

Matemáticas Avanzadas:

de los espacios lineales al análisis vectorial,
con aplicaciones en Máxima

$$\begin{aligned}\varepsilon^{ijk} a_j \varepsilon_{kmn} b^m c^n &= \varepsilon^{ijk} \varepsilon_{kmn} a_j b^m c^n = \varepsilon^{ijk} \varepsilon_{mnk} a_j b^m c^n \\ (\delta_m^i \delta_n^j - \delta_m^j \delta_n^i) a_j b^m c^n &= \delta_m^i \delta_n^j a_j b^m c^n - \delta_m^j \delta_n^i a_j b^m c^n \\ \delta_m^i b^m \delta_n^j a_j c^n - \delta_n^i c^n \delta_m^j a_j b^m &= b^i a_n c^n - c^i a_j b^j\end{aligned}$$

Héctor Hernández

Luis. A. Núñez

19 de septiembre de 2021

Índice general

1. Los vectores de siempre	2
1.1. Vectores geométricos	2
1.2. Vectores en componentes	15
1.3. Aplicaciones del álgebra vectorial	25
1.4. Álgebra vectorial con índices	33
1.5. Un comienzo a la derivación e integración de vectores	42
1.6. Vectores y números complejos	64
2. Espacios vectoriales lineales	81
2.1. Grupos, campos y espacios vectoriales	81
2.2. Espacios métricos, normados y con producto interno	96
2.3. Variedades lineales	112
2.4. Aproximación de funciones	128
3. Vectores duales y tensores	146
3.1. Funcionales lineales	146
3.2. Tensores y producto tensorial	159
3.3. Aplicaciones en Física: Esfuerzos, Inercia y Energía libre de un medio elástico	185
3.4. Vectores, tensores y espacios pseudoeuclidianos	197
3.5. Funcionales lineales y Distribuciones	208
4. Matrices, determinantes y autovectores	226
4.1. Operadores lineales	226
4.2. Tipos de operadores	237
4.3. Representación matricial de operadores	249
4.4. Un zoológico de matrices cuadradas	270
4.5. Sistemas de ecuaciones lineales	278
4.6. Autovectores y autovalores	287
4.7. Autovalores y autovectores de matrices importantes	304

Introducción

Como toda obra, el contenido de este libro tuvo como motivación inicial la insatisfacción con lo que estaba disponible y esa necesidad de discutir un conjunto de conceptos con el matiz personal de los autores. A lo largo de casi 10 años se fue dibujando esa ruta, donde confluyen: una presentación abstracta, una herramienta de cálculo algebraico y variados ejemplos de aplicación proveniente de la más diversas áreas. Hemos tratado de mostrar que los conceptos abstractos son útiles porque engloban, bajo un mismo enfoque, una multiplicidad de aplicaciones que normalmente las percibimos aisladas.

Existe, desde hace mucho tiempo, cierta resistencia en los Departamentos de Física en incluir en sus programas de docencia cursos para la enseñanza de herramientas de computación científica y cálculo numérico, tal vez debido a los altos costos de la mayoría de estos programas, casi todos comerciales. Pensamos que la utilización de herramientas computacionales enriquece enormemente el aprendizaje de los estudiantes ya que los enseñan a abordar los problemas desde diferentes puntos de vista, les ayuda a familiarizarse con determinados lenguajes de programación y los incentiva a desarrollar sus propias técnicas de cálculo.

Hemos decidido utilizar un sistema de computación algebraico de dominio público (o software libre) como **Maxima** porque estos programas tienen la capacidad de ofrecer al estudiante toda una gama de herramientas de cálculo simbólico, numérico y de visualización de muy alto rendimiento. **Maxima** es un programa que no requiere conocimientos previos de lenguajes de programación pero permitirá que el estudiante se familiarice con la sintaxis de programación y la transición a los lenguajes del tipo Fortran, C, o Python será mas sencilla.

Los contenidos que aquí presentamos han sido utilizados en cursos de Métodos Matemáticos para estudiantes de pregrado en Física de la Universidad de los Andes (Mérida-Venezuela) y más recientemente en los cursos de pregrado y posgrado para estudiantes en Física e Ingenierías en la Universidad Industrial de Santander (Bucaramanga-Colombia).

Los vectores de siempre

La ruta de este capítulo

Suponemos que los estudiantes que se acercan a estas notas, no solo conocen parte de los términos aquí expuestos, sino que son capaces de buscar muchos otros en la red. Por lo tanto, concebimos que este capítulo apunte a varios objetivos. Por un lado, buscamos refrescar un conjunto de conceptos básicos que seguramente son conocidos por el lector. Si no lo son, aprovechamos la oportunidad para presentarlos –en el marco de \mathbb{R}^3 , es decir exemplificando con vectores tridimensionales– utilizando el lenguaje abstracto al cual haremos referencia en los próximos capítulos.

Siguiendo esta lógica discutimos las propiedades de los vectores en la próxima sección 1.1; la independencia lineal, bases, producto interno (sección 1.1.3) y los sistemas de coordenadas (sección 1.2). Con la excusa del álgebra vectorial en coordenadas, introducimos en la sección 1.4 algunos elementos de álgebra vectorial con índices. Muchos de estos términos no son cubiertos tan tempranamente en cursos de métodos matemáticos (sección 1.4). Adicionalmente, esta excusa nos sirve de puente para presentar nociones operativas de tensores y de análisis de vectorial que formalizaremos más adelante en los capítulos 3 y ??, respectivamente.

La representación de los números complejos como vectores, con “componentes” reales y complejas justifica la incorporación de este repaso en la última sección 1.6. Finalmente, este capítulo nos sirve para iniciar el uso de la herramienta de cálculo algebraico que nos acompañará en el resto del libro y que se detalla en el apéndice ??.

1.1 Vectores geométricos

Desde los primeros cursos de Física en educación media, venimos hablando de vectores como cantidades que tienen que ser representadas con más de un número. Son varias las razones que obligan a introducir este (y otro) tipo de cantidades “multidimensionales”.

Es común que tropecemos con: escalares, vectores, tensores y espinores, dependiendo del número de cantidades que necesitemos para representar determinado objeto matemático. Podremos constatar que las leyes de la Física vienen escritas en forma vectorial (o tensorial) y, por lo tanto, será la misma ley para la familia de observadores equivalentes.

1.1.1 Escalares y vectores

Dejaremos para más adelante caracterizar objetos como tensores y espinores, por ahora nos contentaremos con refrescar nuestros recuerdos con cantidades como:

- **Escalares:** Serán aquellas cantidades las cuales se representan con UN solo número, una magnitud: temperatura, volumen, masa, entre otras. Es costumbre no denotarlas de manera especial, así $T = 5^{\circ}\text{C}$ representará una temperatura de 5 grados centígrados.
- **Vectores:** Serán cantidades las cuales, para ser representadas por un objeto matemáticos, necesitan más de una cantidad: requieren de UN número, UNA dirección y UN sentido. Entre las cantidades que típicamente reconocemos como vectores están: la velocidad, la aceleración, la fuerza. En términos gráficos podremos decir que un vector será un segmento orientado, en el cual la dimensión del segmento representará su módulo y su orientación la dirección y el sentido. Para diferenciarlos de las cantidades escalares hay una variedad de representaciones, entre ellas: en negrita \mathbf{a} ; con una flecha arriba de la cantidad \vec{a} ; con una tilde arriba o abajo \tilde{a} ; o explicitando el origen del segmento orientado \overrightarrow{OP} . El módulo del vector lo representaremos dentro de la función valor absoluto, o sencillamente sin la flecha arriba $a = |\mathbf{a}| = |\vec{a}|$.

Los vectores son independientes del sistema de coordenadas. Sus características (módulo, dirección y sentido) se preservarán en todos los sistemas de coordenadas. Más aún, habrá vectores que podremos desplazar (conservando su módulo dirección y sentido) paralelos a ellos mismos, y seguirán representando las mismas cantidades. Por ello encontraremos el término de *vectores deslizantes*. Un ejemplo son las fuerzas que actúan en un determinado cuerpo, como se muestra en el cuadrante I en la figura 1.2.

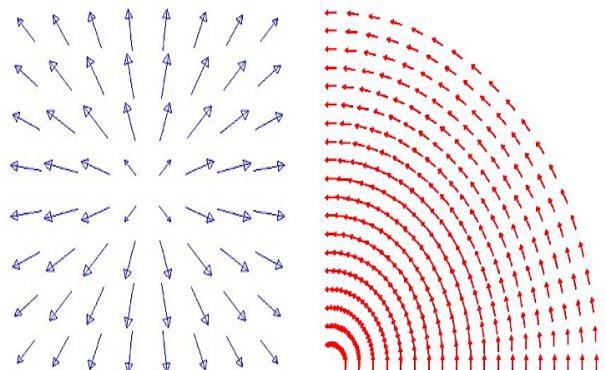


Figura 1.1: Ejemplos de *vectores atados*

También habrá vectores atados a un punto en el espacio, por cuanto representan una de las propiedades de ese punto: la velocidad del viento, el campo eléctrico, o sus variaciones son ejemplos de *vectores atados* (observe la figura 1.1 como ejemplos ilustrativos).

1.1.2 Álgebra de vectores

Enumeraremos rápidamente el álgebra de vectores sin hacer referencia a un sistema de coordenadas en particular. Desde los cursos básicos de matemáticas nos enseñaron a representar gráficamente este álgebra, así tenemos que:

1.1.2.0.1. Vector nulo. Es aquel que tiene por módulo cero y no se le puede asignar dirección ni sentido. El frecuente representar al vector nulo por $\mathbf{0}$.

1.1.2.0.2. Vector unitario. Es aquel que tiene por módulo la unidad, es muy útil por cuanto, para efectos algebraicos, “contiene” únicamente dirección y sentido. Lo denotaremos con un acento circunflejo: $\hat{\mathbf{u}}_a = \mathbf{a}/|\mathbf{a}|$, con lo cual todo vector se podrá expresar por un módulo en la dirección y sentido de un vector unitario: $\mathbf{a} = |\mathbf{a}| \hat{\mathbf{u}}_a = a \hat{\mathbf{u}}_a$.

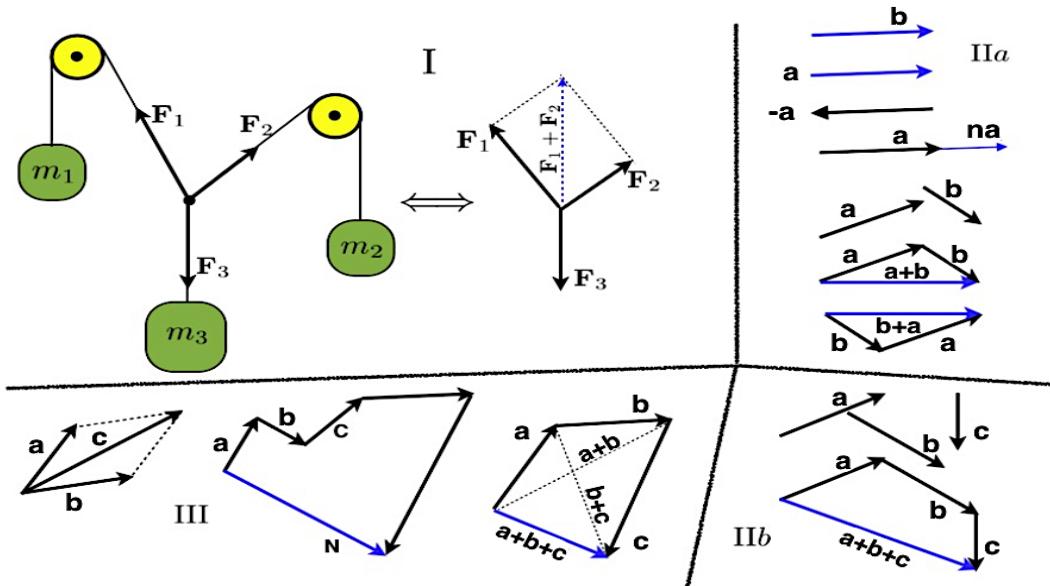


Figura 1.2: Vectores representados como segmentos de flechas orientados, vectores geométricos, y algunas de sus posibles operaciones.

1.1.2.0.3. Comparación de vectores. Dos vectores podrán ser comparados cuando tienen la misma dirección. Al comparar sus módulos diremos que pueden ser mayores, menores o iguales. Por lo tanto, tal y como mostramos en el cuadrante IIa de la figura 1.2, dos vectores serán iguales, $\mathbf{a} = \mathbf{b}$, si tienen la misma dirección y sentido.

1.1.2.0.4. Multiplicación por un número. Un vector multiplicado por un número, α , cambiará su módulo si $\alpha > 0$ y cambiará su sentido, y eventualmente su módulo, si $\alpha < 0$. Tal y como puede apreciarse en el cuadrante IIa de la figura 1.2. Claramente dos vectores proporcionales serán colineales. Diremos además, que el inverso del vector \mathbf{a} será la multiplicación de \mathbf{a} por (-1) . Esto es: $(-1)\mathbf{a} = -\mathbf{a}$.

1.1.2.0.5. Suma de vectores. Para sumar vectores utilizamos la regla del paralelogramo, es decir, desplazamos paralelamente uno de los vectores y lo colocamos a continuación del otro, de tal forma que la diagonal del paralelogramo, que tiene por lados los vectores sumandos, constituye el vector suma, (ver cuadrantes IIa y IIb de la figura 1.2).

Este esquema se puede generalizar para varios vectores tal y como lo mostramos en el cuadrante III de la figura 1.2. Allí construimos un polígono cuyos lados los constituyen los vectores sumandos $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ y \mathbf{n} con $\mathbf{n} = \mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}$.

Igualmente, podemos definir la resta de vectores al sumar el inverso. Esto es:

$$\mathbf{a} - \mathbf{b} \equiv \mathbf{a} + (-\mathbf{b}) \Rightarrow \mathbf{0} = \mathbf{a} - \mathbf{a} \equiv \mathbf{a} + (-\mathbf{a}).$$

En términos gráficos la resta de dos vectores se representa colocando los vectores (minuendo y sustraendo) con el mismo origen y uniendo las cabezas de flecha. Dependiendo de cual vector es el minuendo y cual el sustraendo el vector resta apuntará del sustraendo hacia el minuendo, esto es: $(\mathbf{a} + \mathbf{b} + \mathbf{c}) - \mathbf{a} = \mathbf{b} + \mathbf{c}$.

Claramente, el módulo del vector resta representa la distancia entre los dos extremos de los vectores minuendo y sustraendo.

1.1.2.0.6. Un resumen de propiedades. Las propiedades (obvias) del álgebra de vectores son:

- La suma de vectores:
 - es cerrada $\mathbf{a} + \mathbf{b} = \mathbf{c}$,
 - es comutativa $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$,
 - es asociativa $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c})$,
 - tiene un único elemento neutro $\mathbf{0} + \mathbf{a} = \mathbf{a} + \mathbf{0} = \mathbf{a}, \forall \mathbf{a}$,
 - existe un elemento simétrico $-\mathbf{a}$ (uno para cada vector) tal que $\mathbf{0} = \mathbf{a} - \mathbf{a} \equiv \mathbf{a} + (-\mathbf{a})$,
 - es distributiva respecto a la multiplicación por números: $\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$.
- La multiplicación de números por vectores:
 - es comutativa $\alpha\mathbf{a} = \alpha\mathbf{a}$,
 - es asociativa $\alpha(\beta\mathbf{a}) = (\alpha\beta)\mathbf{a}$,
 - es distributiva $(\alpha + \beta)\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$.

Estas propiedades lucen obvias planteadas en este contexto de los vectores en \mathbb{R}^3 . Sin embargo en la sección 2.1.3 en la página 87 veremos que, son estas propiedades las que nos permitirán definir un espacio vectorial abstractos de manera rigurosa.

Ejemplo 1.1 En un segmento de recta AB ubicamos un punto D de manera que este punto divide al segmento en dos partes, es decir, en la proporción $\alpha : \beta$ ($\overline{AB} = \overline{AD} + \overline{DB} = \alpha + \beta$). Ubiquemos el vector posición del punto D .

A los puntos A y B le hacemos corresponder el vector \mathbf{a} y el vector \mathbf{b} , respectivamente, con un origen común O . De manera que $\mathbf{c} = \mathbf{b} - \mathbf{a}$ es un vector que va desde el punto A al punto B . Entonces, la distancia \overline{OD} , a la que le podemos asociar el vector \mathbf{d} , no es más que:

$$\mathbf{d} = \overline{OD} = \mathbf{a} + \frac{\alpha}{\alpha + \beta}(\mathbf{b} - \mathbf{a}) = \left(1 - \frac{\alpha}{\alpha + \beta}\right)\mathbf{a} + \frac{\alpha}{\alpha + \beta}\mathbf{b} = \frac{\beta}{\alpha + \beta}\mathbf{a} + \frac{\alpha}{\alpha + \beta}\mathbf{b}.$$



Ejemplo 1.2 Hemos definido la posición \mathbf{r} del centro de masa, para un sistema de N partículas, al vector:

$$\mathbf{r} = \frac{\sum_{i=1}^N m_i \mathbf{r}_i}{\sum_{j=1}^N m_j},$$

donde \mathbf{r}_i corresponde con la posición de la i -ésima partícula. Determinemos la posición del centro de masa para un sistema de tres masas, $m_i = 1, 2, 3$, ubicadas en los vértices de un triángulo equilátero de lado $l = 2$.

Colocando el origen de coordenadas en uno de los vértices y uno de los ejes de coordenadas sobre uno de los lados, entonces tenemos:

$$\mathbf{r} = \frac{\sum_{i=1}^3 m_i \mathbf{r}_i}{\sum_{j=1}^3 m_j} = \frac{m_1 \mathbf{r}_1 + m_2 \mathbf{r}_2 + m_3 \mathbf{r}_3}{M_T} = \frac{1 \cdot 2\hat{\mathbf{i}} + 3 \cdot (\hat{\mathbf{i}} + \sqrt{3}\hat{\mathbf{j}})}{6} = \frac{5}{6}\hat{\mathbf{i}} + \frac{\sqrt{3}}{2}\hat{\mathbf{j}}.$$



1.1.3 Vectores linealmente independientes

Comencemos a construir la primera aproximación a uno de los conceptos fundamentales del álgebra lineal: la noción de *independencia* o *dependencia lineal*. Diremos que tres vectores $\mathbf{a}, \mathbf{b}, \mathbf{c}$ son *linealmente independientes* en \mathbb{R}^3 si se cumple que:

$$\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c} = \mathbf{0} \Rightarrow \alpha = \beta = \gamma = 0. \quad (1.1)$$

Es decir, que la única manera que al sumar cualquier múltiplo de \mathbf{a}, \mathbf{b} y \mathbf{c} de modo que la suma se anule es obligando a que los escalares: α, β y γ , sean **necesariamente** nulos. En caso contrario, es decir, si no se cumple

lo anterior diremos que uno de los vectores será *linealmente dependiente* y por lo tanto se podrá expresar como combinación lineal de los otros dos. Si por ejemplo $\gamma \neq 0$, entonces:

$$\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c} = \mathbf{0} \Rightarrow \mathbf{c} = -\frac{\alpha}{\gamma} \mathbf{a} - \frac{\beta}{\gamma} \mathbf{b} \Rightarrow \mathbf{c} = \bar{\alpha} \mathbf{a} + \bar{\beta} \mathbf{b}.$$

Es muy importante señalar que los vectores linealmente independientes formarán una *base* para el espacio donde estos vectores “viven” y el número máximo de vectores linealmente independientes será la dimensión de ese espacio de “residencia”. Más adelante estudiaremos con más detalle el concepto de bases.

Tratemos de concretar algunas de estas afirmaciones.

- *Dos vectores linealmente dependientes son colineales.* Es decir, los vectores están contenidos en la misma línea y es claro que:

$$\alpha \mathbf{a} + \beta \mathbf{b} = \mathbf{0} \quad \text{con alguno de } \begin{cases} \alpha \neq 0 \\ \beta \neq 0 \end{cases} \Rightarrow \begin{cases} \mathbf{a} = -\frac{\beta}{\alpha} \mathbf{b} \\ \mathbf{b} = -\frac{\alpha}{\beta} \mathbf{a} \end{cases}$$

el contrario también será cierto: *si dos vectores son colineales ellos serán linealmente dependientes*.

$$\mathbf{a} = \lambda \mathbf{b} \Rightarrow \alpha \mathbf{a} + \beta \mathbf{b} = \mathbf{0} \Rightarrow \alpha \lambda \mathbf{b} + \beta \mathbf{b} = \mathbf{0} \Rightarrow (\alpha \lambda + \beta) \mathbf{b} = \mathbf{0} \Rightarrow \lambda = -\frac{\beta}{\alpha},$$

con lo cual podremos afirmar que si dos vectores son linealmente independientes ellos **no** son colineales.

- *Tres vectores linealmente dependientes son coplanares.* Por ser los tres vectores *linealmente dependientes* al menos uno de los escalares tiene que ser distinto de cero, digamos γ , esto es:

$$\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c} = \mathbf{0} \Rightarrow \mathbf{c} = -\frac{\alpha}{\gamma} \mathbf{a} - \frac{\beta}{\gamma} \mathbf{b} = \xi^1 \mathbf{a} + \xi^2 \mathbf{b},$$

pero como $\xi^1 \mathbf{a} \propto \mathbf{a}$ y $\xi^2 \mathbf{b} \propto \mathbf{b}$, esto significa que: $\xi^1 \mathbf{a}$ y \mathbf{a} son colineales. De la misma manera que $\xi^2 \mathbf{b}$ y \mathbf{b} , y por lo tanto, la suma estará en el mismo plano.

- *Dos vectores linealmente independientes generan todos los vectores coplanares.* Dado dos vectores \mathbf{a} y \mathbf{b} linealmente independientes, entonces cualquier vector \mathbf{c} , coplanar con \mathbf{a} y \mathbf{b} , podrá expresarse como una combinación lineal de éstos. Diremos que: \mathbf{c} se expresa en términos de \mathbf{a} y \mathbf{b} como: $\mathbf{c} = \xi^1 \mathbf{a} + \xi^2 \mathbf{b}$, y esa expresión es única.

La primera de las afirmaciones es directa por cuanto hemos visto que si \mathbf{a} y \mathbf{b} son linealmente independientes y \mathbf{c} es coplanar con \mathbf{a} y \mathbf{b} , entonces, necesariamente \mathbf{a} , \mathbf{b} y \mathbf{c} son linealmente dependientes. Esto es:

$$\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c} = \mathbf{0} \Rightarrow \mathbf{c} = -\frac{\alpha}{\gamma} \mathbf{a} - \frac{\beta}{\gamma} \mathbf{b} = \xi^1 \mathbf{a} + \xi^2 \mathbf{b}.$$

La demostración de que la expansión es única viene de suponer que existen dos maneras distintas de representar al mismo vector \mathbf{c} . Veamos:

$$\left. \begin{array}{l} \mathbf{c} = \xi^1 \mathbf{a} + \xi^2 \mathbf{b} \\ \mathbf{c} = \zeta^1 \mathbf{a} + \zeta^2 \mathbf{b} \end{array} \right\} \Rightarrow \mathbf{0} = (\xi^1 - \zeta^1) \mathbf{a} + (\xi^2 - \zeta^2) \mathbf{b} \Rightarrow \left\{ \begin{array}{l} \xi^1 - \zeta^1 = 0 \Rightarrow \xi^1 = \zeta^1 \\ \xi^2 - \zeta^2 = 0 \Rightarrow \xi^2 = \zeta^2 \end{array} \right.$$

debido a que \mathbf{a} y \mathbf{b} son linealmente independientes.

La demostración para el caso tridimensional es equivalente. Es decir tres vectores linealmente independientes \mathbf{a} , \mathbf{b} y \mathbf{c} expanden, de manera unívoca, todos los vectores del espacio. Esta demostración queda para el lector.

- *Vectores Base.* Cuando un vector \mathbf{c} se pueda expresar en términos de dos vectores linealmente independientes, \mathbf{a} y \mathbf{b} , por ejemplo: $\mathbf{c} = \xi^1 \mathbf{a} + \xi^2 \mathbf{b}$, diremos que \mathbf{a} y \mathbf{b} forman una base para todos los vectores

coplanares a éstos. Igualmente para el caso tridimensional: tres vectores linealmente independientes \mathbf{a} , \mathbf{b} y \mathbf{c} conformarán una base para los vectores del espacio. Los números ξ^1 y ξ^2 para el caso bidimensional se denominan las componentes de \mathbf{c} a lo largo de \mathbf{a} y \mathbf{b} . Equivalentemente, (ξ^1, ξ^2, ξ^3) serán las componentes de cualquier vector para el caso 3D a lo largo de \mathbf{a} , \mathbf{b} y \mathbf{c} , respectivamente. Esta nomenclatura será más evidente luego de la próxima sección.

Mas adelante, en el próximo capítulo en el desarrollo de la sección 2.3 analizaremos de forma rigurosa el concepto de independencia lineal y de bases de espacios vectoriales abstractos.

Ejemplo 1.3 Dada una base ortonormal $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ y los siguientes vectores:

$$\mathbf{a} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}, \quad \mathbf{b} = 3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}, \quad \mathbf{c} = \hat{\mathbf{i}} - \hat{\mathbf{k}}.$$

Queremos comprobar si $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ forman una base.

Veamos, para que los vectores formen una base tienen que ser linealmente independientes. Esto es: $\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = 0 \Rightarrow \alpha = \beta = \gamma = 0$, con lo cual:

$$\alpha(3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \beta(3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \gamma(\hat{\mathbf{i}} - \hat{\mathbf{k}}) = 0 \Rightarrow \begin{cases} 3\alpha + 3\beta + \gamma = 0 \\ 2\alpha - 2\beta = 0 \\ \alpha + \beta - \gamma = 0 \end{cases}$$

es fácil ver que la solución de este sistema es: $\alpha = \beta = \gamma = 0$, por lo tanto, se demuestra que los vectores son linealmente independientes y por consiguiente forman una base.



Ejemplo 1.4 Si el conjunto de vectores $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ del ejemplo anterior forman una base, podemos expresar otros vectores en términos de esta base. Tomemos, por ejemplo, los vectores: $\mathbf{d} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}$ y $\mathbf{e} = 3\hat{\mathbf{i}} - 2\hat{\mathbf{j}}$, expresemos estos dos vectores en términos de $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

Entonces, para el vector \mathbf{d} tenemos:

$$\mathbf{d} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} = \alpha(3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \beta(3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \gamma(\hat{\mathbf{i}} - \hat{\mathbf{k}}) \Rightarrow \begin{cases} 3\alpha + 3\beta + \gamma = 1 \\ 2\alpha - 2\beta = 2 \\ \alpha + \beta - \gamma = 0 \end{cases}$$

resolviendo el sistema de ecuaciones anterior tendremos que: $\mathbf{d} = \frac{5}{8}\mathbf{a} - \frac{3}{8}\mathbf{b} + \frac{1}{4}\mathbf{c}$.

Seguidamente, para el vector \mathbf{e} se tiene:

$$\mathbf{e} = 3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} = \alpha(3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \beta(3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \gamma(\hat{\mathbf{i}} - \hat{\mathbf{k}}) \Rightarrow \begin{cases} 3\alpha + 3\beta + \gamma = 3 \\ 2\alpha - 2\beta = -2 \\ \alpha + \beta - \gamma = 0 \end{cases}$$

resolviendo el nuevo sistema de ecuaciones resulta que: $\mathbf{e} = -\frac{1}{8}\mathbf{a} + \frac{7}{8}\mathbf{b} + \frac{3}{4}\mathbf{c}$.



Ejemplo 1.5 Consideremos los siguientes tres vectores:

$$\mathbf{w}_1 = \hat{\mathbf{i}} + 3\hat{\mathbf{k}}, \quad \mathbf{w}_2 = 2\hat{\mathbf{i}} - 3\hat{\mathbf{j}} \text{ y } \mathbf{w}_3 = -\hat{\mathbf{j}} + \hat{\mathbf{k}}.$$

¿Formarán una base para \mathbb{R}^3 ?

Veamos entonces si son linealmente independientes:

$$\alpha\mathbf{w}_1 + \beta\mathbf{w}_2 + \gamma\mathbf{w}_3 = 0 \Rightarrow \alpha = \beta = \gamma = 0,$$

La comprobación es directa al resolver el siguiente sistema de ecuaciones:

$$\begin{array}{rcl} \alpha & + 2\beta & = 0 \\ & -3\beta & - \gamma = 0 \\ 3\alpha & + \gamma & = 0 \end{array}$$

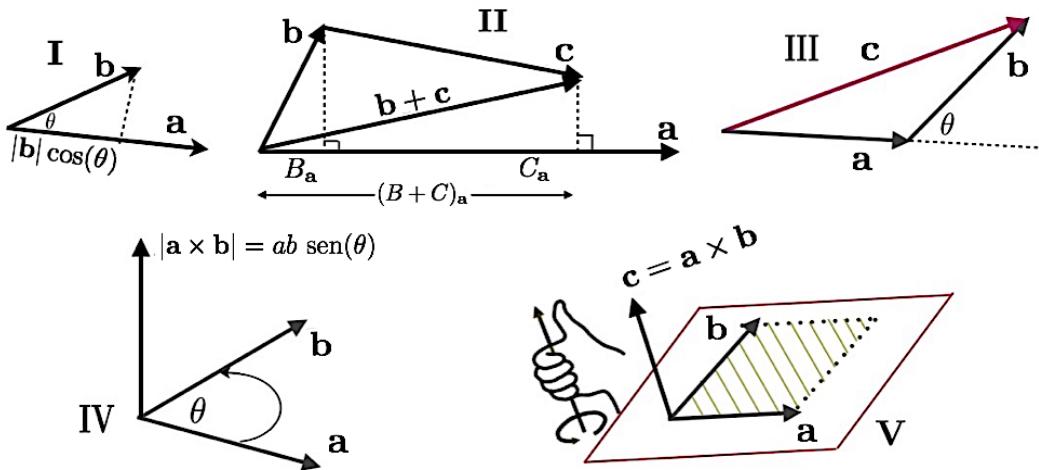


Figura 1.3: Producto escalar y vectorial de vectores geométricos

cuya solución es $\alpha = \beta = \gamma = 0$. Por lo tanto, forman una base para \mathbb{R}^3 .

Como forman base, podemos expresar otro vector, digamos: $\mathbf{a} = \hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$, en término de la base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$, esto es: $\mathbf{a} = \alpha\mathbf{w}_1 + \beta\mathbf{w}_2 + \gamma\mathbf{w}_3$

Si procedemos como en los ejemplos anteriores, tenemos:

$$\mathbf{a} = \alpha(\hat{\mathbf{i}} + 3\hat{\mathbf{k}}) + \beta(2\hat{\mathbf{i}} - 3\hat{\mathbf{j}}) + \gamma(-\hat{\mathbf{j}} + \hat{\mathbf{k}}) \Rightarrow \begin{cases} \alpha + 2\beta &= 1 \\ -3\beta - \gamma &= -3 \\ 3\alpha + \gamma &= 3 \end{cases} \Rightarrow \begin{cases} \alpha = \frac{1}{3} \\ \beta = \frac{1}{3} \\ \gamma = 2 \end{cases}$$

es decir:

$$\mathbf{a} = \frac{1}{3}\mathbf{w}_1 + \frac{1}{3}\mathbf{w}_2 + 2\mathbf{w}_3.$$

△

1.1.4 Productos de vectores

Hemos sumado y restado vectores. El siguiente paso es multiplicarlos. Básicamente existen dos formas de multiplicar vectores: el producto escalar y el producto vectorial. Veremos a continuación de qué se trata y sin especificar un sistema de coordenadas para referirlos.

1.1.4.1 Producto escalar

Denominaremos producto escalar de dos vectores \mathbf{a} y \mathbf{b} a un escalar cuyo valor será igual al producto de los módulos multiplicado por el coseno del ángulo que ellos forman:

$$\zeta = \mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle}. \quad (1.2)$$

El significado geométrico del producto escalar es evidente, cuadrante I de la figura 1.3. El producto escalar representa la proyección de \mathbf{a} sobre \mathbf{b} y equivalentemente la proyección de \mathbf{b} sobre \mathbf{a} .

De esta definición se derivan varias consecuencias las cuales por obvias no dejan de ser importantes:

- *El producto escalar de un vector consigo mismo, siempre es positivo:*

$\zeta = \mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 \geq 0$, y sólo será nulo si \mathbf{a} es el vector nulo. Esto es, $\zeta = 0 \Rightarrow \mathbf{a} = \mathbf{0}$. Con esto podemos concluir que $|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{\zeta}$.

- *El producto escalar es commutativo:*

$\zeta = \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$, ya que el ángulo entre los vectores es el mismo y la multiplicación entre escalares es commutativa.

- *El producto escalar es distributivo:*

Esto es, $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$. La demostración (gráfica) puede apreciarse en el cuadrante II de la figura 1.3.

- *La multiplicación por un número:*

$$\bar{\zeta} = \alpha\zeta = |\alpha|(\mathbf{a} \cdot \mathbf{b}) = (\alpha\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha\mathbf{b}) = |\alpha|\mathbf{a}|\mathbf{b}|\cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} = |\mathbf{a}||\alpha\mathbf{b}|\cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle}.$$

- *Desigualdad de Cauchy-Schwarz.*

A partir de la definición de producto interno es inmediata la comprobación de la siguiente desigualdad:

$$(\mathbf{a} \cdot \mathbf{b})^2 = (|\mathbf{a}||\mathbf{b}|\cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle})^2 \Rightarrow (\mathbf{a} \cdot \mathbf{b})^2 \leq |\mathbf{a}|^2|\mathbf{b}|^2 \Leftrightarrow \mathbf{a} \cdot \mathbf{b} \leq |\mathbf{a}||\mathbf{b}|,$$

ya que: $0 \leq \cos^2(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} \leq 1$.

- *Del producto escalar surge el teorema del coseno.*

Es inmediato calcular el producto escalar de un vector consigo mismo, para ello vamos a suponer que $\mathbf{c} = \mathbf{a} + \mathbf{b}$, con lo cual:

$$\mathbf{c} = \mathbf{a} + \mathbf{b} \Rightarrow \mathbf{c} \cdot \mathbf{c} = (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \Rightarrow |\mathbf{c}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|\mathbf{a}||\mathbf{b}|\cos(\theta),$$

donde θ es el ángulo que forman los vectores \mathbf{a} y \mathbf{b} . Esto no es otra cosa que el teorema del coseno y está ilustrado en el cuadrante III de la figura 1.3.

- *Dos vectores no nulos son ortogonales (perpendiculares) si su producto escalar es nulo.*

Esta afirmación es inmediata:

$$\mathbf{a} \perp \mathbf{b} \Rightarrow \theta_{\langle \mathbf{a}, \mathbf{b} \rangle} = \frac{\pi}{2} \Rightarrow \mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} = 0.$$

El producto escalar lo hemos visto desde nuestras tempranas épocas de colegio. Quizá por eso no apreciamos las consecuencias que su definición nos trae. El producto escalar nos permite definir una geometría, vale decir el ángulo entre vectores. A partir de esta definición sabemos si dos vectores son o no ortogonales. Otra vez, en \mathbb{R}^3 , no nos impactan mucho esa consecuencia. En la sección 2.2.3, veremos como la definición de un producto interno (en \mathbb{R}^3 lo llamamos escalar) nos permitirá definir distancia entre dos puntos, el tamaño de los vectores y el ángulo que ellos mantienen entre sí.

1.1.4.2 Producto vectorial

A diferencia del producto escalar que genera un escalar, el producto vectorial tiene como resultado otro vector: $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ (realmente un pseudovector o vector axial en contraposición a los vectores polares, pero eso lo veremos más adelante en la sección 1.4.4), con las siguientes características:

- El módulo de \mathbf{c} , será:

$$|\mathbf{c}| = |\mathbf{a}||\mathbf{b}|\sin(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle}. \quad (1.3)$$

Es claro que el módulo de \mathbf{c} representa el área del paralelogramo cuyos lados están formados por \mathbf{a} y \mathbf{b} (ver el cuadrante V de la figura 1.3).

- Tal y como muestran los cuadrantes IV y V de la figura 1.3, \mathbf{c} tendrá como dirección la perpendicular al plano que forman \mathbf{a} y \mathbf{b} , y como sentido la regla del pulgar derecho, regla de la mano derecha, o de

manera más elegante, será positiva cuando la multiplicación de $\mathbf{a} \times \mathbf{b}$ corresponda al sentido antihorario.

Podemos deducir algunas consecuencias de esta definición.

- *El producto vectorial es anticomutativo.*

$\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$, y se sigue de la definición que expresa el cuadrante IV de la figura 1.3.

- *El producto vectorial es distributivo respecto a la suma.*

$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}$. La demostración de esto lo dejaremos para más adelante.

- *La multiplicación por un número.*

$$|\mathbf{c}| = |\alpha| |\mathbf{a} \times \mathbf{b}| = |(\alpha \mathbf{a}) \times \mathbf{b}| = |\mathbf{a} \times (\alpha \mathbf{b})| = |\alpha \mathbf{a}| |\mathbf{b}| \sin(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} = |\mathbf{a}| |\alpha \mathbf{b}| \sin(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle}.$$

- *Dos vectores serán colineales si su producto vectorial se anula.*

Como en el caso cuando se anulaba el producto escalar identificábamos a dos vectores ortogonales, cuando se anula el producto vectorial tendremos dos vectores paralelos. Es claro que esto se cumple de inmediato:

$$\mathbf{a} \parallel \mathbf{b} \Rightarrow \theta_{\langle \mathbf{a}, \mathbf{b} \rangle} = 0 \Rightarrow |\mathbf{c}| = |\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} = 0.$$

Si el módulo del vector es cero, obvio que es el vector nulo. Ahora bien, también de aquí deducimos que:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \Rightarrow \mathbf{c} \cdot \mathbf{a} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = \mathbf{c} \cdot \mathbf{b} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0.$$

1.1.5 Producto triple o mixto

Analicemos ahora el número (pseudoescalar) que proviene de la multiplicación:

$$V = |\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})| = |\mathbf{c}| |(\mathbf{a} \times \mathbf{b})| \cos(\theta)_{\langle \mathbf{c}, \mathbf{a} \times \mathbf{b} \rangle}. \quad (1.4)$$

Este producto también cumple con algunas propiedades que enunciaremos ahora y demostraríremos más tarde.

- *El producto mixto representa el volumen del paralelepípedo cuyos lados son los vectores \mathbf{a} , \mathbf{b} y \mathbf{c} .*

$|\mathbf{a} \times \mathbf{b}|$ representa el área de la base y la altura está representada por la proyección del vector \mathbf{c} sobre la perpendicular al plano de la base que es, precisamente $|\mathbf{c}| \cos(\theta)_{\langle \mathbf{c}, \mathbf{a} \times \mathbf{b} \rangle}$.

- *El producto mixto es cíclico respecto a sus factores.*

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = (\mathbf{c} \times \mathbf{a}) \cdot \mathbf{b} = (\mathbf{b} \times \mathbf{c}) \cdot \mathbf{a}. \quad (1.5)$$

Esta afirmación se verá demostrada más adelante.

- *El producto mixto se anula cuando se repite alguno de sus factores.*

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = (\mathbf{a} \times \mathbf{a}) \cdot \mathbf{c} = (\mathbf{b} \times \mathbf{b}) \cdot \mathbf{c} = 0.$$

Claramente, si $(\mathbf{a} \times \mathbf{b}) \perp \mathbf{a} \Rightarrow (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} = 0$.

- *Si los tres vectores \mathbf{a} , \mathbf{b} y \mathbf{c} son coplanares (linealmente dependientes) entonces:*

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = 0.$$

Dicho de manera más elegante, útil e impactante: tres vectores que cumplen con:

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} \neq 0,$$

son linealmente independientes y forman una base para el espacio tridimensional. Esa base se denominará levógira (contraria al giro de las manecillas del reloj) si el producto $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} < 0$ y dextrógira (la convencional base de la mano derecha) si $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} > 0$.

Ejemplo 1.6 Resolvamos el problema del ejemplo 1.3 usando el producto mixto

Debemos calcular: $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) \neq 0$, y efectivamente:

$$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = \begin{vmatrix} 1 & 0 & -1 \\ 3 & 2 & 1 \\ 3 & -2 & 1 \end{vmatrix} = 4 \neq 0,$$

es decir, el conjunto de vectores son linealmente independientes.



Ejemplo 1.7 Ahora bien, si $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ del ejemplo 1.3 forman una base, calculemos al vector $\mathbf{f} = \mathbf{a} \times \mathbf{b}$ y expresémoslo en términos de $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$.

Procedemos a calcular \mathbf{f} :

$$\mathbf{f} = \mathbf{a} \times \mathbf{b} = (3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) \times (3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ 3 & 2 & 1 \\ 3 & -2 & 1 \end{vmatrix} = 4\hat{\mathbf{i}} - 12\hat{\mathbf{k}},$$

con lo cual para el vector \mathbf{f} resulta:

$$\mathbf{f} = 4\hat{\mathbf{i}} - 12\hat{\mathbf{k}} = \alpha(3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \beta(3\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + \hat{\mathbf{k}}) + \gamma(\hat{\mathbf{i}} - \hat{\mathbf{k}}) \Rightarrow \begin{cases} 3\alpha + 3\beta + \gamma = 4 \\ 2\alpha - 2\beta = 0 \\ \alpha + \beta - \gamma = -12 \end{cases}$$

y finalmente, al resolver el sistema anterior, $\mathbf{f} = \mathbf{a} \times \mathbf{b} = -\mathbf{a} - \mathbf{b} + 10\mathbf{c}$.



1.1.6 Ejercicios

1. Dado el triángulo cuyos vértices son los: $(2, 3)$, $(6, 9)$, $(8, 1)$. Encuentre:
 - (a). el baricentro, es decir, el punto donde se interceptan las medianas del triángulo.
 - (b). el circuncentro, es decir, el punto donde se interceptan las mediatrices del triángulo.
 2. Utilice métodos vectoriales para probar que las líneas que unen los puntos medios con las aristas opuestas (bimedias) de un tetraedro $OABC$ se interceptan en un punto, y que ese punto divide en dos partes cada una de las líneas.
 3. Los vértices de un triángulo ABC tienen como vectores posición \mathbf{a} , \mathbf{b} y \mathbf{c} , respectivamente y relativos a un origen común O . Demuestre que el vector posición \mathbf{g} del centroide G del triángulo viene dado por:
- $$\mathbf{g} = \frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c}).$$
4. Un paralelogramo tiene un ángulo agudo de $\pi/3$ y lados de longitud $a = 1$ y $b = 2$. Si pensamos que esos lados como vectores \mathbf{a} y \mathbf{b} encuentre:
 - (a). Los vectores: $\mathbf{a} + \mathbf{b}$ y $\mathbf{a} - \mathbf{b}$.
 - (b). Los vectores: $2\mathbf{a} + 3\mathbf{b}$ y $5\mathbf{a} - 7\mathbf{b}$.
 5. Con la definición de posición de centro de masa del ejemplo 1.2, encuentre el centro de masas para los siguientes sistemas:
 - (a). Masas iguales a: 1, 2, 3, 4 en los vértices de un cuadrado de lados $a = 2$.
 - (b). Masas iguales a: 1, 2, 3, 4 en los vértices inferiores de un cubo cuyos lados son de longitud $a = 2$ y masas iguales a: 5, 6, 7, 8 en la vértices superiores.
 6. ¿Los siguientes vectores son linealmente independientes?

$$\mathbf{a} = (0, 2, -1), \mathbf{b} = (0, 1/2, -1/2), \mathbf{c} = (0, -2/3, -1/3).$$

7. Las componentes de un vector y la regla para sumar vectores se combinan para introducir la forma

más simple de representar un vector como una combinación lineal de los vectores más elementales que podemos tener. Estos vectores forman lo que conocemos la base canónica: $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$, vectores de longitud unitaria que apuntan en la dirección positiva de los ejes x , y y z .

Compruebe, entonces, si los siguientes vectores forman una base:

- (a). $\mathbf{w}_1 = 2\hat{\mathbf{i}} + \hat{\mathbf{j}} - 3\hat{\mathbf{k}}$, $\mathbf{w}_2 = \hat{\mathbf{i}} - 4\hat{\mathbf{k}}$, $\mathbf{w}_3 = 4\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - \hat{\mathbf{k}}$
- (b). $\mathbf{w}_1 = \hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 2\hat{\mathbf{k}}$, $\mathbf{w}_2 = 2\hat{\mathbf{i}} - 4\hat{\mathbf{j}} - \hat{\mathbf{k}}$, $\mathbf{w}_3 = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} - \hat{\mathbf{k}}$

8. Un paralelogramo tiene un ángulo agudo de $\pi/4$ y lados $a = 1, b = 2$. Si consideramos que los lados son vectores, encuentre:

- (a). El área del paralelogramo.
- (b). La proyección de cada lado sobre la dirección del otro.

9. Considere un triángulo cuyos lados están conformados por los vectores \mathbf{a}, \mathbf{b} y $\mathbf{c} = \mathbf{a} + \mathbf{b}$. Con el producto vectorial entre ellos demuestre la ley del seno:

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \beta} = \frac{c}{\sin \gamma}.$$

donde α, β, γ son los ángulos opuestos a los lados a, b, c respectivamente.

10. Demuestre que el volumen de un tetraedro formado a partir de tres vectores \mathbf{a}, \mathbf{b} y \mathbf{c} que coinciden en un mismo origen, puede representarse de la manera siguiente:

$$V = \frac{1}{6} |\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})|.$$

Practicando con Maxima

El programa de manipulación simbólica **Maxima** está diseñado para realizar cálculos algebraicos y numéricos que iremos descubriendo a medida que desarrollemos los diferentes temas de este curso. Es indispensable ir al apéndice ?? para familiarizarnos con la sintaxis básica del programa.

Maxima es una potente calculadora y maneja números de diferentes tipos: enteros, racionales, irracionales, complejos y números con decimales (punto flotante).

Haremos algunos cálculos sencillos para ir familiarizándonos con la sintaxis, por ejemplo, pidamos al programa que nos diga cuánto vale $\ln(20)$. Para eso debemos escribir:

```
(%i1) log(20);
(%o1) log(20)
```

Como podemos ver, el programa nos devuelve el valor exacto de $\ln(20)$, para obtener su valor numérico podemos escribir lo siguiente

```
(%i2) log(20),numer;
(%o2) 2,995732273553991
```

Si tenemos una fracción:

```
(%i3) 420/16000;
(%o3) 21
(%o3) 800
(%i4) 420/16000,numer;
(%o4) 0,02625
```

Podemos utilizar la función **float** para el mismo resultado

```
(%i5) float(log(20)); float(420/16000);
```

(%o5) 2,995732273553991

(%o6) 0,02625

El programa contiene una gran cantidad de funciones matemáticas básicas internas, y por supuesto, entre ellas están las trigonométricas:

(%i7) $\sin(\pi/3); \cos(\pi/3); \tan(\pi/3);$

$$(\%o7) \frac{\sqrt{3}}{2}$$

$$(\%o8) \frac{1}{2}$$

$$(\%o9) \sqrt{3}$$

Ahora recurriremos a una de las facilidades de estructura de datos que nos ofrece el sistema para agrupar objetos matemáticos: las listas.

Las listas son un arreglo de cantidades que se escriben entre corchetes y los objetos deben separarse con comas.

(%i10) $[\sin(\pi/3), \cos(\pi/3), \tan(\pi/3)];$

$$(\%o10) \left[\frac{\sqrt{3}}{2}, \frac{1}{2}, \sqrt{3} \right]$$

(%i11) $[\sin(\pi/3), \cos(\pi/3), \tan(\pi/3)], \text{numer};$

$$(\%o11) [0,8660254037844386, 0,5000000000000001, 1,732050807568877]$$

Cuando necesitemos generar una lista por medio de alguna regla específica o fórmula usamos la función **makelist**. La sintaxis es la siguiente:

(%i12) $\text{makelist}(\exp(t*x), t, 1, 10);$

$$(\%o12) [e^x, e^{2x}, e^{3x}, e^{4x}, e^{5x}, e^{6x}, e^{7x}, e^{8x}, e^{9x}, e^{10x}]$$

Podemos también aplicar una función a cada elemento de la lista, en este caso a cada elemento le aplicaremos la función $\ln(x)$. Para tal fin utilizaremos el comando **map**.

(%i13) $\text{map}(\log, (\text{makelist}(\exp(t*x), t, 1, 10)));$

$$(\%o13) [x, 2x, 3x, 4x, 5x, 6x, 7x, 8x, 9x, 10x]$$

Si queremos, por ejemplo, sumar todos los elementos de la lista anterior utilizamos la función **apply** con la operación que queremos realizar. Aquí aprovecharemos para utilizar un atajo muy práctico que consiste en el uso del símbolo **%**, que toma la última salida del programa para ser usado en la instrucción siguiente, de esta manera nos evitamos volver a escribir toda la instrucción anterior. La sintaxis para todo esto es:

(%i14) $\text{apply}("+, \%);$

$$(\%o14) 55x$$

Otra cosa que podemos hacer es asignarle a una lista el nombre de una variable.

(%i15) $L: [\sin(\pi/3), \log(3), \sqrt{2}, \text{abs}(x), \exp(x^2)];$

$$(\%o15) \left[\frac{\sqrt{3}}{2}, \log(3), \sqrt{2}, |x|, e^{x^2} \right]$$

De manera que para aislar elementos de una lista escribimos:

(%i16) $L[1]; L[4];$

$$(\%o16) \frac{\sqrt{3}}{2}$$

(%o17) $|x|$

Esto nos permite operar con sus elementos.

(%i18) $L[2] * (L[1] + L[4]) / L[5];$

$$(\%o18) \log(3) e^{-x^2} \left(|x| + \frac{\sqrt{3}}{2} \right)$$

La primera utilidad que le podemos dar a las listas es que nos permite definir vectores. Si queremos que el programa interprete los vectores

$$\mathbf{a} = (a^1, a^2, a^3), \mathbf{b} = (b^1, b^2, b^3), \mathbf{c} = (c^1, c^2, c^3), \mathbf{v0} = (0, 0, 0),$$

los podemos escribir como listas:

(%i19) $a:[a_1, a_2, a_3]; b:[b_1, b_2, b_3]; c:[c_1, c_2, c_3]; v0:[0, 0, 0];$

$$(\%o19) [a_1, a_2, a_3]$$

$$(\%o20) [b_1, b_2, b_3]$$

$$(\%o21) [c_1, c_2, c_3]$$

$$(\%o22) [0, 0, 0]$$

Ahora podemos hacer operaciones con los vectores

(%i23) $\mathbf{a} + \mathbf{b};$

$$(\%o23) [b_1 + a_1, b_2 + a_2, b_3 + a_3]$$

(%i24) $(\mathbf{a} + \mathbf{b}) + \mathbf{c};$

$$(\%o24) [c_1 + b_1 + a_1, c_2 + b_2 + a_2, c_3 + b_3 + a_3]$$

(%i25) $\mathbf{a} + \mathbf{v0};$

$$(\%o25) [a_1, a_2, a_3]$$

(%i26) $\mathbf{a} - \mathbf{a};$

$$(\%o26) [0, 0, 0]$$

(%i27) $\alpha * (\mathbf{a} + \mathbf{b});$

$$(\%o27) [\alpha (b_1 + a_1), \alpha (b_2 + a_2), \alpha (b_3 + a_3)]$$

(%i28) $(\alpha + \beta) * \mathbf{a};$

$$(\%o28) [a_1 (\beta + \alpha), a_2 (\beta + \alpha), a_3 (\beta + \alpha)]$$

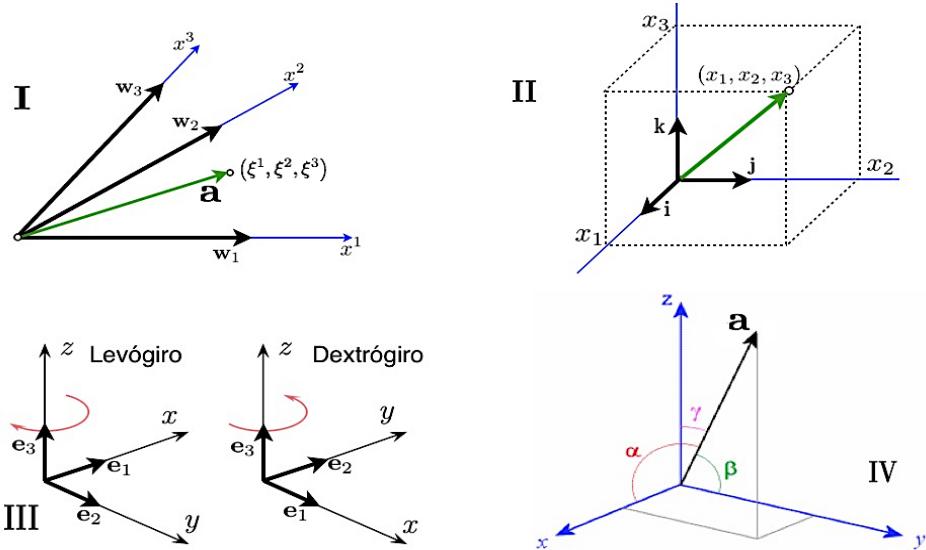


Figura 1.4: Bases vectoriales oblicuas I, bases cartesianas II, las bases orientadas del tipo levógiro y del tipo dextrógiro III y los cosenos directores IV.

1.2 Vectores en componentes

La formulación de las leyes físicas debe hacerse en término de cantidades vectoriales (tensoriales). Esto independiza su formulación de un sistema particular de coordenadas, pero llegado el momento de calcular valores y utilizar estas leyes, es mucho más conveniente referirla a un sistema de coordenadas particularmente adaptado a la geometría del problema. En ese caso, la ecuación vectorial se convertirá en tantas ecuaciones como componentes (referidas al sistema de coordenadas utilizado) tengan los vectores en ese sistema de coordenadas.

1.2.1 Bases y componentes

Tal y como mencionamos anteriormente, tres vectores **no coplanares** cualesquiera son linealmente independientes y constituyen una base para el espacio tridimensional. Denominaremos a estos vectores base como $\{\mathbf{w}_i\}$, y por ser linealmente independientes podremos expresar cualquier vector \mathbf{a} como una combinación lineal única, tal y como lo mostramos en el cuadrante I de la figura 1.4.

Con los vectores base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ podemos construir un sistema (oblicuo en general) de coordenadas al colocarlos con un mismo origen, esto es:

$$\mathbf{a} = a^1 \mathbf{w}_1 + a^2 \mathbf{w}_2 + a^3 \mathbf{w}_3 ,$$

donde las cantidades $\{a^1, a^2, a^3\}$ son números que representan las componentes del vector \mathbf{a} a lo largo de cada uno de los vectores base $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$. Nótese que por costumbre (será evidente más adelante, en la sección 3.1.1) etiquetamos estos números con superíndices y la letra que identifica al vector.

Más aún, cada punto P del espacio viene definido por un radio vector $\mathbf{r}(P) \equiv \overrightarrow{OP}$ que une el origen de coordenadas con el punto P y se le asocian tres números $\{x^1, x^2, x^3\}$, que son las proyecciones a lo largo de cada uno de los ejes coordenados $\{\overline{0x^1}, \overline{0x^2}, \overline{0x^3}\}$. Los números $\{x^1, x^2, x^3\}$ se denominan las componentes de $\mathbf{r}(P)$ en el sistema de referencia $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$.

Existe una familia de sistemas de coordenadas en la cual sus vectores base son ortogonales (o mejor ortonormales), es decir los vectores base $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ son perpendiculares entre si. Tal y como mostraremos

más adelante, siempre se puede construir un sistema ortogonal $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ u ortonormal $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ a partir de una base genérica de vectores linealmente independientes $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$. Cuando el sistema es ortogonal sus componentes se denominarán rectangulares. Dependiendo del signo del triple producto mixto, el sistema de coordenadas será dextrógiro ($(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 > 0$) o levógiro ($(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 < 0$), tal y como se muestra en el cuadrante III de la figura 1.4.

Es costumbre ancestral¹ utilizar la convención dextrógiro donde el producto: $(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 > 0$, y en ese caso utilizamos el bien conocido conjunto de vectores unitarios $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ con los que ya hemos estado familiarizados:

$$\mathbf{a} = a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} + a_z \hat{\mathbf{k}} \quad \text{y} \quad \mathbf{r}(P) = x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}}.$$

También es costumbre representar este sistema de coordenadas cartesiano ortonormal como:

$$\hat{\mathbf{i}} \equiv \hat{\mathbf{i}}_1, \quad \hat{\mathbf{j}} \equiv \hat{\mathbf{i}}_2 \quad \text{y} \quad \hat{\mathbf{k}} \equiv \hat{\mathbf{i}}_3, \quad (1.6)$$

para recordar que estamos en un sistema de coordenadas cartesianas y utilizaremos los superíndices 1, 2, 3 para indicar las componentes del vector.

$$\mathbf{a} = a^1 \hat{\mathbf{i}}_1 + a^2 \hat{\mathbf{i}}_2 + a^3 \hat{\mathbf{i}}_3 \quad \text{y} \quad \mathbf{r}(P) = x^1 \hat{\mathbf{i}}_1 + x^2 \hat{\mathbf{i}}_2 + x^3 \hat{\mathbf{i}}_3.$$

Obviamente el módulo del vector se podrá expresar con la utilización del teorema de Pitágoras:

$$|\mathbf{a}| = \sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2} \quad \text{y} \quad |\mathbf{r}(P)| = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}, \quad (1.7)$$

y la multiplicación por un número será:

$$\alpha \mathbf{a} = \alpha (a^1 \hat{\mathbf{i}}_1 + a^2 \hat{\mathbf{i}}_2 + a^3 \hat{\mathbf{i}}_3) = (\alpha a^1) \hat{\mathbf{i}}_1 + (\alpha a^2) \hat{\mathbf{i}}_2 + (\alpha a^3) \hat{\mathbf{i}}_3 \Rightarrow |\alpha \mathbf{a}| = \alpha \sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2}.$$

Igualmente para un vector unitario:

$$\hat{\mathbf{u}}_a = \frac{\mathbf{a}}{|\mathbf{a}|} = \frac{a^1 \hat{\mathbf{i}}_1 + a^2 \hat{\mathbf{i}}_2 + a^3 \hat{\mathbf{i}}_3}{\sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2}},$$

con lo cual todo vector:

$$\mathbf{a} = |\mathbf{a}| \hat{\mathbf{u}}_a = \sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2} \hat{\mathbf{u}}_a.$$

1.2.2 Cosenos directores

Como se puede apreciar en el cuadrante IV de la figura 1.4, podemos construir tres triángulos rectángulos con el radio vector $\mathbf{a}(P)$ como hipotenusa de cada uno de ellos. Los ángulos que forma el radio vector $\mathbf{a}(P)$ con cada uno de los ejes coordenados $\{x, y, z\}$ son $\{\theta_1, \theta_2, \theta_3\}$, respectivamente, con lo cual:

$$a_x = |\mathbf{a}| \cos \theta_1, \quad a_y = |\mathbf{a}| \cos \theta_2 \quad \text{y} \quad a_z = |\mathbf{a}| \cos \theta_3 \Rightarrow \cos^2 \theta_1 + \cos^2 \theta_2 + \cos^2 \theta_3 = 1, \quad (1.8)$$

pero además:

$$\hat{\mathbf{u}}_a = \frac{\mathbf{a}}{|\mathbf{a}|} = \cos \theta_1 \hat{\mathbf{i}} + \cos \theta_2 \hat{\mathbf{j}} + \cos \theta_3 \hat{\mathbf{k}}.$$

1.2.3 Una división fallida

Uno esperaría que para cada una de las definiciones de productos vectoriales, existiera el vector cociente, es decir, que pudiéramos “despejar” uno de los vectores multiplicados en términos del otro. La situación es que esta operación no está definida únicamente y lo podemos intuir a partir de una de la definición del producto escalar. Supongamos que tenemos que: $\zeta = \mathbf{a} \cdot \mathbf{b}$, con lo cual, si pudiéramos “despejar”, digamos, $\mathbf{b} = \zeta / \mathbf{a}$

¹Quizá por las arraigadas relaciones de dominación de los derechos sobre los izquierdos (en latín e italiano los zurdos son siniestros), o quizá tal vez por conservar la definición de volumen como positivo.

¿Tendríamos entonces definido \mathbf{b} de una manera única? La respuesta es NO, ya que $\zeta = \mathbf{a} \cdot \left(\frac{\zeta}{\mathbf{a}} + \mathbf{d} \right)$, donde $\mathbf{a} \perp \mathbf{d}$, por lo cual existen infinitos $\mathbf{b} = \frac{\zeta}{\mathbf{a}} + \mathbf{d}$ que cumplen $\zeta = \mathbf{a} \cdot \mathbf{b}$.

1.2.4 Algebra vectorial en componentes

Es posible reescribir toda el álgebra vectorial que hemos visto mediante operaciones referidas a sistemas de coordenadas, como mostraremos a continuación. Por simplicidad, anclaremos nuestro sistema de coordenadas a una base ortonormal: $\{\hat{\mathbf{e}}_i\}$.

Para los vectores $\mathbf{a} = a^1 \hat{\mathbf{e}}_1 + a^2 \hat{\mathbf{e}}_2 + a^3 \hat{\mathbf{e}}_3$ y $\mathbf{b} = b^1 \hat{\mathbf{e}}_1 + b^2 \hat{\mathbf{e}}_2 + b^3 \hat{\mathbf{e}}_3$, la suma será representada por:

$$\mathbf{a} + \mathbf{b} = (a^1 \hat{\mathbf{e}}_1 + a^2 \hat{\mathbf{e}}_2 + a^3 \hat{\mathbf{e}}_3) + (b^1 \hat{\mathbf{e}}_1 + b^2 \hat{\mathbf{e}}_2 + b^3 \hat{\mathbf{e}}_3) = (a^1 + b^1) \hat{\mathbf{e}}_1 + (a^2 + b^2) \hat{\mathbf{e}}_2 + (a^3 + b^3) \hat{\mathbf{e}}_3,$$

y obviamente, la resta:

$$\mathbf{a} - \mathbf{b} = (a^1 \hat{\mathbf{e}}_1 + a^2 \hat{\mathbf{e}}_2 + a^3 \hat{\mathbf{e}}_3) - (b^1 \hat{\mathbf{e}}_1 + b^2 \hat{\mathbf{e}}_2 + b^3 \hat{\mathbf{e}}_3) = (a^1 - b^1) \hat{\mathbf{e}}_1 + (a^2 - b^2) \hat{\mathbf{e}}_2 + (a^3 - b^3) \hat{\mathbf{e}}_3,$$

con lo cual la distancia entre dos puntos P y M será:

$$d(P, M) = |(\mathbf{r}(P) = \mathbf{a}) - (\mathbf{r}(M) = \mathbf{b})| = \sqrt{(x^1 - y^1)^2 + (x^2 - y^2)^2 + (x^3 - y^3)^2}.$$

1.2.5 Dependencia e independencia lineal

Ahora es fácil estudiar la dependencia o independencia lineal en coordenadas. Otra vez, tres vectores: $\mathbf{a} = a^1 \hat{\mathbf{e}}_1 + a^2 \hat{\mathbf{e}}_2 + a^3 \hat{\mathbf{e}}_3$, $\mathbf{b} = b^1 \hat{\mathbf{e}}_1 + b^2 \hat{\mathbf{e}}_2 + b^3 \hat{\mathbf{e}}_3$ y $\mathbf{c} = c^1 \hat{\mathbf{e}}_1 + c^2 \hat{\mathbf{e}}_2 + c^3 \hat{\mathbf{e}}_3$, serán *linealmente independientes* si se cumple que:

$$\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c} = \mathbf{0} \Rightarrow \alpha = \beta = \gamma = 0.$$

Por inspección podemos ver que base canónica: $\hat{\mathbf{e}}_1 = \hat{\mathbf{i}} \equiv (1, 0, 0)$, $\hat{\mathbf{e}}_2 = \hat{\mathbf{j}} \equiv (0, 1, 0)$, $\hat{\mathbf{e}}_3 = \hat{\mathbf{k}} \equiv (0, 0, 1)$ conforman un conjunto de vectores claramente linealmente independientes y por lo tanto constituyen una base.

En general tendremos que:

$$\mathbf{0} = \alpha (a^1 \hat{\mathbf{e}}_1 + a^2 \hat{\mathbf{e}}_2 + a^3 \hat{\mathbf{e}}_3) + \beta (b^1 \hat{\mathbf{e}}_1 + b^2 \hat{\mathbf{e}}_2 + b^3 \hat{\mathbf{e}}_3) + \gamma (c^1 \hat{\mathbf{e}}_1 + c^2 \hat{\mathbf{e}}_2 + c^3 \hat{\mathbf{e}}_3)$$

$$\mathbf{0} = (\alpha a^1 + \beta b^1 + \gamma c^1) \hat{\mathbf{e}}_1 + (\alpha a^2 + \beta b^2 + \gamma c^2) \hat{\mathbf{e}}_2 + (\alpha a^3 + \beta b^3 + \gamma c^3) \hat{\mathbf{e}}_3 \Rightarrow \begin{cases} \alpha a^1 + \beta b^1 + \gamma c^1 = 0 \\ \alpha a^2 + \beta b^2 + \gamma c^2 = 0 \\ \alpha a^3 + \beta b^3 + \gamma c^3 = 0 \end{cases}$$

Esto no es otra cosa que un sistema de 3 ecuaciones lineales con 3 incógnitas: $\{\alpha, \beta, \gamma\}$, y la solución que estamos buscando $\alpha = \beta = \gamma = 0$ se cumplirá si:

$$\begin{vmatrix} a^1 & b^1 & c^1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{vmatrix} = a^1 (b^2 c^3 - b^3 c^2) + a^2 (b^3 c^1 - b^1 c^3) + a^3 (b^1 c^2 - b^2 c^1) \neq 0.$$

1.2.6 Productos de vectores en componentes

Exploremos más sobre el álgebra vectorial y al producto de vectores si los tenemos anclados a algún sistema de coordenadas.

1.2.6.1 Producto escalar

Todo lo expresado en la sección 1.1.4.1 lo representaremos ahora en término de una base ortonormal: $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$. Dados los vectores $\mathbf{a} = a^1\hat{\mathbf{e}}_1 + a^2\hat{\mathbf{e}}_2 + a^3\hat{\mathbf{e}}_3$ y $\mathbf{b} = b^1\hat{\mathbf{e}}_1 + b^2\hat{\mathbf{e}}_2 + b^3\hat{\mathbf{e}}_3$, el producto escalar entre ellos es:

$$\mathbf{a} \cdot \mathbf{b} = (a^1\hat{\mathbf{e}}_1 + a^2\hat{\mathbf{e}}_2 + a^3\hat{\mathbf{e}}_3) \cdot (b^1\hat{\mathbf{e}}_1 + b^2\hat{\mathbf{e}}_2 + b^3\hat{\mathbf{e}}_3) = a^1b^1 + a^2b^2 + a^3b^3,$$

por ser ortonormales se tiene que:

$$\hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_1 = \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_2 = \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_3 = 1, \quad \text{y} \quad \begin{cases} \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_2 = \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_1 = 0 \\ \hat{\mathbf{e}}_1 \cdot \hat{\mathbf{e}}_3 = \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_1 = 0 \\ \hat{\mathbf{e}}_2 \cdot \hat{\mathbf{e}}_3 = \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_2 = 0 \end{cases}$$

Las propiedades del producto escalar en coordenadas cartesianas se comprueban fácilmente.

- *El producto escalar de un vector consigo mismo, siempre es positivo.*

$$\zeta = \mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 = (a^1)^2 + (a^2)^2 + (a^3)^2 \geq 0,$$

y

$$(a^1)^2 + (a^2)^2 + (a^3)^2 = 0 \Rightarrow a^1 = a^2 = a^3 = 0 \Leftrightarrow \mathbf{a} = \mathbf{0}.$$

Adicionalmente: $|\mathbf{a}| = \sqrt{\zeta} = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2}$.

- *El producto escalar es commutativo.*

$$\zeta = \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} = a^1b^1 + a^2b^2 + a^3b^3 = b^1a^1 + b^2a^2 + b^3a^3.$$

- *El producto escalar es distributivo.*

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = [a^1\hat{\mathbf{e}}_1 + a^2\hat{\mathbf{e}}_2 + a^3\hat{\mathbf{e}}_3] \cdot [(b^1 + c^1)\hat{\mathbf{e}}_1 + (b^2 + c^2)\hat{\mathbf{e}}_2 + (b^3 + c^3)\hat{\mathbf{e}}_3],$$

por lo tanto:

$$\begin{aligned} a^1(b^1 + c^1) + a^2(b^2 + c^2) + a^3(b^3 + c^3) &= (a^1b^1 + a^1c^1) + (a^2b^2 + a^2c^2) + (a^3b^3 + a^3c^3) \\ &= (a^1b^1 + a^2b^2 + a^3b^3) + (a^1c^1 + a^2c^2 + a^3c^3) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}. \end{aligned}$$

- *La multiplicación por un escalar.*

$$|\alpha|(\mathbf{a} \cdot \mathbf{b}) = (\alpha\mathbf{a}) \cdot \mathbf{b} = \mathbf{a} \cdot (\alpha\mathbf{b}) = (\alpha a^1)b^1 + (\alpha a^2)b^2 + (\alpha a^3)b^3 = a^1(\alpha b^1) + a^2(\alpha b^2) + a^3(\alpha b^3).$$

- *Desigualdad de Cauchy-Schwarz.*

$$\mathbf{a} \cdot \mathbf{b} = a^1b^1 + a^2b^2 + a^3b^3 \leq \sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2} \sqrt{(b^1)^2 + (b^2)^2 + (b^3)^2} = |\mathbf{a}| |\mathbf{b}|.$$

- *Diremos que dos vectores, no nulos son ortogonales (perpendiculares) si su producto escalar es nulo.*

Esta afirmación es inmediata:

$$\mathbf{a} \perp \mathbf{b} \Rightarrow \theta_{\langle \mathbf{a}, \mathbf{b} \rangle} = \frac{\pi}{2} \Rightarrow \mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} = 0,$$

por lo cual:

$$a^1b^1 + a^2b^2 + a^3b^3 = |\mathbf{a}| |\mathbf{b}| \cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} \Rightarrow \cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle} = \frac{a^1b^1 + a^2b^2 + a^3b^3}{\sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2} \sqrt{(b^1)^2 + (b^2)^2 + (b^3)^2}},$$

de donde se deduce que para dos vectores perpendiculares:

$$\mathbf{a} \perp \mathbf{b} \Rightarrow 0 = a^1b^1 + a^2b^2 + a^3b^3.$$

- *Del producto escalar surge el teorema del coseno.*

Es inmediato generalizar el producto escalar de un vector consigo mismo, para ello suponemos que

$\mathbf{c} = \mathbf{a} + \mathbf{b}$, con lo cual:

$$\mathbf{c} = \mathbf{a} + \mathbf{b} \Rightarrow \mathbf{c} \cdot \mathbf{c} = (\mathbf{a} + \mathbf{b}) \cdot (\mathbf{a} + \mathbf{b}) \Rightarrow |\mathbf{c}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 + 2|\mathbf{a}||\mathbf{b}| \cos(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle},$$

que no es otra cosa que el teorema del coseno y está ilustrado en el cuadrante III de la figura 1.3.

1.2.6.2 Producto vectorial

De igual manera, lo que aprendimos en la sección 1.1.4.2 ahora lo expresamos en términos de las componentes de los vectores en una base ortonormal de la forma:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = (a^2 b^3 - a^3 b^2) \hat{\mathbf{e}}_1 + (a^3 b^1 - a^1 b^3) \hat{\mathbf{e}}_2 + (a^1 b^2 - a^2 b^1) \hat{\mathbf{e}}_3,$$

lo anterior se puede organizar como el determinante de la matriz:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{e}}_1 & \hat{\mathbf{e}}_2 & \hat{\mathbf{e}}_3 \\ a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \end{vmatrix},$$

con lo cual:

$$\begin{aligned} |\mathbf{c}| &= \sqrt{(a^2 b^3 - a^3 b^2)^2 + (a^3 b^1 - a^1 b^3)^2 + (a^1 b^2 - a^2 b^1)^2} \\ &= \sqrt{(a^1)^2 + (a^2)^2 + (a^3)^2} \sqrt{(b^1)^2 + (b^2)^2 + (b^3)^2} \operatorname{sen}(\theta)_{\langle \mathbf{a}, \mathbf{b} \rangle}. \end{aligned}$$

1.2.6.3 Triple producto mixto

Finalmente, analicemos el número (pseudoescalar) que proviene de la multiplicación:

$$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = |\mathbf{c}| |\mathbf{a} \times \mathbf{b}| \cos(\theta)_{\langle \mathbf{c}, \mathbf{a} \times \mathbf{b} \rangle} = \begin{vmatrix} c^1 & c^2 & c^3 \\ a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \end{vmatrix}.$$

Obviamente, el valor absoluto de este número, $V = |\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})|$, representa el volumen del paralelepípedo cuyos lados quedan definidos por los vectores: \mathbf{a} , \mathbf{b} y \mathbf{c} .

Ejemplo 1.8 Si tenemos los vectores $\mathbf{a} = \hat{\mathbf{i}} + 3\hat{\mathbf{j}} + 5\hat{\mathbf{k}}$ y $\mathbf{b} = 2\hat{\mathbf{i}} + 4\hat{\mathbf{j}} + 6\hat{\mathbf{k}}$, encontremos el ángulo que forman.

$$\cos \theta = \frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|} = \frac{[\hat{\mathbf{i}} + 3\hat{\mathbf{j}} + 5\hat{\mathbf{k}}] \cdot [2\hat{\mathbf{i}} + 4\hat{\mathbf{j}} + 6\hat{\mathbf{k}}]}{\sqrt{1^2 + 3^2 + 5^2} \sqrt{2^2 + 4^2 + 6^2}} = \frac{2 + 12 + 30}{\sqrt{35} \sqrt{56}} = \frac{44}{\sqrt{35} \sqrt{56}} = \frac{22}{\sqrt{14} \sqrt{35}}.$$

Por lo tanto: $\theta = \arccos\left(\frac{22}{\sqrt{14} \sqrt{35}}\right) = 0,11088$.

Notemos que de la ecuación (1.8) para los vectores \mathbf{a} y \mathbf{b} y su componentes resulta que:

$$\begin{aligned} \cos \alpha &= \frac{a_x}{|\mathbf{a}|} = \frac{1}{\sqrt{35}}, \quad \cos \beta = \frac{a_y}{|\mathbf{a}|} = \frac{3}{\sqrt{35}}, \quad \cos \gamma = \frac{a_z}{|\mathbf{a}|} = \frac{5}{\sqrt{35}} \Rightarrow \left[\frac{1}{\sqrt{35}}\right]^2 + \left[\frac{3}{\sqrt{35}}\right]^2 + \left[\frac{5}{\sqrt{35}}\right]^2 = 1, \\ \cos \alpha &= \frac{b_x}{|\mathbf{b}|} = \frac{2}{\sqrt{56}}, \quad \cos \beta = \frac{b_y}{|\mathbf{b}|} = \frac{4}{\sqrt{56}}, \quad \cos \gamma = \frac{b_z}{|\mathbf{b}|} = \frac{6}{\sqrt{56}} \Rightarrow \left[\frac{2}{\sqrt{56}}\right]^2 + \left[\frac{4}{\sqrt{56}}\right]^2 + \left[\frac{6}{\sqrt{56}}\right]^2 = 1. \end{aligned}$$

Y además podemos ver claramente que el ángulo entre los vectores y los cosenos directores están relacionados, como se muestra a continuación:

$$\cos \theta = \frac{1}{\sqrt{35}} \frac{2}{\sqrt{56}} + \frac{3}{\sqrt{35}} \frac{4}{\sqrt{56}} + \frac{5}{\sqrt{35}} \frac{6}{\sqrt{56}} = \frac{22}{\sqrt{14} \sqrt{35}}.$$



Ejemplo 1.9 Consideremos los vectores:

$$\mathbf{w}_1 = \hat{\mathbf{i}} = (1, 0, 0), \quad \mathbf{w}_2 = \hat{\mathbf{i}} + \hat{\mathbf{j}} = (1, 1, 0), \quad \mathbf{w}_3 = \hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}} = (1, 1, 1).$$

¿Son linealmente independientes?

Tenemos entonces que:

$$\alpha(1, 0, 0) + \beta(1, 1, 0) + \gamma(1, 1, 1) = (0, 0, 0).$$

Al escribir el sistema de ecuaciones resulta:

$$\alpha + \beta + \gamma = 0, \beta + \gamma = 0, \gamma = 0, \Rightarrow \alpha = 0, \beta = 0, \gamma = 0,$$

con lo cual demostramos que son linealmente independientes y por lo tanto constituyen una base para los vectores tridimensionales. Notemos además que esta base: $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$, es una base que no es ortogonal.



Ejemplo 1.10 Dados los vectores:

$$\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}, \mathbf{b} = 4\hat{\mathbf{i}} + 5\hat{\mathbf{j}}, \mathbf{c} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}.$$

Revisemos si forman una base.

Podemos ver que $\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} = \mathbf{0}$ implica que:

$$\alpha [\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}] + \beta [4\hat{\mathbf{i}} + 5\hat{\mathbf{j}}] + \gamma [3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}] = \mathbf{0} \Rightarrow \begin{cases} \alpha + 4\beta + 3\gamma = 0 \\ 2\alpha + 5\beta + 2\gamma = 0 \\ 3\alpha + \gamma = 0 \end{cases}$$

Cuya solución es: $\alpha = \beta = \gamma = 0$. Por lo tanto el conjunto $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ es linealmente independiente, forman una base no ortogonal.

Por otro lado, si queremos calcular el volumen recurrimos al triple producto vectorial:

$$\begin{aligned} V &= |\mathbf{a} \cdot [\mathbf{b} \times \mathbf{c}]| = \left| [\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}] \cdot [4\hat{\mathbf{i}} + 5\hat{\mathbf{j}}] \times [3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}] \right| \\ &= \left| [\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}] \cdot [5\hat{\mathbf{i}} - 4\hat{\mathbf{j}} - 7\hat{\mathbf{k}}] \right| = |-24| = 24. \end{aligned}$$

Analicemos el siguiente cálculo adicional: si tenemos un vector arbitrario, digamos, $\mathbf{d} = d^1\hat{\mathbf{i}} + d^2\hat{\mathbf{j}} + d^3\hat{\mathbf{k}}$, podemos construir las siguientes cantidades:

$$c^1 = -\frac{\mathbf{a} \cdot \mathbf{d}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, \quad c^2 = -\frac{\mathbf{b} \cdot \mathbf{d}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, \quad c^3 = -\frac{\mathbf{c} \cdot \mathbf{d}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}.$$

resultando:

$$c^1 = \frac{d^1 + 2d^2 + 3d^3}{24}, \quad c^2 = \frac{4d^1 + 5d^2}{24}, \quad c^3 = \frac{3d^1 + 2d^2 + d^3}{24},$$

podemos ver que si calculamos la siguiente expresión: $c^1(\mathbf{b} \times \mathbf{c}) + c^2(\mathbf{c} \times \mathbf{a}) + c^3(\mathbf{a} \times \mathbf{b}) + \mathbf{d}$, resulta que es igual a:

$$\begin{aligned} \frac{d^1 + 2d^2 + 3d^3}{24}(5\hat{\mathbf{i}} - 4\hat{\mathbf{j}} - 7\hat{\mathbf{k}}) &+ \frac{4d^1 + 5d^2}{24}(4\hat{\mathbf{i}} - 8\hat{\mathbf{j}} + 4\hat{\mathbf{k}}) + \frac{3d^1 + 2d^2 + d^3}{24}(-15\hat{\mathbf{i}} + 12\hat{\mathbf{j}} - 3\hat{\mathbf{k}}) \\ &+ d^1\hat{\mathbf{i}} + d^2\hat{\mathbf{j}} + d^3\hat{\mathbf{k}} = (0, 0, 0). \end{aligned}$$

Es decir,

$$c^1(\mathbf{b} \times \mathbf{c}) + c^2(\mathbf{c} \times \mathbf{a}) + c^3(\mathbf{a} \times \mathbf{b}) + \mathbf{d} = \mathbf{0},$$

siempre y cuando el conjunto $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ sea linealmente independiente y \mathbf{d} un vector arbitrario.



1.2.7 Ejercicios

1. Con la definición del producto escalar entre vectores, demuestre que si θ es el ángulo entre los vectores \mathbf{a} y \mathbf{b} , entonces:

$$\cos \theta = \frac{a^1 b^1}{ab} + \frac{a^2 b^2}{ab} + \frac{a^3 b^3}{ab}.$$

Donde las cantidades a^i/a y b^i/b son los cosenos directores de \mathbf{a} y \mathbf{b} respectivamente y $a = |\mathbf{a}|$ y $b = |\mathbf{b}|$.

2. Encuentre la distancia del punto P al origen, si P viene dado por el vector posición: $\mathbf{r} = 2\hat{\mathbf{i}} + 4\hat{\mathbf{j}} - 3\hat{\mathbf{k}}$. Y si para un punto arbitrario el vector posición es: $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ ¿Qué superficie describe éste vector cuando $|\mathbf{r}| = 3$?

3. Encuentre los cosenos directores y los correspondientes ángulos para los siguientes vectores:

(a). $\mathbf{a} = \hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}}$.

(b). $\mathbf{b} = \hat{\mathbf{i}} - 2\hat{\mathbf{j}} + 2\hat{\mathbf{k}}$.

(c). $\mathbf{c} = 4\hat{\mathbf{i}} - 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$.

4. Verifique que los vectores:

$$\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}, \quad \mathbf{b} = \hat{\mathbf{i}} + 5\hat{\mathbf{j}}, \quad \mathbf{c} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}.$$

forman una base ¿Esta base será del tipo dextrógiro o levógiro?

5. ¿Son los siguientes conjuntos de vectores una base?

(a). $\mathbf{a}_1 = 2\hat{\mathbf{i}} + \hat{\mathbf{j}} - 3\hat{\mathbf{k}}, \quad \mathbf{a}_2 = \hat{\mathbf{i}} - 4\hat{\mathbf{k}}, \quad \mathbf{a}_3 = 4\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - \hat{\mathbf{k}}$.

(b). $\mathbf{b}_1 = \hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 2\hat{\mathbf{k}}, \quad \mathbf{b}_2 = 2\hat{\mathbf{i}} - 4\hat{\mathbf{j}} - \hat{\mathbf{k}}, \quad \mathbf{b}_3 = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} - \hat{\mathbf{k}}$.

6. Dados los vectores:

$$\mathbf{a} = \hat{\mathbf{e}}_1 + 2\hat{\mathbf{e}}_2 + 3\hat{\mathbf{e}}_3, \quad \mathbf{b} = 4\hat{\mathbf{e}}_1 + 5\hat{\mathbf{e}}_2 + 6\hat{\mathbf{e}}_3, \quad \mathbf{c} = 3\hat{\mathbf{e}}_1 + 2\hat{\mathbf{e}}_2 + \hat{\mathbf{e}}_3, \quad \mathbf{d} = 6\hat{\mathbf{e}}_1 + 5\hat{\mathbf{e}}_2 + 4\hat{\mathbf{e}}_3.$$

- (a). Encuentre:

$$\mathbf{a} + \mathbf{b} + \mathbf{c} + \mathbf{d}, \quad \mathbf{a} + \mathbf{b} - \mathbf{c} - \mathbf{d}, \quad \mathbf{a} - \mathbf{b} + \mathbf{c} - \mathbf{d}, \quad -\mathbf{a} + \mathbf{b} - \mathbf{c} + \mathbf{d}.$$

- (b). El ángulo entre los vectores $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ y los vectores base $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3$.

- (c). La magnitud de los vectores $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$.

- (d). El ángulo entre \mathbf{a} y \mathbf{b} y entre \mathbf{c} y \mathbf{d} .

- (e). La proyección de \mathbf{a} sobre \mathbf{b} .

- (f). ¿Son los vectores $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ coplanares?

- (g). Encuentre $(\mathbf{a} + \mathbf{b}) \cdot (\mathbf{c} + \mathbf{d})$.

- (h). Los productos $\mathbf{a} \times \mathbf{b}, \mathbf{b} \times \mathbf{c}, \mathbf{c} \times \mathbf{d}$ y los ángulos que estos forman con \mathbf{d} .

- (i). $\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$.

7. Verifique la desigualdad triangular: $|\mathbf{a} + \mathbf{b}| \leq |\mathbf{a}| + |\mathbf{b}|$, para los siguientes vectores:

(a). $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$ y $\mathbf{b} = 2\hat{\mathbf{i}} + \hat{\mathbf{j}} + 7\hat{\mathbf{k}}$.

(b). $\mathbf{a} = 2\hat{\mathbf{i}} - \hat{\mathbf{j}} - 2\hat{\mathbf{k}}$ y $\mathbf{b} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$.

8. Si \mathbf{a} y \mathbf{b} son vectores arbitrarios y α y β números, demuestre que:

$$|\alpha \mathbf{a} + \beta \mathbf{b}|^2 \leq \alpha^2 |\mathbf{a}|^2 + 2\alpha\beta(\mathbf{a} \cdot \mathbf{b}) + \beta^2 |\mathbf{b}|^2.$$

9. Si $\mathbf{a}, \mathbf{b}, \mathbf{c}$ y \mathbf{d} son vectores arbitrarios y α, β, γ escalares que satisfacen:

$$\alpha(\mathbf{b} \times \mathbf{c}) + \beta(\mathbf{c} \times \mathbf{a}) + \gamma(\mathbf{a} \times \mathbf{b}) + \mathbf{d} = \mathbf{0},$$

demuestre que si \mathbf{a}, \mathbf{b} y \mathbf{c} son linealmente independientes, entonces:

$$\alpha = -\frac{\mathbf{a} \cdot \mathbf{d}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, \quad \beta = -\frac{\mathbf{b} \cdot \mathbf{d}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, \quad \gamma = -\frac{\mathbf{c} \cdot \mathbf{d}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}.$$

10. Si $\mathbf{a}, \mathbf{b}, \mathbf{c}$ y \mathbf{d} son vectores arbitrarios y α, β, γ escalares que satisfacen:

$$\alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c} + \mathbf{d} = \mathbf{0},$$

demuestre que si \mathbf{a}, \mathbf{b} y \mathbf{c} son linealmente independientes, entonces:

$$\alpha = -\frac{\mathbf{d} \cdot (\mathbf{b} \times \mathbf{c})}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, \quad \beta = -\frac{\mathbf{d} \cdot (\mathbf{c} \times \mathbf{a})}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, \quad \gamma = -\frac{\mathbf{d} \cdot (\mathbf{a} \times \mathbf{b})}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}.$$

Ayuda: tome el producto escalar de la ecuación con $\mathbf{b} \times \mathbf{c}$, $\mathbf{a} \times \mathbf{c}$ y $\mathbf{a} \times \mathbf{b}$.

11. Demuestre que los vectores $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}$, $\mathbf{b} = 2\hat{\mathbf{i}} - \hat{\mathbf{j}} - \hat{\mathbf{k}}$ y $\mathbf{c} = 4\hat{\mathbf{i}} + 3\hat{\mathbf{j}} + \hat{\mathbf{k}}$ son linealmente independientes.

Escoja un vector \mathbf{d} y verifique los resultados de los dos últimos ejercicios.

12. Utilizando **Maxima** realice todos los ejercicios anteriores y compare los resultados.

Practicando con Maxima

Con el programa de manipulación simbólica **Maxima** haremos algunos cálculos sencillos con vectores. Nuevamente recomendamos ver el apéndice ?? como introducción al programa.

Dados los vectores, en coordenadas cartesianas: $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$ y $\mathbf{b} = 7\hat{\mathbf{i}} + 8\hat{\mathbf{j}} + 9\hat{\mathbf{k}}$.

(%i1) $\mathbf{a}: [1, 2, 3];$
 (%o1) $[1, 2, 3]$

(%i2) $\mathbf{b}: [7, 8, 9];$
 (%o2) $[7, 8, 9]$

La multiplicación por escalares y suma es simple, si queremos calcular $\alpha\mathbf{a} + \beta\mathbf{b}$, escribimos:

(%i3) $\alpha\mathbf{a} + \beta\mathbf{b};$
 (%o3) $[7\beta + \alpha, 8\beta + 2\alpha, 9\beta + 3\alpha]$

Para el producto escalar procedemos utilizando el operador punto, como se muestra a continuación.

(%i4) $\mathbf{a} \cdot \mathbf{b};$
 (%o4) 50

El cálculo de producto vectorial no es tan obvio, debemos cargar previamente la biblioteca **vect**.

(%i5) $\text{load}(\text{vect})$$

El operador para el producto vectorial es una tilde \sim y además debemos utilizar la función **express**. Entonces, para calcular $\mathbf{a} \times \mathbf{b}$, ejecutamos los siguientes comandos:

(%i6) $\text{express}(\mathbf{a} \sim \mathbf{b});$
 (%o6) $[-6, 12, -6]$

La norma de un vector, como ya vimos, es: $\sqrt{\mathbf{a} \cdot \mathbf{a}}$

(%i7) $\sqrt{\mathbf{a} \cdot \mathbf{a}};$
 (%o7) $\sqrt{14}$

Si tenemos otro vector, digamos $\mathbf{c} = -4\hat{\mathbf{i}} + 5\hat{\mathbf{j}} - 6\hat{\mathbf{k}}$, el producto triple: $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c}$ se calcula así:

(%i8) $\mathbf{c}: [-4, 5, -6];$
 (%o8) $[-4, 5, -6]$

(%i9) $\mathbf{a} \cdot \mathbf{b} \times \mathbf{c};$
 (%o9) 120

El ángulo entre los vectores **a** y **b**, es:

$$\theta = \arccos \left(\frac{\mathbf{a} \cdot \mathbf{b}}{|\mathbf{a}| |\mathbf{b}|} \right).$$

En **Maxima** usamos la función **acos(x)** para el arcocoseno(*x*). Consultar el manual del programa para ver el resto de las funciones trigonométricas.

```
(%i10) acos((a.b)/(sqrt(a.a)*sqrt(b.b)));
(%o10) acos  $\left( \frac{50}{\sqrt{14} \sqrt{194}} \right)$ 
```

Seguramente lo queremos es el valor numérico, esto se hace agregando la función **float**. Con la siguiente sintaxis logramos el objetivo:

```
(%i11) acos((a.b)/(sqrt(a.a)*sqrt(b.b))),float;
(%o11) 0,2858867976945064
```

Para finalizar, podemos considerar el problema de la independencia lineal de vectores. Tomemos el conjunto de vectores del ejemplo ??, es decir, los vectores: $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$, $\mathbf{b} = 4\hat{\mathbf{i}} + 5\hat{\mathbf{j}}$ y $\mathbf{c} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}$. El sistema de ecuaciones a resolver era:

$$\begin{cases} \alpha + 4\beta + 3\gamma = 0 \\ 2\alpha + 5\beta + 2\gamma = 0 \\ 3\alpha + \gamma = 0 \end{cases}$$

Disponemos de un comando que permite resolver sistemas de ecuaciones lineales, este comando se llama **linsolve**. Podemos escribir el sistema de ecuaciones como una lista y luego resolver:

```
(%i12) ecus:[alpha+4*beta+3*gamma=0,2*alpha+5*beta+2*gamma=0,3*alpha+gamma=0];
(%o12) [3  $\gamma$  + 4  $\beta$  +  $\alpha$  = 0, 2  $\gamma$  + 5  $\beta$  + 2  $\alpha$  = 0,  $\gamma$  + 3  $\alpha$  = 0]
```

```
(%i13) linsolve(ecus,[alpha,beta,gamma]);
(%o13) [ $\alpha$  = 0,  $\beta$  = 0,  $\gamma$  = 0]
```

Repasemos el ejemplo 1.10. Primero escribamos los vectores

```
(%i14) a:[1,2,3];b:[4,5,0];c:[3,2,1];
(%o14) [1, 2, 3]
(%o15) [4, 5, 0]
(%o16) [3, 2, 1]
```

Resolvemos el sistema de ecuaciones

```
(%i17) ecus:[alpha+4*beta+3*gamma=0, 2*alpha+5*beta+2*gamma=0, 3*alpha+gamma=0];
(%o17) [3  $\gamma$  + 4  $\beta$  +  $\alpha$  = 0, 2  $\gamma$  + 5  $\beta$  + 2  $\alpha$  = 0,  $\gamma$  + 3  $\alpha$  = 0]
```

```
(%i18) linsolve(ecus,[alpha,beta,gamma]);
(%o18) [ $\alpha$  = 0,  $\beta$  = 0,  $\gamma$  = 0]
```

El volumen

```
(%i19) V: abs(a.cross(b));
(%o19) 24
```

Ahora vamos a introducir un vector arbitrario

```
(%i20) d:[d1,d2,d3];
```

(%o20) $[d1, d2, d3]$

y calculamos las cantidades c^1 , c^2 , y c^3 del ejemplo

(%o21) $c1:-(a.d)/(a.\text{express}(b \sim c)); c2:-(b.d)/(a.\text{express}(b \sim c)); c3:-(c.d)/(a.\text{express}(b \sim c));$

$$(\%o21) -\frac{-3d3 - 2d2 - d1}{24}$$

$$(\%o22) -\frac{-5d2 - 4d1}{24}$$

$$(\%o23) -\frac{-d3 - 2d2 - 3d1}{24}$$

y ahora la expresión: $c^1(\mathbf{b} \times \mathbf{c}) + c^2(\mathbf{c} \times \mathbf{a}) + c^3(\mathbf{a} \times \mathbf{b}) + \mathbf{d}$:

(%o24) $c1*\text{express}(b \sim c) + c2*\text{express}(c \sim a) + c3*\text{express}(a \sim b) + d;$

$$(\%o24) \left[\frac{5(-d3 - 2d2 - 3d1)}{8} - \frac{5(-3d3 - 2d2 - d1)}{24} - \frac{-5d2 - 4d1}{6} + d1, -\frac{-d3 - 2d2 - 3d1}{2} + \right. \\ \left. \frac{-3d3 - 2d2 - d1}{6} + d2 + \frac{-5d2 - 4d1}{3}, d3 + \frac{-d3 - 2d2 - 3d1}{8} + \frac{7(-3d3 - 2d2 - d1)}{24} - \frac{-5d2 - 4d1}{6} \right]$$

Simplificamos

(%o25) $\text{ratsimp}(%);$

(%o25) $[0, 0, 0]$

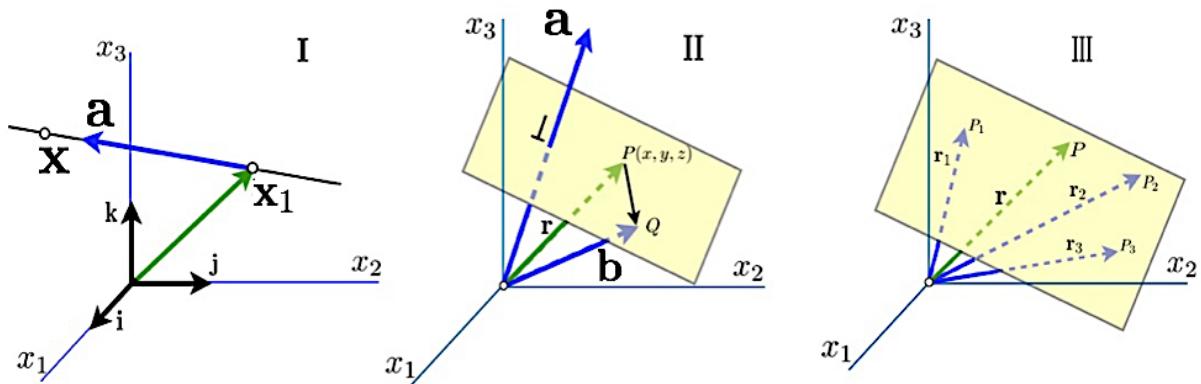


Figura 1.5: Geometría analítica: representación de rectas y planos con vectores cartesianos.

1.3 Aplicaciones del álgebra vectorial

Uno de los terrenos más exitosos de las aplicaciones del álgebra vectorial es la geometría analítica. Esto se realiza en base a la definición que hicimos del radio vector: a cada punto P del espacio le asociamos un radio vector posición, tal y como lo mostramos en el cuadrante I de la figura 1.4.

$$P \longleftrightarrow (x, y, z) \equiv (x^1, x^2, x^3) \Rightarrow \mathbf{r}(P) = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} = x^1\hat{\mathbf{i}}_1 + x^2\hat{\mathbf{i}}_2 + x^3\hat{\mathbf{i}}_3 = \sum_{i=1}^3 x^i\hat{\mathbf{i}}_i.$$

A partir de esta definición todas las propiedades geométricas del espacio las podemos construir con vectores.

1.3.1 Rectas y vectores

La ecuación de la recta en término de vectores la definiremos fijando uno de sus puntos, digamos:

$$\mathbf{r}(P_1) \equiv \mathbf{x}(P_1) = \mathbf{x}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} + z_1\hat{\mathbf{k}} = x_{(1)}^1\hat{\mathbf{i}}_1 + x_{(1)}^2\hat{\mathbf{i}}_2 + x_{(1)}^3\hat{\mathbf{i}}_3,$$

y un vector que indique su dirección, digamos $\mathbf{a} = a^1\hat{\mathbf{i}} + a^2\hat{\mathbf{j}} + a^3\hat{\mathbf{k}}$ (ver cuadrante I de la figura 1.5) con lo cual la ecuación de una recta en lenguaje vectorial será:

$$\mathbf{x} = \mathbf{x}_1 + \lambda\mathbf{a} \Rightarrow x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}} + z_1\hat{\mathbf{k}} + \lambda(a^1\hat{\mathbf{i}} + a^2\hat{\mathbf{j}} + a^3\hat{\mathbf{k}}) \Rightarrow \begin{cases} x = x_1 + \lambda a^1 \\ y = y_1 + \lambda a^2 \\ z = z_1 + \lambda a^3 \end{cases}$$

donde $\mathbf{x} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}$ es el conjunto de puntos genéricos que cumple con la ecuación de la recta en 3D.

Existe una manera más elegante, como veremos en la sección siguiente, de reescribir las ecuaciones anteriores utilizando la notación de índices². Las ecuaciones ahora son más evidentes:

$$\mathbf{x} = \mathbf{x}_1 + \lambda\mathbf{a} \Rightarrow x^i\hat{\mathbf{i}}_i = x_{(1)}^i\hat{\mathbf{i}}_i + \lambda a^i\hat{\mathbf{i}}_i \Rightarrow x^i = x_{(1)}^i + \lambda a^i, \quad \text{para } i = 1, 2, 3.$$

donde: $(x, y, z) \equiv (x^1, x^2, x^3)$ y $(\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}) \equiv (\hat{\mathbf{i}}_1, \hat{\mathbf{i}}_2, \hat{\mathbf{i}}_3)$.

Nótese que efectivamente se cumplen tres ecuaciones escalares y cada una de éstas tiene la forma de una recta. Además, tal y como se muestra la figura 1.5, el punto genérico (x, y, z) lo describe (sobre la recta) la variación del módulo de \mathbf{a} mediante la constante de proporcionalidad λ . Si se requiere describir una recta

²Suprimimos aquí el símbolo de sumatoria, esta convención quedará clara en la siguiente sección, pero mantengamos en mente que tenemos una suma sobre el índice i .

que pase por dos puntos: (x_1, y_1, z_1) y (x_2, y_2, z_2) entonces una vez seleccionado uno de los puntos (digamos (x_1, y_1, z_1)) seleccionamos el vector $\mathbf{a} = \mathbf{r}(P_2) - \mathbf{r}(P_1)$ como la resta de los dos radio vectores a los puntos P_2 y P_1 . Esto es:

$$\mathbf{x} = \mathbf{x}_1 + \lambda (\mathbf{x}_2 - \mathbf{x}_1).$$

Al despejar λ de la ecuaciones de las rectas resulta:

$$x^i = x_{(1)}^i + \lambda a^i \Rightarrow \lambda = \frac{x^i - x_{(1)}^i}{a^i} = \frac{x - x_1}{a^1} = \frac{y - y_1}{a^2} = \frac{z - z_1}{a^3},$$

y de manera equivalente:

$$x^i = x_{(1)}^i + \lambda \left(x_{(2)}^i - x_{(1)}^i \right) \Rightarrow \lambda = \frac{x^i - x_{(1)}^i}{x_{(2)}^i - x_{(1)}^i} = \frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}.$$

1.3.2 Planos y vectores

Ocurre exactamente lo mismo cuando construimos la ecuación vectorial para un plano. En general una superficie la define su vector normal (perpendicular). En el caso de una superficie plana (un plano) tendrá una única normal que lo define, por lo tanto, un plano vendrá definido por su vector perpendicular en un punto, digamos $Q = P_1 : (x_1, y_1, z_1)$. La ecuación vectorial del plano vendrá definida por todos los vectores \overrightarrow{PQ} tales que sean perpendiculares a un determinado vector \mathbf{a} (ver cuadrante II de la figura 1.5). Donde el punto P es un punto genérico (x, y, z) que define un radio vector. La ecuación vectorial del plano será simplemente:

$$\mathbf{a} \cdot \overrightarrow{PQ} = \mathbf{a} \cdot \left[\mathbf{r}(P) - \underbrace{\mathbf{r}(P_1)}_{\mathbf{b}} \right] = 0 \Leftrightarrow \mathbf{a} \cdot (\mathbf{r} - \mathbf{r}_1) = 0 \Leftrightarrow \mathbf{a} \cdot \mathbf{r} = \underbrace{\mathbf{a} \cdot \mathbf{r}_1}_b.$$

Esto es, se tiene que cumplir la condición:

$$\begin{aligned} & \left(a^1 \hat{\mathbf{i}} + a^2 \hat{\mathbf{j}} + a^3 \hat{\mathbf{k}} \right) \cdot \left[(x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}}) - (x_1 \hat{\mathbf{i}} + y_1 \hat{\mathbf{j}} + z_1 \hat{\mathbf{k}}) \right] = 0 \\ & \left(a^1 \hat{\mathbf{i}} + a^2 \hat{\mathbf{j}} + a^3 \hat{\mathbf{k}} \right) \cdot \left[(x - x_1) \hat{\mathbf{i}} + (y - y_1) \hat{\mathbf{j}} + (z - z_1) \hat{\mathbf{k}} \right] = 0 \\ & a^1 (x - x_1) + a^2 (y - y_1) + a^3 (z - z_1) = 0, \end{aligned}$$

con lo cual, la ecuación del plano queda como siempre la hemos conocido:

$$a^1 x + a^2 y + a^3 z - a^1 x_1 - a^2 y_1 - a^3 z_1 = 0 \Rightarrow a^1 x + a^2 y + a^3 z = b = a^1 x_1 + a^2 y_1 + a^3 z_1.$$

Nuevamente, de manera compacta³

$$a^i x_i - a^j x_{j(1)} = 0 \Rightarrow a_k x^k = b = a_l x_{(1)}^l.$$

Es claro que $\mathbf{a} \cdot \mathbf{r}_1 = b$ es la proyección del radio vector $\mathbf{r}(P_1)$ sobre la perpendicular que define al plano. Por lo tanto será la distancia entre el plano y el origen de coordenadas. Si $b = 0$ el plano pasa por el origen de coordenadas.

Consideremos ahora el cuadrante III de la figura 1.5. Allí están especificados tres puntos en el espacio caracterizados por sus correspondientes radio vectores posición: $\mathbf{r}(P_1) = \mathbf{r}_1$, $\mathbf{r}(P_2) = \mathbf{r}_2$ y $\mathbf{r}(P_3) = \mathbf{r}_3$. Estos tres puntos serán coplanares si:

$$(\mathbf{r}_1 - \mathbf{r}_2) \cdot [(\mathbf{r}_2 - \mathbf{r}_3) \times (\mathbf{r}_3 - \mathbf{r}_1)] = 0 \Leftrightarrow \varepsilon_{mnl} (x_1^m - x_2^m) (x_2^n - x_3^n) (x_3^l - x_1^l) = 0,$$

y la ecuación vectorial del plano vendrá dada por:

$$(\mathbf{r} - \mathbf{r}_1) \cdot [(\mathbf{r}_2 - \mathbf{r}_1) \times (\mathbf{r}_3 - \mathbf{r}_1)] = 0.$$

³Los índices repetidos arriba y abajo indican suma

Ejemplo 1.11 Un plano viene determinado por los puntos $A = (1, 1, 1)$, $B = (1, 2, 3)$ y $C = (0, 0, 0)$. Encuentre la ecuación del plano

Para encontrar la ecuación del plano podemos hacer lo siguiente: Encontremos el vector posición de los puntos A y B ,

$$\mathbf{r}_{AB} = B - A = (0, 1, 2), \quad \mathbf{r}_{AC} = C - A = (-1, -1, -1),$$

un vector normal al plano es: $\mathbf{n} = \mathbf{r}_{AB} \times \mathbf{r}_{AC} = (1, -2, 1)$.

Para la ecuación del plano, podemos escoger el vector $\mathbf{a} = (1, 1, 1)$ por lo que tenemos entonces que:

$$\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{a} \Rightarrow (1, -2, 1) \cdot (x, y, z) = (1, -2, 1) \cdot (1, 1, 1) \Rightarrow x - 2y + z = 0.$$

▷

Ejemplo 1.12 Dados los siguientes puntos en el espacio:

$$A = (1, 0, 3), B = (2, -1, 0), C = (0, -1, 1), d = (-1, 0, 1).$$

¿Estos puntos serán coplanares?

Consideremos los tres primeros puntos, que podemos considerar coplanares ya que bastan tres puntos para definir un plano. Estos tres puntos son los vértices de un triángulo cuya área podemos calcular de la siguiente manera:

Primero seleccionamos uno de los puntos como un vértice privilegiado (digamos $(2, -1, 0)$) respecto al cual construiremos dos vectores que representan dos de los lados del triángulo. Esto es:

$$\mathbf{a} = (1, 0, 3) - (2, -1, 0) \Leftrightarrow \mathbf{a} = -\hat{\mathbf{i}} + \hat{\mathbf{j}} + 3\hat{\mathbf{k}}, \quad \mathbf{b} = (0, -1, 1) - (2, -1, 0) \Leftrightarrow \mathbf{b} = -2\hat{\mathbf{i}} + \hat{\mathbf{k}},$$

con lo cual, el área del triángulo será la mitad del área del paralelogramo que tiene por lados estos dos vectores. Es decir:

$$A = \frac{1}{2}|\mathbf{a} \times \mathbf{b}| \Rightarrow \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ -1 & 1 & 3 \\ -2 & 0 & 1 \end{vmatrix} = \hat{\mathbf{i}} - 5\hat{\mathbf{j}} + 2\hat{\mathbf{k}} \Rightarrow A = \frac{1}{2}|\hat{\mathbf{i}} - 5\hat{\mathbf{j}} + 2\hat{\mathbf{k}}| = \frac{\sqrt{30}}{2}.$$

Por otro lado, la ecuación del plano que generan estos tres puntos se calcula con la siguiente ecuación:

$$(\mathbf{r} - \mathbf{r}_1) \cdot ((\mathbf{r}_2 - \mathbf{r}_1) \times (\mathbf{r}_3 - \mathbf{r}_1)) = 0,$$

donde:

$$\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}, \quad \mathbf{r}_1 = \hat{\mathbf{i}} + 3\hat{\mathbf{k}}, \quad \mathbf{r}_2 = 2\hat{\mathbf{i}} - \hat{\mathbf{j}}, \quad \mathbf{r}_3 = -\hat{\mathbf{j}} + \hat{\mathbf{k}},$$

con lo cual la ecuación del plano queda como:

$$\begin{vmatrix} (x-1) & y & (z-3) \\ 1 & -1 & -3 \\ -1 & -1 & -2 \end{vmatrix} = 0 \Rightarrow -(x-1) + 5y - 2(z-3) = 0 \Rightarrow x - 5y + 2z = 7.$$

Podemos verificar si el cuarto punto, $(-1, 0, 1)$, se encuentra en el plano, es decir, debemos verificar que cumple la ecuación que lo define.

$$x - 5y + 2z = 7 \Rightarrow (-1) - 5(0) + 2(1) \neq 7,$$

por lo tanto, los cuatro puntos no son coplanares. Podemos entonces calcular la distancia del cuarto punto al plano construyendo un vector unitario normal al plano.

$$\hat{\mathbf{n}}_P = \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|} = \frac{1}{\sqrt{30}} (\hat{\mathbf{i}} - 5\hat{\mathbf{j}} + 2\hat{\mathbf{k}}),$$

con lo cual la distancia al cuarto punto será:

$$d = |\hat{\mathbf{n}}_P \cdot \mathbf{c}| = \left| \frac{1}{\sqrt{30}} (\hat{\mathbf{i}} - 5\hat{\mathbf{j}} + 2\hat{\mathbf{k}}) \cdot (-3\hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}}) \right| = \left| -\frac{6}{\sqrt{30}} \right| = \frac{6}{\sqrt{30}}.$$



1.3.3 Ejercicios

1. Para las rectas dadas a continuación:

- (a). $L : \frac{3x-1}{4} = \frac{2y+3}{2} = 2 - 3z$.
(b). $L : \frac{2x+1}{3} = \frac{3y+2}{3} = -2 + 4z$.

Encuentre los vectores posición para dos puntos diferentes sobre la recta y un vector unitario paralelo a la recta L .

2. Dada una linea recta L_1 que pasa a través de los puntos $(-2, 3, 1)$ y $(1, 4, 6)$ encuentre:

- (a). El vector posición de un punto sobre la recta y un vector paralelo a ésta.
(b). Una recta L_2 paralela a L_1 y que pase por el punto $(1, 2, 1)$.

3. Una linea recta tiene como ecuación vectorial: $\mathbf{r} = \mathbf{a} + \lambda\mathbf{b}$, donde: $\mathbf{a} = 3\hat{\mathbf{j}} + 2\hat{\mathbf{k}}$ y $\mathbf{b} = 2\hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}$.

Encuentre la ecuación cartesiana de la recta y las coordenadas de tres puntos sobre la recta.

4. Una linea recta pasa por el punto $(3, 2, -3)$ y paralela al vector $\mathbf{a} = 2\hat{\mathbf{i}} + 3\hat{\mathbf{j}} - 3\hat{\mathbf{k}}$. Encuentre la ecuación cartesiana de la recta y las coordenadas de tres puntos sobre la recta.

5. Dado un plano que pasa por el punto $(2, 3, -5)$ y con vector normal $\mathbf{a} = 2\hat{\mathbf{i}} + \hat{\mathbf{k}}$, encuentre la forma cartesiana de la ecuación del plano.

6. Encuentre la ecuación del plano con normal \mathbf{a} y que contiene el punto P cuando:

- (a). $\mathbf{a} = 2\hat{\mathbf{i}} - 3\hat{\mathbf{j}} + \hat{\mathbf{k}}$, $P = (1, 0, 1)$.
(b). $\mathbf{a} = \hat{\mathbf{i}} - 2\hat{\mathbf{j}} + 2\hat{\mathbf{k}}$, $P = (2, -3, 4)$.

7. El ángulo entre dos planos se define como el ángulo entre sus normales. Encuentre el ángulo entre los siguientes planos:

- (a). $x + 3y + 2z = 4$ y $2x - 5y + z = 2$.
(b). $3x + 2y - 2z = 4$ y $2x + y + 2z = 1$.

8. Demuestre que la ecuación de una esfera puede expresarse como:

$$|\mathbf{r} - \mathbf{c}|^2 = (\mathbf{r} - \mathbf{c}) \cdot (\mathbf{r} - \mathbf{c}) = a^2,$$

donde \mathbf{c} es el vector posición del centro de la esfera y a el radio.

9. Auguste Bravais⁴ se dio cuenta que replicando un arreglo geométrico muy simple, se puede describir una estructura cristalina. Dicho de otro modo, que conociendo una celda simple, podemos conocer la estructura cristalina. Esto es, que las posiciones de los átomos en una red cristalina puede ser descrita por un vector:

$$\mathbf{R} = \mathbf{a} + \mathbf{b} + \mathbf{c} = n^1\mathbf{a}_1 + n^2\mathbf{a}_2 + n^3\mathbf{a}_3 = n^i\mathbf{a}_i,$$

donde los \mathbf{a}_i son vectores no coplanares (vectores primitivos o, simplemente en nuestro lenguaje, vectores base). Los n^i son números enteros (negativos, cero o positivos). La posición de cada átomo de un cristal puede ser descrita como reescalamientos (discretos) de este vector genérico o, de manera más precisa, la traslación del origen de coordenadas por un vector. Ese concepto se conoce como redes de Bravais⁵. En

⁴http://en.wikipedia.org/wiki/Auguste_Bravais

⁵http://en.wikipedia.org/wiki/Bravais_lattice

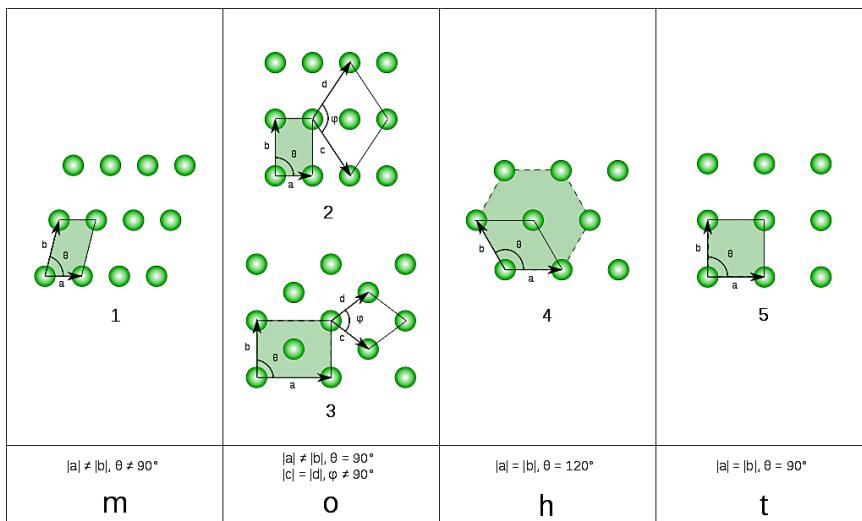


Figura 1.6: Las 5 redes de Bravais bidimensionales fundamentales: 1 Oblicuas, 2 rectangular, 3 rectangular centrada (rómbica), 4 hexagonal, y 5 cuadrada. Figura tomada de http://en.wikipedia.org/wiki/Bravais_lattice.

cada red puede haber varios vectores primitivos⁶. Se puede definir la *celda primitiva* como la estructura mínima que replicada reproduce todo el cristal. Vale decir, la estructura cristalina es invariante bajo traslaciones espaciales del tipo:

$$\mathbf{R}' = \mathbf{R} + \mathbf{T}, \quad \text{con } \mathbf{T} = m^i \mathbf{a}_i.$$

Uno también se puede plantear redes aperiódicas que no presentan simetrías de traslación pero sí de rotación y reflexión. Estas estructuras fueron estudiadas por R. Penrose⁷ en los años 70 como meras curiosidades matemáticas. Existen varias herramientas en la red que les permite construir interesantes figuras de los embaldosados de Penrose.⁸ Luego, a principio de la década de los 80, se descubren materiales cuasi-cristalinos (cuasi-cristales o cristales cuasiperiódicos) que pueden ser modelados con este tipo de estructuras. Este tipo de materiales y los modelos que los describen motivaron un premio Nobel⁹.

(a). **Redes de Bravais bidimensionales.** Tal y como muestra la figura 1.6 existen 5 tipos distintos de redes de Bravais bidimensionales.

- I. Dada la red bidimensional de la figura 1.7 (Izquierda) encuentre todos los posibles vectores primitivos y celdas primitivas asociadas.
- II. La humanidad ha estado seducida por la geometría desde que empezó a representar figuras. A partir de las cuatro imágenes que se ilustran en la figura 1.7 (Centro), encuentre todos los posibles vectores y celdas primitivas asociadas.
- III. Maurits Cornelis Escher¹⁰ fue un fenomenal dibujante holandés, quien se interesó por las simetrías de los grupos de imágenes de papel tapiz. Berend, hermano de Maurits, era cristalógrafo y le mostró la belleza de las simetrías de la naturaleza.

⁶<http://www.engr.sjsu.edu/rkwok/Phys175A/Chapter%201.pdf>

⁷Roger Penrose (Premio Nobel 2020), por descubrir que la formación de agujeros negros es una predicción de la Relatividad General.

⁸<https://misc.0o0.org/penrose/>

⁹El Premio Nobel de Química 2011 fue otorgado a Dan Shechtman por el descubrimiento de los cuasi-cristales.

¹⁰http://en.wikipedia.org/wiki/M._C._Escher



Figura 1.7: A la izquierda red cristalina bidimensional. Al centro cuatro detalles geométricos: mural egipcio, mural asirio, tejido tahití e ilustración en pieza de porcelana china. Tomado de: http://en.wikipedia.org/wiki/Wallpaper_group. Finalmente, a la derecha, teselados de M.C. Escher, tomados de: <http://www.wikipaintings.org/en/paintings-by-genre/tessellation?firstArtist=m-c-escher#artist-m-c-escher>.

En las cuatro obras del género de teselado¹¹ de M.C. Escher, presentadas en la figura 1.7 (Derecha) encuentre todos los posibles vectores y celdas primitivas asociadas.

- (b). **Redes de Bravais tridimensionales.** Este tipo de redes complica un poco más el escenario. Se puede demostrar que existen 14 de estas redes, tal y como se muestran en la figura 1.8. Muestre que los volúmenes de ocupación atómica, para los sistemas: monoclinico, triclinico, ortorombico, tetragonal, romboédrico, hexagonal y cúbico, corresponden a las expresiones que se muestran en la figura 1.8.
- (c). El sistema cúbico, el más simple, corresponde a un sistema con un único parámetro de red $a = |\mathbf{a}|$, ya que $\mathbf{a} = \mathbf{b} = \mathbf{c}$. Además, una posible descripción, para el caso más simple, es $\mathbf{a} = \hat{\mathbf{i}}$, $\mathbf{b} = \hat{\mathbf{j}}$, $\mathbf{c} = \hat{\mathbf{k}}$, los tres vectores cartesianos ortogonales. Existen otros sistemas que también están asociados al cúbico. Estos son el sistema cúbico cara centrada (*fcc* por sus siglas en inglés) y cúbico cuerpo centrado (*bcc*). En el primero existen átomos en el centro de cada una de las caras del cubo definido por la tríada, $\mathbf{a} = \mathbf{b} = \mathbf{c}$. En el sistema *fcc* se añade un átomo en el centro del cubo simple.

I. Muestre que un sistema *bcc* también puede ser descrito por los vectores primitivos:

$$\mathbf{a} = a\hat{\mathbf{i}}, \quad \mathbf{b} = a\hat{\mathbf{j}}, \quad \mathbf{c} = a(\hat{\mathbf{i}} + \hat{\mathbf{j}} + \hat{\mathbf{k}})/2.$$

Dibuje la celda primitiva y calcule su volumen.

II. Muestre que un sistema *bcc* también puede ser descrito por los vectores primitivos:

$$\mathbf{a} = a(\hat{\mathbf{j}} + \hat{\mathbf{k}} - \hat{\mathbf{i}})/2, \quad \mathbf{b} = a(\hat{\mathbf{k}} + \hat{\mathbf{i}} - \hat{\mathbf{j}})/2, \quad \mathbf{c} = a(\hat{\mathbf{i}} + \hat{\mathbf{j}} - \hat{\mathbf{k}})/2.$$

Dibuje la celda primitiva y calcule su volumen.

III. Muestre que un sistema *fcc* también puede ser descrito por los vectores primitivos:

$$\mathbf{a} = a(\hat{\mathbf{j}} + \hat{\mathbf{k}})/2, \quad \mathbf{b} = a(\hat{\mathbf{i}} + \hat{\mathbf{k}})/2, \quad \mathbf{c} = a(\hat{\mathbf{i}} + \hat{\mathbf{j}})/2.$$

Otra vez, dibuje la celda primitiva y calcule su volumen.

- (d). Se puede definir la red recíproca como:

$$\mathbf{a}' = \frac{\mathbf{b} \times \mathbf{c}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}, \quad \mathbf{b}' = \frac{\mathbf{c} \times \mathbf{a}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})} \quad \text{y} \quad \mathbf{c}' = \frac{\mathbf{a} \times \mathbf{b}}{\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})}.$$

De esta manera es claro que, por construcción, $\mathbf{a}' \cdot \mathbf{b} = \mathbf{a}' \cdot \mathbf{c} = 0$ y además $\mathbf{a}' \cdot \mathbf{a} = 1$. Con lo cual podemos generalizarlo como $\hat{\mathbf{e}}^{i'} \cdot \hat{\mathbf{e}}_j = \delta_j^{i'}$. Exprese los vectores y las celdas recíprocas para los

¹¹<http://en.wikipedia.org/wiki/Tessellation>

Crystal family	Lattice system	14 Bravais lattices				
		Primitive	Base-centered	Body-centered	Face-centered	Rhombohedrally-centered
triclinic						
monoclinic		$\beta \neq 90^\circ$ $a \neq c$ 	$\beta = 90^\circ$ $a = c$ 			
orthorhombic		$a \neq b \neq c$ 	$a \neq b \neq c$ 	$a \neq b \neq c$ 	$a \neq b \neq c$ 	
tetragonal		$a \neq c$ 		$a \neq c$ 		
hexagonal	rhombohedral					
	hexagonal	$\gamma = 120^\circ$ 				
cubic						

Figura 1.8: Las 14 Redes de Bravais tridimensionales y las estructuras cristalinas asociadas. Tomado de: http://en.wikipedia.org/wiki/Bravais_lattice.

sistemas cúbico simple, y los distintos *bcc* y *fcc*. Calcule además el volumen de cada celda recíproca.

10. Realice los cálculos anteriores utilizando el programa **Maxima**.

Practicando con Maxima

Consideremos el plano determinado por los puntos $P = (1, 1, 1)$, $Q = (2, 4, 6)$ y $R = (2, 2, 2)$. Para construir la ecuación del plano primero ingresamos los vectores posición de cada punto. También incorporaremos las bibliotecas **vect** y **draw** que necesitaremos.

```
(%ci1) load(vect)$ load(draw)$
(%ci2) a:[1,1,1];
(%o2) [1, 1, 1]
(%ci3) b:[2,4,6];
(%o3) [2, 4, 6]
(%ci4) c:[2,2,2];
(%o4) [2, 2, 2]
```

Podemos verificar que los vectores estén en el mismo plano simplemente calculando el triple producto vectorial: $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ entre ellos. Sabemos que si es nulo es porque los vectores son coplanares.

```
(%ci5) a.express(b ~ c);
(%o5) 0
```

Necesitamos ahora calcular los vectores que van del punto P al punto Q y del punto P al punto R , es decir, los vectores: $\mathbf{ab} = \mathbf{b} - \mathbf{a}$ y $\mathbf{ac} = \mathbf{c} - \mathbf{a}$

(%i6) $\text{ab}: \mathbf{b}-\mathbf{a};$
 (%o6) $[1, 3, 5]$

(%i7) $\text{ac}: \mathbf{c}-\mathbf{a};$
 (%o7) $[1, 1, 1]$

Un vector normal al plano será sencillamente el vector $\mathbf{n} = \mathbf{ab} \times \mathbf{ac}$:

(%i8) $\text{n}: \text{express}(\mathbf{ab} \sim \mathbf{ac});$
 (%o8) $[-2, 4, -2]$

Podemos escoger cualquiera de los vectores originales, en este caso al vector \mathbf{a} , para escribir la ecuación del plano: $\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{a}$.

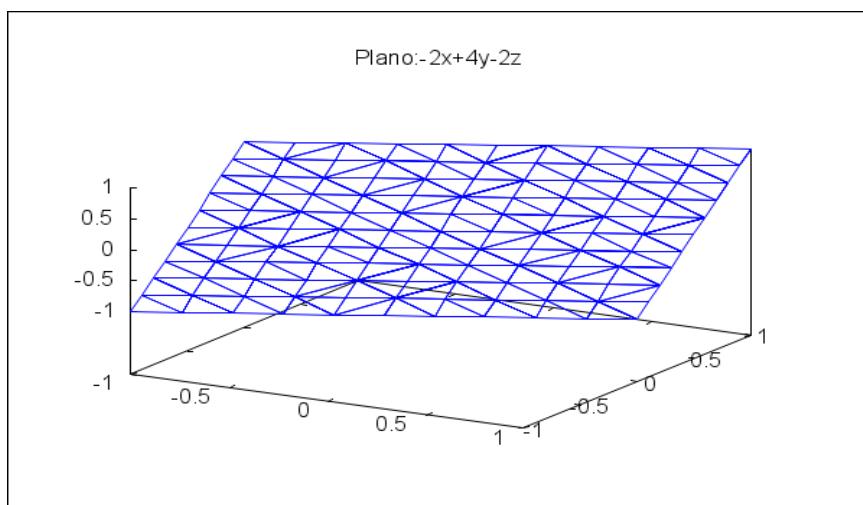
(%i9) $\text{r}: [\mathbf{x}, \mathbf{y}, \mathbf{z}];$
 (%o9) $[x, y, z]$

(%i10) $\text{plano}: \mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{a};$
 (%o10) $-2z + 4y - 2x = 0$

Probemos a graficar el plano. Para saber más sobre las diferentes opciones que incorpora **Maxima** para hacer gráficas, en dos o tres dimensiones, es recomendable consultar el manual de usuario.

Utilizaremos el comando **wxdraw3d** para hacer que la figura aparezca inmersa en nuestra hoja de trabajo. Es recomendable consultar también las funciones **draw** y **gr3d**.

(%i11) $\text{wxdraw3d}(\text{implicit}(\text{plano}, \mathbf{x}, -1, 1, \mathbf{y}, -1, 1, \mathbf{z}, -1, 1));$
 (%o11)



1.4 Álgebra vectorial con índices

Antes de comenzar con la presentación de este esquema de cálculo cabe aclarar algunas costumbres y convenciones con la notación de índices que estaremos utilizando durante el resto de este curso.

1.4.1 Convención de Einstein

El convenio de suma de Einstein, es una simplificación que se utiliza para abreviar la escritura de las sumatorias, en el que se suprime el símbolo de sumatoria y consiste en lo siguiente:

1. Los índices repetidos (arriba y abajo) indicarán suma por los valores que tomen los índices. Las componentes de los vectores tendrán índices arriba y los vectores base abajo:

$$\mathbf{a} = a^1 \mathbf{e}_1 + a^2 \mathbf{e}_2 + a^3 \mathbf{e}_3 = \sum_{m=1}^3 a^m \mathbf{e}_m \Leftrightarrow \mathbf{a} = a^m \mathbf{e}_m = a^i \mathbf{e}_i .$$

2. Los índices repetidos son mudos (no importa las letras que los etiquete) y representan suma. Así:

$$k^j a_j = k^m a_m = k^1 a_1 + k^2 a_2 + k^3 a_3 = b .$$

En este punto del discurso, la posición de los índices (arriba y abajo) solo tiene sentido estético y solo así indican suma. Más adelante veremos que representan cantidades distintas.

3. Llamaremos contracción cuando sumamos respecto a un par de índices, vale decir:

$$A_i^j \Rightarrow \sum_i A_i^i = A_1^1 + A_2^2 + A_3^3 \Rightarrow A_i^i = A_1^1 + A_2^2 + A_3^3 .$$

Las cantidades con dos o más índices las llamaremos componentes de tensores, y deben entenderse como arreglos bidimensionales (tridimensionales, tetradimensionales, según el número de índices). Estas cantidades serán considerados en detalle posteriormente pero por ahora, contentémonos con saber qué cosas son cantidades con dos índices. Es claro que la contracción de índices convierte un conjunto de números $(i \times j) \rightarrow 1$, en un sólo número.

Los índices libres (aquellos que no están sumados) indican el número de objetos disponibles y deben mantenerse a ambos lados de la ecuación. Por ejemplo:

$$B_i = K_i^k A_k \Leftrightarrow \begin{cases} K_1^1 A_1 + K_1^2 A_2 + K_1^3 A_3 = B_1 \\ K_2^1 A_1 + K_2^2 A_2 + K_2^3 A_3 = B_2 \\ K_3^1 A_1 + K_3^2 A_2 + K_3^3 A_3 = B_3 \end{cases}$$

con lo cual $B_i = K_i^k A_k$ representa 3 ecuaciones. La operación $B_{ij} = K_i^k A_{kj}$ representa 9.

La delta de Kronecker¹² es un objeto matemático de dos índices, representa $\delta_i^k = 1$ si $i = k$, y es nula en los otros casos. Por ejemplo:

$$K_{ij}^k \delta_k^i = K_{1j}^1 \underbrace{\delta_1^1}_{=1} + K_{2j}^1 \underbrace{\delta_1^2}_{=0} + K_{3j}^1 \underbrace{\delta_1^3}_{=0} + K_{1j}^2 \underbrace{\delta_2^1}_{=0} + K_{2j}^2 \underbrace{\delta_2^2}_{=1} + K_{3j}^2 \underbrace{\delta_2^3}_{=0} + K_{1j}^3 \underbrace{\delta_3^1}_{=0} + K_{2j}^3 \underbrace{\delta_3^2}_{=0} + K_{3j}^3 \underbrace{\delta_3^3}_{=1},$$

es decir:

$$K_{ij}^k \delta_k^i = K_{kj}^k = K_{ij}^i = K_{1j}^1 + K_{2j}^2 + K_{3j}^3 .$$

¹²Leopold Kronecker (1823-1891, Berlin, Alemania) Matemático polaco con importantes contribuciones en teoría de números, funciones elípticas y álgebra, así como la interrelación entre estas disciplinas. https://es.wikipedia.org/wiki/Leopold_Kronecker

Además de la delta de Kronecker introduciremos el símbolo de permutación de Levi-Civita¹³, ε^{ijk} , para el caso de tres dimensiones: $i, j, k = 1, 2, 3$.

$$\varepsilon_{ijk} = \varepsilon^{ijk} = \begin{cases} +1 & \text{cuando } \{(i, j, k) = (1, 2, 3); (3, 1, 2); (2, 3, 1)\} \text{ permutación cíclica} \\ -1 & \text{cuando } \{(i, j, k) = (1, 3, 2); (3, 2, 1); (2, 1, 3)\} \text{ permutación impar o anticíclica} \\ 0 & \text{cuando } \{i = j, i = k \wedge j = k\} \end{cases}$$

Por lo tanto, es distinto de cero cuando todos los índices son diferentes. Toma el valor 1 si la permutación de índices es cíclica (o par), y toma el valor -1 si la permutación es anticíclica (o impar).

Si queremos calcular, por ejemplo: $c^i = \varepsilon^{ijk} a_j b_k$, entonces resulta:

$$\begin{aligned} c^1 &= \varepsilon^{111} a_1 b_1 + \varepsilon^{112} a_1 b_2 + \varepsilon^{113} a_1 b_3 + \varepsilon^{121} a_2 b_1 + \varepsilon^{122} a_2 b_2 + \varepsilon^{123} a_2 b_3 + \varepsilon^{131} a_3 b_1 + \varepsilon^{132} a_3 b_2 + \varepsilon^{133} a_3 b_3, \\ c^2 &= \varepsilon^{211} a_1 b_1 + \varepsilon^{212} a_1 b_2 + \varepsilon^{213} a_1 b_3 + \varepsilon^{221} a_2 b_1 + \varepsilon^{222} a_2 b_2 + \varepsilon^{223} a_2 b_3 + \varepsilon^{231} a_3 b_1 + \varepsilon^{232} a_3 b_2 + \varepsilon^{233} a_3 b_3, \\ c^3 &= \varepsilon^{311} a_1 b_1 + \varepsilon^{312} a_1 b_2 + \varepsilon^{313} a_1 b_3 + \varepsilon^{321} a_2 b_1 + \varepsilon^{322} a_2 b_2 + \varepsilon^{323} a_2 b_3 + \varepsilon^{331} a_3 b_1 + \varepsilon^{332} a_3 b_2 + \varepsilon^{333} a_3 b_3, \end{aligned}$$

con lo cual:

$$c^i = \varepsilon^{ijk} a_j b_k \Rightarrow \begin{cases} c^1 = \varepsilon^{123} a_2 b_3 + \varepsilon^{132} a_3 b_2 = a_2 b_3 - a_3 b_2 \\ c^2 = \varepsilon^{231} a_3 b_1 + \varepsilon^{213} a_1 b_3 = a_3 b_1 - a_1 b_3 \\ c^3 = \varepsilon^{312} a_1 b_2 + \varepsilon^{321} a_2 b_1 = a_1 b_2 - a_2 b_1 \end{cases}$$

A continuación enumeramos algunas propiedades de la delta de Kronecker y del símbolo de permutación de Levi-Civita, dejamos al lector su demostración. Ellas son:

$$\begin{aligned} \delta_j^j &= 3, \\ \varepsilon_{jkm} \varepsilon^{ilm} &= \delta_j^i \delta_k^l - \delta_k^i \delta_j^l = \delta_j^i \delta_k^l - \delta_j^l \delta_k^i, \\ \varepsilon_{jmn} \varepsilon^{imn} &= 2\delta_j^i, \\ \varepsilon_{ijk} \varepsilon^{ijk} &= 6. \end{aligned}$$

1.4.2 Vectores e índices

Disponemos ahora de una manera más elegante para escribir ecuaciones que involucren vectores. Veamos qué forma toma el álgebra vectorial con esta nueva notación. Recordemos que con la notación de índices podemos escribir un vector como: $\mathbf{a} = a^i \mathbf{e}_i$, donde las cantidades a^i son las componentes del vector en la base $\{\mathbf{e}_i\}$.

1.4.2.1 Sumas de vectores

La suma de vectores será expresada de la siguiente manera:

$$\mathbf{a} + \mathbf{b} = a^i \mathbf{e}_i + b^i \mathbf{e}_i = (a^i + b^i) \mathbf{e}_i = c^i \mathbf{e}_i \Rightarrow c^i = a^i + b^i, \quad \text{con } i = 1, 2, 3. \quad (1.9)$$

1.4.2.2 Producto escalar

A partir da ahora y de forma equivalente, expresaremos el producto escalar en término de los índices. De forma y manera que:

$$\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}| |\mathbf{b}| \cos(\theta)_{\mathbf{ab}} = a^i b_i, \quad \text{con } i = 1, 2, 3. \quad (1.10)$$

¹³Tullio Levi-Civita (1873-1941, Roma, Italia) Geómetra italiano y uno de los desarrolladores del cálculo tensorial que más tarde sería utilizado por Einstein y Weyl como el lenguaje de la Relatividad General. https://es.wikipedia.org/wiki/Tullio_Levi-Civita

En este momento no explicaremos la razón por la que hemos escrito b_i , pero más adelante aclararemos esta situación. Por ahora seamos consistentes con la convención de índices de Einstein.

1.4.2.3 Producto vectorial

En términos de índices, la componente i del producto vectorial se puede expresar como:

$$c^i = (\mathbf{a} \times \mathbf{b})^i = \epsilon^{ijk} a_j b_k, \quad \text{con } i, j, k = 1, 2, 3. \quad (1.11)$$

De esta manera, todas las particularidades de producto vectorial ahora descansan en las propiedades del símbolo de Levi-Civita.

1.4.2.4 Triple producto mixto

Para finalizar, analicemos ahora el número (pseudoescalar) que proviene de la multiplicación mixta:

$$\mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}) = |\mathbf{c}| |\mathbf{a} \times \mathbf{b}| \cos(\theta)_{\langle \mathbf{c}, \mathbf{a} \times \mathbf{b} \rangle} = c^i \epsilon_{ijk} a^j b^k = \epsilon_{ijk} c^i a^j b^k = \begin{vmatrix} c^1 & c^2 & c^3 \\ a^1 & a^2 & a^3 \\ b^1 & b^2 & b^3 \end{vmatrix}. \quad (1.12)$$

Ejemplo 1.13 Mostraremos a continuación dos casos de identidades vectoriales que pueden ser fácilmente demostradas mediante la utilización de índices.

- $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{c} \cdot \mathbf{a}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}$

El resultado será un vector, por lo tanto:

$$\begin{aligned} (\mathbf{a} \times (\mathbf{b} \times \mathbf{c}))^i &= \epsilon^{ijk} a_j (\mathbf{b} \times \mathbf{c})_k = \epsilon^{ijk} a_j \epsilon_{kmn} b^m c^n = \epsilon^{ijk} \epsilon_{kmn} a_j b^m c^n = \epsilon^{ijk} \epsilon_{mnk} a_j b^m c^n \\ &= (\delta_m^i \delta_n^j - \delta_m^j \delta_n^i) a_j b^m c^n = \delta_m^i \delta_n^j a_j b^m c^n - \delta_m^j \delta_n^i a_j b^m c^n \\ &= \delta_m^i b^m \delta_n^j a_j c^n - \delta_n^i c^n \delta_m^j a_j b^m = \underbrace{b^i a_n c^n}_{(\mathbf{c} \cdot \mathbf{a})} - \underbrace{c^i a_j b^j}_{(\mathbf{a} \cdot \mathbf{b})} \\ &= b^i (\mathbf{c} \cdot \mathbf{a}) - c^i (\mathbf{a} \cdot \mathbf{b}) = \mathbf{b} (\mathbf{c} \cdot \mathbf{a}) - \mathbf{c} (\mathbf{a} \cdot \mathbf{b}). \end{aligned}$$

En la segunda línea hemos hecho uso de la identidad: $\epsilon_{jkm} \epsilon^{ilm} = \delta_j^i \delta_k^l - \delta_k^i \delta_j^l = \delta_j^i \delta_k^l - \delta_j^l \delta_k^i$.

- $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c}) (\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d}) (\mathbf{b} \cdot \mathbf{c})$

El lado derecho es un escalar, por lo tanto:

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) &= (\mathbf{a} \times \mathbf{b})^l (\mathbf{c} \times \mathbf{d})_l \\ &= \epsilon^{ljk} a_j b_k \epsilon_{lmn} c^m d^n = \epsilon^{ljk} \epsilon_{lmn} a_j b_k c^m d^n \\ &= \epsilon^{jkl} \epsilon_{mnl} a_j b_k c^m d^n = (\delta_m^j \delta_n^k - \delta_m^k \delta_n^j) a_j b_k c^m d^n \\ &= \delta_m^j \delta_n^k a_j b_k c^m d^n - \delta_m^k \delta_n^j a_j b_k c^m d^n \\ &= \underbrace{\delta_m^j a_j c^m}_{(\mathbf{a} \cdot \mathbf{c})} \underbrace{\delta_n^k b_k d^n}_{(\mathbf{b} \cdot \mathbf{d})} - \underbrace{\delta_m^k b_k c^m}_{(\mathbf{b} \cdot \mathbf{c})} \underbrace{\delta_n^j a_j d^n}_{(\mathbf{a} \cdot \mathbf{d})} \\ &= (\mathbf{a} \cdot \mathbf{c}) (\mathbf{b} \cdot \mathbf{d}) - (\mathbf{b} \cdot \mathbf{c}) (\mathbf{a} \cdot \mathbf{d}). \end{aligned}$$



Ejemplo 1.14 Si tenemos tres vectores $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$ queremos ver si se cumple:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}.$$

En notación de índices resulta:

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \epsilon^{lmi} a_m \epsilon_{ijk} b^j c^k + \epsilon^{lmi} b_m \epsilon_{ijk} c^j a^k + \epsilon^{lmi} c_m \epsilon_{ijk} a^j b^k,$$

con lo cual, arreglando:

$$\begin{aligned} \epsilon^{lmi} \epsilon_{ijk} a_m b^j c^k + \epsilon^{lmi} \epsilon_{ijk} b_m c^j a^k + \epsilon^{lmi} \epsilon_{ijk} c_m a^j b^k &= \left(\delta_j^l \delta_k^m - \delta_j^m \delta_k^l \right) a_m b^j c^k + \left(\delta_j^l \delta_k^m - \delta_j^m \delta_k^l \right) b_m c^j a^k \\ &\quad + \left(\delta_j^l \delta_k^m - \delta_j^m \delta_k^l \right) c_m a^j b^k, \end{aligned}$$

y ahora desarrollando los productos de las δ 's, e identificando término a término, notamos que se anulan:

$$\underbrace{a_k b^l c^k}_I - \underbrace{a_k b^k c^l}_I + \underbrace{b_k c^l a^k}_{II} - \underbrace{b_k c^k a^l}_{III} + \underbrace{c_k a^l b^k}_{III} - \underbrace{c_k a^k b^l}_I = 0.$$



1.4.3 Rotación de coordenadas

Consideremos un sistema de coordenadas cartesiano (x, y, z) y su base canónica $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$, si rotamos el sistema de coordenadas un ángulo ϕ alrededor del eje z tendremos un nuevo sistema de coordenadas (x', y', z') y una nueva base $\{\hat{\mathbf{i}}', \hat{\mathbf{j}}', \hat{\mathbf{k}}'\}$. La regla de transformación que relaciona ambos sistemas de coordenadas es:

$$\begin{cases} x = x' \cos \phi - y' \sin \phi \\ y = x' \sin \phi + y' \cos \phi \\ z = z' \end{cases} \iff \begin{cases} x' = x \cos \phi + y \sin \phi \\ y' = -x \sin \phi + y \cos \phi \\ z' = z \end{cases} \quad (1.13)$$

Mientras que las bases transformarán, como veremos más adelante, como:

$$\begin{cases} \hat{\mathbf{i}}' = \hat{\mathbf{i}} \cos \phi + \hat{\mathbf{j}} \sin \phi \\ \hat{\mathbf{j}}' = -\hat{\mathbf{i}} \sin \phi + \hat{\mathbf{j}} \cos \phi \\ \hat{\mathbf{k}}' = \hat{\mathbf{k}} \end{cases} \quad (1.14)$$

Diremos que un triplete de números (a^1, a^2, a^3) definen las componentes de un vector $\mathbf{a} = a^1 \hat{\mathbf{i}} + a^2 \hat{\mathbf{j}} + a^3 \hat{\mathbf{k}}$ si estas cantidades transforman bajo rotación de la siguiente manera:

$$\begin{cases} a^{1'} = a^1 \cos \phi + a^2 \sin \phi \\ a^{2'} = -a^1 \sin \phi + a^2 \cos \phi \\ a^{3'} = a^3 \end{cases} \quad (1.15)$$

Notemos también lo siguiente, al usar la notación de índices podemos escribir las ecuaciones de transformación de coordenadas de una manera compacta:

$$\begin{cases} x' = x \cos \phi + y \sin \phi \\ y' = -x \sin \phi + y \cos \phi \\ z' = z \end{cases} \Rightarrow \begin{cases} x^{1'} = A_1^{1'} x^1 + A_2^{1'} x^2 + A_3^{1'} x^3 \\ x^{2'} = A_1^{2'} x^1 + A_2^{2'} x^2 + A_3^{2'} x^3 \\ x^{3'} = A_1^{3'} x^1 + A_2^{3'} x^2 + A_3^{3'} x^3 \end{cases} \Rightarrow x^{i'} = A_j^{i'} x^j, \quad (1.16)$$

con: $i', j = 1, 2, 3$.

Se puede ver fácilmente que las cantidades $A_j^{i'}$, en coordenadas cartesianas, vienen dadas por:

$$A_j^{i'} = \frac{\partial \tilde{x}^i}{\partial x^j}. \quad (1.17)$$

Por ejemplo:

$$A_1^{1'} = \frac{\partial x^{1'}}{\partial x^1} = \frac{\partial x'}{\partial x} = \cos \phi, \quad A_2^{1'} = \frac{\partial x^{1'}}{\partial x^2} = \frac{\partial x'}{\partial y} = \sin \phi, \quad A_3^{1'} = \frac{\partial x^{1'}}{\partial x^3} = \frac{\partial x'}{\partial z} = 0 \dots$$

Como la transformación de coordenadas es invertible, se tiene que:

$$x^j = A_{i'}^j x^{i'}, \quad \text{con: } A_{i'}^j = \frac{\partial x^j}{\partial x^{i'}}, \quad (1.18)$$

y se cumple la siguiente condición de ortogonalidad:

$$A_k^{i'} A_{i'}^j = \delta_k^j. \quad (1.19)$$

Por lo tanto, en general, diremos que las componentes de un vector transformarán de la manera siguiente:

$$a^{i'} = \frac{\partial x^{i'}}{\partial x^j} a^j \equiv A_j^{i'} a^j \Leftrightarrow a^i = \frac{\partial x^i}{\partial x^{j'}} a^{j'} \equiv A_{j'}^i a^{j'} . \quad (1.20)$$

$$\begin{pmatrix} a^{1'} \\ a^{2'} \\ a^{3'} \end{pmatrix} = \begin{pmatrix} \cos \phi & \operatorname{sen} \phi & 0 \\ -\operatorname{sen} \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} \Leftrightarrow \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} \cos \phi & -\operatorname{sen} \phi & 0 \\ \operatorname{sen} \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^{1'} \\ a^{2'} \\ a^{3'} \end{pmatrix} .$$

En esta ecuación hemos recurrido a la notación de matrices para representar la transformación de coordenadas.

Por lo tanto, el vector \mathbf{a} se podrá escribir en ambos sistemas de coordenadas:

$$\mathbf{a} = a^1 \hat{\mathbf{i}} + a^2 \hat{\mathbf{j}} + a^3 \hat{\mathbf{k}} = a^{1'} \hat{\mathbf{i}}' + a^{2'} \hat{\mathbf{j}}' + a^{3'} \hat{\mathbf{k}}' \Leftrightarrow \mathbf{a} = a^i \hat{\mathbf{i}}_i = a^i \hat{\mathbf{i}}'_i .$$

1.4.4 Escalares, pseudoescalares, vectores y pseudovectores

Además de las rotaciones podemos considerar otra clase de transformaciones: las reflexiones. Estas transformaciones, como por ejemplo: $(x^1, x^2, x^3) \rightarrow (-x^1, x^2, x^3)$, muestran una sutil diferencia entre dos tipos de cantidades vectoriales: los vectores polares y los axiales, diferencia que no se puede apreciar en las rotaciones.

Una reflexión en el plano yz se puede representar como un cambio de signo en la componente x del vector:

$$(a_x, a_y, a_z) \rightarrow (-a_x, a_y, a_z) , \quad (1.21)$$

o también, en nuestra notación de índices:

$$a^1 \hat{\mathbf{i}}_1 + a^2 \hat{\mathbf{i}}_2 + a^3 \hat{\mathbf{i}}_3 \rightarrow a^1 (-\hat{\mathbf{i}}_1) + a^2 \hat{\mathbf{i}}_2 + a^3 \hat{\mathbf{i}}_3 \equiv (-a^1) \hat{\mathbf{i}}_1 + a^2 \hat{\mathbf{i}}_2 + a^3 \hat{\mathbf{i}}_3 .$$

Diremos que los objetos que transforman de esta manera bajo reflexión son vectores polares o simplemente vectores. Nótese que esa transformación de coordenadas es equivalente a invertir el vector base: $\hat{\mathbf{i}}_1 \rightarrow -\hat{\mathbf{i}}_1$.

Ahora bien, si dos vectores polares \mathbf{a} y \mathbf{b} , son transformados bajo reflexión como en (1.21), entonces un vector axial \mathbf{c} transformará como:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \rightarrow \tilde{\mathbf{a}} \times \tilde{\mathbf{b}} = -\tilde{\mathbf{c}} .$$

Es decir, los vectores axiales o pseudovectores cambian de signo bajo reflexión de los vectores que los generan¹⁴.

De igual manera notamos que $d = \mathbf{a} \cdot \mathbf{b}$ queda invariante bajo la transformación (1.21), mientras que $V = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b})$ cambia de signo. Siguiendo el ejemplo de vectores y pseudovectores, llamaremos escalar a d y pseudoescalar a V .

Existen importantes cantidades físicas que vienen representadas por pseudovectores, entre ellas mencionamos: la cantidad de momento angular, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$; el torque $\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$ y la variación del campo de inducción magnética, $\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}$.

Pseudovectores y vectores representan distintos objetos geométricos. Los primeros se asocian a orientaciones de superficies, mientras que los segundos lo están con segmentos de rectas orientadas. El vector velocidad angular, $\boldsymbol{\omega}$, es un pseudovector por cuanto \mathbf{r} y \mathbf{v} son vectores polares y $\mathbf{v} = \boldsymbol{\omega} \times \mathbf{r}$. Algo equivalente ocurre con la fuerza de Lorentz $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$, donde \mathbf{B} es el campo magnético.

En general, podemos representar la reflexión (1.21) bajo el esquema que presentamos en (1.20), es decir, como transformaciones del tipo:

$$a^{i'} = A_j^{i'} a^j ,$$

¹⁴Esta inversión de vector axial puede comprobarse fácilmente utilizando la definición de producto vectorial (1.2.6.2) en la cual se aprecia que un cambio en el signo en a_x y b_x induce un cambio de signo en c_y y c_z .

O en forma de matrices:

$$\begin{pmatrix} a^{1'} \\ a^{2'} \\ a^{3'} \end{pmatrix} \equiv \begin{pmatrix} -a^1 \\ a^2 \\ a^3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a^1 \\ a^2 \\ a^3 \end{pmatrix}.$$

donde:

$$A_j^{i'} = \frac{\partial \tilde{x}^{i'}}{\partial x^j} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

es la matriz de transformación de coordenadas¹⁵.

Notemos que aquí el determinante de la matriz de transformación tiene como valor -1 ($\det |A_j^{i'}| = -1$) ¿Cuál es el valor del determinante de la matriz que podemos asociar a las rotaciones (1.16)?

El hecho de que el $\det |A_j^{i'}| = 1$ o $\det |A_j^{i'}| = -1$, permite clasificar a las transformaciones de coordenadas en *transformaciones propias* o *transformaciones impropias*, respectivamente.

Entonces, en general, diremos que las componentes de vectores y pseudovectores transformarán bajo reflexión de la siguiente manera:

$$\text{si: } A_j^{i'} = \frac{\partial x^{i'}}{\partial x^j} \Rightarrow \begin{cases} a^{i'} = A_j^{i'} a^j, & \text{vectores polares o simplemente vectores,} \\ p^{i'} = \det |A_j^{i'}| A_j^{i'} p^j, & \text{pseudovectores o vectores axiales.} \end{cases}$$

En el ejercicio 6 de la sección 2.2.4 de la página 103 mostraremos como estos conceptos de: escalares, vectores, pseudo-vectores y pseudo-escalares, se aplican en computación gráfica.

Ejemplo 1.15 Vimos que los vectores polares transforman bajo una reflexión respecto al plano yz de la forma:

$$\hat{\mathbf{i}}_1 \rightarrow -\hat{\mathbf{i}}_1 \Rightarrow (a^1, a^2, a^3) = (-a^1, a^2, a^3).$$

Para saber que pasa con el vector $\mathbf{c} = \mathbf{a} \times \mathbf{b}$, cuando lo sometemos a esta reflexión, estudiemos como se transforman sus componentes directamente de la definición del producto vectorial, miremos con respecto a la componente x de \mathbf{c} .

Entonces, para \mathbf{a} y \mathbf{b} vectores polares se tiene:

$$\begin{aligned} \mathbf{c}' = \mathbf{a}' \times \mathbf{b}' &= \left[(a^2 b^3 - a^3 b^2), (a^3 b^1 - a^1 b^3), (a^1 b^2 - a^2 b^1) \right] \\ &= [(a^2 b^3 - a^3 b^2), (a^3(-b^1) - (-a^1)b^3), ((-a^1)b^2 - a^2(-b^1))] \\ &= [(a^2 b^3 - a^3 b^2), -(a^3 b^1 - a^1 b^3), -(a^1 b^2 - a^2 b^1)], \end{aligned}$$

es decir, \mathbf{c}' es un vector axial por el hecho de obedecer a una ley de transformación diferente:

$$\hat{\mathbf{i}}_1 \rightarrow -\hat{\mathbf{i}}_1 \Rightarrow (c^1, c^2, c^3) = (c^1, -c^2, -c^3).$$

Supongamos que ahora queremos aplicar una reflexión al vector $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ pero con \mathbf{a} axial y \mathbf{b} polar. Veamos lo que sucede:

$$\begin{aligned} \tilde{\mathbf{c}}' = \tilde{\mathbf{a}}' \times \tilde{\mathbf{b}}' &= \left[(a^2 b^3 - a^3 b^2), (a^3 b^1 - a^1 b^3), (a^1 b^2 - a^2 b^1) \right] \\ &= [((-a^2)b^3 - (-a^3)b^2), ((-a^3)(-b^1) - a^1 b^3), (a^1 b^2 - (-a^2)(-b^1))] \\ &= [-(a^2 b^3 - a^3 b^2), (a^3 b^1 - a^1 b^3), (a^1 b^2 - a^2 b^1)]. \end{aligned}$$

¹⁵Una discusión de vectores y pseudovectores puede consultarse en la nota de Quigley, Robert J. "Pseudovectors and Reflections" American Journal of Physics 41, no. 3 (1973): 428-430; y en el capítulo 52 del Vol. 1 de Feynman, R.P., Leighton, R.B. and Sands, M., 2013. The Feynman Lectures on Physics, Desktop Edition. Basic Books. Para algunas de las consecuencias que se presentan cuando se consideran cantidades físicas y transformaciones de reflexión.

Esto es:

$$\hat{\mathbf{i}}_1 \rightarrow -\hat{\mathbf{i}}_1 \Rightarrow (c^1', c^2', c^3') = (-c^1, c^2, c^3),$$

por lo tanto, en este caso \mathbf{c} es un vector polar.



1.4.5 Ejercicios

1. Verifique las siguientes identidades:

- (a). $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}$.
- (b). $(\mathbf{a} \times \mathbf{b}) \times (\mathbf{c} \times \mathbf{d}) = \mathbf{b}[\mathbf{a} \cdot (\mathbf{c} \times \mathbf{d})] - \mathbf{a}[\mathbf{b} \cdot (\mathbf{c} \times \mathbf{d})]$.
- (c). $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) + (\mathbf{b} \times \mathbf{c}) \cdot (\mathbf{a} \times \mathbf{d}) + (\mathbf{c} \times \mathbf{a}) \cdot (\mathbf{b} \times \mathbf{d}) = 0$.
- (d). $\mathbf{a} \cdot (\mathbf{b} \times \mathbf{a}) = 0$.
- (e). $(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{a} \times \mathbf{b}) = a^2 b^2 - (\mathbf{a} \cdot \mathbf{b})^2$.

$$(f). (\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = \begin{vmatrix} \mathbf{a} \cdot \mathbf{c} & \mathbf{a} \cdot \mathbf{d} \\ \mathbf{b} \cdot \mathbf{c} & \mathbf{b} \cdot \mathbf{d} \end{vmatrix}$$

2. Demuestre la siguiente tabla de relaciones:

vector	\cdot	vector	$=$	escalar
vector	\cdot	pseudovector	$=$	pseudoescalar
pseudovector	\cdot	pseudovector	$=$	escalar
vector	\times	vector	$=$	pseudovector
vector	\times	pseudovector	$=$	vector
pseudovector	\times	pseudovector	$=$	pseudovector

3. Considere lo expuesto en la sección 1.4.3 y demuestre que:

$$A_k^{i'} A_{i'}^j = \delta_k^j,$$

y además, como un caso especial, establecer la relación con los cosenos directores que satisfacen:

$$\cos(\alpha)^2 + \cos(\beta)^2 + \cos(\gamma)^2 = 1.$$

4. Considere el radio vector posición $\mathbf{r} = x^i \hat{\mathbf{i}}_i \equiv x\hat{\mathbf{i}} + y\hat{\mathbf{j}}$ en 2 dimensiones. Dado el conjunto de transformaciones que se indican a continuación, demuestre en cuales casos las componentes de \mathbf{r} transforman como verdaderas componentes de vectores.

$$(x, y) \rightarrow (-y, x), \quad (x, y) \rightarrow (x, -y), \quad (x, y) \rightarrow (x - y, x + y), \quad (x, y) \rightarrow (x + y, x - y).$$

Practicando con Maxima

A fines prácticos, las transformaciones de coordenadas del tipo rotaciones o reflexiones es de utilidad representarlas por matrices. Por ejemplo, las rotaciones alrededor del eje z se puede representar como:

```
(%i1) Rz:matrix([cos(theta),sin(theta),0],[-sin(theta),cos(theta),0],[0,0,1]);
(%o1) \left(\begin{array}{ccc} \cos(\theta) & \sin(\theta) & 0 \\ -\sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{array}\right)
```

De manera que un vector \mathbf{a} transformará bajo esta rotación en un nuevo vector \mathbf{a}' .

```
(%i2) a:[a1,a2,a3];
```

(%o2) $[a_1, a_2, a_3]$

(%i3) Rx.a;

$$(\%o3) \begin{pmatrix} a_2 \sin(\theta) + a_1 \cos(\theta) \\ a_2 \cos(\theta) - a_1 \sin(\theta) \\ a_3 \end{pmatrix}$$

Si la rotación se hace al rededor del eje x o y las matrices correspondientes son las siguientes:

(%i4) Rx:matrix([1,0,0],[0,cos(theta),-sin(theta)],[0,sin(theta),cos(theta)]);

$$(\%o4) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & -\sin(\theta) \\ 0 & \sin(\theta) & \cos(\theta) \end{pmatrix}$$

(%i5) Ry:matrix([cos(theta),0,sin(theta)],[0,1,0],[-sin(theta),0,cos(theta)]);

$$(\%o5) \begin{pmatrix} \cos(\theta) & 0 & \sin(\theta) \\ 0 & 1 & 0 \\ -\sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$$

Notemos que el determinante de éstas matrices es:

(%i6) determinant(Rx),trigsimp;

$$(\%o6) \sin^2(\theta) + \cos^2(\theta)$$

(%i7) trigrat(%);

$$(\%o7) 1$$

Si la rotación es en un ángulo de $\theta = 60^\circ$ entonces:

(%i8) sublis([theta= %pi/3], Rx);

$$(\%o8) \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ 0 & \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}$$

Si el eje de rotación se define a través de un vector unitario, digamos, $\hat{\mathbf{u}} = u_x \hat{\mathbf{i}} + u_y \hat{\mathbf{j}} + u_z \hat{\mathbf{k}}$, la matriz de rotación para un ángulo θ es:

(%i9) t:1-cos(theta)\$

(%i10) R:matrix([cos(theta)+ux^2*t, ux*uy*t-uz*sin(theta), ux*uz*t+uy*sin(theta)],

[uy*ux*t+uz*sin(theta),cos(theta)+uy^2*t,uy*uz*t-ux*sin(theta)],

[uz*ux*t*uy*sin(theta),uz*uy*t+ux*sin(theta),cos(theta)+uz^2*t]);

$$(\%o10) \begin{pmatrix} (1 - \cos(\theta)) ux^2 + \cos(\theta) & (1 - \cos(\theta)) ux uy - \sin(\theta) uz & (1 - \cos(\theta)) ux uz + \sin(\theta) uy \\ \sin(\theta) uz + (1 - \cos(\theta)) ux uy & (1 - \cos(\theta)) uy^2 + \cos(\theta) & (1 - \cos(\theta)) uy uz - \sin(\theta) ux \\ (1 - \cos(\theta)) ux uz - \sin(\theta) uy & (1 - \cos(\theta)) uy uz + \sin(\theta) ux & (1 - \cos(\theta)) uz^2 + \cos(\theta) \end{pmatrix}$$

Por ejemplo, si el eje de rotación coincide con el vector $\hat{\mathbf{u}} = \frac{\hat{\mathbf{i}}}{\sqrt{3}} + \frac{\hat{\mathbf{j}}}{\sqrt{3}} + \frac{\hat{\mathbf{k}}}{\sqrt{3}}$ y la rotación es en un ángulo de $\theta = 60^\circ$, entonces:

(%i11) [ux,uy,uz]:1/sqrt(3)*[1,1,1]; theta: %pi/3\$

$$(\%o11) \begin{bmatrix} \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \end{bmatrix}$$

Al evaluar con el comando `ev` podemos ver como queda la matriz R con estos valores definidos

(%i13) Ru:ev(R);

$$(\%o13) \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & \frac{2}{3} \end{pmatrix}$$

De manera que un vector, por ejemplo, $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$ transformará bajo esta rotación de la manera siguiente:

$$(\%o14) \text{ a:[1,2,3];}$$

$$(\%o14) [1, 2, 3]$$

$$(\%o15) \text{ Ru.a;}$$

$$(\%o15) \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$$

Por otro lado, una reflexión en el plano zy es más sencilla de representar porque es la matriz:

$$(\%o16) \text{ Re:matrix([-1,0,0],[0,1,0],[0,0,1]);}$$

$$(\%o16) \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si tenemos los siguientes vectores polares

$$(\%o17) \text{ a:[1,2,3]; b:[4,5,6];}$$

$$(\%o17) [1, 2, 3]$$

$$(\%o18) [4, 5, 6]$$

Bajo la reflexión estos vectores transforman así:

$$(\%o19) \text{ Re.a; Re.b;}$$

$$(\%o19) \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$$

$$(\%o20) \begin{pmatrix} -4 \\ 5 \\ 6 \end{pmatrix}$$

Veamos ahora cómo transforma el vector $\mathbf{c} = \mathbf{a} \times \mathbf{b}$

$$(\%o21) \text{ c:express(a~b);}$$

$$(\%o21) [-3, 6, -3]$$

$$(\%o22) \text{ Re.c;}$$

$$(\%o22) \begin{pmatrix} 3 \\ 6 \\ -3 \end{pmatrix}$$

Es decir, \mathbf{c} es un pseudovector.

