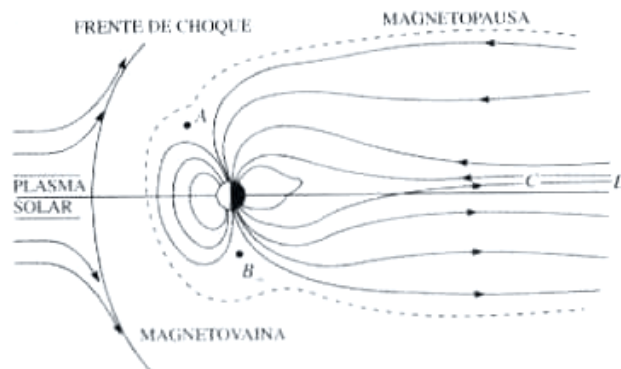


Figura 3. Líneas del campo magnético terrestre y estructura simplificada de la magnetosfera. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

22.5. Anillos de radiaciones y Auroras

Los anillos o cinturones de Van Allen son partículas “atrapadas” en órbitas ecuatoriales alrededor de la Tierra que se distribuyen en capas según su energía y que se generan por la colisión de las partículas de alta energía, bien provenientes de rayos cósmicos o de erupciones solares, con las partículas de la atmósfera produciendo un neutrón que se desintegra en un protón, un electrón y un neutrino, quedando el electrón y el protón capturados por el campo magnético en estos anillos.

Las partículas más energéticas, de energías superiores a 30MeV para los protones y próximas a 2MeV para los electrones, se sitúan cerca de la Tierra, en torno 1,5-2 radios terrestres. Las de menor energía, protones y electrones de menos de 6MeV, se encuentran entre los 4 y los 5 radios terrestres.

Estas partículas giran alrededor de las líneas de fuerza del campo magnético (v_n) con periodos de 1-100 μs entre dos puntos de inflexión o puntos espejo (v_p) (con un periodo de 10 a 30 segundos) y a su vez tardan entre 10 minutos y 24 horas para la revolución completa alrededor de la Tierra ya que se desplazan con un movimiento de deriva v_D alrededor de la Tierra (hacia el este los electrones y hacia el oeste los protones).

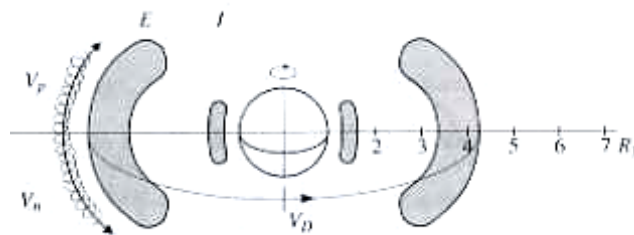


Figura 4. Anillos o cinturones de Van Allen. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

Las auroras son fenómenos asociados con la actividad solar y la penetración de partículas a través de la magnetosfera en la alta atmósfera. Las más conocidas son las llamadas auroras luminosas. Pero también existen las auroras de radar, de rayos X, absorción de ondas de radio, emisión de bajas frecuencias, etc.

Las auroras luminosas se producen porque los átomos de los gases de la alta atmósfera (N_2 , O, H) pasan a un estado excitado al interactuar con las partículas del viento solar que penetran en la atmósfera (principalmente electrones con energías de entre 1 y 10 keV). Al volver los gases a su estado de equilibrio se emite radiación luminosa que es la causante de las auroras.

La radiación emitida es roja o verde cuando las partículas cargadas interactúan con los átomos de oxígeno según sus líneas espectrales, mientras que es violeta, si la interacción ocurre con las moléculas de nitrógeno.

Las auroras ocurren en las altas latitudes, cerca de las regiones polares, ya que las partículas del viento solar entran por una zona que, en periodos de calma solar, es casi circular y tiene un radio de unos 20° alrededor de los polos geomagnéticos. Al aumentar la actividad solar, esta zona se deforma pudiendo extenderse a latitudes medias. La mayor parte de las auroras tienen lugar entre los $67-77^\circ$.

Las auroras luminosas se clasifican según:

- su forma: banda, arco, difusa
- su estructura: homogénea, estriada, rayada
- su comportamiento: tranquilo, activo, pulsante
- su intensidad.

Las auroras se extienden verticalmente entre los 100 y 400 km de altura y horizontalmente pueden ocupar varios miles de km. La zona más visible suele situarse entre 95 y 120 km, es decir, en la capa E.

Bibliografía

- [1] Herráiz, M. Apuntes de Geomagnetismo II: Campos Externos. 5º Curso de Ciencias Físicas, Universidad Complutense de Madrid.
- [2] Udías, A. y Mézcua, J. Fundamentos de Geofísica. Editorial Alianza, 1996. pág . 307-326

Tema 23. Observaciones del campo magnético. Medidas absolutas y relativas. Métodos clásicos y modernos de medidas del campo magnético. Observación desde satélites.

23.1. Observaciones del campo magnético

El campo geomagnético es una complicada función del tiempo y del espacio. Las medidas siempre tienen entre otros fines el estudiar un aspecto particular de la variación temporal o de la variación espacial.

Observatorios magnéticos: en ellos se registran las tres componentes del campo de forma continua. La resolución temporal es del minuto, aunque en la actualidad el registro se hace cada segundo. El tiempo siempre está referido al tiempo universal y la intensidad del campo está referida a un magnetómetro de protones, bien directamente o mediante una intercalibración de instrumentos. Los observatorios son los que aportan una información más precisa sobre la variación secular y la mayor parte de la información de las variaciones temporales en el rango de décadas a minutos.

El lugar elegido para construir un observatorio, debe estar exento de perturbaciones de origen artificial y de anomalías magnéticas producidas por el terreno. En la construcción de los pabellones no deberá emplearse ningún material magnético. En las proximidades de los pabellones, tampoco deberá haber nada de hierro y el observador en el momento de realizar las observaciones deberá ir desprovisto de todo material magnético.

Los observatorios suelen estar compuestos de diversos pabellones: pabellón de absolutas, pabellón de variómetros y otras dependencias. Las salas donde están instalados los instrumentos, suelen ser interiores, rodeadas de un pasillo y mediante un sistema de ventilación facilita la circulación del aire y un cierto aislamiento térmico.

El objetivo principal de INTERMAGNET es establecer una red global de cooperación entre observatorios digitales geomagnéticos, adoptando especificaciones estándar modernas para la medida y registro de los datos geomagnéticos, para facilitar el intercambio de los datos y la producción de productos geomagnéticos en tiempo casi-real.

La adhesión al proyecto Intermagnet supone el adoptar un compromiso de utilizar un instrumental y una metodología que produzcan unos datos de gran calidad y que puedan ser transmitidos lo más rápidamente posible (vía satélite o e-mail) a los distintos Nodos de recepción de datos (GIN), dos en América del Norte, dos en Europa y uno en Japón. Se registran los valores del campo magnético cada minuto.

Los requerimientos mínimos son:

Magnetómetro Vector

Resolución:	0.1 nT
Rango Dinámico:	6000 nT Auroral & Ecuatorial 2000 nT Latitudes medias
Paso de Banda :	DC to 0.1 Hz
Intervalo de muestreo:	.2 Hz (5 sec)
Estabilidad Térmica:	.25 nT/°C
Estabilidad a largo plazo:	5 nT/year
Precisión:	±5 nT datos definitivos

Magnetómetro de Protones o escalar:

Resolución:	0.1 nT
Intervalo de muestreo:	0.033 Hz (30 sec)
Precisión:	1 nT

Desde 1997, el Instituto Geográfico Nacional tiene integrados en la red Intermagnet sus dos Observatorios Geomagnéticos: San Pablo-Toledo (SPT) situado en el centro de la península y Güimar-Tenerife (GUI) en la Isla de Tenerife. Ambos transmiten sus datos vía e-mail al Nodo de París.

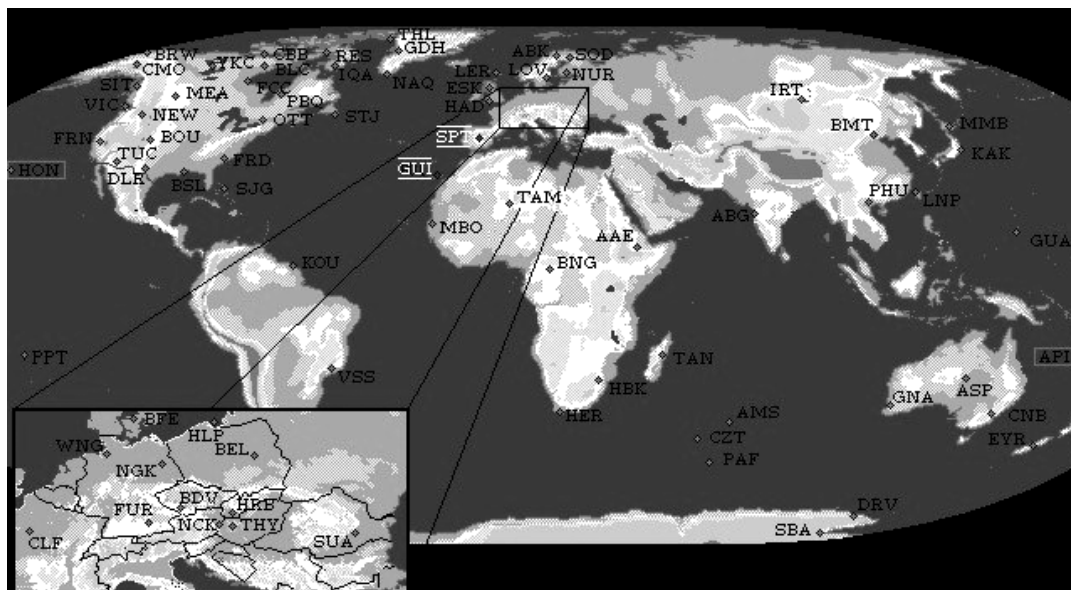


Figura 1. Observatorios

Levantamientos Regionales (Estaciones seculares) El objetivo principal de las Estaciones Seculares o de Repetición es el de poder determinar la Variación Secular en el mayor número posible de puntos y durante un periodo de tiempo lo mas grande posible, así se densifica la labor de los observatorios geomagnéticos, con menor precisión pero con un considerable menor coste. Otras utilizaciones son la producción de modelos regionales y la cartografía magnética.

