2.5 STEINITZ'S THEOREM – THE COMBINATORIAL PART

As we have seen in previous sections, the question whether a given combinatorial (or topological) 3-configuration can be geometrically realized is very hard. This is the reason the 1894 PhD thesis of Ernst Steinitz [S17] is remarkable in its generality: Although Steinitz fails in completely characterizing the realizability of combinatorial or topological 3-configurations, he come as close to doing so as anybody since then.

Steinitz's claim (in our terminology). Every connected combinatorial 3-configuration (n_3) has a geometric realization by points and lines as a #1-subfiguration in the Euclidean plane; moreover, the point and line of the ignored incidence can be arbitrarily chosen.

Recall from Section 1.3 that a *#1-subfiguration* of a combinatorial configuration is a family of points and (straight) lines that satisfies all the incidence requirements except possibly one that is ignored, and has no additional incidences.

In the next section we shall see that this claim is not correct. However, even the weaker result that Steinitz' arguments actually establish (see Theorem 2.6.1) is remarkable in several ways. Steinitz' proof has two parts, a combinatorial and a geometric. The combinatorial part is correct, and was much ahead of its time. However, the geometric part is defective; we discuss this in the next section. We start with the combinatorial part of Steinitz' theorem, and first recall from Section 1.3 a useful definition.

A configuration table for a combinatorial configuration is said to be **orderly** if every row of the table contains all the points (hence contains each precisely once). For example, the configuration table in Table 2.1.1 is orderly, and the configuration tables in Sections 1.3 and 2.2 are not orderly.

The following is a basic result, due to Steinitz [S17].

Theorem 2.5.1. Every combinatorial k-configuration admits an orderly configuration table.

A statement that Theorem 2.5.1 holds for $k = 3$ appears in Martinetti [M2], without any justification or hint of proof. The majority of later writers do not mention the result – much less its proof – although many seem to accept it as selfevident. On the other hand, the statement in Page and Dorwart [P1] regarding this result is incorrect, as are the consequences deduced by them from the erroneous statement. It is interesting, as stressed by Gropp [G26], that Steinitz's result is very well known in combinatorics, but under a different name and credited to other people. It is considered a part of the branch of combinatorics called *matching theory*. (For this discussion, a *matching* in a graph is a collection of disjoint edges that contain all the nodes.) In this guise Steinitz's result from 1894 was independently discovered by König [K13] in 1916; in modern terminology König's theorem can be formulated it as: *Every bipartite graph having all nodes of the same valence has a matching*. This statement is completely equivalent with Theorem 2.5.1, although neither König nor many later writers seem to have been aware of Steinitz's theorem. In still another guise, Steinitz's theorem has been generalized by the theorem of P. Hall [H1] in 1935 concerning the existence of systems of distinct representatives. For details and proofs see, for example, Roberts [R5, Chapter 12] or Brualdi [B29, Chapter 9]. None of these authors is aware of Steinitz either, although the idea of Steinitz's proof is central to the topic.

We shall start by presenting a proof of this result, and then discuss some of its corollaries. Our proof is modeled after Steinitz's presentation, but using what I hope is a better notation. For easier understanding of the proof, a worked-out example is given later in the section. Except for the names of the points and lines, the steps in the example are precisely parallel to those of the proof. In contrast to most of the proofs of the equivalent results mentioned in the preceding paragraph, Steinitz's proof is constructive; it can be used to find effectively an orderly configuration table, convenient for geometric constructions. We shall see such an application in Section 5.2.

Given a fixed combinatorial configuration (n_k), the first goal is to define a 1-to-1 correspondence between points and lines such that each point is incident with (that is, is contained in) the corresponding line. If we have such a correspondence the first step in the proof is complete. We can certainly start constructing the correspondence by a *greedy algorithm:* We pick an arbitrary point and pair it with one of the lines that are incident with it; then we chose a point not on this line, and assign it to one of the lines containing it, then a point on neither of these lines, etc. Continuing with such a selection as long as possible, we find ourselves at the end in the following situation (adjusting the notation as appropriate and convenient):