1.5 Un comienzo a la derivación e integración de vectores

En esta penúltima sección de este capítulo abordaremos una introducción somera a las consecuencias que imponen la variación de un vector. Mostrarímos, de manera funcional y como una ejercitación a los conceptos de derivación e integración de vectores y campos vectoriales¹⁶.

1.5.1 Vectores variables

Los vectores podrán ser constantes o variables. Ahora bien, esta característica se verificará tanto en las componentes como en la base. Esto quiere decir que cuando un vector cambia podrá variar su módulo, su dirección, su sentido, o todo junto o por separado. Obviamente esta variabilidad del vector dependerá de la base en la cual se exprese. Por lo cual un vector podrá tener una componente constante en una base y variable en otra, vale decir:

$$\mathbf{a}(t) = a^i(t)\mathbf{e}_i(t) = \tilde{a}^k\tilde{\mathbf{e}}_k(t) = \bar{a}^m(t)\bar{\mathbf{e}}_m.$$

Nótese que hemos utilizado una base $\{\mathbf{e}_i(t)\}$ de vectores variables a diferencia de la tradicional base de vectores cartesianos $\{\hat{\mathbf{i}}_i\}$, los cuales **son constantes** en módulo, dirección y sentido (ver los cuadrantes I y II de la figura 1.9). Más aún, tal y como se muestra en cuadrante II de la figura 1.9, todo vector variable podrá ser expresado como la suma de uno variable, $\alpha(t)$, más otro constante \mathbf{c} .

$$\mathbf{a}(t) = \alpha(t) + \mathbf{c}.$$

1.5.2 Derivación

De esta manera, cuando uno piensa en un vector variable $\mathbf{a}(t)$ uno rápidamente intenta establecer un cociente incremental:

$$\lim_{\Delta t \rightarrow 0} \frac{\mathbf{a}(t + \Delta t) - \mathbf{a}(t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\Delta \mathbf{a}(t)}{\Delta t} = \frac{d\mathbf{a}(t)}{dt}.$$

El cuadrante IV de la figura 1.9 ilustra gráficamente este cociente incremental.

Como siempre, las propiedades de esta operación derivación serán:

$$\begin{aligned} \frac{d}{dt} [\mathbf{a}(t) + \mathbf{b}(t)] &= \frac{d}{dt}\mathbf{a}(t) + \frac{d}{dt}\mathbf{b}(t), \\ \frac{d}{dt} [\lambda(t)\mathbf{a}(t)] &= \left[\frac{d}{dt}\lambda(t) \right] \mathbf{a}(t) + \lambda(t) \left[\frac{d}{dt}\mathbf{a}(t) \right], \\ \frac{d}{dt} [\mathbf{a}(t) \cdot \mathbf{b}(t)] &= \left[\frac{d}{dt}\mathbf{a}(t) \right] \cdot \mathbf{b}(t) + \mathbf{a}(t) \cdot \left[\frac{d}{dt}\mathbf{b}(t) \right], \\ \frac{d}{dt} [\mathbf{a}(t) \times \mathbf{b}(t)] &= \left[\frac{d}{dt}\mathbf{a}(t) \right] \times \mathbf{b}(t) + \mathbf{a}(t) \times \left[\frac{d}{dt}\mathbf{b}(t) \right]. \end{aligned}$$

Ahora bien, esto implica que:

$$\mathbf{a}(t) = a^i(t)\mathbf{e}_i(t) \Rightarrow \frac{d\mathbf{a}(t)}{dt} = \frac{d[a^i(t)\mathbf{e}_i(t)]}{dt} = \frac{da^i(t)}{dt}\mathbf{e}_i(t) + a^i(t)\frac{d\mathbf{e}_i(t)}{dt}, \quad (1.22)$$

con lo cual hay que tener cuidado al derivar vectores y cerciorarse de la dependencia funcional de la base y componentes.

Habrá sistemas de coordenadas (bases de vectores) que serán constantes y otros en los cuales sus vectores bases cambiarán en su dirección. El primer término de (1.22) representa la variación del módulo, y el segundo muestra la contribución de los cambios en dirección del vector. Más aún, mostraremos apoyándonos en la

¹⁶Un análisis más detallado lo presentaremos en el último capítulo ??

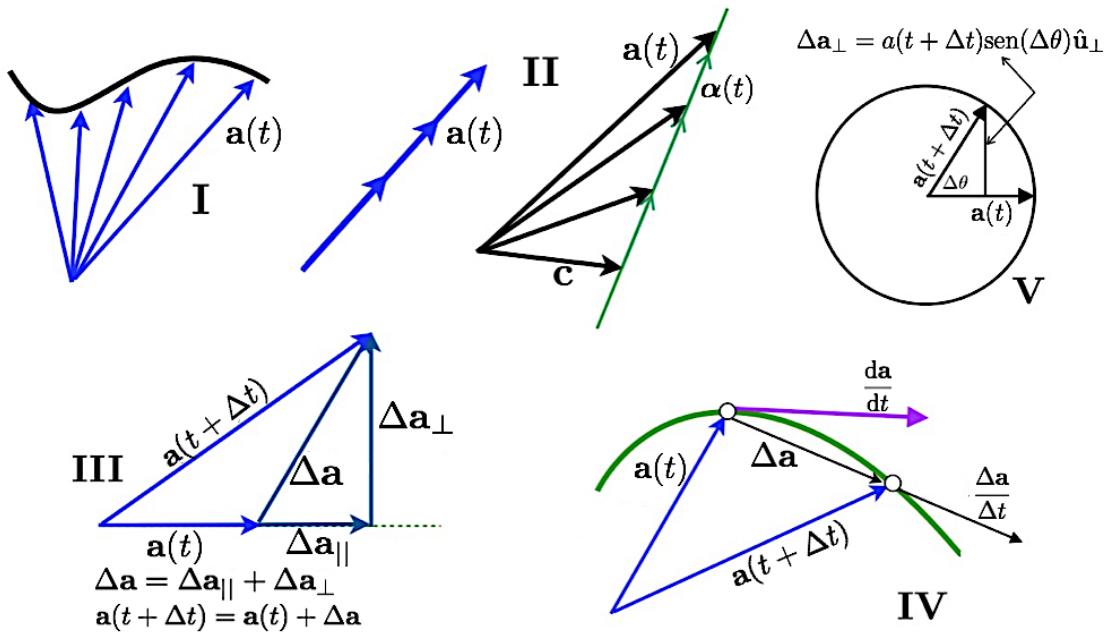


Figura 1.9: Diferentes representaciones para vectores que varían respecto a algún parámetro.

ilustración de el cuadrante III de la figura 1.9 que, independientemente del sistema de coordenadas, el cambio en el módulo apunta en la dirección del vector, mientras que las contribuciones en dirección apuntan en la dirección perpendicular al vector. Esto es:

$$\frac{d\mathbf{a}(t)}{dt} = \frac{d|\mathbf{a}(t)|}{dt}\hat{\mathbf{u}}_{||} + |\mathbf{a}(t)|\hat{\mathbf{u}}_{\perp}, \quad \text{con } \hat{\mathbf{u}}_{||} \cdot \hat{\mathbf{u}}_{\perp} = 0.$$

Es fácil convencernos de la forma del primer término. Siempre podemos representar un vector como su módulo y un vector unitario en la dirección apropiada. Esto es, si:

$$\mathbf{a}(t) = |\mathbf{a}(t)|\hat{\mathbf{u}}(t), \quad (1.23)$$

$$\frac{d\mathbf{a}(t)}{dt} = \frac{d[|\mathbf{a}(t)|\hat{\mathbf{u}}(t)]}{dt} = \frac{d|\mathbf{a}(t)|}{dt}\hat{\mathbf{u}}(t) + |\mathbf{a}(t)|\frac{d\hat{\mathbf{u}}(t)}{dt}. \quad (1.24)$$

Como podemos ver a continuación cuando multiplicamos 1.24 por $\hat{\mathbf{u}}(t)$ resulta:

$$\hat{\mathbf{u}}(t) \cdot \frac{d\mathbf{a}(t)}{dt} = \hat{\mathbf{u}}(t) \cdot \left[\frac{d|\mathbf{a}(t)|}{dt}\hat{\mathbf{u}}(t) + |\mathbf{a}(t)|\frac{d\hat{\mathbf{u}}(t)}{dt} \right] \Rightarrow \begin{cases} \hat{\mathbf{u}}(t) \cdot \frac{d\mathbf{a}(t)}{dt} = \frac{d|\mathbf{a}(t)|}{dt} \\ \hat{\mathbf{u}}(t) \cdot \frac{d\hat{\mathbf{u}}(t)}{dt} = 0 \end{cases}$$

Es decir que el cambio en el módulo de un vector se manifiesta en la dirección del mismo vector, tal y como era intuitivo suponer. Adicionalmente, vemos que el vector $\hat{\mathbf{u}}$ siempre será perpendicular a su derivada. Gráficamente podemos apreciarlo en el cuadrante IV de la figura 1.9, pero también surge analíticamente si derivamos el vector unitario en la dirección de $\mathbf{a}(t)$:

$$\frac{d[\hat{\mathbf{u}}(t) \cdot \hat{\mathbf{u}}(t)]}{dt} \equiv \frac{d(|\hat{\mathbf{u}}(t)|^2)}{dt} = \frac{d(1)}{dt} \equiv 0 = \hat{\mathbf{u}}(t) \cdot \frac{d\hat{\mathbf{u}}(t)}{dt} \Rightarrow \hat{\mathbf{u}}(t) \perp \frac{d\hat{\mathbf{u}}(t)}{dt},$$

es decir:

$$\frac{d\mathbf{a}(t)}{dt} = \frac{d[|\mathbf{a}(t)|\hat{\mathbf{u}}(t)]}{dt} = \frac{d|\mathbf{a}(t)|}{dt}\hat{\mathbf{u}}(t) + |\mathbf{a}(t)|\frac{d\hat{\mathbf{u}}(t)}{dt} = \frac{d|\mathbf{a}(t)|}{dt}\hat{\mathbf{u}}_{||} + |\mathbf{a}(t)|\hat{\mathbf{u}}_{\perp}.$$

Supongamos que ahora definimos un vector $\Delta\theta = \Delta\theta \hat{v}$, con:

$$\left. \begin{array}{l} \hat{v} \perp \hat{u}_{||} \\ \hat{v} \perp \hat{u}_{\perp} \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \hat{v} \times \hat{u}_{||} = \hat{u}_{\perp} \\ \hat{u}_{\perp} \times \hat{v} = \hat{u}_{||} \\ \hat{u}_{||} \times \hat{u}_{\perp} = \hat{v} \end{array} \right.$$

donde $\Delta\theta$ es el ángulo de rotación del vector $\mathbf{a}(t)$ (ver cuadrante V de la figura 1.9). Claramente:

$$\Delta\mathbf{a}_{\perp} = [a(t + \Delta t) \sin(\Delta\theta)] \hat{u}_{\perp} \approx [a(t + \Delta t) \Delta\theta] \hat{u}_{\perp} \Rightarrow \Delta\mathbf{a}_{\perp} = \Delta\theta \times \mathbf{a}(t),$$

entonces:

$$\frac{\Delta\mathbf{a}_{\perp}}{\Delta t} \equiv \left[\frac{\Delta\mathbf{a}}{\Delta t} \cdot \mathbf{a}_{\perp} \right] \mathbf{a}_{\perp} = \frac{\Delta\theta}{\Delta t} \times \mathbf{a}(t) \Rightarrow \left[\frac{d\mathbf{a}(t)}{dt} \cdot \hat{u}_{\perp} \right] \hat{u}_{\perp} = \frac{d\theta(t)}{dt} \hat{v} \times \mathbf{a}(t) = \boldsymbol{\omega} \times \mathbf{a}(t),$$

donde hemos identificado $\boldsymbol{\omega} = \frac{d\theta(t)}{dt} \hat{v}$.

Podemos ir más allá observando el cuadrante V de la figura 1.9, vemos que si suponemos que el módulo del vector \mathbf{a} es constante, entonces:

$$\frac{d|\mathbf{a}(t)|}{dt} = 0 \Rightarrow \frac{d\mathbf{a}(t)}{dt} = |\mathbf{a}(t)| \hat{u}_{\perp} \Rightarrow \left[\frac{d\mathbf{a}(t)}{dt} \cdot \hat{u}_{\perp} \right] \hat{u}_{\perp} = \boldsymbol{\omega} \times \mathbf{a}(t).$$

1.5.3 Velocidades y aceleraciones

El radio vector posición de una partícula genera, como sabemos, los vectores velocidad y aceleración:

$$\mathbf{r} = \mathbf{r}(t) \Rightarrow \mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} \Rightarrow \mathbf{a}(t) = \frac{d\mathbf{v}(t)}{dt} = \frac{d^2\mathbf{r}(t)}{dt^2},$$

ahora bien:

$$\mathbf{r} = r\hat{u}_r = x\hat{i} + y\hat{j} + z\hat{k}, \quad \text{con: } \hat{u}_r = \cos(\theta)\hat{i} + \sin(\theta)\hat{j} + z\hat{k}.$$

Si suponemos que la partícula describe una trayectoria, entonces:

$$\mathbf{r}(t) = r(t)\hat{u}_r(t) = x(t)\hat{i} + y(t)\hat{j} + z(t)\hat{k}, \quad \text{con: } \hat{u}_r(t) = \cos(\theta(t))\hat{i} + \sin(\theta(t))\hat{j} + z(t)\hat{k}.$$

Es muy común denotar a la derivada temporal sobre funciones de una variable con un punto, es decir, podemos utilizar la siguiente notación:

$$\dot{f}(t) \equiv \frac{df(t)}{dt},$$

con lo cual, y en el caso de que $z = z(t) = \text{constante}$, se tiene:

$$\frac{d\hat{u}_r}{dt} = \frac{d[\cos(\theta(t))\hat{i} + \sin(\theta(t))\hat{j}]}{dt} = \dot{\theta}(t) \underbrace{[-\sin(\theta(t))\hat{i} + \cos(\theta(t))\hat{j}]}_{\hat{u}_{\theta}} = \dot{\theta}(t)\hat{u}_{\theta}.$$

Podemos ver que:

$$|\hat{u}_r| = \sqrt{\hat{u}_r \cdot \hat{u}_r} = \sqrt{[\cos(\theta(t))\hat{i} + \sin(\theta(t))\hat{j}] \cdot [\cos(\theta(t))\hat{i} + \sin(\theta(t))\hat{j}]} = 1,$$

$$|\hat{u}_{\theta}| = \sqrt{\hat{u}_{\theta} \cdot \hat{u}_{\theta}} = \sqrt{[-\sin(\theta(t))\hat{i} + \cos(\theta(t))\hat{j}] \cdot [-\sin(\theta(t))\hat{i} + \cos(\theta(t))\hat{j}]} = 1,$$

además:

$$\hat{u}_{\theta} \cdot \hat{u}_r = \hat{u}_r \cdot \hat{u}_{\theta} = [-\sin(\theta(t))\hat{i} + \cos(\theta(t))\hat{j}] \cdot [\cos(\theta(t))\hat{i} + \sin(\theta(t))\hat{j}] = 0.$$

También podemos calcular:

$$\frac{d\hat{u}_{\theta}}{dt} = \frac{d[-\sin(\theta(t))\hat{i} + \cos(\theta(t))\hat{j}]}{dt} = -\dot{\theta}(t) [\cos(\theta(t))\hat{i} + \sin(\theta(t))\hat{j}] = -\dot{\theta}(t)\hat{u}_r.$$

Para una partícula que sigue un movimiento arbitrario, su trayectoria vendrá descrita, en coordenadas

cartesianas, por:

$$\mathbf{r} = x(t) \hat{\mathbf{i}} + y(t) \hat{\mathbf{j}} + z(t) \hat{\mathbf{k}}.$$

- Su velocidad será:

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = \frac{d[x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}]}{dt} = \dot{x}(t)\hat{\mathbf{i}} + \dot{y}(t)\hat{\mathbf{j}} + \dot{z}(t)\hat{\mathbf{k}} = v_x(t)\hat{\mathbf{i}} + v_y(t)\hat{\mathbf{j}} + v_z(t)\hat{\mathbf{k}}.$$

- Y su aceleración:

$$\mathbf{a}(t) = \ddot{x}(t)\hat{\mathbf{i}} + \ddot{y}(t)\hat{\mathbf{j}} + \ddot{z}(t)\hat{\mathbf{k}} = a_x(t)\hat{\mathbf{i}} + a_y(t)\hat{\mathbf{j}} + a_z(t)\hat{\mathbf{k}}.$$

Mientras que en coordenadas polares las ecuaciones son:

$$\mathbf{r}(t) = r(t)\hat{\mathbf{u}}_r(t).$$

- Velocidad:

$$\mathbf{v}(t) = \frac{d[r(t)\hat{\mathbf{u}}_r(t)]}{dt} = \dot{r}(t)\hat{\mathbf{u}}_r(t) + r(t)\frac{d\hat{\mathbf{u}}_r(t)}{dt} = \dot{r}(t)\hat{\mathbf{u}}_r(t) + r(t)\dot{\theta}(t)\hat{\mathbf{u}}_\theta(t),$$

- Aceleración:

$$\begin{aligned} \mathbf{a}(t) &= \frac{d\mathbf{v}(t)}{dt} = \frac{d[\dot{r}(t)\hat{\mathbf{u}}_r(t) + r(t)\dot{\theta}(t)\hat{\mathbf{u}}_\theta(t)]}{dt} = \frac{d[\dot{r}(t)\hat{\mathbf{u}}_r(t)]}{dt} + \frac{d[r(t)\dot{\theta}(t)\hat{\mathbf{u}}_\theta(t)]}{dt} \\ &= \ddot{r}(t)\hat{\mathbf{u}}_r(t) + \dot{r}(t)\frac{d\hat{\mathbf{u}}_r(t)}{dt} + \dot{r}(t)\dot{\theta}(t)\hat{\mathbf{u}}_\theta(t) + r(t)\ddot{\theta}(t)\hat{\mathbf{u}}_\theta(t) + r(t)\dot{\theta}(t)\frac{d\hat{\mathbf{u}}_\theta(t)}{dt} \\ &= \left[\ddot{r}(t) - r(t) (\dot{\theta}(t))^2 \right] \hat{\mathbf{u}}_r(t) + \left[2\dot{r}(t)\dot{\theta}(t) + r(t)\ddot{\theta}(t) \right] \hat{\mathbf{u}}_\theta(t). \end{aligned}$$

Claramente para el caso de un movimiento circular, donde $r = R = \text{const}$, resulta:

$$\frac{dR}{dt} = 0 \quad \Rightarrow \quad \begin{cases} \mathbf{r}(t) = R \hat{\mathbf{u}}_r(t) \\ \mathbf{v}(t) = R \dot{\theta}(t) \hat{\mathbf{u}}_\theta \\ \mathbf{a}(t) = -R \dot{\theta}(t)^2 \hat{\mathbf{u}}_r(t) + R \ddot{\theta}(t) \hat{\mathbf{u}}_\theta(t) \end{cases}$$

De aquí podemos ver que el vector velocidad $\mathbf{v}(t)$ y el vector posición $\mathbf{r}(t)$ son ortogonales. La velocidad, $\mathbf{v}(t)$, siempre es tangente a la trayectoria $\mathbf{r}(t)$ y en este caso la trayectoria es una circunferencia.

En general el vector:

$$\mathbf{r}_{med} = \sum_i \Delta \mathbf{r}(t_i) = \sum_i (\mathbf{r}(t_i + \Delta t_i) - \mathbf{r}(t_i)) \Rightarrow \lim_{\Delta t \rightarrow 0} \sum_i \Delta \mathbf{r}(t_i) = \int d\mathbf{r}(t) = \mathbf{r}(t),$$

es decir, $d\mathbf{r}(t) = \lim_{\Delta t \rightarrow 0} \sum_i \Delta \mathbf{r}(t_i)$ es tangente a la trayectoria. Es claro que:

$$d\mathbf{r}(t) = d[x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}] \equiv \frac{dx(t)}{dt}\hat{\mathbf{i}} + \frac{dy(t)}{dt}\hat{\mathbf{j}} + \frac{dz(t)}{dt}\hat{\mathbf{k}}.$$

Tal y como mencionamos arriba, para el sistema de coordenadas cartesiano podemos definir un vector velocidad angular $\boldsymbol{\omega}$ tal que:

$$\left. \begin{array}{l} \frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|} \times \hat{\mathbf{u}}_r = \hat{\mathbf{u}}_v \\ \hat{\mathbf{u}}_v \times \frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|} = \hat{\mathbf{u}}_r \\ \hat{\mathbf{u}}_r \times \hat{\mathbf{u}}_v = \frac{\boldsymbol{\omega}}{|\boldsymbol{\omega}|} \end{array} \right\} \Rightarrow \mathbf{v}(t) = \boldsymbol{\omega} \times \mathbf{r}(t).$$

Supongamos por simplicidad que elegimos el sistema de coordenadas cartesiano, donde \mathbf{r} está en el plano xy . En este caso es inmediato comprobar que $v^i = \epsilon^{ijk} \omega_j x_k$, y dado que \mathbf{r} y \mathbf{v} tienen únicamente componentes

1 y 2 entonces, necesariamente ω tiene únicamente componente 3, Es decir:

$$\left. \begin{array}{l} \mathbf{r} = r^i \mathbf{e}_i \\ \mathbf{v} = v^i \mathbf{e}_i \end{array} \right\} \Rightarrow \left. \begin{array}{l} v^1 = \varepsilon^{1j2} \omega_j x_2 \\ v^2 = \varepsilon^{2j1} \omega_j x_1 \end{array} \right\} \Rightarrow \omega = |\omega| \mathbf{e}_3 = \omega \hat{\mathbf{k}},$$

como $\mathbf{r} = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}}$, entonces:

$$\mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = v_x(t)\hat{\mathbf{i}} + v_y(t)\hat{\mathbf{j}} = \omega \times \mathbf{r}(t) = \dot{\theta}(t)\hat{\mathbf{k}} \times [x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}}].$$

Se verá más claro en coordenadas polares, esto es:

$$\begin{aligned} \mathbf{v}(t) &= \frac{d\mathbf{r}(t)}{dt} = r(t)\dot{\theta}(t)\hat{\mathbf{u}}_\theta(t) = [|\omega| \hat{\mathbf{u}}_n(t)] \times [r(t)\hat{\mathbf{u}}_r(t)], \quad |\mathbf{r}(t)| = \text{const} \\ &= \underbrace{r(t)\dot{\theta}(t)}_{\mathbf{v}_\perp} \hat{\mathbf{u}}_\theta(t) = |\omega| r(t)\hat{\mathbf{u}}_\theta(t) \Rightarrow \dot{\theta}(t) \equiv |\omega|. \end{aligned}$$

Ejemplo 1.16 Si una partícula se mueve a lo largo de una curva descrita por:

$$x(t) = 3t^2, \quad y(t) = 4t^3 - t, \quad z(t) = t.$$

Calcular las expresiones para los vectores: posición, velocidad y aceleración de esa partícula.

Podemos ver que:

$$\mathbf{r}(t) = 3t^2\hat{\mathbf{i}} + (4t^3 - t)\hat{\mathbf{j}} + t\hat{\mathbf{k}}, \quad \mathbf{v} = 6t\hat{\mathbf{i}} + (12t^2 - 1)\hat{\mathbf{j}} + \hat{\mathbf{k}}, \quad \mathbf{a} = 6\hat{\mathbf{i}} + 24t\hat{\mathbf{j}}.$$

Si nos proponemos encontrar las expresiones, más generales, de los vectores tangentes y perpendiculares a todo punto de la trayectoria de la partícula podemos ver que el vector tangente a todo punto de la trayectoria es el vector velocidad

$$\mathbf{v} = 6t\hat{\mathbf{i}} + (12t^2 - 1)\hat{\mathbf{j}} + \hat{\mathbf{k}},$$

El vector perpendicular a todo punto, será un vector $\mathbf{b} = b_x\hat{\mathbf{i}} + b_y\hat{\mathbf{j}} + b_z\hat{\mathbf{k}}$, tal que:

$$(6t\hat{\mathbf{i}} + (12t^2 - 1)\hat{\mathbf{j}} + \hat{\mathbf{k}}) \cdot (b_x\hat{\mathbf{i}} + b_y\hat{\mathbf{j}} + b_z\hat{\mathbf{k}}) = 6tb_x + (12t^2 - 1)b_y + b_z = 0,$$

con lo cual:

$$\mathbf{b} = b_x\hat{\mathbf{i}} + b_y\hat{\mathbf{j}} - (6tb_x + (12t^2 - 1)b_y)\hat{\mathbf{k}}.$$



Ejemplo 1.17 La trayectoria de un punto en el plano vista por un observador (1) es la siguiente:

$$\mathbf{r}_1(t) = 5 \cos(3t^2) \hat{\mathbf{i}} + 5 \sin(3t^2) \hat{\mathbf{j}}.$$

Calculemos las expresiones para las radiales y tangenciales de esta partícula.

Para ésta partícula ya que es claro que la partícula describe un movimiento circular, donde $\theta(t) = 3t^2$.

Entonces:

$$\mathbf{r}(t) = 5\hat{\mathbf{u}}_r \Rightarrow \mathbf{v}(t) = \frac{d\mathbf{r}(t)}{dt} = 5 \frac{d\theta(t)}{dt} \hat{\mathbf{u}}_\theta = 30t \hat{\mathbf{u}}_\theta \Rightarrow \mathbf{a}(t) = \frac{d\mathbf{a}(t)}{dt} = 30 \hat{\mathbf{u}}_\theta - 30t \hat{\mathbf{u}}_r.$$

Consideremos ahora un segundo observador (2), el cual describe una trayectoria respecto al primero representada por:

$$\mathbf{r}_{21}(t) = (t^3 - 4t)\hat{\mathbf{i}} + (t^2 + 4t)\hat{\mathbf{j}}.$$

Y queremos encontrar las expresiones para los vectores posición, velocidad y aceleración de la partícula medidos respecto al segundo observador.

Por lo tanto, la trayectoria de la partícula respecto al segundo observador será:

$$\mathbf{r}_2(t) = \mathbf{r}_1(t) - \mathbf{r}_{21}(t) = 5 \cos(3t^2) \hat{\mathbf{i}} + 5 \sin(3t^2) \hat{\mathbf{j}} - ((t^3 - 4t)\hat{\mathbf{i}} + (t^2 + 4t)\hat{\mathbf{j}}),$$

con lo cual:

$$\mathbf{r}_2(t) = [5 \cos(3t^2) - t^3 + 4t] \hat{\mathbf{i}} + [5 \sin(3t^2) - t^2 - 4t] \hat{\mathbf{j}},$$

entonces:

$$\mathbf{v}_2(t) = \frac{d\mathbf{r}_2(t)}{dt} = -[30 \sin(3t^2)t + 3t^2 - 4] \hat{\mathbf{i}} + [30 \cos(3t^2)t - 2t - 4] \hat{\mathbf{j}},$$

y

$$\mathbf{a}_2(t) = \frac{d\mathbf{v}_2(t)}{dt} = -6[30 \cos(3t^2)t^2 + 5 \sin(3t^2)t + 1] \hat{\mathbf{i}} - 2[90 \sin(3t^2)t^2 - 15 \cos(3t^2)t + 1] \hat{\mathbf{j}}.$$



1.5.4 Vectores y funciones

Antes de continuar con la integración repensemosemos algunas funciones de tipo $\phi(x^i)$ y $\mathbf{A}(x^i)$. Estas funciones son sin duda funciones de varias variables, en el caso cartesiano:

$$\phi = \phi(x, y, z), \quad \mathbf{A} = \mathbf{A}(x, y, z) = \hat{\mathbf{i}}A_x(x, y, z) + \hat{\mathbf{j}}A_y(x, y, z) + \hat{\mathbf{k}}A_z(x, y, z).$$

Un par de reflexiones se pueden hacer en este punto, primeramente, dado que hemos relacionado un punto del espacio con el radio vector posición, entonces:

$$P_{(x,y,z)} \leftrightarrow (x, y, z) \leftrightarrow \mathbf{r} = x \hat{\mathbf{i}} + y \hat{\mathbf{j}} + z \hat{\mathbf{k}} \Rightarrow \begin{cases} \phi = \phi(x, y, z) \equiv \phi(\mathbf{r}) \\ \mathbf{A} = \mathbf{A}(x, y, z) \equiv \mathbf{A}(\mathbf{r}) \end{cases}$$

La primera función, $\phi(\mathbf{r})$, será una función escalar de argumento vectorial o, simplemente un campo escalar y la segunda, $\mathbf{A}(\mathbf{r})$, se conoce como una función vectorial de argumento vectorial o campo vectorial. Como hemos dicho, este tipo de funciones y las operaciones que pueden ser realizadas con ellas, y su significado, serán analizadas en detalle más adelante durante el desarrollo de este curso.

En segundo lugar, siempre podremos parametrizar las coordenadas y tendremos

$$\phi = \phi(t) = \phi(x(t), y(t), z(t)),$$

$$\mathbf{A} = \mathbf{A}(t) = \mathbf{A}(x(t), y(t), z(t)) = A_x(x(t), y(t), z(t)) \hat{\mathbf{i}} + A_y(x(t), y(t), z(t)) \hat{\mathbf{j}} + A_z(x(t), y(t), z(t)) \hat{\mathbf{k}}.$$

Este caso lo hemos encontrado en montones de situaciones, por ejemplo, el movimiento parabólico viene descrito por vectores velocidad y posición dados por:

$$\mathbf{v}(t) = -gt \hat{\mathbf{k}} + \mathbf{v}_0 = -gt \hat{\mathbf{k}} + (v_{0x} \hat{\mathbf{i}} + v_{0y} \hat{\mathbf{j}} + v_{0z} \hat{\mathbf{k}}) \Rightarrow \begin{cases} v_x = v_{0x} \\ v_y = v_{0y} \\ v_z = v_{0z} - gt \end{cases}$$

$$\mathbf{r}(t) = -\frac{g}{2}t^2 \hat{\mathbf{k}} + \mathbf{v}_0 t = -\frac{g}{2}t^2 \hat{\mathbf{k}} + (v_{0x} \hat{\mathbf{i}} + v_{0y} \hat{\mathbf{j}} + v_{0z} \hat{\mathbf{k}}) t \Rightarrow \begin{cases} x = v_{0x}t \\ y = v_{0y}t \\ z = v_{0z}t - \frac{g}{2}t^2 \end{cases}$$

1.5.4.1 Derivada de funciones del tipo: $\phi(\mathbf{r}(t))$

Al derivar una función de argumento vectorial también se aplica la “regla de la cadena”. Esto es, si

$$\phi(\mathbf{r}(t)) = \phi(x(t), y(t), z(t)),$$

entonces:

$$\begin{aligned}\frac{d\phi(\mathbf{r}(t))}{dt} &= \frac{\partial\phi(x(t), y(t), z(t))}{\partial x} \frac{dx(t)}{dt} + \frac{\partial\phi(x(t), y(t), z(t))}{\partial y} \frac{dy(t)}{dt} + \frac{\partial\phi(x(t), y(t), z(t))}{\partial z} \frac{dz(t)}{dt} \\ &= \left[\frac{\partial\phi(x, y, z)}{\partial x} \hat{\mathbf{i}} + \frac{\partial\phi(x, y, z)}{\partial y} \hat{\mathbf{j}} + \frac{\partial\phi(x, y, z)}{\partial z} \hat{\mathbf{k}} \right] \cdot \left[\frac{dx(t)}{dt} \hat{\mathbf{i}} + \frac{dy(t)}{dt} \hat{\mathbf{j}} + \frac{dz(t)}{dt} \hat{\mathbf{k}} \right] \\ &= \nabla\phi(x(t), y(t), z(t)) \cdot \frac{d\mathbf{r}(t)}{dt},\end{aligned}$$

donde hemos representado:

$$\nabla\phi(\mathbf{r}(t)) \equiv \frac{\partial\phi(x, y, z)}{\partial x} \hat{\mathbf{i}} + \frac{\partial\phi(x, y, z)}{\partial y} \hat{\mathbf{j}} + \frac{\partial\phi(x, y, z)}{\partial z} \hat{\mathbf{k}} = \partial^i\phi(x^j)\hat{\mathbf{i}}_i = \phi^i(x^j)\hat{\mathbf{i}}_i,$$

y lo llamaremos el **gradiente** de la función $\phi(\mathbf{r}(t))$.

El gradiente de un campo escalar es uno de los objetos más útiles que encontraremos en el estudio de problemas de Física-Matemática, el cual lo utilizaremos por ahora de manera operacional. Es bueno recordar que emerge como consecuencia de una derivación contra un parámetro. El gradiente mide el cambio de la función $\phi(x, y, z)$.

La idea de gradiente nos lleva a considerar a ∇ como un operador vectorial (el operador nabla) que actúa sobre la función escalar de variable vectorial $\phi(\mathbf{r}(t))$.

$$\nabla\phi(\mathbf{r}(t)) \equiv \left(\frac{\partial}{\partial x} \hat{\mathbf{i}} + \frac{\partial}{\partial y} \hat{\mathbf{j}} + \frac{\partial}{\partial z} \hat{\mathbf{k}} \right) \phi(x, y, z) = \hat{\mathbf{i}}_i \partial^i\phi(x, y, z).$$

Es decir, y con un poquito de imaginación:

$$\nabla(\circ) = \left(\frac{\partial}{\partial x} \hat{\mathbf{i}} + \frac{\partial}{\partial y} \hat{\mathbf{j}} + \frac{\partial}{\partial z} \hat{\mathbf{k}} \right) (\circ) = \hat{\mathbf{i}}_i \partial^i(\circ).$$

1.5.4.2 Derivada de funciones del tipo: $\mathbf{A}(\mathbf{r}(t))$

De modo que inspirados en la regla de la cadena de una función escalar de variable vectorial podemos comprobar que:

$$\frac{d\mathbf{A}}{dt} = \frac{dA_x(x, y, z)}{dt} \hat{\mathbf{i}} + \frac{dA_y(x, y, z)}{dt} \hat{\mathbf{j}} + \frac{dA_z(x, y, z)}{dt} \hat{\mathbf{k}} = \frac{dA^i(x^j)}{dt} \hat{\mathbf{i}}_i,$$

por consiguiente, si \mathbf{A} , tiene por componentes cartesianas (A_x, A_y, A_z) las componentes del vector derivado serán: $\left(\frac{dA_x}{dt}, \frac{dA_y}{dt}, \frac{dA_z}{dt} \right)$. Con lo cual, para cada componente:

$$\frac{d(A^i(x(t), y(t), z(t)))}{dt} = \frac{d(A^i(x^j(t)))}{dt} = \frac{\partial(A^i(x^j))}{\partial x^k} \frac{dx^k(t)}{dt} = \left(\frac{d\mathbf{r}(t)}{dt} \cdot \nabla \right) A^i(x, y, z).$$

En términos vectoriales es:

$$\frac{d\mathbf{A}}{dt} = \left(\frac{d\mathbf{r}(t)}{dt} \cdot \nabla \right) \mathbf{A} \equiv (\mathbf{v} \cdot \nabla) \mathbf{A} \Rightarrow \frac{d(\circ)}{dt} = (\mathbf{v} \cdot \nabla)(\circ) \equiv v^i \partial_i(\circ),$$

con \mathbf{v} la derivada del radio vector posición $\mathbf{r}(t)$, es decir, la velocidad. Entonces, estamos viendo que el cambio del vector \mathbf{A} respecto al tiempo es el cambio de sus componentes en la dirección de la velocidad.

Si se nos ocurre calcular la derivada del vector velocidad para encontrar la aceleración tendremos que nos quedará expresada como:

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = (\mathbf{v} \cdot \nabla) \mathbf{v} \Rightarrow a^i = (\mathbf{v} \cdot \nabla) v^i,$$

donde las componentes cartesianas de los vectores velocidad y aceleración son: $v^i = v^i(x(t), y(t), z(t))$ y

$a^i = a^i(x(t), y(t), z(t))$, respectivamente.

1.5.5 El operador ∇

El operador vectorial ∇ (\circ) merece un poco de atención en este nivel. Tal y como hemos visto cuando construimos:

El Gradiente:

$$\begin{aligned}\nabla \phi(x, y, z) &= \frac{\partial \phi(x, y, z)}{\partial x} \hat{\mathbf{i}} + \frac{\partial \phi(x, y, z)}{\partial y} \hat{\mathbf{j}} + \frac{\partial \phi(x, y, z)}{\partial z} \hat{\mathbf{k}} \\ &= \partial^1 \phi(x, y, z) \hat{\mathbf{i}}_1 + \partial^2 \phi(x, y, z) \hat{\mathbf{i}}_2 + \partial^3 \phi(x, y, z) \hat{\mathbf{i}}_3 = \partial^i \phi(x^j) \hat{\mathbf{i}}_i.\end{aligned}\quad (1.25)$$

Podemos ver ahora que existen otras posibilidades:

El Rotacional: Se puede construir la siguiente operación: $\nabla \times \mathbf{A}$, que denominaremos **rotacional** de \mathbf{A} , y vendrá dado por la siguiente expresión:

$$\begin{aligned}\nabla \times \mathbf{A} &= \left(\frac{\partial}{\partial x} \hat{\mathbf{i}} + \frac{\partial}{\partial y} \hat{\mathbf{j}} + \frac{\partial}{\partial z} \hat{\mathbf{k}} \right) \times \left(A_x \hat{\mathbf{i}} + A_y \hat{\mathbf{j}} + A_z \hat{\mathbf{k}} \right) \\ &= \left(\frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \right) \hat{\mathbf{i}} + \left(\frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \right) \hat{\mathbf{j}} + \left(\frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \right) \hat{\mathbf{k}} = \varepsilon^{ijk} \partial_j A_k \hat{\mathbf{i}}_i.\end{aligned}\quad (1.26)$$

La Divergencia: También podemos hablar del “producto escalar” de ∇ por un vector \mathbf{A} . A esta operación la llamaremos **divergencia** de \mathbf{A} :

$$\nabla \cdot \mathbf{A} = \frac{\partial A^i(x^j)}{\partial x^i} \equiv \partial_i A^i(x^j) \equiv \frac{\partial A_x(x, y, z)}{\partial x} + \frac{\partial A_y(x, y, z)}{\partial y} + \frac{\partial A_z(x, y, z)}{\partial z}, \quad (1.27)$$

Por ahora consideremos ∇ como un vector.

Existe una gran cantidad de relaciones vectoriales que involucran a ∇ , las cuales se podrán demostrar.

Ejemplo 1.18 Demostrar que:

$$\nabla(\mathbf{a} \cdot \mathbf{b}) = (\mathbf{a} \cdot \nabla) \mathbf{b} + (\mathbf{b} \cdot \nabla) \mathbf{a} + \mathbf{a} \times (\nabla \times \mathbf{b}) + \mathbf{b} \times (\nabla \times \mathbf{a}).$$

El resultado es un gradiente, es decir un vector. El lado izquierdo será:

$$(\nabla(\mathbf{a} \cdot \mathbf{b}))^i = \partial^i(\mathbf{a} \cdot \mathbf{b}) = \partial^i(a_j b^j) = (\partial^i a_j) b^j + (\partial^i b_j) a^j.$$

Mientras que el lado derecho:

$$\begin{aligned}(\nabla(\mathbf{a} \cdot \mathbf{b}))^i &= (a_j \partial^j) b^i + (b_j \partial^j) a^i + \varepsilon^{ijk} a_j (\nabla \times \mathbf{b})_k + \varepsilon^{ijk} b_j (\nabla \times \mathbf{a})_k \\ &= (a_j \partial^j) b^i + (b_j \partial^j) a^i + \varepsilon^{ijk} a_j \varepsilon_{kmn} \partial^m b^n + \varepsilon^{ijk} b_j \varepsilon_{kmn} \partial^m a^n \\ &= (a_j \partial^j) b^i + (b_j \partial^j) a^i + \varepsilon^{ijk} \varepsilon_{mnk} a_j \partial^m b^n + \varepsilon^{ijk} \varepsilon_{mnk} b_j \partial^m a^n \\ &= (a_j \partial^j) b^i + (b_j \partial^j) a^i + (\delta_m^i \delta_n^j - \delta_m^j \delta_n^i) a_j \partial^m b^n + (\delta_m^i \delta_n^j - \delta_m^j \delta_n^i) b_j \partial^m a^n \\ &= a_j \partial^j b^i + b_j \partial^j a^i + \delta_m^i \delta_n^j a_j \partial^m b^n - \delta_m^j \delta_n^i a_j \partial^m b^n + \delta_m^i \delta_n^j b_j \partial^m a^n - \delta_m^j \delta_n^i b_j \partial^m a^n \\ &= a_j \partial^j b^i + b_j \partial^j a^i + a_n \partial^i b^n - a_m \partial^m b^i + b_n \partial^i a^n - b_m \partial^m a^i \\ &= \underbrace{a_j \partial^j b^i}_{=0} - \underbrace{a_m \partial^m b^i}_{=0} + b_j \partial^j a^i - b_m \partial^m a^i + a_n \partial^i b^n + b_n \partial^i a^n \\ &= a_n \partial^i b^n + b_n \partial^i a^n = \partial^i(a_j b^j) = \partial^i(\mathbf{a} \cdot \mathbf{b}).\end{aligned}$$

◀

Ejemplo 1.19 Demostrar la siguiente identidad:

$$\nabla \times (\mathbf{a} \cdot \nabla) \mathbf{a} = (\nabla \cdot \mathbf{a}) (\nabla \times \mathbf{a}) - [\nabla \cdot (\nabla \times \mathbf{a})] \mathbf{a} + (\mathbf{a} \cdot \nabla) (\nabla \times \mathbf{a}) - [(\nabla \times \mathbf{a}) \cdot \nabla] \mathbf{a}.$$

Iniciamos la traducción a índices por el lado izquierdo de la ecuación, así:

$$\begin{aligned}\boldsymbol{\nabla} \times (\mathbf{a} \cdot \boldsymbol{\nabla}) \mathbf{a} &= \epsilon^{ijk} \partial_j (a_m \partial^m) a_k = \epsilon^{ijk} (\partial_j a_m) \partial^m a_k + \epsilon^{ijk} a_m \partial_j \partial^m a_k \\ &= \epsilon^{ijk} (\partial_j a_m) \partial^m a_k + a_m \partial^m (\epsilon^{ijk} \partial_j a_k) .\end{aligned}$$

El lado derecho lo traduciremos término por término:

$$\begin{aligned}(\boldsymbol{\nabla} \cdot \mathbf{a}) (\boldsymbol{\nabla} \times \mathbf{a}) &= (\partial^m a_m) (\epsilon^{ijk} \partial_j a_k) \\ - [\boldsymbol{\nabla} \cdot (\boldsymbol{\nabla} \times \mathbf{a})] \mathbf{a} &= - [\partial_m \epsilon^{mjk} \partial_j a_k] a^i = - [\epsilon^{mjk} \partial_m \partial_j a_k] a^i = 0 \\ (\mathbf{a} \cdot \boldsymbol{\nabla}) (\boldsymbol{\nabla} \times \mathbf{a}) &= a_m \partial^m (\epsilon^{ijk} \partial_j a_k) \\ - [(\boldsymbol{\nabla} \times \mathbf{a}) \cdot \boldsymbol{\nabla}] \mathbf{a} &= - [\epsilon^{mjk} \partial_j a_k] \partial_m a^i .\end{aligned}$$

El segundo término se anula por cuanto ϵ^{mjk} es antisimétrico respecto a los índices m, j , mientras que $\partial_m \partial_j$ es simétrico. El tercer término del desarrollo del lado derecho corresponde con el segundo del desarrollo del lado izquierdo. Por lo tanto, llegamos a la siguiente igualdad:

$$\epsilon^{ijk} (\partial_j a_m) \partial^m a_k = (\partial^m a_m) (\epsilon^{ijk} \partial_j a_k) - [\epsilon^{mjk} \partial_j a_k] \partial_m a^i . \quad (1.28)$$

Para verificar la igualdad tendremos que evaluar componente a componente, esto es, para $i = 1$, el lado izquierdo de (1.28) resulta en:

$$\begin{aligned}\epsilon^{1jk} (\partial_j a_m) \partial^m a_k &= \epsilon^{123} (\partial_2 a_m) \partial^m a_3 + \epsilon^{132} (\partial_3 a_m) \partial^m a_2 = (\partial_2 a_m) \partial^m a_3 - (\partial_3 a_m) \partial^m a_2 \\ &= (\partial_2 a_1) \partial^1 a_3 + (\partial_2 a_2) \partial^2 a_3 + (\partial_2 a_3) \partial^3 a_3 - (\partial_3 a_1) \partial^1 a_2 - (\partial_3 a_2) \partial^2 a_2 - (\partial_3 a_3) \partial^3 a_2 .\end{aligned}$$

Para el primer término del lado derecho de (1.28) :

$$\begin{aligned}(\partial^m a_m) (\epsilon^{1jk} \partial_j a_k) &= (\partial^m a_m) (\epsilon^{123} \partial_2 a_3) + (\partial^m a_m) (\epsilon^{132} \partial_3 a_2) \\ &= \underbrace{\partial_2 a_3 \partial^1 a_1}_{\alpha} + \partial_2 a_3 \partial^2 a_2 + \partial_2 a_3 \partial^3 a_3 - \underbrace{\partial_3 a_2 \partial^1 a_1}_{\beta} - \partial_3 a_2 \partial^2 a_2 - \partial_3 a_2 \partial^3 a_3 ,\end{aligned}$$

y el segundo término de (1.28) se escribe como:

$$\begin{aligned}- [\epsilon^{mjk} \partial_j a_k] \partial_m a^i &= - (\epsilon^{1jk} \partial_j a_k) \partial_1 a^1 - (\epsilon^{2jk} \partial_j a_k) \partial_2 a^1 - (\epsilon^{3jk} \partial_j a_k) \partial_3 a^1 \\ &= - (\partial_2 a_3 - \partial_3 a_2) \partial_1 a^1 - (\partial_3 a_1 - \partial_1 a_3) \partial_2 a^1 - (\partial_1 a_2 - \partial_2 a_1) \partial_3 a^1 \\ &= \underbrace{\partial_3 a_2 \partial_1 a^1}_{\beta} - \underbrace{\partial_2 a_3 \partial_1 a^1}_{\alpha} + \partial_1 a_3 \partial_2 a^1 - \underbrace{\partial_3 a_1 \partial_2 a^1}_{\gamma} + \underbrace{\partial_2 a_1 \partial_3 a^1}_{\gamma} - \partial_1 a_2 \partial_3 a^1 .\end{aligned}$$

Al sumar ambos términos se eliminan los sumandos indicados con letras griegas, y queda como:

$$\begin{aligned}(\partial^m a_m) (\epsilon^{1jk} \partial_j a_k) - [\epsilon^{mjk} \partial_j a_k] \partial_m a^i &= \partial_2 a_3 \partial_2 a_2 + \underbrace{\partial_2 a_3 \partial_3 a_3}_{\Xi} \\ &\quad - \underbrace{\partial_3 a_2 \partial_2 a_2}_{\Omega} - \underbrace{\partial_2 a_2 \partial_3 a_3}_{\Psi} + \underbrace{\partial_1 a_3 \partial_2 a_1}_{\Lambda} - \underbrace{\partial_1 a_2 \partial_3 a_1}_{\Sigma} ,\end{aligned}$$

y al compararlo con el desarrollo del lado derecho de (1.28) e identificar término a término queda demostrada la igualdad:

$$\epsilon^{1jk} (\partial_j a_m) \partial^m a_k = (\partial_2 a_1) \partial_1 a_3 + \underbrace{(\partial_2 a_2) \partial_2 a_3}_{\Xi} + (\partial_2 a_3) \partial_3 a_3 - (\partial_3 a_1) \partial_1 a_2 - (\partial_3 a_2) \partial_2 a_2 - (\partial_3 a_3) \partial_3 a_2 .$$

De igual manera se procede con $i = 2$ e $i = 3$. ◀

Ejemplo 1.20 Utilizando la notación de índices muestre si se cumple la siguiente identidad:

$$\boldsymbol{\nabla} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{a} (\boldsymbol{\nabla} \cdot \mathbf{b}) - \mathbf{b} (\boldsymbol{\nabla} \cdot \mathbf{a}) + (\mathbf{b} \cdot \boldsymbol{\nabla}) \mathbf{a} - (\mathbf{a} \cdot \boldsymbol{\nabla}) \mathbf{b} .$$

Desarrollemos en índices el lado izquierdo:

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = \epsilon^{ijk} \partial_j (\epsilon_{klm} a^l b^m) = (\delta_l^i \delta_m^j - \delta_m^i \delta_l^j) \partial_j (a^l b^m) = \partial_m (a^i b^m) - \partial_l (a^l b^i),$$

expandiendo la derivada:

$$\nabla \times (\mathbf{a} \times \mathbf{b}) = b^m \partial_m (a^i) + a^i \partial_m (b^m) - b^i \partial_l (a^l) - a^l \partial_l (b^i) \equiv (\mathbf{b} \cdot \nabla) \mathbf{a} + (\nabla \cdot \mathbf{b}) \mathbf{a} - (\nabla \cdot \mathbf{a}) \mathbf{b} - (\mathbf{a} \cdot \nabla) \mathbf{b}.$$



Ejemplo 1.21 Tal vez, uno de los problemas que ilustra mejor el uso del álgebra vectorial y la derivación de vectores es el movimiento bajo fuerzas centrales. La ley de gravitación de Newton nos dice que para un sistema de dos masas, m y M se tiene:

$$\sum \mathbf{F} = m \mathbf{a} \Rightarrow mG \frac{M}{r_{mM}^2} \hat{\mathbf{u}}_r = m \frac{d\mathbf{v}}{dt} \Rightarrow \frac{d\mathbf{v}}{dt} = \frac{GM}{r_{mM}^2} \hat{\mathbf{u}}_r.$$

Es costumbre definir la *velocidad areolar*, \mathbf{v}_a , como el área barrida por el radio vector posición, $\mathbf{r}(t)$, que describe la trayectoria de la partícula:

$$2\mathbf{v}_a \equiv \mathbf{r} \times \frac{d\mathbf{r}}{dt} = r \hat{\mathbf{u}}_r \times \frac{d(r \hat{\mathbf{u}}_r)}{dt} = r \hat{\mathbf{u}}_r \times \left(\frac{dr}{dt} \hat{\mathbf{u}}_r + r \frac{d\hat{\mathbf{u}}_r}{dt} \right) = r \hat{\mathbf{u}}_r \times r \frac{d\hat{\mathbf{u}}_r}{dt} = r^2 \hat{\mathbf{u}}_r \times \frac{d\hat{\mathbf{u}}_r}{dt}.$$

Nótese que:

$$\frac{d}{dt} \left(\hat{\mathbf{u}}_r \times \frac{d\hat{\mathbf{u}}_r}{dt} \right) = 0 \Rightarrow \hat{\mathbf{u}}_r \times \frac{d\hat{\mathbf{u}}_r}{dt} = \mathbf{c} \Rightarrow 2\mathbf{v}_a = r^2 \hat{\mathbf{u}}_r \times \frac{d\hat{\mathbf{u}}_r}{dt} = \text{const},$$

donde \mathbf{c} es un vector constante, con lo cual:

$$\begin{aligned} \frac{d}{dt} (\mathbf{v} \times \mathbf{v}_a) &= \frac{d\mathbf{v}}{dt} \times \mathbf{v}_a = \frac{GM}{r_{mM}^2} \hat{\mathbf{u}}_r \times \mathbf{v}_a = \frac{GM}{2} \left\{ \hat{\mathbf{u}}_r \times \left(\hat{\mathbf{u}}_r \times \frac{d\hat{\mathbf{u}}_r}{dt} \right) \right\} \\ &= \frac{GM}{2} \left\{ \left(\hat{\mathbf{u}}_r \cdot \frac{d\hat{\mathbf{u}}_r}{dt} \right) \hat{\mathbf{u}}_r - (\hat{\mathbf{u}}_r \cdot \hat{\mathbf{u}}_r) \frac{d\hat{\mathbf{u}}_r}{dt} \right\} = \frac{GM}{2} \frac{d\hat{\mathbf{u}}_r}{dt}. \end{aligned}$$

Integrando:

$$\mathbf{v} \times \mathbf{v}_a = \frac{GM}{2} \hat{\mathbf{u}}_r + \mathbf{p},$$

donde \mathbf{p} es un vector arbitrario que aparece como constante de integración.

Finalmente nos damos cuenta que:

$$\begin{aligned} \mathbf{r} \cdot (\mathbf{v} \times \mathbf{v}_a) &= r \hat{\mathbf{u}}_r \cdot \left(\frac{GM}{2} \hat{\mathbf{u}}_r + \mathbf{p} \right) = \frac{GM}{2} r + rp \cos(\theta) \\ &= \epsilon^{ijk} r_i v_j v_{ak} \equiv \mathbf{v}_a \cdot (\mathbf{r} \times \mathbf{v}) = \mathbf{v}_a \cdot \mathbf{v}_a = v_a^2, \end{aligned}$$

y entonces:

$$v_a^2 = \frac{GM}{2} r + rp \cos(\theta) \Rightarrow r = \frac{v_a^2}{\frac{GM}{2} + p \cos(\theta)} \equiv \frac{\frac{2v_a^2}{GM}}{1 + \frac{2p}{GM} \cos(\theta)},$$

que constituye la ecuación de una curva cónica ¿Cuál curva?



1.5.6 Integración

Después de haber diferenciado campos escalares y vectoriales, el siguiente paso es integrarlos. Encontraremos algunos objetos vectoriales a integrar y serán:

- Integración de un vector por un escalar:

$$\int \mathbf{A}(u) \, du$$

- Integración de un escalar a lo largo de un vector:

$$\int_C \phi(x, y, z) \, d\mathbf{r}$$

- Integración de un vector a lo largo de otro vector:

$$\int_C \mathbf{A}(x, y, z) \cdot d\mathbf{r}$$

El primero de los casos es el tipo de integral que siempre hemos utilizado para encontrar la posición a partir de la velocidad. Los siguientes tres casos se conocen con el nombre de integrales de línea por cuanto es importante la “ruta” o trayectoria que sigamos al integrar. Esto aparece indicado por la letra C en la integral y será evidente más adelante. En general la integral de línea dependerá de la trayectoria.

1.5.6.1 Un vector por un escalar: $\int \mathbf{A}(u) \, du$

El primer caso de este tipo integrales es el trivial que ya sabemos calcular:

$$\int \mathbf{A}(u) \, du = \hat{\mathbf{i}} \int A_x(u) \, du + \hat{\mathbf{j}} \int A_y(u) \, du + \hat{\mathbf{k}} \int A_z(u) \, du = \left(\int A^i(u) \, du \right) \hat{\mathbf{i}}_i.$$

La integral de un vector (en un sistema de coordenadas cartesianas) por un escalar se convierte en la suma de tres integrales, cada una a lo largo de las componentes cartesianas del vector.

Recordemos que así integramos la aceleración en un movimiento parabólico:

$$\frac{d\mathbf{v}}{dt} = \mathbf{a} = -g\hat{\mathbf{k}} \Rightarrow \mathbf{v} = \int \mathbf{a} \, dt = \hat{\mathbf{k}} \int -g \, dt = -\hat{\mathbf{k}}gt + \mathbf{v}_0 = -\hat{\mathbf{k}}gt + \hat{\mathbf{i}}v_{0x} + \hat{\mathbf{j}}v_{0y} + \hat{\mathbf{k}}v_{0z}.$$

Ahora bien, existen sutilezas en este caso que debemos tener en cuenta. Por ejemplo, considere la integral:

$$\int dt \left(\mathbf{a} \times \frac{d^2\mathbf{a}}{dt^2} \right) = \int dt \left(\frac{d}{dt} \left(\mathbf{a} \times \frac{d\mathbf{a}}{dt} \right) - \frac{d\mathbf{a}}{dt} \times \frac{d\mathbf{a}}{dt} \right) = \int dt \frac{d}{dt} \left(\mathbf{a} \times \frac{d\mathbf{a}}{dt} \right) = \mathbf{a} \times \frac{d\mathbf{a}}{dt} + \mathbf{c}.$$

Pero en general los casos quedan resueltos integrando componente a componente con la ayuda de la notación de índices:

$$\int dt (\mathbf{a} \times \mathbf{b}) = \left[\int dt (\varepsilon^{ijk} a_j b_k) \right] \hat{\mathbf{i}}_i.$$

1.5.6.2 Un escalar a lo largo de un vector: $\int_C \phi(\mathbf{r}) \, dr$

El segundo objeto que “tropezaremos” es la integración de funciones de varias variables a lo largo de una curva determinada. Esto es:

$$\int_C \phi(x, y, z) \, dr = \int_C \phi(x^i) \left(dx\hat{\mathbf{i}} + dy\hat{\mathbf{j}} + dz\hat{\mathbf{k}} \right) = \hat{\mathbf{i}} \int_C \phi(x^i) \, dx + \hat{\mathbf{j}} \int_C \phi(x^i) \, dy + \hat{\mathbf{k}} \int_C \phi(x^i) \, dz.$$

La integral se nos ha convertido en tres integrales, las cuales son ahora componentes de un vector. Esto es posible dado que la base $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ es una base constante. Ahora bien, cada una de estas integrales son interdependientes, dado que hay que seguir la misma curva C .

1.5.6.3 Un vector a lo largo de otro vector: $\int_C \mathbf{F}(\mathbf{r}) \cdot dr$

Quizá la integral de línea más conocida sea una del tipo $\int_C \mathbf{F}(\mathbf{r}) \cdot dr$ por cuanto nos la hemos “tropezado” en el cálculo del trabajo que realiza una fuerza. Todo lo que hemos considerado al parametrizar la curva en el caso anterior, sigue siendo válido.

$$\int_C \mathbf{F}(\mathbf{r}) \cdot dr = \int_C F_x(x, y, z) \, dx + \int_C F_y(x, y, z) \, dy + \int_C F_z(x, y, z) \, dz = \int_C F^i(x^j) \, dx_i.$$

Ejemplo 1.22 Consideremos la siguiente función:

$$\phi(x, y) = 3x^2 + 2y.$$

Queremos calcular la siguiente integral de línea:

$$\int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, d\mathbf{r} = \hat{\mathbf{i}} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, dx + \hat{\mathbf{j}} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, dy.$$

Se requiere especificar la curva C a lo largo de la cual integraremos, en este caso, desde el punto $P_1 \rightarrow (0, 0)$ al punto $P_2 \rightarrow (1, 2)$. Podemos ir por diferentes recorridos:

- Si recorremos la ruta $C_1: (0, 0) \rightarrow (1, 0) \rightarrow (1, 2)$ podemos hacerlo de la manera más sencilla:

$$(0, 0) \rightarrow (1, 0) \Rightarrow y = \text{cte} = 0 \Rightarrow \int_{(0,0)}^{(1,0)} (3x^2 + 2y) \, d\mathbf{r} = \hat{\mathbf{i}} \int_{(0,0)}^{(1,0)} (3x^2 + 2y) \, dx = \hat{\mathbf{i}} \int_0^1 (3x^2) \, dx = \hat{\mathbf{i}}$$

$$(1, 0) \rightarrow (1, 2) \Rightarrow x = \text{cte} = 1 \Rightarrow \int_{(0,0)}^{(1,0)} (3x^2 + 2y) \, d\mathbf{r} = \hat{\mathbf{j}} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, dy = \hat{\mathbf{j}} \int_0^2 (3 + 2y) \, dy = 10\hat{\mathbf{j}}$$

con lo cual:

$$C_1 \longleftrightarrow \underbrace{(0, 0) \rightarrow (1, 0)}_{C_1^A} \underbrace{\rightarrow (1, 2)}_{C_1^B} \Rightarrow \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, d\mathbf{r} = \hat{\mathbf{i}} + 10\hat{\mathbf{j}}.$$

- Si hubiéramos seleccionado la recta que une a estos dos puntos como la curva C_2 entonces:

$$C_2 : y = 2x \Rightarrow dy = 2dx,$$

esto es:

$$\begin{aligned} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, d\mathbf{r} &= \hat{\mathbf{i}} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, dx + \hat{\mathbf{j}} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, dy \\ &= \hat{\mathbf{i}} \int_0^1 (3x^2 + 2(2x)) \, dx + \hat{\mathbf{j}} \int_0^1 (3x^2 + 2(2x)) 2 \, dx = 3\hat{\mathbf{i}} + 6\hat{\mathbf{j}}. \end{aligned}$$

En general la curva C se puede parametrizar y las integrales en varias variables se convertirán en integrales a lo largo del parámetro que caracteriza la curva.

$$C : \{x = x(\tau), y = y(\tau), z = z(\tau)\}$$

Por lo tanto:

$$\begin{aligned} \int_C \phi(x, y, z) \, d\mathbf{r} &= \int_C \phi(x(\tau), y(\tau), z(\tau)) \left(\frac{\partial x(\tau)}{\partial \tau} d\tau \hat{\mathbf{i}} + \frac{\partial y(\tau)}{\partial \tau} d\tau \hat{\mathbf{j}} + \frac{\partial z(\tau)}{\partial \tau} d\tau \hat{\mathbf{k}} \right) \\ &= \hat{\mathbf{i}} \int_C \phi(x(\tau), y(\tau), z(\tau)) \frac{\partial x(\tau)}{\partial \tau} d\tau + \hat{\mathbf{j}} \int_C \phi(x(\tau), y(\tau), z(\tau)) \frac{\partial y(\tau)}{\partial \tau} d\tau \\ &\quad + \hat{\mathbf{k}} \int_C \phi(x(\tau), y(\tau), z(\tau)) \frac{\partial z(\tau)}{\partial \tau} d\tau. \end{aligned}$$

Las parametrización de las curvas anteriores es muy simple:

$$C_1^A = \begin{cases} x = \tau \\ y = 0 \end{cases}; \quad C_1^B = \begin{cases} x = 2 \\ y = \tau \end{cases}; \quad C_2 = \begin{cases} x = \tau \\ y = 2\tau \end{cases}$$

Con esta manera de parametrizar, la integral que resolvimos anteriormente tomando el camino C_2 :

$$\int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, d\mathbf{r} = \hat{\mathbf{i}} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, dx + \hat{\mathbf{j}} \int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, dy.$$

Queda ahora como:

$$\int_{(0,0)}^{(1,2)} (3x^2 + 2y) \, d\mathbf{r} = \hat{\mathbf{i}} \int_0^1 (3\tau^2 + 4\tau) \, d\tau + \hat{\mathbf{j}} \int_0^1 (3\tau^2 + 4\tau) \, 2d\tau = 3\hat{\mathbf{i}} + 6\hat{\mathbf{j}}.$$

Ya que: $0 \leq \tau \leq 1$



Ejemplo 1.23 Consideramos el siguiente campo vectorial:

$$\mathbf{F}(\mathbf{r}) = (3x^2 + 2xy^3) \hat{\mathbf{i}} + 6xy \hat{\mathbf{j}}.$$

Queremos evaluar la siguiente integral:

$$\begin{aligned} \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} &= \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} ((3x^2 + 2xy^3) \hat{\mathbf{i}} + 6xy \hat{\mathbf{j}}) (dx \hat{\mathbf{i}} + dy \hat{\mathbf{j}}) \\ &= \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} (3x^2 + 2xy^3) dx + \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} 6xy dy. \end{aligned}$$

Consideremos que la curva que une esos puntos viene parametrizada por:

$$x = 2\tau^2, \quad y = \tau^3 + \tau \Rightarrow \frac{\partial x(\tau)}{\partial \tau} = 4\tau, \quad \frac{\partial y(\tau)}{\partial \tau} = 3\tau^2 + 1,$$

entonces, la primera de las integrales resulta:

$$\begin{aligned} \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} (3x^2 + 2xy^3) dx &= \int_0^{\frac{\sqrt{2}}{2}} \left(3(2\tau^2)^2 + 2(2\tau^2)(\tau^3 + \tau)^3 \right) (4\tau) d\tau \\ &= \int_0^{\frac{\sqrt{2}}{2}} (16\tau^{12} + 48\tau^{10} + 48\tau^8 + 16\tau^6 + 48\tau^5) d\tau = 1 + \frac{9305}{24024}\sqrt{2}. \end{aligned}$$

Y la segunda:

$$\int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} 6xy dy = \int_0^{\frac{\sqrt{2}}{2}} 6(2\tau^2)(\tau^3 + \tau)(3\tau^2 + 1) d\tau = \frac{65}{32},$$

con lo cual:

$$\int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r} = \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} (3x^2 + 2xy^3) dx + \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} 6xy dy = \frac{97}{32} + \frac{9305}{24024}\sqrt{2}.$$



Ejemplo 1.24 El campo de fuerzas de un oscilador anisótropo bidimensional se escribe como:

$$\mathbf{F} = -k_1 x^2 \hat{\mathbf{i}} + k_2 y \hat{\mathbf{j}}.$$

Encontremos el trabajo realizado

$$\int_{(x_1,y_1)}^{(x_2,y_2)} \mathbf{F} \cdot d\mathbf{r},$$

a lo largo de las siguientes trayectorias:

1. $(1, 1) \rightarrow (4, 1) \rightarrow (4, 4)$

$$\int_{(1,1)}^{(4,1)} (\hat{\mathbf{i}} dx) \cdot (-k_1 x^2 \hat{\mathbf{i}} + k_2 \hat{\mathbf{j}}) + \int_{(4,1)}^{(4,4)} (\hat{\mathbf{j}} dy) \cdot (-k_1 16 \hat{\mathbf{i}} + k_2 y \hat{\mathbf{j}}) = -21k_1 + \frac{15k_2}{2}.$$

2. $(1, 1) \rightarrow (1, 4) \rightarrow (4, 4)$

$$\int_{(1,1)}^{(1,4)} (\hat{\mathbf{j}} dy) \cdot (-k_1 \hat{\mathbf{i}} + k_2 y \hat{\mathbf{j}}) + \int_{(1,4)}^{(4,4)} (\hat{\mathbf{i}} dx) \cdot (-k_1 x^2 \hat{\mathbf{i}} + k_2 4 \hat{\mathbf{j}}) = -21k_1 + \frac{15k_2}{2}.$$

3. $(1, 1) \rightarrow (4, 4)$ siguiendo la recta $x = y$

$$\int_{(1,1)}^{(4,4)} (\hat{\mathbf{i}} dx + \hat{\mathbf{j}} dy) \cdot (-k_1 x^2 \hat{\mathbf{i}} + k_2 x \hat{\mathbf{j}}) = \int_{(1,1)}^{(4,4)} (-k_1 x^2 + k_2 x) dx = -21k_1 + \frac{15k_2}{2}.$$

Dejamos al lector que calcule el trabajo para la trayectoria $y = x^2$.



1.5.7 Ejercicios

1. Demuestre que:

- (a). $\frac{d}{dt} [\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})] = \frac{d\mathbf{a}}{dt} \cdot (\mathbf{b} \times \mathbf{c}) + \mathbf{a} \cdot \left(\frac{d\mathbf{b}}{dt} \times \mathbf{c} \right) + \mathbf{a} \cdot \left(\mathbf{b} \times \frac{d\mathbf{c}}{dt} \right)$
- (b). $\frac{d}{dt} \left[\mathbf{a} \cdot \left(\frac{d\mathbf{a}}{dt} \times \frac{d^2\mathbf{a}}{dt^2} \right) \right] = \mathbf{a} \cdot \left(\frac{d^2\mathbf{a}}{dt^2} \times \frac{d^3\mathbf{a}}{dt^3} \right)$
- (c). $\nabla \times (\nabla \times \mathbf{a}) = \nabla \nabla \cdot \mathbf{a} - \nabla \cdot \nabla \mathbf{a}$
- (d). $\nabla \times (\phi \nabla \phi) = 0$
- (e). $\nabla \times [\mathbf{a} \times (\nabla \times \mathbf{a})] = 0$, si $\mathbf{a} = a_x(y, z)\hat{\mathbf{i}}$.

2. Considere que:

- $\mathbf{r} = x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}} = x^i\hat{\mathbf{i}}_i$,
- $\mathbf{a} = \mathbf{a}(\mathbf{r}) = \mathbf{a}(x, y, z) = a^i(x, y, z)\hat{\mathbf{i}}_i$ y $\mathbf{b} = \mathbf{b}(\mathbf{r}) = \mathbf{b}(x, y, z) = b^i(x, y, z)\hat{\mathbf{i}}_i$,
- $\phi = \phi(\mathbf{r}) = \phi(x, y, z)$ y $\psi = \psi(\mathbf{r}) = \psi(x, y, z)$.

Utilizando la notación de índices, e inspirándose en el ejemplo: 1.13 demuestre las siguientes identidades vectoriales:

- (a). $\nabla(\phi\psi) = \phi\nabla\psi + \psi\nabla\phi$.
- (b). $\nabla \cdot (\phi\mathbf{a}) = \phi\nabla \cdot \mathbf{a} + (\nabla\phi) \cdot \mathbf{a}$.
- (c). $\nabla \times \nabla\phi = 0$.
- (d). $\nabla \cdot (\nabla \times \mathbf{a})$ ¿Qué puede decir de $\nabla \times (\nabla \cdot \mathbf{a})$?
- (e). $\nabla \cdot (\mathbf{a} \times \mathbf{b}) = (\nabla \times \mathbf{a}) \cdot \mathbf{b} - \mathbf{a} \cdot (\nabla \times \mathbf{b})$.
- (f). $\nabla \times (\nabla \times \mathbf{a}) = \nabla(\nabla \cdot \mathbf{a}) - \nabla^2\mathbf{a}$.

3. Para los vectores dados a continuación, encuentre $d\mathbf{r}/ds$

- (a). $\mathbf{r} = t\hat{\mathbf{i}} + 3t^2\hat{\mathbf{j}} - (t-1)\hat{\mathbf{k}}$, y $t = \ln(1+s^2)$.
- (b). $\mathbf{r} = \sin(t)\hat{\mathbf{i}} + \cos(t)\hat{\mathbf{j}} + \tan(t)\hat{\mathbf{k}}$, y $t = 2+s^2$.

4. Una partícula describe un movimiento dado por el vector posición \mathbf{r} . Encuentre la componente de su velocidad en la dirección del vector indicado:

- (a). $\mathbf{r} = t^2\hat{\mathbf{i}} + 4\cos(2t)\hat{\mathbf{j}} + 3\sin(2t)\hat{\mathbf{k}}$, $2\hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}$.
- (b). $\mathbf{r} = 3\cos(t)\hat{\mathbf{i}} + 3\sin(t)\hat{\mathbf{j}} + (t^2 - 2)\hat{\mathbf{k}}$, $\hat{\mathbf{i}} + 2\hat{\mathbf{j}} - \hat{\mathbf{k}}$.

5. Si \mathbf{u} , \mathbf{v} y \mathbf{w} son funciones que dependen del parámetro t , demuestre que:

- (a). $\frac{d}{dt} [\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})] = \mathbf{u} \cdot (\mathbf{v} \times \frac{d\mathbf{w}}{dt}) + \mathbf{u} \cdot (\frac{d\mathbf{v}}{dt} \times \mathbf{w}) + \frac{d\mathbf{u}}{dt} \cdot (\mathbf{v} \times \mathbf{w})$.
- (b). $\frac{d}{dt} [\mathbf{u} \times (\mathbf{v} \times \mathbf{w})] = \mathbf{u} \times (\mathbf{v} \times \frac{d\mathbf{w}}{dt}) + \mathbf{u} \times (\frac{d\mathbf{v}}{dt} \times \mathbf{w}) + \frac{d\mathbf{u}}{dt} \times (\mathbf{v} \times \mathbf{w})$.

6. Si $\mathbf{u} = 2t\hat{\mathbf{i}} - t^2\hat{\mathbf{j}} + \hat{\mathbf{k}}$, $\mathbf{v} = 2\hat{\mathbf{i}} + 3t\hat{\mathbf{j}} + t\hat{\mathbf{k}}$ y $\mathbf{w} = t\hat{\mathbf{i}} + 2t\hat{\mathbf{j}} - t\hat{\mathbf{k}}$. Utilice el resultado del ejercicio (a) anterior para encontrar: $\frac{d}{dt} [\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})]$.

7. Si $\mathbf{u} = t\hat{\mathbf{i}} - t\hat{\mathbf{j}} + t^2\hat{\mathbf{k}}$, $\mathbf{v} = -t\hat{\mathbf{i}} + 2t\hat{\mathbf{j}} - t^2\hat{\mathbf{k}}$ y $\mathbf{w} = 2t\hat{\mathbf{i}} - 2t\hat{\mathbf{j}} + t\hat{\mathbf{k}}$. Utilice el resultado del ejercicio (b) anterior para encontrar: $\frac{d}{dt} [\mathbf{u} \times (\mathbf{v} \times \mathbf{w})]$.

8. Encuentre el gradiente de los siguientes campos:

- (a). $\phi(x, y, z) = x^2 + 3xyz - yz^2$.
- (b). $\phi(x, y, z) = (x^2 + 2y^2 + 4z^2)^{-1}$.

9. Encuentre la divergencia de los siguientes campos:

- (a). $\mathbf{A}(x, y, z) = x^2y\hat{\mathbf{i}} + y^2z^2\hat{\mathbf{j}} + xz^3\hat{\mathbf{k}}$.
- (b). $\mathbf{A}(x, y, z) = (1-x^2)\hat{\mathbf{i}} + \sin(yz)\hat{\mathbf{j}} + e^{xyz}\hat{\mathbf{k}}$.

10. Encuentre el rotacional de los siguientes campos:

(a). $\mathbf{A}(x, y, z) = xyz^2\hat{\mathbf{i}} + x^2yz\hat{\mathbf{j}} + xy^2\hat{\mathbf{k}}$.

(b). $\mathbf{A}(x, y, z) = \operatorname{senh}(xy)\hat{\mathbf{i}} + \cosh(yz)\hat{\mathbf{j}} + xyz\hat{\mathbf{k}}$.

11. Evalúe las siguientes integrales:

(a). $\int \left(t \operatorname{sen}(t)\hat{\mathbf{i}} + 2t^2\hat{\mathbf{j}} - 7t\hat{\mathbf{k}} \right) dt$.

(b). $\int \left(\cosh^2(t)\hat{\mathbf{i}} + 2\operatorname{sen}^2(2t)\hat{\mathbf{j}} - \hat{\mathbf{k}} \right) dt$.

12. Un campo de fuerza:

$$\mathbf{F} = -kx\hat{\mathbf{i}} - ky\hat{\mathbf{j}},$$

actúa sobre un oscilador. Compare el trabajo hecho al moverse en contra de este campo al ir desde el punto $(1, 1)$ al punto $(4, 4)$, siguiendo los siguientes caminos:

a) $(1, 1) \rightarrow (4, 1) \rightarrow (4, 4)$. b) $(1, 1) \rightarrow (1, 4) \rightarrow (4, 4)$. c) $(1, 1) \rightarrow (4, 4)$, siguiendo el camino: $x = y$.

13. Dado el campo de fuerza:

$$\mathbf{F} = -\frac{y}{x^2 + y^2}\hat{\mathbf{i}} + \frac{x}{x^2 + y^2}\hat{\mathbf{j}}.$$

Calcule el trabajo hecho en contra de este campo de fuerza al moverse al rededor de un círculo de radio uno y en el plano $x - y$.

(a). desde 0 a π en sentido contrario a las agujas del reloj.

(b). desde 0 a $-\pi$ en sentido de las agujas del reloj.

14. Evaluar la siguiente integral:

$$\oint \mathbf{r} \cdot d\mathbf{r}.$$

15. Una partícula se mueve bajo la ley $\mathbf{r}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}$, con:

$$x(t) = 2t^2; \quad y(t) = t^2 - 4t; \quad z(t) = 3t - 5.$$

El parámetro t representa el tiempo.

Encuentre las expresiones para la aceleración y la velocidad de la partícula, para $t = 1$ y en la dirección del vector $\hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 2\hat{\mathbf{k}}$.

16. Suponga ahora el caso general de una partícula que se mueve en una curva descrita por:

$$\mathbf{r}(t) = x(t)\hat{\mathbf{i}} + y(t)\hat{\mathbf{j}} + z(t)\hat{\mathbf{k}}.$$

Muestre que el vector velocidad es tangente a la trayectoria descrita.

17. Encuentre la ecuación vectorial para una trayectoria recta que pasa por los puntos $P \rightarrow (1, 2, 3)$ y $Q \rightarrow (1, 1, 1)$.

18. Encuentre el ángulo entre los siguientes planos: $x + y + z = 9$ y $x + y - z = 3$.

19. Un fluido se considera irrotacional si su campo de velocidades $\mathbf{v} = \mathbf{v}(\mathbf{r}) = \mathbf{v}(x, y, z)$ cumple con la ecuación: $\nabla \times \mathbf{v} = 0$. Suponga, ahora que: $\mathbf{v} = (x + 2y + az)\hat{\mathbf{i}} + (bx - 3y - z)\hat{\mathbf{j}} + (4x + cy + 2z)\hat{\mathbf{k}}$.

(a). Encuentre el valor de a, b y c para que este campo de velocidades sea irrotacional.

(b). Es intuitivo convencerte que si $\nabla \times \mathbf{v} = 0 \Rightarrow \mathbf{v} = \nabla\psi$. Encuentre la expresión para la función potencial $\psi = \psi(\mathbf{r}) = \psi(x, y, z)$.

(c). Considere la siguiente integral: $I = \int_{\mathcal{C}} d\mathbf{r} \cdot \mathbf{v}$. Donde \mathcal{C} es el circuito a recorrer.

I. Calcule el valor de la integral I a lo largo del trayecto: $(0, 0, 0) \rightarrow (1, 1, 0)$ mediante una segmento de recta. Luego, de $(1, 1, 0) \rightarrow (2, 0, 0)$ a lo largo de otro segmento de recta. Finalmente regresando $(2, 0, 0) \rightarrow (0, 0, 0)$ también siguiendo una recta.

II. Calcule el valor de la integral I de $(0, 0, 0) \rightarrow (2, 0, 0)$ a lo largo de un arco de circunferencia que cumple con la ecuación: $(x - 1)^2 + y^2 = 1$. Ahora regresando de $(2, 0, 0) \rightarrow (0, 0, 0)$ también a través de una recta.

III. ¿Qué puede concluir del campo \mathbf{v} ?

20. Resuelva todos los problemas anteriores utilizando **Maxima**.

Practicando con Maxima

1. **Maxima** nos permite hacer cálculos con el operador ∇ en coordenadas cartesianas, más adelante veremos que también es posible en otros sistemas de coordenadas. Los operadores que se pueden expresar son: **grad**(gradiente), **div** (divergencia), **curl** (rotacional), **laplacian** (laplaciano). Debemos hacer uso de la biblioteca **vect** para poder utilizar los operadores diferenciales.

Por otro lado, se debe usar también de función **express(expr)** para transformar los nombres de los operadores diferenciales en expresiones que contienen derivadas parciales y con la función **ev** evaluamos la expresión. Veamos como se hace si queremos calcular el gradiente de la siguiente función:

$$f = \frac{x^2 + y^2}{(x^2 + y^2 + z^2)^{1/2}}.$$

```
(%ci1) load(vect)$
(%ci2) f:(x^2+y^2)/(x^2+y^2+z^2)^(1/2);
(%eo2) 
$$\frac{y^2 + x^2}{\sqrt{z^2 + y^2 + x^2}}$$

(%ci3) grad(f);
(%eo3) grad 
$$\left( \frac{y^2 + x^2}{\sqrt{z^2 + y^2 + x^2}} \right)$$

```

Estaremos haciendo uso del símbolo **%** que representa la última expresión de salida, de esta manera nos ahorraremos estar escribiendo nuevamente el último resultado en la nueva linea de entrada.

```
(%ci4) express(%);
(%eo4) 
$$\left[ \frac{d}{dx} \left( \frac{y^2 + x^2}{\sqrt{z^2 + y^2 + x^2}} \right), \frac{d}{dy} \left( \frac{y^2 + x^2}{\sqrt{z^2 + y^2 + x^2}} \right), \frac{d}{dz} \left( \frac{y^2 + x^2}{\sqrt{z^2 + y^2 + x^2}} \right) \right]$$

```

Notemos que el resultado anterior es una lista del tipo $[x, y, z]$, que es la manera como el programa maneja los vectores. En este caso, el primer elemento de la lista es la componente x del vector gradiente. De la misma manera para las componentes y y z . Evaluamos las derivadas de la última salida:

```
(%ci5) ev(% , diff);
(%eo5) 
$$\left[ \frac{2x}{\sqrt{z^2 + y^2 + x^2}} - \frac{x(y^2 + x^2)}{(z^2 + y^2 + x^2)^{3/2}}, \frac{2y}{\sqrt{z^2 + y^2 + x^2}} - \frac{y(y^2 + x^2)}{(z^2 + y^2 + x^2)^{3/2}}, - \frac{(y^2 + x^2)z}{(z^2 + y^2 + x^2)^{3/2}} \right]$$

```

En este punto podemos intentar simplificar las expresiones anteriores, esto lo hacemos con el comando **ratsimp** y así obtener finalmente el vector gradiente:

```
(%ci6) ratsimp(%);
(%eo6) 
$$\left[ \frac{\sqrt{z^2 + y^2 + x^2}(2xz^2 + xy^2 + x^3)}{z^4 + (2y^2 + 2x^2)z^2 + y^4 + 2x^2y^2 + x^4}, \frac{\sqrt{z^2 + y^2 + x^2}(2yz^2 + y^3 + x^2y)}{z^4 + (2y^2 + 2x^2)z^2 + y^4 + 2x^2y^2 + x^4}, - \frac{(y^2 + x^2)z}{(z^2 + y^2 + x^2)^{3/2}} \right]$$

```

$$-\frac{(y^2+x^2)z}{(z^2+y^2+x^2)^{\frac{3}{2}}}$$

Hagamos uso de los otros operadores. Por ejemplo, dado el vector:

$$\mathbf{a} = \frac{x^2}{x^2+y^2}\hat{\mathbf{i}} + \frac{y^2}{x^2+y^2}\hat{\mathbf{j}} + \frac{z^2}{x^2+y^2}\hat{\mathbf{k}},$$

calculemos la divergencia, $\nabla \cdot \mathbf{a}$. Esta vez combinaremos algunos comandos en una misma linea.

```
(%ci7) a:[x^2/(x^2+y^2),y^2/(x^2+y^2),z^2/(x^2+y^2)];
(%o7) [x^2/(y^2+x^2),y^2/(y^2+x^2),z^2/(y^2+x^2)]
(%ci8) express(div(a));
(%o8) d/dz (z^2/(y^2+x^2)) + d/dy (y^2/(y^2+x^2)) + d/dx (x^2/(y^2+x^2))
(%ci9) ev(%, diff), ratsimp;
(%o9) (2 y^2 + 2 x^2) z + 2 x y^2 + 2 x^2 y
y^4 + 2 x^2 y^2 + x^4
```

Calculemos ahora el rotacional de \mathbf{a} , $\nabla \times \mathbf{a}$, pero combinado más comandos en una sola instrucción:

```
(%ci10) ev(express(curl(a)), diff);
(%o10) [-2 y z^2/(y^2+x^2)^2, 2 x z^2/(y^2+x^2)^2, 2 x^2 y/(y^2+x^2)^2 - 2 x y^2/(y^2+x^2)^2]
```

Intentemos simplificar nuevamente para ver lo que ocurre:

```
(%ci11) ratsimp(%);
(%o11) [-2 y z^2/(y^4+2 x^2 y^2+x^4), 2 x z^2/(y^4+2 x^2 y^2+x^4), -2 x y^2-2 x^2 y/(y^4+2 x^2 y^2+x^4)]
```

Finalmente, calculemos el laplaciano de f , $\nabla^2 f$, todo en una sola instrucción

```
(%ci12) ratsimp(ev(express(laplacian(f)), diff));
(%o12) 4 z^2/(z^2+y^2+x^2)^{\frac{3}{2}}
```

Podemos verificar rápidamente que: $\nabla^2 f = \nabla \cdot \nabla f$

```
(%ci13) div(grad(f));
(%o13) div(grad((y^2+x^2)/sqrt(z^2+y^2+x^2)))
(%ci14) ratsimp(ev(express(%), diff));
(%o14) 4 z^2/(z^2+y^2+x^2)^{\frac{3}{2}}
```

2. **Maxima** integra simbólicamente funciones por medio del comando **integrate**(*expr*, *x*, *a*, *b*). Esto es, calcula la integral definida de la *expr* respecto de *x* con los límites de integración *a* y *b*. Si se omiten los límites de integración se calcula la integral indefinida.

Por ejemplo, si queremos calcular la integral:

$$\int a \sin(x+b)^3 dx .$$

Escribimos:

(%ci15) `integrate(a*sin(x+b)^3,x);`
 (%o15) $a \left(\frac{\cos^3(x+b)}{3} - \cos(x+b) \right)$

Calculemos ahora el área de la región comprendida entre las funciones:

$$f(x) = x^3 - 3x^2 + 1 \quad y \quad g(x) = -x + 1.$$

Es decir, vamos a calcular la integral:

$$\int_a^b [f(x) - g(x)] dx.$$

Aprovecharemos este ejercicio para aprender, entre otras cosas, a graficar y definir funciones. Primero que todo, debemos introducir al programa las dos funciones y a diferencia de los cálculos anteriores, usaremos := para definirlas.

(%ci16) `f(x):=x^3-3*x^2+1; g(x):=-x+1;`
 (%o16) $f(x) := x^3 - 3 * x^2 + 1$
 (%o17) $g(x) := -x + 1$

La diferencia de estas dos funciones la podemos llamar la función $d(x)$.

(%ci18) `d(x):=f(x)-g(x);`
 (%o18) $d(x) := f(x) - g(x)$

Para encontrar los puntos donde las funciones se interceptan utilizaremos el comando **solve**, que resuelve la ecuación algebraica y devuelve una lista de igualdades con las variables despejadas. Si la expresión a resolver no es una igualdad, se supone que se quiere resolver la ecuación ya igualada a cero.

Calculemos los puntos donde se interceptan las curvas:

(%ci20) `solve(f(x)-g(x),x);`
 (%o20) $\left[x = -\frac{\sqrt{5} - 3}{2}, x = \frac{\sqrt{5} + 3}{2}, x = 0 \right]$

Los valores numéricos son:

(%ci21) `float(%);`
 (%o21) $[x = 0,3819660112501051, x = 2,618033988749895, x = 0,0]$

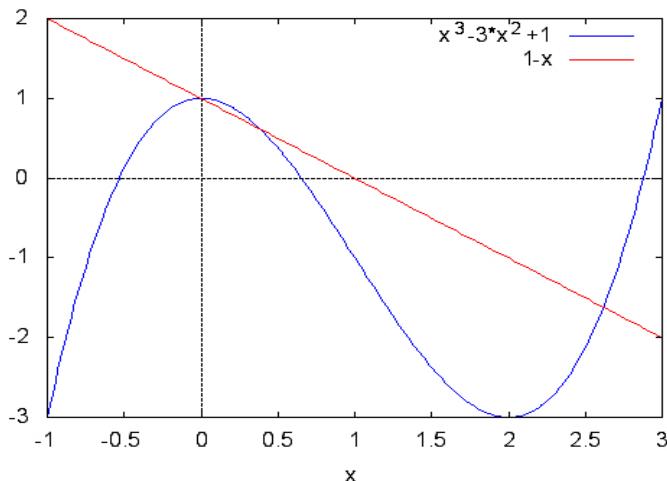
Haremos la gráfica de las dos funciones. **Maxima** permite hacer gráficos en 2D con el comando **plot2d**. Pero aquí utilizaremos una variante y haremos el gráfico con el comando **wxplot2d** para que la figura aparezca dentro de nuestra hoja de cálculo y no nos aparezca una ventana emergente con la gráfica. Se recomienda revisar el Manual del programa para ver las múltiples opciones adicionales que se pueden agregar a los comandos de hacer las gráficas.

En **Maxima** necesario introducir las funciones como una lista, es decir, dentro de corchetes. Por otra parte, le pediremos al programa que grafique para los valores comprendidos de: $-1 \leq x \leq 3$.

Entonces, para graficar las funciones $f(x)$ y $g(x)$ entre $[-1, 3]$ escribimos el siguiente comando:

(%ci19) `wxplot2d([f(x),g(x)],[x,-1,3]);`

(%o19)



Procedemos ahora sí a integrar la función diferencia para encontrar el área contenida dentro de las dos funciones. En este caso será la integral:

$$A = \int_0^{-\frac{\sqrt{5}-3}{2}} d(x) \, dx + \int_{-\frac{\sqrt{5}-3}{2}}^{\frac{\sqrt{5}-3}{2}} (-d(x)) \, dx.$$

(%i22) `integrate(d(x),x,0,-(sqrt(5)-3)/2)+integrate(-d(x),x,-(sqrt(5)-3)/2,(sqrt(5)+3)/2);`

$$(\%o22) \frac{5^{\frac{3}{2}} + 11}{8} + \frac{5^{\frac{3}{2}} - 11}{4}$$

(%i23) `float(%);`

$$(\%o23) 2,817627457812107$$

3. Calculemos ahora las dos integrales que aparecen en el ejemplo 1.23. Es decir las integrales:

$$\int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} (3x^2 + 2xy^3) \, dx, \quad \int_{(0,0)}^{(1,\frac{3}{4}\sqrt{2})} 6xy \, dy.$$

donde hicimos los cambio de variables: $x = 2\tau^2$, $y = \tau^3 + \tau \Rightarrow dx = 4\tau d\tau$, $dy = (3\tau^2 + 1)d\tau$.

Escribiremos primero los dos integrandos:

(%i24) `int1:(3*x^2+2*x*y^3)*dx; int2:6*x*y*dy;`

$$(\%o24) dx (2 x y^3 + 3 x^2)$$

$$(\%o25) 6 dy x y$$

Realizaremos el cambio de variable a través de comando **subst** que nos permite sustituir expresiones dentro de otras:

(%i26) `e1:subst([x=2*t^2,dx=4*t,y=t^3+t,dy=3*t^2+1],int1);`

$$(\%o26) 4 t \left(4 t^2 (t^3 + t)^3 + 12 t^4\right)$$

(%i27) `e2:subst([x=2*t^2,dx=4*t,y=t^3+t,dy=3*t^2+1],int2);`

$$(\%o27) 12 t^2 (3 t^2 + 1) (t^3 + t)$$

Podemos introducir las integrales sin evaluarlas en el momento, esto se consigue mediante el operador

de comilla simple. En este caso la función no es evaluada y devuelve una expresión simbólica o imagen pictórica.

```
(%ci28) int1:'integrate(e1,t,0,sqrt(2)/2);
(%o28) 4 ∫₀^{√²} t ⎛ 4t² (t³ + t)³ + 12t⁴ ⎞ dt
```

Ahora si evaluamos la integral:

```
(%ci29) ev(int1,integrate),expand;
(%o29) 9305
         3003 2² + 1
```

Y para la segunda integral:

```
(%ci30) int2:'integrate(e2,t,0,sqrt(2)/2);
(%o30) 12 ∫₀^{√²} t² (3t² + 1) (t³ + t) dt
(%ci31) ev(int2,integrate),expand;
(%o31) 65
         32
(%ci32) kill(all)$
```

4. Aclaremos la diferencia entre asignar una expresión a una variable y la definición de funciones. Consideremos el movimiento de una partícula que sigue la trayectoria:

$$x = x_0 + v_0 t + \frac{1}{2} a t^2.$$

Primero escribiremos la expresión y se la asignamos a una variable que llamaremos x .

Note que usaremos : para asignar el lado derecho a la variable.

```
(%ci1) x: x0+v0*t+a*t^2/2;
(%o1) x₀ + t v₀ + a t²
```

Si queremos evaluarla para algún valor de la variable t , digamos $t = 1$, entonces podemos escribir:

```
(%ci2) subst(t=1,x);
(%o2) x₀ + v₀ + a
```

Si queremos evaluar la expresión con el resto de parámetros, entonces:

```
(%ci3) subst([t=1,a=-9.8,x0=0,v0=10],x);
(%o3) 5,1
```

Las sustituciones se realizan pero no se asignan a las variables. Podemos ver que x sigue siendo lo que le asignamos originalmente.

```
(%ci4) x;
(%o4) x₀ + t v₀ + a t²
```

Ahora definamos la expresión anterior pero como la función $x(t)$. Ahora utilizaremos := para definir la

función,

$$\begin{aligned} (\%i5) \quad & x(t) := x_0 + v_0 * t + a * t^2 / 2; \\ (\%o5) \quad & x(t) := x_0 + v_0 t + a \frac{t^2}{2} \end{aligned}$$

Evaluárla es ahora más sencillo:

$$\begin{aligned} (\%i6) \quad & x(1); \\ (\%o6) \quad & x_0 + v_0 + \frac{a}{2} \end{aligned}$$

También podemos escribir:

$$\begin{aligned} (\%i7) \quad & \text{subst}([a=-9.8,x0=0,v0=10],x(1)); \\ (\%o7) \quad & 5,1 \end{aligned}$$

Si queremos asignarle los valores a los parámetros de manera global, entonces escribimos:

$$(\%i8) \quad a:-9.8\$ \quad x0:0\$ \quad v0:10\$$$

De manera que:

$$\begin{aligned} (\%i11) \quad & x(1); \\ (\%o11) \quad & 5,1 \end{aligned}$$

Para limpiar los parámetros que acabamos de asignar hacemos lo siguiente:

$$(\%i12) \quad \text{kill}(a,v0,x0)\$$$

Por lo tanto:

$$\begin{aligned} (\%i13) \quad & x(1); \\ (\%o13) \quad & x_0 + v_0 + \frac{a}{2} \end{aligned}$$

La velocidad y la aceleración son funciones fáciles de calcular:

$$\begin{aligned} (\%i14) \quad & \text{'diff}(x(t),t)=\text{diff}(x(t),t); \\ (\%o14) \quad & \frac{d}{dt} \left(x_0 + t v_0 + \frac{a t^2}{2} \right) = v_0 + a t \\ (\%i15) \quad & \text{'diff}(x(t),t,2)=\text{diff}(x(t),t,2); \\ (\%o15) \quad & \frac{d^2}{dt^2} \left(x_0 + t v_0 + \frac{a t^2}{2} \right) = a \end{aligned}$$

En el caso de que estemos interesados en definir una nueva función a partir de algún cálculo anterior podemos utilizar la función **define**.

Entonces, lo que vamos a hacer, en las siguientes líneas de comandos, es definir la función velocidad a partir de la derivada de la posición.

$$\begin{aligned} (\%i16) \quad & \text{diff}(x(t),t); \\ (\%o16) \quad & v_0 + a t \end{aligned}$$

$$\begin{aligned} (\%i17) \quad & \text{define}(v(t), \%); \\ (\%o17) \quad & v(t) := v_0 + a t \end{aligned}$$

(%i18) $v(1);$
 (%o18) $v0 + a$

(%i19) $\text{subst}([a=-9.8,x0=0,v0=10],v(1));$
 (%o19) 0.1999999999999993

5. Cuando necesitemos manejar campo vectoriales podemos utilizar las listas. Veamos el primer ejemplo de 1.16 donde:

$$\mathbf{r} = 3t^2\hat{\mathbf{i}} + (4t^3 - t)\hat{\mathbf{j}} + t\hat{\mathbf{k}}.$$

Escribamos el vector posición:

(%i20) $[x,y,z]:[3*t^2,4*t^3-t,t];$
 (%o20) $[3t^2, 4t^3 - t, t]$

La velocidad:

(%i21) $v:\text{diff}([x,y,z],t);$
 (%o21) $[6t, 12t^2 - 1, 1]$

La aceleración:

(%i22) $a:\text{diff}([x,y,z],t,2);$
 (%o22) $[6, 24t, 0]$

Podemos fácilmente calcular:

(%i23) $v.[bx,by,bz];$
 (%o23) $by (12t^2 - 1) + 6bx t + bz$

(%i24) $\text{solve}(%o,bz);$
 (%o24) $[bz = -12byt^2 - 6bx t + by]$

1.6 Vectores y números complejos

Desde los primeros cursos de matemática nos hemos tropezado con las llamadas raíces imaginarias o complejas de polinomios. La solución a un polinomio cúbico:

$$x^3 - 3x^2 + 4x - 12 = 0 \Rightarrow \left\{ \begin{array}{l} x = 2i \\ x = -2i \\ x = 3 \end{array} \right\} \Rightarrow (x + 2i)(x - 2i)(x - 3) = 0,$$

o cuadrático:

$$x^2 + 4 = 0 \Rightarrow \left\{ \begin{array}{l} x = 2i \\ x = -2i \end{array} \right\} \Rightarrow (x + 2i)(x - 2i) = 0.$$

nos lleva a definir un número $i^2 \equiv -1$.

De lo anterior podemos ver que al multiplicar el número imaginario i por cualquier número real obtendremos el número imaginario puro ib , con $b \in \mathbb{R}$. La nomenclatura de números imaginarios surgió de la idea de que estas cantidades no representaban mediciones físicas. Esa idea ha sido abandonada pero el nombre quedó¹⁷.

Es importante señalar aquí, que suele tomarse de la definición $i^2 \equiv -1$ el hecho de que $i = \sqrt{-1}$. Pero si hacemos esto, sin pensarlo bien, llegamos a que: $-1 = i^2 = ii = \sqrt{-1}\sqrt{-1} = \sqrt{1} = 1$. La contradicción surge por el hecho de que aquí el -1 no es un número real sino un número complejo y será evidente más adelante cuando desarrollemos el álgebra de los números complejos.

1.6.1 Los números complejos y su álgebra

Un número complejo, z , es la generalización de los números imaginarios (puros), ib . Esto es:

$$z = a + ib \quad \text{con } a, b \in \mathbb{R} \Rightarrow \left\{ \begin{array}{l} a \rightarrow \text{parte real} \\ b \rightarrow \text{parte imaginaria} \end{array} \right. \quad (1.29)$$

Obviamente los números reales serán $a + i0$ números complejos con su parte imaginaria nula. Los números imaginarios puros serán números complejos con su parte real nula, esto es, $0 + ib$. Por ello, en general diremos que:

$$z = a + ib \Rightarrow a = \operatorname{Re}(z) \quad \wedge \quad b = \operatorname{Im}(z), \quad (1.30)$$

es decir, a corresponde a la parte real de z y b a su parte imaginaria.

Cada número complejo, z tendrá asociado un número complejo conjugado, z^* tal que:

$$\begin{aligned} z = a + ib &\iff z^* = a - ib, \\ &\Downarrow \\ (z^*)^* &= z \quad \wedge \quad z \cdot z^* = a^2 + b^2, \end{aligned}$$

claramente:

$$z \cdot z^* \geq 0 \Rightarrow |z|^2 = |z^*|^2 = z \cdot z^* = a^2 + b^2.$$

Es importante señalar que, en general, no existe relación de orden entre los números complejos. Vale decir,

¹⁷Los números complejos comienzan a aparecer en los trabajos de Cardano (1501-1576) y Bombelli (1526-1672) mientras estudiaban las raíces de la ecuación cúbica. Pero es a René Descartes (1596-1650) a quién se le atribuye la afirmación: "ciertas ecuaciones algebraicas sólo tienen solución en nuestra imaginación" y utilizó el término "números imaginarios". Caspar Wessel en 1799 y Jean-Robert Argand en 1806 proponen la estructura del plano complejo y la representación de la unidad imaginaria como el punto $(0, 1)$ del eje vertical de dicho plano. Pero el término, hoy usado de "números complejos" se debe a Johann Carl Friedrich Gauss (1777-1855).

que no sabremos si un número complejo es mayor que otro. No está definida esta operación.

$$z_1 \not> z_2 \quad \vee \quad z_1 \not< z_2.$$

Las relaciones de orden sólo se podrán establecer entre módulos de números complejos y no números complejos en general.

Rápidamente recordamos el álgebra de los números complejos:

- Dos números complejos serán iguales si sus partes reales e imaginarios lo son:

$$z_1 = z_2 \Rightarrow (a_1 + ib_1) = (a_2 + ib_2) \Rightarrow a_1 = a_2 \quad \wedge \quad b_1 = b_2.$$

- Se suman dos números complejos sumando sus partes reales y sus partes imaginarias:

$$z_3 = z_1 + z_2 \Rightarrow (a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2) = a_3 + ib_3,$$

- El conjugado de un número complejo cambia el signo a la parte imaginaria

$$(a_1 + ib_1)^* = (a_1 - ib_1),$$

y claramente $z + z^* = 2\operatorname{Re} z$, también $z - z^* = 2\operatorname{Im} z$. Igualmente es inmediato comprobar que: $(z_1 + z_2)^* = z_1^* + z_2^*$.

- Se multiplican números complejos por escalares multiplicando el escalar por sus partes reales e imaginarias:

$$z_3 = \alpha z_1 \Rightarrow \alpha(a_1 + ib_1) = \alpha a_1 + i(\alpha b_1).$$

- Se multiplican números complejos entre si, multiplicando los dos binomios y teniendo cuidado que $i^2 = -1$:

$$z_3 = z_1 z_2 \Rightarrow (a_1 + ib_1) \cdot (a_2 + ib_2) = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2),$$

también es inmediato comprobar que $(z_1 z_2)^* = z_1^* z_2^*$.

- Se dividen números complejos siguiendo la estrategia de racionalización de fracciones irracionales. Esto es:

$$z_3 = \frac{z_1}{z_2} \Rightarrow \frac{a_1 + ib_1}{a_2 + ib_2} = \frac{(a_1 + ib_1)(a_2 - ib_2)}{(a_2 + ib_2)(a_2 - ib_2)} = \frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2} + i \frac{b_1 a_2 - a_1 b_2}{a_2^2 + b_2^2},$$

es claro que el divisor será cualquier número complejo excepto el cero complejo: $0 + i0$.

1.6.2 Vectores y el plano complejo

Mirando con cuidado el álgebra de números complejos nos damos cuenta que un número complejo puede ser representado por una *dupla* de números, es decir:

$$z = (a + ib) \quad \Leftarrow \quad z = (a, b). \quad (1.31)$$

Las propiedades entre números complejos de igualdad, suma y multiplicación por un escalar arriba expuestas se cumplen de forma inmediata con esta nueva representación. Hay que definir las operaciones de multiplicación y división entre números complejos de forma que:

$$(a_1, b_1)(a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + b_1 a_2) \quad \wedge \quad \frac{(a_1, b_1)}{(a_2, b_2)} = \left(\frac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2}, \frac{b_1 a_2 - a_1 b_2}{a_2^2 + b_2^2} \right).$$

Con esta definición para la multiplicación de números complejos podemos ver que si la unidad compleja viene representada por el par $i = (0, 1)$, entonces ahora si:

$$i^2 = ii = (0, 1)(0, 1) = (-1, 0) = -1.$$

La asociación de un número complejo con una pareja de números inmediatamente nos lleva a imaginar un punto (x, y) en un plano (complejo), la primera componente (horizontal) representa la parte real y la segunda componente (vertical) representa la parte imaginaria.

De esta forma asociamos un número complejo a un vector que une ese punto (x, y) con el origen del plano complejo.

Como mencionamos con anterioridad, todo número complejo z tiene asociado su complejo conjugado z^* . La representación geométrica de z^* no es otra cosa que la reflexión de z respecto al eje real. Como puede verse en la figura 1.10.

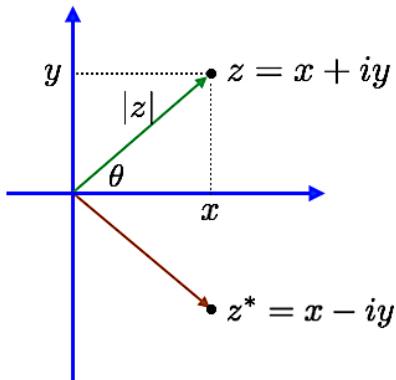


Figura 1.10: Representación del plano complejo. El eje horizontal recibe el nombre de eje real, y el eje vertical el nombre de eje imaginario.

Por otro lado, $|z| = \sqrt{zz^*}$ viene a ser la distancia del punto $(0, 0)$ al punto (x, y) , es decir, la longitud o norma del vector (x, y) . Esta representación de números complejos como vectores en el plano (complejo) se conoce con el nombre de Diagrama de Argand¹⁸ a pesar que no fue Jean Argand, sino Caspar Wessel¹⁹ el primero en proponerlo. Por cierto, esta interpretación fue tres veces redescubierta, primero por Caspar Wessel en 1799, luego por Jean Argand en 1806 y finalmente por Gauss²⁰ en 1831.

De esta manera, como un recordatorio al plano real podemos ver que:

$$z = x + iy \iff z = r(\cos(\theta) + i \sin(\theta)), \quad \text{con: } \begin{cases} r = \sqrt{zz^*} = |z| = \sqrt{x^2 + y^2} \\ \tan(\theta) = \frac{y}{x} \quad \text{donde } -\pi \leq \theta \leq \pi \end{cases}$$

Con esta interpretación tendremos:

$$\begin{aligned} x &= \operatorname{Re} z &\iff &\text{componente real del vector } z \text{ o parte real de } z \\ y &= \operatorname{Im} z &\iff &\text{componente imaginaria del vector } z \text{ o parte imaginaria de } z \\ r &= \sqrt{zz^*} = |z| &\iff &\text{módulo, magnitud o valor absoluto de } z \\ \theta & &\iff &\text{ángulo polar o de fase del número complejo } z \end{aligned}$$

La interpretación vectorial de números complejos permite que la suma de números complejos: $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ sea representada por la “regla del paralelogramo”, es decir, en la representación gráfica de la suma es fácil ver que z_1 y z_2 limitan un paralelogramo cuya diagonal es $z_1 + z_2$.

Mientras que los productos escalar y vectorial nos llevan a:

$$z_1 \cdot z_2 = \operatorname{Re}(z_1 z_2^*) = \operatorname{Re}(z_1^* z_2) \quad \wedge \quad z_1 \times z_2 = \operatorname{Im}(z_1^* z_2) = -\operatorname{Im}(z_1 z_2^*) .$$

Volviendo nuevamente a la relación $|z| = \sqrt{zz^*}$, observemos que:

- $|z_1 z_2| = |z_1| |z_2|$.

¹⁸Jean Robert Argand (1768-1822). Contador pero matemático aficionado, propuso esta interpretación de números complejos como vectores en un plano complejo en un libro autoeditado con sus reflexiones que se perdió y fue rescatado 7 años después, fecha a partir de la cual Argand comenzó a publicar en Matemáticas. https://es.wikipedia.org/wiki/Jean-Robert_Argand

¹⁹Casper Wessel (1745-1818) Matemático que se dedicó principalmente al levantamiento topográfico de Noruega. Su trabajo sobre los números complejos permaneció desconocido por casi 100 años. https://es.wikipedia.org/wiki/Casper_Wessel

²⁰Johann Carl Friedrich Gauss (1777-1855). Uno de los matemáticos más geniales y precoces de la Historia. Desde los 7 años comenzó a mostrar sus condiciones de genialidad. Sus contribuciones en Astronomía y Matemáticas son múltiples y diversas. https://es.wikipedia.org/wiki/Carl_Friedrich_Gauss

Ya que si $|z_1 z_2|$, $|z_1|$ y $|z_2|$ son cantidades positivas, entonces:

$$|z_1 z_2|^2 = (z_1 z_2)(z_1 z_2)^* = z_1 z_2 z_1^* z_2^* = z_1 z_1^* z_2 z_2^* = |z_1|^2 |z_2|^2 = (|z_1| |z_2|)^2 .$$

- $|z_1 + z_2| \leq |z_1| + |z_2|$.

Veamos:

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(z_1 + z_2)^* = (z_1 + z_2)(z_1^* + z_2^*) = z_1 z_1^* + z_2 z_2^* + z_1 z_2^* + z_1^* z_2 = \\ &= |z_1|^2 + |z_2|^2 + 2 \operatorname{Re}(z_1 z_2^*) \leq |z_1|^2 + |z_2|^2 + 2 |\operatorname{Re}(z_1 z_2^*)| \\ &\leq |z_1|^2 + |z_2|^2 + 2 |z_1 z_2^*| = |z_1|^2 + |z_2|^2 + 2 |z_1| |z_2| = |z_1|^2 + |z_2|^2 + 2 |z_1| |z_2| \\ &= (|z_1| + |z_2|)^2 . \end{aligned}$$

Notemos que la igualdad $|z_1 + z_2| = |z_1| + |z_2|$ se satisface si, y sólo si, $\operatorname{Re}(z_1 z_2^*) = |z_1 z_2^*|$.

1.6.3 Producto escalar

El producto escalar definido con anterioridad para vectores puede ser redefinido al caso de vectores con componentes complejas.

Definición 1.1 (Producto escalar complejo)

Dados los vectores: $\mathbf{a} = a^i \hat{\mathbf{i}}_i$ y $\mathbf{b} = b^i \hat{\mathbf{i}}_i$, con $\{a^i\}$ y $\{b^i\} \in \mathbb{C}$, entonces

$$\mathbf{a} \cdot \mathbf{b} = (a_i)^* b^i . \quad (1.32)$$

Notemos que con esta definición, el producto $\mathbf{a} \cdot \mathbf{a} = (a^i)^* a^i$ siempre será un número real. Por lo tanto:

$$|\mathbf{a}| = \sqrt{\mathbf{a} \cdot \mathbf{a}} \in \mathbb{R} . \quad (1.33)$$

Esta definición hace que ahora cambien algunas de las propiedades del producto escalar.

$$\mathbf{a} \cdot \mathbf{b} = (\mathbf{b} \cdot \mathbf{a})^* , \quad (1.34)$$

$$(\lambda \mathbf{a}) \cdot \mathbf{b} = \lambda^* \mathbf{a} \cdot \mathbf{b} , \quad (1.35)$$

$$\mathbf{a} \cdot (\lambda \mathbf{b}) = \lambda \mathbf{a} \cdot \mathbf{b} . \quad (1.36)$$

1.6.4 Fórmulas de Euler y De Moivre

En cursos anteriores, nos hemos tropezado con la expansión en Taylor²¹ de funciones, esta serie permite expresar cualquier función analítica²² alrededor de un punto x_0 como una serie infinita de potencias del argumento de la función, esto es:

$$f(x) = f(x_0) + \frac{df(x)}{dx} \Big|_{x=x_0} (x - x_0) + \frac{1}{2} \frac{d^2 f(x)}{dx^2} \Big|_{x=x_0} (x - x_0)^2 + \frac{1}{3!} \frac{d^3 f(x)}{dx^3} \Big|_{x=x_0} (x - x_0)^3 + \dots \dots$$

O de manera equivalente:

$$f(x) = C_n (x - x_0)^n \quad \text{con } C_n = \frac{1}{n!} \frac{d^n f(x)}{dx^n} \Big|_{x=x_0} \quad \text{y } n = 0, 1, 2, 3, \dots$$

²¹Brook Taylor (1685-1731) Físico y Matemático inglés contemporáneo de Newton y Leibniz y junto con ellos participó profundamente en el desarrollo del cálculo diferencial e integral. Además de sus aportes al estudio del magnetismo, capilaridad y termometría, desarrolló el área de diferencias finitas que hasta hoy utilizamos para cálculos en computación. Inventó la integración por partes y descubrió la serie que lleva su nombre. https://es.wikipedia.org/wiki/Brook_Taylor

²²Básicamente, una función analítica es una función que puede expresarse como una serie de potencias convergente.

Si consideramos que $x_0 = 0$, podremos ver a continuación algunos desarrollos en series de funciones elementales:

$$\begin{aligned} e^x &= 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \frac{1}{24}x^4 + \frac{1}{120}x^5 + \frac{1}{720}x^6 + \frac{1}{5040}x^7 + \dots \\ \cos(x) &= 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + \dots \\ \operatorname{sen}(x) &= x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + \dots \end{aligned}$$

Es fácil convencerse que para la serie de e^x se tiene:

$$e^{i\theta} = 1 + i\theta - \frac{1}{2}\theta^2 - \frac{1}{6}i\theta^3 + \frac{1}{24}\theta^4 + \frac{1}{120}i\theta^5 - \frac{1}{720}\theta^6 - \frac{1}{5040}i\theta^7 + \dots$$

y que puede arreglarse como:

$$e^{i\theta} = \underbrace{\left(1 - \frac{1}{2}\theta^2 + \frac{1}{24}\theta^4 - \frac{1}{720}\theta^6 + \dots\right)}_{\cos(\theta)} + i\underbrace{\left(\theta - \frac{1}{6}\theta^3 + \frac{1}{120}\theta^5 - \frac{1}{5040}\theta^7 + \dots\right)}_{\operatorname{sen}(\theta)}$$

obteniéndose la importante relación:

$$e^{i\theta} = \cos(\theta) + i \operatorname{sen}(\theta), \quad (1.37)$$

conocida como la relación de Euler²³.

La fórmula de Euler, para $-\pi < \theta < \pi$ resulta en los siguientes simpáticos resultados:

$$i = e^{i\frac{\pi}{2}}, \quad -1 = e^{i\pi}, \quad -i = e^{-i\frac{\pi}{2}}, \quad 1 = e^{i2k\pi} \quad \text{con } k = 0, \pm 1, \pm 2, \pm 3 \dots$$

Entonces, tenemos tres formas de representar un número complejo:

$$z = x + iy \quad \Leftarrow \quad z = |z|(\cos(\theta) + i \operatorname{sen}(\theta)) \quad \Leftarrow \quad z = |z|e^{i\theta}.$$

La expresión $z = x + iy$ se conoce como forma cartesiana de representación de un número complejo, la forma $z = r(\cos(\theta) + i \operatorname{sen}(\theta))$ será la forma trigonométrica o polar y la expresión $z = re^{i\theta}$ será la forma de Euler.

Es importante notar una sutileza implícita en esta notación. La forma cartesiana representa únicamente a un número complejo, mientras que la forma polar (y la de Euler), es ambigua, ya que:

$$z = r(\cos(\theta) + i \operatorname{sen}(\theta)) = r(\cos(\theta + 2n\pi) + i \operatorname{sen}(\theta + 2n\pi)), \quad n = 0, \pm 1, \pm 2, \pm 3, \dots$$

Es decir, existen varios valores del argumento que definen el mismo número complejo.

Ejemplo 1.25 El número $z = 1 + \sqrt{3}i$ lo podemos representar de las siguientes maneras:

$$\begin{aligned} z = 1 + \sqrt{3}i &= 2e^{\frac{\pi}{3}i} = 2 \left[\cos\left(\frac{\pi}{3}\right) + i \operatorname{sen}\left(\frac{\pi}{3}\right) \right] \\ &= 2e^{(\frac{\pi}{3}+2n\pi)i} = 2 \left[\cos\left(\frac{\pi}{3} + 2n\pi\right) + i \operatorname{sen}\left(\frac{\pi}{3} + 2n\pi\right) \right]. \end{aligned}$$



Analicemos estas situaciones con más detalle. Primero es importante notar que r es siempre positivo, mientras que el ángulo θ puede tomar infinitos valores incluyendo valores negativos. Como ya vimos, θ se relaciona con las coordenadas cartesianas por la ecuación $\tan(\theta) = y/x$, y donde es necesario especificar el cuadrante para su cálculo. A cada valor del ángulo se le denomina el *argumento de z* .

Por lo tanto, si tenemos un número complejo z denotaremos por $\arg(z)$ a:

$$\arg(z) = \{\theta \in \mathbb{R} / |z|(\cos(\theta) + i \operatorname{sen}(\theta))\}.$$

²³Leonhard Euler (1707-1783). Uno de los matemáticos más prolíficos de todos los tiempos. Desarrolló inmensamente campos como la geometría analítica y trigonometría, siendo el primero que consideró el coseno y el seno como funciones. Hizo aportes significativos en el desarrollo del cálculo diferencial e integral, astronomía, elasticidad y mecánica de medios continuos. https://es.wikipedia.org/wiki/Leonhard_Euler

Esto significa que dos números complejos z_1 y z_2 serán iguales si: $|z_1| = |z_2|$ y $\arg(z_1) = \arg(z_2)$.

Dado un número complejo, $z \neq 0$, existe un único argumento Θ que se encuentra en el intervalo $[-\pi, \pi]$, y lo representaremos por $\text{Arg}(z)$, el *argumento principal* o *valor principal*, de manera que:

$$\arg(z) = \text{Arg}(z) + 2n\pi, \quad n = 0, \pm 1, \pm 2, \pm 3, \dots$$

Otro aspecto a rescatar de la forma de Euler o exponencial, es que la ecuación:

$$z = re^{i\theta}, \quad 0 \leq \theta \leq 2\pi,$$

no es más que una representación paramétrica del círculo $|z| = r$, es decir, de un círculo de radio r y centrado en el origen.

Por lo tanto, un círculo centrado en z_0 y de radio R tendrá como ecuación paramétrica:

$$|z - z_0| = R \Rightarrow z = z_0 + Re^{i\theta},$$

con $0 \leq \theta \leq 2\pi$. Como podemos apreciar en la figura 1.11.

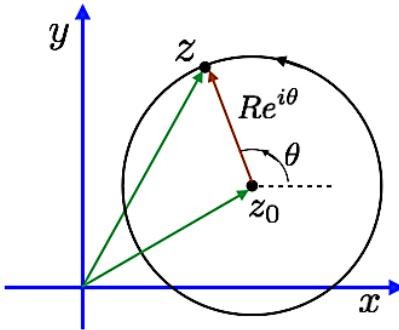


Figura 1.11: Representación en el plano complejo del círculo $|z - z_0| = R$.

Ejemplo 1.26 Un número complejo, digamos $z = -1 - i$ (tercer cuadrante) tendrá como valor principal:

$$\tan(\Theta) = \frac{y}{x} = 1 \Rightarrow \text{Arg}(-1 - i) = -\frac{3\pi}{4}.$$

Nota: no debe tomarse el valor $-\frac{5\pi}{4}$, ya que $-\pi \leq \Theta \leq \pi$.

Lo que si es cierto es que:

$$\arg(-1 - i) = -\frac{3\pi}{4} + 2n\pi, \quad n = 0, \pm 1, \pm 2, \pm 3, \dots$$



Es claro que las sumas de números complejos se plantean más fácilmente en su forma cartesiana. Mientras la multiplicación y división serán directas en la forma de Euler. Si $z_1 = |z_1|e^{i\theta_1}$ y $z_2 = |z_2|e^{i\theta_2}$, entonces:

$$z_1 z_2 = |z_1|e^{i\theta_1}|z_2|e^{i\theta_2} = |z_1||z_2|e^{i(\theta_1+\theta_2)} = |z_1 z_2|(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)). \quad (1.38)$$

Esto significa que para multiplicar dos números complejos se debe, por un lado, multiplicar sus módulos y por el otro, sumar sus argumentos. Geométricamente, al sumarse los argumentos, la multiplicación es en realidad un giro en el plano complejo por el producto de sus módulos, es decir, si multiplicamos un número por i , el resultado es un giro de un cuarto de vuelta hacia la izquierda, es por eso, que $i^2 = -1 = e^{i\pi}$.

Mientras que para la división:

$$\frac{z_1}{z_2} = \frac{|z_1|e^{i\theta_1}}{|z_2|e^{i\theta_2}} = \frac{|z_1|}{|z_2|}e^{i(\theta_1-\theta_2)} = \frac{|z_1|}{|z_2|}(\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)).$$

Podemos notar de esta última expresión que el inverso de un número complejo diferente de cero es:

$$z^{-1} = \frac{1}{z} = \frac{1}{|z|}e^{-i\theta}. \quad (1.39)$$

Se puede mostrar que a partir de (1.38) y de (1.39) resulta que:

$$\arg(z_1 z_2) = \arg(z_1) + \arg(z_2),$$

$$\arg\left(\frac{z_1}{z_2}\right) = \arg(z_1) - \arg(z_2).$$

Ejemplo 1.27 Vimos en el ejemplo 1.26 que para el número complejo $z = -1 - i$, resultaba que

$$\arg(-1 - i) = -\frac{3\pi}{4} + 2n\pi \quad n = 0, \pm 1, \pm 2, \pm 3, \dots$$

Si lo queremos escribir en la forma exponencial tenemos que hacer lo siguiente:

$$|z| = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} \Rightarrow -1 - i = \sqrt{2}e^{-i(\frac{3\pi}{4})}.$$

Pero en realidad hay infinitas posibilidades para la forma exponencial de $z = -1 - i$:

$$-1 - i = \sqrt{2}e^{i(-\frac{3\pi}{4}+2n\pi)}, \quad n = 0, \pm 1, \pm 2, \dots$$



Ejemplo 1.28 Consideremos los números $z = 2 + 2i$ y $z = -\sqrt{3} + i$ y representémoslos en la forma polar.

Para representar $z = 2 + 2i$ en la forma polar debemos calcular primeramente su módulo:

$$|z| = \sqrt{x^2 + y^2} = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2}.$$

y luego el argumento:

$$\tan(\theta) = \frac{y}{x} = \frac{2}{2} \Rightarrow \theta = \arctan(1) = \frac{\pi}{4}.$$

Por lo tanto, $z = 2 + 2i = 2\sqrt{2}(\cos(\pi/4) + i \sin(\pi/4))$, es un punto ubicado en el primer cuadrante cuyo radio vector hace un ángulo de 45° con respecto al eje x .

En cambio, para el número complejo $z = -\sqrt{3} + i$, resulta:

$$|z| = \sqrt{(-\sqrt{3})^2 + 1^2} = \sqrt{4} = 2, \quad \theta = \arctan\left(-\frac{1}{\sqrt{3}}\right) = -\frac{\pi}{6}.$$

Aquí debemos tener cuidado, pues el argumento principal en realidad es:

$$\theta = \pi - \frac{\pi}{6} = \frac{5\pi}{6}.$$

$z = -\sqrt{3} + i = 2(\cos(5\pi/6) + i \sin(5\pi/6))$, es un punto ubicado en el segundo cuadrante cuyo radio vector hace un ángulo de 150° con respecto al eje x .



Por otro lado, notemos lo siguiente: si $z = x + iy$, entonces

$$e^z = e^{(x+iy)} = e^x e^{iy} = e^x (\cos(y) + i \sin(y)),$$

y a partir de la relación o fórmula de Euler se puede demostrar

$$z^n = |z|^n \left(e^{i\theta}\right)^n = |z|^n e^{in\theta} \Rightarrow |z|^n (\cos(\theta) + i \sin(\theta))^n = |z|^n (\cos(n\theta) + i \sin(n\theta)),$$

con n entero. De manera que llegamos a la fórmula de De Moivre²⁴

$$[\cos(\theta) + i \sin(\theta)]^n = \cos(n\theta) + i \sin(n\theta), \quad \text{con } n \text{ entero.} \quad (1.40)$$

1.6.5 Algunas aplicaciones inmediatas

Presentaremos algunas aplicaciones inmediatas la fórmula de De Moivre en diferentes ámbitos, algunas de estas aplicaciones seguramente las recordamos de nuestra educación media.

1. Identidades trigonométricas

La primera de las aplicaciones de la fórmula de De Moivre es para construir identidades trigonométricas en las cuales se expresa el coseno, o el seno, de factores de un ángulo. Veamos las siguientes (nada

²⁴Abraham De Moivre (1667-1754) Matemático francés que tuvo que emigrar a Inglaterra por razones religiosas. Contemporáneo de Newton, Leibniz y Halley, fue pionero con sus contribuciones en geometría analítica y teoría de probabilidades. https://es.wikipedia.org/wiki/Abraham_de_Moivre

triviales) identidades trigonométricas:

$$\cos(3\theta) = 4\cos^3(\theta) - 3\cos(\theta) \quad \text{o} \quad \sin(3\theta) = 3\sin(\theta) - 4\sin^3(\theta).$$

Para demostrar estas (y otras) identidades utilizamos la fórmula de De Moivre, es decir:

$$\begin{aligned}\cos(3\theta) + i\sin(3\theta) &= [\cos(\theta) + i\sin(\theta)]^3 \\ &= \cos^3(\theta) - 3\cos(\theta)\sin^2(\theta) + i[3\cos^2(\theta)\sin(\theta) - \sin^3(\theta)].\end{aligned}$$

Igualando ahora parte real e imaginaria tendremos:

$$\begin{aligned}\cos(3\theta) &= \cos^3(\theta) - 3\cos(\theta)\sin^2(\theta) = \cos^3(\theta) - 3\cos(\theta)[1 - \cos^2(\theta)] = 4\cos^3(\theta) - 3\cos(\theta), \\ \sin(3\theta) &= 3\cos^2(\theta)\sin(\theta) - \sin^3(\theta) = 3[1 - \sin^2(\theta)]\sin(\theta) - \sin^3(\theta) = 3\sin(\theta) - 4\sin^3(\theta).\end{aligned}$$

El método puede extenderse a expresiones de senos y cosenos de $n\theta$.

Igualmente podemos desarrollar un método para encontrar expresiones de potencias de funciones trigonométricas en término de funciones de factores de ángulo del tipo $(\cos(\theta))^n = F(\cos(n\theta), \sin(n\theta))$.

Para empezar, supongamos que tenemos un número complejo de módulo 1, de tal forma que

$$z = e^{i\theta} = \cos(\theta) + i\sin(\theta) \Rightarrow \begin{cases} z^n + \frac{1}{z^n} = 2\cos(n\theta) \\ z^n - \frac{1}{z^n} = 2i\sin(n\theta) \end{cases}$$

Estas identidades surgen de manera inmediata a partir de lo siguiente

$$\begin{aligned}z^n + \frac{1}{z^n} &= [\cos(\theta) + i\sin(\theta)]^n + [\cos(\theta) + i\sin(\theta)]^{-n} \\ &= [\cos(n\theta) + i\sin(n\theta)] + [\cos(-n\theta) + i\sin(-n\theta)] \\ &= \cos(n\theta) + i\sin(n\theta) + \cos(n\theta) - i\sin(n\theta) = 2\cos(n\theta),\end{aligned}$$

igualmente puede demostrarse la segunda de las afirmaciones anteriores.

Supongamos además que $n = 1$, con lo cual se cumple que:

$$z + \frac{1}{z} = e^{i\theta} + e^{-i\theta} = 2\cos(\theta) \quad \text{y} \quad z - \frac{1}{z} = e^{i\theta} - e^{-i\theta} = 2i\sin(\theta),$$

que también lo sabíamos desde la más temprana edad de nuestros cursos de bachillerato.

Ahora bien, lo que quizás no sabíamos en ese entonces (y quizás ahora tampoco) es que a partir de aquí podemos construir, expresiones como:

$$\cos(\theta) = \frac{1}{2} \left(z + \frac{1}{z} \right) \Rightarrow \cos^5(\theta) = \frac{1}{2^5} \left(z + \frac{1}{z} \right)^5 = \frac{1}{2^5} \left[\left(z^5 + \frac{1}{z^5} \right) + \left(5z^3 + \frac{5}{z^3} \right) + \left(10z + \frac{10}{z} \right) \right],$$

es decir:

$$\cos^5(\theta) = \frac{1}{2^5} [2\cos(5\theta) + 10\cos(3\theta) + 20\cos(\theta)].$$

De la misma manera se puede proceder con otras potencias y con potencias de la función seno.

2. Raíces de polinomios

Las raíces de un número complejo se obtienen de la relación:

$$z^{1/n} = [|z|(\cos(\theta) + i\sin(\theta))]^{1/n} = |z|^{1/n} \left[\cos\left(\frac{\theta + 2k\pi}{n}\right) + i\sin\left(\frac{\theta + 2k\pi}{n}\right) \right],$$

donde $k = 0, 1, \dots, n-1$. De manera que la fórmula de De Moivre nos puede ayudar para encontrar raíces de polinomios.

Supongamos, para empezar, que queremos encontrar las n raíces de la ecuación:

$$z^n = 1.$$

Para ello procedemos con el siguiente artificio:

$$z^n = 1 = e^{i(2\pi k)} = \cos(2\pi k) + i \sin(2\pi k), \quad \text{donde } k = 0, 1, 2, \dots$$

con lo cual las n raíces de la ecuación $z^n = 1$ serán:

$$z^n = 1 = e^{i(2\pi k)} \Rightarrow z = e^{i(\frac{2\pi k}{n})}, \quad (1.41)$$

esto es:

$$z_0 = 1; \quad z_1 = e^{2\pi i(\frac{1}{n})}; \quad z_2 = e^{2\pi i(\frac{2}{n})}; \quad z_3 = e^{2\pi i(\frac{3}{n})}; \dots \quad z_{n-2} = e^{2\pi i(\frac{n-2}{n})}; \quad z_{n-1} = e^{2\pi i(\frac{n-1}{n})},$$

es decir, n raíces corresponderán a los n valores de $k = 0, 1, 2, \dots, n-2, n-1$. Mayores valores de k no proveen nuevas raíces.

Las raíces de la ecuación $z^3 = 1$ serán entonces:

$$z = e^{i(\frac{2\pi k}{3})} \Rightarrow z_0 = 1, \quad z_1 = e^{i(\frac{2\pi}{3})}, \quad z_2 = e^{i(\frac{4\pi}{3})}.$$

Como veremos más adelante, estas propiedades pueden extenderse a raíces de polinomios que contengan más términos.

Una afirmación que nos han dicho, y que quizás no sepamos de dónde viene, es que *si un polinomio con coeficientes reales tiene raíces complejas, ellas serán complejas conjugadas unas de otras*. Vale decir, si $z^5 - z^4 + 2z - 2 = 0$ tiene como raíz $(1+i)$, también tendrá como raíz $(1-i)$.

Esta afirmación se prueba de forma general si suponemos que tenemos la siguiente ecuación:

$$a_k z^k = 0, \quad \text{con } k = 0, 1, 2, \dots, n-1, n \Rightarrow a_0 + a_1 z + a_2 z^2 + \dots + a_{n-1} z^{n-1} + a_n z^n = 0,$$

donde los coeficientes $a_0, a_1, a_2, \dots, a_{n-1}, a_n$ los suponemos reales, esto es: $a_k = a_k^*$ para todos los valores del índice k .

Al tomar el complejo conjugado nos queda:

$$a_0^* + a_1^* z^* + a_2^* (z^*)^2 + \dots + a_{n-1}^* (z^*)^{n-1} + a_n^* (z^*)^n = 0,$$

y como los coeficientes son reales tenemos que:

$$a_0 + a_1 z^* + a_2 (z^*)^2 + \dots + a_{n-1} (z^*)^{n-1} + a_n (z^*)^n = 0,$$

esto nos dice que si z es solución también lo será z^* ya que la ecuación es la misma por tener los mismos coeficientes (reales).

Ejemplo 1.29 Supongamos la siguiente ecuación polinómica con sus raíces:

$$z^5 - z^4 + 2z - 2 = 0 \Rightarrow (z^4 + 2)(z - 1) = 0 \Rightarrow \begin{cases} z^4 + 2 = 0 \Rightarrow z^4 = -2 \\ z - 1 = 0 \Rightarrow z = 1 \end{cases}$$

De la ecuación (1.41) podemos ver que:

$$z^4 = -2(1) = -2 \left(e^{i(2\pi k)} \right) \Rightarrow z = \left[-2 \left(e^{i(2\pi k)} \right) \right]^{1/4} = (-2)^{1/4} e^{i(\frac{2\pi k}{4})} = \frac{2^{3/4}}{2} (1+i) e^{i(\frac{2\pi k}{4})},$$

donde hemos utilizado el hecho de que: $(-1)^{1/4} = i^{1/2} = \left(e^{i\frac{\pi}{2}} \right)^{1/2} = e^{i\frac{\pi}{4}} = \frac{\sqrt{2}}{2}(1+i)$.

Por lo tanto:

$$\begin{aligned} z_0 &= \frac{1}{2^{1/4}} (1+i) \\ z_1 &= \frac{1}{2^{1/4}} (1+i) e^{i(\frac{\pi}{2})} = \frac{i}{2^{1/4}} (1+i) \\ z_2 &= \frac{1}{2^{1/4}} (1+i) e^{i(\pi)} = -\frac{1}{2^{1/4}} (1+i) \\ z_3 &= \frac{1}{2^{1/4}} (1+i) e^{i(\frac{3\pi}{2})} = -\frac{i}{2^{1/4}} (1+i). \end{aligned}$$

Entonces, la ecuación $z^5 - z^4 + 2z - 2 = 0$, tendrá las siguientes cinco raíces:

$$z_0 = \frac{1}{2^{1/4}} (1+i), \quad z_1 = -\frac{1}{2^{1/4}} (1-i), \quad z_2 = -\frac{1}{2^{1/4}} (1+i), \quad z_3 = \frac{1}{2^{1/4}} (1-i), \quad z_4 = 1.$$



Ejemplo 1.30 Ahora consideremos el siguiente polinomio complejo:

$$P(z) = z^6 - z^5 + 4z^4 - 6z^3 + 2z^2 - 8z + 8 = 0.$$

Encontremos sus raíces.

Sí por algún método comprobamos que $(z^3 - 2)$ es uno de sus factores, entonces podremos encontrar las raíces del polinomio $P(z)$.

Veamos, claramente si $(z^3 - 2)$ es un factor, entonces podemos expresar:

$$P(z) = z^6 - z^5 + 4z^4 - 6z^3 + 2z^2 - 8z + 8 = (z^3 - 2)(z^3 - z^2 + 4z - 4) = (z^3 - 2)(z - 1)(z^2 + 4),$$

con lo cual, como z es complejo, hay que tener cuidado con las raíces encubiertas. Entonces, las raíces son:

$$z^3 = 2, \quad z = 1, \quad z^2 = -4.$$

- Para $z^2 = -4 \Rightarrow z = \pm 2i$.
- Para $z^3 = 2 \Rightarrow z^3 = 2(e^{i(2\pi k)}) \Rightarrow z = [2(e^{i(2\pi k)})]^{1/3} = 2^{1/3}e^{i(\frac{2\pi k}{3})}$.

Por lo tanto:

$$z_0 = 2^{1/3}, \quad z_1 = 2^{1/3}e^{i(\frac{2\pi}{3})} = -\frac{2^{1/3}}{2}[1 - \sqrt{3}i], \quad z_2 = 2^{1/3}e^{i(\frac{4\pi}{3})} = -\frac{2^{1/3}}{2}[1 + \sqrt{3}i].$$

La ecuación: $z^6 - z^5 + 4z^4 - 6z^3 + 2z^2 - 8z + 8 = 0$, tendrá las siguientes seis raíces:

$$z = \sqrt[3]{2}, \quad z = -\frac{1}{\sqrt[3]{4}}[1 \pm \sqrt{3}i], \quad z = 1, \quad z = \pm 2i.$$



1.6.5.1 Logaritmos y potencias de números complejos

La motivación surge cuando queremos resolver la ecuación:

$$e^w = z = re^{i\Theta}, \quad \text{con } w, z \in \mathbb{C} \quad y \quad -\pi < \Theta < \pi. \quad (1.42)$$

Notemos que al despejar w en realidad lo que tenemos es la función logarítmica, que como veremos en su debido tiempo puede escribirse de la forma $w = u + iv$. Por lo tanto:

$$e^{u+iv} = e^u e^{iv} = re^{i\Theta} \Rightarrow e^u = r \wedge v = \Theta + 2\pi n.$$

donde n es un entero. Por lo tanto, es claro que:

$$e^u = r \Rightarrow u = \ln(r),$$

y que la ecuación (1.42) se satisface si y sólo si:

$$w = \ln(r) + i(\Theta + 2\pi n), \quad \text{con } n = 0, \pm 1, \pm 2, \pm 3, \dots$$

Por lo tanto, si definimos la función multivaluada:

$$\text{Log}(z) \equiv \ln |r| + i(\Theta + 2\pi n), \quad \text{con } n = 0, \pm 1, \pm 2, \pm 3, \dots \quad (1.43)$$

podemos escribir la relación:

$$e^{\text{Log}(z)} = z \quad \text{con } z \neq 0.$$

Llamaremos *valor principal* de $\text{Log}(z)$ al valor que se obtiene cuando $n = 0$ en la ecuación (1.43) y lo denotaremos con $\text{Ln}(z)$.

$$\text{Ln}(z) = \ln |r| + i\Theta. \quad (1.44)$$

Notemos que ésta es una función univaluada cuando $z \neq 0$, y además, si combinamos (1.43) y (1.44) obtenemos:

$$\text{Log}(z) = \text{Ln}(z) + i2\pi n, \quad \text{con } n = 0, \pm 1, \pm 2, \pm 3, \dots$$

Podemos ver que cuando z es un número real positivo, es decir, $z = re^{i0}$, entonces recobramos la función logarítmica usual:

$$\text{Ln}(z) = \text{Ln}(r) = \ln(r).$$

En consecuencia, podemos ver que:

$$\begin{aligned} \text{Log}(1) &= \ln(1) + i(0 + 2\pi n) \quad \wedge \quad \text{Ln}(1) = 0 \\ \text{Log}(-1) &= \ln(1) + i(\pi + 2\pi n) \quad \wedge \quad \text{Ln}(-1) = i\pi \end{aligned}$$

con $n = 0, \pm 1, \pm 2, \pm 3, \dots$

Por otro lado, podemos ver también:

$$\begin{aligned} \ln(z^n) &= \ln[|z|(\cos(\theta) + i\sin(\theta))^n] = n\ln[|z|(\cos(\theta) + i\sin(\theta))] = n\ln[|z|e^{i\theta}] \\ &= n\ln(|z|) + n(\theta + 2k\pi)i = n\ln(|z|) + (n\theta)i + (2nk\pi)i, \quad \text{con } k \text{ entero}. \end{aligned}$$

Ejemplo 1.31 Consideremos algunos ejemplos puntuales:

1. $z = \text{Log}(-3i)$

$$\text{Log}(-3i) = \text{Log}\left[3e^{i(-\frac{\pi}{2}+2n\pi)}\right] = \ln(3) + i\left(-\frac{\pi}{2} + 2n\pi\right) \quad \text{con } n = 0, 1, 2, \dots$$

Decimos que el valor principal de $\text{Log}(-3i)$ será $\text{Ln} = \ln(3) - i\frac{\pi}{2}$.

2. $z = i^{-5i}$

$$i^{-5i} \Rightarrow \text{Log}(z) = \text{Log}(i^{-5i}) = -5i\text{Log}(i) = -5i\text{Log}\left[e^{i(\frac{\pi}{2}+2n\pi)}\right] = 5\left(\frac{\pi}{2} + 2n\pi\right).$$

Con lo cual $z = i^{-5i}$ ¡es un número real!

3. $z = i^i$

$$i^i = \left[e^{i(\frac{\pi}{2}+2n\pi)}\right]^i = e^{i^2(\frac{\pi}{2}+2n\pi)} = e^{-(\frac{\pi}{2}+2n\pi)}.$$

4. $z = \text{Log}$

$$\begin{aligned} \text{Log}\left[\left(\sqrt{3} + i\right)^3\right] &= 3\text{Log}\left[2e^{i\left(\arctan\left(\frac{1}{\sqrt{3}}\right)\right)}\right] = 3\left[\ln(2) + i\left(\arctan\left(\frac{1}{\sqrt{3}}\right) + 2n\pi\right)\right] \\ &= \ln(8) + i\left(\frac{\pi}{2} + 6n\pi\right). \end{aligned}$$



1.6.6 Ejercicios

1. Si los números complejos $z_1 = x_1 + iy_1$ y $z_2 = x_2 + iy_2$ se pueden representar como vectores en el plano $\mathbf{z}_1 = x_1\hat{\mathbf{i}} + y_1\hat{\mathbf{j}}$ y $\mathbf{z}_2 = x_2\hat{\mathbf{i}} + y_2\hat{\mathbf{j}}$, muestre que:

$$z_1^* z_2 = \mathbf{z}_1 \cdot \mathbf{z}_2 + i\hat{\mathbf{k}} \cdot (\mathbf{z}_1 \times \mathbf{z}_2).$$

2. Demuestre:

- (a). $\cos(3\alpha) = \cos^3(\alpha) - 3\cos(\alpha)\sin^2(\alpha)$.
 (b). $\sin(3\alpha) = 3\cos^2(\alpha)\sin(\alpha) - \sin^3(\alpha)$.

3. Demuestre:

- (a). $\cos^4(\alpha) = \frac{1}{8}(3 + 4\cos(2\alpha) + \cos(4\alpha))$.
 (b). $\cos^3(\alpha) + \sin^3(\alpha) = \frac{1}{4}(\cos(3\alpha) + 3\cos(\alpha) - \sin^3(\alpha) + 3\sin(\alpha))$.

4. Demuestre:

$$\left(\frac{ix-1}{ix+1}\right)^{iy} = e^{(-2y\cot^{-1}(x))},$$

donde x y y son números reales.

5. Encuentre las raíces de:

$$a) \quad (2i)^{1/2} \quad b) \quad \left(1 - \sqrt{3}i\right)^{1/2} \quad c) \quad (-1)^{1/3} \quad d) \quad 8^{1/6} \quad e) \quad (-8 - 8\sqrt{3}i)^{1/4}.$$

6. Demuestre que:

- (a). $\text{Log}(-ie) = 1 - \frac{\pi}{2}i$.
 (b). $\text{Log}(1 - i) = \frac{1}{2}\ln(2) - \frac{\pi}{4}i$.
 (c). $\text{Log}(e) = 1 + 2n\pi i$.
 (d). $\text{Log}(i) = \left(2n + \frac{1}{2}\right)\pi i$.

7. Dos funciones complejas $f_1(t)$ y $f_2(t)$ cumplen con las siguientes ecuaciones:

$$\frac{df_1^*}{dt} = \frac{-i}{f_1 - f_2} \quad \text{y} \quad \frac{df_2^*}{dt} = \frac{-i}{f_2 - f_1}$$

Muestre que las siguientes cantidades son constantes:

- (a). $f_1 + f_2$
 (b). $|f_1 - f_2|$
 (c). $|f_1|^2 + |f_2|^2$

8. Considere la siguiente ecuación:

$$z^7 - 4z^6 + 6z^5 - 6z^4 + 6z^3 - 12z^2 + 8z + 4 = 0.$$

Encuentre sus raíces sabiendo que $z^3 = 2$.

9. Muestre que la expansión binomial puede ser escrita como:

$$(1+x)^n = \sum_{m=0}^n A_m(n) x^m, \quad \text{con} \quad A_m(n) = \frac{n!}{m!(n-m)!}.$$

Si está convencido de la expansión anterior, considere ahora una parecida: $(1 + e^{i\theta})^n$ y muestre que:

$$\sum_{m=0}^n A_m(n) \cos(n\theta) = 2^n \cos^n\left(\frac{\theta}{2}\right) \cos\left(\frac{n\theta}{2}\right),$$

y

$$\sum_{m=0}^n A_m(n) \sin(n\theta) = 2^n \cos^n\left(\frac{\theta}{2}\right) \sin\left(\frac{n\theta}{2}\right).$$

10. Las funciones hiperbólicas se definen como:

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \quad \text{y} \quad \operatorname{senh}(x) = \frac{e^x - e^{-x}}{2},$$

y de manera análoga a las funciones trigonométricas tendremos el resto de funciones:

$$\tanh(x) = \frac{\operatorname{senh}(x)}{\cosh(x)}; \quad \operatorname{sech}(x) = \frac{1}{\cosh(x)}; \quad \operatorname{csech}(x) = \frac{1}{\operatorname{senh}(x)}; \quad \operatorname{ctanh}(x) = \frac{1}{\tanh(x)};$$

(a). Muestre las siguientes equivalencias:

$$\cosh(x) = \cos(ix), \quad i \operatorname{senh}(x) = \operatorname{sen}(ix), \quad \cos(x) = \cosh(ix) \quad \text{y} \quad i \operatorname{sen}(x) = \operatorname{senh}(x).$$

(b). Muestre las siguientes identidades:

$$\cosh^2(x) - \operatorname{senh}^2(x) = 1; \quad \operatorname{sech}^2(x) = 1 - \tanh^2(x); \quad \cosh(2x) = \cosh^2(x) + \operatorname{senh}^2(x).$$

(c). Resuelva las siguientes ecuaciones hiperbólicas:

$$\cosh(x) - 5 \operatorname{senh}(x) - 5 = 0, \quad 2 \cosh(4x) - 8 \cosh(2x) + 5 = 0 \quad \text{y} \quad \cosh(x) = \operatorname{senh}(x) + 2 \operatorname{sech}(x).$$

11. Utilizando un programa de manipulación simbólica (ver Apéndice ??) realice las siguientes tareas.

- (a). Calcule la función $f(z) = e^z$ a partir de su expansión en serie que la define. Calcule también $f(z)$ cuando $z = e^{\frac{in\pi}{6}}$ para $n = 0, 1, 2, \dots, 12$. Para los diferentes valores de n haga una tabla con los valores de: $\theta = \frac{n\pi}{6}$, $\operatorname{Re}(z)$, $\operatorname{Im}(z)$, $\operatorname{Re}(e^z)$, $\operatorname{Im}(e^z)$, $|z|$ y la fase de e^z .
- (b). Calcule y haga una tabla para los valores de $(x; y) = (0, 0; 0, 0)(0, 1; 0, 1)(0, 5; 0, 5)(1, 0; 1, 0)$ de: $\operatorname{Re}(\operatorname{senh}(z))$, $\operatorname{Im}(\operatorname{senh}(z))$, $|\operatorname{senh}(z)|$ y la fase de $\operatorname{senh}(z)$.
- (c). Calcule y haga una tabla para los valores de $(x; y) = (0, 0; 0, 0)(0, 1; 0, 1)(0, 5; 0, 5)(1, 0; 1, 0)$ de: $\operatorname{Re}(\cosh(z))$, $\operatorname{Im}(\cosh(z))$, $|\cosh(z)|$ y la fase de $\cosh(z)$.

Practicando con Maxima

Maxima maneja números complejos escritos en la forma: $a+bi$. Donde la unidad imaginaria es interpretada por el programa como $i = \%i$. Si no queremos que las salidas del programa mantenga la notación del $\%i$ podemos ir a la parte superior de la ventana del programa y hacer click en:

wxMaxima > Preferencias > Worksheet

y desactivar el botón que dice

[] Mantener signo porcentual en símbolos especiales: $\%\mathrm{e}$, $\%\mathrm{i}$, etc.

1. Algunos cálculos básicos.

Los números imaginarios aparecen si queremos calcular la raíz cuadrada de un número negativo, por ejemplo, $\sqrt{-7}$

(%i1) $\sqrt{-7};$
 (%o1) $\sqrt{7}i$

El programa nos permite desarrollar toda el álgebra en variable compleja. Si queremos sumar $z_1 = 1 + 2i$ y $z_2 = 3 + 4i$, escribimos:

(%i2) $z1:1+2*\mathrm{i}; z2:3+4*\mathrm{i};$
 (%o2) $2i + 1$
 (%o3) $4i + 3$

(%i4) $z1+z2;$
 (%o4) $6i + 4$

La multiplicación:

```
(%ci5) z1*z2,expand;
(%o5) 10 i - 5
```

Y la división:

```
(%ci6) z1/z2;
(%o6)  $\frac{2i + 1}{4i + 3}$ 
(%ci7) rectform(%);
(%o7)  $\frac{2i}{25} + \frac{11}{25}$ 
```

Este último comando pertenece a una lista de funciones de **Maxima** para el cálculo de números complejos:

- rectform(expresión) expresión en forma cartesiana o binómica
- realpart(expresión) parte real de expresión
- imagpart(expresión) parte imaginaria de expresión
- polarform(expresión) forma polar de expresión
- abs(expresión) módulo o valor absoluto de expresión
- cabs(expresión) módulo de expresión compleja
- carg(expresión) argumento de expresión
- conjugate(expresión) conjugado de expresión
- demoivre(expresión) expresa el número complejo utilizando senos y cosenos
- exponentialize(expresión) expresa el número complejo utilizando exponentiales

Veamos como funcionan algunas de estos comandos:

```
(%ci8) realpart(z1); imagpart(z1);
(%o8) 1
(%o9) 2
```

```
(%ci10) abs(z1);
(%o10)  $\sqrt{5}$ 
```

```
(%ci11) polarform(z1);
(%o11)  $\sqrt{5} e^{i \arctan(2)}$ 
```

```
(%ci12) cabs(z1);carg(z1);
(%o12)  $\sqrt{5}$ 
(%o13) arctan(2)
```

```
(%ci14) conjugate(z1);
(%o14)  $1 - 2i$ 
```

2. Para una función, digamos $\tan(x + iy)$, la podemos escribir como exponentiales o funciones senos y cosenos.

```
(%ci15) tan(x+ %i*y);
(%o15)  $\tan(iy + x)$ 
```

```
(%ci16) exponentialize(%),factor;
(%o16)  $\frac{i(e^y - e^{ix})(e^y + e^{ix})}{e^{2y} + e^{2ix}}$ 
```

(%ci17) demoivre(%),factor;

$$(\%o17) \frac{i (e^y - i \sin(x) - \cos(x)) (e^y + i \sin(x) + \cos(x))}{e^{2y} + i \sin(2x) + \cos(2x)}$$

(%ci18) log(3+ %i*2);

$$(\%o18) \log(2i + 3)$$

(%ci19) cabs(%);

$$(\%o19) \sqrt{\frac{\log^2 13}{4} + \arctan^2\left(\frac{2}{3}\right)}$$

(%ci20) float(%);

$$(\%o20) 1,410846683153171$$

3. Tomemos dos números complejos, o vectores del plano complejo:

(%ci21) z:a+b* %i; w:c+d* %i;

$$(\%o21) ib + a$$

$$(\%o22) id + c$$

Revisemos la llamada igualdad del paralelogramo: la suma de los cuadrados de las longitudes de los cuatro lados de un paralelogramo, es igual a la suma de los cuadrados de las longitudes de las dos diagonales de éste. Es decir:

$$|z + w|^2 + |z - w|^2 = 2(|z|^2 + |w|^2).$$

Pero sabemos que $|z|^2 = zz^*$, y por lo tanto:

$$(w + z)(w + z)^* + (w - z)(w - z)^* = 2(|z|^2 + |w|^2).$$

Calculemos el lado izquierdo de la ecuación anterior:

(%ci23) (w+z)*conjugate(w+z)+(w-z)*conjugate(w-z),expand;

$$(\%o23) 2d^2 + 2c^2 + 2b^2 + 2a^2$$

Ahora el lado derecho, es decir, la suma de las longitudes al cuadrado de los cuatro lados:

(%ci24) 2*cabs(z)^ 2+2*cabs(w)^ 2,expand;

$$(\%o24) 2d^2 + 2c^2 + 2b^2 + 2a^2$$

4. Para encontrar las raíces de un número complejo, debemos declarar a la variable como compleja.

(%ci25) declare(z,complex)\$

De manera que si queremos encontrar las raíces de $z^3 = 1$

(%ci26) ec1:z^ 3=1;

$$(\%o26) z^3 = 1$$

(%ci27) solve(ec1,z);

$$(\%o27) \left[z = \frac{\sqrt{3}i - 1}{2}, z = -\frac{\sqrt{3}i + 1}{2}, z = 1 \right]$$

5. En otro de los ejemplo que discutimos anteriormente teníamos la ecuación:

$$z^6 - z^5 + 4z^4 - 6z^3 + 2z^2 - 8z + 8 = 0,$$

por lo tanto:

(%ci28) $\text{ec2:z}^6-z^5+4\text{z}^4-6\text{z}^3+2\text{z}^2-8\text{z}+8=0;$
(%o28) $z^6 - z^5 + 4z^4 - 6z^3 + 2z^2 - 8z + 8 = 0$

(%ci29) $\text{factor}(\text{ec2});$
(%o29) $(z - 1) (z^2 + 4) (z^3 - 2) = 0$

(%ci30) $\text{solve}(\text{ec2}, \text{z}), \text{factor};$
(%o30) $\left[z = \frac{\sqrt{3}i - 1}{2^{\frac{2}{3}}}, z = -\frac{\sqrt{3}i + 1}{2^{\frac{2}{3}}}, z = 2^{\frac{1}{3}}, z = 1, z = -2i, z = 2i \right]$

Bibliografía

- [1] Arfken, G. B., Weber, H. J., & Harris, F. E. (2011). **Mathematical methods for physicists**. Academic press.
- [2] Riley, K. F., Hobson, M. P., & Bence, S. J. (2006). **Mathematical methods for physics and engineering: a comprehensive guide**. Cambridge University Press.
- [3] Santaló, L.A (1969) **Vectores y Tensores**. Editorial Universitaria, Buenos Aires).
- [4] Spiegel, M. (1959) **Vector Analysis** (*Schaums Outline Series, McGraw Hill New York*).
- [5] Brown, J. W., Churchill, R. V. (2004) **Complex Variables and Applications** (*McGraw Hill*).

Capítulo 2

Espacios vectoriales lineales

La ruta de este capítulo

Este es el capítulo central de esta obra, en el cual discutiremos los conceptos fundamentales sobre espacios vectoriales abstractos, sus propiedades y sus múltiples expresiones: vectores de \mathbb{R}^n , matrices, polinomios, funciones continuas (entre otras). Además, introduciremos la notación de *bra* y *ket* que facilita mucho la organización de los conceptos y, que nos acompañará por el resto de este libro. Haremos una constante referencia a los conceptos que fueron discutidos en el capítulo anterior en el marco de los vectores en \mathbb{R}^3 . Como en el capítulo 1, cada una de las secciones presenta ejemplos con la utilización de la herramienta de cómputo algebraico **Maxima**, que consideramos parte fundamental de estas notas y nos permite mostrar el alcance de los conceptos abstractos.

Para empezar, en la próxima sección, iniciamos con una discusión somera sobre grupos y sus propiedades. Seguimos con la sección 2.1.3, en la cual presentamos el concepto de espacio vectorial lineal para describir varios de sus escenarios que usualmente se presentan desconectados. Hacemos un esfuerzo por ilustrar, bajo un mismo enfoque, su aplicación desde \mathbb{R}^n hasta los espacios vectoriales de funciones continuas $C_{[a,b]}$. Luego, en la sección 2.2.3.1, abordamos los conceptos de distancia (métrica), norma y, finalmente el de producto interno. Esta última definición nos permite equipar los espacios vectoriales, no solo con norma y distancia, sino también dotarlos de geometría, definiendo ángulos entre vectores abstractos. Nos detenemos un momento para discutir el significado de ángulo entre vectores pertenecientes a espacios vectoriales reales y complejos. En la sección 2.3 mostramos el concepto de variedades lineales, discutimos el criterio de independencia lineal y, a partir de éste definimos las bases para los espacios vectoriales. Apoyándonos en el criterio de ortogonalidad construimos bases ortogonales para varios espacios vectoriales, discutimos los subespacios vectoriales ortogonales y las proyecciones ortogonales. El capítulo lo finalizamos mostrando la utilidad de expresar funciones como combinación lineal de vectores ortogonales y, cómo este tipo de expansiones constituye la mejor aproximación a la función.