La distancia media entre estaciones recomendada por la Asociación Internacional de Geomagnetismo y Aeronomía IAGA (200 Km.), para poder detectar rasgos magnéticos con longitudes de onda mayores de 400 Km. y así poder estudiar la relación del núcleo de la Tierra con la variación secular verdadera.

El intervalo de observación es de aproximadamente cada dos años y la duración de la observación es de uno o dos días, compuesta por dos tandas de 6 observaciones completas cada una.

La precisión con que se observa en una estación secular en la actualidad es para la intensidad del campo de 5nT, para declinación 1' y para la inclinación de 0,5'. Los criterios para la elección de la localización de una estación secular son: el campo magnético de la estación no debe estar influenciado por anomalías magnéticas producidas por estructuras geológicas, los valores del campo, deben ser representativos de la región, el subsuelo en los alrededores de la estación deben ser eléctricamente homogéneo. Alejarse de zonas costeras con corrientes de conductividad, el campo magnético ha de ser uniforme en las inmediaciones de la estación, los gradientes tanto horizontales como verticales deberán ser de unas pocas nanoteslas por metro y deberá estar libre de posibles interferencias artificiales tales como: ferrocarriles eléctricos, líneas eléctricas, generadores, etc.

Además de los observatorios magnéticos el campo magnético o algunas de sus componentes también se miden en los levantamientos locales, los levantamientos aéreos o marítimos, en estudios de paleomagnetismo y desde satélites.

23.2. Medidas absolutas y relativas

Debido al carácter vectorial del campo geomagnético, su determinación implica la medida de su magnitud y dirección. Esto se puede hacer por observación del módulo del vector total de campo F y de los ángulos de declinación D e inclinación I , o también por determinación de sus componentes X , Y y Z en la dirección de las coordenadas geográficas, N-S, E-W y vertical.

Como en el caso de las medidas gravimétricas, las medidas magnéticas pueden ser absolutas y relativas, aunque ahora la distinción no es tan clara. De una manera general, se puede decir que las medidas absolutas proporcionan el valor de la magnitud que se quiere medir en un instante

dato. Las medidas relativas a su vez proporcionan las variaciones del campo, de un lugar a otro, o de un instante de tiempo a otro en el mismo lugar. En realidad, las medidas absolutas no son instantáneas, ya que el proceso de medida dura un cierto tiempo, y el valor obtenido representa un valor promedio durante el intervalo de medida.

El fundamento de las medidas del campo magnético se basaba, hasta hace relativamente poco, en la orientación o movimiento de un imán suspendido. Esta técnica se ha visto complementada con la aplicación de los magnetómetros modernos, cuyo funcionamiento se basa en fenómenos de resonancia magnética atómica o nuclear, y en las propiedades magnéticas de ciertos materiales, lo que permite obtener de forma sencilla valores más exactos del campo geomagnético.

Históricamente, las primeras medidas magnéticas son las de la declinación, realizadas con una aguja imanada que se remontan al siglo xv. Hacia 1540, Alonso de Santa Cruz y Joao de Castro realizaron numerosas determinaciones de la declinación durante sus viajes a América y las Indias Orientales. Las primeras medidas de la inclinación con aguja imanada de eje horizontal fueron realizadas por G. Hartmann, en 1544. Las medidas de las componentes de la intensidad del campo magnético terrestre son de época posterior. Las primeras son de la componente horizontal H y fueron realizadas por Humboldt y Gauss, hacia 1830. Para 1837 se describe ya el método de medidas absolutas" de H , Y la construcción de un variómetro para el mismo componente. Entre 1836 y 1840 se realiza, impulsado por Humboldt, Gauss y Weber, un programa de observaciones simultáneas en un número de observatorios magnéticos (Göttingen Magnetische Verein), al que se unieron, en 1839, los instalados en las colonias británicas (British Colonial Observatories), bajo la dirección de E. Sabine. En el mismo año, 1839, Lloyd introdujo la balanza magnética para medir las variaciones de la componente vertical Z . El registro fotográfico de las variaciones fue introducido por Brooke, en 1846, en Greenwich y se ha mantenido hasta nuestros días, en que se ha complementado por métodos modernos de registro digital.

23.2.1. Métodos clásicos de medidas absolutas y relativas

Medidas absolutas

La determinación absoluta de las componentes D , H , Z , ha servido hasta época muy reciente como única forma de la determinación del campo total F y sus componentes. Actualmente se utiliza de forma más extendida la determinación absoluta de F mediante métodos modernos, como veremos más adelante.

La observación de la declinación magnética D , en un lugar determinado, consiste en la determinación del acimut de una aguja imantada, que pueda oscilar libremente en un plano horizontal. Considerando que no existe torsión en el hilo de suspensión, la aguja se orientará en la dirección del meridiano magnético, formando un ángulo α con el origen 0 del limbo graduado situado en la plataforma horizontal. Retirando el imán, el origen 0 del aparato formará con una mira de referencia de acimut β conocido, un ángulo α' .

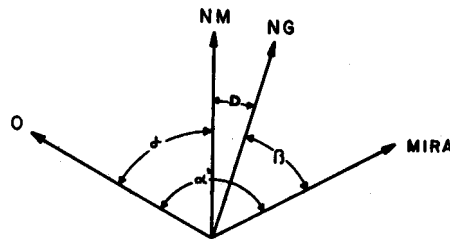


Figura 2. Determinación de la declinación magnética. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

Según puede observarse en la figura 2, se deduce que:

$$\alpha' - \alpha = D + \beta$$

donde α , α' y β son conocidos, por lo que puede deducirse el valor de la declinación D .

Existen, no obstante, errores debidos a la tensión residual del hilo de suspensión que es necesario compensar. Para eliminar este efecto se realizan las medidas con dos imanes de momento magnético fuerte y débil. Por otra, parte, como estas observaciones son muy lentas, es necesario corregirlas por; la propia variación de la declinación durante el tiempo de ejecución de la medida.

Para la determinación del módulo de la componente horizontal H se utiliza un instrumento llamado teodolito magnético. Este instrumento consiste en un imán de momento magnético m suspendido horizontalmente por una fibra de cuarzo. El método en el que se basa es el de Gauss o determinación de mH y m/H , correspondientes al momento magnético del imán y la componente horizontal del campo magnético.

Inicialmente se suspende el imán, cuyo momento de inercia es J , en un plano horizontal, es decir, sometido a la acción de H . Si se desplaza el imán de su posición de equilibrio, comienza a oscilar alrededor del meridiano con un periodo T dado por:

$$T = 2\pi \sqrt{\frac{J}{mH}}$$

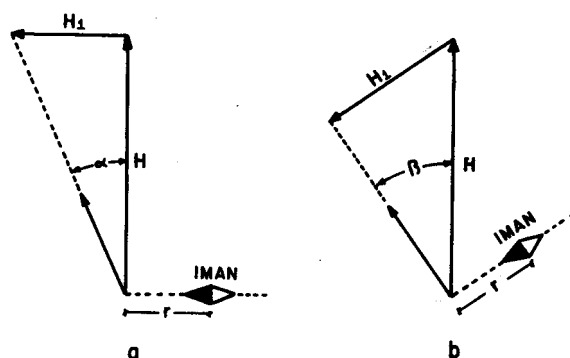


Figura 3. Posiciones en la determinación absoluta de la componente horizontal H. a) Gauss y b) Lamont. Udías, A. y Mézcua, J. Fundamentos de Geofísica

donde se desprecia la acción de la fibra, y la única fuerza actuando es la del campo magnético del imán. Midiendo el periodo T y conocido J obtenemos mH. Para el cálculo del cociente m/H se sustituye el imán suspendido por una aguja imanada, y el imán del momento m se utiliza como imán desviador, el cual se sitúa en el plano horizontal en dos posiciones que se denominan de Gauss y Lamont (J. von Lamont, 1805-1879).

