Chapter 39

Matroids

This chapter gives the basic definitions, examples, and properties of matroids. We use the shorthand notation

$$X + y := X \cup \{y\} \text{ and } X - y := X \setminus \{y\}.$$

39.1. Matroids

A pair $(S, \mathcal{I})$ is called a *matroid* if S is a finite set and $\mathcal{I}$ is a nonempty collection of subsets of S satisfying:

(39.1) (i) if $I \in \mathcal{I}$ and $J \subseteq I$, then $J \in \mathcal{I}$,
(ii) if $I, J \in \mathcal{I}$ and $|I| < |J|$, then $I + z \in \mathcal{I}$ for some $z \in J \setminus I$.

(These axioms are given by Whitney [1935].)

Given a matroid $M = (S, \mathcal{I})$, a subset I of S is called *independent* if I belongs to $\mathcal{I}$, and *dependent* otherwise. For $U \subseteq S$, a subset B of U is called a *base* of U if B is an inclusionwise maximal independent subset of U. That is, $B \in \mathcal{I}$ and there is no $Z \in \mathcal{I}$ with $B \subset Z \subseteq U$.

It is not difficult to see that, under condition (39.1)(i), condition (39.1)(ii) is equivalent to:

(39.2) for any subset U of S, any two bases of U have the same size.

The common size of the bases of a subset U of S is called the *rank* of U, denoted by $r_M(U)$. If the matroid is clear from the context, we write $r(U)$ for $r_M(U)$.

A set is called simply a *base* if it is a base of S. The common size of all bases is called the *rank* of the matroid. A subset of S is called *spanning* if it contains a base as a subset. So bases are just the inclusionwise minimal spanning sets, and also just the independent spanning sets. A *circuit* of a matroid is an inclusionwise minimal dependent set. A *loop* is an element s such that $\{s\}$ is a circuit. Two elements s, t of S are called *parallel* if $\{s, t\}$ is a circuit.

Nakasawa [1935] showed the equivalence of axiom system (39.1) with an ostensibly weaker system, which will be useful in proofs:

Theorem 39.1. *Let S be a finite set and let $\mathcal{I}$ be a nonempty collection of subsets satisfying* (39.1)(i)*. Then* (39.1)(ii) *is equivalent to:*

(39.3) *if $I, J \in \mathcal{I}$ and $|I \setminus J| = 1$, $|J \setminus I| = 2$, then $I + z \in \mathcal{I}$ for some $z \in J \setminus I$.*

Proof. Obviously, (39.1)(ii) implies (39.3). Conversely, (39.1)(ii) follows from (39.3) by induction on $|I \setminus J|$, the case $|I \setminus J| = 0$ being trivial. If $|I \setminus J| \geq 1$, choose $i \in I \setminus J$. We apply the induction hypothesis twice: first to $I - i$ and J to find $j \in J \setminus I$ with $I - i + j \in \mathcal{I}$, and then to $I - i + j$ and J to find $j' \in J \setminus (I + j)$ with $I - i + j + j' \in \mathcal{I}$. Then by (39.3) applied to I and $I - i + j + j'$, we have that $I + j \in \mathcal{I}$ or $I + j' \in \mathcal{I}$. ∎

39.2. The dual matroid

With each matroid M, a dual matroid M^* can be associated, in such a way that $(M^*)^* = M$. Let $M = (S, \mathcal{I})$ be a matroid, and define

(39.4) $\mathcal{I}^* := \{I \subseteq S \mid S \setminus I \text{ is a spanning set of } M\}$.

Then (Whitney [1935]):

Theorem 39.2. *$M^* = (S, \mathcal{I}^*)$ is a matroid.*

Proof. Condition (39.1)(i) trivially holds for $\mathcal{I}^*$. To see (39.1)(ii), consider $I, J \in \mathcal{I}^*$ with $|I| < |J|$. By definition of $\mathcal{I}^*$, $S \setminus J$ contains some base B of M. As also $S \setminus I$ contains some base of M, and as $B \setminus I \subseteq S \setminus I$, there exists a base B' of M with $B \setminus I \subseteq B' \subseteq S \setminus I$. Then $J \setminus I \not\subseteq B'$, since otherwise (as $B \cap I \subseteq I \setminus J$, and as $B \setminus I$ and $J \setminus I$ are disjoint, since $B \cap J = \emptyset$)

(39.5) $$|B| = |B \cap I| + |B \setminus I| \leq |I \setminus J| + |B \setminus I| < |J \setminus I| + |B \setminus I| \leq |B'|,$$

which is a contradiction. As $J \setminus I \not\subseteq B'$, there is a $z \in J \setminus I$ with $z \notin B'$. So B' is disjoint from $I + z$. Hence $I + z \in \mathcal{I}^*$. ∎

The matroid M^* is called the *dual matroid* of M. The bases of M^* are precisely the complements of the bases of M. This implies $(M^*)^* = M$, which justifies the name dual.

Theorem 39.3. *The rank function r_{M^*} of the dual matroid M^* satisfies, for $U \subseteq S$:*

(39.6) $$r_{M^*}(U) = |U| + r_M(S \setminus U) - r_M(S).$$

Proof. Let $\mathcal{B}$ and $\mathcal{B}^*$ denote the collections of bases of M and of M^*, respectively. Then

$$
\begin{aligned}
(39.7)\qquad & r_{M^*}(U) = \max\{|U \cap A| \bigm| A \in \mathcal{B}^*\} = \max\{|U \setminus B| \bigm| B \in \mathcal{B}\} \\
& = |U| - \min\{|B \cap U| \bigm| B \in \mathcal{B}\} \\
& = |U| - r_M(S) + \max\{|B \setminus U| \bigm| B \in \mathcal{B}\} \\
& = |U| - r_M(S) + r_M(S \setminus U). \qquad \blacksquare
\end{aligned}
$$

The circuits of M^* are called the *cocircuits* of M. They are the inclusionwise minimal sets intersecting each base of M (as they are the inclusionwise minimal sets contained in no base of M^*, that is, not contained in the complement of any base of M). The loops of M^* are the *coloops* or *bridges* of M, and parallel elements of M^* are called *coparallel* or *in series* in M.

Let $M = (S, \mathcal{I})$ be a matroid, and suppose that we can test in polynomial time if any subset of S is independent in M (or we have an oracle for that). Then we can calculate, for any subset U of S, the rank $r_M(U)$ of U in polynomial time (by growing an independent set (starting from $\emptyset$) to an inclusionwise maximal independent subset of U). It follows that we can test in polynomial time if any subset U of S in independent in M^*, just by testing if $r_M(S \setminus U) = r_M(S)$.

A matroid $M = (S, \mathcal{I})$ is called *connected* if $r_M(U) + r_M(S \setminus U) > r_M(S)$ for each nonempty proper subset U of S. This is equivalent to: for any two elements $s, t \in S$ there exists a circuit containing both s and t. One may derive from (39.6) that a matroid M is connected if and only if M^* is connected.

39.3. Deletion, contraction, and truncation

We can derive matroids from matroids by 'deletion' and 'contraction'. Let $M = (S, \mathcal{I})$ be a matroid and let $Y \subseteq S$. Define

$$
(39.8)\qquad \mathcal{I}' := \{Z \mid Z \subseteq Y, Z \in \mathcal{I}\}.
$$

Then $M' = (Y, \mathcal{I}')$ is a matroid again, as directly follows from the matroid axioms (39.1). M' is called the *restriction* of M to Y, denoted by $M|Y$. If $Y = S \setminus Z$ with $Z \subseteq S$, we say that M' arises by *deleting* Z, and denote M' by $M \setminus Z$. Clearly, the rank function of $M|Y$ is the restriction of the rank function of M to subsets of Y.

Contraction is the operation dual to deletion. *Contracting* Z means replacing M by $(M^* \setminus Z)^*$. This matroid is denoted by M/Z. If $Y = S \setminus Z$, then we denote $M \cdot Y := M/Z$. Theorem 39.3 implies that the rank function r' of M/Z satisfies

$$
(39.9)\qquad r_{M/Z}(X) = r(X \cup Z) - r(Z)
$$

for $X \subseteq S \setminus Z$.

We can describe contraction as follows. Let $Z \subseteq S$ and let X be a base of Z. Then

(39.10) a subset I of $S \setminus Z$ is independent in M/Z if and only if $I \cup X$ is independent in M.

Note that for disjoint subsets Y, Z of S one has $(M \setminus Y) \setminus Z = M \setminus (Y \cup Z)$ and hence $(M/Y)/Z = M/(Y \cup Z)$. Moreover, deletion and contraction commute, as for any two distinct $x, y \in S$ and any $Z \subseteq S \setminus \{x, y\}$ one has (using (39.9)):

$$(39.11) \qquad \begin{aligned} r_{M\setminus x/y}(Z) &= r_{M\setminus x}(Z \cup \{y\}) - r_{M\setminus x}(\{y\}) = r_M(Z \cup \{y\}) - r_M(\{y\}) \\ &= r_{M/y}(Z) = r_{M/y\setminus x}(Z). \end{aligned}$$

If matroid M' arises from M by a series of deletions and contractions, M' is called a *minor* of M.

The circuits of $M|Y$ are exactly the circuits of M contained in Y, and the circuits of $M \cdot Y$ are exactly the minimal nonempty sets $C \cap Y$, where C is a circuit of M.

Another operation is that of 'truncation'. Let $M = (S, \mathcal{I})$ be a matroid and let k be a natural number. Define $\mathcal{I}' := \{I \in \mathcal{I} \mid |I| \leq k\}$. Then $(S, \mathcal{I}')$ is again a matroid, called the *k-truncation* of M.

39.4. Examples of matroids

We describe some basic classes of matroids.

Uniform matroids. An easy class of matroids is given by the *uniform matroids*. They are determined by a set S and a number k: the independent sets are the subsets I of S with $|I| \leq k$. This trivially gives a matroid, called a *k-uniform matroid* and denoted by U_n^k, where $n := |S|$.

Linear matroids (Grassmann [1862], Steinitz [1913]). Let A be an $m \times n$ matrix. Let $S := \{1, \ldots, n\}$ and let $\mathcal{I}$ be the collection of all those subsets I of S such that the columns of A with index in I are linearly independent. That is, such that the submatrix of A consisting of the columns with index in I has rank $|I|$.

Then $(S, \mathcal{I})$ is a matroid (property (39.1)(ii) was proved by Grassmann [1862] and by Steinitz [1913], and is called *Steinitz' exchange property*). Condition (39.1)(i) is trivial. To see condition (39.1)(ii), let $I, J \in \mathcal{I}$ with $|I| < |J|$. Then I spans an $|I|$-dimensional space $\overline{I}$. So $J \not\subseteq \overline{I}$. Take $j \in J \setminus \overline{I}$. Then $I + j \in \mathcal{I}$ and $j \in J \setminus I$.

Any matroid obtained in this way, or isomorphic to such a matroid, is called a *linear matroid*. If A has entries in a field $\mathbb{F}$, then M is called *representable over* $\mathbb{F}$. We will also say that M is *represented by* (the columns of) A, and A is called a *representation* of M.

Note that the rank $r_M(U)$ of any subset U of S is equal to the rank of the matrix formed by the columns indexed by U.

The dual matroid of a matroid representable over a field $\mathbb{F}$ is again representable over $\mathbb{F}$. Indeed, we can assume that the matrix A is of the form $[I_m \;\; B]$, where I_m is the $m \times m$ identity matrix, and B is an $m \times (n-m)$

matrix. Then the dual matroid can be represented by the matrix $[B^{\mathsf{T}} \ I_{n-m}]$, as follows directly from elementary linear algebra. This implies that the class of matroids representable over $\mathbb{F}$ is closed under taking minors.

MacLane [1936] (and also Lazarson [1958]) showed that nonlinear matroids exist.

Binary matroids. A matroid representable over GF(2) — the field with two elements — is called a *binary matroid*. For later purposes, we give some characterizations of binary matroids. The following is direct (Whitney [1935]):

(39.12) a matroid M is binary if and only if for each choice of circuits $C_1, \ldots, C_t$, the set $C_1 \triangle \cdots \triangle C_t$ can be partitioned into circuits.

In a binary matroid M, disjoint unions of circuits are called the *cycles* of M. Of special interest is the *Fano matroid* F_7, represented by the nonzero vectors in $\mathrm{GF}(2)^3$.

Tutte [1958a,1958b] showed that the unique minor-minimal nonbinary matroid is U_4^2, the 2-uniform matroid on 4 elements. (We follow the proof suggested by A.M.H. Gerards.)

Theorem 39.4. *A matroid is binary if and only if it has no U_4^2 minor.*

Proof. Necessity follows from the facts that the class of binary matroids is closed under taking minors and that U_4^2 is not binary.

To see sufficiency, we first show the following. Let M and N be matroids on the same set S. Call a set *wrong* if it is a base of precisely one of M and N. A *far base* is a common base B of M and N such that there is no wrong set X with $|B \triangle X| = 2$. We first show:

(39.13) if M and N are different and have a far base, then M or N has a U_4^2 minor.

Let M, N form a counterexample with S as small as possible. Let B be a far base and X be a wrong set with $|B \triangle X|$ minimal. Then $B \cup X = S$, since we can delete $S \setminus (B \cup X)$. Similarly (by considering M^* and N^*), $B \cap X = \emptyset$. Then, by the minimality of $|B \triangle X|$, X is the only wrong set. By symmetry, we may assume that X is a base of M. Then M has a base B' with $|B \triangle B'| = 2$. By the uniqueness of X, B' is also a base of N. By the minimality of $|B \triangle X|$, B' is not far. Hence, by the uniqueness of X, $|B' \triangle X| = 2$. So $|S| = 4$.

Let $S = \{a, b, c, d\}$, $B = \{a, b\}$, $X = \{c, d\}$. Since $M \neq U_4^2$ by assumption, we may assume that $\{a, c\}$ is not a base of M. Hence, since $\{a\}$ and $\{c, d\}$ are independent in M, $\{a, d\}$ is a base of M. Similarly, since $\{c\}$ and $\{a, b\}$ are independent in M, $\{b, c\}$ is a base of M.

Since B is far, $\{a, d\}$ and $\{b, c\}$ are bases also of N, and $\{a, c\}$ is not a base of N. So $\{c\}$ is independent in N, implying that $\{c, a\}$ or $\{c, d\}$ is a base of N, a contradiction. This proves (39.13).

Now let M be a nonbinary matroid on a set S. Choose a base B of M. Let $\{x_b \mid b \in B\}$ be a collection of linearly independent vectors over GF(2). For each $s \in S \setminus B$, let C_s be the circuit contained in $B \cup \{s\}$, and define

$$x_s := \sum_{b \in C_s \setminus \{s\}} x_b. \tag{39.14}$$

Let N be the binary matroid represented by $\{x_s \mid s \in S\}$. Now for each $b \in B$ and each $s \in S \setminus B$ one has that $(B \setminus \{b\}) \cup \{s\}$ is a base of M if and only if it is a base of N. So B is a far base. Since N is binary, we know that $N \neq M$ and that N has no U_4^2 minor. Hence, by (39.13), M has a U_4^2 minor. ▮

Regular matroids. A matroid is called *regular* if it is representable over each field. It is equivalent to requiring that it can be represented over $\mathbb{R}$ by the columns of a totally unimodular matrix.

