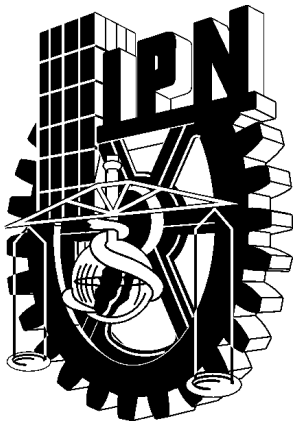
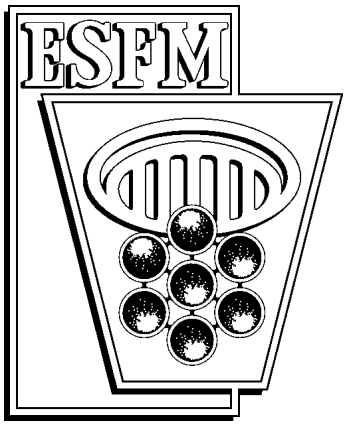




INSTITUTO POLITÉCNICO NACIONAL

Escuela Superior de Física y Matemáticas

Grupos de Mathieu y una Decodificación del Grupo M_{12}

T E S I S

QUE PARA OBTENER EL TÍTULO DE
LICENCIADO EN FÍSICA Y MATEMÁTICAS

PRESENTA
HUGO AARÓN GONZÁLEZ GALINDO

DIRECTOR DE TESIS
DR. PABLO LAM ESTRADA

México, D. F. Agosto de 2006

Introducción

Esta tesis fue desarrollada con el objeto de estudiar cierto tipo de grupos, para lo cual hemos escogido a los grupos de Mathieu. El hecho de que hayamos escogido a tales grupos fue debido a que poseen ciertas propiedades de gran interés, y que en muy pocos casos son estudiados dentro de los cursos de álgebra moderna que se imparten a nivel Licenciatura. No estamos equivocados al pensar esto, además las propiedades no sólo son de interés dentro de la teoría de grupos sino también dentro de otras ramas de la ciencia, como se verá en el último capítulo de la presente tesis.

La primer pregunta que surge es, ¿Cuáles son las propiedades que poseen los grupos de Mathieu y que los hace de interés en su estudio? La respuesta precisa a tal pregunta se contestada, por supuesto, durante el desarrollo de este trabajo, pero demos una pequeña introducción al respecto.

Los grupos de Mathieu se caracterizan por ser los primeros 5 grupos esporádicos descubiertos del total de 26. En la tabla que se muestra abajo se enumeran todos ellos, además de su fecha de descubrimiento y de quien los descubrió.

No.	**Nombre**	**Orden**	**Descubierto por**	**Año**
1	M_{11}	$2^3 \cdot 3^2 \cdot 5 \cdot 11$	Mathieu	1895
2	M_{12}	$2^6 \cdot 3^3 \cdot 5 \cdot 11$	Mathieu	1899
3	M_{22}	$2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11$	Mathieu	1900
4	M_{23}	$2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 23$	Mathieu	1900
5	M_{24}	$2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23$	Mathieu	1900
6	J_1	$2^3 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 19$	Janko	1966
7	$J_2(HaJ)$	$2^7 \cdot 3^3 \cdot 5^2 \cdot 7 \cdot 11 \cdot 19$	Janko, Higman, MacKey	1967

Grupos Esporádicos

No.	Nombre	Orden	Descubierto por	Año
8	$J_3(HJM)$	$2^7 \cdot 3^5 \cdot 5 \cdot 17 \cdot 19$	Hall, Janko	1969
9	HS	$2^9 \cdot 3^2 \cdot 5^3 \cdot 7 \cdot 11$	Higman, Sims	1969
10	McL	$2^7 \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11$	McLaughlin	1969
11	Suz	$2^{13} \cdot 3^7 \cdot 5^2 \cdot 7 \cdot 11 \cdot 13$	Suzuki	1969
12	He	$2^{10} \cdot 3^3 \cdot 5^2 \cdot 7^3 \cdot 17$	Held, Higman, McKay	1969
13	Co_1	$2^{21} \cdot 3^9 \cdot 5^4 \cdot 7^2 \cdot 11 \cdot 13 \cdot 23$	Conway, Leech	1969
14	Co_2	$2^{18} \cdot 3^6 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23$	Conway	1969
15	Co_3	$2^{10} \cdot 3^7 \cdot 5^3 \cdot 7 \cdot 11 \cdot 23$	Conway	1969
16	Fi_{22}	$2^{17} \cdot 3^9 \cdot 5^2 \cdot 7 \cdot 11 \cdot 23$	Fischer	1969
17	Fi_{23}	$2^{18} \cdot 3^{13} \cdot 5^2 \cdot 7 \cdot 11 \cdot 13 \cdot 17 \cdot 23$	Fischer	1969
18	Fi'_{24}	$2^{21} \cdot 3^{16} \cdot 5^2 \cdot 7^3 \cdot 11 \cdot 13 \cdot 17 \cdot$ $\cdot 23 \cdot 29$	Fischer	1969
19	Ly	$2^8 \cdot 3^7 \cdot 5^6 \cdot 7 \cdot 11 \cdot 31 \cdot 37 \cdot 67$	Lyons, Sims	1971
20	Ru	$2^{14} \cdot 3^3 \cdot 5^3 \cdot 7 \cdot 13 \cdot 29$	Rudvalis, Conway, Wales	1972
21	$O'N$	$2^9 \cdot 3^4 \cdot 5 \cdot 7^3 \cdot 11 \cdot 19 \cdot 31$	O'Nan, Sims	1973
22	M	$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot$ $\cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71$	Fischer	1974
23	B	$2^{41} \cdot 3^{13} \cdot 5^6 \cdot 7^2 \cdot 11 \cdot 13 \cdot 17 \cdot$ $\cdot 19 \cdot 23 \cdot 31 \cdot 47$	Fischer	1974
24	F_3	$2^{15} \cdot 3^{10} \cdot 5^3 \cdot 7^2 \cdot 13 \cdot 19 \cdot 31 \cdot 47$	Thompson, Smith	1974
25	F_5	$2^{14} \cdot 3^6 \cdot 5^6 \cdot 7 \cdot 11 \cdot 19$	Fischer, Smith, Harada	1974
26	J_4	$2^{21} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11^3 \cdot 23 \cdot 29 \cdot$ $\cdot 31 \cdot 37 \cdot 43$	Janko	1975

Grupos Esporádicos

La característica de estos grupos esporádicos, y por tanto de los grupos de Mathieu, es el hecho de que son grupos simples finitos no abelianos, distintos a los grupos ya bien conocidos como A_n, S_n y los *grupos de Lie*. Además, los grupos de Mathieu M_{11} y M_{12} poseen otra característica especial la cual los coloca dentro de otra clasificación más reducida de grupos. Esta característica es el que son de los pocos grupos distintos de S_n y A_n que son fielmente transitivos; en particular, son fielmente 4-transitivo y 5-trasitivo, respectivamente, y son los de mayor grado de transitividad distintos de S_n y A_n. Por otro lado, veremos cómo estos pueden ser utilizados en la teoría de códigos, sobre todo M_{12}.

Como su nombre lo indica, los grupos de Mathieu fueron descubiertos por Emile Léonard Mathieu (1835-1890) en los años de 1861 y 1873. Mathieu fue influenciado por Cauchy para trabajar sobre permutaciones. Había investigado las funciones múltiplemente transitivas, así como grupo de permutaciones y grupos de permutaciones múltiplemente transitivas. A lo largo de su trabajo, Mathieu intento extender, por construcciones transitivas, un grupo de permutación $(n-1)$-transitivo a un grupo n-transitivo. De estos trabajos, descubrió el algoritmo para la construcción de estos grupos, cuando esto era posible. La transitividad más alta en los grupos transitivos simples es el grupo 5-transitivo, y Mathieu descubrió el grupo de permutaciones 5-transitivo sobre 12 símbolos y sobre 24 símbolos, los cuales son M_{12} y M_{24}. Los otros grupos de Mathieu surgieron como subgrupos de estos grupos y de un subgrupo M_{23}. Por ejemplo, M_{11} es el subgrupo de M_{12} formado como el estabilizador de un punto de M_{12}. Cada uno de estos grupos son múltiplemente transitivos. La simplicidad y unicidad de los grupos de Mathieu no fue expresada sino hasta la decada de los 1930's por Witt.

La construcción de los grupos de Mathieu es lo que básicamente trataremos, y sobre todo la transitividad y la simplicidad de estos.

La tesis está estructurada de la siguiente manera:

- Capitulo 1. Aquí se dan de una forma rápida y sin demostración los resultados más importantes sobre la teoría de grupos y de campos que utilizamos en el desarrollo de la tesis. El lector que haya llevado un curso o leído un libro sobre teoría de grupos y de campos podrá ver que cada uno de los resultados presentados son los de estándar.

- Capítulo 2. Presentamos la construcción del grupo lineal general y de algunas de sus propiedades. Este grupo será importante en la construcción de los grupos de Mathieu. En particular, se obtendrá el grupo unimodular proyectivo $PSL(m, K)$.

- Capítulo 3. En este capítulo, analizaremos la transitividad de los grupos de permutaciones en detalle, para obtener resultados derivados de la geometría afín y proyectiva. Aquí es donde harán su aparición, por primera vez, los grupos de Mathieu.

- Capítulo 4. Aquí, probaremos la simplicidad de los grupos de Mathieu. Haremos, en particular, una construcción geométrica del grupo de Mathieu M_{12}, y lo usaremos para tener una aplicación a la teoría de códigos. Por supuesto, no desarrollaremos la teoría de códigos pero, en este caso, daremos algunas ideas de cómo se aplica el grupo M_{12} a dicha teoría. De esta manera, finalizaremos dando un algoritmo en el que establece una alternativa para decodificar algunas palabras de permutaciones de M_{12} que sean erróneas en la transmisión de la información de datos y la manera de cómo corregirlos.

Índice general

Capítulo 1

Algunos Resultados de la Teoría de Grupos y Campos

En este capítulo mencionaremos algunos resultados de la teoría de grupos y de campos que serán usados en este trabajo, estos se presentarán sin demostración, salvo algunas excepciones, y pueden ser consultados, por ejemplo, en las referencias bibliográficas [3], [4] y [7].

Por supuesto, supondremos conocidas las notaciones estándar de los la teoría de grupos y de campos, dando énfasis a aquellos temas que más usaremos. Así que, iniciaremos con el tema de acciones de grupo.

1.1. Acciones de Grupo

Sea X un conjunto no vacío. Denotamos por S_X al conjunto de las funciones biyectivas de X en sí mismo. S_X es un grupo con la operación de composición de funciones, el cual es llamado el **grupo de permutaciones de** X. A los elementos de S_X se les llaman **permutaciones** de X. Si el conjunto X es finito de n elementos, entonces S_X es isomorfo al grupo simétrico S_n de grado n, es decir, $S_n = S_Y$ donde $Y = \{1, \ldots, n\}$. Más adelante analizaremos con más detenimiento al grupo S_n.

Definición 1.1.1. *Decimos que un grupo G* ***actúa sobre un conjunto*** *X si existe un homomorfismo $\varphi : G \longrightarrow S_X$.*

Si $\varphi : G \longrightarrow S_X$ es un homomorfismo, entonces para $a \in G$, escribimos φ_a para representar al elemento $\varphi(a)$ de S_X, y si $x \in X$, escribimos ax en lugar de $\varphi_a(x)$.

Notemos que si H es un subgrupo de un grupo G y X es el conjunto de todos los **conjugados** de H, es decir, $X = \{gHg^{-1} \mid g \in G\}$, entonces G actúa sobre X. Pues la función

$$\begin{aligned} \varphi : G &\longrightarrow S_X \\ g &\longmapsto \varphi_g \end{aligned}$$

es un homomorfismo, donde

$$\begin{aligned} \varphi_g : X &\longrightarrow X \\ aHa^{-1} &\longmapsto gaHa^{-1}g^{-1} \end{aligned}$$

es una biyección.

Si G es un grupo, decimos $a, b \in G$ son **conjugados** si existe $g \in G$ tal que $a = gbg^{-1}$. Tenemos que la relación de ser conjugado es una relación de equivalencia sobre G cuyas clases son llamadas **clases de conjugación** de G. Bajo las definiciones anteriores, tenemos que todo grupo G actúa sobre sí mismo por conjugación. Esto es, para cada $g \in G$, sea $\varphi_g : G \longrightarrow G$ dada por la correspondencia $k \longmapsto gkg^{-1}$. Claramente φ_g es una función biyectiva para cada $g \in G$. Además, $\varphi_{g_1} \circ \varphi_{g_2} = \varphi_{g_1 g_2}$ para cada $g_1, g_2 \in G$. Entonces, la función

$$\begin{aligned} \varphi : G &\longrightarrow S_G \\ g &\longmapsto \varphi_g \end{aligned}$$

es un homomorfismo de grupos, la cual es llamada la **acción de conjugación** de G.

Definición 1.1.2. *Si G actúa sobre X y $x \in X$, definimos la **órbita** de x como el conjunto $O_x = \{gx \mid g \in G\}$.*

Teorema 1.1.1. *Si G es un grupo que actúa sobre un conjunto X, entonces:*

(*i*) *La relación sobre X definida por $x \sim x' \iff gx = x'$ para algún $g \in G$ es una relación de equivalencia.*

(*ii*) *Para cada $x \in X$, $G_x = \{g \in G \mid gx = x\}$ es un subgrupo de G.* ■

Bajo la relación de equivalencia sobre X establecida en el teorema anterior, tenemos que la clase de equivalencia de un elemento $x \in X$ es su órbita.

Definición 1.1.3. *Si G es un grupo que actúa sobre un conjunto X y $x \in X$, entonces al subgrupo $G_x = \{g \in G \mid gx = x\}$ de G lo llamaremos el* ***estabilizador*** *de x.*

Teorema 1.1.2. *Si G es un grupo que actúa sobre un conjunto X y $x \in X$, entonces $Card(O_x) = [G : G_x]$.* ■

Corolario 1.1.1. *Si un grupo finito G actúa sobre un conjunto finito X, entonces $Card(O_x) \mid |G|$, para cada $x \in X$.* ■

Corolario 1.1.2. *Si G es un grupo finito y $x \in G$, entonces el número de conjugados de x es $[G : C_G(x)]$, donde $C_G(x)$ es el centralizador de x en G.* ■

Corolario 1.1.3 (Ecuación de Clase). *Si G es un grupo finito y $\{x_1, \ldots, x_n\}$ es un conjunto completo de representantes en las clases de conjugación, entonces*

$$|G| = |Z(G)| + \sum_{x_i \notin Z(G)} [G : C_G(x_i)].$$

Corolario 1.1.4. *Si H es un subgrupo de un grupo finito G, entonces el número de conjugados de H en G es $[G : N_G(H)]$.* ■

Teorema 1.1.3 (Teorema de Caley). *Todo grupo G es isomorfo a un subgrupo de S_G. En particular, todo grupo finito de orden n es isomorfo a un subgrupo de S_n.*

Demostración: Sea a un elemento arbitrario de G. Definimos

$$\begin{aligned} L_a : G &\longrightarrow G \\ x &\longmapsto ax \end{aligned}$$

Tenemos que L_a es una biyección. Además, la función

$$\begin{aligned} \lambda : G &\longrightarrow S_G \\ a &\longmapsto L_a \end{aligned}$$

es un isomorfismo de G en $\lambda(G)$. □

1.2. Los p-Grupos y Teoremas de Sylow

Definición 1.2.1. *Sea p un número primo. Decimos que un grupo G es un p-**grupo** si todo elemento x de G tiene orden igual a una potencia de p.*

Teorema 1.2.1 (Teorema de Cauchy). *Si G es un grupo finito cuyo orden es divisible por un primo p, entonces G contiene un elemento de orden p.* ■

Proposición 1.2.1. *Un grupo finito G es un p-grupo si y sólo si $|G|$ es una potencia de p.* ■

Teorema 1.2.2. *Si G es un p-grupo finito con más de un elemento, entonces $Z(G) \neq \{e\}$.* ■

Corolario 1.2.1. *Si p es un número primo, entonces todo grupo G de orden p^2 es abeliano.* ■

Definición 1.2.2. *Sea p un número primo. Decimos que un subgrupo P de un grupo G es un p-**subgrupo Sylow** de G si P es un p-subgrupo máximal de G (esto es, $P \leq H \leq G$ con H p-subgrupo de G implica $H = P$ ó $H = G$).*

Proposición 1.2.2. *Sea p un número primo fijo. Si G es un grupo que tiene solamente un p-subgrupo Sylow P, entonces $P \lhd G$.* ■

Proposición 1.2.3. *Sea P un p-subgrupo Sylow de G. Entonces,*

1. *$N_G(P)/P$ no tiene elementos distintos de la identidad cuyo orden es una potencia de p.*
2. *Si $a \in G$ tiene orden una potencia de p, entonces $aPa^{-1} = P$ implica $a \in P$.*

■

Teorema 1.2.3 (Primer Teorema de Sylow). *Sea G un grupo de orden $p^k m$, donde p es número primo, $k \geq 1$ y $(p, m) = 1$. Entonces, G contiene un subgrupo de orden p^i para cada $1 \leq i \leq k$, y cada subgrupo de orden p^i, con $i < k$, es normal en algún subgrupo de orden p^{i+1}.* ■

Teorema 1.2.4 (Segundo Teorema de Sylow). *Si H es un p-subgrupo de un grupo finito G, y P es cualquier p-subgrupo Sylow de G, entonces existe $x \in G$ tal que $H < xPx^{-1}$. En particular, cualesquiera dos p-subgrupos Sylow de G son conjugados.* ■

Teorema 1.2.5 (Tercer Teorema de Sylow). *Sea G un grupo finito y p un número primo. Si n_p es el número de p-subgrupos de Sylow de G, entonces $n_p \equiv 1 \bmod p$ y $n_p \mid |G|$.* ■

Teorema 1.2.6 (Burnside, 1900). *Sea G un grupo finito y sea Q un subgrupo de Sylow contenido en el centro de su normalizador; entonces Q tiene un complemento normal K (y K es un subgrupo caracteristico de G).*

1.3. Grupo Simétrico

La herramienta central en este trabajo es el grupo simétrico de grado finito, por lo que en esta sección recordaremos algunas propiedades fundamentales del mismo.

Sea X un conjunto no vacío, y sea

$$S_X = \{f : X \longrightarrow X \mid f \text{ es una función biyectiva}\}.$$

Bajo la operación de composición de funciones, el conjunto S_X forma un grupo, pues la composición de funciones es asociativa, y la composición de funciones biyectivas es biyectiva. La función identidad de S_X, id_X, es una biyección que juega el papel de identidad para el grupo, y toda función biyectiva tiene una función inversa que también es biyectiva.

Definición 1.3.1. *Los elementos de S_X se llaman **permutaciones** y S_X se llama **grupo de permutaciones** sobre el conjunto X. Si $n \in \mathbb{N}$ y $X = \{1, 2, \ldots, n\}$, entonces S_X es llamado **grupo simétrico de grado** n y se representa con el símbolo S_n.*

Se tiene que $|S_n| = n!$. Puesto que un elemento $\sigma \in S_n$ es una función sobre el conjunto $X = \{1, \ldots, n\}$, podemos describir a σ listando todos los elementos de X sobre una línea horizontal y la imagen bajo σ directamente abajo de ésta; por ejemplo,

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n) \end{pmatrix}.$$

Observación 1.3.1. Sean X y Y dos conjuntos tales que existe una biyección $f : X \to Y$. Entonces, la correspondencia $\alpha \mapsto f \circ \alpha \circ f^{-1}$ determina un isomorfismo entre S_X y S_Y.

Definición 1.3.2. *Sean $i \in \{1, 2, \ldots, n\}$ y $\sigma \in S_n$. Decimos que σ **fija** a i si $\sigma(i) = i$, y que σ **mueve** a i si $\sigma(i) \neq i$.*

Definición 1.3.3. *Sean $i_1, i_2, \ldots, i_r$ elementos distintos del conjunto $\{1, 2, \ldots, n\}$. Si $\sigma \in S_n$ deja fijos a todos los elementos del conjunto $\{1, 2, \ldots, n\} \setminus \{i_1, i_2, \ldots, i_r\}$ y $\sigma(i_1) = i_2, \sigma(i_2) = i_3, \sigma(i_3) = i_4, \ldots, \sigma(i_r) = i_1$, entonces decimos que σ es un r**-ciclo** o que σ es un **ciclo de longitud** r y en lugar de representar a σ como*

$$\sigma = \begin{pmatrix} i_1 & i_2 & i_3 & \ldots & i_r & i_{r+1} & i_{r+2} & \ldots & i_n \\ i_2 & i_3 & i_4 & \ldots & i_1 & i_{r+1} & i_{r+2} & \ldots & i_n \end{pmatrix}$$

lo escribiremos como $\sigma = (i_1\ i_2\ i_3\ \ldots\ i_r)$.

Notemos que el orden de un r-ciclo es r.

Definición 1.3.4. *Dos permutaciones β y α en S_n son **disjuntas** si todo elemento movido por una de ellas es fijado por la otra. Más explícitamente, si $\alpha(i) \neq i$, entonces $\beta(i) = i$ y si $\beta(j) \neq j$, entonces $\alpha(j) = j$ (es posible que $\beta(k) = k = \alpha(k)$, para algún $k \in \{1, 2, \ldots, n\}$).*

Teorema 1.3.1. *Sean α y β dos r-ciclos en S_n. Si existe un $x_0 \in \{1, 2, \ldots, n\}$ tal que*

(i) α y β mueven a x_0, y

(ii) $\alpha^t(x_0) = \beta^t(x_0)$, para todo $t \in \mathbb{Z}$,

entonces $\alpha = \beta$. ■

Teorema 1.3.2. *Si* α *y* β *son permutaciones disjuntas, entonces* $\alpha\beta = \beta\alpha$. ■

Si $m, n \in \mathbb{N}$ con $m > n$ y $\alpha \in S_n$, entonces α puede ser considerada como una permutación de S_m que fija a todo elemento del conjunto $\{n+1, n+2, \ldots, m\}$.

Teorema 1.3.3. *Toda permutación* $\alpha \in S_n$ *es un ciclo ó un producto de ciclos disjuntos.* ■

Más aún, tenemos el siguiente:

Teorema 1.3.4. *Sea* $\alpha \in S_n$ *y supongamos que* $\alpha = \beta_1\beta_2 \ldots \beta_t$ *es una factorización en ciclos disjuntos de longitud* ≥ 2. *Entonces, esta factorización es única salvo por el orden en que aparecen los ciclos.* ■

Por medio del teorema anterior, podemos establecer que toda permutación de S_n admite una estructura cíclica, es decir, una permutación esta expresado de acuerdo a su estructura cíclica si α se expresa como producto de ciclos disjuntos, ordenado de acuerdo a la longitud de cada ciclo de menor a mayor ó viceversa.

Observación 1.3.2.

1. Si α es un n-ciclo, entonces α^k es producto de (n, k) ciclos disjuntos, cada uno de longitud $n/(n, k)$.
2. Si p es primo, entonces los únicos elementos de S_n de orden p son los p-ciclos o los productos de p-ciclos disjuntos.

Definición 1.3.5. *Todo* 2*-ciclo recibe el nombre* ***transposición****.*

Observación 1.3.3. El subconjunto $\mathbb{V} = \{e, (1,4)(2,3), (1,2)(3,4), (1,3)(2,4)\}$ de S_4 es un subgrupo normal de éste (el cual llamaremos 4**-grupo**).

Teorema 1.3.5. *Todo $\alpha \in S_n$ es un producto de transposiciones.* ■

Definición 1.3.6. *Decimos que una permutación $\alpha \in S_n$ es **par** (**impar**) si α tiene una factorización como un producto de un número par (impar) de transposiciones.*

Observación 1.3.4. Observemos que si $\alpha = \beta_1\beta_2 \dots \beta_t$ es una factorización en S_n de α como producto de ciclos disjuntos de longitud ≥ 2 y $\Sigma = \sum_{i=1}^{t} longitud(\beta_i)$, entonces el número $\Sigma - t$ depende solamente de α. Aún cuando en la factorización pusieramos los 1-ciclos, el número $\Sigma - t$ no cambiaría, ya que si el número de 1-ciclos es m, entonces $\Sigma' = \Sigma + m$ y $t' = t + m$ implica que $\Sigma' - t' = \Sigma + m - (t + m) = \Sigma - t$.

Definición 1.3.7. *Definimos la función **signo**, denotada por sgn, como sigue:*

$$\begin{aligned} sgn : S_n &\longrightarrow \{\pm 1\} \\ \alpha &\longmapsto (-1)^{\Sigma - t} \end{aligned}$$

Claramente $sgn(id_{S_n}) = 1$ y $sgn(\tau) = -1$ para cada transposición τ.

Teorema 1.3.6. *Si $\alpha \in S_n$ y τ es una transposición, entonces*

$$sgn(\tau\alpha) = -sgn(\alpha).$$

■

Teorema 1.3.7. *La función $sgn : S_n \longrightarrow \{-1, 1\}$ es un homomorfismo de grupos.* ■

Corolario 1.3.1. *Una permutación α de S_n es par si $sgn(\alpha) = 1$ e impar si $sgn(\alpha) = -1$. El número de factores que aparecen en cualquier factorización de α como producto de transposiciones es siempre un número par ó impar.* ■

Definición 1.3.8. *El **grupo alternante** de grado n, denotado por A_n, es el subgrupo de S_n generado por todas las permutaciones pares.*

Veamos que en efecto A_n es un grupo; más aún tenemos el siguiente:

Teorema 1.3.8. *$A_n = \{\sigma \in S_n : sgn(\sigma) = 1\}$ es un subgrupo de S_n.*

Teorema 1.3.9. *Para $n \geq 2$, A_n es un subgrupo normal de S_n cuyo orden es $n!/2$.* ■

Teorema 1.3.10.

(*i*) *Si $\alpha \in S_n$ y $\beta = (i_1\ i_2\ \ldots\ i_r)$ es un r-ciclo, entonces $\alpha\beta\alpha^{-1}$ es el r-ciclo $(\alpha(i_1)\ \ldots\ \alpha(i_r))$*

(*ii*) *Cualesquiera dos r-ciclos en S_n son conjugados.* ■

Corolario 1.3.2. *Dos permutaciones $\alpha, \beta \in S_n$ son conjugadas si y sólo si tienen la misma estructura cíclica.* ■

Sabemos que la relación $\sim$ de S_n de ser conjugados es una relación de equivalencia, en donde $\alpha \sim \alpha'$ si y sólo si poseen la misma estructura cíclica.
Sean $C_1, \ldots, C_m$ son la totalidad de clases de equivalencia, o de clases de conjugación. Entonces, tenemos el siguiente:

Teorema 1.3.11. *El número de clases de conjugación distintas C_i está en correspondencia biunívoca con las sucesiones $\{(\beta_1, \ldots, \beta_n) \in \mathbb{Z}^n \mid \beta_i \geq 0\ y\ \sum_{i=1}^{n} i\beta_i = n\}$.* ■

Teorema 1.3.12. *El número de clases de conjugación C_i es igual al número de particiones de n, es decir, el número de vectores (con un número finito de coordenadas) distintos, digamos $(n_1, n_2, \ldots, n_l)$ donde cada n_i es un entero positivo tal que $n = n_1 + n_2 + \cdots + n_l$ y $n_1 \geq n_2 \geq \cdots \geq n_l$.* ■

Teorema 1.3.13 (Fórmula de Cauchy). *Sea C_i una de las clases de conjugación. Si en la descomposición cíclica de los elementos de C_i aparecen:*

$$\alpha_1 : 1 - ciclos$$
$$\alpha_2 : 2 - ciclos$$
$$\vdots$$
$$\alpha_n : n - ciclos$$

Entonces:

$$Card(C_i) = \frac{n!}{(1^{\alpha_1})(\alpha_1!)(2^{\alpha_2})(\alpha_2!)\cdots(n^{\alpha_n})(\alpha_n!)}.$$

■

Un resultado que es siempre útil recordar es el siguiente:

Teorema 1.3.14. *A_n es un grupo simple si $n \geq 5$.* ■

1.4. Producto Directo

Dados dos grupos A y B, podemos formar, partiendo de ellos, el conjunto de todos los pares ordenados (a, b), $a \in A,\ b \in B$. Estos pares ordenados serán los elementos de un nuevo grupo, el **producto directo** $A \times B$, si definimos nuestro producto por la regla

$$(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2). \tag{1.4.1}$$

La verificación de que la regla del producto (1.4.1) satisface los axiomas de grupo con (e_A, e_B) como elemento identidad es directa, y depende solamente de la validéz de estos axiomas para A y B. Por otra parte, la correspondencia $(a, b) \rightleftarrows (b, a)$ muestra que $A \times B$ y $B \times A$ son isomorfos, de modo que podemos hablar del producto directo de dos grupos sin especificar el orden. La correspondencia $a \rightleftarrows (a, e_B)$ es un isomorfismo entre A y el subgrupo de elementos de $A \times B$ con la identidad como segunda coordenada. Análogamente, $b \rightleftarrows (e_A, b)$ es un isomorfismo entre B y el subgrupo de elementos (e_A, b). Identificaremos A y B con estos subgrupos. De acuerdo con esta identificación decimos que $G = A \times B$ es el producto directo de sus subgrupos A y B. Como $(a, e_B)(e_A, b) = (a, b) = (e_A, b)(a, e_B)$, se sigue que en

$A \times B$ todo elemento de A **conmuta** (o permuta) con todo elemento de B; es decir, $ab = ba$ para $a \in A,\ \ b \in B$.

En el producto directo $(a,b)^{-1} = (a^{-1}, b^{-1})$, de aquí que $(a_1, b_1)(a, e_B)(a_1, b_1)^{-1} = (a_1 a a_1^{-1}, e_B)$, luego A es un subgrupo normal de $A \times B$. Análogamente, B es un subgrupo normal de $A \times B$. El único elemento que es simultáneamente de la forma (a, e_B) y de la forma (e_A, b) es (e_A, e_B), de donde $A \cap B = \{e\}$. Por otra parte, $\langle A \cup B \rangle$ incluye todos los productos de la forma $(a, e_B)(e_A, b) = (a, b)$, de donde $\langle A \cup B \rangle = A \times B$. Estas relaciones entre A y B caracterizan $A \times B$.

Teorema 1.4.1. *Un grupo G es isomorfo al producto directo de dos subgrupos A y B si A y B son subgrupos normales tales que $A \cap B = \{e\}$, $G = \langle A \cup B \rangle$.* ■

Definición 1.4.1. *Al producto directo definido en el teorema anterior lo llamaremos el **producto directo interno** de A y B.*

Podemos generalizar las ideas precedentes para definir un producto de cualquier número de grupos, finito o infinito. Para lo cual daremos la definición de producto cartesiano.

Definición 1.4.2. *Sea $\{A_i \mid i \in I\}$ una familia indexada de conjuntos por un conjunto $I \neq \emptyset$. El **producto cartesiano** de los A_i, es el conjunto de todas las funciones $f : I \longrightarrow \bigcup_{i \in I} A_i$ tal que $f(i) \in A_i,\ \ \forall\, i \in I$. A este conjunto lo denotaremos por $\prod_{i \in I} A_i$.*

Sea $\{G_i \mid i \in I\}$ una familia indexada arbitraria de grupos por un conjunto $I \neq \phi$, entonces definimos en el conjunto $\prod_{i \in I} G_i$ el producto "$*$" de dos elementos de $\prod_{i \in I} G_i$ de la siguiente forma:

Si $f, g \in \prod_{i \in I} G_i$, $f * g = h$, donde $h(i) = f(i)g(i),\ \ \forall\, i \in I$. Es claro que $h \in \prod_{i \in I} G_i$ pues $f(i), g(i) \in G_i\ \ \forall\, i \in I$, por lo que $h(i) = f(i)g(i) \in G_i\ \ \forall\, i \in I$.

El subgrupo $\{f \in \prod_{i \in I} G_i \mid f(g_i) = e_i,\ \ \text{salvo para un número finito de índices}\}$ del producto cartesiano, se llama el **producto directo debíl** de los G_i. Claramente, producto directo debíl y producto directo coinciden cuando el número de factores es finito. En cualquier caso, observemos que $\bar{G}_j = \{f \in \prod_{i \in I} G_i \mid f(g_i) = e_i,\ \ \forall\ \ i \neq j\}$,

forma un subgrupo normal isomorfo a G_j, e identificando a G_j con este subgrupo, notemos que $G_j \cap \langle \bigcup_{i \neq j} G_j \rangle = \{e\}$.

Teorema 1.4.2. *Un grupo G es isomorfo al producto directo débil de los subgrupos A_i, $i \in I$, si*

(1) *Todo A_i es un subgrupo normal;*

(2) *$A_j \cap \langle \cup_{i \neq j} A_i \rangle = \{e\}$ para todo $j \in I$;*

(3) *$G = \langle \cup_{i \in I} A_i \rangle$.*

■

1.5. Grupos Abelianos Finitos

Ya hemos definido en la sección anterior lo que es el producto directo de una familia arbitraria de grupos. Lo que ahora se verá es que cualquier grupo abeliano finito es el producto directo de grupos cíclicos. Esto reduce la mayoría de las interrogantes relativas a grupos abelianos finitos a interrogantes referentes a grupos cíclicos.

Proposición 1.5.1. *Sea G un grupo abeliano finito de orden mn, donde m y n son primos relativos. Si $M = \{x \in G \mid x^m = e\}$ y $N = \{x \in G \mid x^n = e\}$, entonces $G = M \times N$. Además, si m y n son distintos de 1, entonces $M \neq \langle e \rangle$ y $N \neq \langle e \rangle$.* ■

Corolario 1.5.1. *Sea G un grupo abeliano finito y p un primo tal que $p \mid |G|$. Entonces, $G = P \times T$ para algunos subgrupos P y T, donde $|P| = p^m$, $m > 0$, y $|T|$ no es divisible por p.* ■

Ahora se llega al resultado clave que permite abordar la demostración del teorema fundamental de esta sección.

Teorema 1.5.1. *Sea G un grupo abeliano de orden p^n, p primo, y supóngase que $a \in G$ tiene el máximo orden de todos los elementos de G. Entonces $G = A \times Q$, donde A es el subgrupo cíclico de G generado por a.* ■

Teorema 1.5.2 (Teorema Fundamental de los Grupos Abelianos Finitos). *Todo grupo abeliano finito es el producto directo de grupos cíclicos.* ■

Volvemos a los grupos abelianos G de orden p^n. Se tiene ahora a la mano que $G = A_1 \times A_2 \times \cdots \times A_k$, donde los A_i son subgrupos cíclicos de orden p^{n_i}. Se puede arreglar la numeración de manera que $n_1 \geq n_2 \geq \cdots \geq n_k$. Además, $|G| = |A_1 \times A_2 \times \cdots \times A_k| = |A_1||A_2| \cdots |A_k|$, lo cual da que

$$p^n = p^{n_1} p^{n_2} \cdots p^{n_k} = p^{n_1+n_2+\cdots+n_k},$$

por consiguiente $n = n_1 + n_2 + \cdots + n_k$. De esta manera los enteros $n_i \geq 0$ proporcionan una partición de n. Se puede demostrar que estos enteros $n_1, n_2, \ldots, n_k$, los cuales se llaman **invariantes** de G, son únicos. En otras palabras, dos grupos abelianos de orden p^n son isomorfos si y sólo si tienen los mismos invariantes. Dado esto, se sigue que el número de subgrupos abelianos no isomorfos entre sí de orden p^n es igual al número de particiones de n.

1.6. Producto Orlado

Sean G y H grupos de permutaciones sobre conjuntos A y B, respectivamente. Definimos el **producto orlado** de G por H, escrito $G \wr H$, en la siguiente forma: $G \wr H$ es el grupo de todas las permutaciones θ sobre $A \times B$ de la siguiente clase:

$$\theta(a, b) = (\gamma_b a, \eta b), \quad a \in A,\ b \in B, \tag{1.6.1}$$

donde para cada $b \in B$, γ_b es una permutación de G sobre A, pero para diferentes b las elecciones de las permutaciones γ_b son independientes. La permutación η es una permutación de H en B. Las permutaciones θ, con $\eta = 1$, forman un subgrupo normal G^* isomorfo al producto directo de n copias de G, donde n es el número de elementos en el conjunto B. El grupo factor $G \wr H/G^*$ es isomorfo a H, y las permutaciones θ con las $\gamma_b = 1$ forman un subgrupo isomorfo a H, cuyos elementos pueden tomarse como representantes de las clases laterales de G^* en G.