2.1 Grupos, campos y espacios vectoriales

En los cursos básicos de matemáticas nos enseñaron el concepto de conjunto: una colección de elementos de una misma naturaleza, y aprendimos una variedad de operaciones aplicadas a los elementos que conforman los conjuntos y a los conjuntos mismos. Es probable que en esos tiempos nos quedara la sensación de la poca

utilidad de todos esos conceptos de la teoría de conjuntos, pero como veremos en esta sección, la noción de conjuntos es fundamental para desarrollar todas las ideas que conforman lo que se conoce como el *álgebra abstracta*. En el estudio de las estructuras algebraicas se incluyen las teorías sobre: grupos, anillos, campos, espacios vectoriales, redes, álgebras; que merecen ahora nuestra atención. Comenzaremos con la estructura de grupo, y nos daremos cuenta de que una buena cantidad de objetos matemáticos, en apariencia muy diferentes, tienen incorporadas la estructura de grupo. La noción de grupo no llevará entonces al importante concepto de espacios vectoriales abstractos que discutiremos en la sección 2.1.3 y a la definición de los espacios métricos, sección 2.2.1.

2.1.1 Grupos

Considere el siguiente conjunto no vacío $\mathbf{G} = \{g_1, g_2, g_3, \dots, g_n, \dots\}$ y la operación interna \square (la ley del grupo). Entonces los elementos del conjunto forman un grupo respecto a la operación \square si $\forall g_i \in \mathbf{G}$ se cumplen las siguientes condiciones:

1. Cerrada respecto a la operación \square : $\{g_i \in \mathbf{G}, g_j \in \mathbf{G}\} \Rightarrow \exists g_k = g_i \square g_j \in \mathbf{G}$.
 2. Asociativa respecto a la operación \square : $g_k \square (g_i \square g_j) = (g_k \square g_i) \square g_j$.
 3. Existencia de un elemento neutro: $\exists \hat{g} \in \mathbf{G} \Rightarrow g_i \square \hat{g} = g_i = \hat{g} \square g_i$.
 4. Existencia de un elemento inverso: $g_i \in \mathbf{G} \Rightarrow \exists g_i^{-1} \in \mathbf{G} \Rightarrow g_i \square g_i^{-1} = g_i^{-1} \square g_i = \hat{g}$.
- Si adicionalmente se cumple que:
5. Comutativa respecto a la operación \square : $g_i \square g_j \equiv g_j \square g_i$; el grupo se denomina *grupo abeliano*¹

Suele denotarse al grupo \mathbf{G} como (\mathbf{G}, \square) para indicar el tipo de operación.

2.1.1.1 Algunos grupos:

- Los enteros $\mathbb{Z} = \{\dots - 3 - 2, -1, 0, 1, 2, 3, \dots\}$ respecto a la suma.
- Los racionales $\mathbb{Q} = \{a/b\}$ respecto a la suma y a la multiplicación.
- Los números complejos $\mathbb{C} = \{e^{i\theta}\}$ respecto a la multiplicación.
- Las rotaciones en 2 dimensiones (2D), y las rotaciones en 3D (grupo no-abeliano).
- Las matrices de dimensión $n \times m$ respecto a la suma (grupo abeliano).

Consideremos el siguiente grupo de tres elementos: $\mathbf{G} = \{\mathbf{1}, a, b\}$, con $\hat{g} \equiv \mathbf{1}$ y la operación \square . Por construcción, si queremos que la operación de dos de los elementos provea un tercero distinto, entonces la ÚNICA “tabla de multiplicación” posible será:

\square	$\mathbf{1}$	a	b
$\mathbf{1}$	$\mathbf{1}$	a	b
a	a	b	$\mathbf{1}$
b	b	$\mathbf{1}$	a

Notemos que el grupo más simple será aquel conformado únicamente por el elemento neutro: $\mathbf{G} = \{\mathbf{1}\}$ y, como veremos más adelante, se puede definir subgrupos si un subconjunto de los elementos de un grupo $g_i \in \mathbf{G}$ también forman un grupo.

Podemos resumir, sin demostración, las principales propiedades que presenta un grupo (\mathbf{G}, \square) .

1. El elemento identidad o neutro es único.

¹Niels Henrik Abel (1802-1829) Pionero en el desarrollo de diferentes ramas de la matemática moderna, la teoría de las integrales de funciones algebraicas. https://es.wikipedia.org/wiki/Niels_Henrik_Abel

2. Para todo $g \in G$ existe un inverso único $g^{-1} \in G$.
3. Para todo elemento $g \in G$ se cumple que: $(g^{-1})^{-1} = g$.
4. Para cualesquiera g_1 y $g_2 \in G$, se cumple que: $(g_1 \square g_2)^{-1} = g_1^{-1} \square g_2^{-1}$.

El número de los elementos de un grupo puede ser finito o infinito. En el primer caso se denominan *grupos finitos* y el número de elementos que contenga, el cardinal del conjunto, se conoce como el *orden del grupo*. Un grupo finito que se construye a partir de una operación con un único miembro se denomina *grupo cíclico*, y el caso más elemental es $G = \{1, g, g^2, g^3, \dots, g^{n-1}\}$. Obviamente hemos definido aquí: $g^2 = g \square g$ y $g^3 = g^2 \square g = g \square g \square g$ y así consecutivamente hasta ejecutarse $n - 1$ veces, entonces se retoma el elemento identidad, esto es: $g^{n-1} \square g = g^n = 1$.

2.1.1.2 Subgrupos

Dado un grupo (G, \square) y H un subconjunto de G , $H \subseteq G$. Si H es un grupo bajo la operación \square , definida para G , entonces diremos que H es un subgrupo. De manera equivalente: si $H \subseteq G$ entonces H es un subgrupo si para cualesquiera h_1 y $h_2 \in H$ se cumple:

1. $h_1 \square h_2 \in H$.
2. $h^{-1} \in H \forall h \in H$.

Es decir, $1 \in H$ y H es cerrado para la misma operación que define G y para sus inversos.

Existe una condición necesaria para que un subconjunto de un grupo finito G sea un subgrupo, esta condición se conoce como el teorema de Lagrange, que dice lo siguiente:

Teorema 2.1 (Teorema de Lagrange)

Si (G, \square) es un grupo finito y H un subgrupo de G , entonces el orden de H divide al orden de G .

Por otro lado, sea (G, \square) un grupo y $g \in G$, se puede demostrar que el conjunto $H = \{g^n\}$, con n perteneciente al conjunto de los números enteros, es un subgrupo de G . Con la propiedad adicional de que H es el subgrupo, que contiene a g , más pequeño de G .

Como mencionamos con anterioridad, un grupo finito que se construye a partir de operar con un único miembro se le denomina *grupo cíclico*, en este caso se dice que H es un subgrupo cíclico generado por g . A este elemento se le llama el generador de H y se acostumbra a denotarlo con $H = \langle g \rangle$.

Para finalizar, definiremos el *orden de un elemento* de un grupo. Sea un elemento $g \in G$, el orden de g es el menor entero positivo n tal que: $g^n = 1$. Cuando este entero positivo no existe se dice que g tiene un orden infinito.

Ejemplo 2.1 El conjunto de las rotaciones del espacio ordinario cuando rotamos un ángulo ϕ alrededor del eje z , ecuación (1.13), forman un grupo.

Denominaremos a este grupo por $R_{z\phi} = \{\mathcal{G}_{\phi_i}\}$, con $0 \leq \phi \leq 2\pi$. Donde con \mathcal{G}_ϕ estaremos indicando un giro alrededor del eje z y la operación que nos define las rotaciones la definiremos de la siguiente forma:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_2} = \mathcal{G}_{\phi_1 + \phi_2} .$$

Entonces podemos ver que:

1. \square es una operación cerrada ya que dos rotaciones resulta en otra rotación:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_2} = \mathcal{G}_{\phi_1 + \phi_2} = \mathcal{G}_{\phi_3} \in R_{z\phi} .$$

2. \square es asociativa:

$$(\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_2}) \square \mathcal{G}_{\phi_3} = \mathcal{G}_{\phi_1} \square (\mathcal{G}_{\phi_2} \square \mathcal{G}_{\phi_3}) .$$

3. Existe el elemento neutro, \mathcal{G}_{ϕ_0} , que no hace ninguna rotación:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_0} = \mathcal{G}_{\phi_0} \square \mathcal{G}_{\phi_1} = \mathcal{G}_{\phi_1}.$$

4. Existe el elemento inverso, \mathcal{G}_{ϕ_-} , ya que podemos rotar en un sentido y en sentido inverso:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_{-1}} = \mathcal{G}_{\phi_{-1}} \square \mathcal{G}_{\phi_1} = \mathcal{G}_{\phi_0}.$$

5. Es conmutativa ya que las rotaciones no se afectan por el orden en que son producidas:

$$\mathcal{G}_{\phi_1} \square \mathcal{G}_{\phi_2} = \mathcal{G}_{\phi_2} \square \mathcal{G}_{\phi_1}.$$

▷

Ejemplo 2.2 Consideremos el grupo $\mathbf{G} = \mathbb{Z} \times \mathbb{Z}$, es decir, el conjunto de duplas (x_i, y_i) de números enteros, y la operación:

$$(x_1, y_1) \square (x_2, y_2) = (x_1 + x_2, y_1 + y_2).$$

Aquí el símbolo $+$ es la suma convencional de números enteros.

Probemos que tenemos un grupo.

1. La operación \square es cerrada para la suma:

$$(x_1, y_1) \square (x_2, y_2) = (x_1 + x_2, y_1 + y_2) = (x_3, y_3) \in \mathbf{G}.$$

2. La operación \square es asociativa:

$$\begin{aligned} ((x_1, y_1) \square (x_2, y_2)) \square (x_3, y_3) &= (x_1 + x_2, y_1 + y_2) \square (x_3, y_3) = ((x_1 + x_2) + x_3, (y_1 + y_2) + y_3) \\ &= (x_1 + (x_2 + x_3), y_1 + (y_2 + y_3)) = (x_1, y_1) \square ((x_2, y_2) \square (x_3, y_3)). \end{aligned}$$

3. Existe el elemento neutro, $(0, 0)$:

$$(x_1, y_1) \square (0, 0) = (0, 0) \square (x_1, y_1) = (x_1, y_1).$$

4. Existe el elemento inverso, $(-x_1, -y_1)$:

$$(x_1, y_1) \square (-x_1, -y_1) = (-x_1, -y_1) \square (x_1, y_1) = (0, 0).$$

5. La operación \square es conmutativa:

$$(x_1, y_1) \square (x_2, y_2) = (x_2, y_2) \square (x_1, y_1).$$

▷

Ejemplo 2.3 Consideremos el conjunto de las permutaciones de 3 objetos, cuyos elementos pueden ser representados como se muestra a continuación:

$$\begin{aligned} \mathbb{P}_0 &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}; \quad \mathbb{P}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}; \quad \mathbb{P}_2 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}; \\ \mathbb{P}_3 &= \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}; \quad \mathbb{P}_4 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}; \quad \mathbb{P}_5 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}; \end{aligned}$$

y la operación de composición de permutaciones:

$$\mathbb{P}_1 \odot \mathbb{P}_3 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \odot \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \mathbb{P}_4,$$

es decir, luego de intercambiar las primera y segunda posición, intercambio la segunda y la tercera. La tabla de multiplicación del grupo quedará:

\odot	\mathbb{P}_0	\mathbb{P}_1	\mathbb{P}_2	\mathbb{P}_3	\mathbb{P}_4	\mathbb{P}_5
\mathbb{P}_0	\mathbb{P}_0	\mathbb{P}_1	\mathbb{P}_2	\mathbb{P}_3	\mathbb{P}_4	\mathbb{P}_5
\mathbb{P}_1	\mathbb{P}_1	\mathbb{P}_0	\mathbb{P}_5	\mathbb{P}_4	\mathbb{P}_3	\mathbb{P}_2
\mathbb{P}_2	\mathbb{P}_2	\mathbb{P}_4	\mathbb{P}_0	\mathbb{P}_5	\mathbb{P}_1	\mathbb{P}_3
\mathbb{P}_3	\mathbb{P}_3	\mathbb{P}_5	\mathbb{P}_4	\mathbb{P}_0	\mathbb{P}_2	\mathbb{P}_1
\mathbb{P}_4	\mathbb{P}_4	\mathbb{P}_2	\mathbb{P}_3	\mathbb{P}_1	\mathbb{P}_5	\mathbb{P}_0
\mathbb{P}_5	\mathbb{P}_5	\mathbb{P}_3	\mathbb{P}_1	\mathbb{P}_2	\mathbb{P}_0	\mathbb{P}_4

Ver el ejercicio 3 de 2.1.6.



2.1.1.3 Grupos Isomorfos

Muchas veces podemos comparar la estructura entre grupos buscando relaciones entre ellos, esto lo podemos hacer estudiando, por ejemplo, cómo “se multiplican”. La búsqueda de estas relaciones no llevan al concepto de *isomorfismo* de grupos. Podemos decir que dos grupos son isomorfos porque tienen aspectos en común dentro de sus estructuras algebraicas, es decir, tienen las mismas propiedades. De esta manera, los isomorfismos permiten clasificar los grupos por sus relaciones de equivalencia.

A continuación, veremos algunos ejemplos de grupos y sus tablas de “multiplicación”.

Ejemplo 2.4 Consideremos el siguiente conjunto y una operación que define la “multiplicación”² módulo \times , es decir, $(a \cdot b)_{\text{mod } \times}$

1. $G_{\text{mod}8} = \{1, 3, 5, 7\}$. Tenemos la operación multiplicación módulo 8 y la tabla de multiplicación del grupo es

$\times \text{mod}8$	1	3	5	7
1	1	3	5	7
3	3	1	7	5
5	5	7	1	3
7	7	5	3	1

Parece entender como funciona la tabla anterior, recordemos la regla que estamos utilizando. Por ejemplo, para los elementos 3 y 7, tenemos lo siguiente: $3 \cdot 7 = 21$ y el residuo de dividir 21 entre 8 es 5, por lo tanto

$$(3 \cdot 7)_{\text{mod}8} = (21)_{\text{mod}8} = 5 \iff 21 \div 8 = 2 \times 8 + 5.$$

y se denota: $21 \text{ mod } 8 = 5$.

2. $G_{\text{mod}24} = \{1, 5, 7, 11\}$, con la operación multiplicación módulo 24:

²Vamos a considerar lo que se conoce como *Aritmética Modular*. Sabemos que para la división de dos enteros a y b ($b \neq 0$) existe un único entero c tal que $a = cb + r$ con $0 \leq r \leq |b|$. Y como es conocido: a , b , c y r son: el dividendo, el divisor, el cociente y el resto, respectivamente. Cuando $r = 0$ se dice que a es divisible por b . En esta “aritmética” tan particular se establece una relación de congruencia sobre los enteros: para un entero positivo n , dos enteros a y b se llaman congruentes módulo n ($\text{mod } n$) si a y b tienen el mismo resto r cuando se dividen entre n , y se denota $a \equiv b \pmod{n}$.

$\times \text{mod}24$	1	5	7	11
1	1	5	7	11
5	5	1	11	7
7	7	11	1	5
11	11	7	5	1

Diremos que los grupos $G_{\text{mod}8}$ y $G_{\text{mod}24}$ son *isomorfos* porque tienen tablas equivalentes de multiplicación. Esto es, dado un grupo genérico $G = \{1, A, B, C\}$ o de manera equivalente $\{1, \bigcirc, \square, \triangle\}$ sus tablas de multiplicación serán las que se muestran a continuación:

\times	1	A	B	C
1	1	A	B	C
A	A	1	C	B
B	B	C	1	A
C	C	B	A	1

↔

\times	1	\bigcirc	\square	\triangle
1	1	\bigcirc	\square	\triangle
\bigcirc	\bigcirc	1	\triangle	\square
\square	\square	\triangle	1	\bigcirc
\triangle	\triangle	\square	\bigcirc	1

Note que $A^{-1} = A$, de la misma manera $\triangle = \triangle^{-1}$, es decir, que siempre la operación de dos elementos da uno distinto a los operados. Esto último también se puede representar de la manera siguiente: $A \times A^{-1} = 1$, es decir, el resultado es el elemento neutro del grupo. Tenemos entonces una operación sobre los elementos del grupo que es biyectiva.



Consideremos ahora el siguiente ejemplo con otros dos grupos isomorfos

Ejemplo 2.5

1. $G_{\text{mod}5} = \{1, 2, 3, 4\}$, y la operación multiplicación módulo 5:

$\times \text{mod}5$	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

↔

$\times \text{mod}5$	1	2	4	3
1	1	2	4	3
2	2	4	3	1
4	4	3	1	2
3	3	1	2	4

Note que las tablas anteriores se diferencian por el hecho de que se cambió el orden de los elementos, por lo tanto, son dos tablas de multiplicación diferentes porque los elementos de los grupos están dispuestos de manera diferente.

2. $G_x = \{1, i, -1, -i\}$, y la siguiente operación multiplicación:

\times	1.	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

Por lo tanto, los grupos G_x y $G_{\text{mod}5}$ son isomorfos con una tabla de multiplicación como la que se muestra a continuación:

\times	1	A	B	C
1	1	A	B	C
A	A	B	C	1
B	B	C	1	A
C	C	1	A	B



2.1.2 Campo

Definiremos como un campo (o cuerpo) el conjunto $K = \{\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n, \dots\}$ sobre el cual están definidas dos operaciones: suma (+) y multiplicación (\cdot) y que satisfacen las siguientes propiedades:

1. Forman un grupo abeliano respecto a la suma (+) con el elemento neutro representado por el cero: 0.
2. Forman un grupo abeliano respecto a la multiplicación (\cdot). Se excluye el cero y se denota el elemento neutro de la multiplicación como 1.
3. Es distributiva respecto a la suma (+). Dados α_i, α_j y α_k , se tiene que: $\alpha_i \cdot (\alpha_j + \alpha_k) = \alpha_i \cdot \alpha_j + \alpha_i \cdot \alpha_k$.

Ejemplos típicos de campos lo constituyen el conjunto de los números racionales \mathbb{Q} , los números reales \mathbb{R} y los números complejos \mathbb{C} . Normalmente se refiere estos campos como *campos escalares*.

La noción de campo nos permite introducir el importante concepto de espacio vectorial.

2.1.3 Espacios vectoriales lineales

Sea el conjunto de objetos $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_i\rangle, \dots\}$. Se denominará V un espacio vectorial lineal y sus elementos $|v_i\rangle$ vectores, si existe una operación suma, \boxplus , respecto a la cual los elementos $|v_i\rangle \in V$ forman un grupo abeliano y una operación multiplicación por un elemento de un campo, $K = \{\alpha, \beta, \gamma, \dots\}$, tal que³:

1. V es cerrado bajo la operación \boxplus : $|v_k\rangle = |v_i\rangle \boxplus |v_j\rangle \in V \quad \forall |v_i\rangle, |v_j\rangle \in V$.
2. La operación \boxplus es commutativa: $|v_i\rangle \boxplus |v_j\rangle = |v_j\rangle \boxplus |v_i\rangle \quad \forall |v_i\rangle, |v_j\rangle \in V$.
3. La operación \boxplus es asociativa: $(|v_i\rangle \boxplus |v_j\rangle) \boxplus |v_k\rangle = |v_i\rangle \boxplus (|v_j\rangle \boxplus |v_k\rangle) \quad \forall |v_i\rangle, |v_j\rangle, |v_k\rangle \in V$.
4. Existe un único elemento neutro $|0\rangle$: $|0\rangle \boxplus |v_i\rangle = |v_i\rangle \boxplus |0\rangle = |v_i\rangle \quad \forall |v_i\rangle \in V$.
5. Existe un elemento simétrico para cada elemento de V : $\exists | -v_i \rangle / |v_i\rangle \boxplus | -v_i \rangle = |0\rangle \quad \forall |v_i\rangle \in V$.
6. V es cerrado bajo el producto por un número: $\alpha |v_i\rangle \in V \quad \forall \alpha \in K$ y cualquier $|v_i\rangle \in V$,
7. $\alpha(\beta |v_i\rangle) = (\alpha\beta) |v_i\rangle \quad \forall |v_i\rangle \in V$ y $\alpha, \beta \in K$.
8. $(\alpha + \beta) |v_i\rangle = \alpha |v_i\rangle \boxplus \beta |v_i\rangle \quad \forall |v_i\rangle \in V$ y $\alpha, \beta \in K$.
9. $\alpha(|v_i\rangle \boxplus |v_j\rangle) = \alpha |v_i\rangle \boxplus \alpha |v_j\rangle \quad \forall |v_i\rangle, |v_j\rangle \in V$ y $\alpha \in K$.
10. $\mathbf{1} |v_i\rangle = |v_i\rangle \quad \forall |v_i\rangle \in V$ y $\mathbf{1} \in K$.

Es importante resaltar lo siguiente: existen dos objetos neutros, el vector $|0\rangle \in V$ y el elemento $0 \in K$ y también dos operaciones producto diferentes, el producto de dos números dentro de K y el producto de un $\alpha \in K$ por un vector $|v\rangle \in V$.

Notemos también que podemos definir subespacios S vectoriales dentro de los espacios vectoriales. Ellos serán aquellos conjuntos de vectores que cumplan con los requisitos anteriores pero además cerrados dentro de los mismos conjuntos de vectores. Se puede ver entonces que la condición necesaria y suficiente para que

³También suele decirse: un espacio vectorial V sobre K .

$S \subseteq V$ sea un subespacio vectorial de V es que para cualesquier $|u_i\rangle$ y $|v_i\rangle$ de S y para cualesquier α y β de K se tiene que: $\alpha|u_i\rangle + \beta|v_i\rangle \in S$.

2.1.4 Algunos espacios vectoriales

- Los conjunto de los números reales $V = \mathbb{R}$ y el conjunto de los números complejos $V = \mathbb{C}$ con el campo K de reales o complejos y definidas las operaciones ordinarias de suma y multiplicación.

Cuando el campo K es el conjunto de los números reales se dirá que es *un espacio vectorial real de números reales* si $V \equiv \mathbb{R}$, pero si $V \equiv \mathbb{C}$ se dirá *un espacio vectorial real de números complejos*. Por su parte, si $K \equiv \mathbb{C}$ diremos que es un espacio vectorial complejo. Siempre se asociará el campo de escalares al espacio vectorial: se dirá que es un espacio vectorial sobre el campo de los escalares. Es decir, si el campo es real (complejo) se dirá que el espacio vectorial es real (complejo).

- El conjunto $V \equiv \mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$, vale decir el producto cartesiano de \mathbb{R} , cuyos elementos son n -uplas de números, con la operación suma ordinaria de vectores en n -dimensionales y la multiplicación por números.

$$\begin{aligned}|x\rangle &= (x_1, x_2, x_3, \dots, x_n) \quad \wedge \quad |y\rangle = (y_1, y_2, y_3, \dots, y_n), \\ |x\rangle \boxplus |y\rangle &\equiv (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots, x_n + y_n), \\ \alpha|x\rangle &= (\alpha x_1, \alpha x_2, \alpha x_3, \dots, \alpha x_n).\end{aligned}$$

Este espacio vectorial es de dimensión finita.

Igualmente, será un espacio vectorial $V \equiv \mathbb{C}^n = \mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}$, en donde los elementos $x_i \in \mathbb{C}$. Si para este caso el campo, sobre el cual se define el espacio vectorial \mathbb{C}^n es real, tendremos un espacio vectorial real de números complejos.

Es obvio que en el caso $V \equiv \mathbb{R}$, para el cual $|x\rangle_1 = (x_1, 0, 0, \dots, 0)$ y $|y\rangle_1 = (y_1, 0, 0, \dots, 0)$ o cualquier espacio de vectores formados por las componentes, i.e. $|x\rangle_i = (0, 0, 0, \dots, x_i, \dots, 0)$ y $|y\rangle_i = (0, 0, 0, \dots, y_i, \dots, 0)$ formarán subespacios vectoriales dentro de \mathbb{R}^n .

En el caso específico de \mathbb{R}^3 , y en donde hemos desarrollado todo un conjunto de conceptos matemáticos, es bueno comentar sobre la equivalencia que hemos tomado como obvia entre dos definiciones diferentes:

- Los vectores $\mathbf{a} = (a^1, a^2, a^3)$ y $\mathbf{b} = (b^1, b^2, b^3)$ con sus respectivas operaciones para la suma $\mathbf{a} + \mathbf{b} = (a^1 + b^1, a^2 + b^2, a^3 + b^3)$ y multiplicación por un escalar $\lambda\mathbf{a} = (\lambda a^1, \lambda a^2, \lambda a^3)$, con $\{a^i\}$, $\{b^i\}$ y $\lambda \in \mathbb{R}$.
- Los vectores geométricos, es decir, segmentos orientados en el espacio con un origen común y donde la suma se definió mediante la regla del paralelogramo y el producto por un escalar como el alargamiento o acortamiento de los segmentos o flechas con el posible cambio de dirección.

Aunque ambos son vectores de \mathbb{R}^3 es bueno tener claro que se trata de objetos diferentes que viven en espacios vectoriales diferentes. Se puede decir también que son dos representaciones diferentes para los vectores. Es buena la ocasión para señalar que existe una manera de pasar de una representación a otra, esto se hace por medio de una función que asigne a una tríada (x^1, x^2, x^3) una y sólo una flecha (y viceversa), conservando por supuesto las operaciones de suma y multiplicación por números. A este tipo de funciones se les denomina un *isomorfismo*.

Para finalizar, también es posible ver \mathbb{R}^3 simplemente como un conjunto de puntos donde se pueden definir superficies inmersas, como por ejemplo una esfera, \mathbb{S}^2 , centrada en el origen y de radio R . En un punto q sobre la esfera es posible generar un plano tangente a la esfera y en ese punto construir un espacio vectorial con todos los vectores cuyo origen está en q , y por lo tanto, son tangentes a la esfera.

En el lenguaje de la geometría diferencial y las variedades a este conjunto de vectores se le denota con: $T_q\mathbb{S}^2$. Lo anteriormente dicho se puede generalizar para \mathbb{R}^n .

3. El espacio E^∞ constituido por vectores $|x\rangle = (x_1, x_2, x_3, \dots, x_n, \dots)$ contables pero con infinitas componentes.

$$\begin{aligned}|x\rangle &= (x_1, x_2, x_3, \dots, x_n, \dots) \quad \wedge \quad |y\rangle = (y_1, y_2, y_3, \dots, y_n, \dots), \\|x\rangle \boxplus |y\rangle &\equiv (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots, x_n + y_n, \dots) \\ \alpha |x\rangle &= (\alpha x_1, \alpha x_2, \alpha x_3, \dots, \alpha x_n, \dots),\end{aligned}$$

con la siguiente restricción:

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n x_i = L, \quad \text{con } L \text{ finito.}$$

4. El conjunto de las matrices $n \times n$, reales o complejas, con el campo K real o complejo.

$$\begin{aligned}|u\rangle &= M_{ab} \quad \wedge \quad |v\rangle = N_{ab}, \\|u\rangle \boxplus |v\rangle &\equiv M_{ab} + N_{ab} = (M + N)_{ab}, \\ \alpha |u\rangle &= \alpha M_{ab} = (\alpha M)_{ab}.\end{aligned}$$

Es también obvio que se podrán formar subespacios vectoriales cuyos elementos sean matrices de dimensión menor a $n \times n$.

5. El conjunto de los vectores geométricos, o vectores cartesianos, en 2 y 3 dimensiones, con las propiedades habituales de suma vectorial y multiplicación por un escalar.
6. El conjunto de todos los polinomios con coeficientes reales: $\mathcal{P} = \{a_0, a_1x, a_2x^2, \dots, a_nx^n, \dots\}$, con \boxplus la suma ordinaria entre polinomios y la multiplicación ordinaria de polinomios con números.
7. Espacios Funcionales (de los cuales los polinomios son un caso particular). En estos espacios los vectores serán funciones, la suma será la suma ordinaria entre funciones y la multiplicación por un escalar también será la multiplicación ordinaria de una función por un elemento de un campo:

$$\begin{aligned}|f\rangle &= f(x) \quad \wedge \quad |g\rangle = g(x), \\|f\rangle \boxplus |g\rangle &\equiv f(x) + g(x) \equiv (f + g)(x), \\ \alpha |f\rangle &= (\alpha f)(x) \equiv \alpha f(x).\end{aligned}$$

8. El conjunto de todas las funciones continuas e infinitamente diferenciables, definidas en el intervalo $[a, b] : \mathcal{C}_{[a,b]}^\infty$.
9. El conjunto de todas las funciones complejas de variable real, $\psi(x)$, definidas en $[a, b]$, de cuadrado integrable (es decir para las cuales $\int_a^b dx |\psi(x)|^2$ sea finita). Este espacio se denomina comúnmente L^2 y puede ser definido dentro de un rango $[a, b]$, finito o infinito, y para más de una variable.
10. El conjunto de todas las soluciones de un sistema de ecuaciones lineales y homogéneas, por ejemplo:

$$\begin{aligned}a + 3b + c &= 0 \\ 4a + 2b + 2c &= 0\end{aligned}$$

Este sistema tiene como solución al conjunto: $\{a, b = a/2, c = -5a/2\}$. La suma de estos elementos y la multiplicación por un número conforman un espacio vectorial en el sentido de que son soluciones del sistema de ecuaciones.

2.1.4.1 Subespacios

Supongamos que tenemos un conjunto S de elementos de un espacio vectorial lineal V que satisface las siguientes propiedades:

1. El vector neutro de V está en S .
2. Si $|s_1\rangle, |s_2\rangle \in S$, entonces $|s_1\rangle \oplus |s_2\rangle \in S$.
3. Si $|s\rangle \in S$ y α es un elemento del campo K , entonces $\alpha|s\rangle \in S$.

De esta manera, las operaciones que hacen de V un espacio vectorial también están definidas para S . Se puede demostrar que S es también un espacio vectorial lineal, es decir, satisface el conjunto de axiomas 2.1.3. Al conjunto $S \subset V$ que satisface las propiedades 1 y 2 más los axiomas 2.1.3 se le denomina un subespacio vectorial lineal (o simplemente subespacio) del espacio vectorial lineal V .

Notemos que el conjunto conformado con el vector neutro como único elemento: $S = \{|0\rangle\} \in V$ es el subespacio vectorial más pequeño de V . Por otro lado, el espacio completo V es el subespacio más grande posible de V . Se acostumbra llamar a estos dos subespacios los subespacios triviales de V .

Notemos que el espacio vectorial \mathbb{R}^2 no es un subespacio de \mathbb{R}^3 porque \mathbb{R}^2 ni siquiera es un subconjunto de \mathbb{R}^3 . De la misma manera, un plano en \mathbb{R}^3 que no contenga el origen no es un subespacio de \mathbb{R}^3 , porque el plano no contiene el vector cero de \mathbb{R}^3 .

2.1.5 La importancia de la notación

En los ejemplos antes mencionados hemos utilizado para representar un vector abstracto la notación de $|v\rangle$ y con éstos construimos un espacio vectorial abstracto $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$. Un espacio vectorial abstracto será un conjunto de elementos genéricos que satisfacen ciertos axiomas. Dependiendo del conjunto de axiomas tendremos distintos tipos de espacios abstractos, la teoría desarrolla las consecuencias lógicas que resultan de esos axiomas. En matemática el concepto de espacios abstractos es reciente (1928) y, aparentemente, se le debe a Maurice Fréchet⁴.

Los elementos de esos espacios se dejan sin especificar a propósito. Ese vector abstracto puede representar, vectores en \mathbb{R}^n , matrices $n \times n$ o funciones continuas. La notación $|v\rangle$, que se denomina un *ket* y al cual le corresponde un *bra* $\langle u|$ proviene del vocablo inglés *braket* que significa corchete y será evidente más adelante cuando construyamos escalares *braket* $\langle u|v\rangle$. Esta útil notación la ideó Paul Dirac⁵, uno de los físicos más influyentes en el desarrollo de la Física del siglo XX. En Mecánica Cuántica un estado cuántico particular suele representarse por una función de onda $\psi(x)$, que depende de la variable posición x o de una función alternativa que puede depender de la variable momentum p . En la notación de Dirac, el símbolo $|\psi\rangle$ denota el estado cuántico sin referirse a la función en particular y también sirve para distinguir a los vectores de los escalares (números complejos) que vienen a ser los elementos fundamentales en el espacio de Hilbert de la Mecánica Cuántica.

⁴Maurice Fréchet (1878-1973). Versátil matemático francés, con importantes contribuciones en espacios métricos, topología y creador del concepto de espacios abstractos. https://es.wikipedia.org/wiki/Maurice_Fréchet

⁵Paul Adrien Maurice Dirac (1902-1984). Además de contribuir de manera determinante en la comprensión de la Mecánica Cuántica, es uno de los creadores de la Mecánica Cuántica Relativista la cual ayudó a comprender el papel que juega el espín en las partículas subatómicas. https://es.wikipedia.org/wiki/Paul_Dirac

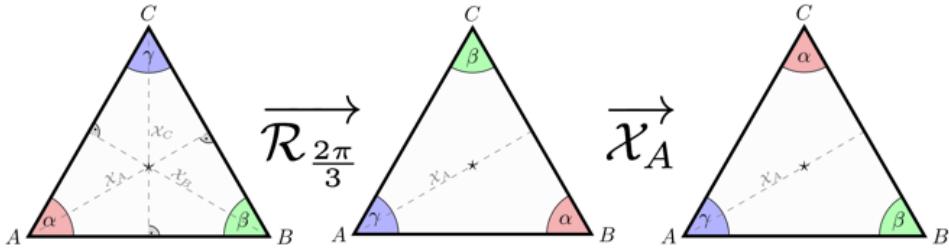


Figura 2.1: Transformaciones que dejan invariante un triángulo equilátero. Concatenación de una rotación, $\mathcal{R}_{\frac{2\pi}{3}}$ con una reflexión, \mathcal{X}_A , respecto a un plano que pasa por A

2.1.6 Ejercicios

- Diga cuales de los siguientes grupos son abelianos: $\{\mathbb{Z}, +\}$; $\{\mathbb{Z}_n, + \forall, n \in \mathbb{N}\}$; $\{\mathbb{N}, +\}$ y el grupo $\{\mathbb{Z}, \cdot\}$.
- Sea S el conjunto de todos los números reales excluyendo -1 y defina la operación \square tal que:

$$a \square b = a + b + ab.$$

Donde $+$ es la suma estándar entre números reales. Entonces:

- (a). Muestre si $[S, \square]$ forman grupo.
- (b). Encuentre la solución en S para la ecuación $2 \square x \square 3 = 7$.
- 3. Considere un triángulo equilátero que se muestra en la figura 2.1. Se pueden identificar operaciones de rotación alrededor de un eje perpendicular a la figura que pasa por su baricentro \star y, reflexiones respecto a planos, \mathcal{X}_A , \mathcal{X}_B y \mathcal{X}_C , que dejan invariante la figura del triángulo. El lector puede consultar los ejemplos 2.1, 2.2 y 2.3 para fijar ideas.

Ahora bien, se puede definir la operación concatenación de rotaciones y reflexiones que dejan igualmente invariante al triángulo, tal y como mostramos en la mencionada figura 2.1. Note que lo ilustrado en la figura, puede esquematizarse como:

$$(A \alpha, B \beta, C \gamma) \quad \xrightarrow{\mathcal{R}_{\frac{2\pi}{3}}} \quad (A \gamma, B \alpha, C \beta) \quad \xrightarrow{\mathcal{X}_A} \quad (A \gamma, B \beta, C \alpha).$$

- (a). Construya la tabla de multiplicación para G_Δ , vale decir $G_\Delta = \{\mathcal{I}, \{\mathcal{R}_i\}, \{\bar{\mathcal{R}}_j\}, \{\mathcal{X}_k\}\}$ y la operación es concatenación tal y como mostramos en la figura 2.1. Donde \mathcal{I} es la operación identidad, $\{\mathcal{R}_i\}$ es un conjunto de rotaciones en sentido horario, mientras que $\{\bar{\mathcal{R}}_j\}$ es un conjunto de rotaciones en el sentido antihorario, y $\{\mathcal{X}_k\}$ el conjunto de las reflexiones que dejan invariante el triángulo.
- (b). Muestre que el conjunto de estas operaciones forman el grupo: G_Δ .
- (c). Identifique cada una de las \mathcal{R}_i y $\bar{\mathcal{R}}_j$, y muestre además, que forman un subgrupo cíclico de orden 3. De igual modo identifique las reflexiones y muestre que, cada una de las reflexiones y la identidad, $\{\mathcal{I}, \mathcal{X}_i\}$, forman también un subgrupo cíclico, pero de orden 2.
- (d). Considere las siguientes matrices:

$$\mathbb{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{A} = \begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix},$$

$$\mathbb{C} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbb{D} = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad \mathbb{E} = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}.$$

Muestre que forman grupo bajo la multiplicación de matrices y que ese grupo es isomorfo a \mathbf{G}_Δ .

- (e). Considere el conjunto de permutaciones de 3 objetos y la operación composición de permutaciones que discutimos como ejemplo en la sección 2.3. ¿Es ese grupo isomorfo a \mathbf{G}_Δ ? Justifique su respuesta.
(f). ¿Qué puede decir de las operaciones simetrías que dejan invariante un triángulo isósceles? ¿formarán grupo? ¿y si el triángulo es escaleno, cuales son las operaciones de simetría que lo dejan invariante?

4. Considere las siguientes funciones:

$$f_1(x) = x, \quad f_2(x) = \frac{1}{x}, \quad f_3(x) = \frac{1}{1-x}, \quad f_4(x) = \frac{x-1}{x}, \quad f_5(x) = 1-x, \quad f_6(x) = \frac{x}{x-1}.$$

Muestre que forman grupo bajo la operación: $f_i(x) \odot f_j(x) = f_i(f_j(x))$, y que ese grupo es isomorfo a \mathbf{G}_Δ , del ejercicio anterior.

5. Definamos una operación binaria \blacksquare como:

$$x \blacksquare y = x + y + \alpha xy,$$

con $x, y, \alpha \in \mathbb{R}$ y además $\alpha \neq 0$.

- (a). Demuestre que \blacksquare es asociativa.
(b). Muestre que \blacksquare genera un grupo en $\{\mathbb{R} - (\frac{-1}{\alpha})\}$. Es decir, $\forall x, y \in \mathbb{R} \wedge x \neq \frac{-1}{\alpha}, y \neq \frac{-1}{\alpha}$, entonces: $x \blacksquare y$ forma un grupo.

6. Muestre que el siguiente conjunto de transformaciones en el plano xy forman un grupo y construya su tabla de multiplicación.

- (a). $I = \{x \rightarrow x \wedge y \rightarrow y\}.$
(b). $I = \{x \rightarrow -x \wedge y \rightarrow -y\}.$
(c). $I_x = \{x \rightarrow -x \wedge y \rightarrow y\}.$
(d). $I_y = \{x \rightarrow x \wedge y \rightarrow -y\}.$

7. Considere un conjunto S conformado únicamente por números reales positivos. Consideremos las siguientes reglas sobre S : Por "suma" de dos números entenderemos su producto en el sentido usual, y el "producto" de un elemento $r \in S$ y un número real λ entenderemos r elevado a la potencia de λ , en el sentido usual. ¿ S es un espacio vectorial?

8. Considere el conjunto de vectores en el plano conformado por vectores localizados en el origen y cuyos puntos finales permanecen siempre en el primer cuadrante. ¿Este conjunto es un espacio vectorial?

9. Muestre que también serán espacios vectoriales:

- (a). El conjunto de todas las funciones $f = f(x)$ definidas en $x = 1$ con $f(1) = 0$. Si $f(1) = c$. ¿Tendremos igual un espacio vectorial? ¿Por qué?
(b). Los vectores $(x, y, z) \in V^3$ tal que sus componentes satisfacen el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} a_{11}x + a_{12}y + a_{13}z &= 0 \\ a_{21}x + a_{22}y + a_{23}z &= 0 \\ a_{31}x + a_{32}y + a_{33}z &= 0. \end{aligned}$$

10. Sea \mathcal{P}_n el conjunto de todos los polinomios de grado n , en x , con coeficientes reales:

$$|p_n\rangle \Rightarrow p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_{n-1}x^{n-1} = \sum_{i=0}^{n-1} a_i x^i.$$

- (a). Demostrar que \mathcal{P}_n es un espacio vectorial respecto a la suma de polinomios y a la multiplicación de polinomios por un número (número real).
- (b). Si los coeficientes a_i son enteros, ¿ \mathcal{P}_n será un espacio vectorial? ¿Por qué?
- (c). ¿Cuál de los siguientes subconjuntos de \mathcal{P}_n es un subespacio vectorial?
 - I. El polinomio cero y todos los polinomios de grado $n - 1$.
 - II. El polinomio cero y todos los polinomios de grado par.
 - III. Todos los polinomios que tienen a x como un factor ($\text{grado } n > 1$).
 - IV. Todos los polinomios que tienen a $x - 1$ como un factor.
- 11. Un subespacio \mathcal{P} es generado por: $|x_1\rangle = x^3 + 2x + 1$, $|x_2\rangle = x^2 - 2$, $|x_3\rangle = x^3 + x$. ¿Cuál(es) de los siguientes polinomios pertenece al subespacio \mathcal{P} ?
 - (a). $x^2 - 2x + 1$.
 - (b). $x^4 + 1$.
 - (c). $-\frac{1}{2}x^3 + \frac{5}{2}x^2 - x - 1$.
 - (d). $x - 5$.
- 12. Resuelva los ejercicios anteriores utilizando **Maxima**.

Practicando con Maxima

En este módulo utilizaremos algunas de las herramientas disponibles para incorporar el álgebra discreta. Comenzaremos con la introducción a los conjuntos. Podemos operar con conjuntos pero primero debemos definirlos. Existen varias posibilidades, como mostramos a continuación:

```
(%i1) A:{1,2,3,4,5,6,7,8,9};
(%o1) {1, 2, 3, 4, 5, 6, 7, 8, 9}

(%i2) B:set(1,3,5,7,9);
(%o2) {1, 3, 5, 7, 9}

(%i3) C:set(2,4,6,8);
(%o3) {2, 4, 6, 8}

(%i4) D:makeset(i/j, [i,j], [[1,3], [2,3], [3,3], [4,3]]);
(%o4) {1/3, 2/3, 1, 4/3}
```

Notemos que es igual definir los conjuntos con llaves, con la función **set** o **makeset**. Podemos preguntar si determinado elemento pertenece, o no, a un conjunto.

```
(%i5) elementp(7,A);
(%o5) true

(%i6) elementp(7,C);
(%o6) false
```

Operaciones elementales con conjuntos:

```
(%i7) UBC:union(B,C);
```

(%o7) $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$

(%i8) is (A=UBC);
 (%o8) true

(%i9) Cv:intersection(B,C);
 (%o9) {}

Para **Maxima** el conjunto vacío es {}.

(%i10) setdifference(A,C);
 (%o10) {1, 3, 5, 7, 9}

Esto es, el conjunto con los elementos del conjunto A que no pertenecen al conjunto C .

Si queremos el conjunto de todos los subconjuntos del conjunto $\{a, b, c\}$ le podemos pedir al programa que nos lo muestre:

(%i11) powerset(a,b,c);
 (%o11) {{}, {a}, {a, b}, {a, b, c}, {a, c}, {b}, {b, c}, {c}}

El producto cartesiano de los conjuntos A y B es el conjunto conformado por los pares (a, b) :

$$A \times B = \{(a, b) / a \in A, b \in B\}.$$

(%i12) cartesian_product(B,C);
 (%o12) {[1, 2], [1, 4], [1, 6], [1, 8], [3, 2], [3, 4], [3, 6], [3, 8], [5, 2], [5, 4], [5, 6], [5, 8], [7, 2],
 , [7, 4], [7, 6], [7, 8], [9, 2], [9, 4], [9, 6], [9, 8]}

Le podemos pedir al programa la suma de los pares del producto cartesiano anterior:

(%i13) makeset(a+b, [a,b], cartesian_product(B,C));
 (%o13) {3, 5, 7, 9, 11, 13, 15, 17}

Maxima trata a los conjuntos y a las listas como objetos de distinta naturaleza, lo que permite trabajar con conjuntos cuyos elementos puedan ser también conjuntos o listas, es decir, subconjuntos.

(%i14) lista:makelist(i,i,1,25);
 (%o14) [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25]

La lista anterior la convertiremos en un conjunto, para este fin debemos utilizar el comando **setify**.

(%i15) E:setify(lista);
 (%o15) {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25}

De este último conjunto podemos construir el subconjunto conformado por los números primos:

(%i16) Primos:subset(E,primep);
 (%o16) {2, 3, 5, 7, 11, 13, 17, 19, 23}

Con **cardinality** podemos saber cuantos elementos contiene un conjunto:

(%i17) cardinality(%);

(%o17) 9

La función de **Maxima** que nos permite calcular las tablas que utilizamos en la sección 2.1 es la función **mod**. Veamos el siguiente ejemplo:

(%i18) `Gm8:1,3,5,7;`
 (%o18) {1, 3, 5, 7}

Todos los productos posibles entre los elementos de este conjunto son:

(%i19) `makeset(a,b, [a,b], cartesian_product(Gm8,Gm8));`
 (%o19) {1, 3, 5, 7, 9, 15, 21, 25, 35, 49}

Los modulo 8 para algunos de los números anteriores son:

(%i20) `mod(21,8); mod(35,8); mod(49,8);`
 (%o20) 5
 (%o21) 3
 (%o22) 1

Para generar el grupo $G_{\text{mod}5}$ escribimos:

(%i23) `setify(makelist(mod(i,5),i,1,4));`
 (%o23) {1, 2, 3, 4}

En la sección 2.1.1.2, definimos el orden de un elemento $g \in G$. Consideremos el siguiente conjunto de números enteros:

$$\mathbb{Z}_{15} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}.$$

Se podría demostrar que el orden de $g \in G$ es igual al número de elementos de $\langle g \rangle$ y lo dejaremos como ejercicio.

Vamos a calcular el orden de todos los elementos de \mathbb{Z}_{15} , sabiendo que el orden de cada uno de esos elementos divide a 15, que es el cardinal de \mathbb{Z}_{15} .

Probemos primero con el número 6 $\in \mathbb{Z}_{15}$.

(%i24) `setify(makelist(mod(6*i,15),i,0,14));`
 (%o24) {0, 3, 6, 9, 12}

El orden de $6 \in \mathbb{Z}_{15}$ es:

(%i25) `length(%);`
 (%o25) 5

En la siguiente instrucción lo haremos para todos los elementos de \mathbb{Z}_{15} .

(%i26) `makelist([j,length(setify(makelist(mod(j*i,15),i,0,14)))],j,0,14);`
 (%o26) [[0, 1], [1, 15], [2, 15], [3, 5], [4, 15], [5, 3], [6, 5], [7, 15], [8, 15], [9, 5], [10, 3], [11, 15], [12, 5], [13, 15], [14, 15]]

La salida no es más que una lista $[x, y]$ con cada elemento $x \in \mathbb{Z}_{15}$ y su orden y . Por lo tanto, el conjunto de órdenes es: {1, 3, 5, 15}, todos divisores de 15, como estipula el teorema de Lagrange.

2.2 Espacios métricos, normados y con producto interno

En esta sección vamos a introducir una función de distancia, de manera que sí tenemos un par de puntos o elementos de un espacio vectorial podemos hablar de que existe una cierta distancia entre ellos. Se dice que la función distancia induce una topología sobre el espacio vectorial. Comenzaremos definiendo el concepto de métrica sobre espacios vectoriales y con esta estructura llegar a la noción de norma.

2.2.1 Métricas y espacios métricos

El concepto de métrica surge de la generalización de la idea de distancia entre dos puntos de la recta real. Un espacio vectorial será métrico si podemos definir una función:

$$d : \mathbf{V} \times \mathbf{V} \rightarrow \mathbb{R} / \forall |x\rangle, |y\rangle, |z\rangle \in \mathbf{V},$$

tal que se cumple lo siguiente:

1. $d(|x\rangle, |y\rangle) \geq 0$, si $d(|x\rangle, |y\rangle) = 0 \Rightarrow |x\rangle \equiv |y\rangle$.
2. $d(|x\rangle, |y\rangle) \equiv d(|y\rangle, |x\rangle)$.
3. $d(|x\rangle, |y\rangle) \leq d(|x\rangle, |z\rangle) + d(|y\rangle, |z\rangle)$ (La desigualdad triangular).

Así, diremos que $(\mathbf{V}, \mathbf{K}, \boxplus, d)$ es un espacio vectorial lineal, métrico.

Podemos enumerar algunos ejemplos de espacios métricos:

1. Espacios reales \mathbb{R}^n .

- (a). Para \mathbb{R} , es decir la recta real, la definición de métrica es: $d(|x\rangle, |y\rangle) \equiv |x - y|$.
- (b). Para \mathbb{R}^2 , es decir el plano, una definición de métrica es: $d(|x\rangle, |y\rangle) \equiv \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$. También podemos construir otra definición de métrica como: $d(|x\rangle, |y\rangle) \equiv |x_1 - y_1| + |x_2 - y_2|$. La primera de estas métricas se conoce como métrica euclídea y la segunda como métrica Manhattan o métrica de taxistas. Es claro como el mismo espacio vectorial genera varios espacios métricos, dependiendo de la definición de métrica. Para estos casos particulares, las métricas “miden” el desplazamiento entre dos puntos de forma distinta: en drones (métrica euclídea) o vehículos terrestre en ciudades.
- (c). En general para espacios reales \mathbb{R}^n una posible definición de métrica será:

$$d(|x\rangle, |y\rangle) \equiv \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \cdots + (x_n - y_n)^2}.$$

Esta definición de métrica no es más que una generalización del teorema de Pitágoras y se denomina “distancia euclíadiana”.

2. Espacios unitarios n -dimensionales, o espacios complejos, \mathbb{C}^n . La definición de distancia puede construirse como:

$$d(|x\rangle, |y\rangle) \equiv \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + \cdots + |x_n - y_n|^2},$$

y es claro que se recobra la idea de distancia en el plano complejo: $d(|x\rangle, |y\rangle) \equiv |x - y|$.

3. Para los espacios de funciones $\mathcal{C}_{[a,b]}^\infty$ una posible definición de distancia sería:

$$d(|f\rangle, |g\rangle) \equiv \max_{t \in [a,b]} |f(t) - g(t)|.$$

4. La métrica trivial o discreta

$$d(|x\rangle, |y\rangle) = \begin{cases} 1 & \text{si } x \neq y \\ 0 & \text{si } x = y \end{cases}$$

Ejemplo 2.6 Como vimos en la sección anterior, en el campo de los números complejos el valor absoluto de

$z = x + iy$ es $|z| = \sqrt{x^2 + y^2}$. La métrica que podemos asociar a este espacio vectorial viene dada por:

$$d(|z_1\rangle, |z_2\rangle) \equiv \| |z_1\rangle - |z_2\rangle \| = \sqrt{\langle z_1 - z_2 | z_1 - z_2 \rangle} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2},$$

con: $|z_1\rangle = x_1 + iy_1$ y $|z_2\rangle = x_2 + iy_2$.



2.2.2 Normas y espacios normados

El concepto de distancia (o métrica) es la propiedad más elemental que se le puede exigir a un espacio vectorial. Mucho más interesante son aquellos espacios vectoriales que están equipados con la idea de norma y, a partir de allí, se obtiene la distancia. La norma tiene que ver con el “tamaño” del vector y la métrica con la distancia entre vectores.

La norma, $\mathcal{N}(|v_i\rangle) \equiv \| |v_i\rangle \|$, de un espacio vectorial $\mathbf{V} = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ será una función:

$$\mathcal{N} : \mathbf{V} \rightarrow \mathbb{R} / \forall |v_i\rangle \in \mathbf{V},$$

que cumple con:

1. $\| |v_i\rangle \| \geq 0$, si $\| |v_i\rangle \| = 0 \Rightarrow |v_i\rangle \equiv |0\rangle$.
2. $\| \alpha |v_i\rangle \| = |\alpha| \| |v_i\rangle \|$.
3. $\| |v_i\rangle + |v_j\rangle \| \leq \| |v_i\rangle \| + \| |v_j\rangle \|$ (Desigualdad triangular).

La definición de norma induce una métrica de la forma:

$$d(|v_i\rangle, |v_j\rangle) \equiv \| |v_i\rangle - |v_j\rangle \| . \quad (2.1)$$

Denotaremos un espacio vectorial normado⁶ como $(\mathbf{V}, \mathbf{K}, \boxplus; \| \cdot \|)$, espacio que también es conocido como un espacio de Banach⁷. Esto significa que todo espacio normado es a su vez un espacio métrico, pero no todo espacio métrico es normado.

De la definición de distancia (2.1) resulta que la métrica así definida es invariante bajo traslaciones de vectores. Esto es, si: $|\tilde{x}\rangle = |x\rangle + |a\rangle \wedge |\tilde{y}\rangle = |y\rangle + |a\rangle$, entonces, $d(|x\rangle, |y\rangle) \equiv d(|\tilde{x}\rangle, |\tilde{y}\rangle)$. Y además es homogénea: $d(\lambda |x\rangle, \lambda |y\rangle) = |\lambda| d(|x\rangle, |y\rangle)$.

Como ejemplos de espacios normados podemos mostrar los siguientes:

1. Los espacios reales, \mathbb{R}^n y los espacios complejos \mathbb{C}^n . Para estos espacios de Banach, la norma se define como:

$$\| |x\rangle \| = \sqrt{|x_1|^2 + |x_2|^2 + |x_3|^2 + \dots + |x_n|^2} = \sqrt{\sum_{i=1}^n |x_i|^2}.$$

Para un espacio en \mathbb{R}^3 se cumple que $\| |x\rangle \| = \sqrt{x_1^2 + x_2^2 + x_3^2}$, por lo tanto, la idea de norma generaliza la noción de “tamaño” del vector $|x\rangle$. Es claro que la definición de distancia se construye a partir de la norma de la forma:

$$d(|x\rangle, |y\rangle) \equiv \| |x\rangle - |y\rangle \| = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + |x_3 - y_3|^2 + \dots + |x_n - y_n|^2}.$$

Este espacio se llama espacio normado euclíadiano de dimensión n .

2. Para el espacio lineal de matrices $n \times n$, reales o complejas, con el campo \mathbf{K} real o complejo, una definición

⁶El concepto de espacio normado fue formulado en 1922 de manera independiente por: S. Banach, H. Hahn y N. Wiener.

⁷Stefan Banach (1892-1945) Matemático polaco, uno de los fundadores del análisis funcional moderno, con sus mayores contribuciones a la teoría de espacios topológicos. https://es.wikipedia.org/wiki/Stefan_Banach

de norma es:

$$\|M\| = \sum_{a=1}^m \sum_{b=1}^n |M_{ab}| ,$$

y la correspondiente definición de distancia:

$$d(|x\rangle, |y\rangle) \equiv \|M - N\| = \sum_{a=1}^m \sum_{b=1}^n |M_{ab} - N_{ab}| .$$

3. Para los espacios funcionales $\mathcal{C}_{[a,b]}^\infty$ una posible definición de norma sería:

$$\|f\| = \max_{t \in [a,b]} |f(t)| ,$$

otra posible definición puede ser:

$$\|f\| = \sqrt{\int_a^b |f(t)|^2 dt} .$$

2.2.3 Espacios euclidianos

El siguiente paso en la construcción de espacios vectoriales más ricos es equiparlo con la definición de producto interno, y a partir de esta definición, construir el concepto de norma y con éste el de distancia. La idea de producto interno generaliza el concepto de producto escalar de vectores en \mathbb{R}^3 e incorpora a los espacios vectoriales abstractos el concepto de ortogonalidad y descomposición ortogonal. Históricamente, la teoría de espacios vectoriales con producto interno es anterior a la teoría de espacios métricos y espacios de Banach y se le debe a D. Hilbert⁸. Adicionalmente, la semejanza entre la geometría eucliana y la geométrica de \mathbb{R}^n ha hecho que espacios en los cuales se puedan definir, distancia, ángulos, a partir de una definición de producto interno, se denominen también espacios euclidianos.

2.2.3.1 Producto interno

En un espacio vectorial $\mathbf{V} = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$, la definición del producto interno de dos vectores la denotaremos como $\langle v_i | v_j \rangle$ y es una aplicación:

$$\mathcal{I}(|v_i\rangle, |v_j\rangle) : \mathbf{V} \times \mathbf{V} \rightarrow \mathbf{K}, \quad \forall |v_i\rangle, |v_j\rangle \in \mathbf{V}.$$

Es decir, asocia a ese par de vectores con un elemento del campo \mathbf{K} .

Las propiedades que definen el producto interno son:

1. $\langle v_i | v_i \rangle \equiv \| |v_i\rangle\|^2 \in \mathbf{K} \wedge \langle v_i | v_i \rangle \geq 0 \quad \forall |v_i\rangle \in \mathbf{V}, \quad \text{si } \langle v_i | v_i \rangle = 0 \Rightarrow |v_i\rangle \equiv |0\rangle.$
2. $\langle v_i | v_j \rangle = \langle v_j | v_i \rangle^* \quad \forall |v_i\rangle, |v_j\rangle \in \mathbf{V}.$
3. $\langle v_i | \alpha v_j + \beta v_k \rangle = \alpha \langle v_i | v_j \rangle + \beta \langle v_i | v_k \rangle \quad \forall |v_i\rangle, |v_j\rangle, |v_k\rangle \in \mathbf{V} \wedge \alpha, \beta \in \mathbf{K}.$
4. $\langle \alpha v_i + \beta v_j | v_k \rangle = \alpha^* \langle v_i | v_k \rangle + \beta^* \langle v_j | v_k \rangle \quad \forall |v_i\rangle, |v_j\rangle, |v_k\rangle \in \mathbf{V} \wedge \alpha, \beta \in \mathbf{K}.$
5. $\langle v_i | 0 \rangle = \langle 0 | v_i \rangle = 0.$

La segunda y cuarta propiedad resultan del hecho de que si el campo es complejo $\mathbf{K} = \mathbb{C}$, entonces:

$$\langle \alpha v_i | \alpha v_i \rangle = \alpha^2 \langle v_i | v_i \rangle = i^2 \langle v_i | v_i \rangle = -\langle v_i | v_i \rangle ,$$

lo cual contradice el hecho de que la norma tiene que ser positiva. Por eso la necesidad de tomar el complejo conjugado. Se dice entonces, que el producto escalar es *antilineal* respecto al primer factor y *lineal* respecto al segundo.

⁸David Hilbert(1862-1943) Matemático alemán defensor de la axiomática como enfoque primordial de los problemas científicos. Hizo importantes contribuciones en distintas áreas de la matemática, como: Invariantes, Campos de Números Algebraicos, Análisis Funcional, Ecuaciones Integrales, Física-Matemática y Cálculo en Variaciones. https://es.wikipedia.org/wiki/David_Hilbert

A partir de la definición de producto interno se construyen los conceptos de norma y distancia:

$$\| |v_i\rangle \| = \sqrt{\langle v_i | v_i \rangle} \quad \text{y} \quad d(|v_i\rangle, |v_j\rangle) \equiv \| |v_i\rangle - |v_j\rangle \| = \sqrt{\langle v_i - v_j | v_i - v_j \rangle}.$$

2.2.3.2 La desigualdad de Cauchy-Schwarz: los ángulos entre vectores reales y complejos

Todo producto interno $\langle v_i | v_j \rangle$ definido en un espacio vectorial normado $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ cumple con la desigualdad de Cauchy-Schwarz:

$$|\langle v_i | v_j \rangle|^2 \leq \langle v_i | v_i \rangle \langle v_j | v_j \rangle \iff |\langle v_i | v_j \rangle| \leq \| |v_i\rangle \| \| |v_j\rangle \| . \quad (2.2)$$

Es claro que si $|v_i\rangle = |0\rangle \wedge |v_j\rangle = |0\rangle$ se cumple la igualdad y es trivial la afirmación.

Para demostrar la desigualdad, tomemos dos vectores $|v_i\rangle \wedge |v_j\rangle$ cualesquiera, entonces podemos construir un tercero: $|v_k\rangle = \alpha |v_i\rangle + \beta |v_j\rangle$ (α y β tendrán valores particulares), por lo tanto:

$$\langle v_k | v_k \rangle \equiv \langle \alpha v_i + \beta v_j | \alpha v_i + \beta v_j \rangle \geq 0,$$

esto significa que:

$$\begin{aligned} \langle \alpha v_i + \beta v_j | \alpha v_i + \beta v_j \rangle &= \langle \alpha v_i | \alpha v_i \rangle + \langle \alpha v_i | \beta v_j \rangle + \langle \beta v_j | \alpha v_i \rangle + \langle \beta v_j | \beta v_j \rangle \geq 0 \\ &= |\alpha|^2 \langle v_i | v_i \rangle + \alpha^* \beta \langle v_i | v_j \rangle + \beta^* \alpha \langle v_j | v_i \rangle + |\beta|^2 \langle v_j | v_j \rangle \geq 0. \end{aligned}$$

Si $\alpha = \langle v_j | v_j \rangle$, se tiene que:

$$\begin{aligned} \langle v_j | v_j \rangle \langle v_i | v_i \rangle + \beta \langle v_i | v_j \rangle + \beta^* \langle v_j | v_i \rangle + |\beta|^2 &\geq 0 \\ \langle v_j | v_j \rangle \langle v_i | v_i \rangle &\geq -\beta \langle v_i | v_j \rangle - \beta^* \langle v_j | v_i \rangle - |\beta|^2, \end{aligned}$$

seguidamente seleccionamos: $\beta = -\langle v_j | v_i \rangle$, y por lo tanto: $\beta^* = -\langle v_i | v_j \rangle$, consecuentemente:

$$\begin{aligned} \langle v_j | v_j \rangle \langle v_i | v_i \rangle &\geq \langle v_j | v_i \rangle \langle v_i | v_j \rangle + \langle v_i | v_j \rangle \langle v_j | v_i \rangle - \langle v_j | v_i \rangle \langle v_i | v_j \rangle \\ \langle v_j | v_j \rangle \langle v_i | v_i \rangle &\geq \langle v_i | v_j \rangle \langle v_j | v_i \rangle = |\langle v_i | v_j \rangle|^2. \quad \blacksquare \end{aligned}$$

Ahora consideremos lo siguiente, de la desigualdad de Cauchy-Schwarz y la definición de norma se desprende que:

$$\frac{|\langle v_i | v_j \rangle|^2}{\| |v_i\rangle \|^2 \| |v_j\rangle \|^2} \leq 1 \Rightarrow -1 \leq \frac{|\langle v_i | v_j \rangle|}{\| |v_i\rangle \| \| |v_j\rangle \|} \leq 1,$$

por lo tanto podemos definir el “ángulo” entre los vectores abstractos $|v_i\rangle \wedge |v_j\rangle$ como:

$$\cos(\Theta_G) = \frac{|\langle v_i | v_j \rangle|}{\| |v_i\rangle \| \| |v_j\rangle \|}, \quad (2.3)$$

donde hemos denotado como Θ_G el ángulo genérico que forman los vectores reales o complejos.

Si estamos considerando espacios vectoriales reales -en los cuales el campo corresponde a los números reales- entonces el ángulo definido entre vectores abstractos reales corresponde al que intuitivamente siempre hemos considerado para los vectores cartesianos y que discutimos en la sección 1.1.4.1,

$$\cos(\Theta_R) = \frac{|\langle v_i | v_j \rangle|}{\| |v_i\rangle \| \| |v_j\rangle \|}, \quad \text{con } 0 \leq \Theta_R \leq \pi,$$

donde se toma $\Theta_R = 0$ para vectores colineales y $\Theta_R = \pi$ para vectores opuestos (antilineales). Si bien es cierto que esta definición coincide con la de los vectores cartesianos, hay que resaltar que la estamos extendiendo para cualquier vector abstracto, vale decir: funciones reales, matrices, y todos los objetos matemáticos que cumplan con las reglas para los espacios vectoriales expuestas en 2.1.3.

Para el caso de espacios vectoriales complejos la situación es más sutil y significa definir un ángulo entre dos vectores abstractos y complejos, sin embargo podemos abordar el problema suponiendo:

1. un espacio complejo n -dimensional de n -uplas de números complejos $|z\rangle \leftrightarrow (z_1, z_2, \dots, z_n)$ asociado (isomorfo) a un espacio vectorial real de $2n$ dimensiones, con vectores representados por $2n$ -uplas de números reales $|w\rangle \leftrightarrow (\operatorname{Re}(z_1), \operatorname{Re}(z_2), \dots, \operatorname{Re}(z_n), \operatorname{Im}(z_1), \operatorname{Im}(z_2), \dots, \operatorname{Im}(z_n))$, donde hemos representado $\operatorname{Re}(z_j)$ y $\operatorname{Im}(z_j)$ como las partes reales e imaginarias de z_j , respectivamente o,

2. directamente a partir de una definición de producto interno entre vectores complejos implementado por:

$$\langle w_i | v_j \rangle = \sum_{j=1}^n w_j^* v_j.$$

Ambas aproximaciones no son del todo independientes pero igualmente justificadas⁹.

Consideremos el segundo caso, esto es: directamente a partir de una definición de producto interno entre vectores complejos. Para este caso consideramos un ángulo complejo, y $\cos(\Theta_C)$ una función de variable compleja, que puede ser expresada en su forma polar como:

$$\cos(\Theta_C) = \frac{|\langle v_i | v_j \rangle|}{\|v_i\| \|v_j\|} \Rightarrow \cos(\Theta_C) = \rho e^\phi, \quad \text{con } \rho = |\cos(\Theta_C)| < 1.$$

Entonces podemos asociar $\rho = \cos(\Theta_H)$ y definir Θ_H , en el rango $0 \leq \Theta_H \leq \pi/2$, como el *ángulo hermítico* entre los vectores complejos $|v_i\rangle$ y $|v_j\rangle$. Mientras que ϕ , definido en $-\pi \leq \phi \leq \pi$, corresponde al pseudo ángulo de Kasner, que representa la orientación o rotación del ángulo hermítico y no tiene mayor significado al cuantificar el ángulo entre esos dos vectores. Esta diferencia de significados puede intuirse cuando multiplicamos $|v_i\rangle$ y $|v_j\rangle$, por una constante compleja: $|\tilde{v}_i\rangle \rightarrow \alpha_i |v_i\rangle$ y comprobamos que el ángulo Θ_H permanece inalterado y no así el ángulo de Kasner¹⁰.

2.2.3.3 Teoremas del coseno y de Pitágoras

A partir de la definición de norma se obtiene:

$$\begin{aligned} \| |v_i\rangle - |v_j\rangle \|^2 &= \langle v_i - v_j | v_i - v_j \rangle = \langle v_i | v_i \rangle - \langle v_i | v_j \rangle - \langle v_i | v_j \rangle^* + \langle v_j | v_j \rangle \\ &= \langle v_i | v_i \rangle + \langle v_j | v_j \rangle - 2 \operatorname{Re}(\langle v_i | v_j \rangle), \end{aligned}$$

con lo cual hemos generalizado el teorema del coseno para un espacio vectorial abstracto:

$$\| |v_i\rangle - |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2 - 2 \| |v_i\rangle \| \| |v_j\rangle \| \cos(\Theta_G). \quad (2.4)$$

De tal forma que para espacios vectoriales reales tendremos:

$$\| |v_i\rangle - |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2 - 2 \| |v_i\rangle \| \| |v_j\rangle \| \cos(\Theta), \quad \text{con } 0 \leq \Theta \leq \pi,$$

y para espacios vectoriales complejos:

$$\| |v_i\rangle - |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2 - 2 \| |v_i\rangle \| \| |v_j\rangle \| \cos(\Theta_H) \cos(\phi), \quad \text{con } 0 \leq \Theta_H \leq \pi/2.$$

Para el caso que los vectores $|v_i\rangle \wedge |v_j\rangle$ sean ortogonales, esto es $\langle v_i | v_j \rangle = 0$, tendremos el teorema de Pitágoras generalizado:

$$\| |v_i\rangle - |v_j\rangle \|^2 \equiv \| |v_i\rangle + |v_j\rangle \|^2 = \| |v_i\rangle \|^2 + \| |v_j\rangle \|^2. \quad (2.5)$$

Veamos algunos ejemplos de espacios vectoriales con producto interno.

1. Espacios euclidianos reales, \mathbb{R}^n y espacios euclidianos complejos \mathbb{C}^n .

Los vectores en estos espacios euclidianos pueden ser representados por

$$|x\rangle = (x_1, x_2, \dots, x_n) \wedge |y\rangle = (y_1, y_2, \dots, y_n),$$

⁹Scharnhorst, K. (2001). Angles in complex vector spaces. *Acta Applicandae Mathematica*, **69** (1), 95-103.

¹⁰Puede consultarse Reju, V. G., Koh, S. N., y Soon, Y. (2009). An algorithm for mixing matrix estimation in instantaneous blind source separation. *Signal Processing*, **89** (9), 1762-1773.

y el **producto interno** queda definido por:

$$\langle x | y \rangle = x_1^* y_1 + x_2^* y_2 + x_3^* y_3 + \cdots + x_n^* y_n = \sum_{i=1}^n x_i^* y_i , \quad (2.6)$$

es claro que esta definición de producto interno coincide, para \mathbb{R}^2 (y \mathbb{R}^3) con la idea de producto escalar convencional que consideramos en las secciones 1.1.4.1 y 1.2.6.1, vale decir:

$$\left. \begin{array}{l} \mathbf{a} = a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} \\ \mathbf{b} = b_x \hat{\mathbf{i}} + b_y \hat{\mathbf{j}} \end{array} \right\} \Rightarrow \mathbf{a} \cdot \mathbf{b} = a_x b_x + a_y b_y .$$

Ahora bien, el lector puede comprobar que para vectores en \mathbb{R}^2 también se puede proveer una definición de producto interno diferente:

$$\mathbf{a} \circledast \mathbf{b} = 2a_x b_x + a_x b_y + a_y b_x + a_y b_y ,$$

igualmente válida, con lo cual es claro que en un mismo espacio vectorial pueden coexistir diferentes productos internos.

Por su parte, la **norma** es:

$$\|x\| = \sqrt{\langle x | x \rangle} = \sqrt{x_1^2 + x_2^2 + x_3^2 + \cdots + x_n^2} = \sqrt{\sum_{i=1}^n x_i^2} .$$

La **distancia** también recupera la idea intuitiva de distancia euclíadiana:

$$\begin{aligned} d(|x\rangle, |y\rangle) &\equiv \| |x\rangle - |y\rangle \| = \sqrt{\langle x - y | x - y \rangle} \\ d(|x\rangle, |y\rangle) &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + \cdots + (x_n - y_n)^2} . \end{aligned}$$

El teorema del coseno queda como:

$$\sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 + 2 \sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2} \cos(\Theta) ,$$

mientras que el teorema de Pitágoras es:

$$\sum_{i=1}^n (x_i + y_i)^2 = \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 ,$$

es obvio que para \mathbb{R}^2 tanto el teorema del coseno como el teorema de Pitágoras retoman su forma tradicional.

Finalmente la desigualdad de Cauchy-Schwarz se expresa de la siguiente manera:

$$|\langle x | y \rangle| \leq \| |x\rangle \| \| |y\rangle \| \Rightarrow \left| \sum_{i=1}^n x_i y_i \right|^2 \leq \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i^2 .$$

2. Para los espacios de funciones continuas $\mathcal{C}_{[a,b]}^\infty$ una posible definición de **producto interno** sería:

$$\langle f | g \rangle = \int_a^b dx f^*(x) g(x) ,$$

de la cual se deriva la siguiente expresión para la **norma**:

$$\| |f\rangle \|^2 = \langle f | f \rangle = \int_a^b dx |f(x)|^2 .$$

La **distanza** entre funciones quedará definida como:

$$\begin{aligned} d(|f\rangle, |g\rangle) &\equiv \||f\rangle - |g\rangle\| \equiv \sqrt{\langle f - g | f - g \rangle} = \sqrt{\langle f | f \rangle - \langle f | g \rangle - \langle f | g \rangle^* + \langle g | g \rangle} \\ d(|f\rangle, |g\rangle) &= \sqrt{\int_a^b dx |f(x) - g(x)|^2} \\ &= \sqrt{\int_a^b dx |f(x)|^2 - 2 \operatorname{Re} \left(\int_a^b dx f^*(x) g(x) \right) + \int_a^b dx |g(x)|^2}. \end{aligned}$$

El teorema del coseno puede ser escritos como:

$$\begin{aligned} \int_a^b dx |f(x) + g(x)|^2 &= \int_a^b dx |f(x)|^2 + \int_a^b dx |g(x)|^2 \\ &\quad + 2 \left(\int_a^b dx |f(x)|^2 \right)^{\frac{1}{2}} \left(\int_a^b dx |g(x)|^2 \right)^{\frac{1}{2}} \cos(\Theta), \end{aligned}$$

donde:

$$\cos(\Theta) = \frac{\int_a^b dx f^*(x) g(x)}{\left(\int_a^b dx |f(x)|^2 \right)^{\frac{1}{2}} \left(\int_a^b dx |g(x)|^2 \right)^{\frac{1}{2}}}.$$

Y como era de esperarse el teorema de Pitágoras queda:

$$\int_a^b dx |f(x) + g(x)|^2 = \int_a^b dx |f(x)|^2 + \int_a^b dx |g(x)|^2,$$

para funciones $f(x)$ y $g(x)$ ortogonales, mientras que para este caso, la desigualdad de Cauchy-Schwarz se expresa:

$$\left| \int_a^b dx f^*(x) g(x) \right|^2 \leq \int_a^b dx |f(x)|^2 \int_a^b dx |g(x)|^2.$$

Ejemplo 2.7 Consideramos el espacio vectorial de polinomios de grado $g \leq n$ definidos en el intervalo $[0, 1]$ o en el intervalo $[-1, 1]$ según el caso. Suponiendo las siguientes definiciones de producto interno en \mathcal{P}^n :

$$\langle q_n | p_n \rangle = \int_{-1}^1 p(x)q(x)dx \quad \text{y} \quad \langle q_n | p_n \rangle = \int_0^1 p(x)q(x)dx.$$

Vamos a encontrar la distancia y el ángulo entre los vectores: $|x_1\rangle = x(x-1)$ y $|x_2\rangle = x$.

En general, la definición de distancia es:

$$d(|x_1\rangle, |x_2\rangle) = \sqrt{\langle x_2 - x_1 | x_2 - x_1 \rangle},$$

por lo tanto para $\langle q_n | p_n \rangle = \int_{-1}^1 p(x)q(x)dx$ la distancia será:

$$\sqrt{\langle x_2 - x_1 | x_2 - x_1 \rangle} = \sqrt{\int_{-1}^1 [x(x-1) - x]^2 dx} = \frac{1}{15}\sqrt{690},$$

y para $\langle q_n | p_n \rangle = \int_0^1 p(x)q(x)dx$, será:

$$\sqrt{\langle x_2 - x_1 | x_2 - x_1 \rangle} = \sqrt{\int_0^1 (x(x-1) - x)^2 dx} = \frac{2}{15}\sqrt{30}.$$

Con respecto a los ángulos:

$$\theta = \arccos \left(\frac{\langle x_1 | x_2 \rangle}{\sqrt{\langle x_1 | x_1 \rangle} \sqrt{\langle x_2 | x_2 \rangle}} \right).$$

Para $\langle q_n | p_n \rangle = \int_{-1}^1 p(x)q(x)dx$ tenemos:

$$\theta = \arccos \left(\frac{\langle x_1 | x_2 \rangle}{\sqrt{\langle x_1 | x_1 \rangle} \sqrt{\langle x_2 | x_2 \rangle}} \right) = \arccos \left(\frac{\int_{-1}^1 (x(x-1)) x \, dx}{\sqrt{\int_{-1}^1 (x(x-1))^2 \, dx} \sqrt{\int_{-1}^1 x^2 \, dx}} \right)$$

$$= \arccos \left(-\frac{1}{12} \sqrt{15} \sqrt{6} \right) = 2,4825 \text{ rad.}$$

Para $\langle q_n | p_n \rangle = \int_0^1 p(x)q(x)dx$

$$\theta = \arccos \left(\frac{\langle x_1 | x_2 \rangle}{\sqrt{\langle x_1 | x_1 \rangle} \sqrt{\langle x_2 | x_2 \rangle}} \right) = \arccos \left(\frac{\int_0^1 (x(x-1)) (x) \, dx}{\sqrt{\int_0^1 (x(x-1))^2 \, dx} \sqrt{\int_0^1 x^2 \, dx}} \right)$$

$$= \arccos \left(-\frac{1}{12} \sqrt{15} \sqrt{2} \right) = 2,4825 \text{ rad} \quad \text{¡El mismo ángulo!}$$

△

2.2.4 Ejercicios

1. Consideremos el espacio vectorial conformado por los vectores geométricos en \mathbb{R}^3 ¿Serán espacios euclidianos para las siguientes definiciones de producto interno?

- (a). El producto de las longitudes de los vectores.
- (b). El producto de las longitudes por el cubo del coseno del ángulo entre ellos.
- (c). El producto como dos veces el producto escalar usual entre vectores.

2. Considerando estas definiciones de producto interno en \mathcal{P}_n :

$$a) \quad \langle q_n | p_n \rangle = \int_{-1}^1 p(x)q(x)dx, \quad b) \quad \langle q_n | p_n \rangle = \int_0^1 p(x)q(x)dx.$$

- (a). Encuentre los ángulos en el “triángulo” formado por los vectores: $|x_1\rangle = 1, |x_2\rangle = t, |x_3\rangle = 1-t$.
- (b). Encuentre la distancia y el ángulo entre los siguientes pares de vectores en \mathcal{P}_3 :
 - I. $|x_1\rangle = 1; |x_2\rangle = x$.
 - II. $|x_1\rangle = 2x; |x_2\rangle = x^2$.

3. Sea \mathbf{E}' un subespacio euclidiano de dimensión k , $\mathbf{E}' \subset \mathbf{E}$, y sea $|v\rangle$ un vector que no necesariamente es un elemento \mathbf{E}' . Podemos plantearnos el problema de representar $|v\rangle$ de la forma: $|v\rangle = |g\rangle + |h\rangle$; donde $|g\rangle \in \mathbf{E}'$ y $|h\rangle$ es ortogonal a $|g\rangle$. La existencia de la expansión anterior nos muestra que el espacio total \mathbf{E} , de dimensión n , es la suma directa de los subespacios \mathbf{E}' y su complemento ortogonal \mathbf{E}'^\perp de dimensión $n-k$.

Nota: El conjunto de los vectores de \mathbf{E} que son ortogonales a todos los de \mathbf{E}' :

$$\mathbf{E}'^\perp = \{|v\rangle \in \mathbf{E} \mid \langle v | g \rangle = 0 \ \forall |g\rangle \in \mathbf{E}'\}$$

será un subespacio de \mathbf{E} y se denomina complemento ortogonal de \mathbf{E}' .

Encuentre el vector $|v\rangle$, como la suma del vector $|g\rangle$, expandido por los vectores $|g_i\rangle$, y el vector perpendicular $|h\rangle$ cuando:

- (a). $|h\rangle = (5, 2, -2, 2), |g_1\rangle = (2, 1, 1, -\alpha), |g_2\rangle = (1, \beta, 3, 0)$.
- (b). $|h\rangle = (-3, 5, 9, 3), |g_1\rangle = (1, 1, 1, \gamma), |g_2\rangle = (2\eta, -1, 1, 1), |g_3\rangle = (2, -7\delta, -1, -1)$.

4. Considere el espacio vectorial de matrices complejas 2×2 con la siguiente definición

$$\langle a | b \rangle \Rightarrow \text{Tr}(\mathbb{A}^\dagger \mathbb{B}) \equiv (A^\dagger)_j^i B_i^j \equiv (A^*)_i^j B_i^j.$$

Compruebe si ésta es una buena definición de producto interno.

Nota: Las matrices complejas 2×2 están conformadas por cuatro números complejos, $z_1 = x_1 + iy_1$, $z_2 = x_2 + iy_2$, $z_3 = x_3 + iy_3$, y $z_4 = x_4 + iy_4$. Entonces, hemos denotado por $(A^\dagger)_j^i$ la matriz adjunta de A_j^i . Vale decir, una matriz adjunta de otra será su traspuesta conjugada. Por traspuesta entendemos cambiar filas por columnas manteniendo intacta la diagonal, $A_j^i \rightarrow A_i^j$ (ver sección 4.4.3). Por otro lado, hemos definido como $\text{Tr}(\mathbb{A})$ la traza de una matriz como la suma de los elementos de la diagonal, $\text{Tr}(\mathbb{A}) = A_i^i$. A continuación ejemplificamos estas dos definiciones:

$$\mathbb{A} \leftrightarrow \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} \Rightarrow \mathbb{A}^\dagger \leftrightarrow \begin{pmatrix} z_1^* & z_3^* \\ z_2^* & z_4^* \end{pmatrix} \quad \text{y} \quad \text{Tr}(\mathbb{A}) = z_1 + z_2.$$

Ambas definiciones serán discutidas con rigor en las secciones 4.2.3 y 4.3.4, respectivamente. Sin embargo, este par de definiciones funcionales –la adjunta de una matriz es su traspuesta conjugada y su traza será la suma de los elementos de la diagonal– serán suficiente para desarrollar este ejercicio.

5. Sean $|p_n\rangle = p(x) = \sum_{i=0}^{n-1} a_i x^i$; $|q_n\rangle = q(x) = \sum_{i=0}^{n-1} b_i x^i \in \mathcal{P}_n$ y considere la siguiente definición:

$$\langle q_n | p_n \rangle \Rightarrow a_0 b_0 + a_1 b_1 + a_2 b_2 + \dots + a_{n-1} b_{n-1} = \sum_{i=0}^{n-1} a_i b_i$$

(a). Muestre que ésta es una buena definición de producto interno.

(b). Con esta definición de producto interior ¿Se puede considerar \mathcal{P}_n un subespacio de $\mathcal{C}_{[a,b]}$? ¿Por qué?

6. Los vectores en \mathbb{R}^3 en coordenadas cartesianas los definimos como $\mathbf{a} = a^i |e_i\rangle = a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} + a_z \hat{\mathbf{k}}$ y definimos una “tabla de multiplicación” entre ellos de la forma $\langle e^i | e_j \rangle = \delta_j^i$ con $i, j = 1, 2, 3$, esto es:

$\langle e^i e_j \rangle$	$\hat{\mathbf{i}}$	$\hat{\mathbf{j}}$	$\hat{\mathbf{k}}$
$\hat{\mathbf{i}}$	1	0	0
$\hat{\mathbf{j}}$	0	1	0
$\hat{\mathbf{k}}$	0	0	1

Un cuaternión cartesiano puede escribirse de manera análoga a los vectores cartesianos, vale decir:

$$|a\rangle = a^\alpha |q_\alpha\rangle = a^0 + a^i |q_i\rangle = a_0 + a_x \hat{\mathbf{i}} + a_y \hat{\mathbf{j}} + a_z \hat{\mathbf{k}},$$

con $\alpha = 0, 1, 2, 3$ y donde las a^i (con $i = 1, 2, 3$) son números reales que representan las componentes vectoriales en coordenadas cartesianas de los cuaterniones, mientras que la a^0 , también un número real se le llama componente escalar¹¹.

Los cuaterniones fueron inventados por el matemático irlandés William Rowan Hamilton a mediados del siglo XIX, y por decirlo de alguna manera, son híbridos o generalizaciones a un plano hipercomplejo.

Un vector cartesiano es un cuaternión con la componente escalar nula. Hoy encontramos aplicaciones del álgebra de cuaterniones en Física¹² y más recientemente ha tenido alguna utilización en computación

¹¹Recuerde que estamos utilizando la convención de Einstein: $c^\alpha |q_\alpha\rangle \equiv c^0 + \sum_{j=1}^3 c^j |q_j\rangle$. Es decir, hemos supuesto que: $|q_0\rangle \equiv 1$, la unidad en los números reales. Adicionalmente, nótese que los índices griegos α, β, \dots toman los valores 0, 1, 2, 3, mientras que los latinos que acompañan a los vectores cartesianos toman los siguientes valores $j, k, l = 1, 2, 3$.

¹²Hace algunas décadas se dio una discusión sobre la importancia de utilizar esta representación en Física Cuántica. Pueden consultar:

- Berezin, A. V., Kurochkin, Y. A., & Tolkachev, E. A. (1989). Quaternions in relativistic physics. Nauka i Tekhnika, Minsk.
- Girard, P. R. (1984). The quaternion group and modern physics. European Journal of Physics, 5(1), 25.
- Horwitz, L. P., & Biedenharn, L. C. (1984). Quaternion quantum mechanics: second quantization and gauge fields. Annals of Physics, 157(2), 432-488.

gráfica que discutiremos en el próximo problema dentro del contexto del álgebra geométrica y las álgebras de Grassman.

Basándonos en este esquema podemos definir la “tabla de multiplicación” para los cuaterniones cartesianos como:

$ q_i\rangle \odot q_j\rangle$	1	$ q_1\rangle$	$ q_2\rangle$	$ q_3\rangle$
1	1	$ q_1\rangle$	$ q_2\rangle$	$ q_3\rangle$
$ q_1\rangle$	$ q_1\rangle$	-1	$ q_3\rangle$	$- q_2\rangle$
$ q_2\rangle$	$ q_2\rangle$	$- q_3\rangle$	-1	$ q_1\rangle$
$ q_3\rangle$	$ q_3\rangle$	$ q_2\rangle$	$- q_1\rangle$	-1

Nótese que por el hecho de que:

$$|q_j\rangle \odot |q_j\rangle = -1 \Rightarrow |q_1\rangle \odot |q_1\rangle = |q_2\rangle \odot |q_2\rangle = |q_3\rangle \odot |q_3\rangle = -1,$$

se puede pensar que un cuaternión es la generalización de los números complejos a más de una dimensión (un número hipercomplejo), donde la parte imaginaria tendría tres dimensiones y no una como es costumbre.

Esto es,

$$|a\rangle = a^\alpha |q_\alpha\rangle = a^0 \underbrace{|q_0\rangle}_1 + a^j |q_j\rangle = a^0 + \underbrace{a^1 |q_1\rangle + a^2 |q_2\rangle + a^3 |q_3\rangle}_{\text{“parte compleja”}}.$$

Siendo consistente con esa visión de generalización de un número complejo, definiremos el conjugado de un cuaternión como:

$$|b\rangle^* = b^0 |q_0\rangle - b^j |q_j\rangle,$$

con $j = 1, 2, 3$.

Es decir, en analogía con los números complejos el conjugado de un cuaternión cambia el signo de su “parte compleja vectorial”.

Igualmente, definiremos la suma entre cuaterniones de la siguiente manera:

$$\left. \begin{array}{l} |a\rangle = a^\alpha |q_\alpha\rangle \\ |b\rangle = b^\alpha |q_\alpha\rangle \end{array} \right\} \Rightarrow |c\rangle = c^\alpha |q_\alpha\rangle = |a\rangle + |b\rangle = (a^\alpha + b^\alpha) |q_\alpha\rangle \Rightarrow c^\alpha = (a^\alpha + b^\alpha).$$

Esto quiere decir que los vectores se suman componente a componente. Mientras que la multiplicación por un escalar queda definida por $\alpha |c\rangle = \alpha c^\alpha |q_\alpha\rangle$, es decir se multiplica el escalar por cada componente. Con la información anterior, responda las siguientes preguntas:

- (a). Compruebe si los cuaterniones, $|a\rangle$, forman un espacio vectorial respecto a esa operación suma y esa multiplicación por escalares, análoga a la de los vectores en \mathbb{R}^3 en coordenada cartesianas.
- (b). Dados dos cuaterniones cualesquiera $|b\rangle \equiv (b^0, \mathbf{b})$ y $|r\rangle \equiv (r^0, \mathbf{r})$, y su tabla de multiplicación, muestre que el producto entre esos cuaterniones $|d\rangle = |b\rangle \odot |r\rangle$ podrá representarse como:

$$|d\rangle = |b\rangle \odot |r\rangle \longleftrightarrow (d^0, \mathbf{d}) = (b^0 r^0 - \mathbf{b} \cdot \mathbf{r}, r^0 \mathbf{b} + b^0 \mathbf{r} + \mathbf{b} \times \mathbf{r}),$$

donde \cdot y \times corresponden con los productos escalares y vectoriales tridimensionales de siempre.

- (c). Ahora con índices: dados $|b\rangle = b^\alpha |q_\alpha\rangle$ y $|r\rangle = r^\alpha |q_\alpha\rangle$, compruebe si el producto $|d\rangle = |b\rangle \odot |r\rangle$ puede ser siempre escrito de la forma:

$$|d\rangle = |b\rangle \odot |r\rangle = a |q_0\rangle + S^{(\alpha j)} \delta_\alpha^0 |q_j\rangle + A^{[jk]i} b_j r_k |q_i\rangle.$$

donde a representa un número, $S^{(\alpha j)} \delta_\alpha^0$ (recuerde que los índices latinos toman los valores $j, k, l = 1, 2, 3$).

1, 2, 3, mientras $\alpha = 0, 1, 2, 3$), donde $S^{(ij)}$ indica $S^{ji} = S^{ij}$, que la cantidad S^{ij} es simétrica, y por lo tanto $(S^{\alpha j}\delta_\alpha^0 + S^{j\alpha}\delta_\alpha^0)|q_j\rangle$.

Mientras $A^{[jk]i}$ representa un conjunto de objetos antisimétricos en j y k :¹³

$$A^{[jk]i} \rightarrow A^{jki} = -A^{kji} \rightarrow (A^{jki}b_jr_k - A^{kji}b_jr_k)|q_i\rangle .$$

(d). Identifique las cantidades: a , $S^{(ij)}$ y $A^{[jk]i}$, en términos de las componentes de los cuaterniones.

¿El producto de cuaterniones $|d\rangle = |a\rangle \odot |r\rangle$ será un vector, pseudovector o ninguna de las anteriores? Explique por qué.

(e). Compruebe si las matrices de Pauli y la identidad

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_0 \equiv \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

pueden representar la base de los cuaterniones $\{|q_1\rangle, |q_2\rangle, |q_3\rangle, |q_0\rangle\}$. Seguidamente muestre que matrices complejas 2×2 del tipo:

$$|b\rangle \longleftrightarrow \begin{pmatrix} z & w \\ -w^* & z^* \end{pmatrix},$$

pueden ser consideradas como cuaterniones, donde $z = x + iy$ y $w = a + ib$ son números complejos.

Las Matrices de Pauli aparecen en mecánica cuántica cuando se tiene en cuenta la interacción del espín de una partícula con un campo electromagnético externo y en estas notas las consideraremos en varios momentos (ver, por ejemplo los ejercicios de las secciones 4.3.9 y 4.7.6)

(f). Muestre que una representación posible para la base de cuaterniones son: la matriz identidad y las matrices reales 4×4 de la forma:

$$|q_1\rangle = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad |q_2\rangle = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad |q_3\rangle = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

(g). Compruebe si la siguiente es una buena definición de producto interno:

$$\widetilde{\langle a | b \rangle} = |a\rangle \boxtimes \odot |b\rangle .$$

(h). Modifique un poco la definición anterior de tal forma que:

$$\langle a | b \rangle = \frac{1}{2} [\widetilde{\langle a | b \rangle} - |q_1\rangle \odot \widetilde{\langle a | b \rangle} \odot |q_1\rangle],$$

y compruebe si esta definición compleja del producto interno cumple con todas las propiedades.

Nótese que un cuaternion de la forma $|f\rangle = f^0 + f^1|q_1\rangle$ es un número complejo convencional.

(i). Compruebe si la siguiente es una buena definición de norma para los cuaterniones:

$$n(|b\rangle) = \| |a\rangle \| = \sqrt{\langle a | a \rangle} = \sqrt{|a\rangle \boxtimes \odot |a\rangle} .$$

(j). Compruebe si un cuaternion definido por:

$$\overline{|a\rangle} = \frac{|a\rangle \boxtimes}{\| |a\rangle \|^2},$$

puede ser considerado como el inverso o elemento simétrico de $|a\rangle$, respecto a la multiplicación \odot .

(k). Compruebe si los cuaterniones $|a\rangle$ forman un grupo respecto a una operación multiplicación \odot .

(l). Los vectores en \mathbb{R}^3 en coordenadas cartesianas, $|v\rangle$, pueden ser representados como cuaterniones, donde la parte escalar es nula $v^0 = 0 \rightarrow |v\rangle = v^j|q_j\rangle$. Compruebe si el siguiente producto conserva

¹³Para familiarizarse con las expresiones vectoriales con la notación de índices puede consultar la sección 1.4.

la norma:

$$|v'\rangle = \overline{|a\rangle} \odot |v\rangle \odot |a\rangle .$$

Estos es: $\|v'\|^2 = (v'^1)^2 + (v'^2)^2 + (v'^3)^2 \equiv (v^1)^2 + (v^2)^2 + (v^3)^2 = \|v\|^2$.

7. En el mismo espíritu de los cuaterniones considerados previamente estudiemos el siguiente problema.

Consideremos otra vez el espacio \mathbb{R}^3 expandido por la base ortonormal estándar $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$. Supongamos en este espacio el producto de dos vectores \mathbf{a} y \mathbf{b} (o $|a\rangle$ y $|b\rangle$) en la notación de vectores abstractos de Dirac) definido a la manera del *álgebra geométrica*. Esto es:

$$|a\rangle \odot |b\rangle \equiv \mathbf{ab} = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \wedge \mathbf{b},$$

con $\mathbf{a} \cdot \mathbf{b}$ el producto escalar estándar de \mathbb{R}^3 , representando la parte comutativa de \mathbf{ab} y $\mathbf{a} \wedge \mathbf{b}$ su parte anticomutativa. Esta última parte se relaciona con el producto vectorial estándar de la representación de Gibbs como: $\mathbf{a} \wedge \mathbf{b} = i\mathbf{a} \times \mathbf{b}$, con $i = \hat{\mathbf{i}} \wedge \hat{\mathbf{j}} \cdot \hat{\mathbf{k}}$ un pseudoescalar.

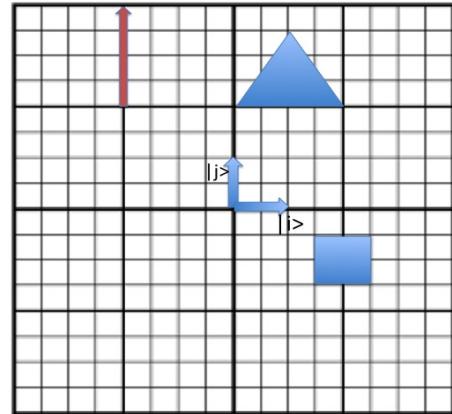


Figura 2.2: Se ilustran los vectores ortonormales estándares $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}\}$ y como siempre $\hat{\mathbf{k}} = \hat{\mathbf{i}} \times \hat{\mathbf{j}}$.

Este formalismo ha tenido cierto impacto en computación gráfica¹⁴ y lo vamos a utilizar para modelar transformaciones de objetos en el espacio. Considere el caso 2D representado en la figura 2.2, en el formalismo de álgebra geométrica el objeto triángulo lo asociamos con vector abstracto $|\Delta\rangle$, mientras que el objeto cuadrado lo representaremos como $|\square\rangle$, y el tercer objeto por $|\uparrow\rangle$.

Entonces:

- (a). Exprese los vectores: $|\Delta\rangle, |\square\rangle, |\uparrow\rangle$ en términos de la base ortonormal estándar $\{\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}\}$ (o $|i\rangle, |j\rangle, |k\rangle$) en el formalismo de álgebra geométrica.
- (b). Exprese $|\Delta\rangle \odot |\uparrow\rangle$ en término de la base geométrica. Vale decir, si $|O_g\rangle$ es un objeto geométrico este podrá representarse en términos de una base: $|\epsilon\rangle, |\sigma\rangle, |B\rangle, |i\rangle$, donde $|\epsilon\rangle$ es un escalar; $|\sigma\rangle$ es un vector, $|B\rangle$ un bivector y, finalmente $|i\rangle$ un pseudoescalar.
- (c). Encuentre la norma de cada uno de ellos y la distancia entre $|\Delta\rangle$ y $|\uparrow\rangle$.
- (d). Considere ahora la operación: $\mathbb{A}|\Delta\rangle = |\tilde{\Delta}\rangle$.

I. Si $\mathbb{A}_{|j\rangle}$ es el operador reflexión respecto $|j\rangle$, exprese en términos geométricos $|\tilde{\Delta}\rangle = \mathbb{A}_{|j\rangle}|\Delta\rangle$ y $\mathbb{A}_{|i\rangle}(|\Delta\rangle \odot |\uparrow\rangle)$.

¹⁴Esta representación para objetos físicos ha tenido cierto éxito en Física clásica y cuántica; se puede consultar:

- Hestenes, D., & Sobczyk, G. (2012). Clifford algebra to geometric calculus: a unified language for mathematics and physics (Vol. 5). Springer Science & Business Media.
 - Hestenes, D. (1971). Vectors, spinors, and complex numbers in classical and quantum physics. Am. J. Phys., 39(9), 1013-1027.
 - Hestenes, D. (2003). Spacetime physics with geometric algebra. American Journal of Physics, 71(7), 691-714.
 - Dressel, J., Bliokh, K. Y., & Nori, F. (2015). Spacetime algebra as a powerful tool for electromagnetism. Physics Reports, 589, 1-71.
 - Goldman, R. (2002). On the algebraic and geometric foundations of computer graphics. ACM Transactions on Graphics (TOG), 21(1), 52-86.
 - Hildenbrand (2011). From Grassmann's vision to geometric algebra computing. Springer Basel.
 - Hildenbrand, D., Fontijne, D., Perwass, C., & Dorst, L. (2004). Geometric algebra and its application to computer graphics. In Tutorial notes of the EUROGRAPHICS conference.
 - Vince, J. (2008). Geometric algebra for computer graphics. Springer Science & Business Media
- y las referencias allí citadas.

- II. Si $\mathbb{A}_{\theta,|B\rangle}$ es el operador de rotación alrededor de un bivector $|B\rangle$ un ángulo θ . Encuentre $\mathbb{A}_{\pi/4,|B\rangle}(|\Delta\rangle \odot |\uparrow\rangle)$ con $|B\rangle = |j\rangle + |k\rangle$.
- III. ¿Cómo interpreta Ud. la ecuación de autovalores $\mathbb{A}|\Delta\rangle = 2|\Delta\rangle$?
- (e). Considere el caso 3D en el cual $|\hat{\Delta}\rangle$ representa un tetraedro regular con la base representada por la figura 2.2 y $|\hat{\square}\rangle$, un cubo, también con su base representada en el plano de la misma figura.
- I. Exprese los vectores: $|\hat{\Delta}\rangle, |\hat{\square}\rangle$ en términos de la base ortonormal estándar $\{\hat{i}, \hat{j}, \hat{k}\}$ y el formalismo de álgebra geométrica.
 - II. Exprese $(|\hat{\Delta}\rangle \odot |\uparrow\rangle) \odot |\Delta\rangle$ en término de la base geométrica.
 - III. Encuentre la norma de $|\hat{\Delta}\rangle$ y la distancia entre $|\hat{\Delta}\rangle$ y $\mathbb{A}_{\pi/2,|-j\rangle}|\hat{\square}\rangle$.
 - IV. Exprese $[\mathbb{A}_{\pi/2,|-j\rangle}, \mathbb{B}_{\pi/4,|i\rangle}]|\hat{\square}\rangle$, con $[\mathbb{A}_{\pi/2,|-j\rangle}, \mathbb{B}_{\pi/4,|i\rangle}] = \mathbb{A}_{\pi/2,|-j\rangle}\mathbb{B}_{\pi/4,|i\rangle} - \mathbb{B}_{\pi/4,|i\rangle}\mathbb{A}_{\pi/2,|-j\rangle}$.
8. En Geometría Diferencial podemos considerar a \mathbb{R}^3 como un conjunto de puntos (una variedad o espacio topológico), es decir, \mathbb{R}^3 no es un espacio vectorial como se definió con anterioridad. En un punto q cualquiera podemos generar un plano \mathbb{R}^2 , entendido también como un conjunto de puntos, y definir el siguiente conjunto $T_q\mathbb{R}^2 = \{\text{el conjunto de todos los vectores geométricos (flechas) con origen en } q \text{ que son tangentes al plano } \mathbb{R}^2\}$. Notemos que la idea la podemos extender para cualquier superficie, por ejemplo, una esfera \mathbb{S}^2 (embebida en \mathbb{R}^3) y sobre la esfera seleccionar un punto arbitrario q . A partir de este punto q generar un plano y construir el conjunto $T_q\mathbb{S}^2 = \{\text{el conjunto de todos los vectores geométricos con origen en } q \text{ que son tangentes a la esfera } \mathbb{S}^2\}$.
- (a). ¿Es $T_q\mathbb{R}^2$ un espacio vectorial?
 - (b). Claramente podemos seleccionar otro punto arbitrario, pero diferente de q , digamos p y construir el conjunto $T_p\mathbb{R}^2$. ¿Son estos dos espacios isomorfos?
9. Consideremos ahora el conjunto de todas las funciones diferenciables $f : \mathbb{R}^3 \rightarrow \mathbb{R}$. Esto es, que existen todas las derivadas parciales en $q \in \mathbb{R}^3$:

$$\partial_x f(x^i), \partial_y f(x^i), \partial_z f(x^i) \in \mathbb{R} \quad \forall (x^i) \in \mathbb{R}^3.$$

Note que hemos supuesto un sistema de coordenadas cartesiano $x^i = (x, y, z)$ en \mathbb{R}^3 y que las derivadas son evaluadas en q .

Consideremos el espacio tangente $T_q\mathbb{R}^3$ en un punto q arbitrario y un vector $|u_q\rangle \in T_q\mathbb{R}^3$ con componentes (u^1, u^2, u^3) , podemos definir el siguiente operador (operador derivada direccional) en q siguiendo a $|u_q\rangle$:

$$|U_q\rangle = (u^1 \partial_x + u^2 \partial_y + u^3 \partial_z)_q.$$

Por lo tanto:

$$|U_q\rangle f \equiv (u^1 \partial_x + u^2 \partial_y + u^3 \partial_z)_q f = u^1 (\partial_x f)_q + u^2 (\partial_y f)_q + u^3 (\partial_z f)_q.$$

Es posible entonces construir el conjunto de los operadores derivadas direccionales que actúan sobre las funciones f :

$$\mathcal{D}_q(\mathbb{R}^3) = \{|U_q\rangle = (u^1 \partial_x + u^2 \partial_y + u^3 \partial_z)_q \quad \forall |u_q\rangle \in T_q\mathbb{R}^3\}.$$

- c) ¿Es \mathcal{D}_q un espacio vectorial?
 d) ¿Los espacios \mathcal{D}_q y $T_q\mathbb{R}^3$ son isomorfos?

Practicando con Maxima

1. Espacios y subespacios vectoriales Sea el espacio vectorial $V = K^n$, definido en $K = \mathbb{R}$ y donde n es un entero positivo. Consideremos el caso $n = 4$. El producto de un elemento de K^4 , digamos $|x\rangle = (x_1, x_2, x_3, x_4)$ por un escalar $\alpha \in K$ resulta en otro elemento de K^4 .

Primero introducimos los elementos como listas:

```
(%ci1) X:[x1,x2,x3,x4]; Y:[y1,y2,y3,y4]; Z:[z1,z2,z3,z4];
(%o1) [x1, x2, x3, x4]
(%o2) [y1, y2, y3, y4]
(%o3) [z1, z2, z3, z4]

(%ci4) alpha*X=Y;
(%o4) [\alpha x1, \alpha x2, \alpha x3, \alpha x4] = [y1, y2, y3, y4]
```

El resultado es un elemento del espacio vectorial K^4 .

La suma de $|x\rangle = (x_1, x_2, x_3, x_4)$ y $|y\rangle = (y_1, y_2, y_3, y_4)$ será:

```
(%ci5) X+Y=Z;
(%o5) [y1 + x1, y2 + x2, y3 + x3, y4 + x4] = [z1, z2, z3, z4]
```

con $(z_1, z_2, z_3, z_4) \in K^4$.

Podemos ver rápidamente que el conjunto de vectores que tienen la forma $(x_1, x_2, x_3, 0)$ conforman un subespacio de K^4 , ya que:

```
(%ci6) map(":",[x4,y4,z4],[0,0,0]);
(%o6) [0, 0, 0]

(%ci7) X:[x1,x2,x3,x4];Y:[y1,y2,y3,y4];Z:[z1,z2,z3,z4];
(%o7) [x1, x2, x3, 0]
(%o8) [y1, y2, y3, 0]
(%o9) [z1, z2, z3, 0]

(%ci10) alpha*X+beta*Y=Z;
(%o10) [\beta y1 + \alpha x1, \beta y2 + \alpha x2, \beta y3 + \alpha x3, 0] = [z1, z2, z3, 0]
```

Para recobrar las variables x_4, y_4, z_4 escribimos:

```
(%ci11) kill(x4,y4,z4)$
(%ci12) X:[x1,x2,x3,x4]; Y:[y1,y2,y3,y4]; Z:[z1,z2,z3,z4];
(%o12) [x1, x2, x3, x4]
(%o13) [y1, y2, y3, y4]
(%o14) [z1, z2, z3, z4]
```

Para calcular el producto interno entre vectores es necesario utilizar la biblioteca **eigen**.

```
(%ci15) load("eigen")$
(%ci16) innerproduct(X,Y);
(%o16) x4 y4 + x3 y3 + x2 y2 + x1 y1
```

Consideremos ahora $\mathbf{V} = \mathbf{K}^n$, definido en $\mathbf{K} = \mathbb{C}$, con $n = 3$. Por lo tanto, los vectores serán ahora de la siguiente forma: $z = (x_1 + iy_1, x_2 + iy_2, x_3 + iy_3)$.

(%ci17) Z1:[x1+ %i*y1,x2+ %i*y2,x3+ %i*y3]; Z2:[u1+ %i*v1,u2+ %i*v2,u3+ %i*v3];
 (%o17) [i y₁ + x₁, i y₂ + x₂, i y₃ + x₃]
 (%o18) [i v₁ + u₁, i v₂ + u₂, i v₃ + u₃]

Y los escalares de la forma $\alpha = a + ib$.

(%ci19) alpha:a+ %i*b;
 (%o19) i b + a

El producto por el escalar α es:

(%ci20) Z3:alpha*Z1;
 (%o20) [(i b + a) (i y₁ + x₁), (i b + a) (i y₂ + x₂), (i b + a) (i y₃ + x₃)]
 (%ci21) map(rectform,Z3);
 (%o21) [i (a y₁ + b x₁) - b y₁ + a x₁, i (a y₂ + b x₂) - b y₂ + a x₂, i (a y₃ + b x₃) - b y₃ + a x₃]
 (%ci22) map(realpart,Z3),factor; map(imagpart,Z3),factor;
 (%o22) [-(b y₁ - a x₁), -(b y₂ - a x₂), -(b y₃ - a x₃)]
 (%o23) [a y₁ + b x₁, a y₂ + b x₂, a y₃ + b x₃]

Calculemos ahora el producto interno:

$$Z_1 \cdot Z_2 = (x_1 + iy_1)^*(u_1 + iv_1) + (x_2 + iy_2)^*(u_2 + iv_2) + (x_3 + iy_3)^*(u_3 + iv_3).$$

(%ci24) Z4:innerproduct(Z1,Z2);
 (%o24) (v₃ - iu₃) y₃ + (v₂ - iu₂) y₂ + (v₁ - iu₁) y₁ + (iv₃ + u₃) x₃ + (iv₂ + u₂) x₂ + (iv₁ + u₁) x₁
 (%ci25) Re:map(realpart,Z4)\$ Im:map(imagpart,Z4)\$
 (%ci26) Re+ %i*Im;
 (%o26) i (-u₃y₃ - u₂y₂ - u₁y₁ + v₃x₃ + v₂x₂ + v₁x₁) + v₃y₃ + v₂y₂ + v₁y₁ + u₃x₃ + u₂x₂ + u₁x₁
 (%ci27) kill(all)\$

2. **Producto de polinomios** Consideremos el siguiente producto escalar entre elementos de un espacio vectorial de polinomios:

$$\langle p_i | p_j \rangle = \int_a^b p_i(x)p_j(x)dx.$$

Vamos a encontrar la distancia y el ángulo entre los vectores $|x_1\rangle = x(x-1)$ y $|x_2\rangle = x$ en dos intervalos diferentes: $[0, 1]$ y $[-1, 1]$

Debemos introducir los objetos a multiplicar:

(%ci1) P1:x*(x-1); P2:x;
 (%o1) (x - 1) x
 (%o2) x

Ahora calculamos las distancias entre los vectores para ambos intervalos. Haremos gala de algunas posibilidades que ofrece el programa para escribir las expresiones.

(%ci3) $\text{sqrt}(\text{integrate}(((P1-P2)^2),x,-1,1))=\text{sqrt}(\text{integrate}(((P1-P2)^2),x,-1,1));$
 (%o3) $\sqrt{\int_{-1}^1 ((x - 1) x - x)^2 dx} = \frac{\sqrt{46}}{\sqrt{15}}$
 (%ci4) $\text{sqrt}(\text{integrate}(((P1-P2)^2),x,0,1))=\text{sqrt}(\text{integrate}(((P1-P2)^2),x,0,1));$
 (%o4) $\sqrt{\int_0^1 ((x - 1) x - x)^2 dx} = \frac{2^{\frac{3}{2}}}{\sqrt{15}}$

El ángulo entre los polinomios definidos en el intervalo $[-1, 1]$ es:

(%ci5) $\text{'integrate}((P1*P2),x,-1,1)/(\text{sqrt}(\text{'integrate}((P1*P1),x,-1,1))*\text{sqrt}(\text{'integrate}((P2*P2),x,-1,1)));$
 (%o5) $\frac{\int_{-1}^1 (x - 1) x^2 dx}{\sqrt{\int_{-1}^1 x^2 dx} \sqrt{\int_{-1}^1 (x - 1)^2 x^2 dx}}$
 (%ci6) $\text{ev}(%,\text{integrate});$
 (%o6) $-\frac{\sqrt{15}}{2^{\frac{3}{2}} \sqrt{3}}$
 (%ci7) $\text{acos}(%),\text{numer};$
 (%o7) 2,482534617763384

Y ahora, el ángulo entre los polinomios definidos en el intervalo $[0, 1]$:

(%ci8) $\text{'integrate}((P1*P2),x,0,1)/(\text{sqrt}(\text{'integrate}((P1*P1),x,0,1))*\text{sqrt}(\text{'integrate}((P2*P2),x,0,1)));$
 (%o8) $\frac{\int_0^1 (x - 1) x^2 dx}{\sqrt{\int_0^1 x^2 dx} \sqrt{\int_0^1 (x - 1)^2 x^2 dx}}$
 (%ci9) $\text{ev}(%,\text{integrate});$
 (%o9) $-\frac{\sqrt{30}}{4 \sqrt{3}}$
 (%ci10) $\text{acos}(%),\text{numer};$
 (%o10) 2,482534617763384

2.3 Variedades lineales

Si tenemos una cantidad de vectores: $\{|w_1\rangle, |w_2\rangle, |w_3\rangle, \dots\} \in V$, entonces una variedad lineal generada por este conjunto se entenderá como el conjunto de todas las combinaciones lineales finitas:

$$\alpha |w_1\rangle + \beta |w_2\rangle + \gamma |w_3\rangle + \dots,$$

donde los coeficientes: $\alpha, \beta, \gamma, \dots$ pertenecen al campo K .

Se puede comprobar que esta variedad lineal es un subespacio de V . Claramente, todo subespacio que contenga los vectores $\{|w_1\rangle, |w_2\rangle, |w_3\rangle, \dots\}$ también contiene todas sus combinaciones lineales, por lo tanto, la variedad lineal generada de esta manera es el subespacio más pequeño que contiene al conjunto de estos vectores.

Una variedad lineal sencilla de construir en $V = \mathbb{R}^3$ la constituye un par de vectores no colineales. La variedad lineal generada por estos dos vectores será el conjunto de todos los vectores paralelos al plano determinado por este par de vectores no colineales. Mientras que la variedad lineal generada por los vectores: $\{\hat{i}, \hat{j}, \hat{k}\} \in \mathbb{R}^3$ es el mismo espacio entero \mathbb{R}^3 .

Pasemos ahora a considerar el problema de construir una base para una variedad lineal y a partir de allí determinar la dimensión de la variedad.

2.3.1 Dependencia/Independencia lineal

Siguiendo la misma línea de razonamiento que en las secciones 1.1.3 y 1.2.5, generalizamos el concepto de dependencia e independencia lineal de \mathbb{R}^2 y \mathbb{R}^3 .

Dada la ecuación

$$|0\rangle = C_1 |v_1\rangle + C_2 |v_2\rangle + C_3 |v_3\rangle + \dots + C_n |v_n\rangle = \sum_{i=1}^n C_i |v_i\rangle, \quad (2.7)$$

las cantidades C_i son llamados los coeficientes de la combinación lineal.

Podemos afirmar que:

- Si esta ecuación se cumple para algún conjunto de $\{C_i\}$ no nulos, se dirá que el conjunto de vectores correspondiente $\{|v_i\rangle\}$ son **linealmente dependientes**.
- Por el contrario, si esta ecuación **sólo** puede ser satisfecha para todos los $C_i = 0$, entonces se dirá que el conjunto de vectores correspondiente $\{|v_i\rangle\}$ son **linealmente independientes**.

Notemos que la suma que aparece aquí es necesariamente una suma finita, y cuando un determinado conjunto de vectores es linealmente dependiente, entonces uno de ellos se puede escribir como combinación lineal de los demás.

2.3.2 Bases de un espacio vectorial

Ahora bien, dado un espacio vectorial $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$, si encontramos que el conjunto de $\{|v_n\rangle\}$ es linealmente dependiente, entonces siempre es posible despejar uno de los vectores en términos de los demás, vale decir:

$$|v_n\rangle = \bar{C}_1 |v_1\rangle + \bar{C}_2 |v_2\rangle + \bar{C}_3 |v_3\rangle + \dots + \bar{C}_{n-1} |v_{n-1}\rangle = \sum_{i=1}^{n-1} \bar{C}_i |v_i\rangle.$$

Seguidamente podemos proceder a comprobar si $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_{n-1}\rangle\}$ es un conjunto de vectores linealmente independientes, es decir: $\bar{C}_1 = \bar{C}_2 = \bar{C}_3 = \dots = \bar{C}_{n-1} = 0$.

En caso de no serlo se procede otra vez a despejar uno de los vectores en términos de los anteriores y aplicar el criterio de independencia lineal:

$$|v_{n-1}\rangle = \tilde{C}_1 |v_1\rangle + \tilde{C}_2 |v_2\rangle + \tilde{C}_3 |v_3\rangle \cdots + \tilde{C}_{n-2} |v_{n-2}\rangle = \sum_{i=1}^{n-2} \tilde{C}_i |v_i\rangle,$$

nuevamente se comprueba si se cumple: $\tilde{C}_1 = \tilde{C}_2 = \tilde{C}_3 = \cdots = \tilde{C}_{n-2} = 0$.

En caso contrario, se repite este procedimiento hasta encontrar un conjunto: $\{|v_1\rangle, |v_2\rangle, |v_3\rangle \cdots, |v_{n-j}\rangle\}$ de vectores linealmente independientes. Esto es: $\check{C}_1 = \check{C}_2 = \check{C}_3 = \cdots = \check{C}_{n-j} = 0$.

Por lo tanto:

$$|v_{n-j+1}\rangle = \check{C}_1 |v_1\rangle + \check{C}_2 |v_2\rangle + \check{C}_3 |v_3\rangle \cdots + \check{C}_{n-j} |v_{n-j}\rangle = \sum_{i=1}^{n-j} \check{C}_i |v_i\rangle.$$

Diremos entonces que $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \cdots, |v_{n-j}\rangle\}$ forman una base para el espacio vectorial V .

Es importante señalar, que la dimensión de V será el número de vectores linealmente independientes, que para este caso será: $\dim V = n - j$.

Entonces, se puede comprobar que, dado un vector arbitrario $|x\rangle \in V$, se tiene que:

$$|x\rangle = \sum_{i=1}^{n-j} C_i |v_i\rangle \quad \forall |x\rangle \in V,$$

y el conjunto $\{C_1, C_2, C_3, \cdots C_{n-j}\}$ será único.

Diremos que el número mínimo de vectores: $|v_1\rangle, |v_2\rangle, |v_3\rangle, \cdots, |v_{n-j}\rangle$ que expanden V conforman una base de ese espacio vectorial, y que el número finito de cantidades $C_1, C_2, C_3, \cdots C_{n-j}$, constituyen las componentes de $|x\rangle$ relativas a la base $\{|v_1\rangle, |v_2\rangle, \cdots, |v_{n-j}\rangle\}$.

Queda claro que el vector cero, $|0\rangle$, es linealmente dependiente, y cualquier conjunto de vectores que lo contenga es un conjunto linealmente dependiente de vectores.

De lo anteriormente expuesto se puede concretar la siguiente definición para una base de un espacio vectorial V :

Definición 2.1 (Bases de un espacio vectorial)

Un conjunto finito de vectores:

$$\mathcal{B} = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \cdots, |v_n\rangle\} \in V,$$

se les denominará base del espacio vectorial V si los vectores: $|v_1\rangle, |v_2\rangle, |v_3\rangle, \cdots, |v_n\rangle$, son linealmente independientes y expanden V . Diremos entonces que \mathcal{B} es un sistema generador de V .

Es fácil darse cuenta que si V lo expanden n vectores linealmente independientes, cualquier otro vector $|x\rangle \in V$ será linealmente dependiente. Igualmente, todas las bases de un espacio vectorial V , de dimensión finita, tendrán el mismo número de elementos y ese número de elemento será la dimensión del espacio, es decir, la $\dim V =$ número de vectores que forman una base del espacio.

La base más familiar en el espacio tridimensional real es el conjunto de vectores ortogonales y unitarios: $\{\hat{i}, \hat{j}, \hat{k}\}$. Por lo tanto, como ya sabemos, la dimensión del espacio vectorial $V = \mathbb{R}^3$ es 3. Al conjunto de vectores $\{\hat{i}, \hat{j}, \hat{k}\}$ le podemos asociar tres ejes coordenados: $\{x, y, z\}$ (decimos que le anclamos un sistema de coordenadas), de manera que las componentes C_1, C_2, C_3 de un vector $|v\rangle$ respecto a esta base son las proyecciones de $|v\rangle$ a lo largo de los ejes coordinados.

El concepto de base de un espacio vectorial es de fundamental importancia, ya que una vez especificada la base las operaciones sobre los elementos del espacio vectorial abstracto se pueden realizar ahora sobre los

números que representan las componentes del vector con respecto a la base. Esto significa que cuando sumamos dos vectores de un espacio vectorial abstracto V , sus componentes (respecto a una base) son sumadas. Cuando multiplicamos un vector de V por un elemento α del campo K , todas sus componentes son multiplicadas por α .

Adicionalmente, dentro de un espacio vectorial V se pueden encontrar subespacios y dentro de esos subespacios un conjunto de vectores base.

Vale decir, si $\forall |x\rangle \in V$:

$$|x\rangle = \underbrace{C_1 |v_1\rangle + \dots + C_{n-j} |v_{n-j}\rangle}_{S_1} + \underbrace{C_{n-j+1} |v_{n-j+1}\rangle + \dots + C_{n-k} |v_{n-k}\rangle}_{S_2} + \underbrace{C_{n-k+1} |v_{n-k+1}\rangle + \dots + C_n |v_n\rangle}_{S_3},$$

con: $|x\rangle = |x_1\rangle + |x_2\rangle + |x_3\rangle$ y $|x_1\rangle \in S_1$; $|x_2\rangle \in S_2$; $|x_3\rangle \in S_3$. Entonces diremos que V es la suma directa de S_1, S_2 y S_3 y lo denotaremos como: $V = S_1 \oplus S_2 \oplus S_3$.

También es bueno señalar que, una vez fijada una base, las componentes de un vector según esa base, son únicas y que dada una base de un espacio vectorial, se pueden construir otras bases diferentes de ésta, como veremos más adelante.

Ejemplo 2.8 Independencia lineal: Presentamos otros ejemplos sobre dependencia/independencia lineal.

- Si consideramos el espacio vectorial $V = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ serán ejemplos de **independencia lineal**:

- $|v_k\rangle \equiv f(t) = t^k$, para $k = 1, 2, 3, \dots$. Es claro que un polinomio de grado $n+1$, no podrá ser expresado en términos un polinomio de grado n , en otras palabras: $t^{n+1} \neq \sum_{i=0}^n C_i t^i$.
- $|v_k\rangle \equiv f(t) = e^{a_k t}$, con a_1, a_2, a_3, \dots coeficientes constantes. También salta a la vista que no podremos expresar una de esas funciones exponenciales como una combinación lineal.

- Si consideramos: $|v_1\rangle = \cos^2(t)$, $|v_2\rangle = \sin^2(t)$ y $|v_3\rangle = 1$, es claro que $|v_1\rangle, |v_2\rangle$ y $|v_3\rangle$ son **linealmente dependientes** por cuanto: $|v_1\rangle + |v_2\rangle = |v_3\rangle$. Nótese que si:

$$|v_1\rangle = \cos(t), |v_2\rangle = \sin(t) \text{ y } |v_3\rangle = 1,$$

entonces $|v_1\rangle, |v_2\rangle$ y $|v_3\rangle$ serán vectores **linealmente independientes**.

- Consideremos ahora otro ejemplo en \mathcal{P}^3 ,

$$|x_1\rangle = 1, \quad |x_2\rangle = x - 1, \quad |x_3\rangle = x^2, \quad |x_4\rangle = x^2 + 2x + 1.$$

Podemos ver que este conjunto es linealmente dependiente ya que siempre podremos expresar:

$$|x_4\rangle = 3|x_1\rangle + 2|x_2\rangle + |x_3\rangle.$$



2.3.3 La matriz y el determinante de Gram

Para un conjunto de vectores: $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\} \in V$, existe una forma directa de comprobar la independencia lineal del conjunto. Dado un vector $|x\rangle \in V$ tendremos que $|x\rangle = C^i |v_i\rangle$, entonces, al multiplicar ambos lados por $\langle v^i |$, resulta:

$$\begin{aligned} \langle v^1 |x\rangle &= C^1 \langle v^1 |v_1\rangle + C^2 \langle v^1 |v_2\rangle + C^3 \langle v^1 |v_3\rangle + \dots + C^n \langle v^1 |v_n\rangle \\ \langle v^2 |x\rangle &= C^1 \langle v^2 |v_1\rangle + C^2 \langle v^2 |v_2\rangle + C^3 \langle v^2 |v_3\rangle + \dots + C^n \langle v^2 |v_n\rangle \\ &\vdots && \vdots \\ \langle v^n |x\rangle &= C^1 \langle v^n |v_1\rangle + C^2 \langle v^n |v_2\rangle + C^3 \langle v^n |v_3\rangle + \dots + C^n \langle v^n |v_n\rangle \end{aligned}$$

donde las $C^1, C^2, C^3, \dots, C^n$ son las incógnitas. Por lo cual, para que este sistema tenga solución se impone que el determinante:

$$\begin{vmatrix} \langle v^1 | v_1 \rangle & \langle v^1 | v_2 \rangle & \langle v^1 | v_3 \rangle & \cdots & \langle v_1 | v_n \rangle \\ \langle v^2 | v_1 \rangle & \langle v^2 | v_2 \rangle & \langle v^2 | v_3 \rangle & \cdots & \langle v_2 | v_n \rangle \\ \vdots & & \ddots & & \vdots \\ \langle v^n | v_1 \rangle & \langle v^n | v_2 \rangle & \langle v^n | v_3 \rangle & \cdots & \langle v_n | v_n \rangle \end{vmatrix} \neq 0.$$

Esto es, que el determinante de Gram¹⁵ sea distinto de cero implica que el conjunto: $\{|v_1\rangle, |v_2\rangle, \dots, |v_n\rangle\}$ es linealmente independiente. La inversa también es cierta.

La matriz de Gram aparece de manera natural cuando estudiamos el producto interno en un espacio vectorial. Sea V^n un espacio vectorial de dimensión n con una base $\mathcal{B} = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$. Consideremos también que en V^n existe un producto escalar $\langle x | y \rangle$. Entonces, para dos vectores cualesquiera $|x\rangle$ y $|y\rangle$ expandidos en esa base \mathcal{B} resulta lo siguiente:

$$\begin{aligned} \langle x | y \rangle &= \left\langle \sum_{i=1}^n x_i v_i \left| \sum_{j=1}^n y_j v_j \right. \right\rangle = \sum_{i=1}^n x_i^* \left[\sum_{j=1}^n y_j \langle v^i | v_j \rangle \right] \\ &= (x_1^* \ x_2^* \ x_3^* \cdots \ x_n^*) \begin{pmatrix} \sum_{j=1}^n y_j \langle v^1 | v_j \rangle \\ \sum_{j=1}^n y_j \langle v^2 | v_j \rangle \\ \vdots \\ \sum_{j=1}^n y_j \langle v^n | v_j \rangle \end{pmatrix} \\ &= (x_1^* \ x_2^* \ x_3^* \cdots \ x_n^*) \begin{pmatrix} \langle v^1 | v_1 \rangle & \langle v^1 | v_2 \rangle & \cdots & \langle v^1 | v_n \rangle \\ \langle v^2 | v_1 \rangle & \langle v^2 | v_2 \rangle & \cdots & \langle v^2 | v_n \rangle \\ \vdots & \vdots & & \vdots \\ \langle v^n | v_1 \rangle & \langle v^n | v_2 \rangle & \cdots & \langle v^n | v_n \rangle \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \end{aligned}$$

Por lo tanto:

$$\langle x | y \rangle = (|x\rangle^t)^* G |y\rangle \quad (2.8)$$

La matriz G se llama la matriz de Gram para el producto escalar respecto a la base \mathcal{B} . En el caso de espacios vectoriales reales la matriz de Gram es simétrica: $G_{ij} = G_{ji}$, y en el caso complejo es hermítica o autoadjunta: $(G_{ij}^t)^* = G_{ij}$, es decir, es igual a su transpuesta compleja conjugada.

Ejemplo 2.9 El espacio vectorial V^n tendrá dimensión n y una de las posibles bases $\{|v_i\rangle\}$ será:

$$|v_1\rangle = (1, 0, 0, \dots, 0), \ |v_2\rangle = (0, 1, 0, \dots, 0), \ |v_3\rangle = (0, 0, 1, \dots, 0), \dots, \ |v_n\rangle = (0, 0, 0, \dots, 1).$$

Esta base se conoce con el nombre de base canónica, la base con la que aprendimos el álgebra vectorial en \mathbb{R}^3 : $\{\hat{i}_i\} = \{\hat{i}, \hat{j}, \hat{k}\}$.



Ejemplo 2.10 Es fácil ver que la matriz de Gram para el producto escalar respecto a la base canónica $\{\hat{i}_i\}$ en

¹⁵Jørgen Pedersen Gram (1850-1916) Es conocido mayormente por el método de ortogonalización, pero aparentemente fue ideado por Laplace y utilizado también por Cauchy en 1836. https://en.wikipedia.org/wiki/Jørgen_Pedersen_Gram

\mathbb{R}^n es la matriz identidad. Para dos vectores $|x\rangle$ y $|y\rangle$ tenemos

$$\begin{aligned}\langle x | y \rangle &= (x_1 \ x_2 \ x_3 \cdots x_n) \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\ &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.\end{aligned}$$



Ejemplo 2.11 El espacio de polinomios, \mathcal{P}^n , de grado $g \leq n$ tendrá como una de las posibles bases al conjunto: $\{1, t, t^2, t^3, \dots, t^n\}$, porque cualquier polinomio de grado $\leq n$ podrá ser expresado como combinación lineal de estos $n+1$ vectores. Más aún, el espacio de **todos** los polinomios, \mathcal{P}^∞ , tendrá como una posible base al conjunto de funciones: $\{1, t, t^2, t^3, \dots, t^n, \dots\}$. En este caso \mathcal{P}^∞ será infinito dimensional.



Ejemplo 2.12 Consideremos el espacio de polinomios \mathcal{P}^3 y la base $\{1, t, t^2\}$ con el producto interno definido como:

$$\langle p, q \rangle = \frac{1}{2} \int_{-1}^1 p(t)q(t)dt.$$

Por lo tanto:

$$\langle 1, 1 \rangle = \frac{1}{2} \int_{-1}^1 dt = 1, \quad \langle 1, t \rangle = \frac{1}{2} \int_{-1}^1 t dt = 0, \quad \langle 1, t^2 \rangle = \frac{1}{2} \int_{-1}^1 t^2 dt = \frac{1}{3}, \dots$$

De esta manera, la matriz de Gram respecto a la base dada es:

$$G_{\{1,t,t^2\}} = \begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{5} \end{pmatrix}.$$



2.3.4 Ortogonalidad y bases ortogonales

En un espacio vectorial con producto interno, dos vectores $|v_1\rangle \wedge |v_2\rangle$ serán ortogonales si su producto interno se anula

$$|v_1\rangle \perp |v_2\rangle \Leftrightarrow \langle v_2 | v_1 \rangle = 0.$$

Se denomina un conjunto **ortogonal** de vectores $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ si:

$$\langle e^i | e_j \rangle = \delta_j^i \|\langle e_j \rangle\|^2, \quad i, j = 1, 2, 3, \dots, n \quad \text{y con} \quad \delta_j^i = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

En este caso, es fácil ver que la matriz de Gram es una matriz en diagonal.

Además, se denominará un conjunto **ortonormal** de vectores $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_n\rangle\}$ si:

$$\langle \hat{e}^i | \hat{e}_j \rangle = \delta_j^i, \quad i, j = 1, 2, 3, \dots, n,$$

es decir, cuando $\|\langle e_j \rangle\|^2 = 1$ y la matriz de Gram será la matriz identidad.

Un conjunto ortogonal de vectores $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\} \in V$ es linealmente independiente. Más aún, para el caso particular de un espacio euclíadiano este conjunto conforma una base ortogonal para V . La demostración es sencilla, para un determinado espacio vectorial V una combinación lineal de los vectores:

$\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ se anula. Veamos:

$$\sum_{i=1}^n C^i |e_i\rangle = |0\rangle \Rightarrow \begin{cases} \langle e^1 | \left[\sum_{i=1}^n C^i |e_i\rangle \right] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^1 = 0 \Rightarrow C^1 = 0 \\ \langle e^2 | \left[\sum_{i=1}^n C^i |e_i\rangle \right] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^2 = 0 \Rightarrow C^2 = 0 \\ \langle e^3 | \left[\sum_{i=1}^n C^i |e_i\rangle \right] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^3 = 0 \Rightarrow C^3 = 0 \\ \vdots & \ddots \quad \vdots \quad \vdots \\ \langle e^n | \left[\sum_{i=1}^n C^i |e_i\rangle \right] = 0 & \Rightarrow \sum_{i=1}^n C^i \delta_i^n = 0 \Rightarrow C^n = 0 \end{cases}$$

entonces, queda claro que: $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ son un conjunto de vectores linealmente independientes.

Si la dimensión de V es n ($\dim V = n$) y tenemos n vectores linealmente independientes, entonces esos n vectores $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ forman una base ortogonal para V , y por lo tanto, las componentes de un vector en esa base se pueden expresar de manera simple:

$$\forall |x\rangle \in V \Rightarrow |x\rangle = \sum_{i=1}^n C^i |e_i\rangle \Rightarrow \langle e^j | x \rangle = \langle e^j | \left[\sum_{i=1}^n C^i |e_i\rangle \right] \Rightarrow C^j = \frac{\langle e^j | x \rangle}{\langle e^j | e_j \rangle} = \frac{\langle e^j | x \rangle}{\|e_j\|^2}.$$

En el caso de un conjunto ortonormal de vectores $\{|\hat{e}_i\rangle\} \in V^n$, las componentes de cualquier vector quedan determinadas de una forma todavía más simple y con consecuencias mucho más impactantes:

$$\text{Si } \|\hat{e}_i\|^2 = 1 \Rightarrow C^i = \langle \hat{e}^i | x \rangle \Rightarrow |x\rangle = \sum_{i=1}^n C^i |\hat{e}_i\rangle = \sum_{i=1}^n \langle \hat{e}^i | x \rangle |\hat{e}_i\rangle \equiv \underbrace{\sum_{i=1}^n |\hat{e}_i\rangle \langle \hat{e}^i |}_{1} |x\rangle.$$

Es bueno recalcar la relación de cierre¹⁶

$$\sum_{i=1}^n |\hat{e}_i\rangle \langle \hat{e}^i| = 1, \quad (2.9)$$

con lo cual es trivial demostrar la fórmula de Parseval:

$$\forall |x\rangle, |y\rangle \in V \Rightarrow \langle y | x \rangle \equiv \langle y | \left(\sum_{i=1}^n |\hat{e}_i\rangle \langle \hat{e}^i| \right) |x\rangle = \sum_{i=1}^n \langle y | \hat{e}_i \rangle \langle \hat{e}^i | x \rangle = \sum_{i=1}^n \langle y | \hat{e}_i \rangle \langle x | \hat{e}_i \rangle^*,$$

para el caso de $|x\rangle \equiv |y\rangle$ se llega a la generalización del teorema de Pitágoras:

$$\langle x | x \rangle \equiv \|x\|^2 = \sum_{i=1}^n |\langle x | \hat{e}_i \rangle|^2.$$

Consideremos un par de ejemplos de bases ortogonales en espacio funcionales:

Ejemplo 2.13 Funciones trigonométricas: Uno de los ejemplos más emblemáticos es el caso de las funciones continuas, reales de variable real y definidas en $[0, 2\pi]$, $C_{[0, 2\pi]}^\infty$, con el producto interno definido por:

$$\langle f | g \rangle = \int_0^{2\pi} dx f(x) g(x), \quad (2.10)$$

esto es, el conjunto de funciones $\{|e_i\rangle\}$ representadas por:

$$|e_0\rangle = 1, \quad |e_{2n-1}\rangle = \cos(nx) \quad \text{y} \quad |e_{2n}\rangle = \sin(nx), \quad \text{con } n = 1, 2, 3, \dots \quad (2.11)$$

¹⁶La relación de cierre expresa una propiedad importante de los vectores base: si los $|\hat{e}_i\rangle$ se multiplican por la derecha por los $\langle \hat{e}^i|$, el resultado, luego de sumar para todos los vectores, es el operador lineal unitario. Se dice también que el conjunto de vectores $\{|\hat{e}_i\rangle\}$ forman un conjunto completo.

Es claro que $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle, \dots\}$ es un conjunto de funciones ortogonales por cuanto:

$$\langle e^n | e_m \rangle = \delta_m^n \| |e_n\rangle \|^2 \Rightarrow \begin{cases} 0 & \text{si } n \neq m \\ \| |e_n\rangle \|^2 & \text{si } n = m \end{cases} \Rightarrow \begin{cases} \int_0^{2\pi} dx \sin(nx) \sin(mx) = 0 \\ \int_0^{2\pi} dx \cos(nx) \sin(mx) = 0 \\ \int_0^{2\pi} dx \cos(nx) \cos(mx) = 0 \end{cases}$$

$$\int_0^{2\pi} dx = 2\pi \quad \text{si } n = m = 0$$

$$\int_0^{2\pi} dx \cos^2(nx) = \pi \quad \text{si } n = m = 2k - 1$$

$$\int_0^{2\pi} dx \sin^2(nx) = \pi \quad \text{si } n = m = 2k$$

con $k = 1, 2, 3, \dots$.

Claramente, podremos construir una base ortonormal de funciones: $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, \dots, |\hat{e}_n\rangle, \dots\}$ de la forma:

$$|\hat{e}_0\rangle = \frac{1}{\sqrt{2\pi}}, \quad |\hat{e}_{2n-1}\rangle = \frac{1}{\sqrt{\pi}} \cos(nx) \quad \text{y} \quad |\hat{e}_{2n}\rangle = \frac{1}{\sqrt{\pi}} \sin(nx).$$

Por lo tanto, dada una función definida en el intervalo $[0, 2\pi]$, podremos expresarla en términos de la base ortogonal como:

$$|f\rangle = \sum_{i=1}^{\infty} C^i |e_i\rangle \Rightarrow C^i = \langle e^i | f \rangle = \begin{cases} \int_0^{2\pi} dx f(x) = C^0 & \text{si } i = 0 \\ \int_0^{2\pi} dx f(x) \cos(nx) = C^{2n-1} & \text{si } i = 2n - 1 \\ \int_0^{2\pi} dx f(x) \sin(nx) = C^{2n} & \text{si } i = 2n \end{cases}$$

donde los C^i son los coeficientes de Fourier.



Ejemplo 2.14 Polinomios de Legendre: Otro de los ejemplos típicos son los llamados polinomios de Legendre, $P_n(x)$, definidos en el intervalo $[-1, 1]$. Estos polinomios pueden ser generados a partir de la Fórmula de Rodrigues¹⁷

$$P_n(x) = \frac{1}{n! 2^n} \frac{d^n}{dx^n} (x^2 - 1)^n, \quad n = 0, 1, 2, \dots \quad (2.12)$$

con $P_0(x) = 1$.

Algunos de estos polinomios son los siguientes:

$$P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1), \quad P_3(x) = \frac{x}{2}(5x^2 - 3), \quad P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3), \dots$$

los polinomios de Legendre son soluciones de la ecuación diferencial:

$$(1 - x^2) y'' - 2x y' + \lambda(\lambda + 1)y = 0.$$

Es fácil comprobar que los polinomios de Legendre $|P_\alpha\rangle = P_\alpha(x)$ son mutuamente ortogonales con un producto interno definido como,

$$\langle P_n | P_m \rangle = \int_{-1}^1 P_n(x) P_m(x) dx = \frac{2}{2n + 1} \delta_{nm}, \quad (2.13)$$

y con una norma definida por,

$$\|P_n\|^2 = \langle P_n | P_n \rangle = \int_{-1}^1 P_n^2(x) dx = \frac{2}{2n + 1}. \quad (2.14)$$

¹⁷Benjamin Olinde Rodrigues (1794-1851). https://es.wikipedia.org/wiki/Olinde_Rodrigues

Por lo tanto, cualquier función en el intervalo $[-1, 1]$ puede ser expresada en esa base.

$$f(x) = |f\rangle = \sum_{k=0}^{\infty} a_k |P_k\rangle = \sum_{k=0}^{\infty} \frac{\langle P_k |F\rangle}{\langle P_k |P_k\rangle} |P_k\rangle . \quad (2.15)$$

Si $f(x)$ es un polinomio, entonces

$$f(x) = \sum_{n=0}^m b_n x^n = \sum_{k=0}^{\infty} a^k |P_k\rangle = \sum_{n=0}^{\infty} a^n P_n(x) ,$$

y los coeficientes a^n se determinan fácilmente a través de un sistema de ecuaciones algebraicas. Por ejemplo, para el caso de $f(x) = x^2$ tendremos

$$\begin{aligned} f(x) = x^2 &= a^0 P_0(x) + a^1 P_1(x) + a^2 P_2(x) \\ &= a^0 + a^1 x + \frac{1}{2} a^2 (3x^2 - 1) = \frac{1}{3} P_0(x) + \frac{2}{3} P_2(x) . \end{aligned}$$

Quedará como ejercicio demostrar que para el caso

$$g(x) = \sqrt{\frac{1-x}{2}} = \sum_{k=0}^{\infty} \frac{\langle P_k |g\rangle}{\langle P_k |P_k\rangle} |P_k\rangle = \frac{2}{3} P_0(x) - 2 \sum_{n=1}^{\infty} \frac{P_n(x)}{(2n-1)(2n+3)} ,$$

con

$$\langle P_k |g\rangle = \int_{-1}^1 g(x) P_k(x) dx = \int_{-1}^1 \sqrt{\frac{1-x}{2}} P_k(x) dx .$$



2.3.5 Ortogonalización

Hemos visto que un conjunto de vectores ortogonales forman una base para un espacio vectorial. Ahora bien, siempre es posible construir un conjunto de vectores ortogonales a partir de un conjunto de vectores linealmente independientes. El método de “ortogonalización” se conoce como el método de Gram-Schmidt¹⁸, en honor de estos dos matemáticos alemanes que no lo inventaron pero hicieron famoso al método que, al parecer, se le debe al matemático francés P.S. Laplace.

Consideremos, por ejemplo, un conjunto de vectores linealmente independientes, $\{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle\}$ que expanden un espacio euclíadiano real de dimensión finita, E^n . Entonces, siempre se podrá construir un conjunto ortogonal de vectores, $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$, que también expandan E^n , de la siguiente forma:

1. Hacemos coincidir uno de los vectores de la base dada con nuestro primer vector ortogonal:

$$|e_1\rangle \equiv |v_1\rangle .$$

2. A partir de $|e_1\rangle$ y otro de los vectores dados construimos un segundo vector:

$$|e_2\rangle \equiv |v_2\rangle - \frac{\langle e^1 | v_2 \rangle}{\langle e^1 | e_1 \rangle} |e_1\rangle \quad \backslash \quad \langle e^2 | e_1 \rangle = 0 .$$

3. De esta manera podemos seguir calculando los otros vectores de la base ortogonal hasta completar los n vectores dados:

¹⁸Erhard Schmidt (1876-1959). Matemático alemán fundador del primer instituto de matemáticas aplicadas de Berlín. Alumno de Hilbert, Schmidt hizo sus mayores contribuciones en el campo de ecuaciones integrales y teoría de funciones en el espacio de Hilbert.
https://es.wikipedia.org/wiki/Erhard_Schmidt

$$\begin{aligned}
 |\mathbf{e}_3\rangle &\equiv |v_3\rangle - \frac{\langle e^2|v_3\rangle}{\langle e^2|e_2\rangle}|\mathbf{e}_2\rangle - \frac{\langle e^1|v_3\rangle}{\langle e^1|e_1\rangle}|\mathbf{e}_1\rangle & \setminus \quad \left\{ \begin{array}{l} \langle e^3|e_1\rangle = 0 \\ \langle e^3|e_2\rangle = 0 \end{array} \right. \\
 |\mathbf{e}_4\rangle &\equiv |v_4\rangle - \frac{\langle e^3|v_4\rangle}{\langle e^3|e_3\rangle}|\mathbf{e}_3\rangle - \frac{\langle e^2|v_4\rangle}{\langle e^2|e_2\rangle}|\mathbf{e}_2\rangle - \frac{\langle e^1|v_4\rangle}{\langle e^1|e_1\rangle}|\mathbf{e}_1\rangle & \setminus \quad \left\{ \begin{array}{l} \langle e^4|e_1\rangle = 0 \\ \langle e^4|e_2\rangle = 0 \\ \langle e^4|e_3\rangle = 0 \end{array} \right. \\
 &\vdots & & \vdots \\
 |\mathbf{e}_n\rangle &\equiv |v_n\rangle - \sum_{i=1}^{n-1} \frac{\langle e^i|v_n\rangle}{\langle e^i|e_i\rangle}|\mathbf{e}_i\rangle & \setminus \quad \left\{ \begin{array}{l} \langle e^n|e_1\rangle = 0 \\ \langle e^n|e_2\rangle = 0 \\ \vdots \\ \langle e^n|e_{n-1}\rangle = 0 \end{array} \right.
 \end{aligned}$$

Entonces, siempre es posible construir una base ortogonal a partir de un conjunto de vectores linealmente independientes y una definición de producto interno. Esta base será única en \mathbf{E}^n –para cada definición de producto interno– y si existe otra, sus vectores serán proporcionales. Más aún, cada espacio vectorial \mathbf{V}^n de dimensión finita tendrá una base ortogonal asociada¹⁹.

Ejemplo 2.15 Consideremos un par de ejemplos del método de ortogonalización sencillos, y uno más elaborado:

1. Para el caso de \mathbb{R}^2 es muy claro. Si tenemos dos vectores $|v_1\rangle$ y $|v_2\rangle$ linealmente independientes:

$$|v_1\rangle = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |v_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

y si elegimos $|\mathbf{e}_1\rangle \equiv |v_2\rangle$, entonces, $|\mathbf{e}_2\rangle$ vendrá dado por:

$$|\mathbf{e}_2\rangle \equiv |v_1\rangle - \frac{\langle e^1|v_1\rangle}{\langle e^1|e_1\rangle}|\mathbf{e}_1\rangle \Rightarrow |\mathbf{e}_2\rangle \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

tal y como se esperaba, el otro vector ortogonal es el canónico.

2. Un subespacio de \mathbf{V}^4 , expandido por los siguientes vectores

$$|v_1\rangle = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix}, \quad |v_2\rangle = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 3 \end{pmatrix}, \quad |v_3\rangle = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

¹⁹Hemos construido la base ortogonal para un espacio de dimensión finita, pero el procedimiento es válido para espacios de dimensión infinita

tendrá una base ortogonal asociada dada por:

$$|e_1\rangle \equiv |v_3\rangle = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}; |e_2\rangle \equiv |v_2\rangle - \frac{\langle e^1 | v_2 \rangle}{\langle e^1 | e_1 \rangle} |e_1\rangle = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 3 \end{pmatrix} - (-1) \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}$$

$$|e_3\rangle \equiv |v_1\rangle - \frac{\langle e^2 | v_1 \rangle}{\langle e^2 | e_2 \rangle} |e_2\rangle - \frac{\langle e^1 | v_1 \rangle}{\langle e^1 | e_1 \rangle} |e_1\rangle = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix} - \left(\frac{9}{12}\right) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix} - (1) \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{5}{4} \\ \frac{5}{4} \\ -\frac{7}{4} \\ -\frac{1}{4} \end{pmatrix}.$$

Y la base ortonormal asociada será:

$$|\hat{e}_1\rangle = \frac{|e_1\rangle}{\sqrt{\langle e^1 | e_1 \rangle}} = \frac{\sqrt{2}}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}; |\hat{e}_2\rangle = \frac{|e_2\rangle}{\sqrt{\langle e^2 | e_2 \rangle}} = \frac{\sqrt{3}}{6} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 3 \end{pmatrix}; |\hat{e}_3\rangle = \frac{|e_3\rangle}{\sqrt{\langle e^3 | e_3 \rangle}} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \\ -\frac{7}{5} \\ -\frac{1}{5} \end{pmatrix}.$$

En este ejemplo hemos mostrado que $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ son linealmente independientes y, por lo tanto, base de un subespacio de V^4 . Cabría preguntarse ¿Cómo construimos un cuarto vector, linealmente independiente a los anteriores, que expanda todo V^4 ?

3. Suponga el espacio de polinomios, \mathcal{P}^n , de grado $g \leq n$ definidos en el intervalo $[-1, 1]$. Este espacio vectorial tendrá como una de las posibles bases al conjunto $\{|\pi_i\rangle\} = \{1, t, t^2, t^3, \dots, t^n\}$ con el producto interno definido por:²⁰

$$\langle \pi^i | \pi_j \rangle = \int_{-1}^1 dt \pi_i(t) \pi_j(t).$$

Procederemos a construir una base ortogonal $\{|P_i\rangle\}$, y tomaremos como vector de inicio a $|\pi_0\rangle$:

$$|P_0\rangle \equiv |\pi_0\rangle = 1.$$

El siguiente vector será:

$$|P_1\rangle \equiv |\pi_1\rangle - \frac{\langle \pi^1 | P_0 \rangle}{\langle P^0 | P_0 \rangle} |P_0\rangle = t \iff \begin{cases} \langle \pi^1 | P_0 \rangle = \int_{-1}^1 dt t = 0 \\ \langle P^0 | P_0 \rangle = \int_{-1}^1 dt = 2 \end{cases}$$

²⁰En este punto es importante señalar la importancia de la definición de producto interno. Si esta definición hubiera sido $\langle f | g \rangle = \int_{-1}^1 dx f(x) g(x) \sqrt{1-x^2}$ o $\langle f | g \rangle = \int_{-1}^1 dx \frac{f(x) g(x)}{\sqrt{1-x^2}}$ las bases correspondientes a estas definiciones de producto interno serían distintas.

El siguiente:

$$|P_2\rangle \equiv |\pi_2\rangle - \frac{\langle\pi^2|P_1\rangle}{\langle P^1|P_1\rangle} |P_1\rangle - \frac{\langle\pi^2|P_0\rangle}{\langle P^0|P_0\rangle} |P_0\rangle = t^2 - \frac{1}{3} \leftarrow \begin{cases} \langle\pi^2|P_0\rangle = \int_{-1}^1 dt t^2 = \frac{2}{3} \\ \langle\pi^2|P_1\rangle = \int_{-1}^1 dt t^3 = 0 \\ \langle P^1|P_1\rangle = \int_{-1}^1 dt t^2 = \frac{2}{3} \end{cases}$$

Para el cuarto:

$$|P_3\rangle \equiv |\pi_3\rangle - \frac{\langle\pi^3|P_2\rangle}{\langle P^2|P_2\rangle} |P_2\rangle - \frac{\langle\pi^3|P_1\rangle}{\langle P^1|P_1\rangle} |P_1\rangle - \frac{\langle\pi^3|P_0\rangle}{\langle P^0|P_0\rangle} |P_0\rangle = t^3 - \frac{3}{5}t$$

$$\begin{cases} \langle\pi^3|P_0\rangle = \int_{-1}^1 dt t^3 = 0, & \langle\pi^3|P_1\rangle = \int_{-1}^1 dt t^4 = \frac{2}{5} \\ \langle\pi^3|P_2\rangle = \int_{-1}^1 dt t^3[t^2 - \frac{1}{3}] = 0, & \langle P^2|P_2\rangle = \int_{-1}^1 dt [t^2 - \frac{1}{3}]^2 = \frac{8}{45} \end{cases}$$

Queda como ejercicio comprobar la ortogonalidad de los vectores recién calculados:

$$\langle P^0|P_1\rangle = \langle P^1|P_2\rangle = \langle P^2|P_3\rangle = \dots = 0.$$

Esta base ortogonal está formada por los polinomios de Legendre, discutidos en la sección 2.3.4. Es decir, si ortogonalizamos una base de monomios $\{|\pi_i\rangle\}$ mediante la definición de producto interno: $\langle f|g\rangle = \int_{-1}^1 dx f(x) g(x)$, obtendremos la base de polinomios ortogonales de Legendre.

Podemos resumir los cálculos anteriores, construyendo también la base ortonormal a partir de los monomios $\{|\pi_i\rangle\}$ como se muestra a continuación:

$ \pi_n\rangle$	$ P_n\rangle$	$ \hat{P}_n\rangle$
1	1	$\frac{1}{\sqrt{2}}$
t	t	$\sqrt{\frac{3}{2}} t$
t^2	$t^2 - \frac{1}{3}$	$\frac{1}{2}\sqrt{\frac{5}{2}}(3t^2 - 1)$
t^3	$t^3 - \frac{3}{5}t$	$\frac{1}{2}\sqrt{\frac{7}{2}}(5t^3 - 3t)$
t^4	$t^4 - \frac{6}{7}t^2 + \frac{3}{35}$	$\frac{3}{8}\sqrt{\frac{1}{2}}(35t^4 - 30t^2 + 3)$
\vdots	\vdots	\vdots



2.3.6 Ejercicios

- Diga si los siguientes conjuntos de vectores en \mathcal{P}^3 son o no linealmente independientes.
 - $|x_1\rangle = 2x$, $|x_2\rangle = x^2 + 1$, $|x_3\rangle = x + 1$, $|x_4\rangle = x^2 - 1$.
 - $|x_1\rangle = x(x - 1)$, $|x_2\rangle = x$, $|x_3\rangle = x^3$, $|x_4\rangle = 2x^3 - x^2$.
- Probar que los polinomios: $|x_1\rangle = 1$, $|x_2\rangle = x$, $|x_3\rangle = \frac{3}{2}x^2 - \frac{1}{2}$ y $|x_4\rangle = \frac{5}{2}x^3 - \frac{3}{2}x$, forman una base en \mathcal{P}_4 . Expresar $|p\rangle = x^2$, $|q\rangle = x^3$ en función de esa base.
- Encontrar la proyección perpendicular de los siguientes vectores en $\mathcal{C}_{[-1,1]}$ (espacio de funciones continuas en el intervalo $[-1,1]$ y con el producto interno definido por: $\langle f|g\rangle = \int_{-1}^1 dx f(x) g(x)$) al subespacio generado por los polinomios: $\{1, x, x^2 - 1\}$. Calcular la distancia de cada una de estas funciones al subespacio mencionado.
 - $f(x) = x^n$, n entero.
 - $f(x) = \sin(x)$.

(c). $f(x) = 3x^2$.

4. Utilizando el método de Gram-Schmidt encuentre una base ortonormal para los siguientes conjuntos de vectores:
 - (a). $|v_1\rangle = (1, 0, 1)$, $|v_2\rangle = (0, 1, 1)$ y $|v_3\rangle = (1, 0, 0)$. En \mathbb{R}^3 .
 - (b). $|v_1\rangle = (2, -4, 5, 2)$, $|v_2\rangle = (-6, 5, -1, -6)$ y $|v_3\rangle = (-10, 13, -4, -3)$. En \mathbb{R}^4 .
5. Considere el espacio vectorial de las matrices complejas 2×2 hermíticas. Tal y como demostraremos con rigor en la sección 4.3.2.2 y lo detallamos en la sección 4.4.9, una matriz hermítica (o autoadjunta) será igual a su adjunta. Esto es, una matriz será igual a su traspuesta conjugada $(A^\dagger)_j^i \rightarrow (A^*)_i^j \equiv A_i^j$:

$$\mathbb{A} \leftrightarrow \begin{pmatrix} z_1 & z_2 \\ z_3 & z_4 \end{pmatrix} = \mathbb{A}^\dagger = \begin{pmatrix} z_1^* & z_3^* \\ z_2^* & z_4^* \end{pmatrix} \quad \text{es decir} \quad \begin{cases} z_1^* = z_1 & \text{real} \\ z_4^* = z_4 & \text{real} \\ z_2^* = z_3 & \text{complejos} \end{cases}.$$

Entonces

- (a). Muestre que las matrices de Pauli $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ presentadas en los ejercicios de la sección 2.2.4 forman una base para ese espacio vectorial.
- (b). Compruebe que esa base es ortogonal bajo la definición de producto interno $\langle a | b \rangle = \text{Tr}(A^\dagger B)$ que introdujimos en los ejercicios de esa misma sección.
- (c). Explore si se pueden construir subespacios vectoriales de matrices reales e imaginarias puras.
6. Utilizando **Maxima** reproduzca el ejemplo 2. que expusimos en la página 120. Es decir, suponga el espacio de polinomios, \mathcal{P}^n , de grado $g \leq n$ definidos en el intervalo $[-1, 1]$. Este espacio vectorial tendrá como una de las posibles bases al conjunto $\{|\pi_i\rangle\} = \{1, t, t^2, t^3, \dots, t^n\}$, pero en este caso con el producto interno definido por:

$$\langle f | g \rangle = \int_{-1}^1 dx f(x) g(x) \sqrt{1 - x^2}.$$

Encuentre la base ortogonal correspondiente. A esta nueva base se le conoce como polinomios de Chebyshev de segunda especie²¹.

7. Otra vez, utilizando **Maxima**, reproduzca el ejercicio anterior, pero con la definición de producto interno:

$$\langle f | g \rangle = \int_{-1}^1 dx \frac{f(x) g(x)}{\sqrt{1 - x^2}}.$$

A esta nueva base se le conoce como polinomios de Chebyshev de primera especie.

Practicando con Maxima

1. **Independencia Lineal:** En 2.3.5 vimos que si en la ecuación

$$|0\rangle = C_1 |v_1\rangle + C_2 |v_2\rangle + C_3 |v_3\rangle \cdots + C_n |v_n\rangle$$

con todos los $C_i = 0$, entonces se dirá que el conjunto de vectores son linealmente independientes. Para el segundo ejemplo de esa sección (ejemplo 2.) se obtuvo el siguiente sistema de ecuaciones:

$$\begin{array}{rcl} C_1 & +2C_2 & -C_3 = 0 \\ 3C_1 & & +C_3 = 0 \\ -C_1 & +C_2 & = 0 \\ 2C_1 & +3C_2 & = 0 \end{array}$$

que procedemos a resolver usando el comando **linsolve**:

²¹https://en.wikipedia.org/wiki/Chebyshev_polynomials

(%i1) linsolve([C1+2*C2-C3=0, 3*C1+C3=0, -C1+C2=0, 2*C1+3*C2=0], [C1,C2,C3]);
 solve: dependent equations eliminated: (4)

(%o1) $[C_1 = 0, C_2 = 0, C_3 = 0]$

2. **Bases para espacios vectoriales:** En este ejercicio aprenderemos a calcular una base a partir de un conjunto de vectores perteneciente a un determinado espacio vectorial. Por ejemplo, si en \mathbb{R}^5 tenemos el siguiente conjunto de vectores:

$$|v_1\rangle = (1, 2, 3, 4, 5), |v_2\rangle = (0, -1, 1, 2, 3), |v_3\rangle = (3, 2, 1, 0, -1), |v_4\rangle = (-4, -3, -2, -1, 0).$$

(%i1) v1:[1,2,3,4,5]; v2:[0,-1,1,2,3]; v3:[3,2,1,0,-1]; v4:[-4,-3,-2,-1,0];
 (%o1) $[1, 2, 3, 4, 5]$

(%o2) $[0, -1, 1, 2, 3]$

(%o3) $[3, 2, 1, 0, -1]$

(%o4) $[-4, -3, -2, -1, 0]$

Con los vectores dados construimos la siguiente matriz

(%i5) M:matrix(v1,v2,v3,v4);
 (%o5)
$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & -1 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 & -1 \\ -4 & -3 & -2 & -1 & 0 \end{pmatrix}$$

Como veremos más adelante, el Rango de una matriz indica el número máximo de vectores fila o columna linealmente independientes.

(%i6) rank(M);
 (%o6) 3

Podemos aplicar el método de eliminación gaussiana a la matriz M para obtener una nueva matriz escalonada. El cálculo además se hace normalizando el primer elemento no nulo de cada fila. Esto se logra con el comando **echelon(M)**.

