Arch. Hist. Exact Sci. 51 (1997) 3–27. © Springer-Verlag 1997

On the Origins of Dénes König's Infinity Lemma[1]

MIRIAM FRANCHELLA

Communicated by U. BOTTAZZINI

KÖNIG's infinity lemma affirms that, if there is no finite upper bound to the length of paths in a finitary tree, then there is at least one infinite path in the tree. This lemma[2] has been a powerful — perhaps over powerful — instrument in classical mathematical proofs, as W. T. TUTTE wrote[3] concerning graph theory: 'The *Unendlichkeitslemma* has been a powerful tool for investigating locally finite graphs. In a sense it has been too powerful, causing some of us to lose interest in that kind of graph. Their theory, we say, is but a succession of exercises on KÖNIG's lemma". Furthermore, its history presents some intriguing situations. As M. DUMMETT pointed out[4], "there is no reason to suppose KÖNIG's lemma to be *constructively* true . . . the mere fact that there is no finite upper bound on the lengths of paths does not supply us with any way of doing this". KÖNIG's lemma appeared for the first time in his constructivist period as the main step in demonstrating a very general form of the BERNSTEIN theorem, at the end of a series of research studies which JULIUS KÖNIG's proof of the BERNSTEIN theorem inspired. This lemma is attributed to KÖNIG, while in its formulation — as DÉNES KÖNIG himself always stressed — it belonged to a proof that was mainly and essentially ascribable to STEPHAN VALKÓ.

The fact that it is a fruitful lemma together with these circumstances justify investigating its history in order to clarify the role of DÉNES KÖNIG and the influence of others in its formation.

---

[1] This work was supported by CNR (contratto nr. CT94.03342.CT08.115.16509).

[2] For the sake of completeness, it is useful to recall that this lemma has to be distinguished from that of DÉNES' father:

"If the single-valued functions f and g assign to each member t of a non-empty set T cardinals $f(t) = \mathbf{a}_t$ and $g(t) = \mathbf{b}_t$ such that $\mathbf{a}_t < \mathbf{b}_t$ for each $t \in T$, then $\Sigma_t \mathbf{a}_t < \Pi_t \mathbf{b}_t$."

Furthermore, it should be noticed that the infinity lemma has also to be distinguished from what is called the EGERVÁRY–KÖNIG theorem [see GALLAI *1978*: *199–200*]:

"In a bipartite graph the minimal number of vertices which exhaust the edges equals the maximal number of edges, no two of which have a common vertex."

[3] TUTTE [*1990*: 25].

[4] DUMMETT [*1977*: 71].

1. Julius and Dénes

DÉNES KÖNIG was born in Budapest on 21 September 1884[5]. In 1899, while still a gymnasium student, his first paper was published in the journal *Mathematikai ès Fizikai Lapok*. In this, he gave an elementary discussion of two extreme-value problems. In the same period, he wrote the first volume of his book *Mathematical Recreations*, which appeared in 1902, followed, in 1905, by the second volume. This was the first high-quality work on recreational mathematics written in Hungarian, and it was very successful. DÉNES attended the first four semesters of his university career at the University of Budapest, and the last five at Göttingen. In 1907, he obtained his Ph.D. with a dissertation on a geometrical topic. From that time until his death in 1944, he remained attached to the Polytechnical University in Budapest, as *Privatdozent*, as extraordinary professor and, from 1935, as professor. On 19 October 1944, he sought refuge from racial persecution by taking his own life[6].

His scientific interests first embraced topology, in the forms of combinatorial topology and graph theory. In writing his book on recreational mathematics he encountered combinatorial topology and he studied it in more detail in Göttingen under MINKOWSKI. He lectured on it at the very beginning of his career, and devoted two papers [*1913a–b*] and a textbook [*1918*] to it.