El producto orlado es asociativo en el sentido de que si K es un tercer grupo de permutaciones sobre un conjunto C, entonces $(G \wr H) \wr K$ y $G \wr (H \wr K)$ son isomorfos, y si identificamos los conjuntos $(A \times B) \times C$ y $A \times (B \times C)$ con $A \times B \times C$, entonces son idénticos.

Los subgrupos de Sylow del grupo simétrico S_n se construyen fácilmente por medio del producto orlado. ¿Cuál es la máxima potencia de p que divide a $n!$? Los factores

de $n!$ divisibles por p son $p,\ 2p,\ 3p,\ldots,kp$, donde $k=[n/p]$ es el mayor entero que no excede a n/p, de donde $n!$ es divisible por p^k y las potencias de p que dividen a $k!$. Haciendo notar que $[k/p]=[n/p^2]$, y así sucesivamente, encontramos que la potencia de p que divide a $n!$ es p^M, de donde

$$M=\left[\frac{n}{p}\right]+\left[\frac{n}{p^2}\right]+\left[\frac{n}{p^3}\right]+\cdots$$

Si expresamos a n en expansión p-ádica,

$$n=a_0p^u+a_1p^{u-1}+\cdots+a_{u-1}p+a_u, \tag{1.6.2}$$

donde cada a_i está en el rango $0\leq a_i\leq p-1$, encontramos que

$$M=a_0(p^{u-1}+p^{u-2}+\cdots+p+1)+a_1(p^{u-2}+\cdots+p+1)+\cdots+a_{u-1}. \tag{1.6.3}$$

En particular, un subgrupo de Sylow del grupo simétrico sobre p^r elementos será de orden p^{N_r}, donde $N_r=p^{r-1}+p^{r-2}+\cdots+1$. Vemos pues que habiendo construido subgrupos de Sylow para grupos simétricos sobre $p,\ p^2,\ldots,p^u$ elementos, podemos fácilmente construir un subgrupo de Sylow para el grupo simétrico sobre n elementos, donde n está dado por (1.6.2). Dividimos los n elementos en a_0 bloques de p^u elementos, a_1 de p^{u-1} elementos,..., a_{u-1} de p elementos, y a_u de un solo elemento. Entonces, si en cada bloque construimos el apropiado subgrupo de Sylow y tomamos el producto directo de ellos, tendremos un grupo P de orden p^M, donde M está dado por (1.6.3). De donde P será un subgrupo de Sylow de S_n.

Un subgrupo de Sylow de S_p sobre $1,2,\ldots,\ p$ será de orden p, por tanto un subgrupo Sylow será un grupo cíclico de orden p generado por $a_1=(1\ \ 2\ \ldots\ \ p)$. S_{p^2} sobre $1,2,\ldots,\ p^2$ tendrá un subgrupo que es producto directo de los subgrupos cíclicos generados por $a_1=(1\ \ 2\ \ldots\ \ p),\ a_2=(p{+}1\ \ p{+}2\ \ldots\ 2p),\ldots,\ a_p=(p^2{-}p{+}1\ \ldots\ p^2)$. Si tomamos otro elemento de orden p, $b=(1\ \ p+1\ \ 2p+1\ \ldots\ p^2-p+1)(2\ \ p+2\ \ldots)\cdots(p\ \ 2p\ \ldots\ p^2)$, entonces $b^{-1}a_ib=a_{i+1}$, donde los subíndices se toman módulo p. Luego b y las $a_i's$ generan un subgrupo P_2 de orden p^{p+1}, que es el producto orlado del primer ciclo de b y el grupo cíclico $\langle a_1\rangle$. Aquí P_2 es un subgrupo Sylow de S_{p^2}. En general, sea P_r un subgrupo de Sylow de S_{p^r} sobre $1,\ldots,\ p^r$. Tómese los elementos $1,\ldots,\ p^r,\ p^r+1,\ldots,\ 2p^r,\ldots,\ p^{r+1}$ como elementos permutables de $S_{p^{r+1}}$. Escojamos entonces el elemento

$$c=(1\ \ p^r+1\ \ 2p^r+1\ \ldots\ (p-1)p^r+1)\cdots(j\ \ p^r+j\ \ 2p^r+j\ \ldots\ (p-1)p^r+j)\cdots,$$

donde j toma los valores de 1 a p^r. Tenemos entonces $P_r^{(i)}=c^{-1}P_rc^i$ como un grupo de orden p^{N_r} sobre los elementos $ip^r+1,\ldots,\ (i+1)p^r$. Como cada $P_r^{(i)}$,

$i = 0, 1, \ldots, p-1$, desplaza un distinto conjunto de elementos, el grupo que ellas generan es su producto directo. Aquí c y P_r generan un grupo que es de orden $p^{p^{N_r}+1}$. Pero $p^{N_r}+1 = p[p^{r-1}+\cdots+(p+1)]+1 = N_{r+1}$, y por tanto c y P_r generan P_{r+1}, un subgrupo de Sylow del grupo simétrico sobre p^{r+1} elementos. Con P_r actuando sobre los elementos $1, \ldots, p^r$ y tomando c como el ciclo $c = (u_0 \;\; u_1 \;\; \ldots \;\; u_{p-1})$, entonces el producto orlado $P_r \wr \langle c \rangle$ permuta símbolos (i, u_j), $i = 1, \ldots, p^r,\ j = 0, \ldots, p-1$. Si identificamos (i, u_j) con $i + jp^r$ vemos que P_{r+1} como acabamos de definir es precisamente el producto orlado $P_r \wr \langle c \rangle$. Hagamos notar que P_r está generado por r elementos de orden p.

1.7. Producto Semidirecto

Teorema 1.7.1. *Sean H y K dos grupos tales que para cada elemento $h \in H$, existe un automorfismo de K representado como sigue:*

$$k \leftrightarrows k^h, \quad \forall\, k \in K, \tag{1.7.1}$$

satisfaciendo la propiedad

$$(k^{h_1})^{h_2} = k^{h_1 h_2}, \quad h_1, h_2 \in H. \tag{1.7.2}$$

Entonces, los símbolos $[h, k],\ h \in H,\ k \in K$, forman un grupo respecto a la regla de producto

$$[h_1, k_1] \cdot [h_2, k_2] = [h_1 h_2, k_1^{h_2} k_2], \tag{1.7.3}$$

Demostración: Como para toda k y toda h, $k^h \in K$, la regla del producto (1.7.3) esta bien definida.

1. La regla del producto (1.7.3) es asociativa, ya que

$$\begin{aligned} ([h_1, k_1] \cdot [h_2, k_2]) \cdot [h_3, k_3] &= [h_1 h_2, k_1^{h_2} k_2] \cdot [h_3, k_3] \\ &= [(h_1 h_2) h_3, (k_1^{h_2} k_2)^{h_3} k_3] \\ &= [h_1 h_2 h_3, k_1^{h_2 h_3} k_2^{h_3} k_3] \end{aligned} \tag{1.7.4}$$

y usando (1.7.1) y (1.7.2)

$$\begin{aligned} [h_1, k_1] \cdot ([h_2, k_2] \cdot [h_3, k_3]) &= [h_1, k_1] \cdot [h_2 h_3, k_2^{h_3} k_3] \\ &= [h_1 h_2 h_3, k_1^{h_2 h_3} k_2^{h_3} k_3]. \end{aligned} \tag{1.7.5}$$

2. El elemento $[1,1]$ es la identidad, ya que

$$\begin{aligned}[1,1][h,k] &= [1h,1^h k] = [h,k],\\ [h,k][1,1] &= [h1,k^1 1] = [h,k].\end{aligned} \tag{1.7.6}$$

   Aquí $k^1 = k$ por (1.7.5).

3. Un $[h,k]$ arbitrario tiene un inverso $[h^{-1},(k^{-1})^{h^{-1}}]$, pues

$$[h^{-1},(k^{-1})^{h^{-1}}]\cdot[h,k] = [h^{-1}h, k^{-1}k] = [1,1] \tag{1.7.7}$$

   y

$$[h,k]\cdot[h^{-1},(k^{-1})^{h^{-1}}] = [hh^{-1}, k^{h^{-1}}(k^{-1})^{h^{-1}}] = [1,1]. \tag{1.7.8}$$

De donde los símbolos $[h,k]$ con la regla del producto (1.7.3) forman un grupo. □

De acuerdo con las notaciones e hipótesis del teorema anterior, denotamos por $H \ltimes K$ al conjunto de símbolos $[h,k]$, con $h \in H$ y $k \in K$, que junto con la operación producto del teorema anterior es llamado el **producto semidirecto** o **producto normal** de K por H.

Teorema 1.7.2. *Si G es el producto semiderecto de K por H, entonces los elementos $[h,1]$ de G forman un subgrupo isomorfo a H y los elementos $[1,k]$ forman un subgrupo normal isomorfo a K. Además, los automorfismos (1.7.1) de K, como un subgrupo de G, están inducidas por los automorfismos de conjugación por los elementos de la forma $h = [h,1]$ de H, como un subgrupo de G, ya que*

$$[h,1]^{-1}[1,k][h,1] = [1,k^h]. \tag{1.7.9}$$

Por otra parte, $G = \langle H \cup K\rangle$, ya que

$$[h,1][1,k] = [h,k]. \tag{1.7.10}$$

Demostración: Tenemos solamente que observar que $h \leftrightarrows [h,1]$ y $k \leftrightarrows [1,k]$ son isomorfismos entre H y K y subgrupos de G usando la regla (1.7.3) y notando que $k_1 = k$. También (1.7.9) y (1.7.10) son consecuencia directa de la regla (1.7.3). La ecuación (1.7.9) nos muestra aquí que K es un subgrupo normal y que el automorfismo determinado por (1.7.1) está inducido por el automorfimismo de conjugación por el elemento $h = [h,1]$. Además, $H \cap K = [1,1] = 1$, y (1.7.10) nos muestra que los elementos de H pueden tomarse como representantes de las clases laterales de K. □

Teorema 1.7.3. *G es el producto normal de K por H si y sólo si K es un subgrupo normal de G y H es el subgrupo de G cuyos elementos pueden ser tomados como representantes de las clases laterales de K. Dicho de otra forma*

(*i*) *K es un subgrupo normal de G;*

(*ii*) *H es un subgrupo de G;*

(*iii*) *$K \cap H = \{e\}$;*

(*iv*) *$HK = G$.*

Demostración: Hemos observado ya que estas propiedades se verifican si G es el producto semidirecto de K por H. Recíprocamente, supongamos que estas propiedades se verifican. Entonces de $K \cap H = \{e\}$ y $HK = G$ con K normal en G, se sigue que todo elemento de G tiene una representación única de la forma

$$g = hk. \tag{1.7.11}$$

Como K es normal,

$$h^{-1}kh = k^h \in K \tag{1.7.12}$$

y claramente $k \leftrightarrows k^h$ es un automorfismo de K. Por otra parte, por (1.7.12) tenemos que

$$(k^{h_1})^{h_2} = k^{h_1 h_2}. \tag{1.7.13}$$

Para el producto de dos elementos de G, $g_1 = h_1 k_1$ y $g_2 = h_2 k_2$, se tiene

$$\begin{aligned} g_1 g_2 &= h_1 k_1 h_2 k_2 \\ &= h_1 h_2 (h_2^{-1} k_1 h_2) k_2 \\ &= h_1 h_2 \cdot k_1^{h_2} k_2 \end{aligned} \tag{1.7.14}$$

y la regla del producto en G es por tanto la misma que la (1.7.3), luego G es el producto semidirecto de K por H. □

Observemos que la asociación de un automorfismo de K con un elemento de H es un homomorfismo de H en el grupo de automorfismos de K. Si H está aplicado en el automorfismo identidad de K, es decir, si $k^h = k$ para todo h, k, entonces la regla (1.7.3) es la del producto directo de H y K.

Definición 1.7.1. *Sea K y Q (no necesariamente normal) subgrupos de un grupo G. Entonces Q es un **complemento** de K si $K \cap Q = \{e\}$ y $KQ = G$.*

Definición 1.7.2. *Una* ***retracción*** *(o* ***proyección****) es un homomorfismo* $\varphi : G \to G$ *con* $\varphi \circ \varphi = \varphi$*. Un* ***retracto*** *de* G *es un subgrupo* Q *de* G *con* $Q = \mathrm{Im}\,(\varphi)$*, donde* $\varphi : G \to G$ *es una retracción.*

Observación 1.7.1. Si $\varphi : G \to G$ es una retracción, entonces $\ker(\varphi)$ es el subgrupo de G generado por $\{g\varphi(g^{-1}) \mid g \in G\}$. Además, un grupo G es un producto semidirecto de K por Q si y sólo si es existe una retracción $\varphi : G \to G$ con $K = \ker(\varphi)$ y $Q = \mathrm{Im}\,(\varphi)$.

1.8. Campos Finitos

Definición 1.8.1. *Un* ***campo primo*** *es un campo que no tiene subcampos propios.*

Teorema 1.8.1. *Cada campo primo* K *es isomorfo a* $\mathbb{Z}_p$ *o a* $\mathbb{Q}$*.* ■

Teorema 1.8.2. *Cada campo* F *contiene un único campo primo* K*.* ■

Definición 1.8.2. *Un campo* F *con campo primo* K *tiene* ***caracteristica*** p*, con* p *número primo, si* $K \cong \mathbb{Z}_p$*; en caso contrario, diremos que* F *tiene* ***caracteristica*** 0*, es decir, cuando* $K \cong \mathbb{Q}$*. Por tanto el campo primo de* $\mathbb{R}$ *es* $\mathbb{Q}$*.*

Observe que si F es de caracteristica p, entonces $pa = 0$ para cada $a \in F$.

En esta sección, p representará un número primo.

Definición 1.8.3. *Todo campo con un número finito de elementos es llamado* ***campo finito****.*

Es claro que si F es un campo finito, entonces la característica de F es p, para algún número primo p.

Un resultado bien conocido de Álgebra Lineal es el hecho de que todo espacio vectorial V sobre un campo F admite una base, y que cualesquiera dos bases de V sobre F tienen la misma cardinalidad, lo cual permite definir la dimensión del espacio vectorial V sobre F. Este hecho permite probar la siguiente proposición.

Proposición 1.8.1. *Si F es un campo finito de caracteristica p, entonces F tiene exactamente p^n elementos, para algún $n \geq 1$.* ■

Observemos que existen campos infinitos de caracteristica p, por ejemplo, todas las funciones racionales con coeficientes en $\mathbb{Z}_p$ (una función racional es un cociente de dos polinomios). Además si la caracteristica de K es cero entonces K es un campo infinito.

Teorema 1.8.3. *Si K es un campo y $f(x) \in K[x]$ es un polinomio no constante, entonces existe un campo F conteniendo a K como subcampo sobre el cual $f(x)$ se factoriza como un producto de factores lineales.* ■

Observación 1.8.1. Si F es un campo de caracteristica $p > 0$, entonces para cada $a, b \in F$,

$$(a+b)^{p^k} = a^{p^k} + b^{p^k}, \ \ \forall\, k > 0.$$

Notación.

1. Sea F un campo, denotaremos por F^* al grupo multiplicativo de los elementos no cero de F.
2. Cuando un campo sea finito con q elementos, entonces denotaremos a tal campo por $\mathbb{F}_q$.

Observación 1.8.2.

1. Cada elemento del campo $\mathbb{F}_q$ es raíz del polinomio $x^q - x$.
2. Todo subgrupo multiplicativo finito de F^* es cíclico. Por tanto $\mathbb{F}_q^*$ es cíclico.

Teorema 1.8.4. *Sean p un número primo y n un entero positivo. Entonces, existe un campo F con exactamente p^n elementos.* ■

Teorema 1.8.5 (E. H. Moore, 1983). *Cualesquiera dos campos con p^n elementos son isomorfos.* ■

Definición 1.8.4. *Un elemento* $\rho \in \mathbb{F}_q$ *es un* ***elemento primitivo*** *si* ρ *es un generador del grupo cíclico* $\mathbb{F}_q^*$.

Teorema 1.8.6. *Para cada número primo* p*, el grupo* $\mathrm{Aut}(\mathbb{F}_{p^n})$ *de todos los automorfismos del campo* $\mathbb{F}_{p^n}$ *es cíclico de orden* n*.* ■

Capítulo 2

Algunos Grupos Lineales Simples

2.1. El Grupo Lineal General

El grupo de matrices no singulares son un objeto natural de estudio, así como el grupo de permutaciones. Investigando la estructura de estos grupos, descubriremos una nueva familia de grupos simples.

Aquí, q denotará la potencia de algún primo p.

Definición 2.1.1. *Sea K un campo. El* ***grupo lineal general*** $GL(m, K)$ *es el grupo multiplicativo de todas las matrices no singulares cuadradas de orden m sobre K.*

Teorema 2.1.1. $|GL(m, \mathbb{F}_q)| = (q^m - 1)(q^m - q) \cdots (q^m - q^{m-1})$.

Demostración: Sea V un espacio vectorial m-dimensional sobre $\mathbb{F}_q$, con base ordenada $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$. Considerando a los elementos de $GL(m, \mathbb{F}_q)$ como transformaciones lineales sobre V. Podemos exhibir una biyección entre $GL(m, \mathbb{F}_q)$ y la familia de todas las bases ordenadas de V. En efecto, si $T \in GL(m, \mathbb{F}_q)$, entonces $\{T\alpha_1, T\alpha_2, \ldots, T\alpha_m\}$ es una base ordenada de V (pues T es no singular); si $\{\beta_1, \beta_2, \ldots, \beta_m\}$ es una base ordenada de V, existe una única $T \in GL(m, \mathbb{F}_q)$ tal que $T\alpha_i = \beta_i$ para cada i.

Una base ordenada de V consiste de los vectores $\{\beta_1, \beta_2, \ldots, \beta_m\}$. Ya que existe q^m vectores en V, existen $q^m - 1$ elecciones para β_1; existe así $q^m - q$ elecciones para β_2. Más generalmente, habiendo sido elegido un conjunto linealmente independiente $\{\beta_1, \beta_2, \ldots, \beta_i\}$ la única restricción sobre β_{i+1} es que no pertenezca al subespacio

generado por $\{\beta_1, \beta_2, \ldots, \beta_i\}$; existe así $q^m - q^i$ elecciones para β_{i+1}. Por lo tanto, existen exactamente $(q^m-1)(q^m-q)\cdots(q^m-q^{m-1})$ bases ordenadas de V. □

Definición 2.1.2. *Si t es un entero no negativo y ρ es un elemento primitivo de $\mathbb{F}_q$, entonces definimos el conjunto*

$$M(t) = \{A \in GL(m, \mathbb{F}_q) \mid \det(A) = \rho^{nt}, \textit{para algún } n \in \mathbb{Z}\}.$$

Lema 2.1.1. *Sea t un divisor de $q-1$. Si $\Omega = |GL(m, \mathbb{F}_q)|$, entonces $M(t)$ es un subgrupo normal de $GL(m, \mathbb{F}_q)$ de orden Ω/t.*

Demostración: Sea

$$\det : GL(m, \mathbb{F}_q) \longrightarrow \mathbb{F}_q^*$$

la función determinante. (Notemos que la función determinante es un homomorfismo de grupos).

Si t es un divisor de $q-1 = |\mathbb{F}_q^*|$, entonces el subgrupo cíclico $\langle \rho^t \rangle$ de $\mathbb{F}_q^*$ tiene orden $(q-1)/t$, y por tanto $[\mathbb{F}_q : \langle \rho^t \rangle] = t$. Además, $\langle \rho^t \rangle \lhd \mathbb{F}_q^*$, pues este grupo es abeliano. Dado que $M(t)$ es el subgrupo de $GL(m, \mathbb{F}_q)$ que corresponde a $\langle \rho^t \rangle$ bajo el homomorfismo determinante, tenemos que $M(t) \lhd GL(m, \mathbb{F}_q)$, de índice t, y de orden Ω/t. □

Teorema 2.1.2. *Sea $q-1 = p_1 p_2 \cdots p_k$, donde los p_i's son primos (no necesariamente distintos). La siguiente serie normal es la parte inicial de una serie de composición:*

$$GL(m, \mathbb{F}_q) = M(1) \supset M(p_1) \supset M(p_1 p_2) \supset \cdots \supset M(q-1).$$

Demostración: Ya hemos visto que cada uno de los términos en esta serie es normal en $GL(m, \mathbb{F}_q)$. Además, si $\Omega = |GL(m, \mathbb{F}_q)|$,

$$\left| \frac{M(p_1 p_2 \cdots p_i)}{M(p_1 p_2 \cdots p_i p_{i+1})} \right| = \frac{\Omega/p_1 p_2 \cdots p_i}{\Omega/p_1 p_2 \cdots p_i p_{i+1}} = p_{i+1},$$

es decir, $M(p_1 p_2 \cdots p_i p_{i+1})$ es un subgrupo maximal de $M(p_1 p_2 \cdots p_i)$. □

El último subgrupo en la cadena, es decir, $M(q-1)$ es de especial interés, el cual consiste de todas las matrices de determinante $\rho^{q-1} = 1$.

Definición 2.1.3.

1. *Una matriz es* ***unimodular*** *si tiene determinante* 1.

2. *Sea K un campo. El* ***grupo lineal especial*** *$SL(m, K)$ es el grupo multiplicativo de todas las matrices unimodulares sobre K.*

Por supuesto, tenemos que $SL(m, K)$ es subgrupo normal de $GL(m, K)$.

Definición 2.1.4. *Sea λ un elemento no cero de K y sean $i, j \in \{1, \ldots, m\}$ tales que $i \neq j$. Una* ***transvección*** *$B_{ij}(\lambda)$ es la matriz que se obtiene de agregarle λ a la matriz identidad I en la (i, j)-ésima entrada.*

Observación 2.1.1.

1. Cada transvección es unimodular.

2. $B_{ij}(\lambda)A$ es la matriz obtenida de la matriz A agregandole λ veces la j-ésima fila a la i-ésima fila de A.

Lema 2.1.2. *Si $A \in GL(m, K)$, entonces $A = UD(\mu)$, donde U es un producto de transvecciones, $D(\mu)$ es la matriz diagonal con entradas $\{1, 1, \ldots, 1, \mu\}$ y $\det(A) = \mu$.*

Demostración: Procederemos por inducción sobre el tamaño de la matriz.
Sea t el tamaño de la matriz. Si $t = 1$, entonces claramente se cumple el teorema. Supongamos que para $t < m$ se cumple el teorema, así que hay que probar que para $t = m$ también se cumple.
Sea $A \in GL(m, K)$. Procedamos por casos sobre el valor de la primer entrada de A.

Caso i) Si $a_{11} = 1$, entonces

$$A = \begin{bmatrix} 1 & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{bmatrix}.$$

Si tomamos $B = \prod_{i=2}^{m} B_{i1}(-a_{i1})$, entonces

$$BA = \begin{bmatrix} 1 & a_{12} & \dots & a_{1m} \\ 0 & \alpha_{22} & \dots & \alpha_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_{m2} & \dots & \alpha_{mm} \end{bmatrix}.$$

Notemos que B es producto de transvecciones si para algún $i \in \{2, \dots, m\}$, $a_{i1} \neq 0$. Pues en caso contrario B es la identidad. Además, también observemos que para toda $j = 2, \dots, m$, existe $i_j \in \{2, \dots, m\}$ tal que $\alpha_{i_j j} \neq 0$, por tanto tomando $C = \prod_{j=2}^{m} B_{1i_j}(-\alpha_{i_j j}^{-1} a_{1j})$ (el cual es producto de transvecciones), obtendremos que

$$CBA = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \alpha\prime_{22} & \dots & \alpha\prime_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha\prime_{m2} & \dots & \alpha\prime_{mm} \end{bmatrix}.$$

Por hipótesis de inducción, existen $U, D(\mu) \in GL(m-1, K)$ tales que la submatriz

$$A_{11} = \begin{bmatrix} \alpha\prime_{22} & \dots & \alpha\prime_{2m} \\ \vdots & \ddots & \vdots \\ \alpha\prime_{m2} & \dots & \alpha\prime_{mm} \end{bmatrix} = UD(\mu),$$

donde U es producto de transvecciones en $GL(m-1, K)$ y $D(\mu)$ es la matriz diagonal $(m-1) \times (m-1)$ con entradas $\{1, \dots, 1, \mu\}$ y con $\mu = \det(A_{11}) = \det(A)$.

Es fácil probar que

$$U\prime = \begin{bmatrix} 1 & 0 \\ 0 & U \end{bmatrix}$$

es producto de transvecciones en $GL(m, K)$ y claramente

$$D\prime(\mu) = \begin{bmatrix} 1 & 0 \\ 0 & D(\mu) \end{bmatrix}$$

es la matriz diagonal $m \times m$ con entradas $\{1, \dots, 1, \mu\}$. Por lo tanto

$$\begin{bmatrix} 1 & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & D(\mu) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & A_{11} \end{bmatrix} = CBA \text{ y } A = B^{-1}C^{-1}U\prime D\prime(\mu),$$

donde $B^{-1}C^{-1}U\prime$ es producto de transvecciones en $GL(m, K)$. Por lo que hemos probado el teorema para este caso.

Caso ii) Si $a_{11} = 0$, entonces existe $i \in \{2, \dots, m\}$ tal que $a_{i1} \neq 0$. Así que multiplicando la matriz A por la transvección $B_{1i}(a_{i1}^{-1})$ llegamos al caso (i).

Caso iii) Si $a_{11} \neq 0, 1$ entonces multipliquemos ahora a la matriz A por $B_{12}(1 - a_{11})B_{21}(-a_{11}^{-1}(1 - a_{21}))$ para obtener el caso (i). □

Teorema 2.1.3. *$GL(m, K)$ es el producto semidirecto de $SL(m, K)$ por K^*, y $SL(m, K)$ es generado por las transvecciones.*

Demostración: Sabemos que $M(q - 1) = SL(m, K) \lhd GL(m, K)$. Si Δ es el conjunto de todas las matrices $D(\mu) = \text{diagonal}\{1, \ldots, 1, \mu\}$, entonces Δ es un subgrupo de $GL(m, K)$ isomorfo a K^*. Dado que $\det(D(\mu)) = \mu$ y $\Delta \cap SL(m, K) = \{I\}$, se tiene que $GL(m, K)$ es el producto semidirecto de $SL(m, K)$ por K^*.

Si $A = UD(\mu)$ es una factorización de un elemento A de $GL(m, K)$, como en el lema anterior, entonces $\det(A) = \mu$; se sigue que si A es unimodular, entonces $D(\mu) = D(1) = I$, y así $A = U$ es un producto de transvecciones. □

Notación. *Si K es un campo, denotaremos por $Z_1(m, K)$ el grupo de todas las matrices cuadradas escalares kI de orden m, con $k \in K$ y $k^m = 1$.*

Teorema 2.1.4. *El centro de $SL(m, K)$ es $Z_1(m, K)$.*

Demostración: Es suficiente probar que si $A \in SL(m, K)$ tal que conmuta con todas las transvecciones, entonces A es una matriz escalar kI, ya que en este caso $k^m = 1$ pues $\det(A) = 1$.

Consideremos a $SL(m, K)$ como transformaciones en un espacio vectorial sobre K con base $\{e_1, \ldots, e_m\}$. La transvección $B_{ij}(1)$ es la transformación B la cual transforma a e_j en $e_i + e_j$ y fija a los otros elementos básicos. Además, consideremos a A como la transformación enviando a e_l en $\sum_{k=1}^{m} a_{kl}e_k$.

Si $l \neq j$, entonces $ABe_l = Ae_l$. En otro sentido,

$$BAe_l = B\left(\sum a_{kl}e_k\right) = \sum a_{kl}Be_k = \left(\sum_{k\neq j} a_{kl}e_k\right) + a_{jl}(e_i + e_j) = Ae_l + a_{jl}e_i$$

Puesto que $AB = BA$, se tiene que $a_{jl} = 0$ siempre que $j \neq l$; por lo tanto A es diagonal, es decir, $Ae_l = a_{ll}e_l$. Claramente $ABe_j = a_{ii}e_i + a_{jj}e_j$ mientras que $BAe_j = a_{jj}(e_i + e_j)$. Entonces $a_{ii} = a_{jj}$. Por lo tanto, si A conmuta con todos los $B_{ij}(1)$, entonces A es escalar. □

Teorema 2.1.5. $|Z_1(m, \mathbb{F}_q)|$=d, *donde* $d = (m, q-1)$.

Demostración: Si $G = \langle \rho \rangle$ es un subgrupo cíclico de orden n y d un divisor de n, entonces es fácil ver que $\{x \in G \mid x^d = 1\}$ es un subgrupo cíclico de G de orden d, dado por $\langle \rho^{n/d} \rangle$. Ahora, el homomorfismo

$$\begin{aligned} Z_1(m, \mathbb{F}_q) &\longrightarrow \mathbb{F}_q^* \\ kI &\longmapsto k \end{aligned}$$

es inyectivo. Por tanto, es suficiente con probar que $k^m = 1$ si y sólo si $k^d = 1$, donde $d = (m, q-1)$. Puesto que $m = dc$ para algún entero c, se tiene que $k^d = 1$ ímplica $k^m = k^{dc} = 1$. Recíprocamente, existen enteros a y b con $d = am + b(q-1)$. Así

$$k^d = k^{am+b(q-1)} = k^{ma} k^{b(q-1)} = k^{ma}$$

y así que $k^m = 1$ implica $k^d = 1$. □

Observación 2.1.2. Sea $H \lhd SL(2, K)$ y sea $A \in H$. Si A es similar a $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, entonces existe $\mu \in K^*$ tal que H contiene a $\begin{bmatrix} a & \mu^{-1}b \\ \mu c & d \end{bmatrix}$.

Demostración: Dado que A es similar a $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ entonces existe una matriz $P \in GL(2, K)$ tal que $P^{-1}AP = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Sea $P = UD(\mu)$ la factorización como en el Lema 2.1.2, donde $\det(P) = \det(D(\mu)) = \mu$ y $D(\mu) = \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix}$, así que $PD^{-1}(\mu) \in SL(2, m)$, pues $\det(PD^{-1}(\mu)) = \det(P)\det(D^{-1}(\mu)) = 1$. De donde se sigue que:

$$\begin{aligned} (PD^{-1}(\mu))^{-1} A (PD^{-1}(\mu)) &= D(\mu)(P^{-1}AP)D^{-1}(\mu) \\ &= \begin{bmatrix} 1 & 0 \\ 0 & \mu \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \mu^{-1} \end{bmatrix} \\ &= \begin{bmatrix} a & \mu^{-1}b \\ \mu c & d \end{bmatrix} \in H. \end{aligned}$$

□

2.2. PSL(2,K)

Ahora, extenderemos la serie normal del Teorema 2.1.2 un eslabón más

$$GL(m,K) \supset M(1) \supset M(p_1) \supset \cdots \supset M(q-1) = SL(m,K) \supset Z_1(m,K).$$

Además $Z_1(m,K)$ es un grupo abeliano. Investigaremos el último grupo factor de esta serie.

Definición 2.2.1. *El **grupo unimodular proyectivo** $\mathbf{PSL(m,K)}$ es el grupo cociente $SL(m,K)/Z_1(m,k)$.*

Teorema 2.2.1. *Si $d = (m, q-1)$, entonces*

$$|PSL(m,\mathbb{F}_q)| = \frac{(q^m-1)(q^m-q)\cdots(q^m-q^{m-1})}{d(q-1)}.$$

Demostración: Si $\Omega = |GL(m,\mathbb{F}_q)|$, entonces $|SL(m,\mathbb{F}_q)| = \Omega/(q-1)$. Por el Lema 2.1.1, el teorema se sigue de los Teoremas 2.1.1 y 2.1.5. □

De aquí en adelante, en esta sección nos concentraremos en el caso $m = 2$, con la aspiración de probar que los grupos $PSL(2,\mathbb{F}_q)$ son simples cuando $q > 3$.

Lema 2.2.1. *Si un grupo normal H de $SL(2,\mathbb{F}_q)$ contiene una transvección $B_{ij}(\lambda)$, entonces $H = SL(2,\mathbb{F}_q)$.*

Demostración: Por el Teorema 2.1.3, es suficiente probar que H contiene cada transvección. Denotaremos por K a $\mathbb{F}_q$.

Si conjugamos $B_{12}(\lambda)$ por una matriz unimodular, tenemos

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} 1-ac\lambda & a^2\lambda \\ -c^2\lambda & 1+ac\lambda \end{bmatrix}.$$

En particular, este conjugado es $B_{12}(\lambda a^2)$ si $c = 0$, y es $B_{21}(-\lambda c^2)$ si $a = 0$. Además, estas matrices pertenecen a H, dado que H es normal.
La aplicación $k \longmapsto k^2$ es un endomorfismo del grupo abeliano K^* cuyo núcleo consiste de todas las k con $k^2 = 1$. Dado que K es un campo, el polinomio $x^2 - 1$ tiene a lo más dos raíces, y así el núcleo tiene orden 1 o 2 (tiene orden 1 cuando K tiene caracteristica 2). Se sigue que al menos la mitad de los elementos de K^* son cuadrados.

Sea $\Gamma = \{\lambda \in K \mid B_{12}(\lambda) \in H\} \cup \{0\}$. Es fácil probar que Γ es un subgrupo de K (donde consideramos a K sólo como un grupo aditivo). Además, sabemos que Γ contiene al 0 y a todos los elementos de la forma λa^2. Por lo tanto, Γ contiene más de la mitad de los elementos de K, y así $\Gamma = K$, por el Teorema de Lagrange. Entonces, H contiene todas las transvecciones de la forma $B_{12}(\lambda)$, y un argumento similar prueba que H contiene todas las transvecciones de la forma $B_{21}(\lambda)$. □

Así, las transvecciones juegan un papel importante en el estudio del grupo lineal especial como los 3-ciclos lo juegan en el estudio del grupo alternante. Ahora probaremos el Teorema principal de esta sección.

Teorema 2.2.2 (Jordan-Moore). *Los grupos $PSL(2, \mathbb{F}_q)$ son simples si y sólo si $q > 3$.*

Demostración: Del Teorema 2.2.1 sabemos que

$$|PSL(2, \mathbb{F}_q)| = \begin{cases} (q+1)(q^2-q), & \text{si } q = 2^n; \\ \frac{1}{2}(q+1)(q^2-q), & \text{si } q = p^n \text{ y } p \text{ número primo impar.} \end{cases}$$

Por lo tanto, $|PLS(2, \mathbb{F}_2)| = 6$ y $|PSL(2, \mathbb{F}_3)| = 12$, así que estos grupos no son simples.

Sea $H \lhd SL(2, \mathbb{F}_q)$ tal que $Z_1(m, K) < H$, es decir, es un subgrupo propio. Por lo cual sabemos que existe una matriz $A \in H$, tal que $A \notin Z_1(m, K)$, así que apartir de la matriz A, construiremos una transvección en H y aplicando el Lema 2.2.1, obtendremos que $H = SL(m, K)$.

Procedamos por casos:

Caso i) Si A es una transvección, se obtiene lo deseado.

Caso ii) Supongamos que A no es una transvección, pero tiene la siguiente forma

$$A = \begin{bmatrix} r & 0 \\ s & t \end{bmatrix},$$

con $r \neq \pm 1$. Si

$$S = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},$$

entonces

$$SAS^{-1}A^{-1} = \begin{bmatrix} 1 & 0 \\ 1-t^2 & 1 \end{bmatrix} \in H.$$

Dado que $\det(A) = rt = 1$, entonces $t \neq \pm 1$ y $1 - t^2 \neq 0$. Así que esta última matriz es una transvección.

Caso iii) Si

$$A = \begin{bmatrix} r & s \\ t & u \end{bmatrix},$$

con $s \neq 0$, entonces tomando a

$$B = \begin{bmatrix} 1 & 0 \\ -s^{-1}u & 1 \end{bmatrix} \in SL(2, K)$$

y conjugando a A por B se obtiene que

$$C = BAB^{-1} = \begin{bmatrix} r+u & s \\ -s^{-1} & 0 \end{bmatrix} \in H.$$

Ahora, consideremos la matriz

$$T = \begin{bmatrix} \alpha^{-1} & 0 \\ 0 & \alpha \end{bmatrix},$$

con lo cual obtenemos que

$$TCT^{-1} = \begin{bmatrix} r+u & -\alpha^2 s \\ \alpha^{-2}s^{-1} & r+u \end{bmatrix} \in H,$$

lo cual implica que

$$U = TCT^{-1}C^{-1} = \begin{bmatrix} \alpha^2 & s(r+u)(\alpha^2 - 1) \\ 0 & \alpha^{-2} \end{bmatrix} \in H.$$

Más aún,

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} U \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} \alpha^2 & 0 \\ s(r+u)(\alpha^2 - 1) & \alpha^{-2} \end{bmatrix} \in H.$$

Entonces si existe $\alpha \in \mathbb{F}_q$ tal que $\alpha^{-2} \neq \pm 1$, es decir, $\alpha^4 \neq 1$, J sería una matriz como en el caso (ii) con lo cual habremos terminado.