(%i7) echelon(M);
 (%o7)
$$\begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} & 0 & -\frac{1}{3} \\ 0 & 1 & -1 & -2 & -3 \\ 0 & 0 & 1 & \frac{5}{3} & \frac{7}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por lo tanto, cada fila de la matriz anterior conformará un conjunto de vectores linealmente independiente. Los podemos aislar de la matriz con el comando **row(M,i)**, que devolverá la i -ésima fila de la matriz M.

(%i8) e1:row(%o7,1);
 (%o8)
$$\begin{pmatrix} 1 & \frac{2}{3} & \frac{1}{3} & 0 & -\frac{1}{3} \end{pmatrix}$$

(%i9) e2:row(%o7,2);
 (%o9)
$$\begin{pmatrix} 0 & 1 & -1 & -2 & -3 \end{pmatrix}$$

(%i10) e3:row(%o7,3);

$$(\%o10) \begin{pmatrix} 0 & 0 & 1 & \frac{5}{3} & \frac{7}{3} \end{pmatrix}$$

Es trivial verificar que el conjunto: $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ es linealmente independientes.

```
(%ci11) solve(alpha*[1,2/3,1/3,0,-1/3]+beta*[0,1,-1,-2,-3]+gamma*[0,0,0,5/3,7/3],[alpha,beta,gamma]);
solve: dependent equations eliminated: (3)
(%o11) [[\alpha = 0, \beta = 0, \gamma = 0]]
```

Consideremos otros ejemplos, dados los vectores $\mathbf{a} = (1, 3)$ y $\mathbf{b} = (-1, 1)$ ¿Serán linealmente independientes?

```
(%ci12) solve(alpha*[1,3]+beta*[-1,1],[alpha,beta]);
(%o12) [[\alpha = 0, \beta = 0]]
```

Los vectores $\mathbf{a} = (1, 2, 3)$ y $\mathbf{b} = (4, 8, 12)$ ¿Serán linealmente independientes?

```
(%ci13) solve(alpha*[1,2,3]+beta*[4,8,12],[alpha,beta]);
solve: dependent equations eliminated: (2 3)
(%o13) [[\alpha = -4 \%r_4, \beta = \%r_4]]
```

Aquí $\alpha = -4\beta$.

Sea ahora $\{\mathbf{e}_i\} = \{(1, 1, 1), (1, 2, 1), (0, 0, 2)\}$ una base para \mathbb{R}^3 . Vamos a calcular las componentes del vector $\mathbf{a} = (3, 2, 1)$ respecto de esa base.

Primero verifiquemos si efectivamente es una base:

```
(%ci14) solve(alpha*[1,1,1]+beta*[1,2,1]+gamma*[0,0,2],[alpha,beta,gamma]);
(%o14) [[\alpha = 0, \beta = 0, \gamma = 0]]

(%ci15) solve([3,2,1]-ax*[1,1,1]-ay*[1,2,1]-az*[0,0,2],[ax,ay,az]);
(%o15) [[ax = 4, ay = -1, az = -1]]
```

Por lo tanto, $\mathbf{a} = 4\mathbf{e}_1 - \mathbf{e}_2 - \mathbf{e}_3$.

En la base canónica las componentes del vector \mathbf{a} serán:

```
(%ci16) solve([3,2,1]-ax*[1,0,0]-ay*[0,1,0]-az*[0,0,1],[ax,ay,az]);
(%o16) [[ax = 3, ay = 2, az = 1]]
```

Es decir, $\mathbf{a} = 3\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}$.

```
(%ci17) kill(all)$
```

3. **Ortogonalización con Maxima** Anteriormente hicimos los cálculos para hallar una base ortogonal a partir del siguiente conjunto de vectores linealmente independientes:

$$|v_1\rangle = \begin{pmatrix} 1 \\ 3 \\ -1 \\ 2 \end{pmatrix}; \quad |v_2\rangle = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 3 \end{pmatrix}; \quad |v_3\rangle = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

La función **gramschmidt(x)** es la que ejecuta el algoritmo de ortogonalización de Gram-Schmidt sobre

el objeto `x`. Si `x` es una matriz el algoritmo actúa sobre sus filas. Antes de usar la función es necesario cargar la biblioteca `eigen`.

```
(%ci1) load("eigen")$  
(%ci2) v:matrix([-1, 1, 0, 0], [2, 0, 1, 3], [1, 3, -1, 2]);  
(%o2) 
$$\begin{pmatrix} -1 & 1 & 0 & 0 \\ 2 & 0 & 1 & 3 \\ 1 & 3 & -1 & 2 \end{pmatrix}$$

```

Notemos que el vector $|v_3\rangle$ es el que hemos puesto en la primera fila de la matriz. Ahora procedemos al cálculo:

```
(%ci3) e:gramschmidt(v);  
(%o3) 
$$\left[ [-1, 1, 0, 0], [1, 1, 1, 3], \left[ \frac{5}{2^2}, \frac{5}{2^2}, -\frac{7}{2^2}, -\frac{1}{2^2} \right] \right]$$

```

El resultado viene en forma de una lista: `[e[1], e[2], e[3]]`. Simplificando resulta:

```
(%ci4) e:expand(%);  
(%o4) 
$$\left[ [-1, 1, 0, 0], [1, 1, 1, 3], \left[ \frac{5}{4}, \frac{5}{4}, -\frac{7}{4}, -\frac{1}{4} \right] \right]$$

```

Podemos verificar que son ortogonales:

```
(%ci5) map(innerproduct, [e[1], e[2], e[3]], [e[2], e[3], e[1]]);  
(%o5) [0, 0, 0]
```

Normalizamos ahora cada uno de los vectores:

```
(%ci6) e[1]/sqrt(innerproduct(e[1], e[1]));  
(%o6) 
$$\left[ -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, 0 \right]  
(%ci7) e[2]/sqrt(innerproduct(e[2], e[2]));  
(%o7) 
$$\left[ \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{\sqrt{3}}{2} \right]  
(%ci8) e[3]/sqrt(innerproduct(e[3], e[3]));  
(%o8) 
$$\left[ \frac{1}{2}, \frac{1}{2}, -\frac{7}{10}, -\frac{1}{10} \right]$$$$$$

```

La función `gramschmidt(x,F)` nos permite definir un producto interno diferente a `innerproduct`. Veamos como se hace para el otro ejemplo donde el conjunto de vectores linealmente independientes estaba dado por: $\{1, t, t^2, t^3, \dots, t^n\}$.

```
(%ci9) producto(f,g):=integrate(f*g, t, a, b);  
(%o9) producto(f,g) := 
$$\int_a^b f g \, dt$$
  
(%ci10) e:gramschmidt([1, t, t^2, t^3, t^4], producto), a=-1, b=1;  
(%o10) 
$$\left[ 1, t, \frac{3t^2 - 1}{3}, \frac{t(5t^2 - 3)}{5}, \frac{35t^4 - 30t^2 + 3}{35} \right]  
(%ci11) e:expand(e);$$

```

$$(\%o11) \quad \left[1, t, t^2 - \frac{1}{3}, t^3 - \frac{3t}{5}, t^4 - \frac{6t^2}{7} + \frac{3}{35} \right]$$

Verifiquemos si son ortogonales:

$$(\%i12) \quad \text{map}(producto, [e[1], e[2], e[3]], [e[2], e[3], e[1]]), a = -1, b = 1; \\ (\%o12) \quad [0, 0, 0]$$

Para normalizar cada uno de los vectores procedemos como en el caso anterior pero primero hacemos $a = -1$ y $b = 1$.

$$(\%i13) \quad a: -1b : 1$$

$$(\%i14) \quad e[1]/\sqrt{\text{producto}(e[1], e[1])}; \\ (\%o14) \quad \frac{1}{\sqrt{2}}$$

$$(\%i15) \quad e[2]/\sqrt{\text{producto}(e[2], e[2])}; \\ (\%o15) \quad \frac{\sqrt{3}t}{\sqrt{2}}$$

$$(\%i16) \quad e[3]/\sqrt{\text{producto}(e[3], e[3])}; \\ (\%o16) \quad \frac{3\sqrt{5}(t^2 - \frac{1}{3})}{2^{\frac{3}{2}}}$$

$$(\%i17) \quad e[4]/\sqrt{\text{producto}(e[4], e[4])}; \\ (\%o17) \quad \frac{5\sqrt{7}(t^3 - \frac{3t}{5})}{2^{\frac{3}{2}}}$$

2.4 Aproximación de funciones

Armados de los conocimientos previos podemos pasar a aplicarlos en un intento de aproximar funciones. La aproximación de una función tiene varias facetas y seguramente en cursos anteriores hemos hecho este tipo de aproximaciones una buena cantidad de veces cuando necesitábamos convertir una expresión, que nos resultaba muy complicada, en otras más sencillas y casi equivalentes. Por ejemplo, cuando aplicamos la aproximación lineal: $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$, con x muy cercano a x_0 . ¿Cuántas veces no hemos utilizado la aproximación: $\sin(x) \approx x$? También aproximamos funciones cuando en el laboratorios nos veíamos en la necesidad de “ajustar con la mejor curva” una serie de puntos experimentales.

Consideremos primero algunos conceptos importantes.

2.4.1 Complementos ortogonales y descomposición ortogonal

Volvamos a retomar la idea de los subespacios y su complemento ortogonal, ejercicio 3. de la sección 2.2. Sea un subespacio $S \subset V$. Un elemento $|\bar{v}_i\rangle \in V$ se dice ortogonal a S si $\langle s^k | \bar{v}_i \rangle = 0 \forall |s_k\rangle \in S$, es decir, es ortogonal a todos los elementos de S . El conjunto $\{|\bar{v}_1\rangle, |\bar{v}_2\rangle, |\bar{v}_3\rangle, \dots, |\bar{v}_m\rangle\}$ de todos los elementos ortogonales a S , se denomina S -perpendicular y se denota como S^\perp . Es fácil demostrar que S^\perp es un subespacio, aún si S no lo es.

Dado un espacio euclíadiano de dimensión infinita $V : \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle, \dots\}$ y un subespacio $S \subset V$ con dimensión finita: $\dim S = m$. Entonces, $\forall |v_k\rangle \in V$ puede expresarse como la suma de dos vectores $|s_k\rangle \in S \wedge |s_k\rangle^\perp \in S^\perp$. Esto es:

$$|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp, \quad |s_k\rangle \in S \quad \wedge \quad |s_k\rangle^\perp \in S^\perp.$$

La norma de $|v_k\rangle$ se calcula a través del teorema de Pitágoras generalizado:

$$\| |v_k\rangle \|^2 = \| |s_k\rangle \|^2 + \| |s_k\rangle^\perp \|^2.$$

La demostración es sencilla. Primero se prueba que la descomposición ortogonal $|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp$ es siempre posible. Para ello recordamos que $S \subset V$ es de dimensión finita, por lo tanto existe una base ortonormal $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_m\rangle\}$ para S . Es decir, dado un $|v_k\rangle$ definimos los elementos $|s_k\rangle$ y $|s_k\rangle^\perp$ como sigue:

$$|s_k\rangle = \sum_{i=1}^m \langle v_k | \hat{e}_i \rangle |\hat{e}_i\rangle \quad \wedge \quad |s_k\rangle^\perp = |v_k\rangle - |s_k\rangle. \quad (2.16)$$

Nótese que $\langle v_k | \hat{e}_i \rangle |\hat{e}_i\rangle$ es la proyección de $|v_k\rangle$ a lo largo de $|\hat{e}_i\rangle$ y $|s_k\rangle$ se expresa como la combinación lineal de la base de S , por lo tanto, está en S . Por otro lado:

$${}^\perp \langle s^k | \hat{e}_i \rangle = \langle v^k - s^k | \hat{e}_i \rangle = \langle v^k | \hat{e}_i \rangle - \langle s^k | \hat{e}_i \rangle = \langle v^k | \hat{e}_i \rangle - \left[\sum_{j=1}^m \langle v^k | \hat{e}_j \rangle \langle \hat{e}^j | \right] |\hat{e}_i\rangle = 0 \Rightarrow |s_k\rangle^\perp \perp |\hat{e}_j\rangle,$$

lo cual indica que $|s_k\rangle^\perp \in S^\perp$.

Podemos ir un poco más allá y ver que la descomposición $|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp$ es única en V . Para ello suponemos que existen dos posibles descomposiciones, vale decir:

$$|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp \wedge |v_k\rangle = |t_k\rangle + |t_k\rangle^\perp, \quad \text{con } |s_k\rangle \wedge |t_k\rangle \in S \wedge |s_k\rangle^\perp \wedge |t_k\rangle^\perp \in S^\perp.$$

Por lo tanto:

$$|v_k\rangle - |v_k\rangle = (|s_k\rangle + |s_k\rangle^\perp) - (|t_k\rangle + |t_k\rangle^\perp) = 0 \Rightarrow |s_k\rangle - |t_k\rangle = |t_k\rangle^\perp - |s_k\rangle^\perp.$$

Pero $|s_k\rangle - |t_k\rangle \in S$, es decir, ortogonal a todos los elementos de S^\perp y $|s_k\rangle - |t_k\rangle = |t_k\rangle^\perp - |s_k\rangle^\perp$.

Con lo cual $|s_k\rangle - |t_k\rangle \equiv |0\rangle$, que es el único elemento que es ortogonal a sí mismo y en consecuencia la descomposición $|v_k\rangle = |s_k\rangle + |s_k\rangle^\perp$ es única.

Finalmente, con la definición de norma:

$$\| |v_k\rangle \|^2 = \| |s_k\rangle + |s_k\rangle^\perp \|^2 = (\langle s^k | + \langle s^k |^\perp) (|s_k\rangle + |s_k\rangle^\perp) = \langle s^k | s_k \rangle + \langle s^k | s_k \rangle^\perp = \| |s_k\rangle \|^2 + \| |s_k\rangle^\perp \|^2.$$

Así, dado S^m un subespacio de V de dimensión finita y dado un $|v_k\rangle \in V$ el elemento:

$$|s_k\rangle \in S \Rightarrow |s_k\rangle = \sum_{i=1}^m \langle v_k | e_i \rangle |e_i\rangle,$$

será la proyección de $|v_k\rangle$ en S .

En general, dado un vector $|x\rangle \in V$ y un subespacio de V con dimensión finita, $S^m \subset V$ y $\dim S = m$, entonces la distancia de $|x\rangle$ a S^m es la norma de la componente de $|x\rangle$, perpendicular a S^m . Más aún, esa distancia será mínima y $|x\rangle_{S^m}$ la proyección de $|x\rangle$, en S^m será el elemento de S^m más próximo a $|x\rangle$ y, por la mejor aproximación.

Ejemplo 2.16 Sea el espacio vectorial $V : \mathbb{R}^3$ con los vectores base: $|x_1\rangle = \hat{i}$, $|x_2\rangle = \hat{k}$ y $|x_3\rangle = \hat{j} + \hat{k}$ y consideramos el subespacio $S : \{x + y = 0\} \subset \mathbb{R}^3$. Podemos calcular una base en S parametrizándola de la manera siguiente:

$$x + y = 0 \Rightarrow \begin{cases} x = -\mu \\ y = \mu \\ z = \nu \end{cases}$$

Por lo que una base en S puede ser: $|e_1\rangle = -\hat{i} + \hat{j}$ y $|e_2\rangle = \hat{k}$. Si tomamos un vector $|x\rangle \in \mathbb{R}^3$ con coordenadas (x, y, z) , para que éste vector $|x\rangle \in S^\perp$ se tiene que cumplir que: $\langle x | e_k \rangle = 0$. Podemos utilizar la matriz de Gram, que conocemos para la base dada, para calcular el producto escalar a través de las ecuaciones:

$$\begin{aligned} \langle x | e_1 \rangle &= (x, y, z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 0 \Rightarrow -x + y + z = 0 \\ \langle x | e_2 \rangle &= (x, y, z) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0 \Rightarrow y + 2z = 0 \end{aligned}$$

El par de ecuaciones: $-x + y + z = 0$ y $y + 2z = 0$ vienen a ser representación cartesiana de S^\perp . La matriz de los coeficientes de ese sistema de ecuaciones es:

$$\begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \Rightarrow \begin{cases} x + z = 0 \\ y + 2z = 0 \end{cases}$$

por lo que una base para esa representación cartesiana puede ser:

$$\begin{cases} x + z = 0 \\ y + 2z = 0 \end{cases} \Rightarrow \begin{cases} x = -\mu \\ y = -2\mu \\ z = \mu \end{cases},$$

es decir, $S^\perp : \{(-1, -2, 1)\}$.

Notemos que partimos de un espacio vectorial V de dimensión 3, la dimensión de S es 2 y por lo tanto, la de S^\perp es 1.

El lector debe comprobar que efectivamente $\{(-1, -2, 1)\}$ es ortogonal a cada uno de los vectores base de $S : \{(-1, 1, 0), (0, 0, 1)\}$.



2.4.2 Condiciones para la aproximación de funciones

Sea $\mathbf{V} = \{|v_1\rangle, |v_2\rangle, |v_3\rangle, \dots, |v_n\rangle, \dots\}$ un espacio euclidiano de dimensión infinita, y un subespacio $\mathbf{S}^m \subset \mathbf{V}$, con dimensión finita, $\dim \mathbf{S} = m$, y sea un elemento $|v_i\rangle \in \mathbf{V}$. La proyección de $|v_i\rangle$ en \mathbf{S}^m , $|s_i\rangle$, será el elemento de \mathbf{S}^m más próximo a $|v_i\rangle$. En otras palabras:

$$\| |v_i\rangle - |s_i\rangle \| \leq \| |v_i\rangle - |t_i\rangle \| \quad \forall |t_i\rangle \in \mathbf{S}.$$

La demostración se sigue así:

$$|v_i\rangle - |t_i\rangle = (|v_i\rangle - |s_i\rangle) + (|s_i\rangle - |t_i\rangle) \Rightarrow \| |v_i\rangle - |t_i\rangle \|^2 = \| |v_i\rangle - |s_i\rangle \|^2 + \| |s_i\rangle - |t_i\rangle \|^2,$$

ya que $|v_i\rangle - |s_i\rangle = |s_k\rangle^\perp \in \mathbf{S}^\perp \wedge |s_i\rangle - |t_i\rangle \in \mathbf{S}$, y vale el teorema de Pitágoras generalizado.

Ahora bien, como:

$$\| |s_i\rangle - |t_i\rangle \|^2 \geq 0 \Rightarrow \| |v_i\rangle - |t_i\rangle \|^2 \geq \| |v_i\rangle - |s_i\rangle \|^2 \Rightarrow \| |v_i\rangle - |t_i\rangle \| \geq \| |v_i\rangle - |s_i\rangle \| . \quad \blacktriangleleft$$

En la sección 2.3.4 consideramos la expansión de funciones continuas, reales de variable real, definidas en $[0, 2\pi]$, $\mathcal{C}_{[0,2\pi]}^\infty$, mediante funciones trigonométricas y con el producto interno definido por:

$$\langle f | g \rangle = \int_0^{2\pi} dx f(x) g(x).$$

En ese entonces consideramos, para ese espacio vectorial, una base ortogonal definida por:

$$|e_0\rangle = 1, \quad |e_{2n-1}\rangle = \cos(nx) \quad y \quad |e_{2n}\rangle = \sin(nx), \quad \text{con } n = 1, 2, 3, \dots$$

Por lo tanto, cualquier función definida en el intervalo $[0, 2\pi]$ puede expresarse en términos de esta base como mostramos a continuación:

$$|f\rangle = \sum_{i=1}^{\infty} C^i |e_i\rangle.$$

Los C^i son los coeficientes de Fourier. Es decir, cualquier función puede ser expresada como una serie de Fourier de la forma:

$$f(x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)],$$

donde:

$$a_k = \frac{1}{\pi} \int_0^{2\pi} dx f(x) \cos(kx) \quad \wedge \quad b_k = \frac{1}{\pi} \int_0^{2\pi} dx f(x) \sin(kx).$$

Es claro que para la aproximación de funciones por funciones trigonométricas cuyos coeficientes son los coeficientes de Fourier constituyen la mejor aproximación. Por lo tanto, de todas las funciones $\mathcal{F}(x) \in \mathcal{C}_{[0,2\pi]}^\infty$ las funciones trigonométricas, $\mathcal{T}(x)$ minimizan la desviación cuadrática media:

$$\int_0^{2\pi} dx (f(x) - \mathcal{P}(x))^2 \geq \int_0^{2\pi} dx (f(x) - \mathcal{T}(x))^2.$$

2.4.3 El método de mínimos cuadrados

Una de las aplicaciones más importantes en la aproximación de funciones es el método de mínimos cuadrados. La idea es determinar el valor más aproximado de una cantidad física, c , a partir de un conjunto de medidas experimentales: $\{x_1, x_2, x_3, \dots, x_n\}$. Para ello asociamos ese conjunto de medidas con las componentes de un vector $|x\rangle \equiv (x_1, x_2, x_3, \dots, x_n)$ en \mathbb{R}^n y supondremos que su mejor aproximación, que llamaremos $c|1\rangle \equiv (c, c, c, \dots, c)$, será la proyección perpendicular de $|x\rangle$ (las medidas) sobre el subespacio generado por

$|1\rangle$. Esto es:

$$|x\rangle = c|1\rangle \Rightarrow c = \frac{\langle x | 1 \rangle}{\langle 1 | 1 \rangle} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n},$$

que no es otra cosa que construir -de una manera sofisticada- el promedio aritmético de las medidas. Es claro que la proyección perpendicular de $|x\rangle$ sobre $|1\rangle$ hace mínima la distancia entre el subespacio perpendicular generado por $|1\rangle$ y el vector $|x\rangle$, con ello también hará mínimo su cuadrado:

$$[d(|x\rangle, c|1\rangle)]^2 = \| |x - c\rangle \|^2 = \langle x - c | x - c \rangle = \sum_{i=1}^n (x_i - c)^2.$$

Esta manera sofisticada, que se deriva del formalismo utilizado, muestra el significado del promedio aritmético como medida más cercana al valor “real” de una cantidad obtenida a partir de una serie de medidas experimentales.

Obviamente, este problema se puede generalizar para el caso de dos (o n) cantidades si extendemos la dimensión del espacio y los resultados experimentales se expresarán como un vector de $2n$ dimensiones

$$|x\rangle = (x_{11}, x_{12}, x_{13}, \dots, x_{1n}, x_{21}, x_{22}, x_{23}, \dots, x_{2n}),$$

mientras que los vectores que representan las cantidades más aproximadas serán:

$$c_1|1_1\rangle = \left(\underbrace{c_1, c_1, c_1, \dots, c_1}_n, \underbrace{0, 0, 0, \dots, 0}_n \right) \quad \wedge \quad c_2|1_2\rangle = (0, 0, 0, \dots, 0, c_2, c_2, c_2, \dots, c_2).$$

Ahora $\{|1_1\rangle, |1_2\rangle\}$ expanden un subespacio vectorial sobre el cual $|x\rangle$ tiene como proyección ortogonal a: $c_1|1_1\rangle + c_2|1_2\rangle$ y consecuentemente $|x - c_1|1_1\rangle - c_2|1_2\rangle$ será perpendicular a $\{|1_1\rangle, |1_2\rangle\}$. Por lo tanto:

$$c_1 = \frac{\langle x | 1_1 \rangle}{\langle 1^1 | 1_1 \rangle} = \frac{x_{11} + x_{12} + x_{13} + \dots + x_{1n}}{n} \quad \wedge \quad c_2 = \frac{\langle x | 1_2 \rangle}{\langle 1^2 | 1_2 \rangle} = \frac{x_{21} + x_{22} + x_{23} + \dots + x_{2n}}{n}.$$

Quizá la consecuencia más conocida de esta forma de aproximar funciones es el “ajuste” de un conjunto de datos experimentales $\{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)\}$ a la ecuación de una recta $y = cx$. En este caso, el planteamiento del problema se reduce a encontrar el vector $c|x\rangle$, en el subespacio $S(|x\rangle)$, que esté lo más cercano posible al vector $|y\rangle$.

Como en el caso anterior, queremos que la distancia entre $|y\rangle$ y su valor más aproximado $c|x\rangle$ sea la menor posible. Por lo tanto, $\|cx - y\|^2$ será la menor posible y $|cx - y\rangle$ será perpendicular a $S(|x\rangle)$, por lo que:

$$\langle x | cx - y \rangle = 0 \Rightarrow c = \frac{\langle x | y \rangle}{\langle x | x \rangle} = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3 + \dots + x_n y_n}{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}.$$

Si la recta a “ajustar” es $y = cx + b$ el procedimiento será uno equivalente: proyectar sobre los vectores y obtener ecuaciones. Si representamos $|b\rangle = b|1\rangle$, tendremos:

$$|y\rangle = c|x\rangle + |b\rangle \Rightarrow \begin{cases} \langle x | y \rangle = c \langle x | x \rangle + \langle x | b \rangle \Rightarrow \sum_{i=1}^n x_i y_i = c \sum_{i=1}^n x_i^2 + b \sum_{i=1}^n x_i \\ \langle b | y \rangle = c \langle b | x \rangle + \langle b | b \rangle \Rightarrow \sum_{i=1}^n y_i = c \sum_{i=1}^n x_i + b n \end{cases}$$

Que no es otra cosa que un sistema de dos ecuaciones con dos incógnitas: c y b . Por simplicidad y conveniencia continuaremos con el análisis del caso $b = 0$, vale decir, approximando las medidas experimentales a una recta que pase por el origen de coordenadas $y = cx$. El lector podrá extender el procedimiento para caso $b \neq 0$.

Para tratar de entender (y extender) lo antes expuesto, consideremos tres ejemplos que muestran la versatilidad del método y la ventaja de disponer de una clara notación. Primeramente, mostraremos el caso más utilizado para construir el mejor ajuste lineal a un conjunto de datos experimentales. Buscaremos la mejor recta que describe ese conjunto de puntos. Luego mostraremos la aplicación del método para buscar la mejor

función multilineal, es decir, que ajustaremos la mejor función de n variables con una contribución lineal de sus argumentos: $y = y(x_1, x_2, \dots, x_n) = c_1x_1 + c_2x_2 + \dots + c_nx_n$. Este caso tendrá una complicación adicional, por cuando realizaremos varias mediciones de las variables. En los ejemplos de la sección 2.18 analizaremos varios casos en los cuales extendemos el método.

2.4.4 Interpolación polinomial de puntos experimentales

Muchas veces nos encontramos con la situación en la cual tenemos un conjunto de (digamos n) medidas o puntos experimentales $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ y para modelar ese experimento quisiéramos una función que ajuste estos puntos, de manera que: $\{(x_1, y_1 = f(x_1)), (x_2, y_2 = f(x_2)), \dots, (x_n, y_n = f(x_n))\}$.

El tener una función nos provee la gran ventaja de poder intuir o aproximar los puntos que no hemos medido. La función candidata más inmediata es un polinomio y debemos definir el grado del polinomio y la estrategia que aproxime esos puntos. Puede ser que el polinomio no sea lineal y entonces queramos ajustar esos puntos a un polinomio tal que éste pase por los puntos experimentales.

Queda entonces por decidir la estrategia, esto es, construimos la función como “trozos” de polinomios que ajusten a subconjuntos: $\{(x_1, y_1 = f(x_1)), (x_2, y_2 = f(x_2)), \dots, (x_m, y_m = f(x_m))\}$ con $m < n$ de los puntos experimentales. En este caso tendremos una función de interpolación para cada conjunto de puntos. También podemos ajustar la función a todo el conjunto de puntos experimentales de manera que el máximo grado del polinomio que los interpole será de grado $n - 1$.

Para encontrar este polinomio, y a manera de ejemplo, lo expresaremos como una combinación lineal de polinomios de Legendre, una base ortogonal que hemos discutido con anterioridad en las secciones 2.3.4 y 2.8. Esto es:

$$f(x) = \sum_{k=0}^{n-1} C^k |P_k\rangle = \sum_{k=0}^{n-1} C^k P_k(x) \Rightarrow \begin{cases} y_1 = f(x_1) = C^0 P_0(x_1) + C^1 P_1(x_1) + \dots + C^{n-1} P_{n-1}(x_1) \\ y_2 = f(x_2) = C^0 P_0(x_2) + C^1 P_1(x_2) + \dots + C^{n-1} P_{n-1}(x_2) \\ \vdots \\ y_n = f(x_n) = C^0 P_0(x_n) + C^1 P_1(x_n) + \dots + C^{n-1} P_{n-1}(x_n) \end{cases}$$

que no es otra cosa que un sistema de n ecuaciones con n incógnitas: los coeficientes $\{C^0, C^1, \dots, C^{n-1}\}$.

Al resolver el sistema de ecuaciones y obtener los coeficientes, podremos obtener la función polinómica que interpola esos puntos. Una expansión equivalente se pudo haber logrado con cualquier otro conjunto de polinomios ortogonales, ya que ellos son base del espacio de funciones.

Ejemplo 2.17

Series de Fourier Un ejemplo sencillo para aprender el mecanismo del cálculo de los coeficientes para la aproximación de funciones en series de Fourier lo podemos hacer para la siguiente función.

$$f(x) = x^2, \quad -\pi \leq x \leq \pi.$$

Los coeficientes serán:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 dx = \frac{2\pi^2}{3}, \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} x^2 \cos(nx) dx = \frac{4 \cos(n\pi)}{n^2} = \frac{4(-1)^n}{n^2}, \quad b_n = 0 \forall n.$$

Se puede verificar que si $f(x)$ es par ($f(-x) = f(x)$) sólo la parte que contiene $\cos(nx)$ contribuirá a la serie.

Si $f(x)$ es impar ($f(-x) = -f(x)$), lo hará sólo la parte que contiene $\sin(nx)$.

Por lo tanto:

$$f(x) = x^2 = \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(nx) = \frac{\pi^2}{3} - 4 \cos(x) + \cos(2x) - \frac{4}{9} \cos(3x) + \frac{1}{4} \cos(4x) - \dots$$

**Ejemplo 2.18****Mínimos cuadrados:**

Mostraremos como se puede utilizar el método de mínimos cuadrados para ajustar un conjunto de datos experimentales a un polinomio de cualquier grado. Veamos estos tres primeros casos:

- La situación más simple será el conjunto de datos experimentales: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, y nos preguntamos ¿Cuál será la recta, $y = cx$, que ajusta más acertadamente a estos puntos? Es inmediato darnos cuenta que la pendiente de esa recta será:

$$|y\rangle = c|x\rangle \Rightarrow c = \frac{\langle x |y\rangle}{\langle x |x\rangle}.$$

Consideremos los puntos experimentales: $\{(1, 2), (3, 2), (4, 5), (6, 6)\}$, entonces:

$$|y\rangle = c|x\rangle \Rightarrow \begin{pmatrix} 2 \\ 2 \\ 5 \\ 6 \end{pmatrix} = c \begin{pmatrix} 1 \\ 3 \\ 4 \\ 6 \end{pmatrix} \Rightarrow c = \frac{\langle x |y\rangle}{\langle x |x\rangle} = \frac{2 + 6 + 20 + 36}{1 + 9 + 16 + 36} = \frac{32}{31} = 1,03226.$$

¿Cuál sería el ajuste para el caso de una recta $y = cx + b$? Queda como ejercicio para el lector utilizar el esquema descrito en las ecuaciones (2.4.3) y encontrar los parámetros c y b para este nuevo ajuste.

- El segundo caso que consideraremos será la generalización a una contribución lineal de varios parámetros, i.e. un “ajuste” multilineal a la mejor función, $y = y(x_1, x_2, \dots, x_m) = c^1 x_1 + c^2 x_2 + \dots + c^m x_m$. El procedimiento anterior puede generalizarse de forma casi inmediata:

$$|y\rangle = \sum_{j=1}^m c^j |x_j\rangle \Rightarrow \begin{cases} \langle x^1 |y\rangle = c^1 \langle x^1 |x_1\rangle + c^2 \langle x^1 |x_2\rangle + \dots + c^m \langle x^1 |x_j\rangle \\ \langle x^2 |y\rangle = c^1 \langle x^2 |x_1\rangle + c^2 \langle x^2 |x_2\rangle + \dots + c^m \langle x^2 |x_j\rangle \\ \vdots \quad \vdots \\ \langle x^m |y\rangle = c^1 \langle x^m |x_1\rangle + c^2 \langle x^m |x_2\rangle + \dots + c^m \langle x^m |x_j\rangle. \end{cases}$$

Es decir, un sistema de m ecuaciones con m incógnitas que corresponden a los “pesos”, c^1, c^2, \dots, c^m , de las contribuciones a la función multilineal.

Supongamos ahora un paso más en este afán de generalizar y consideremos que ejecutamos n experimentos con $n > m$. De esta forma nuestro problema adquiere un grado mayor de riqueza y las medidas experimentales estarán representadas ahora por:

$$(y_1, x_{11}, x_{12}, \dots, x_{1m}; y_2, x_{21}, x_{22}, \dots, x_{2m}; y_3, x_{31}, x_{32}, \dots, x_{3m}; \dots, y_n, x_{n1}, x_{n2}, x_{n3}, \dots, x_{nm}),$$

por lo tanto, podremos construir m vectores:

$|\tilde{x}_1\rangle = (x_{11}, \dots, x_{1n})$; $|\tilde{x}_2\rangle = (x_{21}, \dots, x_{2n})$; \dots ; $|\tilde{x}_m\rangle = (x_{m1}, \dots, x_{mn})$; $|y\rangle = (y_{m1}, \dots, y_n)$, y el conjunto $\{|\tilde{x}_1\rangle, |\tilde{x}_2\rangle, \dots, |\tilde{x}_m\rangle\}$ expande el subespacio $S(|\tilde{x}_1\rangle, |\tilde{x}_2\rangle, \dots, |\tilde{x}_m\rangle)$, que alberga la aproximación de $|y\rangle$. Otra vez, la distancia de este subespacio al vector $|y\rangle$, será mínima, y por lo tanto, será la mejor aproximación:

$$[d(S(\tilde{c}^i|\tilde{x}_i\rangle), |y\rangle)]^2 = \langle S(\tilde{c}^i|\tilde{x}_i\rangle) - y |S(\tilde{c}^i|\tilde{x}_i\rangle) - y \rangle,$$

y $|S(\tilde{c}^i|\tilde{x}\rangle) - y\rangle$ será ortogonal a los $|\tilde{x}_i\rangle$:

$$\langle \tilde{x}^j |S(\tilde{c}^i|\tilde{x}\rangle) - y \rangle \equiv \langle \tilde{x}^j | \sum_{i=1}^m \tilde{c}^i |\tilde{x}_i\rangle - y \rangle = 0 \quad \forall \quad i, j = 1, 2, 3, \dots, m.$$

Finalmente, el problema queda de la misma manera que en el caso multilineal (2.):

$$|y\rangle = \sum_{j=1}^m \tilde{c}^j |\tilde{x}_j\rangle \Rightarrow \begin{cases} \langle \tilde{x}^1 |y\rangle = \tilde{c}^1 \langle \tilde{x}^1 |\tilde{x}_1\rangle + \tilde{c}^2 \langle \tilde{x}^1 |\tilde{x}_2\rangle + \cdots + \tilde{c}^m \langle \tilde{x}^1 |\tilde{x}_j\rangle \\ \langle \tilde{x}^2 |y\rangle = \tilde{c}^1 \langle \tilde{x}^2 |\tilde{x}_1\rangle + \tilde{c}^2 \langle \tilde{x}^2 |\tilde{x}_2\rangle + \cdots + \tilde{c}^m \langle \tilde{x}^2 |\tilde{x}_j\rangle \\ \vdots \\ \langle \tilde{x}^m |y\rangle = \tilde{c}^1 \langle \tilde{x}^m |\tilde{x}_1\rangle + \tilde{c}^2 \langle \tilde{x}^m |\tilde{x}_2\rangle + \cdots + \tilde{c}^m \langle \tilde{x}^m |\tilde{x}_j\rangle \end{cases}$$

3. Se puede extender el razonamiento anterior y generar un ajuste “lineal no lineal”. Esto es: el ajuste lineal es en los coeficientes, pero la funcionalidad de la ley a la cual queremos ajustar los datos puede ser un polinomio de cualquier orden. Ese es el caso de una parábola que ajusta, por ejemplo, al siguiente conjunto de puntos experimentales:

$$\{(0, 1), (1, 3), (2, 7), (3, 15)\} \Leftrightarrow y = ax^2 + bx + c.$$

Las ecuaciones toman la siguiente forma:

$$\begin{aligned} 1 &= 0 + 0 + c \\ 3 &= a + b + c \\ 7 &= 4a + 2b + c \\ 15 &= 9a + 3b + c \end{aligned}$$

y los vectores construidos a partir de los datos experimentales serán:

$$|x_1\rangle = (0, 1, 4, 9), \quad |x_2\rangle = (0, 1, 2, 3), \quad |x_3\rangle = (1, 1, 1, 1), \quad |y\rangle = (1, 3, 7, 15).$$

Una vez más, la ecuación vectorial sería: $|y\rangle = a|x_1\rangle + b|x_2\rangle + c|x_3\rangle$, y las ecuaciones normales (2.) para este sistema se construyen como:

$$\begin{aligned} 136 &= 98a + 36b + 14c \\ 62 &= 36a + 14b + 6c \\ 26 &= 14a + 6b + 4c \end{aligned} \Rightarrow \begin{cases} a = -6 \\ b = \frac{113}{5} \\ c = -\frac{32}{5} \end{cases} \Rightarrow y = -6x^2 + \frac{113}{5}x - \frac{32}{5}.$$

4. Para finalizar analicemos el caso típico de aproximación por mínimos cuadrados. Se sospecha que una determinada propiedad de un material cumple con la ecuación $y = ax_1 + bx_2$ y realizamos 4 mediciones independientes obteniendo:

$$\begin{pmatrix} y_1 \\ x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} 15 \\ 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} y_2 \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 12 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} y_3 \\ x_{31} \\ x_{32} \end{pmatrix} = \begin{pmatrix} 10 \\ 1 \\ 1 \end{pmatrix}; \quad \begin{pmatrix} y_4 \\ x_{41} \\ x_{42} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Es claro que tenemos un subespacio de $m = 2$ dimensiones, que hemos hecho $n = 4$ veces el experimento y que los vectores considerados arriba serán:

$$|x_1\rangle = (1, 2, 1, 1), \quad |x_2\rangle = (2, 1, 1, -1), \quad |y\rangle = (15, 12, 10, 0).$$

Por lo tanto, vectorialmente tendremos: $|y\rangle = a|x_1\rangle + b|x_2\rangle$, es decir, las ecuaciones normales (2.) se escribirán de la siguiente manera:

$$\begin{aligned} 7a + 4b &= 49 \\ 4a + 7b &= 52 \end{aligned} \Rightarrow \begin{cases} a = \frac{45}{11} \\ b = \frac{56}{11} \end{cases} \Rightarrow y = \frac{45}{11}x_1 + \frac{56}{11}x_2.$$



2.4.4.1 Aproximación polinómica

Consideremos los puntos experimentales representados en la figura 2.3:

$$\{(2, 8), (4, 10), (6, 11), (8, 18), (10, 20), (12, 34)\}.$$

Es importante hacer notar que debido a que los polinomios de Legendre están definido en el intervalo $[-1, 1]$ los puntos experimentales deberán reescalarse a ese intervalo para poder encontrar el polinomio de interpolación como combinación lineal de los polinomios de Legendre. Esto se puede hacer con la ayuda del siguiente cambio de variable:

$$x = \frac{(b-a)t + b+a}{2}, \quad dx = \frac{b-a}{2}dt,$$

donde a y b son los valores mínimos y máximos de los datos, respectivamente.

En este caso podemos ver que $a = 2$ y $b = 12$, y por lo tanto:

$$x = 5t + 7 \text{ y } dx = 5dt \Rightarrow t = (x - 7)/5 \text{ y } dt = dx/5.$$

Ahora los datos quedan de la siguiente forma:

$$\left\{(-1, 8), \left(-\frac{3}{5}, 10\right), \left(-\frac{1}{5}, 11\right), \left(\frac{1}{5}, 18\right), \left(\frac{3}{5}, 20\right), (1, 34)\right\}.$$

Ahora bien, necesitamos encontrar los coeficientes de la serie

$$f(x) = \sum_{k=0}^5 C^k P_k(x) = C^0 P_0(x) + C^1 P_1(x) + \cdots + C^5 P_5(x)$$

y el sistema de ecuaciones que resulta es el siguiente:

$$\left\{ \begin{array}{l} (-1, 8) \Rightarrow 8 = C^0 - C^1 + C^2 - C^3 + C^4 - C^5 \\ \left(-\frac{3}{5}, 10\right) \Rightarrow 10 = C^0 - \frac{3}{5}C^1 + \frac{1}{25}C^2 + \frac{9}{25}C^3 - \frac{51}{125}C^4 + \frac{477}{3125}C^5 \\ \left(-\frac{1}{5}, 11\right) \Rightarrow 11 = C^0 - \frac{1}{5}C^1 - \frac{11}{25}C^2 + \frac{7}{25}C^3 + \frac{29}{125}C^4 - \frac{961}{3125}C^5 \\ \left(\frac{1}{5}, 18\right) \Rightarrow 18 = C^0 + \frac{1}{5}C^1 - \frac{11}{25}C^2 - \frac{7}{25}C^3 + \frac{29}{125}C^4 + \frac{961}{3125}C^5 \\ \left(\frac{3}{5}, 20\right) \Rightarrow 20 = C^0 + \frac{3}{5}C^1 + \frac{1}{25}C^2 - \frac{9}{25}C^3 - \frac{51}{125}C^4 - \frac{477}{3125}C^5 \\ (1, 34) \Rightarrow 34 = C^0 + C^1 + C^2 + C^3 + C^4 + C^5 \end{array} \right.$$

Al resolver este sistema obtendremos los siguientes resultados:

$$C^0 = \frac{2249}{144}, \quad C^1 = \frac{3043}{336}, \quad C^2 = \frac{1775}{504}, \quad C^3 = -\frac{175}{216}, \quad C^4 = \frac{625}{336}, \quad C^5 = \frac{14375}{3024},$$

con lo cual:

$$\mathcal{P}(x) = f(x) = \frac{2249}{144} + \frac{3043}{336}x + \frac{1775}{504}P(2, x) - \frac{175}{216}P(3, x) + \frac{625}{336}P(4, x) + \frac{14375}{3024}P(5, x).$$

y la interpolación queda representada en la figura 2.3.

Nótese que mientras más puntos experimentales se incluyan para la interpolación, el polinomio resultante será de mayor grado y, por lo tanto incluirá oscilaciones que distorsionarán una aproximación más razonable. Por ello, la estrategia de hacer la interpolación a trozos, digamos de tres puntos en tres puntos, generará un mejor ajuste, pero será una función (un polinomio) continua a trozos.

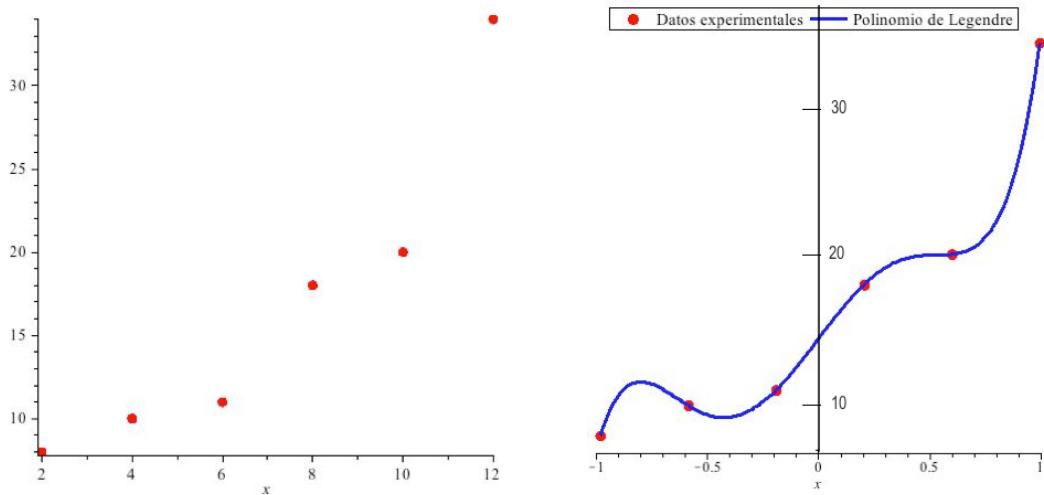


Figura 2.3: A la izquierda los puntos experimentales: $\{(2, 8), (4, 10), (6, 11), (8, 18), (10, 20), (12, 34)\}$ y a la derecha la función polinómica que los interpola.

2.4.5 Ejercicios

- Demuestre que con las identidades siguientes:

$$\cos(kx) = \frac{e^{ikx} + e^{-ikx}}{2}, \quad \sin(kx) = \frac{e^{ikx} - e^{-ikx}}{2i},$$

la serie de Fourier definida para funciones en el intervalo $(t, t + 2\pi)$ se escribe en su forma compleja como:

$$f(x) = \frac{1}{2}a_0 + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] \Rightarrow f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx},$$

con:

$$c_k = \frac{1}{2\pi} \int_t^{t+2\pi} f(x) e^{-ikx} dx.$$

Y que para funciones definidas en el intervalo $(l, l + 2L)$ como:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{\frac{ik\pi x}{L}}, \quad \text{con } c_k = \frac{1}{2L} \int_l^{l+2L} f(x) e^{-\frac{ik\pi x}{L}} dx.$$

Nota: Para los siguientes ejercicios supondremos la utilización del programa **Maxima**.

- Para las siguientes funciones determine la serie de Fourier calculando los coeficientes como en la sección 2.4 y compare los resultados con los cálculos hechos en el ambiente de manipulación simbólica.
 - $f(x) = x \sin(x)$, si $-\pi < x < \pi$.
 - $f(x) = e^x$, si $-\pi < x < \pi$.
 - $f(x) = x$, si $0 < x < 2$.
 - $f(x) = 2x - 1$, si $0 < x < 1$.
- Considere el espacio vectorial, $C_{[-1,1]}^\infty$, de funciones reales, continuas y continuamente diferenciables definidas en el intervalo $[-1, 1]$. Es claro que una posible base de este espacio de funciones la constituye el conjunto de monomios $\{1, x, x^2, x^3, x^4, \dots\}$ por cuanto estas funciones son linealmente independientes.
 - Si suponemos que este espacio vectorial está equipado con un producto interno definido por $\langle f | g \rangle = \int_{-1}^1 dx f(x)g(x)$, muestre que esa base de funciones no es ortogonal.
 - Utilizando la definición de producto interno $\langle f | g \rangle = \int_{-1}^1 dx f(x)g(x)$ ortogonalice la base

$\{1, x, x^2, x^3, x^4, \dots\}$ y encuentre los 10 primeros vectores ortogonales, base para $C_{[-1,1]}^\infty$. Esta nueva base de polinomios ortogonales se conoce como los polinomios de Legendre.

- (c). Modificando un poco la definición de producto interno para que ahora sea

$$\langle f|g \rangle = \int_{-1}^1 dx f(x)g(x)\sqrt{(1-x^2)},$$

ortogonalice la base $\{1, x, x^2, x^3, x^4, \dots\}$ y encuentre otros 10 primeros vectores ortogonales base para el mismo $C_{[-1,1]}^\infty$. Esta nueva base de polinomios ortogonales se conoce como los polinomios de Chebyshev.

- (d). Suponga la función $h(x) = \sin(3x)(1 - x^2)$:

- I. Expanda la función $h(x)$ en términos de la base de monomios y de polinomios de Legendre, grafique, compare y encuentre el grado de los polinomios en los cuales difieren las expansiones.
- II. Expanda la función $h(x)$ en términos de la base de monomios y de polinomios de Chebyshev, grafique, compare y encuentre el grado de los polinomios en los cuales difieren las expansiones.
- III. Expanda la función $h(x)$ en términos de la base de polinomios de Legendre y de Chebyshev, grafique, compare y encuentre el grado de los polinomios en los cuales difieren las expansiones.
- IV. Estime en cada caso el error que se comete como función del grado del polinomio (o monomio) de la expansión.

¿Qué puede concluir respecto a la expansión en una u otra base?

4. Parecido al ejercicio anterior, considere el espacio vectorial, $C_{[0,1]}^\infty$, de funciones reales, continuas y continuamente diferenciables definidas en el intervalo $[0, 1]$. Es claro que otra posible base de este espacio de funciones la constituye el conjunto de funciones exponenciales $\{1, e^x, e^{2x}, e^{3x}, e^{4x}, \dots\}$ por cuanto, al igual que el caso anterior estas funciones también son linealmente independientes. Adicionalmente, considere aquí la función $g(x) = \cos(3x^3)(1 - x^2)$,

- (a). Suponga que este espacio vectorial está equipado con un producto interno definido por

$$\langle f|g \rangle = \int_0^1 dx f(x)g(x),$$

y muestre que esa base de funciones $\{1, e^x, e^{2x}, e^{3x}, e^{4x}, \dots\}$ no es ortogonal.

- (b). Utilizando la definición de producto interno ortogonalice la base $\{1, e^x, e^{2x}, e^{3x}, e^{4x}, \dots\}$ y encuentre los 7 primeros vectores ortogonales, base para $C_{[0,1]}^\infty$, i.e. $\{1, E_1(x), E_2(x), E_3(x), E_4(x), \dots\}$.
- (c). Encuentre el valor de los coeficientes, C_i de la expansión

$$g_M(x) \approx \sum_{i=0}^7 C_i x^i,$$

utilizando la definición de producto interno anterior y estime también el error ϵ_M en esta aproximación

$$g_M(x) = \sum_{i=0}^7 C_i x^i + \epsilon_M.$$

- (d). Expanda en serie de Taylor la función anterior

$$g_T(x) \approx \sum_{n=0}^7 \left. \frac{dg(x)}{dx} \right|_{x=0} \frac{x^n}{n!},$$

estime el error ϵ_T y compárela con el caso anterior.

(e). Ahora encuentre el valor de los coeficientes, \tilde{C}_i de la expansión

$$g_E(x) \approx \sum_{m=0}^7 \tilde{C}_m e^{mx},$$

utilizando la definición de producto interno anterior y, del mismo modo, estime también el error ϵ_E en esta aproximación

$$g_E(x) = \sum_{m=0}^7 \tilde{C}_m e^{mx} + \epsilon_E.$$

¿Qué puede concluir de la comparación de los errores ϵ_M , ϵ_T y ϵ_E ?

(f). A continuación encuentre el valor de los coeficientes, \bar{C}_i de la expansión

$$g_{Eo}(x) \approx \sum_{m=0}^7 \bar{C}_m E_m(x),$$

y estime también el error ϵ_{Eo} en esta aproximación

$$g_{Eo}(x) = \sum_{m=0}^7 \bar{C}_m E_m(x) + \epsilon_{Eo}.$$

. ¿Otra vez, qué puede concluir de la comparación de los errores ϵ_M , ϵ_T , ϵ_E y ϵ_{Eo} ?

(g). Grafique las funciones $g(x)$, $g_M(x)$, $g_T(x)$, $g_E(x)$ y $g_{Eo}(x)$ y compare con los errores.

5. Al medir la temperatura a lo largo de una barra material obtenemos los siguientes valores:

x_i (cm)	1,0	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0
T_i ($^{\circ}$ C)	14,6	18,5	36,6	30,8	59,2	60,1	62,2	79,4	99,9

Encuentre, mediante el método de los mínimos cuadrados los coeficientes que mejor ajustan a la recta $T = ax + b$.

6. Los precios de un determinado producto varían como se muestra a continuación:

Año	2007	2008	2009	2010	2011	2012	2013	2014
Precio	133,5	132,2	138,7	141,5	144,2	144,5	148,6	153,8

Realice una interpolación polinomial que permita modelar estos datos.

Practicando con Maxima

1. Series de Fourier

Existe una biblioteca llamada **fourie** en **Maxima** que contiene instrucciones para el cálculo simbólico de series de Fourier de una función $f(x)$ en el intervalo $[-l, l]$: **fourier** (f, x, l). También encontraremos un conjunto de comandos en el paquete para calcular los coeficientes y para manipular las expresiones resultantes.

(%ci1) load(fourie)\$

En el ejemplo anterior aproximamos la función:

$$f(x) = x^2.$$

Veamos como se trabaja con el programa para calcular la serie de Fourier. Los resultados aparecerán en la forma de listas temporales y entre ellas los coeficientes. Las listas temporales serán indicadas con la notación (%t).

```
(%ci2) fx^2;
(%o2) x^2

(%ci3) fourier(f,x,%pi);
(%ot3) a_0 =  $\frac{\pi^2}{3}$ 
(%ot4) a_n =  $\frac{2 \left( \frac{\pi^2 \sin(\pi n)}{n} - \frac{2 \sin(\pi n)}{n^3} + \frac{2\pi \cos(\pi n)}{n^2} \right)}{\pi}$ 
(%ot5) b_n = 0
(%o5) [%t3, %t4, %t5]
```

Lo anterior se puede simplificar con el comando **foursimp**:

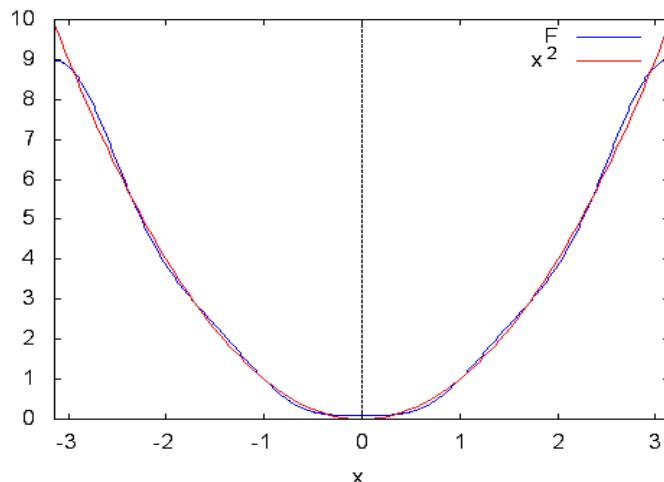
```
(%ci6) foursimp(%);
(%ot6) a_0 =  $\frac{\pi^2}{3}$ 
(%ot7) a_n =  $\frac{4(-1)^n}{n^2}$ 
(%ot8) b_n = 0
(%o8) [%t6, %t7, %t8]
```

Podemos evaluar la lista de los coeficiente hasta el término k . Aquí lo haremos hasta $k = 4$ y el resultado lo lo asignaremos a la variable F . Por otro lado, usaremos (%o8), la última salida, como entrada para siguiente comando.

```
(%ci9) F:fourexpand(%o8,x,%pi,4);
(%o9)  $\frac{\cos(4x)}{4} - \frac{4\cos(3x)}{9} + \cos(2x) - 4\cos x + \frac{\pi^2}{3}$ 
```

Construiremos ahora en un mismo gráfico la función original y los primeros 5 términos de la serie, de esta manera podremos comparar el resultado de la aproximación. Las opciones para realizar los diferentes gráficos en **Maxima** se pueden consultar en el manual de programa.

```
(%ci10) wxplot2d([F,f], [x,-%pi,%pi],[legend, "F", "x^ 2"])$
(%o10)
```



Veamos que sucede si escribimos:

```
(%ci11) totalfourier(f,x,%pi);
(%o11)  $a_0 = \frac{\pi^2}{3}$ 
(%o12)  $a_n = \frac{2 \left( \frac{\pi^2 \sin(\pi n)}{n} - \frac{2 \sin(\pi n)}{n^3} + \frac{2\pi \cos(\pi n)}{n^2} \right)}{\pi}$ 
(%o13)  $b_n = 0$ 
(%o14)  $a_0 = \frac{\pi^2}{3}$ 
(%o15)  $a_n = \frac{4(-1)^n}{n^2}$ 
(%o16)  $b_n = 0$ 
(%o16)  $4 \sum_{n=1}^{\infty} \frac{(-1)^n \cos(nx)}{n^2} + \frac{\pi^2}{3}$ 
```

En este caso fueron aplicados de manera simultánea los comandos **fourier** y **foursimp** para finalmente presentar la serie en forma de una sumatoria.

```
(%ci17) kill(all)$
```

2. Mínimos Cuadrados

Maxima puede estimar los parámetros que mejor se ajusten a una función $f = (x, y)$ para un conjunto de datos, utilizando el método de mínimos cuadrados. El programa buscará primero una solución exacta, si no la encuentra buscará una aproximada. El resultado lo presentará como una lista de ecuaciones. La función a utilizar será **lsquares**.

Vamos a considerar los ejemplos estudiados con anterioridad:

(a). En el primer ejemplo los datos eran los siguientes:

$$(x, y) = (1, 2), (3, 2), (4, 5), (6, 6).$$

Es necesario hacer uso de la biblioteca **lsquares** y los datos deben introducirse en forma de matriz.

```
(%i1) load(lsquares)$
(%i2) datos: matrix([1,2],[3,2],[4,5],[6,6])$
```

Por conveniencia, para más adelante hacer un gráfico, convertimos la matrix “datos” en una lista. Esto es sencillo si utilizamos el comando **args**:

```
(%i3) datosL: args(datos);
(%o3) [[1, 2], [3, 2], [4, 5], [6, 6]]
```

Supondremos entonces que los puntos se ajustan a un polinomio lineal del tipo: $y = ax$. El parámetro a se calcula con la función **lsquares_estimates**. Es importante prestar bastante atención a la sintaxis del siguiente comando.

```
(%i4) param: lsquares_estimates(datos,[x,y],y=a*x,[a]), numer;
(%o4) [[a = 1,032258064516129]]
```

Este será entonces el valor del parámetro a de la ecuación de la recta $y = ax$ que pasa por el origen. Notemos también que le hemos asignado el valor del parámetro a a la variable `param`.

Lo que haremos ahora es escribir la ecuación de dicha recta. Podemos hacer uso de la instrucción `ev` que nos permite evaluar una expresión.

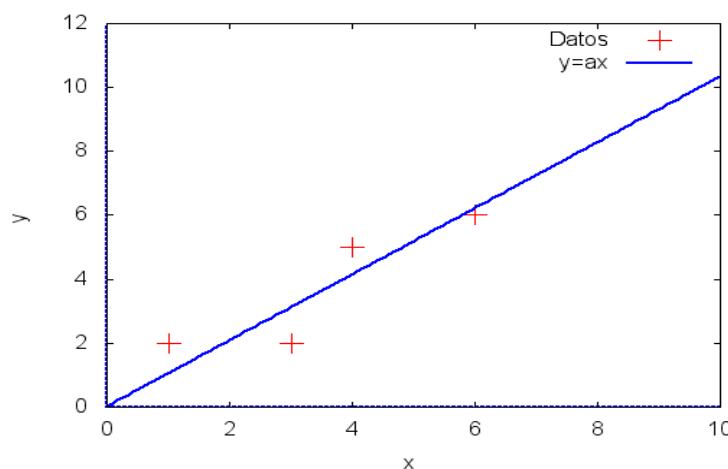
```
(%o15) y:ev(a*x,first(param));
(%o5) 1,032258064516129 x
```

Procederemos ahora a graficar los datos experimentales *vs* el ajuste por mínimos cuadrados en un mismo gráfico.

Recordemos que el conjunto de puntos lo tenemos en la forma de una lista, que hemos denominada más arriba como `datosL`. Mientras que al ajuste que hemos calculado, es decir, la recta: $1,032258064516129 x$ le hemos asignado la variable denominada y .

Es recomendable consultar el manual del programa, en la parte que tiene que ver con gráficos, `plot2d`, `plot3d`, para identificar la sintaxis que aparecerá en la siguiente instrucción.

```
(%o16) wxplot2d([[discrete,datosL], y], [x,0,10], [style, [points,5,2], [lines,2,1]],
[point_type, plus], [legend,"Datos","y=ax"],[xlabel, "x"],[ylabel, " y"]])$
```



Nota: Se deja como ejercicio repetir éste mismo cálculo pero usando un ajuste para los datos de la forma: $y = ax + b$.

- (b). Consideremos el conjunto de datos: $|x_1\rangle = (1, 2, 1, 1)$, $|x_2\rangle = (2, 1, 1, -1)$, $|y\rangle = (15, 12, 10, 0)$. Vamos a suponer que ajustan de la manera siguiente: $|y\rangle = a|x_1\rangle + b|x_2\rangle$.

```
(%o17) datos2: matrix([1,2,15],[2,1,12],[1,1,10],[1,-1,0])$
```

Cambiemos ligeramente la notación por: $z = ax + by$ y calculemos los parámetros a y b .

```
(%o18) param: lsquares_estimates(datos2,[x,y,z], z=a*x+b*y,[a,b]), numer;
(%o8) [[a = 4,090909090909091, b = 5,090909090909091]]
```

- (c). Para el tercer ejemplo se consideraron los siguientes datos:

$$\{(0, 1), (1, 3), (2, 7), (3, 15)\} \Leftrightarrow y = ax^2 + bx + c.$$

Haremos con **Maxima** el cálculo directo usando un ajuste cuadrático para los datos suministrados.

```
(%i9)  datos3: matrix([0,1],[1,3],[2,7],[3,15])$
```

Como hicimos anteriormente, hacemos una lista:

```
(%i10)  datosL3: args(datos3)$
```

Procedemos a calcular los parámetros:

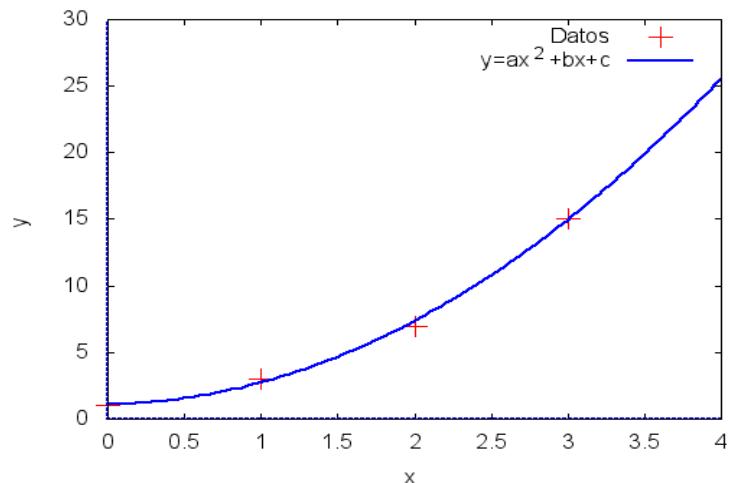
```
(%i11)  param: lsquares_estimates(datos3,[x,y], y=a*x^2+b*x+c,[a,b,c]), numer;
(%o11)  [[a = 1,5, b = 0,1, c = 1,1]]
```

Con los parámetros determinados podemos asignarlos a una nueva variable, en este caso y_2 ,

```
(%i12)  y2:ev(a*x^2+b*x+c,first(param))$
```

Como hicimos anteriormente, graficamos los datos y el ajuste cuadrático en una misma figura.

```
(%i13)  wxplot2d([[discrete,datosL3], y2], [x,0,4], [style, [points,5,2], [lines,2,1]],
[point_type, plus], [legend,"Datos"," y=ax^2+bx+c"], [xlabel, "x"], [ylabel, " y "]])$
```



```
(%i14)  kill(all)$
```

3. Polinomios ortogonales

Maxima contiene la librería **orthopoly** que nos permite acceder a la evaluación simbólica y numérica de los diferentes tipos de polinomios ortogonales: Chebyshev, Laguerre, Hermite, Jacobi, Legendre...

```
(%i1)  load (orthopoly)$
```

Por ejemplo, para obtener los primeros 6 polinomios de Legendre escribimos los siguientes comandos:

```
(%i2)  [legendre_p(0,x),legendre_p(1,x),legendre_p(2,x),
legendre_p(3,x),legendre_p(4,x),legendre_p(5,x)]$
```

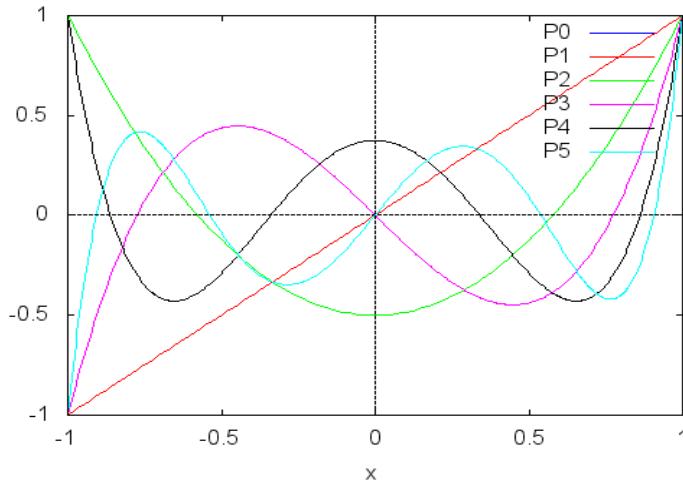
Simplificamos con **ratsimp**:

```
(%i3)  ratsimp (%);
```

$$(\%o3) \quad \left[1, x, \frac{3x^2 - 1}{2}, \frac{5x^3 - 3x}{2}, \frac{35x^4 - 30x^2 + 3}{8}, \frac{63x^5 - 70x^3 + 15x}{8} \right]$$

Los diferentes polinomios de Legendre se pueden visualizar de la manera siguiente:

(%ci4) `wxplot2d(%,[x,-1,1],[legend, "P0", "P1", "P2", "P3", "P4", "P5"]);`
 (%o4)



Ahora bien, con los datos de la figura 2.3 planteamos un sistema de ecuaciones lineales que ahora introduciremos en el programa:

(%ci4) `ecu1:C0-C1+C2-C3+C4-C5=8$`
 (%ci5) `ecu2:C0-3/5*C1+1/25*C2+9/25*C3-51/125*C4+477/3125*C5=10$`
 (%ci6) `ecu3:C0-1/5*C1-11/25*C2+7/25*C3+29/125*C4-961/3125*C5=11$`
 (%ci7) `ecu4:C0+1/5*C1-11/25*C2-7/25*C3+29/125*C4+961/3125*C5=18$`
 (%ci8) `ecu5:C0+3/5*C1+1/25*C2-9/25*C3-51/125*C4-477/3125*C5=20$`
 (%ci9) `ecu6:C0+C1+C2+C3+C4+C5=34$`

Procedemos a resolver el sistema de ecuaciones anterior:

(%ci10) `linsolve([ecu1,ecu2,ecu3,ecu4,ecu5,ecu6], [C0,C1,C2,C3,C4,C5]);`
 (%o10) $\left[C_0 = \frac{2249}{144}, C_1 = \frac{3043}{336}, C_2 = \frac{1775}{504}, C_3 = -\frac{175}{216}, C_4 = \frac{625}{336}, C_5 = \frac{14375}{3024} \right]$

Para asignar cada solución a la variable correspondiente podemos hacer lo siguiente:

(%ci11) `[C0,C1,C2,C3,C4,C5]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4]),rhs(%[5]),rhs(%[6])];`
 (%o11) $\left[\frac{2249}{144}, \frac{3043}{336}, \frac{1775}{504}, -\frac{175}{216}, \frac{625}{336}, \frac{14375}{3024} \right]$

Por lo tanto, la función aproximada será la siguiente:

(%ci13) `f:C0+C1*legendre_p(1,x)+C2*legendre_p(2,x)+C3*legendre_p(3,x)`
`+C4*legendre_p(4,x)+legendre_p(5,x)*C5$`
 (%ci14) `f:expand(%);`

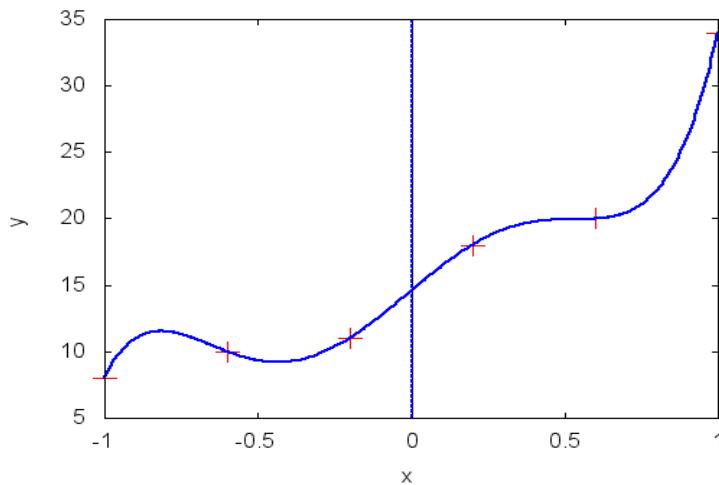
$$(\%o14) \frac{14375 x^5}{384} + \frac{3125 x^4}{384} - \frac{8375 x^3}{192} - \frac{325 x^2}{192} + \frac{7367 x}{384} + \frac{1863}{128}$$

Procedemos a introducir los datos:

$$(\%o15) \text{datos:} [[-1,8],[-3/5,10],[-1/5,11],[1/5,18],[3/5,20],[1,34]]; \\ (\%o15) \left[[-1, 8], \left[-\frac{3}{5}, 10 \right], \left[-\frac{1}{5}, 11 \right], \left[\frac{1}{5}, 18 \right], \left[\frac{3}{5}, 20 \right], [1, 34] \right]$$

Haremos ahora la gráfica con los datos y con la interpolación:

$$(\%o16) \text{wxplot2d}([[discrete,datos],f], [x,-1,1], [\text{style, [points,5,2]}, [\text{lines,2,1}]], [\text{point_type, plus}], [\text{legend, false}], [\text{xlabel, "x"}], [\text{ylabel, "y"}])\$ \\ (\%o16)$$



Bibliografía

- [1] Arfken, G. B., Weber, H., Weber, H.J. (2000) **Mathematical Methods for Physicists** 5ta Edición (Academic Press, Nueva York)
- [2] Scharnhorst, K. (2009) **Angles in complex vector spaces** Acta Applicandae Mathematica), **69**(1), 95-103
- [3] Dennery, P. y Krzywicki, A. (1995) **Mathematics for Physicists** (Dover Publications Inc, Nueva York)
- [4] Harper, C. (1971) **Introduction to Mathematical Physics** (*Prentice Hall, Englewood Cliff, N.J:*)
- [5] Hassani, S. (1991) **Foundations of Mathematical Physics** (*Prentice Hall, International Edition, London:*)
- [6] Riley, K.F., Hobson, M.P. y Bence, S.J. (2002) **Mathematical Methods for Physics and Engineering** (*Cambridge University Press*)

Capítulo 3

Vectores duales y tensores

La ruta de este capítulo

Este capítulo completa el esfuerzo de formalización de conceptos que comenzamos en el capítulo anterior. Iniciamos con el estudio de los funcionales lineales, extendiendo el concepto de función al de una aplicación entre elementos de espacios vectoriales, esta extensión nos llevará a la definición de los espacios vectoriales duales.

Luego incorporamos el concepto de 1-forma como el conjunto de funcionales que forman al espacio vectorial dual. Ahora el producto interno se definirá entre 1-formas y vectores. Y en la sección 3.2 construiremos una nueva clase de objetos matemáticos: los tensores, los cuales pueden verse como arreglos multilineales y constituyen una extensión de objetos ya conocidos: los escalares, los vectores y las 1-formas.

Desarrollaremos el álgebra tensorial y sus leyes de transformación. Ejemplificamos la utilización de este concepto en Física cuando discutimos, en la sección 3.3, el tensor de inercia y el de esfuerzos (*stress*). En la sección 3.4 incursionaremos en los espacios pseudoeuclidianos para mostrar una situación en la cual podemos diferenciar 1-formas de vectores, aprovechamos además para introducir algunas nociones básicas de la teoría especial de la relatividad.

Para finalizar, en la sección 3.5 extenderemos los conceptos de espacios vectoriales y bases discretas a espacios de funciones y bases continuas y en ese contexto discutimos algunos rudimentos de teorías de distribuciones.

3.1 Funcionales lineales

En los cursos más elementales de cálculo se estudiaron funciones de una y varias variables reales. Estas funciones pueden considerarse que actúan sobre vectores en \mathbb{R}^3 y podemos extender esta idea para otras que tengan como argumento vectores de un espacio vectorial abstracto.

Como hemos visto, los espacios vectoriales pueden estar formados por funciones, es decir, por aplicaciones que pueden transformar unas funciones en otras. Comenzaremos con las más sencillas, las lineales, que también son conocidas como *operadores lineales*.

Definición 3.1 (Funcional lineal)

Un funcional lineal es una operación que asocia un número complejo (o real) $\in \mathbf{K}$ a un vector $|v\rangle \in \mathbf{V}$, es decir:

$$\forall |v\rangle \in \mathbf{V} \Rightarrow \mathcal{F}[|v\rangle] \in \mathbb{C},$$

y cumple con:

$$\mathcal{F}[\alpha|v_1\rangle + \beta|v_2\rangle] \equiv \alpha\mathcal{F}[|v_1\rangle] + \beta\mathcal{F}[|v_2\rangle], \quad \forall |v_1\rangle, |v_2\rangle \in \mathbf{V} \text{ y } \forall \alpha, \beta \in \mathbf{K}.$$

En otras palabras, un funcional lineal (o forma lineal) es un *morfismo*¹ del espacio lineal \mathbf{V} a un espacio \mathbf{K} .

Notemos que cuando se escoge una base $\{|e_i\rangle\}$ de un espacio vectorial \mathbf{V} de manera que para cualquier vector $|v\rangle \in \mathbf{V}$ se especifican sus componentes $\{\xi^i\}$ respecto a esa base, es decir, $|v\rangle = \xi^i|e_i\rangle$ (recordemos el uso de la convención de Einstein para las sumas: 1.4.1), entonces, lo que se tiene para cada componente es un funcional lineal, como por ejemplo, $\mathcal{F}[|v\rangle] = \xi^1$. Algo parecido ocurre con el producto interno, cuando en un espacio vectorial \mathbf{V} se define el producto escalar $\langle v_0|v\rangle$, del vector $|v\rangle$ con un vector fijo $|v_0\rangle$, lo que tenemos es un funcional lineal $\mathcal{F}[|v\rangle] = \langle v_0|v\rangle = \alpha \in \mathbf{K}$.

Otro ejemplo sencillo de un funcional lineal es la integral de Riemann que podemos interpretar de la manera siguiente:

$$\mathcal{F}[|f\rangle] = \int_a^b f_0(x)f(x)dx,$$

donde $f(x), f_0(x) \in \mathcal{C}_{[a,b]}$, es decir, pertenece al espacio vectorial de funciones reales y continuas en el intervalo $[a, b]$ y $f_0(x)$ es una función que se toma como fija.

3.1.1 Espacio vectorial dual

El conjunto de funcionales lineales $\{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3, \dots, \mathcal{F}_n, \dots\}$ constituyen a su vez un espacio vectorial, el cual se denomina espacio vectorial dual de \mathbf{V} –que es el espacio directo– y se denotará como \mathbf{V}^* (aquí $*$ no es complejo conjugado al que hemos denotado también con $*$).

Si \mathbf{V} es de dimensión finita n , entonces $\dim \mathbf{V} = \dim \mathbf{V}^* = n$. Es fácil convencerse que los funcionales lineales forman un espacio vectorial ya que, dados $\mathcal{F}_1, \mathcal{F}_2 \in \mathbf{V}^*$ se tiene:

$$\left. \begin{array}{l} (\mathcal{F}_1 + \mathcal{F}_2)[|v\rangle] = \mathcal{F}_1[|v\rangle] + \mathcal{F}_2[|v\rangle] \\ (\alpha \mathcal{F})[|v\rangle] = \alpha \mathcal{F}[|v\rangle] \end{array} \right\} \quad \forall |v\rangle \in \mathbf{V}.$$

A este espacio lineal se le llama espacio de formas lineales y, a los funcionales se les denomina 1-formas o covectores.

Como ya lo mencionamos, en aquellos espacios lineales con producto interno definido, el mismo producto interno constituye la expresión natural del funcional. Así tendremos que:

$$\mathcal{F}_a[|v\rangle] \equiv \langle a|v\rangle \quad \forall |v\rangle \in \mathbf{V} \quad \wedge \quad \forall \langle a| \in \mathbf{V}^*.$$

¹Entenderemos por morfismo a toda regla o mapa que asigne a todo vector $|v\rangle$ de un espacio vectorial \mathbf{V} un vector $|w\rangle$ de un espacio vectorial \mathbf{W} y usualmente se denota por $\mathcal{F} : \mathbf{V} \Rightarrow \mathbf{W}$.

Es claro comprobar que el producto interno garantiza que los $\{\mathcal{F}_a, \mathcal{F}_b, \dots\}$ forman un espacio vectorial:

$$\left. \begin{aligned} (\mathcal{F}_a + \mathcal{F}_b)[|v\rangle] &= \mathcal{F}_a[|v\rangle] + \mathcal{F}_b[|v\rangle] = \langle a|v\rangle + \langle b|v\rangle \\ (\alpha \mathcal{F}_a)[|v\rangle] &= \langle \alpha a|v\rangle = \alpha^* \langle a|v\rangle = \alpha^* \mathcal{F}_a[|v\rangle] \end{aligned} \right\} \quad \forall |v\rangle \in V.$$

Esta última propiedad se conoce como antilinealidad.

Se establece entonces una correspondencia 1 a 1 entre *kets* y *bras*, entre vectores y funcionales lineales (o formas diferenciales):

$$\lambda_1|v_1\rangle + \lambda_2|v_2\rangle \quad \Leftrightarrow \quad \lambda_1^*\langle v_1| + \lambda_2^*\langle v_2|,$$

que ahora podemos precisar de la siguiente forma:

$$\begin{aligned} \langle a|v\rangle &= \langle v|a\rangle^*, \\ \langle a|\lambda_1 v_1 + \lambda_2 v_2\rangle &= \lambda_1 \langle a|v_1\rangle + \lambda_2 \langle a|v_2\rangle, \\ \langle \lambda_1 a_1 + \lambda_2 a_2|v\rangle &= \lambda_1^* \langle a_1|v\rangle + \lambda_2^* \langle a_2|v\rangle. \end{aligned}$$

Más aún, dada una base $\{|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle\}$ para V siempre es posible asociar una base para V^* de tal manera que:

$$|v\rangle = \xi^i|e_i\rangle \Leftrightarrow \langle v| = \xi_i^* \langle e^i|, \quad \text{con } \xi^i = \langle e^i|v\rangle \wedge \xi_i^* = \langle v|e_i\rangle, \quad \text{para } i = 1, 2, \dots, n.$$

En un lenguaje arcaico (y muchos textos de mecánica todavía lo reproducen) se denota a la base del espacio dual $\{\langle e^i|\}$ como la base recíproca de la base $\{|e_i\rangle\}$, este caso lo ilustraremos más adelante.

Se puede ver también que si $|v\rangle = \xi^i|e_i\rangle$, entonces

$$\mathcal{F}[|v\rangle] = \mathcal{F}[\xi^i|e_i\rangle] = \xi^i \mathcal{F}[|e_i\rangle] = \xi^i \omega_i, \quad \text{con } \mathcal{F}[|e_i\rangle] \equiv \omega_i.$$

Nótese que estamos utilizando la notación de Einstein en la que índices repetidos indican suma. También que las bases del espacio dual de formas diferenciales $\{\langle e^k|\}$ llevan los índices arriba, los llamaremos índices contravariantes, mientras que los índices abajo, serán covariantes. Por lo tanto, las componentes de las formas diferenciales en una base dada, llevan índices abajo $\langle a| = a_i \langle e^i|$ mientras que las componentes de los vectores los llevan arriba $|v\rangle = \xi^j|e_j\rangle$.

Observe también que dada una base en el espacio directo $\{|\hat{e}_i\rangle\}$ existe una única base canónica en el dual definida como:

$$\langle \hat{e}^i|\hat{e}_j\rangle = \mathcal{F}^i[|\hat{e}_j\rangle] = \delta_j^i.$$

Esta 1-forma al actuar sobre un vector arbitrario resulta en:

$$\mathcal{F}^i[|v\rangle] = \mathcal{F}^i[\xi^j|\hat{e}_j\rangle] = \xi^j \mathcal{F}^i[|\hat{e}_j\rangle] = \xi^j \delta_j^i = \xi^i$$

su componente contravariante. El conjunto de 1-formas $\{\mathcal{F}^i\}$ será linealmente independientes.

Si $|v\rangle = \xi^i|\hat{e}_i\rangle$ es un vector arbitrario en el espacio directo y $\langle a| = a_i \langle \hat{e}^i|$ un vector en el espacio dual, entonces:

$$\langle a|v\rangle = \langle a_i \langle \hat{e}^i| | \xi^j|\hat{e}_i\rangle \rangle = a_i^* \xi^j \langle \hat{e}^i|\hat{e}_j\rangle = a_i^* \xi^j \delta_j^i = a_i^* \xi^i,$$

y para bases arbitrarias, no ortogonales, $\{|w_i\rangle\}$ de V y $\{\langle w^i|\}$ de V^* se tiene:

$$\langle a|v\rangle = \langle w^i|w_j\rangle a_i^* \xi^j.$$

Nótese que $\langle w^i|w_j\rangle$ son las proyecciones de cada uno de los vectores base en las distintas direcciones.

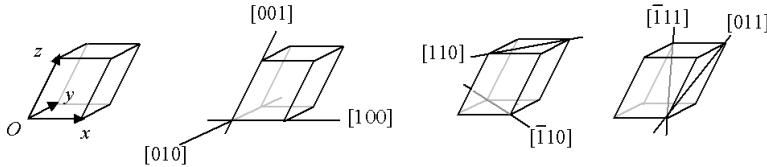


Figura 3.1: Bases directas y recíprocas. Note como los covectores representan planos perpendiculares a dos vectores de la base directa. Figura tomada de <https://commons.wikimedia.org/w/index.php?curid=457850> CC BY-SA 3.0,

3.1.2 Espacios duales y bases recíprocas

Un ejemplo que ilustra las bases duales es la construcción de las llamadas bases recíprocas que mencionamos anteriormente en los ejercicios de aplicaciones del álgebra vectorial en la página 28. Este ejemplo ilustra la pregunta de siempre: si los vectores se representan geométricamente como un segmento orientado, con módulo, dirección y sentido ¿cómo representamos geométricamente una forma lineal o covector? Las formas diferenciales representan vectores perpendiculares a planos formados por dos vectores de la base directa.

Los cristalográficos lograron exemplificar las bases duales y luego la aplicación de la difracción de rayos X a la cristalografía potenció utilidad de este concepto.

3.1.2.1 Bases recíprocas

Consideremos el problema de expandir un vector $|a\rangle$ con respecto a una base oblicua (no es ortogonal), $\{|\mathbf{w}_i\rangle\}$, de tal forma que, $|a\rangle = a^i |\mathbf{w}_i\rangle$. Por simplicidad, tomemos el caso \mathbb{R}^3 , de manera que $\mathbf{a} = a^i \mathbf{w}_i$ ($i = 1, 2, 3$). Al proyectar el vector \mathbf{a} sobre los ejes de algún sistema de coordenadas, es posible resolver el sistema de tres ecuaciones que resulta para las incógnitas a^i . Las bases de vectores \mathbf{w}_i y de formas o covectores \mathbf{w}^i serán duales, y por construcción satisfacen $\mathbf{w}_i \cdot \mathbf{w}^j = \delta_i^j$. Es decir, cada uno de los vectores bases duales es perpendicular a los otros dos de la base dual: \mathbf{w}^1 será perpendicular a \mathbf{w}_2 y \mathbf{w}_3 : $\mathbf{w}^1 = \alpha(\mathbf{w}_2 \times \mathbf{w}_3)$. Como $\mathbf{w}_1 \cdot \mathbf{w}^1 = 1$, entonces:

$$\alpha \mathbf{w}_1 \cdot (\mathbf{w}_2 \times \mathbf{w}_3) = 1 \Rightarrow \alpha = \frac{1}{\mathbf{w}_1 \cdot (\mathbf{w}_2 \times \mathbf{w}_3)} \Rightarrow \mathbf{w}^1 = \frac{\mathbf{w}_2 \times \mathbf{w}_3}{\mathbf{w}_1 \cdot (\mathbf{w}_2 \times \mathbf{w}_3)},$$

y en general, es fácil ver que:

$$\mathbf{w}^i = \frac{\mathbf{w}_j \times \mathbf{w}_k}{\mathbf{w}_i \cdot (\mathbf{w}_j \times \mathbf{w}_k)}, \quad (3.1)$$

donde i, j, k son permutaciones cíclicas de 1, 2, 3.

La definición (3.1) nos permite construir la base de 1-formas o covectores, a partir de su ortogonalidad con los vectores de la base directa. Notemos también que $V = \mathbf{w}_i \cdot (\mathbf{w}_j \times \mathbf{w}_k)$ es el volumen del paralelepípedo que soportan los vectores $\{\mathbf{w}_i\}$ y que además se puede obtener una expresión análoga para los $\{\mathbf{w}_i\}$ en término de los $\{\mathbf{w}^i\}$. Al ser $\{\mathbf{w}_i\}$ y $\{\mathbf{w}^i\}$ duales, $\mathbf{a} = a^i \mathbf{w}_i \Rightarrow \mathbf{w}^i \cdot \mathbf{a} = \mathbf{w}^i \cdot (a^i \mathbf{w}_i) = a^i (\mathbf{w}^i \cdot \mathbf{w}_i) = a^i \delta_i^i = a^i$, con $i = 1, 2, 3$ y equivalentemente, el covector de \mathbf{a} : $\bar{\mathbf{a}} = \bar{a}_j \mathbf{w}^j \Rightarrow \bar{\mathbf{a}} \cdot \mathbf{w}_i = (\bar{a}_j \mathbf{w}^j) \cdot \mathbf{w}_i = \bar{a}_j (\mathbf{w}^j \cdot \mathbf{w}_i) = \bar{a}_j \delta_j^i = \bar{a}_i$, con $i = 1, 2, 3$. Hemos querido enfatizar el carácter dual del covector $\bar{\mathbf{a}}$ al vector \mathbf{a} y obviamente $\bar{\mathbf{a}} \cdot \mathbf{a} = |\mathbf{a}|^2$.

De la misma ecuación (3.1) se puede ver que la representación gráfica de una forma será una superficie orientada. La figura 3.1 muestra como, a partir de una base directa de vectores oblicuos se construyen planos perpendiculares a esos vectores y la orientación (dirección y sentido) representan las direcciones de los vectores duales. En cristalografía esa base dual se le representa como los índices de Miller.

Claramente si la base directa es ortonormal su dual también lo será y, más importante aún, ambas bases

coinciden $\mathbf{e}^i \equiv \mathbf{e}_i$, y si la base original o directa es dextrógira su dual también lo será.

Ejemplo 3.1 Dados los vectores: $\mathbf{u}_1 = \hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}$, $\mathbf{u}_2 = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$, y $\mathbf{u}_3 = \hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 4\hat{\mathbf{k}}$. Revisaremos si estos vectores son mutuamente ortogonales. Encontraremos la base recíproca \mathbf{u}^i , el tensor métrico en ambas bases y para el vector $\mathbf{a} = 3\mathbf{u}_1 + 2\mathbf{u}_2 + \mathbf{u}_3$ encontraremos sus componentes covariantes.

Para saber si son ortogonales simplemente calculamos el producto escalar entre ellos:

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 9, \quad \mathbf{u}_1 \cdot \mathbf{u}_3 = 6 \quad \text{y} \quad \mathbf{u}_2 \cdot \mathbf{u}_3 = 7,$$

por lo tanto no son ortogonales y adicionalmente sabemos que:

$$\mathbf{u}^i = \frac{\mathbf{u}_j \times \mathbf{u}_k}{\mathbf{u}_i \cdot (\mathbf{u}_j \times \mathbf{u}_k)}$$

Procederemos a calcular primero el denominador:

$$V = \mathbf{u}_1 \cdot (\mathbf{u}_2 \times \mathbf{u}_3) \Rightarrow \left(\hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}} \right) \cdot \left(\left[\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}} \right] \times \left[\hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 4\hat{\mathbf{k}} \right] \right) = 6.$$

En general:

$$\mathbf{u}^i = \frac{\mathbf{u}_j \times \mathbf{u}_k}{V} \Rightarrow \begin{cases} \mathbf{u}^1 = \frac{\mathbf{u}_2 \times \mathbf{u}_3}{V} = \frac{17}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} - \frac{5}{6}\hat{\mathbf{k}} \\ \mathbf{u}^2 = \frac{\mathbf{u}_3 \times \mathbf{u}_1}{V} = -\frac{5}{3}\hat{\mathbf{i}} + \frac{1}{3}\hat{\mathbf{j}} + \frac{2}{3}\hat{\mathbf{k}} \\ \mathbf{u}^3 = \frac{\mathbf{u}_1 \times \mathbf{u}_2}{V} = -\frac{1}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} + \frac{1}{6}\hat{\mathbf{k}} \end{cases}$$

Notemos que:

$$\tilde{V} = \mathbf{u}^1 \cdot (\mathbf{u}^2 \times \mathbf{u}^3) \Rightarrow \left(\frac{17}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} - \frac{5}{6}\hat{\mathbf{k}} \right) \cdot \left(\left[-\frac{5}{3}\hat{\mathbf{i}} + \frac{1}{3}\hat{\mathbf{j}} + \frac{2}{3}\hat{\mathbf{k}} \right] \times \left[-\frac{1}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} + \frac{1}{6}\hat{\mathbf{k}} \right] \right) = \frac{1}{6}.$$

▷

3.1.2.2 Bases duales

Si tenemos un espacio vectorial V y una definición de producto interno podremos conectar la base $\{|\hat{\mathbf{e}}_i\rangle\}$ de V con la base $\{|\hat{\mathbf{e}}^j|\}$ de V^* , imponiendo la ortogonalidad entre las bases, i.e. $\langle \hat{\mathbf{e}}^i | \hat{\mathbf{e}}_j \rangle = \delta_j^i$. En general, tal y como hemos dicho al comienzo de este capítulo (sección 3.1.1), una forma diferencial o 1-forma, $\langle b | \in V^*$ es susceptible de expresarse en una base $\{|\hat{\mathbf{e}}^i|\}$ del espacio dual V^* como $b_i |\hat{\mathbf{e}}^i|$ y, como el espacio está equipado con un producto interno, entonces:

$$\langle a | b \rangle = \langle b | a \rangle = b_i a^j \langle \hat{\mathbf{e}}^i | \hat{\mathbf{e}}_j \rangle = b_i a^j \delta_j^i = a^i b_i.$$

3.1.3 Vectores, formas, componentes y leyes de transformación

Tal y como hemos mencionado anteriormente (tempranamente en la sección 1.4.3 y luego en la sección 2.3.4), un vector $|a\rangle \in V$ puede expresarse en una base ortonormal $\{|\hat{\mathbf{e}}_j\rangle\}$ como: $a^j |\hat{\mathbf{e}}_j\rangle$ donde las a^j son las componentes *contravariantes* del vector en esa base. En general, como es muy largo decir “componentes del vector contravariante” uno se refiere (y nos referiremos de ahora en adelante) al conjunto $\{a^j\}$ como un *vector contravariante* obviando la precisión de *componente*, pero realmente las a^j son las componentes del vector.

Adicionalmente, en esta etapa pensaremos a las bases como distintos observadores o sistemas de referencias. Con ello tendremos (algo que ya sabíamos) que un vector se puede expresar en distintas bases y tendrá distintas componentes referidas a esa base

$$|a\rangle = a^j |\hat{\mathbf{e}}_j\rangle = a^{j'} |\hat{\mathbf{e}}_{j'}\rangle. \quad (3.2)$$

Así una misma cantidad física vectorial se verá distinta (tendrá distintas componentes) desde diferentes sistemas de coordenadas. Las distintas “visiones” están conectadas mediante un transformación de sistema de referencia como veremos más adelante. Con lo cual avanzamos otra vez en la interpretación de este tipo de objetos: una cantidad física escalar se verá igual (será invariante) desde distintos sistemas de referencia.

Sabemos que unas y otras componentes se relacionan y por (3.2) tenemos:

$$\left. \begin{array}{l} \langle \hat{e}^i | a \rangle = a^j \langle \hat{e}^i | \hat{e}_j \rangle = a^j \delta_j^i = a^{j'} \langle \hat{e}^i | \hat{e}_{j'} \rangle \\ \langle \hat{e}^{i'} | a \rangle = a^{j'} \langle \hat{e}^{i'} | \hat{e}_{j'} \rangle = a^{j'} \delta_{j'}^{i'} = a^j \langle \hat{e}^{i'} | \hat{e}_j \rangle \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} a^i = A_{j'}^{i'} a^{j'} \\ a^{i'} = A_j^{i'} a^j, \end{array} \right.$$

donde claramente:

$$\langle \hat{e}^i | \hat{e}_{j'} \rangle = A_{j'}^i, \quad \langle \hat{e}^{i'} | \hat{e}_j \rangle = A_j^{i'} \quad \text{y} \quad A_k^i A_j^{k'} = \delta_j^i \iff A_{j'}^i = (A_j^{i'})^{-1}.$$

Diremos entonces que aquellos objetos cuyas componentes transforman como: $a^i = A_{j'}^i a^{j'}$ o, equivalentemente como: $a^{i'} = A_j^{i'} a^j$ serán vectores, o en un lenguaje un poco más antiguo, vectores contravariantes.

Tradicionalmente, e inspirados en la ley de transformación, la representación matricial de las componentes contravariantes de un vector, $\langle \hat{e}^i | a \rangle = a^i$, para una base determinada $\{|\hat{e}_j\rangle\}$ se representan como una columna

$$|a\rangle \Rightarrow a^i = \langle \hat{e}^i | a \rangle \iff \begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^n \end{pmatrix}, \quad (3.3)$$

con: $i = 1, 2, 3, \dots, n$.

De la misma manera, en el espacio dual, V^* , las formas diferenciales se podrán expresar en término de una base de ese espacio vectorial como

$$\langle b | = b_i \langle \hat{e}^i | = b_{i'} \langle \hat{e}^{i'} |. \quad (3.4)$$

Las $\{b_i\}$ serán las componentes de las formas diferenciales o las componentes *covariantes* de un vector $|b\rangle$, o dicho rápidamente un *vector covariante* o *covector*. Al igual que en el caso de las componentes contravariantes las componentes covariantes transforman de un sistema de referencia a otro mediante la siguiente ley de transformación:

$$\left. \begin{array}{l} \langle b | \hat{e}_j \rangle = b_i \langle \hat{e}^i | \hat{e}_j \rangle = b_i \delta_j^i = b_{i'} \langle \hat{e}^{i'} | \hat{e}_j \rangle \\ \langle b | \hat{e}_{j'} \rangle = b_{i'} \langle \hat{e}^{i'} | \hat{e}_{j'} \rangle = b_{i'} \delta_{j'}^{i'} = b_i \langle \hat{e}^i | \hat{e}_{j'} \rangle \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} b_j = b_{i'} A_j^{i'} \\ b_{j'} = b_i A_{j'}^{i'} . \end{array} \right.$$

Otra vez, objetos cuyas componentes transformen como $b_j = b_{i'} A_j^{i'}$ los denominaremos formas diferenciales o *vectores covariantes* o *covectores* y serán representados como matrices en un arreglo tipo fila:

$$\langle b | \Rightarrow b_i = \langle b | \hat{e}_i \rangle \iff \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix}, \quad (3.5)$$

con: $i = 1, 2, 3, \dots, n$.

Quizá hasta este punto la diferencia de formas y vectores, de componentes covariantes y contravariantes, así como sus esquemas de transformación es todavía confusa. No disponemos de ejemplos contundentes que ilustren esa diferencia. Estos serán evidentes cuando nos toque discutir las características de los espacios pseudoeuclidianos en la sección 3.4.

Ejemplo 3.2 Consideremos $V = \mathbb{R}^3$ como el espacio vectorial conformado por todos los vectores columna

$$|v\rangle = \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix},$$

el cual al representarse en su base canónica $\{\hat{i}_i\}$ resulta en:

$$|v\rangle = \xi^1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \xi^2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \xi^3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \xi^1 |\hat{i}_1\rangle + \xi^2 |\hat{i}_2\rangle + \xi^3 |\hat{i}_3\rangle = \xi^i |\hat{i}_i\rangle.$$

Sea un funcional lineal $\mathcal{F} \in V^*$, de manera que los vectores duales $\mathcal{F}[\circ] \equiv \langle F | \leftrightarrow (w_1, w_2, w_3)$, puedan ser representados por “vectores” filas.

Notemos que la base de funcionales lineales $\zeta^i[\circ] \equiv \langle \hat{i}^i |$, la definimos como:

$$\zeta^i [\hat{i}_j] = \langle \hat{i}^i | \hat{i}_j \rangle = \delta_j^i \Rightarrow \langle \hat{i}^1 | \hat{i}_1 \rangle = 1, \langle \hat{i}^1 | \hat{i}_2 \rangle = 0, \langle \hat{i}^1 | \hat{i}_3 \rangle = 0, \langle \hat{i}^2 | \hat{i}_1 \rangle = 0, \langle \hat{i}^2 | \hat{i}_2 \rangle = 1, \dots$$

En este caso: $\zeta^i = \langle \hat{i}_i |$ entonces $\zeta^1 = (1, 0, 0)$, $\zeta^2 = (0, 1, 0)$, $\zeta^3 = (0, 0, 1)$ y además,

$$\zeta^1 [|v\rangle] = (1, 0, 0) \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \xi^1, \quad \zeta^2 [|v\rangle] = (0, 1, 0) \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \xi^2, \quad \zeta^3 [|v\rangle] = (0, 0, 1) \begin{pmatrix} \xi^1 \\ \xi^2 \\ \xi^3 \end{pmatrix} = \xi^3.$$

△

Ejemplo 3.3 Encontremos la base dual para el espacio vectorial $V = \mathbb{R}^3$, con base ortogonal:

$$|e_1\rangle = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, |e_2\rangle = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, |e_3\rangle = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}.$$

Todo vector de ese espacio queda representado en esa base por:

$$|v\rangle = v^i |e_i\rangle = v^1 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + v^2 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} + v^3 \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} v^1 + 2v^2 \\ v^1 - v^2 - v^3 \\ -v^1 + v^2 - v^3 \end{pmatrix}.$$

Sea un funcional lineal $\mathcal{F} \in V^*$ que representa vectores duales $\mathcal{F}[\circ] \equiv \langle F | \leftrightarrow (w_1, w_2, w_3)$ y la base en el dual: $\langle \hat{e}^i | = (a_i, b_i, c_i)$ que supondremos satisface la condición $\langle \hat{e}^i | \hat{e}_j \rangle = \delta_j^i$.

Podemos hacer los siguientes cálculos con los vectores previamente normalizados

$$\langle \hat{e}^1 | \hat{e}_1 \rangle = (a_1, b_1, c_1) \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{3}} (a_1 + b_1 - c_1) = 1,$$

$$\langle \hat{e}^1 | \hat{e}_2 \rangle = (a_1, b_1, c_1) \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} (2a_1 - b_1 + c_1) = 0,$$

$$\langle \hat{e}^1 | \hat{e}_3 \rangle = (a_1, b_1, c_1) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (-b_1 - c_1) = 0.$$

Si resolvemos este sistema de tres ecuaciones resulta: $\left\{ a_1 = \frac{1}{\sqrt{3}}, b_1 = \frac{1}{\sqrt{3}}, c_1 = -\frac{1}{\sqrt{3}} \right\}$.

Repetimos los cálculos para el segundo vector:

$$\begin{aligned}\langle \hat{e}^2 | \hat{e}_1 \rangle &= (a_2, b_2, c_2) \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{3}} (a_2 + b_2 - c_2) = 0, \\ \langle \hat{e}^2 | \hat{e}_2 \rangle &= (a_2, b_2, c_2) \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} (2a_2 - b_2 + c_2) = 1, \\ \langle \hat{e}^2 | \hat{e}_3 \rangle &= (a_2, b_2, c_2) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (-b_2 - c_2) = 0.\end{aligned}$$

Ahora la solución que resulta es: $\left\{ a_2 = \frac{\sqrt{6}}{3}, b_2 = -\frac{1}{\sqrt{6}}, c_2 = \frac{1}{\sqrt{6}} \right\}$.

Para finalizar:

$$\begin{aligned}\langle \hat{e}^3 | \hat{e}_1 \rangle &= (a_3, b_3, c_3) \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{3}} (a_3 + b_3 - c_3) = 0, \\ \langle \hat{e}^3 | \hat{e}_2 \rangle &= (a_3, b_3, c_3) \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} (2a_3 - b_3 + c_3) = 0, \\ \langle \hat{e}^3 | \hat{e}_3 \rangle &= (a_3, b_3, c_3) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} (-b_3 - c_3) = 1.\end{aligned}$$

Con: $\left\{ a_3 = 0, b_3 = -\frac{1}{\sqrt{2}}, c_3 = -\frac{1}{\sqrt{2}} \right\}$.

Por lo tanto, la base del dual es la siguiente:

$$\langle \hat{e}^1 | = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right), \quad \langle \hat{e}^2 | = \left(\frac{\sqrt{6}}{3}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right), \quad \langle \hat{e}^3 | = \left(0, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right),$$

de manera que:

$$\langle F | = w_1 \langle \hat{e}^1 | + w_2 \langle \hat{e}^2 | + w_3 \langle \hat{e}^3 |.$$



Ejemplo 3.4

Consideremos un espacio vectorial complejo $V = \mathbb{C}^3$, vale decir un espacio en el cual los vectores tengan componentes complejas y puedan ser representados por un vector columna de la forma

$$|v\rangle = \begin{pmatrix} x^1 + iy^1 \\ x^2 + iy^2 \\ x^3 + iy^3 \end{pmatrix} \equiv \begin{pmatrix} r_1 e^{\theta_1} \\ r_2 e^{\theta_2} \\ r_3 e^{\theta_3} \end{pmatrix}.$$

Claramente, dependiendo de la representación del número complejo tendremos bases diferentes. La primera la hemos denominado representación vectorial, mientras que la segunda la conocemos como representación polar.

Si utilizamos la representación vectorial pudiéramos tener una posible base de la forma:

$$|w_1\rangle = \begin{pmatrix} i \\ 0 \\ 0 \end{pmatrix}, \quad |w_2\rangle = \begin{pmatrix} 0 \\ i \\ 0 \end{pmatrix}, \quad |w_3\rangle = \begin{pmatrix} 0 \\ 0 \\ i \end{pmatrix}.$$

Dejamos al lector comprobar que ésta es una base para este espacio vectorial complejo y que cualquier vector complejo puede ser expresado como combinación lineal de esta base. Esto es, cualquier número complejo $|z\rangle$ podrá ser expresado como: $|z\rangle = z^i|w_i\rangle$.

Para calcular la base dual necesitaremos la definición de producto interno. Una posible definición es la que adelantamos en el la sección 1.6.3 en la página 67: $\langle a|b\rangle = (a_i)^* b^i$ y si utilizamos esta definición la base dual será:

$$\langle w_1| = (-i, 0, 0), \quad \langle w_2| = (0, -i, 0) \text{ y } \langle w_3| = (0, 0, -i).$$

Si ahora consideramos un espacio vectorial complejo $V = \mathbb{C}^2$

$$|v\rangle = \begin{pmatrix} \tilde{x}^1 + i\tilde{y}^1 \\ \tilde{x}^2 + i\tilde{y}^2 \end{pmatrix} \quad \text{la base sería } |\tilde{w}_1\rangle = \begin{pmatrix} i \\ 0 \end{pmatrix}, \quad |\tilde{w}_2\rangle = \begin{pmatrix} 0 \\ i \end{pmatrix}$$

Si consideramos otra definición de producto interno $\langle \tilde{a}|\tilde{b}\rangle = (a_1)^*b^1 + 2(a_2)^*b^2$ la base dual será:

$$\langle \tilde{w}_1| = (-i, 0) \text{ y } \langle \tilde{w}_2| = (0, -i/2).$$



3.1.4 Ejercicios

1. Encuentre las bases duales para los siguientes espacios vectoriales:

(a). \mathbb{R}^2 , donde: $|e_1\rangle \leftrightarrow \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ y $|e_2\rangle \leftrightarrow \begin{pmatrix} 4 \\ 1 \end{pmatrix}$.

(b). \mathbb{R}^3 , donde: $|e_1\rangle \leftrightarrow \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}$, $|e_2\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ y $|e_3\rangle \leftrightarrow \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$.

2. Si V es el espacio vectorial de todos los polinomios reales de grado $n \leq 1$, y definimos:

$$\zeta^1 [|p\rangle] = \int_0^1 p(x)dx \wedge \zeta^2 [|p\rangle] = \int_0^2 p(x)dx,$$

donde $\{\zeta^1, \zeta^2\} \in V^*$. Encuentre una base $\{|e_1\rangle, |e_2\rangle\} \in V$ que resulte ortogonal a la dual $\{\zeta^1, \zeta^2\}$.

3. Si V es el espacio vectorial de todos los polinomios reales de grado $n \leq 2$, y si además definimos

$$\zeta^1 [|p\rangle] = \int_0^1 p(x)dx = 1, \quad \zeta^2 [|p\rangle] = \frac{dp(x)}{dx} \Big|_{x=2} \wedge \zeta^3 [|p\rangle] = p(1)$$

donde $\{\zeta^1, \zeta^2, \zeta^3\} \in V^*$. Encuentre una base $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\} \in V$ que resulte ortogonal a la dual $\{\zeta^1, \zeta^2, \zeta^3\}$.

4. Sean $|v_1\rangle$ y $|v_2\rangle \in V$ y supongamos que $\mathcal{F}[|v_1\rangle] = 0$ implica que $\mathcal{F}[|v_2\rangle] = 0 \forall \mathcal{F} \in V^*$. Muestre que $|v_2\rangle = \alpha|v_1\rangle$ con $\alpha \in \mathbf{K}$.

5. Sean \mathcal{F}_1 y $\mathcal{F}_2 \in V^*$ y supongamos que $\mathcal{F}_1[|v\rangle] = 0$ implica que $\mathcal{F}_2[|v\rangle] = 0 \forall |v\rangle \in V$. Muestre que $\mathcal{F}_2 = \alpha\mathcal{F}_1$ con $\alpha \in \mathbf{K}$.

6. En el caso 3D tenemos que si $\{\mathbf{e}_i\}$ define un sistema de coordenadas (dextrógiro) no necesariamente ortogonal, entonces, demuestre que:

(a).

$$\mathbf{e}^i = \frac{\mathbf{e}_j \times \mathbf{e}_k}{\mathbf{e}_i \cdot (\mathbf{e}_j \times \mathbf{e}_k)}, \quad i, j, k = 1, 2, 3 \text{ y sus permutaciones cíclicas}$$

(b). si los volúmenes: $V = \mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)$ y $\tilde{V} = \mathbf{e}^1 \cdot (\mathbf{e}^2 \times \mathbf{e}^3)$, entonces $V\tilde{V} = 1$.

(c). ¿Qué vector satisface $\mathbf{a} \cdot \mathbf{e}^i = 1$? Demuestre que \mathbf{a} es único.

(d). Encuentre el producto vectorial de dos vectores \mathbf{a} y \mathbf{b} que están representados en un sistema de

coordenadas oblicuo: Dada la base: $\mathbf{w}_1 = 4\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + \hat{\mathbf{k}}$, $\mathbf{w}_2 = 3\hat{\mathbf{i}} + 3\hat{\mathbf{j}}$, $\mathbf{w}_3 = 2\hat{\mathbf{k}}$. Entonces encuentre:

I. Las bases recíprocas $\{\mathbf{e}^i\}$.

II. Las componentes covariantes y contravariantes del vector $\mathbf{a} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$.

7. Considere una vez más el espacio vectorial de matrices hermíticas 2×2 y la definición de producto interno $\langle a | b \rangle \equiv \text{Tr}(A^\dagger B)$ que introdujimos en los ejercicios de la sección 2.2.4. Hemos comprobado que las matrices de Pauli $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$ –presentadas también en los ejercicios de la sección 2.2.4– forman base para ese espacio (ver ejercicios sección 2.3.6). Encuentre entonces la base dual asociada a la base de Pauli y, adicionalmente dado un vector genérico en este espacio vectorial, encuentre también su 1-forma asociada.

Practicando con Maxima

1. Tenemos para el espacio vectorial $V = \mathbb{R}^3$, la base ortogonal:

$$|e_1\rangle = (1, 1, -1), |e_2\rangle = (2, -1, 1), |e_3\rangle = (0, -1, -1).$$

Con $\mathcal{F} \in V^*$ donde $\mathcal{F}[o] \equiv \langle F | \leftrightarrow (w_1, w_2, w_3)$ y $\langle \hat{e}^i | = (a_i, b_i, c_i)$, con $\langle \hat{e}^i | \hat{e}_j \rangle = \delta_j^i$.

Comencemos introduciendo los vectores en las filas de matriz, pero los vectores ya normalizados.

(%ci1) E:matrix([1,1,-1]/sqrt(3),[2,-1,1]/sqrt(6),[0,-1,-1]/sqrt(2));

$$(\%o1) \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

De manera que podemos construir una nueva matriz pero donde cada fila contenga cada una de las ecuaciones que queremos resolver

(%ci2) M1:E.[a1,b1,c1];

$$(\%o2) \begin{pmatrix} -\frac{c1}{\sqrt{3}} + \frac{b1}{\sqrt{3}} + \frac{a1}{\sqrt{3}} \\ \frac{c1}{\sqrt{6}} - \frac{b1}{\sqrt{6}} + \frac{2a1}{\sqrt{6}} \\ -\frac{c1}{\sqrt{2}} - \frac{b1}{\sqrt{2}} \end{pmatrix}$$

Las ecuaciones son:

(%ci3) ec1:M1[1,1]=1;ec2:M1[2,1]=0;ec3:M1[3,1]=0;

$$(\%o3) -\frac{c1}{\sqrt{3}} + \frac{b1}{\sqrt{3}} + \frac{a1}{\sqrt{3}} = 1$$

$$(\%o4) \frac{c1}{\sqrt{6}} - \frac{b1}{\sqrt{6}} + \frac{2a1}{\sqrt{6}} = 0$$

$$(\%o5) -\frac{c1}{\sqrt{2}} - \frac{b1}{\sqrt{2}} = 0$$

Resolvemos el sistema de ecuaciones

(%ci6) linsolve([ec1,ec2,ec3],[a1,b1,c1]);

$$(\%o6) \left[a1 = \frac{1}{\sqrt{3}}, b1 = \frac{1}{\sqrt{3}}, c1 = -\frac{1}{\sqrt{3}} \right]$$

Debemos repetir los pasos anteriores dos veces más.

(%i7) M2:E.[a2,b2,c2];
(%o7)
$$\begin{pmatrix} -\frac{c_2}{\sqrt{3}} + \frac{b_2}{\sqrt{3}} + \frac{a_2}{\sqrt{3}} \\ \frac{c_2}{\sqrt{6}} - \frac{b_2}{\sqrt{6}} + \frac{2a_2}{\sqrt{6}} \\ -\frac{c_2}{\sqrt{2}} - \frac{b_2}{\sqrt{2}} \end{pmatrix}$$

(%i8) ec1:M2[1,1]=0;ec2:M2[2,1]=1;ec3:M2[3,1]=0;

(%o8) $-\frac{c_2}{\sqrt{3}} + \frac{b_2}{\sqrt{3}} + \frac{a_2}{\sqrt{3}} = 0$

(%o9) $\frac{c_2}{\sqrt{6}} - \frac{b_2}{\sqrt{6}} + \frac{2a_2}{\sqrt{6}} = 1$

(%o10) $-\frac{c_2}{\sqrt{2}} - \frac{b_2}{\sqrt{2}} = 0$

(%i11) linsolve([ec1,ec2,ec3],[a2,b2,c2]);

(%o11)
$$\left[a_2 = \frac{\sqrt{6}}{3}, b_2 = -\frac{1}{\sqrt{6}}, c_2 = \frac{1}{\sqrt{6}} \right]$$

Por último:

(%i12) M3:E.[a3,b3,c3];

(%o12)
$$\begin{pmatrix} -\frac{c_3}{\sqrt{3}} + \frac{b_3}{\sqrt{3}} + \frac{a_3}{\sqrt{3}} \\ \frac{c_3}{\sqrt{6}} - \frac{b_3}{\sqrt{6}} + \frac{2a_3}{\sqrt{6}} \\ -\frac{c_3}{\sqrt{2}} - \frac{b_3}{\sqrt{2}} \end{pmatrix}$$

(%i13) ec1:M3[1,1]=0;ec2:M3[2,1]=1;ec3:M3[3,1]=1;

(%o13) $-\frac{c_3}{\sqrt{3}} + \frac{b_3}{\sqrt{3}} + \frac{a_3}{\sqrt{3}} = 0$

(%o14) $\frac{c_3}{\sqrt{6}} - \frac{b_3}{\sqrt{6}} + \frac{2a_3}{\sqrt{6}} = 0$

(%o15) $-\frac{c_3}{\sqrt{2}} - \frac{b_3}{\sqrt{2}} = 1$

(%i16) linsolve([ec1,ec2,ec3],[a3,b3,c3]);

(%o16)
$$\left[a_3 = 0, b_3 = -\frac{1}{\sqrt{2}}, c_3 = -\frac{1}{\sqrt{2}} \right]$$

Por lo tanto, la base dual es:

(%i17) d1:[1/sqrt(3),1/sqrt(3),-1/sqrt(3)]; d2:[sqrt(6)/3,-1/sqrt(6),1/sqrt(6)]; d3:[0,-1/sqrt(2),-1/sqrt(2)];

(%o17)
$$\left[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right]$$

(%o18)
$$\left[\frac{\sqrt{6}}{3}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right]$$

(%o19)
$$\left[0, -\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right]$$

Que es una base ortonormal, como debería ser

(%i20) ratsimp(d1.d2);d1.d3;d3.d2;

(%o20) 0

(%o21) 0

(%o22) 0

(%i23) d1.d1;d2.d2;d3.d3;

(%o23) 1

(%o24) 1

(%o25) 1

(%i26) kill(all)\$

2. En este ejercicio veremos la manera de construir la matriz de transformación entre bases y el cálculo de las bases recíprocas.

Si tenemos las siguientes bases:

$$|w_1\rangle = |j\rangle + |k\rangle, \quad |w_2\rangle = |i\rangle + |k\rangle, \quad |w_3\rangle = |i\rangle + |j\rangle.$$

Para calcular la matriz de transformación:

$$|w_i\rangle = A_i^{j'} |i_{j'}\rangle,$$

podemos trabajar de la manera siguiente. Primero introducimos los vectores como una matriz y luego calculamos la transpuesta:

(%i1) w1:[0,1,1]\$ w2:[1,0,1]\$ w3:[1,1,0]\$

(%i2) Aij: transpose(matrix(v1,v2,v3));

$$(\%o2) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

La matriz de transformación inversa $|i_{i'}\rangle = A_{i'}^j |w_j\rangle$, es simplemente la matriz inversa:

(%i3) Aji: invert(%);

$$(\%o3) \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

Es claro que $A_{k'}^j A_i^{k'} = \delta_i^j$

(%i4) Aij.Aji;

$$(\%o4) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(%i5) kill(all)\$

3. Con el uso de la librería **vect** podemos hacer algunos cálculos sencillos con vectores, como por ejemplo, el cálculo de las bases recíprocas.

(%i1) load(vect)\$

Dado el siguiente conjunto de vectores:

$$b1 = e_1 = \hat{i} + \hat{j} + 2\hat{k}, \quad b2 = e_2 = -\hat{i} - \hat{j} - \hat{k}, \quad b3 = e_3 = 2\hat{i} - 2\hat{j} + \hat{k}$$

Calcularemos la base recíproca a través de:

$$\mathbf{e}^i = \frac{\mathbf{e}_j \times \mathbf{e}_k}{\mathbf{e}_i \cdot (\mathbf{e}_j \times \mathbf{e}_k)},$$

Primero que todo, debemos introducir los vectores

(%o2) $b1:[1,1,2]; b2:[-1,-1,-1]; b3:[2,-2,1];$
 (%o2) $[1, 1, 2]$
 (%o3) $[-1, -1, -1]$
 (%o4) $[2, -2, 1]$

Podemos comprobar si la base original b_i es ortogonal calculando sus productos escalares:

(%o5) $\{b1.b2, b1.b3, b2.b3\};$
 (%o5) $\{-4, -1, 2\}$

Por lo tanto, no son ortogonales. Ahora, los vectores recíprocos e^i se calculan de la manera siguiente:

(%o6) $e1:express(b2 \sim b3)/(b3.(express(b1 \sim b2)));$
 (%o6) $\left[-\frac{3}{4}, -\frac{1}{4}, 1 \right]$
 (%o7) $e2:express(b3 \sim b1)/(b3.(express(b1 \sim b2)));$
 (%o7) $\left[-\frac{5}{4}, -\frac{3}{4}, 1 \right]$
 (%o8) $e3:express(b1 \sim b2)/(b3.(express(b1 \sim b2)));$
 (%o8) $\left[\frac{1}{4}, -\frac{1}{4}, 0 \right]$

La base recíproca es entonces:

$$e^1 = -\frac{3}{4}\hat{i} - \frac{1}{4}\hat{j} + \hat{k}, \quad e^2 = -\frac{5}{4}\hat{i} - \frac{3}{4}\hat{j} + \hat{k}, \quad e^3 = \frac{1}{4}\hat{i} - \frac{1}{4}\hat{j}$$

Que tampoco es ortogonal:

(%o9) $\{e1.e2, e1.e3, e2.e3\};$
 (%o9) $\left\{ \frac{17}{8}, -\frac{1}{8}, -\frac{1}{8} \right\}$

Como veremos más adelante, podemos construir un objeto con dos índices para la base original: $g_{ij} = e_i \cdot e_j$ de la manera siguiente:

(%o10) $gb:matrix([b1.b1,b1.b2,b1.b3],[b2.b1,b2.b2,b2.b3],[b3.b1,b3.b2,b3.b3]);$
 (%o10) $\begin{pmatrix} 6 & -4 & 2 \\ -4 & 3 & -1 \\ 2 & -1 & 9 \end{pmatrix}$

Para la base recíproca: $g^{ij} = e^i \cdot e^j$

(%o11) $ge:matrix([e1.e1,e1.e2,e1.e3],[e2.e1,e2.e2,e2.e3],[e3.e1,e3.e2,e3.e3]);$
 (%o11) $\begin{pmatrix} \frac{13}{8} & \frac{17}{8} & -\frac{1}{8} \\ \frac{17}{8} & \frac{25}{8} & -\frac{1}{8} \\ -\frac{1}{8} & -\frac{1}{8} & \frac{1}{8} \end{pmatrix}$

De manera que: $e^i \cdot e_j = g_j^i = \delta_j^i$:

(%o12) $gb.ge;$
 (%o12) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

3.2 Tensores y producto tensorial

Las funciones más simples que se pueden definir sobre un espacio vectorial son los funcionales lineales y éstos nos permiten extenderlos al concepto de tensor. Para llegar a la noción de tensores ampliaremos la idea de funcionales lineales, que actúan sobre un único vector, al de funcionales bilineales (o formas bilineales) que tienen dos vectores en su argumento. Como veremos más adelante este tipo de funcionales nos revelarán un contenido geométrico de gran riqueza. Es conocido que presentar el concepto de tensores para estudiantes de pregrado o, estudiantes graduados de ingeniería tiene algunas barreras² y, la principal quizás sea la notación. En este capítulo presentaremos la notación que mas nos sedujo desde las épocas de estudiantes, es tomada de dos libros clásicos³, que han marcado una huella en nuestra generación y creemos permite su aplicación a variados campos.

3.2.1 Tensores, una definición funcional

Definiremos como un tensor a un funcional lineal (bilineal en este caso) que asocia un elemento del campo K , complejo o real, a un vector $|v\rangle \in V$, a una forma $\langle u| \in V^*$, o ambas, y cumple con la linealidad. Esto es:

$$\forall |v\rangle \in V \quad \wedge \quad \langle u| \in V^* \rightarrow \mathcal{T}[\langle u|; |v\rangle] \in \mathbb{C}.$$

Se debe cumplir:

1. $\mathcal{T}[\langle u|; \alpha|v_1\rangle + \beta|v_2\rangle] \equiv \alpha\mathcal{T}[\langle u|; |v_1\rangle] + \beta\mathcal{T}[\langle u|; |v_2\rangle] \quad \forall |v_1\rangle, |v_2\rangle \in V \quad \wedge \quad \langle u| \in V^*$.
2. $\mathcal{T}[\mu\langle u_1| + \nu\langle u_2|; |v\rangle] \equiv \mu^*\mathcal{T}[\langle u_1|; |v\rangle] + \nu^*\mathcal{T}[\langle u_2|; |v\rangle] \quad \forall |v\rangle \in V \quad \wedge \quad \langle u_1|, \langle u_2| \in V^*$.

En pocas palabras: un tensor es un funcional generalizado cuyos argumentos son vectores y/o formas⁴, lo que significa que $\mathcal{T}[\circ; \bullet]$ es una cantidad con dos “puestos” y una vez “cubiertos” se convierte en un número complejo o real.

Las combinaciones son muy variadas:

- Un tensor, con un argumento correspondiente a un vector y un argumento correspondiente a una forma, lo podremos representar de la siguiente manera:

$$\mathcal{T} \left[\begin{smallmatrix} |v\rangle & \langle u| \\ \downarrow & \downarrow \\ \circ & \bullet \end{smallmatrix} \right] \in \mathbb{C} \quad \Rightarrow \quad \text{tensor de tipo } \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

- Un tensor con dos argumentos correspondientes a vectores y uno a una forma sería:

$$\mathcal{T}[\circ, \circ; \bullet] \Rightarrow \mathcal{T} \left[\begin{smallmatrix} |v_1\rangle & |v_2\rangle & \langle u| \\ \downarrow & \downarrow & \downarrow \\ \circ & \circ & \bullet \end{smallmatrix} \right] \in \mathbb{C} \quad \Rightarrow \quad \text{tensor de tipo } \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

- Un tensor con dos argumentos correspondientes a formas y uno a un vector:

$$\mathcal{T}[\circ; \bullet, \bullet] \Rightarrow \mathcal{T} \left[\begin{smallmatrix} |v\rangle & \langle u_1| & \langle u_2| \\ \downarrow & \downarrow & \downarrow \\ \circ & \bullet & \bullet \end{smallmatrix} \right] \in \mathbb{C} \quad \Rightarrow \quad \text{tensor de tipo } \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

²Tal y como lo plantea Battaglia, Franco, y Thomas F. George. "Tensors: A guide for undergraduate students." *American Journal of Physics* **81**(2013): 498-51.

³Misner, C. W., K. S. Thorne, and J. A. Wheeler. Gravitation. Princeton University Press, 2017; y Cohen-Tannoudji, C., Diu, B. and Laloë, F., 1977. Quantum Mechanics vol 1 (Paris: Hermann).

⁴Una presentación interesante y detallada de la utilización del concepto de tensor para el manejo de datos en computación la pueden encontrar en: Lu, H., Plataniotis, K. N., & Venetsanopoulos, A. N. (2011). A survey of multilinear subspace learning for tensor data. *Pattern Recognition*, 44 (7), 1540-1551.

- En general:

$$\mathcal{T} \left[\begin{smallmatrix} |v_1\rangle & |v_2\rangle & & |v_n\rangle & \langle u_1| & \langle u_2| & & \langle u_m| \\ \downarrow \circ & \downarrow \circ & \cdots & \downarrow \circ & \downarrow \bullet & \cdots & \downarrow \bullet & \end{smallmatrix} \right] \in \mathbb{C} \quad \Rightarrow \quad \text{tensor de tipo } \binom{m}{n}.$$

En esta notación el punto y coma (;) separa las “entradas” formas de las “entradas” vectores. Es importante recalcar que **el orden si importa**, no sólo para las cantidades separadas por el punto y coma, sino el orden de los “puestos” vectores y “puestos” formas separados por coma, y repercutirá en las propiedades de los tensores. Por ejemplo: **si el orden de las “entradas” vectores no importa**, podremos permutarlas sin alterar al tensor, **tendremos entonces tensores simétricos respecto a esos “puestos” o “entradas”**; del mismo modo, serán tensores antisimétricos aquellos en los cuales **el orden si importa** y al permutar esos “puestos” o “entradas” hay un cambio de signo en el tensor. Todos estos casos serán tratados con detalle más adelante, pero vale la pena recalcar que en general, para un tensor genérico el orden de la “entradas” o “puestos” si importa pero no necesariamente se comporta como los casos reseñados anteriormente.

Notemos que en el caso general un tensor es básicamente un aplicación multilinear \mathcal{T} sobre $\mathbf{V}^* \times \mathbf{V}$:

$$\mathcal{T} : \mathbf{V}^{*m} \times \mathbf{V}^n = \underbrace{\mathbf{V}^* \times \mathbf{V}^* \cdots \mathbf{V}^* \times \mathbf{V}^*}_{m} \times \underbrace{\mathbf{V} \times \mathbf{V} \cdots \mathbf{V} \times \mathbf{V}}_{n} \Rightarrow \mathbb{C},$$

con n el orden covariante y m el orden contravariante. Hay que hacer notar que por simplicidad hemos construido el espacio tensorial a partir de una solo espacio vectorial \mathbf{V} y su conjugado \mathbf{V}^* , pero muy bien cada espacio vectorial puede ser diferente.

Por lo tanto, **un tensor** $\binom{m}{n}$ es un funcional multilinear que asocia m 1-formas y n vectores con \mathbb{C} .

Un ejemplo sencillo de un funcional bilineal sobre un espacio vectorial \mathbf{V}^n con una base $|e_i\rangle$ es:

$$\mathcal{T}[|v_1\rangle, |v_2\rangle] = a_{ij} \xi^i \zeta^j \quad (i, j = 1, 2, \dots, n),$$

donde:

$$|v_1\rangle = \xi^i |e_i\rangle \wedge |v_2\rangle = \zeta^i |e_i\rangle.$$

son vectores arbitrarios $\in \mathbf{V}^n$ y los a_{ij} son números. Notemos que:

$$\mathcal{T}[|v_1\rangle, |v_2\rangle] = \mathcal{T}[\xi^i |e_i\rangle, \zeta^j |e_j\rangle] = \xi^i \zeta^j \mathcal{T}[|e_i\rangle, |e_j\rangle] = \xi^i \zeta^j a_{ij}.$$

Diremos que ésta será la representación del funcional bilineal para \mathbf{V}^n .

Obviamente las formas pueden ser representadas por tensores ya que son funcionales lineales de vectores. Para finalizar, notemos lo siguiente:

- **Un vector** es un tensor del tipo:

$$\binom{1}{0} \Rightarrow \mathcal{T} \left[\begin{smallmatrix} \langle u| \\ \downarrow \bullet \end{smallmatrix} \right] \in \mathbb{C}.$$

Los vectores constituyen un caso especial de tensores.

- **Una forma** es un tensor del tipo:

$$\binom{0}{1} \Rightarrow \mathcal{T} \left[\begin{smallmatrix} |v\rangle \\ \downarrow \circ \end{smallmatrix} \right] \in \mathbb{C},$$

porque son funcionales lineales para las formas diferenciales.

- **Un escalar** es un tensor del tipo:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathbb{C}.$$

3.2.2 Producto tensorial

Los tensores (*simples*) pueden provenir del *producto tensorial* (exterior o directo) de espacios vectoriales. Esto es, consideraremos \mathbf{E}_1 y \mathbf{E}_2 dos espacios vectoriales con dimensiones n_1 y n_2 , respectivamente y vectores genéricos, $|\varphi(1)\rangle$ y $|\chi(2)\rangle$ pertenecientes a espacios vectoriales: $|\varphi(1)\rangle \in \mathbf{E}_1$ y $|\chi(2)\rangle \in \mathbf{E}_2$.

Definiremos el *producto tensorial* (*exterior o directo*) de espacios vectoriales, $\mathbf{E} = \mathbf{E}_1 \otimes \mathbf{E}_2$, si a cada par de vectores $|\varphi(1)\rangle$ y $|\chi(2)\rangle$ le asociamos un tensor tipo $\begin{pmatrix} 2 \\ 0 \end{pmatrix}$ y si se cumple que:

$$|\varphi(1)\chi(2)\rangle \equiv |\varphi(1)\rangle \otimes |\chi(2)\rangle \Leftrightarrow \mathcal{T} \begin{bmatrix} \langle \psi(1)| & \langle v(2)| \\ \downarrow & \downarrow \end{bmatrix} = \langle \psi(1)|\varphi(1)\rangle \langle v(2)|\chi(2)\rangle \in \mathbb{C},$$

y si además se cumplen las siguientes propiedades:

1. La suma entre tensores de \mathbf{E} viene definida como:

$$|\varphi(1)\chi(2)\rangle + |\zeta(1)\xi(2)\rangle = |\varphi(1)\rangle \otimes |\chi(2)\rangle + |\zeta(1)\rangle \otimes |\xi(2)\rangle = |\varphi(1) + \zeta(1)\rangle \otimes |\chi(2) + \xi(2)\rangle.$$

2. El producto tensorial es lineal respecto a la multiplicación con números reales λ y μ :

$$[|\lambda\varphi(1)\rangle] \otimes |\chi(2)\rangle = [\lambda|\varphi(1)\rangle] \otimes |\chi(2)\rangle = \lambda [|\varphi(1)\rangle \otimes |\chi(2)\rangle] = \lambda|\varphi(1)\chi(2)\rangle,$$

$$|\varphi(1)\rangle \otimes [|\mu\chi(2)\rangle] = |\varphi(1)\rangle \otimes [\mu|\chi(2)\rangle] = \mu [|\varphi(1)\rangle \otimes |\chi(2)\rangle] = \mu|\varphi(1)\chi(2)\rangle.$$

3. El producto tensorial es distributivo respecto a la suma:

$$|\varphi(1)\rangle \otimes [|\chi_1(2)\rangle + |\chi_2(2)\rangle] = |\varphi(1)\rangle \otimes |\chi_1(2)\rangle + |\varphi(1)\rangle \otimes |\chi_2(2)\rangle$$

$$[|\varphi_1(1)\rangle + |\varphi_2(1)\rangle] \otimes |\chi(2)\rangle = |\varphi_1(1)\rangle \otimes |\chi(2)\rangle + |\varphi_2(1)\rangle \otimes |\chi(2)\rangle.$$

Nótese que las etiquetas (1) y (2) denotan la pertenencia al espacio respectivo.

3.2.3 Espacios tensoriales

Es fácil convencerse que los tensores $|\varphi(1)\chi(2)\rangle \in \mathbf{E} = \mathbf{E}_1 \otimes \mathbf{E}_2$ forman un espacio vectorial y la demostración se basa en comprobar los axiomas o propiedades de los espacios vectoriales tal y como lo describimos en la sección 2.1.3:

1. La operación suma \boxplus es cerrada en \mathbf{V} : $\forall |v_i\rangle, |v_j\rangle \in \mathbf{V} \Rightarrow |v_k\rangle = |v_i\rangle \boxplus |v_j\rangle \in \mathbf{V}$.

Esto se traduce en demostrar que sumados dos tensores $|\varphi(1)\chi(2)\rangle$ y $|\zeta(1)\xi(2)\rangle \in \mathbf{E}$ el tensor suma también pertenece a \mathbf{E} , con α y β pertenecientes al campo del espacio vectorial

$$\alpha|\varphi(1)\chi(2)\rangle + \beta|\zeta(1)\xi(2)\rangle = |\alpha\varphi(1) + \zeta(1)\rangle \otimes |\chi(2) + \xi(2)\rangle,$$

y esto se cumple siempre ya que, el producto tensorial es lineal respecto a la multiplicación con números reales, y por ser \mathbf{E}_1 y \mathbf{E}_2 espacios vectoriales se cumple:

$$\left. \begin{array}{l} |\alpha\varphi(1) + \zeta(1)\rangle = \alpha|\varphi(1)\rangle + |\zeta(1)\rangle \in \mathbf{E}_1 \\ |\varphi(2) + \xi(2)\rangle = |\varphi(2)\rangle + \xi|\xi(2)\rangle \in \mathbf{E}_2 \end{array} \right\} \Rightarrow |\varphi(1) + \zeta(1)\rangle \otimes |\chi(2) + \xi(2)\rangle \in \mathbf{E}.$$

2. La operación suma \boxplus es commutativa y asociativa.

Commutativa: $\forall |v_i\rangle, |v_j\rangle \in \mathbf{V} \Rightarrow |v_i\rangle \boxplus |v_j\rangle = |v_j\rangle \boxplus |v_i\rangle$.

Esta primera es clara de la definición de suma:

$$\begin{aligned} |\varphi(1)\chi(2)\rangle + |\zeta(1)\xi(2)\rangle &= |\varphi(1) + \zeta(1)\rangle \otimes |\chi(2) + \xi(2)\rangle \\ |\zeta(1)\xi(2)\rangle + |\varphi(1)\chi(2)\rangle &= |\zeta(1) + \varphi(1)\rangle \otimes |\xi(2) + \chi(2)\rangle, \end{aligned}$$

por ser \mathbf{E}_1 y \mathbf{E}_2 dos espacios vectoriales.

Asociativa: $\forall |v_i\rangle, |v_j\rangle, |v_k\rangle \in \mathbf{V} \Rightarrow (|v_i\rangle \boxplus |v_j\rangle) \boxplus |v_k\rangle = |v_j\rangle \boxplus (|v_i\rangle \boxplus |v_k\rangle)$

Una vez más, esto se traduce en:

$$(|\varphi(1)\chi(2)\rangle + |\zeta(1)\xi(2)\rangle) + |\varkappa(1)\kappa(2)\rangle = |\varphi(1)\chi(2)\rangle + (|\zeta(1)\xi(2)\rangle + |\varkappa(1)\kappa(2)\rangle),$$

con lo cual, por la definición de suma, la expresión anterior queda como:

$$\begin{aligned} (|\varphi(1) + \zeta(1)\rangle \otimes |\xi(2) + \chi(2)\rangle) + |\varkappa(1)\kappa(2)\rangle &= |\varphi(1)\chi(2)\rangle + (|\zeta(1) + \varkappa(1)\rangle \otimes |\xi(2) + \kappa(2)\rangle) \\ (|\varphi(1) + \zeta(1)\rangle + |\varkappa(1)\kappa(2)\rangle) \otimes (|\xi(2) + \chi(2)\rangle + |\kappa(2)\rangle) &= |\varphi(1) + (\zeta(1) + \varkappa(1))\rangle \otimes |\xi(2) + (\chi(2) + \kappa(2))\rangle. \end{aligned}$$

3. Existe un único elemento neutro: $\exists |0\rangle / |0\rangle \boxplus |v_j\rangle = |v_j\rangle \boxplus |0\rangle = |v_j\rangle \forall |v_j\rangle \in \mathbf{V}$.

Es decir:

$$|\varphi(1)\chi(2)\rangle + |0(1)0(2)\rangle = |\varphi(1) + 0(1)\rangle \otimes |\chi(2) + 0(2)\rangle = |\varphi(1)\rangle \otimes |\chi(2)\rangle = |\varphi(1)\chi(2)\rangle.$$

4. Existe un elemento simétrico para cada elemento de \mathbf{V} : $\forall |v_j\rangle \in V \exists |-v_j\rangle / |v_j\rangle \boxplus |-v_j\rangle = |0\rangle$.

$$|\varphi(1)\chi(2)\rangle - |\varphi(1)\chi(2)\rangle = |\varphi(1) - \varphi(1)\rangle \otimes |\chi(2) - \chi(2)\rangle = |0(1)\rangle \otimes |0(2)\rangle = |0(1)0(2)\rangle.$$

5. $\alpha(\beta|v_i\rangle) = (\alpha\beta)|v_i\rangle$:

$$\alpha(\beta|\varphi(1)\chi(2)\rangle) = \alpha(|\beta\chi(2)\rangle \otimes |\varphi(1)\rangle) = |\alpha\beta\chi(2)\rangle \otimes |\varphi(1)\rangle = (\alpha\beta)|\chi(2)\rangle \otimes |\varphi(1)\rangle = (\alpha\beta)|\varphi(1)\chi(2)\rangle.$$

6. $(\alpha + \beta)|v_i\rangle = \alpha|v_i\rangle + \beta|v_i\rangle$:

$$\begin{aligned} (\alpha + \beta)|\varphi(1)\chi(2)\rangle &= |\varphi(1)\rangle \otimes |(\alpha + \beta)\chi(2)\rangle = |\varphi(1)\rangle \otimes |\alpha\chi(2) + \beta\chi(2)\rangle \\ &= |\varphi(1)\rangle \otimes [(\alpha|\chi(2)\rangle + \beta|\chi(2)\rangle)] \\ &= \alpha|\varphi(1)\rangle \otimes |\chi(2)\rangle + \beta|\varphi(1)\rangle \otimes |\chi(2)\rangle. \end{aligned}$$

7. $\alpha(|v_i\rangle \boxplus |v_j\rangle) = \alpha|v_i\rangle \boxplus \alpha|v_j\rangle$:

$$\begin{aligned} \alpha(|\varphi(1)\chi(2)\rangle + |\zeta(1)\xi(2)\rangle) &= \alpha(|\varphi(1) + \zeta(1)\rangle \otimes |\xi(2) + \chi(2)\rangle) = |\alpha(\varphi(1) + \zeta(1))\rangle \otimes |\xi(2) + \chi(2)\rangle \\ &= |\alpha\varphi(1) + \alpha\zeta(1)\rangle \otimes |\xi(2) + \chi(2)\rangle = |\alpha\varphi(1)\chi(2)\rangle + |\alpha\zeta(1)\xi(2)\rangle \\ &= \alpha|\varphi(1)\chi(2)\rangle + \alpha|\zeta(1)\xi(2)\rangle. \end{aligned}$$

Equivalentemente, podemos construir un producto tensorial entre espacios de formas diferenciales. Si \mathbf{E}_1^* y \mathbf{E}_2^* son dos espacios vectoriales duales a \mathbf{E}_1 y \mathbf{E}_2 , con dimensiones n_1 y n_2 , respectivamente. A estos espacios pertenecen las formas diferenciales genéricas $\langle\zeta(1)| \in \mathbf{E}_1^*$ y $\langle\xi(2)| \in \mathbf{E}_2^*$.

Definiremos el producto tensorial de espacios vectoriales duales, $\mathbf{E}^* = \mathbf{E}_1^* \otimes \mathbf{E}_2^*$, si a cada par de formas diferenciales $\langle\zeta(1)| \in \mathbf{E}_1^*$ y $\langle\xi(2)| \in \mathbf{E}_2^*$ le asociamos un tensor tipo $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$. Esto es:

$$\langle\zeta(1)\xi(2)| = \langle\zeta(1)| \otimes \langle\xi(2)|.$$

Es importante aclarar que el número $\langle\tilde{\varphi}(1)\tilde{\chi}(2)|\varphi(1)\chi(2)\rangle = \langle\tilde{\varphi}(1)|\varphi(1)\rangle \cdot \langle\tilde{\chi}(2)|\chi(2)\rangle$, NO representa el producto interno entre los tensores, $\langle\tilde{\varphi}(1)\tilde{\chi}(2)|$ y $|\varphi(1)\chi(2)\rangle$. Tal y como hemos descrito arriba,

$$\mathcal{T}[|\varphi(1)\rangle, |\chi(2)\rangle] \equiv \mathcal{T}^*[\langle\tilde{\varphi}(1)|, \langle\tilde{\chi}(2)|] = \langle\tilde{\varphi}(1)|\varphi(1)\rangle \cdot \langle\tilde{\chi}(2)|\chi(2)\rangle,$$

representa las evaluaciones de los funcionales $\mathcal{T}[\circ]$ y $\mathcal{T}^*[\bullet]$, respectivamente.

3.2.4 Bases para un producto tensorial

Si $\{|u_i(1)\rangle\}$ y $\{|v_i(2)\rangle\}$ son, respectivamente, bases discretas para E_1 y E_2 entonces podremos construir el tensor:

$$|u_i(1)v_j(2)\rangle = |u_i(1)\rangle \otimes |v_j(2)\rangle \in E,$$

el cual funcionará como una base para E

Podremos construir un tensor genérico de E :

$$|\varphi(1)\chi(2)\rangle = |\varphi(1)\rangle \otimes |\chi(2)\rangle = \varphi^i \chi^j |u_i(1)v_j(2)\rangle,$$

donde φ^i y χ^j son las componentes de $|\varphi(1)\rangle$ y $|\chi(2)\rangle$ en sus respectivas bases. En otras palabras, las componentes de un tensor en E corresponden a la multiplicación de las componentes de los vectores en E_1 y E_2 . Recuerde que estamos utilizando la convención de Einstein de suma tácita en índices covariantes y contravariantes, en la cual $c^k|v_k\rangle \equiv \sum_{k=1}^n c^k|v_k\rangle$.

Es importante señalar que si bien un tensor genérico $|\Psi\rangle \in E$ siempre se puede expandir en la base $|u_i(1)v_j(2)\rangle$ no es cierto que todo tensor de E provenga del producto tensorial de E_1 y E_2 . Es decir, E tiene más tensores de los que provienen el producto tensorial. Esta afirmación puede intuirse del hecho que si $|\Psi\rangle \in E$ entonces:

$$|\Psi\rangle = c^{ij}|u_i(1)v_j(2)\rangle,$$

por ser $\{|u_i(1)v_j(2)\rangle\}$ base para E . Es claro que dados dos números α_1 y α_2 habrá c^{ij} que no provienen de la multiplicación de $\alpha_1\alpha_2$.

El conjunto de todas funciones bilineales $\mathcal{T}[\langle u|; |v\rangle]$ forman un espacio vectorial sobre el espacio directo E . Este espacio vectorial de funciones tendrá una base dada por:

$$|u_i(1)v_j(2)\rangle = |u_i(1)\rangle \otimes |v_j(2)\rangle.$$

3.2.5 Tensores, sus componentes y sus contracciones

Hemos mencionado anteriormente, ubicándonos en \mathbb{R}^3 , que un escalar es una cantidad que se puede especificar, independientemente del sistema de coordenadas, por un sólo número. Los vectores geométricos que dibujábamos con flechas los sustituimos ahora por tres números respecto a una base seleccionada, es decir, a través de sus tres componentes. Los escalares y los vectores son casos particulares de objetos más generales que denominamos tensores, tensores de orden o rango k y cuya especificación en cualquier sistema de coordenadas requerirá de 3^k números, llamadas las componentes del tensor. Esto significa que un escalar es un tensor de orden 0 ($3^0 = 1$ componente) y un vector un tensor de orden 1 ($3^1 = 3$ componentes). Si el espacio vectorial es de dimensión n , entonces un tensor de orden k tendrá n^k componentes⁵.

Los tensores son mucho más que simples números respecto a un sistema de coordenadas y la clave radica en la “ley de transformación” de sus componentes, es decir, en la relación que existe entre las componentes de un tensor en un sistema de coordenadas y las componentes del mismo tensor en otro sistema de coordenadas diferente. Lo que hay detrás de todo esto es el hecho que las leyes matemáticas que describen los fenómenos físicos deben ser “invariantes” bajo transformaciones de coordenadas, como por ejemplo: traslaciones (el espacio es homogéneo) y rotaciones (el espacio es isótropo).

⁵Existen varias presentaciones operativas de estos conceptos para Física e Ingeniería, pueden consultar:

- Battaglia, F., & George, T. F. (2013). Tensors: A guide for undergraduate students. American Journal of Physics, 81(7), 498-511.
- Comon, P. (2014). Tensors: a brief introduction. IEEE Signal Processing Magazine, 31(3), 44-53.

3.2.5.1 Componentes de un tensor

Consideremos un espacio tensorial $\mathbf{V} = \mathbf{E}_1^* \otimes \mathbf{E}_2^* \otimes \mathbf{E}_3^* \otimes \mathbf{E}_1 \otimes \mathbf{E}_2$. Denominaremos componentes de un tensor, aquellos números que surgen al evaluar los funcionales (formas diferenciales o vectores) con la bases de los espacios que lo constituyen. Así, si $\{|u_i(1)\rangle\}$ y $\{|x_i(1)\rangle\}$ son base para \mathbf{E}_1 , mientras $\{|v_j(2)\rangle\}$ y $\{|y_j(2)\rangle\}$ lo son para \mathbf{E}_2 y finalmente $\{|t_k(3)\rangle\}$ para \mathbf{E}_3 , entonces las componentes de un tensor $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ serán:

$$S^{mn}_{ijk} = \mathcal{S} \left[\begin{matrix} \langle x^m(1)| & \langle y^n(2)| & |u_i(1)\rangle & |v_j(2)\rangle & |w_k(3)\rangle \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \bullet & \bullet & ; & \circ & , & \circ \end{matrix} \right] \Rightarrow \mathcal{S} = S^{mn}_{ijk} [x_m(1), y_n(2); u^i(1), v^j(2), t^k(3)],$$

donde hemos denotado $[x_m(1), y_n(2); u^i(1), v^j(2), t^k(3)]$ como el tensor base

$$[x_m(1), y_n(2); u^i(1), v^j(2), t^k(3)] = |x_m(1)\rangle \otimes |y_n(2)\rangle \otimes \langle u^i(1)| \otimes \langle v^j(2)| \otimes \langle t^k(3)|.$$

Podemos comentar varias cosas:

- Por simplicidad hemos considerado un espacio tensorial de la forma $\mathbf{V} = \mathbf{E}_1^* \otimes \mathbf{E}_2^* \otimes \mathbf{E}_3^* \otimes \mathbf{E}_1 \otimes \mathbf{E}_2$, pero claramente pudimos haberlo construido mas general, $\mathbf{W} = \mathbf{E}_1^* \otimes \mathbf{E}_2^* \otimes \mathbf{E}_3^* \otimes \mathbf{E}_4 \otimes \mathbf{E}_5$.
- En este ejemplo, en la base del espacio tensorial $[x_m(1), y_n(2); u^i(1), v^j(2), t^k(3)]$, hemos colocado primero su “parte” vectorial y luego su “parte” forma. Los índices de las componentes heredan esta convención S^{mn}_{ijk} : primero irán los contravariantes y luego los covariantes. Pero lo importante es que los índices etiqueten el espacio al cual corresponde.
- Nuestra notación será:

- $|\psi(1), \phi(2)\rangle \equiv |\psi(1)\rangle \otimes |\phi(2)\rangle = C^{ij}|u_i(1), v_j(2)\rangle = \tilde{C}^{ij}|x_i(1), v_j(2)\rangle = \bar{C}^{ij}|u_i(1), y_j(2)\rangle$. Es decir un tensor perteniente a un espacio tensorial conformado por dos espacios vectoriales y que se pueden expresar en varias de las bases de esos espacios.
- $\langle\psi(1), \phi(2)| \equiv \langle\psi(1)| \otimes \langle\phi(2)| = K_{ij}\langle u^i(1), v^j(2)| = \tilde{K}_{ij}\langle x^i(1), v^j(2)| = \bar{K}_{ij}\langle u^i(1), y^j(2)|$. Igual que para el caso anterior pero a partir de 1-formas.
- $[\psi(1), \phi(2)] \equiv \langle\psi(1)| \otimes |\phi(2)\rangle = M_i^j [u^i(1); v_j(2)] = \tilde{M}_i^j [x^i(1); v_j(2)] = \bar{M}_i^j [u^i(1); y_j(2)]$.

Un tensor mixto

Es importante convenir el orden en el cual se presenten los espacios con los cuales se realiza el producto tensorial

Es de hacer notar que la selección de las bases no es arbitraria sino que deben corresponderse, entre el espacio directo, \mathbf{E} , y su dual \mathbf{E}^* , i.e.

$$\{|u_i(1)\rangle, |v_j(2)\rangle, |w_k(3)\rangle\} \otimes \{\langle x^m(1)|, \langle y^n(2)|\} \Leftrightarrow \{|x_p(1)\rangle, |y_q(2)\rangle\} \otimes \{\langle u^a(1)|, \langle v^b(2)|, \langle w^c(2)|\}.$$

Si consideramos un tensor como resultado de un producto tensorial y consideramos las bases: $\{|u_i(1)\rangle, \langle x^m(1)|\}$, sus componentes se pueden expresar $\{\varphi^m(1)\chi_i(1)\}$, vale decir:

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \iff |\varphi(1)\rangle \otimes \langle\Delta(1)| \Rightarrow \langle x^m(1)|\varphi(1)\rangle \otimes \langle\Delta(1)|u_i(1)\rangle \Rightarrow \{\varphi^m(1)\chi_i(1)\}.$$

3.2.5.2 Combinaciones lineales de tensores

Es claro que podremos sumar (componentes) de tensores como lo hemos hecho con la suma de (componentes) de vectores:

$$\mathbf{a} + \mathbf{b} = (a_x + b_x)\hat{\mathbf{i}} + (a_y + b_y)\hat{\mathbf{j}} + (a_z + b_z)\hat{\mathbf{k}} = (a^1 + b^1)\hat{\mathbf{i}} + (a^2 + b^2)\hat{\mathbf{j}} + (a^3 + b^3)\hat{\mathbf{k}} = (a^i + b^i)|\mathbf{i}_i\rangle,$$

esto es:

$$R_{kl}^{ij} = \alpha Q_{kl}^{ij} + \beta P_{kl}^{ij}.$$

3.2.5.3 Producto tensorial de tensores

Podemos extender aún más la idea del producto directo y extenderla para tensores. Así, para dos tensores, uno tipo:

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} \Rightarrow |\varphi(1)\chi(2)\rangle = |\varphi(1)\rangle \otimes |\chi(2)\rangle = \mathcal{T} \begin{bmatrix} |\zeta(1)| & |\xi(2)| \\ \downarrow \bullet & \downarrow \bullet \end{bmatrix},$$

y el otro tipo:

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow |\mu(1)\kappa(2)\Theta(1)\rangle = |\mu(1)\rangle \otimes |\kappa(2)\rangle \otimes \langle\Theta(1)| = \mathcal{P} \begin{bmatrix} |u_i(1)\rangle & |\varepsilon(1)\rangle & |\phi(2)| \\ \downarrow \circ & \downarrow \bullet & \downarrow \bullet \end{bmatrix},$$

el producto directo es:

$$\begin{aligned} |\varphi(1)\chi(2)\rangle \otimes |\mu(1)\kappa(2)\Theta(1)\rangle &= |\varphi(1)\rangle \otimes |\chi(2)\rangle \otimes |\mu(1)\rangle \otimes |\kappa(2)\rangle \otimes \langle\Theta(1)| \\ &= \mathcal{T} \begin{bmatrix} |\zeta(1)| & |\xi(2)| \\ \downarrow \bullet & \downarrow \bullet \end{bmatrix} \otimes \mathcal{P} \begin{bmatrix} |u_i(1)\rangle & |\varepsilon(1)\rangle & |\phi(2)| \\ \downarrow \circ & \downarrow \bullet & \downarrow \bullet \end{bmatrix} \\ &= \mathcal{R} \begin{bmatrix} |u_i(1)\rangle & |\varepsilon(1)\rangle & |\phi(2)| & |\zeta(1)| & |\xi(2)| \\ \downarrow \circ & \downarrow \bullet & \downarrow \bullet & \downarrow \bullet & \downarrow \bullet \end{bmatrix}. \end{aligned}$$

En componentes será como se muestra a continuación:

$$R_{k}^{ij \ l m} = T^{ij} P_k^{lm}.$$

Una vez más, note que los puestos de los índices en las componentes de los tensores heredan las posiciones de los vectores y formas que construyeron el espacio tensorial a partir del producto directo o tensorial.

3.2.5.4 Contracción de un tensor

Denominaremos una contracción cuando sumamos las componentes covariantes y contravariantes, esto es, si tenemos $\varphi^i(1)\chi_i(1)$, entonces se genera un escalar independiente de la base. La situación será más evidente cuando definamos métricas y contracción de tensores. Por analogía y considerando un caso más general, dada las componentes S_{ijk}^{mn} correspondiente a un tensor $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ podremos construir un nuevo tensor $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ a partir de una contracción. Las componentes de este nuevo tensor serán: $S_{ijk}^{mn} \Rightarrow S_{ijk}^{in} \equiv S_{jk}^{n}$.

Del mismo modo, dadas las componentes de dos tensores, P^{lm} y Q_{zk}^{ij} generarán componentes de nuevos tensores $R_k^{lij} = P^{lm}Q_{mk}^{ij}$. Así:

$$\left. \begin{array}{l} \begin{pmatrix} 2 \\ 0 \end{pmatrix} \Rightarrow P^{lm} \\ \begin{pmatrix} 2 \\ 2 \end{pmatrix} \Rightarrow Q_{zk}^{ij} \end{array} \right\} \Rightarrow \begin{pmatrix} 3 \\ 1 \end{pmatrix} \Rightarrow R_k^{lij} = P^{lm}Q_{mk}^{ij}.$$

Es claro que si dos tensores derivan de productos tensoriales y si $\{|u_i(1)\rangle\}, \{|\varepsilon^m(1)\rangle\}$ y $\{|v_i(2)\rangle\}$ son

bases ortonormales para \mathbf{E}_1 , \mathbf{E}_1^* y \mathbf{E}_2 , entonces sus productos tensoriales podrán ser expresados como:

$$|\gamma(1); \delta(2)\rangle = \underbrace{(\gamma^i(1)\delta^j(2))}_{P^{ij}} |u_i(1)\rangle \otimes |v_j(2)\rangle, \quad [\alpha(1); \beta(1)] = \underbrace{(\alpha^l(1)\beta_m(1))}_{Q_m^l} |u_l(1)\rangle \otimes \langle u^m(1)|,$$

entonces:

$$\begin{aligned} & \left((\alpha^l(1)\beta_m(1)) |u_l(1)\rangle \otimes \langle u^m(1)| \right) \left((\gamma^i(1)\delta^j(2)) |u_i(1)\rangle \otimes |v_j(2)\rangle \right) \\ &= \alpha^l(1)\beta_m(1) (\gamma^i(1)\delta^j(2)) \underbrace{\{\langle u^m(1)|u_i(1)\rangle\}}_{\delta_i^m} |v_j(2)\rangle \otimes |u_l(1)\rangle \\ &= \alpha^l(1)\beta_k(1) (\gamma^k(1)\delta^j(2)) |v_j(2)\rangle \otimes |u_l(1)\rangle \\ &= P^{ij} Q_i^l |u_l(1); v_j(2)\rangle \\ &= R^{jl} |u_l(1); v_j(2)\rangle. \end{aligned}$$

Pero más aún, si $|u_i(1)v_j(2)\rangle = |u_i(1)\rangle \otimes |v_j(2)\rangle \in \mathbf{E}$ es base de \mathbf{E} entonces se puede demostrar lo anterior sin circunscribirnos a tensores cuyas componentes provengan de multiplicación de las componentes en cada espacio vectorial.

3.2.5.5 Simetrización de tensores

Un tensor y sus componentes será simétrico respecto a dos de sus índices si su permutación no cambia su valor:

$$S_{ij} = S_{ji}, \quad S^{ij} = S^{ji}, \quad S_{ij...kl...mn} = S_{ij...lk...mn}, \quad S^{ij...kl...mn} = S^{ij...lk...mn},$$

y será antisimétrico si:

$$A_{ij} = -A_{ji}, \quad A^{ij} = -A^{ji}, \quad A_{ij...kl...mn} = -A_{ij...lk...mn}, \quad A^{ij...kl...mn} = -A^{ij...lk...mn}.$$

Un tensor de rango 2, viene representado por una matriz que tendrá $3^2 = 9$ componentes. Si la matriz es simétrica tendrá como máximo 6 componentes distintas.

$$S_j^i = S_i^j = \begin{pmatrix} S_1^1 & S_2^1 & S_3^1 \\ S_1^2 & S_2^2 & S_3^2 \\ S_1^3 & S_2^3 & S_3^3 \end{pmatrix} = \begin{pmatrix} S_1^1 & S_1^2 & S_1^3 \\ S_2^1 & S_2^2 & S_2^3 \\ S_3^1 & S_3^2 & S_3^3 \end{pmatrix}.$$

Mientras que un tensor antisimétrico de segundo orden tendrá, como máximo, tres componentes con valor absoluto distintos de cero,

$$A_j^i = -A_i^j = \begin{pmatrix} 0 & A_2^1 & A_3^1 \\ -A_1^2 & 0 & A_3^2 \\ -A_1^3 & -A_2^3 & 0 \end{pmatrix}.$$

Siempre es posible construir tensores simétricos y antisimétricos a partir de un tensor genérico. Esto es:

$$S_{ij} = \frac{1}{2} (T_{ij} + T_{ji}) \equiv T_{[ij]} \iff S_{ij...kl...mn} = \frac{1}{2} (T_{ij...kl...mn} + T_{ij...lk...mn}) = T_{ij...(kl)...mn}$$

$$A_{ij} = \frac{1}{2} (T_{ij} - T_{ji}) \equiv T_{[ij]} \iff A_{ij...kl...mn} = \frac{1}{2} (T_{ij...kl...mn} - T_{ij...lk...mn}) = T_{ij...[kl]...mn}.$$

Es evidente que las componentes de un tensor genérico, T_{ij} , pueden expresarse como una combinación de su parte simétrica y antisimétrica:

$$T_{ij} = S_{ij} + A_{ij}.$$

Obviamente que algo equivalente se puede realizar para componentes contravariantes de tensores.