Regular matroids are characterized by Tutte [1958a,1958b] as those binary matroids not having an F_7 or F_7^* minor. (Gerards [1989b] gave a short proof.)

A basic decomposition theorem of Seymour [1980a] states that each regular matroid can be obtained by taking 1-, 2-, and 3-sums from graphic and cographic matroids and from copies of a 10-element matroid called R_{10}. (We do not use this theorem in this book. Background can be found in the book of Truemper [1992].)

Algebraic matroids (Steinitz [1910]). Let L be a field extension of a field K and let S be a finite subset of L. Let $\mathcal{I}$ be the collection of all subsets $\{s_1, \ldots, s_n\}$ of S that consist of algebraically independent elements over K. That is, there is no nonzero polynomial $p(x_1, \ldots, x_n) \in K[x_1, \ldots, x_n]$ with $p(s_1, \ldots, s_n) = 0$. Then $(S, \mathcal{I})$ is a matroid, and matroids arising in this way are called *algebraic* (*over* K). (Steinitz [1910] showed that $(S, \mathcal{I})$ satisfies the matroid axioms, although the term matroid was not yet introduced.)

To see that $(S, \mathcal{I})$ is a matroid, we check (39.3). It suffices to show that for all $s_1, \ldots, s_n \in S$ one has:

(39.15) if $\{s_1, s_2, s_3, \ldots, s_{n-1}\} \in \mathcal{I}$ and $\{s_3, \ldots, s_{n-1}, s_n\} \in \mathcal{I}$, then $\{s_1, s_3, \ldots, s_n\} \in \mathcal{I}$ or $\{s_2, s_3, \ldots, s_n\} \in \mathcal{I}$.

Suppose not. Then there exist nonzero polynomials $p(x_1, x_3, \ldots, x_n)$ and $q(x_2, x_3, \ldots, x_n)$ over K with $p(s_1, s_3, \ldots, s_n) = 0$ and $q(s_2, s_3, \ldots, s_n) = 0$. We may assume that p and q are irreducible. Moreover, since $\{s_3, \ldots, s_n\} \in \mathcal{I}$, p and q are relatively prime. Define $F := K(x_1, x_2, \ldots, x_{n-1})$. So p and q belong to the Euclidean ring $F[x_n]$. Let r be the g.c.d. of p and q in $F[x_n]$. As p and q are relatively prime, we know $r \in F$, and hence we may assume $r \in K[x_1, \ldots, x_{n-1}]$. Now $r = \alpha p + \beta q$ for some $\alpha, \beta \in F[x_n]$. So $r(s_1, \ldots, s_{n-1}) = 0$, contradicting the fact that $\{s_1, \ldots, s_{n-1}\} \in \mathcal{I}$. This proves (39.15).

Each linear matroid is algebraic (as we can consider the linear relations between the elements as polynomials of rank 1), while Ingleton [1971] gave an

example of a nonlinear algebraic matroid. Examples of nonalgebraic matroids were given by Ingleton and Main [1975] and Lindström [1984,1986]. The class of algebraic matroids can be easily seen to be closed under taking minors (deletion is direct, while contraction of an element t corresponds to replacing K by $K(t)$), but it is unknown if it is closed under duality.

In fact, for any field K, the class of matroids that are algebraic over K is closed under taking minors, since Lindström [1989] showed that any matroid algebraic over $K(t)$ (for any t), is also algebraic over K.

For an in-depth survey on algebraic matroids, see Oxley [1992].

Graphic matroids (Birkhoff [1935c], Whitney [1935]). Let $G = (V, E)$ be a graph and let $\mathcal{I}$ be the collection of all subsets of E that form a forest. Then $M = (E, \mathcal{I})$ is a matroid. Condition (39.1)(i) is trivial. To see that condition (39.2) holds, let $F \subseteq E$. Then, by definition, each base U of F is an inclusionwise maximal forest contained in F. Hence U forms a spanning tree in each component of the graph (V, F). So U has $|V| - k$ elements, where k is the number of components of (V, F). So each base of F has $|V| - k$ elements, proving (39.2).

The matroid M is called the *cycle matroid* of G, denoted by $M(G)$. Any matroid obtained in this way, or isomorphic to such a matroid, is called a *graphic matroid*.

Trivially, the circuits of $M(G)$, in the matroid sense, are exactly the circuits of G, in the graph sense. The bases of $M(G)$ are exactly the inclusionwise maximal forests F of G. So if G is connected, the bases are the spanning trees.

The rank function of $M(G)$ can be described as follows. For each subset F of E, let $\kappa(V, F)$ denote the number of components of the graph (V, F). Then for each $F \subseteq E$:

$$r_{M(G)}(F) = |V| - \kappa(V, F). \tag{39.16}$$

Note that deletion and contraction in the matroid correspond to deletion and contraction of edges in the graph.

Graphic matroids are regular, that is, representable over any field: orient the edges of G arbitrarily, and consider the $V \times E$ matrix L given by: $L_{v,e} = +1$ if v is the head of e, $L_{v,e} := -1$ if v is the tail of e, and $L_{v,e} := 0$ otherwise (for $v \in V$, $e \in E$). Then a subset F of E is a forest if and only if the set of columns with index in F is linearly independent.

By a theorem of Tutte [1959], the graphic matroids are precisely those regular matroids containing no $M(K_5)^*$ and $M(K_{3,3})^*$ minor. (Alternative proofs were given by Ghouila-Houri [1964] (Chapitre III), Seymour [1980d], Truemper [1985], Wagner [1985], and Gerards [1995b].)

Cographic matroids (Whitney [1935]). The dual of the cycle matroid $M(G)$ of a graph $G = (V, E)$ is called the *cocycle matroid* of G, and denoted by $M^*(G)$. Any matroid obtained in this way, or isomorphic to such a matroid, is called a *cographic matroid*.

So the bases of $M^*(G)$ are the complements of maximal forests of G. (So if G is connected, these are exactly the complements of the spanning trees in G.)

Hence the independent sets are those edge sets F for which $E \setminus F$ contains a maximal forest of G; that is, $(V, E \setminus F)$ has the same number of components as G.

A subset C of E is a circuit of $M^*(G)$ if and only if C is an inclusionwise minimal set with the property that $(V, E \setminus C)$ has more components than G. Hence C is a circuit of $M^*(G)$ if and only if C is an inclusionwise minimal nonempty cut in G.

The rank function of $M^*(G)$ can be described as follows. Again, for each subset F of E, let $\kappa(V, F)$ denote the number of components of the graph (V, F). Then (39.6) and (39.16) give that for each $F \subseteq E$:

$$r_{M^*(G)}(F) = |F| - \kappa(V, E \setminus F) + \kappa(V, E). \tag{39.17}$$

Let G be an (embedded) planar graph, and let G^* be the dual planar graph of G. Then the cycle matroid $M(G^*)$ of G^* is isomorphic to the cocycle matroid $M^*(G)$ of G.

A theorem of Whitney [1933] implies that a matroid is both graphic and cographic if and only if it is isomorphic to the cycle matroid of a planar graph.

Transversal matroids (Edmonds and Fulkerson [1965], Mirsky and Perfect [1967]). Let $\mathcal{X} = (X_1, \ldots, X_n)$ be a family of subsets of a finite set S and let $\mathcal{I}$ be the collection of all partial transversals of $\mathcal{X}$. Then $M = (S, \mathcal{I})$ is a matroid, as follows directly from Corollary 22.4a. Any matroid obtained in this way, or isomorphic to such a matroid, is called a *transversal matroid* (*induced by* $\mathcal{X}$).

The bases of this matroid are the inclusionwise maximal partial transversals. If $\mathcal{X}$ has a transversal, the bases of M are the transversals of $\mathcal{X}$. In fact, Theorem 22.5 implies that we can assume the latter situation:

(39.18) Let M be the transversal matroid induced by the family $\mathcal{X}$. Then $\mathcal{X}$ has a subfamily $\mathcal{Y}$ such that M is equal to the transversal matroid induced by $\mathcal{Y}$ and such that $\mathcal{Y}$ has a transversal.

So we can assume that any transversal matroid has the transversals of a family of sets as bases.

It follows from Kőnig's matching theorem that the rank function r of the transversal matroid induced by $\mathcal{X}$ is given by

$$\begin{aligned} r(U) &= \min_{T \subseteq U} (|U \setminus T| + |\{i \mid X_i \cap T \neq \emptyset\}|) \\ &= \min_{I \subseteq \{1,\ldots,n\}} (n - |I| + \Big|\bigcup_{i \in I} (X_i \cap U)\Big|) \end{aligned} \tag{39.19}$$

for $U \subseteq S$. This follows directly from Theorem 22.2 and Corollary 22.2a, applied to the family $(X_1 \cap U, \ldots, X_n \cap U)$.

Piff and Welsh [1970] (cf. Atkin [1972]) showed that

(39.20) any transversal matroid is representable over all fields, except for finitely many finite fields.

If the sets $X_1, \ldots, X_m$ form a partition of S, one speaks of a *partition matroid*. Trivially, each partition matroid is graphic and cographic (by considering a graph consisting of vertex-disjoint parallel classes of edges). Also uniform matroids are special cases of transversal matroids.

Gammoids (Perfect [1968]). An extension of transversal matroids is obtained by taking a directed graph $D = (V, A)$ and subsets U and S of V. For $X, Y \subseteq V$, call X *linked to* Y if $|X| = |Y|$ and D has $|X|$ vertex-disjoint $X - Y$ paths. (So X is the set of starting vertices of these paths, and Y the set of end vertices.)

Let $\mathcal{I}$ be the collection of subsets I of S such that some subset of U is linked to I. Then $M = (S, \mathcal{I})$ is a matroid. This follows from Theorem 9.11: let $I, J \in \mathcal{I}$ with $|I| < |J|$. Let $T := I \cup J$. Let k be the maximum number of disjoint $U - T$ paths. So $k \geq |J| > |I|$. By Theorem 9.11, there exist k disjoint $U - T$ paths covering I. Hence $I + j \in \mathcal{I}$ for some $j \in J \setminus I$. So M is a matroid.

Matroids obtained in this way are called *gammoids*. If $S = V$, the gammoid is called a *strict gammoid* (*induced by* D, U). Hence:

(39.21) gammoids are exactly the restrictions of strict gammoids.

The bases of the strict gammoid induced by D, U are the subsets B of V such that U is linked to B. In particular, U is a base.

From Menger's theorem (Corollary 9.1a) one easily derives the following formula for the rank function r_M of M:

$$r_M(X) = \min\{|Y| \mid Y \text{ intersects each } U - X \text{ path}\} \tag{39.22}$$

for $X \subseteq S$. (One may prove easily that the right-hand side of (39.22) satisfies Theorem 39.8 below, thus proving again that M is a matroid.)

39.4a. Relations between transversal matroids and gammoids

Ingleton and Piff [1973] showed the following theorem (based on a duality of bipartite graphs and directed graphs similar to that described in Section 16.7c). The proof provides an alternative proof that gammoids are indeed matroids.

Theorem 39.5. *Strict gammoids are exactly the duals of the transversal matroids.*

Proof. Let M be the strict gammoid induced by the directed graph $D = (V, A)$ and $U \subseteq V$. We can assume that $(v, v) \in A$ for each $v \in V$. For each $v \in V$, let

$$X_v := \{u \in V \mid (u, v) \in A\}. \tag{39.23}$$

Let L be the transversal matroid induced by the family $\mathcal{X} := (X_v \mid v \in V \setminus U)$. We show that $L = M^*$.

As $v \in X_v$ for each $v \in V \setminus U$, the set $V \setminus U$ is a transversal of $\mathcal{X}$. Hence the bases of L are the transversals of $\mathcal{X}$. As U is a base of the strict gammoid induced by D, U, it suffices to show, for each $B \subseteq V$:

(39.24) U is linked to B in D if and only if $V \setminus B$ is a transversal of $\mathcal{X}$.

To see necessity in (39.24), let U be linked to B in D and let $\mathcal{P}$ be a set of $|U|$ disjoint $U - B$ paths. Then for each $v \in V \setminus U$, let $x_v := u$ if v is entered by an arc (u, v) in a path P in $\mathcal{P}$ and let $x_v := v$ otherwise. Then:

(39.25) (i) $x_v \in X_v$, (ii) $x_v \neq x_{v'}$ for $v \neq v' \in V \setminus U$, and (iii) $\{x_v \mid v \in V \setminus U\} = V \setminus B$.

So $V \setminus B$ is a transversal of $\mathcal{X}$.

To see sufficiency in (39.24), let $V \setminus B$ be a transversal of $\mathcal{X}$. Hence there exist x_v for $v \in V \setminus U$ satisfying (39.25). Let A' be the set of arcs (x_v, v) of D with $v \in V \setminus U$. Then $V \setminus U$ is the set of vertices entered by an arc in A', and $V \setminus B$ is the set of vertices left by an arc in A'. Hence U is linked to B in D.

This shows (39.24), and hence that $M^* = L$. So the dual of a strict gammoid is a transversal matroid.

To see that each transversal matroid is the dual of a strict gammoid, we show that the construction described above can be reversed. Let L be the transversal matroid induced by the family $\mathcal{X} = (X_i \mid i = 1, \ldots, m)$ of sets. By (39.18) we can assume that $\mathcal{X}$ has a transversal. Hence we can assume that $i \in X_i$ for $i = 1, \ldots, m$ (by renaming). Let $V := X_1 \cup \cdots \cup X_m$ and let

(39.26) $A := \{(u, v) \mid v \in \{1, \ldots, m\}, u \in X_v\}$.

Let $D = (V, A)$ and define $U := V \setminus \{1, \ldots, m\}$. Since D, U and $\mathcal{X}$ are related as in (39.23), we again have (39.25). So L is equal to the dual of the strict gammoid induced by D, U. ∎

This theorem has a number of implications for the interrelations of the classes of transversal matroids and of gammoids. Consider the following class of matroids, introduced by Ingleton and Piff [1973]. Let $G = (V, E)$ be a bipartite graph, with colour classes U and W. Let $M = (V, \mathcal{I})$ be the transversal matroid induced by the family $(\{v\} \cup N(v) \mid v \in U)$ (where $N(v)$ is the set of neighbours of v). So $B \subseteq V$ is a base of M if and only if $(U \setminus B) \cup (W \cap B)$ is matchable in G (that is, it induces a subgraph of G having a perfect matching).

Any such matroid M is called a *deltoid* (*induced by* G, U, W). Then M^* is the deltoid induced by G, W, U. So

(39.27) the dual of a deltoid is a deltoid again.

Now

(39.28) transversal matroids are exactly those matroids that are the restriction of a deltoid.