Si $q > 5$, tal elemento α distinto de cero existe, ya que el polinomio $z^4 - 1$ tiene a lo más 4 raíces en el campo $\mathbb{F}_q$. Si $q = 4$, entonces cada elemento $\alpha \in \mathbb{F}_4$ satisface la ecuación $\alpha^4 = \alpha$, así que si $\alpha \neq 1$, entonces se tiene que $\alpha^4 \neq 1$.

Ahora sólo hace falta el caso para cuando $q = 5$. Consideraremos dos posibilidades.

La primera cuando $r+u \neq 0$, entonces elijamos $\alpha \in \mathbb{F}_5 \cong \mathbb{Z}_5$ tal que $\alpha^2 - 1 \neq 0$; notemos que para $\alpha^2 = -1$ obtenemos que

$$J = \begin{bmatrix} -1 & 0 \\ -2s(r+u) & -1 \end{bmatrix}.$$

Por tanto $J^2 = B_{21}(4s(r+u)) \in H$, es decir, H contiene una transvección. Ahora si $r+u=0$, entonces tomemos

$$C = \begin{bmatrix} 0 & -s^{-1} \\ s & 0 \end{bmatrix} \in H.$$

Por tanto H contiene a

$$\begin{bmatrix} 1 & 2s^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -s^{-1} \\ s & 0 \end{bmatrix} \begin{bmatrix} 1 & -2s^{-1} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ s & -2 \end{bmatrix},$$

es decir, una matriz deseada. Con este último caso concluimos la demostración del teorema. □

2.3. PSL(m,K)

En esta sección se probará que $PSL(m,K)$ es simple para cada campo K y todo $m \geq 3$. Como una consecuencia de esto, seremos capacez de exhibir dos grupos simples finitos del mismo orden y que no sean isomorfos.

Notación. Si V es un espacio vectorial sobre un campo K, entonces $GL(V)$ denota al grupo de todas las transformaciones lineales no singulares. El grupo de todas las transformaciones lineales de determinante 1 es denotado por $SL(V)$.

Claramente, si $\dim(V) = m$, entonces eligiendo una base de V podemos definir un isomorfismo tal que $GL(V) \cong GL(m,K)$ y $SL(V) \cong SL(m,K)$.

Definición 2.3.1. *Sea K un campo y V un espacio vectorial de dimensión m sobre K, decimos que un subespacio H de V es un* ***hiperplano*** *en V si la dimensión de H sobre K es $m-1$.*

Si H es un hiperplano en V y $\alpha \notin H$, entonces cada vector $\beta \in V$ tiene una única expresión de la forma

$$\beta = \mu\alpha + \gamma, \quad \text{con} \quad \mu \in K \quad \text{y} \quad \gamma \in H.$$

(Bajo el epimorfismo canónico

$$\pi : V \longrightarrow V/H \cong K$$
$$\beta \longmapsto \beta + H$$

los múltiplos escalares de α forman el conjunto completo de representantes de clases de H en V.)

Lema 2.3.1. *Sea H un hiperplano en V y sea $A \in GL(V)$ fijando a H (punto por punto). Si $\alpha \in V$ y $\alpha \notin H$, entonces existe $\mu \in K$ tal que*

$$A(\alpha) = \mu\alpha + \gamma, \quad con \quad \gamma \in H.$$

Además, para $\beta \in V$,

$$A(\beta) = \mu\beta + \gamma\prime, \ \ \gamma\prime \in H.$$

Demostración: Sabemos que todo $\beta \in V$ puede ser expresado de la forma

$$\beta = b\alpha + \gamma", \ \ \text{con} \ \ b \in K \ \ \text{y} \ \ \gamma" \in H.$$

Entonces

$$\begin{aligned} A(\beta) &= A(b\alpha + \gamma") \\ &= bA(\alpha) + \gamma" \\ &= b(\mu\alpha + \gamma) + \gamma" \\ &= \mu(b\alpha + \gamma") + [(1-\mu)\gamma" + b\gamma] \\ &= \mu\beta + \gamma\prime, \quad \text{donde} \quad \gamma\prime = (1-\mu)\gamma" + b\gamma \in H. \end{aligned} \tag{2.3.1}$$

□

Definición 2.3.2. *Sea $A \in GL(V)$ una transformación lineal fijando a cada punto de un hiperplano H de V y sea $\mu = \mu(A)$ (es decir, μ depende de A). Si $\mu \neq 1$, entonces A es llamada **dilatación**; si $\mu = 1$ y $A \neq I$, entonces A es llamada **transvección**.*

En el siguiente teorema probaremos que si una transformación A es una transvección, entonces existe una matriz representando a A tal que A es una transvección $B_{ij}(\lambda)$.

Teorema 2.3.1. *Sea $A \in GL(V)$ fijando a cada uno de los puntos del hiperplano H.*

1. *Si A es una dilatación, entonces A puede ser representada (relativa a una base conveniente de V) por una matriz de la forma $D(\mu) = diagonal\{1, \ldots, 1, \mu\}$.*

2. *Si A es una transvección, entonces A puede ser representada por $B_{21}(1)$. Además, una transvección no tiene vectores característicos fuera de H.*

Demostración:

1. Elijamos $\alpha \in V,\ \alpha \notin H$, y sea $A(\alpha) = \mu\alpha + \gamma,\ \mu \in K,\ \gamma \in H$. Claramente, cada vector distinto de cero en H es un vector característico de A. Veamos que si existen otros vectores característicos. Sea $\beta \in V$, entonces $\beta = b\alpha + \gamma\prime,\ \ b \in K,\ \gamma\prime \in H$. De la Ecuación (2.3.1), se obtiene que

$$A(\beta) = c\beta \iff \mu = c \text{ y } b\gamma = (\mu - 1)\gamma\text{".} \tag{2.3.2}$$

   Si A es una dilatación, entonces $\mu - 1 \neq 0$ y podemos resolverlo para $\gamma\text{"} = b(\mu - 1)^{-1}\gamma$.

   Por lo tanto, cada elección de $b \neq 0$ da un vector característico β fuera de H. Si $\{\tau_1, \ldots, \tau_{m-1}\}$ es una base de H, entonces $\{\tau_1, \ldots, \tau_{m-1}, \beta\}$ es una base de V, y escribiendo a A con respecto a esta base obtendremos la matriz $D(\mu)$.

2. Si ahora suponemos que A es una transvección, entonces $\mu = 1$ y, por la Ecuación (2.3.2), se tiene que A tiene un vector característico fuera de H si y sólo si $\gamma = 0$. Esto implica que $A = I$, contradicción.

   Por lo tanto, A no tiene vectores característicos fuera de H. Sea τ_1 un vector fuera de H, entonces definamos $\tau_2 = A\tau_1 - \tau_1$. Notemos que $A\tau_1 = \tau_1 + \gamma\prime$, es decir, $\tau_2 = A\tau_1 - \tau_1 = \gamma\prime = (\tau_1 + \gamma\prime) - \tau \in H$. Así que, el conjunto $\{\tau_1, \tau_2\}$ es linealmente independiente. Extendemos $\{\tau_1, \tau_2\}$ a una base $\{\tau_1, \ldots, \tau_m\}$ de V, donde $\{\tau_3, \ldots, \tau_m\}$ son elementos de H. Relativa a esta base, A tiene la representación matricial $B_{21}(1)$, pues $A\tau_1 = \tau_1 + \tau_2$ y $A\tau_i = \tau_i$ para $i \neq 1$.

□

Corolario 2.3.1. *Todas las matrices $B_{ij}(\lambda)$, donde $i \neq j$ y $\lambda \neq 0$, son conjugados en $GL(m, K)$.*

Demostración: Es suficiente probar que $B_{ij}(\lambda)$ también representa la transvección A. Eligiendo τ_i fuera de H y $\tau_j = \lambda^{-1}(A\tau_i - \tau_i)$; para $l \neq i, j$, eligimos $\tau_l \in H$ tal que $\{\tau_1, \ldots, \tau_m\}$ es una base de V. Relativa a esta base, A tiene la representación matricial $B_{ij}(\lambda)$. □

Corolario 2.3.2. *$GL(V)$ es generado por todas las dilataciones y transvecciones, y $SL(V)$ es generado por todas las transvecciones.*

Demostración: Claramente este corolario se sigue del Lema 2.1.2 pero trasladando las matrices a transformaciones lineales y aplicando el Teorema 2.3.1. □

Podemos ver las transvecciones de una forma mas conveniente. Si A es una transvección de V fijando un hiperplano H y $\alpha \in V - H$, entonces

$$A(\alpha) = \alpha + \gamma, \quad \text{con} \quad \gamma \in H.$$

Si $\beta \in V$, entonces

$$\beta = b\alpha + \gamma\prime, \quad \text{con} \quad b \in K,\ \gamma\prime \in H, \tag{2.3.3}$$

y aplicando la Ecuación (2.3.1) obtenemos que

$$A(\beta) = (b\alpha + \gamma\prime) + [(1-1)\gamma\prime + b\gamma] = \beta + b\gamma.$$

Por lo tanto, definiendo la función

$$\begin{aligned} f : V &\longrightarrow K \\ \beta &\longmapsto b \end{aligned}$$

y usando la Ecuación (2.3.3), se observa que la función es K-lineal (la cual es una funcional sobre V) y tiene núcleo H. Así, A determina una funcional f y un vector $\gamma \in H = \ker(f)$ tal que A satisface la ecuación

$$A(\beta) = \beta + f(\beta)\gamma, \qquad \forall\ \beta \in V.$$

Notación. Dado un funcional f y $\gamma \in \ker(f)$, definimos

$$\begin{aligned} T_{f,\gamma} : V &\longrightarrow V \\ \beta &\longmapsto \beta + f(\beta)\gamma. \end{aligned}$$

Es fácil ver que $T_{f,\gamma}$ es una transvección, donde $f \not\equiv 0$ y $\gamma \neq 0$. Además, cada transvección es de la forma $T_{f,\gamma}$ para algún f y γ.

Observación 2.3.1.

1. Si f, g son funcionales y $f(\gamma) = g(\gamma) = f(\gamma\prime) = 0$, entonces

$$T_{f,\gamma} \circ T_{f,\gamma\prime} = T_{f,\gamma+\gamma\prime} \quad \text{y} \quad T_{f,\gamma} \circ T_{g,\gamma} = T_{f+g,\gamma}.$$

2. $T_{af,\gamma} = T_{f,a\gamma}, \quad \forall\ a \in K.$
3. $T_{f,\gamma} = T_{g,\delta} \iff \exists\ a \in K,$ con $g = af$ y $\gamma = a\delta$.
4. Si $S \in GL(V)$, entonces $ST_{f,\gamma}S^{-1} = T_{fS^{-1},S\gamma}$.

Demostración:

1. Sea $\beta \in V$, entonces

$$\begin{aligned} T_{f,\gamma} \circ T_{f,\gamma\prime}(\beta) &= T_{f,\gamma}(\beta + f(\beta)\gamma\prime) \\ &= (\beta + f(\beta)\gamma\prime) + f(\beta + f(\beta)\gamma\prime)\gamma \\ &= (\beta + f(\beta)\gamma\prime) + (f(\beta) + f(\beta)f(\gamma\prime))\gamma \\ &= (\beta + f(\beta)\gamma\prime) + f(\beta)\gamma \\ &= \beta + f(\beta)(\gamma\prime + \gamma) \\ &= T_{f,\gamma\prime+\gamma}(\beta) \end{aligned}$$

   y

$$\begin{aligned} T_{f,\gamma} \circ T_{g,\gamma}(\beta) &= T_{f,\gamma}(\beta + g(\beta)\gamma) \\ &= (\beta + g(\beta)\gamma) + f(\beta + g(\beta)\gamma)\gamma \\ &= (\beta + g(\beta)\gamma) + f(\beta)\gamma + g(\beta)f(\gamma)\gamma \\ &= (\beta + g(\beta)\gamma) + f(\beta)\gamma \\ &= \beta + (g(\beta) + f(\beta))\gamma \\ &= \beta + (g + f)(\beta)\gamma \\ &= T_{g+f,\gamma}(\beta). \end{aligned}$$

2. Sean $a \in K$ y $\beta \in V$, entonces

$$T_{af,\gamma}(\beta) = \beta + (af)(\beta)\gamma = \beta + af(\beta)\gamma = \beta + f(\beta)(a\gamma) = T_{f,a\gamma}(\beta).$$

3. $\Rightarrow$) Supongamos que $T_{f,\gamma} = T_{g,\delta}$, entonces para cada $\beta \in V$, $f(\beta) \neq 0$;

$$T_{f,\gamma}(\beta) = \beta + f(\beta)\gamma = \beta + g(\beta)\delta = T_{g,\delta},$$

   en particular, tomemos β tal que $g(\beta) \neq 0 \neq f(\beta)$ (el cual sabemos que existe pues en caso contrario $T_{f,\gamma}$ sería la identidad), por tanto $f(\beta)\gamma = g(\beta)\delta$ implica que $\gamma = f^{-1}(\beta)g(\beta)\delta$. Así que, tomando $a = f(\beta)^{-1}g(\beta)$, obtenemos lo deseado.

   Por último, dado lo anterior se tiene que $T_{f,\gamma} = T_{f,a\delta}$, y por el inciso anterior $T_{f,\gamma} = T_{f,a\delta} = T_{af,\delta} = T_{g,\delta}$, es decir, $af = g$.

   $\Leftarrow$) Ahora, si para algún $a \in K$, $af = g$ y $\gamma = a\delta$, entonces $T_{f,\gamma} = T_{a^{-1}g,a\delta} = T_{g,a^{-1}a\delta} = T_{g,\delta}$.

4. Sea $S \in GL(V)$. Entonces

$$\begin{aligned} ST_{f,\gamma}S^{-1}(\beta) &= ST_{f,\gamma}(S^{-1}(\beta)) \\ &= S\{S^{-1}(\beta) + f(S^{-1}(\beta))\gamma\} \\ &= \beta + S\{f(S^{-1}(\beta))\gamma\} \\ &= \beta + (fS^{-1})(\beta)S\gamma \\ &= T_{fS^{-1},S\gamma}(\beta). \end{aligned}$$

□

De la observación anterior, podemos notar que el conjugado de una transvección es también una transveción. Por este motivo se prefiere trabajar con transformaciones lineales en lugar de matrices, puesto que el conjugado de $B_{ij}(\lambda)$ no necesariamente es una matriz de la misma forma.

Teorema 2.3.2. *El subgrupo conmutador de $GL(V)$ es $SL(V)$ a menos que V sea un espacio vectorial de dimensión dos sobre $\mathbb{F}_2$.*

Demostración: Sabemos que $\det : GL(V) \longrightarrow K^*$ es un epimorfismo de grupos con núcleo $SL(V)$ y por lo tanto $GL(V)/SL(V) \cong K^*$; dado que K^* es abeliano, $(GL(V))\prime \subseteq SL(V)$.

Ahora, veamos inclusión contraria. Sea $\pi : GL(V) \longrightarrow GL(V)/(GL(V))\prime$ el epimorfismo canónico. Del Corolario 2.3.1, se sabe que las transvecciones son conjugadas en $GL(V)$, así que $\pi(T) = \pi(T\prime)$ para cualesquiera dos transvecciones $T, T\prime \in GL(V)$. Si $d = \pi(T) \in GL(V)/(GL(V))\prime$, evitamos el caso excepcional de la afirmación, entonces cada hiperplano H contiene vectores $\gamma, \gamma\prime$ distintos de cero (no necesariamente distintos entre sí) tales que $\gamma + \gamma\prime$ también es distinto de cero. Elijamos un hiperplano H, vectores $\gamma, \gamma\prime, \gamma + \gamma\prime \in H$, y un funcional f teniendo núcleo H. Así que, por la Observación 2.3.1, se tiene que

$$T_{f,\gamma} \circ T_{f,\gamma\prime} = T_{f,\gamma+\gamma\prime}. \tag{2.3.4}$$

Dado que los vectores $\gamma, \gamma\prime, \gamma + \gamma\prime$ son distintos de cero, cada uno de estos es una transvección, distinta de la identidad. Aplicando π a la Ecuación (2.3.4), obtenemos que $d^2 = \pi(T_{f,\gamma})\pi(T_{f,\gamma\prime}) = \pi(T_{f,\gamma+\gamma\prime}) = d \in K^*$, por lo tanto $d = 1$. Así, cada transvección está en $\ker(\pi) = GL(V)\prime$. Dado que $SL(V)$ es generado por las transvecciones, $SL(V) \subset GL(V)\prime$, como deseabamos.

Finalmente, cuando V es un espacio vectorial de dimensión dos sobre $\mathbb{F}_2$, entonces $GL(V)$ es una excepción:

$$GL(V) = SL(V) \cong S_3,$$

donde $S_3\prime = A_3$ es un subgrupo propio. □

Hasta este momento, hemos permitido que $m = \dim(V) \geq 2$; la condición $m \geq 3$ será crucial en el siguiente resultado.

Teorema 2.3.3. *Si $m \geq 3$, entonces todas las transvecciones son conjugadas en $SL(V)$.*

Demostración: Sabemos que las transvecciones son conjugadas en $GL(V)$; queremos probar que dicha conjugación se da en $SL(V)$ cuando $m \geq 3$.

Sean $T_{f,\gamma}$ y $T_{f\prime,\gamma\prime}$ transvecciones, y sean H y $H\prime$ hiperplanos fijados por $T_{f,\gamma}$ y $T_{f\prime,\gamma\prime}$ respectivamente. Eligimos vectores $\beta, \beta\prime \in V$ tales que

$$f(\beta) = 1 = f\prime(\beta\prime).$$

Afirmamos que existe una transformación S en $GL(V)$ tal que

$$S(\gamma) = \gamma\prime,\ S(H) = H\prime \ \text{ y } \ S(\beta) = \beta\prime.$$

En efecto, existe una base de H la cual posee al elemento γ, sea ésta $\{\gamma, \gamma_2, \ldots, \gamma_{m-1}\}$; similarmente, existe una base $\{\gamma\prime, \gamma\prime_2, \ldots, \gamma\prime_{m-1}\}$ de $H\prime$ conteniendo a $\gamma\prime$. Dado que H y $H\prime$ son hiperplanos y $\beta \notin H$, $\beta\prime \notin H\prime$, entonces $\{\beta, \gamma, \gamma_2, \ldots, \gamma_{m-1}\}$ y $\{\beta\prime, \gamma\prime, \gamma\prime_2, \ldots, \gamma\prime_{m-1}\}$ son bases de V. Así que, de estas dos últimas bases de V, podemos suponer que $S : V \longrightarrow V$ es la transformación lineal la cual envia la primera de estas bases ordenadas sobre la segunda.

Más aún, si $m \geq 3$, afirmamos que podemos elegir S con determinante 1. Pues si $\det(S) = d$, entonces, puesto que $m \geq 3$, la primera base de V construida anteriormente contiene al menos un vector distinto de β y γ. Así que, podemos suponer, ahora, que la transformación lineal S es tal que $S(\gamma_{m-1}) = d^{-1}\gamma\prime_{m-1}$. Relativa a la base $\{\beta, \gamma, \gamma_2, \ldots, \gamma_{m-1}\}$, la nueva transformación lineal tiene como matriz a la original de S con la última columna mutiplicada por d^{-1}. Esta nueva transformación lineal tiene determinante 1 así como las otras propiedades de S.

Veamos que $ST_{f,\gamma}S^{-1} = T_{f\prime,\gamma\prime}$, con lo cual completaremos la prueba. De la Observación 2.3.1, sabemos que $ST_{f,\gamma}S^{-1} = T_{fS^{-1},S\gamma}$, con $S\gamma = \gamma\prime$; pero además, $fS^{-1} = f\prime$, ya que estas coinciden sobre la base $\{\beta\prime, \gamma\prime, \gamma\prime_2, \ldots, \gamma\prime_{m-1}\}$ de V. □

Necesitamos un poco más de información acerca de las transvecciones, para posteriormente probar el teorema principal.

Notación. Si H es un hiperplano, entonces $\mathbf{T}(H)$ es el conjunto de transvecciones fijando a H junto con la identidad.

[$\mathbf{T}(H)$ es el análogo del subgrupo Γ en el caso 2×2 de la demostración del Lema 2.2.1].

Notación. Para un espacio vectorial V sobre un campo K, denotaremos por $Sc_1(V)$ al subgrupo de todas las transformaciones escalares de determinante 1, y escribiremos $PSL(V) = SL(V)/Sc_1(V)$.

Si V es un espacio vectorial m-dimensional sobre un campo K, entonces eligiendo una base de V podemos definir un isomorfismo $PSL(V) \cong PSL(m, K)$.

Lema 2.3.2. *Sea H un hiperplano de V.*

1. *Existe una funcional f, con $H = \ker(f)$, tal que $\mathbf{T}(H) = \{T_{f,\gamma} \mid \gamma \in H\}$;*
2. *$\mathbf{T}(H)$ es un subgrupo abeliano de $SL(m, K)$; de hecho, $\mathbf{T}(H) \cong H$;*
3. *El centro $C_{SL}(\mathbf{T}(H)) = Sc_1(V) \cdot \mathbf{T}(H)$.*

Demostración:

1. Asumamos que $T_{f,\gamma}$ y $T_{f\prime,\gamma\prime}$ están en $\mathbf{T}(H)$ y que $f \not\equiv 0$. Sea $\{\alpha_1, \ldots, \alpha_m\}$ una base de V cuyos primeros $m-1$ términos pertenecen a H. Si $f(\alpha_m) = a$ y $f\prime(\alpha_m) = a\prime$, entonces sea $b = a\prime a^{-1}$; es fácil probar que $f\prime = bf$. Por la Observación 2.3.1,
$$T_{f\prime,\gamma\prime} = T_{bf,\gamma\prime} = T_{f,b\gamma\prime}.$$
2. Verifiquemos que $\mathbf{T}(H)$ es un subgrupo de $SL(V)$, isomorfismo a H. Sean $T_{f,\gamma}, T_{f,\gamma\prime} \in \mathbf{T}(H)$ dos transvecciones. (Del inciso anterior sabemos que el funcional f es el mismo y $\ker(f) = H$.) Así que, por la Observación 2.3.1, $T_{f,\gamma} \circ T_{f,\gamma\prime} = T_{f,\gamma+\gamma\prime}$ y el inverso de $T_{f,\gamma}$ es $T_{f,-\gamma}$. Con lo que se demuestra que $T(H)$ es un grupo.

   Pero aún más, $T(H)$ es abeliano, pues notemos que
$$T_{f,\gamma} \circ T_{f,\gamma\prime} = T_{f,\gamma+\gamma\prime} = T_{f,\gamma\prime+\gamma} = T_{f,\gamma\prime} \circ T_{f,\gamma}.$$

3. Dado que $\mathbf{T}(H)$ es abeliano, tenemos que $C_{SL}(\mathbf{T}(H)) \supset Sc_1(V) \cdot \mathbf{T}(H)$. Para la inclusión inversa, suponga $A \in SL(V)$ y

$$AT_{f,\gamma}A^{-1} = T_{f,\gamma}, \quad \forall\, T_{f,\gamma} \in \mathbf{T}(H).$$

Por la Observación 2.3.1 se sabe que

$$T_{fA^{-1},A\gamma} = T_{f,\gamma};$$

y existe un escalar $a \in K$, con $fA^{-1} = af$ y $\gamma = aA\gamma$.

Por lo tanto,

$$\begin{aligned} A(\gamma) &= a^{-1}\gamma, \quad \forall\, \gamma \in H \\ \text{y} \qquad fA &= a^{-1}f. \end{aligned}$$

Si $\{\alpha_1, \ldots, \alpha_m\}$ es una base de V cuyos primeros $m-1$ términos pertenecen a H, entonces dado $\beta \in V$ y $a_1, \ldots, a_m \in K$ tales que $\beta = a_1\alpha_1 + \cdots + a_m\alpha_m$ obtenemos por las dos ecuaciones anteriores que:

$$\begin{aligned} A(\beta) &= A(a_1\alpha_1 + \cdots + a_m\alpha_m) \\ &= A(a_1\alpha_1 + \cdots + a_{m-1}\alpha_{m-1}) + a_mA\alpha_m \\ &= a^{-1}(a_1\alpha_1 + \cdots + a_{m-1}\alpha_{m-1}) + a_mA\alpha_m \\ &= a^{-1}(a_1\alpha_1 + \cdots + a_{m-1}\alpha_{m-1}) + (a^{-1}a_m\alpha_m - a^{-1}a_m\alpha_m) + a_mA\alpha_m \\ &= a^{-1}(a_1\alpha_1 + \cdots + a_m\alpha_m) + a_m(-a^{-1}\alpha_m + A\alpha_m), \end{aligned}$$

donde $\gamma\prime := a_m(-a^{-1}\alpha_m + A\alpha_m) \in \ker(f) = H$, pues

$$\begin{aligned} f(a_m(-a^{-1}\alpha_m + A\alpha_m)) &= -a_ma^{-1}f(\alpha_m) + a_mf(A\alpha_m) \\ &= -a_ma^{-1}f(\alpha_m) + a_m(a^{-1}f(\alpha_m)) \\ &= 0. \end{aligned}$$

Por lo que $A(\beta) = a^{-1}\beta + \gamma\prime$ es un múltiplo escalar de una transvección, como deseabamos.

□

Teorema 2.3.4 (Jordan-Dickson). *Los grupos $PSL(V)$ son simples siempre que V sea un espacio vectorial de dimensión ≥ 3 sobre un campo arbitrario K.*

Demostración: Probemos que si $G \lhd SL(V)$ y $G \nsubseteq Sc_1(V)$, entonces $G = SL(V)$. Por el Teorema 2.3.3, será suficiente con probar que G contiene una transvección. Sea $A \in G$, con $A \notin Sc_1(V)$. Entonces, existe una transvección T que no es conmutativa con A, luego

$$B = (T^{-1}AT)A^{-1} \neq I.$$

Notemos que $B \in G$, ya que G es normal. Ahora

$$B = T^{-1}(ATA^{-1}) = T_1T_2,$$

donde T_i es una transvección, $i = 1, 2$. Si $T_i = T_{f_i,\gamma_i}$ y $H_i = \ker(f_i)$, entonces, para cada $\beta \in V$,

$$T_i(\beta) = \beta + f_i(\beta)\gamma_i, \quad \text{donde } \gamma_i \in H_i, \ \ i = 1, 2.$$

Sea W el subespacio de V generado por γ_1 y γ_2, por lo que $\dim(W) \leq 2$. Dado que $\dim(V) \geq 3$, existe un hiperplano H de V conteniendo a W. Ahora

$$B(H) \subset H, \tag{2.3.5}$$

pues si $\eta \in H$, entonces

$$\begin{aligned} B(\eta) &= T_1T_2(\eta) = T_2(\eta) + f_1(T_2(\eta))\gamma_1 \\ &= \eta + f_2(\eta)\gamma_2 + f_1(T_2(\eta))\gamma_1 \in H + W = H. \end{aligned}$$

Afirmamos que $H_1 \cap H_2 \neq \{0\}$. Si $H_1 = H_2$, entonces la afirmación es cierta. Si $H_1 \neq H_2$, entonces $H_1 + H_2 = V$ (los hiperplanos son subespacios maximales) y $\dim(H_1 + H_2) = m$. Se sigue de

$$\dim(H_1 \cap H_2) + \dim(H_1 + H_2) = \dim H_1 + \dim H_2$$

que $\dim(H_1 \cap H_2) = m-2 \geq 1$, de donde $H_1 \cap H_2 \neq \{0\}$. Elijamos $\zeta \in H_1 \cap H_2, \ \zeta \neq 0$. Entonces

$$B(\zeta) = T_1T_2(\zeta) = \zeta. \tag{2.3.6}$$

Podemos asumir que B no es transvección (ya que en caso contrario habremos terminado pues $B \in G$). Por tanto, $B \notin \mathbf{T}(H)$, el cual está totalmente constituído de transvecciones. Si $B = aS$, donde $a \in K$ y $S \in \mathbf{T}(H)$, entonces la Ecuación (2.3.6) establece que ζ es un vector característico de S. Por el Teorema 2.3.1, $\zeta \in H$ y $S(\zeta) = \zeta$. Así,

$$\zeta = B(\zeta) = aS(\zeta) = a\zeta$$

y $a = 1$; entonces $B = S \in \mathbf{T}(H)$, lo cual es una contradicción. Por tanto

$$B \notin Sc_1(V) \cdot \mathbf{T}(H) = C_{SL}(\mathbf{T}(H)),$$

por el Lema 2.3.2. Existe así una transvección $U \in \mathbf{T}(H)$ tal que

$$C = UBU^{-1}B^{-1} \neq I.$$

Claramente $C = (UBU^{-1})B^{-1} \in G$. Además, si $\eta \in H$,

$$C(\eta) = UBU^{-1}B^{-1}(\eta) = UB(B^{-1}(\eta)) = \eta,$$

ya que $B^{-1}(\eta) \in H$ (ver (2.3.5)), y $U \in \mathbf{T}(H)$ fija a cada punto de H. Por lo tanto C fija a H; entonces, C es ya sea una dilatación o una transvección. Pero C no es una dilatación ya que $\det(C) = 1$. Entonces C es una transvección en G, y la prueba queda completada. □

Daremos una prueba diferente del Teorema 2.2.2 y del Teorema 2.3.4 en el siguiente capítulo.

Observemos que $|PSL(3, \mathbb{F}_4)| = 20160 = \frac{1}{2}8!$, por lo tanto $PSL(3, \mathbb{F}_4)$ y A_8 tienen el mismo orden. Probaremos que estos subgrupos simples no son isomorfos.

Lema 2.3.3. *Sean H un subgrupo normal de índice primo en un grupo finito G y $h \in H$. Si existe un elemento que no está en H el cual conmuta con h, entonces un elemento $k \in H$ conjugado de h en G es conjugado a h en H.*

Demostración: Basta probar que la órbita de h en H coincide con la órbita de h en G, es decir, $[H : C_H(h)] = [G : C_G(h)]$. Usando la fórmula del producto, y el Segundo Teorema de Isomorfismo, tenemos que

$$[H : C_H(h)] = [H : C_G(h) \cap H] = |HC_G(h)|/|C_G(h)| = [G : C_G(h)]$$

[esta última ecuación resulta de $HC_G(h) = G$ puesto que $HC_G(h)$ es un subgrupo G ($H \lhd G$) conteniendo propiamente al subgrupo maximal H de G ($[G : H]$ es primo)]. □

Teorema 2.3.5 (Schottenfels, 1900). *A_8 y $PSL(3, \mathbb{F}_4)$ son grupos simples no isomorfos del mismo orden.*

Demostración: Los elementos (1 2)(3 4) y (1 2)(3 4)(5 6)(7 8) son permutaciones pares, y no son conjugados en A_8 (pues estos elementos no tienen la misma estructura cíclica en S_8). En contraste, probaremos que todos los elementos de orden 2 en

$PSL(3,\mathbb{F}_4)$ son conjugados, y esto probará el teorema.

Una matriz no escalar $A \in SL(3,\mathbb{F}_4)$ corresponde a un elemento de orden 2 en $PSL(3,\mathbb{F}_4)$ si y sólo si A^2 es escalar, equivalentemente si $(P^{-1}AP)^2$ es escalar para alguna matriz P no-singular. Así que, A puede ser reemplazada por cualquier matriz similar a ella; por lo tanto, podemos suponer que A es una forma canónica racional. Si A es una suma directa de 1×1 matrices compañeras, entonces $A =$diagonal$\{a,b,c\}$; pero A^2 escalar implica que $a^2 = b^2 = c^2$, con lo cual $a = b = c$, ya que $\mathbb{F}_4$ es de caracteristica 2; luego, A sería escalar, lo cual es absurdo. Si A es una matriz compañera 3×3 con $A^2 = \gamma I$, entonces A satisface la ecuación $x^2-\gamma = 0$; pero el mínimo polinomio de una matriz compañera coincide con su polinomio característico, el cual en este caso tiene grado 3. Concluyendo que A tiene forma

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & 0 & b \\ 0 & 1 & c \end{bmatrix}.$$

Ahora, $1 = \det(A) = -ab = ab$ ($-1 = 1$ en $\mathbb{F}_4$) lo cual implica que $b = a^{-1}$ y, como A^2 escalar, necesariamente $c = 0$. Así,

$$A = \begin{bmatrix} a & 0 & 0 \\ 0 & 0 & a^{-1} \\ 0 & 1 & 0 \end{bmatrix}.$$

Existen solo tres tales matrices; si α es un elemento primitivo de $\mathbb{F}_4$, estas son:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}; \qquad B = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & \alpha^2 \\ 0 & 1 & 0 \end{bmatrix}; \qquad C = \begin{bmatrix} \alpha^2 & 0 & 0 \\ 0 & 0 & \alpha \\ 0 & 1 & 0 \end{bmatrix}.$$

Note que $A^2 = I,\ B^2 = \alpha^2 I$, y $C^2 = \alpha I$. Se sigue que si $M \in SL(3,\mathbb{F}_4)$ tiene orden 2, entonces M es similar a A, es decir, existe $P \in GL(3,\mathbb{F}_4)$, con $M = P^{-1}AP$. En particular, $\alpha^2 B$ y αC tienen orden 2, y existen $P, Q \in GL(3,\mathbb{F}_4)$ tales que

$$P^{-1}AP = \alpha^2 B \quad \text{y} \quad Q^{-1}AQ = \alpha C. \tag{2.3.7}$$

Ahora, $SL(3,\mathbb{F}_4)$ es un subgrupo normal de indice 3 en $GL(3,\mathbb{F}_4)$, ya que

$$GL(3,\mathbb{F}_4)/SL(3,\mathbb{F}_4) \cong \mathbb{F}_4^*.$$

Y la matriz

$$\begin{bmatrix} \alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

conmuta con A. El Lema 2.3.3 establece que podemos asumir que P y Q se pueden tomar en $SL(3, \mathbb{F}_4)$. Finalmente, la Ecuación (2.3.7) prueba que A, B y C son conjugados en $PSL(3, \mathbb{F}_4)$. Hemos probado que todos los elementos de orden 2 en $PSL(3, \mathbb{F}_4)$ son conjugados, y esto prueba que $A_8 \ncong PSL(3, \mathbb{F}_4)$. □

Otra manera para probar el teorema anterior es observando que A_8 contiene un elemento de orden 15, digamos, (1 2 3)(4 5 6 7 8), pero $PSL(3, \mathbb{F}_4)$ no contiene tal elemento.

Existe una infinidad de parejas de grupos simples no isomorfos entre sí teniendo el mismo orden, pero no existen tres grupos simples no isomorfos del mismo orden.

Capítulo 3

Permutaciones. Geometría Afín y Proyectiva

Los grupos de Mathieu son cinco grupos simples notables, descubiertos por Mathieu en 1861. Este capítulo está dedicado a probar la existencia de tales grupos, así como de mostrar la simplicidad de ellos.

3.1. G-conjuntos

En el Capítulo 1 hemos recordado lo que es la acción de un grupo G sobre un conjunto no vacío X, obteniendo algunas propiedades sobre las órbitas y estabilizadores. En esta sección obtendremos nuevos conceptos, asumiendo de aquí en adelante que todos los grupos y conjuntos son finitos. Por supuesto, conservaremos las notaciones establecidas en el Capítulo 1.

Así pues, si G es un grupo actuando en un conjunto X no vacío, es decir, existe φ homomorfismo de G en el grupo de permutaciones S_X, entonces φ induce una función ψ de $G \times X$ en X dada por $\psi(g, x) := \varphi_g(x) = gx$ tal que ψ satisface las siguientes propiedades:

(1) $ex = x \ \ \forall\, x \in X$ y;

(2) $g(hx) = (gh)x \ \ \forall\, g, h \in G$ y $x \in X$.

Recíprocamente, toda función ψ de $G \times X$ en X, con $\psi(g, x) = gx$ para cada $g \in G$ y $x \in X$ satisfaciendo las propiedades (1) y (2), induce una acción de G en X al definir el homomorfismo φ de G en S_X como $\varphi_g(x) = gx$.

Si G actúa en X, entonces en algunas ocasiones se dice que X es un **G-conjunto**. Por esta razón, indistintamente mencionaremos la acción a través de φ o de ψ.

Notemos que si X es un G-conjunto y H es un subgrupo de G, entonces X es también un H-conjunto, por restricción de la acción de $G \times X$ a $H \times X$.