Ejemplo 3.5 Productos tensoriales, vectores y matrices

Para fijar los conceptos que hemos desarrollado en las secciones 3.2.2, y 3.2.5 consideremos dos espacios vectoriales \mathbf{V}_1 , \mathbf{V}_2 y sus duales \mathbf{V}_1^* y \mathbf{V}_2^* cuyos vectores y formas pueden ser representados como

$$|v\rangle = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \in \mathbf{V}_1, \langle v| = (v_1 \ v_2 \ v_3) \in \mathbf{V}_1^* \quad \text{y} \quad |u\rangle = \begin{pmatrix} u^1 \\ u^2 \\ u^3 \\ u^4 \end{pmatrix} \in \mathbf{V}_2, \langle u| = (u_1 \ u_2 \ u_3 \ u_4) \in \mathbf{V}_2^*$$

Para ejemplificar las componentes y bases de los espacios tensoriales a partir de los espacios vectoriales y sus duales previamente identificados, podemos entonces definir la siguiente operación, como el producto tensorial:

$$|v(1)\rangle \otimes \langle u(2)| \equiv [v(1)u^*(2)] = \begin{pmatrix} v^1 \\ v^2 \\ v^3 \end{pmatrix} \otimes (u_1 \ u_2 \ u_3 \ u_4) = \begin{pmatrix} v^1 u_1 & v^1 u_2 & v^1 u_3 & v^1 u_4 \\ v^2 u_1 & v^2 u_2 & v^2 u_3 & v^2 u_4 \\ v^3 u_1 & v^3 u_2 & v^3 u_3 & v^3 u_4 \end{pmatrix} \in \mathbf{V}_3 = \mathbf{V}_1 \otimes \mathbf{V}_2^*.$$

y del mismo modo el dual del producto anterior

$$\langle w(1)| \otimes |z(2)\rangle \equiv [w^*(1)z(2)] = (w_1 \ w_2 \ w_3) \otimes \begin{pmatrix} z^1 \\ z^2 \\ z^3 \\ z^4 \end{pmatrix} = \begin{pmatrix} z^1 w_1 & z^1 w_2 & z^1 w_3 \\ z^2 w_1 & z^2 w_2 & z^2 w_3 \\ z^3 w_1 & z^3 w_2 & z^3 w_3 \\ z^4 w_1 & z^4 w_2 & z^4 w_3 \end{pmatrix} \in \mathbf{V}_3^* = \mathbf{V}_1^* \otimes \mathbf{V}_2.$$

Notemos que tanto $[v(1)u^*(2)]$ como $[v^*(1)u(2)]$ son funcionales bilineales tal y como lo definimos en 3.2.2.

Es fácil convencerse que $[v(1)u^*(2)]$ y $[v^*(1)u(2)]$ son tensores duales uno del otro.

Consideremos que si $\{|e_1(1)\rangle, |e_2(1)\rangle, |e_3(1)\rangle\}$ es una base para \mathbf{V}_1 con $\{\langle e^1(1)|, \langle e^2(1)|, \langle e^3(1)|\}$ y su dual en \mathbf{V}_1^* , entonces, éstas pueden ser definidas como:

$$|e_1(1)\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, |e_2(1)\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, |e_3(1)\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

mientras que por otra lado

$$\langle e^1(1)| = (1 \ 0 \ 0), \langle e^2(1)| = (0 \ 1 \ 0), \langle e^3(1)| = (0 \ 0 \ 1).$$

De igual forma las bases para \mathbf{V}_2 y \mathbf{V}_2^* pueden ser definidas como

$$|e_1(2)\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, |e_2(2)\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, |e_3(2)\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, |e_4(2)\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

y respectivamente:

$$\langle e^1(2)| = (1 \ 0 \ 0 \ 0), \langle e^2(2)| = (0 \ 1 \ 0 \ 0), \langle e^3(2)| = (0 \ 0 \ 1 \ 0), \langle e^4(2)| = (0 \ 0 \ 0 \ 1),$$

Entonces con el producto tensorial previamente definido, podemos construir la siguiente base para el espacio tensorial $\mathbf{V}_3 = \mathbf{V}_1 \otimes \mathbf{V}_2^*$:

$$[e_1(1); e^1(2)] = |e_1(1)\rangle \otimes \langle e^1(2)| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, [e_1(1); e^2(2)] = |e_1(1)\rangle \otimes \langle e^2(2)| = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$[e_1(1) ; e^3(2)] = |e_1(1)\rangle \otimes \langle e^3(2)| = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, [e_1(1) ; e^4(2)] = |e_1(1)\rangle \otimes \langle e^4(2)| = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$[e_2(1) ; e^1(2)] = |e_2(1)\rangle \otimes \langle e^1(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, [e_2(1) ; e^2(2)] = |e_2(1)\rangle \otimes \langle e^2(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$[e_2(1) ; e^3(2)] = |e_2(1)\rangle \otimes \langle e^3(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, [e_2(1) ; e^4(2)] = |e_2(1)\rangle \otimes \langle e^4(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$[e_3(1) ; e^1(2)] = |e_3(1)\rangle \otimes \langle e^1(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, [e_3(1) ; e^2(2)] = |e_3(1)\rangle \otimes \langle e^2(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

$$[e_3(1) ; e^3(2)] = |e_3(1)\rangle \otimes \langle e^3(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, [e_3(1) ; e^4(2)] = |e_3(1)\rangle \otimes \langle e^4(2)| = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Invocando la correspondencia entre *kets* y *bras* que expresamos en 3.1.1 construimos la base tensorial dual $[e^i(1) ; e_j(2)] = \langle e^i(1)| \otimes |e_j(2)\rangle \in \mathbf{V}_3^*$ y tendremos

$$[e^1(1) ; e_1(2)] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, [e^1(1) ; e_2(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, [e^1(1) ; e_3(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$[e^1(1) ; e_4(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, [e^2(1) ; e_1(2)] = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, [e^2(1) ; e_2(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$[e^2(1) ; e_3(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, [e^2(1) ; e_4(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, [e^3(1) ; e_1(2)] = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$[e^3(1) ; e_2(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, [e^3(1) ; e_3(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, [e^3(1) ; e_4(2)] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Como no tenemos definido un producto interno entre tensores y sus duales no podemos afirmar que esta base sea ortogonal, pero es claro que ambos conjuntos de vectores $\{[e_k(1) ; e^m(2)]\}$ y sus duales $\{[e^i(1) ; e_j(2)]\}$ son linealmente independientes en \mathbf{V}_3 y \mathbf{V}_3^* , respectivamente. Por lo tanto, cualquier tensor, $[v(1) ; u^*(2)] \in \mathbf{V}_3$ se puede expresar como combinación lineal de esta base

$$|v(1)\rangle \otimes \langle u(2)| \equiv [v(1) ; u^*(2)] = v^i u_j [e_i(1) ; e^j(2)] \equiv U_j^i [e_i(1) ; e^j(2)],$$

donde los $U_j^i = v^i u_j$ son las componentes del tensor en esta base.

Equivalentemente, el tensor $[w^*(1) ; z(2)] \in \mathbf{V}_3^*$, también podrá expresarse en la base dual

$$\langle w(1) | \otimes | z(2) \rangle \equiv [w^*(1) ; z(2)] = [e^i(1) ; e_j(2)] z^j w_i \equiv [e^i(1) ; e_j(2)] W_i^j.$$

Otra vez, los $W_i^j = z^j w_i$ son las componentes del tensor $[w^*(1) ; z(2)]$ en la base $[e^i(1) ; e_j(2)]$.



3.2.6 Tensor métrico, bases y componentes

Para una base genérica, $\{|u_j\rangle\}$, no necesariamente ortogonal, de un espacio vectorial con producto interno, podemos definir la expresión de un tensor simétrico, $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$ que denominaremos “tensor métrico”, de la siguiente manera:

$$\mathbf{g} \begin{bmatrix} |u_i\rangle & |u_j\rangle \\ \downarrow \circ, & \downarrow \circ, \end{bmatrix} = \mathbf{g} [|u_i\rangle, |u_j\rangle] = g_{|u_i\rangle |u_j\rangle} \equiv g_{|u_j\rangle |u_i\rangle}, \quad (3.6)$$

$$\mathbf{g} \begin{bmatrix} \langle u^i | & \langle u^j | \\ \downarrow \bullet, & \downarrow \bullet \end{bmatrix} = \mathbf{g} [\langle u^i |, \langle u^j |] = g^{\langle u^i | \langle u^j |} \equiv g^{\langle u^j | \langle u^i |}, \quad (3.7)$$

con $g^{\langle u^i | \langle u^j |} = (g_{|u_i\rangle |u_j\rangle})^{-1}$.

Nótese que las $g_{|u_i\rangle |u_j\rangle} \equiv g_{|u_j\rangle |u_i\rangle}$ son las componentes del tensor $\mathbf{g} [\circ, \circ]$ una vez que la base $\{|u_j\rangle\}$ ha actuado. Esto hace que podamos definir una forma a partir del tensor métrico como $\mathbf{g} [|u_i\rangle, \circ] = \langle u^i |$ y equivalentemente un vector como $\mathbf{g} [\langle u^j |, \bullet] = |u_j\rangle$.

Ambas asociaciones nos llevan a pensar que el tensor métrico está relacionado a una función del producto interno, vale decir $\mathbf{g} [\langle u^i |, \langle u^j |] = g^{\langle u^i | \langle u^j |} = g^{\langle u^j | \langle u^i |} = \mathcal{F}[\langle u^i | u_j \rangle]$, que preserve la idea de métrica: $\mathcal{F} \in \mathbb{R}$, que sea simétrico respecto a los índices i, j y que cumpla con la desigualdad triangular. Entonces, la denominación de tensor **métrico**, no es gratuita, \mathbf{g} cumple con todas las propiedades de la métrica definida para un **espacio vectorial euclíadiano** expuestas en la sección 2.2.1, vale decir:

1. $\mathbf{g} [|u_i\rangle, |u_j\rangle] = g_{ij} \equiv g_{ji} \geq 0 \quad \forall |u_j\rangle, \text{ y si } \mathbf{g} [|u_i\rangle, |u_j\rangle] = 0 \Rightarrow i = j$.
2. $\mathbf{g} [|u_i\rangle, |u_j\rangle] = \mathbf{g} [|u_j\rangle, |u_i\rangle] \Rightarrow g_{ij} \equiv g_{ji}$.
3. $\mathbf{g} [|u_i\rangle, |u_j\rangle] \leq \mathbf{g} [|u_i\rangle, |u_k\rangle] + \mathbf{g} [|u_k\rangle, |u_j\rangle]$: La desigualdad Triangular.

Si la base genérica es ortonormal, $\{|u_i\rangle\} \rightarrow \{|\hat{e}_i\rangle\}$, entonces tendremos, de manera natural:

$$\mathbf{g} [\circ, \circ] \equiv g_{ij} \langle \hat{e}^i | \otimes \langle \hat{e}^j | \equiv g_{ji} \langle \hat{e}^j | \otimes \langle \hat{e}^i | \quad \text{y} \quad \mathbf{g} [\bullet, \bullet] \equiv g^{ij} |\hat{e}_i\rangle \otimes |\hat{e}_j\rangle \equiv g^{ji} |\hat{e}_j\rangle \otimes |\hat{e}_i\rangle, \quad (3.8)$$

claramente sus componentes tensoriales serán simétricas $g_{ij} = g_{ji}$, (igualmente las componentes contravariantes $g^{ij} = g^{ji}$) y

$$(g_{ij} \langle \hat{e}^i | \otimes \langle \hat{e}^j |) (g^{km} |\hat{e}_k\rangle \otimes |\hat{e}_m\rangle) = g_{ij} g^{km} \langle \hat{e}^i | \hat{e}_k \rangle \langle \hat{e}^j | \hat{e}_m \rangle = g_{ij} g^{km} \delta_k^i \delta_m^j = g_{ij} g^{ji} = \delta_i^j = n,$$

ya que $i, j = 1, 2, 3, \dots, n$.

Por lo tanto, g_{ij} es la matriz inversa de g^{ij} , es decir, hemos definido las componentes covariantes y contravariantes del tensor de modo que cumplen con $g_{ik} g^{kj} = \delta_i^j$.

También es claro que si $|a\rangle = a^k |\hat{e}_k\rangle$, entonces:

$$(g_{ij} \langle \hat{e}^i | \otimes \langle \hat{e}^j |) |a\rangle = a^k (g_{ij} \langle \hat{e}^i | \otimes \langle \hat{e}^j |) |\hat{e}_i\rangle = a^k g_{ij} \langle \hat{e}^j | \hat{e}_k \rangle \langle \hat{e}^i | = a^k g_{ij} \delta_k^j \langle \hat{e}^i | = a^k g_{ik} \langle \hat{e}^i | \equiv a_i \langle \hat{e}^i |,$$

con lo cual $a_i = a^k g_{ik}$. De la misma forma:

$$\langle a | (g^{ij} |\hat{e}_i\rangle \otimes |\hat{e}_j\rangle) = \langle a | (g^{ij} |\hat{e}_i\rangle \otimes |\hat{e}_j\rangle) = g^{ij} \langle a |\hat{e}_i\rangle \otimes |\hat{e}_j\rangle = a_k g^{ij} \langle \hat{e}^k |\hat{e}_i\rangle |\hat{e}_j\rangle = a_k g^{kj} |\hat{e}_j\rangle \equiv a^j |\hat{e}_j\rangle ,$$

otra vez $a^j = a_k g^{kj}$, ahora subimos el índice correspondiente.

De esta manera, el tensor métrico nos permite asociar formas con vectores, componentes covariantes (de las formas) a componentes contravariantes (de los vectores). Dicho rápido y feo, pero de uso muy frecuente: el tensor métrico nos permite “subir y bajar índices”.

Otra forma de verlo es combinando las propiedades del producto directo de tensores y contracción de índices:

$$g^{ij} |\hat{e}_i\rangle \otimes |\hat{e}_j\rangle \otimes P_k{}^{lmn} |\hat{e}_l\rangle \otimes |\hat{e}_m\rangle \otimes |\hat{e}_n\rangle \otimes \langle \hat{e}^k | \Rightarrow g^{ij} P_k{}^{lmn} |\hat{e}_j\rangle \otimes P_k{}^{lmn} |\hat{e}_l\rangle \otimes |\hat{e}_m\rangle \otimes |\hat{e}_n\rangle \otimes \langle \hat{e}^k | \hat{e}_i\rangle$$

$$g^{ij} P_k{}^{lmn} |\hat{e}_j\rangle \otimes |\hat{e}_l\rangle \otimes |\hat{e}_m\rangle \otimes |\hat{e}_n\rangle \cdot \underbrace{\langle \hat{e}^k | \hat{e}_i\rangle}_{\delta_i^k} = P^{jlmn} |\hat{e}_j\rangle \otimes |\hat{e}_l\rangle \otimes |\hat{e}_m\rangle \otimes |\hat{e}_n\rangle \Rightarrow g^{ij} P_i{}^{lmn} \equiv P^{jlmn} .$$

Adicionalmente, el tensor métrico permite la contracción de índices. Así, dado un producto tensorial de dos vectores que se pueden expresar en una base ortonormal $\{|\hat{e}_i\rangle\}$:

$$|a, b\rangle = |a\rangle \otimes |b\rangle = a^k b^m |\hat{e}_k\rangle \otimes |\hat{e}_m\rangle$$

↓

$$(g_{ij} \langle \hat{e}^i | \otimes \langle \hat{e}^j |) (a^k |\hat{e}_k\rangle \otimes b^m |\hat{e}_m\rangle) = a^k b^m g_{ij} \delta_k^i \delta_m^j = a^k b^m g_{km} = a^k b_k = \langle b | a \rangle = \langle a | b \rangle .$$

Es decir, el producto interno de dos vectores involucra, de manera natural, la métrica del espacio,

$$\langle b | a \rangle = \langle a | b \rangle = a^k b_k = a_k b^k = a^k b^m g_{km} = a_k b_m g^{km} .$$

Obviamente la norma de un vector, también incluirá al tensor métrico:

$$\| |a\rangle \|^2 = \langle a | a \rangle = a_i a^j \langle \hat{e}^i | \hat{e}_j \rangle = a_i a^i = a_i a_j g^{ij} = a^i a^j g_{ij} .$$

Si partimos de la definición general de un tensor, cuando es evaluado en dos vectores tendremos

$$\mathbf{g} \begin{bmatrix} |a\rangle & |b\rangle \\ \downarrow & \downarrow \\ \circ & \circ \end{bmatrix} = \mathbf{g} [|a\rangle, |b\rangle] = g_{ij} a^k b^m \langle \hat{e}^i | \hat{e}_k \rangle \langle \hat{e}^j | \hat{e}_m \rangle = a^k b^m g_{ij} \delta_k^i \delta_m^j = a^k b^m g_{km} = a^k b_k = \langle b | a \rangle = \langle a | b \rangle .$$

Consideramos la definición (3.6) para una base genérica, $\{|u_j\rangle\}$, no necesariamente ortogonal, tendremos

$$\mathbf{g} \begin{bmatrix} |a\rangle & |b\rangle \\ \downarrow & \downarrow \\ \circ & \circ \end{bmatrix} = \langle b | a \rangle = \langle a | b \rangle \Rightarrow \mathbf{g} \begin{bmatrix} |u_i\rangle & |u_j\rangle \\ \downarrow & \downarrow \\ \circ & \circ \end{bmatrix} = g_{|u_i\rangle |u_j\rangle} \Leftrightarrow \langle u^i | u_i \rangle . \quad (3.9)$$

Estamos asociando la componente del tensor métrico al producto interno de la base de vectores (genérica, no necesariamente ortonormal). Hay que aclarar que utilizamos la palabra asociar para no violentar nuestra notación de base directas y duales. Adicionalmente, debemos garantizar que $g^{\langle u^i | \langle u^j |} = (g_{|u_i\rangle |u_j\rangle})^{-1}$. Analizaremos algunos ejemplos concretos en la sección 3.5.

3.2.7 Métrica, elemento de línea y factores de escala

El caso más emblemático lo constituye la norma de un desplazamiento infinitesimal. Para una base genérica, $\{|u_i\rangle\}$, no necesariamente ortogonal de un espacio vectorial con producto interno, el desplazamiento infinitesimal puede expresarse como:

$$ds^2 \equiv \langle dr | dr \rangle = (dx_k \langle u^k |) (dx^m | u_m \rangle) = \langle u^k | u_m \rangle dx_k dx^m = dx_m dx^m = g_{km} dx^k dx^m . \quad (3.10)$$

Si las bases de formas y vectores son ortogonales, $\{|e_i\rangle\}$, (cosa más o menos común pero no necesariamente cierta siempre) la métrica será diagonal y como en general: $\| |e_i\rangle \| \neq 1$, entonces surgen los llamados factores de escala $h_i = \sqrt{g_{ii}}$:

$$g_{ii} = \frac{1}{g^{ii}} \Rightarrow (ds)^2 = (h_1 dx^1)^2 + (h_2 dx^2)^2 + (h_3 dx^3)^2, \quad (3.11)$$

donde $h_i = \sqrt{g_{ii}}$, con $i, j = 1, 2, 3$ (aquí no hay suma).

De esta manera, las componentes covariantes y contravariantes estarán relacionadas, a través de los factores de escala como:

$$a_j = g_{jk} a^k \Rightarrow a_i = h_{[i]} a^{[i]}. \quad (\text{Aquí } h_{[i]} a^{[i]} \text{ NO indica suma}).$$

En otras palabras, en aquellos sistemas de coordenadas en los cuales la métrica es diagonal pero no viene representada por la matriz unidad, subir y bajar indices puede incluir los cambios de escala. Obviamente, si la base $\{|i_i\rangle\}$ es la canónica, es fácil ver que:

$$(ds)^2 \equiv \langle dr | dr \rangle = \delta_m^k dx_k dx^m = dx_m dx^m.$$

es decir: $g_{11} = g_{22} = g_{33} = 1$, $g_{ij} = 0$ si $i \neq j$, esto significa que en coordenadas cartesianas el desplazamiento infinitesimal es la ya conocida expresión: $ds^2 = dx^2 + dy^2 + dz^2$.

En este punto analizaremos dos ejemplos en los cuales construiremos los tensores métricos a partir de dos bases para \mathbb{R}^3 .

Ejemplo 3.6 Bases oblicuas, bases recíprocas y tensores métricos.

Consideremos una base genérica oblicua para \mathbb{R}^3 , formada por los vectores:

$$\mathbf{u}_1 = \hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}, \mathbf{u}_2 = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}} \text{ y } \mathbf{u}_3 = \hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 4\hat{\mathbf{k}}.$$

Revisaremos si estos vectores son mutuamente ortogonales. Encontraremos la base recíproca \mathbf{u}^i , el tensor métrico en ambas bases y para el vector $\mathbf{a} = 3\mathbf{u}_1 + 2\mathbf{u}_2 + \mathbf{u}_3$ encontraremos sus componentes covariantes.

Para saber si son ortogonales simplemente calculamos el producto escalar entre ellos:

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 9, \mathbf{u}_1 \cdot \mathbf{u}_3 = 6 \text{ y } \mathbf{u}_2 \cdot \mathbf{u}_3 = 7,$$

por lo tanto no son ortogonales y adicionalmente sabemos que (ver 3.1)

$$\langle \mathbf{u}^i | = \frac{|\mathbf{u}_j\rangle \times |\mathbf{u}_k\rangle}{|\mathbf{u}_i\rangle \cdot (|\mathbf{u}_j\rangle \times |\mathbf{u}_k\rangle)}.$$

A pesar que hemos exagerado con la notación, esto no es otra cosa que la receta para construir la base recíproca, que en el lenguaje de formas y vectores coincide con la base dual, es decir la base del espacio de 1-formas.

En general la base dual será:

$$\langle \mathbf{u}^i | \Rightarrow \mathbf{u}^i = \frac{\mathbf{u}_j \times \mathbf{u}_k}{V} \Rightarrow \begin{cases} \mathbf{u}^1 = \frac{\mathbf{u}_2 \times \mathbf{u}_3}{V} = \frac{17}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} - \frac{5}{6}\hat{\mathbf{k}} \\ \mathbf{u}^2 = \frac{\mathbf{u}_3 \times \mathbf{u}_1}{V} = -\frac{5}{3}\hat{\mathbf{i}} + \frac{1}{3}\hat{\mathbf{j}} + \frac{2}{3}\hat{\mathbf{k}} \\ \mathbf{u}^3 = \frac{\mathbf{u}_1 \times \mathbf{u}_2}{V} = -\frac{1}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} + \frac{1}{6}\hat{\mathbf{k}}, \end{cases}$$

donde $V = \mathbf{u}_1 \cdot (\mathbf{u}_2 \times \mathbf{u}_3) \Rightarrow (\hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}) \cdot ((\hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}) \times (\hat{\mathbf{i}} - 3\hat{\mathbf{j}} + 4\hat{\mathbf{k}})) = 6$.

Notemos que el volumen unitario del espacio dual se construye:

$$\tilde{V} = \mathbf{u}^1 \cdot (\mathbf{u}^2 \times \mathbf{u}^3) \Rightarrow \left(\frac{17}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} - \frac{5}{6}\hat{\mathbf{k}} \right) \cdot \left(\left[-\frac{5}{3}\hat{\mathbf{i}} + \frac{1}{3}\hat{\mathbf{j}} + \frac{2}{3}\hat{\mathbf{k}} \right] \times \left[-\frac{1}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} + \frac{1}{6}\hat{\mathbf{k}} \right] \right) = \frac{1}{6}.$$

Construimos entonces el tensor métrico para la base recíproca implementando la ecuación (3.9):

$$g^{ij} = g^{ji} \Leftrightarrow \begin{pmatrix} \mathbf{u}^1\mathbf{u}^1 & \mathbf{u}^1\mathbf{u}^2 & \mathbf{u}^1\mathbf{u}^3 \\ \mathbf{u}^2\mathbf{u}^1 & \mathbf{u}^2\mathbf{u}^2 & \mathbf{u}^2\mathbf{u}^3 \\ \mathbf{u}^3\mathbf{u}^1 & \mathbf{u}^3\mathbf{u}^2 & \mathbf{u}^3\mathbf{u}^3 \end{pmatrix} = \begin{pmatrix} \frac{35}{4} & -\frac{16}{3} & -\frac{7}{12} \\ -\frac{16}{3} & \frac{10}{3} & \frac{1}{3} \\ -\frac{7}{12} & \frac{1}{3} & \frac{1}{12} \end{pmatrix},$$

mientras que para la base directa:

$$g_{ij} = g_{ji} \Leftrightarrow \begin{pmatrix} \mathbf{u}_1\mathbf{u}_1 & \mathbf{u}_1\mathbf{u}_2 & \mathbf{u}_1\mathbf{u}_3 \\ \mathbf{u}_2\mathbf{u}_1 & \mathbf{u}_2\mathbf{u}_2 & \mathbf{u}_2\mathbf{u}_3 \\ \mathbf{u}_3\mathbf{u}_1 & \mathbf{u}_3\mathbf{u}_2 & \mathbf{u}_3\mathbf{u}_3 \end{pmatrix} = \begin{pmatrix} 6 & 9 & 6 \\ 9 & 14 & 7 \\ 6 & 7 & 26 \end{pmatrix}.$$

En ambos casos hemos denotado $\mathbf{u}^i\mathbf{u}^j$ como el producto escalar entre los vectores \mathbf{u}^i y \mathbf{u}^j y su equivalente en la base recíproca. Y, como debe ser,

$$g_{ia}g^{aj} = \delta_i^j = \begin{pmatrix} 6 & 9 & 6 \\ 9 & 14 & 7 \\ 6 & 7 & 26 \end{pmatrix} \begin{pmatrix} \frac{35}{4} & -\frac{16}{3} & -\frac{7}{12} \\ -\frac{16}{3} & \frac{10}{3} & \frac{1}{3} \\ -\frac{7}{12} & \frac{1}{3} & \frac{1}{12} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Entonces, para un vector dado por: $\mathbf{a} = 3\mathbf{u}_1 + 2\mathbf{u}_2 + \mathbf{u}_3$, podemos calcular sus componentes covariantes de la manera siguiente:

$$a_i = g_{ij}a^j = \begin{cases} a_1 = g_{11}a^1 + g_{12}a^2 + g_{13}a^3 = (6)(3) + (9)(2) + (6)(1) = 42 \\ a_2 = g_{21}a^1 + g_{22}a^2 + g_{23}a^3 = (9)(3) + (14)(2) + (7)(1) = 62 \\ a_3 = g_{31}a^1 + g_{32}a^2 + g_{33}a^3 = (6)(3) + (7)(2) + (26)(1) = 58 \end{cases}$$

Esto es: $\mathbf{a} = a_1\hat{\mathbf{i}} + a_2\hat{\mathbf{j}} + a_3\hat{\mathbf{k}} = 42\mathbf{u}^1 + 62\mathbf{u}^2 + 58\mathbf{u}^3$.

△

Ejemplo 3.7 Bases ortogonales, bases recíprocas y tensores métricos.

El segundo ejemplos será construir una base dual a una base de vectores mutuamente ortogonales, para ello consideraremos los siguientes vectores

$$\mathbf{w}_1 = \hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}, \quad \mathbf{w}_2 = -\hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}}, \quad \mathbf{w}_3 = 3\hat{\mathbf{i}} - 3\hat{\mathbf{j}}.$$

Ortogonalidad: $\mathbf{w}_1 \cdot \mathbf{w}_2 = 0$, $\mathbf{w}_1 \cdot \mathbf{w}_3 = 0$ y $\mathbf{w}_2 \cdot \mathbf{w}_3 = 0$. Son ortogonales.

La base recíproca se construye a partir de:

$$\mathbf{w}^i = \frac{\mathbf{w}_j \times \mathbf{w}_k}{\mathbf{w}_i \cdot (\mathbf{w}_j \times \mathbf{w}_k)},$$

donde el denominador es:

$$V = \mathbf{w}_1 \cdot (\mathbf{w}_2 \times \mathbf{w}_3) \Rightarrow (\hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}}) \cdot ((-\hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}}) \times [3\hat{\mathbf{i}} - 3\hat{\mathbf{j}}]) = 18,$$

y por lo tanto:

$$\mathbf{w}^i = \frac{\mathbf{w}_j \times \mathbf{w}_k}{V} \Rightarrow \begin{cases} \mathbf{w}^1 = \frac{\mathbf{w}_2 \times \mathbf{w}_3}{V} = \frac{1}{6}\hat{\mathbf{i}} + \frac{1}{6}\hat{\mathbf{j}} + \frac{1}{3}\hat{\mathbf{k}} \\ \mathbf{w}^2 = \frac{\mathbf{w}_3 \times \mathbf{w}_1}{V} = -\frac{1}{3}\hat{\mathbf{i}} - \frac{1}{3}\hat{\mathbf{j}} + \frac{1}{3}\hat{\mathbf{k}} \\ \mathbf{w}^3 = \frac{\mathbf{w}_1 \times \mathbf{w}_2}{V} = \frac{1}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} \end{cases}$$

El volumen recíproco, como era de esperarse:

$$\tilde{V} = \mathbf{w}^1 \cdot (\mathbf{w}^2 \times \mathbf{w}^3) \Rightarrow \left(\frac{1}{6}\hat{\mathbf{i}} + \frac{1}{6}\hat{\mathbf{j}} + \frac{1}{3}\hat{\mathbf{k}} \right) \cdot \left(\left[-\frac{1}{3}\hat{\mathbf{i}} - \frac{1}{3}\hat{\mathbf{j}} + \frac{1}{3}\hat{\mathbf{k}} \right] \times \left[\frac{1}{6}\hat{\mathbf{i}} - \frac{1}{6}\hat{\mathbf{j}} \right] \right) = \frac{1}{18},$$

mientras que el tensor métrico para la base recíproca es:

$$\bar{g}^{ij} = \bar{g}^{ji} = \mathbf{w}^i \otimes \mathbf{w}^j = \begin{pmatrix} \mathbf{w}^1 \mathbf{w}^1 & \mathbf{w}^1 \mathbf{w}^2 & \mathbf{w}^1 \mathbf{w}^3 \\ \mathbf{w}^2 \mathbf{w}^1 & \mathbf{w}^2 \mathbf{w}^2 & \mathbf{w}^2 \mathbf{w}^3 \\ \mathbf{w}^3 \mathbf{w}^1 & \mathbf{w}^3 \mathbf{w}^2 & \mathbf{w}^3 \mathbf{w}^3 \end{pmatrix} = \begin{pmatrix} \frac{1}{6} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{18} \end{pmatrix}$$

y para la base original

$$g_{ij} = g_{ji} = \mathbf{w}_i \otimes \mathbf{w}_j = \begin{pmatrix} \mathbf{w}_1 \mathbf{w}_1 & \mathbf{w}_1 \mathbf{w}_2 & \mathbf{w}_1 \mathbf{w}_3 \\ \mathbf{w}_2 \mathbf{w}_1 & \mathbf{w}_2 \mathbf{w}_2 & \mathbf{w}_2 \mathbf{w}_3 \\ \mathbf{w}_3 \mathbf{w}_1 & \mathbf{w}_3 \mathbf{w}_2 & \mathbf{w}_3 \mathbf{w}_3 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 18 \end{pmatrix}.$$

Entonces el vector $\mathbf{a} = 3\mathbf{w}_1 + 2\mathbf{w}_2 + \mathbf{w}_3 = a^1\mathbf{w}_1 + a^2\mathbf{w}_2 + a^3\mathbf{w}_3$, tendrá como componentes covariantes

$$a_i = g_{ij}a^j = \begin{cases} a_1 = g_{11}a^1 = (6)(3) = 18 \\ a_2 = g_{22}a^2 = (3)(2) = 6 \\ a_3 = g_{33}a^3 = (18)(1) = 18 \end{cases}$$

Esto es: $\mathbf{a} = a_1\hat{\mathbf{i}} + a_2\hat{\mathbf{j}} + a_3\hat{\mathbf{k}} = 18\mathbf{w}^1 + 6\mathbf{w}^2 + 18\mathbf{w}^3$.

Notemos que si la base original hubiese sido ortonormal:

$$\mathbf{w}_1 = (\hat{\mathbf{i}} + \hat{\mathbf{j}} + 2\hat{\mathbf{k}})/\sqrt{6}, \quad \mathbf{w}_2 = (-\hat{\mathbf{i}} - \hat{\mathbf{j}} + \hat{\mathbf{k}})/\sqrt{3}, \quad \mathbf{w}_3 = (3\hat{\mathbf{i}} - 3\hat{\mathbf{j}})/\sqrt{18},$$

entonces:

$$g_{ij} = \bar{g}^{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow \begin{cases} a_1 = a^1 \\ a_2 = a^2 \\ a_3 = a^3 \end{cases}$$



3.2.8 Vectores, formas, tensores y leyes de transformación

En esta sección discutiremos la importancia de caracterizar objetos identificando las leyes de transformación. La invariancia bajo determinadas leyes de transformación es clave en Física y nos permite identificar propiedades fundamentales. Ya hemos visto, en la sección 1.4.4, como los esquemas de transformación de coordenadas nos han permitido diferenciar escalares y pseudoscalares y vectores de pseudovectores. En esta sección puntualizaremos como transforman vectores y formas y como esos esquemas de transformaciones los heredan los tensores.

En general las afirmaciones anteriores se pueden generalizar considerando que las coordenadas que definen un determinado punto, P , expresado en un sistema de coordenadas particular, son (x^1, x^2, \dots, x^n) y las coordenadas de ese mismo punto P , expresado en otro sistema de coordenadas son $(x^{1'}, x^{2'}, \dots, x^{n'})$. Ambas coordenadas estarán relacionadas por:

$$\left. \begin{array}{l} x^{1'} = x^{1'}(x^1, x^2, \dots, x^n) \\ x^{2'} = x^{2'}(x^1, x^2, \dots, x^n) \\ \vdots \\ x^{n'} = x^{n'}(x^1, x^2, \dots, x^n) \end{array} \right\} \iff \left. \begin{array}{l} x^1 = x^1(x^{1'}, x^{2'}, \dots, x^{n'}) \\ x^2 = x^2(x^{1'}, x^{2'}, \dots, x^{n'}) \\ \vdots \\ x^n = x^n(x^{1'}, x^{2'}, \dots, x^{n'}) \end{array} \right\}$$

En una notación más compacta lo que tenemos es:

$$x^{i'} = x^{i'}(x^j) \iff x^i = x^i(x^{j'}), \quad (3.12)$$

con: $i, j = 1, 2, 3, \dots, n$ y $i', j' = 1, 2, 3, \dots, n$.

Retomemos lo que discutimos en la sección 3.1.3, pero ahora en el lenguaje de coordenadas. En esa

oportunidad mostramos como transformaban las bases y las componentes de formas y vectores. Ahora lo generalizaremos a los tensores. Otra vez, sólo exigiremos (y es bastante) que:

1. Las funciones $x^i = x^i(x^{m'})$ y $x^{j'} = x^{j'}(x^m)$ sean al menos C^2 (función y derivada continua)
2. Que el determinante de la matriz jacobiana sean finito y distinto de cero, esto es

$$\det \left| \frac{\partial x^i(x^{m'})}{\partial x^{j'}} \right| \neq 0 \Rightarrow \begin{vmatrix} \frac{\partial x^1}{\partial x^{1'}} & \frac{\partial x^1}{\partial x^{2'}} & \cdots & \frac{\partial x^1}{\partial x^{n'}} \\ \frac{\partial x^2}{\partial x^{1'}} & \frac{\partial x^2}{\partial x^{2'}} & \cdots & \frac{\partial x^2}{\partial x^{n'}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x^n}{\partial x^{1'}} & \frac{\partial x^n}{\partial x^{2'}} & \cdots & \frac{\partial x^n}{\partial x^{n'}} \end{vmatrix} \neq 0 \Rightarrow x^i = x^i(x^{m'}) \iff x^{j'} = x^{j'}(x^m).$$

Ahora bien, una vez más, derivando y utilizando la regla de la cadena:

$$x^i = x^i(x^{j'}(x^m)) \Rightarrow \frac{\partial x^i}{\partial x^m} = \frac{\partial x^i}{\partial x^{j'}} \frac{\partial x^{j'}}{\partial x^m} = \delta_m^i \Rightarrow dx^i = \frac{\partial x^i}{\partial x^{j'}} dx^{j'},$$

como hemos comprobado para los dos casos particulares estudiados con anterioridad.

De ahora en adelante tendremos las siguientes redefiniciones:

1. Un conjunto de cantidades $\{a^1, a^2, \dots, a^n\}$ se denominarán componentes *contravariantes* de un vector $|a\rangle \in \mathbf{V}$ en un punto P de coordenadas (x^1, x^2, \dots, x^n) , si:
 - (a). dada dos bases ortonormales de vectores coordinados: $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, \dots, |\hat{e}_n\rangle\}$ y $\{|\hat{e}_{1'}\rangle, |\hat{e}_{2'}\rangle, \dots, |\hat{e}_{n'}\rangle\}$, se cumple que:

$$|a\rangle = a^i |\hat{e}_i\rangle = a^{i'} |\hat{e}_{i'}\rangle \Rightarrow \begin{cases} \langle \hat{e}^i | a \rangle = a^i \\ \langle \hat{e}^{i'} | a \rangle = a^{i'} \end{cases} \Rightarrow a^{i'} = a^j \langle \hat{e}^{i'} | \hat{e}_j \rangle.$$

- (b). o equivalentemente, bajo una transformación de coordenadas: $x^i = x^i(x^{j'})$ (con $i, j = 1, 2, 3, \dots, n$) y $i', j' = 1, 2, 3, \dots, n$), estas cantidades transforman como:

$$a^{i'} = \frac{\partial x^{i'}}{\partial x^k} a^k \iff a^i = \frac{\partial x^i}{\partial x^{k'}} a^{k'}, \text{ con: } \frac{\partial x^i}{\partial x^{k'}} \frac{\partial x^{k'}}{\partial x^l} = \delta_l^i, \quad (3.13)$$

2. Un conjunto de cantidades $\{b_1, b_2, \dots, b_n\}$ se denominarán componentes *covariantes* de un vector $\langle b | \in \mathbf{V}^*$ en un punto P de coordenadas (x^1, x^2, \dots, x^n) , si:

- (a). dada dos bases de formas: $\{ \langle \hat{e}^1 |, \langle \hat{e}^2 |, \dots, \langle \hat{e}^n | \}$ y $\{ \langle \hat{e}^{1'} |, \langle \hat{e}^{2'} |, \dots, \langle \hat{e}^{n'} | \}$ se cumple que:

$$\langle b | = b_i \langle \hat{e}^i | = b_{i'} \langle \hat{e}^{i'} | \Rightarrow \begin{cases} \langle b | \hat{e}^i = b^i \\ \langle b | \hat{e}^{i'} = b^{i'} \end{cases} \Rightarrow b^{i'} = b^j \langle \hat{e}_j | \hat{e}^{i'} \rangle.$$

- (b). o equivalentemente, bajo una transformación de coordenadas $x^i = x^i(x^{j'})$ (con $i, j = 1, 2, 3, \dots, n$) y $i', j' = 1, 2, 3, \dots, n$) estas cantidades transforman como:

$$b_{k'} = \frac{\partial x^i}{\partial x^{k'}} b_i \iff b_k = \frac{\partial x^i}{\partial x^k} b_{i'} \text{ con: } \frac{\partial x^i}{\partial x^{k'}} \frac{\partial x^{k'}}{\partial x^l} = \delta_l^i, \quad (3.14)$$

Es bueno recalcar que cantidades $\frac{\partial x^{i'}}{\partial x^k}$ y $\frac{\partial x^i}{\partial x^{k'}}$ deben ser evaluadas en el punto P .

Podemos generalizar los conceptos anteriores:

1. Dado un conjunto bases para las formas diferenciales $\{ \langle x^m(1) |, \langle y^n(2) | \}$, hemos definido las componentes *contravariantes* de un tensor:

$$T^{ij} = \mathcal{T} \left[\begin{array}{c|c} \langle x^i(1) | & \langle y^j(2) | \\ \downarrow \bullet & \downarrow \bullet \end{array} \right] \iff \{T^{ij}\} \equiv \{T^{11}, T^{12}, \dots, T^{1n}, T^{21}, T^{22}, \dots, T^{2n}, \dots, T^{nn}\},$$

Las componentes contravariantes en un punto P de coordenadas (x^1, x^2, \dots, x^n) $\Leftrightarrow x^i = x^i(x^{j'})$ (con

$i, j = 1, 2, 3, \dots, n$ y $i', j' = 1, 2, 3, \dots, n$) transforman como:

$$T^{i'j'} = \frac{\partial x^{i'}}{\partial x^k} \frac{\partial x^{j'}}{\partial x^m} T^{km} \iff T^{ij} = \frac{\partial x^i}{\partial x^{k'}} \frac{\partial x^j}{\partial x^{m'}} T^{k'm'}.$$

Esto nos permite construir el caso más general.

2. Si $\{|t_i(1)\rangle, |u_j(2)\rangle, \dots, |v_k(m)\rangle\}$ y $\{|\langle x^e(1)|, \langle y^f(2)|, \dots, \langle z^g(n)|\}$ son bases para los vectores y las formas, respectivamente, las componentes de un tensor:

$$T_{ijk}{}^{mn} = \mathcal{T} \begin{bmatrix} |t_i(1)\rangle & |u_j(2)\rangle & |v_k(m)\rangle & |\langle x^e(1)| & |\langle y^f(2)| & |\langle z^g(n)| \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \circ & \circ & \dots & \circ & \bullet & \bullet & \dots & \bullet & \bullet \end{bmatrix},$$

serán un conjunto de cantidades: $\{T_{ijk}{}^{mn}\}$ que se denominarán las componentes *contravariantes* y *covariantes* respectivamente, de un tensor mixto en un punto P de coordenadas (x^1, x^2, \dots, x^n) .

Bajo una transformación de coordenadas $x^i = x^i(x^{j'})$, estas cantidades transforman como:

$$T_{e' \dots g'}{}^{i' \dots k'} = \frac{\partial x^{i'}}{\partial x^p} \dots \frac{\partial x^{k'}}{\partial x^q} \frac{\partial x^a}{\partial x^{e'}} \dots \frac{\partial x^d}{\partial x^{g'}} T_{a \dots d}{}^{p \dots q} \iff T_{e' \dots g'}{}^{i \dots k} = \frac{\partial x^i}{\partial x^{p'}} \dots \frac{\partial x^k}{\partial x^{q'}} \frac{\partial x^{a'}}{\partial x^e} \dots \frac{\partial x^{d'}}{\partial x^g} T_{a' \dots d'}{}^{p' \dots q'}. \quad (3.15)$$

Recordemos nuevamente que

$$\frac{\partial x^i}{\partial x^{k'}} \frac{\partial x^{k'}}{\partial x^l} = \delta_l^i,$$

y las cantidades $\frac{\partial x^{i'}}{\partial x^k}$ y $\frac{\partial x^i}{\partial x^{k'}}$ deben ser evaluadas en el punto P .

Ejemplo 3.8 Un ejemplo detallado de transformación de tensores

Ilustremos ahora las transformaciones de tensores bajo cambios de la base del espacio tensorial. Esta vez, consideraremos que un tensor tendrá componentes T_j^i cuando se expresa como combinación lineal de la base $[\hat{e}_i(1); \hat{e}_j(2)] = |\hat{e}_i(1)\rangle \otimes \langle \hat{e}_j(2)|$ donde la base $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle\} \equiv \{|i\rangle, |j\rangle, |k\rangle\}$ son los vectores unitarios cartesianos para el espacio vectorial \mathbb{R}^3 . Esto es

$$|\hat{e}_1\rangle = |i\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; |\hat{e}_2\rangle = |j\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, |\hat{e}_3\rangle = |k\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

y su base dual

$$\langle \hat{e}_1| = \langle i| = (1 \ 0 \ 0), \langle \hat{e}_2| = \langle j| = (0 \ 1 \ 0), \langle \hat{e}_3| = \langle k| = (0 \ 0 \ 1).$$

Consideremos un tensor en la base tensorial $[\hat{e}_i(1); \hat{e}_j(2)]$ se expresa como:

$$T_j^i = \begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 2 \end{pmatrix}.$$

Este ejemplo consiste en ver cómo calcular las expresiones de los siguientes tensores: $\tilde{T}_{j'}^{i'}, \tilde{T}_{i'j'}^i$ y $\tilde{T}^{i'j'}$, si consideramos una nueva base:

$$\left\{ \begin{array}{l} |w_{1'}\rangle = |i\rangle \\ |w_{2'}\rangle = |i\rangle + |j\rangle \\ |w_{3'}\rangle = |i\rangle + |j\rangle + |k\rangle \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \langle w_{1'}| = \langle i| \\ \langle w_{2'}| = \langle i| + \langle j| \\ \langle w_{3'}| = \langle i| + \langle j| + \langle k| \end{array} \right\}$$

Nótese que la nueva base **no es ortogonal**, $\langle w^{k'}|w_{i'}\rangle \neq \delta_{i'}^{k'}$, con lo cual no se cumplen muchas cosas, entre ellas: $|w_{k'}\rangle\langle w^{k'}| \neq 1$. Pero sin duda se cumple que para transformar componentes de un tensor tendremos

siempre la relación

$$T_{m'}^{k'} = \frac{\partial x^{k'}}{\partial x^i} \frac{\partial x^j}{\partial x^{m'}} T_j^i.$$

Tenemos que identificar las matrices de transformación $\frac{\partial x^{k'}}{\partial x^i}$ y $\frac{\partial x^j}{\partial x^{m'}}$. Estas matrices de transformación son las mismas que transforman componentes (contravariantes) de vectores y las componentes (covariantes) de covectores o formas, i.e.

$$a^{i'} = \frac{\partial x^{i'}}{\partial x^k} a^k \Leftrightarrow a_{k'} = \frac{\partial x^{m'}}{\partial x^{k'}} a_m$$

y además son las mismas matrices que devuelven el cambio de un sistema de coordenadas a otro:

$$a^i = \frac{\partial x^i}{\partial x^{k'}} a^{k'} \Leftrightarrow a_k = \frac{\partial x^m}{\partial x^k} a_{m'}.$$

Adicionalmente, estas matrices son una la inversa de la otra $\frac{\partial x^i}{\partial x^{k'}} \frac{\partial x^{k'}}{\partial x^j} = \delta_j^i$.

Para identificar esas matrices recordamos que un vector genérico transforma de la siguiente manera:
 $|a\rangle = a^j |\hat{e}_j\rangle = a^{j'} |w_{j'}\rangle$, por lo tanto:

$$|a\rangle = a^j |\hat{e}_j\rangle = a^{1'} |w_{1'}\rangle + a^{2'} |w_{2'}\rangle + a^{3'} |w_{3'}\rangle = a^{1'} |\hat{e}_1\rangle + a^{2'} (|\hat{e}_1\rangle + |\hat{e}_2\rangle) + a^{3'} (|\hat{e}_1\rangle + |\hat{e}_2\rangle + |\hat{e}_3\rangle),$$

donde hemos sustituido la expresión de la base $\{|\tilde{w}_{i'}\rangle\}$ en términos de la base $\{|\hat{e}_j\rangle\}$ con lo cual:

$$|a\rangle = a^1 |\hat{e}_1\rangle + a^2 |\hat{e}_2\rangle + a^3 |\hat{e}_3\rangle = (a^{1'} + a^{2'} + a^{3'}) |\hat{e}_1\rangle + (a^{2'} + a^{3'}) |\hat{e}_2\rangle + a^{3'} |\hat{e}_3\rangle.$$

Lo que nos lleva al siguiente sistema de ecuaciones:

$$\left. \begin{array}{l} a^1 = a^{1'} + a^{2'} + a^{3'} \\ a^2 = a^{2'} + a^{3'} \\ a^3 = a^{3'} \end{array} \right\} \Rightarrow a^i = \frac{\partial x^i}{\partial x^{k'}} a^{k'} \Rightarrow \left\{ \begin{array}{l} \frac{\partial x^1}{\partial x^{1'}} = 1; \quad \frac{\partial x^1}{\partial x^{2'}} = 1; \quad \frac{\partial x^1}{\partial x^{3'}} = 1 \\ \frac{\partial x^2}{\partial x^{1'}} = 0; \quad \frac{\partial x^2}{\partial x^{2'}} = 1; \quad \frac{\partial x^2}{\partial x^{3'}} = 1 \\ \frac{\partial x^3}{\partial x^{1'}} = 0; \quad \frac{\partial x^3}{\partial x^{2'}} = 0; \quad \frac{\partial x^3}{\partial x^{3'}} = 1 \end{array} \right.$$

Es de hacer notar que dado que la base: $\{|\hat{e}_i\rangle\}$ es ortonormal, se tiene que:

$$|a\rangle = a^j |\hat{e}_j\rangle = a^{i'} |w_{i'}\rangle \Rightarrow \langle \hat{e}^i | a \rangle = a^j \langle \hat{e}^i | \hat{e}_j \rangle = a^j \delta_j^i = a^i = a^{k'} \langle \hat{e}^i | w_{k'} \rangle \Rightarrow \frac{\partial x^i}{\partial x^{k'}} = \langle \hat{e}^i | w_{k'} \rangle.$$

Este mismo procedimiento se puede aplicar para expresar el vector $|a\rangle$ como una combinación lineal de los vectores $|w_{j'}\rangle$:

$$|a\rangle = a^{j'} |\hat{e}_{j'}\rangle = a^j |\hat{e}_j\rangle = a^1 |\hat{e}_1\rangle + a^2 |\hat{e}_2\rangle + a^3 |\hat{e}_3\rangle = a^1 |w_{1'}\rangle + a^2 (|w_{2'}\rangle - |w_{1'}\rangle) + a^3 (|w_{3'}\rangle - |w_{2'}\rangle),$$

esto es:

$$\left. \begin{array}{l} a^{1'} = a^1 - a^2 \\ a^{2'} = a^2 - a^3 \\ a^{3'} = a^3 \end{array} \right\} \Rightarrow a^{k'} = \frac{\partial x^{k'}}{\partial x^i} a^i \Rightarrow \left\{ \begin{array}{l} \frac{\partial x^{1'}}{\partial x^1} = 1; \quad \frac{\partial x^{1'}}{\partial x^2} = -1; \quad \frac{\partial x^{1'}}{\partial x^3} = 0 \\ \frac{\partial x^{2'}}{\partial x^1} = 0; \quad \frac{\partial x^{2'}}{\partial x^2} = 1; \quad \frac{\partial x^{2'}}{\partial x^3} = -1 \\ \frac{\partial x^{3'}}{\partial x^1} = 0; \quad \frac{\partial x^{3'}}{\partial x^2} = 0; \quad \frac{\partial x^{3'}}{\partial x^3} = 1 \end{array} \right.$$

Nótese que, como era de esperarse:

$$\frac{\partial x^i}{\partial x^{k'}} \frac{\partial x^{k'}}{\partial x^j} = \delta_j^i \Rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Con las expresiones matriciales para las transformaciones, estamos en capacidad de calcular, componente

a componente, las representación del tensor dado en la nueva base:

$$T_{m'}^{k'} = \frac{\partial x^{k'}}{\partial x^i} \frac{\partial x^j}{\partial x^{m'}} T_j^i.$$

Veamos:

$$\begin{aligned} T_{1'}^{1'} &= \frac{\partial x^{1'}}{\partial x^i} \frac{\partial x^j}{\partial x^{1'}} T_j^i \\ &= \frac{\partial x^{1'}}{\partial x^1} \left(\frac{\partial x^1}{\partial x^{1'}} T_1^1 + \frac{\partial x^2}{\partial x^{1'}} T_2^1 + \frac{\partial x^3}{\partial x^{1'}} T_3^1 \right) + \frac{\partial x^{1'}}{\partial x^2} \left(\frac{\partial x^1}{\partial x^{1'}} T_1^2 + \frac{\partial x^2}{\partial x^{1'}} T_2^2 + \frac{\partial x^3}{\partial x^{1'}} T_3^2 \right) \\ &\quad + \frac{\partial x^{1'}}{\partial x^3} \left(\frac{\partial x^1}{\partial x^{1'}} T_1^3 + \frac{\partial x^2}{\partial x^{1'}} T_2^3 + \frac{\partial x^3}{\partial x^{1'}} T_3^3 \right), \end{aligned}$$

es decir:

$$\begin{aligned} T_{1'}^{1'} &= 1 \cdot (1 T_1^1 + 0 T_2^1 + 0 T_3^1) - 1 \cdot (1 T_1^2 + 0 T_2^2 + 0 T_3^2) + 0 (1 T_1^3 + 0 T_2^3 + 0 T_3^3) \\ &= T_1^1 - T_1^2 = 2 - 2 = 0. \end{aligned}$$

Del mismo modo para $T_{2'}^{1'}$:

$$\begin{aligned} T_{2'}^{1'} &= \frac{\partial x^{1'}}{\partial x^i} \frac{\partial x^j}{\partial x^{2'}} T_j^i \\ &= \frac{\partial x^{1'}}{\partial x^1} \left(\frac{\partial x^1}{\partial x^{2'}} T_1^1 + \frac{\partial x^2}{\partial x^{2'}} T_2^1 + \frac{\partial x^3}{\partial x^{2'}} T_3^1 \right) + \frac{\partial x^{1'}}{\partial x^2} \left(\frac{\partial x^1}{\partial x^{2'}} T_1^2 + \frac{\partial x^2}{\partial x^{2'}} T_2^2 + \frac{\partial x^3}{\partial x^{2'}} T_3^2 \right) \\ &\quad + \frac{\partial x^{1'}}{\partial x^3} \left(\frac{\partial x^1}{\partial x^{2'}} T_1^3 + \frac{\partial x^2}{\partial x^{2'}} T_2^3 + \frac{\partial x^3}{\partial x^{2'}} T_3^3 \right), \end{aligned}$$

resultando:

$$\begin{aligned} T_{2'}^{1'} &= 1 \cdot (1 T_1^1 + 1 T_2^1 + 0 T_3^1) - 1 \cdot (1 T_1^2 + 1 T_2^2 + 0 T_3^2) + 0 (1 T_1^3 + 1 T_2^3 + 0 T_3^1) \\ &= (T_1^1 + T_2^1) - (T_1^2 + T_2^2) = (2 + 1) - (2 + 3) = -2. \end{aligned}$$

Se puede continuar término a término (el lector debe terminar los cálculos) o realizar la multiplicación de las matrices provenientes de la transformación de componentes de tensores. Vale decir:

$$T_{m'}^{k'} = \left(\frac{\partial x^{k'}}{\partial x^i} \right) (T_j^i) \left(\frac{\partial x^j}{\partial x^{m'}} \right) \Leftrightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -3 \\ 1 & 2 & 4 \\ 1 & 3 & 5 \end{pmatrix}.$$

Hay que resaltar el especial cuidado que se tuvo en la ubicación de las matrices para su multiplicación. Si bien en la expresión $T_{m'}^{k'} = \frac{\partial x^{k'}}{\partial x^i} \frac{\partial x^j}{\partial x^{m'}} T_j^i$ las cantidades $\frac{\partial x^{k'}}{\partial x^i}$ son números y no importa el orden con el cual se multipliquen, cuando se escriben como matrices debe respetarse la “concatenación interna de índices”. Esto es, cuando queremos expresar $T_{m'}^{k'}$ como una matriz, donde el índice contravariante k' indica filas y el índice covariante m' las columnas, fijamos primero estos índices y luego respetamos la “concatenación de índices” covariantes con los contravariantes. Esta es la convención para expresar la multiplicación de matrices en la notación de índices⁶. Esto es:

$$T_{m'}^{k'} = \frac{\partial x^{k'}}{\partial x^i} \frac{\partial x^j}{\partial x^{m'}} T_j^i \Rightarrow \left(T_{m'}^{k'} \right) = \left(\frac{\partial x^{k'}}{\partial x^i} \right) (T_j^i) \left(\frac{\partial x^j}{\partial x^{m'}} \right).$$

Los objetos $\frac{\partial x^{k'}}{\partial x^i}, T_j^i$ y $\frac{\partial x^j}{\partial x^{m'}}$ fueron sustituidos (en sus puestos correspondientes) por su representación matricial. Con lo cual hemos encontrado la representación matricial $T_{j'}^{i'}$ de las componentes del tensor \mathbf{T} en la

⁶Quizá una forma de comprobar si los índices están bien concatenados se observa si se “bajan” los índices contravariantes pero se colocan antes que los covariantes. Esto es, $T_j^i \rightarrow T_{ij}$. Así, la multiplicación de matrices queda representada de la siguiente forma: $C_j^i = A_k^i B_j^k \rightarrow C_{ij} = A_{ik} B_{kj}$ y aquí es claro que índices consecutivos están “concatenados” e indican multiplicación.

base $\{|w_{1'}\rangle, |w_{2'}\rangle, |w_{3'}\rangle\}$

$$T_{j'}^{i'} = \begin{pmatrix} T_{1'}^{1'} & T_{2'}^{1'} & T_{3'}^{1'} \\ T_{1'}^{2'} & T_{2'}^{2'} & T_{3'}^{2'} \\ T_{1'}^{3'} & T_{2'}^{3'} & T_{3'}^{3'} \end{pmatrix} = \begin{pmatrix} 0 & -2 & -3 \\ 1 & 2 & 4 \\ 1 & 3 & 5 \end{pmatrix}.$$

Para encontrar la expresión de $T_{i'j'} (T_{i'j'} = g_{i'k'} T_{j'}^{k'})$, requeriremos las componentes covariantes y contravariantes del tensor métrico $g_{i'k'}$ que genera esta base. Para ello recordamos que para una base genérica, $\{|w_{j'}\rangle\}$, no necesariamente ortogonal, de un espacio vectorial con producto interno, podemos definir la expresión de un tensor que denominaremos tensor métrico como:

$$\text{un tensor del tipo } \begin{pmatrix} 0 \\ 2 \end{pmatrix} \Rightarrow g_{i'j'} = \frac{\partial x^m}{\partial x^{i'}} \frac{\partial x^n}{\partial x^{j'}} g_{mn} \equiv \langle \hat{e}^m | w_{i'} \rangle \langle \hat{e}^n | w_{j'} \rangle g_{mn}.$$

Recordemos que la métrica covariante g_{ij} generada por una base ortonormal $\{|\hat{e}_i\rangle\} \equiv \{|i_i\rangle\}$ es:

$$g_{11} = 1, g_{22} = 1, g_{33} = 1, g_{12} = g_{21} = g_{13} = g_{31} = g_{23} = g_{32} = 0.$$

Es decir:

$$g_{ij} = g^{ij} = g_j^i \iff \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Con lo cual, para el caso de la base genérica no ortonormal $\{|w_{j'}\rangle\}$ tenemos dos formas de calcular las componentes covariantes y contravariantes del tensor métrico.

La primera es la forma directa: $g_{i'j'} = \langle \hat{e}^m | w_{i'} \rangle \langle \hat{e}^n | w_{j'} \rangle g_{mn}$. Esto es:

$$g_{1'1'} = \langle \hat{e}^n | w_{1'} \rangle \langle \hat{e}^m | w_{1'} \rangle g_{nm} = \langle \hat{e}^1 | w_{1'} \rangle \langle \hat{e}^1 | w_{1'} \rangle + \langle \hat{e}^2 | w_{1'} \rangle \langle \hat{e}^2 | w_{1'} \rangle + \langle \hat{e}^3 | w_{1'} \rangle \langle \hat{e}^3 | w_{1'} \rangle = \langle \hat{e}^1 | w_{1'} \rangle^2 = 1$$

$$g_{1'2'} = \langle \hat{e}^n | w_{1'} \rangle \langle \hat{e}^m | w_{2'} \rangle g_{nm} = \langle \hat{e}^1 | w_{1'} \rangle \langle \hat{e}^1 | w_{2'} \rangle + \langle \hat{e}^2 | w_{1'} \rangle \langle \hat{e}^2 | w_{2'} \rangle + \langle \hat{e}^3 | w_{1'} \rangle \langle \hat{e}^3 | w_{2'} \rangle = \langle \hat{e}^1 | w_{1'} \rangle \langle \hat{e}^1 | w_{2'} \rangle = 1$$

$$g_{1'3'} = \langle \hat{e}^n | w_{1'} \rangle \langle \hat{e}^m | w_{3'} \rangle g_{nm} = \langle \hat{e}^1 | w_{1'} \rangle \langle \hat{e}^1 | w_{3'} \rangle + \langle \hat{e}^2 | w_{1'} \rangle \langle \hat{e}^2 | w_{3'} \rangle + \langle \hat{e}^3 | w_{1'} \rangle \langle \hat{e}^3 | w_{3'} \rangle = \langle \hat{e}^1 | w_{1'} \rangle \langle \hat{e}^1 | w_{3'} \rangle = 1$$

$$g_{2'1'} = \langle \hat{e}^n | w_{2'} \rangle \langle \hat{e}^m | w_{1'} \rangle g_{nm} = \langle \hat{e}^1 | w_{2'} \rangle \langle \hat{e}^1 | w_{1'} \rangle + \langle \hat{e}^2 | w_{2'} \rangle \langle \hat{e}^2 | w_{1'} \rangle + \langle \hat{e}^3 | w_{2'} \rangle \langle \hat{e}^3 | w_{1'} \rangle = \langle \hat{e}^1 | w_{2'} \rangle \langle \hat{e}^1 | w_{1'} \rangle = 1$$

$$g_{2'2'} = \langle \hat{e}^n | w_{2'} \rangle \langle \hat{e}^m | w_{2'} \rangle g_{nm} = \langle \hat{e}^1 | w_{2'} \rangle \langle \hat{e}^1 | w_{2'} \rangle + \langle \hat{e}^2 | w_{2'} \rangle \langle \hat{e}^2 | w_{2'} \rangle + \langle \hat{e}^3 | w_{2'} \rangle \langle \hat{e}^3 | w_{2'} \rangle = \langle \hat{e}^1 | w_{2'} \rangle \langle \hat{e}^1 | w_{2'} \rangle \langle \hat{e}^2 | w_{2'} \rangle \langle \hat{e}^2 | w_{2'} \rangle = 2$$

$$g_{2'3'} = \langle \hat{e}^n | w_{2'} \rangle \langle \hat{e}^m | w_{3'} \rangle g_{nm} = \langle \hat{e}^1 | w_{2'} \rangle \langle \hat{e}^1 | w_{3'} \rangle + \langle \hat{e}^2 | w_{2'} \rangle \langle \hat{e}^2 | w_{3'} \rangle + \langle \hat{e}^3 | w_{2'} \rangle \langle \hat{e}^3 | w_{3'} \rangle = \langle \hat{e}^1 | w_{2'} \rangle \langle \hat{e}^1 | w_{3'} \rangle + \langle \hat{e}^2 | w_{2'} \rangle \langle \hat{e}^2 | w_{3'} \rangle = 2$$

$$g_{3'1'} = \langle \hat{e}^n | w_{3'} \rangle \langle \hat{e}^m | w_{1'} \rangle g_{nm} = \langle \hat{e}^1 | w_{3'} \rangle \langle \hat{e}^1 | w_{1'} \rangle + \langle \hat{e}^2 | w_{3'} \rangle \langle \hat{e}^2 | w_{1'} \rangle + \langle \hat{e}^3 | w_{3'} \rangle \langle \hat{e}^3 | w_{1'} \rangle = \langle \hat{e}^1 | w_{3'} \rangle \langle \hat{e}^1 | w_{1'} \rangle = 1$$

$$g_{3'2'} = \langle \hat{e}^n | w_{3'} \rangle \langle \hat{e}^m | w_{2'} \rangle g_{nm} = \langle \hat{e}^1 | w_{3'} \rangle \langle \hat{e}^1 | w_{2'} \rangle + \langle \hat{e}^2 | w_{3'} \rangle \langle \hat{e}^2 | w_{2'} \rangle + \langle \hat{e}^3 | w_{3'} \rangle \langle \hat{e}^3 | w_{2'} \rangle = \langle \hat{e}^1 | w_{3'} \rangle \langle \hat{e}^1 | w_{2'} \rangle + \langle \hat{e}^2 | w_{3'} \rangle \langle \hat{e}^2 | w_{2'} \rangle = 2$$

$$g_{3'3'} = \langle \hat{e}^n | w_{3'} \rangle \langle \hat{e}^m | w_{3'} \rangle g_{nm} = \langle \hat{e}^1 | w_{3'} \rangle \langle \hat{e}^1 | w_{3'} \rangle + \langle \hat{e}^2 | w_{3'} \rangle \langle \hat{e}^2 | w_{3'} \rangle + \langle \hat{e}^3 | w_{3'} \rangle \langle \hat{e}^3 | w_{3'} \rangle = \langle \hat{e}^1 | w_{3'} \rangle \langle \hat{e}^1 | w_{3'} \rangle + \langle \hat{e}^2 | w_{3'} \rangle \langle \hat{e}^2 | w_{3'} \rangle + \langle \hat{e}^3 | w_{3'} \rangle \langle \hat{e}^3 | w_{3'} \rangle = 3$$

De manera que:

$$g_{i'j'} \Rightarrow g_{1'1'} = 1, g_{1'2'} = 1, g_{1'3'} = 1, g_{2'1'} = 1, g_{2'2'} = 2, g_{2'3'} = 2, g_{3'1'} = 1, g_{3'2'} = 2, g_{3'3'} = 3.$$

Consecuentemente podemos “arreglarlo como una matriz”⁷ de la siguiente forma:

$$g_{i'j'} \equiv g_{j'i'} \iff \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} \Rightarrow g^{i'j'} \equiv g^{j'i'} = (g_{i'j'})^{-1} \iff \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Nótese que también podríamos haber procedido, y en términos “matriciales”, de la siguiente forma:

$$g_{k'm'} = \frac{\partial x^i}{\partial x^{k'}} g_{ij} \frac{\partial x^j}{\partial x^{m'}} \Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

Nuevamente es bueno aclarar que para conservar la convención de índices y poder representar la multiplicación de matrices, los índices deben estar consecutivos, por lo tanto, hay que trasponer la representación matricial para multiplicarlas. Es por eso que en el cálculo anterior aparece como primera matriz la transpuesta de la última. El cálculo se debe hacer como se muestra a continuación:

$$g_{k'm'} = \frac{\partial x^i}{\partial x^{k'}} g_{ij} \frac{\partial x^j}{\partial x^{m'}} \rightarrow g_{k'm'} = \Pi_{ik'} g_{ij} \Pi_{jm'} \rightarrow g_{k'm'} = \bar{\Pi}_{k'i} g_{ij} \Pi_{jm'}.$$

Finalmente, estamos en capacidad de obtener las representaciones matriciales para los tensores: $T_{i'}^{j'}, T_{i'j'}, T^{i'j'}$.

$$T_{i'}^{j'} = (T_{j'}^{i'})^T \iff T_{i'}^{j'} = \begin{pmatrix} 0 & -2 & -3 \\ 1 & 2 & 4 \\ 1 & 3 & 5 \end{pmatrix}^T = \begin{pmatrix} 0 & 1 & 1 \\ -2 & 2 & 3 \\ -3 & 4 & 5 \end{pmatrix}.$$

$$T_{k'm'} = g_{k'n'} T_{m'}^{n'} \iff T_{k'm'} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 0 & -2 & -3 \\ 1 & 2 & 4 \\ 1 & 3 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 6 \\ 4 & 8 & 15 \\ 5 & 11 & 20 \end{pmatrix}.$$

$$T^{k'n'} = T_{m'}^{n'} g^{m'k'} \iff T^{k'n'} = \begin{pmatrix} 0 & -2 & -3 \\ 1 & 2 & 4 \\ 1 & 3 & 5 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 & -1 \\ 0 & -1 & 2 \\ -1 & 0 & 2 \end{pmatrix}.$$

Quisiéramos ejemplificar una forma “rápida y furiosa” (pero sucia) de calcular la métrica generada por una determinada base genérica. La idea es que, violentando toda nuestra notación e idea de tensores, construimos la métrica a partir de los vectores base definiéndola como $g_{i'j'} = \langle w_{i'} | w_{i'} \rangle$, de esta manera:

$$\begin{aligned} \tilde{g}_{ij} &= \begin{pmatrix} \langle w_1' | w_1' \rangle & \langle w_1' | w_2' \rangle & \langle w_1' | w_3' \rangle \\ \langle w_2' | w_1' \rangle & \langle w_2' | w_2' \rangle & \langle w_2' | w_3' \rangle \\ \langle w_3' | w_1' \rangle & \langle w_3' | w_2' \rangle & \langle w_3' | w_3' \rangle \end{pmatrix} \\ &= \begin{pmatrix} \langle i | i \rangle & \langle i | [i] + [j] \rangle & \langle i | [i] + [j] + [k] \rangle \\ [\langle i \rangle + \langle j \rangle] | i \rangle & [\langle i \rangle + \langle j \rangle] | i \rangle + [j] \rangle & [\langle i \rangle + \langle j \rangle] | i \rangle + [j] + [k] \rangle \\ [\langle i \rangle + \langle j \rangle + \langle k \rangle] | i \rangle & [\langle i \rangle + \langle j \rangle + \langle k \rangle] | i \rangle + [j] \rangle & [\langle i \rangle + \langle j \rangle + \langle k \rangle] | i \rangle + [j] + [k] \rangle \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}. \end{aligned}$$

Resultado que coincide con la métrica que calculamos transformando el tensor métrico.



A continuación vamos a considerar otros ejemplos con tensores cartesianos.

Ejemplo 3.9 Dado los tensores:

$$T_{ij} = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 1 \\ 1 & 3 & 4 \end{pmatrix}, \quad E^i = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}.$$

⁷Recordemos que hemos insistido que las matrices representan tensores mixtos

La parte simétrica de T_{ij} es:

$$S_{ij} = \frac{1}{2} [T_{ij} + T_{ji}] = \frac{1}{2} \left[\begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 1 \\ 1 & 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 3 & 1 \\ 0 & 4 & 3 \\ 2 & 1 & 4 \end{pmatrix} \right] = \frac{1}{2} \begin{pmatrix} 2 & 3 & 3 \\ 3 & 8 & 4 \\ 3 & 4 & 8 \end{pmatrix}.$$

Para la parte antisimétrica de T_{ij} :

$$A_{ij} = \frac{1}{2} [T_{ij} - T_{ji}] = \frac{1}{2} \left[\begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 1 \\ 1 & 3 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 3 & 1 \\ 0 & 4 & 3 \\ 2 & 1 & 4 \end{pmatrix} \right] = \frac{1}{2} \begin{pmatrix} 0 & -3 & 1 \\ 3 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix}.$$

Por lo tanto:

$$T_{ij} = S_{ij} + A_{ij} = \frac{1}{2} \left[\begin{pmatrix} 2 & 3 & 3 \\ 3 & 8 & 4 \\ 3 & 4 & 8 \end{pmatrix} + \begin{pmatrix} 0 & -3 & 1 \\ 3 & 0 & -2 \\ -1 & 2 & 0 \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 4 & 1 \\ 1 & 3 & 4 \end{pmatrix}.$$

Para calcular $T_{ij}E^j$ hacemos lo siguiente:

$$C_i = T_{i1}E^1 + T_{i2}E^2 + T_{i3}E^3 = \begin{cases} C_1 = T_{11}E^1 + T_{12}E^2 + T_{13}E^3 = (1)(1) + (0)(4) + (2)(3) = 7 \\ C_2 = T_{21}E^1 + T_{22}E^2 + T_{23}E^3 = (3)(1) + (4)(4) + (1)(3) = 22 \\ C_3 = T_{31}E^1 + T_{32}E^2 + T_{33}E^3 = (1)(1) + (3)(4) + (4)(3) = 25 \end{cases}$$

◀

Ejemplo 3.10 Dados T_j^i , a^i y $b^i \in \mathbb{R}^3$ con:

$$T_j^i = \begin{pmatrix} 1 & 0 & 3 \\ 3 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix} \quad \text{y} \quad \mathbf{b} = \begin{pmatrix} -2 \\ 5 \\ 4 \end{pmatrix}.$$

Supongamos que T_j^i , \mathbf{a} y \mathbf{b} están expresados en coordenadas cartesianas.

Vamos a calcular: $A_{ij}a^i b^j$, $S_{ij}a^i a^j$ y $A_{ij}b^i b^j$, donde S_{ij} y A_{kl} son, respectivamente, las partes simétrica y antisimétrica del tensor T_j^i .

La métrica en coordenadas cartesianas viene dada por:

$$g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Tendremos que:

$$S_{ij} = \frac{1}{2} (T_{ij} + T_{ji}) = \frac{1}{2} \left(\begin{pmatrix} 1 & 0 & 3 \\ 3 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 3 & 3 \\ 0 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} \right) = \frac{1}{2} \begin{pmatrix} 2 & 3 & 6 \\ 3 & 8 & 2 \\ 6 & 2 & 8 \end{pmatrix}.$$

Nótese que la expresión matricial para $T_{ij} \equiv g_{ik}T_j^k$ es la misma que para T_j^i debido a la forma de la métrica en este espacio. Del mismo modo:

$$A_{ij} = \frac{1}{2} (T_{ij} - T_{ji}) = \frac{1}{2} \left(\begin{pmatrix} 1 & 0 & 3 \\ 3 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} - \begin{pmatrix} 1 & 3 & 3 \\ 0 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} \right) = \frac{1}{2} \begin{pmatrix} 0 & -3 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

por lo tanto:

$$a^i A_{ij} b^j = \frac{1}{2} \begin{pmatrix} 5 & 2 & -5 \end{pmatrix} \begin{pmatrix} 0 & -3 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -2 \\ 5 \\ 4 \end{pmatrix} = -\frac{87}{2},$$

$$a^i S_{ij} a^j = \frac{1}{2} \begin{pmatrix} 5 & 2 & -5 \end{pmatrix} \begin{pmatrix} 2 & 3 & 6 \\ 3 & 8 & 2 \\ 6 & 2 & 8 \end{pmatrix} \begin{pmatrix} 5 \\ 2 \\ -5 \end{pmatrix} = 1,$$

$$b^i A_{ij} b^j = 0.$$

△

Ejemplo 3.11 Sistema de coordenadas genérico.

Dado un sistema genérico de coordenadas oblicuas:

$$|w_1\rangle = a|i\rangle + b|j\rangle, \quad |w_2\rangle = c|i\rangle + d|j\rangle.$$

- Para una base arbitraria, $\{|w_i\rangle\}$, la métrica viene definida por:

$$g_{ij} \equiv g_{ji} = \mathbf{g}[|w_i\rangle, |w_j\rangle] \equiv \langle w^i | w_j \rangle$$

$$g_{ij} = \begin{pmatrix} \langle w^1 | w_1 \rangle & \langle w^1 | w_2 \rangle \\ \langle w^2 | w_1 \rangle & \langle w^2 | w_2 \rangle \end{pmatrix} = \begin{pmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{pmatrix}.$$

- Un vector genérico: $|v\rangle = v_x|i\rangle + v_y|j\rangle$, en estas coordenadas se escribe como:

$$\left. \begin{array}{l} |w_1\rangle = a|i\rangle + b|j\rangle \\ |w_2\rangle = c|i\rangle + d|j\rangle \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} |i\rangle = \frac{1}{\Delta} (d|w_1\rangle - b|w_2\rangle) \\ |j\rangle = \frac{1}{\Delta} (-c|w_1\rangle + a|w_2\rangle) \end{array} \right.$$

con $\Delta = ad - bc$, por lo cual:

$$|v\rangle = v_x|i\rangle + v_y|j\rangle = \frac{v_x}{\Delta} (d|w_1\rangle - b|w_2\rangle) + \frac{v_y}{\Delta} (-c|w_1\rangle + a|w_2\rangle)$$

$$= \left(d\frac{v_x}{\Delta} - c\frac{v_y}{\Delta} \right) |w_1\rangle - \left(b\frac{v_x}{\Delta} - a\frac{v_y}{\Delta} \right) |w_2\rangle.$$

- Consideremos ahora una base y un tensor de manera concreta.

$$|w_1\rangle = |i\rangle, \quad |w_2\rangle = \frac{\sqrt{2}}{2}|i\rangle + \frac{\sqrt{2}}{2}|j\rangle, \quad T_j^i = \begin{pmatrix} 4 & 2 \\ 1 & 4 \end{pmatrix}.$$

Encontremos primero la expresión matricial para el tensor \tilde{T}_{ij} ⁸.

En general:

$$T_{i'j'} = g_{ik} T_{j'}^{k'} = g_{ik} \frac{\partial x^{k'}}{\partial x^m} T_n^m \frac{\partial x^n}{\partial x^{j'}},$$

Identificando:

$$v_{x'} = v^{1'} = \frac{\partial x^{1'}}{\partial x^j} v^j = \frac{\partial x^{1'}}{\partial x^1} v^1 + \frac{\partial x^{1'}}{\partial x^2} v^2 = \frac{d}{\Delta} v_x - \frac{c}{\Delta} v_y,$$

$$v_{y'} = v^{2'} = \frac{\partial x^{2'}}{\partial x^j} v^j = \frac{\partial x^{2'}}{\partial x^1} v^1 + \frac{\partial x^{2'}}{\partial x^2} v^2 = -\frac{b}{\Delta} v_x + \frac{a}{\Delta} v_y,$$

⁸Ayuda: dada una matriz genérica $A_j^i = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, su inversa será $\begin{pmatrix} \frac{D}{AD-BC} & -\frac{B}{AD-BC} \\ -\frac{C}{AD-BC} & \frac{A}{AD-BC} \end{pmatrix}$.

donde:

$$\frac{\partial x^{1'}}{\partial x^1} = \frac{d}{\Delta}, \frac{\partial x^{1'}}{\partial x^2} = -\frac{c}{\Delta}, \frac{\partial x^{2'}}{\partial x^1} = -\frac{b}{\Delta} \text{ y } \frac{\partial x^{2'}}{\partial x^2} = \frac{a}{\Delta}.$$

Como $a = 1, b = 0, c = \frac{\sqrt{2}}{2}, d = \frac{\sqrt{2}}{2}$, entonces:

$$g_{ik} \Rightarrow \begin{pmatrix} 1 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 1 \end{pmatrix}, \quad \frac{\partial x^{k'}}{\partial x^m} \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & \sqrt{2} \end{pmatrix}, \quad \frac{\partial x^n}{\partial x^{j'}} = \left(\frac{\partial x^{n'}}{\partial x^j} \right)^{-1} \Rightarrow \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Finalmente:

$$T_{i'j'} = \begin{pmatrix} 1 & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 4 & 3\sqrt{2} \\ \frac{5}{\sqrt{2}} & \frac{11}{2} \end{pmatrix}.$$

△

Ejemplo 3.12 Cartesianas y polares, otra vez

El ejemplo más simple, y por ello clásico y emblemático de cambio de coordenadas, lo constituye las expresiones de un mismo vector en dos sistemas de coordenadas en el plano: Cartesianas $\{|i\rangle, |j\rangle\}$ y polares $\{|u_r\rangle, |u_\theta\rangle\}$. Esto es, para un vector $|a\rangle$:

$$|a\rangle = a_x|i\rangle + a_x|j\rangle = a^1|\hat{e}_1\rangle + a^2|\hat{e}_2\rangle \quad \text{y} \quad |a\rangle = a_r|u_r\rangle + a_\theta|u_\theta\rangle = a^{1'}|\hat{e}_{1'}\rangle + a^{2'}|\hat{e}_{2'}\rangle.$$

Al expresar una base en términos de la otra obtenemos:

$$|u_r\rangle = \cos(\theta)|i\rangle + \sin(\theta)|j\rangle \quad \text{y} \quad |u_\theta\rangle = -\sin(\theta)|i\rangle + \cos(\theta)|j\rangle,$$

con lo cual:

$$\langle \hat{e}^{i'} | e_j \rangle = A_j^{i'} \iff A_j^{i'} = \begin{pmatrix} \langle u_r | i \rangle & \langle u_r | j \rangle \\ \langle u_\theta | i \rangle & \langle u_\theta | j \rangle \end{pmatrix} \equiv \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Para la transformación inversa se tiene:

$$\langle e^i | \hat{e}_{j'} \rangle = A_{j'}^i \iff A_{j'}^i = \begin{pmatrix} \langle i | u_r \rangle & \langle i | u_\theta \rangle \\ \langle j | u_r \rangle & \langle j | u_\theta \rangle \end{pmatrix} \equiv \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

y por lo tanto:

$$|i\rangle = \cos(\theta)|u_r\rangle - \sin(\theta)|u_\theta\rangle \quad \text{y} \quad |j\rangle = \sin(\theta)|u_r\rangle + \cos(\theta)|u_\theta\rangle.$$

Cumpliendo además con:

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \iff A_{k'}^i A_j^{k'} = \delta_j^i.$$

Por otro lado, si:

$$|a\rangle = a_r|u_r\rangle + a_\theta|u_\theta\rangle \equiv a^{1'}|\hat{e}_{1'}\rangle + a^{2'}|\hat{e}_{2'}\rangle = a_x|i\rangle + a_x|j\rangle \equiv a^1|e_1\rangle + a^2|e_2\rangle,$$

tendremos que:

$$a^{i'} = A_j^{i'} a^j \iff \begin{pmatrix} a_r \\ a_\theta \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} a_x \\ a_y \end{pmatrix} = \begin{pmatrix} a_x \cos(\theta) + a_y \sin(\theta) \\ -a_x \sin(\theta) + a_y \cos(\theta) \end{pmatrix},$$

con lo cual:

$$a_r = a_x \cos(\theta) + a_y \sin(\theta) \quad \text{y} \quad a_\theta = -a_x \sin(\theta) + a_y \cos(\theta).$$

Del mismo modo:

$$a^i = A_{j'}^i a^{j'} \iff \begin{pmatrix} a_x \\ a_y \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} a_r \\ a_\theta \end{pmatrix} = \begin{pmatrix} a_r \cos(\theta) - a_\theta \sin(\theta) \\ a_r \sin(\theta) + a_\theta \cos(\theta) \end{pmatrix}$$

y

$$a_x = a_r \cos(\theta) - a_\theta \sin(\theta) \quad \text{y} \quad a_y = a_r \sin(\theta) + a_\theta \cos(\theta).$$



En el ejemplo anterior consideramos los cambios de base entre cartesianas y polares. A partir de cambios de la base construimos la matriz de transformación entre sus componentes. Ahora pondremos la atención en las componentes. En general, podemos pensar que las componentes de los vectores pueden ser funciones de las otras. Consideremos el ejemplo anterior con esta visión para estudiarlo con más detalle en dos y tres dimensiones.

Ejemplo 3.13 Caso bidimensional

Tendremos que un punto en el plano viene representado en coordenadas cartesianas por dos números (x, y) y en coordenadas polares por otros dos números (r, θ) . Siguiendo el ejemplo anterior un punto P , en el plano lo describimos como:

$$|P\rangle = r_P |u_r\rangle = x_P |i\rangle + y_P |j\rangle .$$

Veamos como están relacionadas estas dos descripciones, para este caso en el cual las ecuaciones de transformación son:

$$\begin{aligned} x_P &= x_P(r, \theta) = x^1 = x^1(x^{1'}, x^{2'}) \\ y_P &= y_P(r, \theta) = x^2 = x^2(x^{1'}, x^{2'}) \end{aligned} \quad \iff \quad \begin{cases} r_P = r_P(x, y) = x^{1'} = x^1(x^1, x^2) \\ \theta = \theta_P(x, y) = x^{2'} = x^2(x^1, x^2) \end{cases}$$

y explícitamente:

$$\begin{aligned} x_P = r_P \cos(\theta) &\Rightarrow x^1 = x^{1'} \cos(x^{2'}) \\ y_P = r_P \sin(\theta) &\Rightarrow x^2 = x^{1'} \sin(x^{2'}) \end{aligned} \quad \text{y} \quad \begin{cases} r_P = \sqrt{x_P^2 + y_P^2} \Rightarrow x^{1'} = \sqrt{(x^1)^2 + (x^2)^2} \\ \theta = \arctan\left(\frac{y_P}{x_P}\right) \Rightarrow x^{2'} = \arctan\left(\frac{x^2}{x^1}\right) \end{cases}$$

Es claro que ambas coordenadas están relacionadas y que se puede invertir la relación.

$$\begin{cases} x^{1'} = x^{1'}(x^1, x^2) \\ x^{2'} = x^{2'}(x^1, x^2) \end{cases} \quad \iff \quad \begin{cases} x^1 = x^1(x^{1'}, x^{2'}) \\ x^2 = x^2(x^{1'}, x^{2'}) \end{cases} .$$

Se debe pedir, eso si, dos cosas razonables:

1. que las funciones $x^i = x^i(x^{m'})$ y $x^{j'} = x^{j'}(x^m)$ sean al menos \mathcal{C}^2 (función y derivada continua)
2. que el determinante de la matriz Jacobiana sea finito,

$$\det \left| \frac{\partial x^i(x^{k'})}{\partial x^{j'}} \right| \neq 0 .$$

Más aún, si:

$$x^i = x^i(x^{j'}(x^k)) \Rightarrow \frac{\partial x^i}{\partial x^k} = \frac{\partial x^i}{\partial x^{j'}} \frac{\partial x^{j'}}{\partial x^k} = \delta_k^i \Rightarrow dx^i = \frac{\partial x^i}{\partial x^{j'}} dx^{j'} ,$$

con lo cual intuimos dos cosas:

1. que las componentes de un vector, $|P\rangle$, deben transformar bajo un cambio de coordenadas como:

$$p^i = \frac{\partial x^i(x^{k'})}{\partial x^{j'}} p^{j'} .$$

2. Las matrices jacobianas $\frac{\partial x^i}{\partial x^{k'}}$ y $\frac{\partial x^{i'}}{\partial x^k}$ son una la inversa de la otra.

Veamos si es cierto para el caso de vectores en el plano. Para ello calculamos la matriz jacobiana:

$$\frac{\partial x^i(x^{1'}, x^{2'})}{\partial x^{j'}} \Rightarrow \begin{pmatrix} \frac{\partial x^1(x^{1'}, x^{2'})}{\partial x^{1'}} & \frac{\partial x^1(x^{1'}, x^{2'})}{\partial x^{2'}} \\ \frac{\partial x^2(x^{1'}, x^{2'})}{\partial x^{1'}} & \frac{\partial x^2(x^{1'}, x^{2'})}{\partial x^{2'}} \end{pmatrix} = \begin{pmatrix} \cos(x^{2'}) & -x^{1'} \sin(x^{2'}) \\ \sin(x^{2'}) & x^{1'} \cos(x^{2'}) \end{pmatrix} ,$$

y seguidamente, identificando $|P\rangle = x_P|i\rangle + y_P|j\rangle = r_P|u_r\rangle$

$$p^i = \frac{\partial x^i(x^{1'}, x^{2'})}{\partial x^{j'}} p^{j'} \Rightarrow \begin{pmatrix} p^1 \\ p^2 \end{pmatrix} = \begin{pmatrix} \cos(x^{2'}) & -x^{1'} \sin(x^{2'}) \\ \sin(x^{2'}) & x^{1'} \cos(x^{2'}) \end{pmatrix} \begin{pmatrix} \tilde{p}^1 \\ 0 \end{pmatrix}.$$

Igualmente, si calculamos la inversa de la matriz jacobiana:

$$\left(\frac{\partial x^i(x^{1'}, x^{2'})}{\partial x^{j'}} \right)^{-1} = \begin{pmatrix} \cos(x^{2'}) & \sin(x^{2'}) \\ \frac{-\sin(x^{2'})}{x^{1'}} & \frac{\cos(x^{2'})}{x^{1'}} \end{pmatrix} = \begin{pmatrix} \frac{x^1}{\sqrt{(x^1)^2 + (x^2)^2}} & \frac{x^2}{\sqrt{(x^1)^2 + (x^2)^2}} \\ \frac{-x^2}{(x^1)^2 + (x^2)^2} & \frac{x^1}{(x^1)^2 + (x^2)^2} \end{pmatrix},$$

tendremos:

$$\begin{pmatrix} p^{1'} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{x^1}{\sqrt{(x^1)^2 + (x^2)^2}} & \frac{x^2}{\sqrt{(x^1)^2 + (x^2)^2}} \\ \frac{-x^2}{(x^1)^2 + (x^2)^2} & \frac{x^1}{(x^1)^2 + (x^2)^2} \end{pmatrix} \begin{pmatrix} p^1 \\ p^2 \end{pmatrix} \Rightarrow p^{i'} = \frac{\partial x^{i'}(x^1, x^2)}{\partial x^j} p^j.$$

◇

Ejemplo 3.14 Caso tridimensional

Consideremos nuevamente dos sistemas de coordenadas: uno cartesiano ($x^1 = x, x^2 = y, x^3 = z$) y otro esféricico ($x^{1'} = r, x^{2'} = \theta, x^{3'} = \phi$).

Tal y como hemos supuesto anteriormente el punto P vendrá descrito por:

$$|P\rangle = r_P|u_r\rangle = x_P|i\rangle + y_P|j\rangle + z_P|k\rangle,$$

de nuevo

$$\left. \begin{array}{l} x = x(r, \theta, \phi) = x^1 = x^1(x^{1'}, x^{2'}, x^{3'}) \\ y = y(r, \theta, \phi) = x^2 = x^2(x^{1'}, x^{2'}, x^{3'}) \\ z = z(r, \theta, \phi) = x^3 = x^3(x^{1'}, x^{2'}, x^{3'}) \end{array} \right\} \iff \left. \begin{array}{l} r = r(x, y, z) = x^{1'} = x^{1'}(x^1, x^2, x^3) \\ \theta = \theta(x, y, z) = x^{2'} = x^{2'}(x^1, x^2, x^3) \\ \phi = \phi(x, y, z) = x^{3'} = x^{3'}(x^1, x^2, x^3) \end{array} \right\}$$

Las ecuaciones de transformación serán:

$$\left\{ \begin{array}{l} x_P = r_P \sin(\theta) \cos(\phi) \Rightarrow x^1 = x^{1'} \sin(x^{2'}) \cos(x^{3'}) \\ y_P = r_P \sin(\theta) \sin(\phi) \Rightarrow x^2 = x^{1'} \sin(x^{2'}) \sin(x^{3'}) \\ z_P = r_P \cos(\theta) \Rightarrow x^3 = x^{1'} \cos(x^{2'}) \end{array} \right.$$

y las transformaciones inversas:

$$\left\{ \begin{array}{l} r_P = \sqrt{x_P^2 + y_P^2 + z_P^2} \Rightarrow x^{1'} = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2} \\ \phi = \arctan\left(\frac{y_P}{x_P}\right) \Rightarrow x^{2'} = \arctan\left(\frac{x^2}{x^1}\right) \\ \theta = \arctan\left(\frac{\sqrt{x_P^2 + y_P^2}}{z_P}\right) \Rightarrow x^{3'} = \arctan\left(\frac{\sqrt{(x^1)^2 + (x^2)^2}}{x^3}\right) \end{array} \right.$$

Con lo cual la matriz de las derivadas para esta transformación en particular será:

$$\frac{\partial x^i(x^{1'}, x^{2'}, x^{3'})}{\partial x^{j'}} = \begin{pmatrix} \text{sen}(\theta) \cos(\phi) & -r \text{sen}(\theta) \text{sen}(\phi) & r \cos(\theta) \cos(\phi) \\ \text{sen}(\theta) \text{sen}(\phi) & r \text{sen}(\theta) \cos(\phi) & r \cos(\theta) \text{sen}(\phi) \\ \cos(\theta) & 0 & -r \text{sen}(\theta) \\ \text{sen}(x^{2'}) \cos(x^{3'}) & -\tilde{x}^1 \text{sen}(x^{2'}) \text{sen}(x^{3'}) & \tilde{x}^1 \cos(x^{2'}) \cos(x^{3'}) \\ \text{sen}(x^{2'}) \text{sen}(x^{3'}) & x^{1'} \text{sen}(x^{2'}) \cos(x^{3'}) & x^{1'} \cos(x^{2'}) \text{sen}(x^{3'}) \\ \cos(x^{2'}) & 0 & -x^{1'} \text{sen}(x^{2'}) \end{pmatrix}.$$

Mientras que la inversa es:

$$\frac{\partial x^{i'}(x^1, x^2, x^3)}{\partial x^j} = \begin{pmatrix} \text{sen}(\theta) \cos(\phi) & \text{sen}(\theta) \text{sen}(\phi) & \cos(\theta) \\ -\frac{\text{sen}(\phi)}{r \text{sen}(\theta)} & \frac{\cos(\phi)}{r \text{sen}(\theta)} & 0 \\ \frac{\cos(\theta) \cos(\phi)}{r} & \frac{\cos(\theta) \text{sen}(\phi)}{r} & -\frac{\text{sen}(\theta)}{r} \\ \frac{x}{\sqrt{x^2+y^2+z^2}} & \frac{y}{\sqrt{x^2+y^2+z^2}} & \frac{z}{\sqrt{x^2+y^2+z^2}} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \\ \frac{xz}{(x^2+y^2+z^2)\sqrt{x^2+y^2}} & \frac{yz}{(x^2+y^2+z^2)\sqrt{x^2+y^2}} & \frac{-\sqrt{x^2+y^2}}{(x^2+y^2+z^2)} \end{pmatrix}.$$

Dejaremos al lector comprobar que, efectivamente,

$$p^i = \frac{\partial x^i(x^{1'}, x^{2'}, x^{3'})}{\partial x^{j'}} p^{j'} \iff p^{i'} = \frac{\partial x^{i'}(x^1, x^2, x^3)}{\partial x^j} p^j.$$

▷

3.3 Aplicaciones en Física: Esfuerzos, Inercia y Energía libre de un medio elástico

En este punto vamos a ejemplificar la utilización de los tensores en varios ámbitos de la Física, en particular de la Mecánica. Para comenzar consideraremos el tensor de esfuerzos para describir las tensiones internas de cuerpos sometidos a fuerzas externas⁹. Haremos el análisis tanto para el caso de dos como de tres dimensiones. Luego a continuación consideraremos el tensor de inercia y su impacto en la dinámica de cuerpos en movimiento.

3.3.1 El tensor de esfuerzos o tensor stress (caso 2D)

Supongamos un cuerpo que se encuentra en equilibrio y está sometido a un conjunto de fuerzas externas. Para facilitar las cosas consideraremos el efecto de esas fuerzas sobre un plano que contiene a un determinado punto P (ver figura 3.2 cuadrante Ia). Es decir, vamos a considerar los efectos de las componentes de todas las fuerzas sobre ese plano y obviaremos el efecto del resto de las componentes.

⁹Una presentación más contemporánea del tensor de esfuerzos la pueden encontrar en De Prunelé, E. (2007). Linear strain tensor and differential geometry. American Journal of Physics, 75(10), 881-887.

Como observamos en la figura 3.2 Ib y Ic, si cortamos la superficie en dos líneas (AB y $A'B'$), podemos ver que el efecto del conjunto de fuerzas externas es distinto sobre P , en la dirección perpendicular a cada una de esas líneas. De hecho, al “cortar” la superficie las fuerzas que aparecen sobre las líneas AB (y $A'B'$) eran fuerzas internas y ahora son externas al nuevo cuerpo “cortado”. Así, estas fuerzas por unidad de longitud¹⁰ sobre el punto P existen como un conjunto de fuerzas que generan esfuerzos (*stress*). Por lo tanto, es claro que los esfuerzos sobre un punto dependen del punto, de las fuerzas externas y de la dirección del efecto.

Para irnos aclarando consideremos un elemento de área infinitesimal ds sobre la cual actúa un conjunto de fuerzas externas, las cuales podemos descomponer como normales y tangenciales a la línea sobre la cual están aplicadas (ver figura 3.2 cuadrante II). Es costumbre denotar los esfuerzos normales y tangenciales por σ y τ respectivamente.

$$dA = dx dy \Rightarrow \left\{ \begin{array}{l} Y_2 = \sigma_2 dx \\ Y_3 = \tau_3 dy \\ X_3 = \sigma_3 dy \\ \uparrow Y_4 = \sigma_4 dx \end{array} \right. \quad \begin{array}{l} \rightarrow X_2 = \tau_2 dx \\ dx \\ dy \quad ds \quad dy \\ dx \\ \rightarrow X_1 = \sigma_1 dy \end{array}$$

Consideramos la segunda ley de Newton aplicada a cada diferencial de masa dm y obtendremos:

$$\sum \mathbf{F}_i^{ext} = dm \mathbf{a} = 0 \Rightarrow \left\{ \begin{array}{l} \tau_1 dy + \sigma_2 dx + \tau_3 dy + \sigma_4 dx = 0 = (\sigma_2 + \sigma_4) dx + (\tau_1 + \tau_3) dy \\ \sigma_1 dy + \tau_2 dx + \sigma_3 dy + \tau_4 dx = 0 = (\tau_2 + \tau_4) dx + (\sigma_1 + \sigma_3) dy \end{array} \right.$$

con lo cual:

$$\sigma_2 = -\sigma_4, \quad \tau_1 = -\tau_3, \quad \tau_2 = -\tau_4, \quad \sigma_1 = -\sigma_3.$$

Como se trata de una situación en equilibrio, también la sumatoria de torques se tendrá que anular. Esto significa que:

$$\left. \begin{array}{l} (\tau_1 dy) \frac{dx}{2} - (\tau_2 dx) \frac{dy}{2} = 0 \\ (\tau_3 dy) \frac{dx}{2} - (\tau_4 dx) \frac{dy}{2} = 0 \end{array} \right\} \Rightarrow \tau_1 = \tau_2 = \tau_3 = \tau_4,$$

entonces, nos damos cuenta que existen sólo tres cantidades independientes: dos esfuerzos normales σ_1 y σ_2 ; y un esfuerzo tangencial τ_1 . Adicionalmente, notamos que los esfuerzos tienen que ver con la dirección de la fuerza y la superficie sobre la cual va aplicada. Con ello podemos diseñar la siguiente notación para los esfuerzos: σ_{ij} . El primer índice indica la dirección de la fuerza y el segundo la dirección de la normal de la superficie donde está aplicada. Así, tal y como muestra la figura (ver figura 3.2 cuadrante II)

$$\sigma_1 \equiv \sigma_{xx}, \quad -\sigma_4 \equiv \sigma_{yy}, \quad \tau_2 \equiv \sigma_{xy} \equiv \sigma_{yx}.$$

El cambio de signo se debe a lo incómodo de la notación: $\sigma_4 \equiv \sigma_{y-y}$ ya que la normal de lado 4 apunta en la dirección $-y$. Es importante también señalar que los esfuerzos en cualquier punto contenido en el diferencial de área $dA = dx dy$ deben ser considerados constantes, o lo que es lo mismo, que podemos hacer tender a cero el área del diferencial y con ello asociar los esfuerzos σ_{ij} a un punto P contenido en dA sobre la cual hemos calculado los esfuerzos.

En esta misma línea de razonamiento, nos podemos preguntar cuál es la expresión de los esfuerzos cuando se miden respecto a una superficie genérica, definida por un vector normal \mathbf{n} (ver figura 3.2 cuadrante III). Es

¹⁰En el caso tridimensional, las fuerzas que generan los esfuerzos serán definidas como fuerzas por unidad de área. Ese caso lo veremos en la próxima sección.

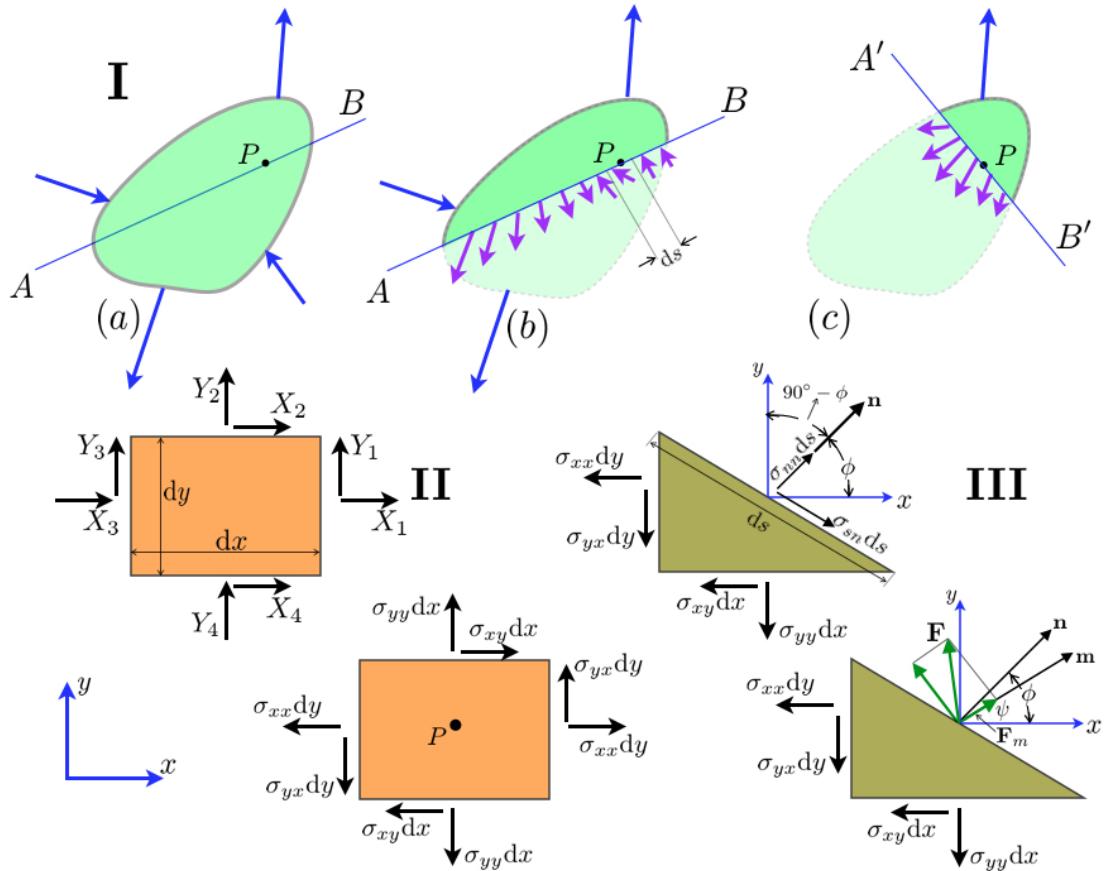


Figura 3.2: Tensor de Esfuerzos (stress) en 2 dimensiones

dicho, queremos conocer los esfuerzos medidos en el punto P y en la dirección \mathbf{n} , es decir, σ_{nn} .

Por lo tanto, tendremos que:

$$x \rightarrow \sigma_{xx}dy + \sigma_{xy}dx = \sigma_{nn}ds \cos(\phi) + \sigma_{sn}ds \sin(\phi)$$

$$y \rightarrow \sigma_{yy}dx + \sigma_{yx}dy = \sigma_{nn}ds \sin(\phi) - \sigma_{sn}ds \cos(\phi)$$

Ahora bien, dado que $dy = ds \cos(\phi)$ y $dx = ds \sin(\phi)$, entonces podemos expresar:

$$\sigma_{nn} = \sigma_{xx} \cos^2(\phi) + \sigma_{xy} \sin(\phi) \cos(\phi) + \sigma_{yx} \sin(\phi) \cos(\phi) + \sigma_{yy} \sin^2(\phi)$$

$$\sigma_{sn} = \sigma_{xx} \sin(\phi) \cos(\phi) + \sigma_{xy} \sin^2(\phi) - \sigma_{yx} \cos^2(\phi) - \sigma_{yy} \sin(\phi) \cos(\phi)$$

y si ahora nos damos cuenta que si construimos una matriz:

$$A_j^i = \begin{pmatrix} A_n^x & A_s^x \\ A_n^y & A_s^y \end{pmatrix} = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ \sin(\phi) & -\cos(\phi) \end{pmatrix},$$

entonces:

$$\sigma_{nn} = A_n^x A_n^x \sigma_{xx} + A_n^x A_n^y \sigma_{xy} + A_n^y A_n^x \sigma_{yx} + A_n^y A_n^y \sigma_{yy} \Rightarrow \sigma_{nn} = A_n^i A_n^j \sigma_{ij} \text{ con } i, j = x, y$$

$$\sigma_{sn} = A_s^x A_n^x \sigma_{xx} + A_s^x A_n^y \sigma_{xy} + A_s^y A_n^x \sigma_{yx} + A_s^y A_n^y \sigma_{yy} \Rightarrow \sigma_{sn} = A_s^i A_n^j \sigma_{ij} \text{ con } i, j = x, y$$

es decir:

$$\sigma_{kl} = A_k^i A_l^j \sigma_{ij}, \text{ con } i, j = x, y; k, l = n, s.$$

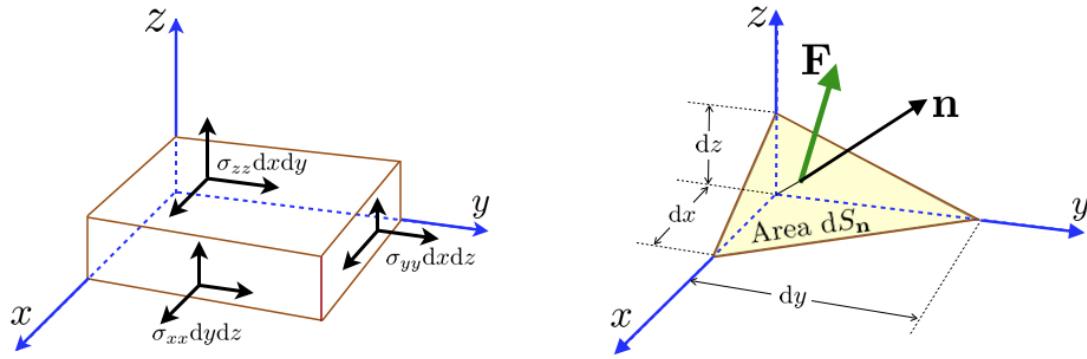


Figura 3.3: Tensor de Esfuerzos en 3 dimensiones

Como ya hemos mencionado, y veremos más adelante con más detalle, cualquier objeto que transforme como $\sigma_{kl} = A_k^i A_l^j \sigma_{ij}$ lo llamaremos tensor de segundo orden.

3.3.2 El tensor de esfuerzos o tensor stress (caso 3D)

Podemos proceder como en el caso anterior estableciendo las siguientes condiciones de equilibrio:

$$\sum \mathbf{F}_i^{ext} = 0 \quad \text{y} \quad \sum \boldsymbol{\tau}_i^{ext} = 0,$$

con ello construimos un volumen (cúbico) diferencial y construimos los esfuerzos normales y tangenciales, los cuales serán:

$$\sigma_{xx}dydz, \quad \sigma_{yy}dxdz, \quad \sigma_{zz}dxdy, \quad \sigma_{xz}dxdy, \quad \sigma_{yz}dxdy, \quad \sigma_{xy}dxdz.$$

Siguiendo el mismo proceso que involucra imponer el equilibrio, es fácil demostrar que al igual que el caso anterior, el tensor de esfuerzos σ_{ij} cumple con:

$$\sigma_{xz} = \sigma_{zx}, \quad \sigma_{yz} = \sigma_{zy}, \quad \sigma_{xy} = \sigma_{yx}.$$

Tendremos seis componentes (tres normales y tres tangenciales) independientes. Es decir, si bien el tensor de esfuerzos σ_{ij} viene representado por una matriz 3×3 y por lo tanto tiene 9 elementos, sólo seis son independientes.

Vayamos ahora el caso general para un tensor de esfuerzos en un medio elástico. Para ello construimos un tetraedro regular tal y como muestra la figura 3.3, y sobre su cara genérica asociada a un vector normal \mathbf{n} una fuerza \mathbf{F} :

$$\mathbf{F} = F^i \hat{\mathbf{i}}_i = F_x \hat{\mathbf{i}} + F_y \hat{\mathbf{j}} + F_z \hat{\mathbf{k}} \Rightarrow \left\{ \begin{array}{l} F_x = \sigma_{xn} dS_n \\ F_y = \sigma_{yn} dS_n \\ F_z = \sigma_{zn} dS_n \end{array} \right\} \Rightarrow F^i = \sigma_j^i n^j dS \Rightarrow \mathbf{F} = \boldsymbol{\sigma} \cdot d\mathbf{S}.$$

De esta manera se especifica como la fuerza actúa sobre un determinado elemento de superficie. Es claro que

la condición de equilibrio se traduce en:

$$\sum F_{xi} = 0 \Rightarrow \sigma_{xn} dS_n - \frac{1}{2} \sigma_{xx} dy dz - \frac{1}{2} \sigma_{xy} dx dz - \frac{1}{2} \sigma_{xz} dx dy = 0,$$

$$\sum F_{yi} = 0 \Rightarrow \sigma_{yn} dS_n - \frac{1}{2} \sigma_{yx} dy dz - \frac{1}{2} \sigma_{yy} dx dz - \frac{1}{2} \sigma_{yz} dx dy = 0,$$

$$\sum F_{zi} = 0 \Rightarrow \sigma_{zn} dS_n - \frac{1}{2} \sigma_{zx} dy dz - \frac{1}{2} \sigma_{zy} dx dz - \frac{1}{2} \sigma_{zz} dx dy = 0.$$

Si consideramos la proyección de dS_n sobre cada uno de los planos del sistema cartesiano tendremos:

$$\left. \begin{aligned} dS^n \cos(\mathbf{i}; \mathbf{n}) &= \frac{1}{2} dy dz = dS^n A_n^x \\ dS^n \cos(\mathbf{j}; \mathbf{n}) &= \frac{1}{2} dx dz = dS^n A_n^y \\ dS^n \cos(\mathbf{k}; \mathbf{n}) &= \frac{1}{2} dx dy = dS^n A_n^z \end{aligned} \right\} \Rightarrow \sigma_{xn} = \sigma_{xx} A_n^x + \sigma_{xy} A_n^y + \sigma_{xz} A_n^z,$$

y equivalentemente:

$$\sigma_{yn} = \sigma_{yx} A_n^x + \sigma_{yy} A_n^y + \sigma_{yz} A_n^z \quad y \quad \sigma_{zn} = \sigma_{zx} A_n^x + \sigma_{zy} A_n^y + \sigma_{zz} A_n^z,$$

las cuales se conocen como las relaciones de Cauchy. Estas relaciones representan los esfuerzos sobre la superficie con normal \mathbf{n} .

Ahora bien, dado que: $\mathbf{F} = \boldsymbol{\sigma} dS$ es una relación vectorial podemos proyectar en la dirección $\hat{\mathbf{u}}_m$:

$$\hat{\mathbf{u}}_m \cdot \mathbf{F} = \hat{\mathbf{u}}_m \cdot (\boldsymbol{\sigma} dS) \Rightarrow F^m = \sigma_n^m dS^n = (\sigma_i^m A_n^i) dS^n = (\sigma_i^m A_n^i) dS^n,$$

por lo tanto:

$$\sigma_{mn} dS^n = (\sigma_{mi} A_n^i) dS^n \Rightarrow \sigma_{mnd} dS^n = (\sigma_{ki} A_m^k A_n^i) dS^n, \quad \text{con } i, j = x, y, z.$$

Una vez más, podemos ver que transforma como un tensor.

3.3.3 El Tensor de inercia

Consideremos el caso de un sistema de n partículas. La cantidad de movimiento angular para este sistema vendrá dada por:

$$\mathbf{L} = \sum_i m_{(i)} (\mathbf{r}_{(i)} \times \mathbf{v}_{(i)}) ,$$

donde hemos indicado que la i -ésima partícula que está en la posición $\mathbf{r}_{(i)}$ tiene una velocidad $\mathbf{v}_{(i)}$.

Si las distancias entre las partículas, y entre las partículas con el origen de coordenadas, son constantes, podemos expresar la velocidad de cada una de ellas como:

$$\mathbf{v}_{(i)} = \boldsymbol{\omega} \times \mathbf{r}_{(i)}, \quad ? \text{Por qué?}$$

Donde $\boldsymbol{\omega}$ es la velocidad angular instantánea del sistema. Entonces tendremos que:

$$\mathbf{L} = \sum_i m_{(i)} [\mathbf{r}_{(i)} \times (\boldsymbol{\omega} \times \mathbf{r}_{(i)})] = \sum_i m_{(i)} [\boldsymbol{\omega} (\mathbf{r}_{(i)} \cdot \mathbf{r}_{(i)}) - \mathbf{r}_{(i)} (\boldsymbol{\omega} \cdot \mathbf{r}_{(i)})] ,$$

y para cada partícula se cumple que las componentes de la cantidad de movimiento angular serán:

$$L^k = \sum_i m_{(i)} \left[\omega^k (r_{(i)}^m r_{(i)m}) - r_{(i)}^k (\omega^m r_{(i)m}) \right] .$$

Si vemos que $\omega_{(i)}^k = \delta_l^k \omega_{(i)}^l$ entonces:

$$L^k = \sum_i m_{(i)} \left[\delta_l^k \omega^l \left(r_{(i)}^m r_{(i)m} \right) - r_{(i)}^k \left(\omega^m r_{(i)m} \right) \right] = \omega_{(i)}^l \underbrace{\left[\sum_i m_{(i)} \left(\delta_l^k \left(r_{(i)}^m r_{(i)m} \right) - r_{(i)}^k \left(r_{(i)l} \right) \right) \right]}_{I_l^k}$$

es decir:

$$L^k = \omega_{(i)}^l I_l^k, \quad \text{donde: } I_l^k = \sum_i m_{(i)} \left(\delta_l^k \left(r_{(i)}^m r_{(i)m} \right) - r_{(i)}^k \left(r_{(i)l} \right) \right).$$

El objeto I_l^k se conoce como el tensor de inercia y corresponde a nueve cantidades (sólo seis son independientes porque es un tensor simétrico). En coordenadas cartesianas se verá de la siguiente forma:

$$I_l^k = \begin{pmatrix} I_x^x & I_x^y & I_x^z \\ I_y^x & I_y^y & I_y^z \\ I_z^x & I_z^y & I_z^z \end{pmatrix} = \begin{pmatrix} \sum_i m_{(i)} \left(y_{(i)}^2 + z_{(i)}^2 \right) & -\sum_i m_{(i)} \left(x_{(i)} y_{(i)} \right) & -\sum_i m_{(i)} \left(x_{(i)} z_{(i)} \right) \\ -\sum_i m_{(i)} \left(x_{(i)} y_{(i)} \right) & \sum_i m_{(i)} \left(x_{(i)}^2 + z_{(i)}^2 \right) & -\sum_i m_{(i)} \left(y_{(i)} z_{(i)} \right) \\ -\sum_i m_{(i)} \left(x_{(i)} z_{(i)} \right) & -\sum_i m_{(i)} \left(y_{(i)} z_{(i)} \right) & \sum_i m_{(i)} \left(x_{(i)}^2 + y_{(i)}^2 \right) \end{pmatrix}$$

Por ahora, nos contentaremos con esta construcción y en los ejercicios le propondremos que muestre que es un tensor.

3.3.4 Energía libre para un medio elástico

Consideremos el siguiente par de tensores provenientes de la teoría de elasticidad:

$$u_{ik} = \frac{1}{2} (\partial_k u_i + \partial_i u_k) \equiv \frac{1}{2} \left(\frac{\partial u_i}{\partial x^k} + \frac{\partial u_k}{\partial x^i} \right), \quad (u_k^i)^0 = u_k^i - \frac{1}{3} u_m^m \delta_k^i,$$

y construyamos el tensor de esfuerzos como: $p_j^i = 2\lambda (u_j^i)^0 + K u_l^l \delta_j^i$.

Calculemos la energía libre para el medio elástico, definida como: $F = \frac{1}{2} p_j^i u_i^j$.

Tenemos que:

$$p_j^i = 2\lambda (u_j^i)^0 + K u_l^l \delta_j^i = 2\lambda \left(u_j^i - \frac{1}{3} u_m^m \delta_j^i \right) + K u_l^l \delta_j^i = 2\lambda u_j^i + u_l^l \delta_j^i \left(K - \frac{2\lambda}{3} \right),$$

donde: $u_m^m = \frac{1}{2} (\partial^m u_m + \partial_m u^m) = \partial^m u_m$, con lo cual:

$$\begin{aligned} F &= \frac{1}{2} p_j^i u_i^j = \frac{1}{2} \left(2\lambda u_j^i + u_l^l \delta_j^i \left(K - \frac{2\lambda}{3} \right) \right) u_i^j = \left(\lambda u_j^i u_i^j + u_l^l \delta_j^i u_i^j \left(\frac{1}{2} K - \frac{\lambda}{3} \right) \right) \\ &= \lambda (u_1^i u_i^1 + u_2^i u_i^2 + u_3^i u_i^3) + \left(\frac{1}{2} K - \frac{\lambda}{3} \right) (u_l^l)^2 \\ &= \lambda ((u_1^1 u_1^1 + u_1^2 u_1^2 + u_1^3 u_1^3) + (u_2^1 u_2^1 + u_2^2 u_2^2 + u_2^3 u_2^3) + (u_3^1 u_3^1 + u_3^2 u_3^2 + u_3^3 u_3^3)) + \left(\frac{1}{2} K - \frac{\lambda}{3} \right) (u_l^l)^2 \\ &= \lambda \left((u_1^1)^2 + (u_2^2)^2 + (u_3^3)^2 + 2(u_2^1 u_1^2 + u_3^1 u_2^3 + u_1^3 u_3^1) \right) + \left(\frac{1}{2} K - \frac{\lambda}{3} \right) (u_l^l)^2 \\ &= \left(\frac{1}{2} K + \frac{2\lambda}{3} \right) (u_l^l)^2 + 2\lambda (u_2^1 u_1^2 + u_3^1 u_2^3 + u_1^3 u_3^1) \\ &= \left(\frac{1}{2} K + \frac{2\lambda}{3} \right) (\partial_x u_x + \partial_y u_y + \partial_z u_z)^2 + 2\lambda (\partial_x u_y \partial_y u_x + \partial_y u_z \partial_z u_y + \partial_z u_x \partial_x u_z). \end{aligned}$$

3.3.5 Ejercicios

1. Si A_{ijk} es un tensor covariante de orden 3 y B^{lmnp} un tensor contravariante de orden 4, pruebe que $A_{ijk}B^{jkn}$ es un tensor mixto de orden 3.

2. Dados los tensores:

$$R_j^i = \begin{pmatrix} 1/2 & 1 & 3/2 \\ 2 & 5/2 & 3 \\ 7/2 & 4 & 9/2 \end{pmatrix}, \quad T^i = \begin{pmatrix} 1/3 \\ 2/3 \\ 1 \end{pmatrix}, \quad g^{ij} = g_{ij} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Encuentre:

(a). La parte simétrica S_j^i y antisimétrica A_j^i de R_j^i .

(b). $R_{kj} = g_{ik}R_j^i$, $R^{ki} = g^{jk}R_j^i$, $T_j = g_{ij}T^i$ ¿Qué se concluye de estos cálculos?

(c). $R_j^i T_i$, $R_j^i T^j$, $R_j^i T_i T^j$.

(d). $R_j^i S_i^j$, $R_j^i A_i^j$, $A_i^j T^i$, $A_i^j T^i T_j$.

(e). $R_j^i - 2\delta_j^i R_l^l$, $(R_j^i - 2\delta_j^i R_l^l)T_i$, $(R_j^i - 2\delta_j^i R_l^l)T_i T^j$.

3. Demuestre que si S_j^i representa un tensor simétrico y A_j^i uno antisimétrico, entonces, $S_j^i A_i^j = 0$.

4. Dado F_{ijk} un tensor totalmente antisimétrico respecto a sus índices ijk , demuestre que el rotor de F_{ijk} definido como sigue, también será un tensor.

$$\text{rot}[F_{ijk}] = \partial_m F_{ijk} - \partial_i F_{jkm} + \partial_j F_{kmi} - \partial_k F_{mij} \equiv \frac{\partial F_{ijk}}{\partial x^m} - \frac{\partial F_{jkm}}{\partial x^i} + \frac{\partial F_{kmi}}{\partial x^j} - \frac{\partial F_{mij}}{\partial x^k}.$$

5. Consideremos un tensor **B** con componentes B_i^j , un tensor **A** con componentes A_i y el producto:

$$B_i^j A_j = C_i.$$

Esta última expresión se puede interpretar como la acción de **B** sobre **A** que consiste en producir un nuevo tensor **C** que tendrá una magnitud y una dirección diferente a **A**. Podremos estar interesados en encontrar todos los vectores que NO son rotados por **B**, es decir, que estaríamos interesados en resolver la ecuación:

$$B_i^j A_j = \lambda A_i, \text{ donde } \lambda \text{ es un escalar.}$$

Si estos vectores existen se denominan vectores característicos o “autovectores” de **B** y sus direcciones: direcciones principales o características. Más aún, los ejes determinados por las direcciones principales se denominan ejes principales de **B**. Los valores de las componentes B_i^j en el sistema de coordenadas determinado por los ejes principales se denominan valores característicos o “autovalores” de **B**.

Ahora, como lo ilustramos en los ejemplos, consideremos un sistema conformado por n partículas de masas iguales: $m_1, m_2, \dots, m_n = m$, distribuidas en el plano xy , y sea I_{ij} el tensor de inercia con respecto a un sistema rectangular de coordenadas. Por ser un caso en 2D, el tensor de inercia tendrá solamente cuatro componentes.

(a). Encuentre: $I_{11} = I_{xx}$, $I_{22} = I_{yy}$ y $I_{12} = I_{xy} = I_{21} = I_{yx}$.

(b). Si un vector **A** coincide con el eje principal de I_i^j entonces debe satisfacer la ecuación:

$$I_i^j A_j = \lambda A_i \Rightarrow (I_i^j - \lambda \delta_i^j) A_j = 0 \Rightarrow \begin{cases} (I_{11} - \lambda) A_1 + I_{12} A_2 = 0 \\ I_{12} A_1 + (I_{22} - \lambda) A_2 = 0 \end{cases}$$

Encuentre la solución (no trivial) para λ , es decir, resuelva: $\lambda^2 - \lambda(I_{11} + I_{22}) + I_{11}I_{22} - (I_{12})^2 = 0$.

(c). ¿Cómo se interpreta el hecho de que $I_{12} = 0$?

- (d). Si $I_{12} \neq 0$ y $\lambda_1 \neq \lambda_2$, entonces, para cada valor de λ se puede encontrar un vector (autovector) $\mathbf{A}^{(\lambda_1)}$ y $\mathbf{A}^{(\lambda_2)}$ resolviendo el sistema de dos ecuaciones. Demuestre que las direcciones de estos vectores tienen pendientes, respecto al sistema de coordenada, dadas por:

$$\tan(\theta_1) = \frac{\lambda_1 - I_{11}}{I_{12}}, \quad \tan(\theta_2) = \frac{\lambda_2 - I_{11}}{I_{12}}.$$

- (e). Demuestre que:

$$\tan(2\theta_1) = \tan(2\theta_2) = \frac{2I_{12}}{I_{11} - I_{22}}, \text{ donde } \theta_2 = \theta_1 + \frac{\pi}{2},$$

es decir: $\mathbf{A}^{(\lambda_1)} \perp \mathbf{A}^{(\lambda_2)}$.

- (f). ¿Cuáles son las componentes del tensor de inercia en el sistema de coordenadas determinado por los ejes principales?

6. En este ejercicio generalizamos la definición del momento de inercia para cuerpos continuos, el cual se define como:

$$I_j^i = \int_V dv \rho(\mathbf{r}) \left(\delta_j^i \left(x^k x_k \right) - x^i x_j \right), \quad \text{con } x^i = \{x, y, z\} \text{ y } dv = dx dy dz.$$

- (a). Muestre que I_j^i es un tensor.

- (b). Considere un cubo de lado l y masa total M tal que tres de sus aristas coinciden con un sistema de coordenadas cartesiano. Encuentre el tensor momento de inercia, I_j^i .

7. Dados dos sistemas de coordenadas ortogonales $O \rightleftharpoons (x, y, z)$ y $O' \rightleftharpoons (x', y', z')$, donde el sistema de coordenadas O' se obtiene rotando a O , $\pi/6$ alrededor del eje z y $\pi/2$ alrededor del eje x' , con lo cual los ejes y' y z coinciden.

- (a). Si tenemos los vectores:

$$\mathbf{A} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}, \quad \mathbf{B} = 2\hat{\mathbf{i}} + \hat{\mathbf{j}} + 3\hat{\mathbf{k}}.$$

Expréselos en el sistema de coordenadas $O' \rightleftharpoons (x', y', z')$.

- (b). El tensor de esfuerzos (tensiones normales y tangenciales a una determinada superficie) se expresa en el sistema $O \rightleftharpoons (x, y, z)$ como:

$$P_j^i = \begin{pmatrix} P_1 & 0 & P_4 \\ 0 & P_2 & 0 \\ 0 & 0 & P_3 \end{pmatrix}.$$

¿Cuál será su expresión en el sistema de coordenadas $O' \rightleftharpoons (x', y', z')$?

8. Suponga un sistema de coordenadas ortogonales generalizadas (q^1, q^2, q^3) las cuales tienen la siguiente relación funcional con las coordenadas cartesianas¹¹:

$$q^1 = x + y; \quad q^2 = x - y; \quad q^3 = 2z.$$

- (a). Compruebe que el sistema (q^1, q^2, q^3) conforma un sistema de coordenadas ortogonales

- (b). Encuentre los vectores base para este sistema de coordenadas.

- (c). Encuentre el tensor métrico y el elemento de volumen en estas coordenadas.

- (d). Encuentre las expresiones en el sistema (q^1, q^2, q^3) para los vectores:

$$\mathbf{A} = 2\hat{\mathbf{j}}, \quad \mathbf{B} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}}, \quad \mathbf{C} = \hat{\mathbf{i}} + 7\hat{\mathbf{j}} + 3\hat{\mathbf{k}}.$$

- (e). Encuentre en el sistema (q^1, q^2, q^3) las expresiones para las siguientes relaciones vectoriales :

$$\mathbf{A} \times \mathbf{B}, \quad \mathbf{A} \cdot \mathbf{C}, \quad (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}.$$

¹¹Estos tipo de transformaciones a coordenadas generalizadas será analizado, con todo detalle, en la sección ?? de la página ??.

¿Qué puede decir si compara esas expresiones en ambos sistemas de coordenadas?

(f). Considere los siguientes tensores y vectores en coordenadas cartesianas :

$$R_j^i = \begin{pmatrix} 1/2 & 1 & 3/2 \\ 2 & 5/2 & 3 \\ 7/2 & 4 & 9/2 \end{pmatrix}, \quad T^i = \begin{pmatrix} 1/3 \\ 2/3 \\ 1 \end{pmatrix}, \quad g^{ij} = g_{ij} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

y encuentre sus expresiones para el nuevo sistema de coordenadas (q^1, q^2, q^3) .

Practicando con Maxima

1. Tensores y Maxima

Maxima presenta dos bibliotecas para la manipulación de tensores, bibliotecas en constante desarrollo todavía. Estas son **ctensor**, manipulación de componentes y **itensor**, manipulación indexada. En la manipulación por componentes los tensores se representan por arreglos o matrices. En la manipulación indexada los tensores se representan por funciones de sus índices covariantes, contravariantes y de derivadas.

Los índices covariantes se especifican mediante una lista que será el primer argumento del objeto indexado, siendo los índices contravariantes otra lista que será el segundo argumento del mismo objeto indexado. Si al objeto indexado le falta cualquiera de estos grupos de índices, entonces se le asignará al argumento correspondiente la lista vacía `[]`. Así, $g([a, b], [c])$ representa un objeto indexado llamado g , el cual tiene dos índices covariantes (a, b) , un índice contravariante (c) y no tiene índices de derivadas, es decir, g_{ab}^c

```
(%ci1) load(itensor)$
(%ci2) imetric(g);
(%ci3) ishow(g([j,k],[l])*g([i,l],[l])*a([l],[i,j]))$ 
(%ot3) a^i_j g_{j k} g_{i l}
(%ci4) ishow(contract(%))$ 
(%ot4) a_{k l}
(%ci5) ishow(kdelta([i],[k])*a([k,l],[l]))$ 
(%ot5) delta_i^k a_{k l}
(%ci6) ishow(contract(%))$ 
(%ot6) a_{i l}
```

La identidad $\delta_i^j \delta_j^i = 3$

```
(%ci7) ishow(contract(kdelta([i], [j])*kdelta([j], [i])))$ 
(%ot7) kdelta
(%ci8) ishow(ev(% ,kdelta))$ 
(%ot8) 3
```

El símbolo de Levi-Civita:

```
(%ci9) levi_civita([1, 2, 3]); levi_civita([1, 3, 2]); levi_civita([3, 3, 2]);
(%o9) 1
(%o10) - 1
(%o11) 0
```

(%ci12) `expr:ishow('levi_civita([], [i,j,k])*'levi_civita([i,j,k], []));`
 (%o12) $\varepsilon^{ijk} \varepsilon_{ijk}$

Aquí es necesario utilizar la función **lc2kdt()** que simplifica expresiones que contengan el símbolo de Levi-Civita, para convertirlas en expresiones con la delta de Kronecker siempre que esto sea posible.¹²

(%ci13) `ishow(lc2kdt(expr))$`
 (%o13) $\left(\delta_i^k \delta_j^i - 3 \delta_j^k \right) \delta_k^j + \left(\delta_i^j \delta_j^k - 3 \delta_i^k \right) \delta_k^i - 3 \delta_i^j \delta_j^i + 27$

(%ci14) `ishow(contract(expand(%)))$`
 (%o14) $27 - 7\delta$

(%ci15) `ishow(ev(%kdelta))$`
 (%o15) 6

(%ci16) `kill(all)$`

2. Transformación de tensores y Maxima

Como ya mencionamos, **Maxima** contiene una librería para manipular componentes de tensores: **ctensor**. La manipulación en componentes se realiza en términos de matrices. La métrica se almacena en la matriz lg y la métrica inversa se obtiene y almacena en la matriz ug.

En un ejemplo visto con anterioridad se especificaba las componentes de un tensor en coordenadas cartesianas, esto es:

$$T_0 = T_j^i = \begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 2 \end{pmatrix}, \quad \text{en la base: } \{\hat{e}_1, \hat{e}_2, \hat{e}_3\} \equiv \{|i\rangle, |j\rangle, |k\rangle\}.$$

Para luego representarlo en la nueva base: $|w_{1'}\rangle = |i\rangle$, $|w_{2'}\rangle = |i\rangle + |j\rangle$ y $|w_{3'}\rangle = |i\rangle + |j\rangle + |k\rangle$.

Entonces necesitamos calcular:

$$T_{m'}^{k'} = \frac{\partial x^{k'}}{\partial x^i} \frac{\partial x^j}{\partial x^{m'}} T_j^i \Rightarrow T_n = \alpha \beta T_0 =$$

Como vimos cuando hacíamos los cálculos:

$$\alpha = \frac{\partial x^{k'}}{\partial x^i} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \beta = \frac{\partial x^i}{\partial x^{k'}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Respetando “la concatenación interna de índices” podemos realizar la multiplicación de matrices.

(%ci1) `To:matrix([2,1,3],[2,3,4],[1,2,2]);`
 (%o1) $\begin{pmatrix} 2 & 1 & 3 \\ 2 & 3 & 4 \\ 1 & 2 & 2 \end{pmatrix}$
 (%ci2) `alpha:matrix([1,-1,0],[0,1,-1],[0,0,1]);`
 (%o2) $\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$
 (%ci3) `beta:matrix([1,1,1],[0,1,1],[0,0,1]);`

¹²En ocasiones, la función **lc2kdt()** hace uso del tensor métrico. Si el tensor métrico no fue previamente definido con **imetric**, se obtiene un mensaje de error.

$$(\%o3) \quad \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

(%i4) `Tn:=alpha.To.beta;`

$$(\%o4) \quad \begin{pmatrix} 0 & -2 & -3 \\ 1 & 2 & 4 \\ 1 & 3 & 5 \end{pmatrix}$$

Una vez que hemos calculado el nuevo tensor $T_n = \tilde{T}_j^i$ en la nueva base, podemos calcular: \tilde{T}^{ij} , \tilde{T}_{ij} . Pero podemos utilizar la métrica para las coordenadas nuevas y con el paquete **ctensor** hacer las respectivas contracciones.

(%i5) `load(ctensor)$`

(%i6) `ct_coordsys(cartesian3d)$`

Introducimos la métrica \tilde{g}_{ik} :

(%i7) `lg:=matrix([1,1,1],[1,2,2],[1,2,3]);`

$$(\%o7) \quad \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

Para tener la métrica inversa \tilde{g}^{ik} escribimos:

(%i8) `cmetric()$`

(%i9) `ug;`

$$(\%o9) \quad \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

Para calcular $\tilde{T}_{ij} = \tilde{g}_{ik}\tilde{T}_j^k$ hacemos lo siguiente:

(%i10) `lg.Tn;`

$$(\%o10) \quad \begin{pmatrix} 2 & 3 & 6 \\ 4 & 8 & 15 \\ 5 & 11 & 20 \end{pmatrix}$$

Y para calcular $\tilde{T}^{ij} = \tilde{T}_k^i\tilde{g}^{kj}$ procedemos así:

(%i11) `Tn.ug;`

$$(\%o11) \quad \begin{pmatrix} 2 & -1 & -1 \\ 0 & -1 & 2 \\ -1 & 0 & 2 \end{pmatrix}$$

(%i12) `kill(all)$`

Es posible, y a manera de completar esta primer acercamiento con la biblioteca de tensores, calcular la métrica a partir de una transformación de coordenadas, como vimos en la sección anterior, ejemplo 3.2.8. Escribimos las coordenadas que queremos utilizar, en este caso esféricas, utilizando la función **ct coordsys**.

(%i1) `load(ctensor)$`

(%i2) `ct_coordsys([r*sin(theta)*cos(phi),r*sin(theta)*sin(phi),r*cos(theta)],[r,theta,phi]])$`

Y le pedimos al programa que calcule el tensor métrico, al que le asigna el nombre **lg**.

(%i3) `cmetric()$`

El tensor métrico g_{ij} es entonces

(%i4) `lg:trigsimp(lg);`

$$(\%o4) \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2 \theta \end{pmatrix}$$

Y la métrica inversa g^{ij} , el programa la identifica con la etiqueta **ug**

(%i5) `trigsimp(ug);`

$$(\%o5) \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & \frac{1}{r^2 \sin^2 \theta} \end{pmatrix}$$

3.4 Vectores, tensores y espacios pseudoeuclidianos

Hasta este punto la descripción de formas representadas por un *bra*: $\langle a| \equiv a_k \langle e^k|$ ha sido casi estética. Hemos insistido que las componentes de las formas tienen subíndices, mientras que sus vectores bases, $\langle e^k|$, deben tener superíndices, pero no hemos visto clara la necesidad de esa definición. Un ejemplo, un tanto tímido lo desarrollamos en la sección 3.1.2, en la página 149, donde mencionamos que las bases recíprocas de vectores podrían jugar el papel de bases para el espacio dual. Quizá el ejemplo más significativo y simple, donde se observa la diferencia entre formas (*bras*) y vectores (*kets*) es el caso de los espacios minkowskianos. Estos espacios, también llamados pseudoeuclidianos, presentan una variante en la definición de producto interno, de tal forma que: $\langle x|x\rangle$ no es necesariamente positivo, y si $\langle x|x\rangle = 0$ no necesariamente implica que $|x\rangle \equiv |0\rangle$.

La consecuencia inmediata es que la definición de norma $\mathcal{N}(|v_i\rangle) \equiv \||v_i\rangle\|$, que vimos anteriormente, no es necesariamente positiva. Vale decir que tendremos vectores con norma positiva, $\||v_i\rangle\| > 0$, pero también vectores con norma negativa o cero: $\||v_i\rangle\| \leq 0$. Con lo cual la definición de distancia, entendida como la norma de la resta de vectores, $d(|x\rangle, |y\rangle) \equiv \||x\rangle - |y\rangle\|$, tampoco será necesariamente positiva. Esto es, que las distancias serán negativas, positivas o nulas: $d(|x\rangle, |y\rangle) < 0$, $d(|x\rangle, |y\rangle) = 0$ y $d(|x\rangle, |y\rangle) > 0$.

Si extendemos la noción de distancia para que albergue las posibilidades de distancias nula y negativas, entonces la definición del tensor métrico para espacios pseudoeuclidianos también debe cambiar.

$$\mathbf{g}[|x_i\rangle, |x_j\rangle] = g_{ij} \equiv g_{ji} \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases}$$

En resumen

$$\langle x|x\rangle = \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases} \Rightarrow d(|x\rangle, |y\rangle) = \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases} \Rightarrow \mathbf{g}[|x_i\rangle, |x_j\rangle] = \begin{cases} < 0 \\ = 0 \\ > 0 \end{cases}$$

Este tipo de espacios luce como un excentricidad más de los matemáticos y una curiosidad de estudio es ver como organizar los conceptos que aprendimos de los espacios euclidianos y extenderlos a otros espacios. Quizá se hubiera quedado así, como una curiosidad matemática si los físicos no hubieran sacado partido de estas particularidades para describir el comportamiento de la naturaleza. En la próxima sección analizaremos el caso de espacios minkowskianos de dimensión 4, que denominaremos \mathbb{M}^4 .

3.4.1 Espacios minkowskianos

Consideremos un espacio tetradimensional expandido por una base ortonormal: $\{|\hat{e}_0\rangle, |\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle\}$. Los vectores $\{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle\}$ corresponden con la base canónica de \mathbb{R}^3 .

Este espacio vectorial \mathbb{M}^4 tendrá asociado un espacio dual: $\{\langle \hat{e}^0|, \langle \hat{e}^1|, \langle \hat{e}^2|, \langle \hat{e}^3|\}$ a través de una métrica:

$$\eta_{\alpha\beta} \langle \hat{e}^\alpha | \otimes \langle \hat{e}^\beta | \equiv \eta_{\beta\alpha} \langle \hat{e}^\beta | \otimes \langle \hat{e}^\alpha | \quad y \quad \eta^{\alpha\beta} |\hat{e}_\alpha\rangle \otimes |\hat{e}_\beta\rangle \equiv \eta^{\beta\alpha} |\hat{e}_\beta\rangle \otimes |\hat{e}_\alpha\rangle,$$

con $\alpha, \beta = 0, 1, 2, 3$, y donde: $\eta_{00} = \eta^{00} = 1$, $\eta_{11} = \eta^{11} = -1$, $\eta_{22} = \eta^{22} = -1$, $\eta_{33} = \eta^{33} = -1$ (con $\eta_{\alpha\beta} = 0$ para $\alpha \neq \beta$). Se dice que η tiene signo -2 .¹³

Tal y lo como presentamos en la sección 3.8, podemos asociar componentes covariantes y contravariantes

¹³Realmente el signo -2 es una convención, se puede también considerar $\eta_{\mu\nu}$ de signo $+2$, con $\eta_{00} = -1$, $\eta_{11} = +1$, $\eta_{22} = +1$, $\eta_{33} = +1$.

a través de la métrica $|a\rangle = a^\sigma |\hat{e}_\sigma\rangle$, entonces:

$$\left(\eta_{\alpha\beta} \langle \hat{e}^\alpha | \otimes \langle \hat{e}^\beta | \right) |a\rangle = a^\sigma \left(\eta_{\alpha\beta} \langle \hat{e}^\alpha | \otimes \langle \hat{e}^\beta | \right) |\hat{e}_\sigma\rangle = a^\sigma \eta_{\alpha\beta} \langle \hat{e}^\beta | \hat{e}_\sigma \rangle \langle \hat{e}^\alpha | = a^\sigma \eta_{\alpha\beta} \delta_\sigma^\beta \langle \hat{e}^\alpha | = a^\sigma \eta_{\alpha\sigma} \langle \hat{e}^\alpha | \equiv a_\alpha \langle \hat{e}^\alpha | .$$

Lo interesante del caso es que:

$$a_\alpha = a^\sigma \eta_{\sigma\alpha} \Rightarrow a^0 = a_0, \quad a^1 = -a_1, \quad a^2 = -a_2, \quad a^3 = -a_3.$$

Es decir, en este caso, porque la métrica tiene signo -2 , bajar los índices espaciales ($\mu = i = 1, 2, 3$) es cambiar el signo a las componentes¹⁴. Dicho con más propiedad, las componentes espaciales contravariantes ($\mu = i = 1, 2, 3$) tienen signos contrarios a las componentes covariantes.

De la misma manera que se expuso anteriormente en la sección 3.8:

$$\langle a | \left(\eta^{\alpha\beta} |\hat{e}_\alpha\rangle \otimes |\hat{e}_\beta\rangle \right) = a_\sigma \langle \hat{e}^\sigma | \left(\eta^{\alpha\beta} |\hat{e}_\alpha\rangle \otimes |\hat{e}_\beta\rangle \right) = a_\sigma \eta^{\alpha\beta} \langle \hat{e}^\sigma | \hat{e}_\alpha \rangle |\hat{e}_\beta\rangle = a_\sigma \eta^{\alpha\beta} \delta_\alpha^\sigma |\hat{e}_\beta\rangle = a_\sigma \eta^{\sigma\beta} |\hat{e}_\beta\rangle \equiv a^\beta |\hat{e}_\beta\rangle .$$

y otra vez, $a^\sigma = \eta^{\sigma\alpha} a_\alpha$, y habría cambio de signo cuando se bajan los índices $1, 2, 3$ para la métrica con signo -2 que hemos considerado anteriormente.

Del mismo modo se “suben” y se “bajan” índices para componentes de tensores:

$$\eta^{\alpha\beta} P_\alpha^{\gamma\sigma\epsilon} \equiv P^{\beta\gamma\sigma\epsilon} .$$

Por su parte, el producto interno de dos vectores en un espacio de Minkowski involucra, de manera natural, la métrica del espacio. Esto es:

$$\langle a|b\rangle = \langle b|a\rangle = a^\alpha b_\alpha = b^\alpha a_\alpha = a^\alpha b^\beta \eta_{\alpha\beta} = a_\alpha b_\beta \eta^{\alpha\beta} = a^0 b^0 - a^1 b^1 - a^2 b^2 - a^3 b^3 = a_0 b_0 - a_1 b_1 - a_2 b_2 - a_3 b_3 .$$

Una vez más, la norma de un vector, también incluirá al tensor métrico:

$$\|a\rangle\|^2 = \langle a|a\rangle = a_\alpha a^\beta \langle \hat{e}^\alpha | \hat{e}_\beta \rangle = a_\alpha a^\alpha = a_\alpha a_\beta \eta^{\alpha\beta} = a^\alpha a^\beta \eta_{\alpha\beta} = a^0 a^0 - a^1 a^1 - a^2 a^2 - a^3 a^3 .$$

El caso más conocido lo constituye la norma de un desplazamiento infinitesimal, en un espacio tetradimensional, el cual puede expresarse como:

$$ds^2 \equiv \langle dr|dr\rangle = (dx_\alpha \langle w^\alpha |) \left(dx^\beta |w_\beta \rangle \right) = dx_\beta dx^\beta = \eta_{\alpha\beta} dx^\alpha dx^\beta = dt^2 - dx^2 ,$$

con: $dx^2 = (dx^1)^2 + (dx^2)^2 + (dx^3)^2$.

3.4.2 Un toque de Relatividad Especial

La genialidad de Albert Einstein fue haber entendido que tenía que incorporar el tiempo como otra coordenada más, vale decir, que los eventos que ocurren en la naturaleza están etiquetados por cuatro números: $(t, x, y, z) \equiv (x^0, x^1, x^2, x^3)$ ¹⁵. El rápido desarrollo de la comprensión de las ideas relativistas muestra que estaban en el ambiente de la época de comienzos de 1900, y una vez más la simplicidad como prejuicio se impuso.

Sólo dos suposiciones están en el corazón de la Relatividad Especial:

1. *El principio de la Relatividad:* Las leyes de la Física son idénticas en todos los sistemas de referencia inerciales.
2. *La universalidad de la velocidad de la luz en el vacío:* La velocidad de la luz en el vacío es siempre la misma, y es independiente de la velocidad de la fuente de luz respecto a un observador en particular.

¹⁴Otra vez, para la métrica con signo -2 , el cambio de signo entre componentes covariantes y contravariantes se da para la componente, $\mu = 0$

¹⁵Una discusión sobre la necesidad de incorporar los conceptos de relatividad especial en los programas de estudio de Física mediante la utilización del álgebra geométrica la pueden encontrar en Baylis, W. E. (2004). Relativity in introductory physics. Canadian journal of physics, 82(11), 853-873.

En términos matemáticos estas dos audaces suposiciones se concretan en una simple propuesta matemática: el producto interno entre dos elementos de este espacio tetradimensional, debe conservarse para una familia de vectores base. Luego vendrá la asociación de observadores físicos –o sistemas de coordenadas– con los miembros de la familia de vectores base, pero la idea es la misma que planteamos para los espacios euclidianos en 2.2.3.1: el producto interno –y consecuentemente, la norma de los elementos del espacio vectorial y la distancia entre éstos– es el mismo independientemente de la base en la cual expanda el espacio vectorial.

La primera de las interpretaciones es el cómo representamos los eventos en el espacio-tiempo. Supongamos el caso unidimensional en el espacio, vale decir los eventos ocurren en un punto de la recta real $x = x^1$, y en un tiempo determinado, por lo tanto podremos asociar al evento un vector evento: $\rightarrow (x^0, x^1)$.

A continuación nos preguntamos que representan las distancias (espacio-temporales) entre estos dos eventos. Tal y como vimos, las distancias entre dos elementos de un espacio vectorial puede ser construida a partir de la norma (de la resta de coordenadas) y la norma a partir del producto interno:

$$\| |y - x\rangle\|^2 \equiv \langle y - x | y - x \rangle \begin{cases} < 0 & \text{conexión tipo espacio: eventos desconectados causalmente.} \\ = 0 & \text{conexión tipo luz: posible conexión causal a través de rayos de luz.} \\ > 0 & \text{conexión tipo tiempo: posible conexión causal.} \end{cases}$$

Con esta primera interpretación de los valores de la norma y la visión tetradimensional, el espacio-tiempo, dividido en pasado, presente y futuro, se puebla de eventos que pueden estar o no relacionados causalmente tal y como muestra la figura 3.4.

La preservación del producto interno para todos los observadores era intuitiva en los espacios euclidianos y, al mantenerla para los pseudoeuclidianos nos traerá consecuencias nada intuitivas en nuestra idea intuitiva de “realidad”.

Para el caso de la formulación de la Relatividad Especial, añadimos un supuesto más: las componentes del tensor métrico son invariantes bajo transformaciones de coordenadas, esto es:

$$g[|\hat{e}_\mu\rangle, |\hat{e}_\nu\rangle] \equiv g'[|\hat{e}_{\mu'}\rangle, |\hat{e}_{\nu'}\rangle] \Leftrightarrow \eta_{\alpha\beta} = \eta_{\alpha'\beta'},$$

con: $\{|e_\mu\rangle\}$ y $\{|\hat{e}_{\mu'}\rangle\}$ dos bases que se conectan a través de una transformación de coordenadas:

$$x^\mu = x^\mu(x^{\alpha'}) \Leftrightarrow x^{\mu'} = x^{\mu'}(x^\alpha).$$

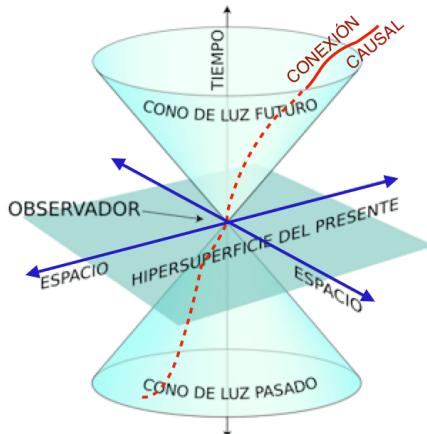


Figura 3.4: Cono de luz, espacio-tiempo y eventos.

Construyamos ahora el tipo de transformación de coordenadas que mantiene estos dos supuestos¹⁶:

1. El producto interno de dos vectores es independiente de la base que expande el espacio vectorial.
2. Las componentes del tensor métricos son invariantes bajo transformaciones de coordenadas.

Si el producto interno de dos vectores es independiente de la base que expande el espacio vectorial, tendremos:

$$\langle x | y \rangle = (x_\alpha \langle \hat{e}^\alpha |) (y^\beta | \hat{e}_\beta \rangle) \equiv \left(x_{\sigma'} \langle \hat{e}^{\sigma'} | \right) \left(y^{\alpha'} | \hat{e}_{\alpha'} \rangle \right) \Leftrightarrow x^\alpha y_\alpha \equiv x^{\beta'} y_{\beta'} \Leftrightarrow x^\alpha y^\beta \eta_{\alpha\beta} \equiv x^{\alpha'} y^{\beta'} \eta_{\alpha'\beta'},$$

y como lo vimos en 3.2.8 las componentes de vectores, bajo cambio de coordenadas, transforman como:

$$a^{i'} = \frac{\partial x^{i'}}{\partial x^k} a^k \Rightarrow x^\alpha y_\alpha = x^{\alpha'} y_{\alpha'} \Leftrightarrow x^\alpha y^\beta \eta_{\alpha\beta} = \frac{\partial x^{\nu'}}{\partial x^\alpha} x^\alpha \frac{\partial x^{\mu'}}{\partial x^\beta} y^\beta \eta_{\nu'\mu'} = x^\alpha y^\beta \frac{\partial x^{\nu'}}{\partial x^\alpha} \frac{\partial x^{\mu'}}{\partial x^\beta} \eta_{\nu'\mu'},$$

¹⁶Estamos suponiendo que observadores, sistemas de coordenadas y sistemas de referencia son conceptos equivalentes.

con lo cual concluimos que:

$$\eta_{\alpha\beta} = \frac{\partial x^{\nu'}}{\partial x^\alpha} \frac{\partial x^{\mu'}}{\partial x^\beta} \eta_{\nu'\mu'} \equiv \frac{\partial x^{\nu'}}{\partial x^\alpha} \frac{\partial x^{\mu'}}{\partial x^\beta} \eta_{\nu\mu}. \quad (3.16)$$

Ahora bien, si derivamos (3.16) respecto a x^γ tendremos que:

$$\eta_{\nu'\mu'} \left[\frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\gamma} \frac{\partial x^{\mu'}}{\partial x^\beta} + \frac{\partial x^{\nu'}}{\partial x^\alpha} \frac{\partial^2 x^{\mu'}}{\partial x^\beta \partial x^\gamma} \right] = 0.$$

Como la cantidad dentro del paréntesis se anula podemos jugar con ésta para descubrir algunas consecuencias ocultas. Es de hacer notar que esa cantidad tiene tres índices libres y por lo tanto son 64 ecuaciones que se anulan. Eso significa que le podemos añadir y sustraer cualesquiera otras con los índices intercambiados.

Supongamos que al paréntesis anulado le añadimos una con los índices α y γ intercambiados y, adicionalmente, le sustraemos una con los índices γ y β intercambiados. Claramente, estamos añadiendo y sustrayendo ceros.

$$\eta_{\nu'\mu'} \left[\frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\gamma} \frac{\partial x^{\mu'}}{\partial x^\beta} + \frac{\partial x^{\nu'}}{\partial x^\alpha} \frac{\partial^2 x^{\mu'}}{\partial x^\beta \partial x^\gamma} + \frac{\partial^2 x^{\nu'}}{\partial x^\gamma \partial x^\alpha} \frac{\partial x^{\mu'}}{\partial x^\beta} + \frac{\partial x^{\nu'}}{\partial x^\gamma} \frac{\partial^2 x^{\mu'}}{\partial x^\beta \partial x^\alpha} - \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\beta} \frac{\partial x^{\mu'}}{\partial x^\gamma} - \frac{\partial x^{\nu'}}{\partial x^\alpha} \frac{\partial^2 x^{\mu'}}{\partial x^\gamma \partial x^\beta} \right] = 0.$$

Con este “truco”, vemos que el último término anula el segundo y el penúltimo el cuarto, de forma y manera que nos queda:

$$2\eta_{\nu'\mu'} \frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\gamma} \frac{\partial x^{\mu'}}{\partial x^\beta} = 0,$$

Con lo cual la única posibilidad que resulta es la siguiente:

$$\frac{\partial^2 x^{\nu'}}{\partial x^\alpha \partial x^\gamma} = 0 \Rightarrow x^{\nu'} = \Lambda_\mu^{\nu'} x^\mu + a^{\nu'}, \quad (3.17)$$

con: $\Lambda_\mu^{\nu'}$ y $a^{\nu'}$ constantes.

Las transformaciones lineales (3.17) se conocen como *las transformaciones, inhomogéneas, de Lorentz* o también las transformaciones de Poincaré. Estas transformaciones forman un grupo y, uno de los posibles subgrupos lo constituye el conjunto de transformaciones propias de Lorentz de la forma:

$$\Lambda_0^{0'} = 1, \quad \Lambda_0^{i'} = \Lambda_j^{0'} = 0, \quad y \quad \Lambda_j^{i'} = R_j^{i'},$$

con: $i, j = 1, 2, 3$, y donde $R_j^{i'}$ es una matriz de rotación.

Supongamos el caso más sencillo de este grupo de transformaciones: $a^{\nu'} = 0$, en la ecuación (3.17). Explícitamente hemos identificado una transformación de la forma:

$$x^{\alpha'} = \Lambda_0^{\alpha'} x^0 + \Lambda_1^{\alpha'} x^1 + \Lambda_2^{\alpha'} x^2 + \Lambda_3^{\alpha'} x^3,$$

la cual, por construcción, deja invariante el intervalo tetradimensional:

$$ds^2 = c^2 dt^2 - dx^1 - dx^2 - dx^3 = \eta_{\mu\nu} dx^\mu dx^\nu,$$

con $\eta_{\mu\nu}$ el tensor métrico. Aquí c es la constante que representa la velocidad de la luz en el vacío. Por razones de conveniencia se escoge un sistema de unidades donde $c = 1$ ¹⁷.

Es inmediato demostrar que este tipo de transformaciones deja invariante el intervalo. Primero, notemos que:

$$\eta^{\mu'\nu'} = \Lambda_\alpha^{\mu'} \Lambda_\beta^{\nu'} \eta^{\alpha\beta} \Rightarrow \eta^{\mu'\nu'} \eta_{\nu'\gamma'} = \delta_{\gamma'}^{\mu'} = \Lambda_\alpha^{\mu'} \Lambda_\beta^{\nu'} \eta^{\alpha\beta} \eta_{\nu'\gamma'} \Rightarrow \Lambda_\alpha^{\mu'} \Lambda_\gamma^\alpha = \delta_{\gamma'}^{\mu'},$$

y como $dx^{\mu'} = \Lambda_\alpha^{\mu'} dx^\alpha$, entonces:

$$ds'^2 = \eta_{\mu'\nu'} dx^{\mu'} dx^{\nu'} \equiv \eta_{\mu'\nu'} \Lambda_\alpha^{\mu'} dx^\alpha \Lambda_\beta^{\nu'} dx^\beta = \eta_{\alpha\beta} dx^\alpha dx^\beta = ds^2.$$

¹⁷Un sistema de unidades denominado *sistema de unidades geometrizado* en el cual la velocidad de la luz y la constante de gravedad universal se toman como la unidad: $c = G = 1$.

Para construir una de las expresiones más utilizadas del grupo de Lorentz consideramos la siguiente situación: un observador, $x^{\mu'}$, ve moverse una partícula con una velocidad \mathbf{v} , mientras que un segundo observador, x^μ , la percibe en reposo. Entonces, para el observador, que registra la partícula en reposo, resulta que $dx^i = 0$, y:

$$dx^{\mu'} = \Lambda_\alpha^{\mu'} dx^\alpha \Rightarrow \begin{cases} dt' &= \Lambda_0^{0'} dt \\ dx^{i'} &= \Lambda_\alpha^{i'} dx^\alpha = \Lambda_0^{i'} dt \quad \text{con } i' = 1, 2, 3. \end{cases}$$

Ahora bien, las ecuaciones anteriores nos imponen:

$$\mathbf{v}' = \frac{d\mathbf{x}'}{dt'} \Rightarrow v^{i'} = \frac{dx^{i'}}{dt'} \Rightarrow \Lambda_0^{i'} = v^{i'} \Lambda_0^{0'}.$$

Además, tenemos que:

$$\eta_{\alpha'\beta'} = \Lambda_\alpha^{\mu'} \Lambda_\beta^{\nu'} \eta_{\mu\nu} \Rightarrow 1 = \Lambda_0^\mu \Lambda_0^\nu \eta_{\mu\nu} = (\Lambda_0^0)^2 - (\Lambda_0^1)^2 - (\Lambda_0^2)^2 - (\Lambda_0^3)^2,$$

con una solución de la forma:

$$\Lambda_0^0 = \gamma, \quad \Lambda_0^i = \gamma v^i,$$

donde:

$$\gamma = \frac{1}{\sqrt{1 - \mathbf{v}^2}} \equiv \frac{1}{\sqrt{1 - v^i v_i}} \equiv \frac{1}{\sqrt{1 - [(v^1)^2 + (v^2)^2 + (v^3)^2]}},$$

los otros términos $\Lambda_{j'}^i$ no quedan únicamente determinados porque está de por medio la arbitrariedad de una rotación $R_{j'}^i$. Por ello, una selección arbitraria pero razonable de todos los términos $\Lambda_{j'}^i$ es:

$$\Lambda_{j'}^i = \delta_{j'}^i + v^i v_{j'} \frac{\gamma - 1}{\mathbf{v}^2} = \delta_{j'}^i + v^i v_{j'} \frac{\gamma - 1}{v^k v_k}.$$

De esta forma quedan determinados todos los elementos de las transformaciones de Lorentz.

Los observadores lorentzianos son los equivalentes a los observadores galileanos en las teorías newtonianas: son observadores que se mueven uno respecto al otro con una velocidad constante y, desempeñan el mismo papel que los observadores inerciales. Quizá la consecuencia más impactante de la necesidad de vincular mediciones de distintos observadores lorentzianos a través de transformaciones de Lorentz, lo ilustra la evolución distinta del tiempo medido por los diferentes observadores. Un observador en reposo respecto a un reloj, ve avanzar el tiempo con *tic* separados por $dt = \Delta t$, ya que su reposo respecto al reloj implica: $dx^i = 0$, por lo tanto la separación espacio temporal será:

$$ds^2 = dt^2 - (dx^i)^2 = (\Delta t)^2 \Rightarrow dt = \Delta t,$$

mientras que un segundo observador, en movimiento, tendrá el mismo elemento de línea pero expresado como:

$$ds'^2 = dt'^2 - (dx^{i'})^2 = (1 - \mathbf{v}^2) dt' \Rightarrow dt' = \frac{\Delta t}{\sqrt{1 - \mathbf{v}^2}},$$

y ésta última ecuación claramente indica que el tiempo evoluciona más lento para relojes en movimiento.

Ejemplo 3.15 En este ejemplo vamos a repetir, para el caso bidimensional, lo expuesto de manera general en la sección anterior. Como vimos, las transformaciones de Lorentz relacionan las coordenadas del espacio-tiempo medidas por un observador O de un evento, con las coordenadas medidas por otro observador O' del mismo evento. El observador O lo representaremos por las coordenadas $\{t, x, y, z\} = \{x^0, x^1, x^2, x^3\}$, mientras que el observador O' por $\{t', x', y', z'\} = \{x^{0'}, x^{1'}, x^{2'}, x^{3'}\}$. Para lo que sigue, supondremos que cuando los observadores coincidan los relojes de ambos marcarán $t = t' = 0$.