The points in a subset $\mathcal{A} = \{a_1, a_2, \dots, a_p\}$ of the set of configuration points have been assigned to the lines of the subset $\mathsf{A} = \{A_1, A_2, \dots, A_p\}$ of the set of configuration lines, so that $a_j \in A_j$ for $j = 1, 2, \dots, p$. We can assume that $p < n$, since otherwise we would be done with the first part of the proof. Hence there is a set $\mathcal{B} = \{b_1, b_2, \dots, b_q\}$ of points of the configuration, and a set $\mathsf{B} = \{B_1, B_2, \dots, B_q\}$ of lines of the configuration, such that no point in $\mathcal{B}$ is incident with any line in B; clearly, $q = n - p$. Now we shall describe a procedure by which we shall change some of the assignments between points in $\mathcal{A}$ and lines in A, so that it will be possible to modify and extend the assignment to include one point in $\mathcal{B}$ and one line in B.

Let B be an arbitrarily chosen line in B, and let $\mathcal{A}_0$ be the subset of $\mathcal{A}$ consisting of the points of B. We denote by A_0 the set of lines in A that are associated with the points of $\mathcal{A}_0$. Let $\mathcal{A}_1 \subseteq \mathcal{A} \setminus \mathcal{A}_0$ be the set of points of $\mathcal{A}$ not in $\mathcal{A}_0$ that are on lines of A_0, and let A_1 be the set of lines associated with the points in $\mathcal{A}_1$. Next, let $\mathcal{A}_2 \subseteq \mathcal{A} \setminus (\mathcal{A}_0 \cup \mathcal{A}_1)$ be the set of points of $\mathcal{A}$ not in $\mathcal{A}_0 \cup \mathcal{A}_1$ that are on lines of A_1, and let A_2 be the set of lines associated with the points in $\mathcal{A}_2$. We continue with assignments of this kind till we reach an r such that $\mathcal{A}_{r+1}$ is empty. This clearly has to happen due to the finiteness of the configuration. Let now $\mathcal{A}^* = \mathcal{A}_0 \cup \mathcal{A}_1 \cup \dots \cup \mathcal{A}_r$ and $\mathcal{A}^{**} = \mathcal{A} \setminus \mathcal{A}^*$. Note that $\mathcal{A}^*$ is the **disjoint** union of the sets $\mathcal{A}_0, \mathcal{A}_1, \dots, \mathcal{A}_r$. Let A^* and A^{**} be the sets of lines associated with the points in $\mathcal{A}^*$ and $\mathcal{A}^{**}$, respectively.

We now pick a line $L_0 \in \mathsf{A}_0 \cup \mathsf{A}_1 \cup \dots \cup \mathsf{A}_r$ such that L_0 is incident with at least one point b of $\mathcal{B}$, so that $b \in L_0$. (Such a line always exists, by a simple counting argument that will be given below.) Let p_0 be the point of $\mathcal{A}^*$ that corresponds to L_0; then p_0 belongs to a well-determined set $\mathcal{A}_s$ for some $s \in \{0, 1, \dots, r\}$. Then $p_0 \in L_1$ for some $L_1 \in \mathsf{A}_{s-1}$, and let p_1 be the point of $\mathcal{A}_{s-1}$ that corresponds to L_1. Continuing in this way we reach a line $L_S \in A_0$ and the corresponding point $p_S \in \mathcal{A}_0$. Finally, there is a line $B \in \mathsf{B}$ such that $p_S \in B$. Notice that we have the chain of incidences and correspondences

$$b \in L_0 \leftrightarrow p_0 \in L_1 \leftrightarrow p_1 \in ... \in L_S \leftrightarrow p_S \in B.$$

Next, we change *for the points of this chain* the assignments with which we started by making b correspond to L_0, p_0 to L_1, p_{S-1} to L_S, and p_S to B. Thus we have now a new 1-to-1 correspondence which decreased the size of the sets $\mathcal{B}$ and B.

Repeating the procedure a finite number of times leads to an assignment of every point of the configuration to a line that is incident with it; this completes the first step of the proof, except for the demonstration of the assertion that we always can pick a line $L_0 \in \mathsf{A}^*$ which contains a point of $\mathcal{B}$. If this were not the case then all points of $\mathcal{B}$ would have to belong to A^{**}, since they do not belong to lines in B either. But this is not possible, since the cardinalities of $\mathcal{A}^{**}$ and A^{**} are the same due to the correspondence established at the beginning, and all incidences of points in $\mathcal{A}^{**}$ are with lines in A^{**} and vice versa — implying that no line in A^{**} can be incident with any point of $\mathcal{B}$.