Situando el imán de momento perpendicular al campo H (posición de Gauss), la aguja imanada girará un ángulo α . Este ángulo dependerá de H_1 , componente radial del campo creado por el imán y cuyo valor es:

$$H_1 = \frac{2Cm}{r^3}$$

siendo r la distancia del imán a la aguja suspendida. De la figura se desprende que:

$$\operatorname{tg} \alpha = \frac{H_1}{H} = \frac{2Cm}{Hr^3}$$

Si se sitúa el imán de tal forma que permanezca normal a la aguja imanada en la posición de equilibrio (posición de Lamont) el ángulo girado, cumplirá:

$$\operatorname{sen} \beta = \frac{H_1}{H} = \frac{2Cm}{Hr^3}$$

con la observación de $\alpha\beta$ y/o β se deduce el cociente m/H . A partir de esta medida y la de mH , se pueden obtener los valores de m y H . El valor de H obtenido de esta forma es su valor absoluto.

Medidas relativas

Uno de los instrumentos clásicos más usados para la determinación de la componente horizontal del campo es el magnetómetro QHM (Quartz Horizontal Magnetometer), desarrollado por D. La Cour, en 1930. Consiste en un imán de momento m , suspendido de un hilo de cuarzo que puede oscilar en un plano horizontal. Todo el conjunto puede girar sobre un plano horizontal graduado. Si en una posición P_1 cualquiera de la base cercana al norte magnético, se libera el imán, éste no se orienta en la dirección del norte magnético, sino que debido a una torsión inicial β , forma un ángulo α_1 , cumpliéndose entonces:

$$mH \sin \alpha_1 = \tau \beta$$

siendo τ el coeficiente de torsión del hilo de suspensión.

Si aumentamos el giro del imán en 2π , éste no quedará en P_1 , sino en una nueva posición P_2 cumpliéndose:

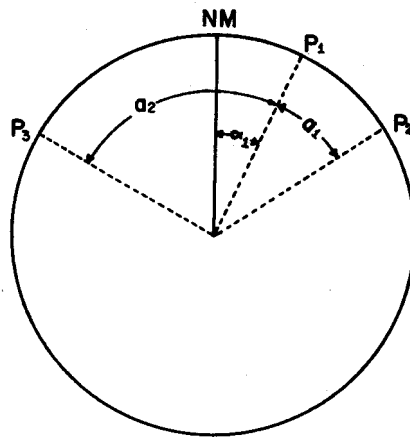


Figura 4. Medida relative de H. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

$$mH \sin(\alpha_1 + a_1) = \tau(\beta + 2\pi)$$

Si volvemos atrás con un giro de -4π se colocará en otra posición P_3 cumpliéndose, por tanto:

$$mH \sin(\alpha_1 - a_2) = \tau(\beta - 2\pi)$$

Sumando las expresiones y sustituyendo el valor de β .

$$\operatorname{tg} \alpha_1 = \frac{\operatorname{sen} \alpha_1 - \operatorname{sen} \alpha_2}{2 - \cos \alpha_1 - \cos \alpha_2}$$

Por otra parte, restando obtenemos:

$$H = \frac{4\pi\tau}{m} [\operatorname{sen}(\alpha_1 + \alpha_2) - \operatorname{sen}(\alpha_1 - \alpha_2)]^{-1}$$

Si el instrumento está bien ajustado de forma que β y α_1 son pequeños, entonces α_1 y α_2 diferirán poco de su valor medio ϕ y, por tanto:

$$H = \frac{2\pi\tau}{m \operatorname{sen} \phi}$$

Esta medida depende de los valores de m y τ que han de contrastarse en continuas calibraciones, ya que varían con el tiempo. Utilizando el mismo aparato se puede determinar D , aunque con mucha menor precisión que con el método descrito anteriormente.

En cuanto a la componente vertical Z , se determina de forma relativa con el aparato BMZ (Magnetometer Zero Balance), basado en lo que se llama método de cero. El instrumento consiste en un imán móvil, que puede girar alrededor de un eje horizontal astatizado con su masa. El conjunto imán e hilo de suspensión puede girar en el plano horizontal, de forma que la aguja esté sometida a la componente vertical del campo magnético.

Cuando el imán móvil forma un cierto ángulo α con la horizontal, la ecuación de equilibrio es:

$$mga \operatorname{sen} \alpha + MH \cos t \operatorname{sen} \alpha = Z \cos \alpha$$

donde t es el acimut del eje magnético del imán con el norte magnético; m , la masa, y M , el momento magnético de dicho imán. Puede apreciarse que el ángulo α depende del valor del acimut t . Si mediante imanes auxiliares se elimina la componente vertical Z , podemos determinar su valor. Para ello se colocan imanes auxiliares de forma que α es cero, cumpliéndose entonces que:

$$Z_c + Z_t + Z_s = 0$$

donde Z_c es la componente vertical del campo geomagnético y Z_t , Z_s los campos producidos por los imanes auxiliares. Este instrumento, cuando se utiliza en prospección, mide las variaciones locales de la componente vertical del campo magnético Z , al moverse de una posición a otra.

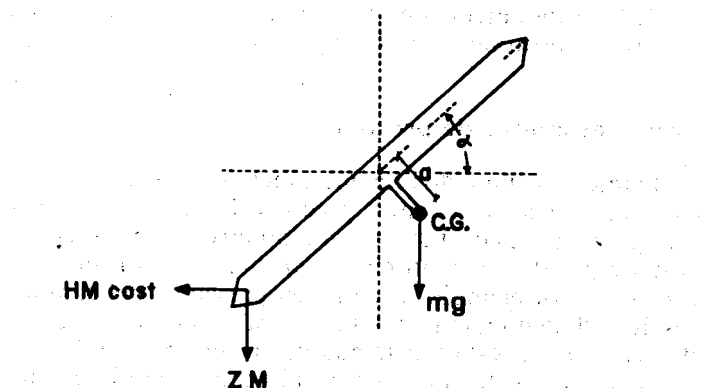


Figura 5. Composición de fuerzas en el plano vertical actuando sobre una balanza BMZ. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

Para medidas rápidas de trabajos de prospección, dentro de los instrumentos de metodología clásica, se usa también el magnetómetro de torsión. Este instrumento se basa en la determinación de un par de torsión ejercido por una fibra de cuarzo, que compensa el par debido al campo magnético terrestre en un imán suspendido. Variando la posición de giro del imán se puede medir con el mismo instrumento las componentes H y Z.

23.2.2. Métodos modernos de medida del campo magnético

Bajo el título de métodos modernos incluimos los que se basan bien en fenómenos de resonancia magnética atómica, o de saturación en núcleos de alta permeabilidad. Este tipo de instrumentación se ha desarrollado prácticamente en los últimos treinta años. El primer magnetómetro de protones fue desarrollado por M. E. Packard y R. H. Varian, en 1954. Una adaptación de este magnetómetro para medir las componentes del campo geomagnético se debe a L. Hurwitz y J. H. Nelson, en 1960. De la misma época, hacia 1958, data el desarrollo de los magnetómetros de bombeo óptico, entre otros por W. E. Bell, T. L. Skillman y P. L. Bender. El uso de la variación de la permeabilidad magnética de núcleos ferromagnéticos, como el instrumental, para medir pequeñas variaciones del campo magnético se remonta a 1936 por Aschen-brenner, G. W. Elmen y H. Antranikian y su aplicación a la prospección de minerales se generaliza a partir de 1949.

Entre las ventajas que han proporcionado estos métodos está la precisión en las medidas, en las que es posible llegar a la décima de nT y la rapidez en las medidas que permite apreciar con precisión variaciones rápidas del campo magnético. Entre los distintos procedimientos actuales, daremos una descripción corta de los fundamentos en que se basan los tres tipos más importantes, magnetómetro de protones, de bombeo óptico y de núcleo saturado.

Magnetómetro de protones. Los instrumentos llamados magnetómetros de protones constituyen en la actualidad una de las formas más precisas de determinar la intensidad total del campo geomagnético. Este magnetómetro se basa en el fenómeno de precesión del espín de los protones de un material. En un líquido cualquiera rico en protones, por ejemplo, el agua, en presencia del campo geomagnético, que no es lo suficientemente intenso para orientarlos en su dirección, los espines de los protones están distribuidos al azar. Si se aplica un fuerte campo magnético polarizador, los protones se alinean en la dirección de este campo. Cuando de forma repentina cesa este campo intenso, los protones tienden a orientarse primero en la dirección del campo geomagnético, produciéndose un movimiento de precesión en torno a él por un breve tiempo, antes de distribuirse de nuevo al azar.