Indeed, each deltoid is a transversal matroid, and hence the restriction of any deltoid is a transversal matroid (as the class of transversal matroids is closed under taking restrictions). Conversely, any transversal matroid, induced by (say) $X_1, \ldots, X_m$ is

the restriction to W of the deltoid induced by the bipartite graph G with colour classes $U := \{1, \dots, m\}$ and $W := X_1 \cup \cdots \cup X_m$, with $i \in U$ and $x \in W$ adjacent if and only if $x \in X_i$. (Assuming without loss of generality that $U \cap W = \emptyset$.) This shows (39.28).

Then (39.27) and (39.28) give with Theorem 39.5:

(39.29) the strict gammoids are exactly the contractions of the deltoids.

Indeed, the strict gammoids are the duals of transversal matroids, hence the duals of restrictions of deltoids, and therefore the contractions of (the duals of) deltoids.

This gives:

Corollary 39.5a. *The gammoids are exactly the contractions of the transversal matroids.*

Proof. Gammoids are the restrictions of strict gammoids, hence the restrictions of contractions of deltoids, hence the contractions of restrictions of deltoids, therefore the contractions of transversal matroids. ▮

Similarly:

(39.30) the gammoids are exactly the minors of deltoids,

which implies (with (39.27)) a result of Mason [1972]:

(39.31) the class of gammoids is closed under taking minors and duals.

Theorem 39.5 also implies, with (39.20), that gammoids are representable over all fields, except for a finite number of finite fields (Mason [1972]). In fact, Lindström [1973] showed that any gammoid $(S, \mathcal{I})$ is representable over each field with at least $2^{|S|}$ elements.

Edmonds and Fulkerson [1965] showed that one gets a transversal matroid as follows. Let $G = (V, E)$ be an undirected graph and let $S \subseteq V$. Let $\mathcal{I}$ be the collection of subsets of S which are covered by some matching in G. Then $M = (S, \mathcal{I})$ is a matroid (which is easy to show), called the *matching matroid* of G. In fact, any matching matroid is a transversal matroid. To prove this, we may assume $S = V$. Let $D(G)$, $A(G)$, $C(G)$ form the Edmonds-Gallai decomposition of G (Section 24.4b). Let $\mathcal{K}$ be the collection of components of $G[D(G)]$. Let $\mathcal{X}$ be the family of sets

(39.32)
$$\begin{array}{ll} \{v\} & \text{for each } v \in A(G) \cup C(G), \\ N(v) \cap D(G) & \text{for each } v \in A(G), \\ K, \text{ repeated } |K| - 1 \text{ times}, & \text{for each } K \in \mathcal{K}. \end{array}$$

Then M is equal to the transversal matroid induced by $\mathcal{X}$, as is easy to derive from the properties of the Edmonds-Gallai decomposition. A min-max relation for the rank function is given by Theorem 24.6.

It is straightforward to see that, conversely, each transversal matroid is a matching matroid, by taking G bipartite.

39.5. Characterizing matroids by bases

In Section 39.1, the notion of matroid is defined by 'axioms' in terms of the independent sets. There are several other axiom systems that characterize matroids. In this and the next sections we give a number of them.

Clearly, a matroid is determined by the collection of its bases, since a set is independent if and only if it is contained in a base. Conditions characterizing a collection of bases of a matroid are given in the following theorem (Whitney [1935]).

Theorem 39.6. *Let S be a set and let $\mathcal{B}$ be a nonempty collection of subsets of S. Then the following are equivalent:*

(39.33)
- (i) *$\mathcal{B}$ is the collection of bases of a matroid;*
- (ii) *if $B, B' \in \mathcal{B}$ and $x \in B' \setminus B$, then $B' - x + y \in \mathcal{B}$ for some $y \in B \setminus B'$;*
- (iii) *if $B, B' \in \mathcal{B}$ and $x \in B' \setminus B$, then $B - y + x \in \mathcal{B}$ for some $y \in B \setminus B'$.*

Proof. (i)⇒(ii): Let $\mathcal{B}$ be the collection of bases of a matroid $(S, \mathcal{I})$. Then all sets in $\mathcal{B}$ have the same size. Now let $B, B' \in \mathcal{B}$ and $x \in B' \setminus B$. Since $B' - x \in \mathcal{I}$, there exists a $y \in B \setminus B'$ with $B'' := B' - x + y \in \mathcal{I}$. Since $|B''| = |B'|$, we know $B'' \in \mathcal{B}$.

(iii)⇒(i): (iii) directly implies that no set in $\mathcal{B}$ is contained in another. Let $\mathcal{I}$ be the collection of sets I with $I \subseteq B$ for some $B \in \mathcal{B}$. We check (39.3). Let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Let $I \setminus J = \{x\}$.

Consider sets $B, B' \in \mathcal{B}$ with $I \subseteq B$, $J \subseteq B'$. If $x \in B'$, we are done. So assume $x \notin B'$. Then by (iii), $B' - y + x \in \mathcal{B}$ for some $y \in B' \setminus B$. As $|J \setminus I| = 2$, there is a $z \in J \setminus I$ with $z \neq y$. Then $I + z \subseteq B' - y + x$, and so $I + z \in \mathcal{I}$.

(ii)⇒(iii): By the foregoing we know that (iii) implies (ii). Now axioms (ii) and (iii) interchange if we replace $\mathcal{B}$ by the collection of complements of sets in $\mathcal{B}$. Hence also the implication (ii)⇒(iii) holds. ∎

The equivalence of (ii) and (iii) also follows from the fact that the collection of complements of bases of a matroid is the collection of bases of the dual matroid. Conversely, Theorem 39.6 implies that the dual indeed is a matroid.

39.6. Characterizing matroids by circuits

A matroid is determined by the collection of its circuits, since a set is independent if and only if it contains no circuit. Conditions characterizing a collection of circuits of a matroid are given in the following theorem (Whitney

[1935] proved (i)⇔(iii), and Robertson and Weston [1958] (and also Lehman [1964] and Asche [1966]) proved (i)⇔(ii)).

Theorem 39.7. *Let S be a set and let $\mathcal{C}$ be a collection of nonempty subsets of S, such that no two sets in $\mathcal{C}$ are contained in each other. Then the following are equivalent:*

(39.34)
- (i) *$\mathcal{C}$ is the collection of circuits of a matroid;*
- (ii) *if $C, C' \in \mathcal{C}$ with $C \neq C'$ and $x \in C \cap C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$;*
- (iii) *if $C, C' \in \mathcal{C}$, $x \in C \cap C'$, and $y \in C \setminus C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$ containing y.*

Proof. (i)⇒(iii): Let $\mathcal{C}$ be the collection of circuits of a matroid $(S, \mathcal{I})$ and let $\mathcal{B}$ be its collection of bases. Let $C, C' \in \mathcal{C}$, $x \in C \cap C'$, and $y \in C \setminus C'$. We can assume that $S = C \cup C'$. Let $B, B' \in \mathcal{B}$ with $B \supseteq C - y$ and $B' \supseteq C' - x$. Then $y \notin B$ and $x \notin B'$ (since $C \not\subseteq B$ and $C' \not\subseteq B'$).

We can assume that $y \notin B'$. Otherwise, $y \in B' \setminus B$, and hence by (ii) of Theorem 39.6, there exists a $z \in B \setminus B'$ with $B'' := B' - y + z \in \mathcal{B}$. Then $z \neq x$, since otherwise $C' \subseteq B''$. Hence, replacing B' by B'' gives $y \notin B'$.

As $y \notin B'$, we know $B' \cup \{y\} \notin \mathcal{I}$, and hence there exists a $C'' \in \mathcal{C}$ contained in $B' \cup \{y\}$. As $C'' \not\subseteq B'$, we know $y \in C''$. Moreover, as $x \notin B'$ we know $x \notin C''$.

(iii)⇒(ii): is trivial.

(ii)⇒(i): Let $\mathcal{I}$ be the collection of sets containing no set in $\mathcal{C}$ as a subset. We check (39.3). Let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Assume that $I + z \notin \mathcal{I}$ for each $z \in J \setminus I$. Let y be the element of $I \setminus J$. If $J + y \in \mathcal{I}$, then $I \cup J \in \mathcal{I}$, contradicting our assumption. So $J + y$ contains a set $C \in \mathcal{C}$. Then C is the unique set in $\mathcal{C}$ contained in $J + y$. For suppose that there is another, C' say. Again, $y \in C'$, and hence by (39.34)(ii) there exists a $C'' \in \mathcal{C}$ contained in $(C \cup C') \setminus \{y\}$. But then $C'' \subseteq J$, a contradiction.

As $C \not\subseteq I$, C intersects $J \setminus I$. Choose $x \in C \cap (J \setminus I)$. Then $X := J + y - x$ contains no set in $\mathcal{C}$ (as C is the only set in $\mathcal{C}$ contained in $J + y$). So $X \in \mathcal{I}$, implying that $I + z \in \mathcal{I}$ for the $z \in J \setminus I$ with $z \neq x$. ∎

This theorem implies the following important property for a matroid $M = (S, \mathcal{I})$:

(39.35) for any independent set I and any $s \in S \setminus I$ there is at most one circuit contained in $I \cup \{s\}$.

39.6a. A characterization of Lehman

Lehman [1964] showed that the cocircuits of a matroid M are exactly the inclusionwise minimal nonempty subsets D of S with $|D \cap C| \neq 1$ for each circuit C of M.

To show this, it suffices to show that

(39.36) (i) $|D \cap C| \neq 1$ for each cocircuit D and circuit C,
(ii) for each nonempty $D \subseteq S$, if $|D \cap C| \neq 1$ for each circuit C, then D contains a cocircuit; that is, then D is dependent in M^*.

To see (i), suppose that $D \cap C = \{s\}$ for some circuit C and cocircuit D. As $D - s$ is independent in M^*, M has a base B disjoint from $D - s$. Since $C - s$ is disjoint from $D - s$ and since $C - s \in \mathcal{I}$, we can assume that $C - s \subseteq B$. Then $s \notin B$, and so B is disjoint from D. This implies that D is independent in M^*, contradicting the fact that D is a circuit in M^*. This shows (i).

To see (ii), let $\emptyset \neq D \subseteq S$ with $|D \cap C| \neq 1$ for each circuit C. We show that D is dependent in M^*. Suppose not. Then M has a base B disjoint from D. Choose $s \in D$. Then $B + s$ contains a circuit C with $s \in C$. Hence $D \cap C = \{s\}$, contradicting our assumption, thus showing (ii).

39.7. Characterizing matroids by rank functions

The *rank function* of a matroid $M = (S, \mathcal{I})$ is the function $r_M : \mathcal{P}(S) \to \mathbb{Z}_+$ given by:

$$r_M(U) := \max\{|Z| \mid Z \in \mathcal{I}, Z \subseteq U\} \tag{39.37}$$

for $U \subseteq S$. Again, a matroid is determined by its rank function, as a set U is independent if and only if $r(U) = |U|$. Conditions characterizing a rank function are given by the following theorem (Whitney [1935]; necessity was also shown (in a different terminology) by Bergmann [1929] and Nakasawa [1935]):

Theorem 39.8. *Let S be a set and let $r : \mathcal{P}(S) \to \mathbb{Z}_+$. Then r is the rank function of a matroid if and only if for all $T, U \subseteq S$:*

(39.38) (i) $r(T) \leq r(U) \leq |U|$ *if* $T \subseteq U$,
(ii) $r(T \cap U) + r(T \cup U) \leq r(T) + r(U)$.

Proof. *Necessity.* Let r be the rank function of a matroid $(S, \mathcal{I})$. Choose $T, U \subseteq S$. Clearly (39.38)(i) holds. To see (ii), let I be an inclusionwise maximal set in $\mathcal{I}$ with $I \subseteq T \cap U$ and let J be an inclusionwise maximal set in $\mathcal{I}$ with $I \subseteq J \subseteq T \cup U$. Since $(S, \mathcal{I})$ is a matroid, we know that $r(T \cap U) = |I|$ and $r(T \cup U) = |J|$. Then

$$\begin{aligned} r(T) + r(U) &\geq |J \cap T| + |J \cap U| = |J \cap (T \cap U)| + |J \cap (T \cup U)| \\ &\geq |I| + |J| = r(T \cap U) + r(T \cup U); \end{aligned} \tag{39.39}$$

that is, we have (39.38)(ii).

Sufficiency. Let $\mathcal{I}$ be the collection of subsets I of S with $r(I) = |I|$. We show that $(S, \mathcal{I})$ is a matroid, with rank function r.

Trivially, $\emptyset \in \mathcal{I}$. Moreover, if $I \in \mathcal{I}$ and $J \subseteq I$, then

(39.40) $$r(J) \geq r(I) - r(I \setminus J) \geq |I| - |I \setminus J| = |J|.$$

So $J \in \mathcal{I}$.

In order to check (39.3), let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Let $J \setminus I = \{z_1, z_2\}$. If $I + z_1, I + z_2 \notin \mathcal{I}$, we have $r(I + z_1) = r(I + z_2) = |I|$. Then by (39.38)(ii),

(39.41) $$r(J) \leq r(I + z_1 + z_2) \leq r(I + z_1) + r(I + z_2) - r(I) = |I| < |J|,$$

contradicting the fact that $J \in \mathcal{I}$.

So $(S, \mathcal{I})$ is a matroid. Its rank function is r, since $r(U) = \max\{|I| \mid I \subseteq U, I \in \mathcal{I}\}$ for each $U \subseteq S$. Here $\geq$ follows from (39.38)(i), since if $I \subseteq U$ and $I \in \mathcal{I}$, then $r(U) \geq r(I) = |I|$. Equality can be shown by induction on $|U|$, the case $U = \emptyset$ being trivial. If $U \neq \emptyset$, choose $y \in U$. By induction, there is an $I \subseteq U - y$ with $I \in \mathcal{I}$ and $|I| = r(U - y)$. If $r(U) = r(U - y)$ we are done, so assume $r(U) > r(U - y)$. Then $I + y \in \mathcal{I}$, since $r(I + y) \geq r(I) + r(U) - r(U - y) \geq |I| + 1$. Moreover, $r(U) \leq r(U - y) + r(\{y\}) \leq |I| + 1$. This proves equality for U. ∎

Set functions satisfying condition (39.38)(ii) are called *submodular*, and will be studied in Chapter 44.

Whitney [1935] also showed that (39.38) is equivalent to:

(39.42)
(i) $r(\emptyset) = 0$,
(ii) $r(U) \leq r(U + s) \leq r(U) + 1$ for $U \subseteq S$, $s \in S \setminus U$,
(iii) for all $U \subseteq S$, $s, t \in S \setminus U$, if $r(U + s) = r(U + t) = r(U)$, then $r(U + s + t) = r(U)$.

The proof above in fact uses only these properties of r.

The following equivalent form of Theorem 39.8 will be useful.