Dado X no vacío, denotemos el conjunto de todas las funciones de X sobre sí mismo por X^X. Uno observa fácilmente que X^X bajo la composición de funciones es un semigrupo cuya identidad es la función identidad id_X sobre X; además, es claro que $S_X \subseteq X^X$. Si G es un grupo y $\theta : G \to X^X$ es un homomorfismo (de semigrupos), con $\theta(e) = id_X$, entonces $\operatorname{Im}\theta \subseteq S_X$ [pues para cada $g \in G$, $1_X = \theta(1) = \theta(gg^{-1}) = \theta(g)\theta(g^{-1})$ y, similarmente, $1_X = \theta(g^{-1})\theta(g)$, esto es $\theta(g)$ es una biyección]. Se sigue que tal homomorfismo θ hace de X un G-conjunto.

Definición 3.1.1. *Un G-conjunto X es **fiel** si su acción $\varphi : G \to S_X$ es inyectiva.*

Si X es un G-conjunto fiel, podemos identificar a G con un subgrupo de S_X vía φ, y G llegando a ser un grupo de permutaciones sobre X. Precisamente, el Teorema de Caley afirma que cada grupo G de orden n es un G-conjunto fiel de grado n, esto es, G está encajado en el grupo de permutaciones S_G de grado n.

Generalizaremos el concepto de subgrupo estabilizador de G para un número finito de elementos de X.

Definición 3.1.2. *Si X es un G-conjunto y $x_1, \ldots, x_k \in X$, entonces el **estabilizador** $G_{x_1,\ldots,x_k}$ de los elementos $x_1, \ldots, x_k \in X$ es el subgrupo*

$$G_{x_1,\cdots,x_k} = \{g \in G \mid gx_i = x_i, \;\; i = 1, \ldots, k\}.$$

Así $G_{x_1,\cdots,x_k}$ es el conjunto de todos los elementos $g \in G$ fijando a cada x_i, $i = 1, \ldots, k$. Podemos también observar que

$$G_{x_1,\cdots,x_k} = \bigcap_{i=1}^{k} G_{x_i}.$$

Por otro lado, si X es un G-conjunto y $x, y \in X$, entonces X es también un G_x-conjunto y un G_y-conjunto, y

$$(G_x)_y = G_{x,y} = (G_y)_x.$$

Ya hemos visto que si X es un G-conjunto y $x \in X$, el índice $[G : G_x]$ es el número de elementos de la órbita O_x (Teorema 1.1.2).

Ahora, examinaremos la estructura de los G-conjuntos.

Definición 3.1.3. *Un G-conjunto X es **transitivo** si para cada $x, y \in X$, existe $g \in G$, con $gx = y$.*

Claramente, un G-conjunto X es transitivo si y sólo si tiene una sola órbita.

Teorema 3.1.1. *Cada G-conjunto X es particionado en sus G-órbitas, cada una de las cuales es un G-conjunto transitivo. Inversamente, si un G-conjunto X está particionado en G-conjuntos transitivos $\{X_i \mid i \in I\}$, entonces los X_i son las G-órbitas de X.*

Demostración: Cada G-órbita Gx es un G-conjunto transitivo. Definiendo la relación $x \sim y$ (para $x, y \in X$) la cual significa que para algún $g \in G$, $\ gx = y$, es fácil ver que $\sim$ es una relación de equivalencia en X cuyas clases de equivalencia son precisamente las G-órbitas.

Para probar la recíproca, es suficiente con probar que para cada $i = 1, \ldots, n$, $X_i = Gx_i$, donde $x_i \in X_i$. Ahora, $Gx_i \subseteq X_i$ puesto que X_i es un G-conjunto. Para la inclusión inversa, si $y \in X_i$, entonces transitivamente tenemos $y = gx_i$ para algún $g \in G$. Por lo tanto $y \in Gx_i$ y $X \subseteq Gx_i$. □

Esta descomposición única permite que uno se centre en los G-conjuntos transitivos.

Teorema 3.1.2. *Si X es un G-conjunto transitivo de grado n, es decir, $card(X) = n$ y $x \in X$, entonces*

$$|G| = n|G_x|.$$

Si X es un G-conjunto fiel, entonces $|G_x|$ es un divisor de $(n-1)!$.

Demostración: Sabemos que $|G| = [G : G_x]|G_x|$. Pero el índice $[G : G_x]$ es el cardinal de la órbita Gx, y $Gx = X$ ya que G actúa transitivamente sobre X. La segunda afirmación se sigue de que G es isomorfo a un subgrupo de S_n. □

Teorema 3.1.3. *Sea X un G-conjunto transitivo y sean $x, y \in X$.*

1. *Si $tx = y$, para algún $t \in G$, entonces $G_y = tG_xt^{-1}$.*
2. *X tiene el mismo número de G_x-órbitas en cuanto a G_y-órbitas.*

Demostración:

1. Si g fija a x, entonces $tgt^{-1}y = tgx = tx = y$ y tgt^{-1} fija a y. El resultado se sigue inmediatamente.

2. Ya que G actúa transitivamente, tenemos que existe $t \in G$ tal que $tx = y$. Si las G_x-órbitas están dadas por la familia $\{G_x a_i \mid i \in I\}$, donde $a_i \in X$ para cada i, definimos $b_i = ta_i \in X$. Claramente el conjunto $G_y b_i$, $i \in I$, determinan las G_y-órbitas de X, y cada uno de estos G_y-conjuntos son transitivos. Además,

$$G_y b_i = tG_x t^{-1} b_i = tG_x a_i.$$

Dado que t actúa como una permutación de X, lleva una partición de X en otra partición; por lo tanto, los subconjuntos $G_y b_i$ particionan a X, y por el Teorema 3.1.1 se tiene la demostración.

□

Definición 3.1.4. *Si X es un G-conjunto transitivo, entonces el **rango** de X es el número de G_x-órbitas de X.*

Notemos que el Teorema 3.1.3 establece que la definición de rango es independiente de la elección del estabilizador G_x.

Teorema 3.1.4. *Sea X un G-conjunto transitivo y sea $x \in X$. Entonces, el rango de X es el número de G_x-G_x clases laterales dobles en G.*

Demostración: Definamos la función

$$\begin{aligned} f : \{G_x - \text{órbitas}\} &\to \{(G_x - G_x) - \text{clases laterales dobles}\} \\ G_x y &\mapsto G_x g G_x \end{aligned}$$

donde $gx = y$.

f está bien definida:
Si $hx = y$, entonces $gx = hx$, $g^{-1}h \in G_x$, $h = eg(g^{-1}h) \in G_x g G_x$, y $G_x g G_x = G_x h G_x$.

f es inyectiva:
Si $f(G_x y) = G_x g G_x = G_x h G_x = f(G_x z)$, donde $gx = y$ y $hx = z$, entonces existen $a, b \in G_x$ tales que $g = ahb$. Ahora, $y = gx = ahbx = ahx = az \in G_x z$; se sigue que $G_x y = G_x z$.

f es suprayectiva:
Si $g \in G$, entonces $gx = y$ y $f(G_x y) = G_x g G_x$. □

Obsérvese que X es transitivo y $card(X) \geq 2$, entonces el rango de X es ≥ 2, ya que en caso contrario $G = G_x G_x$ y cada $g \in G$ fija a x, lo cual contradice la transitividad. Si consideramos el caso en el que el rango de X es 2, entonces las G_x-órbitas son $\{x\}$ y $X \setminus \{x\}$, con lo cual G_x actúa transitivamente sobre $X \setminus \{x\}$.

Definición 3.1.5. *Sea X un G-conjunto de grado n y sea $k \leq n$ un entero positivo. Decimos que que X es k-**transitivo** si, para cada pareja de k-uplas $(x_1, \ldots, x_k)$ y $(y_1, \ldots, y_k)$ teniendo distintas entradas en X, existe $g \in G$ con $gx_i = y_i$, para $i = 1, \ldots, k$.*

Por supuesto, 1-transitivo es ordinariamente transitivo, y si $k > 1$, entonces k-transitivo implica $(k-1)$-transitivo. Un G-conjunto X el cual es k-transitivo es llamado **doblemente transitivo** si $k = 2$, **triplemente transitivo** si $k = 3, \ldots$, y **múltiplemente transitivo** si $k > 1$.

Teorema 3.1.5. *Cada G-conjunto X múltiplemente transitivo tiene rango 2, y si $x \in X$ y $g \notin G_x$, entonces $G = G_x \cup G_x g G_x$.*

Demostración: Dado que G actúa k-transitivamente sobre X para algún $k > 1$, G_x actúa transitivamente sobre $X \setminus \{x\}$, así que $X \setminus \{x\}$ tiene justamente una G_x-órbita. Entonces, X tiene rango 2, y el Teorema 3.1.4 prueba que existen solo dos $G_x - G_x$ clases laterales dobles. □

Observación 3.1.1. *Sean X un G-conjunto transitivo y $k \geq 2$. Entonces, X es k-transitivo si y sólo si para cada $x \in X$, el G_x-conjunto $X \setminus \{x\}$ es $(k-1)$-transitivo.*

Teorema 3.1.6. *Si X es un G-conjunto k-transitivo de grado n, entonces*

$$|G| = n(n-1)\cdots(n-k+1)|G_{x_1,\ldots,x_k}|$$

para cada elección de k elementos distintos $x_1, \ldots, x_k$ en X. Si G actúa fielmente, entonces $|G_{x_1,\ldots,x_k}|$ es un divisor de $(n-k)!$.

Demostración: Por el Teorema 3.1.1, si $x_1 \in X$, entonces $|G| = n|G_{x_1}|$. Pero, por la Observación 3.1.1, G_{x_1} actúa $(k-1)$-transitivamente sobre $X \setminus \{x_1\}$. Por lo tanto, si $x_2, \ldots, x_k$ son elementos distintos de $X \setminus \{x_1\}$, inductivamente se tiene que

$$|G_{x_1}| = (n-1)\cdots(n-k+1)|G_{x_1,\ldots,x_k}|.$$

Si G actúa fielmente, entonces $|G_{x_1,\ldots,x_k}|$ es un divisor de $(n-k)!$, ya que G está encajado en $S_{X\setminus\{x_1,\ldots,x_k\}}$. □

Definición 3.1.6. *Sea X un G-conjunto k-transitivo. Decimos que X es* ***simplemente k-transitivo*** *si la identidad de G es el único elemento que fija a k elementos distintos de X.*

Corolario 3.1.1. *Sea X un G-conjunto fielmente k-transitivo de grado n. Entonces, los siguientes enunciados son equivalentes:*

(*i*) *X es simplemente k-transitivo;*

(*ii*) *Si $(x_1, \ldots, x_k)$ y $(y_1, \ldots, y_k)$ son k-tuplas con entradas distintas en X, entonces existe un único $g \in G$ tal que $gx_i = y_i$, para $i = 1, \ldots, k$;*

(*iii*) $|G| = n(n-1)\cdots(n-k+1)$;

(*iv*) *Cada estabilizador de k elementos distintos en X es $\{e\}$.*

Si $k \geq 2$, los enunciados anteriores son equivalentes a

(*v*) *Para cada $x \in X$, los G_x-conjuntos $X \setminus \{x\}$ son simplemente $(k-1)$-transitivos.*

Demostración: La demostración se realiza de manera directa usando la Definición 3.1.6 y el Teorema 3.1.6. □

Teorema 3.1.7. *Para cada n, el grupo simétrico S_n de grado n actúa simplemente n-transitivamente sobre el conjunto $X = \{1, \ldots, n\}$; para cada $n \geq 3$, el grupo alternante A_n de grado n actúa simplemente $(n-2)$-transitivamente sobre X.*

Demostración: La primera afirmación es obvia dado que S_n contiene cada permutación de X.
Probaremos por inducción sobre $n \geq 3$ que A_n actúa simplemente $(n-2)$-transitivamente sobre X. Cuando $n = 3$, entonces $A_3 = \langle (1\ 2\ 3) \rangle$ actúa transitivamente sobre $X = \{1, 2, 3\}$. Si $n > 3$, notemos que el estabilizador de cada i, con $1 \leq i \leq n$, $(A_n)_i$, es isomorfo a A_{n-1}, el cual actúa simplemente $(n-3)$-transitivamente sobre el conjunto $X \setminus \{i\}$, por inducción. Por el Teorema 3.1.1, se tiene completo el proceso inductivo y la demostración del teorema. □

Definición 3.1.7. *Todo G-conjunto simplemente 1-transitivo es llamado* ***regular****.*

Si X es un G conjunto fiel, entonces X es regular si y sólo si X es transitivo y únicamente la identidad de G fija a cada elemento de X. De aquí se sigue que la acción de Cayley de un grupo G hace de G mismo un G-conjunto regular.

Nuestra discusión de los G-conjuntos simplemente 2-transitivos comienza con la siguiente definición.

Definición 3.1.8. *Si X es un G-conjunto, entonces el* ***núcleo de Frobenius*** *N de G es el subconjunto*

$$N = \{e\} \cup \{g \in G \mid g \text{ no tiene puntos fijos }\}.$$

En general, el núcleo de Frobenius no necesariamente es un subgrupo de G, ya que puede no ser cerrado bajo la multiplicación.

Lema 3.1.1. *Si X es un G-conjunto fiel simplemente 2-transitivo de grado n, entonces el núcleo de Frobenius N de G tiene exactamente n elementos.*

Demostración: Por el Teorema 3.1.1(iii), tenemos que $|G| = n(n-1)$. Para cada $x \in X$, el estabilizador G_x tiene orden $n-1$, ya que $|G| = n|G_x|$, en consecuencia $|G_x^*| = n-2$, donde $G_x^* = G_x \setminus \{e\}$. Si $x \neq y$, entonces $G_x \cap G_y = G_{x,y} = \{e\}$, por el Teorema 3.1.1(iv). De aquí que, $\{G_x^* \mid x \in X\}$ es una familia disjunta, y $|\cup_{x\in X} G_x^*| = n(n-2)$. Pero N es justamente el complemento de esta unión, así que $|N| = n(n-1) - n(n-2) = n$. □

Ahora clasificaremos los grupos G teniendo un G-conjunto X como sucede en el Lema 3.1.1 para el caso especial cuando el grado $n = |X|$ es impar (la clasificación cuando n es par es también conocida, pero es más dificil). La idea es de que el conocimiento de las representaciones de la permutación de un grupo G nos puedan dar cierta información importante sobre G.

Teorema 3.1.8. *Sea X un G-conjunto fiel simplemente 2-transitivo de grado n impar.*

(*i*) *Cada G_x contiene un único elemento de orden 2.*

(*ii*) *El núcleo de Frobenius N de G es un subgrupo normal de G.*

(*iii*) *El grado n es una potencia de un primo impar p.*

(*iv*) *G es un producto semidirecto de N por G_x.*

Demostración:

(*i*) Dado que $|G| = n(n-1)$ es par, G contiene un elemento g de orden 2. Ahora g, siendo una permutación de los n elementos de X, tiene una descomposición cíclica, exactamente como un 2-ciclo ó como producto de 2-ciclos disjuntos. Digamos $g = \tau_1 \cdots \tau_m$. Puesto que $|X| = n$ es impar, g debe fijar algún $x \in X$, es decir, $g \in G_x$; además, dado que X es simplemente 2-transitivo, g nada

más fija a x, es decir, g mueve a todos los elementos de X salvo a x, de donde $m = \frac{1}{2}(n-1)$. Si h es algún otro elemento de orden 2 en G, entonces $h \in G_y$ para algún $y \in X$ (quizás $y = x$) y h es también un producto de transposiciones disjuntas: $h = \tau\prime_1 \cdots \tau\prime_m$. Note que g y h puede que no tengan factores en común: si $\tau_i = \tau\prime_j = (a\ b)$, entonces podemos asumir $\tau_i = \tau_m$ y $\tau\prime_j = \tau\prime_m$ (pues los ciclos son disjuntos y conmutan); entonces, gh^{-1} fija a a y a b, de aquí que gh^{-1} es la identidad, y $g = h$.

Suponga que G_x tiene t elementos de orden 2. Como X es simplemente 2-transitivo, $\{G_x^* \mid x \in X\}$ es una familia subconjuntos disjuntos, así que G contiene nt de tales elementos, de los cuales envuelve $m = \frac{1}{2}(n-1)$ transposiciones; luego, existen $\frac{1}{2}nt(n-1)$ transposiciones distintas que son factores de estos elementos de orden 2. Como $S_X \cong S_n$ contiene sólo $\frac{1}{2}n(n-1)$ transposiciones, tenemos necesariamente que $t = 1$.

(*ii*) Si T es el conjunto de todos los elementos de G de orden 2, entonces $TT \subset N$, pues en caso contrario, existen $g, h \in T$ tales que $gh \neq e$ y gh fija a algún elemento $y \in X$. Definimos $z = hy$. Ambos g y h tienen la transposición $(y\ z)$ como factor: $hy = z$ y $hz = h^2y = y$; $gy = gghy = hy = z$ y $gz = ggy = y$. Por el inciso (i), tendríamos que $g = h$, con lo cual $gh = e$, contradicción.

Para $g \in T$ fijo, las funciones $T \to N$ definidas por $h \mapsto gh$ y $h \mapsto hg$ son inyectivas. Por el inciso (i), sabemos que $|T| = n$; por el Lema 3.1.1, sabemos que $|N| = n$. Así que las dos funciones deben ser suprayectivas, es decir, $gT = N = Tg$ para todo $g \in T$.

El núcleo de Frobenius N contiene a e y es cerrado bajo inversos y conjugación por elementos en G. Notemos que en este caso, N es cerrado bajo multiplicación: eligiendo $g \in T$ y observando que

$$NN = (Tg)(gT) = Tg^2T = TT \subset N.$$

Por lo tanto, N es subgrupo normal de G.

(*iii*) Elijamos un elemento $h \in N$ de orden primo p (p es impar puesto que divide a $|N| = n$). Suponga que $f \in G$ conmuta con h, es decir, $hfh^{-1} = f$. Si $f \in G_x$ para algún $x \in X$, entonces

$$f \in G_x \cap hG_xh^{-1} = G_x \cap G_{hx} = G_{x,hx} = \{e\},$$

ya que $hx \neq x$. Así, si $f \neq e$, tenemos que $f \notin \bigcup_{x \in X} G_x$, luego $f \in N$. Concluimos que $C_G(h)$, el centralizador de h en G, está contenido en N, lo cual implica que $[G : C_G(h)] = [G : N][N : C_G(h)] \geq n - 1$. Pero $[G : C_G(h)]$ es el número de conjugados de h, y estos conjugados pertenecen al subgrupo

normal N de orden n. Siguiéndose que $N \setminus \{e\}$ consiste precisamente de los conjugados de h, de donde N es un p-grupo de exponente p. Por tanto, $n = |N|$ es una potencia de p.

(*iv*) El orden de N y G_x son enteros consecutivos, por tanto sus órdenes son primos relativos. Puesto que N es normal en G, $N \cap G_x = \{e\}$ y $NG_x = G$, concluimos que G es el producto semidirecto de N por G_x.

□

Definición 3.1.9. *Un* ***bloque*** *de un G-conjunto X es un subconjunto B de X tal que para cada $g \in G$, se tiene que $gB = B$ o $gB \cap B = \emptyset$ (por supuesto $gB = \{gb \mid b \in B\}$).*

Claramente $B = \emptyset$ y $B = X$ son bloques, al igual que cada subconjunto de un solo elemento de X; estos bloques son llamados **triviales**, y cualquier otro bloque es llamado **no trivial**.

Definición 3.1.10. *Un G-conjunto transitivo X es* ***primitivo*** *si no contiene bloques no triviales; en caso contrario, decimos que es* ***imprimitivo****.*

Teorema 3.1.9. *Sea B un bloque no trivial en un G-conjunto X transitivo de grado n. Entonces,*

(*i*) *Si $g \in G$, entonces gB es un bloque.*

(*ii*) *Existen elementos $g_1, g_2, \ldots, g_m \in G$ tales que $Y = \{g_1B, g_2B, \ldots, g_mB\}$ es una partición de X.*

(*iii*) *$|B|$ divide $|X|$.*

(*iv*) *Y es un G-conjunto transitivo de grado $n/|B|$.*

Demostración:

(*i*) Suponga que $gB \cap hgB \neq \emptyset$ para algún $h \in G$. Entonces, $B \cap g^{-1}hgB \neq \emptyset$, de donde $g^{-1}hgB = B$ y $hgB = gB$.

(*ii*) Definimos $g_1 = e$ y eligimos $b \in B$ y $x_1 \notin B$. Dado que G actúa transitivamente sobre X, existe $g_2 \in G$ tal que $g_2b = x_1$. Así que, $B \neq g_2B$. Puesto que B es un bloque de X, tenemos que B y g_2B son disjuntos. Si $X = B \cup g_2B$, entonces habremos terminado. En otro caso, eligimos $x_2 \notin B \cup g_2B$ y $g_3 \in G$ tales que $g_3b = x_2$. Puesto que B y g_2B son bloques, tenemos que g_3B es un bloque disjunto a B y a g_2B. La prueba finaliza por iteración de este procedimiento.

(*iii*) Si $g \in G$, entonces $|B| = |gB|$ y el resultado ahora se sigue de (*ii*).

(*iv*) Sea $g \in G$. Tenemos que para cada $i \in \{1, \ldots, m\}$, existe $j \in \{1, \ldots, m\}$ tal que $g_jB \cap gg_iB \neq \emptyset$, donde g_jB es bloque y $gg_iB = (gg_ig_j^{-1})g_jB$; de aquí que, $g_jB = gg_iB$. Además, si $gg_iB = gg_jB$ para $i, j \in \{1, \ldots, m\}$, entonces $g_iB = g_jB$. Así, la función de Y en sí mismo dada por $g_iB \longmapsto gg_iB$ es biyección. Puesto que $g(hg_iB) = ghg_iB$ para cada $i \in \{1, \ldots, m\}$, tenemos que la función $\varphi : G \longrightarrow S_Y$ dada por $\varphi_g(g_iB) = gg_iB$ para cada $i \in \{1, \ldots, m\}$ es un homomorfismo y, en consecuencia, Y es un G conjunto de grado $|Y| = n/|B|$, ya que $|B| = |g_iB|$ para cada $i \in \{1, \ldots, m\}$. Finamente, Y es un G-conjunto transitivo puesto que $(g_jg_i^{-1})g_iB = g_jB$ para cada $i, j \in \{1, \ldots, m\}$.

□

Definición 3.1.11. *El conjunto Y en el Teorema* 3.1.9 (*ii*) *es llamado un* ***sistema imprimitivo*** (*generado por B*).

Corolario 3.1.2. *Un G-conjunto transitivo de grado primo es primitivo.*

Demostración: Se sigue inmediatamente del Teorema 3.1.9 (*iv*). □

Observación 3.1.2. Se tiene lo siguiente:

(*i*) Suponga que X es un G-conjunto transitivo imprimitivo y B es un bloque no trivial maximal, es decir, B no está contenido propiamente en un bloque más grande. Si Y es el sistema imprimitivo generado por B, entonces Y es un G-conjunto primitivo.

(*ii*) Sea X un G-conjunto con $x, y \in X$. Si H es un subgrupo de G, entonces $Hx \cap Hy \neq \emptyset$ implica que $Hx = Hy$. Si $H \lhd G$ (así que $gH = Hg$, para cada $g \in G$), entonces los conjuntos Hx son bloques de X.

Teorema 3.1.10. *Cada G-conjunto X k-transitivo, con $k \geq 2$, es primitivo.*

Demostración: Suponga que X tiene un bloque B no trivial. Sean $x, y \in B$ distintos y $z \notin B$. Entonces, existe $g \in G$ tal que $gx = x$ y $gy = z$, y así que B y gB son distintos y se intersectan, lo cual es una contradicción. □

Damos una caracterización de los G-conjuntos primitivos.

Teorema 3.1.11. *Sea X un G-conjunto transitivo. Entonces, X es primitivo si y sólo si, para cada $x \in X$, el estabilizador G_x es un subgrupo maximal de G.*

Demostración:[1] $\Rightarrow$) Suponga que X es primitivo y que existe un subgrupo de H tal que $G_x \subsetneq H \subsetneq G$. Definamos $B = Hx$. Tenemos que B es un bloque; en efecto, si $g \in G$ y $B \cap gB \neq \emptyset$, entonces $hx = gh\prime x$ para algunos $h, h\prime \in H$, así que $h^{-1}gh\prime \in G_x \subset H$ y $g \in H$, luego $gB = gHx = Hx = B$. Veamos que B es un bloque no trivial para llegar a una contradicción. Claramente $B = Hx \neq \emptyset$. Si $Hx = X$, entonces para $g \notin H$ se tiene que $gx = hx$ para algún $h \in H$, de donde $h^{-1}g \in G_x \subset H$ y $g \in H$, lo cual es absurdo. Finalmente, Hx no es un conjunto de un sólo punto, pues $G_x \subsetneq H$.

$\Leftarrow$) Suponga que cada G_x es un subgrupo maximal, pero existe un bloque B no trivial en X. Definamos el subgrupo H de G como sigue:

$$H = \{g \in G \mid gB = B\}$$

Sea $x \in B$. Como B es un bloque, $G_x \subset H$. Ya que B es no trivial, existe $y \in B$ tal que $y \neq x$. Sea $g \in G$ tal que $gx = y$. Tenemos que $g \notin G_x$ pero $g \in H$ (pues $B \cap gB \neq \emptyset$); luego, $G_x \subsetneq H$. Finalmente, $H \neq G$, ya que en caso contrario $B = X$. Por lo tanto, no existe tal bloque B, y X es primitivo. □

Observación 3.1.3. Suponga X es un G-conjunto transitivo y $H \lhd G$. En general, X no necesariamente es H-conjunto transitivo. Por ejemplo, si V es un espacio vectorial de dimensión finita sobre un campo K, entonces $V^* = V \setminus \{0\}$ es un $GL(V)$-conjunto transitivo; si H es el centro de $GL(V)$, es decir, el subgrupo de todas las tranformaciones escalares, entonces V^* es un H-conjunto no transitivo.

Teorema 3.1.12. *Sea X un G-conjunto fiel y primitivo de grado $n \geq 2$ y sea $H \lhd G$ tal que $H \neq \{e\}$. Entonces, X es un H-conjunto transitivo.*

Demostración: Hemos notado en la Observación 3.1.2 que el conjunto Hx es un bloque, para cada $x \in X$. Puesto que G actúa primitivamente sobre X, tenemos que para cada $x \in X$ o bien $Hx = \emptyset$ (lo cual es imposible) o $Hx = X$ o $Hx = \{x\}$. Si $Hx = \{x\}$ para algún $x \in X$, entonces $H \subseteq G_x$. Por normalidad de H, $H = gHg^{-1} \subseteq gG_xg^{-1} = G_{gx}$, para cada $g \in G$; puesto que X es G-conjunto transitivo fiel, tenemos que $H \subseteq \cap_{y \in X} G_y = \{e\}$, lo cual contradice el hecho de que $H \neq \{e\}$. Por lo tanto, $Hx = X$ para cada $x \in X$ y X es un H-conjunto transitivo. □

Usando el teorema anterior, tenemos que el $GL(V)$-conjunto V^* tratado en la Observación 3.1.3 es transitivo pero no primitivo.

[1] Daremos una prueba distinta en el Teorema 3.1.15

Definición 3.1.12. *Un subgrupo normal H de G para el cual un G-conjunto X es un H-conjunto regular es llamado un* ***subgrupo normal regular*** *de G.*

Observemos que todos los subgrupos normales regulares tienen el mismo orden $|X|$.

Teorema 3.1.13. *Suponga X es un G-conjunto fiel primitivo con estabilizador G_x simple. Entonces, o bien G es simple o cada subgrupo normal no trivial H de G es un subgrupo normal regular.*

Demostración: Suponga que $H \neq \{e\}$ es un subgrupo normal de G. Por el Teorema 3.1.12, X es H-transitivo. Por tanto, cada H actúa regularmente o $H \cap G_x \neq \{e\}$ para algún $x \in X$. Pero $H \cap G_x \lhd G_x$, así la simplicidad de G_x da $H \cap G_x = G_x$, es decir, $G_x \subset H$. Por el Teorema 3.1.11, G_x es maximal de G, así que $H = G_x$ o $H = G$, donde el primer caso no puede ocurrir porque G actúa transitivamente. Luego, G es simple. □

Definición 3.1.13. *Si X y Y son G-conjuntos, una función $\theta : X \to Y$ es una* **G-*función*** *si*

$$\theta(gx) = g\theta(x) \ \ \forall\, g \in G, \ \ y \, \forall\, x \in X.$$

Si θ es también una función biyectiva, entonces θ es un ***G-isomorfismo****. Dos G-conjuntos X y Y son* ***isomorfos*** *(**G-isomorfos**), denotado por $X \cong Y$, si existe un G-isomorfismo de X en Y.*

Usualmente, cuando estamos trabajando con un G-conjunto X no hacemos mención de la acción φ que sobre X actúa G. Puesto que ahora estamos relacionando dos G-conjuntos X y Y, en ocaciones haremos la distinción de las acciones entre X y Y, es decir, haremos mención de las parejas (X, φ) y (Y, ψ) para establecer que X y Y son G conjuntos. En consecuencia, una G-función es una función $\theta : (X, \varphi) \longrightarrow (Y, \psi)$ tal que para cada $g \in G$ y para cada $x \in X$

$$\theta(\varphi_g(x)) = \psi_g(\theta(x)).$$

Teorema 3.1.14. *Cada G-conjunto X transitivo es isomorfo al conjunto $Y = \{gG_x \mid g \in G\}$ de clases laterales izquierdas de G_x en G, en donde G actúa por multiplicación izquierda sobre los elementos de Y.*

Demostración: Sean $X = \{x_1, \ldots, x_n\}$, $H = G_{x_1}$ y, para cada $i \in \{1, \ldots, n\}$, sea $g_i \in G$ tal que $g_i x_1 = x_i$ (el cual existe porque G actúa transitivamente sobre X). El argumento rutinario demuestra que la función $\theta : X \to Y$, donde Y es el conjunto de clases laterales izquierdas de H en G actuando G por multiplicación izquierda, definida por $\theta(x_i) = g_i H$, para cada i, está bien definida y es una biyección (recordemos que $n = |Gx_1| = [G : H]$). Veamos que θ es un G-isomorfismo. Si $g \in G$, entonces para cada i existe un j tal que $gx_i = x_j$, y así que

$$\theta(gx_i) = \theta(x_j) = g_j H.$$

Por otro lado,

$$g\theta(x_i) = gg_i H.$$

Pero $gg_i x_1 = gx_i = x_j = g_j x_1$, implica $g_j^{-1} g g_i \in G_{x_1} = H$, entonces $g_j H = gg_i H$. □

Teorema 3.1.15.

(*i*) *Si H y K son subgrupos de un grupo G, entonces Y_H y Y_K, donde Y_H y Y_K son el conjunto de clases laterales izquierdas de H en G y de K en G, respectivamente, actuando G (en ambos) por multiplicación izquierda, son G-isomorfos si y sólo si H y K son subgrupos conjugados.*

(*ii*) *Dos G-conjuntos transitivos (X, φ) y (Y, ψ) son G-isomorfos si y sólo si sus estabilizadores en cada punto son subgrupos conjugados de G.*

Demostración:

(*i*) $\Rightarrow$) Suponga que $\theta : Y_H \to Y_K$ es un G-isomorfismo. Entonces, en particular, existe $g \in G$ tal que $\theta(H) = gK$. Si $h \in H$, entonces

$$gK = \theta(H) = \theta(hH) = h\theta(H) = hgK.$$

Por tanto, $g^{-1}hg \in K$ y $g^{-1}Hg \subset K$. Dado que $\theta(g^{-1}H) = g^{-1}\theta(H) = g^{-1}gK = K$, vemos que $\theta^{-1}(K) = g^{-1}H$. La discusión de arriba ahora da que $gKg^{-1} \subset H$, de donde $g^{-1}Hg = K$.

$\Leftarrow$) Para la recíproca, elijamos $g \in G$ con $g^{-1}Hg = K$. Observe que las siguientes relaciones son equivalencias: para $a, b \in G$, $aH = bH$; $a^{-1}b \in H$; $g^{-1}a^{-1}bg \in g^{-1}Hg = K$; $agK = bgK$. Concluimos que la función $\theta : Y_H \to Y_K$ dada por $\theta(aH) = agK$ está bien definida y es inyectiva. Claramente θ es sobre, pues para $b \in G$ se tiene que $bK = \theta(bg^{-1}H)$. Finalmente, θ es una G-función, ya que $\theta(abH) = (ab)gK$ y $a\theta(bH) = a(bgK)$.

(*ii*) Sean H y K estabilizadores de elementos en (X, φ) y (Y, ψ), respectivamente. Por Teorema 3.1.14, (X, φ) es G-isomorfo a Y_H y (Y, ψ) es G-isomorfo a Y_K. El resultado ahora se sigue fácilmente de la parte (i).

□

Es ahora fácil exhibir G-conjuntos transitivos no isomorfos del mismo grado; basta elegir dos subgrupos no conjugados de G teniendo el mismo índice.

Lema 3.1.2. *Sean X un G-conjunto transitivo y H un subgrupo normal regular de G. Sea $x \in X$ y sea G_x actuando sobre $H^* = H \setminus \{e\}$ por conjugación. Entonces los G_x-conjuntos H^* y $X \setminus \{x\}$ son isomorfos.*

Demostración: Definimos $\theta : H^* \to X \setminus \{x\}$ dada por $\theta(h) = hx$. Si $\theta(h) = \theta(k)$, entonces $h^{-1}k \in H_x = \{e\}$ (ya que H es subgrupo normal regular), luego θ es inyectiva. También, por la regularidad, se tiene $|H| = |X|$ y, en consecuencia, $|H^*| = |X \setminus \{x\}|$, por lo tanto θ debe ser sobre. Finalmente, probaremos que θ es una G-función. Si $g \in G_x$ y $h \in H^*$, denotamos la acción de g sobre h por $g * h$: así, $g * h = ghg^{-1}$. Por tanto,

$$\theta(g * h) = \theta(ghg^{-1}) = ghg^{-1}x = ghx$$

porque $g^{-1} \in G_x$; por otro lado, $g\theta(h) = g(hx)$. y así que $\theta(g * h) = g\theta(h)$. □

Teorema 3.1.16. *Si X es un G-conjunto k-transitivo ($k \geq 2$) de grado n y si G tiene un subgrupo normal regular H, entonces $k \leq 4$. Además,*

(*i*) *Si $k \geq 2$, entonces H es un p-grupo abeliano elemental para algún primo p y $n = p^m$;*

(*ii*) *Si $k \geq 3$, entonces o bien $H \cong \mathbb{Z}_3$ y $n = 3$ o H es un 2-grupo abeliano elemental y $n = 2^m$;*

(*iii*) *Si $k \geq 4$, entonces $H \cong \mathbb{V}$ y $n = 4$, donde $\mathbb{V}$ es el 4-grupo.*

Demostración: Si X es un G-conjunto k-transitivo con $k \geq 2$, entonces para $x \in X$ fijo, tenemos que $X \setminus \{x\}$ es un G_x-conjunto $(k-1)$-transitivo. Por el Lema 3.1.2, H^* es un G_x-conjunto $(k-1)$-transitivo, donde G_x actúa sobre H^* por conjugación.

(*i*) Suponga que $k \geq 2$. Entonces, tenemos que H^* es un G_x conjunto transitivo. El estabilizador G_x actúa por conjugación, el cual es un automorfismo (interno), se sigue que todos los elementos de H^* tienen el mismo orden, el cual debe ser un número primo p. Así que H es un grupo de exponente p. Luego, $Z(H)$ es un subgrupo característico no trivial de G, con lo que $|X| = |Z(H)| = |H|$, pues $Z(H)$ y H son subgrupos normales regulares de G'. Por lo tanto, $Z(H) = H$, H es abeliano elemental y $|X|$ es una potencia de p.

(*ii*) Sea $k \geq 3$. Así que H^* es múltiplemente transitivo y, en consecuencia, un G_x-conjunto primitivo. Si $h \in H^*$, es fácil ver que $\{h, h^{-1}\}$ es un bloque. Por la primitividad, se tiene que o bien cualquier $\{h, h^{-1}\} = H^*$ o $\{h, h^{-1}\} = \{h\}$. Si $\{h, h^{-1}\} = H^*$, es decir, H^* tiene dos elementos, entonces $H \cong \mathbb{Z}_3$ y $n = 3$. Si $\{h, h^{-1}\} = \{h\}$, es decir, h tiene orden 2, entonces, como 3-transitivo implica 2-transitivo, el resultado del caso (i) implica que el número primo p debe de ser 2. Por lo tanto, en este caso, H es un 2-grupo abeliano elemental y $n = |X| = 2^m$.