For the second step we rewrite the configuration table in such a way that for each line (that is, each column) the point assigned to it is in the first row. Then the first row contains all the points, each once. The other rows of the configuration table form now a configuration (n_{k-1}), for which we repeat the steps we just did for the original configuration. Continuing in this way, we clearly reach an orderly configuration table in a finite number of steps. □

It may be mentioned that when we have only two rows to deal with, a simple interchange of the order of the entries in some columns may be used instead of the more complicated procedure used in the general case.

We next illustrate the algorithm used in the proof of Theorem 2.5.1 by an example, the construction of an orderly configuration table for the combinatorial configuration (14_4) given below.

A	B	C	D	E	F	G	H	J	K	L	M	N	P.
a	a	a	a	b	b	b	c	c	c	d	d	d	e
b	f	g	h	g	h	e	h	e	f	e	f	g	f
c	k	n	p	k	m	p	k	m	n	k	q	m	g
d	m	r	q	q	n	r	r	q	p	n	r	p	h

We select the starting assignments as follows:

A	B	C	D	E	F	G	H	J	K	L	M	N	A
a	f	g	h	b	m	e	c	q	n	d	r	p	$\mathcal{A}$

and rewrite the table as

A	B	C	D	E	F	G	H	J	K	L	M	N	P
a	f	g	h	b	m	e	c	q	n	d	r	p	e
b	a	a	a	g	b	b	h	c	c	e	d	d	f
c	k	n	p	k	h	p	k	e	f	k	f	g	g
d	m	r	q	q	n	r	r	m	p	n	q	m	h

so that the assigned points are in the first row for better visibility. We are left with

$\{k\} = \mathcal{B}$ $\{P\} = \mathsf{B}$.

We put:

$\mathcal{A}_0 = \{e, f, g, h\}$ = set of points on P, which happens to be the only line of B. Then
$\mathsf{A}_0 = \{G, B, C, D\}$ = associated set of lines of A.
$\mathcal{A}_1 = \{b, p, r, a, m, n, q\}$ = points of $\mathcal{A} \setminus \mathcal{A}_0$ on lines of A_0.
$\mathsf{A}_1 = \{E, N, M, A, F, K, J\}$ = associated set of lines of A.
$\mathcal{A}_2 = \{d, c\}$ = points of $\mathcal{A} \setminus (\mathcal{A}_0 \cup \mathcal{A}_1)$ on lines of A_1.
$\mathsf{A}_2 = \{L, H\}$ = associated set of lines of A. Finally
$\mathcal{A}_3$ = empty.
Hence we have

$$\mathcal{A}^* = \mathcal{A}_0 \cup \mathcal{A}_1 \cup \mathcal{A}_2 = \{a, b, c, d, e, f, g, h, m, n, p, q, r\}$$

$\mathcal{A}^{**} = \mathcal{A} \setminus \mathcal{A}^{*}$ in this case empty but need not be empty in general.

Now we pick a line of $\mathsf{A}_0 \cup \mathsf{A}_1 \cup \mathsf{A}_2$ that contains an element of $\mathcal{B}$. In our case there is only one such element, k, and we have a choice of lines: B, E, or L. In each case we can form a chain:

$k \in B \leftrightarrow f \in P$ or

$k \in E \leftrightarrow b \in G \leftrightarrow e \in P$ or

$k \in L \leftrightarrow d \in N \leftrightarrow p \in G \leftrightarrow e \in P$ and use it to change the assignments.

We use the last, and it leads to a rewritten table:

A	B	C	D	E	F	G	H	J	K	L	M	N	P
a	f	g	h	b	m	p	c	q	n	k	r	d	e
b	a	a	a	g	b	e	h	c	c	d	d	p	f
c	k	n	p	k	h	b	k	e	f	e	f	g	g
d	m	r	q	q	n	r	r	m	p	n	q	m	h

Now we deal in the same way with the last three rows.