Corollary 39.8a. *Let S be a finite set and let $\mathcal{I}$ be a nonempty collection of subsets of S, closed under taking subsets. For $U \subseteq S$, let $r(U)$ be the maximum size of a subset of U that belongs to $\mathcal{I}$. Then $(S, \mathcal{I})$ is a matroid if and only if r satisfies* (39.38)(ii) *for all $T, U \subseteq S$.*

Proof. Necessity follows directly from Theorem 39.8. To see sufficiency, it is easy to see that r satisfies (39.38)(i). So by Theorem 39.8, r is the rank function of some matroid $M = (S, \mathcal{J})$. Now: $I \in \mathcal{J} \iff r(I) = |I| \iff I \in \mathcal{I}$. Hence $\mathcal{I} = \mathcal{J}$, and so $(S, \mathcal{I})$ is a matroid. ∎

Note that if we can test in polynomial time if a given set is independent, we can also test in polynomial time if a given set is a base, or a circuit, and we can determine the rank of a given set in polynomial time.

39.8. The span function and flats

With any matroid $M = (S, \mathcal{I})$ we can define the *span function* $\mathrm{span}_M : \mathcal{P}(S) \to \mathcal{P}(S)$ as follows:

$$\mathrm{span}_M(T) := \{s \in S \mid r_M(T \cup \{s\}) = r_M(T)\} \tag{39.43}$$

for $T \subseteq S$. If the matroid M is clear from the context, we write $\mathrm{span}(T)$ for $\mathrm{span}_M(T)$. Note that $T \subseteq \mathrm{span}_M(T)$ and that

$$r_M(\mathrm{span}_M(T)) = r_M(T). \tag{39.44}$$

This follows directly from the fact that if $r_M(Y) > r_M(T)$, then $r_M(T \cup \{s\}) > r_M(T)$ for some $s \in Y$.

Note also that

$$T \text{ is spanning} \iff \mathrm{span}_M(T) = S \tag{39.45}$$

for any $T \subseteq S$. To see $\Longrightarrow$, let T be spanning. Then for each $s \in T$: $r_M(T + s) \leq r_M(S) = r_M(T)$. To see $\Longleftarrow$, suppose $\mathrm{span}_M(T) = S$. Then $r_M(T) = r_M(\mathrm{span}_M(T)) = r_M(S)$.

A *flat* in a matroid $M = (S, \mathcal{I})$ is a subset F of S with $\mathrm{span}_M(F) = F$. A matroid is determined by its collection of flats, as is shown by:

(39.46) a subset I of S is independent if and only if for each $y \in I$ there is a flat F with $I - y \subseteq F$ and $y \notin F$.

Indeed, if I is independent and $y \in I$, let $F := \mathrm{span}_M(I - y)$. Then F is a flat containing $I - y$, but not y, since $r_M(F + y) \geq r_M(I) > r_M(I - y) = r_M(F)$. Conversely, if I is not independent, then $y \in \mathrm{span}_M(I - y)$ for some $y \in I$, and hence each flat containing $I - y$ also contains y.

39.8a. Characterizing matroids by span functions

It was observed by Mac Lane [1938] that the following characterizes span functions of matroids (sufficiency was shown by van der Waerden [1937]).

Theorem 39.9. *Let S be a finite set. A function* $\mathrm{span} : \mathcal{P}(S) \to \mathcal{P}(S)$ *is the span function of a matroid if and only if:*

(39.47)
- (i) *if $T \subseteq S$, then $T \subseteq \mathrm{span}(T)$;*
- (ii) *if $T, U \subseteq S$ and $U \subseteq \mathrm{span}(T)$, then $\mathrm{span}(U) \subseteq \mathrm{span}(T)$;*
- (iii) *if $T \subseteq S$, $t \in S \setminus T$, and $s \in \mathrm{span}(T + t) \setminus \mathrm{span}(T)$, then $t \in \mathrm{span}(T + s)$.*

Proof. *Necessity.* Let span be the span function of a matroid $M = (S, \mathcal{I})$ with rank function r. Clearly, (39.47)(i) is satisfied. To see (39.47)(ii), let $U \subseteq \mathrm{span}(T)$ and $s \in \mathrm{span}(U)$. We show $s \in \mathrm{span}(T)$. We can assume $s \notin T$. Then, by the submodularity of r,

$$\begin{aligned} &r(T \cup \{s\}) \leq r(T \cup U \cup \{s\}) \leq r(T \cup U) + r(U \cup \{s\}) - r(U) \\ &= r(T \cup U) = r(T). \end{aligned} \tag{39.48}$$

(The last equality follows from (39.44).) This shows that $s \in \operatorname{span}(T)$.

To see (39.47)(iii), note that $s \in \operatorname{span}(T+t) \setminus \operatorname{span}(T)$ is equivalent to: $r(T+t+s) = r(T+t)$ and $r(T+s) > r(T)$. Hence

$$r(T+t+s) = r(T+t) \leq r(T)+1 \leq r(T+s), \tag{39.49}$$

that is, $t \in \operatorname{span}(T+s)$. This shows necessity of the conditions (39.47).

Sufficiency. Let a function span satisfy (39.47), and define

$$\mathcal{I} := \{I \subseteq S \mid s \notin \operatorname{span}(I-s) \text{ for each } s \in I\}. \tag{39.50}$$

We first show the following:

$$\text{if } I \in \mathcal{I}, \text{ then } \operatorname{span}(I) = I \cup \{t \mid I+t \notin \mathcal{I}\}. \tag{39.51}$$

Indeed, if $t \in \operatorname{span}(I) \setminus I$, then $I+t \notin \mathcal{I}$, by definition of $\mathcal{I}$. Conversely, $I \subseteq \operatorname{span}(I)$ by (39.47)(i). Moreover, if $I+t \notin \mathcal{I}$, then by definition of $\mathcal{I}$, $s \in \operatorname{span}(I+t-s)$ for some $s \in I+t$. If $s = t$, then $t \in \operatorname{span}(I)$ and we are done. So assume $s \neq t$; that is, $s \in I$. As $I \in \mathcal{I}$, we know that $s \notin \operatorname{span}(I-s)$. So by (39.47)(iii) (for $T := I-s$), $t \in \operatorname{span}(I)$, proving (39.51).

We now show that $M = (S, \mathcal{I})$ is a matroid. Trivially, $\emptyset \in \mathcal{I}$. To see that $\mathcal{I}$ is closed under taking subsets, let $I \in \mathcal{I}$ and $J \subseteq I$. We show that $J \in \mathcal{I}$. Suppose to the contrary that $s \in \operatorname{span}(J-s)$ for some $s \in J$. By (39.47)(ii), $\operatorname{span}(J-s) \subseteq \operatorname{span}(I-s)$. Hence $s \in \operatorname{span}(I-s)$, contradicting the fact that $I \in \mathcal{I}$.

In order to check (39.3), let $I, J \in \mathcal{I}$ with $|I \setminus J| = 1$ and $|J \setminus I| = 2$. Let $I \setminus J = \{i\}$ and $J \setminus I = \{j_1, j_2\}$. Assume that $I + j_1 \notin \mathcal{I}$. That is, $J + i - j_2 \notin \mathcal{I}$, and so, by (39.51) applied to $J - j_2$, $i \in \operatorname{span}(J-j_2)$. Therefore, $I \subseteq \operatorname{span}(J-j_2)$, and so $\operatorname{span}(I) \subseteq \operatorname{span}(J-j_2)$. So $j_2 \notin \operatorname{span}(I)$ (as $J \in \mathcal{I}$), and therefore, by (39.51) applied to I, $I + j_2 \in \mathcal{I}$.

So M is a matroid. We finally show that $\operatorname{span} = \operatorname{span}_M$. Choose $T \subseteq S$. To see that $\operatorname{span}(T) = \operatorname{span}_M(T)$, let I be a base of T (in M). Then (using (39.51)),

$$\operatorname{span}_M(T) = I \cup \{x \mid I + x \notin \mathcal{I}\} = \operatorname{span}(I) \subseteq \operatorname{span}(T). \tag{39.52}$$

So we are done by showing $\operatorname{span}(T) \subseteq \operatorname{span}(I)$; that is, by (39.47)(ii), $T \subseteq \operatorname{span}(I)$. Choose $t \in T \setminus I$. By the maximality of I, we know $I + t \notin \mathcal{I}$, and hence, by (39.51), $t \in \operatorname{span}(I)$. ∎

39.8b. Characterizing matroids by flats

Conditions characterizing collections of flats of a matroid are given in the following theorem (Bergmann [1929]):

Theorem 39.10. *Let S be a set and let $\mathcal{F}$ be a collection of subsets of S. Then $\mathcal{F}$ is the collection of flats of a matroid if and only if:*

(39.53)
- (i) $S \in \mathcal{F}$;
- (ii) *if $F_1, F_2 \in \mathcal{F}$, then $F_1 \cap F_2 \in \mathcal{F}$;*
- (iii) *if $F \in \mathcal{F}$ and $t \in S \setminus F$, and F' is the smallest flat containing $F+t$, then there is no flat F'' with $F \subset F'' \subset F'$.*

Proof. *Necessity.* Let $\mathcal{F}$ be the collection of flats of a matroid $M = (S, \mathcal{I})$. Condition (39.53)(i) is trivial, and condition (39.53)(ii) follows from $\operatorname{span}_M(F_1 \cap F_2) \subseteq \operatorname{span}_M(F_1) \cap \operatorname{span}_M(F_2) = F_1 \cap F_2$. To see (39.53)(iii), suppose that such an F'' exists. Choose $s \in F'' \setminus F$. So $s \notin \operatorname{span}_M(F)$. As $F' \not\subseteq F''$, we have $t \notin \operatorname{span}_M(F+s)$. Therefore, by (39.47)(iii) for $T := F$, $s \notin \operatorname{span}_M(F) = F'$, a contradiction.

Sufficiency. Let $\mathcal{F}$ satisfy (39.53). For $Y \subseteq S$, let $\operatorname{span}(Y)$ be the smallest set in $\mathcal{F}$ containing Y. Since $F \in \mathcal{F} \iff \operatorname{span}(F) = F$, it suffices to show that span satisfies the conditions (39.47). Here (39.47)(i) and (ii) are trivial. To see (39.47)(iii), let $T \subseteq S$, $t \in S \setminus T$, and $s \in \operatorname{span}(T+t) \setminus \operatorname{span}(T)$. Then $\operatorname{span}(T) \subset \operatorname{span}(T+s) \subseteq \operatorname{span}(T+t)$. Hence, by (39.53)(iii), $\operatorname{span}(T+s) = \operatorname{span}(T+t)$, and hence $t \in \operatorname{span}(T+s)$. ∎

39.8c. Characterizing matroids in terms of lattices

Bergmann [1929] and Birkhoff [1935a] characterized matroids in terms of lattices. A partially ordered set $(L, \leq)$ is called a *lattice* if

(39.54)
- (i) for all $A, B \in L$ there is a unique element, called $A \wedge B$, satisfying $A \wedge B \leq A, B$ and $C \leq A \wedge B$ for all $C \leq A, B$;
- (ii) for all $A, B \in L$ there is a unique element, called $A \vee B$, satisfying $A \vee B \geq A, B$ and $C \geq A \vee B$ for all $C \geq A, B$.

$A \wedge B$ and $A \vee B$ are called the *meet* and *join* respectively of A and B. Here we assume lattices to be finite. Then a lattice has a unique minimal element, denoted by 0. The *rank* of an element A is the maximum number n of elements $x_1, \ldots, x_n$ with $0 < x_1 < \cdots < x_n = A$. An element of rank 1 is called a *point* or *atom*.

Call a lattice a *point lattice* if each element is a join of points, and a *matroid lattice* (or a *geometric lattice*) if it is isomorphic to the lattice of flats of a matroid. Trivially, each matroid lattice is a point lattice. Moreover, a matroid without loops and parallel elements is completely determined by the lattice of flats.

In the following theorem, the equivalence of (i) and (ii), and the implication (ii)⇒(iv) are due (in a different terminology) to Bergmann [1929]; the equivalence of (iii) and (iv) was shown by Birkhoff [1933], and the implication (iii)⇒(i) was shown by Birkhoff [1935a].

In a partially ordered set $(L, \leq)$ an element y is said to *cover* an element x if $x < y$ and there is no z with $x < z < y$.

Theorem 39.11. *For any finite point lattice $(L, \leq)$, with rank function r, the following are equivalent:*

(39.55)
- (i) *L is a matroid lattice;*
- (ii) *for each $a \in L$ and each point p, if $p \not\leq a$, then $a \vee p$ covers a;*
- (iii) *for each $a, b \in L$, if a and b cover $a \wedge b$, then $a \vee b$ covers a and b;*
- (iv) *$r(a) + r(b) \geq r(a \vee b) + r(a \wedge b)$ for all $a, b \in L$.*

Proof. (i)⇒(iv): Let L be the lattice of flats of a matroid $M = (S, \mathcal{I})$, with rank function r_M. We can assume that M has no loops and no parallel elements. Then for any flat F we have $r(F) = r_M(F)$, since $r_M(F)$ is equal to the maximum number k of nonempty flats $F_1 \subset \cdots \subset F_k$ with $F_k = F$. So (iv) follows from Theorem 39.8.

(iv)⇒(iii): We first show that (iv) implies that if b covers a, then $r(b) = r(a)+1$. As b is a join of points, and as b covers a, we know that $b = a \vee p$ for some point p with $p \not\leq a$. Hence $r(b) = r(a \vee p) \leq r(a) + r(p) - r(a \wedge p) = r(a) + r(p) - r(0) = r(a) + 1$. As $r(b) > r(a)$, we have $r(b) = r(a) + 1$.

To derive (iii) from (iv), let a and b cover $a \wedge b$. Then $r(a) = r(b) = r(a \wedge b) + 1$. Hence $r(a \vee b) \leq r(a) + r(b) - r(a \wedge b) = r(a) + 1$. Hence $a \vee b$ covers a. Similarly, $a \vee b$ covers b.

(iii)⇒(ii): We derive (ii) from (iii) by induction on $r(a)$. If $a = 0$, the statement is trivial. If $a > 0$, let a' be an element covered by a. Then, by induction, $a' \vee p$ covers a'. So $a' = a \wedge (a' \vee p)$. Hence by (iii), $a \vee (a' \vee p) = a \vee p$ covers a.

(ii)⇒(i): Let S be the set of points of L, and for $f \in L$ define $F_f := \{s \in S \mid s \leq f\}$. Let $\mathcal{F} := \{F_f \mid f \in L\}$. Then for all $f_1, f_2 \in L$ we have:

$$f_1 \leq f_2 \iff F_{f_1} \subseteq F_{f_2}. \tag{39.56}$$

Here $\Longrightarrow$ is trivial, while $\Longleftarrow$ follows from the fact that for each $f \in L$ we have $f = \bigvee F_f$, as L is a point lattice.

By (39.56), $(L, \leq)$ is isomorphic to $(\mathcal{F}, \subseteq)$. Moreover, by (39.54)(i), $F_{f_1 \wedge f_2} = F_{f_1} \cap F_{f_2}$. So $\mathcal{F}$ is closed under intersections, implying (39.53)(ii), while (39.53)(i) is trivial. Finally, (39.53)(iii) follows from (39.55)(ii). ∎

Lattices satisfying (39.55)(iii) are called *upper semimodular*.