Graph theory was, however, his main subject. As he himself affirmed [*1926: 127–128*], he had studied GORDAN and HILBERT on invariant theory and they referred to PETERSEN's article on graph theory [*1879*], that arose from the GORDON-HILBERT theorem stating the finiteness of the binary form of the system of invariants corresponding to a certain class of diophantine equations. DÉNES immediately appreciated graph theory as a tool allowing problems (and solutions) to become evident, or visible: hence, he engaged himself in making it a self-standing, respectable discipline (it had previously been considered "the science of trivial and amusing problems for children"[7]). His first published works that used graphs had been [*1911a–b*]. Later, he translated problems coming from other areas of mathematics into graph-theoretical terminology, and solved them in that (more visible) form. He contributed to graph theory with many articles [*1915, 1916, 1923, 1926, 1927*] and his famous 1936 book, in which he touched on all the relevant items in the literature, presented the entire history of graph theory (in his footnotes) and considered many applications of the subject. He was the first to study infinite graphs.

Finally, DÉNES KÖNIG was interested in set theory. His work in this field took place in the period 1908–1926, and was also carried out utilizing tools from graph theory. His lemma was indispensable for this work and appeared for the first time. DÉNES KÖNIG's attention was directed towards set theory by his father JULIUS, an eminent mathematician engaged in big, open mathematical

[5] The source for these biographical notes has been GALLAI [*1978*].
[6] Although he was born a Christian, his parents were of Jewish origin.
[7] TUTTE [*1992*: *1*].

questions. In 1904, JULIUS presented a communication to the third International congress of mathematicians at Heidelberg, where he had refuted the continuum hypothesis by showing that the continuum could not be well-ordered, so the power of the continuum was not an aleph. *For this reason*, his initial position was against ZERMELO's axiom of choice, by which ZERMELO had proved that every set could be well-ordered. However, in two steps of his proof, JULIUS essentially made use of arbitrary choices[8] and relied on an equivalence result that FELIX BERNSTEIN had proved in his recent dissertation for finite ordinals and which JULIUS considered valid for any ordinal. It was ZERMELO himself who discovered the error. Then, in the proceedings of the congress, JULIUS limited his result to the fact that *if* the BERNSTEIN theorem were valid in general, then the continuum could not be well ordered and vice versa. Yet, the proof of this last fact was also wrong, because it utilized a form of trichotomy, which was sufficient to guarantee that the continuum can be well-ordered. Later, in 1905 and 1906, he tried again to refute the continuum hypothesis, but unsuccessfully. In the following years, he abandoned this direction of research, went nearer to POINCARÉ's thought and directed himself to justifying axioms and theorems that seemed to be unacceptable from a constructivistic point of view. In particular, he gave a proof of the CANTOR-BERNSTEIN-SCHRÖDER theorem without the use of induction[9] and wrote a foundational book — published soon later after his death [*1914*] —, where he developed a constructivist view of mathematics based on mental evidence and *on these grounds* he arrived at the conviction that the axiom of choice was acceptable.

The whole story of KÖNIG's infinity lemma comes about because JULIUS KÖNIG's use of an equivalence theorem by FELIX BERNSTEIN in his 1904 attempt to refute the continuum hypothesis, and his later work on the CANTOR–BERNSTEIN–SCHRÖDER theorem, drew DÉNES' attention toward the series of

---

[8] For all details about JULIUS' position on the axiom of choice, see MOORE [*1982*].

[9] What is commonly known as the CANTOR–BERNSTEIN–SCHRÖDER theorem is the affirmation that if each of two sets is equivalent to a proper subset of the other, then they are equivalent to one another. (In symbols: "If $A \sim B_1$ and $B \sim A_1$, then $A \sim B$", where $A_1 \subset A$ and $B_1 \subset B$.) H. POINCARÉ [*1906*: *27–29*], by polemizing with L. COUTURAT's opinion that a theory about infinite sets requires no use of induction rules, pointed out that the main theorem of infinite sets was the CANTOR–BERNSTEIN–SCHRÖDER theorem whose proof uses induction. In [*1906*: *110*] JULIUS KÖNIG affirmed that he shared the critical part of H. POINCARÉ's 1906 article, promised to express his own constructive considerations in a future book [*1914*] and devoted himself at that moment to a proof of the CANTOR–BERNSTEIN–SCHRÖDER theorem that avoided using induction. He did not split the sets at issue, but just used the laws of correspondence between A and B_1(I), and between B and A_1 (II) expressed in the statement of the theorem, and showed that these laws also allowed the formation of a bijective correspondence between A and B. J. KÖNIG used sequences, indexed by natural numbers, but he stressed that "1", "2", . . . , "n" were referred to as signs and not as numbers so that induction did not appear even as a way of defining. He also added that its presentation still needed further "adjustments", in particular a "logical definition" for "successor" and "antecedent", in order to be sure that he avoided inductive procedures.