(*iii*) Si $k \geq 4$. En este caso, $k - 1 \geq 3$ y $|H^*| \geq 3$ [el cual excluye los casos $H \cong \mathbb{Z}_3$ y $H \cong \mathbb{Z}_2$]. Se sigue que H contiene una copia de $\mathbb{V}$; digamos, $\{1, h, k, hk\}$. Ahora, $(G_x)_h$ actúa 2-transitivamente, y por lo tanto primitivamente, sobre $H^* \setminus \{h\}$. Pero uno puede ver fácilmente que $\{k, hk\}$ es un bloque, y así que $H^* \setminus \{h\} = \{k, hk\}$; concluyendo que $H = \{1, h, k, hk\} \cong \mathbb{V}$ y $n = |X| = 4$.

Finalmente, no podemos tener $k \geq 5$ para $n = 4 < k$. □

Por supuesto, el caso $k = 4$ actualmente ocurre cuando $G = S_4$ y $H = \mathbb{V}$. El caso $k = 2$ debe ser comparado con el Teorema 3.1.8.

Corolario 3.1.3. *Sea X un G-conjunto k-transitivo y fiel, con $k \geq 2$, y suponga que G_x es simple para algún $x \in X$. Entonces,*

(*i*) *Si $k \geq 4$, entonces G es simple.*

(*ii*) *Si $k \geq 3$ y $|X|$ no es una potencia de 2, entonces $G \cong S_3$ o G es simple.*

(*iii*) *Si $k \geq 2$ y $|X|$ no es una potencia de un primo, entonces G es simple.*

Demostración: Por el Teorema 3.1.13, o bien G es simple o G contiene un subgrupo H normal regular. Si H existe, el Teorema 3.1.16 implica que $k \leq 4$ y, si $k = 4$ entonces $H \cong \mathbb{V}$ y $|X| = 4$. Ahora, el único subgrupo 4-transitivo de S_4 es el mismo S_4, pero el estabilizador de un elemento es S_3 el cual es no simple. Por lo tanto, tal H no existe. Esto prueba (*i*). Las otras dos afirmaciones son también consecuencia inmediata del Teorema 3.1.16 [note que el estabilizador de un elemento de un S_3-conjunto es el grupo simple $S_2 \cong \mathbb{Z}_2$ así que S_3 es una excepción natural en la parte (*ii*)]. □

Por supuesto, la suposición de que el estabilizador de un elemento es simple no se tiene en general; sin embargo veremos que el Corolario 3.1.3 es útil.

Teorema 3.1.17. *A_n es simple para $n \geq 5$.*

Demostración: Procederemos por inducción sobre $n \geq 5$. Si $n = 5$, entonces sabemos que A_5 es simple. Suponga $n \geq 6$. Por el Teorema 3.1.7, A_n actúa k-transitivamente sobre $X = \{1, 2, \ldots, n\}$, donde $k = n - 2 \geq 4$. Ahora, $(A_n)_n$, el estabilizador de A_n en n, es justamente A_{n-1} (porque consiste de todas las permutaciones pares de $\{1, 2, \ldots, n-1\}$), y entonces es simple, por inducción. Por tanto, A_n es simple, por el Corolario 3.1.3 (i). □

Daremos otro criterio para la simplicidad, precedido por el siguiente lema. Más adelante lo utilizaremos para dar una prueba de la simplicidad de los PSL.

Lema 3.1.3. *Sea X un G-conjunto y sea N un subgrupo de G actuando transitivamente sobre X. Para cada $x \in X$, tenemos que $G = NG_x$.*

Demostración: Si $g \in G$, la transitividad de N nos da la existencia de $n \in N$ tal que $nx = gx$. Entonces, $n^{-1}g = h \in G_x$ y $g = nh \in NG_x$. □

Teorema 3.1.18 (Iwasawa, 1941). *Sea X un G-conjunto. Suponga que existe $x \in X$ y un subgrupo H de G_x tal que*

(i) X es un G-conjunto fiel primitivo;

(ii) H es un subgrupo normal abeliano de G_x;

(iii) Los conjugados $\{gHg^{-1} \mid g \in G\}$ generan a G;

*(iv) G es **perfecto**, es decir, $G = G\prime$.*

Entonces G es un grupo simple.

Demostración: Sea $N \neq \{e\}$ un subgrupo normal de G. Si $g \in G$, la condición (iii) nos dice que $g = \prod g_i h_i g_i^{-1}$, donde $g_i \in G$ y $h_i \in H$, con producto finito. Ahora, N actúa transitivamente sobre X, por el Teorema 3.1.12; de modo que, el Lema 3.1.3 se puede aplicar a cada g_i, y entonces $g_i = n_i s_i$ para $n_i \in N$ y $s_i \in G_x$. Por lo tanto, la normalidad de H en G_x nos da que

$$g = \prod n_i s_i h_i s_i^{-1} n_i^{-1} \in NHN \subset HN,$$

y $G = HN$. Ya que H es abeliano, $G/N = HN/N \cong H/(H \cap N)$ es abeliano. Por lo tanto $N \supset G\prime = G$, por la condición (iv).[2] □

[2]Uno necesita suponer solamente que H es soluble, porque la prueba mostraría que G/H es soluble; esto junto con que $G = G\prime$ implica $G = N$.

3.2. Geometría Afín

En esta sección, consideramos grupos actuando sobre espacios vectoriales y espacios afines. Todos los espacios vectoriales son de dimensión finita.

Notación. Si K es un campo, el espacio vectorial n-dimensional de todas las n-tuplas sobre K es denotado por $V(n, K)$.

Observamos que existen dos subgrupos actuando sobre un espacio vectorial V los cuales son: $GL(V) \cong GL(n, K)$ y $SL(V) \cong SL(n, K)$; además, cada uno de estos actúa también sobre $V^* = V \setminus \{0\}$.

Teorema 3.2.1. *Se cumple lo siguiente:*

(*i*) *Sea V un espacio vectorial de dimensión n sobre un campo K. Entonces V^* es un $GL(V)$-conjunto transitivo que es regular cuando $n = 1$.*

(*ii*) *Si $n \geq 2$, entonces V^* es doblemente transitivo si y sólo si $K = \mathbb{F}_2$.*

Demostración:

1. Que $GL(V)$ actúa transitivamente sobre V^*, se sigue de los siguientes hechos: (1) cada vector distinto de cero es parte de una base; (2) existe una transformación lineal no singular aplicando una base de V sobre alguna otra base de V. Si $n = 1$, sólo la identidad fija un vector distinto de cero, así $GL(V)$ actúa regularmente.

2. $\Rightarrow$) Suponga que $n \geq 2$. Si $K \neq \mathbb{F}_2$, existe un conjunto dependiente conteniendo dos vectores distintos de cero, digamos $\{x, \lambda x\}$, donde $x \in V^*$ y $\lambda \in K$, $\lambda \neq 0, 1$. Si $\{y, z\}$ es un conjunto independiente en V, no existe $g \in GL(V)$ con $g(x) = y$ y $g(\lambda x) = z$. Por lo tanto, V^* no es $GL(V)$-conjunto doblemente transitivo.

   $\Leftarrow$) Si $K = \mathbb{F}_2$, entonces cualesquier pareja de vectores no cero forman un conjunto independiente, y podemos extenderla a una base; por el hecho (2) citado en el primer párrafo, se obtiene lo deseado.

□

Definiremos un nuevo grupo actuando sobre un espacio vectorial.

Definición 3.2.1. *Sea V es un espacio vectorial.*

(*i*) *Si $y \in V$, entonces la* ***traslación*** *por y es la función $t_y : V \to V$ definida por*

$$t_y(x) = x + y,$$

para toda $x \in V$. Sea $\mathrm{Tr}(V)$ denotando el grupo (bajo composición de funciones) de todas las traslaciones de V.

(*ii*) *El conjunto de todas funciones $a : V \to V$ de la forma $a = t_{x_0}g$, es decir,*

$$a(x) = g(x) + x_0,$$

donde $g \in GL(V)$ y $x, x_0 \in V$, forma un grupo bajo composición de funciones, el cual denotaremos por $\mathrm{Aff}(V)$, y lo llamaremos ***grupo afín****. Cuando $V = V(n, K)$ escribiremos $\mathrm{Aff}(n, K)$.*

Observación 3.2.1. $\mathrm{Tr}(V)$ es un subgrupo normal abeliano de $\mathrm{Aff}(V)$, y $\mathrm{Tr}(V)$ es isomorfo al grupo aditivo V.

Demostración: Notemos que si $a = t_{x_0}g$, entonces $a^{-1} = g^{-1}t_{-x_0} = t_{-g^{-1}x_0}g^{-1}$. Por lo tanto si $t_y \in \mathrm{Tr}(V)$, entonces

$$\begin{aligned} at_ya^{-1}(x) &= t_{x_0}g \circ t_y \circ t_{-g^{-1}x_0}g^{-1}(x) \\ &= t_{x_0}g \circ t_y(g^{-1}x - g^{-1}x_0) \\ &= t_{x_0}g(g^{-1}x - g^{-1}x_0 + y) \\ &= x - x_0 + gy + x_0 \\ &= x - gy = t_{-gy}(x). \end{aligned}$$

Ahora veamos que el grupo es abeliano. Esto es claro ya que $t_y \circ t_z = t_{y+z} = t_{z+y} = t_z \circ t_y$, por lo tanto, si definimos la función

$$\begin{aligned} f : \mathrm{Tr}(V) &\to V \\ t_y &\mapsto y \end{aligned}$$

esta función es un isomorfismo entre los grupos abelianos $\mathrm{Tr}(V)$ a V (este último visto como grupo aditivo).

□

Teorema 3.2.2. *Un espacio vectorial V de dimensión n sobre un campo K es un* $\mathrm{Aff}(V)$*-conjunto doblemente transitivo que es simplemente* 2*-transitivo cuando* $n = 1$.

Demostración: Antes de todo, es fácil ver que $\mathrm{Tr}(V)$ y $\mathrm{Aff}(V)$ actúan transitivamente sobre V. Elijamos $w \in V^*$. Si u y v son vectores distintos en V, será suficiente encontrar $a \in \mathrm{Aff}(V)$ con $a(w) = u$ y $a(0) = v$. Ahora a tiene la forma: $a(x) = g(x) + x_0$. Definamos $x_0 = v$ y elijamos $g \in GL(V)$ con $g(w) = u - v$ (tal g existe porque $w \neq 0$ y $u - v \neq 0$). Por lo tanto, $\mathrm{Aff}(V)$ actúa doblemente transitivivamente sobre V. Finalmente, si $n = 1$, la elección de g es única y la acción es simple. □

Es claro que V nunca es un $\mathrm{Aff}(V)$-conjunto triplemente transitivo.

Por el momento divagaremos al definir un isomorfismo afín [sin embargo $\mathrm{Aff}(V)$ consiste de ciertos objetos, llamados afinidades] pues de cualquier modo no tenemos definido lo que es un espacio afín. Después daremos la definición, observaremos que el adjetivo "afín" significa "finito" y su significado llegará a ser claro cuando discutamos espacios proyectivos.

Vamos ahora a considerar subconjuntos "lineales" de un espacio vectorial V. Claramente no debemos restringir nuestra atención a subespacios de V; siendo, por ejemplo, los subespacios lineales unidimensional todas las rectas que pasen a través del origen.

Definición 3.2.2. *Sea V un espacio vectorial sobre un campo K y sea S un subespacio de dimensión m. Un m-**subespacio afín** de V es un subconjunto de la forma $S + v$ para algún $v \in V$. Diremos que la **dimensión** de $S + v$ es m.*

Existe un nombre especial para ciertos m-subconjuntos afines: si $m = 0, 1,$ o 2, entonces los m-subconjuntos afines son llamados **puntos**, **rectas** o **planos afines**, respectivamente. Si V tiene dimensión n, un $(n-1)$-subconjunto afín es llamado **hiperplano afín**.

Uno puede considerar un subespacio afín $S + v$ de dos forma: como la traslación de un subespacio $[S + v = t_v(S)]$ o como una clase lateral de un subespacio S (con representante v).

Lema 3.2.1. *Sea V un espacio vectorial de dimensión $n \geq 2$ sobre un campo K.*

(*i*) *Dos rectas distintas $L_i = Ku_i + v_i,\ i = 1, 2$, son o bien disjuntas o se intersectan en un único punto.*

(*ii*) *Dos puntos distintos* $u, v \in V$ *pertenecen a una única recta* L*, digamos,* $L = K(u - v) + v$*.*

(*iii*) (*Ley del paralelogramo*) *Si* $\{u, v\}$ *es indenpendiente, entonces*

$$\{u + v\} = (Ku + v) \cap (Kv + u).$$

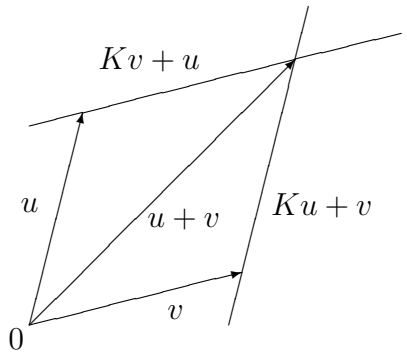
(*iv*) *Dos hiperplanos distintos* $H + x$ *y* $J + y$ *son disjuntos si y sólo si* $H = J$*. En particular, si* $n = 2$*, las distintas rectas* $Ku_1 + v_1$ *y* $Ku_2 + v_2$ *son disjuntas si y sólo si* $Ku_1 = Ku_2$*.*

Demostración:

(*i*) Suponga $z \in L_1 \cap L_2$. Las rectas son clases laterales, por lo tanto tenemos que $L_1 \cap L_2 = (Ku_1 \cap Ku_2) + z$. Ahora $Ku_1 = Ku_2$ no puede ocurrir ya que las rectas distintas L_1 y L_2 son clases laterales distintas de Ku_1, entonces son disjuntas. Por tanto $Ku_1 \cap Ku_2 = \{0\}$ y $L_1 \cap L_2 = \{z\}$.

(*ii*) Es claro que $K(u-v)+v$ es una recta conteniendo a u y v; una segunda recta que intersecte a una de estas en al menos dos puntos, contradice a (*i*).

(*iii*) La afirmación sugiere por la "ley del paralelogramo" representar la adición de vectores en el plano.



Para probar (*iii*), primero notemos que la independencia de $\{u, v\}$ implica que las lineas $Ku+v$ y $Kv+u$ son distintas. Claramente, $u+v \in (Ku+v) \cap (Kv+u)$, así el resultado se sigue de (*i*).

(*iv*) $\Rightarrow$) Si $H = J$, los distintitos hiperplanos son distintas clases laterales de H, entonces son disjuntos.

$\Leftarrow$) Si $H \neq J$, entonces $H + J = V$ (por que H y J son hiperplanos). Luego $y-x = h+j$ para $h \in H$ y $j \in J$. Por lo tanto $h+x = -j+y \in (H+x)\cap(J+y)$.

La última afirmación se sigue de que un hiperplano es un espacio dos-dimensional es una recta.

□

Es necesario enfocar nuestra atención en un subespacio afín de un espacio vectorial, ignorando la presencia de la suma vectorial y la multiplicación escalar (su "geometría").

Definición 3.2.3. *Sea V un espacio vectorial n-dimensional sobre un campo K, sea A un conjunto no vacío y para cada m, con $0 \leq m \leq n$, $\mathcal{L}_m(A)$ una familia de subconjuntos de A. Un n-**espacio afín sobre** K **con espacio vectorial asociado** V es el triplete $(A, \mathcal{L}_*(A), \alpha)$, donde*

(*i*) *$\alpha : V \to A$ es una función biyectiva.*

(*ii*) *Un subconjunto S de V es un m-subespacio afín si y sólo si $\alpha(S) \in \mathcal{L}_m(A)$.*

*Un subconjunto S en $\mathcal{L}_m(A)$ es llamado un m-**subespacio afín**.*

Inmediatamente uno piensa, de manera intuitiva, que $\mathcal{L}_0(A)$ son los puntos de A, $\mathcal{L}_1(A)$ las rectas en A, y así sucesivamente.

Observación 3.2.2. Sea V un espacio vectorial. Entonces,

(*i*) V es un espacio afín asociado a sí mismo vía $\alpha = Id_V$, y en donde $\mathcal{L}_m(V)$ es el conjunto de todos los m-subespacios afines de V.

(*ii*) Similarmente que en (*i*), V es un espacio afín asociado a sí mismo vía al tomar $\alpha = t_x$, la translación en x y, de nuevo, $\mathcal{L}_m(V)$ el conjunto de todos los m-subespacios afines de V.

Por simplicidad, nos referiremos a la pareja (A, α) en lugar del triplete $(A, \mathcal{L}_*(A), \alpha)$.

Definición 3.2.4. *Sean K un campo, y (A, α) y (B, β) espacios afines sobre K asociados a ciertos espacios vectoriales. Una función $h : A \to B$ es un **isomorfismo afín** si h es una función biyectiva tal que para cada subconjunto S de A, S pertenece a $\mathcal{L}_m(A)$, para algún m, si y sólo si $h(S) \in \mathcal{L}_m(B)$. Decimos que (A, α) y (B, β) son **isomorfos** si existe un isomorfismo afín entre ellos.*

Definición 3.2.5. *El conjunto de todos los isomorfismos afines de* (A, α) *en sí mismo* (*llamados* ***automorfismos afines***) *forman un grupo bajo la composición, denotado por* $\mathrm{Aut}\,(A)$*, y es llamado el* ***grupo de automorfismos****.*

Si (A, α) es un espacio afín sobre un campo K asociado con un espacio vectorial V, y si V es construido como un espacio afín (V, Id_V), Observación 3.2.2, entonces es fácil ver que la función $\mathrm{Aut}\,(A) \to \mathrm{Aut}\,(V)$ definida por la correspondencia $h \mapsto \alpha^{-1} h \alpha$ es un isomorfismo de grupos.

Notación. Si $V = V(n, K)$, entonces escribiremos $\mathrm{Aut}\,(n, K)$ en lugar de $\mathrm{Aut}\,(V)$.

Definición 3.2.6. *Sean* (A, α) *y* (B, β) *espacios afines sobre un campo* K *con espacios vectoriales asociados* V *y* W*, respectivamente. Una función* $f : A \to B$ *es una* ***función afín*** *si* $f = \beta g \alpha^{-1}$ *para alguna transformación lineal no singular* $g : V \to W$*. Todas las funciones afines de* (A, α) *en sí mismo forman un grupo bajo la composición, denotado por* $GL(A)$*.*

Tres observaciones son necesarias. Primero, cada función afín es un isomorfismo afín, por lo cual $GL(A)$ es un subgrupo de $\mathrm{Aut}\,(A)$. Segundo, la función $GL(A) \to GL(V)$ definida por la correspondencia $f \mapsto \alpha^{-1} f \alpha$ es un isomorfismo de grupos tal que la notación $GL(A)$ coincide con la notación usual donde ambas están definidas, es decir, donde A es un espacio vectorial. Y tercero, uno puede usar la Observación 3.2.2 para construir cada $a \in \mathrm{Aff}\,(V)$ como una función afín $(V, t_{-y}) \to (V, Id_V)$; por supuesto, t_y depende de a.

Teorema 3.2.3. *Dos espacios afines sobre un campo* K *son isomorfos si y sólo si tienen la misma dimensión.*

Demostración: $\Rightarrow$) Sean (A, α) y (B, β) espacios afines sobre el campo K con espacios vectoriales asociados V y W, respectivamente. Si estos son isomorfos, entonces es obvio que los espacios afines tienen la misma dimensión, es decir, los espacios vectoriales V y W son isomorfos.

$\Leftarrow$) Si (A, α) y (B, β) tienen dimensión n, entonces V y W tienen dimensión n, y existe una transformación lineal no singular $g : V \to W$ tal que la función $f : A \to B$ definida por $f = \beta g \alpha^{-1}$ es un isomorfismo, es decir, es una función afín, y (A, α) y (B, β) son isomorfos. $\square$

Podemos ahora retomar los espacios vectoriales, pues cada espacio afín es isomorfo a un espacio vectorial (viéndolo como un espacio afín sobre sí mismo). Veamos isomorfismos afines que no son transformaciones lineales ni traslaciones.

Suponga que $\sigma : K \to K$ es un automorfismo de un campo K (así, σ es una permutación con $\sigma(\lambda + \mu) = \sigma(\lambda) + \sigma(\mu)$, $\sigma(\lambda\mu) = \sigma(\lambda)\sigma(\mu)$ y $\sigma(1) = 1$). Si $K = \mathbb{F}_q = \mathbb{F}_{p^r}$, el campo finito de q elementos, entonces hemos visto en el Teorema 1.8.6 que $\mathrm{Aut}\,(K)$ es un grupo cíclico de orden r con generador $\sigma : \lambda \mapsto \lambda^p$. Observemos que el campo $\mathbb{R}$ no tiene otro isomorfismo más que la identidad. El campo $\mathbb{C}$ tiene automorfismos distintos de identidad, por ejemplo, la conjugación compleja.

Definición 3.2.7. *Sean V y W espacios vectoriales sobre un campo K. Una función $g : V \to W$ es una* ***transformación semilineal*** *si existe $\sigma \in \mathrm{Aut}\,(K)$ tal que*

$$g(x + y) = g(x) + g(y)$$

y

$$g(\lambda x) = \sigma(\lambda)g(x),$$

para cada $x, y \in V$ y $\lambda \in K$.

Por supuesto, que cada transformación lineal es semilineal (con σ la identidad).

Observación 3.2.3. *Si eligimos una base de V tal que cada $x \in V$ tenga coordenadas: $x = (x_1, \ldots, x_n)$, y $\sigma \in \mathrm{Aut}\,(K)$, entonces al definir $\sigma_* : V \to V$ por*

$$\sigma_*(x_1, \ldots, x_n) = (\sigma(x_1), \ldots, \sigma(x_n)),$$

tendremos que σ_* será una transformación semilineal.

Definición 3.2.8. *Una transformación semilineal es* ***no singular*** *si es una función inyectiva.*

Definición 3.2.9. *$\Gamma L(V)$ es el grupo de todas las transformaciones semilineales de V en sí mismo bajo la composición. Si $V = V(n, K)$, escribiremos $\Gamma L(V) = \Gamma L(n, K)$.*

Observación 3.2.4. La composición de transformaciones semilineales es semilineal (si f y g son transformaciones semilineales con automorfismos σ y τ respectivamente, entonces gf es tranformación semilineal con automorfismo $\tau\sigma$). Si una transformación semilinal es no singular, entonces su inversa también es semilineal.

Que $\Gamma L(V)$ es un grupo se sigue de inmediato de la Observación 3.2.4.

Obviamente, cada transformación semilineal no singular es un isomorfismo afín. De aquí que, si $h : V \to W$ es función biyectiva de la forma

$$h(x) = g(x) + w_0,$$

donde $x, w_0 \in V$ y g es transformación semilineal no singular, entonces h es un isomorfismo afín. De hecho, la observación es que cada isomorfismo afín tiene esta forma (ver Teorema 3.2.4).

Denotaremos al subespacio de un espacio vectorial V generado por un subconjunto X por $\langle X\rangle$.

Lema 3.2.2. *Sean V y Y espacios vectoriales sobre un campo K, $h : V \to Y$ un isomorfismo afín con $h(0) = 0$, y W un subespacio de V con base $\{u, v\}$. Entonces,*

(*i*) $h(Ku) = Kh(u)$*;*

(*ii*) $\{h(u), h(v)\}$ *es linealmente independiente;*

(*iii*) $h(W) = \langle h(u), h(v)\rangle$*;*

(*iv*) $h|W : W \to h(W)$ *es un isomorfismo afín.*

Demostración:

(*i*) Ku es una recta conteniendo a u y a 0; así que, $h(Ku)$ es la línea conteniendo $h(u)$ y $h(0) = 0$. Por lo tanto, $h(Ku) = Kh(u)$.

(*ii*) Si $\{h(u), h(v)\}$ son dependientes, entonces los puntos $h(u), h(v), 0$ son colineales; aplicando el isomorfismo afín h^{-1}, tenemos la contradicción de que $u, v, 0$ colineales.

(*iii*) Para cada pareja ordenada $\lambda, \mu \in K$, (no ambos 0), sea $L(\lambda, \mu)$ la recta conteniendo λu y μv. Es fácil ver que

$$W = \bigcup_{\lambda,\mu} L(\lambda, \mu);$$

por tanto

$$h(W) = \bigcup_{\lambda,\mu} h(L(\lambda,\mu)).$$

Consideremos una recta $L(\lambda,\mu)$ fija. Por (i), $h(\lambda\mu) = \alpha h(u)$ y $h(\mu v) = \beta h(v)$ para algunos $\alpha, \beta \in K$. Pero $h(\lambda u)$ y $h(\mu v)$ pertenecen a $\langle h(u), h(v)\rangle$ para cada λ y μ, por lo tanto $h(L(\lambda,\mu)) \subseteq \langle h(u), h(v)\rangle$; así que, $h(W) \subseteq \langle h(u), h(v)\rangle$. Para la inclusión inversa, aplicando h^{-1} a $\langle h(u), h(v)\rangle$ [notando que, por (ii), $\{h(u), h(v)\}$ es independiente] y teniendo que $h^{-1}(\langle h(u), h(v)\rangle) \subset \langle u, v\rangle = W$, se obtiene que $\langle h(u), h(v)\rangle \subset h(\langle u, v\rangle) = h(W)$.

(iv) Esto es inmediato dado que ahora sabemos que $h(W)$ es un espacio vectorial teniendo la misma dimensión de W.

□

Teorema 3.2.4. *Sean S y T espacios vectoriales isomorfos de dimensión de al menos 2 sobre un campo K. Entonces, cada isomorfismo afín $h : S \to T$ tiene la forma $h = t_y g$ para algún $y \in T$ y $g : S \to T$ una transformación semilineal no singular, es decir,*

$$h(x) = g(x) + y,$$

para cada $x \in S$.

Nota. Uno debe asumir que $\dim S \geq 2$, pues cada función biyectiva entre espacios unidimensionales es un isomorfismo afín.

Demostración: Componiendo h con la traslación $x \mapsto x - h(0)$, podemos suponer que $h(0) = 0$, y ahora es suficiente probar que h es semilineal. Puesto que h es una función biyectiva, h preserva intersección; en particular, $h(L_1 \cap L_2) = h(L_1) \cap h(L_2)$ si L_1 y L_2 son rectas.

Suponga que $\{u, v\}$ es independiente; claramente

$$h(Ku + v) = Kh(u) + h(v).$$

Por Lema 3.2.2 (iii), tanto $h(Ku+v)$ como $h(Ku)$ son lineas contenidas en $\langle h(u), h(v)\rangle$; de hecho, son disjuntas porque $Ku + v$ y Ku son disjuntos. Dado que $h(Ku) = Kh(u)$, por Lema 3.2.2 (i), podemos aplicar el Lema 3.2.1 (iv) a $\langle h(u), h(v)\rangle$ para obtener que

$$h(Ku + v) = Kh(u) + y,$$

para algún y. En particular, existen escalares $\alpha, \beta \in K$ tales que $h(v) = \alpha h(u) + y$ y $h(\lambda\mu + v) = \beta h(u) + y$, donde $\lambda \in K$ y β depende de λ. Así,

$$\begin{aligned} h(\lambda u + v) &= \beta h(u) + h(v) - \alpha h(u) \\ &= (\beta - \alpha)h(u) + h(v) \in Kh(u) + h(v). \end{aligned}$$

Por tanto $h(Ku+v) \subset Kh(u)+h(v)$ e igualmente se sigue de que ambas son rectas.

Primero probaremos que $h(u+v) = h(u)+h(v)$, donde $\{u, v\}$ es independiente. Por Lema 3.2.1 (iii), sabemos que $\{u+v\} = (Ku+v)\cap(Kv+u)$. Puesto que h preserva intersecciones, tenemos que

$$\begin{aligned} \{h(u+v)\} &= h(Ku+v) \cap h(Kv+u) \\ &= [Kh(u)+h(v)] \cap [Kh(v)+h(u)] \\ &= \{h(u)+h(v)\}, \end{aligned}$$

la última igualdad se tiene porque $\{h(u), h(v)\}$ es independiente [Lema 3.2.2 (ii)]. Resta por evaluar $h(\lambda u+\mu u)$; hacemos esto en dos pasos. Elijamos w tal que $\{u, w\}$ sea independiente. Entonces, $\{u+w, -u\}$ es también independiente, y

$$h(w) = h((u+w)-u) = h(u+w) + h(-u) = h(u) + h(w) + h(-u).$$

Se sigue que $h(-u) = -h(u)$. En consecuencia, si $\lambda + \mu = 0$, tenemos $h(\lambda u + \mu u) = 0 = h(\lambda u) + h(\mu u)$. Finalmente, supongamos $\lambda + \mu \neq 0$. Entonces, tenemos que $\{\lambda u + w, \mu u - w\}$ es independiente y

$$\begin{aligned} h(\lambda u + \mu u) &= h(\lambda u + w) + h(\mu u - w) \\ &= h(\lambda u) + h(w) + h(\mu u) + h(-w) = h(\lambda u) + h(\mu u). \end{aligned}$$

Hemos probado que h es aditivo.

Si $u \neq 0$ y $\lambda \in K$, entonces $h(Ku) = Kh(u)$ implica que existe $\sigma_u(\lambda) \in K$ con $h(\lambda u) = \sigma_u(\lambda)h(u)$. Obviamente $\sigma_u(1) = 1$. La aditivamente de h implica

$$\begin{aligned} \sigma_u(\lambda+\mu)h(u) &= h((\lambda+\mu)u) = h(\lambda u + \mu u) \\ &= h(\lambda u) + h(\mu u) = [\sigma_u(\lambda) + \sigma_u(\mu)]h(u); \end{aligned}$$

dado que $h(u) \neq 0$, vemos que σ_u es aditivo.

A continuación probaremos que σ_u no depende de u. Eligiendo w tal que $\{u, w\}$ es independiente. Ahora

$$h(\lambda u + \lambda w) = h(\lambda u) + h(\lambda w) = \sigma_u(\lambda)h(u) + \sigma_w(\lambda)h(w).$$

Por otro lado,

$$h(\lambda u + \lambda w) = \sigma_{u+w}(\lambda)h(u+w) = \sigma_{u+w}(\lambda)[h(u) + h(w)].$$

Igualando coeficientes,

$$\sigma_u(\lambda) = \sigma_{u+w}(\lambda) = \sigma_w(\lambda).$$

Por último, si $\mu \in K^*$, entonces $\{\mu u, w\}$ es independiente y tenemos, con μu en lugar de u, que $\sigma_{\mu u}(\lambda) = \sigma_w(\lambda)$.

Resta por probar $\sigma : K \to K$ es multiplicativo (el subíndice puede ser ahora omitido). Pero

$$h(\lambda\mu u) = \sigma(\lambda\mu)h(u)$$

y también

$$h(\lambda\mu u) = \sigma(\lambda)h(\mu u) = \sigma(\lambda)\sigma(\mu)h(u).$$

Comparando las dos últimas igualdades, tenemos que $\sigma \in \mathrm{Aut}\,(K)$, y h es semilineal. □

La demostración anterior se utiliza el hecho de que $h(L)$ es una recta siempre que L lo sea, pero no que $h(S)$ sea subespacio afín para algún subespacio afín S de dimensión mayor. La hipótesis del Teorema 3.2.4 puede así debilitarse.

Corolario 3.2.1. *Sean (A, α) y (B, β) espacios afines sobre K de dimensión ≥ 2 con espacios vectoriales asociados V y W, respectivamente. Si $f : A \to B$ es un isomorfismo afín, entonces*

$$f = \beta t_z g \alpha^{-1},$$

donde $g : V \to W$ es una transformación semilineal no singular y $t_z : W \to W$ es una traslación por $z = \beta^{-1} f \alpha(0)$.

Demostración: Si $g\prime : V \to W$ está definida por $g\prime = \beta^{-1} f \alpha$, entonces $g\prime$ es un isomorfismo afín; si $g = t_{-z} g\prime$, entonces $g : V \to W$ es un isomorfismo afín con $g(0) = 0$. Por Teorema 3.2.4, g es una transformación semilineal no singular. Por lo tanto

$$g = t_{-z} g\prime = t_{-z} \beta^{-1} f \alpha$$

y el resultado se sigue. □

Teorema 3.2.5. *Sea V un espacio vectorial de dimensión finita n sobre un campo K. Entonces,*

(*i*) $\mathrm{Aff}(V) \subset \mathrm{Aut}(V)$, *y* $\mathrm{Aut}(V)$ *actúa doblemente transitivamente sobre* V.

(*ii*) *Si* $\mathrm{Aut}(K) = \{1\}$, *entonces* $\mathrm{Aff}(V) = \mathrm{Aut}(V)$.

(*iii*) $\mathrm{Aut}(V)$ *es un producto semidirecto de* $\mathrm{Tr}(V)$ *por* $\Gamma L(V)$.

(*iv*) *Si* $V = V(n, \mathbb{F}_q)$, *entonces* $|\mathrm{Aut}(n, \mathbb{F}_q)| = q^n |\Gamma L(n, \mathbb{F}_q)|$.

Demostración:

(*i*) Que $\mathrm{Aff}(V) \subset \mathrm{Aut}(V)$ es claro por su definición; la segunda afirmación resulta de la doble transitividad de la acción de $\mathrm{Aff}(V)$ (Teorema 3.2.2).

(*ii*) Si $\mathrm{Aut}(K) = \{1\}$, entonces cada transformación semilineal es lineal.

(*iii*) Si $h \in \mathrm{Aut}(V)$, entonces h tiene la forma $h(x) = g(x) + x_0$ para algún $g \in \Gamma L(V)$ y para algún $x_0 \in V$. Definimos $\pi : \mathrm{Aut}(V) \to \Gamma L(V)$ por $\pi(h) = g$. Se tiene que $\ker(\pi) = \mathrm{Tr}(V)$ y que π fija a cada elemento de $\Gamma L(V)$, de donde π es una retracción (Observación 1.7.1). Por lo tanto, $\mathrm{Aut}(V)$ es el producto semidirecto de $\mathrm{Tr}(V)$ por $\Gamma L(V)$.

(*iv*) Por la Observación 3.2.1, se tiene que $\mathrm{Tr}(V) \cong V$, así que si $V = V(n, \mathbb{F}_q)$, entonces $|\mathrm{Aut}(n, \mathbb{F}_q)| = |\mathrm{Tr}(V)||\Gamma L(n, \mathbb{F}_q)| = q^n |\Gamma L(n, \mathbb{F}_q)|$.

□

Teorema 3.2.6. *Si V es un espacio vectorial de dimensión n sobre un campo K, entonces $\Gamma L(V)$ es un producto semidirecto de $GL(V)$ por $\mathrm{Aut}(K)$. Si $V = V(n, \mathbb{F}_q)$, con $q = p^r$, entonces*

$$|\Gamma L(n, \mathbb{F}_q)| = r|GL(n, \mathbb{F}_q)|.$$

Demostración: Cada transformación semilineal no singular h determina un único automorfismo σ de K. La función $\theta : \Gamma L(V) \to \mathrm{Aut}(K)$ definida por la correspondencia $h \mapsto \sigma$ es un epimorfismo cuyo núcleo es $GL(V)$. Como en la Observación 3.2.3, eligiendo una base de V, tenemos que para cada $\sigma \in \mathrm{Aut}(K)$, podemos considerar la transformación semilineal σ_*. Así, el conjunto de todas las σ_*, con $\sigma \in \mathrm{Aut}(K)$, es un subgrupo de $\Gamma L(V)$ cuya intersección con $GL(V)$ es trivial. Por lo tanto, $\Gamma L(V)$ es el producto semidirecto de $GL(V)$ por $\mathrm{Aut}(K)$.
Cuando $V = V(n, \mathbb{F}_q)$, entonces

$$|\Gamma L(n, \mathbb{F}_q)| = |\mathrm{Aut}\,(\mathbb{F}_{p^r})||GL(n, \mathbb{F}_q)| = r|GL(n, \mathbb{F}_q)|,$$

por el Teorema 1.8.6. □

Observemos que $|\mathrm{Aut}\,(n, \mathbb{F}_q)| = rq^n|GL(n, \mathbb{F}_q)|$. Dado que $|GL(n, \mathbb{F}_q)|$ ha sido calculado en el Teorema 2.1.1, tenemos fórmulas explicitas para $|\mathrm{Aut}\,(n, \mathbb{F}_q)|$ y $|\Gamma L(n, \mathbb{F}_q)|$.

3.3. Geometría Proyectiva

Sea V un espacio vectorial sobre un campo K. Definimos una relación de equivalencia sobre $V^* = V \setminus \{0\}$ por $x \sim y$ si existe $\lambda \in K^*$ tal que $y = \lambda x$. Si $x \in V^*$, denotamos su clase de equivalencia por $[x]$.