Si $\mathbf{L}$ es el momento angular cinético del protón y $\boldsymbol{\mu}$ su momento magnético que forman un ángulo θ con la dirección del campo magnético externo $\mathbf{F}$, el protón estará sometido a un par dado por $\boldsymbol{\mu} \times \mathbf{F}$ y el momento angular magnético $\mathbf{L}$ adquiere un movimiento de precesión:

$$d\mathbf{L}/dt = \boldsymbol{\mu} \times \mathbf{F}$$

Por otra parte, si ω_p es la velocidad angular de la precesión se tendrá:

$$d\mathbf{L}/dt = \boldsymbol{\omega}_p \times \mathbf{L}$$

de donde resulta:

$$\mu F \sin \theta = \omega_p pL \sin \theta$$

$$\boldsymbol{\omega}_p = \mu F/L = \gamma_p F$$

La cantidad γ se la conoce como relación giromagnética del protón, y es una constante conocida por experimentación en el laboratorio. Un valor reciente de esta constante es $\gamma_p = 2,67513 \times 1,08 \text{ T}^{-1} \text{ s}^{-1}$. Por tanto, para medir F se debe detectar la corriente inducida en la bobina de polarización por efecto de la precesión, y cuya frecuencia es idéntica a la de precesión de los protones ω_p . La duración de este fenómeno es tan sólo de unos cinco segundos después de cesar la polarización, por lo que se obtiene es un valor prácticamente instantáneo. Las frecuencias obtenidas son de alrededor 2000 Hz para un valor típico del campo total F . La relación entre el valor de la frecuencia $f = \omega_p/2\pi$ medida en Hz y la intensidad total del campo F en nanoteslas es:

$$F = 23,4874 f$$

La frecuencia se determina durante los pocos segundos que dura el fenómeno, por lo que estos instrumentos no dan una medida continua de F , sino medidas a intervalos discretos. Debido al hecho de dar medidas absolutas de la intensidad total del campo y a su independencia con respecto a la orientación, estos instrumentos son muy utilizados, tanto para medidas en superficie como en aviones y satélites artificiales.

Magnetómetro de bombeo óptico- Este tipo de magnetómetro está basado en el efecto Zeeman, de forma análoga al funcionamiento de un maser, cuya frecuencia de autoexcitación es una función del campo geomagnético. Si un gas alcalino, por ejemplo, rubidio, se somete a la acción de un campo magnético, cada nivel de energía se subdivide en subniveles ligeramente diferentes. Supongamos ahora que luz polarizada procedente de una lámpara de rubidio pasa a través de un contenedor con vapor de rubidio. Se produce una absorción de fotones que en el vapor excita los átomos del Rb^{87} a un nivel superior. Los átomos excitados vuelven al estado inicial, pero, a causa de la luz polarizada, a un subnivel superior, donde son atrapados.

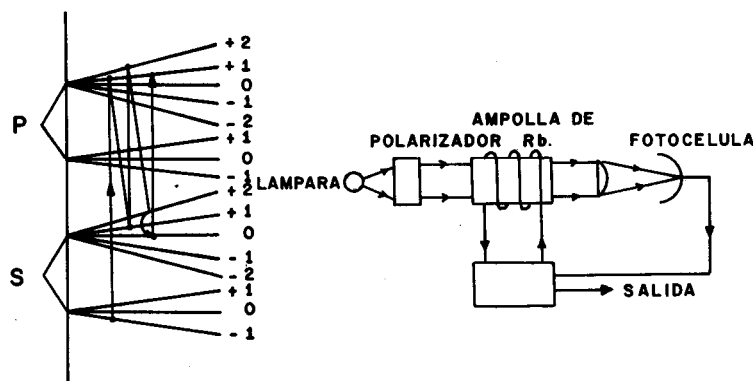


Figura 6. Esquema de un magnetómetro de bombeo óptico. Fuente: Udías, A. y Mézcua, J. Fundamentos de Geofísica

El paso de electrones de un nivel a otro se pone de manifiesto haciendo incidir sobre una célula fotoeléctrica la luz que atraviesa la ampolla. La absorción de energía se traduce en una disminución de la intensidad luminosa que recibe la célula. Al volver los átomos a su estado inicial, la ampolla se hace totalmente transparente y la corriente de salida de la célula es máxima. Esta corriente amplificada pasa por un arrollamiento alrededor de la ampolla de gas, creando un campo magnético variable capaz de redistribuir los electrones en su estado original, cuando la frecuencia sea idéntica a la de transición entre dos subniveles Zeeman. En este momento la ampolla se hace opaca, y la corriente de la célula es mínima. La frecuencia del sistema es del, orden de 200 a 300 kHz, y depende de la diferencia entre los niveles de energía. Por tanto, la frecuencia estará finalmente relacionada con el campo magnético terrestre que actúa sobre él.

En este tipo de magnetómetros se obtienen medidas de la intensidad total del campo geomagnético con gran precisión, influyendo muy poco su orientación. Son además más sensibles que los magnetómetros de protones y se pueden emplear para la observación de las variaciones temporales del campo.

Magnetómetros de núcleo saturado (Fluxgate) -Con el desarrollo de la electrónica se han ido estudiando diversas aplicaciones de ésta a la determinación de las variaciones del campo magnético de la Tierra, para sustituir a los variógrafos clásicos. Uno de los tipos más empleados es el magnetómetro de núcleo saturado. Este tipo de magnetómetros permite el registro simultáneo de las tres componentes geográficas del campo. Esencialmente, el sensor de un magnetómetro de saturación consiste en un núcleo de un material ferromagnético de alta permeabilidad magnética, rodeado por un arrollamiento receptor. El campo magnético terrestre a lo largo del eje del núcleo produce en éste un flujo magnético. Si se hace variar la permeabilidad del núcleo, variará su flujo y como consecuencia se inducirá un voltaje en el arrollamiento receptor que proporciona unas medidas del campo externo. El funcionamiento del magnetómetro está, por tanto, basado en la variación con el tiempo de la permeabilidad del núcleo. Para conseguir esta variación se han desarrollado diferentes tipos de sensores.

Uno de los tipos más utilizados es el llamado de Vacquier, que consiste en dos núcleos paralelos con arrollamientos primarios, encargados de variar la permeabilidad de los núcleos con el tiempo, de sentido contrario y conectados en serie. A los primarios se conecta la señal de un oscilador de una frecuencia de excitación determinada, que produce la saturación de los núcleos. Rodeando ambos núcleos hay un tercer arrollamiento secundario, que actúa como receptor, acoplado a un amplificador. En ausencia de un campo magnético externo, los núcleos se comportan de forma que la señal de salida en el secundario es nula. Si existe un campo externo se produce un desfase entre las señales en ambos núcleos, con una corriente resultante en el secundario, que previa amplificación resulta en una señal proporcional a dicho campo. La mayoría de los magnetómetros posee además un sistema de retroalimentación negativa con un cuarto arrollamiento, que ajustado convenientemente, produce un nuevo campo con el que se consigue la anulación del campo externo. En estas condiciones, la corriente del cuarto arrollamiento proporciona una medida del campo externo que está actuando sobre el sensor. Las medidas dependen de cómo se sitúen los núcleos, paralelos al campo total o en cualquiera de las tres direcciones geográficas, obteniéndose en cada caso los componentes correspondientes. Esta operación se realiza por medio de giróscopos o sensores en tres direcciones perpendiculares.

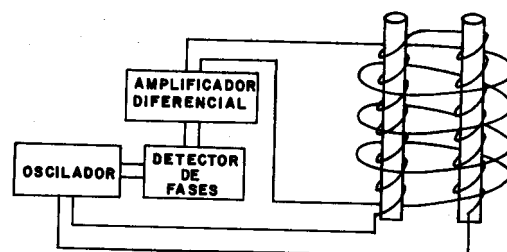


Figura 7. Esquema simplificado de un magnetómetro de núcleo saturado. Fuente: Udías, A. y Mézcua, J.

Errores de este método son debidos a faltas de orientación con respecto al campo, cambios de temperatura, etc. Sin embargo, los magnetómetros de núcleo saturado son de gran utilidad para medir tanto las variaciones del campo geomagnético con el tiempo, como en el espacio, bien en superficie o en aviones.

23.3. Observaciones magnéticas por satélites

Muchos de los resultados recientes más importantes sobre la estructura del campo magnético terrestre han sido debidos a las observaciones obtenidas desde satélites artificiales. El primer satélite artificial, equipado con instrumentos para la medida del campo magnético, fue el Sputnik 3, en 1958, que portaba un magnetómetro triaxial de núcleo saturado. El primero que llevaba un magnetómetro de protones fue el Vanguard 3, en 1959. En 1961, con el Explorer 10, se introduce el uso en los satélites de magnetómetros de bombeo óptico de rubidio, que, combinado con uno de núcleo saturado triaxial, permitió observaciones importantes de las variaciones tanto espaciales como temporales del campo magnético en el espacio exterior. Esta combinación se ha utilizado en muchos otros satélites, como los de las series COSMOS, paGO y OGO. Las observaciones de estos satélites han servido para establecer la estructura de la magnetopausa, la influencia en ella de las radiaciones solares y la determinación de la estructura de la región de transición entre la magnetosfera y el medio interplanetario.

Uno de los programas más importantes de medida del campo magnético con satélites es el programa ØRSTED cuya principal misión es proveer un mapa preciso del campo magnético terrestre. En principio se planeó registrar datos durante 14 meses y las medidas se utilizarían para mejorar los modelos existentes del campo magnético fundamental y determinar los cambios tanto en el campo fundamental como en el más débil, resultante de la interacción entre las partículas iónicas procedentes del Sol (viento solar) y la magnetosfera terrestre en él. También estudiará la transferencia de energía desde el viento solar a la magnetosfera y más abajo hasta las capas más bajas de la atmósfera. Todos estos estudios se beneficiarán no solo de las medidas magnéticas sino de las medidas del flujo de partículas energéticas alrededor del satélite.

Un levantamiento global del campo magnético de la tierra es de gran interés para estudios geofísicos. Levantamientos equivalentes se han realizado solo otra vez, con el satélite americano MAGSAT (1979-80), así con los nuevos datos del Oersted será posible por primera vez tener un levantamiento del cambio en el campo magnético terrestre. Estos cambios están relacionados con los flujos dentro del núcleo terrestre pero el mecanismo exacto todavía no es conocido.