equivalence theorems which BERNSTEIN had presented. Along the initial lines of his father's research, he tried in 1908 to give a proof of two of these theorems without the well-ordering principle. Then, when his father changed his mind about the acceptability of this principle and the axiom of choice, he went on to study equivalence theorems, because he considered his father's proof of the CANTOR-BERNSTEIN-SCHRÖDER theorem as an example of the possibility of simplifying BERNSTEIN's proofs in general. At this point, his familiarity with PETERSEN's graph-theoretical terminology first let him see the possibility of translating his father's proof (and the statements of other equivalence theorems) into graph-theoretical terminology. Second, it seemed to him to suggest new generalizations, and hence new equivalence theorems to be proved. It was one of these that required the infinity lemma and that could not be proved until the lemma had been introduced.

2. First attempts

His first article devoted to this subject was "Sur un problème de la théorie générale des ensembles et la théorie des graphes", communicated in 1914 at the congress of mathematical philosophy in Paris and published nine years later. At the beginning [*1923: 444–446*] he proved the equivalence theorem:

If $3M = 3N$ *then* $N = M$,

where M and N are two (finite or denumerable) sets, by adding that he referred to the number "3" just to fix ideas, and that it would be enough to change some words in order to get the proof for any natural number n. That is to say, he aimed to prove:

[BERNSTEIN Theorem] *If* $nM = nN$ *then* $N = M$

(from now on we will call this the BERNSTEIN theorem *tout court*).

He noticed that even BERNSTEIN's attempts to prove "If $2M = 2N$ then $N = M$" had produced a very complicated proof. He himself wanted to give a simpler one. He expressed the theorem as follows:

If $M_1 \sim M_2 \sim M_3,\ N_1 \sim N_2 \sim N_3$ (1)

and $M_1 + M_2 + M_3 \sim N_1 + N_2 + N_3$ (2)

then $M_1 \sim N_1$

He defined:

$$M_1 = (a_1, b_1, c_1, \ldots);\quad N_1 = (x_1, y_1, z_1, \ldots);$$
$$M_2 = (a_2, b_2, c_2, \ldots);\quad N_2 = (x_2, y_2, z_2, \ldots);$$
$$M_3 = (a_3, b_3, c_3, \ldots);\quad N_3 = (x_3, y_3, z_3, \ldots),$$

where the elements labelled with the same letter were those corresponding to each other according to the equivalence (1) of the premises of the theorem.

He formed the sets M and N, equivalent to M_i and N_i respectively. Then, he expressed M and N as a sum of subsets M^* and N^*, each of them being at most denumerable. Then, he showed that M^* and N^* were equivalent, so that M and N were also equivalent by composition.

M and N were built as follows:

$$M = [(a_1, a_2, a_3), (b_1, b_2, b_3), \ldots]$$

$$N = [(x_1, x_2, x_3), (y_1, y_2, y_3), \ldots]$$

These sets were clearly equivalent to M_i and N_i respectively, since they contained the same number of elements.

Then, König defined a *correspondence* between an element of M and an element of N: an element (a_1, a_2, a_3) corresponded to the element (x_1, x_2, x_3) if at least one of the three a_i $(i = 1,2,3)$ corresponded to one of the x_i $(i = 1,2,3)$.

$M + N$ could be expressed as the sum of the subsets $M^* + N^*$, where M^* was a subset of M; N^* was a subset of N; and every $M^* + N^*$ contained only elements of M and N that could be arranged into a *chain*, that is to say, a sequence such that any two consecutive terms (that belonged respectively to M and N) corresponded to each other. In practice, two elements of $M + N$ belonged to a certain subset $M^* + N^*$ if they could be joined by such a chain.

Every $M^* + N^*$ was either finite or denumerable. That is to say, every element of $M^* + N^*$ (which was a triple) corresponded at most to three elements of M (or of N, according to the fact that the element belonged to N or to M), which were in $M^* + N^*$. On their side, to each of these elements of M (or N) there corresponded three or more elements of $N(M)$, which were in $M^* + N^*$ (at most nine: 3^2). By repeating this reasoning k times (where k represented any finite number), 3^k elements were always obtained. This was the maximum number of elements of $M^* + N^*$, which was therefore at most denumerable. *A fortiori* both M^* and N^* separately were at most denumerable.