Definición 3.3.1. *Sea V un espacio vectorial de dimensión $n+1$ sobre un campo K. El conjunto de clases de equivalencia*

$$P(V) = \{[x] \mid x \in V^*\}$$

*es llamado el **n-espacio proyectivo**; decimos que $P(V)$ tiene **dimensión proyectiva** n.*

Como con los espacios afines, en los espacios proyectivos no tenemos operaciones algebraicas como en los espacios vectoriales: no podremos sumar puntos de $P(V)$ ni multiplicarlos por escalares. Estas operaciones pueden ser restituidas en un espacio afín (A, α) vía la función α. Esto no es posible con los espacios proyectivos.

Definición 3.3.2. *Sea $\pi : V^* \to P(V)$ la función suprayectiva canónica $x \mapsto [x]$. Un m-**subespacio proyectivo** es un subconjunto S de $P(V)$ de la forma $\pi(W^*)$, donde W es un subespacio $(m+1)$-dimensional de V. Decimos que S tiene **dimensión** m.*

Aquí también, existen nombres especiales para ciertos subespacios: si $m = 0, 1$ o 2, un m-subespacio proyectivo es un **punto proyectivo**, **recta proyectiva**, o **plano proyectivo**, respectivamente. Si $P(V)$ tiene dimensión n, un $(n-1)$-subespacio proyectivo es llamado un **hiperplano proyectivo**.

La razón de disminuir la dimensión pasando de V a $P(V)$ por ahora no es aparente: una recta en V (atravesando por el origen) se convierte en un punto proyectivo;

un plano en V (atravesando por el origen) se convierte en una recta proyectiva y así sucesivamente.

El siguiente resultado sencillo es fundamental.

Teorema 3.3.1. *Si V es un espacio vectorial sobre un campo K con $n = \dim(V) \geq 2$.*

(*i*) *Si $x, y \in V^*$, entonces $[x] \neq [y]$ si y sólo si $\{x, y\}$ es independiente.*

(*ii*) *Cada dos puntos distintos en $P(V)$ pertenecen a una única recta proyectiva.*

(*iii*) *Si Ω es un hiperplano proyectivo en $P(V)$ y L es una recta proyectiva no contenida en Ω, entonces Ω y L se intesecta en un único punto proyectivo.*

Demostración:

(*i*) Si $[x] \neq [y]$, entonces $x \neq \lambda y$ para $\lambda \in K^*$ y $\{x, y\}$ es independiente; la recíproca es obvia.

(*ii*) Sean $x, y \in V^*$ tales que $[x] \neq [y]$ en $P(V)$. Una recta proyectiva L conteniendo $[x]$ y $[y]$ debe tener la forma $\pi(W^*)$, donde W es un subespacio 2-dimensional de V conteniendo a x e y. El conjunto $\{x, y\}$ es independiente, por (i), por lo cual $\langle x, y\rangle$ es un subespacio de V 2-dimensional. Esto prueba la unicidad de L.

(*iii*) Sean $L = \pi(W^*)$ y $\Omega = \pi(H^*)$, donde $\dim(W) = 2$ y $\dim(H) = n - 1$. Dado que L no está contenida en Ω, el subespacio W no está contenido en H. Por lo tanto, $W + H = V$ y

$$\begin{aligned}\dim(W \cap H) &= \dim(W) + \dim(H) - \dim(W + H)\\ &= 2 + (n-1) - n = 1.\end{aligned}$$

Entonces $\pi((W \cap H)^*)$ es un punto proyectivo.

□

Veremos conceptos análogos a los isomorfismos afines los cuales podemos llamarlos isomorfismos proyectivos.

Definición 3.3.3. *Sean V y W espacios vectoriales sobre un campo K. Una **colineación** o un **isomorfismo proyectivo** de V en W es una función biyectiva $\theta : P(V) \to P(W)$ tal que el subconjunto S de $P(V)$ es un m-subespacio proyectivo si y sólo si $\theta(S)$ es un m-subespacio de $P(W)$. Dos espacios proyectivos son **isomorfos** si existe un isomorfismo proyectivo entre estos.*

Nota. Usaremos la palabra "isomorfismo" en lugar de colineación proyectiva o isomorfismo proyectivo.

Observación 3.3.1. Si $h : V \to W$ es una transformación semilineal no singular, entonces h induce un isomorfismo $P(h) : P(V) \to P(W)$ dada por la correspondencia $[x] \mapsto [h(x)]$.

Definición 3.3.4. *Un isomorfismo de la forma $P(h)$ descrita en la observación anteriormente se dice ser una* ***proyectividad*** *si h es una transformación lineal no singular.*

Aquí, las proyectividades son análogas a las funciones afines en el sentido de que cada una de ellas surge de transformaciones lineales. Veremos que sólo otros isomorfismos surgen de transformaciones semilineales (ver Teorema 3.3.6).

Teorema 3.3.2. *Dos espacios proyectivos $P(V)$ y $P(W)$ son isomorfos si y sólo si tienen la misma dimensión.*

Demostración:
$\Rightarrow$) Es obvia.

$\Leftarrow$) La igualdad de las dimensiones de $P(V)$ y $P(W)$ implica la igualdad de las dimensiones de V y W. Existe así un isomorfismo (lineal) $h : V \to W$ induciendo un isomorfismo $P(h)$ (el cual es igual a una proyectividad) entre $P(V)$ y $P(W)$. □

Notación. *Si $V = V(n+1, K)$, escribiremos $P^n(K)$ para $P(V)$.*

Teorema 3.3.3.

(*i*) *Para cada $n \geq 0$ y cada potencia q del primo p,*

$$|P^n(\mathbb{F}_q)| = q^n + q^{n-1} + \cdots + q + 1.$$

En particular, cada recta proyectiva tiene exactamente $q + 1$ puntos.

(*ii*) *El número de lineas proyectivas en el plano proyectivo $P^2(\mathbb{F}_q)$ es el mismo que el número de puntos en $P^2(\mathbb{F}_q)$, es decir, $q^2 + q + 1$.*

Demostración:

(*i*) Al calcular $|P^n(\mathbb{F}_q)|$, notemos que si $V = V(n+1, \mathbb{F}_q)$, entonces $|V^*| = q^{n+1} - 1$. Dado que V^* está particionado en clases de la forma $[x]$ cada una de las cuales tiene $q-1$ elementos ($[x] = \{\lambda x \mid \lambda \neq 0\}$), tenemos que

$$|P^n(\mathbb{F}_q)| = (q^{n+1} - 1)/(q-1) = q^n + q^{n-1} + \cdots + q + 1.$$

En caso de que $n = 1$ esta fórmula da el número de puntos sobre una recta proyectiva.

(*ii*) Sean L una recta y $[x] \notin L$. Para cada punto $[y] \in L$, existe una recta uniendo a $[x]$ con $[y]$; así hemos exhibido $q+1$ rectas que pasan por $[x]$. Estas son todas las rectas que pasan por $[x]$, ya que cada recta que contenga a $[x]$ debe de intersectar a L [Teorema 3.3.1 (*iii*)].

Para cada uno de los $q+1$ puntos en L, existe q rectas atravesando los distintos puntos de L. Esto exhibe $(q+1)q + 1$ rectas en $P^2(\mathbb{F}_q)$ y son todas.

□

Usaremos el siguiente lema elemental sobre espacios vectoriales para probar un teorema básico relacionando espacios proyectivos y afines y el cual afirma que un espacio proyectivo surge de un espacio afín por adjuntar un "hiperplano al infinito".

Lema 3.3.1. *Sean H un subespacio de un espacio vectorial V de dimensión finita y $y \notin H$. Si T es un subespacio de H con $T + u \subset H$ para algún $u \in H$, entonces*

(*i*) $\dim\langle T + u + y\rangle = 1 + \dim(T)$;

(*ii*) $\langle T + u + y\rangle \cap (H + y) = T + u + y$.

Demostración:

(*i*) Dado que $u \in H$ y $y \notin H$, tenemos $u + y \notin H$ y $\langle T + u + y\rangle = T \oplus \langle u + y\rangle$.

(*ii*) Claramente $T + u + y \subseteq \langle T + u + y\rangle \cap (H + y)$. Para la inclusión inversa, supongamos $t + \lambda(u + y) = h + y$, donde $t \in T$ y $h \in H$. Entonces $(1 - \lambda)y = t + \lambda u - h \in H$. Dado que $y \notin H$, tenemos que $\lambda = 1$ y así $h + y = t + u + y \in T + u + y$.

□

Teorema 3.3.4. *Sea $P(V)$ un n-espacio proyectivo sobre K. Suponga que Ω es un hiperplano proyectivo en $P(V)$, y sea H el subespacio lineal de V con $\pi(H^*) = \Omega$. Si $A = P(V) - \Omega$, el complemento de Ω, entonces existe $\alpha : H \to A$ tal que (A, α), es un n-espacio afín sobre K.*
Además, cada recta (afín) L en A pertenece a una única recta proyectiva L^* en $P(V)$, y L^* intersecta a Ω en un único punto.

Demostración: Recordemos que la función canónica $\pi : V^* \to P(V)$ está definida por la correspondencia $x \mapsto [x]$. Supongamos que B es un subconjunto no vacío de V tal que la relación $b \in B$ implica $\lambda b \in B$ siempre que $\lambda \in K^*$ (por ejemplo, $B = W^*$ tiene esta propiedad para cualquier subespacio lineal W de V). Se cumple que $\pi(B \cap C) = \pi(B) \cap \pi(C)$ para cada subconjunto C de V; en efecto, es claro que el lado izquierdo está contenido en el lado derecho. Para la inclusión recíproca, suponga $[b] = [c] \in \pi(B) \cap \pi(C)$, entonces $c = \lambda b \in B \cap C$ y $[c] \in \pi(B \cap C)$.
Sea $y \in V$ tal que $y \notin H$. Definimos $\alpha : H \to A$ por

$$\alpha(h) = \pi(h + y) = [h + y]$$

[note que $h + y \neq 0$ dado que $y \notin H$, así que $\pi(h + y)$ está bien definido; además, $\pi(h + y) \in A$ dado que $\pi(h + y) \notin \Omega = \pi(H^*)$].

Ahora α es inyectiva, pues dos elementos distintos de la forma $h + y$ (con $h \in H$) no son equivalentes. Veamos que α es sobre. Tomemos $[x] \in A$. Dado que H es un hiperplano lineal en V, tenemos que $x = h + \lambda y$ para algún $h \in H$ y $\lambda \in K$; además, $\lambda \neq 0$, ya que en caso contrario $x \in H$ y $[x] \in \Omega$. Por lo tanto, $x \sim \lambda^{-1}h + y$ y $[x] = \alpha(\lambda^{-1}h)$.

Para $0 \leq m \leq n$, definimos

$$\mathcal{L}_m(A) = \{A \cap \pi(W^*) : W \text{es un subespacio lineal } (m+1)\text{-dimensional de } V\}.$$

Así, un subconjunto está en $\mathcal{L}_m(A)$ si es la intersección de A con un m-subespacio proyectivo de $P(V)$. Es suficiente probar que un conjunto S de H es un m-subespacio afín si y sólo si $\alpha(S) \in \mathcal{L}_m(A)$.

Suponga que $S = T + u$ es un m-espacio afín de H (con $u \in H$ y T subespacio lineal m-dimensional de H). Por el Lema 3.3.1,

$$T + u + y = \langle T + u + y \rangle \cap (H + y).$$

Usando el desarrollo anterior y el hecho de que $0 \notin H + y$, tenemos que

$$\begin{aligned}\alpha(S) &= \pi(T+u+y)\\ &= \pi(\langle T+u+y\rangle \cap (H+y))\\ &= \pi(\langle T+u+y\rangle^*) \cap \pi(H+y)\\ &= \pi(\langle T+u+y\rangle^*) \cap A.\end{aligned}$$

Dado que $\dim\langle T+u+y\rangle = m+1$, se sigue que $\alpha(S) \in L_m(A)$.

Suponga ahora que S es un subconjunto no vacío de H con $\alpha(S) = \pi(S+y) \in \mathcal{L}_m(A)$. Por definición de $\mathcal{L}_m(A)$, existe un subespacio lineal W $(m+1)$-dimensional de V tal que

$$\pi(S+y) = \pi(W^*) \cap A.$$

Pero $\pi(W^*) \cap A = \pi(W^*) \cap \pi(H+y) = \pi(W^* \cap (H+y)) = \pi(W \cap (H+y))$, debido al desarrollo anterior y el hecho de que $0 \notin H+y$. Dado que α es inyectiva, $\pi \mid (H+y)$ es inyectiva y la relación $\pi(S+y) = \pi(W \cap (H+y))$ implica que

$$S+y = W \cap (H+y).$$

En consecuencia, $W \not\subset H$, a no ser que el lado derecho sea el vacío. También tenemos que $W \cap (H+y) = (W \cap H) + z$, donde $z = u+y \in H+y$. Por lo tanto, $S+y = (W \cap H)+u+y$ y $S = (W \cap H)+u$. Hemos probado que S es un subespacio afín de H. Sólo resta calcular su dimensión. Dado que $W \subset H$ y H es un hiperplano lineal de V, tenemos que $V = W+H$. Por lo tanto

$$\begin{aligned}\dim(W \cap H) &= \dim(W) + \dim(H) - \dim(W+H)\\ &= (m+1) + n - (n+1) = m,\end{aligned}$$

y S es un m-subespacio afín de H, como deseábamos.

Así, $L = \alpha(\langle u-v\rangle + v) = \pi(\langle u-v\rangle + v + y)$. Por el Teorema 3.2.4 (ii), existe una única recta proyectiva L^* conteniendo a $[u+y]$ y a $[v+y]$, digamos $L^* = \pi(W^*)$, donde $W = \langle u+y, v+y\rangle$. Dado que $\langle u-v\rangle + v + y \subseteq \langle u+v, v+y\rangle$, se sigue que $L \subset L^*$. Pero L^* no está contenida en Ω con lo cual, por el Teorema 3.3.1 (iii), $L^* \cap \Omega$ es un único punto. □

Existe otra forma de ver el Teorema 3.3.4, usando **coordenadas homogéneas**. Suponga que una base de V ha sido elegida tal que cada $x \in V$ tiene coordenadas $x = (\lambda_0, \lambda_1, \ldots, \lambda_n)$. Así, $[x] \in P(V)$ tiene una familia de coordenadas $(\lambda\lambda_0, \lambda\lambda_1, \ldots, \lambda\lambda_n)$ para $\lambda \neq 0$. Si $\lambda_0 = 0$, entonces la primer coordenada en $[x]$ no es cero. Definiendo

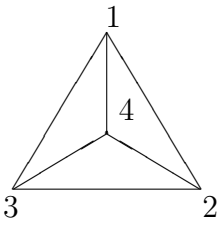
$\Omega = \{[x] \in P(V) | \lambda_0 = 0\}$; uno puede probar fácilmente que Ω es un hiperplano proyectivo. Definiendo $A = \{[x] \in P(V) \mid \lambda_0 \neq 0\}$. Cada $[x] \in A$ tiene un único conjunto de coordenadas de la forma $(1, \lambda_0^{-1}\lambda_1, \ldots, \lambda_0^{-1}\lambda_n)$, y uno puede convertir a A es un espacio afín definiendo $\alpha : V(n, K) \to A$ por $(\mu_1, \ldots, \mu_n) \mapsto [(1, \mu_1 \ldots, \mu_n)]$. En adelante, escribiremos $[\lambda_0, \ldots, \lambda_n]$ en lugar de $[(\lambda_0, \ldots, \lambda_n)]$.

Antes de seguir discutiremos el Teorema 3.3.4, vamos a probar una consecuencia inmediata.

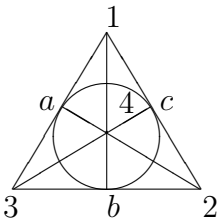
Teorema 3.3.5. *Sean V y W espacios vectoriales de dimensión ≥ 2 y sea Ω un hiperplano proyectivo en $P(V)$ con $A = P(V) - \Omega$. Si θ_1 y θ_2 son isomorfismos $P(V) \to P(W)$ coincidiendo sobre A (es decir, $\theta_1|A = \theta_2|A$), entonces $\theta_1 = \theta_2$.*

Demostración: Sean L una recta afín en A y x e y elementos distintos de L. Puesto que $P(V)$ es un espacio proyectivo, existe una única recta proyectiva L^* a x y a y, por lo tanto a L. Puesto que $\theta_1(x) = \theta_2(x)$ y $\theta_1(y) = \theta_2(y)$, se sigue que $\theta_1(L^*) = \theta_2(L^*)$. Por el Teorema 3.3.1 (iii), $\Omega \cap L^*$ es un punto proyectivo z. Luego, $\theta_1(z) = \theta_1(\Omega \cap L^*) = \theta_1(\Omega) \cap \theta_1(L^*) = \theta_2(\Omega) \cap \theta_2(L^*) = \theta_2(z)$, así que θ_1 y θ_2 también coinciden en todos los puntos de H. □

Ilustraremos el Teorema 3.3.4 para un caso interesante. Sea $V = V(3, \mathbb{F}_2)$ tal que (A, α) es un plano afín sobre $\mathbb{F}_2$. Dibujemos los 4 puntos de A y las 6 rectas entre ellos.



Existen 3 parejas de rectas paralelas (por ejemplo, $\{1, 2\}$ y $\{3, 4\}$), así que $P(V)$ requiere de 3 puntos al infinito. Por el Teorema 3.3.3 (i), $|P^2(\mathbb{F}_2)| = 2^2 + 2 + 1 = 7$. La figura de $P^2(\mathbb{F}_2)$ es



Existen ahora 7 rectas en lugar de 6, cada una teniendo 3 puntos en lugar de 2. [El Teorema 3.3.3 (ii) afirma que el número de rectas es el mismo que el número de puntos.] Hemos adjuntado un punto "al infinito" a cada una de las viejas rectas, así forzando a rectas paralelas a unirse. El conjunto de puntos infinitos $\{a, b, c\}$ es una nueva recta, "la recta infinito".

Hemos dicho que si $g : V \to W$ es una transformación semilineal no singular, entonces g induce un isomorfismo $P(g) : P(V) \to P(W)$ por $[x] \mapsto [g(x)]$. Ahora probemos que salvo una excepción obvia, cada isomorfismo surge de esta forma.

Teorema 3.3.6 (Teorema Fundamental de la Geometría Proyectiva). *Si V y W son espacios de dimensión ≥ 3 sobre un campo K, entonces cada isomorfismo $\theta : P(V) \to P(W)$ tiene la forma $\theta = P(f)$ para alguna transformación semilineal no singular $f : V \to W$.*

Nota. Uno necesita que $\dim(V) \geq 3$, en caso contrario $P(V)$ y $P(W)$ son rectas proyectivas y cada función biyectiva entre ellas es un isomorfismo.

Demostración: Sean $\pi_V \colon V \to P(V)$ y $\pi_W \colon W \to P(W)$ los respectivos epimorfismos canónicos. Elijamos un hiperplano proyectivo Ω en $P(V)$. Dado que θ es isomorfismo, $\theta(\Omega)$ es un hiperplano proyectivo en $P(W)$. Sea H un hiperplano lineal en V con $\varphi(H^*) = \Omega$, y sea M un hiperplano lineal en W con $\psi(M^*) = \theta(\Omega)$. Tomemos $x \in V$ tal que $x \notin H$ y $y \in W$ tal que $[y] = \theta([x])$. Con esta notación, $P(g) = \pi_W g \pi_V^{-1}$ siempre que $g : V \to W$ sea una transformación semilineal.

Así que, existen espacios afines (A, α) y (B, β), donde $A = P(V) - \Omega$, $B = P(W) - \theta(\Omega)$, $\alpha = \pi_V t_x : H \to A$, y $\beta = \pi_W t_y : M \to B$ (ver la demostración del Teorema 3.3.4). La restricción $\theta|A$, denotada por f, es un isomorfismo afín $f : A \to B$ (porque θ es una isomorfismo). Por el Corolario 3.2.1, tenemos $f = \beta t_z g \alpha^{-1}$, donde $g : H \to M$ es una transformación semilineal no singular y $z = \beta^{-1} f \alpha(0)$. Pero,

$$\begin{aligned} z = \beta^{-1} f \alpha(0) &= t_{-y} \pi_W^{-1} f \pi_V t_x(0) \\ &= t_{-y} \pi_W^{-1} f[x] \\ &= t_{-y} \pi_W^{-1}[y] \\ &= y - y = 0 \end{aligned}$$

(recordemos que π_W es inyectiva sobre $M + y$). Por lo tanto, $t_z = t_0$ es la identidad y $\theta|A = f = \beta g \alpha^{-1}$.

Podemos escribir $V = \langle x \rangle \oplus H$, porque H es un hiperplano lineal y $x \notin H$. Definimos $\tilde{g} : V \to W$ por

$$\tilde{g}(\lambda x + u) = \sigma(\lambda)y + g(u),$$

donde $\lambda \in K$, $u \in H$, y $\sigma \in \mathrm{Aut}\,(K)$ es determinado por g. Tenemos que $\tilde{g}$ es una transformación semilineal que extiende a g y es no singular porque g lo es y $y \notin M$. Tenemos que

$$\begin{aligned} t_y \tilde{g} t_{-x}(\lambda x + u) &= t_y \tilde{g}((\lambda - 1)x + u) \\ &= \sigma(\lambda - 1)y + g(u) + y \\ &= \sigma(\lambda)y + g(u) \\ &= \tilde{g}(\lambda x + u). \end{aligned}$$

Por lo tanto,

$$\beta \tilde{g} \alpha^{-1} = \pi_W t_y \tilde{g} t_{-x} \pi_V^{-1} = \pi_W \tilde{g} \pi_V^{-1} = P(\tilde{g}).$$

Dado que $\theta | A = f = \beta g \alpha^{-1} = P(\tilde{g}) | A$, por el Teorema 3.3.5, obtenemos que $\theta = P(\tilde{g})$. □

Notación. Si V es un espacio vectorial sobre K, entonces $Sc(V)$ es el grupo de todas las transformaciones escalares distintas de cero $x \mapsto \lambda x$ con $\lambda \in K^*$.

Teorema 3.3.7. *Si* $\dim(V) \geq 3$*, entonces* $\Gamma L(V)/Sc(V)$ *es isomorfismo al grupo de todas los automorfismos de* $P(V)$*. Si* $\dim(V) = 2$*, entonces* $\Gamma L(V)/Sc(V)$ *es un subgrupo del grupo simétrico sobre* $P(V)$ *(en este caso, el grupo simétrico es el grupo de todos los automorfismo de* $P(V)$*).*

Demostración: Cuando $\dim(V) \geq 2$, hemos definido un homomorfismo π de $\Gamma L(V)$ en el grupo de isomorfismos de $P(V)$ por la relación $\pi(g) = P(g)$. Note que π es suprayectiva cuando $\dim(V) \geq 3$, por el Teorema 3.3.6. Calculemos $\ker(\pi)$. Suponga $g \in \Gamma L(V)$ con $P(g) = Id$, es decir, $[g(x)] = [x]$ para cada $[x] \in P(V)$. Así, $g(x) = \lambda_x x$ para toda $x \in V$, donde $\lambda_x \in K$.
La demostración estará completa, como en el Teorema 3.2.4, si probamos que λ_x no depende de x. Elegimos $w \in V$ tal que $\{x, w\}$ es independiente, pues $\dim(V) \geq 2$. Como g es aditivo,

$$g(x + w) = g(x) + g(w) = \lambda_x x + \lambda_w w;$$

por otro lado,

$$g(x+w) = \lambda_{x+w}(x+w).$$

Igualando los coeficientes, obtenemos que $\lambda_x = \lambda_{x+w} = \lambda_w$. Si $\mu \in K^*$, entonces $\{\mu x, w\}$ es independiente y $\lambda_{\mu x} = \lambda_w$. Por lo tanto, g es la transformación escalar $x \mapsto \lambda x$, donde λ es el valor común de los λ_x. □

Definición 3.3.5. *Para un espacio vectorial V, escribimos*

$$P\Gamma L(V) = \Gamma L(V)/Sc(V).$$

Cuando $V = V(n, K)$, escribiremos $P\Gamma L(V) = P\Gamma L(n, K)$.

Bajo la definición anterior, el Teorema 3.3.7 se puede enunciar a través del siguiente:

Corolario 3.3.1. *Si $n \geq 2$, $P\Gamma L(n+1, K)$ es el grupo de todos los automorfismos de $P^n(K)$. El grupo $P\Gamma L(2, K)$ es un subgrupo del grupo simétrico sobre $P^1(K)$.* ■

Ahora, diremos que una proyectividad es un isomorfimo de la forma $P(g)$, donde g es una transformación lineal no singular.

Definición 3.3.6. *Para un espacio vectorial V, escribiremos*

$$PGL(V) = GL(V)/Sc(V),$$

donde $V = V(n, K)$, escribiremos $PGL(V) = PGL(n, K)$.

Corolario 3.3.2. *Para toda $n \geq 1$, $PGL(n+1, K)$ es el grupo de todas proyectividades de $P^n(K)$.*

Demostración: Sea π_1 la restricción a $GL(n+1, K)$ del homomorfismo π del Teorema 3.3.7; así, $\pi_1(g) = P(g)$. La imagen de π_1 es el grupo de proyectividades y $\ker(\pi_1) = \ker(\pi) = Sc(V)$. □

Los grupos PSL discutidos en el capítulo anterior entran adecuadamente en el esquema notacional anterior y

$$PSL(V) \subset PGL(V) \subset P\Gamma L(V),$$

ya que $PSL(V) = SL(V)/Sc_1(V)$, donde $Sc_1(v) = Sc(V) \cap SL(V)$, y así $PSL(V) \cong SL(V)Sc(V)/Sc(V)$. Además, para toda $n \geq 1$, este grupo actúa fielmente sobre $P^n(K)$, por el Teorema 3.3.7, donde $V = V(n+1, K)$.

Hemos visto que $P\Gamma L(n+1, K)$, y cada uno de sus subgrupos, actúa fielmente sobre $P^n(K)$ para toda $n \geq 1$. Usaremos este hecho para probar que $P\Gamma L(2, \mathbb{F}_4) \cong S_5$ y $PGL(2, \mathbb{F}_4) \cong A_5$. Por el Teorema 3.2.6 y el Teorema 2.1.1, se sigue que $|P\Gamma L(2, \mathbb{F}_4)| = 120$ y $|PGL(2, \mathbb{F}_4)| = 60$. Dado que ambos grupos actúan fielmente sobre $P^1(4)$, un conjunto con 5 elementos, concluimos que cada uno es un subgrupo de S_5. Por esta razón, tenemos el siguiente resultado.

Teorema 3.3.8. *Para toda $n \geq 1$ y cada campo K, $PSL(n+1, K)$ (y cada uno de los grupos más grandes $PGL(n+1, K)$ y $P\Gamma L(n+1, K)$) actúa 2-transitivamente sobre $P^n(K)$.*

Demostración: Consideremos $P^n(K) = P(V)$, donde $V = V(n+1, K)$. Si $[x_1], [x_2]$ y $[y_1], [y_2]$ son parejas ordenadas de distintos puntos en $P(V)$, entonces $\{x_1, x_2\}$ y $\{y_1, y_2\}$ son subconjuntos independientes de V. Existen bases $\{x_1, x_2, \ldots, x_{n+1}\}$ y $\{y_1, y, \ldots, y_{n+1}\}$ de V, y una transformación lineal $g \in GL(n+1, K)$ tal que $g(x_i) = y_i$, para toda i. Entonces, $P(g)([x_i]) = [y_i]$ para $i = 1, 2$, y $PGL(n+1, K)$ actúa 2-transitivamente sobre $P^n(K)$.
Suponga $\det(g) = \lambda$. Definamos $h \in GL(n+1, K)$ por $h(x_1) = \lambda^{-1} y_1$ y $h(x_i) = y_i$ para $i \geq 2$. Entonces, $\det(h) = 1$, es decir, $h \in SL(n+1, K)$, y $P(h)([x_1]) = [\lambda^{-1} y_1] = [y_1]$ y $P(h)([x_2]) = [y_2]$. Así $PSL(n+1, K)$ también actúa 2-transitivamente sobre $P^n(K)$. □

Observación 3.3.2. *$PSL(n+1, K)$ actúa fielmente y transitivamente sobre el conjunto de todas las rectas proyectivas en $P^n(K)$.*

Una diferencia entre $P\Gamma L(2, K)$ y $P\Gamma L(m, K)$ para $m \geq 3$ es que el primero no es completamente el grupo de colineaciones de su espacio proyectivo. Pero veremos que este grupo tiene otras características de interés.

Definición 3.3.7. *Sea K un campo y $\sigma \in \mathrm{Aut}(K)$. Una **transformación fraccional semilineal** es una función $f : K \cup \{\infty\} \to K \cup \{\infty\}$ de la forma*

$$f(\lambda) = \frac{a\sigma(\lambda) + b}{c\sigma(\lambda) + d},$$

*donde $ad - bc \neq 0$ e ∞ es un nuevo símbolo tal que si $c = 0$, definimos $f(\infty) = \infty$; si $c \neq 0$, definimos $f(\infty) = ac^{-1}$; si $c\sigma(\lambda) + d = 0$, $f(\lambda) = \infty$. Si σ es la identidad, entonces f es llamada una **transformación fraccional lineal**.*

Definición 3.3.8. *Toda transformación fraccional semilineal forma un grupo bajo composición, denotado por $\Gamma LF(K)$; toda transformación fracciona lineal forma un subgrupo, denotado por $LF(K)$.*

Teorema 3.3.9. *Para cada campo K,*

$$P\Gamma L(2,K) \cong \Gamma LF(K) \quad y \quad PGL(2,K) \cong LF(K).$$

Demostración: Sea V un espacio vectorial de dimensión 2 sobre K. Al elegir un base de V y al usar el Teorema 3.2.6, tenemos que cada $h \in \Gamma L(2,K)$ tiene una única factorización en la forma $h = g\sigma_*$, donde $\sigma \in \mathrm{Aut}\,(K)$ y g es una transformación no singular de $GL(2,K)$ con matriz

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Definamos $\psi : \Gamma L(2,K) \to \Gamma LF(K)$ por $\psi(g\sigma_*)$ la cual está dada por la correspondencia $\lambda \mapsto (a\sigma(\lambda) + b)/(c\sigma(\lambda) + d)$. Entonces, ψ es un epimorfismo tal que $\ker(\psi)$ consiste de las matrices escalares distintas de cero. El segundo isomorfismo es justamente la restricción del primero. □

Vamos a desplegar los grupos que han surgido de un espacio vectorial V de dimensión m sobre K:

$$\begin{array}{ccccccc} SL(V) & \subset & GL(V) & \subset & \Gamma L(V) & \subset & \mathrm{Aut}\,(V) \\ \cup & & \cup & & & & \cup \\ Sc_1(V) & \subset & Sc(V) & & & & \mathrm{Aff}\,(V) \\ & & & & & & \cup \\ & & & & & & Tr(V); \end{array}$$

$$PSL(V) \subset PGL(V) \subset P\Gamma L(V).$$

Recordemos que el producto semidirecto de un grupo A por un grupo B es denotado por $A \rtimes B$ y que $\mathrm{Aut}\,(K)$ es el grupo de automorfismos de un campo K [no debemos confundir con $\mathrm{Aut}\,(V)$ cuando V es unidimensional].

Las siguientes relaciones establecen los isomorfismos que relacionan estos grupos:

$$\begin{aligned} Z_1(m,K) &\cong Sc_1(V) = Sc(V) \cap SL(V); \\ SL(m,K) &\cong SL(V); \\ GL(m,K) &\cong GL(V) \cong SL(V) \rtimes K^*; \\ \Gamma L(m,K) &\cong \Gamma L(V) \cong GL(V) \rtimes \mathrm{Aut}\,(K); \\ \mathrm{Aut}\,(m,K) &\cong \mathrm{Aut}\,(V) \cong \Gamma L(V) \rtimes \mathrm{Tr}\,(V); \\ \mathrm{Aff}\,(m,K) &\cong \mathrm{Aff}\,(V) \cong GL(V) \rtimes \mathrm{Tr}\,(V). \end{aligned}$$

Cuando $\dim V = 2$, existen otros dos isomorfismos:

$$\begin{aligned} P\Gamma L(V) &\cong P\Gamma L(2,K) \cong \Gamma LF(K); \\ PGL(V) &\cong PGL(2,K) \cong LF(K). \end{aligned}$$

Suponga que (X, ρ_1) es un G-conjunto, (Y, ρ_2) es un H-conjunto, y $\psi : G \to H$ es un isomorfismo; ¿Cuándo es razonable identificar estos conjuntos equipados con sus respectivas acciones de grupos? Recordemos (Observación 1.3.1) que una función biyectiva $\theta : X \to Y$ induce un homomorfismo θ_* de grupos simétricos $S_X \to S_Y$ dada por la correspondencia $\alpha \mapsto \theta\alpha\theta^{-1}$. Notemos que la acción es justamente un homomorfismo de grupos simétricos. De aquí que, podemos ver Y como un G-conjunto de dos maneras distintas, vía $\theta_*\rho_1$ o $\rho_2\psi$.

$$\begin{array}{ccc} G & \xrightarrow[\rho_1]{} & S_X \\ {\scriptstyle\psi}\downarrow & & \downarrow{\scriptstyle\theta_*} \\ H & \xrightarrow[\rho_2]{} & S_Y \end{array}$$

Decimos que el G-conjunto X es **isomorfo** al H-conjunto Y si el diagrama anterior es conmutativo.
Así que, en este caso, para cada $g \in G$, esto dice que $\theta\rho_1(g)\theta^{-1} = \theta_*\rho_1(g) = \rho_2\psi(g)$; para cada $g \in G$ y $x \in X$, esto dice

$$\theta\rho_1(g)\theta^{-1}(x) = \rho_2\psi(g)(x).$$

(Cuando $G = H$ y la función ψ es la identidad, esta definición se vuelve la de G-isomorfismo).

Consideremos el $\Gamma LF(K)$-conjunto $K \cup \{\infty\}$ con acción natural inducida por la correspondencia $\lambda \mapsto (a\sigma(\lambda) + b)/(c\sigma(\lambda) + d)$. Consideremos también el $P\Gamma L(V)$-conjunto $P^1(K)$, donde V es un espacio vectorial de dimensión 2 sobre K, con acción

$P(h)[v] = [h(v)]$. Conservaremos las notaciones de la demostración del Teorema 3.3.9. Eligiendo una base de V, expresamos a cada $[v] \in P^1(K)$ en coordenadas homogéneas $[v] = [\lambda, \mu]$, y a cada transformación semilineal no singular h en la forma $h = g\sigma_*$, donde g tiene por matriz

$$g = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

y

$$\sigma_*[\lambda, \mu] = [\sigma(\lambda), \sigma(\mu)].$$

La acción de $P\Gamma L(2, K)$ sobre $P^1(K)$ es así dada por

$$P(h) \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \sigma(\lambda) \\ \sigma(\mu) \end{bmatrix} = \begin{bmatrix} a\sigma(\lambda) + b\sigma(\mu) \\ c\sigma(\lambda) + d\sigma(\mu) \end{bmatrix}.$$

Ahora definimos $\theta : P^1(K) \to K \cup \{\infty\}$ por $\theta[0, 1] = \infty$ y $\theta[\lambda, 1] = \lambda$. Si $\lambda \in K$ y $c\sigma(\lambda) + d \neq 0$, entonces

$$\begin{aligned} \theta g \sigma_* \theta^{-1}(\lambda) &= \theta g \sigma_*[\lambda, 1] \\ &= \theta[a\sigma(\lambda) + b, c\sigma(\lambda) + d)] \\ &= \theta[(a\sigma(\lambda) + b)/(c\sigma(\lambda) + d), 1] \\ &= (a\sigma(\lambda) + b)/(c\sigma(\lambda) + d). \end{aligned}$$

Una ecuación similar se tiene cuando $c\sigma(\lambda) + d = 0$ o cuando λ es reemplazado por ∞. Por lo tanto, podemos identificar a $P^1(K)$ y $K \cup \{\infty\}$ como conjunto con acción de grupos.