Formaciones geológicas de la corteza terrestre pueden también obtenerse y caracterizarse a partir de estos mapas magnéticos.

Comparando la misión del MAGSAT, el proyecto ØRSTED mejora en diferentes aspectos. La precisión, por ejemplo de los magnetómetros es mayor, la órbita tendrá una orientación diferente con referencia al Sol y por tanto aportará nueva información sobre la distribución global de las corrientes eléctricas en la magnetosfera terrestre.

El satélite porta un total de cinco instrumentos científicos, tres de los sensores están montados al final de un aguijón de 8 metros de longitud para minimizar las perturbaciones generadas por los sistemas eléctricos del satélite. Estos son:

- Magnetómetro CSC flux-gate. Mide el vector campo magnético (intensidad y dirección), tiene una estabilidad de 0.5 nT durante unos días y fue construido en la Danish Technical University.
- Star-imager que determina la orientación del anterior y del satélite. Su precisión es superior a 5 segundos de arco y fabricado por la misma universidad.
- Magnetómetro Overhauser mide la intensidad del campo magnético de la Tierra, tiene una precisión de 0.5 nT. La misión fundamental de este instrumento es la de calibrar a una escala absoluta las medidas del fluxgate.

Los siguientes dos instrumentos están situados en el satélite

- Detector de Partículas- para medir el flujo de electrones rápidos (0.03-1 MeV), protones (0.2-30 MeV), and partículas-alpha (1-100 MeV) alrededor del satélite.
- Receptor Turbo-Rogue GPS- para determinar con precisión la posición del satélite. Periódicamente este instrumento se utiliza para investigar la presión atmosférica, temperatura y humedad por debajo del satélite.

Software. Los datos se registran de todos los instrumentos, los comandos que llegan de la tierra son almacenados y realizados cuando se requiere. Se detectan una serie de errores automáticamente y, en caso necesario, se adopta la configuración mínima (CRI).

Órbita y Operación. Utiliza una órbita elíptica de altura entre los 500 y 850 km y un periodo de 100 minutos.

El receptor principal está en el DMI (Danish Meteorological Institute) de Copenhagen.

Otras misiones de las mismas características han sido CHAMP (lanzado en julio del 2000) y SAC-C (noviembre de 2000) que continúan aportando datos geomagnéticos de alta precisión.

Satélite warm – Estas misiones han sido concebidas como misiones de un único satélite, con diferentes instrumentos, diseños y órbitas. Los progresos en la investigación del campo geomagnético han evidenciado que la precisión en los modelos actuales del campo está ligada al comportamiento dinámico de la configuración de las corrientes externas. Las misiones con un único satélite han conseguido precisiones de unos pocos nT y por tanto no pueden tomar ventaja del gran avance que la instrumentación ha sufrido en los últimos años. Son las misiones compuestas de múltiples satélites midiendo simultáneamente diferentes regiones de la tierra las que pueden sacar ventaja de esta nueva generación de instrumentos. La constelación Swarm consiste de tres satélites en tres órbitas polares diferentes a unas altitudes entre 400 y 550 km.

Objetivos de Swarm-

- estudios de la dinámica del núcleo, procesos de la geodinamo e interacción núcleo-manto.
- mapeado de la magnetización de litosfera y su interpretación geológica.
- determinación del modelo de conductividad eléctrica del manto.
- investigación de las corrientes eléctricas que fluyen en la magnetosfera e ionosfera.

Bibliografía

- [1] Miguel, L de (1980): *Geomagnetismo*. I.G.N. Madrid
- [2] Parkinson, W.D. (1983): *Introduction to Geomagnetism*. Elsevier, Amsterdam.
- [3] Udías, A. y J.Mézcua(1997): *Fundamentos de Geofísica*. Alianza Universidad. Madrid.
- [4] www.intermagnet.org
- [5] www.esa.int

Tema 24. Radioactividad de la Tierra. Elementos radioactivos. Leyes de la desintegración radioactiva. Principios de geocronología. Edad de la Tierra. Evolución térmica de la Tierra.

24.1 Elementos radioactivos

Existe un cierto número de elementos naturales en la Tierra que son radiactivos, es decir, que están formados de átomos inestables que por pérdida de protones o neutrones y/o captura de electrones se convierten en otros elementos, llegando finalmente a convertirse en átomos estables. Esta propiedad es de singular importancia en Geofísica, pues además de proporcionar información sobre el tiempo de formación de las rocas, suministra una fuente de energía calorífica que condiciona las características termodinámicas de la Tierra.

En Física nuclear se denominan isótopos a aquellos elementos que poseen el mismo número atómico y distinta masa atómica. Se consideran radiactivos aquellos elementos que espontáneamente se desintegran en otros que pueden ser también inestables. La primera indicación de que la proporción entre isótopos radiactivos está en relación con la edad de los minerales que los contienen fue hecha por lord Rayleigh, en 1899, estudiando la proporción helio/uranio en diversos minerales. En 1905, Rutherford estudió las posibilidades del fenómeno de la desintegración para determinar las edades de las rocas, deduciendo para el mineral objeto de estudio una edad no inferior a quinientos millones de años. Esta edad venía a ampliar la duración, no superior a cien millones de años, de las eras geológicas, que se admitía a principios de siglo por la mayoría de los geólogos.

Los mecanismos más importantes de desintegración radiactiva son los siguientes: unos elementos radiactivos se transforman en otros por la emisión de una partícula α (núcleo del átomo de helio formado por dos protones y dos neutrones) de su núcleo. Otros emiten electrones (partículas β), resultando en este caso un elemento radiactivo que posee un número atómico superior (en una unidad) con la misma masa. Existen elementos cuyos átomos son capaces de atrapar uno de sus propios electrones, fenómeno opuesto a la emisión β , resultando un nuevo elemento con número atómico disminuido en uno y la misma masa atómica. En este proceso se produce una emisión de rayos X. Por último, otro mecanismo establece la posibilidad de que el elemento radiactivo emita rayos X (radiación electromagnética), transformándose en un elemento no radiactivo. Aunque existen otros tipos de actividad radiactiva natural, los mecanismos descritos corresponden a la gran mayoría de los procesos espontáneos de desintegración.

24.2 Leyes de desintegración radioactiva

En una muestra de material radiactivo con n átomos, el número de átomos desintegrados en un tiempo dado es proporcional al número total de átomos. Llamando λ a la constante de desintegración de este material, el número de átomos desintegrados, dn , en un intervalo de tiempo, dt , viene dado por:

$$\frac{dn}{dt} = \lambda n$$

Integrando se obtiene el número de átomos que quedan sin desintegrar en la muestra después de un tiempo t :

$$n_t = n_0 e^{-\lambda t}$$

donde n_0 es el número de átomos que existían inicialmente. Esta expresión recibe el nombre de *ley de desintegración radioactiva*. También es costumbre llamar actividad R a la relación dn/dt y, como consecuencia de la definición establecida, también puede escribirse en forma análoga a la segunda expresión:

$$R_t = R_0 e^{-\lambda t}$$

Donde R_0 es la actividad inicial (para $t = 0$).

Se suele denominar al elemento inicial de una serie radioactiva elemento Padre y al elemento final de la desintegración, elemento Hija de tal forma que cuando el número de elementos Padre disminuye, el número de elementos Hija H aumentan tal que:

$$H = n_0 - n_t = n_0 (1 - e^{-\lambda t})$$

Puesto que n_0 es difícil de conocer, podremos eliminarlo teniendo:

$$H = n_t (e^{\lambda t} - 1)$$

Es un hecho comprobado experimentalmente que, independientemente del momento en el que comience la observación, se puede obtener un tiempo $T_{1/2}$ en el que la actividad R se convierte en la mitad. Este tiempo se denomina periodo o vida media de la muestra, y depende únicamente del material de que se trate. La relación entre el periodo y la constante de desintegración λ puede establecerse sin más en la relación que nos da la actividad:

$$\frac{R_0}{2} = R_0 e^{-\lambda T_{1/2}}; e^{\lambda T_{1/2}} = 2$$

Tomando logaritmos neperianos en la última expresión:

$$T_{1/2} = \frac{\ln 2}{\lambda} = 0.693 / \lambda$$

La constante de desintegración λ se puede determinar a partir de la observación del periodo o vida media. A continuación presentamos algunos valores de $T_{1/2}$

Te-135	I-135	Xe-135	Cs-135	Ba-135
$T_{1/2}=19s$	$T_{1/2}=6.6h$	$T_{1/2}=9.1h$		$T_{1/2}=2.3 \times 10^6 y$

Si n es el número de átomos de una muestra radiactiva que se han desintegrado en dt años, multiplicando por el tiempo t en el que se ha desintegrado cada uno de ellos, se obtiene la vida de la muestra desintegrada:

$$t = \int_0^t \lambda n dt = \int_0^t n_0 e^{-\lambda t} (\lambda t) dt$$

Si se hace $t = \infty$, se obtiene la vida total de la muestra:

$$t_T = \int_0^{\infty} n_0 t e^{-\lambda t} d(\lambda t) = \frac{n_0}{\lambda}$$

Luego el valor medio de la vida $\bar{t}$ de uno de los átomos que existían primitivamente será:

$$\bar{t} = \frac{t_T}{n_0} = \frac{1}{\lambda}$$

que relaciona una característica como el valor medio de la vida de los átomos y la constante de desintegración de la sustancia.