A	B	C	D	E	F	G	H	J	K	L	M	N	A
b	a	n	p	g	h	e	k	c	f	d	q	m	$\mathcal{A}$

Then we are left with

$\{r\} = \mathcal{B}$ $\{P\} = \mathsf{B}$

This time we put:

$\mathcal{A}_0 = \{f, g, h\}$ = set of points on a line (P) of B.
$\mathsf{A}_0 = \{K, E, F\}$ = associated set of lines in A.
$\mathcal{A}_1 = \{c, p, k, q, b, n\}$ = points of $\mathcal{A} \setminus \mathcal{A}_0$ on lines of A_0.
$\mathsf{A}_1 = \{J, D, H, M, A, C\}$ = associated set of lines in A.
$\mathcal{A}_2 = \{e, m, d, a\}$ = points of $\mathcal{A} \setminus (\mathcal{A}_0 \cup \mathcal{A}_1)$ on lines of A_1.

A_2 = {G, N, L, B} = corresponding set of lines in A.
$\mathcal{A}_3$ = empty.

Then we put:

$\mathcal{A}^* = \mathcal{A}_0 \cup \mathcal{A}_1 \cup \mathcal{A}_2$ = {a, b, c, d, e, f, g, h, m, n, p, q, r}

$\mathcal{A}^{**} = \mathcal{A} \backslash \mathcal{A}^*$ = in this case empty (need not be empty in general) .

Now we pick a line of $\mathsf{A}_0 \cup \mathsf{A}_1 \cup \mathsf{A}_2$ that contains an element of $\mathcal{B}$.

In our case there is only one such element, r, and we have a choice of the lines C and G. In each case we can form a chain:

$r \in C \leftrightarrow n \in F \leftrightarrow h \in P$

$r \in G \leftrightarrow e \in J \leftrightarrow c \in K \leftrightarrow f \in P$.

We shall use the former to change the assignments.

A	B	C	D	E	F	G	H	J	K	L	M	N	P
a	f	g	h	b	m	p	c	q	n	k	r	d	e
b	a	r	p	g	n	e	k	c	f	d	q	m	h
c	k	a	a	k	b	b	h	e	c	e	d	p	f
d	m	n	q	q	h	r	r	m	p	n	f	g	g

Making interchanges in columns C, E, G, J, K, N we finally reach the orderly table

A	B	C	D	E	F	G	H	J	K	L	M	N	P
a	f	g	h	b	m	p	c	q	n	k	r	d	e
b	a	r	p	g	n	e	k	c	f	d	q	m	h
c	k	n	a	q	b	r	h	m	p	e	d	g	f
d	m	a	q	k	h	b	r	e	c	n	f	p	g

in which each point appears in every row.

* * * * *

Before proceeding with the next step in our study of Steinitz's theorem and its ramifications, we recall from Section 1.3 the concept of "multilaterals". A **multilateral**