Teorema 3.3.10. *Para cada campo K, $P^1(K)$ es un $PGL(2, K)$-conjunto fiel simplemente 3-transitivo.*

Demostración: Como en el Teorema 3.3.9 y la discusión anterior, podemos considerar $PGL(2, K) \cong LF(K)$ actuando sobre $K \cup \{\infty\}$ como una transformación fraccional lineal. Considerando a K como un espacio vectorial unidimensional sobre sí mismo, tenemos que $\mathrm{Aff}\,(K) = \{\lambda \mapsto a\lambda + b\}$ [el subgrupo de $LF(K)$ consistente de todos los "numeradores"] es el estabilizador de ∞. Pero, por el Teorema 3.2.2, K es un $\mathrm{Aff}\,(K)$-conjunto simplemente 2-transitivamente; así que, con la Observación 3.1.1 y el Corolario 3.1.1 (v) obtenemos lo desado una vez que se haya probado que $LF(K)$ actúa transitivamente sobre $K \cup \{\infty\}$. Si $\lambda \in K$, entonces $f(x) = x + \lambda$, envia 0 a λ; si $\lambda = \infty$, entonces $f(x) = 1/x$ envia 0 a ∞. □

Si $K = \mathbb{F}_q$, entonces, con el Teorema 3.1.1 y el teorema anterior, obtenemos otra demostración de que $|PGL(2, \mathbb{F}_q)| = (q+1)q(q-1)$. Observemos también que $P\Gamma(2, \mathbb{F}_q)$ actúa 3-transitivamente sobre $P^1(\mathbb{F}_q)$ (porque su subgrupo PGL lo hace), pero la acción de $P\Gamma L$ no es simple.

Ahora exhibamos una segunda familia de G-conjuntos simplemente 3-transitivos. Si $h \in \Gamma LF(q)$, entonces $h(\lambda) = (a\sigma(\lambda) + b)/(c\sigma(\lambda) + d)$, donde $\sigma \in \mathrm{Aut}\,(\mathbb{F}_q)$ y $ad - bc \neq 0$. Multiplicando numerador y denominador por $\mu \in \mathbb{F}_q^*$, tenemos que h no cambia, pero su "determinante" cambia por $\mu^2(ad-bc)$. Si q es una potencia de 2, entonces cada elemento en $\mathbb{F}_q$ es un cuadrado; además, si q es una potencia de un primo impar p y si α es un elemento primitivo de $\mathbb{F}_q$, entonces los cuadrados no nulos forman un subgrupo de índice 2 en $\mathbb{F}_q^*$, a saber, $\langle \alpha^2 \rangle$. Así, en el segundo caso, tiene sentido decir si $\det h$ es o no es un cuadrado.

Un segundo ingrediente en la definición que sigue, y que es un análogo a la conjugación compleja, es la de un automorfismo σ de $\mathbb{F}_q$ teniendo orden 2, en la cual σ existe (y es único) cuando $q = p^{2n}$, en cuyo caso $\sigma(\lambda) = \lambda^{p^n}$.

Definición 3.3.9. *Sea $q = p^{2n}$, donde p es un primo impar, y sea $\sigma \in \mathrm{Aut}\,(\mathbb{F}_q)$ teniendo orden 2 [es decir, $\sigma(\lambda) = \lambda^{p^n}$]. Definamos $Sh(\mathbb{F}_q)$ como el subconjunto $A \cup B$ de $\Gamma LF(\mathbb{F}_q)$, donde*

$$A = \{h : \lambda \mapsto (a\lambda + b)/(c\lambda + d) \mid ad - bc \text{ es un cuadrado } \}$$

y

$$B = \{h : \lambda \mapsto (a\lambda + b)/(c\lambda + d) \mid ad - bc \text{ no es un cuadrado } \}.$$

Tenemos que $Sh(\mathbb{F}_q)$ es un subgrupo de $\Gamma LF(\mathbb{F}_q)$, con A subgrupo de índice 2 en $Sh(\mathbb{F}_q)$ y B la otra clase lateral. [La Observación 3.3.3 prueba que $Sh(\mathbb{F}_q)$ no es el producto semidirecto de A por $\mathbb{Z}_2$.] Puesto que $Sh(\mathbb{F}_q)$ es un grupo de transformaciones de fracciones lineales, tenemos que $Sh(\mathbb{F}_q)$ actúa fielmente sobre $P^1(\mathbb{F}_q)$.

Teorema 3.3.11. *Si p es un primo impar y $q = p^{2n}$, entonces $P^1(\mathbb{F}_q)$ es un $Sh(\mathbb{F}_q)$-conjunto fiel simplemente 3-transitivo.*

Demostración: Sea $K = \mathbb{F}_q$ e identifiquemos a $P^1(\mathbb{F}_q)$ con $K \cup \{\infty\}$. Escribamos G en lugar de $Sh(\mathbb{F}_q)$. Ahora $G_\infty = A_\infty \cup B_\infty$, donde

$$A_\infty = \{h : \lambda \mapsto a\lambda + b \mid a \in (K^*)^2\}$$

y

$$B_\infty = \{h : \lambda \mapsto a\lambda + b \mid a \notin (K^*)^2\}.$$

Si a y b son elementos distintos de K y $a-b$ es un cuadrado, definimos $h \in A_\infty$ por $h(\lambda) = (a-b)\lambda + b$; en caso de que $a-b$ no sea un cuadrado, definimos $h \in B_\infty$ por $h(\lambda) = (a-b)\sigma(\lambda) + b$. Entonces, en cualquier caso, se tiene que $h(0) = b$ y $h(1) = a$. Se sigue que G_∞ actúa doblemente transitivamente sobre K. Pero en A_∞ o B_∞, existen q elecciones para b y $\frac{1}{2}(q-1)$ elecciones para a, así que $|G_\infty| = q(q-1)$ y, por el Teorema 3.1.1, la acción de G_∞ sobre K es simple. Finalmente, G actúa transitivamente sobre $K \cup \{\infty\}$, por $\lambda \mapsto -1/\lambda$ el cual está en G (el signo negativo da determinante 1, y 1 es un cuadrado) y $-1/\lambda$ intercambia 0 e ∞. Concluimos, por el Corolario 3.1.1 (v), que $G = Sh(\mathbb{F}_q)$ actúa fiel y simplemente 3-transitivamente sobre $K \cup \{\infty\}$. □

Zassenhaus (1936) probó que la acción de $PGL(2, \mathbb{F}_q)$ y $Sh(\mathbb{F}_q)$ sobre $P^1(\mathbb{F}_q)$ son los únicos G-conjuntos fieles simplemente 3-transitivos. La primera familia de grupos está definida para toda portencia q de un primo; la segunda está definida para toda potencia de exponente par de primos impares. No es obvio que $PGL(2, \mathbb{F}_q) \ncong Sh(\mathbb{F}_q)$ cuando $q = p^{2n}$ y p impar, lo cual es cierto. Aquí está un pequeño ejemplo de este hecho.

Definición 3.3.10. *El grupo $Sh(\mathbb{F}_9)$ es usualmente denotado por M_{10} y es llamado* ***grupo de Mathieu*** *de grado* 10*.*

Teorema 3.3.12. *$PGL(2, \mathbb{F}_9)$ y M_{10} son grupos de orden* 720 *no isomorfos, cada uno de los cuales actúa simplemente* 3*-transitivamente sobre $P^1(\mathbb{F}_9)$.*

Demostración: Ya hemos visto que cada grupo actúa sobre $P^1(\mathbb{F}_9) = \mathbb{F}_9 \cup \{\infty\}$ simplemente y 3-transitivamente, de donde que el orden de cada grupo es $10 \cdot 9 \cdot 8 = 720 = 16 \cdot 45$. Sea $G = LF(9) \cong PGL(2, \mathbb{F}_9)$ actuando sobre $\mathbb{F}_9 \cup \{\infty\}$. Ahora, el doble estabilizador $G_{\infty,0} = \{h \mid \lambda \mapsto a\lambda \text{ con } a \neq 0\} \cong \mathbb{F}_9^* \cong \mathbb{Z}_8$; en efecto, un generador de $G_{\infty,0}$ es $g : \lambda \mapsto \alpha\lambda$, donde α es un elemento primitivo de $\mathbb{F}_9$. Si $\tau(\lambda) = \lambda^{-1}$, entonces τ es un elemento de orden 2 en G tal que $\tau g \tau^{-1} = g^{-1}$. Siguiéndose que $\langle G_{\infty,0}, \tau \rangle$ es el grupo diédrico de orden 16 y es un 2-subgrupo Sylow de G.
Escribimos $H = M_{10}$, el cual actúa sobre $\mathbb{F}_9 \cup \{\infty\}$. Dado que $q = 3^2$, el automorfismo σ es justamente $\sigma(\lambda) = \lambda^3$. Ahora, el doble estabilizador está dado como

$$\begin{aligned} H_{\infty,0} &= A_{\infty,0} \cup B_{\infty,0} \\ &= \{h : \lambda \mapsto a^2\lambda \text{ con } a \neq 0\} \cup \{h : \lambda \mapsto a\lambda^3 \text{ con } a \text{ no siendo cuadrado}\}. \end{aligned}$$

Este grupo es no abeliano de orden 8 teniendo un sólo elemento de orden 2, de aquí que $H_{\infty,0} \cong Q$, el grupo de los cuaternios. Dado que el grupo diédrico de orden 16 no tiene subgrupos cuaternionicos, se sigue que los 2-subgrupos de Sylow de G y H no son isomorfos, por lo que G y H no son isomorfos. □

Observación 3.3.3.

1. $(M_{10})_\infty$ es un subgrupo de orden 72 teniendo un 3-subgrupo Sylow normal. Por tanto $(M_{10})_\infty$ es el producto semidirecto de $\mathbb{F}_3 \times \mathbb{F}_3$ por Q.

2. Existe exactamente 8 elementos en $(M_{10})_\infty$ de orden 3 y estos son conjugados a algún otro de $(M_{10})_\infty$.

3. M_{10} no es producto semidirecto de A por $\mathbb{Z}_2$.

4. $M_{10} = \langle \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5 \rangle$, donde $\sigma_1(\lambda) = -\lambda^{-1}$, $\sigma_2(\lambda) = \lambda + 1$, $\sigma_3(\lambda) = \lambda + \alpha$, $\sigma_4(\lambda) = \alpha^2\lambda$, y $\sigma_5(\lambda) = \alpha\lambda^3$.

5. $\mathbb{F}_9$ puede ser considerado como un espacio vectorial de dimensión dos sobre $\mathbb{F}_3$ con base $\{1, \alpha\}$. Verificando las siguientes coordenadas [alguna raíz de un polinomio cuadrático irreducible con coeficientes en $\mathbb{F}_3$ sirve como α; además supongamos $\alpha^2 + \alpha = 1$]:

$$\begin{aligned} 1 &= (1,0); & \alpha^4 &= (-1,0); \\ \alpha &= (0,1); & \alpha^5 &= (0,-1); \\ \alpha^2 &= (1,-1); & \alpha^6 &= (-1,1); \\ \alpha^3 &= (-1,-1); & \alpha^7 &= (1,1). \end{aligned}$$

6. M_{10} consiste de permutaciones pares de $P^1(\mathbb{F}_9)$.

Finalizamos esta sección con una segunda prueba de la simplicidad de los $PSL's$. Recordemos la definición: Sea $V = V(n, \mathbb{F}_q)$ y sea $f : V \to \mathbb{F}_q$ un funcional. Si $x \in \ker(f)$, entonces la transvección $T_{f,x}$ es una función $V \to V$ definida por

$$T_{f,x}(\nu) = \nu + f(\nu)x,$$

para todo $\nu \in V$. Digamos que cada transvección pertenece a $SL(V)$ y que el conjunto de estas genera a $SL(V)$.

Teorema 3.3.13. *$PSL(n, \mathbb{F}_q)$ es simple si $(n, \mathbb{F}_q) \neq (2, \mathbb{F}_2), (2, \mathbb{F}_3)$.*

Demostración: Usaremos el Teorema de Iwasawa (Teorema 3.1.18). Si $G = PSL(n, \mathbb{F}_q)$ y $P^{n-1}(\mathbb{F}_q) = P(V)$, donde $V = V(n, \mathbb{F}_q)$, entonces sabemos que $P(V)$ es fiel 2-transitivo, y por lo tanto G-conjunto primitivo. Así, la condición (i) del Teorema 3.1.18 se cumple.
Sea $x \in V$. Definimos H el subconjunto del estabilizador $G_{[x]}$ por

$$H = \{P(T_{f,x}) \mid f \text{ es un funcional con } f(x) = 0\}.$$

Recordemos la Observación 2.3.1: Si S es una transformación lineal no singular [la cual la elegimos estando en $SL(V)$],

$$ST_{f,x}S^{-1} = T_{fS^{-1},Sx}.$$

Aplicando P a la fórmula $T_{f,x}T_{g,x} = T_{f+g,x}$ y usando la Observación 2.3.1, se tiene que H es un subgrupo abeliano de $G_{[x]}$.
Ahora, $P(S) \in G_{[x]}$ si y sólo si $Sx = \lambda x$ para algún $\lambda \in \mathbb{F}_q$. Pero $T_{g,\lambda x} = T_{\lambda g,x}$ (Observación 2.3.1) y esto prueba que H es un subgrupo normal de $G_{[x]}$.

Dado que las transvecciones generan a $SL(V)$, será suficiente la condición (3) para probar que $P(T_{g,y})$ es el conjugado de algún $P(T_{f,x})$. Elijamos $S \in SL(V)$ con $Sx = y$. Entonces $P(S)P(T_{f,x})P(S)^{-1} = P(T_{fs^{-1},y})$. Pero, como f varia sobre todos los funcionales con $f(x) = 0$, fS^{-1} varía sobre todos los funcionales g tales que $g(y) = 0$.

Resta probar que G es perfecto, es decir, $G = G'$. Suponga que tenemos alguna transvección T en el conmutador G': así que $T = [A, B]$ para algunos $A, B \in SL(V)$. Para alguna otra transvección $T\prime$, el Corolario 2.3.1 establece que T y $T\prime$ son conjugadas en $GL(V)$; digamos que $T\prime = T^U$ para algún $U \in GL(V)$. Ya que la conjugación por U es un homomorfismo, $T\prime = [A, B]^U = [A^U, B^U]$; dado que $SL(V) \lhd GL(V)$, ambos A^U y B^U pertenecen a $SL(V)$. Por lo tanto, cada transvección sería un conmutador; ya que las transvecciones generan a $SL(V)$, tendríamos que $SL(V)$ y su imagen canónica homomorfa $PSL(V) = G$ son perfectos.

Así que, suponga $n \geq 3$, y sea $\{e_1, \ldots, e_n\}$ una base de V. Sea T la transvección $T = T_{f,x}$ con $x = -e_2 - e_1$ y f el funcional seleccionando la tercer coordenada: $f(\sum \lambda_i e_i) = \lambda_3$ (es decir, $T(e_i) = e_i$ para $i \neq 3$ y $T(e_3) = e_3 - e_2 - e_1$). Utilizando el funcional f anterior, definimos $A = T_{f,-e_1}$ (o sea, $Ae_i = e_i$ para $i \neq 3$ y $Ae_3 = e_3 - e_1$), y definimos B por $B(e_1) = -e_2$, $B(e_2) = e_1$, y $B(e_i) = e_i$ para $i \geq 3$. Notemos que $A, B \in SL(V)$, y satisfacen que $T = [A, B] = ABA^{-1}B^{-1}$.
Finalmente, suponga $n = 2$ y $q > 3$. Entonces, existe $\lambda \in \mathbb{F}_q$ con $\lambda^2 \neq 1$. Pero

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda^{-1} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \lambda^{-1} & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \lambda^2 - 1 \\ 0 & 1 \end{bmatrix},$$

así la transvección $B_{12}(\lambda^2 - 1)$ es exhibida como un conmutador.
Así pues, todas las condiciones del Teorema 3.1.18 han sido verificadas, y concluimos que $PSL(n, \mathbb{F}_q)$ es simple salvo dos excepciones. □

Observación 3.3.4. Notemos que la demostración del teorema anterior es válida para cualquier campo K, salvo las excepciones consideradas.

Capítulo 4

Grupos de Mathieu

En este capítulo daremos una aplicación del grupo de Mathieu M_{12} a la teoría de códigos. Probaremos primeramente la simplicidad de los grupos de Mathieu.

4.1. Simplicidad de los Grupos de Mathieu

Ya hemos visto algunos resultados de G-conjuntos doblemente y triplemente transitivos. Generalmente, a nivel de ejemplos, uno considera a los G-conjuntos obtenidos de grupos simétricos o alternantes como triviales, no porque sea fácil exhibirlos, sino porque la pregunta en sí sería: ¿Cómo debe ser la transitividad de un subgrupo de S_n antes de que uno pueda concluir de que o bien el subgrupo sea S_n ó A_n? Los grupos de Mathieu son los otros grupos G que nos proveen ejemplos de G-conjuntos fieles y 4-transitivos (dos de ellos son realmente 5-transitivos). Y los analizaremos en esta sección.

En 1873, Jordan probó que no existen G-conjuntos simplemente 6-transitivos (a menos de que G sea el grupo simétrico ó alternante). En 1981, la clasificación de grupos finitos simples no abelianos fue completada: uno de tales grupos es o bien un grupo alternante, un grupo de "tipo Lie" (grupos proyectivos unimodulares son lo ejemplos más sencillo), ó uno de los veintiseis grupos "esporádicos" (cinco de los cuales son los grupos de Mathieu). Esta clasificación puede ser usada para probar que no existen G-conjuntos no triviales 6-transitivos; más aún, se tiene una clasificación de todos los G-conjuntos fieles doblemente transitivos, así como prueba de que los grupos de Mathieu son los únicos grupos fieles 4-transitivos.

Existe una técnica más sencilla para la transitividad en baja dimensión. Si X es un G-conjunto transitivo y $x \in X$, entonces el G_x-conjunto $X \setminus \{x\}$ es (simplemente)

$(k-1)$-transitivo si y sólo si X es un G-conjunto (simplemente) k-transitivo (Observación 3.1.1 y Corolario 3.1.1). ¿Podemos revertir este proceso, empezando con los conjuntos G_x y construyendo G?

G siempre será un grupo finito y todos los G-conjuntos en esta sección son fieles, y llamaremos a los grupos G, de aquí en adelante, por **grupos de permutaciones**. Al final, llega uno a sucumbir a la irresistible urgencia de aplicar a los grupos estos adjetivos hasta ahora reservados a G-conjuntos, por ejemplo, hableremos de un G grupo transitivo múltiple de grado n para significar que existe algún conjunto $X = \{x_1, \ldots, x_n\}$ que es un G-conjunto transitivo múltiple.

Definición 4.1.1. *Sea G un grupo de permutaciones sobre un conjunto no vacío X y sea $\tilde{X} = X \cup \{\infty\}$, donde ∞ es un objeto que no está en X. Un grupo $\tilde{G}$ de permutaciones transitivo sobre $\tilde{X}$ es una **extensión transitiva** de G si contiene a G como subgrupo y el estabilizador $\tilde{G}_\infty$ es G.*

Observación 4.1.1. Las extensiones transitivas pueden no existir. Por ejemplo, el 4-grupo $\mathbb{V}$ no tiene extensión transitiva. Realmente, las extensiones transitivas raramente existen. Sin embargo, si tal extensión existe, y X es un G-conjunto k-transitivo, entonces $\tilde{X}$ es un $\tilde{G}$-conjunto $(k+1)$-transitivo.

Teorema 4.1.1 (Witt, 1938). *Sea G un grupo de permutaciones múltiple transitivo actuando sobre un conjunto X. Supongamos que existe un elemento $\infty \notin X$, una permutación h de $\tilde{X} = X \cup \{\infty\}$, un elemento $x \in X$ y un elemento $g \in G$ tales que:*

(*i*) $g \notin G_x$;

(*ii*) $h(\infty) \in X$;

(*iii*) $h^2 \in G$ *y* $(gh)^3 \in G$;

(*iv*) $hG_xh = G_x$.

Entonces $\tilde{G} = \langle G, h\rangle$ es una extensión transitiva de G.

Demostración: Como G actúa transitivamente sobre X, la condición (*ii*) prueba que $\tilde{G}$ actúa transitivamente sobre $\tilde{X}$. Probaremos que $\tilde{G} = G \cup GhG$, donde $\tilde{G} = \langle G, h\rangle$, como lo predice el Teorema 3.1.5, para concluir que $\tilde{G}_\infty = G$ dado que todo elemento en G fija a ∞ y cada elemento en la clase lateral doble GhG mueve a ∞.

Veamos que $G \cup GhG$ es grupo, por lo que necesitamos ver que $G \cup GhG$ es cerrado bajo multiplicación. Ya que $GG = G$, es suficiente probar que $(GhG)(GhG) = GhGhG \subset G \cup GhG$, y esto se tiene si probamos que $hGh \subset G \cup GhG$.
Dado que G actúa múltiplemente transitivamente sobre X, el Teorema 3.1.5 establece que $G = G_x \cup G_x g G_x$ (ya que $g \notin G_x$). Notemos que la relación $h^2 = \gamma_1 \in G$ implica que $h\gamma_1^{-1} = h^{-1} = \gamma_1^{-1}h$ y que $(gh)^3 = \gamma_2 \in G$ implica que $hgh = g^{-1}h^{-1}g^{-1}\gamma_2$. Ahora

$$\begin{aligned}
hGh &= h(G_x \cup G_x g G_x)h \\
&= hG_xh \cup hG_xgG_xh \\
&= hG_xh \cup (hG_xh)h^{-1}gh^{-1}(hG_xh) \\
&= G_x \cup G_xh^{-1}gh^{-1}G_x \qquad [\text{condición } (iv)] \\
&= G_x \cup G_x(\gamma_1^{-1}h)g(h\gamma_1^{-1})G_x \\
&= G_x \cup G_x\gamma_1^{-1}(g^{-1}h^{-1}g^{-1}\gamma_2)\gamma_1^{-1}G_x \\
&\subset G \cup Gh^{-1}G \\
&= G \cup G\gamma_1^{-1}hG = G \cup GhG.
\end{aligned}$$

□

Las condiciones en el teorema anterior nos da información sobre la estructura cíclica de la permutación h. Si $h(\infty) = a \in X$, entonces la relación $h^2 \in G = \tilde{G}_\infty$ implica que $h(a) = h^2(\infty) = \infty$. Por lo tanto, $h = (\infty, a)h\prime$, donde $h\prime \in G_a$ es disjunta de (∞, a). Similarmente, uno puede ver que gh tiene un factor que involucra a ∞ el cual es un 3-ciclo.

Existen cinco grupos más de Mathieu además de M_{10}, normalmente, M_{11}, M_{12}, M_{22}, M_{23}, M_{24}; los subíndices indican el grado de cada uno de los grupos en la representación usual de un grupo de permutaciones. Presentamos los resultados que nos permitan construir extensiones transitivas por el método establecido en el Teorema 4.1.1.

Teorema 4.1.2. *Existe un grupo simplemente 4-transitivo M_{11} de grado 11 y orden $7920 = 11 \cdot 10 \cdot 9 \cdot 8 = 2^4 \cdot 3^2 \cdot 5 \cdot 11$ tal que el estabilizador de un punto es M_{10}.*

Demostración: Sabemos que M_{10} actúa (simplemente) 3-transitivamente sobre $X = \mathbb{F}_9 \cup \{\infty\}$. Construyamos una extensión transitiva de M_{10} actuando sobre $\tilde{X} = X \cup \{\omega\}$. Para α elemento primitivo de $\mathbb{F}_9$, definimos

$$x = \infty,$$
$$g = (0,\infty)(\alpha,\alpha^7)(\alpha^2,\alpha^6)(\alpha^3,\alpha^5)$$
$$y$$
$$h = (\omega,\infty)(\alpha,\alpha^2)(\alpha^3,\alpha^7)(\alpha^5,\alpha^6).$$

Notemos que $g \in M_{10}$, pues $g(\lambda) = \lambda^{-1}$ y $\det(g) = -1 = \alpha^4$, el cual es un cuadrado en $\mathbb{F}_9$. Usando la Observación 3.3.3, se tiene que $(\alpha,\alpha^2)(\alpha^3,\alpha^7)(\alpha^5,\alpha^6)$ es la permutación de $\mathbb{F}_9$ dada por $\lambda \mapsto \alpha^2\lambda + \alpha\lambda^3$.

Es evidente que $g \notin (M_{10})_\infty$, porque $g(\infty) = 0$, y que $h(\omega) = \infty \in X$. Ahora $h^2 = 1$, y $gh = (\omega, 0, \infty)(\alpha,\alpha^6,\alpha^3)(\alpha^2,\alpha^7,\alpha^5)$ tiene orden 3. Para satisfacer la última condición del Teorema 4.1.1, obsérvese que si $f \in (M_{10})_\infty$, entonces

$$hfh(\infty) = hf(\omega) = h(\omega) = \infty.$$

Así, $h(M_{10})_\infty h = (M_{10})_\infty$ si cada $hfh \in M_{10}$. Escribiendo $(M_{10})_\infty = A_\infty \cup B_\infty$ (como en la definición de M_{10}), o bien $f = \alpha^{2i}\lambda + b$ o $f = \alpha^{2i+1}\lambda^3 + b$, donde $i \geq 0$ y $b \in \mathbb{F}_9$. En el primer caso (calculando con $\alpha^2\lambda + \alpha\lambda^3$),

$$hfh(\lambda) = (\alpha^{2i+4} + \alpha^{6i+4})\lambda + (\alpha^{2i+3} + \alpha^{6i+7})\lambda^3 + \alpha^2 b + \alpha b^3.$$

Los coeficientes de λ y λ^3 son $\alpha^{2i+4}(1+\alpha^{4i})$ y $\alpha^{2i+3}(1+\alpha^{4i+4})$, respectivamente. Cuando i es par, el segundo coeficiente es 0 y el primer coeficiente es $2\alpha^{2i+4}$; pero $2 = -1 = \alpha^4$, así que este coeficiente es un cuadrado y $hfh \in A_\infty$. Cuando i es impar, el primer coeficiente es 0 y el segundo coeficiente es $2\alpha^{2i+3} = \alpha^{2i+7}$ (el cual no es un cuadrado), de donde $hfh \in B_\infty$. El segundo caso ($f(\lambda) = \alpha^{2i+1}\lambda^3 + b$) es similar, y se tiene que

$$hfh(\lambda) = \alpha^{2i+6}(1+\alpha^{4i})\lambda + \alpha^{2i+1}(1+\alpha^{4i+4})\lambda^3 + \alpha^2 b + \alpha b^3,$$

una expresión de la misma forma así como la tratada en el primer caso.
Se sigue del Corolario 3.1.5, (v), que $M_{11} = \langle M_{10}, h\rangle$ actúa simplemente 4-transitivamente sobre $\tilde{X}$, pues M_{10} actúa simplemente 3-transitivamente sobre X, y así $|M_{11}| = 11 \cdot 10 \cdot 9 \cdot 8 = 7920$. □

El procedimiento puede ser repetido. De nuevo, la parte difícil es el de descubrir una buena permutación para adjuntar.

Teorema 4.1.3. *Existe un grupo M_{12} simplemente 5-transitivo de grado 12 orden $12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95,040 = 2^6 \cdot 3^3 \cdot 5 \cdot 11$ tal que el estabilizador de un punto es M_{11}.*

Demostración: Sabemos que M_{11} actúa (simplemente) 4-transitivamente sobre $Y = \{\mathbb{F}_9, \infty, \omega\}$; construiremos una extensión transitiva de M_{11} actuando sobre $\tilde{Y} = Y \cup \{\Omega\}$. Para α elemento primitivo de $\mathbb{F}_9$, definimos

$$\begin{aligned} & x = \omega, \\ & h = (\omega, \infty)(\alpha, \alpha^2)(\alpha^3, \alpha^7)(\alpha^5, \alpha^6) \\ y & \\ & k = (\omega, \Omega)(\alpha, \alpha^3)(\alpha^2, \alpha^6)(\alpha^5, \alpha^7). \end{aligned}$$

Notemos que $h \in M_{11}$ es el mismo h que se propone en el teorema anterior y aquel "factor" $(\alpha, \alpha^3)(\alpha^2, \alpha^6)(\alpha^5, \alpha^7)$ de k es la permutación de $\mathbb{F}_9$ dada por $\lambda \mapsto \lambda^3$. Claramente $k(\Omega) = w \in Y$ y $h \notin (M_{11})_\omega = M_{10}$. Luego, $k^2 = 1$ y $hk = (\omega, \Omega, \infty)(\alpha, \alpha^7, \alpha^6)(\alpha^2, \alpha^5, \alpha^3)$ tiene orden 3. Para satisfacer la última condición del Teorema 4.1.1, observemos primero que si $f \in (M_{11})_\omega = M_{10} = A \cup B$, entonces kfk también fija a ω. Finalmente $kfk \in M_{11}$: si $f(\lambda) = (a\lambda + b)/(c\lambda + d) \in A$, entonces $kfk(\lambda) = (a^3\lambda + b^3)/(c^3\lambda + d^3)$ tiene determinante $a^3d^3 - b^3c^3 = (ad - bc)^3$, el cual es un cuadrado porque $ad - bc$ lo es; un argumento similar se tiene cuando $f \in B$. Así, $kM_{10}k = M_{10}$.

Se sigue que $M_{12} = \langle M_{11}, k \rangle$ actúa simplemente 5-transitivamente sobre $\tilde{Y}$, dado que M_{11} actúa simplemente 4-transitivamente sobre Y, y $|M_{12}| = 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95,040$. □

Observación 4.1.2. Sea $W = \{g \in M_{12} \mid g(\{\infty, \omega, \Omega\}) = \{\infty, \omega, \Omega\}\}$. Entonces, existe un homomorfismo de W sobre S_3 con núcleo $(M_{12})_{\infty,\omega,\Omega}$. Por lo tanto, $|W| = 6 \cdot 72$.

Observación 4.1.3. Si W está dado como en el Observación 4.1.2, entonces el grupo Aut $(2, \mathbb{F}_3)$, de todos los automorfismos afines de $\mathbb{F}_9$ visto como un plano afín sobre $\mathbb{F}_9$, es isomorfo al subgrupo W de M_{12} .

El teorema de Jordan mencionado al inicio de la sección establece que S_5, S_6, A_7 y M_{11} son los únicos grupos simplemente 5-transitivos, y que S_k, S_{k+1} y A_{k+2} son los únicos grupos simplemente k-transitivos para $k \geq 6$. Zassenhaus clasificó todos los grupos simplemente 3-transitivos; tales grupos son o bien $PGL(2, \mathbb{F}_q)$ o $Sh(\mathbb{F}_p^{2n})$ para p impar. Mientras que Thompson completó la clasificación de los grupos simplemente 2-transitivos (estos son ciertos grupos de Frobenius). La clasificación de

todos los grupos simplemente 1-transitivos, es decir, de todos los grupos regulares, es equivalente a la clasificación de todos los grupos, debido al Teorema de Caley.

Los grupos de Mathieu "grandes" son construidos como una secuencia de extensiones transitivas iniciando con $PSL(3,\mathbb{F}_4)$, el cual actúa doblemente transitivamente sobre $P^2(\mathbb{F}_4)$. Puesto que $|P^2(\mathbb{F}_4)| = 4^2+4+1 = 21$, uno comienza con un grupo de permutaciones de grado 21; recordemos que $|PSL(3,\mathbb{F}_4)| = 20,160$. Si describimos los puntos de $P^2(\mathbb{F}_4)$ por coordenadas homogéneas $[\lambda,\mu,\nu]$, de donde $\lambda,\mu,\nu \in \mathbb{F}_4$, entonces la acción de $PSL(3,\mathbb{F}_4)$ sobre un punto es justamente la multiplicación matricial sobre el vector columna $[\lambda,\mu,\nu]$.

Observación 4.1.4. Sean $f_1, f_2, f_3 : P^2(\mathbb{F}_4) \to P^2(\mathbb{F}_4)$ dadas por

$$\begin{aligned} f_1[\lambda,\mu,\nu] &= [\lambda^2+\mu\nu, \mu^2, \nu^2], \\ f_2[\lambda,\mu,\nu] &= [\lambda^2, \mu^2, \beta\nu^2], \qquad \beta \in \mathbb{F}_4 \quad \text{elemento primitivo.} \\ f_3[\lambda,\mu,\nu] &= [\lambda^2, \mu^2, \nu^2]. \end{aligned}$$

Entonces, se tiene lo siguiente:

1. f_1 es una permutación de $P^2(\mathbb{F}_4)$ de orden 2 fijando a $[1,0,0]$, ya que $\lambda \mapsto \lambda^2$ es un automorfismo de $\mathbb{F}_4$.
2. f_2 es una permutación de $P^2(\mathbb{F}_4)$ de orden 2 fijando a $[1,0,0]$.
3. f_3 es una permutación de $P^2(\mathbb{F}_4)$ de orden 2 fijando a $[1,0,0]$.
4. $\langle PSL(3,\mathbb{F}_4), f_2, f_3\rangle = P\Gamma L(3,\mathbb{F}_4)$.

Teorema 4.1.4. *Existe un grupo M_{22} de grado 22 y orden $443,520 = 22\cdot 21\cdot 20\cdot 48 = 2^7\cdot 3^2\cdot 5\cdot 7\cdot 11$ tal que es un grupo 3-transitivo y el estabilizador de un punto es $PSL(3,\mathbb{F}_4)$.*

Demostración: Probemos que el $G = PSL(3,\mathbb{F}_4)$ actuando sobre $X = P^2(\mathbb{F}_4)$ tiene una extensión transitiva. Usando la notación del Teorema 4.1.1, sean

$$\begin{aligned} x &= [1,0,0], \\ g[\lambda,\mu,\nu] &= [\lambda,\mu,\nu] \end{aligned}$$

y

$$h_1 = (\infty, [1,0,0])f_1.$$

En forma matricial:

$$g = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

así que $\det(g) = -1 = 1 \in \mathbb{F}_4$ y g representa un elemento en $PSL(3, \mathbb{F}_4)$. Es claro que g no fija a $x = [1, 0, 0]$ y, usando la Observación 4.1.4 se tiene que $h_1^2 = 1$. Veamos que $(gh_1)^3 = 1$. Para $[\lambda, \mu, \nu] \neq [1, 0, 0], [0, 1, 0], \infty$, se tiene que

$$(gh_1)^3[\lambda, \mu, \nu] = [\lambda\nu + \mu^2(\nu^3 + 1), \mu\nu + \lambda^2(\nu^3 + 1), \nu^2].$$

Si $\nu \neq 0$, entonces $\nu^3 = 1$ y $\nu^3 + 1 = 0$, así que el lado derecho es $[\lambda\nu, \mu\nu, \nu^2] = [\lambda, \mu, \nu]$ (por definición de coordenadas homogéneas). Si $\nu = 0$ y $\lambda\mu \neq 0$, entonces el lado derecho es $[\mu^2, \lambda^2, 0] = [(\lambda\mu)\mu^2, (\lambda\mu)\lambda^2, 0] = [\lambda, \mu, 0]$, como deseabamos. Los casos restantes $[1, 0, 0], [0, 1, 0]$, y ∞ puede fácilmente manejarse de manera directa.

Finalmente, suponga que $k \in G_x$. Así que, $k \in PSL(3, \mathbb{F}_4)$ es representado por una matriz de la forma

$$k = \begin{bmatrix} 1 & * & * \\ 0 & a & b \\ 0 & c & d \end{bmatrix}$$

(porque k fija a $x = [1, 0, 0]$) y $\det(k) = ad - bc = 1$. Además, $h_1 k h_1$ es representado por la matriz

$$h_1 k h_1 = \begin{bmatrix} 1 & * & * \\ 0 & a^2 & b^2 \\ 0 & c^2 & d^2 \end{bmatrix}$$

la cual fija a $[1, 0, 0]$ y cuyo determinante es $a^2d^2 - b^2c^2 = (ad - bc)^2 = 1$. Así que, $h_1 G_x h_1 = G_x$, y el teorema está probado. □

Teorema 4.1.5. *Existe un grupo M_{23} 4-transitivo de grado 23 y orden $10,200,960 = 23 \cdot 22 \cdot 21 \cdot 20 \cdot 48 = 2^7 \cdot 3^2 \cdot 5 \cdot 7 \cdot 11 \cdot 23$ tal que el estabilizador de un punto es M_{22}.*

Demostración: La prueba se sigue del teorema anterior, y así que sólo daremos un bosquejo de la prueba. Al conjunto $P^2(\mathbb{F}_4) \cup \{\infty\}$ le adjuntamos el nuevo símbolo ω. Sea

$$\begin{aligned} x &= \infty, \\ g &= (\infty, [1,0,0])f_1 = \text{ la anterior } h_1 \end{aligned}$$

y

$$h_2 = (\omega, \infty)f_2.$$

Entonces, usando el Teorema 4.1.1, se tiene que $M_{23} = \langle M_{22}, h_2 \rangle$ es una extensión transitiva de M_{22}. □

Teorema 4.1.6. *Existe un grupo M_{24} el cual es 5-transitivo de grado 24 y orden $244,823,040 = 24 \cdot 23 \cdot 22 \cdot 21 \cdot 20 \cdot 48 = 2^{10} \cdot 3^3 \cdot 5 \cdot 7 \cdot 11 \cdot 23$ tal que el estabilizador de un punto es M_{23}.*

Demostración: Al conjunto $P^2(\mathbb{F}_4) \cup \{\infty, \omega\}$ le adjuntamos un nuevo símbolo Ω. Definimos

$$\begin{aligned} x &= \omega, \\ g &= (\omega, \infty)f_2 = \text{ el anterior } h_2 \end{aligned}$$

y

$$h_3 = (\Omega, \omega)f_3.$$

Entonces, usando el Teorema 4.1.1, se tiene que $M_{24} = \langle M_{23}, h_3 \rangle$ es una extensión transitiva de M_{23}. □

Observación 4.1.5. Se tiene que $\langle PSL(3, \mathbb{F}_4), h_2, h_3 \rangle$ es un subgrupo de M_{24} isomorfo a $P\Gamma L(3, \mathbb{F}_4)$.