Un problema que se plantea aquí es el de poder asegurar que todos los átomos del elemento Hija de la muestra proceden por desintegración radiactiva. Si por algún método esto puede asegurarse, y previo conocimiento de λ , se tiene la posibilidad de obtener la vida de la muestra. Esto tiene una aplicación muy importante en la determinación de las edades de las rocas de la Tierra.

24.3 Principios de geocronología

Las edades de ciertos minerales, deducidas por métodos radiactivos, se situaron en la primera década del siglo, en aparente contradicción con las supuestas a partir de métodos puramente geológicos. Esta controversia se prolongó durante algunos años hasta los trabajos de A. Holmes, en 1911, y J. Barrell, en 1917, en los que de forma explícita se establece que la radiactividad se comporta como un verdadero reloj de medición de tiempos para los procesos geológicos. Los perfeccionamientos en la espectrometría de masas en los años 1940 ayudaron de manera decisiva al desarrollo de estas técnicas.

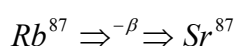
En principio, estas técnicas se basan en la obtención de la relación entre NR del elemento cabeza de serie (Padre), y NE del elemento final del proceso de desintegración (Hija). Los distintos métodos se conocen por los nombres del radioisótopo natural y su producto final.

Como ya se ha dicho, los primeros estudios se realizaron con el uranio y helio, pero pronto se abandonó este método, ya que el helio, por ser un gas, se desprendía del mineral, no pudiendo medirse su contenido con precisión. Entre las familias más usadas actualmente están los productos de desintegración del uranio, por ejemplo, U^{238} , que produce un isótopo intermedio de Th^{234} y finalmente el elemento estable Pb^{206} , y el isótopo menos abundante U^{235} , que produce finalmente Pb^{207} . Otras familias utilizadas son el Rb^{87} , que produce Sr^{87} , el K^{40} , que se descompone en A^{40} y Ca^{40} , y el C^{14} , que finalmente produce N^{14} . La vida media de los procesos indicados son del orden de 10^9 años, menos para el C^{14} , que sólo es de cinco mil setecientos cincuenta años. Por tanto, estos procesos son muy adecuados para la medida de edades en la Tierra.

La exactitud de las medidas está condicionada a la de los métodos analíticos para medir la proporción del elemento radiogénico, y la contaminación con átomos del mismo elemento de otro origen. El valor de λ , la constante de desintegración, se conoce bien por los análisis de laboratorio, y se supone constante a lo largo del tiempo geológico. El tiempo t_0 que se obtiene se refiere al de la formación del mineral en cuestión, y también puede responder a los diversos acontecimientos geológicos, como la cristalización de las rocas ígneas, recristalización de rocas preexistentes, etc. Las limitaciones e incertidumbres a las que están sujetos estos métodos hacen que se esté abandonando en general el apelativo de «edades absolutas», utilizado en el pasado para los métodos radiactivos. Veamos algunos procesos de desintegración utilizados para la datación de los materiales terrestres.

- Método rubidio-estroncio (Rb-Sr)

El isótopo Rb^{87} abunda en la naturaleza con una proporción del 27,2 por 100 del Rb natural. Este isótopo es radiactivo y se convierte, por emisión de una partícula β , en Sr^{87} :



con una vida media $t_{1/2} = 4,9 \times 10^{10}$ años, aproximadamente.

Supongamos una roca ígnea que en el momento de solidificación ($t = 0$) posee una cierta cantidad de Rb^{87} . Al cabo de un cierto tiempo t , después de haberse solidificado, existirá una cierta cantidad de Sr^{87} , como consecuencia de la desintegración radiactiva.

$$(Sr^{87})_t = (Rb^{87})_0 (1 - e^{-\lambda t})$$

y expresando en ella todas las cantidades en términos de sus valores en el instante actual t , que es el que podemos medir:

$$(Sr^{87})_t = (Rb^{87})_t (e^{\lambda t} - 1)$$

por lo que t puede deducirse inmediatamente de esta expresión.

Uno de los problemas que se plantean, como ya se ha indicado, es que en la determinación del Sr^{87} deben eliminarse las contaminaciones de su isótopo no radiactivo Sr^{86} , que existe con el 9,8

por 100 en el estroncio natural. Para corregir esta indeterminación se calculan las relaciones $\text{Sr}^{87}/\text{Sr}^{86}$ para cada mineral de la roca que se analiza, y que será la suma del inicial más el que se ha producido por desintegración del Rb^{87} ;

$$\left. \frac{\text{Sr}^{87}}{\text{Sr}^{86}} \right|_{\text{total}} = \left. \frac{\text{Sr}^{87}}{\text{Sr}^{86}} \right|_{\text{inicial}} + \frac{\text{Rb}^{87}}{\text{Sr}^{86}} (e^{\lambda t} - 1)$$

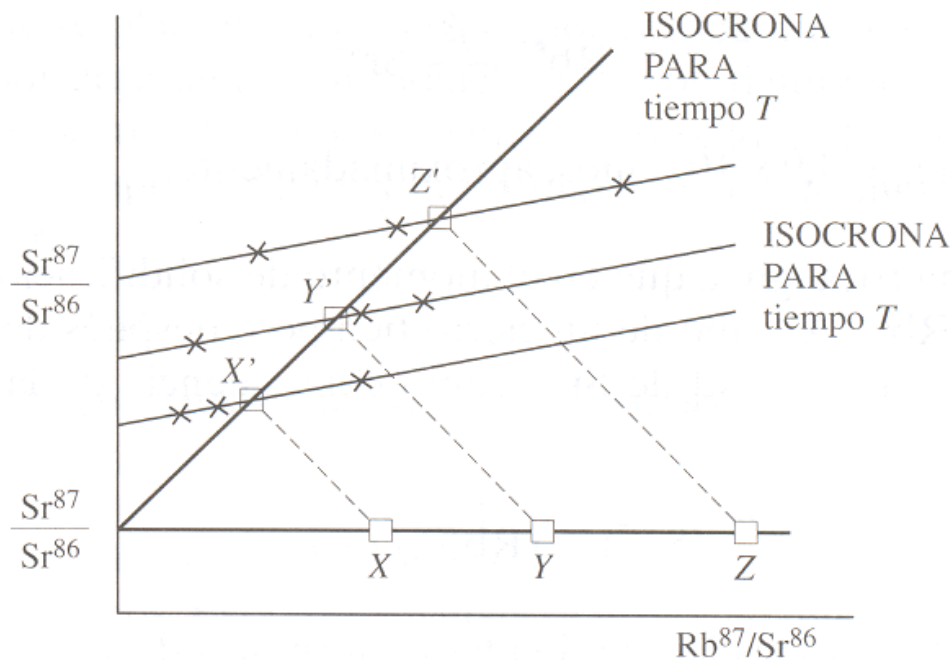


Figura 1. Isocrona de rocas con una fase metamórfica en el tiempo t, en una representación debida a Nycolaysen.

Fuente: Udías y Mezcuá, 1997.

Esta expresión corresponde a una recta de pendiente $e^{\lambda t} - 1$, cuya ordenada en el origen es el contenido de $\text{Sr}^{87}/\text{Sr}^{86}$, existente en el momento inicial de formación de la roca. Por tanto, una representación del contenido total en $\text{Sr}^{87}/\text{Sr}^{86}$, en función del $\text{Rb}^{87}/\text{Sr}^{86}$ presentes actualmente para cada mineral de que se compone la muestra, viene dada por un punto de la recta buscada, denominada isocrona.

En este método existe gran imprecisión, porque el valor $e^{\lambda t} - 1 \approx 0$. Sin embargo, tiene la ventaja que los elementos que han de determinarse son sólidos y, por tanto, existe sólo una pequeña probabilidad de haber perdido elementos resultantes de la serie radiactiva.

- Método potasio-argón (K^{40} - A^{40})

De los tres isótopos del potasio únicamente el K^{40} es radiactivo, existiendo en una proporción del 0,0117 por 100 de los átomos del potasio natural. La desintegración se realiza bajo dos

formas completamente diferentes; una por emisión de partícula β , convirtiéndose en Ca^{40} , y otra por captura de electrones, resultando A^{40} . Si se expresan ambas desintegraciones conjuntamente:

$$\frac{d}{dt}K^{40} = (\lambda_{\beta} + \lambda_e)K^{40} = \lambda K^{40}$$

siendo λ la constante de desintegración promedio. Considerando que en el instante $t = 0$ los elementos fin de desintegración comienzan a acumularse sin existencia de pérdidas, se tiene:

$$\frac{\text{Ca}^{40}}{K^{40}} = \frac{\lambda_{\beta}}{\lambda} (e^{\lambda t} - 1)$$

$$\frac{\text{A}^{40}}{K^{40}} = \frac{\lambda_e}{\lambda} (e^{\lambda t} - 1)$$

En cuanto al conocimiento de λ_{β} , puede determinarse mediante contabilización de partículas β , mientras que para λ_e se obtiene de forma indirecta mediante la emisión de rayos X que van asociados con la captura de electrones. Valores de $\lambda_{\beta} = 4,78 \times 10^{-10} \text{ año}^{-1}$ y $\lambda_e = 0,588 \times 10^{-10} \text{ año}^{-1}$ son generalmente utilizados.