La ecuación (3.17),

$$x^{\nu'} = \Lambda_{\mu}^{\nu'} x^{\mu},$$

la podemos expandir en lo que realmente es:

$$\begin{aligned} t' &= x^{0'} = \Lambda_0^{0'} x^0 + \Lambda_1^{0'} x^1 + \Lambda_2^{0'} x^2 + \Lambda_3^{0'} x^3 \\ x' &= x^{1'} = \Lambda_0^{1'} x^0 + \Lambda_1^{1'} x^1 + \Lambda_2^{1'} x^2 + \Lambda_3^{1'} x^3 \\ y' &= x^{2'} = \Lambda_0^{2'} x^0 + \Lambda_1^{2'} x^1 + \Lambda_2^{2'} x^2 + \Lambda_3^{2'} x^3 \\ z' &= x^{3'} = \Lambda_0^{3'} x^0 + \Lambda_1^{3'} x^1 + \Lambda_2^{3'} x^2 + \Lambda_3^{3'} x^3 \end{aligned}$$

Vamos a suponer que O' se mueve respecto a O , con velocidad v , únicamente en la dirección x , es decir, $y' = y$ y $z' = z$. Se supone entonces que t' no dependerá ni de y ni de z y que cuando $x' = 0$ entonces $x = vt$.

Todo esto hace que el sistema de ecuaciones anterior se simplifique de manera significativa:

$$\begin{aligned} t' &= x^{0'} = \Lambda_0^{0'} x^0 + \Lambda_1^{0'} x^1 = \Lambda_0^{0'} t + \Lambda_1^{0'} x \\ x' &= x^{1'} = \Lambda_1^{1'} (x^1 - vt) = \Lambda_1^{1'} (x - vt) \\ y' &= x^{2'} = x^2 = y \\ z' &= x^{3'} = x^3 = z \end{aligned}$$

Debemos determinar los coeficiente $\Lambda_0^{0'}$, $\Lambda_1^{0'}$ y $\Lambda_1^{1'}$. Para tal fin vamos a suponer que cuando $t' = t = 0$ un pulso de luz es emitido desde el origen de coordenadas. Recordemos que en ese instante ambos observadores coinciden. Como suponemos además que la velocidad de la luz es constante, la onda de luz se propagará en todas las direcciones de manera que cada observador podrá describirla mediante la ecuación de una esfera cuyo radio aumenta con el tiempo a velocidad c . Esto es:

$$x'^2 + y'^2 + z'^2 = c^2 t'^2 \quad \text{y} \quad x^2 + y^2 + z^2 = c^2 t^2.$$

Por lo tanto:

$$[\Lambda_1^{1'}(x - vt)]^2 + y^2 + z^2 = c^2 [\Lambda_0^{0'}t + \Lambda_1^{0'}x]^2.$$

Al desarrollar esta última ecuación obtenemos:

$$\left(\Lambda_1^{1'}\right)^2 (x^2 - 2xvt + v^2 t^2) + y^2 + z^2 = c^2 \left[\left(\Lambda_0^{0'}\right)^2 t^2 + 2\Lambda_0^{0'} \Lambda_1^{0'} xt + \left(\Lambda_1^{0'}\right)^2 x^2 \right],$$

por lo tanto:

$$\left[\left(\Lambda_1^{1'}\right)^2 - c^2 \left(\Lambda_1^{0'}\right)^2\right] x^2 + y^2 + z^2 - 2xt \left[\left(\Lambda_1^{1'}\right)^2 v + c^2 \Lambda_0^{0'} \Lambda_1^{0'}\right] = \left[c^2 \left(\Lambda_0^{0'}\right)^2 - v^2 \left(\Lambda_1^{1'}\right)^2\right] t^2.$$

Como ambos observadores registran la misma onda de luz, entonces:

$$\begin{aligned} \left(\Lambda_1^{1'}\right)^2 - c^2 \left(\Lambda_1^{0'}\right)^2 &= 1 \\ \left(\Lambda_1^{1'}\right)^2 v + c^2 \Lambda_0^{0'} \Lambda_1^{0'} &= 0 \\ c^2 \left(\Lambda_0^{0'}\right)^2 - v^2 \left(\Lambda_1^{1'}\right)^2 &= c^2. \end{aligned}$$

Al resolver este sistema de tres ecuaciones con tres incógnitas resulta:

$$\Lambda_1^{1'} = \Lambda_0^{0'} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad \Lambda_1^{0'} = \frac{v}{c^2 \sqrt{1 - \frac{v^2}{c^2}}}.$$

Si definimos $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$, las transformaciones de Lorentz quedan entonces de la forma:

$$t' = \left(t - \frac{v}{c^2} x\right) \gamma, \quad x' = (x - vt) \gamma, \quad y' = y, \quad z' = z.$$

Existen las transformaciones inversas que se pueden obtener de la misma manera que las anteriores.

$$t = \left(t' + \frac{v}{c^2} x' \right) \gamma, \quad x = \left(x' + vt' \right) \gamma, \quad y = y', \quad z = z'.$$

Notemos que cuando el observador O' se mueve a velocidades muy pequeñas en comparación a c , entonces el factor γ tiende a la unidad:

$$v \ll c \Rightarrow \gamma \rightarrow 1,$$

por lo tanto:

$$t' = t, \quad x' = x - vt, \quad y' = y, \quad z' = z.$$

que no son más que la transformación de Galileo.



Ejemplo 3.16

Si un observador O' determina que dos eventos ocurren en el mismo lugar, pero separados en el tiempo; entonces, un observador O ¿cómo los apreciaría? Veamos dos situaciones diferentes.

- Consideremos a una persona sentada en un tren en movimiento, observador O' , esta persona enciende y apaga rápidamente una linterna dos veces, con una diferencia de 15 minutos entre un evento y el otro. Entonces para O' será dos eventos que ocurren en el mismo lugar. Pero para el observador O que se encuentra en tierra los eventos habrán ocurrido en lugares diferentes, y esto no parece tener ningún conflicto con nuestra experiencia e intuición.
- Si O' sigue en su asiento del tren y observa que dos personas, uno a cada extremo del vagón (a unos 30 mts), encienden las linternas al mismo tiempo, entonces para O' será un evento simultáneo, pero para el observador en tierra O que ve al tren alejarse determinará que la persona de la parte trasera encendió la linterna un poco antes que la persona apostada en la parte delantera. Esto tal vez si nos resulte algo extraño.

Consideremos que el tren viaja a 36 m/s (≈ 130 Km/h) entonces, esta pequeña velocidad se refleja en el factor γ de la manera siguiente:

$$\frac{v}{c} = 1,2 \times 10^{-7} \rightarrow \frac{v^2}{c^2} = 1,44 \times 10^{-14} \rightarrow \gamma = \frac{1}{\sqrt{1 - 1,44 \times 10^{-14}}} \approx 1,0000000000000007.$$

Cuando consideramos la primera situación, la diferencia entre un evento y el otro era de 15 minutos (900 seg), y para el observador en tierra ambos eventos ocurren en dos lugares diferentes: x_1 y x_2 ,

$$x_1 = (x'_1 + vt'_1) \gamma, \quad x_2 = (x'_2 + vt'_2) \gamma,$$

de manera que:

$$x_2 - x_1 = (x'_2 + vt'_2) \gamma - (x'_1 + vt'_1) \gamma = (x'_2 - x'_1) \gamma + (t'_2 - t'_1) v \gamma.$$

Pero: $x'_2 = x'_1$ y $t'_2 - t'_1 = 15$ min. Por lo tanto:

$$x_2 - x_1 = (900 \text{ s}) (36 \text{ m.s}^{-1}) \gamma = 32400 \text{ m} = 32,4 \text{ Km}.$$

Para la segunda situación, el observador en tierra puede dar fe que los dos eventos ocurren en tiempos diferentes: t_1 y t_2 ,

$$t_1 = \left(t'_1 + \frac{v}{c^2} x'_1 \right) \gamma, \quad t_2 = \left(t'_2 + \frac{v}{c^2} x'_2 \right) \gamma,$$

de manera que:

$$t_2 - t_1 = \left(t'_2 + \frac{v}{c^2} x'_2 \right) \gamma - \left(t'_1 + \frac{v}{c^2} x'_1 \right) \gamma = (t'_2 - t'_1) \gamma + (x'_2 - x'_1) \frac{v}{c^2} \gamma,$$

con: $t'_2 = t'_1$ y $x'_2 - x'_1 = 30$ metros, resulta:

$$t_2 - t_1 = (30 \text{ m}) \left[\frac{36 \text{ m.s}^{-1}}{(3 \times 10^8 \text{ m.s}^{-1})^2} \right] \gamma = 1,2 \times 10^{-14} \text{ s}.$$

Intervalo de tiempo muy pequeño pero diferente de cero: el observador en tierra no aprecia que los eventos fueron simultáneos.

Este valor tan pequeño, obviamente, es debido a que el factor γ es casi la unidad. Notemos que si el tren viajara a una velocidad cercana a c , digamos $v = 0,99999999 c$, entonces:

$$\frac{v}{c} = 0,99999999 \rightarrow \frac{v^2}{c^2} = 0,99999998 \rightarrow \gamma = \frac{1}{\sqrt{1 - 0,99999998}} \approx 7071,06780.$$

por lo que:

$$t_2 - t_1 = (30 \text{ m}) \left[\frac{36 \text{ m.s}^{-1}}{(3 \times 10^8 \text{ m.s}^{-1})^2} \right] \gamma = 8,49 \times 10^{-11} \text{ s}.$$

En vista de lo anterior, se podría esperar que ocurran cosas curiosas con las longitudes de los objetos. Supongamos que el observador que viaja en el tren coloca su laptop en el piso y mide su longitud. Esto significa que la computadora está en reposo para O' y, en la dirección del eje x , la laptop mide $x'_2 - x'_1$ según O' . Para el observador que está en tierra la laptop viaja a una velocidad vt , de manera que:

$$x'_2 - x'_1 = [(x_2 - x_1) - v(t_2 - t_1)] \gamma,$$

Para el observador que se encuentra en reposo la laptop mide $x_2 - x_1$, medida en el instante $t_2 = t_1$. Por lo tanto:

$$x_2 - x_1 = \frac{x'_2 - x'_1}{\gamma}.$$

¡La laptop se contrae por el factor γ !



3.4.3 Ejercicios

1. En el espacio euclíadiano 3D, y en coordenadas cartesianas, no distinguimos entre vectores y 1-formas debido a que sus componentes transforman de la misma manera. Demuestre que:
 - (a). $a^{i'} = \Lambda_j^{i'} a^j \wedge b_{j'} = \Lambda_j^i b_i$, son la misma transformación si la matriz $\Lambda_j^{i'}$ es igual a la transpuesta de su inversa, es decir, si es ortogonal.
 - (b). Considere dos observadores $O : x, y \leftrightarrow x^1, x^2$ y $O' : x', y' \leftrightarrow x'^1, x'^2$ y sus sistemas de coordenadas asociados.

I. Considere la siguiente transformación de coordenadas de Galileo:

$$x^{1'} = v^1 t + \frac{\sqrt{2}}{2} x^1 - \frac{\sqrt{2}}{2} x^2 \quad \text{y} \quad x^{2'} = \frac{\sqrt{2}}{2} x^1 + \frac{\sqrt{2}}{2} x^2,$$

con v^1 una constante que representa la velocidad relativa entre $O - O'$, y t al tiempo (parámetro de esta transformación).

A continuación suponga una partícula que describe un movimiento respecto a O siguiendo una trayectoria recta, esto es $x^2 = \alpha x^1$, donde α es una constante. Encuentre cómo lo describiría el observador O' respecto a sus coordenadas (x'^1, x'^2) .

II. Considere ahora la generalización de la transformación de coordenadas anterior:

$$x^{1'} = v^1 t + \frac{\sqrt{2}}{2} x^1 - \frac{\sqrt{2}}{2} x^2 \quad \text{y} \quad x^{2'} = v^2 t + \frac{\sqrt{2}}{2} x^1 + \frac{\sqrt{2}}{2} x^2,$$

con v^1 y v^2 las componentes de una velocidad relativa entre $O - O'$, y t al tiempo. Muestre que

la norma de cualquier vector queda invariante respecto a una transformación de coordenadas como la anterior y encuentre la matriz de transformación.

2. Un acelerador produce partículas que tienen una vida promedio de $5\mu s$. Estas logran alcanzar una velocidad de $0,6c$.
 - (a). Un observador que se encuentra en reposo en el laboratorio ¿qué tiempo de vida media le atribuirá a las partículas?
 - (b). ¿Qué distancia promedio viajan las partículas en el acelerador?
 - (c). Si suponemos que existe un observador en reposo con respecto a la partícula ¿Qué distancia se desplazará éste observador antes de que la partícula se desintegre?
3. Un elemento de área cuadrada se encuentra en reposo para un observador O . Encuentre el área medida por un observador en movimiento O' si éste lleva una velocidad de $0,85c$ a lo largo de la diagonal del cuadrado.
4. Consideremos un vehículo se mueve con una velocidad V con respecto a un observador fijo en tierra O , y un pasajero dentro del vehículo que se mueve con una velocidad v' con respecto al vehículo. Sabemos de los cursos básicos que la velocidad del pasajero respecto al observador fijo en tierra es simplemente la suma: $v = v' + V$. Nos podemos preguntar sobre la forma en que se suman las velocidades en la teoría especial de la relatividad, es decir, ¿Cómo se suman las velocidades si consideramos las transformaciones de Lorentz? Demuestre que se suman como:

$$v = \frac{v' + V}{1 + v' \frac{V}{c^2}}.$$

5. Dado un espacio minkowskiano y un observador O que describe los eventos en el espacio-tiempo respecto a un sistema de coordenadas $\{x^\alpha\}$, donde $\alpha = 0, 1, 2$, y $\eta = \text{diag}[-1, 1, 1]$ el tensor métrico. Considere entonces la siguiente transformación de coordenadas:

$$x^{0'} = \gamma(x^0 - \beta x^1), \quad x^{1'} = \gamma(x^1 - \beta x^0) \quad \text{y} \quad x^{2'} = x^2, \quad \text{con } \gamma = \frac{1}{\sqrt{1 - \beta^2}}.$$

Donde $\beta = v/c$ es la velocidad relativa entre O y O' .

- (a). Otra vez, suponga que una partícula describe una linea recta respecto a O : $x^2 = \alpha x^1$, con α igual a una constante. Encuentre cómo lo describiría el otro observador O' respecto a sus coordenadas $(x^{0'}, x^{1'}, x^{2'})$.
- (b). Encuentre la expresión para la transformación de coordenadas, $\frac{\partial x^{\alpha'}}{\partial x^\beta} = \Lambda_\beta^{\alpha'}$ (transformación de Lorentz) entre estos sistemas relativistas y muestre como la norma: $x^\alpha x_\alpha = x^\alpha x^\beta \eta_{\alpha\beta}$, de cualquier vector se conserva.
6. Considere el tensor de Maxwell definido como:

$$F_{\mu\alpha} = \begin{pmatrix} 0 & E^x & E^y & E^z \\ -E^x & 0 & -cB^z & cB^y \\ -E^y & cB^z & 0 & -cB^x \\ -E^z & cB^y & cB^x & 0 \end{pmatrix}, \quad \text{otra vez con: } \eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

donde $\mathbf{E} = (E^x, E^y, E^z)$ y $\mathbf{B} = (B^x, B^y, B^z)$ son los campos eléctricos y magnéticos (respectivamente), medidos por un observador O .

- (a). Si un observador mide un campo eléctrico $\mathbf{E} = E^x \hat{\mathbf{i}}$ y ningún campo magnético ¿Cuáles campos, $F_{\mu\alpha}$, medirá otro observador que viaja con una velocidad respecto al primero de $\beta = v \hat{\mathbf{i}}$?
- (b). Muestre que las ecuaciones de Maxwell: $\nabla \times \mathbf{B} - \frac{\partial}{\partial t} \mathbf{E} = 4\pi \mathbf{J}$, y $\nabla \cdot \mathbf{E} = 4\pi \rho$ se pueden expresar

como:

$$\frac{\partial}{\partial x^\nu} F^{\mu\nu} \equiv F^{\mu\nu}_{,\nu} = 4\pi J^\mu, \quad \text{donde } J^\mu = (c\rho, J^1, J^2, J^3) \quad \text{y } \mathbf{J} = (J^1, J^2, J^3),$$

(c). Considere la identidad de Bianchi de la forma:

$$\frac{\partial F_{\mu\nu}}{\partial x^\gamma} + \frac{\partial F_{\nu\gamma}}{\partial x^\mu} + \frac{\partial F_{\gamma\mu}}{\partial x^\nu} \equiv \partial_\gamma F_{\mu\nu} + \partial_\mu F_{\nu\gamma} + \partial_\nu F_{\gamma\mu} \equiv F_{\mu\nu,\gamma} + F_{\nu\gamma,\mu} + F_{\gamma\mu,\nu} = 0,$$

y demuestre que las otras dos ecuaciones $\nabla \cdot \mathbf{B} = 0$ y $\nabla \times \mathbf{E} - \frac{\partial}{\partial t} \mathbf{B} = 0$, también están contenidas en las expresiones $F^{\mu\nu}_{,\nu} = 4\pi J^\mu$

7. El dual de un tensor \mathbf{B} de rango 2, cuadridimensional, puede ser definido de manera que sus componentes vienen dadas por:

$$\mathcal{B}^{ij} = \frac{1}{2!} \epsilon^{ijkl} B_{kl}.$$

Muestre que \mathcal{B} transforma como:

- (a). Un tensor de rango 2 bajo rotaciones.
- (b). Un pseudotensor bajo inversiones.

8. Construya $\mathcal{F}^{\mu\alpha}$, el dual del tensor de Maxwell $F^{\mu\alpha}$.
 9. Demuestre que $c^2 \mathbf{B}^2 - \mathbf{E}^2$ es un escalar de Lorentz.
 10. Resuelva los ejercicios anteriores utilizando **Maxima**.

Practicando con Maxima

1. En los grandes aceleradores de partículas se puede producir piones cargados cuando se hace colisionar protones de gran energía con algún blanco preparado para este fin. Se conoce que los piones tienen una vida media muy corta: $1,77 \times 10^{-8}$ seg. Lo que esto significa es, que una vez que se producen los piones, la mitad de ellos se habrá desintegrado en un tiempo de $1,77 \times 10^{-8}$ seg.

Los experimentos en el acelerador reflejan el hecho de que la intensidad del haz de piones que emerge del blanco, y que viaja a $0,99 c$, se reduce a la mitad una vez que recorren 37,2 mts. Pero sucede que si los piones viajan a $0,99 c = 2,968 \times 10^8$ m/s cuando estos hayan decaído a la mitad la distancia que habrán recorrido es $d = vt$:

$$(\%i1) \quad d:(2.968*10^8*(m/s))*(1.77*10^{-8}*s); \\ (\%o1) \quad 5,25336 \text{ m}$$

¿Y entonces por qué el experimento mide que viajaron 37,2 mts? Bueno, no estamos considerando los efectos relativistas que están contenidas en las transformaciones de Lorentz:

$$t_2 - t_1 = (t'_2 - t'_1)\gamma \Rightarrow \Delta t = \Delta t' \gamma,$$

donde $\Delta t'$ es el tiempo propio, el tiempo medido por un reloj respecto a la partícula en movimiento, es decir, el sistema de referencia donde el pion se encuentra en reposo.

Calculemos, primeramente el factor γ :

$$(\%i2) \quad \text{gamma}:1/sqrt(1-(0.99)^2); \\ (\%o2) \quad 7,088812050083354$$

Por lo tanto, la vida media de los piones medida desde el laboratorio es:

$$(\%i3) \quad \text{Delta_t}:(1.77*10^{-8}*s)*\text{gamma}; \\ (\%o3) \quad 1,254719732864754 \times 10^{-7} \text{ s}$$

Según las personas que miden el experimento, y que están en el sistema en reposo, los piones viven $\approx 1,255 \times 10^{-7}$ s, y por lo tanto viajan una distancia: $x = (0,99c)\Delta t$

(%ci4) $(2.968*10^8*(m/s))*\text{Delta_t};$
 (%o4) $37,24008167142589 \text{ m}$

2. Procederemos ahora a resolver con **Maxima** el sistema que nos condujo a las transformaciones impropias de Lorentz.

Primero, escribamos una lista que contiene el sistema de ecuaciones:

(%ci5) $\text{sis}:[L11^2 - c^2 L01^2 = 1, L11^2 + c^2 L01 * L00 = 0, c^2 L00^2 - v^2 L11^2 = c^2];$
 (%o5) $[L_{11}^2 - c^2 L_{01}^2 = 1, v L_{11}^2 + c^2 L_{00} L_{01} = 0, c^2 L_{00}^2 - v^2 L_{11}^2 = c^2]$

Es importante agregar la información de que $v > 0$, y que además $v < c$, para evitar soluciones complejas. El comando **assume** nos permite restringir el valor de los parámetros.

(%ci6) $\text{assume}(v>0,c>v);$
 (%o6) $[v > 0, c > v]$

Procedemos a resolver el sistema de ecuaciones, lo podemos hacer con ayuda del comando **solve**:

(%ci7) $S:\text{solve}(\text{sis},[L11,L00,L01]);$
 (%o7) $\left[\begin{array}{l} \left[L_{11} = \frac{c}{\sqrt{c^2 - v^2}}, L_{00} = \frac{c}{\sqrt{c^2 - v^2}}, L_{01} = \frac{\sqrt{c - v} v \sqrt{v + c}}{c v^2 - c^3} \right], \\ \left[L_{11} = \frac{c}{\sqrt{c^2 - v^2}}, L_{00} = -\frac{c}{\sqrt{c^2 - v^2}}, L_{01} = -\frac{\sqrt{c - v} v \sqrt{v + c}}{c v^2 - c^3} \right], \\ \left[L_{11} = -\frac{c}{\sqrt{c^2 - v^2}}, L_{00} = \frac{c}{\sqrt{c^2 - v^2}}, L_{01} = \frac{\sqrt{c - v} v \sqrt{v + c}}{c v^2 - c^3} \right], \\ \left[L_{11} = -\frac{c}{\sqrt{c^2 - v^2}}, L_{00} = -\frac{c}{\sqrt{c^2 - v^2}}, L_{01} = -\frac{\sqrt{c - v} v \sqrt{v + c}}{c v^2 - c^3} \right] \end{array} \right]$

Seleccionamos la primera de las listas anteriores y factorizamos con la rutina **factor**:

(%ci8) $\text{map}(\text{factor},S[1]);$
 (%o8) $\left[L_{11} = \frac{c}{\sqrt{c^2 - v^2}}, L_{00} = \frac{c}{\sqrt{c^2 - v^2}}, L_{01} = \frac{\sqrt{c - v} v}{c (v - c) \sqrt{v + c}} \right]$

3.5 Funcionales lineales y Distribuciones

En esta sección discutiremos el concepto de distribución utilizando el lenguaje de formas lineales que hemos expuesto en la sección 3.1 y luego haremos una extensión de las bases discretas a bases continuas. Las distribuciones son objetos que generalizan la noción convencional de funciones en el análisis matemático y permiten diferenciar funciones cuyas derivadas no existen en el sentido tradicional del cálculo diferencial.

Las distribuciones también son importantes en Física e Ingeniería, donde muchos problemas conducen a ecuaciones diferenciales cuyas soluciones (con condiciones iniciales o de frontera) son distribuciones. Consideré, por ejemplo el concepto de impulso

$$\Delta \vec{p} = \vec{p}_{t_0+\tau} - \vec{p}_{t_0} = \int_{t_0}^{t_0+\tau} F(t) dt \quad \Rightarrow \mathcal{I} = \vec{p}_\tau = \int_0^\tau F(t) dt,$$

donde una fuerza $F(t)$ actúa sobre un cuerpo durante un intervalo muy corto de tiempo y cambia su cantidad de movimiento $\vec{p}_{t_0} \rightarrow \vec{p}_{t_0+\tau}$. La cantidad físicamente importante es el valor de la integral y puede asociarse al impulso que se le provee al cuerpo para sacarlo del reposo.

El caso más emblemático de las distribuciones lo constituye la función delta de Dirac, una “función” que es cero en todo punto menos en uno que infinita. Si bien ese tipo de funciones toma el nombre de Dirac, fueron propuestas a finales del siglos XIX por Fourier y Cauchy a finales del siglo XIX.

A continuación las presentamos siguiendo el esquema de Sovolev haciendo uso del concepto de funcional lineal que discutimos en la sección 3.1.

3.5.1 Funciones de prueba

Como hemos dicho las distribuciones de Schwartz, funciones generalizadas o, simplemente distribuciones, son objetos que generalizan la noción clásica de funciones en el análisis matemático. Se utilizan ampliamente en la teoría de las ecuaciones diferenciales parciales, donde puede ser más fácil establecer la existencia de soluciones distribucionales que de soluciones clásicas. Comenzaremos con definir funciones de prueba y el soporte para estas funciones.

Consideremos la siguiente función

$$\phi(x) = \begin{cases} 0 & \text{si } |x| \geq 1 \\ \exp\left(-\frac{1}{1-x^2}\right) & \text{si } |x| < 1 \end{cases}.$$

Diremos que una función (o campo escalar), $\phi(\mathbf{x}) = \phi(x_1, x_2, x_3, \dots, x_n)$, es de prueba con soporte $|x| \leq 1$ en \mathbb{R}^n cuando: sea infinitamente derivable en un subconjunto cerrado K de \mathbb{R}^n e idénticamente nula fuera éste. Mas aún, hay funciones cuyo soporte puede ser un solo punto en el dominio. Además, nos podemos convencer fácilmente que estas funciones $\psi = c^1\phi_1 + c^2\phi_2$, forman un espacio vectorial sobre \mathbb{R}^n lo representaremos como V y a K se le denomina el soporte de ϕ .

3.5.2 Distribuciones y Deltas de Dirac

Utilizaremos el concepto de funcional que discutimos en la sección 3.1 para definir una distribución de la forma

$$\mathcal{F}[|\phi\rangle] \in \mathbb{C} \quad \Leftrightarrow \quad \mathcal{F}[|\phi\rangle]_f \equiv \langle f | \phi \rangle = \int_{\mathbb{R}^n} f(\mathbf{x}) \phi(\mathbf{x}) d^n \mathbf{x},$$

donde $\phi(\mathbf{x})$ es una función de prueba.

Siguiendo este esquema podemos definir la distribución $\delta_{(0)}$ definida en \mathbb{R} , como

$$\mathcal{F}[\langle \phi \rangle]_{\delta_{(0)}} \equiv \langle \delta_{(0)} | \phi \rangle = \int_{\mathbb{R}} \delta_{(0)} \phi(x) dx \equiv \int_{\mathbb{R}} \delta(x) \phi(x) dx = \phi(0), \quad \forall \phi(x_0) \in \mathbb{R}.$$

El soporte de $\delta_{(0)}$ es el punto $x = 0$ de \mathbb{R} . Esto significa que **únicamente** en ese punto la función es distinta de cero. En Física este tipo de objetos se conoce como **Deltas de Dirac** y los encontramos frecuentemente, cuando consideramos una masa (o una carga) puntual su función densidad de masa (o de carga) asociada: es infinita en un punto y cero en el resto del espacio. La **delta de Dirac** generaliza la delta de Kronecker δ_{ij} que definimos en la sección 1.4.1 como una cantidad que toma valores discretos: $\delta_{ij} = 1$ si $i = j$ y $\delta_{ij} = 0$ si $i \neq j$. La Delta de Dirac, $\delta(x - x_0) \equiv \delta(x, x_0)$, es la extensión de este objeto matemático cuando los índices toman valores continuos. En la próxima sección 3.5.6, haremos uso de esta propiedad para construir bases continuas de un espacio vectorial.

Para convencernos de la existencia de este tipo de funciones en Matemáticas podemos considerar una sucesión funciones gaussianas:

$$\delta_n(x) = \sqrt{\frac{n}{\pi}} e^{-nx^2},$$

tal que:

$$\int_{-\infty}^{\infty} \delta_n(x) dx = 1, \quad n > 0.$$

A medida que el parámetro n aumenta su valor, la curva gaussiana $\delta_n(x)$ será cada vez más alta y más angosta, y si $n \rightarrow \infty$ la curva será entonces infinitamente más alta e infinitamente más angosta, pero siempre conservando que el área bajo la curva será igual a la unidad. Cada una de las funciones $\delta_n(x)$ es infinitamente derivable.

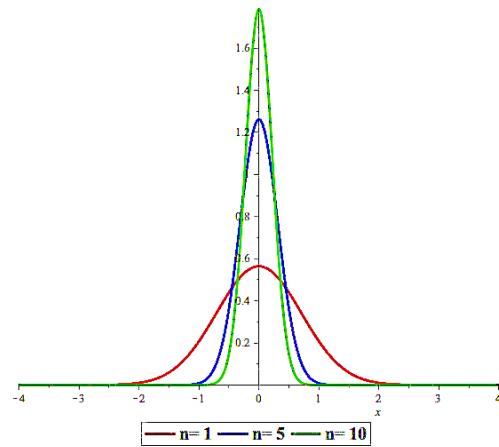


Figura 3.5: $\delta_n(x)$ para tres valores diferentes de n .

Esto nos permite definir la “función” Delta de Dirac como el límite de una sucesión de la forma

$$\delta(x) = \lim_{n \rightarrow \infty} \delta_n \equiv \lim_{n \rightarrow \infty} \sqrt{\frac{n}{\pi}} e^{-nx^2} \Rightarrow \begin{cases} \delta_{(0)} \equiv \delta(x) \rightarrow \infty & \text{si } x \rightarrow 0 \\ \delta(x) = 0 & \text{si } x \neq 0 \end{cases}$$

En el límite $n \rightarrow \infty$ tendremos una “función” que se anula en todos los puntos $x \neq 0$ e infinitamente derivable. Nótese que, siguiendo la costumbre en Física, aquí estamos escribiendo $\delta_{(0)} = \delta(x)$.

3.5.3 Distribuciones y sucesiones

La forma que le dimos a la sucesión δ_j es muy particular. En general uno podría definir las distribuciones como el límite de una familia, $\{f_j(x)\}$, de funciones localmente integrables en \mathbb{R}^n de la forma

$$f(x) = \lim_{j \rightarrow j_0} f_j \Rightarrow \lim_{j \rightarrow j_0} \mathcal{F}[\langle \phi \rangle]_{f_j} \equiv \lim_{j \rightarrow j_0} \langle f_j | \phi \rangle = \lim_{j \rightarrow j_0} \int_{\mathbb{R}^n} f_j(\mathbf{x}) \phi(\mathbf{x}) d^n \mathbf{x}.$$

Así, una expresión equivalente a la que ensayamos anteriormente

$$f_j(x) = \frac{e^{-\frac{x^2}{4j}}}{\sqrt{4\pi j}} \quad \text{para } t \rightarrow 0,$$

se conoce con el nombre de kernel de Gauss. Mientras que la familia

$$\{f_r(\theta)\} \Rightarrow f_r(\theta) = \begin{cases} \frac{1}{2\pi} \left(\frac{1-r^2}{1+r^2-2r \cos \theta} \right) & \text{si } |\theta| \leq \pi \\ 0 & \text{si } |\theta| > \pi \end{cases} \quad \text{para } r \rightarrow 1,$$

se le conoce como el kernel de Poisson. Finalmente la familia $\{f_k(x)\}$

$$f_k(x) = \begin{cases} \sum_{m=-k}^k \frac{1}{2\pi} e^{imx} & \text{si } |x| \leq \pi \\ 0 & \text{si } |x| > \pi \end{cases}, \quad \text{para } k \rightarrow \infty,$$

constituirá el kernel de Dirichlet:

De igual forma podremos construir la distribución $\delta_{(x_0)}$ con soporte en el punto x_0 de \mathbb{R} , definida por

$$\mathcal{F}[|\phi\rangle]_{\delta_{(x_0)}} \equiv \langle \delta_{(x_0)} | \phi \rangle = \int_{\mathbb{R}} \delta_{(x_0)} \phi(x) dx \equiv \int_{\mathbb{R}} \delta(x - x_0) \phi(x) dx = \phi(x_0), \quad \forall \phi(x_0) \in \mathbb{R}.$$

Igual que en caso anterior, en este caso podemos imaginar una sucesión de funciones gausianas donde x_0 es el punto donde alcanza su valor máximo. Entonces

$$\delta_n(x - x_0) = \sqrt{\frac{n}{\pi}} e^{-n(x-x_0)^2} \Rightarrow \delta(x - x_0) = \lim_{n \rightarrow \infty} \sqrt{\frac{n}{\pi}} e^{-n(x-x_0)^2}.$$

3.5.4 Derivadas de distribuciones

Si consideramos una distribución $\mathcal{F}[\phi(x)]_f$ y su derivada $\mathcal{F}[\phi(x)]_{f'}$, entonces se cumple que

$$\mathcal{F}[\phi(x)]_{f'} = -\mathcal{F}[\phi'(x)]_f \Leftrightarrow \int_{\mathbb{R}} f'(x) \phi(x) dx = - \int_{\mathbb{R}} f(x) \phi'(x) dx.$$

La demostración es clara cuando utilizamos la integración por parte el hecho de que las funciones de prueba tienen soporte acotado. De esta manera

$$\int_{\mathbb{R}} f'(x) \phi(x) dx = \underbrace{\int_{-\infty}^{\infty} f'(x) \phi(x) dx}_{=0} - \int_{\mathbb{R}} f(x) \phi'(x) dx \Rightarrow \mathcal{F}[\phi(x)]_{f'} = -\mathcal{F}[\phi'(x)]_f.$$

El caso mas representativo de esta propiedad es cuando nos planteamos calcular la derivada para la función escalón, $\Theta(x)$ de *Heaviside* en \mathbb{R} . Una función localmente integrable, tal que $\Theta(x) = 0$ cuando $x < 0$ y $\Theta(x) = 1$ cuando $x > 0$, entonces

$$\mathcal{F}[\phi(x)]_{\Theta(x)'} = -\mathcal{F}[\phi'(x)]_{\Theta(x)} = - \int_{-\infty}^{\infty} \Theta(x) \frac{d\phi(x)}{dx} dx = - \int_0^{\infty} \frac{d\phi(x)}{dx} dx = \phi(0) \Leftrightarrow \frac{d\Theta(x)}{dx} = \delta(x),$$

en el sentido de las distribuciones, porque $\phi(x)$ es una función de prueba con soporte compacto: $\lim_{x \rightarrow \infty} \phi(x) = 0$.

3.5.5 Algunas propiedades de las distribuciones

Como construimos las distribuciones como funcionales y demostramos en la sección 3.1.1, las distribuciones formarán un espacio vectorial dual. De esta forma

$$\mathcal{F}[|\phi\rangle]_f + \mathcal{F}[|\phi\rangle]_g = \mathcal{F}[|\phi\rangle]_{f+g} \Leftrightarrow \mathcal{F}[|\phi\rangle]_f + \mathcal{F}[|\xi\rangle]_f = \mathcal{F}[|\phi + \xi\rangle]_f.$$

En el primer caso mostramos que la suma de distribuciones es una distribución y en el segundo caso que la suma de dos funciones de prueba es también una función de prueba.

Hay algunas propiedades escalamiento para las distribuciones que se pueden derivar de manera inmediata, y las mostraremos para el caso particular la Delta de Dirac en \mathbb{R} . Así tendremos

$$\langle \delta(x/\alpha) | \phi(x) \rangle = \int_{\mathbb{R}} \delta(x/\alpha) \phi(x) dx = |\alpha| \int_{\mathbb{R}} \delta(y) \phi(\alpha y) dy = |\alpha| \langle \delta(x) | \phi(x) \rangle,$$

que se puede escribir como $\delta(x/\alpha) = |\alpha| \delta(x/\alpha)$ en el sentido distribucional.

Como hemos construido la noción de distribución a partir del concepto de funcionales lineales, podemos también utilizar el producto tensorial definido en la sección 3.2.2 para extender las distribuciones a varias dimensiones. Así, en coordenadas cartesianas

$$\mathcal{F}[\langle \phi \rangle]_{\delta_{(x_0)}} \otimes \mathcal{F}[\langle \phi \rangle]_{\delta_{(y_0)}} \otimes \mathcal{F}[\langle \phi \rangle]_{\delta_{(z_0)}} = \int_{\mathbb{R}^3} \delta(x - x_0) \delta(y - y_0) \delta(z - z_0) \phi(x, y, z) dx dy dz = \phi(x_0, y_0, z_0),$$

Resumimos aquí algunas de las propiedades de la delta de Dirac expresadas en el sentido distribucional:

- $\delta(x - x_0) = \delta(x_0 - x)$
- $x\delta(x) = 0$
- $\delta(x^2 - a^2) = \frac{1}{2|a|} [\delta(x - a) + \delta(x + a)]$
- $\delta(\mathbf{r} - \mathbf{r}_0) = \delta(x - x_0)\delta(y - y_0)\delta(z - z_0)$
- $\bullet \delta(ax) = \frac{\delta(x)}{|a|}$
- $f(x)\delta(x - x_0) = f(x_0)\delta(x - x_0)$
- $x^{n+1} \frac{d^n \delta(x)}{dx^n} = 0$
- $\int f(\mathbf{r})\delta(\mathbf{r} - \mathbf{r}_0)dV = f(\mathbf{r}_0)$

Ejemplo 3.17 Una distribución de carga eléctrica puede escribirse de manera sencilla utilizando la delta de Dirac a través de la siguiente integral:

$$q = \int_V \rho(\mathbf{r}) dV,$$

de manera que para una carga puntual localizada en $\mathbf{r} = \mathbf{r}_0$ se tiene que:

$$q = q \underbrace{\int \delta(\mathbf{r} - \mathbf{r}_0) dV}_{=1} = \int_V \rho(\mathbf{r}) dV \Rightarrow \rho(\mathbf{r}) = q\delta(\mathbf{r} - \mathbf{r}_0).$$



Ejemplo 3.18 Demostremos que:

$$\delta(\alpha x) = \frac{\delta(x)}{\alpha}, \quad \alpha > 0.$$

Entonces, haremos el siguiente cambio de variable: $y = \alpha x$.

$$\int_{-\infty}^{\infty} f(x)\delta(\alpha x) dx = \int_{-\infty}^{\infty} f\left(\frac{y}{\alpha}\right) \delta(y) \frac{1}{\alpha} dy = \frac{1}{\alpha} \int_{-\infty}^{\infty} f\left(\frac{y}{\alpha}\right) \delta(y) dy = \frac{f(0)}{\alpha} = \int_{-\infty}^{\infty} f(x)\delta(x) dx,$$

por lo tanto: $\delta(\alpha x) = \frac{\delta(x)}{\alpha}$.

La condición $\alpha > 0$ es porque se debe tener en cuenta que la función delta de Dirac es par, es decir, $\delta(\alpha x) = \delta(-\alpha x)$. De manera que resulta más apropiado escribir:

$$\delta(\alpha x) = \frac{\delta(x)}{|\alpha|}.$$



Ejemplo 3.19 Demostremos ahora que:

$$\delta(g(x)) = \sum_{\alpha} \frac{\delta(x - \alpha)}{|g'(\alpha)|}, \quad g(\alpha) = 0 \wedge g'(\alpha) \neq 0.$$

Podemos ver que para la función $g(x)$, en los intervalos donde se hace cero, se tiene:

$$g(x) \approx g(\alpha) + (x - \alpha)g'(\alpha), \quad -\varepsilon < \alpha < \varepsilon.$$

De manera que si sumamos para todos esos intervalos:

$$\int_{-\infty}^{\infty} f(x)\delta(g(x)) dx = \sum_{\alpha} \int_{\alpha-\varepsilon}^{\alpha+\varepsilon} f(x)\delta((x - \alpha)g'(\alpha)) dx = \int_{-\infty}^{\infty} f(x) \sum_{\alpha} \frac{\delta(x - \alpha)}{|g'(\alpha)|} dx,$$

donde hemos utilizado el resultado del ejemplo anterior. Por lo tanto:

$$\delta(g(x)) = \sum_{\alpha} \frac{\delta(x - \alpha)}{|g'(\alpha)|}.$$

△

3.5.6 Bases continuas

Tal y como vimos anteriormente, la representación de un vector $|a\rangle$ en un espacio vectorial abstracto V puede darse en término de una base ortonormal de vectores (discreta y finita $B_{DF} = \{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_n\rangle\}$ o discreta e infinita $B_{DI} = \{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_n\rangle, \dots\}$) de la forma:

$$|a\rangle = \begin{cases} c^i |\hat{e}_i\rangle = \langle \hat{e}^i | a \rangle |\hat{e}_i\rangle & \Leftarrow B_{DF} = \{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_n\rangle\} \\ c^i |\hat{e}_i\rangle = \langle \hat{e}^i | a \rangle |\hat{e}_i\rangle & \Leftarrow B_{DI} = \{|\hat{e}_1\rangle, |\hat{e}_2\rangle, |\hat{e}_3\rangle, \dots, |\hat{e}_n\rangle, \dots\} \end{cases}$$

donde en ambos casos:

$$c^i = \langle \hat{e}^i | a \rangle = c^j \langle \hat{e}^i | \hat{e}_j \rangle = c^j \delta_j^i.$$

Recapitulemos ahora algunos puntos que hemos tratado con anterioridad, y con el fin de aclarar conceptos, para el caso de bases discretas de funciones $\{|e_i(x)\rangle\}$.

La relación de ortogonalidad viene dada, como hemos mencionado, por:

$$\langle e^n(x) | e_m(x) \rangle = \int_a^b e_n^*(x) e_m(x) dx = \Theta_n \delta_m^n, \quad (3.18)$$

y con una norma definida como:

$$\| |e_i(x)\rangle \| ^2 = \langle e_n(x) | e_n(x) \rangle = \int_a^b |e_n(x)|^2 dx = \Theta_n. \quad (3.19)$$

La base será ortonormal si $\Theta_n = 1$.

Con una base $\{|e_i(x)\rangle\}$ definida, es posible representar cualquier función $|f(x)\rangle$ cuadrado integrable, es decir que cumpla con:

$$\int_a^b |f(x)|^2 dx \neq \infty, \quad (3.20)$$

en una combinación lineal de la base dada:

$$|f(x)\rangle = C^i |e_i(x)\rangle \quad (3.21)$$

Las cantidades C^i , las componentes de $|f(x)\rangle$, se obtienen de la siguiente manera:

$$\langle e^i(x) | f(x) \rangle = C^j \langle e^i(x) | e_j(x) \rangle = C^j \delta_j^i = C^i = \int_a^b f(x) e_i^*(x) dx, \quad (3.22)$$

Resulta conveniente para los cálculos que vienen a continuación cambiar $i \rightarrow n$ y $x \rightarrow x'$, de manera que:

$$C^n = \int_a^b f(x') e_n^*(x') dx'. \quad (3.23)$$

Por lo tanto:

$$\begin{aligned} |f(x)\rangle = C^n |e_n(x)\rangle \Rightarrow |f(x)\rangle &= \sum_n \left[\int_a^b f(x') e_n^*(x') dx' \right] |e_n(x)\rangle = \sum_n \int_a^b f(x') e_n^*(x') e_n(x) dx' \\ &= \int_a^b f(x') \sum_n e_n^*(x') e_n(x) dx'. \end{aligned}$$

Utilizando la delta de Dirac:

$$|f(x)\rangle = \int_a^b f(x')\delta(x-x')dx' = \int_a^b f(x') \left[\sum_n e_n^*(x')e_n(x) \right] dx' \Rightarrow \delta(x-x') = \sum_n e_n^*(x')e_n(x).$$

Esta última relación viene a ser una generalización de la relación de cierre mencionada en la sección 2.3.4, y se conoce como *condición de completitud* para la base $\{|e_i(x)\rangle\}$, además, también resulta ser una representación en serie para la delta de Dirac.

De todo esto surgen algunas propiedades fundamentales:

- Se dice que el conjunto $\{|e_i(x)\rangle\}$ es *completo* si para $|f(x)\rangle = C^i|e_i(x)\rangle$, entonces se tiene que:

$$\int_a^b |f(x)|^2 = \sum_n |C^n|^2.$$

- Si $|f(x)\rangle = C^i|e_i(x)\rangle$ y $|g(x)\rangle = E^i|e_i(x)\rangle$, entonces $\langle f(x)|g(x)\rangle = C_n^*E^n$.

Es posible pasar de estos espacios vectoriales discretos, donde las bases son un conjunto numerable de elementos $\{e_n(x)\}$, a espacios vectoriales de funciones con las bases dadas por funciones y donde el índice n pasa a ser una variable continua.

Recordemos que en la sección 2.3.4 hablábamos de un conjunto de funciones $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle, \dots\}$ definidas por:

$$|e_0\rangle = 1, \quad |e_{2n-1}\rangle = \cos(nx) \quad y \quad |e_{2n}\rangle = \sin(nx), \quad \text{con } n = 1, 2, 3, \dots,$$

la base discreta de Fourier. Base que también podemos escribir de la forma compleja, como se muestra a continuación:

$$|\hat{e}_n(x)\rangle = \frac{1}{\sqrt{2L}}e^{i\frac{n\pi}{L}x} = \frac{1}{\sqrt{2L}}e^{ikx}, \quad \text{con } n = -\infty, \dots, \infty \text{ y } -L \leq x \leq L.$$

Cuando n aumenta indefinidamente, también lo hará la cantidad $k = \frac{n\pi}{L}$, convirtiéndose entonces en una variable continua, es decir, $\Delta k = \frac{\Delta n\pi}{L} \rightarrow 0$. Notemos que los índices de la delta de Kronecker $\delta_{nm} = \delta_{nn'}$ pueden tomar los valores: $n = \frac{Lk}{\pi}$, $n' = \frac{Lk'}{\pi}$, y en el proceso de $L \rightarrow \infty$ se convertirán en:

$$\delta_{nn'} \rightarrow \delta\left(\frac{Lk}{\pi} - \frac{Lk'}{\pi}\right) = \delta\left(\frac{L}{\pi}(k - k')\right) = \frac{\pi}{L}\delta(k - k').$$

De esta manera, la base ortonormal quedará escrita en función de las variables continuas k y x , como se muestra a continuación:

$$|\hat{e}(k, x)\rangle = \frac{1}{\sqrt{2\pi}}e^{ikx}.$$

El hecho de que sean ortonormales se refleja en la siguiente condición (que es una extensión de la expresión para el producto interno de bases discretas)

$$\langle \hat{e}(k, x)|\hat{e}(k, x)\rangle = \int_a^b \hat{e}(k, x)^*\hat{e}(k, x) dx = \delta(k - k'), \quad \text{con } \alpha \leq k \leq \beta, \quad a \leq x \leq b$$

Al ser el conjunto de funciones $\{|\hat{e}(k, x)\rangle\}$ una base, toda función del mismo espacio vectorial se puede expandir en esa base como:

$$|f(x)\rangle = \int_\alpha^\beta C(k)\hat{e}(k, x) dk.$$

En correspondencia con el caso de las bases discretas, el coeficiente $C(k)$ viene a ser algo parecido a las componentes de $|f(x)\rangle$. Y para calcularlas se procede como se muestra a continuación:

$$\langle \hat{e}(k', x)|f(x)\rangle = \int_a^b \hat{e}(k', x)^*f(x) dx = \int_\alpha^\beta C(k) \left[\int_a^b \hat{e}(k', x)^*\hat{e}(k, x) dx \right] dk = \int_\alpha^\beta C(k)\delta(k - k') dk = C(k').$$

Volvamos a cambiar la notación para los índices: $k' \rightarrow k$, $x \rightarrow x'$ en la última ecuación:

$$C(k) = \int_a^b \hat{e}(k, x')^* f(x') dx',$$

por lo tanto:

$$|f(x)\rangle = \int_{\alpha}^{\beta} C(k) |\hat{e}(k, x)\rangle dk = \int_a^b f(x') \left[\int_{\alpha}^{\beta} \hat{e}(k, x')^* \hat{e}(k, x) dk \right] dx' = \int_a^b f(x') \delta(x - x') dx',$$

y la relación:

$$\int_{\alpha}^{\beta} \hat{e}(k, x')^* \hat{e}(k, x) dk = \delta(x - x'),$$

será la relación de cierre para el conjunto no numerable de funciones ortonormales.

Una integral como la que nos permitió definir $C(k)$, digamos, del tipo:

$$F(k) = \int_a^b f(x') \hat{e}(k, x')^* dx' \Rightarrow |f(x)\rangle = \int_{\alpha}^{\beta} F(k) \hat{e}(k, x) dk,$$

es lo que se denomina una transformada de $f(x)$. Más adelante veremos que si la base, es por ejemplo $e^{ikx}/\sqrt{2\pi}$, la transformada será la de Fourier.

En el espacio vectorial de funciones de cuadrado integrable \mathcal{L}^2 , definidas en \mathbb{R}^3 , tendremos que:

$$|F\rangle = c^i |e_i\rangle \equiv \langle e^i | F \rangle |e_i\rangle = \sum_{i=0}^{\infty} \left(\int_{-\infty}^{\infty} d^3 r' \xi_i^* (\mathbf{r}') f(\mathbf{r}') \right) |e_i\rangle,$$

que se reescribe en términos de funciones como:

$$f(\mathbf{r}) = \sum_{i=0}^{\infty} \left(\int_{-\infty}^{\infty} d^3 r' \xi_i^* (\mathbf{r}') f(\mathbf{r}') \right) \xi_i(\mathbf{r}).$$

Es claro que se pueden intercambiar los símbolos de \int y \sum , por lo cual:

$$f(\mathbf{r}) = \int_{-\infty}^{\infty} d^3 r' f(\mathbf{r}') \underbrace{\left[\sum_{i=0}^{\infty} \xi_i^*(\mathbf{r}') \xi_i(\mathbf{r}) \right]}_{\mathcal{G}(\mathbf{r}', \mathbf{r})},$$

la función $\mathcal{G}(\mathbf{r}', \mathbf{r})$, que depende de los argumentos \mathbf{r}' y \mathbf{r} , vive dentro de las integrales y convierte:

$$f(\mathbf{r}) = \int_{-\infty}^{\infty} d^3 r' f(\mathbf{r}') \mathcal{G}(\mathbf{r}', \mathbf{r}).$$

Podemos ver entonces que:

$$f(\mathbf{r}) = \int_{-\infty}^{\infty} d^3 r' f(\mathbf{r}') \delta(\mathbf{r}' - \mathbf{r}).$$

En resumen, la generalización de bases discretas a continuas se hace transformando el índice de la sumatoria en la variable de una integral:

$$|\Psi\rangle = \int d\alpha c(\alpha) |w_{\alpha}\rangle \Rightarrow c(\beta) = \langle w^{\beta} | \Psi \rangle = \int d\alpha c(\alpha) \langle w^{\beta} | w_{\alpha}\rangle = \int d\alpha c(\alpha) \delta(\alpha - \beta).$$

Así, para los conceptos expresados hasta ahora se tiene el siguiente tabla resumen:

Propiedad \ Base	Discreta	Continua
Ortogonalidad	$\langle u^i u_j \rangle = \delta_j^i$	$\langle w^{\beta} w_{\alpha} \rangle = \delta(\alpha - \beta)$
Cierre	$1 = \sum_{j=0}^{\infty} u_j\rangle \langle u^j $	$1 = \int d\alpha w_{\alpha}\rangle \langle w^{\alpha} $
Expansión	$ F\rangle = \sum_{i=0}^{\infty} c^i u_i\rangle$	$ \Psi\rangle = \int d\alpha c(\alpha) w_{\alpha}\rangle$
Componentes	$c^i = \langle u^i F \rangle$	$c(\beta) = \langle w^{\beta} \Psi \rangle$
Producto Interno	$\langle G F \rangle = \sum_{i=0}^{\infty} g^{i*} f_i$	$\langle G F \rangle = \int d\alpha g^*(\alpha) f(\alpha)$
Norma	$\langle F F \rangle = \sum_{i=0}^{\infty} f_i ^2$	$\langle F F \rangle = \int d\alpha f(\alpha) ^2$

3.5.7 Bases de ondas planas y la transformada de Fourier

En las teorías de oscilaciones es de gran importancia considerar el problema de la transformada de Fourier. Recientemente vimos que una función $f(x)$ puede representarse de la forma:

$$|f(x)\rangle = \int_{\alpha}^{\beta} C(k) e(k, x) dk,$$

si la función $f(x)$ está definida en $(-\infty, \infty)$ y si tomamos al conjunto de vectores base al conjunto de funciones continuas y ortonormales $\{\hat{e}(k, x)\}$ como $\{e^{ikx}/\sqrt{2\pi}\}$, entonces la integral anterior la podemos escribir de la forma:

$$|f(x)\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{ikx} dk, \quad (3.24)$$

Si utilizamos la relación de ortogonalidad: $\langle e(k, x)|e(k, x)\rangle = \int e(k, x)e(k', x)^* dx = \delta(k - k')$, resulta:

$$\int_{-\infty}^{\infty} e^{i(k-k')x} dx = 2\pi\delta(k - k').$$

Donde:

$$|F(k)\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx.$$

A las variables x y k se les denominan variables conjugadas de Fourier (no conjugadas como en variable compleja) y a $F(k)$ la transformada de Fourier de $f(x)$ (y viceversa).

Notemos que si sustituimos $F(k)$ en la otra integral, ecuación (3.24), entonces:

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) e^{ik(x-x')} dk dx' \Rightarrow |f(x')\rangle = \int_{-\infty}^{\infty} f(x) \delta(x - x') dx'.$$

Por otra parte, podemos ver que la integral para $F(k)$ se puede hacer en dos partes:

$$|F(k)\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx \Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 f(x) e^{-ikx} dx + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) e^{-ikx} dx.$$

Si sustituimos $x \rightarrow -x$, en la primera integral de la ecuación anterior resulta:

$$|F(k)\rangle = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(-x) e^{ikx} dx + \frac{1}{\sqrt{2\pi}} \int_0^{\infty} f(x) e^{-ikx} dx.$$

Dos casos podemos considerar:

- Si la función es par, $f(x) = f(-x)$:

$$|F(k)\rangle = \frac{1}{\sqrt{2\pi}} \left[\int_0^{\infty} f(x) e^{ikx} dx + \int_0^{\infty} f(x) e^{-ikx} dx \right] = \sqrt{\frac{2}{\pi}} \int_0^{\infty} f(x) \cos(kx) dx,$$

entonces, de la transformada al hacer $F(k) = F(-k)$ se tiene:

$$|f(x)\rangle = \sqrt{\frac{2}{\pi}} \int_0^{\infty} F(k) \cos(kx) dk.$$

Estas dos funciones se conocen como las transformadas coseno de Fourier.

- Si la función es impar, $f(x) = -f(-x)$:

$$|F(k)\rangle = \frac{1}{\sqrt{2\pi}} \left[\int_0^{\infty} f(-x) e^{ikx} dx - \int_0^{\infty} f(-x) e^{-ikx} dx \right] = i\sqrt{\frac{2}{\pi}} \int_0^{\infty} f(x) \sin(kx) dx.$$

Ahora al hacer $F(k) = -F(-k)$ se tiene:

$$|f(x)\rangle = \sqrt{\frac{2}{\pi}} \int_0^{\infty} \tilde{F}(k) \sin(kx) dk,$$

donde $\tilde{F}(k) = iF(k)$. Tenemos ahora las transformadas seno de Fourier.

La generalización a \mathbb{R}^n es obvia. Si $f = f(\mathbf{r})$, entonces:

$$|f(\mathbf{r})\rangle = \left[\frac{1}{\sqrt{2\pi}} \right]^{\frac{1}{n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} F(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r})} d\mathbf{k},$$

donde $d\mathbf{k} = dk_1 dk_2 dk_3 \dots dk_n$. Para la transformada:

$$|F(\mathbf{k})\rangle = \left[\frac{1}{\sqrt{2\pi}} \right]^{\frac{1}{n}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\mathbf{r}) e^{-i(\mathbf{k}\cdot\mathbf{r})} d\mathbf{r},$$

donde $d\mathbf{r} = dx_1 dx_2 dx_3 \dots dx_n$, representa el elemento de volumen en \mathbb{R}^n .

3.5.7.1 Ondas planas

Como un ejemplo de lo anterior, consideraremos la base de las ondas planas. Si a k la llamaremos s y a la variable x el tiempo t , vale decir:

$$F(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dt e^{ist} f(t) \iff f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} ds e^{-ist} F(s).$$

De esta manera, a la función $F(s)$ se le denomina la distribución espectral de $f(t)$ y a $|F(s)|^2$ la densidad espectral, es decir, la intensidad de la onda en el intervalo $[s, s + \Delta s]$. Esto significa que la energía total es:

$$E = \int_{-\infty}^{\infty} |F(s)|^2 ds.$$

También podemos reescribir las transformadas en términos de la posición y el momento:

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp e^{i(px/\hbar)} \bar{\psi}(p) \iff \bar{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx e^{-i(px/\hbar)} \psi(x).$$

Hemos tenido cuidado de incluir los factores de normalización adecuados para el caso de las descripciones en Mecánica Cuántica. Estas fórmulas pueden ser reinterpretadas en función de los conceptos anteriormente expuestos y podemos definir una base continua de la forma:

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp \underbrace{\left(\frac{1}{\sqrt{2\pi\hbar}} e^{i(px/\hbar)} \right)}_{v_p(x)} \bar{\psi}(p) \iff \bar{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \underbrace{\left(\frac{1}{\sqrt{2\pi\hbar}} e^{-i(px/\hbar)} \right)}_{v_p^*(x)} \psi(x),$$

por lo cual:

$$\psi(x) = \int_{-\infty}^{\infty} dp v_p(x) \bar{\psi}(p) \iff \bar{\psi}(p) = \int_{-\infty}^{\infty} dx v_p^*(x) \psi(x).$$

Diremos que la función $\psi(x)$ está expresada en la base de ondas planas $v_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{i(px/\hbar)}$.

Nótese lo siguiente:

- El índice p de $v_p(x)$ varía de forma continua entre $-\infty$ a ∞ .
- Que $v_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{i(px/\hbar)} \notin L^2$, es decir, no pertenece al espacio vectorial de funciones de cuadrado integrable ya que su norma diverge:

$$\langle v_p | v_p \rangle = \int_{-\infty}^{\infty} dx |v_p(x)|^2 = \int_{-\infty}^{\infty} dx \frac{1}{2\pi\hbar} \rightarrow \infty.$$

- Que las proyecciones de $\psi(x)$ sobre la base de ondas planas es: $\bar{\psi}(p) = \langle v_p | \psi \rangle$.

- La relación de cierre para esta base se expresa como:

$$1 = \int d\alpha |v_\alpha\rangle \langle v_\alpha| \iff \int_{-\infty}^{\infty} dp v_p^*(x') v_p(x) = \int_{-\infty}^{\infty} dp \frac{1}{2\pi\hbar} e^{i[p(x'-x)/\hbar]} = \delta(x' - x),$$

mientras que de la definición de producto interno se obtiene:

$$\langle v_{p'} | v_p \rangle = \int_{-\infty}^{\infty} dx v_{p'}^*(x) v_p(x) = \int_{-\infty}^{\infty} dx \frac{1}{2\pi\hbar} e^{i[x(p'-p)/\hbar]} = \delta(p' - p).$$

En este mismo orden de ideas, podemos construir otra base continua $\xi_{\mathbf{r}_0}(\mathbf{r})$ a partir de la utilización de las

propiedades de la delta de Dirac. Esto es:

$$\psi(\mathbf{r}) = \int_{-\infty}^{\infty} d^3 r_0 \psi(\mathbf{r}_0) \underbrace{\delta(\mathbf{r}_0 - \mathbf{r})}_{\xi_{\mathbf{r}_0}(\mathbf{r})} \Leftrightarrow \psi(\mathbf{r}_0) = \int_{-\infty}^{\infty} d^3 r \psi(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}_0),$$

por lo cual la reinterpretación es inmediata:

$$\psi(\mathbf{r}) = \int_{-\infty}^{\infty} d^3 r_0 \psi(\mathbf{r}_0) \xi_{\mathbf{r}_0}(\mathbf{r}), \quad \text{con:} \quad \psi(\mathbf{r}_0) = \langle \xi_{\mathbf{r}_0} | \psi \rangle = \int_{-\infty}^{\infty} d^3 r \xi_{\mathbf{r}_0}^*(\mathbf{r}) \psi(\mathbf{r}),$$

más aún, la ortogonalidad queda garantizada por la relación de cierre:

$$\langle \xi_{\mathbf{r}_0} | \xi_{\mathbf{r}'_0} \rangle = \int_{-\infty}^{\infty} d^3 r_0 \xi_{\mathbf{r}_0}^*(\mathbf{r}) \xi_{\mathbf{r}'_0}(\mathbf{r}') = \int_{-\infty}^{\infty} d^3 r_0 \delta(\mathbf{r} - \mathbf{r}_0) \delta(\mathbf{r}' - \mathbf{r}_0) = \delta(\mathbf{r}' - \mathbf{r}),$$

al igual que:

$$\langle \xi_{\mathbf{r}_0} | \xi_{\mathbf{r}'_0} \rangle = \int_{-\infty}^{\infty} d^3 r \xi_{\mathbf{r}_0}^*(\mathbf{r}) \xi_{\mathbf{r}'_0}(\mathbf{r}) = \int_{-\infty}^{\infty} d^3 r \delta(\mathbf{r} - \mathbf{r}_0) \delta(\mathbf{r} - \mathbf{r}'_0) = \delta(\mathbf{r}'_0 - \mathbf{r}_0).$$

Ejemplo 3.20 Podemos ver cómo es la transformada de Fourier de la derivada de una función. Si

$$|f(x)\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{ikx} dk \Rightarrow \frac{d}{dx^n} |f(x)\rangle = \frac{(ik)^n}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{ikx} dk.$$



Ejemplo 3.21 Encontremos la transformada de Fourier de la siguiente función:

$$f(t) = \begin{cases} \sin(\omega_0 t) & -\frac{n\pi}{\omega_0} < t < \frac{n\pi}{\omega_0} \\ 0 & t < -\frac{n\pi}{\omega_0} \wedge t > \frac{n\pi}{\omega_0} \end{cases}$$

La función $f(t)$ representa un tren de ondas finito con n ciclos en el intervalo dado, como se aprecia en la figura 3.6.

Por el hecho de ser $f(t)$ una función impar, podemos utilizar la transformada seno de Fourier:

$$\begin{aligned} |F(\omega)\rangle &= \frac{2}{\sqrt{2\pi}} \int_0^{\frac{n\pi}{\omega_0}} f(t) \sin(\omega t) dt \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\frac{n\pi}{\omega_0}} \sin(\omega_0 t) \sin(\omega t) dt. \end{aligned}$$

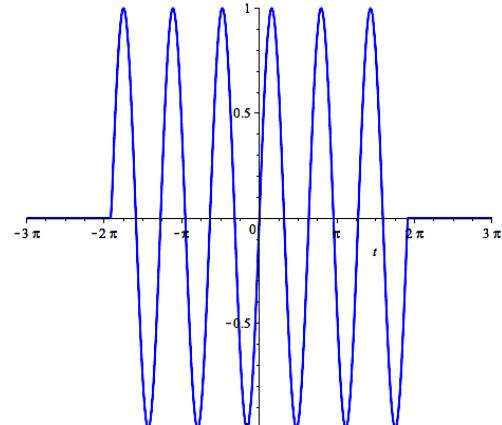


Figura 3.6: $f(t)$ con: $n = 6$ y $\omega_0 = \pi$.

Resolviendo la integral obtenemos la distribución espectral:

$$|F(\omega)\rangle = \frac{2 [\omega_0 \sin(\tau\omega) \cos(n\pi) - \omega \cos(\tau\omega) \sin(n\pi)]}{\sqrt{2\pi} (\omega^2 - \omega_0^2)} = \frac{1}{\sqrt{2\pi}} \left[\frac{\sin[\tau(\omega_0 - \omega)]}{2(\omega_0 - \omega)} - \frac{\sin[\tau(\omega_0 + \omega)]}{2(\omega_0 + \omega)} \right],$$

donde $\tau = \frac{n\pi}{\omega_0}$.



3.5.7.2 Las representaciones $|r\rangle$ y $|p\rangle$

A partir de las bases de ondas planas $v_{p_0}(x)$, y de distribuciones, $\xi_{\mathbf{r}_0}(\mathbf{r})$, construimos las llamadas representaciones $|r\rangle$ y $|p\rangle$ de la forma siguiente.

Primero asociamos:

$$\xi_{\mathbf{r}_0}(\mathbf{r}) \Leftrightarrow |r_0\rangle \wedge v_{p_0}(x) \Leftrightarrow |p_0\rangle.$$

Se puede demostrar que los límites:

$$\lim_{\omega_0 \rightarrow \infty} |F(\omega)\rangle = \lim_{\omega \rightarrow \infty} |F(\omega)\rangle = 0.$$

Se puede apreciar también que el primer término, de la expresión entre corchetes, es el de mayor relevancia debido a que en el denominador aparece la diferencia: $\omega_0 - \omega$.

En la Figura 3.7 se muestra la función de distribución y en ella podemos notar que a medida que n crece la función $|F(\omega)\rangle$ es una delta de Dirac en $\omega = \omega_0 = \pi$.

Por otro lado, es fácil ver que los ceros ocurren cuando:

$$\frac{\omega_0 - \omega}{\omega_0} = \pm \frac{1}{n}, \pm \frac{2}{n}, \pm \frac{3}{n} \dots$$

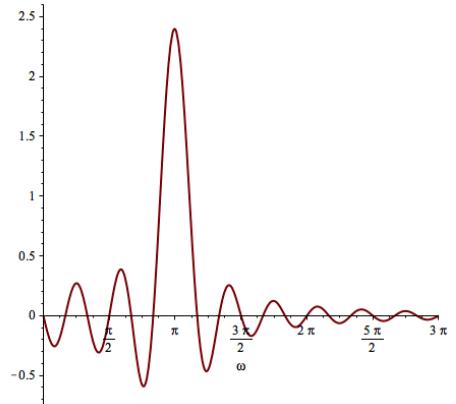


Figura 3.7: $F(\omega)$ con: $n = 6$ y $\omega_0 = \pi$.

De esta forma, dada las bases $\{\xi_{r_0}(\mathbf{r})\}$ y $\{v_{p_0}(x)\}$ para el espacio vectorial \mathbf{V} definiremos dos “representaciones”: la representación de coordenadas, $|r_0\rangle$, y la representación de momentos $|p_0\rangle$ de \mathbf{V} , respectivamente. De tal modo que:

$$\begin{aligned} \langle r_0 | r'_0 \rangle &= \int_{-\infty}^{\infty} d^3 r \xi_{r_0}^*(\mathbf{r}) \xi_{r'_0}(\mathbf{r}) = \delta(r'_0 - r_0) \Rightarrow 1 = \int d^3 r_0 |r_0\rangle \langle r_0|, \\ \langle p_0 | p'_0 \rangle &= \int_{-\infty}^{\infty} d^3 r v_{p'_0}^*(\mathbf{r}) v_{p_0}(\mathbf{r}) = \int_{-\infty}^{\infty} d^3 r \frac{1}{2\pi\hbar} e^{-i(\mathbf{r}_0 \cdot \mathbf{p}_0 / \hbar)} = \delta(\mathbf{p}'_0 - \mathbf{p}_0) \Rightarrow 1 = \int d^3 p_0 |p_0\rangle \langle p_0|. \end{aligned}$$

Podemos entonces expresar el producto interno para la representación de coordenadas como:

$$\langle \Phi | \Psi \rangle = \langle \Phi | \underbrace{\left(\int d^3 r_0 |r_0\rangle \langle r_0| \right)}_1 | \Psi \rangle = \int d^3 r_0 \phi^*(\mathbf{r}_0) \psi(\mathbf{r}_0),$$

y equivalentemente para la representación de momentos:

$$\langle \Phi | \Psi \rangle = \langle \Phi | \underbrace{\left(\int d^3 p_0 |p_0\rangle \langle p_0| \right)}_1 | \Psi \rangle = \int d^3 p_0 \phi^*(\mathbf{p}_0) \psi(\mathbf{p}_0),$$

por lo cual hemos encontrado que:

$$\begin{aligned} |\Psi\rangle &= \int d^3 r_0 |r_0\rangle \langle r_0| \Psi \rangle = \int d^3 p_0 |p_0\rangle \langle p_0| \Psi \rangle, \\ \psi(\mathbf{r}_0) &= \langle r_0 | \Psi \rangle \quad \text{y} \quad \psi(\mathbf{p}_0) = \langle p_0 | \Psi \rangle, \end{aligned}$$

que es la representación de $|\Psi\rangle$ en coordenadas, $\psi(r_0)$, y en momentos, $\psi(p_0)$.

Adicionalmente cuando $|\Psi\rangle = |p\rangle$ tendremos que:

$$\langle r_0 | p_0 \rangle = \langle r_0 | \underbrace{\left(\int d^3 r'_0 |r'_0\rangle \langle r'_0| \right)}_1 | p_0 \rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3 r'_0 \delta(\mathbf{r}'_0 - \mathbf{r}_0) e^{\frac{i}{\hbar}(\mathbf{p}_0 \cdot \mathbf{r}_0)} = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i}{\hbar}(\mathbf{p}_0 \cdot \mathbf{r}_0)},$$

con lo cual $\psi(p_0)$ puede considerarse la transformada de Fourier de $\psi(r_0)$, y denotaremos de ahora en adelante las bases $|r_0\rangle \equiv |r\rangle$ y $|p_0\rangle \equiv |p\rangle$.

Estos índices continuos, \mathbf{r}_0 y \mathbf{p}_0 , representan tres índices continuos $\mathbf{r} = (x, y, z)$ y $\mathbf{p} = (p_x, p_y, p_z)$. La proyección de un vector abstracto $|\Psi\rangle$ en la representación $|r\rangle$ será considerada como su expresión en el espacio de coordenadas, igualmente su proyección $\langle p | \Psi \rangle$ será su expresión en el espacio de los momentos. Eso nos permitirá hacer corresponder los elementos de espacios vectoriales abstractos con elementos de un espacio

vectorial de funciones. Por lo tanto todas las fórmulas de proyección quedan como:

$$\langle r|\Psi\rangle = \psi(\mathbf{r}) \quad \text{y} \quad \langle p|\Psi\rangle = \psi(\mathbf{p}).$$

mientras que las relaciones de cierre y ortonormalización son:

$$\begin{aligned}\langle r|r'\rangle &= \delta(\mathbf{r}' - \mathbf{r}) \quad \text{y} \quad 1 = \int d^3r |r\rangle\langle r|, \\ \langle p|p\rangle &= \delta(\mathbf{p}' - \mathbf{p}) \quad \text{y} \quad 1 = \int d^3p |p\rangle\langle p|.\end{aligned}$$

Por su parte, la relación de cierre hará corresponder a la expresión del producto interno de dos vectores, tanto en la representación de las coordenadas como en la representación de momentos, en una de la forma:

$$\langle \Phi | \left(\int d^3r |r\rangle\langle r| \right) |\Psi\rangle = \int d^3r \phi^*(\mathbf{r}) \psi(\mathbf{r}) \quad \wedge \quad \langle \Phi | \left(\int d^3p |p\rangle\langle p| \right) |\Psi\rangle = \int d^3p \bar{\phi}^*(\mathbf{p}) \bar{\psi}(\mathbf{p}),$$

donde $\bar{\phi}^*(\mathbf{p})$ y $\bar{\psi}(\mathbf{p})$ son las transformadas de Fourier de $\phi^*(\mathbf{r})$ y $\psi(\mathbf{r})$, respectivamente. La afirmación anterior queda evidentemente demostrada del cambio entre las bases $|r\rangle$ y $|p\rangle$. Esto es:

$$\langle r|p\rangle = \langle p|r\rangle^* = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r})},$$

por lo cual:

$$\psi(\mathbf{r}) = \langle r|\Psi\rangle = \langle r| \left(\int d^3p |p\rangle\langle p| \right) |\Psi\rangle = \int d^3p \langle r|p\rangle\langle p|\Psi\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3p e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r})} \bar{\psi}(\mathbf{p}),$$

e inversamente:

$$\psi(\mathbf{p}) = \langle p|\Psi\rangle = \langle p| \left(\int d^3r |r\rangle\langle r| \right) |\Psi\rangle = \int d^3r \langle p|r\rangle\langle r|\Psi\rangle = \frac{1}{(2\pi\hbar)^{3/2}} \int d^3r e^{-\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{r})} \psi(\mathbf{r}).$$

3.5.8 Ejercicios

1. Si:

$$\delta_n(x) = \begin{cases} 0 & x < -\frac{1}{2n} \\ n & -\frac{1}{2n} < x < \frac{1}{2n} \\ 0 & x > \frac{1}{2n} \end{cases}$$

demuestre que:

$$f(0) = \lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f(x) \delta_n(x) dx,$$

para $f(x)$ continua en $x = 0$.

2. Para:

$$\delta_n(x) = \frac{n}{\pi} \frac{1}{1 + n^2 x^2}.$$

Demuestre lo siguiente:

$$\int_{-\infty}^{\infty} \delta_n(x) dx = 1.$$

3. Si:

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}.$$

Demuestre que:

$$x \frac{d\delta_n(x)}{dx} = -\delta_n(x).$$

4. Demuestre la siguiente relación:

$$\int_{-\infty}^{\infty} \delta'(x) f(x) dx = -f(0).$$

5. Dadas las siguientes funciones ortonormales:

$$\begin{aligned} \text{a) } |\hat{e}(k, x)\rangle &= \left\{ \frac{1}{\sqrt{\pi}} \sin(kx), \frac{1}{\sqrt{\pi}} \cos(kx) \right\}, \text{ con: } 0 \leq k < \infty, -\infty \leq x < \infty. \\ \text{b) } |\hat{e}(k, x)\rangle &= \left\{ \frac{1}{\sqrt{2\pi}} e^{ikx} \right\}, \text{ con: } -\infty \leq k < \infty, -\infty \leq x < \infty. \end{aligned}$$

(a). Escriba las condiciones de ortogonalidad y cierre.

(b). Demuestre:

$$\text{I. } \int_{-\infty}^{\infty} \sin(kx) \sin(k'x) dx = \pi \delta(k - k').$$

$$\text{II. } \int_{-\infty}^{\infty} \cos(kx) \cos(k'x) dx = \pi \delta(k - k').$$

$$\text{III. } \int_{-\infty}^{\infty} \sin(kx) \cos(k'x) dx = 0.$$

6. Demuestre que si:

$$\delta_n(x) = \frac{n}{\pi(1 + n^2 x^2)},$$

entonces:

$$\int_{-\infty}^{\infty} \delta_n dx = 1.$$

7. Dada la siguiente sucesión de funciones:

$$\delta_n(x) = \frac{1}{2n\pi} \left(\frac{\sin(nx/2)}{\sin(x/2)} \right)^2.$$

Demuestre que $\delta_n(x)$ es una distribución delta de Dirac al calcular

$$\lim_{n \rightarrow \infty} \left[\frac{1}{2n\pi} \int_{-\infty}^{\infty} f(x) \left(\frac{\sin(nx/2)}{\sin(x/2)} \right)^2 dx \right] = f(0).$$

8. Encuentre las transformadas de Fourier de las siguientes funciones:

$$a) \quad f(x) = \begin{cases} e^{-x}, & x > 0 \\ 0, & x < 0 \end{cases} \quad b) \quad f(x) = \begin{cases} h(1 - a|x|), & |x| < \frac{1}{a} \\ 0, & |x| > \frac{1}{a} \end{cases}$$

9. Dad $F(\mathbf{k})$ como la transformada de Fourier, en tres dimensiones, de $f(\mathbf{r})$ y $F_d(\mathbf{k})$ la transformada de Fourier, tridimensional, de $\nabla f(\mathbf{r})$. Demuestre que:

$$F_d(\mathbf{k}) = -i\mathbf{k}(\mathbf{k}).$$

10. Resuelva los ejercicios anteriores con **Maxima**.

Practicando con Maxima

Estudiaremos en este módulo como calcular transformadas integrales de Fourier con **Maxima**. Disponemos de las siguientes funciones: **fourint(f,x)** que calcula y devuelve la lista de los coeficientes integrales de Fourier de $f(x)$ definida en el intervalo $(-\infty, \infty)$. La función **fourintcos(f,x)** que devuelve los coeficientes integrales de los cosenos $f(x)$ en $(0, \infty)$ y la función **fourintsin(f,x)** que retorna los coeficientes integrales de los senos $f(x)$ en $(0, \infty)$

Vamos a aclarar a que se refieren cuando se habla de los coeficientes integrales.

Para **Maxima** si $f(x)$ es una función definida en $(-\infty, \infty)$ la transformada de Fourier exponencial es:

$$F(f, z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{izx} dx \Leftrightarrow f(x) = \int_{-\infty}^{\infty} F(f, z) e^{-izx} dz.$$

Como vimos anteriormente podemos tener los siguientes casos:

- La función es par: $f(-x) = f(x)$, entonces: $F(f, -z) = F(f, z)$ y podemos utilizar la representación en transformadas coseno de Fourier:

$$F(f, z) = \frac{1}{2} F_{\cos}(f, z).$$

- La función es impar: $f(-x) = -f(x)$, entonces: $F(f, -z) = -F(f, z)$ y podemos utilizar la representación en transformadas seno de Fourier:

$$F(f, z) = \frac{i}{2} F_{\sin}(f, z).$$

- La función no es par ni impar. Pero $f(x)$ siempre se puede escribir como la suma de una función par $f_+(x)$ y una impar $f_-(x)$:

$$f(x) = f_+(x) + f_-(x) = \frac{1}{2} [f(x) + f(-x)] + \frac{1}{2} [f(x) - f(-x)],$$

de manera equivalente:

$$F(f, z) = F(f_+ + f_-, z) = \frac{1}{2} F_{\cos}(f_+, z) + \frac{i}{2} F_{\sin}(f_-, z).$$

Entonces, la librería **fourie** lo que genera es la transformada de Fourier de la forma:

$$F(f, z) = \frac{1}{2} a_z + \frac{i}{2} b_z.$$

Veamos algunos ejemplos:

1. Consideremos la siguiente función par:

$$f = x^2 e^{-|x|}, \quad \text{con } x \in (-\infty, \infty).$$

Podemos intentar resolver la integral que define la transformada de Fourier de manera “manual” o directa:

$$\int_{-\infty}^{\infty} x^2 e^{-|x|} e^{-ikx} dx.$$

(%ci1) $f:x^2*exp(-abs(x));$
(%o1) $x^2 e^{-|x|}$

(%ci2) $\text{integrate}(exp(-%i*k*x)*f,x,minf,inf);$
(%o2) $\int_{-\infty}^{\infty} x^2 e^{-|x|-i k x} dx$

Lo que significa que el programa no supo resolverla. Intentemos ahora utilizar las funciones propias del programa.

(%ci3) $\text{load(fourie)}$$
(%ci4) $\text{fourint}(f,x);$
(%ot4) $a_z = \frac{2 \left(\frac{2}{z^6+3 z^4+3 z^2+1} - \frac{6 z^2}{z^6+3 z^4+3 z^2+1} \right)}{\pi}$
(%ot5) $b_z = 0$
(%o5) $[\%t4, \%t5]$
(%ci6) $\text{ratsimp}(\%t4),\text{factor};$
(%o6) $a_z = -\frac{4 (3 z^2 - 1)}{\pi (z^2 + 1)^3}$

Esta es la notación que el programa utiliza para decirnos que la transformada es

$$F(k) = \frac{1}{2} \left[-\frac{4 (3 k^2 - 1)}{\pi (k^2 + 1)^3} \right] = -\frac{2 (3 k^2 - 1)}{\pi (k^2 + 1)^3}.$$

Notemos que el cálculo anterior se corresponde a la transformada coseno de Fourier:

```
(%ci7) fourintcos(f,x);
(%ot7)  $a_z = \frac{2 \left( \frac{2}{z^6+3z^4+3z^2+1} - \frac{6z^2}{z^6+3z^4+3z^2+1} \right)}{\pi}$ 
(%oo7) [%t7]

(%ci8) ratsimp(%ot7),factor;
(%oo8)  $a_z = -\frac{4(3z^2-1)}{\pi(z^2+1)^3}$ 
```

Podemos intentar resolver directamente la integral para la transformada coseno de Fourier. Para tal fin escribimos toda la expresión a resolver.

```
(%ci9) Fz:ratsimp(1/%pi*integrate(cos(x*z)*f,x,0,inf)),factor;
(%oo9)  $-\frac{2(3z^2-1)}{\pi(z^2+1)^3}$ 
```

Incluso podemos probar calculando la transformada inversa y verificar que los cálculos son correctos.

```
(%ci10) integrate(exp(-%i*z*x)*Fz,z,minf,inf);
Is x positive, negative or zero? p;
(%oo10)  $x^2 e^{-x}$ 
```

```
(%ci11) integrate(exp(-%i*z*x)*Fz,z,minf,inf);
Is x positive, negative or zero? n;
(%oo11)  $x^2 e^x$ 
```

2. Consideremos la siguiente función impar:

$$f = xe^{-|x|}.$$

```
(%ci12) f:x*exp(-abs(x));
(%oo12)  $x e^{-|x|}$ 
```

```
(%ci13) fourint(f,x);
(%ot13)  $a_z = 0$ 
(%ot14)  $b_z = \frac{4z}{\pi(z^4+2z^2+1)}$ 
(%oo14) [%t13, %t14]
```

```
(%ci15) ratsimp(%ot14),factor;
(%oo15)  $b_z = \frac{4z}{\pi(z^2+1)^2}$ 
```

```
(%ci16) fourintsin(f,x);
(%ot16)  $b_z = \frac{4z}{\pi(z^4+2z^2+1)}$ 
(%oo16) [%t16]
```

```
(%ci17) ratsimp(%ot16),factor;
(%oo17)  $b_z = \frac{4z}{\pi(z^2+1)^2}$ 
```

Esto significa que la transformada seno de Fourier es:

$$F(k) = \frac{i}{2} \left[\frac{4k}{\pi (k^2 + 1)^2} \right] = \frac{2ik}{\pi (k^2 + 1)^2}.$$

Realicemos nuevamente el cálculo pero integrando para comparar.

```
(%ci18) Fz: %i*ratsimp(1/%pi*integrate(sin(x*z)*f,x,0,inf)),factor;
(%o18)  $\frac{2iz}{\pi(z^2+1)^2}$ 
```

Vamos a calcular la transformada inversa:

```
(%ci19) integrate(exp(-%i*z*x)*Fz,z,minf,inf);
Is x positive, negative or zero? p;
(%o19)  $x e^{-x}$ 

(%ci20) integrate(exp(-%i*z*x)*Fz,z,minf,inf);
Is x positive, negative or zero? n;
(%o20)  $x e^x$ 
```

3. Dada la siguiente función que no es par ni impar:

$$f(x) = (x - 1)e^{(-|x|)}.$$

```
(%ci21) f:(x-1)*exp(-abs(x));
(%o21)  $(x - 1) e^{-|x|}$ 

(%ci22) fourint(f,x);
(%o22)  $a_z = -\frac{2}{\pi(z^2+1)}$ 
(%o23)  $b_z = \frac{4z}{\pi(z^4+2z^2+1)}$ 
(%o23) [%t22, %t23]

(%ci24) az:ratsimp(%t22),factor; bz:ratsimp(%t23),factor;
(%o24)  $a_z = -\frac{2}{\pi(z^2+1)}$ 
(%o25)  $b_z = \frac{4z}{\pi(z^2+1)^2}$ 
(%ci26) Fz:1/2*rhs(az)+%i/2*rhs(bz),factor;
(%o26)  $-\frac{z^2-2iz+1}{\pi(z^2+1)^2}$ 
```

De nuevo, la transformada inversa para verificar:

```
(%ci27) integrate(exp(-%i*z*x)*Fz,z,minf,inf);
Is x positive, negative or zero? p;
(%o27)  $(x - 1) e^{-x}$ 

(%ci28) integrate(exp(-%i*z*x)*Fz,z,minf,inf);
Is x positive, negative or zero? n;
(%o28)  $(x - 1) e^x$ 

(%ci29) kill(all)$
```

4. Resolvamos ahora el ejemplo 3.21 donde la función era:

$$f(t) = \sin(\omega_0 t).$$

```
(%ci1) load(fourie)$
(%ci2) f:sin(w0*t);
(%o2) sin (w0 t)
```

Como la función es impar, podemos intentar:

```
(%ci3) fourintsin(f,t);
(%t3) b_z =  $\frac{2 \left( \lim_{t \rightarrow \infty} -\frac{\sin(tz+w_0 t)}{2z+2w_0} - \frac{\sin(tz-w_0 t)}{2w_0-2z} \right)}{\pi}$ 
(%o3) [%t3]
```

¡Intento fallido! Podemos probar resolviendo la integral.

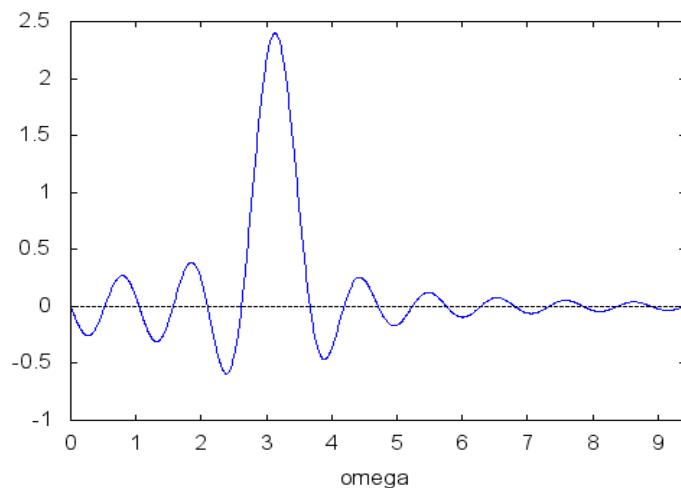
```
(%ci4) tau:n* %pi/w0;
(%ci5) Fw: sqrt(2/%pi)*ratsimp(integrate(sin(t*omega)*f,t,0,tau));
Is nw0 positive, negative or zero? p;
(%o5)  $-\frac{\sqrt{2} \left( (w_0 - \omega) \sin \left( \frac{\pi n w_0 + \pi n \omega}{w_0} \right) + (-w_0 - \omega) \sin \left( \frac{\pi n w_0 - \pi n \omega}{w_0} \right) \right)}{\sqrt{\pi} (2 w_0^2 - 2 \omega^2)}$ 
```

Asignemos los valores numéricos a los parámetros:

```
(%ci6) n:6w0 : %pi
(%ci7) Fw:ev(Fw);
(%o7)  $-\frac{\sqrt{2} \left( (\pi - \omega) \sin \left( \frac{6\pi w_0 + 6\pi^2}{\pi} \right) + (-\omega - \pi) \sin \left( \frac{6\pi^2 - 6\pi \omega}{\pi} \right) \right)}{\sqrt{\pi} (2\pi^2 - 2\omega^2)}$ 
```

Grafiquemos ahora la transformada $F(w)$:

```
(%ci8) wxplot2d([Fw], [omega,0,3* %pi])$
```



Bibliografía

- [1] Arfken, G. B., Weber, H., Weber, H.J. (2000) **Mathematical Methods for Physicists** 5ta Edición (Academic Press, Nueva York)
- [2] Borisenko, A.I, y Tarapov I.E. (1968) **Vector and Tensor Analisys** (Dover Publications Inc, Nueva York)
- [3] Dennery, P. y Krzywicki, A. (1995) **Mathematics for Physicists** (Dover Publications Inc, Nueva York)
- [4] Harper, C. (1971) **Introduction to Mathematical Physics** (*Prentice Hall, Englewood Cliff, N.J:*)
- [5] Hassani, S. (1991) **Foundations of Mathematical Physics** (*Prentice Hall, International Edition, London:*)
- [6] Hauser, W (1971) **Introduction to Principles of Electromagnetism** (*Addison-Wesley Pub Co Reading*)
- [7] Riley, K.F., Hobson, M.P. y Bence, S.J. (2002) **Mathematical Methods for Physics and Engineering** (*Cambridge University Press*)
- [8] Santaló, L.A (1969) **Vectores y Tensores** (*Editorial Universitaria, Buenos Aires*)
- [9] Schutz, B. (1980) **Geometrical Methods in Mathematical Physics** (*Cambridge University Press, Londres*)
- [10] Spiegel, M. (1959) **Vector Analysis** (*Schaums Outline Series, McGraw Hill New York*)

Capítulo 4

Matrices, determinantes y autovectores

La ruta de este capítulo

En el capítulo anterior definimos los funcionales lineales como un *morfismo* de un espacio vectorial lineal \mathbf{V} a un espacio unidimensional \mathbf{K} . Esta misma idea se puede extender a morfismos de un espacio vectorial \mathbf{V}_1 a un espacio vectorial \mathbf{V}_2 , sobre el mismo campo \mathbf{K} . Desarrollaremos estos conceptos a través de los operadores lineales en la sección 4.1, y en la sección 4.2 señalaremos algunos de los operadores lineales de mayor relevancia. En la sección 4.3 estableceremos una correspondencia uno-a-uno entre operadores lineales y matrices y la dependencia de ésta correspondencia con las bases del espacio vectorial. Luego, en la sección 4.4 presentaremos un conjunto de matrices fundamentales junto con sus operaciones. En la sección 4.5, veremos algunos de los métodos más comunes para la resolución de sistemas de ecuaciones lineales. Y para finalizar, las secciones 4.6 y 4.7 estarán dedicadas al importante tema de la representación espectral de un operador: el problema de autovalores y autovectores.

4.1 Operadores lineales

El primer paso es definir lo que en matemáticas se denomina una aplicación lineal, operador lineal o transformación lineal. Esto no es más que una aplicación que preservará la suma de vectores y la multiplicación por escalares.

Definición 4.1 (Operador Lineal)

Se define como un operador lineal a una aplicación que asocia un vector $|v\rangle \in \mathbf{V}_1$ un vector $|v'\rangle \in \mathbf{V}_2$ y que respeta la linealidad, es decir, una función $\mathbb{T} : \mathbf{V}_1 \rightarrow \mathbf{V}_2$ que cumple con:

$$|v'\rangle = \mathbb{T}|v\rangle \quad / \quad \mathbb{T}[\alpha |v_1\rangle + \beta |v_2\rangle] = \alpha \mathbb{T}|v_1\rangle + \beta \mathbb{T}|v_2\rangle \quad \forall \quad |v_1\rangle \text{ y } |v_2\rangle \in \mathbf{V}_1. \quad (4.1)$$

Sencillamente, algo que actúe sobre una suma de vectores en \mathbf{V}_1 y que es equivalente a la suma de sus actuaciones sobre los vectores suma, es decir, a la suma en \mathbf{V}_2 .

A continuación desarrollaremos diferentes ejemplos de transformaciones lineales

Ejemplo 4.1

1. Las siguientes transformaciones claramente son lineales

$$|x'\rangle = \mathbb{T}|x\rangle \rightarrow (x', y', z') = \mathbb{T}\{(x, y, z)\},$$

(a). $\mathbb{T}[(x, y, z)] = (x, 2y, 3z)$

$$\mathbb{T}[\alpha(x_1, y_1, z_1) + \beta(x_2, y_2, z_2)] = \alpha\mathbb{T}[(x_1, y_1, z_1)] + \beta\mathbb{T}[(x_2, y_2, z_2)] ,$$

entonces:

$$\mathbb{T}[(\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2, \alpha z_1 + \beta z_2)] = \alpha(x_1, 2y_1, 3z_1) + \beta(x_2, 2y_2, 3z_2)$$

$$(\alpha x_1 + \beta x_2, 2[\alpha y_1 + \beta y_2], 3[\alpha z_1 + \beta z_2]) = (\alpha x_1 + \beta x_2, 2[\alpha y_1 + \beta y_2], 3[\alpha z_1 + \beta z_2]) .$$

(b). $\mathbb{T}\{(x, y, z)\} = (z, y, x)$

$$\mathbb{T}[\alpha(x_1, y_1, z_1) + \beta(x_2, y_2, z_2)] = \alpha\mathbb{T}[(x_1, y_1, z_1)] + \beta\mathbb{T}[(x_2, y_2, z_2)] ,$$

igual que en el caso anterior:

$$\mathbb{T}\{(\alpha x_1 + \beta x_2, \alpha y_1 + \beta y_2, \alpha z_1 + \beta z_2)\} = \alpha(z_1, y_1, x_1) + \beta(z_2, y_2, x_2)$$

$$(\alpha z_1 + \beta z_2, \alpha y_1 + \beta y_2, \alpha x_1 + \beta x_2) = (\alpha z_1 + \beta z_2, \alpha y_1 + \beta y_2, \alpha x_1 + \beta x_2) .$$

2. Operaciones tan sencillas como la multiplicación por un número es una transformación (u operador) lineal, esto es, una transformación $\mathbb{T} : \mathbf{V} \rightarrow \mathbf{V}$ tal que

$$\mathbb{T}|v\rangle = |v'\rangle = \lambda|v\rangle \Rightarrow \mathbb{T}[\alpha|v\rangle + \beta|w\rangle] = \alpha\mathbb{T}|v\rangle + \beta\mathbb{T}|w\rangle = \alpha\lambda|v\rangle + \beta\lambda|w\rangle .$$

Obviamente, si $\lambda = 1$ tenemos la transformación identidad que transforma todo vector en sí mismo; si $\lambda = 0$ tendremos la transformación cero, vale decir que lleva a todo $|v\rangle \in \mathbf{V}$ a al elemento cero $|0\rangle$.

3. La definición de producto interno también puede ser vista como una transformación (operador) lineal $\mathbb{T} : \mathbf{V} \rightarrow \mathbb{R}$

$$\mathbb{T}|v\rangle = \lambda \Rightarrow \langle u|v\rangle \equiv \lambda .$$

Otra vez:

$$\mathbb{T}[\alpha|v\rangle + \beta|w\rangle] = \langle u|[\alpha|v\rangle + \beta|w\rangle] = \alpha\langle u|v\rangle + \beta\langle u|w\rangle ,$$

por lo tanto es lineal. Esto implica que también la proyección de un determinado $|v\rangle \in \mathbf{V}$ sobre un subespacio S es un operador lineal, y lo denotaremos como:

$$[|s\rangle\langle s|]|v\rangle = \langle s|v\rangle|s\rangle = |v_s\rangle , \quad \text{con } |s\rangle \text{ y } |v_s\rangle \in S .$$

Esta idea se extiende fácilmente para un proyector $\mathbb{T} : \mathbf{V}_m \rightarrow \mathbf{S}_n$ con $m > n$, de tal modo que para un vector $|v\rangle \in \mathbf{V}_m$

$$\mathbb{P}_m|v\rangle \equiv (|u_i\rangle\langle u^i|_m)|v\rangle = \langle u^i|v\rangle_m|u_i\rangle = |v_m\rangle ,$$

con $\{\langle u^i|\}$ una base de \mathbf{S}_n . Es claro que estamos utilizando la convención de Einstein para la suma de índices.

4. Las ecuaciones lineales también pueden verse como transformaciones lineales. Esto es, considere una transformación lineal $\mathbb{T} : \mathbf{V}_n \rightarrow \mathbf{V}_m$. Por lo tanto asociaremos

$$|y\rangle = \mathbb{T}|x\rangle \Rightarrow (y^1, y^2, y^3, \dots, y^m) = \mathbb{T}[(x^1, x^2, x^3, \dots, x^n)] ,$$

a través de $n \times m$ números, a_j^i , organizados de la siguiente forma:

$$y^i = a_j^i x^j \quad \text{con } \begin{cases} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n \end{cases}$$

una vez más,

$$\begin{aligned}\mathbb{T}[\alpha|v\rangle + \beta|w\rangle] &= \mathbb{T}[\alpha(v^1, v^2, v^3, \dots, v^n) + \beta(w^1, w^2, w^3, \dots, w^n)] = \alpha a_j^i v^j + \beta a_j^i w^j \\ &= \mathbb{T}[(\alpha v^1 + \beta w^1, \alpha v^2 + \beta w^2, \alpha v^3 + \beta w^3, \dots, \alpha v^n + \beta w^n)] \\ &= a_j^i (\alpha v + \beta w)^j = \alpha a_j^i v^j + \beta a_j^i w^j = a_j^i (\alpha v^j + \beta w^j).\end{aligned}$$

5. La derivada es un operador lineal

$$|v'\rangle = \mathbb{T}|v\rangle \rightarrow |y'\rangle = \mathbb{D}|y\rangle \rightarrow \mathbb{D}[y(x)] \equiv \frac{d}{dx}[y(x)] \equiv \frac{dy(x)}{dx} \equiv y'(x),$$

es claro que

$$\mathbb{D}[\alpha f(x) + \beta g(x)] = \alpha \mathbb{D}[f(x)] + \beta \mathbb{D}[g(x)] \equiv \alpha f'(x) + \beta g'(x).$$

Igualmente podemos asociar un operador diferencial de cualquier orden a una derivada del mismo orden, esto es

$$|y^{(n)}\rangle = \mathbb{D}^n|y\rangle \rightarrow \mathbb{D}^n[y(x)] \equiv \frac{d^n y(x)}{dx^n} \equiv y^{(n)}(x).$$

6. Del mismo modo, cualquier ecuación diferencial lineal es un ejemplo de operador lineal, digamos

$$y'' - 3y' + 2y = (\mathbb{D}^2 - 3\mathbb{D} + 2)y(x),$$

es claro que si $y(x) = \alpha f(x) + g(x)$ la linealidad es evidente:

$$\begin{aligned}(\alpha f(x) + g(x))'' - 3(\alpha f(x) + g(x))' + 2(\alpha f(x) + g(x)) &= \alpha(f'' - 3f' + 2f) + g'' - 3g' + 2g \\ &\quad \uparrow \downarrow \\ (\mathbb{D}^2 - 3\mathbb{D} + 2)(\alpha f(x) + g(x)) &= (\mathbb{D}^2 - 3\mathbb{D} + 2)\alpha f(x) + (\mathbb{D}^2 - 3\mathbb{D} + 2)g(x);\end{aligned}$$

7. La integral también es un operador lineal:

$$\begin{aligned}\int_a^x f(t)dt &\Leftarrow \mathbb{T}[f(t)], \\ \int_a^x [\alpha f(t) + \beta g(t)]dt &= \alpha \int_a^x \alpha f(t)dt + \beta \int_a^x g(t)dt.\end{aligned}$$



Considerando este último ejemplo, y como estudiaremos más adelante, existen otros ejemplos típicos de operadores de transformaciones integrales, es decir, transformaciones del tipo:

$$F(s) = \int_a^b \mathcal{K}(s, t) f(t)dt \Leftarrow \mathbb{T}[f(t)],$$

donde $\mathcal{K}(s, t)$ es una función conocida de s y t , denominada el *núcleo* de la transformación. Si a y b son finitos la transformación se dirá finita, de lo contrario infinita.

Así, si

$$f(t) = \alpha f_1(t) + f_2(t),$$

con $f_1(t)$ y $f_2(t) \in \mathcal{C}_{[a,b]}^\infty$, es obvio que:

$$F(s) = \int_a^b \mathcal{K}(s, t) [\alpha f_1(t) + f_2(t)]dt = \alpha \int_a^b \mathcal{K}(s, t) f_1(t)dt + \beta \int_a^b \mathcal{K}(s, t) f_2(t)dt,$$

$$F(s) = \alpha F(s_1) + \beta F(s_2) \Leftarrow \mathbb{T}[\alpha f_1(t) + f_2(t)] = \alpha \mathbb{T}[f_1(t)] + \beta \mathbb{T}[f_2(t)].$$

Dependiendo de la selección del núcleo y los límites tendremos distintas transformaciones integrales, en Física las más comunes son:

Nombre	$F(s) = \mathbb{T}\{f(t)\}$	$f(t) = \mathbb{T}^{-1}\{F(s)\}$
Laplace	$F(s) = \int_0^\infty e^{-st} f(t) dt$	$f(t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{st} F(s) ds$
Fourier de senos cosenos	$F(s) = \int_0^\infty \frac{\sin(st)}{\cos(st)} f(t) dt$	$f(t) = \int_0^\infty \frac{\sin(st)}{\cos(st)} F(s) ds$
Fourier compleja	$F(s) = \int_{-\infty}^\infty e^{ist} f(t) dt$	$f(t) = \int_{-\infty}^\infty e^{-ist} F(s) ds$
Hankel	$F(s) = \int_0^\infty t J_n(st) f(t) dt$	$f(t) = \int_0^\infty s J_n(ts) F(s) ds$
Mellin	$F(s) = \int_0^\infty t^{s-1} f(t) dt$	$f(t) = \frac{1}{2\pi i} \int_{\gamma-i\infty}^{\gamma+i\infty} s^{-t} F(s) ds$

4.1.1 Espacio vectorial de operadores lineales

Es posible definir un conjunto de operadores lineales $\{\mathbb{A}, \mathbb{B}, \mathbb{C}, \dots\} : \mathbf{V}_1 \rightarrow \mathbf{V}_2$ y constituir un espacio vectorial lineal si se dispone entre ellos de la operación suma y la multiplicación por un número. Así, claramente, dado $\{\mathbb{A}, \mathbb{B}, \mathbb{C}, \dots\}$, y definida

$$(\lambda \mathbb{A} + \mathbb{B}) |v\rangle \equiv \lambda \mathbb{A} |v\rangle + \mathbb{B} |v\rangle \quad / \quad \begin{cases} \mathbb{A} [\alpha |v_1\rangle + \beta |v_2\rangle] = \alpha \mathbb{A} |v_1\rangle + \beta \mathbb{A} |v_2\rangle \\ \mathbb{B} [\alpha |v_1\rangle + \beta |v_2\rangle] = \alpha \mathbb{B} |v_1\rangle + \beta \mathbb{B} |v_2\rangle \end{cases}$$

es directo comprobar que:

$$\begin{aligned} (\lambda \mathbb{A} + \mathbb{B}) [\alpha |v_1\rangle + \beta |v_2\rangle] &= \lambda \mathbb{A} [\alpha |v_1\rangle + \beta |v_2\rangle] + \mathbb{B} [\alpha |v_1\rangle + \beta |v_2\rangle] \\ &= \lambda (\alpha \mathbb{A} |v_1\rangle + \beta \mathbb{A} |v_2\rangle) + \alpha \mathbb{B} |v_1\rangle + \beta \mathbb{B} |v_2\rangle \\ &= \lambda (\alpha \mathbb{A} |v_1\rangle + \alpha \mathbb{B} |v_1\rangle) + \beta \mathbb{A} |v_2\rangle + \beta \mathbb{B} |v_2\rangle \\ &= \lambda \alpha (\mathbb{A} |v_1\rangle + \mathbb{B} |v_1\rangle) + \beta (\mathbb{A} |v_2\rangle + \mathbb{B} |v_2\rangle) . \end{aligned}$$

Igualmente, se cumple que $(\mathbb{A} + \mathbb{B}) + \mathbb{C} = \mathbb{A} + (\mathbb{B} + \mathbb{C})$, con \mathbb{A}, \mathbb{B} y \mathbb{C} lineales en \mathbf{V} ,

$$\begin{aligned} [(\mathbb{A} + \mathbb{B}) + \mathbb{C}] |v\rangle &= (\mathbb{A} + \mathbb{B}) |v\rangle + \mathbb{C} |v\rangle \quad \forall |v\rangle \in \mathbf{V}_1 \\ &= \mathbb{A} |v\rangle + \mathbb{B} |v\rangle + \mathbb{C} |v\rangle \\ &= \mathbb{A} |v\rangle + (\mathbb{B} + \mathbb{C}) |v\rangle \\ &= [\mathbb{A} + (\mathbb{B} + \mathbb{C})] |v\rangle , \end{aligned}$$

del mismo modo se puede comprobar fácilmente $\mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A}$.

Ahora bien, si definimos la transformación cero de $\mathbf{V}_1 \rightarrow \mathbf{V}_2$ tal que $\mathbb{O} |v\rangle = |0\rangle \quad \forall |v\rangle \in \mathbf{V}_1$, que asigna el vector $|0\rangle \in \mathbf{V}_2 \quad \forall |v\rangle \in \mathbf{V}_1$, entonces el operador lineal \mathbb{O} será el elemento neutro respecto a la suma de operadores.

Finalmente, el elemento simétrico queda definido por

$$(-\mathbb{A}) |v\rangle = -\mathbb{A} |v\rangle \Rightarrow (\mathbb{A} - \mathbb{A}) |v\rangle = \mathbb{O} |v\rangle = |0\rangle .$$

Con ello queda demostrado que los operadores lineales forman un espacio vectorial el cual de ahora en adelante denominaremos $\mathcal{L}(\mathbf{V}_1, \mathbf{V}_2)$.

4.1.2 Composición de operadores lineales

El producto o composición de dos operadores lineales, \mathbb{A} y \mathbb{B} se denotará \mathbb{AB} y significará que primero se aplica \mathbb{B} y al resultado se aplica \mathbb{A} . Esto es:

$$\mathbb{AB}|v\rangle = \mathbb{A}(\mathbb{B}|v\rangle) = \mathbb{A}|\tilde{v}\rangle = |\tilde{v}'\rangle . \quad (4.2)$$

La composición de operadores cumple con las siguientes propiedades:

$$\begin{aligned} (\mathbb{AB})\mathbb{C} &= \mathbb{A}(\mathbb{BC}); & \alpha(\mathbb{AB}) &= (\alpha\mathbb{A})\mathbb{B} = \mathbb{A}(\alpha\mathbb{B}); \\ (\mathbb{A}_1 + \mathbb{A}_2)\mathbb{B} &= \mathbb{A}_1\mathbb{B} + \mathbb{A}_2\mathbb{B}; & \mathbb{A}(\mathbb{B}_1 + \mathbb{B}_2) &= \mathbb{AB}_1 + \mathbb{AB}_2. \end{aligned}$$

Es decir, que la composición de operadores es asociativa y distributiva a la suma y que commuta respecto a la multiplicación por números.

Por otro lado si \mathbb{I} es el operador identidad: $\mathbb{I}|v\rangle = |v\rangle \Rightarrow \mathbb{AI} = \mathbb{IA} = \mathbb{A}$.

En general $\mathbb{AB} \neq \mathbb{BA}$ y podemos construir el comutador de estos operadores como:

$$[\mathbb{A}, \mathbb{B}] = \mathbb{AB} - \mathbb{BA}, \quad (4.3)$$

por lo tanto:

$$[\mathbb{AB} - \mathbb{BA}]|v\rangle = \mathbb{AB}|v\rangle - \mathbb{BA}|v\rangle .$$

Algunas de las propiedades más útiles de los comutadores son:

$$[\mathbb{A}, \mathbb{B}] = -[\mathbb{B}, \mathbb{A}] \quad (4.4)$$

$$[\mathbb{A}, (\mathbb{B} + \mathbb{C})] = [\mathbb{A}, \mathbb{B}] + [\mathbb{A}, \mathbb{C}] \quad (4.5)$$

$$[\mathbb{A}, \mathbb{BC}] = [\mathbb{A}, \mathbb{B}]\mathbb{C} + \mathbb{B}[\mathbb{A}, \mathbb{C}] \quad (4.6)$$

$$[\mathbb{A}, [\mathbb{B}, \mathbb{C}]] = -[\mathbb{B}, [\mathbb{C}, \mathbb{A}]] - [\mathbb{C}, [\mathbb{A}, \mathbb{B}]] . \quad (4.7)$$

Un par de resultados inmediatos se derivan de la composición de operadores:

- Potencias de operadores.** Uno de los ejemplos más útiles en la composición de operadores lo constituyen las potencias de los operadores, las cuales provienen de la aplicación consecutiva de un mismo operador,

$$\mathbb{A}^0 = \mathbb{I}; \quad \mathbb{A}^1 = \mathbb{A}; \quad \mathbb{A}^2 = \mathbb{AA}; \quad \mathbb{A}^3 = \mathbb{A}^2\mathbb{A} = \mathbb{AAA} \dots$$

es claro que las potencias de operadores cumplen las propiedades estándares de las potencias de números

$$\mathbb{A}^{n+m} = \mathbb{A}^n\mathbb{A}^m; \quad (\mathbb{A}^n)^m = \mathbb{A}^{nm}.$$

Llamaremos *operadores nilpotentes de grado n* a los operadores $\mathbb{A}^n \neq 0$, del tipo $\mathbb{A}^n|v\rangle = |0\rangle \forall |v\rangle \in \mathbf{V}_1$ y $|0\rangle \in \mathbf{V}_2$. Es decir, un operador que lleva cualquier vector $|v\rangle$ al elemento neutro de \mathbf{V}_2 . El ejemplo más notorio es el operador diferencial

$$\mathbb{D}^n |P^{n-1}\rangle = |0\rangle \iff \frac{d^n}{dx^n} P_{n-1}(x) = \frac{d^n}{dx^n} [a_i x^i] = 0 ,$$

con $|P^{n-1}\rangle$ perteneciente al espacio de polinomios de grado $n-1$.

- Operador ecuaciones diferenciales.** Si consideramos el espacio de funciones $f(x) \in \mathcal{C}_{[a,b]}^\infty$ podemos construir un operador diferencial

$$[a_0 + a_1\mathbb{D} + a_2\mathbb{D}^2 + \dots + a_n\mathbb{D}^n] |f\rangle \Rightarrow \left(a_0 + a_1 \frac{d}{dx} + a_2 \frac{d^2}{dx^2} + \dots + a_n \frac{d^n}{dx^n} \right) f(x) ,$$

con $\{a_0, a_1, a_2, \dots, a_n\}$ coeficientes constantes. De este modo, por ejemplo:

$$(\mathbb{D}^2 - 3\mathbb{D} + 2) y = (\mathbb{D} - 1)(\mathbb{D} - 2)y \Rightarrow \left(\frac{d^2}{dx^2} - 3\frac{d}{dx} + 2 \right) y(x) = y'' - 3y' + 2y.$$

Ejemplo 4.2 Dados dos operadores lineales tales que:

$$\mathbb{A}\mathbb{B} = -\mathbb{B}\mathbb{A}, \quad \mathbb{A}^2 = \mathbb{I}, \quad \mathbb{B}^2 = \mathbb{I} \text{ y } [\mathbb{A}, \mathbb{B}] = 2i\mathbb{C}.$$

- Mostraremos que: $\mathbb{C}^2 = \mathbb{I}$

$$[\mathbb{A}, \mathbb{B}] = \mathbb{A}\mathbb{B} - \mathbb{B}\mathbb{A} = 2i\mathbb{C} \Rightarrow 2\mathbb{A}\mathbb{B} = 2i\mathbb{C} \Rightarrow \mathbb{A}\mathbb{B}\mathbb{A}\mathbb{B} = -\mathbb{C}^2 \Rightarrow -\mathbb{A}\mathbb{A}\mathbb{B}\mathbb{B} = -\mathbb{C}^2 \Rightarrow \mathbb{I} = \mathbb{C}^2$$

- Y que: $[\mathbb{B}, \mathbb{C}] = 2i\mathbb{A}$

$$[\mathbb{B}, \mathbb{C}] = -i(\mathbb{B}\mathbb{A}\mathbb{B} - \mathbb{A}\mathbb{B}\mathbb{B}) = 2i\mathbb{A} \Rightarrow -i(\mathbb{B}\mathbb{A}\mathbb{B} - \mathbb{A}) = 2i\mathbb{A} \Rightarrow -i(-\mathbb{B}\mathbb{B}\mathbb{A} - \mathbb{A}) = 2i\mathbb{A}.$$

- Evaluemos: $[[\mathbb{A}, \mathbb{B}], [\mathbb{B}, \mathbb{C}]], [\mathbb{A}, \mathbb{B}]$.

$$\begin{aligned} [[\mathbb{A}, \mathbb{B}], [\mathbb{B}, \mathbb{C}]], [\mathbb{A}, \mathbb{B}] &= [[2i\mathbb{C}, 2i\mathbb{A}], 2i\mathbb{C}] = 8[[\mathbb{A}\mathbb{B}, i\mathbb{A}], \mathbb{A}\mathbb{B}] = 8i[(\mathbb{A}\mathbb{B}\mathbb{A} - \mathbb{A}\mathbb{A}\mathbb{B}), \mathbb{A}\mathbb{B}] \\ &= 8i[-2\mathbb{B}, \mathbb{A}\mathbb{B}] = 16i(\mathbb{B}\mathbb{A}\mathbb{B} - \mathbb{A}\mathbb{B}\mathbb{B}) = 32i\mathbb{A}. \end{aligned}$$

▷

Ejemplo 4.3 Dados tres operadores vectoriales:

$$\vec{\mathbb{A}} = \mathbb{A}^{(i)} |e_i\rangle \equiv \mathbb{A}^{(1)} |e_1\rangle + \mathbb{A}^{(2)} |e_2\rangle + \mathbb{A}^{(3)} |e_3\rangle, \quad \vec{\mathbb{B}} = \mathbb{B}^{(i)} |e_i\rangle \text{ y } \vec{\mathbb{L}} = \mathbb{L}^{(i)} |e_i\rangle,$$

que conmutan entre ellos, tales que:

$$[\mathbb{A}^{(i)}, \mathbb{B}^{(j)}] = [\mathbb{A}^{(k)}, \mathbb{L}^{(l)}] = [\mathbb{L}^{(m)}, \mathbb{B}^{(n)}] = 0, \quad \text{con } i, j, k, l, m, n = 1, 2, 3.$$

Donde los superíndices etiquetan distintos operadores, vale decir las componentes de un “operador vectorial” y no la potencia del operador. Adicionalmente, las componentes $\mathbb{L}^{(i)}$ cumplen con el álgebra $\varepsilon_{(k)(l)(m)}\mathbb{L}^{(k)} = [\mathbb{L}_{(l)}, \mathbb{L}_{(m)}]$

La notación vectorial de índices y el tensor de Levi-Civita, $\varepsilon_{(k)(l)(m)}$, fueron discutidos anteriormente en la sección 1.4.

Entonces demostramos la siguiente relación:

$$[\vec{\mathbb{A}} \cdot \vec{\mathbb{L}}, \vec{\mathbb{B}} \cdot \vec{\mathbb{L}}] = (\vec{\mathbb{A}} \times \vec{\mathbb{B}}) \cdot \vec{\mathbb{L}}.$$

Donde $\vec{\mathbb{A}} \times \vec{\mathbb{B}}$ define el “producto vectorial” de los “operadores vectoriales” $\vec{\mathbb{A}}$ y $\vec{\mathbb{B}}$.

Veamos:

$$\begin{aligned} [\vec{\mathbb{A}} \cdot \vec{\mathbb{L}}, \vec{\mathbb{B}} \cdot \vec{\mathbb{L}}] &= \mathbb{A}^{(i)} \mathbb{L}_{(i)} \mathbb{B}^{(j)} \mathbb{L}_{(j)} - \mathbb{B}^{(j)} \mathbb{L}_{(j)} \mathbb{A}^{(i)} \mathbb{L}_{(i)} \\ &\equiv \mathbb{A}^{(i)} \mathbb{B}^{(j)} \mathbb{L}_{(i)} \mathbb{L}_{(j)} - \mathbb{A}^{(i)} \mathbb{B}^{(j)} \mathbb{L}_{(j)} \mathbb{L}_{(i)} \equiv \mathbb{A}^{(i)} \mathbb{B}^{(j)} ([\mathbb{L}_{(i)}, \mathbb{L}_{(j)}]), \end{aligned}$$

donde las dos últimas equivalencias las hemos construido a partir de las propiedades de conmutación de los operadores $\mathbb{A}^{(i)}$, $\mathbb{B}^{(j)}$ y $\mathbb{L}^{(k)}$.

Seguidamente, invocamos el álgebra de los operadores $\mathbb{L}_{(l)}$ y $\mathbb{L}_{(m)}$ y se obtiene

$$\mathbb{A}^{(i)} \mathbb{B}^{(j)} ([\mathbb{L}_{(i)}, \mathbb{L}_{(j)}]) = \mathbb{A}^{(i)} \mathbb{B}^{(j)} (\varepsilon_{(i)(j)(k)} \mathbb{L}^{(k)}) = \varepsilon_{(i)(j)(k)} \mathbb{A}^{(i)} \mathbb{B}^{(j)} \mathbb{L}^{(k)} = (\vec{\mathbb{A}} \times \vec{\mathbb{B}}) \cdot \vec{\mathbb{L}}.$$

▷

4.1.3 Proyectores

La notación de Dirac se hace particularmente conveniente para representar proyectores. Hasta ahora, hemos relacionado un funcional lineal, un *bra* $\langle w |$ del espacio dual V^* , con un vector *ket* $|v\rangle$ del espacio vectorial directo V a través de su producto interno $\langle w | v \rangle$, el cual es, en general, un número complejo. Ahora escribiremos

esta relación entre vectores y formas diferenciales de una manera diferente: la relación entre $\langle w |$ y $|v\rangle$ un *ket* $|\Psi\rangle$ o un *bra* $\langle\Phi|$ arbitrarios puede ser:

$$|v\rangle \langle w| \Rightarrow \begin{cases} |v\rangle \langle w| \Psi \rangle \\ \langle \Phi | v \rangle \langle w| \end{cases}$$

La primera será la multiplicación del vector $|v\rangle$ por el número complejo $\langle w | \Psi \rangle$, mientras que la segunda relación será la multiplicación de la forma $\langle w |$ por el complejo $\langle \Phi | v \rangle$. Es imperioso señalar que el orden en la escritura de los vectores y formas es crítico, sólo los números complejos λ se pueden mover con impunidad a través de estas relaciones.

$$\begin{aligned} \lambda |v\rangle &= |\lambda v\rangle = |v\rangle \lambda, \quad \lambda \langle w| = \langle \lambda w| = \langle w| \lambda \\ \langle w| \lambda |v\rangle &= \lambda \langle w| v\rangle = \langle w| v\rangle \lambda \quad \text{y} \quad \mathbb{A} |v\rangle = \mathbb{A} \lambda |v\rangle = \lambda \mathbb{A} |v\rangle . \end{aligned}$$

Por lo tanto, dado un vector $|v\rangle$, podemos construir un proyector $\mathbb{P}_{|v\rangle}$ a lo largo del vector $|v\rangle$,

$$\mathbb{P}_{|v\rangle} \equiv |v\rangle \langle v| , \quad \text{con } \langle v| v\rangle = 1 ,$$

siempre y cuando este operador lineal cumpla:

$$\begin{aligned} \mathbb{P}_{|v\rangle} [\alpha |z_1\rangle + \beta |z_2\rangle] &= \alpha \mathbb{P}_{|v\rangle} |z_1\rangle + \beta \mathbb{P}_{|v\rangle} |z_2\rangle , \\ |v\rangle \langle v| [\alpha |z_1\rangle + \beta |z_2\rangle] &= |v\rangle \langle v| \alpha |z_1\rangle + |v\rangle \langle v| \beta |z_2\rangle = \alpha |v\rangle \langle v| z_1\rangle + \beta |v\rangle \langle v| z_2\rangle , \end{aligned}$$

además:

$$\begin{aligned} \mathbb{P}_{|v\rangle}^2 &= \mathbb{P}_{|v\rangle} \iff (\langle v| v\rangle) (\langle v| v|) = |v\rangle \langle v| \\ \mathbb{P}_{|v\rangle} \mathbb{P}_{|v\rangle} |z\rangle &= (\langle v| v\rangle) (\langle v| v|) |z\rangle = |v\rangle \underbrace{\langle v| v\rangle}_{1} \langle v| z\rangle = |v\rangle \langle v| z\rangle = \mathbb{P}_{|v\rangle} |z\rangle . \end{aligned}$$

Así, el operador $\mathbb{P}_{|v\rangle}$ actuando sobre el vector $|\Psi\rangle$ representará la proyección de $|\Psi\rangle$ a lo largo de $|v\rangle$

$$\mathbb{P}_{|v\rangle} |\Psi\rangle = |v\rangle \langle v| \Psi\rangle \equiv \langle v| \Psi\rangle |v\rangle .$$

Es inmediato construir un proyector de un vector sobre un subespacio S_q .

Sea: $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_q\rangle\}$ un conjunto ortonormal de vectores que expande S_q , por lo tanto, definiremos el proyector \mathbb{P}_q al proyector sobre el subespacio S_q de la forma:

$$\mathbb{P}_q = |e_i\rangle \langle e^i|_q ,$$

es claro que $\forall |v\rangle \in V$ se tiene:

$$\mathbb{P}_q^2 |v\rangle = \mathbb{P}_q \mathbb{P}_q |v\rangle \Rightarrow \mathbb{P}_q^2 |v\rangle = \left(|e_i\rangle \langle e^i|_q \right) \left(|e_j\rangle \langle e^j|_q \right) |v\rangle = |e_i\rangle \underbrace{\langle e^i|}_{\delta^i_j} |e_j\rangle \langle e^j|_q |v\rangle = |e_j\rangle \langle e^j|_q |v\rangle = \mathbb{P}_q |v\rangle ,$$

es decir, $\mathbb{P}_q^2 = \mathbb{P}_q$.

Ejemplo 4.4 En Mecánica Clásica la cantidad de momento angular viene definida como $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Para pasar a Mecánica Cuántica se asocia \mathbf{r} y \mathbf{p} con los operadores posición y cantidad de movimiento los cuales, al operar

sobre la función de onda nos proveen:

$$\langle r | \mathbb{X} | \psi \rangle = x \langle r | \psi \rangle = x \psi(\mathbf{r}) \quad \langle r | \mathbb{P}_x | \psi \rangle = \left(-i\hbar \frac{\partial}{\partial x} \right) \langle r | \psi \rangle = -i\hbar \frac{\partial}{\partial x} \psi(\mathbf{r})$$

$$\langle r | \mathbb{Y} | \psi \rangle = y \langle r | \psi \rangle = y \psi(\mathbf{r}) \quad \langle r | \mathbb{P}_y | \psi \rangle = \left(-i\hbar \frac{\partial}{\partial y} \right) \langle r | \psi \rangle = -i\hbar \frac{\partial}{\partial y} \psi(\mathbf{r})$$

$$\langle r | \mathbb{Z} | \psi \rangle = z \langle r | \psi \rangle = z \psi(\mathbf{r}) \quad \langle r | \mathbb{P}_z | \psi \rangle = \left(-i\hbar \frac{\partial}{\partial z} \right) \langle r | \psi \rangle = -i\hbar \frac{\partial}{\partial z} \psi(\mathbf{r}).$$

En coordenadas cartesianas, en la representación de coordenadas $\{|r\rangle\}$ tendremos que:

$$\langle r | \mathbb{R} | \psi \rangle = \mathbf{r} \psi(\mathbf{r}) \quad \text{y} \quad \langle r | \mathbb{P}_x | \psi \rangle = -i\hbar \nabla \psi(\mathbf{r}).$$

De forma que en Mecánica Cuántica las componentes cartesianas del operador cantidad de movimiento angular son

$$\langle r | \mathbb{L} | \psi \rangle = -i\hbar (\mathbf{r} \times \nabla) \psi(\mathbf{r}),$$

$$\langle r | \mathbb{L} | \psi \rangle = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \psi(\mathbf{r}) \hat{\mathbf{i}} - i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \psi(\mathbf{r}) \hat{\mathbf{j}} - i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \psi(\mathbf{r}) \hat{\mathbf{k}}.$$

Utilizando las definiciones anteriores mostraremos que el comutador de las componentes cartesianas de la cantidad de movimiento angular cumple con:

$$[\mathbb{L}_x, \mathbb{L}_y] | \psi \rangle = i\hbar \mathbb{L}_z | \psi \rangle,$$

donde: $\mathbb{L}^1 = \mathbb{L}_1 = \mathbb{L}_x$; $\mathbb{L}^2 = \mathbb{L}_2 = \mathbb{L}_y$; $\mathbb{L}^3 = \mathbb{L}_3 = \mathbb{L}_z$. En general: $[\mathbb{L}_l, \mathbb{L}_m] = i\hbar \varepsilon_{lmn} \mathbb{L}^n$.

Dado que:

$$[\mathbb{L}_1, \mathbb{L}_2] | \psi \rangle = [\mathbb{L}_x, \mathbb{L}_y] | \psi \rangle = (\mathbb{L}_x \mathbb{L}_y - \mathbb{L}_y \mathbb{L}_x) | \psi \rangle,$$

entonces:

$$\begin{aligned} [\mathbb{L}_x, \mathbb{L}_y] | \psi \rangle &= \left[\left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) - \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \right] \psi(\mathbf{r}) \\ &= \left[y \frac{\partial}{\partial z} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) - z \frac{\partial}{\partial y} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \right. \\ &\quad \left. - z \frac{\partial}{\partial x} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) + x \frac{\partial}{\partial z} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \right] \psi(\mathbf{r}). \end{aligned}$$

Con lo cual:

$$\begin{aligned} [\mathbb{L}_x, \mathbb{L}_y] | \psi \rangle &= \left[\left(\left(yz \frac{\partial}{\partial z \partial x} + y \frac{\partial}{\partial x} \right) - xy \frac{\partial^2}{\partial z^2} \right) - \left(z^2 \frac{\partial}{\partial y \partial x} - zx \frac{\partial}{\partial y \partial z} \right) \right. \\ &\quad \left. - \left(\left(zy \frac{\partial}{\partial x \partial z} - z^2 \frac{\partial}{\partial y \partial x} \right) - \left(yx \frac{\partial}{\partial z^2} - zx \frac{\partial}{\partial z \partial y} - x \frac{\partial}{\partial y} \right) \right) \right] \psi(\mathbf{r}) \\ &= \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) \psi(\mathbf{r}). \end{aligned}$$



4.1.4 Ejercicios

1. Diga si las siguientes transformaciones, $|x'\rangle = \mathbb{T}|x\rangle$ son lineales

- (a). $\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $\mathbb{T}[(x, y, z)] = (x + y, x + z)$.
- (b). $\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $\mathbb{T}[(x, y, z)] = (x, y, y + z)$.

- (c). $\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \mathbb{T}[(x, y, z)] = (x, x + y, x - y)$.
 (d). $\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^4, \mathbb{T}[(x, y, z)] = (x + y, x + z, 2x + y + z, y - z)$.
 (e). $\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \mathbb{T}[(x, y, z)] = (\operatorname{sen}(x), \cos(y), 0)$.
2. Cuál de las siguientes transformaciones definidas sobre \mathbb{V}^3 son lineales
- (a). $\mathbb{T}|x\rangle = |x\rangle + |a\rangle$ donde $|a\rangle$ es un vector constante diferente de cero.
 (b). $\mathbb{T}|x\rangle = |a\rangle$.
 (c). $\mathbb{T}|x\rangle = \langle a |x\rangle |a\rangle$.
 (d). $\mathbb{T}|x\rangle = \langle a |x\rangle |x\rangle$.
3. Considere las siguientes operaciones en el espacio de los polinomios en x y diga si corresponden a transformaciones lineales:
 a) La multiplicación por x . b) La multiplicación por x^2 . b) La diferenciación.
4. Suponga que $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A}$. Demuestre que:
- (a). $(\mathbb{A} + \mathbb{B})^2 = \mathbb{A}^2 + 2\mathbb{A}\mathbb{B} + \mathbb{B}^2$.
 (b). $(\mathbb{A} + \mathbb{B})^3 = \mathbb{A}^3 + 3\mathbb{A}^2\mathbb{B} + 3\mathbb{A}\mathbb{B}^2 + \mathbb{B}^3$.
- ¿Cómo cambian las fórmulas anteriores si $\mathbb{A}\mathbb{B} \neq \mathbb{B}\mathbb{A}$?
5. Suponga que un operador \mathbb{L} puede ser escrito como la composición de otros dos operadores $\mathbb{L} = \mathbb{L}_-\mathbb{L}_+$ con $[\mathbb{L}_-, \mathbb{L}_+] = \mathbb{I}$. Demostrar que:

$$\text{Si } \mathbb{L}|x\rangle = \lambda|x\rangle \quad \text{y} \quad |y\rangle = \mathbb{L}_+|x\rangle \quad \text{entonces} \quad \mathbb{L}|y\rangle = (\lambda + 1)|y\rangle$$

y, del mismo modo, demuestre que:

$$\text{Si } \mathbb{L}|x\rangle = \lambda|x\rangle \quad \text{y} \quad |z\rangle = \mathbb{L}_-|x\rangle \quad \text{entonces} \quad \mathbb{L}|z\rangle = (\lambda - 1)|z\rangle .$$

Por ello es costumbre denominar a \mathbb{L}_+ y \mathbb{L}_- los operadores de “subidas” y de “bajada” respectivamente, ya que ellos construyen otros vectores con autovalores mayores (menores) en una unidad al vector operado.

Practicando con Maxima

1. Si tenemos la siguiente transformación, $|x'\rangle = \mathbb{T}|x\rangle$:

$$\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^4, \quad \mathbb{T}[(x, y, z)] = (2x + y, x - z, x + y + z, y + z) ,$$

entonces, para que sea una transformación lineal debe satisfacer lo siguiente:

- $\mathbb{T}[|v_1\rangle + |v_2\rangle] = \mathbb{T}[|v_1\rangle] + \mathbb{T}[|v_2\rangle]$.
- $\mathbb{T}[\alpha|v\rangle] = \alpha\mathbb{T}[|v\rangle]$.

(%o1) $f(x,y,z):=[2*x+y,x-z,x+y+z,y+z];$
 (%o1) $f(x, y, z) := [2x + y, x - z, x + y + z, y + z]$

El lado derecho de la igualdad es:

(%o1) $ld:f(x1+x2,y1+y2,z1+z2),factor;$
 (%o2) $[y_2 + y_1 + 2x_2 + 2x_1, -(z_2 + z_1 - x_2 - x_1), z_2 + z_1 + y_2 + y_1 + x_2 + x_1, z_2 + z_1 + y_2 + y_1]$

El lado izquierdo,

(%o1) $li:f(x1,y1,z1)+f(x2,y2,z2),factor;$
 (%o3) $[y_2 + y_1 + 2x_2 + 2x_1, -(z_2 + z_1 - x_2 - x_1), z_2 + z_1 + y_2 + y_1 + x_2 + x_1, z_2 + z_1 + y_2 + y_1]$

Por lo tanto:

(%o1) $\text{ld-li,ratsimp};$
 (%o4) $[0, 0, 0, 0]$

Para la segunda condición

(%o1) $f(\alpha*x, \alpha*y, \alpha*z) - \alpha*f(x, y, z);$
 (%o5) $[-\alpha(y + 2x) + \alpha y + 2\alpha x, -\alpha z - \alpha(x - z) + \alpha x, -\alpha(z + y + x) + \alpha z + \alpha y + \alpha x, -\alpha(z + y) + \alpha z + \alpha y]$

(%o1) $\text{factor}(%);$
 (%o6) $[0, 0, 0, 0]$

2. Consideremos ahora la siguiente transformación, $|x'\rangle = T|x\rangle$:

$$T : \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad T[(x, y, z)] = (x^2, y + z, z^2).$$

(%o1) $f(x, y, z) := [x^2, y + z, z^2];$
 (%o7) $f(x, y, z) := [x^2, y + z, z^2];$

El lado derecho:

(%o1) $\text{ld:f}(x1+x2, y1+y2, z1+z2), \text{factor};$
 (%o8) $[(x_2 + x_1)^2, z_2 + z_1 + y_2 + y_1, (z_2 + z_1)^2]$

El lado izquierdo:

(%o1) $\text{li:f}(x1, y1, z1) + f(x2, y2, z2), \text{factor};$
 (%o9) $[x_2^2 + x_1^2, z_2 + z_1 + y_2 + y_1, z_2^2 + z_1^2]$

Por lo tanto:

(%o1) $\text{ld-li,ratsimp};$
 (%o10) $[2x_1 x_2, 0, 2z_1 z_2]$

No es una transformación lineal.

3. Dadas las siguientes transformaciones de $\mathbb{R}^3 \rightarrow \mathbb{R}^3$:

$$\mathbb{A}[(x, y, z)] = (2x - z, x + z, x), \quad \mathbb{B}[(x, y, z)] = (z, x, y).$$

Veamos si comutan, es decir, $[\mathbb{A}, \mathbb{B}] = \mathbb{A}\mathbb{B} - \mathbb{B}\mathbb{A} = 0$. Podemos ver que $\mathbb{A}\mathbb{B} = \mathbb{A}[\mathbb{B}|x\rangle] = (2z - y, z + y, z)$ y $\mathbb{B}\mathbb{A} = \mathbb{B}[\mathbb{A}|x\rangle] = (x, 2x - z, x + z)$.

(%o1) $A(x, y, z) := [2x - z, x + z, x];$
 (%o11) $A(x, y, z) := [2x - z, x + z, x]$

(%o1) $B(x, y, z) := [z, x, y];$
 (%o12) $B(x, y, z) := [z, x, y]$

(%o1) $[x, y, z] : [z, x, y];$
 (%o13) $[x, y, z] : [z, x, y]$

(%o1) $\mathbb{AB} := A(x, y, z);$
 (%o14) $\mathbb{AB} := [2z - y, z + y, z]$

(%i1) kill(x,y,z);
(%o15) done

(%i1) [x,y,z]:[2*x-z, x+z,x];
(%o16) [2x - z, z + x, x]

(%i1) BA:B(x,y,z);
(%o17) [x, 2x - z, z + x]

(%i1) AB-BA;
(%o18) [2z - y - x, 2z + y - 2x, -x]

Por lo tanto no comutan.

4.2 Tipos de operadores

Discutiremos en esta sección algunas de las propiedades que caracterizan a los operadores lineales. Además, tal y como están definidas las transformaciones lineales tiene sentido estudiar si ellas poseen la característica de ser: inyectivas, sobrejetivas y biyectivas. Las transformaciones lineales tendrán nombres particulares para cada uno de estos casos.

Como ya fue señalado, una transformación lineal es una función, aplicación, operador o mapeo cuyos dominios y rangos son subconjuntos de espacios vectoriales y que hemos simbolizado como:

$$\mathbb{A} : \mathbf{V}_1 \rightarrow \mathbf{V}_2 \Rightarrow \mathbb{A}|v\rangle = |v'\rangle, \text{ con } |v\rangle \in \mathbf{V}_1 \text{ y } |v'\rangle \in \mathbf{V}_2.$$

En el lenguaje de la teoría de conjuntos a los elementos $|v'\rangle \in \mathbf{V}_2$ se les denomina la imagen de $|v\rangle$ debido a \mathbb{A} . Si S es un subconjunto de \mathbf{V}_1 , entonces al conjunto de todas las imágenes, que denotaremos por $\mathbb{A}(S)$, se le denomina la imagen de S debido a la aplicación de \mathbb{A} en \mathbf{V}_1 . A la imagen del dominio de \mathbf{V}_1 , $\mathbb{A}\{\mathbf{V}_1\}$, se le llama el rango de \mathbb{A} y es un subespacio de \mathbf{V}_2 .

4.2.1 Espacio nulo e imagen de un operador

El conjunto de todos los $|v\rangle \in \mathbf{V}_1 / \mathbb{A}|v\rangle = |0\rangle$, se denomina espacio nulo, núcleo o *kernel* (núcleo en alemán) de la transformación \mathbb{A} y lo denotaremos como $\mathfrak{N}(\mathbb{A})$. En símbolos diremos que:

$$\mathfrak{N}(\mathbb{A}) = \{|v\rangle \in \mathbf{V}_1 \wedge \mathbb{A}|v\rangle = |0\rangle\}.$$

Adicionalmente, $\mathfrak{N}(\mathbb{A}) \subset \mathbf{V}_1$, es decir, será un subespacio de \mathbf{V}_1 . La prueba de esta afirmación es inmediata. Dados $|v_1\rangle, |v_2\rangle \in \mathfrak{N}(\mathbb{A})$, con \mathbb{A} un operador lineal, es claro que:

$$\left. \begin{array}{l} \mathbb{A}|v_1\rangle = |0\rangle \\ \mathbb{A}|v_2\rangle = |0\rangle \end{array} \right\} \Rightarrow \alpha_1 \mathbb{A}|v_1\rangle + \alpha_2 \mathbb{A}|v_2\rangle = |0\rangle = \mathbb{A}(\alpha_1|v_1\rangle + \alpha_2|v_2\rangle),$$

por la misma razón se tiene que el elemento neutro está contenido en $\mathfrak{N}(\mathbb{A})$, esto es:

$$\mathbb{A}|\alpha v\rangle = |0\rangle \quad \forall |v\rangle \in \mathbf{V}_1 \wedge \forall \alpha \therefore \mathbb{A}|0\rangle = |0\rangle \quad \text{si } \alpha = 0,$$

por lo tanto, queda demostrado que $\mathfrak{N}(\mathbb{A})$ es un subespacio de \mathbf{V}_1 .

Por otra parte, definiremos la imagen (rango o recorrido) de \mathbb{A} , y la denotaremos como:

$$\mathbb{A}\{\mathbf{V}\} = \{|v'\rangle \in \mathbf{V}_2 \wedge \mathbb{A}|v\rangle = |v'\rangle\},$$

igualmente $\mathfrak{I}(\mathbb{A}) \subset \mathbf{V}_2$ también será un subespacio de \mathbf{V}_2 ya que si $|v\rangle = \alpha_1|v_1\rangle + \alpha_2|v_2\rangle$ y dado que \mathbb{A} es un operador lineal, se cumple que:

$$\mathbb{A}\left(\underbrace{\alpha_1|v_1\rangle + \alpha_2|v_2\rangle}_{|v\rangle}\right) = \alpha_1 \underbrace{\mathbb{A}|v_1\rangle}_{|v'_1\rangle} + \alpha_2 \underbrace{\mathbb{A}|v_2\rangle}_{|v'_2\rangle} = \underbrace{\alpha_1|v'_1\rangle + \alpha_2|v'_2\rangle}_{|v'\rangle}.$$

Es claro que si \mathbf{V} es de dimensión finita n , el rango $\mathbb{A}\{\mathbf{V}\}$ también será de dimensión finita n y tendremos:

$$\dim[\mathfrak{N}(\mathbb{A})] + \dim[\mathbb{A}\{\mathbf{V}\}] = \dim[\mathbf{V}], \tag{4.8}$$

vale decir, que la dimensión del núcleo más la dimensión del recorrido o imagen de una transformación lineal es igual a la dimensión del dominio.

Para demostrar esta afirmación supongamos que $\dim[\mathbf{V}] = n$ y que $\{|e_1\rangle, |e_2\rangle, |e_3\rangle \dots |e_k\rangle\} \in \mathbf{V}$ es una base de $\mathfrak{N}(\mathbb{A})$, donde $k = \dim[\mathfrak{N}(\mathbb{A})] \leq n$.

Como el conjunto: $\{|e_1\rangle, |e_2\rangle, |e_3\rangle \dots |e_k\rangle\} \in \mathbf{V}$, entonces estos elementos forman una base y por lo tanto,

son linealmente independientes. Ademas, necesariamente estos vectores formarán parte de una base mayor de \mathbf{V} , esto es: $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_k\rangle, |e_{k+1}\rangle, \dots, |e_{k+r-1}\rangle, |e_{k+r}\rangle\} \in \mathbf{V}$ será una base de \mathbf{V} , donde $k + r = n$.

El esquema de la demostración será el siguiente:

- primero probaremos que los r elementos $\{\mathbb{A}\{|e_{k+1}\rangle\}, \mathbb{A}\{|e_{k+2}\rangle\}, \dots, \mathbb{A}\{|e_{k+r-1}\rangle\}, \mathbb{A}\{|e_{k+r}\rangle\}\}$ forman una base para $\mathbb{A}\{\mathbf{V}\}$. Esto significa que estaríamos probando que: $\dim\{\mathbb{A}\{\mathbf{V}\}\} = r$, $k + r = n$.
- luego hay que demostrar que los elementos $\{\mathbb{A}\{|e_{k+1}\rangle\}, \mathbb{A}\{|e_{k+2}\rangle\}, \dots, \mathbb{A}\{|e_{k+r-1}\rangle\}, \mathbb{A}\{|e_{k+r}\rangle\}\}$ son linealmente independientes.

Si los r elementos $\{\mathbb{A}\{|e_{k+1}\rangle\}, \mathbb{A}\{|e_{k+2}\rangle\}, \dots, \mathbb{A}\{|e_{k+r-1}\rangle\}, \mathbb{A}\{|e_{k+r}\rangle\}\}$ expanden $\mathbb{A}\{\mathbf{V}\}$ entonces cualquier elemento

$$|w\rangle \in \mathbb{A}\{\mathbf{V}\} / |w\rangle = \mathbb{A}|v\rangle = C^i|\mathbb{A}e_i\rangle, \quad \text{con } |\mathbb{A}e_i\rangle = \mathbb{A}|e_i\rangle.$$

Ahora bien, analicemos con cuidado los límites de la suma implícita del índice $i = 1, 2, \dots, k + r$

$$|w\rangle = C^i|\mathbb{A}e_i\rangle = \underbrace{C^1|\mathbb{A}e_1\rangle + C^2|\mathbb{A}e_2\rangle + \dots + C^k|\mathbb{A}e_k\rangle}_{=|0\rangle \text{ ya que } \mathbb{A}|e_1\rangle=\mathbb{A}|e_2\rangle=\mathbb{A}|e_3\rangle\dots=\mathbb{A}|e_k\rangle=|0\rangle} + C^{k+1}|\mathbb{A}e_{k+1}\rangle + \dots + C^{k+r}|\mathbb{A}e_{k+r}\rangle.$$

Por lo tanto $\{\mathbb{A}\{|e_{k+1}\rangle\}, \mathbb{A}\{|e_{k+2}\rangle\}, \dots, \mathbb{A}\{|e_{k+r-1}\rangle\}, \mathbb{A}\{|e_{k+r}\rangle\}\}$ expanden $\mathbb{A}\{\mathbf{V}\}$.

Ahora bien, para demostrar que son base, demostraremos que son linealmente independientes. Para ello supondremos que:

$$\exists \quad \left\{C^{k+1}, C^{k+2}, \dots, C^{k+r}\right\} / C^i|\mathbb{A}e_i\rangle = 0, \quad \text{con } i = k + 1, k + 2, \dots, k + r,$$

y tenemos que demostrar que $C^{k+1} = C^{k+2} = \dots = C^{k+r} = 0$.

Entonces

$$C^i|\mathbb{A}e_i\rangle = C^i\mathbb{A}|e_i\rangle = \mathbb{A}(C^i|e_i\rangle) = 0, \quad \text{con } i = k + 1, k + 2, \dots, k + r,$$

esto significa que el elemento $|v\rangle = C^i|e_i\rangle \in \mathbb{N}(\mathbb{A})$, con $i = k + 1, k + 2, \dots, k + r$.

Con lo cual, dado que:

$$\forall |v\rangle \in \mathbb{N}(\mathbb{A}), |v\rangle = C^i|e_i\rangle, \text{ con } i = 1, 2, \dots, r,$$

entonces se puede hacer la siguiente resta:

$$|v\rangle - |v\rangle = \sum_{i=1}^k C^i|e_i\rangle - \sum_{i=k+1}^{k+r} C^i|e_i\rangle = |0\rangle,$$

y como los $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_k\rangle, |e_{k+1}\rangle, \dots, |e_{k+r-1}\rangle, |e_{k+r}\rangle\}$ son una base de \mathbf{V} entonces resulta que los coeficientes: $C^{k+1} = C^{k+2} = \dots = C^{k+r} = 0$.

Si el espacio vectorial \mathbf{V} es de dimensión infinita, entonces al menos uno de los dos subespacios $\mathbb{N}(\mathbb{A})$ o $\mathbb{A}\{\mathbf{V}\}$ será de dimensión infinita.

A continuación ejemplificaremos algunos casos que ilustran lo que representa un espacio nulo.

1. Transformación identidad: Sea $\mathbb{I} : \mathbf{V}_1 \rightarrow \mathbf{V}_2$, la transformación identidad, entonces

$$\forall |v\rangle \in \mathbf{V}_1 / \mathbb{I}|v\rangle = |v\rangle \Rightarrow \mathbb{N}(\mathbb{I}) = \{|0\rangle\} \subset \mathbf{V}_1 \wedge \mathbb{A}\{\mathbf{V}\} \equiv \mathbf{V}_1.$$

2. Sistemas de ecuaciones lineales: En \mathbf{V}^n las soluciones a los sistemas de ecuaciones lineales representan el espacio nulo, $\mathbb{N}(\mathbb{A})$, para vectores de \mathbf{V}^n

$$\mathbb{A}|x\rangle = |0\rangle \iff \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & & \cdots \\ \vdots & & \ddots & \\ A_{n1} & A_{n2} & & A_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \iff A_j^i x_i = 0,$$

con j ecuaciones ($j = 1, 2, \dots, n$). Recordemos que estamos utilizando la convención de Einstein para suma de índices.

3. **Ecuaciones diferenciales ordinarias:** Sea $\mathcal{C}_{[-\infty, \infty]}^2$ el espacio vectorial de todas las funciones continuas doblemente diferenciables. Definimos $\mathbb{A} : \mathcal{C}_{[-\infty, \infty]}^2 \rightarrow \mathcal{C}_{[-\infty, \infty]}$ como la transformación lineal $(\mathbb{D}^2 - 1)$ tal que para todas las $y(x) \in \mathcal{C}_{[-\infty, \infty]}^2$ se cumple:

$$\mathbb{A}|x\rangle = |0\rangle \iff (\mathbb{D}^2 - 1)y(x) = 0 \iff \left(\frac{d^2}{dx^2} - 1 \right)y(x) = y'' - y = 0.$$

El núcleo o espacio nulo de \mathbb{A} , $\mathfrak{N}(\mathbb{A})$, lo constituye el conjunto de soluciones para la mencionada ecuación diferencial. Por lo tanto, el problema de encontrar las soluciones de la ecuación diferencial es equivalente a encontrar los elementos del núcleo de \mathbb{A} .

4.2.2 Operadores biyectivos e inversos

Se dice que $\mathbb{A} : \mathbf{V}_1 \rightarrow \mathbf{V}_2$ es biyectiva (uno a uno o biunívoco) si dados $|v_1\rangle, |v_2\rangle \in \mathbf{V}_1$, \wedge $|v'\rangle \in \mathbf{V}_2$, se tiene que:

$$\mathbb{A}|v_1\rangle = |v'\rangle \wedge \mathbb{A}|v_2\rangle = |v'\rangle \Rightarrow |v_1\rangle = |v_2\rangle,$$

es decir, será biyectiva si \mathbb{A} transforma vectores distintos de \mathbf{V}_1 en vectores distintos de \mathbf{V}_2 .

Más aún, se puede afirmar que una transformación lineal \mathbb{A} , será biyectiva si y sólo si $\mathfrak{N}(\mathbb{A}) = \{|0\rangle\}$. Vale decir, si el subespacio nulo está constituido, únicamente por el elemento neutro del espacio vectorial.

La demostración es sencilla. Supongamos que \mathbb{A} es biyectiva y que $\mathbb{A}|v\rangle = |0\rangle$, entonces $|v\rangle = |0\rangle$, es decir, $\mathbb{A}|0\rangle = |0\rangle$, por consiguiente $\mathfrak{N}(\mathbb{A}) = \{|0\rangle\}$. Recíprocamente, si

$$\left. \begin{array}{l} \mathfrak{N}(\mathbb{A}) = \{|0\rangle\} \\ \wedge \\ \mathbb{A}|v_1\rangle = \mathbb{A}|v_2\rangle \end{array} \right\} \Rightarrow \mathbb{A}|v_1\rangle - \mathbb{A}|v_2\rangle = |0\rangle = \mathbb{A}\underbrace{(|v_1\rangle - |v_2\rangle)}_{|v_1\rangle - |v_2\rangle = 0} \Rightarrow |v_1\rangle = |v_2\rangle.$$

La importancia de las transformaciones lineales biyectivas reside en la posibilidad de definir inversa, debido a que siempre existe en \mathbf{V}_2 un vector $|v'\rangle$ asociado a través de \mathbb{A} con un vector $|v\rangle \in \mathbf{V}_1$. Diremos que $\mathbb{A}^{-1} : \mathbf{V}_2 \rightarrow \mathbf{V}_1$ es el inverso de \mathbb{A} , si $\mathbb{A}^{-1}\mathbb{A} = \mathbb{I} = \mathbb{A}\mathbb{A}^{-1}$.

Habrá que hacer un par de comentarios al respecto. El primero es que, tal y como hemos enfatizado arriba, en general, los operadores no comutan entre si, y los inversos no son una excepción. Es decir, debería existir (y de hecho existen) inversas por la izquierda $\mathbb{A}^{-1}\mathbb{A}$ e inversas por la derecha $\mathbb{A}\mathbb{A}^{-1}$. Por simplicidad e importancia en Física obviaremos esta dicotomía y supondremos que $\mathbb{A}^{-1}\mathbb{A} = \mathbb{I} = \mathbb{A}\mathbb{A}^{-1}$. El segundo comentario tiene que ver con la existencia y unicidad del inverso de un operador lineal. Algunos operadores tienen inverso, otros no, pero aquellos que tienen inverso, ese inverso será único.

Supongamos:

$$\left. \begin{array}{l} \mathbb{A}_1^{-1}\mathbb{A}|v\rangle = |v\rangle \\ \wedge \\ \mathbb{A}_2^{-1}\mathbb{A}|v\rangle = |v\rangle \end{array} \right\} \Rightarrow \mathbb{A}_1^{-1}\mathbb{A}|v\rangle - \mathbb{A}_2^{-1}\mathbb{A}|v\rangle = |0\rangle = \underbrace{(\mathbb{A}_1^{-1} - \mathbb{A}_2^{-1})\mathbb{A}|v\rangle}_{\mathbb{A}_1^{-1} = \mathbb{A}_2^{-1}} \Rightarrow \mathbb{A}_1^{-1} = \mathbb{A}_2^{-1}.$$

Ahora bien, un operador lineal \mathbb{A} tendrá inverso si y sólo si para cada vector $|v'\rangle \in \mathbf{V}_2 \exists |v\rangle \in \mathbf{V}_1 / \mathbb{A}|v\rangle = |v'\rangle$. Es decir, cada vector $|v'\rangle$ está asociado con uno y sólo un vector $|v\rangle$ a través de la transformación lineal \mathbb{A} . Dejaremos sin demostración esta afirmación pero lo importante es recalcar que para que exista inverso la transformación lineal \mathbb{A} , tiene que ser biyectiva y esto implica que se asocia uno y sólo un vector de \mathbf{V}_1 con otro de \mathbf{V}_2 .

Esto nos lleva a la definición de espacios vectoriales isomorfos¹:

Definición 4.2 (Espacios Vectoriales Isomorfos)

Dos espacios vectoriales $\mathbf{V}_1, \mathbf{V}_2$ son *isomorfos* si existe una transformación

$$\mathbb{A} : \mathbf{V}_1 \rightarrow \mathbf{V}_2 \quad \Rightarrow \quad |v'\rangle = \mathbb{A}|v\rangle ,$$

denominada *isomorfismo*, tal que:

1. \mathbb{A} es biyectiva.
2. \mathbb{A} es lineal: $\mathbb{A}[\alpha|v_1\rangle + \beta|v_2\rangle] = \alpha\mathbb{A}|v_1\rangle + \beta\mathbb{A}|v_2\rangle \quad \forall |v_1\rangle \text{ y } |v_2\rangle \in \mathbf{V}_1$.

El que \mathbb{A} sea biyectiva nos garantiza que existe la transformación inversa \mathbb{A}^{-1} , que también será biyectiva. Es decir, podríamos hablar de manera equivalente de una transformación $\mathbb{F} : \mathbf{V}_2 \rightarrow \mathbf{V}_1$. Volveremos a este tema más adelante. Todavía podemos añadir algunas demostraciones que resultan ser consecuencias de las afirmaciones anteriores.

Sea la transformación lineal $\mathbb{A} : \mathbf{V}_1 \rightarrow \mathbf{V}_2$ y supongamos además que $\mathbb{A} \in \mathcal{L}(\mathbf{V}_1, \mathbf{V}_2)$. Entonces las siguientes afirmaciones son válidas y equivalentes:

1. \mathbb{A} es biyectiva en \mathbf{V}_1 .
2. \mathbb{A} es invertible y su inversa $\mathbb{A}^{-1} : \mathbb{A}\{\mathbf{V}_1\} \rightarrow \mathbf{V}_1$ es lineal.
3. $\forall |v\rangle \in \mathbf{V}_1, \mathbb{A}\{|v\rangle\} = |0\rangle \Rightarrow |v\rangle = |0\rangle$. Esto es, el espacio nulo $\aleph(\mathbb{A})$ únicamente contiene al elemento neutro de \mathbf{V}_1 .

Si ahora suponemos que \mathbf{V}_1 tiene dimensión finita, digamos $\dim[\mathbf{V}_1] = n$, las siguientes afirmaciones serán válidas y equivalentes:

1. \mathbb{A} es biyectiva en \mathbf{V}_1 .
2. Si $\{|u_1\rangle, |u_2\rangle, |u_3\rangle, \dots, |u_n\rangle\} \in \mathbf{V}_1$ son linealmente independientes, entonces, el conjunto de transformaciones lineales: $\{\mathbb{A}\{|u_1\rangle\}, \mathbb{A}\{|u_2\rangle\}, \mathbb{A}\{|u_3\rangle\}, \dots, \mathbb{A}\{|u_n\rangle\}\} \in \mathbb{A}\{\mathbf{V}_1\}$ también será linealmente independiente.
3. $\dim[\mathbb{A}\{\mathbf{V}_1\}] = n$.
4. Si $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\} \in \mathbf{V}_1$ es una base de \mathbf{V}_1 , entonces $\{\mathbb{A}\{|e_1\rangle\}, \mathbb{A}\{|e_2\rangle\}, \mathbb{A}\{|e_3\rangle\}, \dots, \mathbb{A}\{|e_n\rangle\}\} \in \mathbb{A}\{\mathbf{V}_1\}$ es una base de $\mathbb{A}\{\mathbf{V}_1\}$.

4.2.3 Operadores adjuntos

En la sección 4.1, definimos la acción de un operador lineal $\mathbb{A} : \mathbf{V} \rightarrow \mathbf{W}$ de tal forma que $\mathbb{A}|v\rangle = |v'\rangle$. En esta sección definiremos el operador:

$$\mathbb{A}^\dagger : \mathbf{W}^* \rightarrow \mathbf{V}^* \text{ de tal forma que } \langle v' | = \langle v | \mathbb{A}^\dagger ,$$

donde \mathbf{V}^* y \mathbf{W}^* son los duales de \mathbf{V} y \mathbf{W} , respectivamente. Diremos que el operador \mathbb{A}^\dagger es el adjunto de \mathbb{A} ².

Entonces, discutimos en la sección 3.1.1, a cada vector (*ket*) $|v\rangle$ le está asociado una forma lineal (*bra*) $\langle v|$, a cada *ket* transformado $\mathbb{A}|v\rangle = |v'\rangle$ le corresponderá un *bra* transformado $\langle v' | = \langle v | \mathbb{A}^\dagger$. En pocas palabras:

$$|v\rangle \iff \langle v | \quad \Rightarrow \quad |v'\rangle = \mathbb{A}|v\rangle \iff \langle v' | = \langle v | \mathbb{A}^\dagger .$$

Si \mathbb{A} es lineal, \mathbb{A}^\dagger también lo será, dado que a un vector $|w\rangle = \lambda_1|z_1\rangle + \lambda_2|z_2\rangle$ le corresponde un *bra*

¹Se dice que una transformación lineal es un *monomorfismo* si \mathbb{A} es inyectiva, un *epimorfismo* si \mathbb{A} es sobreyectiva y un *isomorfismo* si es biyectiva.

²En la literatura también encontrarán que \mathbb{A}^\dagger es el hermítico conjugado de \mathbb{A} , pero hemos creído conveniente llamarlo únicamente el adjunto de \mathbb{A} para evitar confusiones con operadores hermíticos.

$\langle w| = \lambda_1^* \langle z_1| + \lambda_2^* \langle z_2|$.³ Por lo tanto, $|w'\rangle = \mathbb{A}|w\rangle = \lambda_1 \mathbb{A}|z_1\rangle + \lambda_2 \mathbb{A}|z_2\rangle$, por ser \mathbb{A} lineal, entonces

$$|w'\rangle \iff \langle w'| \equiv \langle w| \mathbb{A}^\dagger = (\lambda_1^* \langle z_1| + \lambda_2^* \langle z_2|) \mathbb{A}^\dagger \equiv \lambda_1^* \langle z'_1| + \lambda_2^* \langle z'_2| = \lambda_1^* \langle z_1| \mathbb{A}^\dagger + \lambda_2^* \langle z_2| \mathbb{A}^\dagger.$$

Es claro que a partir de la definición de producto interno tendremos:

$$\langle \tilde{x}|y\rangle = \langle y|\tilde{x}\rangle^* \quad \forall \quad |\tilde{x}\rangle = \mathbb{A}|x\rangle, |y\rangle \in \mathbf{V} \Rightarrow \langle x|\mathbb{A}^\dagger|y\rangle = \langle y|\mathbb{A}|x\rangle^* \quad \forall \quad |x\rangle, |y\rangle \in \mathbf{V}.$$

Esta última relación

$$\langle x|\mathbb{A}^\dagger|y\rangle = \langle y|\mathbb{A}|x\rangle^* \quad \forall \quad |x\rangle, |y\rangle \in \mathbf{V}, \quad (4.9)$$

nos permite asociar \mathbb{A}^\dagger con \mathbb{A} . Esto es, conociendo las propiedades de \mathbb{A} las extendemos a \mathbb{A}^\dagger y es a partir de esta relación que podemos deducir las propiedades de los operadores adjuntos:

$$(\mathbb{A}^\dagger)^\dagger = \mathbb{A}, \quad (\lambda\mathbb{A})^\dagger = \lambda^*\mathbb{A}^\dagger, \quad (\mathbb{A} + \mathbb{B})^\dagger = \mathbb{A}^\dagger + \mathbb{B}^\dagger, \quad (\mathbb{A}\mathbb{B})^\dagger = \mathbb{B}^\dagger\mathbb{A}^\dagger.$$

Esta última propiedad es fácilmente demostrable y es educativa su demostración. Dado que $|v'\rangle = \mathbb{A}\mathbb{B}|v\rangle$, y además se tiene que:

$$\left. \begin{array}{l} |\bar{v}\rangle = \mathbb{B}|v\rangle \\ |v'\rangle = \mathbb{A}|\bar{v}\rangle \end{array} \right\} \Rightarrow \langle v'| = \langle \bar{v}|\mathbb{A}^\dagger = \langle v|\mathbb{B}^\dagger\mathbb{A}^\dagger = \langle v|(\mathbb{A}\mathbb{B})^\dagger.$$

A partir de las propiedades anteriores se deriva una más útil relacionada con el comutador,

$$[\mathbb{A}, \mathbb{B}]^\dagger = - [\mathbb{A}^\dagger, \mathbb{B}^\dagger] = [\mathbb{B}^\dagger, \mathbb{A}^\dagger].$$

Las conclusiones a las que llegamos resultan ser que, para obtener el adjunto de una expresión, se debe proceder de la siguiente manera:

- Cambiar constantes por sus complejas conjugadas $\lambda \iff \lambda^*$.
- Cambiar los *kets* por sus *bras* asociados y viceversa (*bras* por *kets*): $|v\rangle \iff \langle v|$.
- Cambiar operadores lineales por sus adjuntos $\mathbb{A}^\dagger \iff \mathbb{A}$.
- Invertir el orden de los factores.