39.9. Further exchange properties

In this section we prove a number of exchange properties of bases, as a preparation to the forthcoming sections on matroid intersection algorithms.

An exchange property of bases, stronger than given in Theorem 39.6, is (Brualdi [1969c]):

Theorem 39.12. *Let $M = (S, \mathcal{I})$ be a matroid. Let B_1 and B_2 be bases and let $x \in B_1 \setminus B_2$. Then there exists a $y \in B_2 \setminus B_1$ such that both $B_1 - x + y$ and $B_2 - y + x$ are bases.*

Proof. Let C be the unique circuit in $B_2 + x$ (cf. (39.35)). Then $(B_1 \cup C) - x$ is spanning, since $x \in \operatorname{span}_M(C - x) \subseteq \operatorname{span}_M((B_1 \cup C) - x)$, implying $\operatorname{span}((B_1 \cup C) - x) = \operatorname{span}(B_1 \cup C) = S$.

Hence there is a base B_3 with $B_1 - x \subseteq B_3 \subseteq (B_1 \cup C) - x$. So $B_3 = B_1 - x + y$ for some y in $C - x$. Therefore, $B_2 - y + x$ is a base, as it contains no circuit (since C is the only circuit in $B_2 + x$). ∎

Let $M = (S, \mathcal{I})$ be a matroid. For any $I \in \mathcal{I}$ define the (bipartite) directed graph $D_M(I) = (S, A_M(I))$, or briefly $(S, A(I))$, by:

$$A(I) := \{(y, z) \mid y \in I, z \in S \setminus I, I - y + z \in \mathcal{I}\}. \tag{39.57}$$

Repeated application of the exchange property described in Theorem 39.12 gives (Brualdi [1969c]):

Corollary 39.12a. *Let $M = (S, \mathcal{I})$ be a matroid and let $I, J \in \mathcal{I}$ with $|I| = |J|$. Then $A(I)$ contains a perfect matching on $I \triangle J$.*[1]

Proof. By truncating M, we can assume that I and J are bases of M. We prove the lemma by induction on $|I \setminus J|$. We can assume $|I \setminus J| \geq 1$. Choose $y \in I \setminus J$. By Theorem 39.12, $I - y + z \in \mathcal{I}$ and $J - z + y \in \mathcal{I}$ for some $z \in J \setminus I$. By induction, applied to I and $J' := J - z + y$, $A(I)$ has a perfect matching N on $I \triangle J'$. Then $N \cup \{(y, z)\}$ is a perfect matching on $I \triangle J$. ∎

Corollary 39.12a implies the following characterization of maximum-weight bases:

Corollary 39.12b. *Let $M = (S, \mathcal{I})$ be a matroid, let B be a base of M, and let $w : S \to \mathbb{R}$ be a weight function. Then B is a base of maximum weight $\iff$ $w(B') \leq w(B)$ for every base B' with $|B' \setminus B| = 1$.*

Proof. Necessity being trivial, we show sufficiency. Suppose to the contrary that there is a base B' with $w(B') > w(B)$. Let N be a perfect matching in $A(B)$ covering $B \triangle B'$. As $w(B') > w(B)$, there is an edge (y, z) in N with $w(z) > w(y)$, where $y \in B \setminus B'$ and $z \in B' \setminus B$. Hence $w(B - y + z) > w(B)$, contradicting the condition. ∎

The following forms a counterpart to Corollary 39.12a (Krogdahl [1974, 1976,1977]):

Theorem 39.13. *Let $M = (S, \mathcal{I})$ be a matroid and let $I \in \mathcal{I}$. Let $J \subseteq S$ be such that $|I| = |J|$ and such that $A(I)$ contains a unique perfect matching N on $I \triangle J$. Then J belongs to $\mathcal{I}$.*

Proof. Since N is unique, we can order N as $(y_1, z_1), \ldots, (y_t, z_t)$ such that $(y_i, z_j) \notin A(I)$ if $1 \leq i < j \leq t$. Suppose that $J \notin \mathcal{I}$, and let C be a circuit contained in J. Choose the smallest i with $z_i \in C$. Then $(y_i, z) \notin A(I)$ for all $z \in C - z_i$ (since $z = z_j$ for some $j > i$). Therefore, $z \in \operatorname{span}(I - y_i)$ for all $z \in C - z_i$. So $C - z_i \subseteq \operatorname{span}(I - y_i)$, and therefore $z_i \in C \subseteq \operatorname{span}(C - z_i) \subseteq \operatorname{span}(I - y_i)$, contradicting the fact that $I - y_i + z_i$ is independent. ∎

This implies:

Corollary 39.13a. *Let $M = (S, \mathcal{I})$ be a matroid and let $I \in \mathcal{I}$. Let $J \subseteq S$ be such that $|I| = |J|$ and $r_M(I \cup J) = |I|$, and such that $A(I)$ contains a unique perfect matching N on $I \triangle J$. Let $s \notin I \cup J$ with $I + s \in \mathcal{I}$. Then $J + s \in \mathcal{I}$.*

Proof. Let t be a new element and let $M' = (S \cup \{t\}, \mathcal{I}')$ be the matroid with $F \in \mathcal{I}'$ if and only if $F \setminus \{t\} \in \mathcal{I}$. Then $N' := N \cup \{(t, s)\}$ forms a

[1] A *perfect matching on* a vertex set U in a digraph is a set of vertex-disjoint arcs such that U is the set of tails and heads of these arcs.

unique perfect matching on $(I \triangle J) \cup \{s,t\}$ in $D_{M'}(I \cup \{t\})$ (since there is no arc from t to $J \setminus I$, as $I + j \notin \mathcal{I}$ for all $j \in J \setminus I$, since $r_M(I \cup J) = |I|$). So by Theorem 39.13, $J \cup \{s\}$ is independent in M', and hence in M. ▮

39.9a. Further properties of bases

Bases satisfy the following exchange property, stronger than that described in Theorem 39.12 (conjectured by G.-C. Rota, and proved by Brylawski [1973], Greene [1973], Woodall [1974a]):

(39.58) if B_1 and B_2 are bases and B_1 is partitioned into X_1 and Y_1, then B_2 can be partitioned into X_2 and Y_2 such that $X_1 \cup Y_2$ and $Y_1 \cup X_2$ are bases.

This will be proved in Section 42.1a (using the matroid union theorem).

Other exchange properties of bases were given by Greene [1974a] and Kung [1978a]. Decomposing exchanges was studied by Gabow [1976b].

In Schrijver [1979c] it was shown that the exchange property described in Corollary 16.8b for bipartite graphs and, more generally, in Theorem 9.12 for directed graphs, in fact characterizes systems that correspond to matroids.

To this end, let U and W be disjoint sets and let Λ be a collection of pairs (X,Y) with $X \subseteq U$ and $Y \subseteq W$. Call (U,W,Λ) a *bimatroid* (or *linking system*) if:

(39.59)
(i) $(\emptyset,\emptyset) \in \Lambda$;
(ii) if $(X,Y) \in \Lambda$ and $x \in X$, then $(X-x, Y-y) \in \Lambda$ for some $y \in Y$;
(iii) if $(X,Y) \in \Lambda$ and $y \in Y$, then $(X-x, Y-y) \in \Lambda$ for some $x \in X$;
(iv) if $(X_1,Y_1),(X_2,Y_2) \in \Lambda$, then there is an $(X,Y) \in \Lambda$ with $X_1 \subseteq X \subseteq X_1 \cup X_2$ and $Y_2 \subseteq Y \subseteq Y_1 \cup Y_2$.

Note that (ii) and (iii) imply that $|X| = |Y|$ for each $(X,Y) \in \Lambda$.

To describe the relation with matroids, define:

(39.60) $\mathcal{B} := \{(U \setminus X) \cup Y \mid (X,Y) \in \Lambda\}$.

So $\mathcal{B}$ determines Λ. Then (Schrijver [1979c]):

(39.61) (U,W,Λ) is a bimatroid if and only if $\mathcal{B}$ is the collection of bases of a matroid on $U \cup W$, with $U \in \mathcal{B}$.

So bimatroids are in one-to-one correspondence with pairs (M,B) of a matroid M and a base B of M, and the conditions (39.59) yield a characterization of matroids. An equivalent axiom system characterizing matroids was given by Kung [1978b].

(Bapat [1994] gave an extension of Kőnig's matching theorem to bimatroids.)

39.10. Further results and notes

39.10a. Further notes

Dilworth [1944] showed that if $r : \mathcal{P}(S) \to \mathbb{Z}$ satisfies (39.38) and $r(U) \geq 0$ if $U \neq \emptyset$, then

(39.62) $$\mathcal{I} := \{I \subseteq S \mid \forall \text{ nonempty } U \subseteq I : |U| \leq r(U)\}$$

is the collection of independent sets of a matroid M. Its rank function satisfies:

(39.63) $$r_M(U) = \min(r(U_1) + \cdots + r(U_t)),$$

where the minimum ranges over partitions of U into nonempty subsets $U_1, \ldots, U_t$ ($t \geq 0$). If $G = (V, E)$ is a graph, and we define $r(F) := \left|\bigcup F\right| - 1$ for $F \subseteq E$, we obtain the cycle matroid of G (this also was shown by Dilworth [1944]).[2]

Conforti and Laurent [1988] showed the following sharpening of Corollary 39.8a. Let $\mathcal{C}$ be a collection of subsets of a set S and let $f : \mathcal{C} \to \mathbb{Z}_+$. Let $\mathcal{I}$ be the collection of subsets T of S with $|T \cap U| \leq f(U)$ for each $U \in \mathcal{C}$. For $T \subseteq S$, let $r(T)$ be the maximum size of a subset of T that belongs to $\mathcal{I}$. Then $(S, \mathcal{I})$ is a matroid if and only if r satisfies the submodular inequality (39.38)(ii) for all $Y, Z \in \mathcal{C}$ with $Y \cap Z \neq \emptyset$. In fact, in the right-hand side of this inequality, r may be replaced by f.

Jensen and Korte [1982] showed that there is no polynomial-time algorithm to find the minimum size of a circuit of a matroid, if the matroid is given by an oracle for testing independence. For binary matroids (represented by binary vectors), the problem of finding a minimum-size circuit was shown by Vardy [1997] to be NP-complete (solving a problem of Berlekamp, McEliece, and van Tilborg [1978], who showed the NP-completeness of finding the minimum size of a circuit containing a given element of the matroid, and of finding a circuit of given size). If we know that a matroid is binary, a vector representation can be derived by a polynomially bounded number of calls from an independence testing oracle.

For further studies of the complexity of matroid properties, see Hausmann and Korte [1978], Robinson and Welsh [1980], and Jensen and Korte [1982].

Extensions of matroid theory to infinite structures were considered by Rado [1949a], Bleicher and Preston [1961], Johnson [1961], and Dlab [1962,1965].

Standard references on matroid theory are Welsh [1976] and Oxley [1992]. The book by Truemper [1992] focuses on decomposition of matroids. Earlier texts were given by Tutte [1965a,1971]. Elementary introductions to matroids were given by Wilson [1972b,1973], and a survey with applications to electrical networks and statics by Recski [1989]. Bixby [1982], Faigle [1987], Lee and Ryan [1992], and Bixby and Cunningham [1995] survey matroid optimization and algorithms. White [1986, 1987,1992] offers a collection of surveys on matroids, and Kung [1986] is a source book on matroids. Stern [1999] focuses on semimodular lattices. Books discussing matroid optimization include Lawler [1976b], Papadimitriou and Steiglitz [1982], Gondran and Minoux [1984], Nemhauser and Wolsey [1988], Parker and Rardin [1988], Cook, Cunningham, Pulleyblank, and Schrijver [1998], and Korte and Vygen [2000].

39.10b. Historical notes on matroids

The idea of a matroid, that is, of abstract dependence, seems to have been developed historically along a number of independent lines during the period 1900-1935. Independently, different axiom systems were given, each of which is equivalent to

[2] $\bigcup F$ denotes the union of the edges (as sets) in F.

that of a matroid. It indicates the naturalness of the concept. Only at the end of the 1930s a synthesis of the different streams was obtained.

There is a line, starting with the *Dualgruppen* (dual groups = lattices) of Dedekind [1897,1900], introduced in order to study modules (= additive subgroups) of numbers. They give rise to lattices satisfying what Dedekind called the *Modulgesetz* (module law). Later, independently, Birkhoff [1933] studied such lattices, calling them initially B-lattices, and later (after he had learned about Dedekind's earlier work), *modular lattices*. Both Dedekind and Birkhoff considered, in their studies of modular lattices, an auxiliary property that characterizes so-called *semimodular lattices*. If the lattice is a point lattice (that is, each element of the lattice is a join of atoms (points)), then such semimodular lattices are exactly the lattices of flats of a matroid. This connection was pointed out by Birkhoff [1935a] directly after Whitney's introduction of matroids.

A second line concerns exchange properties of bases. It starts with the new edition of the *Ausdehnungslehre* of Grassmann [1862], where he showed that each linearly independent set can be extended to a bases, using elements from a given base. Next Steinitz [1910], in his fundamental paper *Algebraische Theorie der Körper* (Algebraic Theory of Fields), showed that algebraic dependence has a number of basic properties, which makes it into a matroid (like the equicardinality of bases), and he derived some other properties from these basic properties (thus deriving essentially properties of matroids). In a subsequent paper, Steinitz [1913] gave, as an auxiliary result, the property that is now called *Steinitz' exchange property* for linearly independent sets of vectors. Steinitz did not mention the similarities to his earlier results on algebraic dependence. These similarities were observed by Haupt [1929a] and van der Waerden [1930] in their books on 'modern' algebra. They formulated properties shared by linear and algebraic dependence that are equivalent to matroids. In the second edition of his book, van der Waerden [1937] condensed these properties to three properties, and gave a unified treatment of linear and algebraic dependence. Mac Lane [1938] observed the relation of this work to the work on lattices and matroids.

A third line pursued the axiomatization of geometry, which clearly can be rooted back to as early as Euclid. At the beginning of the 20th century this was considered by, among others, Hilbert and Veblen. Bergmann [1929] aimed at giving a lattice-theoretical basis for affine geometry, and from lattice-theoretical conditions equivalent to matroids (cf. Theorem 39.11 above) he derived a number of properties, like the equicardinality of bases and the submodularity of the rank function. In their book *Grundlagen der Mathematik I* (Foundations of Mathematics I), Hilbert and Bernays [1934] gave axioms for the collinearity of triples of points, amounting to the fact that any two distinct points belong to exactly one line. A direct extension of these axioms to general dimensions gives the axioms described by Nakasawa [1935], that are again equivalent to the matroid axioms. He introduced the concept of a $\mathcal{B}_1$-space, equivalent to a matroid. In fact, the only reference in Nakasawa [1935] is to the book *Grundlagen der Elementargeometrie* (Foundations of Elementary Geometry) of Thomsen [1933], in which a different axiom system, the *Zyklenkalkül* (cycle calculus), was given (not equivalent to matroids). Nakasawa only gave subsets of linear spaces as an example. In a sequel to his paper, Nakasawa [1936b] observed that his axioms are equivalent to those of Whitney. The same axiom system as Nakasawa's, added with a continuity axiom, was given by Pauc [1937]. In Haupt,

Nöbeling, and Pauc [1940] the concept of an *Abhängigkeitsraum* (dependence space) based on these axioms was investigated.