König called m^* the power of M^* and n^* the power of N^*. The power of the set of the pairs $[a, x]$ where the element a of M corresponded to the element x of N was therefore $3m^*$, if the same pair were counted 1, 2, 3 times, according to the fact that one, two or three a_i's corresponded to one x_i. For the same reason, this power was also equal to $3n^*$. Hence, König could write:

$$3m^* = 3n^*.$$

Since both m^* and n^* were at most denumerable, in this case, the Bernstein theorem obviously held and, hence, could be applied in order to obtain

$$m^* = n^*.$$

From this, König obtained

$$M^* = N^*$$

and by summing

$$M = N.$$

The theorem proved, KÖNIG translated this first set-theoretical theorem into graph-theoretical terms. In particular, he utilized three concepts introduced by PETERSEN: regular graph, factor, degree. It should be noted that, in the papers we are considering, KÖNIG admitted the possibility of two vertices of a graph being joined by more than one edge and considered only undirected graphs.

KÖNIG considered the elements of the sets M and N (that is to say, their triples) as the vertices of the graph, and every edge of the graph as representing the correspondence between two triples of M and N: for every correspondence starting from an element of a triple of M there was an edge; the same for every correspondence starting from an element of a triple of N. Therefore, in the case above, every vertex had three edges going to it: hence, according to PETERSEN's definitions, the graph was *regular* (because every vertex had the same number of edges going to it) and *of degree* 3 (because the number of edges going to every vertex was three). The number of vertices was either finite or denumerable as the number of elements of M and N was such. The graph also had two further characteristics. First, it was *bipartite*[10], that is, each of its cycles contained an even number of edges, since, given the endpoints of any edge, one belonged to N and the other to M. Second, the graph was formed by *connected* (sub)graphs corresponding to the subsets $M^* + N^*$ of $M + N$, where a (sub)graph was said to be connected if for any vertices P and Q there was a path joining P and Q. The BERNSTEIN theorem statement became: "If the vertices of a regular bipartite graph are alternately attached to the letters M and N, the number of the vertices M and of the vertices N is the same."

KÖNIG did not give the proof of the theorem expressed with the graph-theoretical terminology but introduced a stronger theorem from which the BERNSTEIN theorem could be deduced. It was a generalization of one of PETERSEN's theorems[11] and used a further PETERSEN term: *factor* of *degree* n of a graph as the subset G_n of edges of a graph G such that every vertex of G has exactly n edges of G_n going to it. The factor of degree n was itself a regular graph of degree n. Some edges of G formed a *factor of degree* 1 of G if every vertex of G occurred one and only once as an endpoint of these edges. The theorem (called "theorem A") affirms that:

[A] *Every regular bipartite graph has a factor of degree* 1,

where the factor of degree 1 in the specific case at issue was a subset of the graph in which every vertex M was connected by just one edge to a vertex N and vice versa.

---

[10] As KÖNIG himself remarked [*1923*: *449*], PETERSEN had not individuated this notion.

[11] KÖNIG [*1915*: *455*] explained that this theorem was a generalisation of PETERSEN's theorem stating that a second order graph had a factor of degree 1 if and only if every cycle contained a pair number of edges.

In effect, he only had partial results with theorem A: he was able to prove it for graphs of degree 2 and individuated a further theorem (theorem B) as a corollary of theorem A:

> [B] *Every regular bipartite graph is the product of factors of degree* 1; *the number of these factors is equal to the degree of the graph itself.*

As for the case of graphs of degree 2, the proof was simple: in this case, the factor of degree 1 was obtained by removing each second edge going to the vertices. KÖNIG also readily showed that theorem B was a corollary of theorem A.

Theorem A allowed the splitting of the graph G into the graph G' of degree 1 and its complementary graph G_{n-1} of degree $n-1$, which were still regular and bipartite. That is, on the one hand, KÖNIG showed that it was always possible to consider a graph of degree 1, which was obviously regular, as a bipartite graph; on the other hand, the complementary graph above was regular and bipartite because removing an edge of each vertex did not make the graph lose its property of being bipartite and regular.