De este modo $(|v\rangle\langle w|)^\dagger = |w\rangle\langle v|$, que se deduce fácilmente de la consecuencia de la definición de producto interno

$$\langle x| \left(\underline{|v\rangle\langle w|} \right)^\dagger |y\rangle = \langle y|(|v\rangle\langle w|)|x\rangle^* = \langle y| |v\rangle^* \langle w| |x\rangle^* = \langle x| \underline{|w\rangle\langle v|} |y\rangle.$$

4.2.4 Operadores hermíticos

Existe un conjunto de operadores que se denominan hermíticos o autoadjuntos. Un operador hermítico (o autoadjunto) será aquel para el cual su adjunto es el mismo operador: $\mathbb{A}^\dagger = \mathbb{A}$.⁴ Entonces, a partir de (4.9) tendremos:

$$\langle x|\mathbb{A}^\dagger|y\rangle \equiv \langle x|\mathbb{A}|y\rangle = \langle y|\mathbb{A}|x\rangle^*.$$

Estos operadores juegan el rol de los números reales en el sentido de que son “iguales a su propio complejo conjugado” y se utilizan como los observables en la Mecánica Cuántica. Claramente los proyectores son operadores autoadjuntos por construcción: $\mathbb{P}_{|v\rangle}^\dagger \equiv (|v\rangle\langle v|)^\dagger = |v\rangle\langle v|$.

³La correspondencia es antilineal, note la conjugación de los números λ_1 y λ_2 .

⁴Si $\mathbb{A}^\dagger = -\mathbb{A}$ el operador lineal se denomina antihermítico.

4.2.5 Operadores unitarios

Por definición, un operador será unitario si su inversa es igual a su adjunto:

$$\mathbb{U}^{-1} = \mathbb{U}^\dagger \Rightarrow \mathbb{U}^\dagger \mathbb{U} = \mathbb{U} \mathbb{U}^\dagger = \mathbb{I}.$$

De estos operadores podemos decir varias cosas:

- Las transformaciones unitarias dejan invariante al producto interno y consecuentemente la norma de vectores y esto se demuestra fácilmente. Dados dos vectores $|x\rangle, |y\rangle$ sobre los cuales actúa un operador unitario

$$\left. \begin{array}{l} |x'\rangle = \mathbb{U}|x\rangle \\ |y'\rangle = \mathbb{U}|y\rangle \end{array} \right\} \Rightarrow \langle y' | x' \rangle = \langle y | \mathbb{U}^\dagger \mathbb{U} | x \rangle = \langle y | x \rangle .$$

Es claro que si \mathbb{A} es hermítico, $\mathbb{A}^\dagger = \mathbb{A}$, el operador $\mathbb{U} = e^{i\mathbb{A}}$ es unitario.

$$\mathbb{U} = e^{i\mathbb{A}} \Rightarrow \mathbb{U}^\dagger = e^{-i\mathbb{A}^\dagger} = e^{-i\mathbb{A}} \Rightarrow \mathbb{U}\mathbb{U}^\dagger = e^{i\mathbb{A}}e^{-i\mathbb{A}} = \mathbb{I} = \mathbb{U}^\dagger\mathbb{U} = e^{-i\mathbb{A}}e^{i\mathbb{A}}.$$

- El producto de dos operadores unitarios también es unitario. Esto es, si \mathbb{U} y \mathbb{V} son unitarios entonces:

$$(\mathbb{U}\mathbb{V})^\dagger (\mathbb{U}\mathbb{V}) = \mathbb{V}^\dagger \underbrace{\mathbb{U}^\dagger \mathbb{U}}_{\mathbb{I}} \mathbb{V} = \mathbb{V}^\dagger \mathbb{V} = \mathbb{I} \Leftrightarrow (\mathbb{U}\mathbb{V})(\mathbb{U}\mathbb{V})^\dagger = \mathbb{U}^\dagger \underbrace{\mathbb{V}^\dagger \mathbb{V}}_{\mathbb{I}} \mathbb{U} = \mathbb{U}^\dagger \mathbb{U} = \mathbb{I}$$

- La invariancia del producto interno implica que los operadores unitarios aplican una base ortogonal en otra: $\{|e_i\rangle\} \xrightarrow{\mathbb{U}} \{|\hat{e}_{i'}\rangle\}$. Veamos, si $\{|\hat{e}_i\rangle\}$ es una base ortonormal para \mathbb{V} entonces:

$$|\hat{e}_{j'}\rangle = \mathbb{U}|\hat{e}_j\rangle \Rightarrow \langle \hat{e}^{i'} | \hat{e}_{j'} \rangle = \langle \hat{e}^{i'} | \mathbb{U} | \hat{e}_j \rangle = \langle \hat{e}^i | \mathbb{U}^\dagger \mathbb{U} | \hat{e}_j \rangle = \langle \hat{e}^i | \hat{e}_j \rangle = \delta_j^i.$$

4.2.6 Funciones de operadores

Para construir funciones de operadores lineales, procedemos por analogía, basándonos en el primero de los ejemplos de la sección 4.1.2. Vale decir que se puede construir un “polinomio” en potencias de los operadores a partir de la idea:

$$P_n(x) = a_0 + a_1 x + \cdots + a_n x^n = a_i x^i \Leftrightarrow P_n(\mathbb{A}) |v\rangle = [a_0 + a_1 \mathbb{A} + \cdots + a_n \mathbb{A}^n] |v\rangle = [a_i \mathbb{A}^i] |v\rangle ,$$

$$\forall |v\rangle \in \mathbb{V}_1.$$

Si nos saltamos todos los detalles de convergencia de la serie anterior –los cuales dependerán de los autovalores de \mathbb{A} y de su radio de convergencia– y nos inspiramos en el desarrollo de una función $F(z)$ como una serie de potencias de z en un cierto dominio, es posible expresar la función de un operador, $F(\mathbb{A})$, como una serie de potencias del operador \mathbb{A} , esto es:

$$F(z) = a_i z^i \Leftrightarrow F(\mathbb{A}) |v\rangle = [a_i \mathbb{A}^i] |v\rangle .$$

Tal y como se hace en el caso de funciones, “desarrollamos por Taylor” la función como una serie de potencias del operador:

$$F(z) = \sum_{n=0}^{\infty} f^{(n)}(0) \frac{z^n}{n!} \Leftrightarrow F(\mathbb{A}) |v\rangle = \left[\sum_{n=0}^{\infty} f^{(n)}(0) \frac{\mathbb{A}^n}{n!} \right] |v\rangle , \quad (4.10)$$

de esta manera podemos expresar la exponencial de un operador \mathbb{A} , como

$$e^{\mathbb{A}} |v\rangle = \left[\sum_{n=0}^{\infty} \frac{\mathbb{A}^n}{n!} \right] |v\rangle = \left[\mathbb{I} + \mathbb{A} + \frac{\mathbb{A}^2}{2!} + \cdots + \frac{\mathbb{A}^n}{n!} \cdots \right] |v\rangle .$$

En este caso hay que hacer una acotación, dado que, en general, $[\mathbb{A}, \mathbb{B}] \neq 0 \Rightarrow e^{\mathbb{A}}e^{\mathbb{B}} \neq e^{\mathbb{B}}e^{\mathbb{A}} \neq e^{\mathbb{A}+\mathbb{B}}$. Esta

afirmación se corrobora de manera inmediata al desarrollar las exponenciales:

$$e^{\mathbb{A}} e^{\mathbb{B}} |v\rangle = \left[\sum_{n=0}^{\infty} \frac{\mathbb{A}^n}{n!} \right] \left[\sum_{m=0}^{\infty} \frac{\mathbb{B}^m}{m!} \right] |v\rangle = \left[\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\mathbb{A}^n \mathbb{B}^m}{n! m!} \right] |v\rangle ,$$

$$e^{\mathbb{B}} e^{\mathbb{A}} |v\rangle = \left[\sum_{n=0}^{\infty} \frac{\mathbb{B}^n}{n!} \right] \left[\sum_{m=0}^{\infty} \frac{\mathbb{A}^m}{m!} \right] |v\rangle = \left[\sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \frac{\mathbb{B}^n \mathbb{A}^m}{n! m!} \right] |v\rangle ,$$

y sólo en el caso en que $[\mathbb{A}, \mathbb{B}] = 0$ se tiene

$$e^{\mathbb{A}+\mathbb{B}} |v\rangle = \left[\sum_{n=0}^{\infty} \frac{(\mathbb{A} + \mathbb{B})^n}{n!} \right] |v\rangle ,$$

es decir, si $[\mathbb{A}, \mathbb{B}] = 0 \Rightarrow e^{\mathbb{A}} e^{\mathbb{B}} = e^{\mathbb{B}} e^{\mathbb{A}} = e^{\mathbb{A}+\mathbb{B}}$. La demostración no es inmediata y la haremos al final de la próxima sección en la cual desarrollaremos el concepto de derivada de operadores.

4.2.7 Diferenciación de operadores

Distinguiremos dos casos en la diferenciación de operadores: uno cuando el operador depende de una variable y diferenciamos respecto esa variable, $\frac{d e^{\mathbb{A}(t)}}{dt} |v\rangle$ y otro cuando diferenciamos respecto al operador mismo: $\frac{d F(\mathbb{B})}{d \mathbb{B}}$.

Empecemos por considerar operadores $\mathbb{A}(t)$, es decir que pueden depender de una variable arbitraria t . Podremos entonces definir la derivada como:

$$\frac{d \mathbb{A}(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{A}(t + \Delta t) - \mathbb{A}(t)}{\Delta t} .$$

Como era de esperarse, con esta definición se cumplirán todas las propiedades de las derivadas de funciones⁵.

Empecemos por considerar el caso más simple: $\mathbb{A}(t) = \mathbb{A}t$, el operador dependen linealmente de la variable t . Si queremos conocer la expresión para $\frac{d e^{\mathbb{A}t}}{dt}$, para este caso elemental recordemos que

$$e^{\mathbb{A}t} |v\rangle = \left[\sum_{n=0}^{\infty} \frac{(\mathbb{A}t)^n}{n!} \right] |v\rangle = \left[\mathbb{I} + \mathbb{A}t + \frac{(\mathbb{A}t)^2}{2!} + \cdots + \frac{(\mathbb{A}t)^n}{n!} \cdots \right] |v\rangle , \quad (4.11)$$

por lo tanto, tendremos:

$$\frac{d e^{\mathbb{A}t}}{dt} |v\rangle = \frac{d}{dt} \left[\sum_{n=0}^{\infty} \frac{(\mathbb{A}t)^n}{n!} \right] |v\rangle = \left[\sum_{n=0}^{\infty} \frac{d}{dt} \left(\frac{(\mathbb{A}t)^n}{n!} \right) \right] |v\rangle = \left[\sum_{n=0}^{\infty} \frac{n t^{n-1} \mathbb{A}^n}{n!} \right] |v\rangle = \underbrace{\left[\sum_{n=0}^{\infty} \frac{t^{n-1} \mathbb{A}^{n-1}}{(n-1)!} \right]}_{e^{\mathbb{A}t}} \mathbb{A} |v\rangle .$$

Nótese que la suma es hasta infinito, por lo tanto, al cambiar de índice $p = n - 1$, p sigue variando hasta infinito y la serie es la misma que la anterior. Finalmente, obtendremos⁶:

$$\frac{d e^{\mathbb{A}t}}{dt} |v\rangle = e^{\mathbb{A}t} \mathbb{A} |v\rangle \equiv \mathbb{A} e^{\mathbb{A}t} |v\rangle \Rightarrow [e^{\mathbb{A}t}, \mathbb{A}] = 0 .$$

En general también es fácil demostrar que $[F(\mathbb{A}), \mathbb{A}] = 0$ ya que a partir del desarrollo (4.10) se hace evidente

$$\left(\sum_{n=0}^{\infty} f^{(n)}(0) \frac{\mathbb{A}^n}{n!} \right) \mathbb{A} \equiv \mathbb{A} \left(\sum_{n=0}^{\infty} f^{(n)}(0) \frac{\mathbb{A}^n}{n!} \right) ,$$

Ahora bien, cuando se presenta la siguiente situación:

$$\frac{d (e^{\mathbb{A}t} e^{\mathbb{B}t})}{dt} |v\rangle = \frac{d e^{\mathbb{A}t}}{dt} e^{\mathbb{B}t} |v\rangle + e^{\mathbb{A}t} \frac{d e^{\mathbb{B}t}}{dt} |v\rangle = \mathbb{A} e^{\mathbb{A}t} e^{\mathbb{B}t} |v\rangle + e^{\mathbb{A}t} \mathbb{B} e^{\mathbb{B}t} |v\rangle = e^{\mathbb{A}t} \mathbb{A} e^{\mathbb{B}t} |v\rangle + e^{\mathbb{A}t} e^{\mathbb{B}t} \mathbb{B} |v\rangle ,$$

⁵Más adelante, en la sección 4.3.6 consideraremos la expresión matricial de los operadores y serán evidentes estas propiedades que aquí presentamos sin mayores justificaciones.

⁶Es inmediato comprobarlo si consideramos la expansión 4.11

hay que cuidar el orden en el cual se presentan los operadores. En particular en la expresión anterior hemos utilizado que siempre se cumple $[e^{\mathbb{B}t}, \mathbb{B}] = 0$, pero no hemos supuesto (ni conocemos) nada de $[\mathbb{A}, \mathbb{B}]$. Si, adicionalmente, $[\mathbb{A}, \mathbb{B}] = 0$ podremos factorizar $e^{\mathbb{A}t}e^{\mathbb{B}t}$ y tendremos

$$\frac{d(e^{\mathbb{A}t}e^{\mathbb{B}t})}{dt}|v\rangle = (\mathbb{A} + \mathbb{B})e^{\mathbb{A}t}e^{\mathbb{B}t}|v\rangle,$$

pero si $[\mathbb{A}, \mathbb{B}] \neq 0$, el orden de aparición de los operadores es muy importante.

Para construir la expresión de la derivada de una función de operador respecto a su argumento, probaremos la siguiente afirmación:

$$\text{Si } [\mathbb{A}, [\mathbb{A}, \mathbb{B}]] = [\mathbb{B}, [\mathbb{A}, \mathbb{B}]] = 0 \Rightarrow [\mathbb{A}, F(\mathbb{B})] = [\mathbb{A}, \mathbb{B}] \frac{dF(\mathbb{B})}{d\mathbb{B}}. \quad (4.12)$$

Esta relación es fácilmente demostrable si suponemos $[\mathbb{A}, \mathbb{B}] = \mathbb{I}$, el operador identidad. Obviamente aquí se cumple que: $[\mathbb{A}, \mathbb{B}] = \mathbb{I} \Rightarrow [\mathbb{A}, [\mathbb{A}, \mathbb{B}]] = [\mathbb{B}, [\mathbb{A}, \mathbb{B}]] = 0$.

En ese caso es fácil demostrar que: $\mathbb{A}\mathbb{B}^n - \mathbb{B}^n\mathbb{A} = n\mathbb{B}^{n-1}$:

$$\begin{aligned} \mathbb{A}\mathbb{B}^n - \mathbb{B}^n\mathbb{A} &= \underbrace{\mathbb{A}\mathbb{B}\cdots\mathbb{B}}_n - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n = (\mathbb{I} + \mathbb{B}\mathbb{A})\underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}}_{n-1} - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n \\ &= \mathbb{I}\mathbb{B}^{n-1} + \mathbb{B}(\mathbb{I} + \mathbb{B}\mathbb{A})\underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}}_{n-2} - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n \\ &= 2\mathbb{B}^{n-1} + \mathbb{B}^2(\mathbb{I} + \mathbb{B}\mathbb{A})\underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}}_{n-3} - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n = \dots = n\mathbb{B}^{n-1}. \end{aligned}$$

Para demostrar la relación (4.12) “desarrollamos en serie de Taylor” la función $F(\mathbb{B})$ en el conmutador

$$\begin{aligned} [\mathbb{A}, F(\mathbb{B})] &= \left[\mathbb{A}, \sum_{n=0}^{\infty} f_n \frac{\mathbb{B}^n}{n!} \right] = \sum_{n=0}^{\infty} f^{(n)}(0) \frac{[\mathbb{A}, \mathbb{B}^n]}{n!} = [\mathbb{A}, \mathbb{B}] \sum_{n=0}^{\infty} f^{(n)}(0) \frac{n\mathbb{B}^{n-1}}{n!} \\ &= [\mathbb{A}, \mathbb{B}] \sum_{n=0}^{\infty} f^{(n)}(0) \frac{\mathbb{B}^{n-1}}{(n-1)!} = [\mathbb{A}, \mathbb{B}] \frac{dF(\mathbb{B})}{d\mathbb{B}}. \end{aligned}$$

Para el caso más general: si

$$[\mathbb{A}, \mathbb{C}] = [\mathbb{B}, \mathbb{C}] = 0 \text{ con } \mathbb{C} = [\mathbb{A}, \mathbb{B}] \Rightarrow [\mathbb{A}, F(\mathbb{B})] \stackrel{?}{=} [\mathbb{A}, \mathbb{B}] \frac{dF(\mathbb{B})}{d\mathbb{B}},$$

se procede del mismo modo.

Probamos primero que: si

$$[\mathbb{A}, \mathbb{C}] = [\mathbb{B}, \mathbb{C}] = 0, \quad \text{con } \mathbb{C} = [\mathbb{A}, \mathbb{B}] \Rightarrow [\mathbb{A}, \mathbb{B}^n] = \mathbb{A}\mathbb{B}^n - \mathbb{B}^n\mathbb{A} = n[\mathbb{A}, \mathbb{B}]\mathbb{B}^{n-1}.$$

Tendremos:

$$\begin{aligned} \mathbb{A}\mathbb{B}^n - \mathbb{B}^n\mathbb{A} &= \underbrace{\mathbb{A}\mathbb{B}\cdots\mathbb{B}}_n - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n = (\mathbb{C} + \mathbb{B}\mathbb{A})\underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}}_{n-1} - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n \\ &= \mathbb{C}\mathbb{B}^{n-1} + \mathbb{B}(\mathbb{C} + \mathbb{B}\mathbb{A})\underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}}_{n-2} - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n \\ &= 2\mathbb{C}\mathbb{B}^{n-1} + \mathbb{B}^2(\mathbb{C} + \mathbb{B}\mathbb{A})\underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}}_{n-3} - \underbrace{\mathbb{B}\mathbb{B}\cdots\mathbb{B}\mathbb{A}}_n = \dots = n\mathbb{C}\mathbb{B}^{n-1} = n[\mathbb{A}, \mathbb{B}]\mathbb{B}^{n-1}, \end{aligned}$$

con lo cual es inmediato demostrar que:

$$\begin{aligned} [\mathbb{A}, F(\mathbb{B})] &= \left[\mathbb{A}, \sum_{n=0}^{\infty} f^{(n)}(0) \frac{\mathbb{B}^n}{n!} \right] = \sum_{n=0}^{\infty} f^{(n)}(0) \frac{[\mathbb{A}, \mathbb{B}^n]}{n!} = [\mathbb{A}, \mathbb{B}] \sum_{n=0}^{\infty} f^{(n)}(0) \frac{n\mathbb{B}^{n-1}}{n!} \\ &= [\mathbb{A}, \mathbb{B}] \sum_{n=0}^{\infty} f^{(n)}(0) \frac{\mathbb{B}^{n-1}}{(n-1)!} = [\mathbb{A}, \mathbb{B}] \frac{dF(\mathbb{B})}{d\mathbb{B}}. \end{aligned}$$

4.2.7.1 La fórmula de Glauber

Ahora estamos en capacidad de demostrar limpiamente la fórmula de Glauber:

$$e^{\mathbb{A}} e^{\mathbb{B}} = e^{\mathbb{A} + \mathbb{B}} e^{\frac{1}{2}[\mathbb{A}, \mathbb{B}]}$$

Para demostrarla, procedemos a considerar un operador $\mathbb{F}(t) = e^{\mathbb{A}t} e^{\mathbb{B}t}$, por lo tanto:

$$\begin{aligned} \frac{d\mathbb{F}(t)}{dt} |v\rangle &= \frac{de^{\mathbb{A}t}e^{\mathbb{B}t}}{dt} |v\rangle = \mathbb{A}e^{\mathbb{A}t}e^{\mathbb{B}t} |v\rangle + e^{\mathbb{A}t}\mathbb{B}e^{\mathbb{B}t} |v\rangle = (\mathbb{A} + e^{\mathbb{A}t}\mathbb{B}e^{-\mathbb{A}t}) e^{\mathbb{A}t}e^{\mathbb{B}t} |v\rangle \\ &= (\mathbb{A} + e^{\mathbb{A}t}\mathbb{B}e^{-\mathbb{A}t}) \mathbb{F}(t) |v\rangle . \end{aligned}$$

Ahora bien, dado que: $[\mathbb{A}, [\mathbb{A}, \mathbb{B}]] = [\mathbb{B}, [\mathbb{A}, \mathbb{B}]] = 0 \Rightarrow [\mathbb{A}, F(\mathbb{B})] = [\mathbb{A}, \mathbb{B}] \frac{dF(\mathbb{B})}{d\mathbb{B}}$,

entonces: $[e^{\mathbb{A}t}, \mathbb{B}] = t[\mathbb{A}, \mathbb{B}]e^{\mathbb{A}t} \Rightarrow e^{\mathbb{A}t}\mathbb{B} = \mathbb{B}e^{\mathbb{A}t} + t[\mathbb{A}, \mathbb{B}]e^{\mathbb{A}t}$,

es decir: $\frac{d\mathbb{F}(t)}{dt} |v\rangle = (\mathbb{A} + e^{\mathbb{A}t}\mathbb{B}e^{-\mathbb{A}t}) \mathbb{F}(t) |v\rangle = (\mathbb{A} + \mathbb{B} + t[\mathbb{A}, \mathbb{B}]) \mathbb{F}(t) |v\rangle$,

por tanto uno puede darse cuenta que: $\mathbb{F}(t) = e^{\left\{(\mathbb{A}+\mathbb{B})t + \frac{t^2}{2}[\mathbb{A}, \mathbb{B}]\right\}}$,

cumple con la ecuación anterior, por lo tanto absorbiendo t en los operadores correspondientes llegamos a la fórmula de Glauber: $e^{\mathbb{A}} e^{\mathbb{B}} = e^{\mathbb{A} + \mathbb{B}} e^{\frac{1}{2}[\mathbb{A}, \mathbb{B}]}$.

4.2.8 Ejemplos

1. Sean \mathbb{A} y \mathbb{B} dos operadores hermíticos y un operador unitario definido como: $\mathbb{U} = \mathbb{A} + i\mathbb{B}$.

Mostraremos que $[\mathbb{A}, \mathbb{B}] = 0$ y $\mathbb{A}^2 + \mathbb{B}^2 = \mathbb{I}$.

Comenzamos con $[\mathbb{A}, \mathbb{B}] = 0$:

$$\begin{aligned} \mathbb{U}\mathbb{U}^\dagger &= \mathbb{U}^\dagger\mathbb{U} = (\mathbb{A} + i\mathbb{B})(\mathbb{A} + i\mathbb{B})^\dagger = (\mathbb{A} + i\mathbb{B})^\dagger(\mathbb{A} + i\mathbb{B}) \Rightarrow (\mathbb{A} + i\mathbb{B})(\mathbb{A} - i\mathbb{B}) = (\mathbb{A} - i\mathbb{B})(\mathbb{A} + i\mathbb{B}) \\ \mathbb{B}\mathbb{A} - \mathbb{A}\mathbb{B} &= -\mathbb{B}\mathbb{A} + \mathbb{A}\mathbb{B} \Rightarrow [\mathbb{B}, \mathbb{A}] = -[\mathbb{B}, \mathbb{A}] \Rightarrow [\mathbb{B}, \mathbb{A}] = 0 . \end{aligned}$$

Continuamos con la segunda de las afirmaciones, que se puede demostrar, a partir de:

$$\mathbb{U}\mathbb{U}^\dagger = \mathbb{I} \Rightarrow (\mathbb{A} + i\mathbb{B})(\mathbb{A} + i\mathbb{B})^\dagger = (\mathbb{A} + i\mathbb{B})(\mathbb{A} - i\mathbb{B}) = (\mathbb{A}^2 + \mathbb{B}^2 + i(\mathbb{B}\mathbb{A} - \mathbb{A}\mathbb{B})) \Rightarrow \mathbb{I} = \mathbb{A}^2 + \mathbb{B}^2 .$$

2. Un operador cantidad de movimiento generalizado se define como aquel conjunto de operadores hermíticos que cumplen con:

$$[\mathbb{J}_x, \mathbb{J}_y] = i\hbar\mathbb{J}_z \quad [\mathbb{J}_y, \mathbb{J}_z] = i\hbar\mathbb{J}_x \quad [\mathbb{J}_z, \mathbb{J}_x] = i\hbar\mathbb{J}_y , \quad \text{es decir} \quad [\mathbb{J}_i, \mathbb{J}_j] = i\hbar\epsilon_{ijk}\mathbb{J}_k ,$$

con ϵ_{ijk} el símbolo de Levi-Civita (Aquí los índices repetidos NO indican suma).

Adicionalmente, definimos los siguientes operadores:

$$\mathbb{J}^2 = \mathbb{J}_x^2 + \mathbb{J}_y^2 + \mathbb{J}_z^2 ; \quad \mathbb{J}_+ = \mathbb{J}_x + i\mathbb{J}_y \quad \mathbb{J}_- = \mathbb{J}_x - i\mathbb{J}_y .$$

Demostremos que: $[\mathbb{J}^2, \mathbb{J}_+] = [\mathbb{J}^2, \mathbb{J}_-] = [\mathbb{J}^2, \mathbb{J}_z] = 0$.

Para probar esta propiedad se puede demostrar de forma genérica que $[\mathbb{J}_k^2, \mathbb{J}_m] = 0$, donde los índices $k, m = 1, 2, 3 \equiv x, y, z$, esto es

$$[\mathbb{J}_k^2, \mathbb{J}_m] = [\mathbb{J}_k\mathbb{J}_k, \mathbb{J}_m] = \mathbb{J}_k\mathbb{J}_k\mathbb{J}_m - \mathbb{J}_m\mathbb{J}_k\mathbb{J}_k = \mathbb{J}_k\mathbb{J}_k\mathbb{J}_m - (i\hbar\epsilon_{mkl}\mathbb{J}_l + \mathbb{J}_k\mathbb{J}_m)\mathbb{J}_k .$$

Notemos que $[\mathbb{J}_i, \mathbb{J}_j] = i\hbar\epsilon_{ijk}\mathbb{J}_k \Rightarrow \mathbb{J}_m\mathbb{J}_k = i\hbar\epsilon_{mkl}\mathbb{J}_l + \mathbb{J}_k\mathbb{J}_m$, y que hemos realizado esa sustitución en

el segundo de los términos de arriba. Desarrollando y realizando una sustitución equivalente obtenemos:

$$[\mathbb{J}_k^2, \mathbb{J}_m] = \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m - i\hbar \epsilon_{mkl} \mathbb{J}_l \mathbb{J}_k - \mathbb{J}_k (i\hbar \epsilon_{mkn} \mathbb{J}_n + \mathbb{J}_k \mathbb{J}_m) ,$$

y claramente $[\mathbb{J}_k^2, \mathbb{J}_m] = 0$, por cuanto los índices no suman pero si son mudos, y $\epsilon_{mkl} = -\epsilon_{mlk}$.

$$[\mathbb{J}_k^2, \mathbb{J}_m] = \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m - i\hbar \epsilon_{mkl} \mathbb{J}_l \mathbb{J}_k - i\hbar \epsilon_{mkn} \mathbb{J}_k \mathbb{J}_n - \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m ,$$

al conmutar los cuadrados de las componentes con cualquiera de las componentes, y dado que los conmutadores son lineales, entonces queda demostrado que:

$$[\mathbb{J}^2, \mathbb{J}_\pm] = [\mathbb{J}_x^2 + \mathbb{J}_y^2 + \mathbb{J}_z^2, \mathbb{J}_x \pm i\mathbb{J}_y] = [\mathbb{J}_y^2, \mathbb{J}_x] + [\mathbb{J}_z^2, \mathbb{J}_x] \pm i [\mathbb{J}_x^2, \mathbb{J}_y] \pm i [\mathbb{J}_z^2, \mathbb{J}_y] = 0 .$$

4.2.9 Practicando con Maxima

Consideremos la siguiente transformación lineal:

$$\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^4, \quad \mathbb{T}[(x, y, z)] = (2x + y, x - z, x + y + z, y + z) .$$

Podemos preguntarnos si es una transformación biyectiva, para averiguarlo calculemos el núcleo de la transformación. Hagamos entonces lo siguiente:

```
(%i1) solve([2*x+y=0,x-z=0,x+y+z=0,y+z=0],[x,y,z]);
(%o1) [[x = 0, y = 0, z = 0]]
```

De manera que $\mathcal{N}(\mathbb{T}) = \{(0, 0, 0) \in \mathbb{R}^3\}$ y la transformación es biyectiva. Consideremos ahora la transformación lineal:

$$\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^4, \quad \mathbb{T}[(x, y, z)] = (x + y, x + z, 2x + y + z, y - z)$$

```
(%i2) solve([x+y=0,x+z=0,2*x+y+z=0,y-z=0],[x,y,z]);
(%o2) [[x = -%r1, y = %r1, z = %r1]]
```

Por lo tanto $\mathcal{N}(\mathbb{T}) = \{(-\lambda, \lambda, \lambda), \lambda \in \mathbb{R}\}$. En este caso, la transformación no es biyectiva. Notemos que el vector $|e\rangle = (-1, 1, 1)$ resulta ser el vector base para subespacio vectorial núcleo, por lo tanto, la dimensión de este espacio vectorial es: $\dim[\mathcal{N}(\mathbb{T})] = 1$. Como veremos más adelante una base en el espacio imagen o rango es el conjunto: $\{\mathbb{T}[(1, 0, 0)], \mathbb{T}[(0, 1, 0)], \mathbb{T}[(0, 0, 1)]\}$.

Por lo tanto, podemos hacer los siguientes cálculos:

```
(%i3) T(x,y,z):=[x+y,x+z,2*x+y+z,y-z];
(%o3) T(x,y,z):=[x+y,x+z,2x+y+z,y-z]

(%i4) M:matrix(T(1,0,0),T(0,1,0),T(0,0,1));
(%o4) \left(\begin{array}{rrrr} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right)
```

La función **triangularize** devolverá una matriz equivalente, pero en la forma triangular superior de la matriz M, obtenida por eliminación gaussiana.

```
(%i5) triangularize(M);
(%o5) \left(\begin{array}{rrrr} 1 & 1 & 2 & 0 \\ 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)
```

Eso significa que una base para el rango $\mathbb{T}[\mathbb{R}^4]$ será: $\{(1, 1, 2, 0), (0, -1, -1, 1)\}$, es decir, la dimensión se éste subespacio vectorial es $\dim[\mathbb{T}[\mathbb{R}^4]] = 2$.

Por lo tanto:

$$\dim[\mathcal{N}(\mathbb{T})] + \dim[\mathbb{T}[\mathbb{R}^4]] = \dim[\mathbb{R}^3] \Rightarrow 1 + 2 = 3 .$$

Consideremos ahora el siguiente espacio vectorial $\mathbb{V}^3 = \{(x, y, z) : x + y + z = 0\}$ en \mathbb{R}^3 , y la siguiente transformación lineal de $\mathbb{V}^3 \rightarrow \mathbb{V}^3$

$$F(x, y, z) = (x + y + z, x + y - z, 2z) ,$$

```
(%i6) F(x,y,z):=[x+y+z,x+y-z,2*z];
(%o6) [x + y + z, x + y - z, 2z]
```

Ahora bien, para obtener el núcleo de la transformación resolvemos

```
(%i7) linsolve([x+y+z=0,x+y-z=0,2*z=0],[x,y,z]);
solve: dependent equations eliminated: (1)
(%o7) [x = %r1, y = -%r1, z = 0]
```

La solución es: $x = C$, $y = -C$ y $z = 0$, por lo tanto, una base para el espacio $\mathbb{N}[F]$ es: $\{(1, -1, 0)\}$.

Si ahora evaluamos la transformación F para la base canónica resulta lo siguiente:

```
(%i8) Img:matrix(F(1,0,0),F(0,1,0),F(0,0,1));
(%o8) \left(\begin{array}{ccc} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 2 \end{array}\right)
```

Reduciremos la matriz anterior pero esta vez con la función **echelon**, que al igual que **triangularize** devuelve la forma escalonada de la matriz M , obtenida por eliminación gaussiana, pero normalizando el primer elemento no nulo de cada fila.

```
(%i9) echelon(Img);
(%o9) \left(\begin{array}{ccc} 1 & 1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{array}\right)
```

Es decir, una base para el espacio imagen o rango es el conjunto: $\{(1, 1, 0), (0, 1, -1)\}$.

Podemos también calcular una base para $V^3 = \{(x, y, z) : x + y + z = 0\}$

```
(%i10) linsolve(x+y+z=0,[x,y,z]);
(%o10) [x = -%r3 - %r2, y = %r3, z = %r2]
```

Es una solución del tipo: $x = -C_3 - C_2$, $y = C_3$ y $z = C_2$. Una base vendrá dada por: $\{(-1, 1, 0), (-1, 0, 1)\}$.

Si evaluamos la transformación dada en esta base, resulta:

```
(%i11) matrix(F(-1,1,0),F(-1,0,1));
(%o11) \left(\begin{array}{ccc} 0 & 0 & 0 \\ 0 & -2 & 2 \end{array}\right)
```

El espacio vectorial tiene como base $\{(0, -2, 2)\}$.

4.2.10 Ejercicios

1. Considere los siguientes operadores: $\mathbb{A} = \mathbb{A}^\dagger$ hermítico, $\mathbb{K} = -\mathbb{K}^\dagger$ antihermítico; $\mathbb{U}^{-1} = \mathbb{U}^\dagger$ unitario, \mathbb{P} y \mathbb{Q} dos operadores genéricos. Pruebe las siguientes afirmaciones:

(a). En general:

$$\text{I. } (\mathbb{P}^\dagger)^{-1} = (\mathbb{P}^{-1})^\dagger.$$

- II. $(\mathbb{P}\mathbb{Q})^{-1} = \mathbb{Q}^{-1}\mathbb{P}^{-1}$
- III. Si $[\mathbb{P}, \mathbb{Q}] = 0$, entonces $\mathbb{P}(\mathbb{Q})^{-1} = (\mathbb{Q})^{-1}\mathbb{P}$
- IV. $(e^{\mathbb{P}})^\dagger = e^{\mathbb{P}^\dagger}$
- V. $\mathbb{P}e^{\mathbb{Q}\mathbb{P}^{-1}} = e^{\mathbb{P}\mathbb{Q}\mathbb{P}^{-1}}$
- (b). Si \mathbb{A} es hermítico entonces $\tilde{\mathbb{A}} = \mathbb{U}^{-1}\mathbb{A}\mathbb{U}$ también será un operador hermítico.
- (c). Si \mathbb{A} es hermítico entonces $e^{i\mathbb{A}}$ es unitario.
- (d). Si \mathbb{K} es antihermítico entonces $\tilde{\mathbb{K}} = \mathbb{U}^{-1}\mathbb{K}\mathbb{U}$ es también lo será. En particular eso se cumple para $\tilde{\mathbb{K}} = i\mathbb{A}$. Es decir, podemos construir un operador antihermítico a partir de uno hermítico.
- (e). Dados dos operadores \mathbb{A} y \mathbb{B} , hermíticos, su composición $\mathbb{A}\mathbb{B}$, será hermítica *si y sólo si* \mathbb{A} y \mathbb{B} communtan.
- (f). Si \mathbb{S} es un operador real y antisimétrico⁷ y \mathbb{I} el operador unidad, pruebe:
- I. Los operadores $(\mathbb{I} - \mathbb{S})$ y $(\mathbb{I} + \mathbb{S})$ commutan.
 - II. El operador $(\mathbb{I} - \mathbb{S})(\mathbb{I} + \mathbb{S})$ es simétrico, mientras que $(\mathbb{I} - \mathbb{S})(\mathbb{I} + \mathbb{S})^{-1}$ es ortogonal.⁸
- (g). Considere una matriz ortogonal de la forma $\mathbb{R} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$, encuentre la expresión para \mathbb{S} que reproduce $\mathbb{R} = (\mathbb{I} - \mathbb{S})(\mathbb{I} + \mathbb{S})^{-1}$.
2. Si un operador lineal \mathbb{C} genérico, entonces pruebe que $(\mathbb{C} + \mathbb{C}^\dagger)$ y $i(\mathbb{C} - \mathbb{C}^\dagger)$, con $i = \sqrt{-1}$, serán hermíticos. Esto, obviamente implica que siempre podremos separar un operador lineal como:
- $$\mathbb{C} = \frac{1}{2}(\mathbb{C} + \mathbb{C}^\dagger) + \frac{1}{2}(\mathbb{C} - \mathbb{C}^\dagger).$$
- donde $(\mathbb{C} + \mathbb{C}^\dagger)$ representa su parte hermítica y $(\mathbb{C} - \mathbb{C}^\dagger)$ su parte antihermítica.

4.3 Representación matricial de operadores

Dados dos vectores $|v_1\rangle$ y $|v_2\rangle$ definiremos como el elemento de matriz del operador \mathbb{A} al producto interno de dos vectores

$$\langle v_2 | (\mathbb{A} |v_1\rangle) \equiv A_{(|v_1\rangle, |v_2\rangle)}, \quad (4.13)$$

es claro que $A_{(|v_1\rangle, |v_2\rangle)}$ será en general un número complejo, pero además el valor de ese número dependerá de los vectores $|v_1\rangle$ y $|v_2\rangle$ con los cuales se haga la operación (4.13).

El paso más importante para asociar un conjunto de números a un operador lo constituye realizar la operación (4.13) con los elementos de una base del espacio vectorial donde opera \mathbb{A} .

Supongamos un operador lineal \mathbb{A} en el espacio vectorial de transformaciones lineales $\mathcal{L}(\mathbf{V}, \mathbf{W})$ donde $\dim(\mathbf{V}) = n$ y $\dim(\mathbf{W}) = m$, y sean $\{|e_1\rangle, |e_2\rangle, |e_3\rangle, \dots, |e_n\rangle\}$ y $\{|\tilde{e}_1\rangle, |\tilde{e}_2\rangle, |\tilde{e}_3\rangle, \dots, |\tilde{e}_m\rangle\}$ bases ortonormales⁹ para \mathbf{V} y \mathbf{W} respectivamente.

Entonces en la ecuación (4.13) cada uno de los vectores $\mathbb{A}|e_i\rangle \in \mathbf{W}$ nos conduce a:

$$\left\langle e^{\tilde{\beta}} \right| \mathbb{A} |e_i\rangle = A_i^\beta \quad \text{con } i = 1, 2, \dots, n \text{ y } \alpha, \beta = 1, 2, \dots, m. \quad (4.14)$$

Las cantidades A_j^β son la representación del operador \mathbb{A} respecto a las bases $\{|e_n\rangle\}$ y $\{|\tilde{e}_m\rangle\}$ de \mathbf{V} y \mathbf{W} respectivamente. Es decir, definiremos una matriz A_j^β como un arreglo de números donde el superíndice, β ,

⁷Esto es $\mathbb{S}^\dagger \equiv \mathbb{S}^T = -\mathbb{S}$ con \mathbb{S}^T el traspuesto de \mathbb{S} .

⁸Esto es $\mathbb{A}^T = \mathbb{A}^{-1}$ con \mathbb{A}^T el traspuesto de \mathbb{A} .

⁹Como hemos mencionado con anterioridad, las bases no necesariamente deben ser ortogonales (o mejor ortonormales), pero uno siempre puede ortogonalizarlas (y ortonormalizarlas)

indica fila y el subíndice, j , columna:

$$A_j^\beta = \begin{pmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & & A_n^2 \\ \vdots & \ddots & & \\ A_1^m & A_2^m & & A_n^m \end{pmatrix}, \quad \begin{pmatrix} A_1^1 \\ A_1^2 \\ \vdots \\ A_1^n \end{pmatrix}, \quad \left(\begin{array}{cccc} A_1^1 & A_2^1 & \cdots & A_n^1 \end{array} \right).$$

Es importante señalar que cambiando el orden de los vectores dentro de la base cambia la representación matricial del operador. Esto significa que la organización de los número A_j^β dependerá del orden que le demos a los vectores en las bases $\{|e_i\rangle\}$ y $\{|\tilde{e}_i\rangle\}$.

Definitivamente, las matrices son uno de los objetos más útiles de las matemáticas que permiten “aterrizar” conceptos y calcular cantidades. La palabra matriz fue introducida en 1850 por James Joseph Sylvester¹⁰ y su teoría desarrollada por Hamilton¹¹ y Cayley¹². Si bien los físicos las consideramos indispensables, no fueron utilizadas de manera intensiva hasta el aparición de la Mecánica Cuántica alrededor de 1925.

4.3.1 Álgebra elemental de matrices

El álgebra de operadores que definimos en la sección 4.1.1, puede ser traducida al lenguaje de matrices. Por comodidad supondremos un único espacio vectorial, $V \equiv W$ y, por lo tanto nos basta una base ortogonal $\{|e_n\rangle\}$. De este modo, es claro que se obtienen nuevamente las conocidas relaciones para matrices cuadradas

$$\langle e^i | \mathbb{A} + \mathbb{B} | e_j \rangle = \langle e^i | \mathbb{A} | e_j \rangle + \langle e^i | \mathbb{B} | e_j \rangle = A_j^i + B_j^i.$$

Con lo cual tenemos la suma de matrices que todos hemos visto en los cursos básicos

$$\begin{pmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & & A_n^2 \\ \vdots & \ddots & & \\ A_1^n & A_2^n & & A_n^n \end{pmatrix} + \begin{pmatrix} B_1^1 & B_2^1 & \cdots & B_n^1 \\ B_1^2 & B_2^2 & & B_n^2 \\ \vdots & \vdots & & \vdots \\ B_1^n & B_2^n & & B_n^n \end{pmatrix} = \begin{pmatrix} A_1^1 + B_1^1 & A_2^1 + B_2^1 & \cdots & A_n^1 + B_n^1 \\ A_1^2 + B_1^2 & A_2^2 + B_2^2 & & A_n^2 + B_n^2 \\ \vdots & \vdots & & \ddots \\ A_1^n + B_1^n & & & A_n^n + B_n^n \end{pmatrix}.$$

De igual modo, para la representación de composición de operadores que consideramos en la sección 4.1.2 tendremos:

$$\langle e^i | \mathbb{A} \mathbb{B} | e_j \rangle = \langle e^i | \mathbb{A} | \mathbb{B} | e_j \rangle = \langle e^i | \mathbb{A} \left(|e_k\rangle \langle e^k| \right) \mathbb{B} | e_j \rangle = \langle e^i | \mathbb{A} | e_k \rangle \langle e^k | \mathbb{B} | e_j \rangle = A_k^i B_j^k,$$

que se traduce en la tradicional multiplicación de matrices:

$$\begin{pmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & & A_n^2 \\ \vdots & \ddots & & \\ A_1^n & A_2^n & & A_n^n \end{pmatrix} \times \begin{pmatrix} B_1^1 & B_2^1 & \cdots & B_n^1 \\ B_1^2 & B_2^2 & & B_n^2 \\ \vdots & \vdots & & \vdots \\ B_1^n & B_2^n & & B_n^n \end{pmatrix} = \begin{pmatrix} A_k^1 B_1^k & A_k^1 B_2^k & \cdots & A_k^1 B_n^k \\ A_k^2 B_1^k & A_k^2 B_2^k & & A_k^2 B_n^k \\ \vdots & \vdots & & \ddots \\ A_k^n B_1^k & & & A_k^n B_n^k \end{pmatrix},$$

como ya sabíamos $\mathbb{A} \mathbb{B} \neq \mathbb{B} \mathbb{A} \rightarrow A_k^i B_j^k \neq B_k^i A_j^k$.

¹⁰James Joseph Sylvester (1814-1897 Londres, Inglaterra) Además de sus aportes con Cayley a la teoría de las matrices, descubrió la solución a la ecuación cúbica y fue el primero en utilizar el término discriminante para categorizar cada una de las raíces de la ecuación. Para vivir tuvo que ejercer de abogado durante una década. Por fortuna, otro matemático de la época (Arthur Cayley) frecuentaba los mismos juzgados y tribunales y pudieron interactuar. Por ser judío tuvo cantidad de dificultades para conseguir trabajo en la Academia.

¹¹Sir William Rowan Hamilton (1805 - 1865, Dublin, Irlanda) Sus contribuciones en el campo de la óptica, dinámica del cuerpo rígido, teoría de ecuaciones algebraicas y teoría de operadores lineales.

¹²Arthur Cayley (1821, Richmond, 1895, Cambridge, Inglaterra) En sus cerca de 900 trabajos cubrió casi la totalidad de las áreas de las matemáticas de aquel entonces. Sus mayores contribuciones se centran en la teoría de matrices y la geometría no euclídea. No consiguió empleo como matemático y tuvo que graduarse de abogado para ejercer durante más de 15 años, durante los cuales publicó más de 250 trabajos en matemáticas.

Finalmente, es claro que la multiplicación de un número por una matriz es la multiplicación de todos sus elementos por ese número

$$\langle e^i | \alpha \mathbb{A} | e_j \rangle = \alpha \langle e^i | \mathbb{A} | e_j \rangle = \alpha A_j^i.$$

4.3.2 Bases y la representación matricial de operadores

Como es claro de la ecuación (4.14) la representación matricial de un operador depende de las bases de \mathbf{V} y \mathbf{W} , respectivamente. A continuación discutiremos algunos ejemplos particulares.

4.3.2.1 Representación diagonal

Dado un operador lineal $\mathbb{A} \in \mathcal{L}(\mathbf{V}, \mathbf{W})$, donde $\dim(\mathbf{V}) = \dim(\mathbf{W}) = n$, y sea $\{|e_i\rangle\}$ una base ortonormal para \mathbf{V} y \mathbf{W} . Entonces, si adicionalmente se da el caso que $\mathbb{A}|e_i\rangle = \lambda_i|e_i\rangle$, la representación matricial será diagonal

$$\langle e^j | \mathbb{A} | e_i \rangle = A_i^j = \lambda_i \langle e^j | e_i \rangle = \lambda_i \delta_i^j.$$

Es importante señalar que $\lambda_i \delta_i^j$, en la ecuación anterior, no se puede interpretar como suma, sino sencillamente que se repite para cada índice i . Adicionalmente, que esta afirmación también es válida para $\dim(\mathbf{V}) \neq \dim(\mathbf{W})$ pero por simplicidad seguimos trabajando con matrices cuadradas.

4.3.2.2 Representación matricial de operadores adjuntos y hermíticos

Consideremos la representación matricial de un operador adjunto, tal y como lo desarrollamos en la sección 4.2.3,

$$(A^\dagger)_j^i = \langle e^i | \mathbb{A}^\dagger | e_j \rangle = \langle e^j | \mathbb{A} | e_i \rangle^* = (A_i^j)^*,$$

vale decir: la matriz que representa el operador adjunto \mathbb{A}^\dagger , es la traspuesta conjugada de la matriz que representa al operador \mathbb{A} .

Si el operador es hermítico: $\mathbb{A}^\dagger = \mathbb{A} \Rightarrow (A^\dagger)_j^i = A_j^i$. Por lo tanto, las matrices hermíticas son simétricas respecto a la diagonal y los elementos de la diagonal son números reales.

Aquí vale la pena reexpresar, con matrices, algunas de las propiedades que arriba expusimos:

$$\begin{aligned} (\mathbb{A}^\dagger)^\dagger &\rightarrow \left(\langle e^i | \mathbb{A}^\dagger | e_j \rangle \right)^\dagger = \left((A^\dagger)_j^i \right)^\dagger = \left((A_i^j)^* \right)^\dagger = A_j^i \quad \text{y} \\ (\lambda \mathbb{A})^\dagger &\rightarrow \langle e^i | \lambda \mathbb{A}^\dagger | e_j \rangle = \langle e^j | \lambda \mathbb{A} | e_i \rangle^* = \lambda^* \langle e^j | \mathbb{A} | e_i \rangle^* = \lambda^* \langle e^i | \mathbb{A}^\dagger | e_j \rangle = \lambda^* \mathbb{A}^\dagger, \end{aligned}$$

pero más interesante es:

$$(\mathbb{A}\mathbb{B})^\dagger \rightarrow \langle e^i | (\mathbb{A}\mathbb{B})^\dagger | e_j \rangle = (A_k^i B_j^k)^\dagger = (A_k^i)^*(B_j^k)^* = (A^\dagger)_j^k (B^\dagger)_k^i = (B^\dagger)_k^i (A^\dagger)_j^k = \mathbb{B}^\dagger \mathbb{A}^\dagger.$$

4.3.3 Representación matricial y transformaciones

Hemos dicho que dada una base particular en el espacio de acción de un operador, éste quedará representado por una matriz adaptada a esa base. Por lo tanto, si cambiamos la base que genera la representación, ese mismo operador tendrá otra matriz como representación.

Por comodidad volvamos a suponer un único espacio vectorial, $\mathbf{V} \equiv \mathbf{W}$, pero ahora supondremos que este espacio vectorial \mathbf{V} tiene dos bases discretas ortonormales $\{|e_i\rangle\}$ y $\{|\tilde{e}_i\rangle\}$. Entonces las representaciones

matriciales de \mathbb{A} : $\tilde{A}_j^i = \langle \tilde{e}^i | \mathbb{A} | \tilde{e}_j \rangle$ y $A_m^i = \langle e^k | \mathbb{A} | e_m \rangle$, están relacionadas por:

$$\langle \tilde{e}^i | \mathbb{A} | \tilde{e}_j \rangle = \langle \tilde{e}^i | \left(|e_k\rangle \langle e^k| \right) \mathbb{A} (|e_m\rangle \langle e^m|) | \tilde{e}_j \rangle = \underbrace{\langle \tilde{e}^i | e_k \rangle}_{S_k^i} \underbrace{\langle e^k | \mathbb{A} | e_m \rangle}_{\tilde{S}_j^m} \underbrace{\langle e^m | \tilde{e}_j \rangle}_{\tilde{S}_j^m} \Leftrightarrow \tilde{A}_j^i = S_k^i A_m^k \tilde{S}_j^m , \quad (4.15)$$

donde \tilde{A}_j^i es la representación del operador \mathbb{A} en la base $\{|\tilde{e}_j\rangle\}$ y A_m^k en la base $\{|e_m\rangle\}$.

Más aún, siempre podremos expresar unos vectores base en términos de los otros de tal forma que:

$$|\tilde{e}_j\rangle = \tilde{S}_j^m |e_m\rangle = \tilde{S}_j^m (S_m^n |\tilde{e}_n\rangle) \Rightarrow \langle \tilde{e}^n | \tilde{e}_j \rangle = \delta_j^n = \tilde{S}_j^m S_m^n \equiv S_m^n \tilde{S}_j^m \Rightarrow \tilde{S}_j^i = (S_j^i)^{-1} , \quad (4.16)$$

con lo cual la relación (4.15), $\tilde{A}_j^i = S_k^i A_m^k \tilde{S}_j^m$, puede ser reescrita como

$$\tilde{A}_j^i = S_k^i A_m^k (S_j^m)^{-1} \Leftrightarrow \tilde{\mathbb{A}} = \mathbb{S} \mathbb{A} \mathbb{S}^{-1} \Rightarrow \mathbb{A} = \mathbb{S}^{-1} \tilde{\mathbb{A}} \mathbb{S}. \quad (4.17)$$

Diremos que dos representaciones matriciales A_j^i y \tilde{A}_m^k , de un mismo operador \mathbb{A} , son similares si están relacionadas entre si por (4.17), donde la matriz de transformación S_k^i y su inversa se construyen a partir de los productos internos de las bases. Esta misma afirmación se puede refrasear por el hecho que dos operadores $\tilde{\mathbb{A}}$ y \mathbb{A} están relacionados por una transformación de similaridad (4.17) constituyen distintas representaciones de un mismo operador.

Adicionalmente, por la definición de producto interno, $\langle e^k | \tilde{e}_m \rangle = \langle \tilde{e}^m | e_k \rangle^*$ y tomando en cuenta (4.16), tendremos:

$$\tilde{S}_m^k = (S_k^m)^* \equiv (S^\dagger)_m^k = (S_m^k)^{-1} ,$$

es decir, las matrices de productos internos que transforman las representaciones matriciales de los operadores, son unitarias y la relación (4.15) puede escribirse también como:

$$\tilde{A}_j^i = S_k^i A_m^k \tilde{S}_j^m \Leftrightarrow \tilde{A}_j^i = S_k^i A_m^k (S^\dagger)_j^m . \quad (4.18)$$

En este caso, diremos que las representaciones matriciales de \mathbb{A} están relacionadas a una transformación unitaria del tipo (4.18).

4.3.4 Traza de operadores

La traza, $\text{Tr}(\mathbb{A})$, de un operador \mathbb{A} es la suma de los elementos diagonales de su representación matricial \mathbb{A} . Esto es, dado un operador \mathbb{A} y una base ortogonal $\{|e_i\rangle\}$ para V^n , entonces:

$$\text{Tr}(\mathbb{A}) = \langle e^k | \mathbb{A} | e_k \rangle = A_k^k .$$

La traza de la representación matricial de un operador es, por ejemplo:

$$A_j^i = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \Rightarrow \text{Tr}(\mathbb{A}) = A_i^i = 15 .$$

4.3.4.1 Propiedades de la Traza

Claramente la traza es lineal: $\text{Tr}(\mathbb{A} + \lambda \mathbb{B}) = \text{Tr}(\mathbb{A}) + \lambda \text{Tr}(\mathbb{B})$, ya que

$$\text{Tr}(\mathbb{A} + \lambda \mathbb{B}) = \langle e^k | \mathbb{A} + \lambda \mathbb{B} | e_k \rangle = \langle e^k | \mathbb{A} | e_k \rangle + \lambda \langle e^k | \mathbb{B} | e_k \rangle = \text{Tr}(\mathbb{A}) + \lambda \text{Tr}(\mathbb{B}) .$$

La traza de un producto commuta, esto es, $\text{Tr}(\mathbb{A}\mathbb{B}) = \text{Tr}(\mathbb{B}\mathbb{A})$, y es fácilmente demostrable:

$$\text{Tr}(\mathbb{A}\mathbb{B}) = \langle e^k | \mathbb{A}\mathbb{B} | e_k \rangle = \langle e^k | \mathbb{A} | e_m \rangle \underbrace{\langle e^m | \mathbb{B} | e_k \rangle}_{\mathbb{I}} = \langle e^k | \mathbb{B} | e_m \rangle \underbrace{\langle e^m | \mathbb{A} | e_k \rangle}_{\mathbb{I}} = \text{Tr}(\mathbb{B}\mathbb{A}) .$$

Recuerde que $\langle e^k | \mathbb{B} | e_m \rangle$ y $\langle e^k | \mathbb{A} | e_k \rangle$ son números que pueden ser reordenados. Donde una vez más hemos utilizado las dos relaciones de cierre $|e_m\rangle\langle\tilde{e}^m| = |e_k\rangle\langle e^k| = \mathbb{I}$.

Del mismo modo es fácil demostrar que la traza de un triple producto de matrices respeta la ciclicidad del orden de las matrices en el producto: $\text{Tr}(\mathbb{A}\mathbb{B}\mathbb{C}) = \text{Tr}(\mathbb{B}\mathbb{C}\mathbb{A}) = \text{Tr}(\mathbb{C}\mathbb{A}\mathbb{B})$.

4.3.4.2 Invariancia de la Traza

La traza de una matriz no depende de la base que seleccionemos, es un invariante que caracteriza al operador independientemente de la base en la cual se represente. Entonces, a partir de (4.15) tendremos:

$$A_k^k = \langle e^k | \mathbb{A} | e_k \rangle = \underbrace{\langle e^k |}_{\mathbb{I}} \underbrace{\langle \tilde{e}_m \rangle}_{\mathbb{I}} \langle \tilde{e}^m | \mathbb{A} | e_k \rangle = \langle \tilde{e}^m | \mathbb{A} | e_k \rangle \underbrace{\langle e^k |}_{\mathbb{I}} \langle \tilde{e}_m \rangle = \langle \tilde{e}^m | \mathbb{A} | \tilde{e}_m \rangle = \tilde{A}_m^m.$$

Nótese que primero hemos introducido $|\tilde{e}_m\rangle\langle\tilde{e}^m|$, luego hemos reubicado el término $\langle e^k | \tilde{e}_m \rangle$, para el lado derecho y, finalmente, hemos identificado el término $|e_k\rangle\langle e^k|$ como el operador unidad. Es claro que la traza caracteriza al operador independiente de su representación matricial.

4.3.5 Un paréntesis determinante

El determinante de un operador, \mathbb{A} , se define como una aplicación de su representación matricial, $\langle e^i | \mathbb{A} | e_j \rangle$ en \mathbb{R} , es decir, asocia un número real con la representación matricial del operador, y se define como:

$$\det |\mathbb{A}| = \varepsilon^{ijk\dots} A_i^1 A_j^2 A_k^3 \dots \equiv \begin{vmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & & A_n^2 \\ \vdots & & \ddots & \\ A_1^n & A_2^n & & A_n^n \end{vmatrix},$$

donde los elementos de la representación matricial se expresan como $A_j^i = \langle e^i | \mathbb{A} | e_j \rangle$ y hemos generalizado el tensor de Levi-Civita que presentamos en la sección 1.4.1, de tal forma que:

$$\varepsilon^{ijk\dots} = \varepsilon_{ijk\dots} = \begin{cases} 0, & \text{si cualesquiera dos índices son iguales} \\ 1, & \text{si los índices } i, j, k \dots \text{ constituyen una permutación cíclica de } 1, 2, 3 \dots n \\ -1, & \text{si los índices } i, j, k \dots \text{ constituyen una permutación antícíclica de } 2, 1, 3 \dots n \end{cases}$$

Un ejemplo simplificado de esta asociación para el caso de representaciones matriciales 3×3 es el siguiente:

$$\langle e^i | \mathbb{A} | e_j \rangle = \begin{pmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{pmatrix} \Rightarrow \det |\mathbb{A}| = \varepsilon^{ijk} A_i^1 A_j^2 A_k^3 = \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix},$$

con lo cual

$$\begin{aligned} \det |\mathbb{A}| &= \varepsilon^{123} A_1^1 A_2^2 A_3^3 + \varepsilon^{312} A_3^1 A_1^2 A_2^3 + \varepsilon^{231} A_2^1 A_3^2 A_1^3 + \varepsilon^{132} A_1^1 A_3^2 A_2^3 + \varepsilon^{321} A_3^1 A_2^2 A_1^3 + \varepsilon^{213} A_2^1 A_1^2 A_3^3 \\ &= A_1^1 A_2^2 A_3^3 + A_3^1 A_1^2 A_2^3 + A_2^1 A_3^2 A_1^3 - A_1^1 A_3^2 A_2^3 - A_3^1 A_2^2 A_1^3 - A_2^1 A_1^2 A_3^3. \end{aligned}$$

4.3.5.1 Propiedades de los determinantes

1. $\det |\mathbb{A}| = \det |\mathbb{A}^T|$, donde \mathbb{A}^T es el operador traspuesto de \mathbb{A} , esto es claro para las representaciones matriciales $\langle e^i | \mathbb{A} | e_j \rangle = A_j^i = (\langle e^j | \mathbb{A} | e_i \rangle)^T$.

Esta propiedad proviene de la definición del índice de Levi-Civita: $\det |\mathbb{A}| = \varepsilon^{ijk\dots} A_i^1 A_j^2 A_k^3 \dots = \varepsilon_{ijk\dots} A_1^i A_2^j A_3^k \dots = \det |\mathbb{A}^T|$, que se traduce en que si se intercambian filas por columnas el determinante

no se altera.

$$\begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix} = \begin{vmatrix} A_1^1 & A_1^2 & A_1^3 \\ A_2^1 & A_2^2 & A_2^3 \\ A_3^1 & A_3^2 & A_3^3 \end{vmatrix}.$$

2. Si dos filas o dos columnas son idénticas el determinante se anula

$$\varepsilon^{iik\cdots} A_i^1 A_i^2 A_k^3 \cdots = \varepsilon_{iik\cdots} A_1^i A_2^i A_3^k \cdots = 0, \Leftrightarrow \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^1 & A_2^1 & A_3^1 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix} = 0.$$

3. Si multiplicamos una fila o una columna por un número, el determinante queda multiplicado por el número

$$\begin{aligned} \varepsilon^{ijk\cdots} A_i^1 (\lambda A_j^2) A_k^3 \cdots &= \lambda \varepsilon^{ijk\cdots} A_i^1 A_j^2 A_k^3 \cdots = \lambda \det |\mathbb{A}| \\ \varepsilon_{ijk\cdots} A_1^i A_2^j (\lambda A_3^k) \cdots &= \lambda \varepsilon_{ijk\cdots} A_1^i A_2^j A_3^k \cdots = \lambda \det |\mathbb{A}|, \end{aligned}$$

de aquí claramente se desprende que si una fila o una columna es cero ($\lambda = 0$) el determinante se anula.

Más aún, si dos filas o dos columnas son proporcionales $A_i^1 = \lambda A_j^1$ el determinante se anula, por cuanto se cumple la propiedad anterior:

$$\begin{vmatrix} A_1^1 & \lambda A_2^1 & A_3^1 \\ A_1^2 & \lambda A_2^2 & A_3^2 \\ A_1^3 & \lambda A_2^3 & A_3^3 \end{vmatrix} = \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ \lambda A_1^3 & \lambda A_2^3 & \lambda A_3^3 \end{vmatrix} = \lambda \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix}.$$

Resulta obvio que:

$$\begin{vmatrix} A_1^1 & 0 & A_3^1 \\ A_1^2 & 0 & A_3^2 \\ A_1^3 & 0 & A_3^3 \end{vmatrix} = \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ 0 & 0 & 0 \end{vmatrix} = 0,$$

al igual que:

$$\begin{vmatrix} A_1^1 & \lambda A_1^1 & A_3^1 \\ A_1^2 & \lambda A_1^2 & A_3^2 \\ A_1^3 & \lambda A_1^3 & A_3^3 \end{vmatrix} = \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ \lambda A_1^1 & \lambda A_2^1 & \lambda A_3^1 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix} = \lambda \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix} = \lambda \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^1 & A_2^1 & A_3^1 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix} = 0.$$

4. Si se intercambian dos filas o dos columnas cambia de signo el determinante.

$$\det |\mathbb{A}| = \varepsilon^{ijk\cdots} A_i^1 A_j^2 A_k^3 \cdots = \begin{vmatrix} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & & A_n^2 \\ \vdots & & \ddots & \\ A_1^n & A_2^n & & A_n^n \end{vmatrix} \Rightarrow \varepsilon^{ijk\cdots} A_j^1 A_i^2 A_k^3 \cdots = -\det |\langle e^i | \bar{\mathbb{A}} | e_j \rangle|,$$

donde en la representación matricial $\langle e^i | \bar{\mathbb{A}} | e_j \rangle$ se han intercambiando un par de columnas. Claramente las propiedades del índice de Levi-Civita, obliga al cambio de signo $\det |\tilde{\mathbb{A}}| = -\det |\mathbb{A}|$.

Nótese que las propiedades anteriores nos lleva a reescribir el determinante de un operador de la siguiente manera:

$$\det |\mathbb{A}| = \varepsilon^{\alpha\beta\gamma\cdots} \det |\mathbb{A}| = \varepsilon_{ijk\cdots} A_\alpha^i A_\beta^j A_\gamma^k \cdots \Leftrightarrow \det |\mathbb{A}| = \varepsilon_{\alpha\beta\gamma\cdots} \det |\mathbb{A}| = \varepsilon^{ijk\cdots} A_i^\alpha A_j^\beta A_k^\gamma \cdots$$

claramente, si $\alpha\beta\gamma\cdots \Leftrightarrow 123\cdots$ obtenemos nuevamente la definición anterior. Si se intercambian dos filas o dos columnas, el determinante cambia de signo debido al intercambio de dos índices griegos.

Si dos filas o dos columnas son iguales el determinante se anula debido a la propiedad de símbolo de Levi-Civita con índices griegos.

5. El determinante de la composición de operadores es el producto de los determinantes

$$\det |\mathbb{A}\mathbb{B}| = \det |\mathbb{A}| \det |\mathbb{B}|.$$

Antes de proceder a la demostración de esta importante propiedad jugaremos un poco más con las propiedades de las matrices. Si \mathbb{A} es un operador con una representación matricial $m \times n$ y \mathbb{B} es uno con una representación matricial $n \times p$, entonces tendremos $(\mathbb{A}\mathbb{B})^\alpha = \mathbb{A}^\alpha \mathbb{B}$, esto es, la α -ésima fila de $\mathbb{A}\mathbb{B}$ es igual a la multiplicación de la α -ésima fila de \mathbb{A} , por toda la matriz \mathbb{B} .

Veamos: $C_j^\alpha = (\mathbb{A}\mathbb{B})_j^\alpha = A_l^\alpha B_j^l$, por lo tanto la α -ésima fila:

$$C_j^\alpha = A_l^\alpha B_j^l \Rightarrow C_j^\alpha = (A_1^\alpha, A_2^\alpha, A_3^\alpha, \dots, A_n^\alpha) \begin{pmatrix} B_1^1 & B_2^1 & \cdots & B_n^1 \\ B_1^2 & B_2^2 & & B_n^2 \\ \vdots & & \ddots & \\ B_1^n & B_2^n & & B_n^n \end{pmatrix}.$$

Tendremos entonces:

$$\det |\mathbb{A}| \det |\mathbb{B}| = \det |\mathbb{A}| \left(\varepsilon_{ijk} \dots B_1^i B_2^j B_3^k \dots \right) = \left(\varepsilon_{ijk} \dots A_\alpha^i A_\beta^j A_\gamma^k \dots \right) \left(\varepsilon_{abc} \dots B_1^a B_2^b B_3^c \dots \right),$$

que siempre puede ser rearreglado como:

$$\left(\varepsilon^{ijk} \dots A_i^\alpha A_j^\beta A_k^\gamma \dots \right) \left(\varepsilon_{ijk} \dots B_1^i B_2^j B_3^k \dots \right) = A_i^\alpha B_1^i A_j^\beta B_2^j A_k^\gamma B_3^k \dots = \det |\mathbb{A}\mathbb{B}|.$$

Veamos este desarrollo para el caso de matrices 3×3 :

$$\begin{aligned} \det |\mathbb{A}| \det |\mathbb{B}| &= (\varepsilon^{123} A_1^1 A_2^2 A_3^3 + \varepsilon^{312} A_3^1 A_1^2 A_2^3 + \varepsilon^{231} A_2^1 A_3^2 A_1^3 + \varepsilon^{132} A_1^1 A_3^2 A_2^3 \\ &+ \varepsilon^{321} A_3^1 A_2^2 A_1^3 + \varepsilon^{213} A_2^1 A_1^2 A_3^3) \cdot (\varepsilon^{123} B_1^1 B_2^2 B_3^3 + \varepsilon^{312} B_3^1 B_1^2 B_2^3 + \varepsilon^{231} B_2^1 B_3^2 B_1^3 \\ &+ \varepsilon^{132} B_1^1 B_3^2 B_2^3 + \varepsilon^{321} B_3^1 B_2^2 B_1^3 + \varepsilon^{213} B_2^1 B_1^2 B_3^3), \end{aligned}$$

con lo cual:

$$\begin{aligned} &= A_1^1 A_2^2 A_3^3 (B_1^1 B_2^2 B_3^3 + B_3^1 B_1^2 B_2^3 + B_2^1 B_3^2 B_1^3 - B_1^1 B_3^2 B_2^3 - B_3^1 B_2^2 B_1^3 - B_2^1 B_1^2 B_3^3) \\ &+ A_3^1 A_1^2 A_2^3 (B_1^1 B_2^2 B_3^3 + B_3^1 B_1^2 B_2^3 + B_2^1 B_3^2 B_1^3 + B_1^1 B_3^2 B_2^3 + B_3^1 B_2^2 B_1^3 + B_2^1 B_1^2 B_3^3) \\ &+ A_2^1 A_3^2 A_1^3 (B_1^1 B_2^2 B_3^3 + B_3^1 B_1^2 B_2^3 + B_2^1 B_3^2 B_1^3 + B_1^1 B_3^2 B_2^3 + B_3^1 B_2^2 B_1^3 + B_2^1 B_1^2 B_3^3) \\ &- A_1^1 A_3^2 A_2^3 (B_1^1 B_2^2 B_3^3 + B_3^1 B_1^2 B_2^3 + B_2^1 B_3^2 B_1^3 + B_1^1 B_3^2 B_2^3 + B_3^1 B_2^2 B_1^3 + B_2^1 B_1^2 B_3^3) \\ &- A_3^1 A_2^2 A_1^3 (B_1^1 B_2^2 B_3^3 + B_3^1 B_1^2 B_2^3 + B_2^1 B_3^2 B_1^3 + B_1^1 B_3^2 B_2^3 + B_3^1 B_2^2 B_1^3 + B_2^1 B_1^2 B_3^3) \\ &- A_2^1 A_1^2 A_3^3 (B_1^1 B_2^2 B_3^3 + B_3^1 B_1^2 B_2^3 + B_2^1 B_3^2 B_1^3 + B_1^1 B_3^2 B_2^3 + B_3^1 B_2^2 B_1^3 + B_2^1 B_1^2 B_3^3). \end{aligned}$$

Como son números se pueden arreglar de la siguiente forma:

$$\begin{aligned} &= A_1^1 A_2^2 A_3^3 (B_1^1 B_2^2 B_3^3 + B_1^2 B_2^3 B_3^1 + B_3^1 B_2^1 B_3^2 - B_1^1 B_3^2 B_2^3 - B_3^1 B_2^2 B_1^3 - B_2^1 B_1^2 B_3^3) \\ &+ A_3^1 A_1^2 A_2^3 (B_1^1 B_2^2 B_3^3 + B_1^2 B_3^1 B_2^3 + B_3^1 B_2^1 B_2^3 - B_1^1 B_3^2 B_2^3 - B_3^1 B_2^2 B_1^3 - B_2^1 B_1^2 B_3^3) \\ &+ A_2^1 A_3^2 A_1^3 (B_1^1 B_2^2 B_3^3 + B_1^2 B_3^1 B_3^2 + B_3^1 B_2^1 B_3^2 - B_1^1 B_3^2 B_3^2 - B_3^1 B_2^2 B_3^1 - B_2^1 B_1^2 B_3^3) \\ &- A_1^1 A_3^2 A_2^3 (B_1^1 B_2^2 B_3^3 + B_1^2 B_3^1 B_3^2 + B_3^1 B_2^1 B_3^2 - B_1^1 B_3^2 B_3^2 - B_3^1 B_2^2 B_3^1 - B_2^1 B_1^2 B_3^3) \\ &- A_3^1 A_2^2 A_1^3 (B_1^1 B_2^2 B_3^3 + B_1^2 B_3^1 B_3^2 + B_3^1 B_2^1 B_3^2 - B_1^1 B_3^2 B_3^2 - B_3^1 B_2^2 B_3^1 - B_2^1 B_1^2 B_3^3) \\ &- A_2^1 A_1^2 A_3^3 (B_1^1 B_2^2 B_3^3 + B_1^2 B_3^1 B_3^2 + B_3^1 B_2^1 B_3^2 - B_1^1 B_3^2 B_3^2 - B_3^1 B_2^2 B_3^1 - B_2^1 B_1^2 B_3^3) \\ &= \varepsilon_{ijk} A_\alpha^i B_1^\alpha A_\lambda^j B_2^\lambda A_\mu^k B_3^\mu. \end{aligned}$$

Por lo tanto:

$$\varepsilon_{ijk} A_\alpha^i B_1^\alpha A_\lambda^j B_2^\lambda A_\mu^k B_3^\mu = \det |\mathbb{A}| \det |\mathbb{B}| = \det |\mathbb{A}\mathbb{B}|. \quad (4.19)$$

4.3.5.2 La invariancia del determinante

El resultado expresado en la ecuación (4.19) tiene varias consecuencias importantes:

1. El determinante del operador inverso es el inverso del determinante, es decir:

$$\det |\mathbb{A}^{-1}| = \frac{1}{\det |\mathbb{A}|} = \det |\mathbb{A}|^{-1}.$$

Esta afirmación es fácilmente demostrable

$$\mathbb{I} = \mathbb{A}^{-1}\mathbb{A} \Rightarrow \det |\mathbb{I}| = \det |\mathbb{A}\mathbb{A}^{-1}| = \det |\mathbb{A}| \det |\mathbb{A}^{-1}| = 1 \Rightarrow \det |\mathbb{A}^{-1}| = \frac{1}{\det |\mathbb{A}|}.$$

2. La segunda consecuencia es todavía más importante: el determinante de un operador no depende de su representación matricial. Es decir, el determinante, al igual que la traza, asocia un número real al operador, independientemente de su representación matricial:

$$\det |\mathbb{A}| \equiv \det |\langle e^i | \mathbb{A} | e_j \rangle| = \det |\langle \tilde{e}^i | \mathbb{A} | \tilde{e}_j \rangle| \equiv \det |\tilde{\mathbb{A}}|.$$

Es inmediato darse que cuenta que al aplicar la función determinante a la ecuación (4.17) se tiene:

$$\det |\tilde{\mathbb{A}}_j^i| = \det |S_k^i A_m^k (S_j^m)^{-1}| \equiv \det |S_k^i| \det |A_m^k| \det |(S_j^m)^{-1}| = \det |S_k^i| \det |A_m^k| \frac{1}{\det |S_j^m|} = \det |A_m^k|.$$

Con lo cual queda demostrada la invariancia del determinante de representaciones matriciales de operadores en bases ortonormales.

4.3.5.3 Fórmula de Laplace

La fórmula de Laplace permite expresar el determinante de una matriz en términos de sus matrices menores o cofactores, matrices que definiremos en la siguiente sección.

$$\det |\mathbb{A}| = \sum_{j=1}^n A_j^i (A^c)_j^i, \quad \text{para cualquier } i. \quad (4.20)$$

4.3.6 Diferenciación de operadores y representación matricial

Retomemos, ahora en el lenguaje de representaciones matriciales de operadores, lo que desarrollamos en la sección 4.2.7. Dado un operador $\mathbb{A}(t)$, el cual supondremos dependiente de una variable arbitraria t , podremos definir la derivada como:

$$\frac{d\mathbb{A}(t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{A}(t + \Delta t) - \mathbb{A}(t)}{\Delta t},$$

por lo tanto, si $\langle u^k | \mathbb{A} | u_i \rangle = A_i^k$ entonces:

$$\left\langle u^k \left| \frac{d\mathbb{A}(t)}{dt} \right| u_i \right\rangle = \left(\frac{d\mathbb{A}(t)}{dt} \right)_i^k = \frac{d}{dt} \left\langle u^k \left| \mathbb{A}(t) \right| u_i \right\rangle = \frac{dA_i^k}{dt} = \begin{pmatrix} \frac{dA_1^1}{dt} & \frac{dA_2^1}{dt} & \dots & \frac{dA_n^1}{dt} \\ \frac{dA_1^2}{dt} & \frac{dA_2^2}{dt} & & \frac{dA_n^2}{dt} \\ \vdots & & \ddots & \\ \frac{dA_1^n}{dt} & \frac{dA_2^n}{dt} & & \frac{dA_n^n}{dt} \end{pmatrix}.$$

La regla es simple, la representación matricial de la derivada de un operador será la derivada de cada uno de sus elementos. Por ejemplo:

$$\frac{d}{dx} \begin{pmatrix} x & x^2 & 2 \\ 1 & e^{-x} & 5x \\ 3x^3 & 3 & \cos(x) \end{pmatrix} = \begin{pmatrix} 1 & 2x & 0 \\ 0 & -e^{-x} & 5 \\ 9x^2 & 0 & -\operatorname{sen}(x) \end{pmatrix}.$$

De igual forma tendremos las reglas usuales de la diferenciación y por lo tanto se cumplirá

$$\frac{d(\mathbb{A}(t) + \mathbb{B}(t))}{dt} = \frac{d(\mathbb{A}(t))}{dt} + \frac{d(\mathbb{B}(t))}{dt},$$

es decir:

$$\begin{aligned} \left\langle u^k \middle| \frac{d(\mathbb{A}(t) + \mathbb{B}(t))}{dt} \right| u_i \rangle &= \frac{d}{dt} \left\langle u^k \middle| (\mathbb{A}(t) + \mathbb{B}(t)) \right| u_i \rangle = \frac{d}{dt} \left(\left\langle u^k \middle| \mathbb{A}(t) \right| u_i \rangle + \left\langle u^k \middle| \mathbb{B}(t) \right| u_i \rangle \right) \\ &= \frac{d}{dt} \left\langle u^k \middle| \mathbb{A}(t) \right| u_i \rangle + \frac{d}{dt} \left\langle u^k \middle| \mathbb{B}(t) \right| u_i \rangle \\ &= \left\langle u^k \middle| \frac{d\mathbb{A}(t)}{dt} \right| u_i \rangle + \left\langle u^k \middle| \frac{d\mathbb{B}(t)}{dt} \right| u_i \rangle = \frac{d(\mathbb{A}(t))}{dt} + \frac{d(\mathbb{B}(t))}{dt}. \end{aligned}$$

Del mismo modo se cumplirá:

$$\frac{d(\mathbb{A}(t)\mathbb{B}(t))}{dt} = \frac{d\mathbb{A}(t)}{dt}\mathbb{B}(t) + \mathbb{A}(t)\frac{d\mathbb{B}(t)}{dt},$$

con la precaución que no se puede modificar el orden de aparición de los operadores. Es fácil ver que:

$$\begin{aligned} \left\langle u^k \middle| \frac{d(\mathbb{A}(t)\mathbb{B}(t))}{dt} \right| u_i \rangle &= \frac{d}{dt} \left\langle u^k \middle| \mathbb{A}(t)\mathbb{B}(t) \right| u_i \rangle = \frac{d}{dt} \left\langle u^k \middle| \mathbb{A}(t)\mathbb{B}(t) \right| u_i \rangle \\ &= \left(\left\langle u^k \middle| \mathbb{A}(t) \right| u_m \rangle \langle u^m \middle| \mathbb{B}(t) \right| u_i \rangle \right) \\ &= \frac{d \langle u^k \middle| \mathbb{A}(t) \rangle u_m}{dt} \langle u^m \middle| \mathbb{B}(t) \rangle u_i \rangle + \left\langle u^k \middle| \mathbb{A}(t) \right| u_m \rangle \frac{d \langle u^m \middle| \mathbb{B}(t) \rangle u_i}{dt} \\ &= \left\langle u^k \middle| \frac{d\mathbb{A}(t)}{dt} \right| u_m \rangle \langle u^m \middle| \mathbb{B}(t) \rangle u_i \rangle + \left\langle u^k \middle| \mathbb{A}(t) \right| u_m \rangle \langle u^m \middle| \frac{d\mathbb{B}(t)}{dt} \rangle u_i \rangle. \end{aligned}$$

4.3.7 Ejemplos

- Si tenemos un matriz \mathbb{B} , 2×3 , de la forma

$$\mathbb{B} = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & 4 \end{pmatrix},$$

y suponemos las bases canónicas para \mathbf{V}^3 y \mathbf{V}^2 : $\{|i_1\rangle, |i_2\rangle, |i_3\rangle\}$ y $\{|i_1\rangle, |i_2\rangle\}$, respectivamente, entonces la matriz \mathbb{B} representa la transformación $\mathbb{B} : \mathbf{V}^3 \rightarrow \mathbf{V}^2$ que lleva un vector genérico $|x\rangle = (x_1, x_2, x_3)$ a un vector genérico $|y\rangle = (y_1, y_2)$ tal que:

$$\mathbb{B} = \begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & 4 \end{pmatrix} \Rightarrow \mathbb{B}|x\rangle = |y\rangle \Rightarrow \begin{pmatrix} 3 & 1 & -2 \\ 1 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix},$$

esto es:

$$y_1 = 3x_1 + x_2 - 2x_3$$

$$y_2 = x_1 + 0x_2 + 4x_3.$$

Es claro que la representación matricial dependerá de la base en la cual se exprese.

- Si suponemos el operador diferencial $\mathbb{D}(\cdot) = \frac{d(\cdot)}{dx}$ cuyo dominio está conformado por el espacio vectorial de los polinomios de grado ≤ 3 , entonces tendremos que: $\mathbb{D}(\cdot) : \mathbf{P}^3 \rightarrow \mathbf{P}^2$. En este caso las bases podrían ser: $|p_i\rangle = \{1, x, x^2, x^3\}$ y $|\tilde{p}_j\rangle = \{1, x, x^2\}$ de \mathbf{P}^3 y \mathbf{P}^2 respectivamente.

Al operar (diferenciar) sobre $|p_j\rangle \in \mathbf{P}^3$ y expresar ese resultado en la base de \mathbf{P}^2 tenemos:

$$\mathbb{D}|p_j\rangle \Rightarrow \begin{cases} \frac{d(1)}{dx} = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ \frac{d(x)}{dx} = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ \frac{d(x^2)}{dx} = 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 \\ \frac{d(x^3)}{dx} = 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2 \end{cases} \Rightarrow \langle \tilde{p}^\alpha | \mathbb{D}|p_j\rangle = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Los coeficientes de esa expansión serán las columnas de la matriz que los representa. Para enfatizar, los elementos de una matriz, no sólo dependen de la base sino del orden en el cual la base se presente.

Consideremos que la base de \mathbf{P}^2 viene representada por $|\hat{p}_j\rangle = \{x^2, x, 1\}$, la representación matricial del operador $\mathbb{D}(\cdot) = \frac{d(\cdot)}{dx}$ será:

$$\mathbb{D}|p_j\rangle \Rightarrow \begin{cases} \frac{d(1)}{dx} = 0 = 0 \cdot x^2 + 0 \cdot x + 0 \cdot 1 \\ \frac{d(x)}{dx} = 1 = 0 \cdot x^2 + 0 \cdot x + 1 \cdot 1 \\ \frac{d(x^2)}{dx} = 2x = 0 \cdot x^2 + 2 \cdot x + 0 \cdot 1 \\ \frac{d(x^3)}{dx} = 3x^2 = 3 \cdot x^2 + 0 \cdot x + 0 \cdot 1 \end{cases} \Rightarrow \langle \hat{p}^\alpha | \mathbb{D}|p_j\rangle = \begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Se puede ver que $\mathbb{D}|p_j\rangle$ es:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix} \Rightarrow 1 + 2x + 3x^2,$$

y equivalentemente:

$$\begin{pmatrix} 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \Rightarrow 3x^2 + 2x + 1,$$

¡Es el mismo polinomio! Recuerde que las componentes del vector multiplican a los vectores bases en el mismo orden.

3. Construyamos la representación para el mismo operador \mathbb{D} del ejemplo anterior, pero ahora en las siguientes bases: $|p_i\rangle = \{1, 1+x, 1+x+x^2, 1+x+x^2+x^3\}$ y $|\tilde{p}_i\rangle = \{1, x, x^2\}$ de \mathbf{P}^3 y \mathbf{P}^2 , respectivamente. Entonces, nuevamente vemos que $\mathbb{D}|p_j\rangle$ implica:

$$\begin{cases} \frac{d(1)}{dx} = 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ \frac{d(1+x)}{dx} = 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ \frac{d(1+x+x^2)}{dx} = 1+2x = 1 \cdot 1 + 2 \cdot x + 0 \cdot x^2 \\ \frac{d(1+x+x^2+x^3)}{dx} = 1+2x+3x^2 = 1 \cdot 1 + 2 \cdot x + 3 \cdot x^2 \end{cases} \Rightarrow \langle \tilde{p}^\alpha | \mathbb{D}|p_j\rangle = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

4. Dada la siguiente transformación:

$$\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^4, \quad \mathbb{T}[(x, y, z)] = (x + y, x - 2z, x + 2y + 3z, y - 2z).$$

y tomemos como bases los conjuntos:

$$|e_i\rangle = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\} \in \mathbb{R}^3, \quad |\tilde{e}_i\rangle = \{(1, 1, 1, 1), (0, 1, 1, 1), (0, 0, 1, 1), (0, 0, 0, 1)\} \in \mathbb{R}^4.$$

Queremos encontrar la representación matricial del operador \mathbb{T} , esto es

$$T_j^i = \langle \tilde{e}^i | \mathbb{T} | e_j \rangle = \langle \tilde{e}^i | (C_j^k | \tilde{e}_k \rangle) \equiv C_j^k \langle \tilde{e}^i | \tilde{e}_k \rangle,$$

por lo tanto los C_j^k son los coeficientes de la expansión de los vectores base de \mathbb{R}^3 transformados. Es decir

$$\mathbb{T} |e_1\rangle = C_1^k | \tilde{e}_k \rangle, \quad \mathbb{T} |e_2\rangle = C_2^k | \tilde{e}_k \rangle, \quad \mathbb{T} |e_3\rangle = C_3^k | \tilde{e}_k \rangle, \quad \text{y} \quad \mathbb{T} |e_4\rangle = C_4^k | \tilde{e}_k \rangle.$$

Entonces, podemos hacer las siguientes expansiones:

$$\mathbb{T}[(1, 1, 0)] = (2, 1, 3, 1) = C_1^1(1, 1, 1, 1) + C_1^2(0, 1, 1, 1) + C_1^3(0, 0, 1, 1) + C_1^4(0, 0, 0, 1),$$

y al resolver para $\{C_1^1, C_1^2, C_1^3, C_1^4\}$ obtener: $\{C_1^1 = 2, C_1^2 = -1, C_1^3 = 2, C_1^4 = -2\}$.

Para la siguiente base de \mathbb{R}^3

$$\mathbb{T}[(1, 0, 1)] = (1, -1, 4, -2) = C_2^1(1, 1, 1, 1) + C_2^2(0, 1, 1, 1) + C_2^3(0, 0, 1, 1) + C_2^4(0, 0, 0, 1),$$

resolviendo para $\{C_2^1, C_2^2, C_2^3, C_2^4\}$ se obtiene: $\{C_2^1 = 1, C_2^2 = -2, C_2^3 = 5, C_2^4 = -6\}$.

Y finalmente, para:

$$\mathbb{T}[(0, 1, 1)] = (1, -2, 5, -1) = C_3^1(1, 1, 1, 1) + C_3^2(0, 1, 1, 1) + C_3^3(0, 0, 1, 1) + C_3^4(0, 0, 0, 1),$$

resulta: $\{C_3^1 = 1, C_3^2 = -3, C_3^3 = 7, C_3^4 = -6\}$.

Por lo tanto, la representación matricial que hemos obtenido será:

$$\begin{aligned} T_1^1 &= C_1^k \langle \tilde{e}^1 | \tilde{e}_k \rangle = C_1^1 \langle \tilde{e}^1 | \tilde{e}_1 \rangle + C_1^2 \langle \tilde{e}^1 | \tilde{e}_2 \rangle + C_1^3 \langle \tilde{e}^1 | \tilde{e}_3 \rangle + C_1^4 \langle \tilde{e}^1 | \tilde{e}_4 \rangle, \\ T_2^1 &= C_2^k \langle \tilde{e}^1 | \tilde{e}_k \rangle = C_2^1 \langle \tilde{e}^1 | \tilde{e}_1 \rangle + C_2^2 \langle \tilde{e}^1 | \tilde{e}_2 \rangle + C_2^3 \langle \tilde{e}^1 | \tilde{e}_3 \rangle + C_2^4 \langle \tilde{e}^1 | \tilde{e}_4 \rangle, \dots \end{aligned}$$

Vale decir:

$$T_1^1 = 2 \cdot 4 + (-1) \cdot 3 + 2 \cdot 2 + (-1) \cdot 1 = 7,$$

$$T_2^1 = 1 \cdot 4 + (-2) \cdot 3 + 5 \cdot 2 + (-6) \cdot 1 = 2, \quad \dots$$

Si continuamos haciendo todas las cuentas, para todas las demás componentes, tendremos la siguiente matriz:

$$\langle \tilde{e}_j | \mathbb{T} | e_i \rangle = \begin{pmatrix} 7 & 2 & 3 \\ 5 & 1 & 2 \\ 4 & 2 & 4 \\ 1 & -2 & -1 \end{pmatrix}.$$

La ecuación para la transformación de cualquier vector es entonces:

$$\mathbb{T} | \tilde{x}_i \rangle = T_i^j | x_j \rangle \Rightarrow \begin{pmatrix} \tilde{x}^1 \\ \tilde{x}^2 \\ \tilde{x}^3 \\ \tilde{x}^4 \end{pmatrix} = \begin{pmatrix} 7 & 2 & 3 \\ 5 & 1 & 2 \\ 4 & 2 & 4 \\ 1 & -2 & -1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix}.$$

Pedimos al lector repetir los cálculos pero considerando las bases:

$$|e_i\rangle = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\} \in \mathbb{R}^3, \quad |\tilde{e}_i\rangle = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\} \in \mathbb{R}^4;$$

y también las siguientes bases canónicas

$$|\mathbf{e}_i\rangle = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\} \in \mathbb{R}^3, \quad |\tilde{\mathbf{e}}_i\rangle = \{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 0, 0, 1)\} \in \mathbb{R}^4.$$

¿Qué puede concluir de los resultados obtenidos de este último par de bases?

5. Para la matriz

$$\mathbb{A} = \begin{pmatrix} 3 & -2 & 2 \\ 1 & 2 & -3 \\ 4 & 1 & 2 \end{pmatrix},$$

el desarrollo de Laplace (4.20) para la primera fila ($i = 1$) es:

$$\begin{aligned} \det |\mathbb{A}| &= A_1^1(A^c)_1^1 + A_2^1(A^c)_2^1 + A_3^1(A^c)_3^1 \\ &= 3 \begin{vmatrix} 2 & -3 \\ 1 & 2 \end{vmatrix} - (-2) \begin{vmatrix} 1 & -3 \\ 4 & 2 \end{vmatrix} + 2 \begin{vmatrix} 1 & 2 \\ 4 & 1 \end{vmatrix} \\ &= 3(7) + 2(14) + 2(-7) = 35. \end{aligned}$$

6. En Mecánica Clásica la cantidad de momento angular viene definida como $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Para pasar a Mecánica Cuántica se asocia \mathbf{r} y \mathbf{p} con los operadores posición y cantidad de movimiento los cuales, al operar sobre la función de onda nos proveen lo siguiente:

$$\langle r | \mathbb{X} | \psi \rangle = x \langle r | \psi \rangle = x \psi(\mathbf{r}) \quad \langle r | \mathbb{P}_x | \psi \rangle = \left(-i\hbar \frac{\partial}{\partial x} \right) \langle r | \psi \rangle = -i\hbar \frac{\partial}{\partial x} \psi(\mathbf{r})$$

$$\langle r | \mathbb{Y} | \psi \rangle = y \langle r | \psi \rangle = y \psi(\mathbf{r}) \quad \langle r | \mathbb{P}_y | \psi \rangle = \left(-i\hbar \frac{\partial}{\partial y} \right) \langle r | \psi \rangle = -i\hbar \frac{\partial}{\partial y} \psi(\mathbf{r})$$

$$\langle r | \mathbb{Z} | \psi \rangle = z \langle r | \psi \rangle = z \psi(\mathbf{r}) \quad \langle r | \mathbb{P}_z | \psi \rangle = \left(-i\hbar \frac{\partial}{\partial z} \right) \langle r | \psi \rangle = -i\hbar \frac{\partial}{\partial z} \psi(\mathbf{r})$$

En coordenadas cartesianas, en la representación de coordenadas $\{|r\rangle\}$ tendremos que

$$\langle r | \mathbb{R} | \psi \rangle = \mathbf{r} \psi(\mathbf{r}) \quad \text{y} \quad \langle r | \mathbb{P}_x | \psi \rangle = -i\hbar \nabla \psi(\mathbf{r}).$$

De forma que en Mecánica Cuántica las componentes cartesianas del operador cantidad de movimiento angular son:

$$\langle r | \mathbb{L} | \psi \rangle = -i\hbar (\mathbf{r} \times \nabla) \psi(\mathbf{r})$$

$$\langle r | \mathbb{L} | \psi \rangle = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \psi(\mathbf{r}) \mathbf{i} - i\hbar \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \psi(\mathbf{r}) \mathbf{j} - i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right) \psi(\mathbf{r}) \mathbf{k}.$$

Utilizando las definiciones anteriores vamos a mostrar que el conmutador de las componentes cartesianas de la cantidad de movimiento angular cumple la siguiente relación

$$[\mathbb{L}_x, \mathbb{L}_y] | \psi \rangle = i\hbar \mathbb{L}_z | \psi \rangle,$$

con: $\mathbb{L}^1 = \mathbb{L}_1 = \mathbb{L}_x$; $\mathbb{L}^2 = \mathbb{L}_2 = \mathbb{L}_y$; $\mathbb{L}^3 = \mathbb{L}_3 = \mathbb{L}_z$. En general, $[\mathbb{L}_l, \mathbb{L}_m] = i\hbar \varepsilon_{lmn} \mathbb{L}^n$.

Dado que:

$$\begin{aligned} [\mathbb{L}_1, \mathbb{L}_2] | \psi \rangle &= [\mathbb{L}_x, \mathbb{L}_y] | \psi \rangle = (\mathbb{L}_x \mathbb{L}_y - \mathbb{L}_y \mathbb{L}_x) | \psi \rangle \\ &= -i\hbar \left(\left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) - \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \right) \psi(\mathbf{r}) \\ &= \left[y \frac{\partial}{\partial z} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) - z \frac{\partial}{\partial y} \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) - \left(z \frac{\partial}{\partial x} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) - x \frac{\partial}{\partial z} \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \right) \right] \psi \end{aligned}$$

Con lo cual:

$$\begin{aligned}
 &= \left[\left(\left(yz \frac{\partial}{\partial z \partial x} + y \frac{\partial}{\partial x} \right) - xy \frac{\partial^2}{\partial z^2} \right) - \left(z^2 \frac{\partial}{\partial y \partial x} - zx \frac{\partial}{\partial y \partial z} \right) \right. \\
 &\quad \left. - \left(zy \frac{\partial}{\partial x \partial z} - z^2 \frac{\partial}{\partial y \partial x} \right) - \left(yx \frac{\partial}{\partial z^2} - zx \frac{\partial}{\partial z \partial y} - x \frac{\partial}{\partial y} \right) \right] \psi(\mathbf{r}) \\
 &= \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) \psi(\mathbf{r}) .
 \end{aligned}$$

7. Consideremos que el espacio de estados para un determinado sistema físico viene expandido por la base ortonormal $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$. Definamos dos operadores \mathbb{L}_z y \mathbb{S} de la siguiente manera:

$$\begin{aligned}
 \mathbb{L}_z |u_1\rangle &= |u_1\rangle, \quad \mathbb{L}_z |u_2\rangle = 0, \quad \mathbb{L}_z |u_3\rangle = -|u_3\rangle, \\
 \mathbb{S} |u_1\rangle &= |u_3\rangle, \quad \mathbb{S} |u_2\rangle = |u_2\rangle, \quad \mathbb{S} |u_3\rangle = |u_1\rangle .
 \end{aligned}$$

- (a). Calculemos la representación matricial en la base $\{|u_1\rangle, |u_2\rangle, |u_3\rangle\}$ del operador: $[\mathbb{L}_z, \mathbb{S}]$.

La matriz será:

$$\begin{pmatrix} \langle u^1 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_1 \rangle & \langle u^1 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_2 \rangle & \langle u^1 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_3 \rangle \\ \langle u^2 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_1 \rangle & \langle u^2 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_2 \rangle & \langle u^2 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_3 \rangle \\ \langle u^3 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_1 \rangle & \langle u^3 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_2 \rangle & \langle u^3 | \mathbb{L}_z \mathbb{S} - \mathbb{S} \mathbb{L}_z | u_3 \rangle \end{pmatrix}$$

con lo cual:

$$\begin{pmatrix} \langle u^1 | \mathbb{L}_z \mathbb{S} | u_1 \rangle - \langle u^1 | \mathbb{S} \mathbb{L}_z | u_1 \rangle & \langle u^1 | \mathbb{L}_z \mathbb{S} | u_2 \rangle - \langle u^1 | \mathbb{S} \mathbb{L}_z | u_2 \rangle & \langle u^1 | \mathbb{L}_z \mathbb{S} | u_3 \rangle - \langle u^1 | \mathbb{S} \mathbb{L}_z | u_3 \rangle \\ \langle u^2 | \mathbb{L}_z \mathbb{S} | u_1 \rangle - \langle u^2 | \mathbb{S} \mathbb{L}_z | u_1 \rangle & \langle u^2 | \mathbb{L}_z \mathbb{S} | u_2 \rangle - \langle u^2 | \mathbb{S} \mathbb{L}_z | u_2 \rangle & \langle u^2 | \mathbb{L}_z \mathbb{S} | u_3 \rangle - \langle u^2 | \mathbb{S} \mathbb{L}_z | u_3 \rangle \\ \langle u^3 | \mathbb{L}_z \mathbb{S} | u_1 \rangle - \langle u^3 | \mathbb{S} \mathbb{L}_z | u_1 \rangle & \langle u^3 | \mathbb{L}_z \mathbb{S} | u_2 \rangle - \langle u^3 | \mathbb{S} \mathbb{L}_z | u_2 \rangle & \langle u^3 | \mathbb{L}_z \mathbb{S} | u_3 \rangle - \langle u^3 | \mathbb{S} \mathbb{L}_z | u_3 \rangle \end{pmatrix} ,$$

de donde:

$$\langle u^i | [\mathbb{L}_z, \mathbb{S}] | u_j \rangle = \begin{pmatrix} 0 - 0 & 0 - 0 & 1 - (-1) \\ 0 - 0 & 0 - 0 & 0 - 0 \\ (-1) - 1 & 0 - 0 & 0 - 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix} .$$

- (b). ¿ \mathbb{L}_z, \mathbb{S} y $[\mathbb{L}_z, \mathbb{S}]$ serán biyectivas?

Veamos. Por definición \mathbb{S} es biyectiva ya que cada vector tiene su imagen, \mathbb{L}_z no lo es por cuanto el vector $|u_2\rangle$ no tiene imagen y, finalmente $[\mathbb{L}_z, \mathbb{S}]$ tampoco será biyectiva dado que $\langle u^i | [\mathbb{L}_z, \mathbb{S}] | u_j \rangle$ no tiene inversa ya que $\det |\langle u^i | [\mathbb{L}_z, \mathbb{S}] | u_j \rangle| = 0$.

- (c). Calculemos la dimensión del dominio, del rango y del núcleo de las transformaciones \mathbb{L}_z, \mathbb{S} y $[\mathbb{L}_z, \mathbb{S}]$.

Dominio Rango Núcleo

\mathbb{L}_z	3	2	1
\mathbb{S}	3	3	0
$[\mathbb{L}_z, \mathbb{S}]$	3	2	2

dado que $[\mathbb{L}_z, \mathbb{S}] | u_2 \rangle = 0$.

8. Buscaremos la expresión matricial para los operadores lineales de Pauli: $\mathbb{R}^2 \rightarrow \mathbb{R}^2$ sabiendo que actúan

como:

$$\sigma_z |+\rangle = |+\rangle, \quad \sigma_z |-\rangle = -|-\rangle, \quad \sigma_x |+\rangle_x = |+\rangle_x, \quad \sigma_x |-\rangle_x = -|-\rangle_x, \quad \sigma_y |+\rangle_y = |+\rangle_y, \quad \sigma_y |-\rangle_y = -|-\rangle_y$$

con:

$$|+\rangle \Leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle \Leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |+\rangle_x = \frac{1}{\sqrt{2}} [|+\rangle + |-\rangle], \quad |-\rangle_x = \frac{1}{\sqrt{2}} [|+\rangle - |-\rangle],$$

$$|+\rangle_y = \frac{1}{\sqrt{2}} [|+\rangle + i |-\rangle], \quad |-\rangle_y = \frac{1}{\sqrt{2}} [|+\rangle - i |-\rangle].$$

Ahora bien:

$${}_x \langle + |+\rangle_x = 1, \quad {}_x \langle + |-\rangle_x = {}_x \langle - |+\rangle_x = 0, \quad {}_x \langle - |-\rangle_x = 1,$$

$${}_y \langle + |+\rangle_y = 1, \quad {}_y \langle + |-\rangle_y = {}_y \langle - |+\rangle_y = 0, \quad {}_y \langle - |-\rangle_y = 1.$$

Es decir, los vectores $\{|+\rangle_x, |-\rangle_x\}$ y $\{|+\rangle_y, |-\rangle_y\}$ forman bases ortonormales, por lo que los vectores $\{|+\rangle, |-\rangle\}$ se pueden expresar en término de esas bases como:

$$|+\rangle = \frac{1}{\sqrt{2}} [|+\rangle_x + |-\rangle_x], \quad |-\rangle = \frac{1}{\sqrt{2}} [|+\rangle_x - |-\rangle_x],$$

$$|+\rangle = \frac{1}{\sqrt{2}} [|+\rangle_y + |-\rangle_y], \quad |-\rangle = \frac{-i}{\sqrt{2}} [|+\rangle_y - |-\rangle_y].$$

Así las expresiones matriciales serán:

$$(\sigma_z)_j^i = \begin{pmatrix} \langle + | \sigma_z |+ \rangle & \langle + | \sigma_z |-\rangle \\ \langle - | \sigma_z |+ \rangle & \langle - | \sigma_z |-\rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Para $(\sigma_x)_j^i$:

$$(\sigma_x)_j^i = \begin{pmatrix} \langle + | \sigma_x |+ \rangle & \langle + | \sigma_x |-\rangle \\ \langle - | \sigma_x |+ \rangle & \langle - | \sigma_x |-\rangle \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} [{}_x \langle + | + {}_x \langle - |] \sigma_x [|+\rangle_x + |-\rangle_x] & [{}_x \langle + | + {}_x \langle - |] \sigma_x [|+\rangle_x - |-\rangle_x] \\ [{}_x \langle + | - {}_x \langle - |] \sigma_x [|+\rangle_x + |-\rangle_x] & [{}_x \langle + | - {}_x \langle - |] \sigma_x [|+\rangle_x - |-\rangle_x] \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} [{}_x \langle + | + {}_x \langle - |] [|+\rangle_x - |-\rangle_x] & [{}_x \langle + | + {}_x \langle - |] [|+\rangle_x + |-\rangle_x] \\ [{}_x \langle + | - {}_x \langle - |] [|+\rangle_x - |-\rangle_x] & [{}_x \langle + | - {}_x \langle - |] [|+\rangle_x + |-\rangle_x] \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Y finalmente, para $(\sigma_y)_j^i$:

$$\begin{aligned}
 (\sigma_y)_j^i &= \begin{pmatrix} \langle + | \sigma_y | + \rangle & \langle + | \sigma_y | - \rangle \\ \langle - | \sigma_y | + \rangle & \langle - | \sigma_y | - \rangle \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} [y \langle + | +_y \langle - |] \sigma_y [|+ \rangle_y + |- \rangle_y] & -i [y \langle + | +_y \langle - |] \sigma_y [|+ \rangle_y - |- \rangle_y] \\ i [y \langle + | -_y \langle - |] \sigma_y [|+ \rangle_y + |- \rangle_y] & [y \langle + | -_y \langle - |] \sigma_y [|+ \rangle_y - |- \rangle_y] \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} [y \langle + | +_y \langle - |] [|+ \rangle_y - |- \rangle_y] & -i [y \langle + | +_y \langle - |] [|+ \rangle_y + |- \rangle_y] \\ i [y \langle + | -_y \langle - |] [|+ \rangle_y - |- \rangle_y] & -[y \langle + | -_y \langle - |] [|+ \rangle_y + |- \rangle_y] \end{pmatrix} \\
 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.
 \end{aligned}$$

4.3.8 Practicando con Maxima

1. Consideremos la transformación lineal del ejercicio resuelto con anterioridad:

$$\mathbb{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^4, \quad \mathbb{T}[(x, y, z)] = (x + y, x - 2z, x + 2y + 3z, y - 2z).$$

```
(%i1) load(vec)$
(%i2) T(x,y,z):=[x+y,x-2*z,x+2*y+3*z,y-2*z];
(%o2) T(x,y,z):=[x+y,x-2z,x+2y+3z,y-2z]
```

Escribimos ahora los vectores de la base en \mathbb{R}^4 :

```
(%i3) e1:[1,1,1,1];e2:[0,1,1,1];e3:[0,0,1,1]; e4:[0,0,0,1];
(%o3) [1,1,1,1]
(%o4) [0,1,1,1]
(%o5) [0,0,1,1]
(%o6) [0,0,0,1]
```

Entonces:

```
(%i7) linsolve(T(1,1,0)-C11*e1-C12*e2-C13*e3-C14*e4,[C11,C12,C13,C14]);
(%o7) [[C11 = 2, C12 = -1, C13 = 2, C14 = -2]]
(%i8) [C11,C12,C13,C14]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4])];
(%o8) [2, -1, 2, -2]
(%i9) linsolve(T(1,0,1)-C21*e1-C22*e2-C23*e3-C24*e4,[C21,C22,C23,C24]);
(%o9) [[C21 = 1, C22 = -2, C23 = 5, C24 = -6]]
(%i10) [C21,C22,C23,C24]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4])];
(%o10) [1, -2, 5, -6]
(%i11) linsolve(T(0,1,1)-C31*e1-C32*e2-C33*e3-C34*e4,[C31,C32,C33,C34]);
(%o11) [[C31 = 1, C32 = -3, C33 = 7, C34 = -6]]
(%i12) [C31,C32,C33,C34]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4])];
(%o12) [1, -3, 7, -6]
```

Calculemos ahora las componentes de la representación matricial:

```
(%i13) T11: C11*(e1.e1) + C12*(e1.e2) +C13*(e1.e3) + C14*(e1.e4);
(%o13) 7
(%i14) T21: C21*(e1.e1) + C22*(e1.e2) +C23*(e1.e3) + C24*(e1.e4);
(%o14) 2
(%i15) T31: C31*(e1.e1) + C32*(e1.e2) +C33*(e1.e3) + C34*(e1.e4);
(%o15) 3
(%i16) T12: C11*(e2.e1) + C12*(e2.e2) +C13*(e2.e3) + C14*(e2.e4);
(%o16) 5
(%i17) T22: C21*(e2.e1) + C22*(e2.e2) +C23*(e2.e3) + C24*(e2.e4);
(%o17) 1
(%i18) T32: C31*(e2.e1) + C32*(e2.e2) +C33*(e2.e3) + C34*(e2.e4);
```

```
(%o18) 2
(%i19) T13: C11*(e3.e1) + C12*(e3.e2) +C13*(e3.e3) + C14*(e3.e4);
(%o19) 4
(%i20) T23: C21*(e3.e1) + C22*(e3.e2) +C23*(e3.e3) + C24*(e3.e4);
(%o20) 2
(%i21) T33: C31*(e3.e1) + C32*(e3.e2) +C33*(e3.e3) + C34*(e3.e4);
(%o21) 4
(%i22) T14: C11*(e4.e1) + C12*(e4.e2) +C13*(e4.e3) + C14*(e4.e4);
(%o22) 1
(%i23) T24: C21*(e4.e1) + C22*(e4.e2) +C23*(e4.e3) + C24*(e4.e4);
(%o23) - 2
(%i24) T34: C31*(e4.e1) + C32*(e4.e2) +C33*(e4.e3) + C34*(e4.e4);
(%o24) - 1
```

Por lo tanto, la matriz que representa a la transformación es la siguiente:

```
(%i25) matrix([T11,T21,T31],[T12,T22,T32],[T13,T23,T33],[T14,T24,T34]);
(%o25) 
$$\begin{pmatrix} 7 & 2 & 3 \\ 5 & 1 & 2 \\ 4 & 2 & 4 \\ 1 & -2 & -1 \end{pmatrix}$$

```

Podemos repetir los cálculos pero ahora con la base canónica:

```
(%i26) e1: [1,0,0,0]; e2: [0,1,0,0]; e3: [0,0,1,0]; e4: [0,0,0,1];
(%o26) [1, 0, 0, 0]
(%o27) [0, 1, 0, 0]
(%o28) [0, 0, 1, 0]
(%o29) [0, 0, 0, 1]
```

Por ser la base ortonormal el cálculo de la representación matricial de la transformación es directa:

```
(%i30) linsolve(T(1,1,0)-C11*e1-C12*e2-C13*e3-C14*e4,[C11,C12,C13,C14]);
(%o30) [C11 = 2, C12 = 1, C13 = 3, C14 = 1]
(%i31) [C11,C12,C13,C14]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4])];
(%o31) [2, 1, 3, 1]
(%i32) linsolve(T(1,0,1)-C21*e1-C22*e2-C23*e3-C24*e4,[C21,C22,C23,C24]);
(%o32) [C21 = 1, C22 = -1, C23 = 4, C24 = -2]
(%i33) [C21,C22,C23,C24]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4])];
(%o33) [1, -1, 4, -2]
(%i34) linsolve(T(0,1,1)-C31*e1-C32*e2-C33*e3-C34*e4,[C31,C32,C33,C34]);
(%o34) [C31 = 1, C32 = -2, C33 = 5, C34 = -1]
(%i35) [C31,C32,C33,C34]:[rhs(%[1]),rhs(%[2]),rhs(%[3]),rhs(%[4])];
(%o35) [1, -2, 5, -1]
```

La representación matricial que resulta ahora es:

```
(%i36) matrix([2,1,1],[1,-1,-2],[3,4,5], [1,-2,-1]);
(%o36)

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & -2 \\ 3 & 4 & 5 \\ 1 & -2 & -1 \end{pmatrix}$$

(%i37) kill(all)$
```

2. Dada una transformación $\mathbb{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, donde una base para \mathbb{R}^3 es:

$$|e_i\rangle = \{(1, 2, 1), (1, 1, 0), (0, 0, 3)\} \in \mathbb{R}^3,$$

y las imágenes de esta base respecto a la base canónica en \mathbb{R}^2 son:

$$\mathbb{T}[|e_1\rangle] = (3, 3), \quad \mathbb{T}[|e_2\rangle] = (6, 3), \quad \mathbb{T}[|e_3\rangle] = (3, 6).$$

¿Qué transformación está asociada a las coordenadas canónicas en \mathbb{R}^3 ?

Escribamos los vectores.

```
(%i1) e1:[1,2,1];e2:[1,1,0];e3:[0,0,3];
(%o1) [1, 2, 1]
(%o2) [1, 1, 0]
(%o3) [0, 0, 3]
```

Ahora introducimos las imágenes:

```
(%i4) Te1:[3,3];Te2:[6,3];Te3:[3,6];
(%o4) [3, 3]
(%o5) [6, 3]
(%o6) [3, 6]
```

Respecto a estas coordenadas la transformación tiene la siguiente representación vectorial:

```
(%i7) A:matrix(Te1,Te2,Te3);
(%o7)

$$\begin{pmatrix} 3 & 3 \\ 6 & 3 \\ 3 & 6 \end{pmatrix}$$

```

Lo que queremos es buscar la representación matricial de la transformación pero respecto a la base canónica, por lo tanto debemos buscar primero el cambio de base a la base canónica en \mathbb{R}^3

```
(%i8) E:invert(matrix([1,2,1],[1,1,0],[0,0,3]));
(%o8)

$$\begin{pmatrix} -1 & 2 & \frac{1}{3} \\ 1 & -1 & -\frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$$

```

En la base canónica la transformación tiene la siguiente representación:

```
(%i9) AE:transpose(E.A);
(%o9)

$$\begin{pmatrix} 10 & -4 & 1 \\ 5 & -2 & 2 \end{pmatrix}$$

```

Por lo tanto la transformación es:

```
(%i10) T(x,y,z):=[10*x-4*y+z,5*x-2*y+2*z];
(%o10) T(x,y,z):=[10x - 4y + z, 5x - 2y + 2z]
```

Podemos verificar si el resultado es el correcto.

```
(%i11) solve(T(1,2,1)-a*[1,0]-b*[0,1],[a,b]);
(%o11) [[a = 3, b = 3]]
(%i12) solve(T(1,1,0)-a*[1,0]-b*[0,1],[a,b]);
(%o12) [[a = 6, b = 3]]
(%i13) solve(T(0,0,3)-a*[1,0]-b*[0,1],[a,b]);
(%o13) [[a = 3, b = 6]]
```

4.3.9 Ejercicios

1. Considere el espacio $\mathbf{P}_{(t)N}$ de polinomios de grado N en t , vale decir, $|f\rangle_t \leftrightarrow \sum_{n=0}^N a_n t^n$, considere además un operador $\mathbb{T} = e^{x\mathbb{D}} \equiv \exp(x\mathbb{D})$, con $\mathbb{D} = \frac{d}{dt}$.
 - (a). Muestre que $\mathbb{T}|p\rangle_t = |p\rangle_{t+x}$, esto es que el operador \mathbb{T} puede ser considerado un operador traslación espacial para los polinomios $\mathcal{P}(t)$ de grado N .
 - (b). Considere que el espacio de polinomios está definido en el intervalo $[-1, 1]$, que definimos un producto interno de la forma $\langle f | g \rangle = \int_{-1}^1 f g dt$ y un espacio de polinomios de grado $N = 2$. ¿Cuál es la representación matricial de \mathbb{T} en la base de polinomios de Legendre $\{P_0, P_1, P_2\}$?
2. Considere ahora el espacio vectorial $\mathbf{P}_{(t)4}$ de polinomios en t de grado 4, definidos en el intervalo $[-1, 1]$. Esto es $|f\rangle_t \leftrightarrow \sum_{n=0}^4 a_n t^n$ y este espacio está equipado con un producto interno de la forma $\langle f | g \rangle = \int_{-1}^1 f(t)g(t) dt$. Considere además para este espacio vectorial un operador lineal representado por $\mathbb{T} = e^{\mathbb{D}} \equiv \exp(\mathbb{D})$, donde $\mathbb{D} = \frac{d}{dt}$. Claramente, podemos encontrar dos bases para ese espacio vectorial: $\{1, t, t^2, t^3, t^4\}$ y la base de polinomios de Legendre $\{P_0, P_1, P_2, P_3, P_4\}$
 - (a). Considere el polinomio $|f\rangle_t \leftrightarrow f(t) = 5t + 3t^2 + 4t^3$. Encuentre la expresión de este polinomio en términos de las bases arriba mencionadas. ¿Cuál es la matriz de transformación de las componentes de ese vector entre ambas bases?
 - (b). Construya un proyector sobre el subespacio $\mathbf{P}_{(t)2}$, de polinomios de grado 2 y encuentre la proyección de $|f\rangle_t$ en ese subespacio. Discuta las diferencias y semejanzas entre las proyecciones de $|f\rangle_t$ en ese subespacio expresado en las bases $\{1, t, t^2\}$ y $\{P_0, P_1, P_2\}$.
 - (c). Para $\mathbf{P}_{(t)4}$, construya el operador inverso \mathbb{T}^{-1} , el adjunto del operador \mathbb{T}^\dagger y precise si \mathbb{T} , es Hermítico o unitario.
 - (d). ¿Cuáles son las representaciones matriciales de \mathbb{T} en $\mathbf{P}_{(t)4}$, para cada una de las bases mencionadas? ¿Cómo transforman las representaciones? Compruebe que la traza y el determinante de ambas representaciones matriciales coinciden.
 - (e). ¿Cuáles son las representaciones matriciales de \mathbb{T}^{-1} y \mathbb{T}^\dagger en $\mathbf{P}_{(t)4}$ para cada una de las bases mencionadas? ¿Cómo transforman esas representaciones?
3. Heredando el formalismo de la Mecánica Clásica, se construye en Mecánica Cuántica el operador hamiltoniano, hermítico, para un sistema unidimensional como:

$$\mathbb{H} = \frac{\mathbb{P}^2}{2m} + \mathbb{V}(\mathbb{X}),$$

donde \mathbb{H} , \mathbb{P} y \mathbb{X} son operadores, y $\mathbb{V}(\mathbb{X})$ un operador que es una función de otro operador. Adicionalmente, uno puede construir el siguiente operador: $[\mathbb{X}, \mathbb{P}] = i\hbar\mathbb{I}$.

- (a). Determine los siguientes conmutadores: $[\mathbb{H}, \mathbb{P}]$, $[\mathbb{H}, \mathbb{X}]$ y $[\mathbb{H}, \mathbb{XP}]$.
 - (b). Suponga que $\mathbb{H}|\psi_n\rangle = E_n|\psi_n\rangle$ y entonces calcule la representación matricial del siguiente operador $\langle\psi^m|[\mathbb{A}, \mathbb{H}]|\psi_n\rangle$, para un operador arbitrario \mathbb{A} .
4. Considere la representación matricial de un operador en la base canónica de la forma:

$$\mathbb{A} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

¿Existirá algún operador \mathbb{B} , no singular, tal que $\mathbb{D} = \mathbb{B}^{-1}\mathbb{AB}$, con \mathbb{D} un operador diagonal? Si ese fuera el caso ¿Cuál sería la representación matricial de ese operador \mathbb{B} , en la base canónica? Justifique su respuesta.

5. Si

$$M_i^j(t) = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix}.$$

Demuestre que:

$$\frac{d^2\mathbb{M}(t)}{dt^2} = -\omega^2\mathbb{M}(t).$$

6. Dada a matriz:

$$A_i^j = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Evalúe: \mathbb{A}^2 , \mathbb{A}^3 , $\mathbb{A}^{\mathbb{A}x}$.

7. Dada la matriz:

$$A_i^j = \begin{pmatrix} 0 & 0 & i \\ -i & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}.$$

Demuestre que $\mathbb{A}^n = \mathbb{AAA}\dots = \mathbb{I}$, con $n \neq 0$.

8. Considere el siguiente “operador vectorial” $\boldsymbol{\sigma} = \sigma_x\mathbf{i} + \sigma_y\mathbf{j} + \sigma_z\mathbf{k}$, donde las matrices $\sigma_x, \sigma_y, \sigma_z$ son los operadores de Pauli que discutimos en la sección 4.3.7 y su representación matricial en la base canónica es:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

El vector dirección puede escribirse como:

$$\mathbf{n} = \sin(\theta)\cos(\phi)\mathbf{i} + \sin(\theta)\sin(\phi)\mathbf{j} + \cos(\theta)\mathbf{k} = n_x\mathbf{i} + n_y\mathbf{j} + n_z\mathbf{k},$$

con lo cual: $\boldsymbol{\sigma} \cdot \mathbf{n} = n_x\sigma_x + n_y\sigma_y + n_z\sigma_z$.

A partir de todo lo anterior calcule la representación matricial del siguiente operador en la base canónica

$$\exp\left(\frac{i}{2}\psi \boldsymbol{\sigma} \cdot \mathbf{n}\right).$$

9. Dadas las siguientes matrices:

$$\mathbb{L}_x = \begin{pmatrix} 0 & \sqrt{2}/2 & 0 \\ \sqrt{2}/2 & 0 & \sqrt{2}/2 \\ 0 & \sqrt{2}/2 & 0 \end{pmatrix}, \quad \mathbb{L}_y = \begin{pmatrix} 0 & i\sqrt{2}/2 & 0 \\ -i\sqrt{2}/2 & 0 & i\sqrt{2}/2 \\ 0 & -i\sqrt{2}/2 & 0 \end{pmatrix}, \quad \mathbb{L}_z = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Demuestre las siguientes reglas de commutación:

- (a). $\mathbb{L}_x \mathbb{L}_y - \mathbb{L}_y \mathbb{L}_x = i\mathbb{L}_z$.
- (b). $\mathbb{L}_y \mathbb{L}_z - \mathbb{L}_z \mathbb{L}_y = i\mathbb{L}_x$.
- (c). $\mathbb{L}_z \mathbb{L}_x - \mathbb{L}_x \mathbb{L}_z = i\mathbb{L}_y$.
- (d). $\mathbb{L}_x^2 + \mathbb{L}_y^2 + \mathbb{L}_z^2 = 2$.

10. Calcule los siguientes determinantes:

$$\begin{vmatrix} 2 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 1 & 2 & 3 \\ 1 & 2-x^2 & 2 & 3 \\ 2 & 3 & 1 & 5 \\ 2 & 3 & 1 & 9-x^2 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -2 & 1 \\ 3 & -3 & 4 & -2 \\ -2 & 1 & -2 & 0 \end{vmatrix}, \quad \begin{vmatrix} -2 & 5 & 0 & -1 & 3 \\ 1 & 0 & 3 & 7 & -2 \\ 3 & -1 & 0 & 5 & -5 \\ 2 & 6 & -4 & 1 & 2 \\ 0 & -3 & -1 & 2 & 3 \end{vmatrix}.$$

11. Utilizando las propiedades de los determinantes resuelva para x

$$\begin{vmatrix} x+2 & x+4 & x-3 \\ x+3 & x & x+5 \\ x-2 & x-1 & x+1 \end{vmatrix} = 0.$$

12. Utilizando la propiedad lineal de los determinantes calcule:

$$\Delta = \begin{vmatrix} a+b & n+p \\ c+d & q+r \end{vmatrix}.$$

13. Encuentre el determinante de la matriz 4×4 definida por la función:

$$A_j^i = \frac{1}{i+j-x}.$$

14. Muestre que si las filas de un determinante de orden n son linealmente independientes, entonces sus columnas son también linealmente independientes.

15. Dada la representación matricial de un operador:

$$\begin{pmatrix} 1 & 1 & 3 \\ 1 & 1 & -3 \\ 3 & -3 & -3 \end{pmatrix}.$$

- (a). Encuentre la representación diagonal y la base que diagonaliza esta matriz.
- (b). Muestre que la traza y el determinante de ambas representaciones matriciales coinciden.
- (c). Encuentre la matriz de transformación a la representación diagonal.

16. Pruebe que la representación matricial, 2×2 más general para un operador unitario y simétrico (vale decir: $\mathbb{S}\mathbb{S}^\dagger = \mathbb{I}$ y $\mathbb{S}^T = \mathbb{S}$) puede ser escrita como:

$$\mathbb{S} = \begin{pmatrix} \alpha e^{2i\beta} & i\sqrt{1-\alpha^2}e^{i(\beta+\gamma)} \\ i\sqrt{1-\alpha^2}e^{i(\beta+\gamma)} & \alpha e^{2i\gamma} \end{pmatrix},$$

con α, β y γ parámetros reales y, adicionalmente $0 \leq \alpha \leq 1$.

17. Resuelva los ejercicios anteriores utilizando **Maxima**.

4.4 Un zoológico de matrices cuadradas

A continuación presentaremos un conjunto de matrices importantes y de operaciones que serán de gran utilidad para el resto de desarrollo del tema¹³. Es bueno tener en mente que las matrices pueden tener todas sus elementos reales, diremos que son matrices pertenecientes al espacio vectorial de matrices reales $\mathbb{R}_{n \times m}$ o tener como elementos números complejos, en este caso diremos que pertenecen al espacio vectorial de matrices complejas $\mathbb{C}_{n \times m}$.

4.4.1 Matriz unidad y la matriz nula

La matriz unitaria es aquella que sólo tiene elementos en la diagonal iguales a la unidad $\mathbb{I} \Rightarrow I_j^i = \delta_j^i$, y el resto cero. Mientras que la matriz nula es aquella con todos los elementos iguales a cero $\mathbb{O} \Rightarrow 0_j^i = 0$.

4.4.2 Matriz diagonal y diagonal a bloques

Una matriz diagonal es aquella que contiene elementos únicamente en la diagonal las demás componentes son cero: $A_j^i = 0 \forall i \neq j$. Como veremos más adelante, existen algunos métodos para reducir una matriz a su forma diagonal. Una propiedad importante de estas matrices es que comutan entre si, es decir, $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A}$, si \mathbb{A} y \mathbb{B} son diagonales. Podemos tener matrices diagonales a bloques, vale decir:

$$D_j^i = \begin{pmatrix} D_1^1 & D_2^1 & 0 & 0 \\ D_1^2 & D_2^2 & 0 & 0 \\ 0 & 0 & D_3^3 & D_4^3 \\ 0 & 0 & D_3^4 & D_4^4 \end{pmatrix}.$$

También tenemos las triangulares superior e inferior:

$$\check{D}_j^i = \begin{pmatrix} \check{D}_1^1 & \check{D}_2^1 & \check{D}_3^1 & \check{D}_4^1 \\ 0 & \check{D}_2^2 & \check{D}_3^2 & \check{D}_4^2 \\ 0 & 0 & \check{D}_3^3 & \check{D}_4^3 \\ 0 & 0 & 0 & \check{D}_4^4 \end{pmatrix}, \quad \text{y} \quad \hat{D}_j^i = \begin{pmatrix} \hat{D}_1^1 & 0 & 0 & 0 \\ \hat{D}_1^2 & \hat{D}_2^2 & 0 & 0 \\ \hat{D}_1^3 & \hat{D}_2^3 & \hat{D}_3^3 & 0 \\ \hat{D}_1^4 & \hat{D}_2^4 & \hat{D}_3^4 & \hat{D}_4^4 \end{pmatrix}.$$

4.4.3 Matriz traspuesta

Tal y como discutimos en la sección 4.3.2.2, cuando desarrollamos el concepto de matrices hermíticas conjungadas, la matriz traspuesta es aquella que se obtiene intercambiando filas por columnas: $\mathbb{A} = A_j^i \Rightarrow \mathbb{A}^T = A_i^j$. Si $\mathbb{A} = \mathbb{A}^T$, ($A_j^i = A_i^j$), se dice que \mathbb{A} es simétrica. Y si $\mathbb{A} = -\mathbb{A}^T$, ($A_j^i = -A_i^j$), es llamada antisimétrica. Por lo tanto, una matriz cuadrada se puede representar por la suma de una matriz simétrica y una antisimétrica

$$\mathbb{A} = \frac{1}{2} [\mathbb{A} + \mathbb{A}^T] + \frac{1}{2} [\mathbb{A} - \mathbb{A}^T].$$

¹³Para mayores detalles sobre éstas y otros tipos de matrices pueden consultar Antony, R. and Alemayehu, H., A NOTE ON SPECIAL MATRICES, *Italian journal of pure and applied mathematics*, (2015), 587-604.

4.4.4 Matriz de cofactores

A una matriz \mathbb{A} le podemos asociar una de cofactores \mathbb{A}^c de la manera siguiente:

$$A_j^i = \begin{pmatrix} a_1^1 & a_2^1 & a_3^1 \\ a_1^2 & a_2^2 & a_3^2 \\ a_1^3 & a_2^3 & a_3^3 \end{pmatrix} \Rightarrow (A^c)_j^i = \begin{pmatrix} (A^c)_1^1 & (A^c)_2^1 & (A^c)_3^1 \\ (A^c)_1^2 & (A^c)_2^2 & (A^c)_3^2 \\ (A^c)_1^3 & (A^c)_2^3 & (A^c)_3^3 \end{pmatrix}.$$

Donde los $(A^c)_j^i$ forman la matriz de cofactores, estos cofactores son:

$$(A^c)_1^1 = (-1)^{1+1} \begin{vmatrix} a_2^2 & a_3^2 \\ a_2^3 & a_3^3 \end{vmatrix}, \quad (A^c)_2^1 = (-1)^{1+2} \begin{vmatrix} a_1^2 & a_3^2 \\ a_1^3 & a_3^3 \end{vmatrix}, \quad (A^c)_3^1 = (-1)^{1+3} \begin{vmatrix} a_1^2 & a_2^2 \\ a_1^3 & a_2^3 \end{vmatrix},$$

$$(A^c)_1^2 = (-1)^{2+1} \begin{vmatrix} a_2^1 & a_3^1 \\ a_2^3 & a_3^3 \end{vmatrix}, \quad (A^c)_2^2 = (-1)^{2+2} \begin{vmatrix} a_1^1 & a_3^1 \\ a_1^3 & a_3^3 \end{vmatrix}, \quad (A^c)_3^2 = (-1)^{2+3} \begin{vmatrix} a_1^1 & a_2^1 \\ a_1^3 & a_2^3 \end{vmatrix},$$

$$(A^c)_1^3 = (-1)^{3+1} \begin{vmatrix} a_2^1 & a_3^1 \\ a_2^2 & a_3^2 \end{vmatrix}, \quad (A^c)_2^3 = (-1)^{3+2} \begin{vmatrix} a_1^1 & a_3^1 \\ a_1^2 & a_3^2 \end{vmatrix}, \quad (A^c)_3^3 = (-1)^{3+3} \begin{vmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{vmatrix}.$$

4.4.5 Matriz adjunta

Llamaremos matriz adjunta, $\text{adj}[\mathbb{A}]$, a la traspuesta de la matriz de cofactores de una determinada matriz:

$$\text{adj}[\mathbb{A}] = (\mathbb{A}^c)^T \Rightarrow \text{adj}[A_j^i] = ((A^c)_j^i)^T = (A^c)_i^j.$$

Por ejemplo: $\mathbb{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \Rightarrow \text{adj}[\mathbb{A}] = \begin{pmatrix} -3 & 6 & -3 \\ 6 & -12 & 6 \\ -3 & 6 & -3 \end{pmatrix}.$

Una matriz será autoadjunta si $\text{adj}[\mathbb{A}] = \mathbb{A}$.

4.4.6 Matriz singular

\mathbb{A} es singular si $\det|\mathbb{A}| = 0$. Si $\det|\mathbb{A}| \neq 0$, existirá un número r máximo de vectores filas, o columnas, linealmente independientes, llamaremos a ese número, $0 \leq r \leq n$, el rango de la matriz.

4.4.7 Matriz inversa

Hemos visto que dada una transformación lineal biyectiva, podemos definir una inversa para esa transformación lineal. Esa transformación lineal tendrá como representación un matriz. Por lo tanto, dado un operador lineal \mathbb{A} diremos que otro operador lineal \mathbb{B} será su inverso (por la derecha) si

$$\mathbb{AB} = \mathbb{I} \rightarrow \langle e^i | \mathbb{AB} | e_j \rangle = \delta_j^i \rightarrow A_k^i B_j^k = \delta_j^i.$$

Ahora bien, como conocemos la matriz A_k^i y la suponemos no singular ($\det|A_k^i| \neq 0$) al tomar un j fijo tendremos un sistema de n ecuaciones lineales inhomogéneo con n incógnitas: $B_j^1, B_j^2, B_j^3, \dots, B_j^n$. Al resolver el sistema tendremos la solución.

Un procedimiento para encontrar la inversa es el método de eliminación de Gauss-Jordan. Veamos como

funciona para una matriz 3×3 :

$$\left(\begin{array}{ccc|ccc} A_1^1 & A_2^1 & A_3^1 & 1 & 0 & 0 \\ A_1^2 & A_2^2 & A_3^2 & 0 & 1 & 0 \\ A_1^3 & A_2^3 & A_3^3 & 0 & 0 & 1 \end{array} \right) \xrightarrow{\text{Gauss-Jordan}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & B_1^1 & B_2^1 & B_3^1 \\ 0 & 1 & 0 & B_1^2 & B_2^2 & B_3^2 \\ 0 & 0 & 1 & B_1^3 & B_2^3 & B_3^3 \end{array} \right).$$

Algunas propiedades importantes:

$$(\mathbb{A}^{-1})^{-1} = \mathbb{A}, \quad (\mathbb{A}^T)^{-1} = (\mathbb{A}^{-1})^T, \quad (\mathbb{A}^\dagger)^{-1} = (\mathbb{A}^{-1})^\dagger, \quad (\mathbb{A}\mathbb{B})^{-1} = \mathbb{B}^{-1}\mathbb{A}^{-1}.$$

4.4.8 Matriz ortogonal

Anteriormente mencionamos, en la sección 1.4.3, que en el espacio real \mathbb{R}^3 podemos tener una transformación de coordenadas cartesianas: $(x, y, z) \rightarrow (\tilde{x}, \tilde{y}, \tilde{z})$ que consiste en rotar, un ángulo θ , uno de los sistemas respecto al otro. Si la rotación se hace, digamos alrededor del eje z , la relación entre las coordenadas es la siguiente:

$$\begin{cases} \tilde{x} = x \cos(\theta) + y \sen(\theta) \\ \tilde{y} = -x \sen(\theta) + y \cos(\theta) \\ \tilde{z} = z \end{cases} \Rightarrow \begin{cases} \tilde{x}^1 = x^1 \cos(\theta) + x^2 \sen(\theta) \\ \tilde{x}^2 = -x^1 \sen(\theta) + x^2 \cos(\theta) \\ \tilde{x}^3 = x^3 \end{cases} \Rightarrow \tilde{x}^i = \tilde{\alpha}_j^i x^j.$$

Si la longitud desde el origen a algún punto P del espacio es la misma para ambos sistemas de coordenadas

$$x^i x_i = \tilde{x}^i \tilde{x}_i = (\tilde{\alpha}_k^i x^k)(\tilde{\alpha}_i^j x_j) = x^k x_j \tilde{\alpha}_i^j \tilde{\alpha}_k^i.$$

entonces se tiene que cumplir que: $\tilde{\alpha}_i^j \tilde{\alpha}_k^i = \delta_k^j$.

Este resultado es una consecuencia de haber impuesto la condición para que las longitudes sean las mismas en ambos sistemas de coordenadas y se conoce como la *condición de ortogonalidad*. En el lenguaje de la matrices reales la condición de ortogonalidad se puede enunciar de diferentes formas, todas ellas equivalentes: $\mathbb{A}^T \mathbb{A} = \mathbb{A} \mathbb{A}^T = \mathbb{I}$, $\mathbb{A}^T = \mathbb{A}^{-1}$, $\sum_i A_i^j A_k^i = \delta_k^j$. y diremos que las matrices que las satisfacen son matrices ortogonales.

4.4.9 Matrices complejas

Las matrices con todas sus componentes reales no permite estudiar problemas que pueden aparecer en algunas ramas de la Física, como en la Mecánica Cuántica. Por lo tanto, de necesita generalizar nuestro estudio a espacios vectoriales a el de espacios vectoriales de matrices complejas: $\mathbb{C}_{n \times n}$, donde en las componentes de las matrices aparecen números complejos. Cuando las matrices son complejas surgen generalizaciones que podemos considerar, como mostramos a continuación.

- **La matriz adjunta o matriz traspuesta conjugada.** Estas matrices se construyen tomando el complejo conjugado de cada elemento y luego la traspuesta de la matriz. Se acostumbra denotarlas con $\mathbb{A}^\dagger = (\mathbb{A}^*)^T = (\mathbb{A}^T)^*$.
- **La matriz hermítica.** Son aquellas matrices que satisfacen la siguiente ecuación: $\mathbb{A}^\dagger = \mathbb{A}$. Notemos que si \mathbb{A} es real, entonces $\mathbb{A}^\dagger = \mathbb{A}^T$, y por lo tanto las matrices hermíticas reales son matrices simétricas.
- **La matriz normal.** Se llama matriz normal a aquella que satisface $\mathbb{A}^\dagger \mathbb{A} = \mathbb{A} \mathbb{A}^\dagger$.
- **La matriz unitaria.** La matriz \mathbb{A} es unitaria si: $\mathbb{A}^\dagger = \mathbb{A}^{-1}$. Si \mathbb{A} es real, entonces $\mathbb{A}^{-1} = \mathbb{A}^T$, de manera que las matrices unitarias reales son ortogonales. Es decir, las matrices unitarias son el análogo a las matrices ortogonales reales y son de fundamental importancia en la Mecánica Cuántica porque dejan la norma de los vectores invariantes.

Existe una relación importante entre matrices normales y matrices unitarias y es el hecho de que una matriz \mathbb{A} es normal, si y sólo si, existe una matriz diagonal \mathbb{D} y una unitaria \mathbb{U} tal que: $\mathbb{A} = \mathbb{U}\mathbb{D}\mathbb{U}^\dagger$.

Más adelante veremos que los valores en la diagonal de \mathbb{D} se denominan los *autovalores* de \mathbb{A} y las columnas de \mathbb{U} son los *autovectores* de \mathbb{A} .

4.4.10 Ejemplos

1. Queremos hallar la matriz inversa de $\mathbb{A} = \begin{pmatrix} 2 & 3 & 4 \\ 2 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix}$.

Utilizando el método de Gauss-Jordan. Entonces, escribimos la siguiente matriz aumentada:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 4 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array} \right).$$

Si multiplicamos la segunda fila por -1 y la sumamos con la primera fila obtenemos:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 4 & 1 & 0 & 0 \\ 2 & 1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & -1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array} \right).$$

Ahora multiplicamos la tercera fila por 2 y sumamos con la primera fila

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & -1 & 0 \\ -1 & 1 & 2 & 0 & 0 & 1 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & -1 & 0 \\ 0 & 5 & 8 & 1 & 0 & 2 \end{array} \right).$$

Multiplicamos la primera fila por -2 y sumamos con la tercera

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 4 & 1 & 0 & 0 \\ 0 & 2 & 3 & 1 & -1 & 0 \\ 0 & 5 & 8 & 1 & 0 & 2 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} -4 & -1 & 0 & -1 & 0 & 2 \\ 0 & 2 & 3 & 1 & -1 & 0 \\ 0 & 5 & 8 & 1 & 0 & 2 \end{array} \right).$$

Multiplicamos la segunda fila por $-8/3$ y sumamos con la tercera, luego multiplicamos esa segunda fila por -3

$$\left(\begin{array}{ccc|ccc} -4 & -1 & 0 & -1 & 0 & 2 \\ 0 & 2 & 3 & 1 & 1 & 0 \\ 0 & 5 & 8 & 1 & 0 & 2 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} -4 & -1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 5 & -8 & -6 \\ 0 & 5 & 8 & 1 & 0 & 2 \end{array} \right).$$

Multiplicamos la tercera fila por $-1/5$ y sumamos con la segunda fila, luego multiplicamos esa tercera fila por $-5/8$

$$\left(\begin{array}{ccc|ccc} -4 & -1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 5 & -8 & -6 \\ 0 & 5 & 8 & 1 & 0 & 2 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} -4 & -1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 5 & -8 & -6 \\ 0 & 0 & 1 & -3 & 5 & 4 \end{array} \right).$$

Se suma la primera fila con la segunda y luego se multiplica esta primera fila que resulta por $-1/4$.

$$\left(\begin{array}{ccc|ccc} -4 & -1 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 5 & -8 & -6 \\ 0 & 0 & 1 & -3 & 5 & 4 \end{array} \right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 2 & 1 \\ 0 & 1 & 0 & 5 & -8 & -6 \\ 0 & 0 & 1 & -3 & 5 & 4 \end{array} \right).$$

Por lo tanto, la matriz inversa es: $\mathbb{A}^{-1} = \begin{pmatrix} -1 & 2 & 1 \\ 5 & -8 & -6 \\ -3 & 5 & 4 \end{pmatrix}$.

2. También se puede obtener la matriz inversa de la siguiente manera: $\mathbb{A}^{-1} = \frac{\text{adj}[\mathbb{A}]}{\det[\mathbb{A}]}$.

Dada la siguiente matriz: $\mathbb{A} = \begin{pmatrix} 3 & -2 & 2 \\ 1 & 2 & -3 \\ 4 & 1 & 2 \end{pmatrix}$. Primero procedemos a calcular el determinante

$$\det[\mathbb{A}] = \begin{bmatrix} 3 & -2 & 2 \\ 1 & 2 & -3 \\ 4 & 1 & 2 \end{bmatrix} = 35.$$

Calculemos ahora la matriz cofactor:

$$(A^c)_1^1 = (-1)^2 \begin{vmatrix} 2 & -3 \\ 1 & 2 \end{vmatrix} = 7, \quad (A^c)_2^1 = (-1)^3 \begin{vmatrix} 1 & -3 \\ 4 & 2 \end{vmatrix} = 14, \quad (A^c)_3^1 = (-1)^4 \begin{vmatrix} 1 & 2 \\ 4 & 1 \end{vmatrix} = -7,$$

$$(A^c)_1^2 = (-1)^3 \begin{vmatrix} -2 & 2 \\ 1 & 2 \end{vmatrix} = 6, \quad (A^c)_2^2 = (-1)^4 \begin{vmatrix} 3 & 2 \\ 4 & 2 \end{vmatrix} = -2, \quad (A^c)_3^2 = (-1)^5 \begin{vmatrix} 3 & -2 \\ 4 & 1 \end{vmatrix} = -11,$$

$$(A^c)_1^3 = (-1)^4 \begin{vmatrix} -2 & 2 \\ 2 & -3 \end{vmatrix} = 2, \quad (A^c)_2^3 = (-1)^5 \begin{vmatrix} 3 & 2 \\ 1 & -3 \end{vmatrix} = 11, \quad (A^c)_3^3 = (-1)^6 \begin{vmatrix} 3 & -2 \\ 1 & 2 \end{vmatrix} = 8.$$

Por lo tanto, la matriz cofactor y la adjunta son:

$$\mathbb{A}^c = \begin{pmatrix} 7 & 14 & -7 \\ 6 & -2 & -11 \\ 2 & 11 & 8 \end{pmatrix} \Rightarrow \text{adj}[\mathbb{A}] = \begin{pmatrix} 7 & 6 & 2 \\ 14 & -2 & 11 \\ -7 & -11 & 8 \end{pmatrix}.$$

Y ahora si, la matriz inversa:

$$\mathbb{A}^{-1} = \frac{1}{35} \begin{pmatrix} 7 & 6 & 2 \\ 14 & -2 & 11 \\ -7 & -11 & 8 \end{pmatrix}.$$

4.4.11 Practicando con Maxima

En este ejercicio vamos a introducir algunos comandos básicos que nos permitirán realizar operaciones típicas con matrices. Dada la siguiente matriz:

```
(%i1) A:matrix([1,2,3],[4,8,5],[9,5,4]);
(%o1) ⎛1 2 3⎞
      ⎝4 8 5⎠
      ⎝9 5 4⎠
```

La traspuesta es:

```
(%i2) transpose(A);
(%o2) ⎛1 4 9⎞
      ⎝2 8 5⎠
      ⎝3 5 4⎠
```

El comando **adjoint(A)** calcula el adjunto de la matriz \mathbb{A} . La matriz adjunta es la traspuesta de la matriz de cofactores de \mathbb{A} .

```
(%i3) adjoint(A);
(%o3) ⎛ 7 7 -14⎞
      ⎝29 -23 7 ⎠
      ⎝-52 13 0 ⎠
```

Para la matriz inversa:

```
(%i4) invert(A);
(%o4) ⎛ -1/13 -1/13 2/13 ⎞
      ⎝ -29/91 23/91 -1/13 ⎠
      ⎝ 4/7 -1/7 0 ⎠
(%i5) determinant(A);
(%o5) -91
```

Si se quiere, se puede calcular la inversa con el determinante afuera:

```
(%i6) invert(A),detout;
(%o6) - ⎛ 7 7 -14 ⎞
      ⎝29 -23 7 ⎠
      ⎝-52 13 0 ⎠
      91
```

Podemos comprobar que $\mathbb{A}^{-1}\mathbb{A} = \mathbb{I}$

```
(%i7) invert(A).A;
(%o7) ⎛1 0 0⎞
      ⎝0 1 0⎠
      ⎝0 0 1⎠
```

Para calcular la menor (i, j) de la matriz, es decir, eliminar la fila i y la columna j de una matriz se debe escribir:

```
(%i8) minor(A,2,2);
(%o8)  \begin{pmatrix} 1 & 3 \\ 9 & 4 \end{pmatrix}
```

Para generar una matriz triangular superior a partir de una matriz dada se debe usar el siguiente comando: **triangularize()**. El comando **echelon()** es equivalente pero normaliza a 1 el primer elemento no nulo de cada fila utilizando el método de eliminación de Gauss-Jordan.

```
(%i9) triangularize(A);
(%o9)  \begin{pmatrix} 1 & 2 & 3 \\ 0 & -13 & -23 \\ 0 & 0 & 91 \end{pmatrix}
(%i10) echelon(A);
(%o10) \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & \frac{23}{13} \\ 0 & 0 & 1 \end{pmatrix}
```

Con el paquete **nchrpl** se puede calcular la traza de una matriz

```
(%i11) load(nchrpl)$
(%i12) mattrace(A);
(%o12) 13
```

Consideremos ahora la matriz

```
(%i13) M:matrix ([1, %i, 0], [1, 2 ,-%i], [%i, 1, 2] );
(%o13) \begin{pmatrix} 1 & i & 0 \\ -1 & 2 & -i \\ i & 1 & 2 \end{pmatrix}
```

Para el exponente de matrices tenemos la opción **domxexpt** (true o false). Cuando es *true* el $\exp(M)$ se interpreta como la matriz cuyo elemento i,j es igual a $\exp(M[i,j])$. En el otro caso, $\exp(M)$ se interpreta como $(x)^M$.

```
(%i14) (1-x)^M;
(%o14) \begin{pmatrix} 1-x & (1-x)^i & 1 \\ \frac{1}{1-x} & (1-x)^2 & \frac{1}{(1-x)^i} \\ (1-x)^i & 1-x & (1-x)^2 \end{pmatrix}
(%i15) domxexpt:false$
(%i16) (1-x)^M;
(%o16) (1-x) \begin{pmatrix} 1 & i & 0 \\ -1 & 2 & -i \\ i & 1 & 2 \end{pmatrix}
```

Y en cuanto al cálculo del rango de una matriz

```
(%i17) rank(M);
(%o17) 3
```

(%i18) **kill(all)**\$

4.4.12 Ejercicios

1. Considere las matrices:

$$\mathbb{A} = \begin{pmatrix} 0 & -i & i \\ i & 0 & -i \\ -i & i & 0 \end{pmatrix}, \quad \mathbb{B} = \frac{1}{\sqrt{8}} \begin{pmatrix} \sqrt{3} & -\sqrt{2} & -\sqrt{3} \\ 1 & \sqrt{6} & -1 \\ 2 & 0 & 2 \end{pmatrix}.$$

Diga si son: (a) reales, (b) diagonales, (c) simétricas, (d) antisimétricas, (e) singulares, (f) ortogonales, (g) hermíticas, (h) antihermíticas, (i) unitarias o (j) normales.

2. Dada una matriz hermítica $\mathbb{H} = \begin{pmatrix} 10 & 3i \\ -3i & 0 \end{pmatrix}$, construya la matriz unitaria \mathbb{U} , tal que $\mathbb{U}^\dagger \mathbb{H} \mathbb{U} = \mathbb{D}$ donde \mathbb{D} es una matriz real y diagonal.

3. Utilizando el método de eliminación de Gauss-Jordan encuentre la inversa de la matriz:

$$\mathbb{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 8 & 5 \\ 9 & 5 & 4 \end{pmatrix}$$

4. Encuentre la matriz inversa de las siguientes matrices:

$$\mathbb{A} = \begin{pmatrix} 2 & 4 & 3 \\ 1 & -2 & -2 \\ -3 & 3 & 2 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad \mathbb{C} = \begin{pmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{pmatrix}.$$

Además, calcule: $\mathbb{B}^T (\mathbb{A})^{-1} \mathbb{B}$ y $(2\mathbb{A} + (\mathbb{B})^{-1} \mathbb{B}^T) \mathbb{A}^T$.

5. Considere matrices ortogonales, reales 3×3 las cuales, adicionalmente cumplen con: $\det |\mathbb{M}| = 1$.
- (a). ¿Cuántos parámetros reales son necesarios para caracterizar únicamente a este tipo de matrices?
 - (b). ¿Este tipo de matrices forman grupo bajo la multiplicación de matrices?
 - (c). Si ahora considera la matrices ortogonales reales con $\det |\mathbb{M}| = -1$ ¿Este tipo de matrices formarán grupo bajo la multiplicación? Justifique su respuesta.
6. Si \mathbb{A} y \mathbb{B} son dos matrices hermíticas que no comutan: $\mathbb{A}\mathbb{B} - \mathbb{B}\mathbb{A} = i\mathbb{C}$, entonces pruebe que \mathbb{C} es hermítica.
7. Si $\mathbb{A} = \exp(i\alpha\mathbb{B})$, $\alpha \in \mathbb{R}$. Demuestre que si \mathbb{B} es hermítica entonces \mathbb{A} es unitaria.
8. Pruebe que el producto directo de dos matrices unitarias resulta en una matriz unitaria.
9. Resuelva los problemas anteriores utilizando **Maxima**.

4.5 Sistemas de ecuaciones lineales

Una de las aplicaciones más útiles del álgebra de matrices es la resolución de sistemas de ecuaciones lineales.

$$\begin{aligned} A_1^1 x^1 + A_2^1 x^2 + \cdots + A_n^1 x^n &= c^1 \\ A_1^2 x^1 + A_2^2 x^2 + \cdots + A_n^2 x^n &= c^2 \\ &\vdots \\ A_1^m x^1 + A_2^m x^2 + \cdots + A_n^m x^n &= c^m, \end{aligned}$$

Este sistema puede ser expresado de una forma más compacta:

$$A_i^\alpha x^i = c^\alpha \quad \text{con } i = 1, 2, \dots, n \quad \text{y } \alpha = 1, 2, \dots, m,$$

por lo tanto, tendremos m ecuaciones lineales para n incógnitas (x^1, x^2, \dots, x^n). Las cantidades A_i^α resultan ser las componentes de la matriz de los coeficientes.

Si todos los c^i son cero el sistema se denomina homogéneo, en caso contrario inhomogéneo. Puede resultar que el conjunto de las incógnitas $\{x^i\}$ represente una solución, infinitas soluciones o que simplemente no exista una solución para el sistema.

Este problema puede ser pensado como un problema de un operador \mathbb{A} en el espacio vectorial de transformaciones lineales $\mathcal{L}(\mathbf{V}, \mathbf{W})$, donde $\dim(\mathbf{V}) = n$ y $\dim(\mathbf{W}) = m$, con las c^α las componentes del vector transformado:

$$|c\rangle = \mathbb{A}|x\rangle \rightarrow c^\alpha = A_i^\alpha x^i.$$

El operador \mathbb{A} aplicará todo vector de \mathbf{V} en algún subespacio (o todo el espacio) de \mathbf{W} . A este subespacio se le denomina el rango o recorrido de \mathbb{A} y su dimensión es igual al rango de la matriz \mathbb{A} . Si \mathbb{A} es no singular, entonces existe algún subespacio de \mathbf{V} que es aplicado al vector cero de \mathbf{W} , es decir, se cumple que $\mathbb{A}|x'\rangle = |0\rangle$, donde al conjunto de vectores $\{|x'\rangle\}$ se le llama el espacio nulo de \mathbb{A} .

A continuación, mostraremos algunos métodos para la resolución de sistemas de ecuaciones lineales.

4.5.1 Eliminación de Gauss-Jordan

Uno de los métodos más utilizados es el de la eliminación de *Gauss-Jordan*, el cual se basa en el intercambio de ecuaciones y la multiplicación apropiada e inteligente por constantes y resta de ecuaciones. La idea es construir una matriz triangular superior para poder luego despejar desde abajo.

Veamos como se aplica el método para resolver el sistema de ecuaciones. Primeramente escribimos el siguiente arreglo:

$$\begin{array}{r} a \\ b \\ c \end{array} \left[\begin{array}{ccc|c} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{array} \right],$$

entonces para eliminar x de la fila c (o la ecuación c) sumamos la fila a con la c , $a + c$ y esta nueva ecuación será la nueva c

$$\begin{array}{r} a \\ b \\ c' \end{array} \left[\begin{array}{ccc|c} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ 0 & 6 & -2 & 6 \end{array} \right],$$

ahora $-2a + b$ será la nueva b

$$\begin{array}{l} a \\ b' \\ c' \end{array} \left[\begin{array}{ccc|c} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 6 & -2 & 6 \end{array} \right],$$

finalmente $3b' + c'$

$$\begin{array}{l} a \\ b' \\ c'' \end{array} \left[\begin{array}{ccc|c} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 0 & -5 & -15 \end{array} \right].$$

Este sistema es equivalente al primer sistema de ecuaciones, el lector puede verificar que poseen el mismo determinante. Por lo tanto, la solución emerge rápidamente:

$$-5z = -15 \rightarrow z = 3, \quad -2y - z = -7 \rightarrow -2y - 3 = -7 \rightarrow y = 2, \quad 2x + 3(2) - 3 = 5 \rightarrow x = 1.$$

Es bueno recalcar que los sistemas de ecuaciones lineales no necesariamente tienen solución y a veces tienen más de una solución.

4.5.2 El método de la matriz inversa

El análisis matricial nos puede ayudar a investigar sobre las posibles soluciones de un sistema de ecuaciones lineales, como vimos, en notación de matrices el sistema se puede escribir de la manera siguiente:

$$\left(\begin{array}{cccc} A_1^1 & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 & \cdots & A_n^2 \\ \vdots & \vdots & & \vdots \\ A_1^m & A_2^m & \cdots & A_n^m \end{array} \right) \left(\begin{array}{c} x^1 \\ x^2 \\ \vdots \\ x^n \end{array} \right) = \left(\begin{array}{c} c^1 \\ c^2 \\ \vdots \\ c^m \end{array} \right) \Rightarrow \mathbb{A}\mathbf{x} = \mathbb{C}.$$

Podemos considerar a cada columna de la matriz \mathbb{A} como un vector, entonces el *rango* r de la matriz \mathbb{A} será igual al número de vectores linealmente independientes, este número es también la dimensión del espacio vectorial que expanden estos vectores.

Con respecto a las posibles soluciones del sistema podemos decir:

1. Si \mathbb{C} pertenece al rango de \mathbb{A} y además $r = n$, entonces todos los vectores del conjunto son linealmente independientes y el sistema tendrá como única solución al conjunto $\{x^1, x^2, \dots, x^n\}$.
2. Si \mathbb{C} pertenece al rango de \mathbb{A} y además $r < n$, entonces únicamente r vectores serán linealmente independientes. Esto significa que podremos escoger los coeficientes de los $n - r$ vectores de manera arbitraria sin dejar de satisfacer el sistema. Por tanto, existirá un número infinito de soluciones, que expanden un espacio vectorial de dimensión $n - r$.
3. La otra posibilidad es que el sistema no tenga solución, en este caso \mathbb{C} no pertenece al rango de \mathbb{A} .

Cuando el sistema es homogéneo, $\mathbb{C} = \mathbb{O}$, claramente existe una solución que viene a ser la trivial: $\{x^1 = x^2 = \dots = x^n = 0\}$, además, si $r = n$ ésta será la única solución. Es bueno anotar que si hay menos ecuaciones que incógnitas ($m < n$) entonces automáticamente $r < n$, por lo tanto un conjunto de ecuaciones lineales homogéneas con menos ecuaciones que incógnitas siempre tiene una infinidad de soluciones.

Debemos considerar el importante caso cuando $m = n$, es decir, la matriz \mathbb{A} es cuadrada y existe igual número de ecuaciones como de incógnitas. Al ser cuadrada la matriz se tiene que la condición $r = n$ implica que la matriz es no singular ($\det |\mathbb{A}| \neq 0$). El caso $r < n$ se corresponde a que $\det |\mathbb{A}| = 0$.

Para resolver el sistema de ecuaciones consideraremos lo siguiente: dada la matriz \mathbb{A} , $n \times n$ no singular,

entonces \mathbb{A} tiene una matriz inversa, tal que:

$$\mathbb{A}\mathbb{X} = \mathbb{C} \Rightarrow \mathbb{X} = \mathbb{A}^{-1}\mathbb{C}.$$

4.5.3 Factorización LU

El método de la matriz inversa es relativamente simple, pero laborioso cuando se trata de resolver sistemas de ecuaciones muy grandes. La factorización o descomposición LU (Lower-Upper) es una técnica, existen varias de este tipo, para factorizar una matriz como el producto de una matriz triangular inferior y una matriz triangular superior. Si \mathbb{A} es un matriz cuadrada no singular, entonces:

$$\mathbb{A}\mathbb{X} = \mathbb{C} \Rightarrow \mathbb{A} = \mathbb{L}\mathbb{U},$$

las matrices \mathbb{L} y \mathbb{U} serán únicas.

Este método es básicamente una versión de el método de Gauss-Jordan, pero es más eficiente a la hora de ser implementado en algoritmos computacionales tanto algebraicos como numéricos.