The fourth 'line' was that of Whitney [1935], who introduced the notion of a matroid as a concept by itself. He was motivated by generalizing certain separability and duality phenomena in graphs, studied by him before. This led him to show that each matroid has a dual. While Whitney showed the equivalence of several axiom systems for matroids, he did not consider an axiom system based on a closure operation or on flats. Whitney gave linear dependence as an example, but not algebraic dependence. In a paper in the same year and journal, Birkhoff [1935a] showed the relation of Whitney's work with lattices.

We now discuss some historical papers more extensively, in a more or less chronological order.

1894-1900: Dedekind: lattices

In the supplements to the fourth edition of *Vorlesungen über Zahlentheorie* (Lectures on Number Theory) by Lejeune Dirichlet [1894], R. Dedekind introduced the notion of a *module* as any nonempty set of (real or complex) numbers closed under addition and subtraction, and he studied the lattice of all modules ordered by inclusion. He called A *divisible by* B if $A \subseteq B$. Trivially, the lattice operations are given by $A \wedge B = A \cap B$ and $A \vee B = A + B$. In fact, Dedekind denoted $A \cap B$ by $A - B$.

He gave the following 'charakteristischen Satz' (characteristic theorem):

> *Ist m theilbar durch d, und a ein beliebiger Modul, so ist*
>
> $$m + (a - d) = (m + a) - d. \quad [3]$$

In modern notation, for all a, b, c:

$$\text{if } a \leq c, \text{ then } a \vee (b \wedge c) = (a \vee b) \wedge c, \tag{39.64}$$

which is now known as the *modular law*, and lattices obeying it are called *modular lattices*.

Next, Dedekind [1897] introduced the notion of a lattice under the name *Dualgruppe* (dual group), motivated by similarities observed by him between operations on modules and those for logical statements as given in the book *Algebra der Logik* (Algebra of Logic) by Schröder [1890]. Dedekind mentioned, as examples, subsets of a set, modules, ideals in a finite field, subgroups of a group, and all fields, and he introduced the name *module law* for property (39.64):

> ich will es daher das Modulgesetz nennen, und jede Dualgruppe, in welcher es herrscht, mag eine Dualgruppe vom Modultypus heißen.[4]

[3] *If m is divisible by d, and a is an arbitrary module, then*

$$m + (a - d) = (m + a) - d.$$

[4] I will therefore call it the module law, and every dual group in which it holds, may be called a dual group of module type.

Dedekind [1900] continued the study of modular lattices, and showed that each modular lattice allows a rank function $r : M \to \mathbb{Z}_+$ with the property that for all a, b:

(39.65) (i) $r(0) = 0$;
(ii) $r(b) = r(a) + 1$ if b covers a;
(iii) $r(a \wedge b) + r(a \vee b) = r(a) + r(b)$.

In fact, this characterizes modular lattices.

In proving (39.65), Dedekind showed that each modular lattice satisfies

(39.66) if a and b cover c, and $a \neq b$, then $a \vee b$ covers a and b,

which is the property characterizing *upper semimodular lattices*, a structure equivalent to matroids.

1862-1913: Grassmann, Steinitz: linear and algebraic dependence

The basic exchange property of linear independence was formulated by Grassmann [1862], in his book *Die Ausdehnungslehre*, as follows (in his terminology, vectors are quantities):

> **20.** *Wenn m Grössen* $a_1, \dots a_m$*, die in keiner Zahlbeziehung zu einander stehen, aus n Grössen* $b_1, \dots b_n$ *numerisch ableitbar sind, so kann man stets zu den m Grössen* $a_1, \dots a_m$ *noch* $(n - m)$ *Grössen* $a_{m+1}, \dots a_n$ *von der Art hinzufügen, dass sich die Grössen* $b_1, \dots b_n$ *auch aus* $a_1, \dots a_n$ *numerisch ableiten lassen, und also das Gebiet der Grössen* $a_1, \dots a_n$ *identisch ist dem Gebiete der Grössen* $b_1, \dots b_n$*; auch kann man jene* $(n - m)$ *Grössen aus den Grössen* $b_1, \dots b_n$ *selbst entnehmen.*[5]

This property was also given by Steinitz [1913] (see below), but before that, Steinitz proved it for algebraic independence. In his fundamental paper *Algebraische Theorie der Körper* (Algebraic Theory of Fields), Steinitz [1910] studied, in § 22, algebraic dependence in field extensions. The statements proved are as follows, where L is a field extension of field K. Throughout, a is *algebraically dependent* on S if a is algebraic with respect to the field extension $K(S)$; in other words, if there is a nonzero polynomial $p(x) \in K(S)[x]$ with $p(a) = 0$.

Calling a set a *system*, he first observed:

> 1. *Hängt das Element a vom System S algebraisch ab, so gibt es ein endliches Teilsystem* S' *von S, von welchem a algebraisch abhängt.*[6]

and next he showed:

> 2. *Hängt* S_3 *von* S_2*,* S_2 *von* S_1 *algebraisch ab, so ist* S_3 *algebraisch abhängig von* S_1*.*[7]

---

[5] **20.** *If m quantities* $a_1, \dots a_m$*, that stand in no number relation to each other, are numerically derivable from n quantities* $b_1, \dots b_n$*, then one can always add to the m quantities* $a_1, \dots a_m$ *another* $(n - m)$ *quantities* $a_{m+1}, \dots a_n$ *such that the quantities* $b_1, \dots b_n$ *can also be derived numerically from* $a_1, \dots a_n$*, and that hence the domain of the quantities* $a_1, \dots a_n$ *is identical to the domain of the quantities* $b_1, \dots b_n$*; one also can take those* $(n - m)$ *quantities from the quantities* $b_1, \dots b_n$ *themselves.*

[6] 1. *If element a depends algebraically on the system S, then there is a finite subsystem* S' *of S on which a depends algebraically.*

[7] 2. *If* S_3 *depends algebraically on* S_2*, and* S_2 *on* S_1*, then* S_3 *is algebraically dependent on* S_1*.*

He called two sets S_1 and S_2 *equivalent* if S_1 depends algebraically on S_2, and conversely. A set is *reducible* if it has a proper subset equivalent to it. He showed:

> 3. *Jedes Teilsystem eines irreduziblen Systems ist irreduzibel.*
> 4. *Jedes reduzible System enthält ein endliches reduzibles Teilsystem.*[8]

and (after statement 5, saying that any two field extensions by equicardinal irreducible systems are isomorphic):

> 6. *Wird ein irreduzibles System S durch Hinzufügung eines Elementes a reduzibel, so ist a von S algebraisch abhängig.*[9]

From these properties, Steinitz derived:

> 7. *Ist S ein (in bezug auf K) irreduzibles System, das Element a in bezug auf K transzendent, aber von S algebraisch abhängig, so enthält S ein bestimmtes endliches Teilsystem T von folgender Beschaffenheit: a ist von T algebraisch abhängig; jedes Teilsystem von S, von welchem a algebraisch abhängt, enthält das System T; wird irgendein Element aus T durch a ersetzt, so geht S in ein äquivalentes irreduzibles System über; keinem der übrigen Elemente von S kommt diese Eigenschaft zu.*[10]

Steinitz proved this using only the properties given above (together with the fact that any $s \in S$ is algebraically dependent on S). Moreover, he derived from 7, (what is now called) *Steinitz' exchange property* for algebraic dependence:

> 8. *Es seien U und B endliche irreduzible Systeme von m bzw. n Elementen; es sei $n \leq m$ und B algebraisch abhängig von U. Dann sind im Falle $m = n$ die Systeme U und B äquivalent, im Falle $n < m$ aber ist U einem irreduziblen System äquivalent, welches aus B und $m - n$ Elementen aus U besteht.*[11]

This in particular implies that any two equivalent irreducible systems have the same size, and that the properties are equivalent to that determining a matroid.

In a subsequent paper, Steinitz [1913] proved a number of auxiliary statements on linear equations. Among other things, he showed (in his terminology, vectors are numbers, and a vector space is a module):

> *Besitzt der Modul* M *eine Basis von p Zahlen, und enthält er r linear unabhängige Zahlen $\beta_1, \ldots, \beta_r$, so besitzt er auch eine Basis von p Zahlen, unter denen die Zahlen $\beta_1, \ldots, \beta_r$ sämtlich vorkommen.*[12]

---

[8] 3. *Every subsystem of an irreducible system is irreducible.*
4. *Every reducible system contains a finite reducible subsystem.*

[9] 6. *If an irreducible system S becomes reducible by adding an element a, then a is algebraically dependent on S.*

[10] 7. *If S is an irreducible system (with respect to K), [and] the element a transcendent with respect to K, but algebraically dependent on S, then S contains a certain finite subsystem T with the following quality: a is algebraically dependent on T; every subsystem of S on which a depends algebraically, contains the system T; if any element from T is replaced by a, then S passes into an equivalent irreducible system; this property belongs to none of the other elements of S .*

[11] 8. *Let U and B be finite irreducible systems of m and n elements respectively; let $n \leq m$ and let B be algebraically dependent on U. Then, in case $m = n$, the systems U and B are equivalent, but in case $n < m$, U is equivalent to an irreducible system which consists of B and $m - n$ elements from U.*

[12] *If a module* M *possesses a base of p numbers, and it contains r linearly independent numbers $\beta_1, \ldots, \beta_r$, then it possesses also a base of p numbers, among which the numbers $\beta_1, \ldots, \beta_r$ all occur.*

Steinitz' proof of this in fact gives a stronger result, known as *Steinitz' exchange property*: the new base is obtained by extending $\beta_1, \ldots, \beta_r$ with vectors from the given base. So Steinitz came to the same result as Grassmann [1862] quoted above. In his paper, Steinitz [1913] did not make a link with similar earlier results in Steinitz [1910] on algebraic dependence.

1929: Bergmann

Inspired by Menger [1928a], who aimed at giving an axiomatic foundation for projective geometry on a lattice-theoretical basis, Bergmann [1929] gave an axiomatic foundation of affine geometry, again on the basis of lattices. Bergmann's article contains a number of proofs that in fact concern matroids, while he assumed, but not used, a complementation axiom (since he aimed at characterizing full affine spaces, not subsets of it): for each pair of elements $A \leq B$ there exist C_1 and C_2 with $A \vee C_1 = B$, $A \wedge C_1 = 0$, $B \wedge C_2 = A$, and $B \vee C_2 = 1$. This obviously implies (in the finite case) that

(39.67) each element of the lattice is a join of points.

(A *point* is a minimal nonzero element.) It is property (39.67) that Bergmann uses in a number of subsequent arguments (and not the complementation axiom). His further axiom is:

(39.68) for any element A and any point P of the lattice, there is no element B with $A < B < A \vee P$.

He called an ordered sequence $(P_1, \ldots, P_n)$ of points a *chain* (Kette) (*of* an element A), if $P_i \not\leq P_1 \vee \cdots \vee P_{i-1}$ for $i = 1, \ldots, n$ (and $A = P_1 \vee \cdots \vee P_n$). He derived from (39.67) and (39.68) that being a chain is independent of the order of the elements in the chain, and that any two chains of an element A have the same length:

Satz: Alle Ketten eines Elementes A haben dieselbe Gliederzahl.[13]

He remarked that under condition (39.67), this in turn implies (39.68).

Denoting the length of any chain of A by $|A|$, Bergmann showed that it is equal to the rank of A in the lattice, and he derived the submodular inequality:

$$|A| + |B| \geq |A + B| + |A \cdot B|.$$

(Bergmann denoted $\vee$ and $\wedge$ by $+$ and $\cdot$.) Thus he proved the submodularity of the rank function of a matroid. These results were also given by Alt [1936] in Menger's *mathematischen Kolloquium* in Vienna on 1 March 1935 (cf. Menger [1936a,1936b]).

1929-1937: Haupt, van der Waerden

Inspired by the work of Steinitz, in the books *Einführung in die Algebra* (Introduction to Algebra) by Haupt [1929a,1929b] and *Moderne Algebra* (Modern Algebra) by van der Waerden [1930], the analogies between proof methods for linear and algebraic dependence were observed.

Haupt mentioned in his preface (after saying that his book will contain the modern developments of algebra):

[13] *Theorem: All chains of an element A have the same number of members.*

> Demgemäß ist das vorliegende Buch durchweg beeinflußt von der bahnbrechenden „Algebraischen Theorie der Körper“ von Herrn E. Steinitz, was hier ein für allemal hervorgehoben sei. Ferner stützt sich die Behandlung der linearen Gleichungen (vgl. *9,1* bis *9,4*), einer Anregung von Frl. E. Noether folgend, auf die von Herrn E. Steinitz gegebene Darstellung (vgl. das Zitat in *9,0*).[14]

(The quotation in Haupt's '*9,0*' is to Steinitz [1910,1913].)

A number of theorems on algebraic dependence were proved in Chapter 23 of Haupt [1929b] by referring to the proofs of the corresponding results on linear dependence in Chapter 9 of Haupt [1929a]. In the introduction of his Chapter 9, Haupt wrote:

> Die Behandlung der linearen Gleichungen ist (soweit es geht) so angelegt, daß sich ein Teil der dabei gewonnenen Sätze auf *Systeme von algebraisch abhängigen Elementen überträgt*, was später (*23,6*) dargelegt wird.[15]

In the first edition of his book, van der Waerden [1930] listed the properties of algebraic dependence:

> Die Relation der algebraischen Abhängigkeit hat demnach die folgenden Eigenschaften:
> 1. a ist abhängig von sich selbst, d.h. von der Menge $\{a\}$.
> 2. Ist a abhängig von M, so hängt es auch von jeder Obermenge von M ab.
> 3. Ist a abhängig von M, so ist a schon von einer endlichen Untermenge $\{m_1, \ldots, m_n\}$ von M (die auch leer sein kann) abhängig.
> 4. Wählt man diese Untermenge minimal, so ist jedes m_i von a und den übrigen m_j abhängig.
> Weiter gilt:
> 5. Ist a abhängig von M und jedes Element von M abhängig von N, so ist a abhängig von N.[16]

Following Steinitz, van der Waerden called two sets *equivalent* if each element of the one set depends algebraically on the other set, and vice versa, while a set is *irreducible* if no element of it depends algebraically on the remaining.

Using only the properties 1-5, van der Waerden derived that each set contains an irreducible set equivalent to it, and that if $M \subseteq N$, then each irreducible subset of M equivalent to M can be extended to an irreducible subset of N equivalent to N — in other words, inclusionwise minimal subsets of M equivalent to M are

---

[14] Accordingly, the present book is invariably influenced by the pioneering 'Algebraic Theory of Fields' by Mr E. Steinitz, which be emphasized here once and for all. Further, following a suggestion by Miss E. Noether, the treatment of linear equations (cf. *9,1* to *9,4*) leans on the presentation by Mr E. Steinitz (cf. the quotation in *9,0*).