Es evidente que los grupos de Mathieu M_{22}, M_{23} y M_{24} no son simplemente k-transitivos para $k = 3, 4, 5$, respectivamente.

Teorema 4.1.7 (Miller, 1900). *Los grupos de Mathieu M_{22}, M_{23} y M_{24} son simples.*

Demostración: Dado que M_{22} es 3-transitivo de grado 22 (y 22 no es potencia de 2) y ya que el estabilizador de un punto es el grupo simple $PSL(3, \mathbb{F}_4)$, el Corolario 3.1.3 (ii) prueba que M_{22} es simple. El grupo M_{23} es 4-transitivo y el estabilizador de un punto es el grupo simple M_{22}, así que la simplicidad de M_{23} se sigue por el Corolario 3.1.3 (i). Finalmente, M_{24} es 5-transitivo y el estabilizador de un punto es el grupo simple M_{23}, así que el Corolario 3.1.3 (i) demuestra que M_{24} es simple. □

Teorema 4.1.8. *Los grupos de Mathieu M_{11} y M_{12} son simples.*

Demostración: Usando el Corolario 3.1.3 (i), es claro que la simplicidad de M_{11} implica la simplicidad de M_{12}.

Suponga que H es un subgrupo normal de M_{11} con $H \neq \{e\}$. Por el Teorema 3.1.12, H es transitivo de grado 11, y así $|H|$ es divisible por 11. Sea P un 11-subgrupo de Sylow de M_{11} el cual es cíclico de orden 11.

Afirmamos que $P \neq N_H(P)$. Si ocurre lo contrario, es decir, $P = N_H(P)$, entonces P está contenido en el centro de su normalizador, pues P es abeliano. El Teorema 1.2.6 (Burnside's), se tiene que P admite un complemento Q que está en H. Por lo tanto, $(11, |Q|) = 1$ y Q es de característica par en H; como $H \lhd M_{11}$, tenemos que $Q \lhd M_{11}$. Por el Teorema 3.1.12, tenemos que Q es transitivo de grado 11, de donde $|Q|$ es divisible por 11, una contradicción.

En lo que resta de la demostración, usaremos los grupos M_{11} y S_{11} como indexaciones del grupo $N_{\_}(P)$ en el cual omitiremos el subíndice 11. Así, probemos que $N_H(P) = N_M(P)$. En S_{11}, existen $11!/11 = 10!$ 11-ciclos, y de aquí $9!$ subgrupos cíclicos de orden 11, cada uno de los cuales consiste de la identidad y de 10 11-ciclos. Por lo tanto, $[S_{11} : N_S(P)] = 9!$ y $|N_S(P)| = 110$. Existe un elemento τ de orden 2 invirtiendo un 11-ciclo σ: si $\sigma = (1, 2, 3, \ldots, 11)$, entonces $\sigma^{-1} = (11, 10, 9, \ldots, 1)$ y $\tau\sigma\tau = \sigma^{-1}$, donde $\tau = (1, 11)(2, 10)(3, 9)(4, 8)(5, 7)$; note que τ es impar. Dado que $M_{11} \subset A_{11}$ (la Observación 3.3.3 prueba que $M_{10} \subset A_{10}$ y $M_{11} = \langle M_{10}, h\rangle$, donde h es par), $N_M(P) = N_S(P) \cap M_{11} \subset N_S(P) \cap A_{11}$; se sigue por ser τ impar que $|N_M(P)| = 11$ ó 55. Finalmente, observe que $P \subseteq N_H(P) \subseteq N_M(P)$, donde $P \neq N_H(P)$, así que $N_H(P) = N_M(P)$; incidentalmente ambos tienen orden 55. El argumento de Frattini da $M_{11} = HN_M(P) = HN_H(P) = H$, ya que $N_H(P) \subset H$, y esto prueba que M_{11} es simple. □

Los grupos de Mathieu fueron los primeros ejemplos de grupos simples "esporadicos", es decir, grupos simples que no aparecen dentro de alguna clase infinita de grupos simples, como ejemplo, $\mathbb{Z}_p$, A_n o $PSL(n, \mathbb{F}_q)$.

4.2. Conceptos Básicos de la Teoría de Corrección de Códigos

Es sabido que los canales de transmisión de mensajes no son totalmente fidedignos, por lo cual a veces el mensaje enviado no es exactemente igual al mensaje recibido.

Veremos en esta sección, hasta cierto punto, cómo corregir ciertos errores debidos a los canales transmisión por los cuales se propaga el mensaje.

Definición 4.2.1.

(*i*) *Una* ***cadena*** *es una sucesión finita de elementos de un conjunto finito* $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ *llamado* ***alfabeto*** *y a los elementos* a_i *los llamaremos* ***símbolos del alfabeto****.*

(*ii*) *La* ***longitud*** *de una cadena es la cantidad de símbolos que la forman.*

(*iii*) *Un* ***código corrector de errores*** *es un conjunto* $\mathcal{C}$ *de cadenas de un alfabeto, a las cuales llamaremos* ***palabras del código****.*

Notación. *En ocasiones para representar una palabra del código podemos usar la notación usual, la cual consiste en la concatenación de los símbolos del alfabeto, o simplemente como un vector de n coordenadas, donde n es la longitud del vector y las entradas son los distintos símbolos que conforman a la cadena.*

Ejemplo 4.2.1. Algunos ejemplos de códigos son los siguientes:

1. Tomando $\mathcal{A} = \{0, 1\}$ y $\mathcal{C} = \{000, 001, 010, 011, 100, 101, 110, 111\}$ tenemos que este tipo de código es uno de los más conocidos y usados en computación.
2. Si $\mathcal{A} = \{a, b, c\}$, entonces podemos tomar a
$$\mathcal{C} = \{a, b, c, aa, ab, ac, ba, bb, bc, ca, cb, cc, aaa, abc\}.$$
3. Si $\mathcal{A} = \{1, 2, \ldots, n\}$, podemos tomar $\mathcal{C} = S_n$.
4. Usando el alfabeto anterior, tenemos que A_n también es un código.

Notación. Tal vez los dos últimos ejemplo no sean tan claros, pero tenemos que una permutación se puede denotar en alguna de las siguientes formas:

$$\text{Dos líneas :} \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 8 & 1 & 9 & 3 & 6 & 4 & 10 & 7 & 2 \end{pmatrix}$$

$$\text{Ciclos :} \quad (1\ 5\ 3)(2\ 8\ 10)(4\ 9\ 7)$$

Sin embargo, una tercera forma de denotar a una permutación se muestra a continuación:

$$\text{Pasiva}: \quad 5\ 8\ 1\ 9\ 3\ 6\ 4\ 10\ 7\ 2$$

Enfatizamos el uso de la notación pasiva (en comparación con la notación cíclica) cuando quitamos los paréntesis.

Podemos observar en los Ejemplos 1, 3 y 4 que en cada código sus respectivas palabras poseen longitud constante. Esta situación nos permite introducir la siguiente:

Definición 4.2.2. *Sea $\mathcal{C}$ un código con todas sus palabras teniendo la misma longitud.*

(*i*) *Sean $x = x_1x_2 \dots x_n$ y $y = y_1y_2 \dots y_n$ dos palabras del código. Entonces definimos la* ***distancia de Hamming*** *como*

$$d_H(x, y) = card(\{i \mid x_i \neq y_i\}).$$

(*ii*) *La* ***distancia mínima*** *del código $\mathcal{C}$ es*

$$d(\mathcal{C}) = \min_{\substack{x,y \in C \\ x \neq y}} d_H(x, y).$$

Observación 4.2.1. Notemos que la distancia de Hamming puede ser extendida a $\mathcal{A}^n$, y no nada más al código $\mathcal{C}$ teniendo todas sus palabras la misma longitud. Además, la distancia de Hamming es en realidad una distancia (es decir, una métrica) en $\mathcal{A}^n$ y, por lo tanto, en $\mathcal{C}$, pues cumple con las tres axiomas de métrica, o sea:

(a) Positiva definida

$$d_H(x, y) \geq 0 \quad \text{y} \quad d_H(x, y) = 0 \Leftrightarrow x = y$$

(b) Simétrica

$$d_H(x, y) = d_H(y, x)$$

(c) Desigualdad del triángulo

$$d_H(x, z) \leq d_H(x, y) + d_H(y, z)$$

Ejemplo 4.2.2. Podemos notar que en el caso de los ejemplos 1 y 2, del Ejemplo 4.2.1, se tiene que $d(\mathcal{C}) = 1$, mientras que para los ejemplos 3 y 4 pareciera que no es sencillo calcular su distancia mínima, pero como veremos más adelante se puede dar una fórmula al respecto.

Dado que de aquí en adelante trabajaremos con códigos los cuales sus palabras son de longitud constante, supondremos esta situación sin mención alguna. Así que, tenemos la siguiente:

Definición 4.2.3. *Sea C es un código sobre un alfabeto $\mathcal{A}$, y sean $n = card(\mathcal{A})$, $M = card(\mathcal{C})$ y $d = d(\mathcal{C})$. Bajo estas condiciones decimos que $\mathcal{C}$ es un (n, M, d)-código.*

Ahora, antes de seguir, cabe aclarar que en general cuando se envía un mensaje, el canal no siempre es seguro en cuestiones de fidelidad de transmisión, como ya se había dicho; pues a veces dependiendo del medio en que se propaguen los mensajes, puede que se introduzca cierta cantidad de errores. Así que, lo que nos gustaría realizar es poder corregir dichos errores. Pero, en la actualidad los canales de transmisión tienen una gran calidad de transmisión, luego entonces los errores que poseen nuestros mensajes son relativamente pocos. Por lo tanto, nuestro método de corrección de errores puede corregir hasta tres errores.

Por esta razón, hemos introducido la definición de lo que es la distancia mínima del un código. El siguiente teorema es un resultado dentro de la teoría de códigos, lo presentamos sin demostración, y nos aclara un poco más esta conexión.

Teorema 4.2.1. *Sea $\mathcal{C}$ un (n, M, d)-código. Entonces, $\mathcal{C}$ es detector de exactamente $d-1$ errores y corrector de exactamente $\left\lfloor \frac{d-1}{2} \right\rfloor$.*

De acuerdo con el teorema anterior, podemos observar que mientras mayor sea la distancia mínima de un código, mayor será la cantidad de errores que uno pueda detectar y corregir. Pero como anteriormente habíamos dicho, no siempre es sencillo calcular la distancia mínima, por lo que nos concentraremos en los grupos de permutaciones simplemente k-transitivos, dado que en estos grupo de permutaciones es sencillo calcular la distancia mínima; en este caso, las permutaciones son consideradas en forma pasiva. Así, el siguiente teorema nos muestra esta situación.

Teorema 4.2.2. *Sea G un grupo de permutaciones simplemente k-transitivo actuando sobre un conjunto X de grado n. Entonces, la distancia mínima de G es $n-k+1$.*

Demostración: Puesto que el grupo G es simplemente k-transitivo, sólo la identidad fija k puntos, así que el número máximo de puntos fijados por un elemento distinto de la identidad es $k-1$. Entonces, dos permutaciones deben de coincidir en a lo más $k-1$ posiciones; en otro caso se tendría una contradicción a la simplicidad de transitividad. Por lo tanto, dos permutaciones deben diferir en al menos $n-k+1$ posiciones, es decir, la distancia mínima entre dos permutaciones que se tenga es $n-k+1$. □

Así, de acuerdo con el teorema anterior, podemos usar los elementos de G para formar un $(n, |G|, n-k+1)$-código.

Afortunadamente, todos los grupos simplemente k-transitivos son conocidos, para cuando $k \geq 2$. Los ejemplos más fáciles de encontrar son los grupos simétricos S_n, los cuales son simplemente n y $(n-1)$-transitivos y los grupos alternantes A_n los cuales son simlplemente $(n-2)$-transitivo. Sin embargo, estos no los hacen particularmente interesantes (o útiles) como códigos. Por ejemplo, S_n produce un $(n, n!, 1)$-código pero no puede corregir errores y A_n produce un $(n, n!/2, 3)$-código el cual puede corregir sólo un error.

Hay también familias infinitas de grupos simplemente 2-transitivo; por ejemplo, se puede tomar el conjunto de transformaciones afines sobre un campo finito. Esto es, el grupo

$$AGL(1, \mathbb{F}_q) = \{\tau : x \mapsto ax + b \mid a, b, \in \mathbb{F}_q, a \neq 0\}$$

actúa simplemente 2-transitivamente sobre $\mathbb{F}_q$. Este grupo tiene grado q y orden $q(q-1)$, así que podemos usarlo para contruir un $(q, q(q-1), q-1)$-código.

Para grupos simplemente 3-transitivos podemos considerar los grupos de transformaciones racionales. Es decir, tomemos el grupo

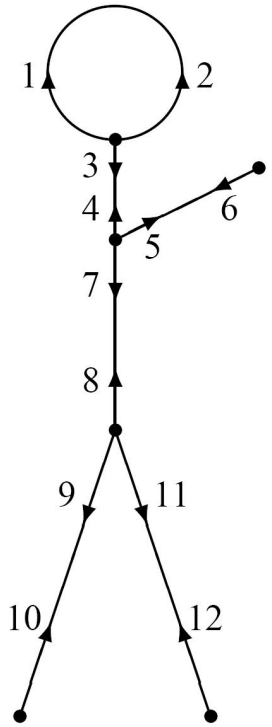
$$\left\{ \tau : x \mapsto \frac{ax+b}{cx+d} \;\middle|\; a, b, c, d \in \mathbb{F}_q, ad - bc \neq 0 \right\},$$

el cual actúa sobre $\mathbb{F}_q \cup \{\infty\}$, donde el punto ∞ se definido por satisfacer $\tau(\infty) = ac^{-1}$ y $\tau(-dc^{-1}) = \infty$. Sabemos que éste es isomorfo al grupo lineal general proyectivo $PGL(2, \mathbb{F}_q)$. Es de grado $q+1$ y de orden $(q+1)q(q-1)$, así puede ser usado para construir un $(q+1, (q+1)q(q-1), q-1)$-código.

Para el caso $k = 4$ y $k = 5$, los únicos ejemplos (distintos de los grupos simétricos y alternantes) son los grupos de Mathieu M_{11}, el cual es simplemente 4-transitivo, y M_{12}, el cual es simplemente 5-transitivo. Este último es al que pondremos más atención en lo que sigue.

4.3. Una Representación de los Grupos M_{11} y M_{12}

Los grupos de Mathieu M_{11} y M_{12} son dos de los 26 grupos simples "esporádicos". Existen varias formas de construir estos, ya hemos visto una construcción formal, pero la siguiente es una forma sencilla de recordarlos, sobre todo el grupo M_{12}. Consideremos el siguiente diagrama:



Primero, tomemos la permutación dadas por el producto de los 2-ciclos de los niveles de cada orilla, es decir, (1 2)(3 4)(5 6)(7 8)(9 10)(11 12). Ahora, en cada vértice, formamos un ciclo, considerándolo en sentido contrario a las manecillas del reloj, ignorando los ciclos triviales, para formar la permutación (1 3 2)(4 7 5)(8 9 11). Entonces, estas dos permutaciones generan a M_{12}, el cual sabemos que es un grupo simple de grado 12 y orden 95040. M_{11} es el estabilizador de un punto de M_{12}.

Ya hemos expuesto que M_{11} y M_{12} son simplemente 4-transitivo y 5-transitivo, respectivamente. Así que, de acuerdo con la sección anterior, tenemos que la mínima distancia entre dos permutaciones de M_{12} es $12 - 5 + 1 = 8$, así que podemos usarlo para formar un $(12, 95040, 8)$-código. Puesto que M_{12} es transitivo, el tamaño de una órbita de uno de sus puntos es exactamente el grado, así que el órden de M_{11}, el estabilizador de M_{12} en un punto, es $95040/12 = 7920$. Puesto que M_{11} es simplemente 4-transitivo de grado 11, podemos usar a M_{11} como un $(11, 7920, 8)$-código.

4.4. Un Método de Decodificación

En cualquier (n, M, d)-código, el número máximo de errores que puede ser corregido es $r = \left\lfloor \frac{d-1}{2} \right\rfloor$ (Teorema 4.2.1). Así que, en el método de decodificación a considerar, supondremos que a lo más r errores han de ocurrir; en consecuencia, hay al menos $n - r$ símbolos que son corregidos. Ahora, puesto que nuestras palabras del código son permutaciones de un grupo simplemente k-transitivo G, cualquier k-tupla de estos $n - r$ símbolos a corregir, ocurre en exactamente una palabra del código. Luego, códigos con distancia mínima $q - 1$ puede corregir a $\lfloor \frac{q-2}{2} \rfloor$ y un código con distancia mínima 8 puede corregir 3 errores. Pero, puesto que no sabemos en cuáles posiciones los r errores aparecen, necesitamos encontrar una forma de elegir un conjunto de k-tuplas tales que pueda ser cierta que al menos una no contiene errores. Es decir, necesitamos un conjunto de k-subconjuntos de $X = \{1, \ldots, n\}$ tal que cualquier r-subconjunto de X sea disjunto de al menos un k-conjunto. Llamaremos a este conjunto de k-conjuntos un (n, k, r)-descubierta (debido a que cualquier r-conjunto es dejado "descubierto" por al menos un k-conjunto). Esto será descrito en la próxima sección.

Una vez que tengamos una descubierta, podemos aplicar el siguiente procedimiento a decodificar una palabra recibida $\mathbf{w} = w_1 w_2 \cdots w_n$. Tomamos un k-conjunto S y observamos las entradas w_i, para $i \in S$.

Caso (i): No hay símbolos repetidos en estas posiciones, es decir, tenemos una permutación. Entonces, puesto que G es simplemente k-transitivo, existe una única permutación que aplica cada i de S a w_i; es decir, existe una única palabra del código con entrada w_i en posición i para $i \in S$. Así, buscamos la única permutación que cumpla con la condición anterior. Luego, checamos la distancia entre esta palabra del código encontrada y la palabra recibida $\mathbf{w}$. Si está dentro de la distancia d (recordemos que $d = n - k + 1$), entonces asumimos que esta debe de ser la palabra

transmitida y paramos el procedimiento. En caso contrario, consideramos el próximo k-conjunto de posición de coordenadas y repetimos el procedimiento.

Caso (ii): Si en alguna etapa una k-tupla de entradas incluye símbolos repetidos, entonces sabemos que debe de haber un error e inmediatamente procedemos al próximo k-conjunto.

4.5. Descubiertas

Formalmente, definimos una (n, k, r)-descubierta como sigue. Supóngase que X es un conjunto de tamaño n, es decir, de cardinalidad n; denotamos por $\mathcal{P}(X)$ al conjunto potencia del conjunto X, y por $\binom{X}{k}$ al conjunto de subconjuntos de X de tamaño k (el conjunto de todos los k-subconjuntos de X). Entonces, una (n, k, r)-descubierta es un conjunto $\mathcal{M}$ de subconjunto de X de tamaño k, esto es,

$$\mathcal{M} \subseteq \binom{X}{k} \subset \mathcal{P}(X)$$

tal que para cada r-subconjunto R de X, con $r < k$, existe $S \in \mathcal{M}$ tal que $R \cap S = \emptyset$. Cuando $\mathcal{M}$ sea una descubierta de menor tamaño, entonces diremos que $\mathcal{M}$ es una (n, k, r)-descubierta mínimal.

Daremos algunos ejemplos que son de interés para nosostros, sobre todo relacionados con el grupo M_{12}, donde en este caso particular $X = \{1, 2, \ldots, 12\}$.

1. Una $(12, 5, 3)$-descubierta para el grupo de Mathieu M_{12}:

1	2	3	4	5
1	2	6	11	12
1	3	7	8	9
1	4	6	7	10
1	5	8	9	11
2	4	8	9	12
2	5	7	10	11
3	4	7	11	12
3	5	6	10	12
3	6	8	9	11
6	7	8	9	10

2. Una $(10, 5, 2)$-descubierta para el grupo de Mathieu M_{12}:

1	2	3	4	5
1	2	7	8	10
1	5	6	7	9
2	3	6	8	9
3	4	7	9	10
4	5	6	8	10

3. Una $(9, 5, 1)$-descubierta para el grupo de Mathieu M_{12}:

1	2	3	4	5
1	6	7	8	9
5	6	7	8	9

4. Una $(8, 5, 1)$-descubierta para el grupo de Mathieu M_{12}:

1	2	3	4	5
1	2	6	7	8
4	5	6	7	8

4.6. M_{12} en Detalle

Existen diferentes posibilidades para que ocurran como máximo tres errores. Si la palabra recibida es una permutación, entonces no tenemos otra elección más que proceder de la manera descrita anteriormente. Sin embargo, si no es una permutación entonces debemos de mejorar el algoritmo. Por ejemplo, si los r errores son r repeticiones del mismo símbolo, entonces sabemos que el resto de los $n - r - 1$ símbolos deben ser correctos y podemos inmediatamente encontrar la única palabra del código que corresponda a alguna k-tupla de estas. En este caso, se tiene que $k < n - r - 1$. Así, tenemos decodificada la palabra recibida en un sólo paso, una mejora considerable.

En el caso en donde es usado M_{12}, existe un número suficientemente pequeño de posibilidades para poder determinarlos todos explícitamente. En cada caso, podemos aislar algunos de los dígitos donde sabemos que existe un error, así aplicando un algoritmo similar para el resto de los dígitos para localizar el número más pequeño de errores, los cuales deberían de requerir un número menor de pasos.

Si la palabra recibida no es una permutación, entonces existen varias posibilidades para esto, como en el listado de abajo. Cada ejemplo está basado en la palabra transmitida, siendo la palabra identidad 1 2 3 4 5 6 7 8 9 10 11 12.

- *Tres errores*

- Un símbolo repetido, dos símbolos movidos.
  Por ejemplo, 1 1 4 3 5 6 7 8 9 10 11 12.
  Aquí necesitamos una $(10, 5, 2)$-descubierta, ya que sabemos que al menos uno de los dos 1's es un error.
- Dos símbolos diferentes repetidos, un símbolo movido.
  Por ejemplo, 1 1 3 3 2 6 7 8 9 10 11 12.
  Aquí necesitamos una $(8, 5, 1)$-descubierta.
- Tres diferentes símbolos repetidos $(*)$.
  Por ejemplo, 1 1 3 3 5 5 7 8 9 10 11 12
  Aquí sabemos que hay tres errores en los primeros seis lugares, así que podemos examinar cualquiera de los cinco de los seis últimos lugares para poder encontrar la única permutación en M_{12} que coincida en estos lugares.
- Un símbolo repetido tres veces $(*)$.
  Por ejemplo, 1 1 1 1 5 6 7 8 9 10 11 12
  Aquí, podemos elegir alguno de los cinco de los últimos ocho lugares y procedemos como antes.
- Un símbolo repetido dos veces, y algún otro símbolo repetido $(*)$.
  Por ejemplo, 1 1 1 4 4 6 7 8 9 10 11 12
  Aquí, podemos elegir cualquiera de los cinco de los últimos siete lugares y procedemos como antes.
- Un símbolo repetido dos veces y otro símbolo movido.
  Por ejemplo, 1 1 1 2 5 6 7 8 9 10 11 12.
  Aquí necesitamos una $(9, 5, 1)$-descubierta.
- Un símbolo repetido dos veces, ambos ocurren en lugares incorrectos, otro símbolo movido.
  Por ejemplo, 3 3 1 4 5 6 7 8 9 10 11 12.
  Aquí necesitamos una $(10, 5, 2)$-descubierta y la aplicamos a los lugares 3 y 12.

Notemos que si no se necesita ninguna descubierta, entonces significa que nos encontramos en alguna de las situaciones marcadas con $(*)$. Por lo que sólo se necesita tomar cinco dígitos de los restantes, es decir, de los que no poseen dígitos repetidos, y entonces se procede a encontrar la permutación.

- *Dos errores*

  - Dos símbolos repetidos.
    Por ejemplo, 1 1 3 3 5 6 7 8 9 10 11 12.

Aquí, necesitamos una $(8, 5, 1)$-descubierta. (Notemos que aunque sólo dos errores han actualmente ocurrido, tenemos que continuar con la suposición de que pueden haber tres errores).

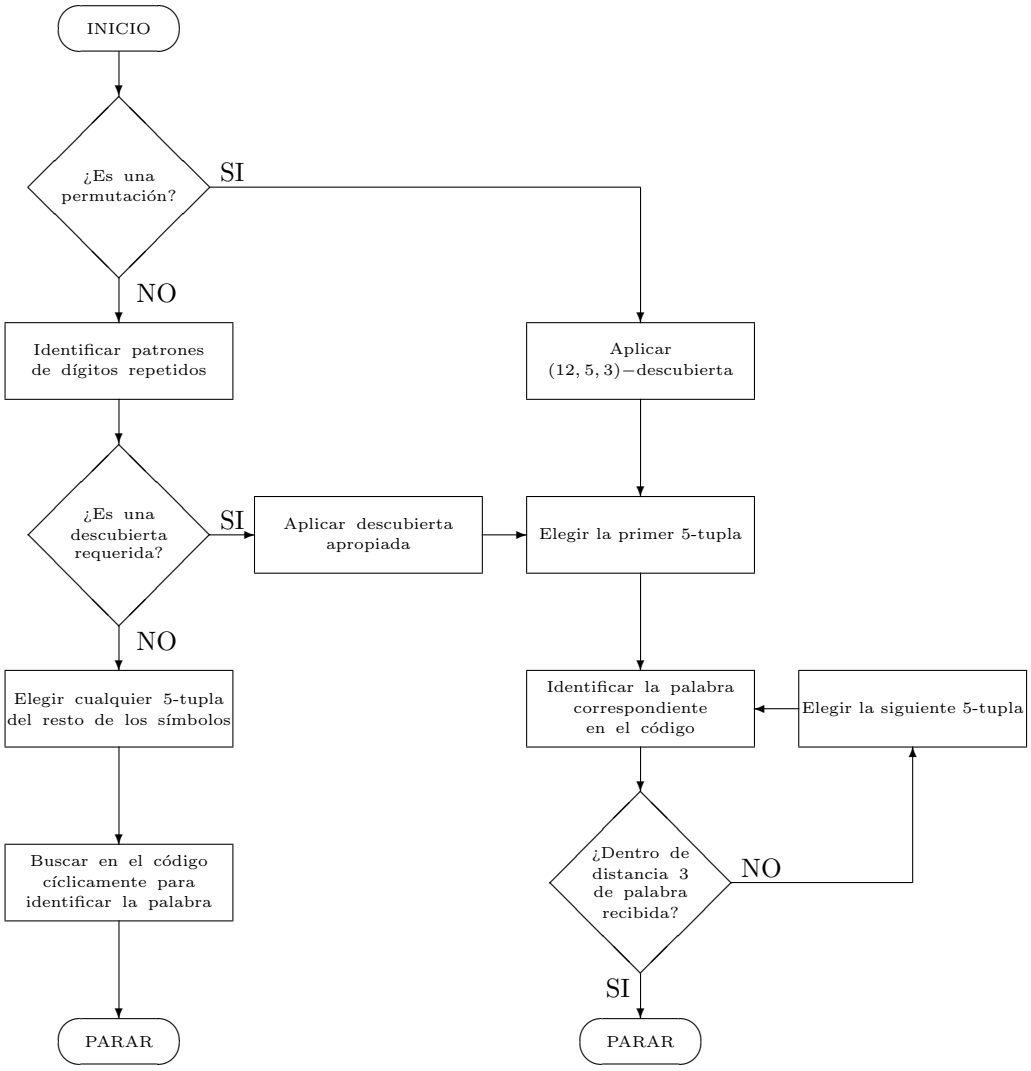
- Un símbolo repetido dos veces.
  Por ejemplo, 1 1 1 4 5 6 7 8 9 10 11 12.
  Aquí, necesitamos una $(9, 5, 1)$-descubierta.

- *Un error*

  - Un símbolo repetido.
    Por ejemplo, 1 1 3 4 5 6 7 8 9 10 11 12.
    Aquí, necesitamos una $(10, 5, 2)$-descubierta.

4.7. El Algoritmo

La forma como trabaja el algoritmo de decodificación para una permutación de M_{12} es mejor descrito por el siguiente diagrama:



Por "identificar patrones de dígitos repetidos", significa determinar cuáles de las situasiones listadas anteriormente corresponden a la palabra recibida. Ahora, como descubrimos, alguna de estas situaciones (etiquetadas con $(*)$) restringen las posiciones de los errores significativamente para conocer más de cinco símbolos correctos. Estas son exactamente aquellas palabras para las cuales la respuesta a la pregunta "¿Es una descubierta requerida?" es "NO". En este caso, elegimos alguna 5-tupla de las posiciones donde el símbolo correcto ocurra, identificándolo de la lista de permutaciones en el grupo la permutación la cual tiene aquellos cinco símbolos en estos lugares. Entonces, esto debe de ser la palabra transmitida.

Ejemplo 4.7.1. Suponga que recibimos la palabra 8 8 11 12 3 2 1 6 9 10 1 2. Esta tiene los símbolos 1, 2 y 8 repetidos, ocurriendo en posiciones 1, 2, 6, 7, 11 y 12.

Así, elegimos alguna 5-tupla del resto de las posiciones, tal como $\{3, 4, 5, 8, 9\}$, y encontramos la permutación $\sigma \in M_{12}$ tal que $\{3, 4, 5, 8, 9\}^\sigma = \{11, 12, 3, 6, 9\}$. Miramos la lista de permutaciones en el grupo para ver que $\sigma = 7\ 8\ 11\ 12\ 3\ 4\ 5\ 6\ 9\ 10\ 1\ 2$ debe ser la palabra transmitida.

La respuesta a que "¿Es una descubierta requerida?" es "SI" en los otros casos donde la palabra recibida no es una permutación. Así determinamos cuál cubierta necesitamos, entonces lo reetiquetamos, luego los números $1,\ 2,\ 3, \ldots$ corresponden a la 1a., 2da., 3ra.,... posición donde un símbolo repetido no ocurre.

Índice de Notación

$\langle x_i : i \in I \rangle$	grupo generad por los elementos x_i
K, $\mathbb{F}_q$	campo, campo con q elementos
G/H	grupo de clases laterales del subgrupo normal H
$G//H$	conjunto de clases laterales del subgrupo H
A_n	grupo alternante de orden $n!/2$
$G^* = G - \{1\}$	cuando G es un grupo multiplicativo con identidad 1
$K^* = K - \{0\}$	cuando K es un campo
$V^* = V - \{0\}$	cuando V es un espacio vectorial
$\mathrm{Aff}\,(V) \cong \mathrm{Aff}\,(n, K)$	grupo afín
$\mathrm{Aut}\,(G)$	grupo de automorfismo del grupo G
$\mathrm{Aut}\,(K)$	grupo de automorfismo del camo K
$\mathrm{Aut}\,(V) \cong \mathrm{Aut}\,(n, K)$	Grupo de automorfismo afines
$B_{ij}(\lambda)$	matriz transvección
$\mathbb{C}$	números complejos
$C_G(x)$	centralizador en G del elemento x
$C_G(H)$	centralizador en G del subgrupo H
I	matriz identidad
G'	subgrupo conmutador de G
$G^{(i)}$	i-ésimo subgrupo conmutador de G
G_x	estabilizador de x
$GL(V) \cong GL(n, K)$	grupo general lineal
$\Gamma L(V) \cong \Gamma L(n, K)$	grupo de transformasiones semilineales no singulares
$\Gamma LF(V)$	grupo de transformasiones fraccionales semilineales no singulares
$Inn(G)$	Grupo de automorfismos del grupo G
$LF(K)$	grupo de transformasiones fraccionales lineales
M_n	grupo de Mathieu ($n = 10, 11, 22, 23, 24$)
$\mathbb{N}$	números naturales
$N_G(H)$	normalizador de G de un subgrupo H
O_x	órbita de x

$P^n(K)$	el n-espacio proyectivo
$PGL(V) \cong PGL(n,K)$	grupo de todas las proyectividades
$PSL(V) \cong PSL(n,K)$	grupo unimodular proyectivo
$P\Gamma L(V) \cong P\Gamma L(n,K)$	grupo de colineaciones
$\mathbb{Q}$	números racionales
$\mathbb{R}$	números reales
S_n	grupo simétrico de orden $n!$
S_X	todas las permutaciones del conjunto X
$Sc(V)$	transformasiones escalares distintas de cero
$Sc_1(V)$	todas las transformasiones escalares teniendo determinante 1
$Sh(K)$	grupo simplemente 3-transivo
$SL(V) \cong SL(n,K)$	grupo especial lineal: todas las transformasiones de determinante 1.
$T_{f,\gamma}$	transvección (transformasión)
$Tr(V)$	grupo de traslaciones
$\mathbb{V}$	4-grupo
$V(n,K)$	espacio vectorial de dimensión n sobre K
$\mathbb{Z}$	los números enteros
$Z(G)$	centro del grupo G
$\mathbb{Z}_n$	grupo cíclico de orden n
$Z(n,K)$	grupo de todas las matrices escalares $n \times n$ distintas de cero
$Z_1(n,K)$	grupo de todas las matrices escalares $n \times n$ con determinante 1
σ_*	colineación inducida por el automorfismo de campo σ
$A \times B$	producto directo
$\prod_{i=1}^{n} A_i$	producto directo
$\prod_{i \in I} A_i$	producto directo
$A \oplus B$	suma directa
$\sum_{i=1}^{n} A_i$	suma directa
$\sum_{i \in I} A_i$	suma directa
$A \rtimes_\theta B \ o \ A \rtimes B$	producto semidirecto
$A \wr B$	producto orlado

Conclusiones

En el transcurso de la tesis hemos notado algunos hechos de interés, como son:

- La teoría de grupos no sólo es una mera abstracción del pensamiento humano como son el caso de algunas ramas de la matématica, existen varios grupos importantes que aparecen en la física o en otras áreas. Algunos de estos grupos son el grupo de Möbius, el grupo de Lie, el grupo de matrices invertible $PSL(m, \mathbb{R})$, etc.

- También se vio una nueva forma de construir grupos (Teorema de Witt) bajo ciertas características a partir de otros, y los cuales nos van a dar un grupo transitivo. Es decir, dimos un método diferente al de los productos de grupos.

- Una conclusión que obtenemos de la parte de la aplicación del grupo de Mathieu M_{12} es el de que este grupo puede corregir hasta tres errores en una palabra de longitud 12 con todos sus símbolos distintos, aunque durante la teoría pudimos observar que hay otros grupos que corrigen una mayor cantidad de éstos, pero no fueron usados por tener una cantidad muy pequeña de palabras, lo cual no convenía debido a que a mayor cantidad de palabras podemos construir un mejor lenguaje. Además, dado que en la actualidad los medios de transmisión digitales son muy fieles, y los errores que suceden en una palabra son mucho menores a los que hemos tratado aquí, por lo general son de un error por cada 1000 bits transmitidos, podemos estar seguros de que el algoritmo presentado será útil.

Bibliografía

[1] Cameron, P. J., and J. H. Van Lint, *Desings, Graphs, Codes and Their Links*, London Mathematical Society Student Texts (22), Cambridge University Press, 1991.

[2] Conway, J. H., R. T. Curtis, S. P. Norton, R. A. Parker and R. A. Wilson *An Atlas of Finite Simple Groups*, Oxford University Press, 1985.

[3] Hungerford, T. W., *Algebra*, Springer-Verlag New York Inc., 1974.

[4] Lang, S., *Algebra*, Addison-Wesley. Reading, MA, 1969.

[5] Bailey, R. F., *Permutation Groups, Error-Correcting Codes and Uncovering*, Thesis Submitted to the University of London for the Degree of Doctor of Philosophy, 2005.

[6] Rodríguez-Cruz, A., *El Grupo Lineal General* (*Obtención de una Familia de Grupos Simples*), Tesis de Licenciatura. ESFM-IPN, México, 2002.

[7] Rotman, J., *An Introduction to the Theory of Groups*, Springer-Verlag, GTM 148, 3rd ed., 1984.

Índice alfabético