La cuantificación de Ca^{40} es muy difícil, dado que forma el 97 por 100 del calcio presente en las rocas, y es prácticamente imposible separar éste del Ca^{40} de origen radiactivo. Por esta razón, las dataciones se realizan por la medida de la concentración de K y A. En la medida del contenido de A^{40} hay que tener en cuenta que se trata de un gas que puede perderse y, por otro lado, es preciso también corregirlo por la contaminación procedente de la atmósfera.

En definitiva, la precisión del método dependerá de la que se tenga en determinar los elementos fin de la serie radiactiva. Esto está determinado, por una parte, por las características radiactivas del propio elemento, y, por otra, del tipo de minerales que se utiliza, es decir, si favorecen o no el aislamiento del elemento radiactivo con relación al medio que le rodea. Así, la temperatura a que esté sometida una roca después de su formación influirá en la pérdida y en la contaminación procedente de la atmósfera que sufre el A^{40} de origen radiactivo.

Este método ha sido extensamente empleado en la determinación de la edad de meteoritos llegados a la Tierra y muy escasamente en aplicaciones petrológicas de minerales terrestres. Por este método se han determinado edades de meteoritos comprendidas entre $4,0\text{--}4,5 \times 10^9$ años. En este tipo de aplicaciones, es necesario tener en cuenta las altas temperaturas a que han estado sometidos los meteoritos, con la consiguiente pérdida de algunos elementos.

24.4 Edad de la Tierra

La primera estimación de la edad de la Tierra propuesta por lord Kelvin en 1864 daba un resultado de solamente entre 20 y 80 Ma y se basaba en el tiempo necesario para su enfriamiento a partir de un estado inicial de material fundido y el hecho de que la Tierra todavía no se ha enfriado del todo. Este valor contradecía la duración de más de 100 Ma exigida por los procesos geológicos, problema que no fue resuelto hasta el descubrimiento de la radiactividad. La presencia de material radiactivo en la Tierra es en efecto la responsable de un calentamiento adicional que extendería en el tiempo la propuesta inicial de Kelvin.

Además, la desintegración de estos materiales ha proporcionado la metodología adecuada para medir su edad. Por otro lado, los estudios modernos se inclinan por las hipótesis sobre la formación de la Tierra por acreción de planetesimales o asteroides de naturaleza condritica situados en la parte interior del sistema solar. Éstos a su vez se formaron por acreción del polvo en la nebulosa solar. El calentamiento, fusión, segregación gravitacional y diferenciación magmática del material de la Tierra con la formación del núcleo son procesos subsiguientes. La formación de la Luna se asigna, también, a un impacto de un asteroide de gran tamaño y posiblemente también la inclinación del eje de rotación de la Tierra con respecto al plano de la eclíptica y la pérdida de material gaseoso.

En la determinación de la edad de la Tierra son muy importantes las observaciones realizadas sobre meteoritos. La razón fundamental es que, suponiendo que las relaciones de isótopos de plomo en el momento inicial son idénticas en la Tierra que en los meteoritos, en éstos no se han producido fenómenos de diferenciación del material radiactivo como en la Tierra. Esta diferenciación es la causa de que se haya borrado en la Tierra la historia ocurrida en los primeros ochocientos a mil millones de años, pues debido a una intensa actividad ígnea, el material correspondiente a las rocas más antiguas se ha transformado en otras completamente nuevas.

El modelo de Tierra propuesto por Holmes y Houtermans, en 1946, puede servir como modelo de partida en la discusión de la evolución de radioisótopos, en particular la correspondiente al uranio/plomo. Se supone que inmediatamente después de la formación de la Tierra se diferenciaron el manto y el núcleo, con un contenido en U/Pb distintos, permaneciendo éstos como sistemas cerrados. Sin embargo, el muestreo de minerales de estos sistemas es hoy imposible, por lo que se recurre a la datación de meteoritos, que pueden establecer edades comparables a las correspondientes a la Tierra. Los valores comúnmente establecidos siguiendo este método proporcionan una edad para la Tierra de $4,54 \times 10^9$ años.

Por otra parte, considerando esta misma relación para material terrestre, se obtiene un valor promedio de $2,7 \times 10^9$ años, claramente inferior al ya establecido previamente. Sin embargo, si

tenemos en cuenta que con anterioridad a este momento de aparente formación existió el material terrestre con composición radio-isotópica idéntica a la de los meteoritos estudiados, puede deducirse que el contenido en uranio/plomo hace $2,7 \times 10^9$ años provenía de una desintegración iniciada $1,84 \times 10^9$ años antes. La suma de estos dos periodos es de $4,54 \times 10^9$ años, que corresponde exactamente con la datación realizada directamente para material proveniente de meteoritos. De las observaciones realizadas en diversos materiales terrestres, tales como sedimentos marinos, basaltos, etc., se infieren diferentes tiempos, considerando el manto y núcleo como sistemas cerrados, lo que sugiere a su vez que la segregación del núcleo pudo realizarse en un periodo de tiempo bastante grande.

Si se utiliza el método Rb/Sr sobre meteoritos, resultan valores del orden de $4,53 \times 10^9$ años, que están en completo acuerdo con el valor proporcionado por el método U/Pb. Sin embargo, existen observaciones sobre meteoritos utilizando el método Rb/Sr que proporcionan valores muy discordantes, lo que puede explicarse por una más reciente homogeneización en el contenido Rb/Sr por efectos de procesos térmicos no conocidos. Sin embargo, estos valores no contradicen el valor establecido como inicio del sistema solar en $4,55 \times 10^9$ años, prácticamente coincidente con la edad de la Tierra.

Procediendo con los métodos geocronológicos anteriormente descritos, es posible realizar dataciones de las distintas épocas geológicas sobre los materiales de la corteza terrestre, y así establecer una relación entre las divisiones geológicas y tiempos de su formación. La utilización de la geocronología ha supuesto un gran avance, dado que las rocas correspondientes a estas orogenias precámbricas no disponen de fósiles que permitan ordenarlas cronológicamente. Así, la edad conocida como Fanerozoico, en la que aún se hace registro claro de fósiles, y que establece el límite inferior del Cámbrico, se sitúa en seiscientos millones de años. Podemos pues establecer el Precámbrico dividido de la siguiente manera:

<i>Proterozoico</i>	m.a.
— tardío	950
— medio	1.750
— antiguo	2.500
Arqueozoico	> 3.200

Esta división ha sido posible mediante datación radiactiva, que ha permitido distinguir distintos periodos para aquellas épocas que la geología era incapaz de subdividir por carecer de fósiles.

24.5 Evolución térmica de la Tierra

La evolución térmica de la Tierra, propuesta por lord Kelvin en 1864, con un enfriamiento progresivo desde un estado de fusión total hasta la situación actual, resultaba, como ya se ha dicho, en una edad muy corta. El descubrimiento de la radiactividad ha permitido establecer una fuente calorífica en la Tierra, adecuada para explicar el flujo geotérmico observado en la superficie, sin recurrir necesariamente a un modelo de Tierra en enfriamiento progresivo. La idea más reciente es que la Tierra se encuentra en la actualidad en un proceso de ligero enfriamiento como resultado de los procesos de convección en el manto. Así, la mayor parte del calor que ésta desprende proviene de la desintegración de radioisótopos de vida media elevada, en la corteza y manto superior, siendo una parte muy pequeña de este calor la procedente del enfriamiento del planeta. En efecto, cerca del 80% de la energía térmica de la Tierra es producida por la desintegración de isótopos radiactivos, en especial de Th^{232} , U^{238} , K^{40} y U^{235} . El resto representa remanentes del calor originado durante la formación y diferenciación del material de la Tierra. El mecanismo de producción de calor por desintegración se explica de forma sencilla por las partículas emitidas por los átomos radiactivos, que al entrar en colisión con los átomos circundantes transforman su energía cinética en calor (vibración de los átomos en la estructura cristalina). Por ejemplo, un gramo de U^{238} genera $9,4 \times 10^{-8}$ W, y uno de K^{40} , $2,8 \times 10^{-8}$ W. Las rocas ígneas de la corteza generan calor debido a la proporción de elementos radiactivos en ella. En el granito, por ejemplo, la proporción de elementos radiactivos produce un calor de $3 \mu\text{W}/\text{m}^3$, de forma que una columna de 1 m^2 de sección y 30 km de espesor genera 90 mW. Esta cifra es mayor que el flujo medio de calor observado en la superficie terrestre. Por otro lado, las observaciones en el fondo del mar para las regiones oceánicas, cuya corteza no tiene capa granítica, dan un flujo de calor semejante al de los continentes, lo que indica que la situación es más compleja.

El modelo de la evolución térmica de la Tierra puede establecerse en una serie de etapas, comenzando con el proceso inicial del calentamiento producido durante su formación como consecuencia de la energía gravitacional liberada de las partículas preexistentes, juntamente con una fuerte compresión adiabática. En esta primera etapa también pudo funcionar como fuente de calor la desintegración de isótopos de muy corto periodo, pudiendo durar en total este periodo alrededor de 10^6 años.