[15] The treatment of linear equations is (as far as it goes) made such that a part of the theorems obtained therewith *transfers to systems of algebraically dependent elements*, which will be discussed later (*23,6*).

[16] The relation of algebraic dependence has therefore the following properties:
1. a is dependent on itself, that is, on the set $\{a\}$.
2. If a is dependent on M, then it also depends on every superset of M.
3. If a is dependent on M, then a is dependent already on a finite subset $\{m_1, \ldots, m_n\}$ of M (that can also be empty).
4. If one chooses this subset minimal, then every m_i is dependent on a and the remaining m_j.
Further it holds:
5. If a is dependent on M and every element of M is dependent on N, then a is dependent on N.

independent, and inclusionwise maximal independent subsets of M are equivalent to M.

Van der Waerden [1930] also showed that two equivalent irreducible systems have the same size, but in the proof he uses polynomials. This is not necessary, since the properties 1-5 determine a matroid.

Van der Waerden noticed the analogy with linear dependence, treated in his § 28, where he uses specific facts on linear equations:

> Tatsächlich gelten für die dort betrachtete lineare Abhängigkeit dieselben Regeln 1 bis 5, die für die algebraische Abhängigkeit in § 61 aufgestellt wurden; man kann also alle Beweise wörtlich übertragen.[17]

In the second edition of his book, van der Waerden [1937] gave a unified treatment of linear and algebraic dependence, slightly different from the first edition. As for linear dependence he stated in § 33:

> Drei Grundsätze genügen. Der erste ist ganz selbstverständlich.
> Grundsatz 1. *Jedes* u_i $(i = 1, \ldots, n)$ *ist von* $u_1, \ldots, u_n$ *linear abhängig.*
> Grundsatz 2. *Ist* v *linear abhängig von* $u_1, \ldots, u_n$*, aber nicht von* $u_1, \ldots, u_{n-1}$*, so ist* u_n *linear abhängig von* $u_1, \ldots, u_{n-1}, v$*.*
> [· · ·]
> Grundsatz 3. *Ist* w *linear abhängig von* $v_1, \ldots, v_s$ *und ist jedes* v_j $(j = 1, \ldots, s)$ *linear abhängig von* $u_1, \ldots, u_n$*, so ist* w *linear abhängig von* $u_1, \ldots, u_n$*.*[18]

The same axioms are given in § 64 of van der Waerden [1937], with 'linear' replaced by 'algebraisch'.

Next, van der Waerden called elements $u_1, \ldots, u_n$ (linearly or algebraically) independent if none of them depend on the rest of them. Among the consequences of these principles, he mentioned that if $u_1, \ldots, u_{n-1}$ are independent but $u_1, \ldots, u_{n-1}, u_n$ are not, then u_n is dependent on $u_1, \ldots, u_{n-1}$, and that each finite system of elements $u_1, \ldots, u_n$ contains a (possibly empty) independent subsystem on which each u_i is dependent. He called two systems $u_1, \ldots, u_n$ and $v_1, \ldots, v_s$ *equivalent* if each v_k depends on $u_1, \ldots, u_n$ and each u_i depends on $v_1, \ldots, v_s$, and he now derived from the three principles that two equivalent independent systems have the same size.

Mac Lane [1938] observed that the axioms introduced by Whitney [1935] and those by van der Waerden [1937] determine equivalent structures.

1934: Hilbert, Bernays: collinearity axioms

Axiom systems for points and lines in a plane were given by Hilbert [1899] in his book *Grundlagen der Geometrie* (Foundations of Geometry), and by Veblen [1904].

---

[17] In fact, the same rules 1 to 5, that were formulated for algebraic dependence in § 61, hold for the linear dependence considered there; one can transfer therefore all proofs word for word.

[18] Three principles suffice. The first one is fully self-evident.
Principle 1. *Every* u_i $(i = 1, \ldots, n)$ *is linearly dependent on* $u_1, \ldots, u_n$*.*
Principle 2. *If* v *is linearly dependent on* $u_1, \ldots, u_n$*, but not on* $u_1, \ldots, u_{n-1}$*, then* u_n *is linearly dependent on* $u_1, \ldots, u_{n-1}, v$*.*
[· · ·]
Principle 3. *If* w *is linearly dependent on* $v_1, \ldots, v_s$ *and every* v_j $(j = 1, \ldots, s)$ *is linearly dependent on* $u_1, \ldots, u_n$*, then* w *is linearly dependent on* $u_1, \ldots, u_n$*.*

Basis is the axiom that any two distinct points are in exactly one line. Note that this axiom determines precisely all matroids of rank at most 3 with no parallel elements (by taking the lines as maximal flats).

One of the axioms of Veblen is:

> Axiom VI. If points C and D ($C \neq D$) lie on the line AB, then A lies on the line CD.

This axiom corresponds to axiom 3) in the book *Grundlagen der Mathematik* (Foundations of Mathematics) of Hilbert and Bernays [1934], who aim to make an axiom system based on points only:

> Dabei empfiehlt es sich für unseren Zweck, von dem HILBERTschen Axiomensystem darin abzuweichen, daß wir nicht die Punkte und die Geraden als zwei Systeme von Dingen zugrunde legen, sondern *nur die Punkte als Individuen nehmen.*[19]

The axiom system of Hilbert and Bernays is in terms of a relation Gr to describe collinearity of triples of points (where (x) stands for $\forall x$, (Ex) for $\exists x$, and $\overline{P}$ for the negation of P):

> I. Axiome der Verknüpfung.
> 1) $(x)(y)Gr(x,x,y)$
> „x, x, y liegen stets auf einer Geraden."
> 2) $(x)(y)(z)(Gr(x,y,z) \rightarrow Gr(y,x,z)\&Gr(x,z,y))$.
> „Wenn x, y, z auf einer Geraden liegen, so liegen stets auch y, x, z sowie auch x, z, y auf einer Geraden."
> 3) $(x)(y)(z)(u)(Gr(x,y,z)\&Gr(x,y,u)\&x \neq y \rightarrow Gr(x,z,u))$.
> „Wenn x, y, verschiedene Punkte sind und wenn x, y, z sowie x, y, u auf einer Geraden liegen, so liegen stets auch x, z, u auf einer Geraden."
> 4) $(Ex)(Ey)(Ez)\overline{Gr(x,y,z)}$.
> „Es gibt Punkte x, y, z, die nicht auf einer Geraden liegen."[20]

The axioms 1) and 2) in fact tell that the relation Gr is determined by unordered triples of distinct points. The exchange axiom 3) is a special case of the matroid axiom for circuits in a matroid.

Hilbert and Bernays extended the system by axioms for a betweenness relation Zw for ordered triples of points, and a parallelism relation Par for ordered quadruples of points.

---

[19] At that it is advisable for our purpose to deviate from HILBERT's axiom system in that we do not lay the points and the lines as two systems of things as base, but *take only the points as individuals.*

[20] I. Axioms of connection.
1) $(x)(y)Gr(x,x,y)$
'x, x, y always lie on a line.'
2) $(x)(y)(z)(Gr(x,y,z) \rightarrow Gr(y,x,z)\&Gr(x,z,y))$.
'If x, y, z lie on a line, then also y, x, z as well as x, z, y always lie on a line.'
3) $(x)(y)(z)(u)(Gr(x,y,z)\&Gr(x,y,u)\&x \neq y \rightarrow Gr(x,z,u))$.
'If x, y, are different points and if x, y, z as well as x, y, u lie on a line, then also x, z, u always lie on a line.'
4) $(Ex)(Ey)(Ez)\overline{Gr(x,y,z)}$.
'There are points x, y, z, that do not lie on a line.'

1933–1935: Birkhoff: Lattices

In his paper 'On the combination of subalgebras', Birkhoff [1933] ('Received 15 May 1933') wrote:

> The purpose of this paper is to provide a point of vantage from which to attack combinatorial problems in what may be termed modern, synthetic, or abstract algebra. In this spirit, a research has been made into the consequences and applications of seven or eight axioms, only one [V] of which itself is new.

The axioms are those for a lattice, added with axiom V, that amounts to (39.64) above. Any lattice satisfying this condition is called by Birkhoff in this paper a 'B-lattice'. In an addendum, Birkhoff [1934b] mentioned that O. Ore had informed him that part of his results had been obtained before by Dedekind [1900]. Therefore, Birkhoff [1935b] renamed it to *modular lattice*.

Birkhoff [1933] mentioned, as examples, the classes of normal subgroups and of characteristic subgroups of a group. Other examples mentioned are the ideals of a ring, and the linear subspaces of Euclidean space. (Both examples actually give sublattices of the lattice of all normal subgroups of the corresponding groups.)

Like Dedekind, Birkhoff [1933] showed that (39.64) implies (39.66). Lattices satisfying (39.66) are called (upper) *semimodular*. Birkhoff showed that any upper semimodular lattice has a rank function satisfying (39.65)(i) and (ii) and satisfying the submodular law:

(39.69) $$r(a \cap b) + r(a \cup b) \leq r(a) + r(b).$$

This characterizes upper semimodular lattices.

Birkhoff noticed that this implies that the modular lattices are exactly those lattices satisfying both (39.66) and its symmetric form:

(39.70) if c covers a and b and $a \neq b$, then a and b cover $a \wedge b$.

Birkhoff [1935c] showed that the partition lattice is upper semimodular, that is, satisfies (39.66), and hence has a rank function satisfying the submodular inequality[21]. Thus the complete graph, and hence any graph, gives a geometric lattice (and hence a matroid — however, Whitney's work seems not to have been known yet to Birkhoff at the time of writing this paper).

In a number of other papers, Birkhoff [1934a,1934c,1935b] made a further study of modular lattices, and gave relations to projective geometries (in which the collection of all flats gives a modular lattice). Klein-Barmen [1937] further investigated semimodular lattices (called by him *Birkhoffsche Verbände* (Birkhoff lattices)), of which he found several lattice-theoretical characterizations.

1935: Whitney: Matroids

Whitney [1935] (presented to the American Mathematical Society, September 1934) introduces the notion of matroid as follows:

[21] In fact, Birkhoff [1935c] claimed the modular equality for the rank function of a partition lattice (page 448), but this must be a typo, witness the formulation of, and the reference in, the first footnote on that page.

> Let $C_1, C_2, \cdots, C_n$ be the columns of a matrix M. Any subset of these columns is either linearly independent or linearly dependent; the subsets thus fall into two classes. These classes are not arbitrary; for instance, the two following theorems must hold:
>
> (a) Any subset of an independent set is independent.
>
> (b) If N_p and N_{p+1} are independent sets of p and $p+1$ columns respectively, then N_p together with some column of N_{p+1} forms an independent set of $p+1$ columns.
>
> There are other theorems not deducible from this; for in § 16 we give an example of a system satisfying these two theorems but not representing any matrix. Further theorems seem, however, to be quite difficult to find. Let us call a system obeying (a) and (b) a "matroid." The present paper is devoted to a study of the elementary properties of matroids. The fundamental question of completely characterizing systems which represent matrices is left unsolved. In place of the columns of a matrix we may equally well consider points or vectors in a Euclidean space, or polynomials, etc.

In the paper, Whitney observed that forests in a graph form the independent sets of a matroid, for which reason he carried over various terms from graphs to matroids.

Whitney described several equivalent axiom systems for the notion of matroid. First, he showed that the rank function is characterized by (39.42), and he derived that it is submodular. Next, he showed that the collection of bases is characterized by (39.33)(ii), and the collection of circuits by (39.34)(iii). Moreover, he showed that complementing all bases gives again a matroid, the dual matroid, and that the dual of a linear matroid is again a linear matroid. In the paper, he also studied separability and representability of matroids. The example given in Whitney's § 16 (mentioned in the above quotation), is in fact the well-known Fano matroid — he apparently did not consider matrices over GF(2). However, in an appendix of the paper, he characterized the matroids representable by a matrix 'of integers mod 2': a matroid is representable over GF(2) if and only if any sum (mod 2) of circuits can be partitioned into circuits.

In a subsequent paper 'Abstract linear independence and lattices', Birkhoff [1935a] pointed out the relations of Whitney's work with Birkhoff's earlier work on semimodular lattices. He stated:

> In a preceding paper, Hassler Whitney has shown that it is difficult to distinguish theoretically between the properties of linear dependence of ordinary vectors, and those of elements of a considerably wider class of systems, which he has called "matroids."
>
> Now it is obviously impossible to incorporate all of the heterogeneous abstract systems which are constantly being invented, into a body of systematic theory, until they have been classified into two or three main species. The purpose of this note is to correlate matroids with abstract systems of a very common type, which I have called "lattices."

Birkhoff showed that a lattice is isomorphic to the lattice of flats of a matroid if and only if the lattice is semimodular, that is, satisfies (39.66), and each element is a join of atoms.

In the paper 'Some interpretations of abstract linear dependence in terms of projective geometry', MacLane [1936] gave a geometric interpretation of matroids. He introduced the notion of a 'schematic n-dimensional figure', consisting of 'k-dimensional planes' for $k = 1, 2, \ldots$. Each such plane is a subset of an (abstract) set of 'points', with the following axioms (for any appropriate k):

(39.71) (i) any k points belonging to no $k-1$-dimensional plane, belong to a unique k-dimensional plane; moreover, this plane is contained in any plane containing these k points;
(ii) every k-dimensional plane contains k points that belong to no $k-1$-dimensional plane.

MacLane mentioned that there is a 1-1 correspondence between schematic figures and the collections of flats of matroids. As a consequence he mentioned that a schematic n-dimensional figure is completely determined by its collection of $n-1$-dimensional planes (as a matroid is determined by its hyperplanes = complements of cocircuits).

1935: Nakasawa: Abhängigkeitsräume

In the paper *Zur Axiomatik der linearen Abhängigkeit. I* (On the axiomatics of linear dependence. I) in *Science Reports of the Tokyo Bunrika Daigaku* (Tokyo University of Literature and Science), Nakasawa [1935] introduced an axiom system for dependence, that he proved to be equivalent to matroids (in a different terminology).