In its turn, this graph of degree $n-1$, according to theorem A, could be split into a factor of degree 1 and its complementary graph of degree $n-2$, and so on. In the end, KÖNIG obtained the splitting:

$$G = G_1 + G_1' + G_1'' + \ldots$$

where each member of the splitting was a factor of degree 1 (that is to say, he obtained theorem B).

We meet this theorem again later, when KÖNIG realised that it is equivalent to theorem A and not only one of its corollaries. However, the problem at that point was to prove theorem A. He just gave some hints: in each case to try to find the factor of degree 2, in order to use the above proof.

3. A further step

In his article "Über Graphen und ihre Anwendung auf Determinantentheorie und Mengenlehre" [*1916*], KÖNIG left set-theoretical terminology totally aside in favour of graph-theoretical terminology, gave a proof of the graph-theoretical version of BERNSTEIN's theorem, introduced a further theorem C, and realised that theorems A, B, and C were equivalent to each other and to another theorem I.

As for the proof of the graph-version of BERNSTEIN's theorem, first of all, he still expressed it in set-theoretical terms but without using the notion of power (theorem H):

> [H] *If an invertible* $(1-n)$ *correspondence exists between the sets M and N, then there is an invertible* $(1-1)$ *correspondence between M and N.*

It must be specified that "invertible" meant that the "multiplicity" of the correspondence of an element a of one set M to an element b of the other set

N was the same as that of b with respect to a, where *multiplicity* meant the number of times in which one element corresponded to another. In other words, "invertible" meant that if an element a of M corresponded n times to the same element b of N, then b also corresponded n times to a.

Given this expression of the BERNSTEIN theorem, it was possible for KÖNIG to translate it into graph-theoretical terms, as we have seen in § 2. He specified that he had to put a restriction on the number of edges (that is to say, each vertex had only a finite number of edges going to it), owing to the fact that he had to prove "If $nM = nN$ then $M = N$", where n was a natural number, that is, a finite number.

In proving the theorem in graph-theoretical terms, KÖNIG first introduced a theorem (theorem G) that reduced the general case to the case where the graph was at most denumerable, when the graph has the above restriction on the number of edges going to each vertex:

> [G] *If the number of the edges starting from each point of a graph remains under a certain finite limit h, then the graph splits into two parts, a finite one and an infinite one.*

This theorem did not require the graph to be connected. Yet, since the graph at issue was also connected, for this case, KÖNIG proved directly that the whole graph was either finite or denumerable. The proof was very simple: the number of vertices of the graph G was the number of vertices that could be reached by a finite number of edges, that is to say, by a finite number of steps. By one step from a point P, at most h points could be reached; by two steps, at most h^2 points could be reached; . . . hence, in general, by n steps at most h^n steps could be reached: this was the number of vertices of G, which were therefore at most denumerable.

Given theorem G, the proof of theorem H (BERNSTEIN's theorem) was even easier than the proof KÖNIG himself used in 1914 for the case $n = 3$. Namely, he let G_1 be one of the (finite/denumerable) parts in which the graph could be split. Its vertices represented the subsets M_1 and N_1 of M and N whose powers were m_1 and n_1 respectively. As G_1 was at most denumerable, m_1 and n_1 were also at most denumerable. The same held for every part of the graph. Therefore, KÖNIG could apply the "obvious" version of BERNSTEIN's theorem to them and obtain the final result by addition.

He could not prove any of the other theorems: in this case, too, he could only give partial results. He stated theorems A and B as in the preceding article, then he introduced theorem C:

> [C] *If every vertex of a regular bipartite graph has at most k edges going to it, then it is possible to assign one of k indices to each edge in such a way that two edges which meet at the same vertex never receive the same index.*

KÖNIG affirmed that:

1) It was apparent that A followed from B as a specific case.
2) Also the inverse held, as he had shown in 1914.

3) B was a consequence of C, where the vertices with the same index formed a factor of degree 1.
4) These relationships held both for finite and infinite graphs.