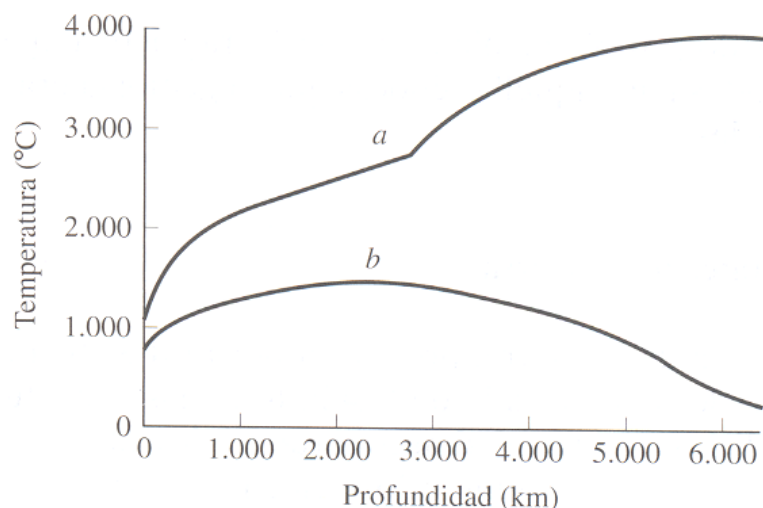


Figura 2. Perfiles de temperatura en el interior de la tierra. A en la situación actual y b antes de la formación del Núcleo. Fuente: Udías y Mezcuá, 1997.

A continuación comienza la etapa de diferenciación del núcleo terrestre, produciendo una fuerte liberación de calor. Esta etapa, cuya duración se estima entre 1 y 1.000×10^6 años, produjo una elevación acusada de temperaturas, llegando a producir una extensa fusión parcial en el manto superior. Posteriormente, durante un tiempo de varios cientos de millones de años, se establece un equilibrio térmico entre el calor producido por radioisótopos de gran periodo, enfriamiento progresivo y flujo geotérmico en la superficie.

En este periodo, al existir una fuerte convección en el manto, el flujo geotérmico en la superficie se considera cinco veces superior a la actual. Por último, se inicia la etapa reciente hace cuatro mil millones de años, con un proceso de enfriamiento de unos cientos de grados, decreciendo el flujo geotérmico en un factor de tres, al tiempo que el calor de desintegración ha ido también disminuyendo. El resultado de este proceso es un engrosamiento de la litosfera, siendo el 75 por 100 de las pérdidas de calor producidas por el enfriamiento concomitante a la creación de nuevo suelo oceánico.

En cuanto al estado actual, la distribución de temperaturas en el interior no puede realizarse a partir de los datos de flujo geotérmico, ya que es necesario conocer la concentración de radioisótopos en la corteza y manto superior, y el mecanismo de convección actuante en el manto. Por estas razones, el perfil de temperaturas es mucho más impreciso que el correspondiente a otros datos, como puede ser la densidad o constantes elásticas.

Consideremos únicamente el manto y núcleo en la investigación del perfil de temperaturas. Como el material del manto es sólido, su temperatura de fusión parcial proporciona un límite superior no alcanzable a cualquier profundidad. Por tanto, es necesario un conocimiento de la variación de la temperatura de fusión parcial del material con la profundidad, existiendo gran cantidad de teorías al respecto. Entre las relaciones empíricas más satisfactorias, puede citarse la

correspondiente a Fraut y Kennedy, propuesta en 1966, que relaciona la temperatura de fusión parcial T_m , la correspondiente a presión cero T_0 y la reducción volumétrica a la presión considerada $\Delta V/V_0$:

$$T_m / T_0 = 1 + b\Delta V / V_0$$

siendo b una constante. Esta expresión, sin embargo, tiende a sobrestimar el punto de fusión parcial de los minerales que componen el manto. Por ejemplo, la fosforita funde a 2.160 °K en la superficie, lo que proporciona un gradiente de 1,5 °K/km con la profundidad.

Para las partes más profundas del manto se utiliza el estudio de peridotitas, conteniendo ciertos compuestos como ortopiroxeno, que proporcionan temperaturas a 130-170 km del orden de 1.170-1.290 °K, respectivamente. La temperatura a 400 km, tomando en consideración el olivino y la relación presión-temperatura, se estima en 1.700-1.900 °K.

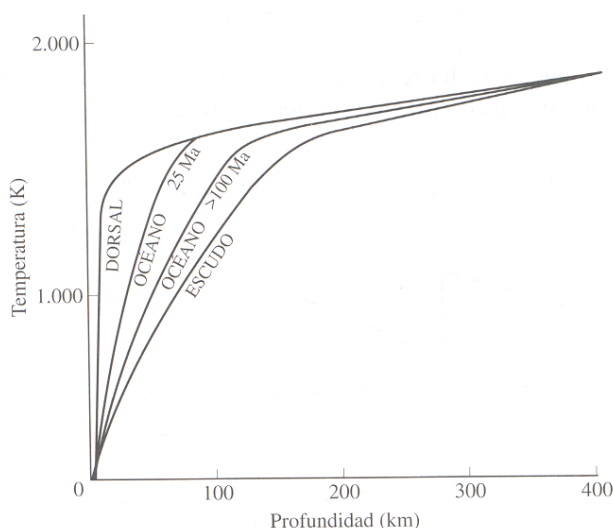


Figura 3. Distribución de temperaturas en el interior del Manto para distintas zonas.

Fuente: Udias y Mezcuá, 1997.

Mediante estos criterios es posible, con grandes reservas, establecer tres o cuatro temperaturas en el manto, siendo el resto del perfil una consecuencia de las hipótesis sobre distribución de radioisótopos y de los mecanismos de transferencia de calor aplicables. En la figura 4 se muestran las distribuciones de temperatura estimadas para el manto superior para distintos tipos de regiones.

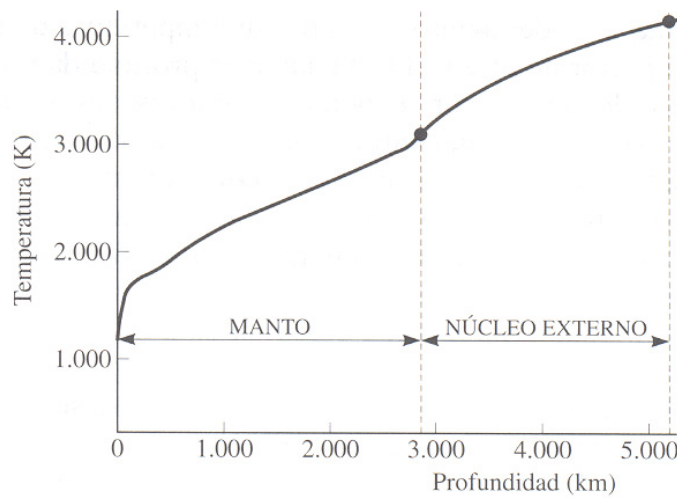


Figura 4. Distribución de Temperaturas en el interior del Manto y del Núcleo externo.

Fuente: Udías y Mezcuá, 1997.

En la figura 5 se presenta de forma esquemática la distribución en el manto inferior, incluyendo además la correspondiente al núcleo externo. La estimación de temperaturas en el núcleo se ha realizado por extrapolación de la temperatura de fusión parcial del hierro a gran presión, aunque la presencia de otros minerales más ligeros complica aún más el análisis. Por otra parte, la presencia de un núcleo externo líquido presupone que la temperatura en esta zona está por encima de la temperatura de fusión parcial. Este hecho ha permitido establecer una hipótesis plausible para el origen del núcleo interno, como consecuencia de una solidificación, al irse enfriando progresivamente el núcleo interior.

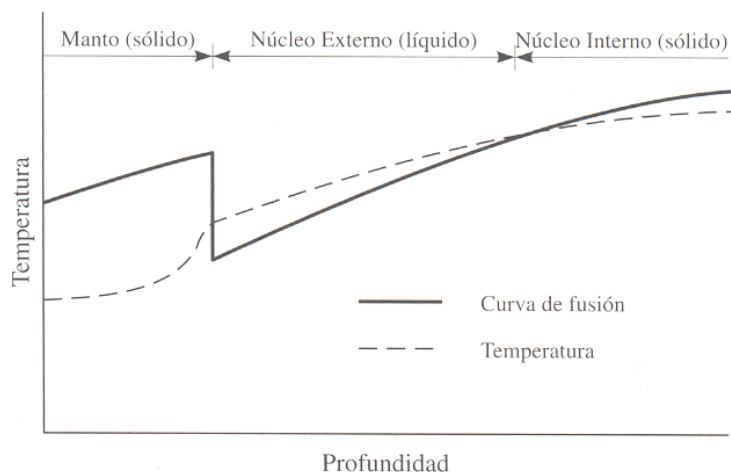


Figura 5. Distribución de temperatura en el Núcleo Externo e Interno y curva de fusión del material.

Fuente: Udías y Mezcuá, 1997.

Bibliografía

- [1] A. Udías y J.Mezcua. (1997). Fundamentos de Geofísica. Ed. Alianza Universidad. CAP.21.
- [2] W. Lowrie. (2000). Fundamentals of Geophysics. Ed. Cambridge Press.CAP.4.