He was motivated by an axiom system described by Thomsen [1933] in his book *Grundlagen der Elementargeometrie* (Foundations of Elementary Geometry). Thomsen's 'cycle calculus' is an attempt to axiomatize relations (like coincidence, orthogonality, parallelism) between geometric objects (points, lines, etc.). Thomsen emphasized that existence questions often are inessential in elementary geometry:

> In der Tat erscheinen uns ja auch die Existenzaussagen als ein verhältnismäßig unwesentliches Beiwerk der Elementargeometrie. Ohne Zweifel empfinden wir als die eigentlich inhaltsvollsten und die wichtigsten Einzelaussagen der Elementargeometrie die von der folgenden reinen Form: ,,Wenn eine Reihe von geometrischen Gebilden, d.h. eine Anzahl von Punkten, Geraden, usw., gegeben vorliegt, und zwar derart, daß zwischen den gegebenen Punkten, Geraden usw. die und die geometrischen Lagebeziehungen bestehen (Koinzidenz, Senkrechtstehen, Parallellaufen, ,,Mittelpunkt sein" und anderes mehr), dann ist eine notwendige Folge dieser Annahme, daß auch noch diese bestimmte weitere geometrische Lagebeziehung gleichzeitig besteht." In Sätzen dieser Form kommt nichts von Existenzaussagen vor. Was das Wichtigste ist, nicht in den Folgerungen. Dann aber auch nicht in den Annahmen. Wir nehmen an: *Wenn* die und die Dinge in den und den Beziehungen gegeben vorliegen..., usw. Wir machen aber keinerlei Voraussetzungen darüber, ob eine solche Konfiguration in unserer Geometrie existieren kann. Der Schluß ist nur: Wenn sie existieren, dann .... Falls die Konfiguration gar nicht existiert, der Satz also gegenstandslos wird, betrachten wir ihn nach der üblichen Konvention ,,gegenstandslos, also richtig" als richtig.[22]

[22] Indeed, also the existence statements seem to us a relatively inessential side issue of elementary geometry. Undoubtedly, we find as the really most substantial and most important special statements of elementary geometry those of the following pure form: 'If a sequence of geometric creations, that is, a number of points, lines etc., are given to us, and that in such a way, that those and those geometric position relations exist between the given points, lines etc. (coincidence, orthogonality, parallelism, "being a centre", and other), then a necessary consequence of this assumption is that also this certain further geometric position relation exists at the same time.' In theorems of this form, no existence statements occur. What is most important: not in the consequences. But then neither in the assumptions. We assume: *If* those and those things are given

Thomsen aimed at founding axiomatically 'the partial geometry of all elementary geometric theorems without existence statements'. To that end, he introduced the concept of a *cycle*, which is an ordered finite sequence of abstract objects, which can be thought of as points, lines, etc. Certain cycles are 'correct' and the other 'incorrect' (essentially they represent a system of relations defining any binary group):

A) *Axiom der Grundzyklen:* Der Zyklus $\alpha\alpha$ ist für jedes α richtig, der Zyklus α für kein α.
B) *Axiom des Löschens:* $\beta_1\beta_2\dots\beta_n\alpha\alpha \to \beta_1\beta_2\dots\beta_n$; in Worten: Aus der Richtigkeit des Zyklus $\beta_1\beta_2\dots\beta_n\alpha\alpha$ folgt auch die des Zyklus $\beta_1\beta_2\dots\beta_n$.
C) *Axiom des Umstellens:* $\beta_1\beta_2\dots\beta_n \to \beta_2\beta_3\dots\beta_n\beta_1$.
D) *Axiom des Umkehrens:* $\beta_1\beta_2\dots\beta_{n-1}\beta_n \to \beta_n\beta_{n-1}\dots\beta_2\beta_1$.
E) *Axiom des Anfügens:* $\beta_1\beta_2\dots\beta_n$ und $\gamma_1\gamma_2\dots\gamma_r \to \beta_1\beta_2\dots\beta_n\gamma_1\gamma_2\dots\gamma_r$.[23]

Axiom B) can be considered as a variant of Steinitz' exchange property. With the other axioms it implies that if $\beta_1\cdots\beta_n\alpha$ and $\gamma_1\cdots\gamma_r\alpha$ are cycles, then $\beta_1\cdots\beta_n\gamma_1\cdots\gamma_r$ is a cycle. Therefore, the set of all inclusionwise minimal nonempty sets containing a cycle form the circuits of a matroid.

The purpose of Nakasawa [1935] is to generalize Thomsen's axiom system:

> In der vorliegenden Untersuchung soll ein Axiomensystem für eine neue Formulierung der linearen Abhängigkeit des n-dimensionalen projektiven Raumes angegeben werden, indem wir hauptsächlich den *Zyklenkalkül*, den Herr G. Thomsen bei seiner Grundlegung der elementaren Geometrie hergestellt hat, hier in einem noch abstrakteren Sinne verwenden.[24]

While Thomsen's cycles relate to unions of circuits in a matroid, those of Nakasawa form the dependent sets of a matroid. His axiom system can be considered as a direct extension to higher dimensions of the collinearity axioms of Hilbert and Bernays given above.

He called the structure *der erste Verknüpfungsraum* (the first connection space), or a $\mathcal{B}_1$*-Raum* ($\mathcal{B}_1$-space), writing $a_1\cdots a_s$ for $a_1\cdots a_s = 0$:

> **Grundannahme:** Wir denken uns eine gewisse Menge der Elementen; $\mathcal{B}_1 \ni a_1, a_2, \cdots, a_s, \cdots$. Für gewisse Reihen der Elementen, die wir *Zyklen* nennen wollen, denken wir dazu die Relationen "*gelten*" oder "*gültig sein*", in Zeichen $a_1\cdots a_s = 0$, bzw. "*nicht gelten*" oder "*nicht gültig sein*", in Zeichen $a_1\cdots a_s \neq 0$. Diese Relationen sollen nun folgenden Axiomen genügen;

---

to us in those and those relations..., etc. We do not make any assumption on the fact if such a configuration can exist in our geometry. The conclusion is only: If they exist, then .... In case the configuration does not exist at all, and the theorem thus becomes meaningless, we consider it by the usual convention 'meaningless, hence correct' as correct.

[23]
A) *Axiom of ground cycles:* The cycle $\alpha\alpha$ is correct for each α, the cycle α for no α.
B) *Axiom of solving:* $\beta_1\beta_2\dots\beta_n\alpha\alpha \to \beta_1\beta_2\dots\beta_n$; in words: From the correctness of the cycle $\beta_1\beta_2\dots\beta_n\alpha\alpha$ follows that of the cycle $\beta_1\beta_2\dots\beta_n$.
C) *Axiom of transposition:* $\beta_1\beta_2\dots\beta_n \to \beta_2\beta_3\dots\beta_n\beta_1$.
D) *Axiom of inversion:* $\beta_1\beta_2\dots\beta_{n-1}\beta_n \to \beta_n\beta_{n-1}\dots\beta_2\beta_1$.
E) *Axiom of addition:* $\beta_1\beta_2\dots\beta_n$ and $\gamma_1\gamma_2\dots\gamma_r \to \beta_1\beta_2\dots\beta_n\gamma_1\gamma_2\dots\gamma_r$.

[24] In the present research, an axiom system for a new formulation of linear dependence of the n-dimensional projective space should be indicated, while we use here mainly the *cycle calculus*, which Mr G. Thomsen has constructed in his foundation of elementary geometry, in a still more abstract sense.

Axiom 1. (Reflexivität) : aa.
Axiom 2. (Folgerung) : $a_1 \cdots a_s \to a_1 \cdots a_s x$, $(s = 1, 2, \cdots)$.
Axiom 3. (Vertauschung) : $a_1 \cdots a_i \cdots a_s \to a_i \cdots a_1 \cdots a_s$, $(s = 2, 3, \cdots; i = 2, \cdots, s)$.
Axiom 4. (Transitivität) : $a_1 \cdots a_s \neq 0$, $xa_1 \cdots a_s$, $a_1 \cdots a_s y \to xa_1 \cdots a_{s-1} y$, $(s = 1, 2, \cdots)$.

Definition I. Eine solche Menge $\mathcal{B}_1$ heisst der erste Verknüpfungsraum, in kurzen Worten, $\mathcal{B}_1$-**Raum**.[25]

Axiom 3 corresponds to condition (39.3).

Nakasawa introduced the concept of span, and he derived that any two independent sets having the same span, have the same size. It implies that $\mathcal{B}_1$-spaces are the same structures as matroids. Moreover, he gave a submodular law for a rank concept.

In a second paper, Nakasawa [1936a] added a further axiom on intersections of subspaces, yielding a '$\mathcal{B}_2$-space', which corresponds to a projective space (in which the rank is modular), and in a third paper, Nakasawa [1936b] observed that his $\mathcal{B}_1$-spaces form the same structure as the matroids of Whitney.

1937-1940: Pauc, Haupt, Nöbeling

The axioms presented by Nakasawa were also given by Pauc [1937], added with an axiom describing the limit behaviour of dependence, if the underlying set is endowed with a topology:

Introduction axiomatique d'une notion de dépendance sur une classe limite. — Soit D un prédicat relatif aux systèmes finis non ordonnés de points d'une classe limite $\mathcal{L}$, assujetti aux axiomes (notation d'Hilbert-Bernays)

(A_1) $(x_1)(x_2)(D[x_1, x_2] \sim (x_1 = x_2))$,

(A_2) $(x_1)(x_2)\dots(x_p)(y)(D[x_1, x_2, \dots, x_p] \to D[x_1, x_2, \dots, x_p, y])$,

(A_3) $(x_1)(x_2)\dots(x_p)(y)(z)(\overline{D[x_1, \dots, x_p]} \,\&\, D[x_1, \dots, x_p, y] \,\&\, D[x_1, \dots, x_p, z] \to D[x_2, \dots, x_p, y, z])$,

(A_4) Quels que soient les points $x_1, x_2, \dots, x_p$ et la suite $y_1, y_2, \dots, y_q, \dots$ de $\mathcal{L}$
$(\lim_{q\to\infty} y_q = y) \& (q) D[x_1, x_2, \dots, x_p, y_q] \to D[x_1, x_2, \dots, x_p, y]$.[26]

In a subsequent paper, Haupt, Nöbeling, and Pauc [1940] studied systems, called *A-Mannigfaltigkeit*, (*A*-manifolds) that satisfy the axioms A_1-A_3. They mentioned

[25] **Basic assumption:** We imagine ourselves a certain set of elements; $\mathcal{B}_1 \ni a_1, a_2, \cdots, a_s, \cdots$. For certain sequences of the elements, which we want to call *cycles*, we think the relations on them '*to hold*' or '*to be valid*', in notation $a_1 \cdots a_s = 0$, and '*not to hold*' or '*not to be valid*', in notation $a_1 \cdots a_s \neq 0$, respectively. These relations now should satisfy the following axioms;

Axiom 1. (reflexivity) : aa.
Axiom 2. (deduction) : $a_1 \cdots a_s \to a_1 \cdots a_s x$, $(s = 1, 2, \cdots)$.
Axiom 3. (exchange) : $a_1 \cdots a_i \cdots a_s \to a_i \cdots a_1 \cdots a_s$, $(s = 2, 3, \cdots; i = 2, \cdots, s)$.
Axiom 4. (transitivity) : $a_1 \cdots a_s \neq 0$, $xa_1 \cdots a_s$, $a_1 \cdots a_s y \to xa_1 \cdots a_{s-1} y$, $(s = 1, 2, \cdots)$.

Definition I. Such a set $\mathcal{B}_1$ is called the first connection space, in short, $\mathcal{B}_1$-**space**.

[26] Axiomatic introduction of a notion of dependence on a limit class. — Let D be a predicate relative to the finite unordered systems of points from a limit class $\mathcal{L}$, subject to the axioms (notation of Hilbert-Bernays)

that this axiom system was indeed inspired by those for collinearity of Hilbert-Bernays quoted above. They commented that its relation with Birkhoff's lattices, is analogous to the relation of the Hilbert-Bernays collinearity axioms with those of Hilbert for points and lines.

Haupt, Nöbeling, and Pauc [1940] gave, as examples, linear and algebraic dependence, and derived several basic facts (all bases have the same size, each independent set is contained in a base, for each pair of bases B, B' and $x \in B \setminus B'$ there is a $y \in B' \setminus B$ such that $B - x + y$ is a base, and the rank is submodular).

The authors mentioned that they were informed by G. Köthe about the relations of their work with the lattice formulation of algebraic dependence of Mac Lane [1938], but no connection is made with Whitney's matroid.

Among the further papers related to matroids are Menger [1936b], giving axioms for (full) affine spaces, and Wilcox [1939,1941,1942,1944] and Dilworth [1941a, 1941b,1944] on matroid lattices. The notion of M-symmetric lattice introduced by Wilcox [1942] was shown in Wilcox [1944] to be equivalent to upper semimodular lattice.

Rado

Rado was one of the first to take the independence structure as a source for further theorems, and to connect it with matching type theorems and combinatorial optimization. He had been interested in Kőnig-Hall type theorems (Rado [1933,1938]), and in his paper Rado [1942], he extended Hall's marriage theorem to transversals that are independent in a given matroid — a precursor of matroid intersection. In fact, with an elementary construction, Rado's theorem implies the matroid union theorem, and hence also the matroid intersection theorem (to be discussed in Chapters 41 and 42).

Rado [1942] did not refer to any earlier literature when introducing the concept of an *independence relation*, but the axioms are similar to those of Whitney for the independent sets in a matroid. Rado mentioned only linear independence as a special case.

He proved that a family of subsets of a matroid has an independent transversal if and only if the union of any k of the subsets contains an independent set of size k, for all k. Rado also showed that this theorem characterizes matroids.

Rado [1949a] extended the concept of matroid to infinite matroids, where he says that he extends the axioms of Whitney [1935].

Rado [1957] showed that if the elements of a matroid are linearly ordered by $\leq$, there is a unique minimal base $\{b_1, \ldots, b_r\}$ with $b_1 < b_2 < \cdots < b_r$ such that for each $i = 1, \ldots, r$ all elements $s < b_i$ belong to $\mathrm{span}(\{b_1, \ldots, b_{i-1}\})$. Rado derived that for any independent set $\{a_1, \ldots, a_k\}$ with $a_1 < \cdots < a_k$ one has $b_i \leq a_i$ for $i = 1, \ldots, k$. Therefore, the greedy method gives an optimum solution when

$$
\begin{array}{ll}
(\mathrm{A}_1) & (x_1)(x_2)(D[x_1, x_2] \sim (x_1 = x_2)), \\
(\mathrm{A}_2) & (x_1)(x_2)\ldots(x_p)(y)(D[x_1, x_2, \ldots, x_p] \to D[x_1, x_2, \ldots, x_p, y]), \\
(\mathrm{A}_3) & (x_1)(x_2)\ldots(x_p)(y)(z)(\overline{D[x_1, \ldots, x_p]}\,\&\, D[x_1, \ldots, x_p, y]\,\& \\
 & \quad D[x_1, \ldots, x_p, z] \to D[x_2, \ldots, x_p, y, z]), \\
(\mathrm{A}_4) & \left\{ \begin{array}{l} \text{Whatever are the points } x_1, x_2, \ldots, x_p \text{ and the sequence } y_1, y_2, \ldots, y_q, \\ \ldots \text{ from } \mathcal{L} \\ (\lim\limits_{q \to \infty} y_q = y)\&(q)D[x_1, x_2, \ldots, x_p, y_q] \to D[x_1, x_2, \ldots, x_p, y]. \end{array} \right.
\end{array}
$$

applied to find a minimum-weight base. Rado mentioned that it extends the work of Borůvka and Kruskal on finding a shortest spanning tree in a graph.

For notes on the history of matroid union, see Section 42.6f. For an excellent survey of early literature on matroids, with reprints of basic articles, see Kung [1986].