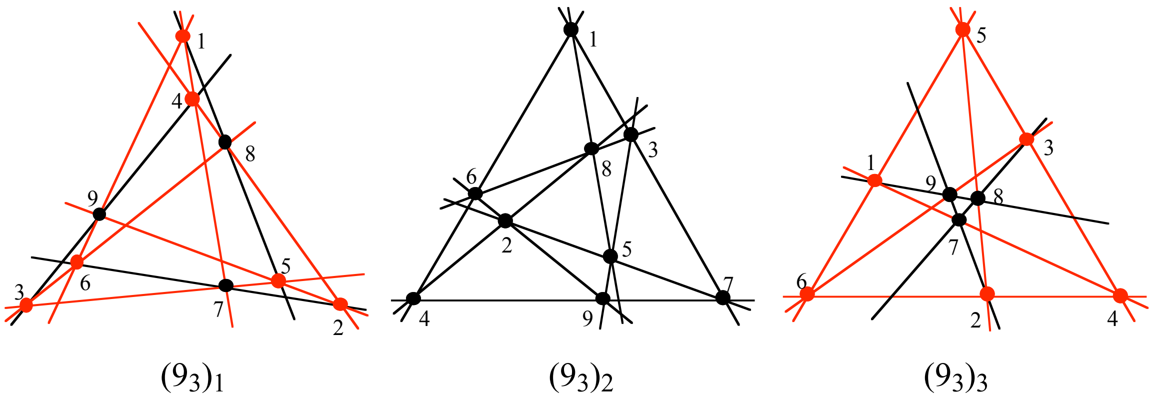

Figure 2.5.1. Some examples of multilaterals, in the three configurations (9_3) shown in Figure 2.2.1. The first is a 6-lateral with sequence of points 1,4,2,5,3,6 (1) in the configuration $(9_3)_1$. The second is 9-lateral, with sequence of points 1,8,3,7,5,9,4,2,6 (1); since this multilateral involves all points (and hence also all lines) it is a Hamiltonian multilateral. The last diagram shows a trilateral 7,8,9 (7), and a 6-lateral 1,6,3,5,2,4 (1). Note that another 6-lateral is 1,5,2,6,3,4 (1). The 3-lateral and either of the 6-laterals taken together form a multilateral decomposition of the configuration $(9_3)_3$.

(often inconsistently called "polygon" in the literature) is any sequence of distinct points and distinct lines of a configuration that can be written as P_0, L_0, P_1, L_1, … , P_{r-1}, L_{r-1}, P_r (= P_0), with each L_i incident with P_i and P_{i+1} (all subscripts understood mod r). Some examples of multilaterals are shown in Figure 2.5.1. If the last point is not required to coincide with the first one, we are dealing with a **multilateral** path. A family of multilaterals in a configuration, that contains all points and all lines but each just once, is called a **multilateral decomposition** of the configuration. We shall return to the topic of multilaterals later (for example, in Chapter 5).

Our next aim is to modify an orderly configuration table in a way that will preserve its orderly character but will be useful for the geometric steps. We assume that a line and one if its points are selected to be ignored in the geometric implementation, and that, as before, the configuration is connected. We also assume that we are concerned with a 3-configuration.

First, the rows are permuted so that the selected point of the selected line is in the first row. Note that since the table is orderly, this yields a correspondence (possibly different from the one we started with) in which each point is associated with a line that contains it. As mentioned earlier, and as is easily seen, by possibly interchanging the order of the columns (that is, lines) the orderly configuration table can be rearranged to show the multilateral decomposition in such a way that the lines of each constituent multilateral occur consecutively. In each of the multilaterals we can assume that the point in the last position in one column is in the middle position in the following column (understood modulo length of the multilateral). We rearrange the columns in such a way that the multilateral that contains the chosen line is placed last, and the selected line is chosen as the last line in the multilateral. If the multilateral is Hamiltonian, this part of the proof is completed. Otherwise, since the configuration is connected, at least one of the points of the last multilateral must be associated to (that is, be in the first row of) a line which is not in the multilateral. Choose the multilateral containing this line to be the next-to-last, and the line in question to be its last line. Then some point of this multilateral must be associated with another multilateral not used so far, and we continue in the same way. At the end we reach what we may call an **arranged** configuration table. This proves

Theorem 2.5.2. Every connected 3-configuration has an arranged configuration table.

As an illustration, we show in Table 2.5.2 an orderly configuration table of a configuration (14_3). Rearranging the columns so as to make the multilateral decomposition visible, we obtain the arranged configuration table, Table 2.5.3.

A	B	C	D	E	F	G	H	J	K	L	M	N	P
c	k	n	a	q	b	r	h	m	p	e	d	g	f
d	m	a	q	k	h	b	r	e	c	n	f	p	g
b	a	r	p	g	n	e	k	c	f	d	q	m	h

Table 2.5.2. An orderly configuration table of a connected combinatorial configuration (14_3).

A	M	P	N	K	B	J	L	C	D	E	F	H	G
b	q	h	m	f	a	c	d	r	p	g	n	k	e
c	d	f	g	p	k	m	e	n	a	q	b	h	r
d	f	g	p	c	m	e	n	a	q	k	h	r	b

Table 2.5.3. A rearranged configuration table of the (14_3) configuration of Table 2.5.2, in which the lines of each multilateral appear as consecutive columns. The point e of line G was chosen as the exceptional point, so its row is the first one. The line G is the last line of its multilateral, which is the last multilateral. Each multilateral is specified by rows 2 and 3 of the table. The table is arranged, in the sense described earlier.

We shall see in the next section how such an arranged multilateral decomposition can be used geometrically. Here we shall conclude the section by discussing certain ramifications of the results we have seen so far.

Corollary 2.5.3. Every connected k-configuration C, with $k \geq 2$, admits multilateral decompositions.