Para matrices 3×3 el método se implementa de la siguiente manera.

$$\mathbb{A} = \begin{pmatrix} 1 & 0 & 0 \\ L_1^2 & 1 & 0 \\ L_1^3 & L_2^3 & 1 \end{pmatrix} \begin{pmatrix} U_1^1 & U_2^1 & U_3^1 \\ 0 & U_2^2 & U_3^2 \\ 0 & 0 & U_3^3 \end{pmatrix} = \begin{pmatrix} U_1^1 & U_2^1 & U_3^1 \\ L_1^2 U_1^1 & L_1^2 U_2^1 + U_2^2 & L_1^2 U_3^1 + U_3^2 \\ L_1^3 U_1^1 & L_1^3 U_2^1 + L_2^3 U_2^2 & L_1^3 U_3^1 + L_2^3 U_3^2 + U_3^3 \end{pmatrix},$$

las nueve incógnitas se obtienen igualando con las componentes conocidas de la matriz \mathbb{A} . Con \mathbb{L} y \mathbb{U} determinadas tenemos entonces que

$$\mathbb{L}\mathbb{U}\mathbb{X} = \mathbb{C},$$

por lo tanto, se debe realizar los cálculos en dos partes:

$$\text{primero } \mathbb{L}\mathbb{Y} = \mathbb{C} \text{ y luego } \mathbb{U}\mathbb{X} = \mathbb{Y}.$$

Con la matrices \mathbb{L} y \mathbb{U} se puede calcular el determinante de la matriz \mathbb{A} de una manera más directa, es fácil ver que:

$$\det |\mathbb{A}| = \det |\mathbb{L}| \det |\mathbb{U}| = \det |\mathbb{U}| = \prod_{i=1}^n U_i^i.$$

4.5.4 Método de Cramer

Un método alternativo para resolver sistemas de ecuaciones lineales es la conocida *regla de Cramer*. El funcionamiento del método lo podemos ilustrar con un sistema de tres ecuaciones y tres incógnitas:

$$\begin{aligned} A_1^1 x^1 + A_2^1 x^2 + A_3^1 x^3 &= c^1 \\ A_1^2 x^1 + A_2^2 x^2 + A_3^2 x^3 &= c^2 \\ A_1^3 x^1 + A_2^3 x^2 + A_3^3 x^3 &= c^3, \end{aligned}$$

Como mencionamos anteriormente, el determinante de una matriz, $\det |\mathbb{A}| = |\mathbb{A}|$,

$$|\mathbb{A}| = \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix},$$

no cambia bajo la operación, por ejemplo, de sumarle a la primera columna las cantidades:

$$|\mathbb{A}| = \begin{vmatrix} A_1^1 & A_2^1 & A_3^1 \\ A_1^2 & A_2^2 & A_3^2 \\ A_1^3 & A_2^3 & A_3^3 \end{vmatrix} = \begin{vmatrix} A_1^1 + (x^2/x^1)A_2^1 + (x^3/x^1)A_3^1 & A_2^1 & A_3^1 \\ A_1^2 + (x^2/x^1)A_2^2 + (x^3/x^1)A_3^2 & A_2^2 & A_3^2 \\ A_1^3 + (x^2/x^1)A_2^3 + (x^3/x^1)A_3^3 & A_2^3 & A_3^3 \end{vmatrix}$$

lo cual es igual a:

$$|\mathbb{A}| = \frac{1}{x^1} \begin{vmatrix} c^1 & A_2^1 & A_3^1 \\ c^2 & A_2^2 & A_3^2 \\ c^3 & A_2^3 & A_3^3 \end{vmatrix} \equiv \frac{1}{x^1} |\Delta_1|.$$

Se puede hacer lo mismo con las otras dos columnas para obtener:

$$x^1 = \frac{|\Delta_1|}{|\mathbb{A}|}, \quad x^2 = \frac{|\Delta_2|}{|\mathbb{A}|}, \quad x^3 = \frac{|\Delta_3|}{|\mathbb{A}|}.$$

Donde los determinantes, $|\Delta_n|$, son los determinantes de Cramer:

$$|\Delta_1| = \begin{vmatrix} c^1 & A_2^1 & A_3^1 \\ c^2 & A_2^2 & A_3^2 \\ c^3 & A_2^3 & A_3^3 \end{vmatrix}, \quad |\Delta_2| = \begin{vmatrix} A_1^1 & c^1 & A_3^1 \\ A_1^2 & c^2 & A_3^2 \\ A_1^3 & c^3 & A_3^3 \end{vmatrix}, \quad |\Delta_3| = \begin{vmatrix} A_1^1 & A_2^1 & c^1 \\ A_1^2 & A_2^2 & c^2 \\ A_1^3 & A_2^3 & c^3 \end{vmatrix}.$$

Se puede demostrar que si $|\mathbb{A}| \neq 0$, entonces la solución así obtenida es única.

4.5.5 Ejemplos

- Resolver el sistema:

$$2x^1 + 4x^2 + 3x^3 = 4, \quad x^1 - 2x^2 - 2x^3 = 0, \quad -3x^1 + 3x^2 + 2x^3 = -7.$$

Utilizando el método de la matriz inversa.

Este sistema de ecuaciones se puede escribir como:

$$\mathbb{A}\mathbb{X} = \mathbb{C} \Rightarrow \begin{pmatrix} 2 & 4 & 3 \\ 1 & -2 & -2 \\ -3 & 3 & 2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -7 \end{pmatrix},$$

donde \mathbb{A} es la matriz de los coeficientes.

Luego de calcular la matriz inversa de \mathbb{A} entonces resulta que

$$\mathbb{X} = \mathbb{A}^{-1}\mathbb{C} \Rightarrow \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \frac{1}{11} \begin{pmatrix} 2 & 1 & -2 \\ 4 & 13 & 7 \\ -3 & -18 & -8 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \\ -7 \end{pmatrix} = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}.$$

Existe entonces una única solución al sistema, que es: $\{x^1 = 2, x^2 = -3, x^3 = 4\}$.

- Resolver el mismo sistema de ecuaciones anteriormente estudiado:

$$2x^1 + 4x^2 + 3x^3 = 4, \quad x^1 - 2x^2 - 2x^3 = 0, \quad -3x^1 + 3x^2 + 2x^3 = -7.$$

Pero utilizando el método de factorización $\mathbb{L}\mathbb{U}$.

La matriz de los coeficientes es:

$$\mathbb{A} = \begin{pmatrix} 2 & 4 & 3 \\ 1 & -2 & -2 \\ -3 & 3 & 2 \end{pmatrix}$$

- Cálculo de \mathbb{L} y \mathbb{U}

$$\begin{pmatrix} 2 & 4 & 3 \\ 1 & -2 & -2 \\ -3 & 3 & 2 \end{pmatrix} = \begin{pmatrix} U_1^1 & U_2^1 & U_3^1 \\ L_1^2 U_1^1 & L_1^2 U_2^1 + U_2^2 & L_1^2 U_3^1 + U_3^2 \\ L_1^3 U_1^1 & L_1^3 U_2^1 + L_2^3 U_2^2 & L_1^3 U_3^1 + L_2^3 U_3^2 + U_3^3 \end{pmatrix}$$

Con los valores de inicio: $U_1^1 = 2$, $U_2^1 = 4$, $U_3^1 = 3$ se procede a resolver las ecuaciones:

$$\begin{array}{ll} L_1^2 U_1^1 = 1 & L_1^3 U_1^1 = -3 \\ L_1^2 U_2^1 + U_2^2 = -2 & L_1^2 U_3^1 + U_3^2 = -2 \\ L_1^3 U_2^1 + L_2^3 U_2^2 = 3 & L_1^3 U_3^1 + L_2^3 U_3^2 + U_3^3 = 2 \end{array}$$

por lo tanto

$$\mathbb{A} = \mathbb{L}\mathbb{U} = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ -3/2 & -9/4 & 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 3 \\ 0 & -4 & -7/2 \\ 0 & 0 & -11/8 \end{pmatrix}$$

- Resolvemos primero $\mathbb{L}\mathbb{Y} = \mathbb{C}$, lo cual es bastante fácil

$$\begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1 & 0 \\ -3/2 & -9/4 & 1 \end{pmatrix} \begin{pmatrix} y^1 \\ y^2 \\ y^3 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -7 \end{pmatrix} \Rightarrow \begin{cases} y^1 = 4 \\ y^2 = -2 \\ y^3 = -\frac{11}{2} \end{cases}$$

- Con este resultado resolvemos: $\mathbb{U}\mathbb{X} = \mathbb{Y}$, y obtenemos la solución

$$\begin{pmatrix} 2 & 4 & 3 \\ 0 & -4 & -7/2 \\ 0 & 0 & -11/8 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \\ -11/2 \end{pmatrix} \Rightarrow \begin{cases} x^1 = 2 \\ x^2 = -3 \\ x^3 = 4 \end{cases}$$

3. Resolver, nuevamente, el sistema de ecuaciones:

$$2x^1 + 4x^2 + 3x^3 = 4, \quad x^1 - 2x^2 - 2x^3 = 0, \quad -3x^1 + 3x^2 + 2x^3 = -7.$$

Utilizando ahora los determinantes de Cramer.

Esto es:

$$\mathbb{A} = \mathbb{X}\mathbb{C} \Rightarrow \begin{pmatrix} 2 & 4 & 3 \\ 1 & -2 & -2 \\ -3 & 3 & 2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ -7 \end{pmatrix} \Rightarrow |\mathbb{A}| = 11.$$

Los diferentes determinantes de Cramer son:

$$|\Delta_1| = \begin{vmatrix} 4 & 4 & 3 \\ 0 & -2 & -2 \\ -7 & 3 & 2 \end{vmatrix} = 22, \quad |\Delta_2| = \begin{vmatrix} 2 & 4 & 3 \\ 1 & 0 & -2 \\ -3 & -7 & 2 \end{vmatrix} = -33, \quad |\Delta_3| = \begin{vmatrix} 2 & 4 & 4 \\ 1 & -2 & 0 \\ -3 & 3 & -7 \end{vmatrix} = 44.$$

de manera que

$$x^1 = \frac{22}{11} = 2, \quad x^2 = -\frac{33}{11} = -3, \quad x^3 = \frac{44}{11} = 4.$$

4.5.6 Practicando con Maxima

1. Resolvemos el siguiente sistema de ecuaciones lineales por el método de Gauss-Jordan

$$\begin{aligned} 2x^1 + 4x^2 + 3x^3 &= 4 \\ x^1 - 2x^2 - 2x^3 &= 0 \\ -3x^1 + 3x^2 + 2x^3 &= -7 \end{aligned}$$

Introducimos el sistema de ecuaciones:

(%i1) `ecu1:2*x1+4*x2+3*x3=4; ecu2:x1-2*x2-2*x3=0; ecu3:-3*x1+3*x2+2*x3=-7;`
 (%o1) $3x_3 + 4x_2 + 2x_1 = 4$

(%o2) $-2x_3 - 2x_2 + x_1 = 0$

(%o3) $2x_3 + 3x_2 - 3x_1 = -7$

Construimos la matriz de coeficientes y la denominaremos A

(%i4) `A:coefmatrix([ecu1,ecu2,ecu3],[x1,x2,x3]);`
 (%o4)
$$\begin{pmatrix} 2 & 4 & 3 \\ 1 & -2 & -2 \\ -3 & 3 & 2 \end{pmatrix}$$

Le agregamos la columna con los elementos independientes

(%i5) `M:addcol(A,[4,0,-7]);`
 (%o5)
$$\begin{pmatrix} 2 & 4 & -3 & 4 \\ 1 & -2 & -2 & 0 \\ -3 & 3 & 2 & -7 \end{pmatrix}$$

Reducimos la matriz M utilizando el algoritmo de Gauss-Jordan.

(%i6) `echelon(M);`
 (%o6)
$$\begin{pmatrix} 1 & -2 & -2 & 0 \\ 0 & 1 & \frac{7}{8} & \frac{1}{2} \\ 0 & 0 & 1 & 4 \end{pmatrix}$$

De la última fila es fácil determinar que $x_3 = 4$. Los otros dos valores que resultan son $x_2 = -3$ y $x_1 = 2$.

Por supuesto podemos recurrir al cálculo directo:

(%i7) `linsolve([ecu1,ecu2,ecu3],[x1,x2,x3]);`
 (%o7) $[x_1 = 2, x_2 = -3, x_3 = 4]$

2. Resolvamos nuevamente el mismo sistema

$$\begin{aligned} 2x^1 + 4x^2 + 3x^3 &= 4 \\ x^1 - 2x^2 - 2x^3 &= 0 \\ -3x^1 + 3x^2 + 2x^3 &= -7 \end{aligned}$$

pero ahora por el método de la matriz inversa.

(%i8) `ecu1:2*x1+4*x2+3*x3=4; ecu2:x1-2*x2-2*x3=0; ecu3:-3*x1+3*x2+2*x3=-7;`
 (%o8) $3x_3 + 4x_2 + 2x_1 = 4$

(%o9) $-2x_3 - 2x_2 + x_1 = 0$

$$(\%o10) 2x_3 + 3x_2 - 3x_1 = -7$$

La matriz de cofactores:

$$(\%i11) A:\text{coefmatrix}([\text{ecu1}, \text{ecu2}, \text{ecu3}], [x1, x2, x3]);$$

$$(\%o12) \begin{pmatrix} 2 & 4 & 3 \\ 1 & -2 & -2 \\ -3 & 3 & 2 \end{pmatrix}$$

La matriz inversa de A la denominaremos $Ainv$, y la calcularemos con el determinante afuera:

$$(\%i13) Ainv:\text{invert}(A), \text{detout};$$

$$(\%o13) \frac{\begin{pmatrix} 2 & 1 & -2 \\ 4 & 13 & 7 \\ -3 & -18 & -8 \end{pmatrix}}{11}$$

La matriz con los términos inhomogéneos lo llamaremos C :

$$(\%i14) C:\text{matrix}([4, 0, -7]);$$

$$(\%o14) \begin{pmatrix} 4 & 0 & -7 \end{pmatrix}$$

Para que finalmente podamos hacer la siguiente multiplicación de matrices:

$$(\%i15) X:Ainv.C;$$

$$(\%o15) \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}$$

Por lo tanto: $x_1 = 2$, $x_2 = -3$, $x_3 = 4$.

3. Vamos ahora a resolver el mismo sistema pero usando la técnica de factorización LU.

Le podemos pedir al programa que factorize la matriz A por medio de la función **lu_factor(A)**, consultar el manual de **Maxima** para más detalles.

$$(\%i16) U:\text{lu_factor}(A)$$$

$$(\%i17) F:\text{get_lu_factors}(U);$$

$$(\%o17) \left[\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ -\frac{3}{2} & -\frac{9}{4} & 1 \end{pmatrix}, \begin{pmatrix} 2 & 4 & 3 \\ 0 & -4 & -\frac{7}{2} \\ 0 & 0 & -\frac{11}{8} \end{pmatrix} \right]$$

Por lo tanto, aquí tendremos que: $\mathbb{L} = F[2]$ y $\mathbb{U} = F[3]$.

Podemos resolver entonces la primera parte, $\mathbb{L}\mathbb{Y} = \mathbb{C}$:

$$(\%i18) Y:\text{invert}(F[2]).C;$$

$$(\%o18) \begin{pmatrix} 4 \\ -2 \\ -\frac{11}{2} \end{pmatrix}$$

Y ahora $\mathbb{U}\mathbb{X} = \mathbb{Y}$:

$$(\%i19) X:\text{invert}(F[3]).Y;$$

$$(\%o19) \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}$$

Por lo que resulta que: $x_1 = 2$, $x_2 = -3$, $x_3 = 4$.

Notemos que:

$$(\%i20) \text{determinant}(A) = \text{determinant}(F[2]) * \text{determinant}(F[3]);$$

$$(\%o20) 11 = 11$$

4. Utilicemos ahora los determinantes de Cramer. Para tal fin, construiremos las matrices eliminando las columnas correspondientes.

$$(\%i21) A1:\text{col}(A,1)\$ A2:\text{col}(A,2)\$ A3:\text{col}(A,3)\$ C:\text{transpose}(\text{matrix}([4,0,-7]))\$$$

Las tres matrices son entonces:

$$(\%i22) D1:\text{addcol}(C,A2,A3); D2:\text{addcol}(A1,C,A3); D3:\text{addcol}(A1,A2,C);$$

$$(\%o22) \begin{pmatrix} 4 & 4 & 3 \\ 0 & -2 & -2 \\ -7 & 3 & 2 \end{pmatrix}$$

$$(\%o23) \begin{pmatrix} 2 & 4 & 3 \\ 1 & 0 & -2 \\ -3 & -7 & 2 \end{pmatrix}$$

$$(\%o24) \begin{pmatrix} 2 & 4 & 4 \\ 1 & -2 & 0 \\ -3 & 3 & -7 \end{pmatrix}$$

Por lo tanto, los determinantes generan la solución a problema:

$$(\%i25) x1:\text{determinant}(D1)/\text{determinant}(A); x2:\text{determinant}(D2)/\text{determinant}(A);$$

$$\quad x3:\text{determinant}(D3)/\text{determinant}(A);$$

$$(\%o25) 2$$

$$(\%o26) -3$$

$$(\%o27) 4$$

5. Existen algunas consideraciones a tomar en cuenta.

- El sistema puede contener más incógnitas que ecuaciones, por ejemplo:

$$(\%i28) \text{ecu1}:x+y+z=1; \text{ecu2}:x+2*y+z=0;$$

$$(\%o28) z + y + x = 1$$

$$(\%o29) z + 2y + x = 0$$

$$(\%i30) \text{linsolve}([\text{ecu1}, \text{ecu2}, \text{ecu3}], [x, y, z]);$$

$$(\%o30) [x = 2 - \%r_1, y = -1, z = \%r_1]$$

El programa nos está señalando que las soluciones, en este caso infinitas, dependerán de un parámetro que aquí es indicado con el símbolo $\%r_1$. Una solución puede ser entonces la siguiente: $\{x = 2, y = -1, z = 0\}$.

- El sistema tiene más ecuaciones que incógnitas, por ejemplo:

$$(\%i31) \text{ecu1}:x+y=5; \text{ecu2}:x+2*y=8; \text{ecu3}:3*x+y=9;$$

(%o31) $y + x = 5$
 (%o32) $2y + x = 8$
 (%o33) $y + 3x = 9$

(%i34) `linsolve([ecu1,ecu2,ecu3],[x,y]);`
 solve: dependent equations eliminated: (3)

(%o34) $[x = 2, y = 3]$

Aquí, **Maxima** encuentra una solución al eliminar una de las ecuaciones dependientes.

- El sistema no tiene solución, por ejemplo:

(%i35) `ecu1:x+y=5; ecu2:x+2*y=8; ecu3:3*x+y=10;`
 (%o35) $y + x = 5$
 (%o36) $2y + x = 8$
 (%o37) $y + 3x = 10$

(%i38) `linsolve([ecu1,ecu2,ecu3],[x,y]);`
 (%o38) $[]$

4.5.7 Ejercicios

1. Resuelva, utilizando dos de los métodos anteriormente vistos, el siguiente sistema de ecuaciones:

$$\begin{aligned} 2x + 3y + z &= 11 \\ x + y + z &= 6 \\ 5x - y + 10z &= 34 \end{aligned}$$

2. Resuelva, utilizando el método de la matriz inversa, el siguiente sistema de ecuaciones:

$$\begin{aligned} x^1 + 2x^2 + 3x^3 &= 1 \\ 3x^1 + 4x^2 + 5x^3 &= 2 \\ x^1 + 3x^2 + 4x^3 &= 3 \end{aligned}$$

3. Demuestre que el siguiente sistema sólo tiene solución si $\eta = 1$ o $\eta = 2$

$$\begin{aligned} x + y + z &= 1 \\ x + 2y + 4z &= \eta \\ x + 4y + 10z &= \eta^2 \end{aligned}$$

4. Encuentre las condiciones sobre η para que al resolver el sistema

$$\begin{aligned} x^1 + \eta x^2 &= 1 \\ x^1 - x^2 + 3x^3 &= -1 \\ 2x^1 - 2x^2 + \eta x^3 &= -2 \end{aligned}$$

- (a). tenga una solución
- (b). no tenga solución
- (c). tenga infinitas soluciones

Encuentre todas las soluciones que puedan existir.

5. Encuentre la factorización LU de las siguientes matrices:

$$\mathbb{A} = \begin{pmatrix} 3 & 6 & 9 \\ 1 & 0 & 5 \\ 2 & -2 & 16 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 2 & -3 & 1 & 3 \\ 1 & 4 & -3 & -3 \\ 5 & 3 & -1 & -1 \\ 3 & -6 & -3 & 1 \end{pmatrix}.$$

- Resuelva $\mathbb{A}\mathbb{X} = \mathbb{C}$, cuando: $\mathbb{C} = (21 \ 9 \ 28)^T$ y $\mathbb{C} = (21 \ 7 \ 22)^T$.
- Resuelva $\mathbb{B}\mathbb{X} = \mathbb{C}$ cuando: $\mathbb{C} = (-4 \ 1 \ 8 \ -5)^T$ y $\mathbb{C} = (-10 \ 0 \ -3 \ -24)^T$.

6. Utilizando la regla de Cramer resuelva el siguiente sistema:

$$\begin{aligned} 13x^1 + 22x^2 - 13x^3 &= 4 \\ 10x^1 - 8x^2 - 10x^3 &= 44 \\ 9x^1 - 18x^2 - 9x^3 &= 72. \end{aligned}$$

7. Utilizando cualquiera de los métodos anteriores resuelva:

$$\begin{aligned} x^1 + x^2 + x^3 + x^4 + x^5 &= 7 \\ 3x^1 + 2x^2 + x^3 + x^4 - 3x^5 &= -2 \\ x^2 + 2x^3 + 2x^4 + 6x^5 &= 23 \\ 5x^1 + 4x^2 + 3x^3 + 3x^4 - x^5 &= 12. \end{aligned}$$

8. Determine el rango de la siguiente matriz:

$$\begin{pmatrix} 1 & -2 & 3 & -1 & -1 & -2 \\ 2 & -1 & 1 & 0 & -2 & -2 \\ -2 & -5 & 8 & -4 & 3 & -1 \\ 6 & 0 & -1 & 2 & -7 & -5 \\ -1 & -1 & 1 & -1 & 2 & 1 \end{pmatrix}$$

9. ¿Cuál es la condición para que las siguientes rectas se intercepten en un mismo punto?

$$a_1x + b_1y + c_1 = 0, \quad a_2x + b_2y + c_2 = 0, \quad a_3x + b_3y + c_3 = 0,$$

10. Resuelva los problemas anteriores utilizando **Maxima**.

4.6 Autovectores y autovalores

La ecuación de Schrödinger independiente del tiempo de la Mecánica Cuántica es una ecuación, que escrita en la forma de operadores, tiene la siguiente forma:

$$\mathbb{H}|\Psi_E\rangle = E|\Psi_E\rangle,$$

donde \mathbb{H} es un operador diferencial de segundo orden llamado el hamiltoniano y Ψ_E la función de onda, en este caso representada por $|\Psi_E\rangle$, *un estado propio* de \mathbb{H} que corresponde a *un valor propio* de la energía E .

Por tratarse de un espacio de Hilbert, es decir, un espacio vectorial con un producto interno bien definido, resulta de interés buscar soluciones en lo que se denomina el estado ligado de la ecuación de Schrödinger. Esto significa buscar los $|\Psi_E\rangle$ en el espacio de las funciones de cuadrado integrable introduciendo una base unidimensional para representar Ψ_E y una matriz para representar \mathbb{H} , en otras palabras: buscar una representación matricial para la ecuación de Schrödinger.

Los vectores propios, $|\Psi_E\rangle$, autovectores o “eigenvectores” de un operador lineal \mathbb{H} son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo

que no cambian su dirección. Este escalar E recibe el nombre de valor propio, autovalor, valor característico o “eigenvalor”. De manera que una transformación queda completamente determinada por sus vectores propios y valores propios. Un espacio propio, autoespacio, “eigenespacio” o subespacio fundamental asociado al valor propio E es el conjunto de vectores propios con un valor propio común.

De manera general, llamaremos a $|\psi\rangle$ un autovector del operador \mathbb{A} si se cumple que

$$\mathbb{A}|\psi\rangle = \lambda|\psi\rangle , \quad (4.21)$$

en este caso λ (que, en general será un número complejo) se denomina el autovalor correspondiente al autovector $|\psi\rangle$. La ecuación (4.21) es conocida en la literatura como la ecuación de autovalores y se cumple para algunos valores particulares de los autovalores λ y, obviamente esto conduce a algunos vectores $|\psi\rangle$, los autovectores. Al conjunto de los autovalores se le denomina el espectro del operador \mathbb{A} .

Nótese que si $|\psi\rangle$ es autovector de \mathbb{A} para un determinado autovalor λ entonces $|\phi\rangle = \alpha|\psi\rangle$ (un vector proporcional a $|\psi\rangle$, con α un número complejo) también es un autovector para el mismo autovalor. Esto representa una incómoda ambigüedad: dos autovectores que corresponden al mismo autovalor. Un intento de eliminarla es **siempre** considerar vectores $|\psi\rangle$ normalizados, i.e. $\langle\psi|\psi\rangle = 1$. Sin embargo, no deja de ser un intento que no elimina la ambigüedad del todo porque siempre queda el ángulo de fase arbitrario. Esto es, el vector $e^{i\theta}|\psi\rangle$, con θ un número real arbitrario, tiene la misma norma del vector $|\psi\rangle$. Sin embargo esta arbitrariedad es inofensiva y en Mecánica Cuántica las predicciones obtenidas con $|\psi\rangle$ son las mismas que con $e^{i\theta}|\psi\rangle$.

4.6.1 Autovalores, autovectores e independencia lineal

La relación que existe entre los autovalores e independencia lineal de los autovectores asociados con cada autovalor es una de las herramientas más útiles que disponemos. Expondremos a continuación tres teoremas que establecen esta relación.

Teorema 1: Sean $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, \dots, |\psi_k\rangle\}$ autovectores del operador $\mathbb{A} : \mathbf{V}^m \rightarrow \mathbf{V}^n$. Supongamos que existen k autovalores: $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$, distintos correspondientes a cada uno de los autovectores $|\psi_j\rangle$, entonces los $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_k\rangle\}$ son linealmente independientes.

Demostración La demostración de este teorema es por inducción y resulta elegante y sencilla.

- Primeramente demostramos para $j = 1$.

Obvio que el resultado se cumple y es trivial para el caso $k = 1$ (un autovector $|\psi_1\rangle$ que corresponde a un autovalor λ_1 es obvia y trivialmente linealmente independiente).

- Seguidamente supondremos que se cumple para $j = k - 1$.

Si existen $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, \dots, |\psi_{k-1}\rangle\}$ autovectores de \mathbb{A} correspondientes a $\{\lambda_1, \lambda_2, \dots, \lambda_{k-1}\}$ entonces los $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, \dots, |\psi_{k-1}\rangle\}$ son linealmente independientes.

- Ahora lo probaremos para $j = k$.

Por lo cual si tenemos k autovectores $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, \dots, |\psi_k\rangle\}$, podremos construir una combinación lineal con ellos, y si esa combinación lineal se anula serán linealmente independientes.

$$c^j|\psi_j\rangle = 0 \quad \text{con } j = 1, 2, \dots, k,$$

al aplicar el operador \mathbb{A} a esa combinación lineal, obtenemos: $c^j\mathbb{A}|\psi_j\rangle = 0 \Rightarrow c^j\lambda_j|\psi_j\rangle = 0$, multiplicando por λ_k y restando miembro a miembro resulta:

$$c^j(\lambda_j - \lambda_k)|\psi_j\rangle = 0 \quad \text{con } j = 1, 2, \dots, k - 1,$$

(nótese que el último índice es $k-1$) pero, dado que los $k-1$ vectores $|\psi_j\rangle$ son linealmente independientes, entonces tendremos $k-1$ ecuaciones $c^j(\lambda_j - \lambda_k) = 0$, una para cada $j = 1, 2, \dots, k-1$. Dado que $\lambda_j \neq \lambda_k$ necesariamente llegamos a que $c^j = 0$ para $j = 1, 2, \dots, k-1$ y dado que:

$$c^j |\psi_j\rangle = 0 \quad \text{con } j = 1, 2, \dots, k \Rightarrow c^j \neq 0,$$

con lo cual si $c^j |\psi_j\rangle = 0 \Rightarrow c^j = 0$ con $j = 1, 2, \dots, k$, y los $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, \dots, |\psi_k\rangle\}$ son linealmente independientes y queda demostrado el teorema. ◀

Es importante acotar que este teorema nos recalca que si $\mathbb{A} : \mathbf{V}^m \rightarrow \mathbf{V}^n$ y \mathbb{A} tiene $k \leq n$ autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ entonces existirán, **cuando menos**, $k \leq n$ autovectores $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_k\rangle\}$ linealmente independientes, uno para cada autovalor. Hacemos énfasis en ese **cuando menos**, $k \leq n$ autovectores linealmente independientes, porque significa que el espacio \mathbf{V}^m no podrá ser expandido por los autovectores de \mathbb{A} . Ese punto lo analizaremos en la sección 4.24.

El inverso de este teorema no se cumple. Esto es, si $\mathbb{A} : \mathbf{V}^m \rightarrow \mathbf{V}^n$ tiene $\{|\psi_1\rangle, |\psi_2\rangle, |\psi_3\rangle, \dots, |\psi_n\rangle\}$ autovectores linealmente independientes, no se puede concluir que existan n autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ distintos correspondientes a cada uno de los autovectores $|\psi_j\rangle$.

El teorema anterior lo complementa el siguiente que lo presentaremos sin demostración.

Teorema 2: Cualquier operador lineal $\mathbb{A} : \mathbf{V}^n \rightarrow \mathbf{V}^n$ tendrá un máximo de n autovalores distintos. Adicionalmente, si \mathbb{A} tiene *precisamente* n autovalores: $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, entonces los n autovectores, $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$ (uno para cada autovalor), forman una base para \mathbf{V}^n y la representación matricial, en esa base, del operador será diagonal

$$\langle \psi^i | \mathbb{A} | \psi_j \rangle = A_j^i = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n). \quad (4.22)$$

Es muy importante recalcar el significado de este teorema. Si la dimensión del espacio vectorial $\dim(\mathbf{V}^n) = n$ y por consiguiente la representación matricial será $n \times n$, entonces \mathbb{A} tendrá un MÁXIMO de n autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. Por otro lado, si tenemos n autovalores distintos $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ TENDREMOS n autovectores LINEALMENTE INDEPENDIENTES y la representación matricial será diagonal con los autovalores en la diagonal: $\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

Una versión complementaria al teorema anterior se obtiene si consideramos que además existe una base ortogonal, $\{|e_i\rangle\}$, para \mathbf{V} . Por lo tanto, la representación matricial de la expresión de la ecuación (4.22) es la siguiente:

$$\mathbb{A} |\psi\rangle = \lambda |\psi\rangle \xrightarrow{\{|e_i\rangle\}} \langle e^i | \mathbb{A} | e_j \rangle \langle e^j | \psi \rangle = \langle e^i | \lambda | \psi \rangle = \lambda \langle e^i | \psi \rangle \Rightarrow A_j^i c^j = \lambda c^i, \quad (4.23)$$

claramente, la base ortonormal, $\{|e_i\rangle\}$, genera una representación diagonal de \mathbb{A} , entonces $A_j^i \propto \delta_j^i$, y esto lo podemos resumir en el siguiente teorema que presentaremos sin demostración para concluir esta sección.

Teorema 3: Dado un operador lineal $\mathbb{A} : \mathbf{V}^n \rightarrow \mathbf{V}^n$, si la representación matricial de \mathbb{A} es diagonal, $\langle e^i | \mathbb{A} | e_j \rangle = A_j^i \propto \delta_j^i$, entonces existe una base ortogonal $\{|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle\}$ y un conjunto de cantidades $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ tales que se cumple $\mathbb{A} |e_i\rangle = \lambda_i |e_i\rangle$ con $i = 1, 2, \dots, n$.

4.6.2 El polinomio característico, autovalores y autovectores de un operador

Nos toca ahora generar un método para calcular los autovalores $\{\lambda_j\}$, con $k = 1, 2, \dots, k \leq n$ para un operador lineal $\mathbb{A} : \mathbf{V}^n \rightarrow \mathbf{V}^n$ suponiendo que existe una base ortonormal $\{|e_1\rangle, |e_2\rangle, \dots, |e_n\rangle\}$. Entonces la

ecuación (4.23) nos ilustra la representación matricial de la ecuación de autovalores:

$$\langle e^i | \mathbb{A} | e_j \rangle \langle e^j | \psi \rangle = \lambda \langle e^i | \psi \rangle \Rightarrow A_j^i c^j = \lambda c^i \Rightarrow (A_j^i - \lambda \delta_j^i) c^j = 0,$$

con $j = 1, 2, \dots, n$. El conjunto de ecuaciones: $(A_j^i - \lambda \delta_j^i) c^j = 0$, puede ser considerado un sistema (lineal y homogéneo) de ecuaciones con n incógnitas c^j , el cual tendrá solución si el determinante de los coeficientes se anula. Tendremos entonces que $(A_j^i - \lambda \delta_j^i) c^j = 0 \Rightarrow \det |\mathbb{A} - \lambda \mathbb{I}| = 0$, es decir:

$$P(\lambda) = \det |A_j^i - \lambda \delta_j^i| = 0. \quad (4.24)$$

Esta ecuación se denomina ecuación característica (o secular) y a partir de ella emergen los autovalores (el espectro) del operador \mathbb{A} , esto es:

$$\det |A_j^i - \lambda \delta_j^i| = \begin{vmatrix} A_1^1 - \lambda & A_2^1 & \cdots & A_n^1 \\ A_1^2 & A_2^2 - \lambda & & A_n^2 \\ \vdots & & \ddots & \\ A_1^n & A_2^n & & A_n^n - \lambda \end{vmatrix} = 0.$$

Es importante señalar que el polinomio característico será independiente de la base a la cual esté referida la representación matricial $\langle e^i | \mathbb{A} | u_j \rangle$ del operador \mathbb{A} , porque hereda del determinante, esa invariancia de la representación matricial tal y como vimos en la página 256.

El resultado será un polinomio de grado n (el polinomio característico). Las raíces de este polinomio serán los autovalores que estamos buscando. Es claro que estas raíces podrán ser reales y distintas, algunas reales e iguales y otras imaginarias.

Para el caso de raíces reales, tendremos el mismo número de raíces que el grado del polinomio, generado por la representación matricial del operador con ese mismo grado. Es decir, tendremos un operador \mathbb{A} con una representación matricial de dimensión $n \times n$, con n autovalores distintos, que estará asociados a n autovectores que serán linealmente independientes y que generarán una representación matricial diagonal.

4.6.3 El caso degenerado

Sea el operador lineal $\mathbb{A} : \mathbf{V}^n \rightarrow \mathbf{V}^n$ y calculamos el polinomio característico de grado n a partir de (4.24). Se puede dar el caso que al menos una de las raíces del polinomio característico presenten algún grado de multiplicidad. Entonces el polinomio característico se podrá factorizar de la forma:

$$P(\lambda) = \det |A_j^i - \lambda \delta_j^i| = (\lambda - \lambda_1)^k (\lambda - \lambda_2) \cdots (\lambda - \lambda_m). \quad (4.25)$$

Entonces existirán $m = n - k$ raíces simples que podrán ser asociadas con $n - k$ autovectores linealmente independientes y una raíz, λ_1 , con multiplicidad k que podrá ser asociada con $1, 2, \dots$ hasta k autovectores linealmente independientes con los anteriores. Es decir, ese autovalor estará asociado a un subespacio vectorial, denominado autoespacio S_{λ_1} tal que $\dim(S_{\lambda_1}) \leq$ grado de multiplicidad del autovalor λ_1 ¹⁴. La demostración general de la afirmación anterior queda fuera de los alcances de este trabajo, pero en la próxima sección ilustraremos la multiplicidad algebraica vs multiplicidad geométrica con varios ejemplos. Más adelante, cuando analicemos el caso particular de los autovalores para las matrices hermíticas en la sección 4.7.1, retomaremos la relación de la multiplicidad del autovalor y la dimensión del autoespacio que genera este autovalor degenerado.

¹⁴Este problema se conoce en la literatura como la relación entre la multiplicidad algebraica del autovalor y la multiplicidad geométrica del autoespacio

4.6.4 Ejemplos

1. **Reflexión respecto al plano xy .** Si $\mathbb{R} : \mathbf{V}^3 \rightarrow \mathbf{V}^3$ es tal que $\mathbb{R}|\psi\rangle = |\tilde{\psi}\rangle$ donde se ha realizado una reflexión en el plano xy . Esto es

$$\mathbb{R}|i\rangle = |i\rangle; \quad \mathbb{R}|j\rangle = |j\rangle; \quad \mathbb{R}|k\rangle = -|k\rangle,$$

con $|i\rangle, |j\rangle, |k\rangle$ los vectores unitarios cartesianos. Es claro que cualquier vector en el plano xy será autovector de \mathbb{R} con un autovalor $\lambda = 1$, mientras que cualquier otro vector $|\psi\rangle \in \mathbf{V}^3$ y que no esté en el mencionado plano cumple con $|\psi\rangle = c|k\rangle$ y también será autovector de \mathbb{R} pero esta vez con un autovalor $\lambda = -1$.

2. **Dos visiones de rotaciones de ángulo fijo θ .** La rotaciones de un vector en el plano pueden verse de dos maneras.

- (a). Se considera el plano como un espacio vectorial *real* \mathbf{V}^2 con una base cartesiana canónica:

$$|i\rangle = (1, 0), \quad |j\rangle = (0, 1), \text{ esto es, si:}$$

$$\mathbb{R}|a\rangle = \lambda|a\rangle \Rightarrow \text{el ángulo de rotación} = n\pi, \quad \text{con } n \text{ entero.}$$

- (b). Igualmente si consideramos el plano complejo unidimensional, expresemos cualquier vector en el plano en su forma polar $|z\rangle = re^{i\theta}$ por lo cual:

$$\mathbb{R}|z\rangle = re^{i(\theta+\alpha)} = e^{i\alpha}|z\rangle,$$

si queremos $\lambda = e^{i\alpha}$ reales, necesariamente $\alpha = n\pi$ con n entero.

3. **Proyectores: autovalores y autovectores.** Es interesante plantearse la ecuación de autovalores con la definición del proyector para un determinado *autoespacio*. Esto es, dado $P_\psi = |\psi\rangle\langle\psi|$ si este proyector cumple con una ecuación de autovalores para un $|\varphi\rangle$ supuestamente arbitrario

$$P_\psi|\varphi\rangle = \lambda|\varphi\rangle \Rightarrow P_\psi|\varphi\rangle = (|\psi\rangle\langle\psi|)|\varphi\rangle \Rightarrow |\varphi\rangle \propto |\psi\rangle,$$

es decir, necesariamente $|\varphi\rangle$ es colineal con $|\psi\rangle$. Más aún, si ahora el $|\varphi\rangle$ no es tan arbitrario sino que es ortogonal a $|\psi\rangle$, $\langle\psi|\varphi\rangle = 0 \Rightarrow \lambda = 0$, entonces el espectro del operador $P_\psi = |\psi\rangle\langle\psi|$ es 0 y 1, el primero de los cuales es infinitamente degenerado y el segundo es simple. Esto nos lleva a reflexionar que si existe un autovector de un determinado operador, entonces su autovalor es distinto de cero, pero pueden existir autovalores nulos que generan un autoespacio infinitamente degenerado.

4. **El operador diferenciación.** $\mathbb{D}|f\rangle \rightarrow D(f) = f'$. Los autovectores del operador diferenciación necesariamente deben satisfacer la ecuación:

$$\mathbb{D}|f\rangle = \lambda|f\rangle \rightarrow D(f)(x) = f'(x) = \lambda f(x),$$

la solución a esta ecuación será una exponencial. Esto es, $|f\rangle \rightarrow f(x) = ce^{\lambda x}$, con $c \neq 0$, y donde las $f(x)$ se denominarán *autofunciones* del operador.

5. **Ejemplos de autovalores y autovectores matrices reales.** El procedimiento para el cálculo de los autovectores y autovalores es el siguiente. Una vez obtenidas las raíces del polinomio característico (los autovalores), se procede a determinar el autovector, $|\psi_j\rangle$, correspondiente a cada autovalor. Distinguiremos en esta determinación casos particulares dependiendo del tipo de raíz del polinomio característico. Ilustraremos estos casos con ejemplos para el caso específico de matrices reales 3×3 .

- (a). **Todos los autovalores reales son distintos.** Hemos presentado y analizado en la sección 4.6.1 tres teoremas que discuten la relación entre los autovalores, y la independencia lineal de los autovectores asociados con éstos. En esta sección ilustraremos el caso que surge cuando las raíces del polinomio característico (4.24) son reales y distintas.

Consideremos la siguiente representación matricial de un operador \mathbb{A} en la base canónica

$$\langle e^i | \mathbb{A} | e_j \rangle = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{pmatrix} \Rightarrow \det |A_j^i - \lambda \delta_j^i| = \begin{vmatrix} 2 - \lambda & 1 & 3 \\ 1 & 2 - \lambda & 3 \\ 3 & 3 & 20 - \lambda \end{vmatrix} = 0,$$

con lo cual el polinomio característico (4.24) queda expresado como:

$$\lambda^3 - 24\lambda^2 + 65\lambda - 42 = (\lambda - 1)(\lambda - 2)(\lambda - 21) = 0,$$

y es claro que tiene 3 raíces distintas. Para proceder a calcular los autovectores correspondientes a cada autovalor resolvemos la ecuación de autovalores (4.23) para cada autovalor. Esto es:

- $\lambda_1 = 1$

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = 1 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 2x^1 + x^2 + 3x^3 & = & x^1 \\ x^1 + 2x^2 + 3x^3 & = & x^2 \\ 3x^1 + 3x^2 + 20x^3 & = & x^3 \end{array},$$

que constituye un sistema de ecuaciones algebraicas de 3 ecuaciones con 3 incógnitas. Resolviendo el sistema tendremos el primer autovector:

$$|\psi\rangle_{\lambda_1=1} = |\psi_1\rangle = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix},$$

con α un número distinto de cero. Este α indica la indeterminación que discutimos a comienzos de la sección 4.6 correspondiente a esa ambigüedad que impone la presencia de una constante arbitraria de proporcionalidad.

- $\lambda_2 = 2$

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = 2 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 2x^1 + x^2 + 3x^3 & = & 2x^1 \\ x^1 + 2x^2 + 3x^3 & = & 2x^2 \\ 3x^1 + 3x^2 + 20x^3 & = & 2x^3 \end{array}.$$

Resolviendo el sistema se tiene el segundo autovector

$$|\psi\rangle_{\lambda_2} = |\psi_2\rangle = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \beta \begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix}.$$

- $\lambda_3 = 21$

$$\begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = 21 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 2x^1 + x^2 + 3x^3 & = & 21x^1 \\ x^1 + 2x^2 + 3x^3 & = & 21x^2 \\ 3x^1 + 3x^2 + 20x^3 & = & 21x^3 \end{array}.$$

Al resolver este sistema el tercer autovector

$$|\psi\rangle_{\lambda_3} = |\psi_3\rangle = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \gamma \begin{pmatrix} 1 \\ 1 \\ 6 \end{pmatrix}.$$

Como hemos dicho, podemos eliminar la ambigüedad de la fase arbitraria si normalizamos los autovectores:

$$|\hat{\psi}_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad |\hat{\psi}_2\rangle = \frac{1}{\sqrt{19}} \begin{pmatrix} -3 \\ -3 \\ 1 \end{pmatrix}, \quad |\hat{\psi}_3\rangle = \frac{1}{\sqrt{38}} \begin{pmatrix} 1 \\ 1 \\ 6 \end{pmatrix}.$$

Notemos que $\langle \hat{\psi}^i | \hat{\psi}_j \rangle = \delta_j^i$, es decir la base de autovectores es, necesariamente, ortonormal.

- (b). **Autovalores degenerados.** En esta sección discutiremos algunos ejemplos de autovalores degenerados, vale decir cuando alguna de las raíces del polinomio característico (4.24) tiene una multiplicidad. Arriba en la página 290 discutimos el caso de la diferencia de la multiplicidad algebraica y la multiplicidad geométrica. Esto es que la dimensión del autoespacio siempre es menor o igual a la multiplicidad de la raíz degenerada del polinomio característico $\dim(S_{\lambda_1}) \leq$ grado de multiplicidad del autovalor.

I. Multiplicidad geométrica igual a multiplicidad algebraica. Consideremos una matriz real

3×3

$$\langle e^i | A | e_j \rangle = \begin{pmatrix} 4 & -3 & 1 \\ 4 & -1 & 0 \\ 1 & 7 & -4 \end{pmatrix} \Rightarrow \det |A_j^i - \lambda \delta_j^i| = \begin{vmatrix} 4 - \lambda & -3 & 1 \\ 4 & -1 - \lambda & 0 \\ 1 & 7 & -4 - \lambda \end{vmatrix} = 0,$$

con lo cual el polinomio característico queda expresado como:

$$\lambda^3 + \lambda^2 - 5\lambda - 3 = (\lambda + 3)(\lambda - 1)^2 = 0,$$

y es claro que tiene 2 raíces iguales y una distinta. En este caso $\lambda = 1$ es un autovalor degenerado de orden 2.

Ahora resolveremos la ecuación de autovalores para cada autovalor.

- $\lambda_1 = -3$

$$\begin{pmatrix} 4 & -3 & 1 \\ 4 & -1 & 0 \\ 1 & 7 & -4 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = -3 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 4x^1 - 3x^2 + x^3 & = & -3x^1 \\ 4x^1 - x^2 & = & -3x^2 \\ x^1 + 7x^2 - 4x^3 & = & -3x^3. \end{array}$$

$$\text{Resolviendo } |\psi\rangle_{\lambda_1} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} -1 \\ 2 \\ 13 \end{pmatrix} \Rightarrow |\hat{\psi}_1\rangle = \frac{1}{\sqrt{174}} \begin{pmatrix} -1 \\ 2 \\ 13 \end{pmatrix}$$

- $\lambda_2 = 1$ (autovalor degenerado de orden 2)

$$\begin{pmatrix} 4 & -3 & 1 \\ 4 & -1 & 0 \\ 1 & 7 & -4 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 4x^1 - 3x^2 + x^3 & = & x^1 \\ 4x^1 - x^2 & = & x^2 \\ x^1 + 7x^2 - 4x^3 & = & x^3. \end{array}$$

Resolviendo el sistema tendremos el segundo autovector

$$|\psi\rangle_{\lambda_2} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \Rightarrow |\hat{\psi}_2\rangle = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Claramente $\{|\hat{\psi}_1\rangle, |\hat{\psi}_2\rangle\}$ son linealmente independientes como nos los había anunciado el teorema 1 en la sección 4.6.1, sin embargo, esos autovectores no presentan ninguna relación de ortogonalidad $\langle \hat{\psi}^i | \hat{\psi}_j \rangle \neq \delta_j^i$.

Más, aún, este autovector puede descomponerse en infinitas parejas de vectores linealmente independientes que dependen el subespacio $\mathbb{S}_{\lambda=1}$, asociado con $\lambda = 1$, el autovalor degenerado de multiplicidad $k = 2$. Una posible combinación lineal podría ser:

$$|\psi\rangle_{\lambda_2} = \alpha |\psi_1\rangle_{\lambda_2} + \beta |\psi_2\rangle_{\lambda_2} = \alpha \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} + \beta \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}.$$

Por lo tanto, la multiplicidad geométrica –la dimensión del autoespacio, $S_{\lambda_2=1}$, asociado con el autovalor degenerado $\lambda = 1$ – coincide con la multiplicidad algebraica del autovalor $\lambda = 1$ como raíz del polinomio característico. Es decir, podemos determinar la dimensión del subespacio que alberga al vector $|\psi\rangle_{\lambda_2}$. Es importante recalcar que ninguno de los (infinitos) dos vectores linealmente independientes, será autovector de \mathbb{A} por separado. Unicamente lo será la combinación lineal en la que resulta $|\psi\rangle_{\lambda_2}$.

II. Multiplicidad geométrica menor a multiplicidad algebraica. Consideremos, para ilustrar este caso la siguiente matriz real 3×3

$$\langle e^i | \mathbb{A} | e_j \rangle = \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} \Rightarrow \det |A_j^i - \lambda \delta_j^i| = \begin{vmatrix} 1 - \lambda & 1 & 2 \\ 0 & 1 - \lambda & 3 \\ 0 & 0 & 2 - \lambda \end{vmatrix} = 0.$$

En este caso, el polinomio característico es:

$$\lambda^3 - 4\lambda^2 + 5\lambda - 2 = (\lambda - 2)(\lambda - 1)^2 = 0.$$

Una vez más, el polinomio característico tiene 2 raíces iguales y una distinta y, como en el caso anterior $\lambda = 1$ es un autovalor degenerado de orden 2. Pasemos a resolver la ecuación de autovalores para cada autovalor.

- $\lambda_1 = 2$

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = 2 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{rcl} x^1 + x^2 + 2x^3 & = & 2x^1 \\ x^2 + 3x^3 & = & 2x^2 \\ 2x^3 & = & 2x^3. \end{array}$$

Resolviendo el sistema obtenemos al primer autovector: $|\psi_1\rangle = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix}$.

- $\lambda_2 = 1$

$$\begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{rcl} x^1 + x^2 + 2x^3 & = & x^1 \\ x^2 + 3x^3 & = & x^2 \\ 2x^3 & = & x^3 \end{array}$$

Con el correspondiente segundo autovector: $|\psi_2\rangle = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Una vez más los dos vectores NO son ortogonales, $\langle \psi^i | \psi_j \rangle \neq \delta_j^i$, pero SI linealmente independientes. En este caso, asociado al autovalor $\lambda_2 = 1$, tenemos además del autovector $|\psi_2\rangle$, el autovector nulo y esto genera una imposibilidad de determinar la dimensión del autoespacio. Los autovectores siguen siendo linealmente independientes, pero aquí la dimensión del autoespacio es NECESARIAMENTE 1, $\dim(S_{\lambda_2=1}) = 1$, por lo tanto ilustra que la multiplicidad geométrica es menor que la multiplicidad aritmética de las raíces del polinomio característico.

III. Vamos ilustrar un tercer ejemplo que se presenta para uno de los autovalores degenerados. Consideremos entonces otra matriz 3×3 con autovalores repetidos

$$\langle e^i | \mathbb{A} | e_j \rangle = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \Rightarrow \det |A_j^i - \lambda \delta_j^i| = \begin{vmatrix} 2 - \lambda & 1 & 1 \\ 2 & 3 - \lambda & 2 \\ 3 & 3 & 4 - \lambda \end{vmatrix} = 0,$$

con lo cual el polinomio característico es:

$$\lambda^3 + \lambda^2 - 5\lambda - 3 = (\lambda - 7)(\lambda - 1)^2 = 0,$$

que tiene 2 raíces iguales y una distinta. En este caso $\lambda = 1$ vuelve a ser un autovalor degenerado de orden 2. Volvemos a calcular los autovectores correspondientes a cada autovalor

- $\lambda_1 = 7$

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = 7 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 2x^1 + x^2 + x^3 & = & 7x^1 \\ 2x^1 + 3x^2 + 3x^3 & = & 7x^2 \\ 3x^1 + 3x^2 + 4x^3 & = & 7x^3. \end{array}$$

Al resolver el sistema

$$|\psi\rangle_{\lambda_1} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

- $\lambda_2 = 1$. En este caso el autovalor degenerado de orden 2 presenta una pequeña patología.

Veamos

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 2x^1 + x^2 + x^3 & = & x^1 \\ 2x^1 + 3x^2 + 2x^3 & = & x^2 \\ 3x^1 + 3x^2 + 4x^3 & = & x^3. \end{array}$$

Resolviendo

$$|\psi\rangle_{\lambda_2} = \begin{cases} |\psi_1\rangle_{\lambda_2} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ |\psi_2\rangle_{\lambda_2} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \beta \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \end{cases}$$

con lo cual el autovector $|\psi\rangle_{\lambda_2}$ correspondiente al autovalor $\lambda_2 = 1$, está asociado con a dos vectores linealmente independientes $\{|\psi_1\rangle_{\lambda_2}, |\psi_2\rangle_{\lambda_2}\}$ y, por lo tanto aquí también la multiplicidad aritmética coincide con la multiplicidad geométrica. Una vez más los autovectores, correspondientes a distintos autovalores no son ortogonales pero si linealmente independientes.

- (c). Finalmente, consideremos una matriz 3×3 con 1 autovalor real y dos autovalores complejos

$$\langle e^i | A | e_j \rangle = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \Rightarrow \det |A_j^i - \lambda \delta_j^i| = \begin{vmatrix} 1-\lambda & 2 & 3 \\ 3 & 1-\lambda & 2 \\ 2 & 3 & 1-\lambda \end{vmatrix} = 0,$$

con lo cual el polinomio característico queda expresado como:

$$\lambda^3 - 3\lambda^2 - 15\lambda - 18 = (\lambda - 6)(\lambda^2 + 3\lambda + 3) = 0.$$

En este caso $\lambda_1 = 6$ es un autovalor real. Adicionalmente existen dos autovalores complejos, uno el complejo conjugado del otro: $\lambda_2 = -\frac{1}{2}(3 + i\sqrt{3})$ y $\lambda_3 = -\frac{1}{2}(3 - i\sqrt{3})$. Para proceder a calcular los autovectores correspondientes a cada autovalor resolvemos la ecuación de autovalores para cada

autovalor real. En este caso existe **un único** autovalor real $\lambda = 6$.

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = 6 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \iff \begin{array}{lcl} 4x^1 - 3x^2 + x^3 & = & 6x^1 \\ 4x^1 - x^2 & = & 6x^2 \\ x^1 + 7x^2 - 4x^3 & = & 6x^3. \end{array}$$

Tendremos que para $\lambda = 6$

$$|\psi\rangle_{\lambda_1} = \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

6. Sean \mathbb{A} y \mathbb{B} dos operadores hermíticos, con autovalores no degenerados y un operador unitario definido como: $\mathbb{U} = \mathbb{A} + i\mathbb{B}$. Vamos a mostrar que:

(a). Si \mathbb{A} y \mathbb{B} comutan, $[\mathbb{B}, \mathbb{A}] = 0$, los autovectores de \mathbb{A} también lo son de \mathbb{B} .

Si $\{|u_i\rangle\}$ son autovectores de \mathbb{A} entonces

$$\mathbb{A}|u_i\rangle = \lambda_i|u_i\rangle \Rightarrow \mathbb{B}\mathbb{A}|u_i\rangle = \lambda_i\mathbb{B}|u_i\rangle, \text{ como } [\mathbb{B}, \mathbb{A}] = 0, \text{ entonces } \mathbb{A}\mathbb{B}|u_i\rangle = \lambda_i\mathbb{B}|u_i\rangle.$$

Por lo tanto, $\mathbb{B}|u_i\rangle$ es un autovector de \mathbb{A} . Pero la solución para la ecuación de autovectores $(\mathbb{A} - \lambda_i\mathbb{I})|u_i\rangle = 0$ es única, por lo cual todos los autovectores de \mathbb{A} son proporcionales. Esto es: $\mathbb{B}|u_j\rangle = \mu_j|u_j\rangle$, con lo cual queda demostrado que los autovectores de \mathbb{A} son autovectores de \mathbb{B} .

(b). Si $\mathbb{U}|v_i\rangle = \mu_i|v_i\rangle$, entonces $|\mu_i| = 1$.

$$\text{Es claro que: } \langle v^j | \mathbb{U}^\dagger \mathbb{U} | v_i \rangle = \langle v^j | \mathbb{I} | v_i \rangle \Rightarrow \mu_j^* \mu_i \langle v^j | v_i \rangle = \langle v^j | \mathbb{I} | v_i \rangle \Rightarrow \mu_i^2 = 1.$$

7. Dada una matriz de la forma

$$\mathbb{A} = \begin{pmatrix} 1 & \alpha & 0 \\ \beta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{y} \quad \mathbb{A}|v_i\rangle = \lambda_i|v_i\rangle$$

con α y β números complejos distintos de cero. Vamos a encontrar:

(a). Las relaciones que deben cumplir α y β para que λ_i sea real.

El polinomio característico y la condición para que λ sea real es:

$$(1 - \lambda)(1 - 2\lambda + \lambda^2 - \alpha\beta) = 0 \Rightarrow \lambda = 1 \pm \sqrt{\alpha\beta} \Rightarrow \alpha\beta > 0 \wedge \alpha\beta \in \mathbb{R}.$$

(b). Las relaciones que deben cumplir α y β para que $\langle v^j | v_i \rangle = \delta_i^j$.

Los autovalores y autovectores para esta matriz serán:

$$\lambda_1 = 1 \Rightarrow |v_1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_2 = 1 + \sqrt{\alpha\beta} \Rightarrow |v_2\rangle = \begin{pmatrix} \frac{\beta}{\sqrt{\alpha\beta}} \\ 1 \\ 0 \end{pmatrix}, \quad \lambda_3 = 1 - \sqrt{\alpha\beta} \Rightarrow |v_3\rangle = \begin{pmatrix} \frac{-\beta}{\sqrt{\alpha\beta}} \\ 1 \\ 0 \end{pmatrix},$$

con lo cual: $\langle v^2 | v_3 \rangle = 0 \Rightarrow \frac{\beta^2}{\alpha\beta} = 1 \Rightarrow |\alpha| = |\beta|$.

(c). Supongamos que \mathbb{A} es hermítica, encontraremos las relaciones que deben cumplir α y β .

Si \mathbb{A} es hermítica, entonces $\alpha^* = \beta$, con lo cual se cumplen automáticamente ambas aseveraciones.

8. Dadas las siguientes matrices:

$$\mathbb{A} = \begin{pmatrix} 6 & -2 \\ -2 & 9 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 1 & 8 \\ 8 & -11 \end{pmatrix}, \quad \mathbb{C} = \begin{pmatrix} -9 & -10 \\ -10 & 5 \end{pmatrix}, \quad \mathbb{D} = \begin{pmatrix} 14 & 2 \\ 2 & 11 \end{pmatrix}.$$

Determinemos cuales comutan entre ellas y busquemos la base de autovectores comunes.

Notamos que: $[\mathbb{A}, \mathbb{B}] = [\mathbb{A}, \mathbb{D}] = [\mathbb{D}, \mathbb{B}] = 0$, y

$$[\mathbb{A}, \mathbb{C}] = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix}, \quad [\mathbb{B}, \mathbb{C}] = \begin{pmatrix} 0 & -8 \\ 8 & 0 \end{pmatrix}, \quad [\mathbb{D}, \mathbb{C}] = \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix}$$

Los autovectores comunes a $\mathbb{A}, \mathbb{B}, \mathbb{D}$, serán:

$$|u_1\rangle = \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix}, \quad |u_2\rangle = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

9. Dada la representación matricial de dos operadores

$$\mathbb{A} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{y} \quad \mathbb{B} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

(a). Evaluemos $[\mathbb{A}, \mathbb{B}]$.

$$\mathbb{A}\mathbb{B} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = \mathbb{B}\mathbb{A} \Rightarrow [\mathbb{A}, \mathbb{B}] = 0$$

(b). Vamos a mostrar que \mathbb{A} tiene por autovalores $\lambda_1 = 1$ y $\lambda_2 = -1$, con λ_1 un autovalor degenerado. Y construyamos luego la base de autovectores para \mathbb{A} .

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \begin{pmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Entonces para:

$$\begin{vmatrix} -\lambda & 0 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 0 & -\lambda \end{vmatrix} = \lambda(1-\lambda)\lambda - (1-\lambda) = (\lambda^2 - 1)(1-\lambda) = 0,$$

se tienen dos autovalores: $\lambda = 1$ y $\lambda = -1$. Para el caso de $\lambda = -1$ se cumple que:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = - \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \begin{array}{l} z = -x \\ y = -y \\ x = -z \end{array}$$

con lo cual el autovector asociado con el autovalor $\lambda = -1$ tendrá la forma de:

$$|u\rangle_{-1} = \alpha \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Para $\lambda = 1$ se cumple:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \begin{array}{l} z = x \\ y = y \\ x = z \end{array}$$

hay dos vectores linealmente independientes asociados con $\lambda = 1$, a saber:

$$|u\rangle_{1a} = \beta \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{y} \quad |u\rangle_{1b} = \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix} \quad \text{con } y \text{ arbitrario.}$$

Nótese que estos tres autovectores: $\{|u\rangle_{1a}, |u\rangle_{1b}, |u\rangle_{-1}\}$ son ortogonales entre si.

(c). ¿Cuál es la representación matricial de \mathbb{A} en la base de autovectores?

Veamos lo siguiente:

$$\tilde{A}_j^i = \begin{pmatrix} \langle u^1 | \mathbb{A} | u_1 \rangle & \langle u^1 | \mathbb{A} | u_2 \rangle & \langle u^1 | \mathbb{A} | u_3 \rangle \\ \langle u^2 | \mathbb{A} | u_1 \rangle & \langle u^2 | \mathbb{A} | u_2 \rangle & \langle u^2 | \mathbb{A} | u_3 \rangle \\ \langle u^3 | \mathbb{A} | u_1 \rangle & \langle u^3 | \mathbb{A} | u_2 \rangle & \langle u^3 | \mathbb{A} | u_3 \rangle \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

ya que los autovectores forman una base ortogonal. Obviamente se cumple que:

$$\det|\mathbb{A}| = \det|\tilde{\mathbb{A}}| = -1 \quad \text{y} \quad \text{Tr}(\mathbb{A}) = \text{Tr}(\tilde{\mathbb{A}}) = 1.$$

- (d). A partir de los autovectores de \mathbb{A} vamos a calcular los autovalores y autovectores de \mathbb{B} .

Claramente $\mathbb{B} |u_{-1}\rangle = -|u_{-1}\rangle$ con lo cual tenemos el primer autovector de \mathbb{B} asociado al autovalor $\lambda = -1$. Para encontrar los otros autovectores tendremos:

$$\begin{pmatrix} -\lambda & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{pmatrix} \begin{pmatrix} 1 \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

4.6.5 Practicando con Maxima

En este ejercicio aprenderemos a resolver el problema de autovalores y autovectores. Primeros lo haremos de manera indirecta, realizando los cálculos por pasos, y luego utilizando las funciones específicas del programa.

1. Dada la matriz A del primer ejemplo anteriormente resuelto:

```
(%i1) A:matrix([2,1,3], [1,2,3], [3,3,20]);
(%o1) ⎛2 1 3⎞
      ⎝1 2 3⎠
            3 3 20
```

Generamos la matriz identidad 3×3 con el comando **ident(n)**.

```
(%i2) I:ident(3);
(%o2) ⎛1 0 0⎞
      ⎝0 1 0⎠
            0 0 1
```

Escribimos la ecuación de autovalores de manera matricial y la llamaremos M .

```
(%i3) M:A-lambda*I;
(%o3) ⎛2 - λ 1 3⎞
      ⎝1 2 - λ 3⎠
            3 3 20 - λ
```

El determinante nos dará el polinomio característico, luego resolvemos para λ .

```
(%i4) factor(determinant(M));
(%o4) -(λ - 21)(λ - 2)(λ - 1)

(%i5) solve(%=0);
(%o5) [λ = 1, λ = 2, λ = 21]
```

Los autovalores son: $\lambda_1 = 1$, $\lambda_2 = 2$ y $\lambda_3 = 21$. Para cada autovalor evaluaremos una matriz correspondiente:

```
(%i6) M1:M,lambda=1; M2:M,lambda=2; M3:M,lambda=21;
(%o6) ⎛1 1 3⎞
      ⎝1 1 3⎠
            3 3 19

(%o7) ⎛0 1 3⎞
      ⎝1 0 3⎠
            3 3 18

(%o8) ⎛-19 1 3⎞
      ⎝1 -19 3⎠
            3 3 -1
```

Necesitamos ahora resolver, para cada autovalor, la ecuación de autovectores: $AX = \lambda_i X$. Así que podemos escribir las siguientes matrices:

```
(%i9) M1:M1.[x,y,z]; M2:M2.[x,y,z]; M3:M3.[x,y,z];
```

$$(\%o9) \begin{pmatrix} 3z + y + x \\ 3z + y + x \\ 19z + 3y + 3x \end{pmatrix}$$

$$(\%o10) \begin{pmatrix} 3z + y \\ 3z + x \\ 18z + 3y + 3x \end{pmatrix}$$

$$(\%o11) \begin{pmatrix} 3z + y - 19x \\ 3z - 19y + x \\ -z + 3y + 3x \end{pmatrix}$$

Iremos ahora resolviendo cada uno de los sistemas de ecuaciones. Para $\lambda_1 = 1$:

(%i12) `ec1:M1[1,1]=0; ec2:M1[2,1]=0; ec3:M1[3,1]=0;`

$$(\%o12) 3z + y + x = 0$$

$$(\%o13) 3z + y + x = 0$$

$$(\%o14) 19z + 3y + 3x = 0$$

(%i15) `solve([ec1,ec2,ec3],[x,y,z]);`

`solve: dependent equations eliminated: (2)`

$$(\%o15) [[x = -\%r_4, y = \%r_4, z = 0]]$$

Recordemos que **Maxima** nos indica con el símbolo $\%r_n$ que se trata de una constante. Así que de las infinitas soluciones escogemos alguna de las más simples: $(1, -1, 0)$.

Repetimos los cálculos para el segundo autovalor, es decir, $\lambda_2 = 2$:

(%i16) `ec1:M2[1,1]=0; ec2:M2[2,1]=0; ec3:M2[3,1]=0;`

$$(\%o16) 3z + y = 0$$

$$(\%o17) 3z + x = 0$$

$$(\%o18) 18z + 3y + 3x = 0$$

(%i19) `solve([ec1,ec2,ec3],[x,y,z]);`

`solve: dependent equations eliminated: (3)`

$$(\%o19) [[x = -3\%r_5, y = -3\%r_5, z = \%r_5]]$$

Por lo tanto un autovector puede ser: $(1, 1, -1/3)$.

Finalmente, para $\lambda_3 = 21$:

(%i20) `ec1:M3[1,1]=0; ec2:M3[2,1]=0; ec3:M3[3,1]=0;`

$$(\%o20) 3z + y - 19x = 0$$

$$(\%o21) 3z - 19y + x = 0$$

$$(\%o22) -z + 3y + 3x = 0$$

(%i23) `solve([ec1,ec2,ec3],[x,y,z]);`

`solve: dependent equations eliminated: (3)`

$$(\%o23) \left[\left[x = \frac{\%r_6}{6}, y = \frac{\%r_6}{6}, z = \%r_6 \right] \right]$$

Podemos tomar como autovector a: $(1, 1, 6)$.

2. **Maxima** ofrece la posibilidad de resolver el problema de autovalores a través de un determinado número

de funciones. Por ejemplo, la función **charpoly** nos dará el polinomio característico directamente de la matriz que queremos estudiar y respecto de la variable que seleccionemos, es este caso será λ .

```
(%i24) charpoly(A,lambda);
(%o24)  $\lambda + ((2 - \lambda)(20 - \lambda) - 9)(2 - \lambda) + 3(3 - 3(2 - \lambda)) - 11$ 

(%i25) factor(%);
(%o25)  $-(\lambda - 21)(\lambda - 2)(\lambda - 1)$ 

(%i26) solve(%=0);
(%o26)  $[\lambda = 1, \lambda = 2, \lambda = 21]$ 
```

El programa también permite obtener los autovalores directamente de la matriz problema. Esto se hace con el comando **eigenvalues(M)**. El resultado será una lista conformada a su vez por dos listas: la primera sublistas la forman los valores propios de la matriz M y la segunda con su multiplicidad correspondientes. Veamos:

```
(%i27) eigenvalues(A);
(%o27)  $[[1, 2, 21], [1, 1, 1]]$ 
```

La función para calcular los autovectores es **eigenvectors(M)**. El resultado será una lista con dos elementos; el primero está formado por dos listas: la primera con los valores propios de M y la segunda con su respectiva multiplicidad, el segundo elemento es una lista de listas de vectores propios, una por cada valor propio.

```
(%i28) eigenvectors(A);
(%o28)  $\left[ [[1, 2, 21], [1, 1, 1]], \left[ [[1, -1, 0]], \left[ \left[ 1, 1, -\frac{1}{3} \right], [[1, 1, 6]] \right] \right] \right]$ 
```

3. Consideremos otro de los ejemplos tratado anteriormente. Dada la matriz:

```
(%i29) B:matrix([4,-3,1], [4,-1,0], [1,7,-4]);
(%o29)  $\begin{pmatrix} 4 & -3 & 1 \\ 4 & -1 & 0 \\ 1 & 7 & -4 \end{pmatrix}$ 
(%i30) eigenvectors(B);
(%o30)  $[[[-3, 1], [1, 2]], [[[1, -2, -13]], [[1, 2, 3]]]]$ 
```

Notemos que sólo se obtienen dos autovectores.

4. Consideremos la matriz:

```
(%i31) C:matrix([2,1,1], [2,3,2], [3,3,4]);
(%o31)  $\begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix}$ 
(%i32) eigenvectors(C);
(%o32)  $[[[7, 1], [1, 2]], [[[1, 2, 3]], [[1, 0, -1], [0, 1, -1]]]]$ 
```

5. Y finalmente, dada la matriz:

```
(%i33) D:matrix([1,2,3], [3,1,2], [2,3,1]);
```

(%o33)
$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}$$

(%i34) `eigenvectors(D);`

(%o34)
$$\left[\left[\left[\left[-\frac{\sqrt{3}i+3}{2}, \frac{\sqrt{3}i-3}{2}, 6 \right], [1, 1, 1] \right], \left[\left[\left[1, \frac{\sqrt{3}i-1}{2}, -\frac{\sqrt{3}i+1}{2} \right] \right] \right], \left[\left[\left[1, -\frac{\sqrt{3}i+1}{2}, \frac{\sqrt{3}i-1}{2} \right] \right], [[1, 1, 1]] \right]$$

4.6.6 Ejercicios

1. Si $|v_1\rangle$ y $|v_2\rangle$ son autovectores del operador lineal \mathbb{A} que corresponden a distintos autovalores. Muestre que $\alpha|v_1\rangle + \beta|v_2\rangle$ ($\alpha \neq 0, \beta \neq 0$) no puede ser un autovector de \mathbb{A} .
2. Demuestre que si todo vector de un espacio vectorial V es un autovector del operador lineal \mathbb{A} , entonces $\mathbb{A} = \lambda\mathbb{I}$. Donde \mathbb{I} es el operador identidad.
3. Demuestre que si el operador lineal \mathbb{A} conmuta con todos operadores que actúan en un determinado espacio vectorial, entonces $\mathbb{A} = \lambda\mathbb{I}$.
4. Si un operador lineal \mathbb{A} tiene un autovector $|v_0\rangle$ con autovalor λ_0 . Demuestre que $|v_0\rangle$ es también un autovector del operador \mathbb{A}^2 con autovalor λ_0^2 .
5. Aun si un operador lineal \mathbb{A} no tiene autovectores el operador \mathbb{A}^2 puede llegar a tenerlos. Demuestre que si \mathbb{A}^2 tiene un autovector con un autovalor no degenerado $\lambda_0 = \mu^2$, entonces \mathbb{A} tiene un autovector.
6. Encuentre los autovalores y autovectores de las siguientes matrices:

$$\mathbb{A} = \begin{pmatrix} 1 & 3 & -1 \\ 3 & 4 & -2 \\ -1 & -2 & 2 \end{pmatrix}, \quad \mathbb{B} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -7 & 1 \end{pmatrix}, \quad \mathbb{C} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathbb{D} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

y la matriz $\mathbb{E} = \text{diag}(1, 1, -1, -1)$. Verifique si los autovectores son ortogonales entre ellos.

7. Demuestre que la matriz

$$\mathbb{A} = \begin{pmatrix} 2 & 0 & 0 \\ -6 & 4 & 4 \\ 3 & -1 & 0 \end{pmatrix},$$

no tiene tres autovectores linealmente independientes y que cualquier autovector tiene la forma:

$$\begin{pmatrix} \lambda \\ 3\lambda - 2\nu \\ \nu \end{pmatrix}.$$

8. Construimos un sistema con tres partículas, de masa m_i rígidamente unidas, colocadas en tres puntos distintos de la siguiente forma:

$$m_1 = 1 \rightarrow \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \quad m_2 = 2 \rightarrow \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \quad \text{y} \quad m_3 = 1 \rightarrow \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}.$$

- (a). Encuentre la matriz del tensor de inercia.
- (b). Diagonalice esa matriz y encuentre los ejes principales de inercia.

9. En Mecánica Cuántica, el problema del oscilador armónico simple puede ser representado por un problema de autovalores

$$\mathbb{L}|\psi\rangle = \lambda|\psi\rangle \Rightarrow \langle x|\mathbb{L}|\psi\rangle = \lambda\langle x|\psi\rangle \Leftrightarrow \mathbb{L}\psi(x) = \lambda\psi(x), \quad \text{donde : } \mathbb{L} \leftrightarrow \frac{d^2}{dx^2} + \frac{x^2}{4} + \frac{1}{2}.$$

Si construimos un par de operadores:

$$\mathbb{L}_+ = \frac{x}{2} + \frac{d}{dx} \quad \text{y} \quad \mathbb{L}_- = \frac{x}{2} - \frac{d}{dx},$$

utilice los resultados del problema anterior y construya las nuevas autofunciones de \mathbb{L} con sus autovalores.

10. Considere la siguiente representación matricial para dos operadores:

$$\langle u^i | \mathbb{A} | u_j \rangle = A_j^i = \begin{pmatrix} 1 & \delta \\ \delta & 1 \end{pmatrix} \quad \text{y} \quad \langle u^m | \mathbb{B} | u_n \rangle = B_n^m = \begin{pmatrix} 1 & 1 \\ \delta^2 & 1 \end{pmatrix}.$$

- (a). Muestre como actúan \mathbb{A} y \mathbb{B} sobre un vector genérico $|\psi\rangle$ expandido en esa base $\{|u_1\rangle, |u_2\rangle\}$ (vale decir $|\psi\rangle = \alpha|u_1\rangle + \beta|u_2\rangle$).
- (b). Muestre que los autovalores de \mathbb{A} son degenerados para $\delta = 0$ y que sus autovectores son ortogonales para $\delta \neq 0$ e incluso para $\delta \rightarrow 0$.
- (c). Muestre que también \mathbb{B} tiene autovalores degenerados para $\delta = 0$ y encuentre la expresión (en función de $0 \leq \delta \leq 1$) para el coseno del ángulo entre dos autovectores.

11. Dada la siguiente representación matricial de un operador en la base canónica:

$$\langle u^i | \mathbb{M} | u_j \rangle = M_j^i = \begin{pmatrix} 2 & i\sqrt{2} \\ -i\sqrt{2} & 2 \end{pmatrix}, \quad |u_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |u_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

- (a). Encuentre los autovectores $\{|\varphi_1\rangle, |\varphi_2\rangle\}$ para ese operador en la base canónica.
- (b). Encuentre las representaciones matriciales de los operadores proyección sobre los autoespacios, $\mathbb{P}_{|\varphi_i\rangle} = |\varphi_i\rangle \langle \varphi^i|$, en esa misma base canónica.
- (c). Encuentre las representaciones matriciales de los operadores proyección sobre los complementos ortogonales de los autoespacios $U_m^n = |\varphi_m\rangle \langle \varphi^n|$ en esa misma base y con ella calcule $M = M_j^i U_i^j$.

12. Las matrices de Pauli :

$$\boldsymbol{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

las hemos visto en varios momentos de estas notas: como cuaterniones en los ejercicios 2.2.4 y como operadores lineales en los ejercicios 4.7.6.

- (a). Muestre si las matrices de Pauli $\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z$, conjuntamente con la matriz identidad, \mathbf{I} forman un grupo respecto a la siguiente operación:

$$\boldsymbol{\sigma}_j \odot \boldsymbol{\sigma}_k \equiv \boldsymbol{\sigma}_j \boldsymbol{\sigma}_k = i\epsilon_{jkm} \boldsymbol{\sigma}^m + \delta_{jk} \mathbf{I} \quad \text{con } j, k, m = x, y, z$$

donde ϵ_{jkm} es el símbolo de Levi-Civita y $i = \sqrt{-1}$.

- (b). Muestre si las matrices de Pauli $\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y, \boldsymbol{\sigma}_z$, conjuntamente con la matriz identidad \mathbf{I} son linealmente independientes.
- (c). ¿Las matrices de Pauli forman base para un espacio vectorial de matrices complejas 2×2 ? ¿Por qué? Si forman una base exprese la matriz

$$\begin{pmatrix} 3 & i \\ 5 & 1 \end{pmatrix},$$

en términos de esa base.

- (d). Derive la expresión general para $[\boldsymbol{\sigma}_j, \boldsymbol{\sigma}_k]$.

(e). Suponga ahora que σ_z actúa de la siguiente forma: $\sigma_z |+\rangle = |+\rangle$, $\sigma_z |-\rangle = -|-\rangle$, con:

$$|+\rangle \Leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |-\rangle \Leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Encuentre la expresión para los autovalores y autovectores de las otras matrices de Pauli:

$$\sigma_x |+\rangle_x = \lambda_{+x} |+\rangle_x, \quad \sigma_x |-\rangle_x = \lambda_{-x} |-\rangle_x, \quad \sigma_y |+\rangle_y = \lambda_{+y} |+\rangle_y, \quad \sigma_y |-\rangle_y = \lambda_{-y} |-\rangle_y.$$

(f). Muestre que cualquier representación matricial de un operador genérico \mathbb{M} puede ser expresado como combinación lineal de las matrices de Pauli.

(g). El polinomio característico para ese operador genérico \mathbb{M} se puede expresar como

$$P_\lambda = \lambda^2 - \lambda \text{Tr}(\mathbb{M}) + \det|\mathbb{M}|.$$

Donde los λ son sus autovalores.

13. Resuelva con **Maxima** los ejercicios anteriores.

4.7 Autovalores y autovectores de matrices importantes

Consideremos la ecuación de autovalores (4.21) y la definición de operadores adjuntos que hemos discutido en la sección 4.2.3. Podemos afirmar sin perder generalidad que dado un conjunto de vectores $\{|u_1\rangle, |u_2\rangle, \dots, |u_n\rangle\}$ y de cantidades $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ que cumplan con $\mathbb{A}|u_i\rangle = \lambda_i|u_i\rangle$ con $i = 1, 2, \dots, n$, entonces

$$\left. \begin{array}{l} \mathbb{A}|u_i\rangle = \lambda_i|u_i\rangle \Rightarrow \langle u^j | \mathbb{A}|u_i\rangle = \lambda_i \langle u^j | u_i \rangle \\ \langle u^j | \mathbb{A}^\dagger = \lambda_j^* \langle u^j | \end{array} \right\} \Rightarrow \langle u^j | \mathbb{A} - \mathbb{A}^\dagger | u_i \rangle = (\lambda_i - \lambda_j^*) \langle u^j | u_i \rangle. \quad (4.26)$$

Vale decir, hemos proyectado las ecuaciones de autovalores de un operador genérico y su adjunto a lo largo de los vectores y las formas correspondientes, $\langle u^j |$ y $|u_i\rangle$, para luego restar esas expresiones miembro a miembro.

Es evidente que si no conocemos el tipo de operador, poco se puede decir de sus autovectores y autovalores. Hay que imponer algunas restricciones sobre el tipo de operador para poder sacar algunas conclusiones respecto al tipo de autovalores y autovectores que ese operador pueda tener. Por ejemplo:

- Si \mathbb{A} es hermítico, $\mathbb{A} = \mathbb{A}^\dagger$ y autovectores son distintos, ($i \neq j$) entonces de la ecuación (4.26) se deduce que esos autovectores serán ortogonales, $\langle u^j | u_i \rangle \propto \delta_i^j$.
- Si \mathbb{A} es hermítico, $\mathbb{A} = \mathbb{A}^\dagger$ y autovectores son los mismos, ($i = j$) entonces los autovalores son reales: $\lambda_i = \lambda_i^*$.

Este caso lo analizaremos en la próxima sección.

4.7.1 Autovalores y autovectores de matrices hermíticas y unitarias

Tal y como mencionamos en la sección 4.2.4 un operador hermítico o autoadjunto cumple con $\mathbb{A}^\dagger = \mathbb{A}$ y luego en la sección 4.3.2.2, comprobamos que su representación matricial es igual a su traspuesta conjugada. Por lo tanto, la representación matricial de un operador autoadjunto es una matriz simétrica con números reales en su diagonal. Veamos ahora cómo se comportan sus autovalores.

Teorema 4: Supongamos un operador lineal $\mathbb{A} : V^n \rightarrow V^n$, hermítico: $\mathbb{A} = \mathbb{A}^\dagger$, con autovalores: $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. Entonces:

- Los autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ son reales.

- Los autovectores $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$, correspondientes a cada uno de los autovalores, serán ortogonales.

Demostración:

- Para demostrar que los autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ son reales, proyectamos la ecuación de autovalores en cada uno de los autovectores:

$$\mathbb{A}|\psi\rangle = \lambda|\psi\rangle \Rightarrow \langle\psi|\mathbb{A}|\psi\rangle = \lambda\langle\psi|\psi\rangle .$$

Ahora bien, dado que $\langle\psi|\psi\rangle$ es real, si demostramos que $\langle\psi|\mathbb{A}|\psi\rangle$ es real, estará demostrado que λ lo será también. Pero como \mathbb{A} es hermítico:

$$\langle\psi|\mathbb{A}|\psi\rangle^* = \langle\psi|\mathbb{A}^\dagger|\psi\rangle = \langle\psi|\mathbb{A}|\psi\rangle \Rightarrow \langle\psi|\mathbb{A}|\psi\rangle \in \mathbb{R} ,$$

por consiguiente los autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ son reales. Más aún, si \mathbb{A} es hermítico, y como sus autovalores son reales entonces:

$$\langle\psi|\mathbb{A}^\dagger = \lambda^* \langle\psi| = \lambda\langle\psi| \Rightarrow \langle\psi|\mathbb{A}|\phi\rangle = \lambda\langle\psi|\phi\rangle .$$

- Para demostrar que los autovectores $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$ son ortogonales, consideremos dos autovectores con sus correspondientes autovalores de tal forma que se cumplen las siguientes ecuaciones:

$$\mathbb{A}|\psi\rangle = \lambda|\psi\rangle \quad \text{y} \quad \mathbb{A}|\varphi\rangle = \mu|\varphi\rangle ,$$

pero como \mathbb{A} es hermítico entonces se cumple que: $\langle\varphi|\mathbb{A} = \mu\langle\varphi|$, multiplicando a la izquierda por $|\psi\rangle$ y a $\langle\psi|\mathbb{A} = \lambda\langle\psi|$ por $\langle\varphi|$ a la derecha:

$$\left. \begin{array}{l} (\langle\varphi|\mathbb{A} = \mu\langle\varphi|)|\psi\rangle \\ \langle\varphi|(\mathbb{A}|\psi\rangle = \lambda|\psi\rangle) \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \langle\varphi|\mathbb{A}|\psi\rangle = \mu\langle\varphi|\psi\rangle \\ \langle\varphi|\mathbb{A}|\psi\rangle = \langle\varphi|\psi\rangle \end{array} \right\} \Rightarrow (\lambda - \mu)\langle\varphi|\psi\rangle = 0 ,$$

y como hemos supuesto que $\lambda \neq \mu$ con lo cual $\langle\varphi|\psi\rangle = 0$, los autovectores correspondientes a dos autovalores son ortogonales. ◀

Es importante señalar el hecho de que si la matriz \mathbb{A} es real, entonces $\mathbb{A} = \mathbb{P}\tilde{\mathbb{A}}\mathbb{P}^T$, con $\tilde{\mathbb{A}}$ diagonal y \mathbb{P} ortogonal: $\mathbb{P}^T = \mathbb{P}^{-1}$.

En resumen, si $\mathbb{A} : \mathbf{V}^n \rightarrow \mathbf{V}^n$ es un operador lineal hermítico entonces:

- Sus autovalores son reales.
- Los autovectores correspondientes a cada autovalor son ortogonales.
- Los autovectores de \mathbb{A} resultan ser una base ortonormal del espacio vectorial \mathbf{V}^n .

4.7.1.1 Autoespacios y autovalores degenerados

Una vez más, consideraremos el caso autovalores degenerados, para operadores hermíticos o autoadjuntos. El siguiente teorema garantiza la existencia de al menos un subespacio $S(\lambda_0) \subset \mathbf{V}^n$.

Teorema 5: Sea un operador lineal $\mathbb{A} : \mathbf{V}^n \rightarrow \mathbf{V}^n$, hermítico, con una representación matricial $n \times n$, tal que su polinomio característico $\mathcal{P}(\lambda) = \det |A_j^i - \lambda\delta_j^i| = 0$ tiene al menos una raíz degenerada $\lambda = \lambda_0$, de orden $k \leq n$. Entonces existen k autovectores, no triviales, que cumplen con: $\mathbb{A}|\psi_j\rangle = \lambda_0|\psi_j\rangle$ para $j = 1, 2, \dots, k$.

Demostración: La demostración también emerge de una variante del *Método de Inducción Completa*. Para ello, probamos que se cumple para $j = 1$. Esta afirmación es obvia. Si existe un $\lambda = \lambda_0$ existe un $|\psi_j\rangle$, tal que cumple con la ecuación anterior y es linealmente independiente con él mismo.

Suponemos que se cumple para $1 \leq j = m \leq k$. Es decir, existen m autovectores $|\psi_j\rangle$ de \mathbb{A} para el autovalor λ_0 . Definamos un subespacio $S_{\lambda_0} = S(\lambda_0) \subset V^n$ donde:

$$|\psi_j\rangle \in S_{\lambda_0} \quad / \quad \mathbb{A}|\psi_j\rangle = \lambda_0|\psi_j\rangle \Rightarrow \mathbb{A}|\psi_j\rangle \in S_{\lambda_0} \quad \text{con } j = 1, 2, \dots, m,$$

por lo tanto, podremos separar V^n como una suma directa entre el subespacio S_{λ_0} y \mathcal{N} , su complemento ortogonal:

$$V^n = S_{\lambda_0} \oplus \mathcal{N} \quad / \quad \mathbb{A}|\psi_j\rangle = \lambda_0|\psi_j\rangle \quad \wedge \quad |\phi\rangle \in \mathcal{N} \Rightarrow \langle\phi|\psi_j\rangle = 0,$$

claramente S_{λ_0} es un subespacio invariante de \mathbb{A} por cuanto su acción se circumscribe dentro del mismo subespacio S_{λ_0} .

Mostraremos que se cumple para operadores hermíticos, por cuanto no es verdad en general, entonces:

$$\left. \begin{array}{l} \langle\phi|\psi_j\rangle = 0 \\ \wedge \\ \mathbb{A}|\psi_j\rangle = \lambda_0|\psi_j\rangle \end{array} \right\} \Rightarrow \langle\psi_j|\phi\rangle = 0 = \langle\psi_j|\mathbb{A}^\dagger|\phi\rangle = \langle\psi_j|\mathbb{A}|\phi\rangle,$$

de donde se concluye que el vector es ortogonal a S_{λ_0} y por lo tanto está en el complemento ortogonal $\mathbb{A}|\phi\rangle \in \mathcal{N}$ (por hipótesis $|\phi\rangle \in \mathcal{N}$). Esto implica que \mathcal{N} también es un espacio invariante del operador hermítico \mathbb{A} .

Entonces, el espacio V^n puede expresarse como una suma directa de los dos subespacios invariantes respecto al operador lineal \mathbb{A} y su representación matricial en la base de autovectores tendrá la forma de una matriz diagonal a bloques:

$$\langle\psi^j|\mathbb{A}|\psi_i\rangle = A_i^j \rightarrow \begin{pmatrix} Q_1^1 & \cdots & Q_m^1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ Q_1^m & & Q_m^m & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & R_{m+1}^{m+1} & \cdots & R_n^{m+1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & R_{m+1}^n & \cdots & R_n^n \end{pmatrix},$$

donde Q_β^α y R_v^μ son matrices $m \times m$ y $(n-m) \times (n-m)$, respectivamente. La matriz Q_β^α opera en S_{λ_0} mientras que R_v^μ actúa sobre el complemento ortogonal \mathcal{N} .

El polinomio característico de \mathbb{A} puede expresarse como:

$$\mathcal{P}(\lambda) = \det |A_j^i - \lambda\delta_j^i| = 0 \Rightarrow \mathcal{P}(\lambda) = \det |Q_j^i - \lambda\delta_j^i| \det |R_j^i - \lambda\delta_j^i| = 0,$$

y como $\lambda = \lambda_0$ es la raíz múltiple del polinomio característico que anula el $\det |Q_j^i - \lambda\delta_j^i|$, tendremos que

$$\det |Q_j^i - \lambda_0\delta_j^i| = 0 \Rightarrow \mathcal{P}(\lambda) = (\lambda - \lambda_0)^m \mathcal{F}(\lambda), \quad \text{con } \mathcal{F}(\lambda_0) \neq 0,$$

donde λ_0 no es raíz del polinomio $\mathcal{F}(\lambda)$. Ahora bien, para que se cumpla cuando $j = k$, el polinomio característico es

$$j = k \Rightarrow \mathcal{P}(\lambda) = (\lambda - \lambda_0)^k \mathcal{R}(\lambda) \Rightarrow (\lambda - \lambda_0)^m \mathcal{F}(\lambda) = (\lambda - \lambda_0)^k \mathcal{R}(\lambda).$$

Otra vez, λ_0 no es raíz del polinomio $\mathcal{R}(\lambda)$. La ecuación anterior se cumple para todo λ , en particular para $\lambda = \lambda_0$. Por lo tanto:

$$1 = (\lambda - \lambda_0)^{k-m} \frac{\mathcal{R}(\lambda)}{\mathcal{F}(\lambda)}.$$

Es claro que $\lambda = \lambda_0$ obliga a que $k = m$. ◀

4.7.1.2 Autovalores y autovectores de matrices unitarias

Para finalizar esta sección volvamos a considerar la ecuación de autovalores (4.21) y la definición de operadores adjuntos que hemos discutido en la sección 4.2.3. Entonces, si un operador \mathbb{U} es unitario, se cumple que $\mathbb{U}^\dagger = \mathbb{U}^{-1}$ entonces si $|\psi_j\rangle$ es un autovector, normalizado del operador \mathbb{U} , correspondiente a un autovalor λ_j tendremos que la norma al cuadrado de $\mathbb{U}|\psi_j\rangle$ será igual a:

$$\mathbb{U}|\psi_j\rangle = \lambda_j |\psi_j\rangle \Rightarrow \langle \psi^j | \mathbb{U}^\dagger \mathbb{U} |\psi_j\rangle = 1 = \lambda_j^* \lambda_j \langle \psi^j | \psi_j \rangle = \lambda_j^* \lambda_j \Rightarrow \lambda_j = e^{i\varphi_j},$$

con φ_u una función real.

Con lo cual podemos concluir que, necesariamente, los autovalores de los operadores unitarios serán números complejos de módulo 1. Cuando los autovalores son diferentes, digamos $k \neq j$, entonces esta condición implica que $\langle \psi^k | \psi_j \rangle = 0$, con lo cual los autovectores de un operador unitario son ortogonales.

4.7.2 Autovalores y autovectores de matrices similares

Como discutimos en la sección 4.3.3, dos operadores $\tilde{\mathbb{A}}$ y \mathbb{A} que están relacionados por una transformación de similaridad. La ecuación (4.17), $\tilde{\mathbb{A}} = \mathbb{S}^{-1} \mathbb{A} \mathbb{S}$, conecta las distintas representaciones matriciales de un mismo operador, las cuales tendrán la misma traza y el mismo determinante, independientemente de su representación matricial. Tal y como se desprende de ecuación (4.15) la representación matricial de los operadores \mathbb{S} corresponde a la matriz de productos internos de los vectores base $\langle \tilde{\mathbf{e}}^i | \mathbf{e}_k \rangle$.

Ahora nos toca identificar otra propiedad fundamental inherente al operador y no a su representación matricial. Para ello complementaremos los teoremas expuestos allá con el siguiente.

Teorema 6: Dos matrices, A_l^k y \tilde{A}_j^i , $n \times n$, similares tienen el mismo polinomio característico y con ello el mismo conjunto de autovalores.

Demostración: Es inmediato verificar que: $\tilde{\mathbb{A}} - \lambda \mathbb{I} = \mathbb{S}^{-1} \mathbb{A} \mathbb{S} - \lambda \mathbb{I} = \mathbb{S}^{-1} (\mathbb{A} - \lambda \mathbb{I}) \mathbb{S}$, y dado que

$$\det |\tilde{\mathbb{A}} - \lambda \mathbb{I}| = \det |\mathbb{S}^{-1} (\mathbb{A} - \lambda \mathbb{I}) \mathbb{S}| = \det |\mathbb{S}^{-1}| \det |\mathbb{A} - \lambda \mathbb{I}| \det |\mathbb{S}| = \det |\mathbb{A} - \lambda \mathbb{I}|.$$

Por lo tanto, ambos operadores $\tilde{\mathbb{A}}$ y \mathbb{A} , tendrán el mismo polinomio característico y con ello el mismo conjunto de autovalores. ◀

Haremos una lista de todas las propiedades para los autovalores y autovectores de operadores, expuestas en ésta y en la sección anterior.

Sea un operador lineal $\mathbb{A} : \mathbb{V}^n \rightarrow \mathbb{V}^n$ con un polinomio característico que tiene n raíces distintas: $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. Entonces tendremos que:

- Los diferentes autovectores $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$ correspondientes a los $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$, formarán una base para \mathbb{V} .
- La representación matricial del operador $\langle \psi^k | \mathbb{A} | \psi_m \rangle$ en la base de autovectores $\{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_n\rangle\}$, será diagonal:

$$D_m^k = \langle \psi^k | \mathbb{A} | \psi_m \rangle = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n).$$

- Cualquier otra representación matricial, $\langle u^k | \mathbb{A} | u_m \rangle$, del operador \mathbb{A} en otra base de \mathbb{V} , estará relacionada con la representación diagonal mediante una transformación de similaridad:

$$\mathbb{D} = \mathbb{S} \mathbb{A} \mathbb{S}^{-1} \iff \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = S_k^i \langle u^k | \mathbb{A} | u_m \rangle (S^{-1})_j^m, \quad (4.27)$$

donde S_j^m es una matriz, no singular y por lo tanto invertible, de cantidades que relacionan ambas bases, vale decir la matriz de productos internos entre ambas bases.

- Considere $|\tilde{\phi}_i\rangle = \mathbb{U}|\phi_i\rangle$, es decir un vector transformado con un operador unitario $\mathbb{U}^\dagger = \mathbb{U}^{-1}$. Además suponga $\mathbb{A}|\phi_i\rangle = \lambda_i|\phi_i\rangle$ entonces $\tilde{\mathbb{A}}|\tilde{\phi}_i\rangle = \lambda_i|\tilde{\phi}_i\rangle$, con $\tilde{\mathbb{A}} = \mathbb{U}\mathbb{A}\mathbb{U}^\dagger$. Los operadores transformados mediante matrices unitarias tienen los mismos autovalores y autovectores transformados un del otro.

4.7.3 Conjunto completo de observables que comutan

Definición: Diremos que un operador $\mathbb{A} : V^n \rightarrow V^n$ es un *observable* si el conjunto de autovectores $\{|u_{i(\mu)}\rangle\}$ de un operador hermítico \mathbb{A} , forman una base de V^n .

$$\mathbb{A}|u_{i(\mu)}\rangle = a_i|u_{i(\mu)}\rangle \Rightarrow |u_{i(\mu)}\rangle \langle u^{i(\mu)}| = 1 \iff \langle u^{i(\mu)}|u_{j(\nu)}\rangle = \delta_j^i \delta_\nu^\mu,$$

donde el índice μ indica el grado de degeneración del autovalor λ_i .

Un ejemplo trivial de un observable lo constituyen los proyectores $\mathbb{P}_{|\psi\rangle} = |\psi\rangle\langle\psi|$, con $\langle\psi|\psi\rangle = 1$. Claramente, la ecuación de autovalores para un proyector obliga a que tenga dos autovalores 0 y 1. El autovalor nulo es infinitamente degenerado y está asociado a todos los vectores ortogonales a $|\psi\rangle$, mientras que el autovalor 1 corresponde a un autovalor simple y está asociado a todos los vectores colineales al mismo vector $|\psi\rangle$. Esto es:

$$\mathbb{P}_{|\psi\rangle}|\psi\rangle = |\psi\rangle \quad \text{y} \quad \mathbb{P}_{|\psi\rangle}|\phi\rangle = 0 \quad \text{si } \langle\psi|\phi\rangle = 0.$$

Más aún, sea un vector arbitrario $|\varphi\rangle \in V^n$, siempre se podrá expresar como:

$$|\varphi\rangle = \mathbb{P}_{|\psi\rangle}|\varphi\rangle + (\mathbb{I} - \mathbb{P}_{|\psi\rangle})|\varphi\rangle \Rightarrow \mathbb{P}_{|\psi\rangle}(|\varphi\rangle = \mathbb{P}_{|\psi\rangle}|\varphi\rangle + (\mathbb{I} - \mathbb{P}_{|\psi\rangle})|\varphi\rangle),$$

por lo tanto:

$$\mathbb{P}_{|\psi\rangle}|\varphi\rangle = \mathbb{P}_{|\psi\rangle}(\mathbb{P}_{|\psi\rangle}|\varphi\rangle) + (\mathbb{P}_{|\psi\rangle} - \mathbb{P}_{|\psi\rangle}^2)|\varphi\rangle = \mathbb{P}_{|\psi\rangle}|\varphi\rangle \implies \mathbb{P}_{|\psi\rangle}(\mathbb{P}_{|\psi\rangle}|\varphi\rangle) = \mathbb{P}_{|\psi\rangle}|\varphi\rangle,$$

ya que $\mathbb{P}_{|\psi\rangle}^2 = \mathbb{P}_{|\psi\rangle}$, por definición de proyector. Entonces, se deduce que $\mathbb{P}_{|\psi\rangle}|\varphi\rangle$ es un autovector de $\mathbb{P}_{|\psi\rangle}$ con autovalor 1. Igualmente $(\mathbb{I} - \mathbb{P}_{|\psi\rangle})|\varphi\rangle$ es un autovector de $\mathbb{P}_{|\psi\rangle}$ con autovalor 0, y la demostración es inmediata:

$$\mathbb{P}_{|\psi\rangle}(\mathbb{I} - \mathbb{P}_{|\psi\rangle})|\varphi\rangle = (\mathbb{P}_{|\psi\rangle} - \mathbb{P}_{|\psi\rangle}^2)|\varphi\rangle = 0.$$

Para el caso de autoespacios correspondientes a autovalores degenerados se puede definir un observable \mathbb{A} de la forma:

$$\mathbb{A} = \sum_i a_i \mathbb{P}_i \quad \text{con: } \mathbb{P}_i = \left(|\psi_{i(\mu)}\rangle \langle \psi^{i(\mu)}| \right)_i \quad \text{para: } \mu = 1, 2, \dots, k.$$

Para los observables que comutan se tienen los siguientes teoremas:

Teorema 7: Si dos operadores lineales \mathbb{A} y \mathbb{B} , hermíticos, comutan, $[\mathbb{A}, \mathbb{B}] = 0$, y $|\psi\rangle$ es autovector de \mathbb{A} con autovalor σ , entonces $\mathbb{B}|\psi\rangle$ también será autovector de \mathbb{A} con el mismo autovalor σ .

Demostración: La demostración es sencilla:

$$\mathbb{A}|\psi\rangle = a|\psi\rangle \Rightarrow \mathbb{B}(\mathbb{A}|\psi\rangle = \sigma|\psi\rangle) \Rightarrow \mathbb{B}\mathbb{A}|\psi\rangle = \mathbb{A}(\mathbb{B}|\psi\rangle) = \sigma(\mathbb{B}|\psi\rangle). \quad \blacktriangleleft$$

Ahora bien, de esta situación se puede distinguir un par de casos:

- si el autovalor σ es no degenerado los autovectores asociados con este autovalor son, por definición, colineales con $|\psi\rangle$. Por lo tanto $\mathbb{B}|\psi\rangle$, será necesariamente colineal con $|\psi\rangle$. La conclusión a esta afirmación es que NECESARIAMENTE $|\psi\rangle$ es autovector de \mathbb{B} .
- si el autovalor σ es degenerado, $\mathbb{B}|\psi\rangle \in S_\sigma$, es decir $\mathbb{B}|\psi\rangle$ está en el autoespacio S_σ con lo cual S_σ es globalmente invariante bajo la acción de \mathbb{B} .

Teorema 8: Si dos observables \mathbb{A} y \mathbb{B} comutan, $[\mathbb{A}, \mathbb{B}] = 0$, y si $|\psi_1\rangle$ y $|\psi_2\rangle$ son autovectores de \mathbb{A} para autovalores distintos, entonces el elemento de matriz $\langle\psi^1| \mathbb{B} |\psi_2\rangle = 0$

Demostración: Si $\mathbb{A} |\psi_1\rangle = \sigma_1 |\psi_1\rangle$ y $\mathbb{A} |\psi_2\rangle = \sigma_2 |\psi_2\rangle$ entonces:

$$\begin{aligned} 0 &= \langle\psi^1| [\mathbb{A}, \mathbb{B}] |\psi_2\rangle = \langle\psi^1| \mathbb{A}\mathbb{B} - \mathbb{B}\mathbb{A} |\psi_2\rangle = (\langle\psi^1| \mathbb{A}) \mathbb{B} |\psi_2\rangle - \langle\psi^1| \mathbb{B} (\mathbb{A} |\psi_2\rangle) \\ &= \sigma_1 \langle\psi^1| \mathbb{B} |\psi_2\rangle - \sigma_2 \langle\psi^1| \mathbb{B} |\psi_2\rangle = (\sigma_1 - \sigma_2) \langle\psi^1| \mathbb{B} |\psi_2\rangle \Rightarrow \langle\psi^1| \mathbb{B} |\psi_2\rangle = 0. \blacksquare \end{aligned}$$

Teorema 9: Si dos observables \mathbb{A} y \mathbb{B} , son hermíticos, y comutan, $[\mathbb{A}, \mathbb{B}] = 0$, los autovectores $\{|\psi_i\rangle\}$ comunes a \mathbb{A} y \mathbb{B} constituyen una base ortonormal para \mathbf{V}^n .

Demostración: Denotemos los autovectores de \mathbb{A} como $|\psi_{i(\mu)}\rangle$, de tal modo que:

$$\mathbb{A} |\psi_{i(\mu)}\rangle = \sigma_i |\psi_{i(\mu)}\rangle \quad \text{donde } i = 1, 2, \dots, n - k_n + 1 \quad \text{y} \quad \mu = 1, 2, \dots, k_n,$$

k_n indica el orden de la degeneración de un determinado autovalor σ_n .

Dado que \mathbb{A} es un observable, los $|\psi_{i(\mu)}\rangle$ forman una base, claramente:

$$\langle\psi^{i(\mu)}| \psi_{j(\nu)}\rangle = \delta_j^i \delta_\nu^\mu.$$

Dado que los elementos de la matriz $\langle\psi^{i(\nu)}| \mathbb{B} |\psi_{j(\nu)}\rangle = \delta_j^i$, entonces esto quiere decir que los elementos $B_{j(\nu)}^{i(\mu)} = \langle\psi^{i(\mu)}| \mathbb{B} |\psi_{j(\nu)}\rangle$ serán nulos para $i \neq j$, pero no podemos decir nada *a priori* para el caso $\mu \neq \nu$ y $i = j$. En general, al ordenar la base:

$$|\psi_{1(1)}\rangle, |\psi_{1(2)}\rangle, \dots, |\psi_{1(k_1)}\rangle, |\psi_{2(1)}\rangle, |\psi_{2(2)}\rangle, \dots, |\psi_{2(k_2)}\rangle, \dots, |\psi_{3(1)}\rangle, \dots, |\psi_{n-k_n(1)}\rangle,$$

para el caso que consideraremos será:

$$|\psi_{1(1)}\rangle, |\psi_{1(2)}\rangle, |\psi_{1(3)}\rangle, |\psi_{2(1)}\rangle, |\psi_{2(2)}\rangle, |\psi_{3(1)}\rangle, |\psi_{4(1)}\rangle, |\psi_{4(2)}\rangle, |\psi_{5(1)}\rangle.$$

La representación matricial de \mathbb{B} en esa base, $\langle\psi^{i(\mu)}| \mathbb{B} |\psi_{j(\nu)}\rangle$, tendrá la forma de una matriz diagonal a bloques:

$$\left(\begin{array}{ccc|ccccc} B_1^{1(1)} & B_1^{1(1)} & B_1^{1(1)} & 0 & 0 & 0 & 0 & 0 & 0 \\ B_1^{1(2)} & B_1^{1(2)} & B_1^{1(2)} & 0 & 0 & 0 & 0 & 0 & 0 \\ B_1^{1(3)} & B_1^{1(3)} & B_1^{1(3)} & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline B_1^{1(1)} & B_1^{1(2)} & B_1^{1(3)} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_2^{2(1)} & B_2^{2(1)} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & B_2^{2(2)} & B_2^{2(2)} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & B_3^{3(1)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & B_4^{4(1)} & B_4^{4(1)} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & B_4^{4(2)} & B_4^{4(2)} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_5^{5(1)} \end{array} \right)$$

Tal y como hemos mencionado los subespacios: E_1, E_2 , y E_4 corresponden a los autovalores degenerados σ_1, σ_2 , y σ_4 (de orden 3, 2 y 2 respectivamente).

Una vez más surgen dos casos a analizar:

- Si σ_n es un autovalor no degenerado, entonces existe un único autovector asociado a este autovalor (la dimensión del autoespacio es $1 \rightarrow k_j = 1$, y no hace falta). Esto corresponde al ejemplo hipotético anterior para los autovalores simples σ_3 , y σ_5 .
- Si σ_n es un autovalor degenerado, entonces existe un conjunto de autovectores asociados a este autovalor σ_n (en este caso la dimensión del autoespacio es k_n). Como los $|\psi_{j(\mu)}\rangle$ son autovectores de \mathbb{A} su

representación matricial será diagonal a bloques. Ahora bien, como el autoespacio S_a es globalmente invariante bajo la acción de \mathbb{B} y $B_j^{i(\mu)} = \langle \psi^{i(\mu)} | \mathbb{B} | \psi_{j(\mu)} \rangle$ es hermítico, por ser \mathbb{B} hermítico, entonces \mathbb{B} es diagonalizable dentro del bloque que la define. Es decir, se podrá conseguir una base $|\chi_{j(\mu)}\rangle$ tal que la representación matricial de \mathbb{B} en esa base es diagonal

$$B_j^{i(\mu)} = \langle \psi^{i(\mu)} | \mathbb{B} | \psi_{j(\mu)} \rangle \implies \langle \chi^{i(\mu)} | \mathbb{B} | \chi_{j(\mu)} \rangle = \tilde{B}_{j(\mu)}^{i(\mu)} = \beta_{j(\mu)} \delta_j^i,$$

que no es otra cosa que los vectores $|\chi_{j(\mu)}\rangle$ serán autovectores de \mathbb{B}

$$\mathbb{B} |\chi_{j(\mu)}\rangle = \beta_{j(\mu)} |\chi_{j(\mu)}\rangle. \blacksquare$$

Es importante recalcar que los autovectores $|\psi_{j(\mu)}\rangle$ de \mathbb{A} asociados con un autovalor degenerado NO son necesariamente autovectores de \mathbb{B} . Sólo que como \mathbb{B} es hermítico puede ser diagonalizado dentro del autoespacio.

De ahora en adelante denotaremos los autovectores comunes a dos operadores \mathbb{A} y \mathbb{B} con distintos autovalores como $|u_{i|j(\mu)}\rangle$ tal que

$$\mathbb{A} |u_{n|m(\mu)}\rangle = \sigma_n |u_{n|m(\mu)}\rangle \quad \text{y} \quad \mathbb{B} |u_{n|m(\mu)}\rangle = \beta_m |u_{n|m(\mu)}\rangle,$$

donde hemos dejado “espacio” para permitir la degeneración la cual será indicada por el índice μ .

La prueba del inverso del teorema anterior es bien simple.

Teorema 10: Si existe una base de autovectores $\{|u_{j(\mu)}\rangle\}$ comunes a \mathbb{A} y \mathbb{B} , entonces \mathbb{A} y \mathbb{B} commutan, $[\mathbb{A}, \mathbb{B}] = 0$.

Demostración: Es claro que:

$$\begin{aligned} \mathbb{A}\mathbb{B} |u_{n|m(\mu)}\rangle &= \beta_m \mathbb{A} |u_{n|m(\mu)}\rangle = \beta_m \sigma_n |u_{n|m(\mu)}\rangle, \\ \mathbb{B}\mathbb{A} |u_{n|m(\mu)}\rangle &= \sigma_n \mathbb{B} |u_{n|m(\mu)}\rangle = \sigma_n \beta_m |u_{n|m(\mu)}\rangle, \end{aligned}$$

restando miembro a miembro obtenemos de manera inmediata

$$(\mathbb{A}\mathbb{B} - \mathbb{B}\mathbb{A}) |u_{n|m(\mu)}\rangle = [\mathbb{A}, \mathbb{B}] |u_{n|m(\mu)}\rangle = (\beta_m \sigma_n - \sigma_n \beta_m) |u_{n|m(\mu)}\rangle = 0. \blacksquare$$

Definición: Los operadores: $\{\mathbb{A}, \mathbb{B}, \mathbb{C}, \mathbb{D}, \dots\}$ constituye un conjunto completo de observables que commutan si:

1. Los operadores del conjunto commutan entre ellos:

$$[\mathbb{A}, \mathbb{B}] = [\mathbb{A}, \mathbb{C}] = [\mathbb{A}, \mathbb{D}] = [\mathbb{B}, \mathbb{C}] = [\mathbb{B}, \mathbb{D}] = [\mathbb{C}, \mathbb{D}] = \dots = 0$$

2. Al determinar el conjunto de autovalores para los operadores

$$\{\alpha_n, \beta_m, \gamma_k, \delta_l, \dots\}$$

se especifica **de manera única** un único autovector común a todos estos operadores

$$\{\alpha_n, \beta_m, \gamma_k, \delta_l, \dots\} \Rightarrow |u_{n|m|k|l\dots(\mu)}\rangle.$$

4.7.4 Ejemplos

1. Consideremos una vez más la transformación $\mathbb{D} = \mathbb{S}^{-1}\mathbb{A}\mathbb{S}$, descrita en la ecuación (4.27), donde \mathbb{D} es un operador diagonal cuya representación matricial será

$$\mathbb{D} = \mathbb{S}^{-1}\mathbb{A}\mathbb{S} \iff \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) \equiv \lambda_j \delta_j^i = (S^{-1})_k^i A_m^k S_j^m,$$

con $A_m^k = \langle u^k | \mathbb{A} | u_m \rangle$. Tal y como hemos comentado en la sección 4.7, debe haber restricciones sobre el operador \mathbb{A} para que sea diagonalizable. En particular, allí pudimos comprobar que si \mathbb{A} era hermítico,

su representación matricial en la base de autovectores era diagonal. Entonces para este caso supondremos que \mathbb{A} es diagonalizable mediante la transformación (4.27). Claramente esto implica que

$$\mathbb{D} = \mathbb{S}^{-1}\mathbb{A}\mathbb{S} \Rightarrow \mathbb{S}\mathbb{D} = \mathbb{A}\mathbb{S} \iff S_m^i \lambda_j \delta_j^m = A_m^i S_j^m \Rightarrow A_m^i S_j^m = \lambda_j S_j^i,$$

y esta última ecuación sugiere una ecuación de autovalores fijando un valor particular para j . Esto es:

$$A_m^i S_1^m = \lambda_1 S_1^i; \quad A_m^i S_2^m = \lambda_2 S_2^i; \quad A_m^i S_3^m = \lambda_3 S_3^i; \quad \dots A_m^i S_j^m = \lambda_j S_j^i \dots \quad A_m^i S_n^m = \lambda_n S_n^i.$$

Cada una de estas ecuaciones es una ecuación de autovalores para autovectores $S_1^i, S_2^i, S_3^i, \dots, S_n^i$. Vale decir la matriz de transformación S_j^m está construida por columnas de autovectores. Con lo cual al resolver la ecuación de autovalores para la matriz \mathbb{A} es inmediato construir la matriz de transformación \mathbb{S} a partir de los autovectores de \mathbb{A} .

2. Para exemplificar numéricamente el ejemplo anterior, consideremos la siguiente matriz real y simétrica:

$$\mathbb{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

Los autovalores y autovectores para \mathbb{A} son respectivamente:

$$P(\lambda) = -(\lambda - 4)(\lambda + 2)^2 = 0 \Rightarrow \lambda_1 = 4, \lambda_2 = -2, \lambda_3 = -2,$$

$$|u_1\rangle = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad |u_2\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |u_3\rangle = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

El lector debe verificar que este conjunto de vectores son mutuamente ortogonales porque claramente la matriz es simétrica. Por lo tanto, la matriz \mathbb{A} se puede diagonalizar a través de la transformación $\mathbb{C}^{-1}\mathbb{A}\mathbb{C}$, donde \mathbb{C} se construye con los autovectores normalizados de \mathbb{A} como columnas.

$$\mathbb{C} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Entonces, a pesar que los autovalores de \mathbb{A} son degenerados ($\lambda = 4, -2, -2$) sus tres autovectores son linealmente independientes, y se tiene que:

$$\mathbb{C}^{-1}\mathbb{A}\mathbb{C} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \\ 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

3. Dada la siguiente matriz:

$$\mathbb{A} = \begin{pmatrix} 2 & 2 & -2 \\ 2 & -1 & 4 \\ -2 & 4 & -1 \end{pmatrix},$$

Aquí también se cumple que $\mathbb{A} = \mathbb{A}^T$, con lo cual se puede diagonalizar. Calculemos nuevamente una matriz \mathbb{C} que permita diagonalizar a la matriz \mathbb{A} .

Primeramente procedemos a calcular los autovalores de \mathbb{A} .

$$P(\lambda) = \begin{vmatrix} 2 - \lambda & 2 & -2 \\ 2 & -1 - \lambda & 4 \\ -2 & 4 & -1 - \lambda \end{vmatrix} = -(\lambda - 3)^2(\lambda + 6) = 0.$$

- Para $\lambda = -6$:

$$\begin{pmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = -6 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \Rightarrow \begin{cases} 14x^1 + 2x^2 - 2x^3 = 0 \\ 2x^1 + 11x^2 + 4x^3 = 0 \\ -2x^1 + 4x^2 + 11x^3 = 0 \end{cases}$$

Un autovector puede ser:

$$|u_1\rangle = \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \Rightarrow |\hat{u}_1\rangle = \frac{1}{3} \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix}.$$

- Para el autovalor degenerado $\lambda = 3$, tenemos:

$$\begin{pmatrix} -1 & 2 & -2 \\ 2 & -4 & 4 \\ -2 & 4 & -4 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} = 3 \begin{pmatrix} x^1 \\ x^2 \\ x^3 \end{pmatrix} \Rightarrow \begin{cases} -4x^1 + 2x^2 - 2x^3 = 0 \\ 2x^1 - 7x^2 + 4x^3 = 0 \\ -2x^1 + 4x^2 - 7x^3 = 0 \end{cases}$$

En este caso podemos tomar como autovectores:

$$|u_2\rangle = \begin{pmatrix} 1 \\ 0 \\ -1/2 \end{pmatrix} \Rightarrow |\hat{u}_2\rangle = \frac{2}{\sqrt{5}} \begin{pmatrix} 1 \\ 0 \\ -1/2 \end{pmatrix}, |u_3\rangle = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \Rightarrow |\hat{u}_3\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

El lector debe comprobar que en este caso los autovectores NO son mutuamente ortogonales. Construimos la matriz \mathbb{C} :

$$\mathbb{C} = \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{3} & 0 & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

De manera que:

$$\mathbb{C}^{-1} \mathbb{A} \mathbb{C} = \begin{pmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{4\sqrt{5}}{9} & \frac{\sqrt{5}}{9} & -\frac{\sqrt{5}}{9} \\ \frac{2\sqrt{2}}{9} & \frac{5\sqrt{2}}{9} & \frac{4\sqrt{2}}{9} \end{pmatrix} \begin{pmatrix} 2 & 2 & -2 \\ 2 & -1 & 4 \\ -2 & 4 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & 0 \\ -\frac{2}{3} & 0 & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & -\frac{1}{\sqrt{5}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

4. La siguiente matriz es un ejemplo de una matriz hermítica o autoadjunta:

$$\mathbb{A} = \begin{pmatrix} 1 & -1+2i & i \\ -1-2i & 2 & -1 \\ -i & -1 & 3 \end{pmatrix} \Rightarrow \mathbb{A}^\dagger = \mathbb{A}$$

Los autovalores de esta matriz son:

$$P(\lambda) = -(\lambda - 4)(\lambda^2 - 2\lambda - 4) = 0 \Rightarrow \lambda_1 = 4, \lambda_2 = 1 + \sqrt{5}, \lambda_3 = 1 - \sqrt{5}.$$

Los tres autovalores *reales* generan los siguientes autovectores:

$$\lambda_1 = 4 \Rightarrow |u_1\rangle = \begin{pmatrix} 1 \\ -1-i \\ 1 \end{pmatrix},$$

$$\lambda_2 = 1 + \sqrt{5} \Rightarrow |u_2\rangle = \frac{1}{3} \begin{pmatrix} 3 \\ 1 - \sqrt{5}i \\ -\sqrt{5} - 2 - (1 + \sqrt{5})i \end{pmatrix}, \quad \lambda_3 = 1 - \sqrt{5} \Rightarrow |u_3\rangle = \frac{1}{3} \begin{pmatrix} 3 \\ 1 + \sqrt{5}i \\ \sqrt{5} - 2 - (1 - \sqrt{5})i \end{pmatrix}.$$

Estos vectores normalizados son:

$$|\hat{u}_1\rangle = \frac{|u_1\rangle}{\sqrt{\langle u_1 | u_1 \rangle}} = \frac{1}{2} \begin{pmatrix} 1 \\ -1-i \\ 1 \end{pmatrix}, \quad |\hat{u}_2\rangle = \frac{|u_2\rangle}{\sqrt{\langle u_2 | u_2 \rangle}} = \frac{3}{\sqrt{30+6\sqrt{5}}} \begin{pmatrix} 3 \\ 1-\sqrt{5}i \\ -\sqrt{5}-2-(1+\sqrt{5})i \end{pmatrix},$$

$$|\hat{u}_3\rangle = \frac{|u_3\rangle}{\sqrt{\langle u_3 | u_3 \rangle}} = \frac{3}{\sqrt{30-6\sqrt{5}}} \begin{pmatrix} 3 \\ 1+\sqrt{5}i \\ \sqrt{5}-2-(1-\sqrt{5})i \end{pmatrix}.$$

Se puede demostrar que los autovectores son ortogonales:

$$\langle u_1 | u_2 \rangle = (u_1)^\dagger u_2 = 0 \Rightarrow \begin{pmatrix} 1 & -1+i & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1-\sqrt{5}i \\ -\sqrt{5}-2-(1+\sqrt{5})i \end{pmatrix} = 0$$

$$\langle u_1 | u_3 \rangle = (u_1)^\dagger u_3 = 0 \Rightarrow \begin{pmatrix} 1 & -1+i & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 1+\sqrt{5}i \\ \sqrt{5}-2-(1-\sqrt{5})i \end{pmatrix} = 0$$

$$\langle u_2 | u_3 \rangle = (u_2)^\dagger u_3 = 0 \Rightarrow \begin{pmatrix} 3 & 1+\sqrt{5}i & -\sqrt{5}-2+(1+\sqrt{5})i \end{pmatrix} \begin{pmatrix} 3 \\ 1+\sqrt{5}i \\ \sqrt{5}-2-(1-\sqrt{5})i \end{pmatrix} = 0$$

La matriz \mathbb{A} es entonces diagonalizable si construimos la siguiente matriz a partir de los autovectores normalizados:

$$\mathbb{P} = \begin{pmatrix} \frac{1}{2} & \frac{9}{\sqrt{30+6\sqrt{5}}} & \frac{9}{\sqrt{30-6\sqrt{5}}} \\ \frac{-1-i}{2} & \frac{3(1-\sqrt{5}i)}{\sqrt{30+6\sqrt{5}}} & \frac{3(1+\sqrt{5}i)}{\sqrt{30-6\sqrt{5}}} \\ \frac{1}{2} & \frac{3(-\sqrt{5}-2-(1+\sqrt{5})i)}{\sqrt{30+6\sqrt{5}}} & \frac{3(\sqrt{5}-2-(1-\sqrt{5})i)}{\sqrt{30-6\sqrt{5}}} \end{pmatrix}.$$

Si elegimos la matriz diagonal a partir de los autovalores:

$$\mathbb{D} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1+\sqrt{5} & 0 \\ 0 & 0 & 1-\sqrt{5} \end{pmatrix},$$

entonces resulta que es posible factorizar la matriz \mathbb{A} de la siguiente forma (Para el lector se deja la comprobación)

$$\mathbb{A} = \mathbb{P}\mathbb{D}\mathbb{P}^\dagger.$$

5. Dada la matriz:

$$\mathbb{A} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{i}{\sqrt{2}} & \frac{i}{\sqrt{2}} & 0 \\ 0 & 0 & i \end{pmatrix}.$$

Esta matriz es unitaria, ya que: $\mathbb{A}^\dagger = \mathbb{A}^{-1}$.

Los autovalores se obtienen de la manera usual:

$$P(\lambda) = (\lambda - i) \left[2\lambda^2 - \sqrt{2}(1+i)\lambda + 2i \right] = 0,$$

es decir:

$$\lambda_1 = \frac{\sqrt{2}(1+i) - 2\sqrt{-3i}}{4}, \lambda_2 = \frac{\sqrt{2}(1+i) + 2\sqrt{-3i}}{4}, \lambda_3 = i.$$

Notemos que los valores propios están normalizados a la unidad:

$$\begin{aligned} \lambda_1 \lambda_1^* &= \left(\frac{\sqrt{2}(1+i) - 2\sqrt{-3i}}{4} \right) \left(\frac{\sqrt{2}(1-i) - 2\sqrt{-3i}}{4} \right) = 1 \\ \lambda_2 \lambda_2^* &= \left(\frac{\sqrt{2}(1+i) + 2\sqrt{-3i}}{4} \right) \left(\frac{\sqrt{2}(1-i) + 2\sqrt{-3i}}{4} \right) = 1 \\ \lambda_3 \lambda_3^* &= i(-i) = 1 \end{aligned}$$

Estos tres autovalores generan los siguientes autovectores:

$$|u_1\rangle_{\lambda_1} = \begin{pmatrix} 1 \\ \frac{i-\sqrt{-6i}-1}{2} \\ 0 \end{pmatrix}, |u_2\rangle_{\lambda_2} = \begin{pmatrix} 1 \\ \frac{i+\sqrt{-6i}-1}{2} \\ 0 \end{pmatrix}, |u_3\rangle_{\lambda_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

Se puede demostrar que los autovectores son ortogonales:

$$\begin{aligned} \langle u_1 | u_2 \rangle &= (u_1)^\dagger u_2 = 0 \Rightarrow \begin{pmatrix} 1 & \frac{-i-\sqrt{6i}-1}{2} & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \frac{i+\sqrt{-6i}-1}{2} \\ 0 \end{pmatrix} = 0, \\ \langle u_1 | u_3 \rangle &= (u_1)^\dagger u_3 = 0 \Rightarrow \begin{pmatrix} 1 & \frac{-i-\sqrt{6i}-1}{2} & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0, \\ \langle u_2 | u_3 \rangle &= (u_2)^\dagger u_3 = 0 \Rightarrow \begin{pmatrix} 1 & \frac{-i+\sqrt{6i}-1}{2} & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0. \end{aligned}$$

Con los autovectores normalizados construimos la matriz \mathbb{U} , que también será unitaria, y la matriz diagonal \mathbb{D} con los autovalores.

$$\mathbb{U} = \begin{pmatrix} \frac{1}{\sqrt{3+\sqrt{3}}} & \frac{1}{\sqrt{3-\sqrt{3}}} & 0 \\ -\frac{(1-i)(\sqrt{3}+1)}{2\sqrt{3+\sqrt{3}}} & \frac{(1-i)(\sqrt{3}-1)}{2\sqrt{3-\sqrt{3}}} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbb{D} = \begin{pmatrix} \frac{\sqrt{2}(1+i)-2\sqrt{-3i}}{4} & 0 & 0 \\ 0 & \frac{\sqrt{2}(1+i)+2\sqrt{-3i}}{4} & 0 \\ 0 & 0 & i \end{pmatrix}.$$

6. Considere que el espacio de estados para un determinado sistema físico viene expandido por una base ortonormal $\{|\xi_1\rangle, |\xi_2\rangle, |\xi_3\rangle\}$. Definimos dos operadores \mathbb{L}_z y \mathbb{S} de la siguiente manera:

$$\mathbb{L}_z |\xi_1\rangle = |\xi_1\rangle, \mathbb{L}_z |\xi_2\rangle = 0, \mathbb{L}_z |\xi_3\rangle = -|\xi_3\rangle, \mathbb{S} |\xi_1\rangle = |\xi_3\rangle, \mathbb{S} |\xi_2\rangle = |\xi_2\rangle, \mathbb{S} |\xi_3\rangle = |\xi_1\rangle.$$

En la base ortonormal $\{|\xi_1\rangle, |\xi_2\rangle, |\xi_3\rangle\}$ las representaciones matriciales para $\mathbb{L}_z, \mathbb{L}_z^2, \mathbb{S}$ y \mathbb{S}^2 serán las

siguientes:

$$\langle \xi^i | \mathbb{L}_z | \xi_j \rangle = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \langle \xi^i | \mathbb{L}_z^2 | \xi_j \rangle = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\langle \xi^i | \mathbb{S} | \xi_j \rangle = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \langle \xi^i | \mathbb{S}^2 | \xi_j \rangle = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Es claro que estas matrices son reales y simétricas y, por lo tanto, son hermíticas y, al ser el espacio de dimensión finita, deben ser diagonalizables y sus autovectores formarán base para ese espacio. Por lo tanto, \mathbb{L}_z , \mathbb{L}_z^2 , \mathbb{S} y \mathbb{S}^2 son observables. Ahora bien: ¿Cuál será la forma más general de una representación matricial de un operador que commute con \mathbb{L}_z ?

Notamos que los vectores de la base ortonormal $\{| \xi_1 \rangle, | \xi_2 \rangle, | \xi_3 \rangle\}$ son autovectores para \mathbb{L}_z con autovalores $\{1, 0, -1\}$, con lo cual su representación matricial tiene que ser diagonal. Recuerde que si dos observables \mathbb{A} y \mathbb{B} comutan, $[\mathbb{A}, \mathbb{B}] = 0$, y si $|\psi_1\rangle$ y $|\psi_2\rangle$ son autovectores de \mathbb{A} para autovalores distintos, entonces el elemento de matriz $\langle \psi^1 | \mathbb{B} | \psi_2 \rangle = 0$, con lo cual:

$$[\mathbb{M}, \mathbb{L}_z] = 0 \iff \langle \xi^i | \mathbb{M} | \xi_j \rangle = \begin{pmatrix} M_1^1 & 0 & 0 \\ 0 & M_2^2 & 0 \\ 0 & 0 & M_3^3 \end{pmatrix}.$$

Esto se desprende de manera directa de:

$$0 = \langle \xi^i | [\mathbb{M}, \mathbb{L}_z] | \xi_j \rangle = \langle \xi^i | \mathbb{M} \mathbb{L}_z - \mathbb{L}_z \mathbb{M} | \xi_j \rangle = (\lambda_j - \lambda_i) \langle \xi^i | \mathbb{M} | \xi_j \rangle, \quad \text{con } (\lambda_j - \lambda_i) \neq 0 \quad \text{para } i \neq j.$$

Si nos planteamos la misma pregunta para \mathbb{L}_z^2 , vemos que sus autovalores son $\{1, 0\}$. Esto es:

$$\mathbb{L}_z^2 | \xi_1 \rangle = | \xi_1 \rangle; \quad \mathbb{L}_z^2 | \xi_2 \rangle = 0; \quad \mathbb{L}_z^2 | \xi_3 \rangle = | \xi_3 \rangle,$$

con lo cual tendremos que la representación matricial de ese operador que commute con \mathbb{L}_z^2 , no será diagonal, es decir:

$$[\mathbb{N}, \mathbb{L}_z^2] = 0 \iff \langle \xi^i | \mathbb{N} | \xi_j \rangle = \begin{pmatrix} N_1^1 & 0 & N_3^1 \\ 0 & N_2^2 & 0 \\ N_1^3 & 0 & N_3^3 \end{pmatrix},$$

ya que:

$$0 = \langle \xi^1 | [\mathbb{N}, \mathbb{L}_z^2] | \xi_3 \rangle \Rightarrow \langle \xi^1 | \mathbb{N} | \xi_3 \rangle = \langle \xi^1 | \mathbb{N} | \xi_3 \rangle,$$

y vale para cualquier elemento N_3^1 (y equivalentemente para N_1^3).

Adicionalmente, si ordenamos la base de autovectores de \mathbb{L}_z , como $\{| \xi_1 \rangle, | \xi_3 \rangle, | \xi_2 \rangle\}$, tendremos entonces como representación matricial diagonal a bloques, correspondiente a un autovalor degenerado 1, a:

$$\langle \xi^i | \tilde{\mathbb{N}} | \xi_j \rangle = \begin{pmatrix} N_1^1 & N_3^1 & 0 \\ N_1^3 & N_2^2 & 0 \\ 0 & 0 & N_3^3 \end{pmatrix}.$$

Autovectores	Autovalor \mathbb{L}_z^2	Autovalor \mathbb{S}
$ q_1 \rangle = \xi_2 \rangle$	0	1
$ q_2 \rangle = \frac{1}{\sqrt{2}} (\xi_1 \rangle + \xi_3 \rangle)$	1	1
$ q_3 \rangle = \frac{1}{\sqrt{2}} (\xi_1 \rangle - \xi_3 \rangle)$	1	-1

Figura 4.1: Dado que no hay líneas repetidas \mathbb{L}_z^2 y \mathbb{S} forman un CCOC.

Finalmente, la representación matricial, más general, de un operador que commute con \mathbb{S}^2 es:

$$[\mathbb{P}, \mathbb{S}^2] = 0 \Leftrightarrow \langle \xi^i | \mathbb{P} | \xi_j \rangle = \begin{pmatrix} P_1^1 & P_2^1 & P_3^1 \\ P_1^2 & P_2^2 & P_3^2 \\ N_1^3 & P_2^3 & P_3^3 \end{pmatrix}.$$

Ahora intentaremos construir una base común de autovectores para \mathbb{L}_z^2 y \mathbb{S} . Para ello notamos que $|\xi_2\rangle$ es un autovector común a \mathbb{L}_z^2 y \mathbb{S} , por lo tanto, existirá un subespacio expandido por: $\{|\xi_1\rangle, |\xi_3\rangle\}$. En ese subespacio las representaciones matriciales para \mathbb{L}_z^2 y \mathbb{S} , serán:

$$\langle \xi^i | \mathbb{L}_z^2 | \xi_j \rangle_{S_{13}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \langle \xi^i | \mathbb{S} | \xi_j \rangle_{S_{13}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Acto seguido planteamos el problema de autovalores para \mathbb{S} , utilizando la base $\{|\xi_1\rangle, |\xi_3\rangle\}$, esto es:

$$\mathbb{S} |q_j\rangle = \lambda_j |u_j\rangle \Rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \lambda \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \Rightarrow \begin{cases} |q_2\rangle = \frac{1}{\sqrt{2}}(|\xi_1\rangle + |\xi_3\rangle) \\ |q_3\rangle = \frac{1}{\sqrt{2}}(|\xi_1\rangle - |\xi_3\rangle) \end{cases}$$

con lo cual tendremos los resultados mostrados en la tabla 4.1. Es decir, hemos expresado los autovectores de \mathbb{S} , $|q_2\rangle$ y $|q_3\rangle$, como combinación lineal de la base $\{|\xi_1\rangle, |\xi_3\rangle\}$.

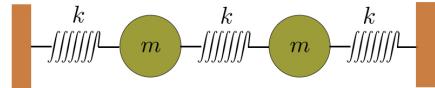
Es costumbre representar esta base común de autovectores de \mathbb{L}_z^2 y \mathbb{S} haciendo referencia a los autovalores asociados a los operadores

$$|q_1\rangle \Leftrightarrow |0;1\rangle, \quad |q_2\rangle \Leftrightarrow |1;1\rangle, \quad |q_3\rangle \Leftrightarrow |1;-1\rangle.$$

7. Consideremos otro ejemplo proveniente de la Mecánica Clásica. Se trata de dos osciladores armónicos, de igual masa, acoplados con resortes con la misma constante elástica k . La ecuaciones de movimiento para este sistema son:

$$m\ddot{x}_1 + kx_1 - k(x_2 - x_1) = 0,$$

$$m\ddot{x}_2 + kx_2 + k(x_2 - x_1) = 0,$$



Podremos expresar estas ecuaciones en forma de operadores:

$$\mathbb{D}|x\rangle = 0 \Leftrightarrow \begin{pmatrix} m \frac{d^2}{dt^2} + 2k & -k \\ -k & m \frac{d^2}{dt^2} + 2k \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = 0.$$

Si pensamos esta ecuación como una ecuación de autovalores, el autovalor es claramente $\lambda = 0$, y como las masas y las constantes elásticas son iguales podemos intercambiar las partículas y la física (las ecuaciones de movimiento) no cambian. Esto se puede expresar matemáticamente como el operador permutación de las partículas:

$$\mathbb{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} x^2 \\ x^1 \end{pmatrix}.$$

Es inmediato comprobar que $[\mathbb{D}, \mathbb{P}] = 0$, con lo cual existirá una combinación lineal de autovectores de \mathbb{D} (asociados con el autovalor $\lambda = 0$) los cuales también serán autovectores de \mathbb{P} . Para ello procedamos a calcular los autovalores y autovectores de \mathbb{P} :

$$\mathbb{P}|x\rangle = \lambda|x\rangle \Rightarrow \begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda \pm 1 \Leftrightarrow |\hat{e}_1\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \quad |\hat{e}_2\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Fácilmente podemos expresar el vector posición como una combinación lineal de estos dos autovectores

de \mathbb{P} , esto es:

$$\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \frac{\xi_1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{\xi_2}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Rightarrow \begin{cases} \xi_1 = \frac{1}{\sqrt{2}}(x_1 + x_2) \\ \xi_2 = \frac{1}{\sqrt{2}}(x_1 - x_2) \end{cases}$$

Es claro que

$$|u_1\rangle = \frac{1}{\sqrt{2}}(x_1 + x_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{y} \quad |u_2\rangle = \frac{1}{\sqrt{2}}(x_1 - x_2) \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

son autovectores de \mathbb{P} y \mathbb{D} .

8. Un operador cantidad de movimiento generalizado se define como aquel conjunto de operadores hermíticos que cumplen con:

$$[\mathbb{J}_x, \mathbb{J}_y] = i\hbar \mathbb{J}_z, \quad [\mathbb{J}_y, \mathbb{J}_z] = i\hbar \mathbb{J}_x, \quad [\mathbb{J}_z, \mathbb{J}_x] = i\hbar \mathbb{J}_y, \quad \text{es decir: } [\mathbb{J}_i, \mathbb{J}_j] = i\hbar \epsilon_{ijk} \mathbb{J}_k,$$

con ϵ_{ijk} el símbolo de Levi-Civita (Aquí los índices repetidos NO indican suma).

Adicionalmente, definimos los siguientes operadores:

$$\mathbb{J}^2 = \mathbb{J}_x^2 + \mathbb{J}_y^2 + \mathbb{J}_z^2, \quad \mathbb{J}_+ = \mathbb{J}_x + i\mathbb{J}_y, \quad \mathbb{J}_- = \mathbb{J}_x - i\mathbb{J}_y.$$

Queremos demostrar que:

$$[\mathbb{J}^2, \mathbb{J}_+] = [\mathbb{J}^2, \mathbb{J}_-] = [\mathbb{J}^2, \mathbb{J}_z] = 0.$$

Para probar esta propiedad se puede demostrar de forma genérica que $[\mathbb{J}_k^2, \mathbb{J}_m] = 0$, con $k, m = 1, 2, 3 \equiv x, y, z$, esto es:

$$[\mathbb{J}_k^2, \mathbb{J}_m] = [\mathbb{J}_k \mathbb{J}_k, \mathbb{J}_m] = \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m - \mathbb{J}_m \mathbb{J}_k \mathbb{J}_k = \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m - (i\hbar \epsilon_{mkl} \mathbb{J}_l + \mathbb{J}_k \mathbb{J}_m) \mathbb{J}_k,$$

con lo cual:

$$[\mathbb{J}_k^2, \mathbb{J}_m] = \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m - i\hbar \epsilon_{mkl} \mathbb{J}_l \mathbb{J}_k - \mathbb{J}_k (i\hbar \epsilon_{mkn} \mathbb{J}_n + \mathbb{J}_k \mathbb{J}_m),$$

y claramente se anula por cuanto los índices no suman pero si son mudos, y $\epsilon_{mkl} = -\epsilon_{mlk}$.

$$[\mathbb{J}_k^2, \mathbb{J}_m] = \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m - i\hbar \epsilon_{mkl} \mathbb{J}_l \mathbb{J}_k - i\hbar \epsilon_{mkn} \mathbb{J}_k \mathbb{J}_n - \mathbb{J}_k \mathbb{J}_k \mathbb{J}_m,$$

al comutar los cuadrados de las componentes con cualquiera de las componentes, y dado que los conmutadores son lineales entonces queda demostrado que:

$$[\mathbb{J}^2, \mathbb{J}_\pm] = [\mathbb{J}_x^2 + \mathbb{J}_y^2 + \mathbb{J}_z^2, \mathbb{J}_x \pm i\mathbb{J}_y] = [\mathbb{J}_y^2, \mathbb{J}_x] + [\mathbb{J}_z^2, \mathbb{J}_x] \pm i [\mathbb{J}_x^2, \mathbb{J}_y] \pm i [\mathbb{J}_z^2, \mathbb{J}_y] = 0.$$

9. Si definimos los autovectores comunes a \mathbb{J}^2 y \mathbb{J}_z como $|j, m\rangle$ de la siguiente manera:

$$\mathbb{J}^2 |j, m\rangle = j(j+1)\hbar^2 |j, m\rangle, \quad \mathbb{J}_z |j, m\rangle = m\hbar |j, m\rangle, \quad \text{con: } \langle j, m | j', m' \rangle = \delta_{jj'} \delta_{mm'},$$

y adicionalmente tenemos que:

$$\mathbb{J}_- |j, m\rangle = \hbar \sqrt{j(j+1) - m(m-1)} |j, m-1\rangle, \quad \mathbb{J}_+ |j, m\rangle = \hbar \sqrt{j(j+1) - m(m+1)} |j, m+1\rangle.$$

Si se supone (es fácil demostrarlo) que $-j \leq m \leq j$. Esto quiere decir que dado algún valor j, m varían entre $-j$ y j de uno en uno, esto es: $m = -j, -j+1, -j+2, \dots, j-2, j-1, j$. Supongamos ahora que $j = \frac{1}{2}$.

Busquemos:

- (a). La representación matricial para: $\mathbb{J}_z, \mathbb{J}_-, \mathbb{J}_+, \mathbb{J}^2$, en la base de autovectores de \mathbb{J}_z y \mathbb{J}^2 .

Si $|j, m\rangle$ son autovectores de \mathbb{J}^2 y \mathbb{J}_z su representación matricial será diagonal y como m varía entre $-j$ y j con incrementos de 1 tendremos que serán matrices 2×2 . La base ortogonal de autovectores

será: $\{ |\frac{1}{2}, -\frac{1}{2}\rangle, |\frac{1}{2}, \frac{1}{2}\rangle \}$.

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}_z | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}_z | \frac{1}{2}, -\frac{1}{2} \rangle \\ \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}_z | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}_z | \frac{1}{2}, -\frac{1}{2} \rangle \end{pmatrix} \equiv \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}^2 | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}^2 | \frac{1}{2}, -\frac{1}{2} \rangle \\ \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}^2 | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}^2 | \frac{1}{2}, -\frac{1}{2} \rangle \end{pmatrix} \equiv \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

La representación matricial para $\mathbb{J}_-, \mathbb{J}_+$ obviamente no será diagonal:

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}_+ | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}_+ | \frac{1}{2}, -\frac{1}{2} \rangle \\ \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}_+ | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}_+ | \frac{1}{2}, -\frac{1}{2} \rangle \end{pmatrix} \equiv \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

$$\begin{pmatrix} \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}_- | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, \frac{1}{2} | \mathbb{J}_- | \frac{1}{2}, -\frac{1}{2} \rangle \\ \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}_- | \frac{1}{2}, \frac{1}{2} \rangle & \langle \frac{1}{2}, -\frac{1}{2} | \mathbb{J}_- | \frac{1}{2}, -\frac{1}{2} \rangle \end{pmatrix} \equiv \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

(b). Calculemos los autovalores y autovectores para: $\mathbb{J}_z, \mathbb{J}_-, \mathbb{J}_+, \mathbb{J}^2$.

Otra vez, $\{ |\frac{1}{2}, -\frac{1}{2}\rangle, |\frac{1}{2}, \frac{1}{2}\rangle \}$ son autovectores de \mathbb{J}^2 y \mathbb{J}_z . En el caso de \mathbb{J}^2 con un autovalor de $\frac{3}{4}\hbar^2$ para ambos autovectores y en el caso de \mathbb{J}_z los autovalores serán $\pm\frac{\hbar}{2}$ respectivamente. Para $\mathbb{J}_-, \mathbb{J}_+$ no tendrán autovalor distinto de cero en esta base.

4.7.5 Practicando con Maxima

Consideremos uno de los ejemplos anteriores donde:

```
(%i1) A:matrix([1,0,3], [0,-2,0], [3,0,1]);
(%o1) ⓥ
(%i2) eigenvectors(A);
(%o2) [[[[-2,4],[2,1]],[[[1,0,-1],[0,1,0]]],[[1,0,1]]]]
```

El resultado es una lista con los autovalores y su multiplicidad, y para cada autovalor los autovectores como sublistas. Es necesario manipular estas sublistas para obtener los autovectores. Entonces, es mejor escribir:

```
(%i3) [val,vec]:eigenvectors(A);
(%o3) [[[[-2,4],[2,1]],[[[1,0,-1],[0,1,0]]],[[1,0,1]]]]
```

Por lo tanto, una lista con los autovalores es la siguiente:

```
(%i4) autovalores:val[1];
(%o4) [-2,4]
```

Notemos el orden: los dos primeros vectores corresponden al autovalor -2 y el tercero al autovalor 4 .

Los autovectores como listas son:

```
(%i5) L1:vec[1];L2:vec[2];
(%o5) [[1,0,-1],[0,1,0]]
(%o5) [[1,0,1]]
```

Pero resulta conveniente separar cada autovector.

```
(%i6) L1[1];L1[2];L2[1];
(%o6) [1,0,-1]
(%o7) [0,1,0]
(%o8) [1,0,1]
```

Vamos ahora a normalizar los autovectores:

```
(%i9) V1:L1[1]/sqrt(L1[1].L1[1]);V2:L1[2]/sqrt(L1[2].L1[2]);V3:L2[1]/sqrt(L2[1].L2[1]);
(%o9) ⓥ
(%o10) [0,1,0]
(%o11) ⓥ
```

Hacemos una lista con los vectores normalizados:

```
(%i12) Lvec:[V1,V2,V3];
(%o12) ⓥ
```

Y construimos la matriz \mathbb{C} :

$$(\%i13) \mathbf{C}:\text{transpose}(\text{apply}('matrix,Lvec));$$

$$(\%o13) \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

Para finalmente comprobar:

$$(\%i14) \text{transpose}(\mathbf{C}) . \mathbf{A} . \mathbf{C};$$

$$(\%o14) \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

Consideremos ahora otro de los ejemplos, donde teníamos la siguiente matriz:

$$(\%i15) \mathbf{A}:\text{matrix}([1,-1+2*i,%i],[-1-2*i,2,-1],[-%i,-1,3]);$$

$$(\%o15) \begin{pmatrix} 1 & 2i-1 & i \\ -2i-1 & 2 & -1 \\ -i & -1 & 3 \end{pmatrix}$$

Es fácil ver que se trata de una matriz hermítica

$$(\%i16) \text{transpose}(\text{conjugate}(\mathbf{A}))=\mathbf{A};$$

$$(\%o16) \begin{pmatrix} 1 & 2i-1 & i \\ -2i-1 & 2 & -1 \\ -i & -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2i-1 & i \\ -2i-1 & 2 & -1 \\ -i & -1 & 3 \end{pmatrix}$$

Calculamos los autovalores y autovectores:

$$(\%i17) [\mathbf{val},\mathbf{vec}]:\text{eigenvectors}(\mathbf{A})$$$

Al primer elemento de la lista, le asignaremos la variable `autovalores`, como se muestra a continuación:

$$(\%i18) \text{autovalores}:\mathbf{val}[1];$$

$$(\%o18) [1 - \sqrt{5}, \sqrt{5} + 1, 4]$$

Ahora procedemos a aislar los diferentes autovectores, es importante tener en mente el orden para utilizar las etiquetas apropiadas.

$$(\%i19) \mathbf{L1}:\mathbf{vec}[1]\$\mathbf{L1}[1]; \mathbf{L2}:\mathbf{vec}[2]\$\mathbf{L2}[1]; \mathbf{L3}:\mathbf{vec}[3]\$\mathbf{L3}[1];$$

$$(\%o20) \left[1, \frac{\sqrt{5}i + 1}{3}, \frac{(\sqrt{5}-1)i + \sqrt{5}-2}{3} \right]$$

$$(\%o22) \left[1, -\frac{\sqrt{5}i - 1}{3}, -\frac{(\sqrt{5}+1)i + \sqrt{5}+2}{3} \right]$$

$$(\%o24) [1, -i - 1, 1]$$

Los autovectores normalizados son:

$$(\%i25) \mathbf{u1}:\mathbf{L1}[1]/\text{sqrt}(\mathbf{L1}[1].\mathbf{L1}[1]),\text{ratsimp};$$

```
(%o25) 
$$\left[ \frac{3}{\sqrt{-(4\sqrt{5}-14)i-2\sqrt{5}+8}}, \frac{\sqrt{5}i+1}{\sqrt{-(4\sqrt{5}-14)i-2\sqrt{5}+8}}, \frac{(\sqrt{5}-1)i+\sqrt{5}-2}{\sqrt{-(4\sqrt{5}-14)i-2\sqrt{5}+8}} \right]$$

(%i26) u2:L2[1]/sqrt(L2[1].L2[1]),ratsimp;
(%o26) 
$$\left[ \frac{3}{\sqrt{(4\sqrt{5}+14)i+2\sqrt{5}+8}}, -\frac{\sqrt{5}i-1}{\sqrt{(4\sqrt{5}+14)i+2\sqrt{5}+8}}, -\frac{(\sqrt{5}+1)i+\sqrt{5}+2}{\sqrt{(4\sqrt{5}+14)i+2\sqrt{5}+8}} \right]$$

(%i27) u3:L3[1]/sqrt(L3[1].L3[1]),ratsimp;
(%o27) 
$$\left[ \frac{1}{\sqrt{2i+2}}, -\frac{i+1}{\sqrt{2i+2}}, \frac{1}{\sqrt{2i+2}} \right]$$

```

Estos vectores serán ortogonales, como podemos ver:

```
(%i28) conjugate(u1).u2,ratsimp;conjugate(u1).u3,ratsimp;conjugate(u3).u2,ratsimp;
(%o28) 0
(%o29) 0
(%o30) 0
```

Para construir la matriz C preparamos primero el siguiente arreglo:

```
(%i31) Lvec:[u1,u2,u3]$
```

Con cada autovector como columna construimos la matriz C :

```
(%i32) C:transpose(apply('matrix,Lvec));
(%o32) 
$$\begin{pmatrix} \frac{3}{\sqrt{-(4\sqrt{5}-14)i-2\sqrt{5}+8}} & \frac{3}{\sqrt{(4\sqrt{5}+14)i+2\sqrt{5}+8}} & \frac{1}{\sqrt{2i+2}} \\ \frac{\sqrt{5}i+1}{\sqrt{-(4\sqrt{5}-14)i-2\sqrt{5}+8}} & -\frac{\sqrt{5}i-1}{\sqrt{(4\sqrt{5}+14)i+2\sqrt{5}+8}} & -\frac{i+1}{\sqrt{2i+2}} \\ \frac{(\sqrt{5}-1)i+\sqrt{5}-2}{\sqrt{-(4\sqrt{5}-14)i-2\sqrt{5}+8}} & -\frac{(\sqrt{5}+1)i+\sqrt{5}+2}{\sqrt{(4\sqrt{5}+14)i+2\sqrt{5}+8}} & \frac{1}{\sqrt{2i+2}} \end{pmatrix}$$

```

La inversa de la matriz C se obtiene como ya sabemos:

```
(%i33) Cinv:invert(C),ratsimp$
```

Aquí aplicamos una rutina de simplificación a cada uno de los elementos de la matriz. Esto lo hicimos con el comando `ratsimp`.

Para finalmente poder calcular $C^{-1}AC$ y obtener:

```
(%i34) Cinv.A.C,ratsimp;
(%o34) 
$$\begin{pmatrix} \frac{\sqrt{5}-5}{\sqrt{5}} & 0 & 0 \\ 0 & \frac{\sqrt{5}+5}{\sqrt{5}} & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

```

4.7.6 Ejercicios

1. Encuentre los autovalores y autovectores de las matrices:

(a).

$$\mathbb{A} \Leftrightarrow A_j^i = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad \mathbb{B} \Leftrightarrow B_j^i = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix};$$

(b).

$$\mathbb{A} \Leftrightarrow A_j^i = \begin{pmatrix} 2 & -i \\ i & 1 \end{pmatrix}, \quad \mathbb{B} \Leftrightarrow B_j^i = \begin{pmatrix} 1 & 1+i & 2i \\ 1-i & 5 & -3 \\ -2i & -3 & 0 \end{pmatrix};$$

(c).

$$\mathbb{A} \Leftrightarrow A_j^i = \begin{pmatrix} 1 & -3 \\ -3 & 1 \end{pmatrix}, \quad \mathbb{B} \Leftrightarrow B_j^i = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

Entonces

- (a). En cada caso, encuentre los autovalores y autovectores
 - (b). Encuentre también, para cada caso, las matrices de transformación que las lleven a su representación diagonal, cada una de
 - (c). Para la primera pareja de operadores lineales, muestre si es posible, encontrar una base de autovectores comunes que los diagonalice.
2. Demuestre que la matriz

$$\mathbb{C} \Rightarrow C_j^i = \begin{pmatrix} x & x - iy \\ x + iy & -z \end{pmatrix}$$

es igual a $\mathbb{C} = x\sigma_1 + y\sigma_2 + z\sigma_3$. Donde las matrices σ_i son las matrices de Pauli. Encuentre los autovalores y autovectores de esta matriz.

3. Considere dos operadores genéricos que conmutan: \mathbb{A} y \mathbb{B} tal que $[\mathbb{A}, \mathbb{B}] = 0$, definidos $\mathbb{C}^3 \rightarrow \mathbb{C}^3$ cuya representaciones matriciales en la base canónica $\{|e_1\rangle, |e_2\rangle, |e_3\rangle\}$ son

$$\mathbb{A} \Leftrightarrow A_j^i = \begin{pmatrix} 2 & 0 & i \\ 0 & 1 & 0 \\ -i & 0 & 2 \end{pmatrix} \quad \mathbb{B} \Leftrightarrow B_j^i = \frac{1}{2} \begin{pmatrix} 3 & -i\sqrt{2} & i \\ i\sqrt{2} & 2 & \sqrt{2} \\ -i & \sqrt{2} & 3 \end{pmatrix},$$

y determine si es posible encontrar una base común de autovectores.

Para ello, repase la sección 4.7.3 e insípírese en el ejercicio de la página 314 y siga los siguientes pasos

- (a). Resuelva el problema de autovalores $\mathbb{A}|u_i\rangle = \lambda_i|u_i\rangle$
- (b). Si alguno de los autovalores λ_i es degenerado utilice el hecho que existe un segundo operador \mathbb{B} que conmuta con \mathbb{A} y encuentre la representación matricial, diagonal a bloques, del operador \mathbb{B} en la base de autovectores $\{|u_i\rangle\}$.
- (c). Resuelva el problema de autovalores para el autoespacio degenerado. Esto es diagonalice el bloque 2×2 , para encontrar una base común de autovectores que diagonalice a ambos, \mathbb{A} y \mathbb{B} .
- (d). Identifique la base de los tres autovectores comunes y las tres parejas de autovalores de $\mathbb{A} \leftrightarrow \lambda_i$ y $\mathbb{B} \leftrightarrow \mu_i$ que etiquetan cada uno de esos autoestados.

4. Las transformaciones de Lorentz se pueden escribir de manera matricial como

$$\begin{pmatrix} \gamma & 0 & 0 & -i\gamma v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ i\gamma v/c & 0 & 0 & \gamma \end{pmatrix}$$

con $\gamma = (\sqrt{1 - v^2/c^2})^{-1}$. ¿será ortogonal? ¿será unitaria? Encuentre sus autovalores y autovectores.

5. Dado un observable \mathbb{A} y un vector de estado $|\psi\rangle$ general, definiremos el valor esperado de \mathbb{A} a la cantidad $\langle \mathbb{A} \rangle = \langle \psi | \mathbb{A} | \psi \rangle$, y la relación de dispersión de \mathbb{A} como:

$$\langle (\Delta \mathbb{A})^2 \rangle = \langle (\mathbb{A} - \langle \mathbb{A} \rangle \mathbb{I})^2 \rangle = \langle \mathbb{A}^2 \rangle - \langle \mathbb{A} \rangle^2 \equiv \langle \psi | \mathbb{A}^2 | \psi \rangle - \langle \psi | \mathbb{A} | \psi \rangle^2 ,$$

donde \mathbb{I} es el operador identidad. Nótese que el valor esperado es un número que representa la dispersión de un observable y tiene la misma estructura e interpretación de la varianza en estadística.

- (a). Muestre que la dispersión siempre es positiva, i.e $\langle (\Delta \mathbb{A})^2 \rangle \geq 0$. Para ello:

I. Inicie mostrando que para cualquier operador hermítico \mathbb{C} se cumple $\langle \mathbb{C}^2 \rangle \geq 0$.

II. Termine mostrando que $\mathbb{A} - \langle \mathbb{A} \rangle \mathbb{I}$ es un operador hermítico.

- (b). Muestre que la dispersión se anula para el caso en que $|\psi\rangle$ es autovector de \mathbb{A} con autovalor $\langle \mathbb{A} \rangle$.

- (c). Utilizando la desigualdad de Cauchy-Schwarz muestre que las relaciones de dispersión entre dos observables \mathbb{A} y \mathbb{B} siempre cumplen con:

$$\langle (\Delta \mathbb{A})^2 \rangle \langle (\Delta \mathbb{B})^2 \rangle \geq \frac{1}{4} |\langle [\mathbb{A}, \mathbb{B}] \rangle|^2 \quad \text{con } [\mathbb{A}, \mathbb{B}] = \mathbb{A}\mathbb{B} - \mathbb{B}\mathbb{A} .$$

Esta es la forma general de la relación de incertidumbre.¹⁵

- (d). En Mecánica Cuántica se define el operador de spin como $\mathbb{S}_i = \frac{\hbar}{2} \sigma_i$, donde las σ_i son las matrices de Pauli y los valores de $i = 1, 2, 3$ representan las direcciones x, y, z , respectivamente.

I. Encuentre la expresión para el conmutador: $[\mathbb{S}_i, \mathbb{S}_j]$, con $i, j = 1, 2, 3$.

II. Considere un vector de estado general $|\psi\rangle = a|+\rangle + b|-\rangle$, donde a y b son números complejos que cumplen con: $a^2 + b^2 = 1$ y $\{|+\rangle, |-\rangle\}$ la base de autovectores de \mathbb{S}_z . Muestre que:

$$\langle (\Delta \mathbb{S}_z)^2 \rangle \langle (\Delta \mathbb{S}_x)^2 \rangle \geq \hbar^4 [\text{Im}(ab^*)]^2 ,$$

con $\text{Im}(\circ)$ la parte imaginaria del argumento.

6. Dada la matriz:

$$\begin{pmatrix} 0 & q & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 2q & 2^2 & q & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 4(2^2) & q & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 4(N-3)^2 & q & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & q & 4(N-2)^2 & q \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & q & 4(N-1)^2 \end{pmatrix}$$

Encuentre los autovalores y autovectores cuando $q = 1$ y $N = 10$.

7. Realice los ejercicios anteriores utilizando el programa **Maxima**.

¹⁵Para detalles de las implicaciones de este problema se puede consultar Dumitru, S. "On the uncertainty relations and quantum measurements: conventionalities, short comings, reconsideration. arXiv preprint quant-ph/0504058 (2005). Y también Dumitru, S. "A possible general approach regarding the conformability of angular observables with mathematical rules of Quantum Mechanics. arXiv preprint quant-ph/0602147 (2006).

Bibliografía

- [1] Arfken, G. B., Weber, H., Weber, H.J. (2000) **Mathematical Methods for Physicists** 5ta Edición (Academic Press, Nueva York)
- [2] Borisenko, A.I, y Tarapov I.E. (1968) **Vector and Tensor Analisys** (Dover Publications Inc, Nueva York)
- [3] Dennery, P. y Krzywicki, A. (1995) **Mathematics for Physicists** (Dover Publications Inc, Nueva York)
- [4] Harper, C. (1971) **Introduction to Mathematical Physics** (*Prentice Hall, Englewood Cliff, N.J:*)
- [5] Hassani, S. (1991) **Foundations of Mathematical Physics** (*Prentice Hall, International Edition, London:*)
- [6] Hauser, W (1971) **Introduction to Principles of Electromagnetism** (*Addison-Wesley Pub Co Reading*)
- [7] Riley, K.F., Hobson, M.P. y Bence, S.J. (2002) **Mathematical Methods for Physics and Engineering** (*Cambridge University Press*)
- [8] Santaló, L.A (1969) **Vectores y Tensores** (*Editorial Universitaria, Buenos Aires*)
- [9] Schutz, B. (1980) **Geometrical Methods in Mathematical Physics** (*Cambridge University Press, Londres*)
- [10] Spiegel, M. (1959) **Vector Analysis** (*Schaums Outline Series, McGraw Hill New York*)