Indeed, any two rows of an orderly configuration table determine, by the above, a multilateral decomposition of C.

Corollary 2.5.3. Every connected k-configuration C, with $k \geq 2$, is 2-connected.

Proof. Assume that C is a connected k-configuration such that, without loss of generality, there is a line L for which for suitable elements R' and R" there is no R'-to-R" multilateral path that misses L. By the connectedness of C, there is a multilateral path M that uses L, that is, there are two points Q' and Q" of L that are part of this path M. In an orderly configuration table of C, permuting the rows if necessary, we may put Q' and Q" in the last rows of the block L. Let S be a multilateral decomposition of C determined by the last two rows of this orderly configuration table. Then one of the multilaterals of this decompositions uses the points Q' and Q". But since the multilateral is a circuit, there is a multilateral path (formed by the lines other than L) that connects Q' and Q". Substituting this path for the one that originally connected R' and R" eliminates the use of L. Hence the assumption that each path between R' and R" uses L is incorrect, and so C is 2-connected.

We shall discuss additional connectedness results in Section 5.1.

Exercises and problems 2.5.

1. Use the procedure applied in the proof of Theorem 2.5.1 to replace the configuration table in Table 2.5.4 by an orderly configuration table.

a	b	c	d	e	f	g	h	i	j	k	l	m	n
1	1	1	1	2	2	2	3	3	3	4	4	6	7
2	5	6	10	3	5	8	4	5	11	5	9	7	8
4	8	9	13	9	6	12	6	7	12	10	11	10	9
7	11	12	14	10	14	13	8	13	14	12	13	11	14

Table 2.5.4. A (14_4) configuration table.

2. Find orderly configuration tables for the two (12_3) configurations in Figure 2.5.2.

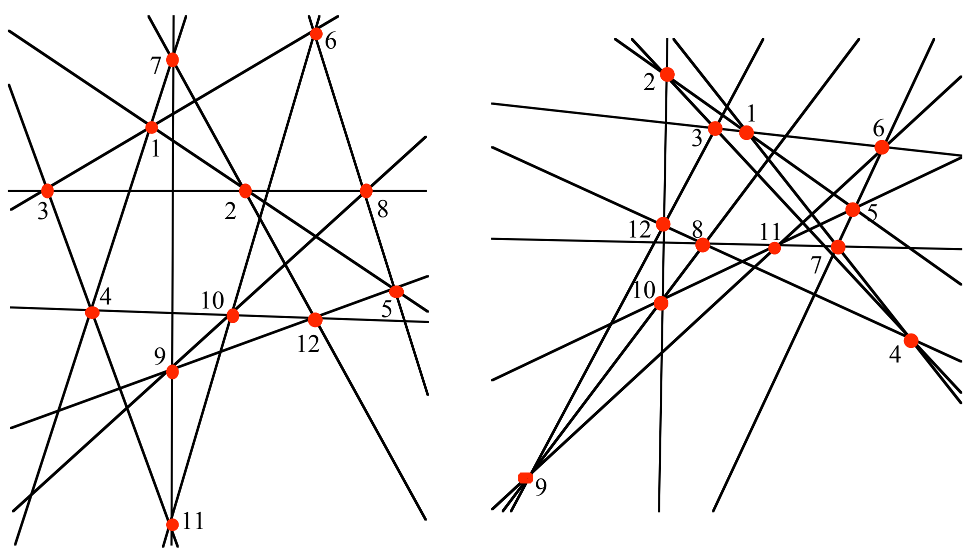

Figure 2.5.2. Two (12_3) configurations.

3. Justify the statement: *Any selection of two rows of an orderly configuration table defines a multilateral decomposition of the configuration.* List all multilateral decompositions resulting from possible choices of the rows in the orderly configuration tables found in Exercise 2.5.2 for the configurations (12_3) shown in Figure 2.5.2.

4. Justify the statement: For every k-configuration **C** and every multilateral decomposition of **C**, there is an orderly configuration table in which the multilateral decomposition can be obtained from the first two rows of the table.

5. Modify the proof of Theorem 2.5.1 to establish the following strengthening: *Every combinatorial k-configuration admits an orderly configuration table in which an arbitrarily chosen line is the last line of the table.*