Finally, he proved that C was also a consequence of B. The proof ran as follows. He began with a graph G such that each of its vertices has *at most* k edges going to it. If he could use theorem B, he could split the graph into k factors and assign a different index to each factor. Therefore, he had to find a method for applying theorem B which referred to graphs of degree k, such that each of their vertices had *exactly* k edges going to it. In practice, he had to "fill up" ("ergänzen") G. The method consisted of exactly reproducing G (calling the new graph G') and then joining each vertex of G with the correspondent of G' by $k - \alpha$ edges, where α ($\leq k$) was the number of edges going (in G) to that vertex: in this way, every vertex of the whole graph had $\alpha + (k - \alpha)$ edges going to it, that is, exactly k edges. This whole graph was obviously still regular and bipartite, because it was possible to divide its vertices exactly into the same two groups into which they were divided in G as it was bipartite. At this point KÖNIG could use theorem B to draw the result.

Having stated that theorems B and C were equivalent, it was sufficient to prove C in order to obtain the other. The proof *for a finite graph* was by induction. KÖNIG supposed the theorem to hold for any graph of degree $k - 1$. Then, he removed an arbitrary edge AB of the k edges (that is to say, an edge connecting an arbitrary vertex A and another arbitrary vertex B), obtaining a graph G', still bipartite, whose edges could be assigned indices from the k numbers in a way corresponding to the theorem. There were two possibilities:

1) There was an index that did not occur either in an edge going to A or in an edge going to B. In this case this index was assigned to the removed edge.
2) There was an index (call it 1) that did not occur in any edge going to B, but occurred in an edge going to A (call it AA_1). The trick consisted in eliminating it also from this last edge, so that this second case was reduced to the first one. In order to eliminate it, KÖNIG started from AA_1 and followed a path $A\ A_1\ A_2 \ldots$ whose edges had alternate indices (for example 1 and 2). The path had these characteristics:

a) It was finite, since the whole graph was such.

b) No vertex could occur twice; for, if A_i were the first vertex, which had already occurred once, then two edges with index 1 or two edges with index 2 would have to occur at A_i. This was impossible. It was also impossible to return to A again, since 2 did not occur at A and 1 was already used (with the edge AA_1).

c) One never got to B on this path, for this could only happen with an edge of index 2 (1 did not occur at B), and then the path from A to B would contain an even number of edges; combined with the omitted edge AB, this path — contrary to the assumptions — would give a cycle with an odd number of vertices.

The path characterized by a)-b)-c) could be considered as detachable from the rest of the graph. Therefore, KÖNIG was allowed to switch the indexes of the path, without changing the indexes of the other edges. The distribution of the indexes also satisfied the requirements after this switching. This is seen not only for the first vertex A, from which no edge with index 2 went out originally, but for the interior vertices A_i as well, and for the endpoint A_r, from which no second edge with index 1 (or 2 respectively) could go out, since otherwise the path could not end in A_r. In this way KÖNIG arranged to reduce the second case to the first.

According to KÖNIG himself [*1936*: *172*], there was an influence of KEMPE's work [*1879*] in this proof: it concerned a procedure of switching district-colours that suggested to him the idea of switching edge-indexes. The procedure was expressed in the following passage [*1879*: *194*]:

"Suppose that we have the surface divided into districts coloured with four colours, viz: blue, yellow, red and green; and suppose that the districts are so coloured. Now if we draw our attention to those districts which are coloured red and green, we shall find that they form one or more detached regions, that is to say, regions which have no boundary in common, though possibly they may meet at a point or points. These regions will be surrounded by and surround other regions composed of blue and yellow districts, the two sets of regions making up the whole surface. It will readily be seen that we can switch the colours of the districts in one or more of the red and green regions without doing so in any others, and the map will still be properly coloured. The same remarks apply to the regions composed of districts of any other pair of colours."

Furthermore, KÖNIG [*1936*: 170] pointed out another similarity between his work and KEMPE's: the use of bipartite graphs. Namely, he individuated in KEMPE's article a structure analogous to a bipartite graph in this passage [*1879*: *200*]: "If, excluding island and peninsula districts from the computation, every district is in contact with an even number of others along every cycle formed by its boundaries, then three colours will suffice to colour the map".

So far, we have only seen KÖNIG's proof of theorem C for finite graphs. Then KÖNIG *tried* to generalize his results onto graphs that were infinite (in the sense that they contained an infinite number of vertices), but with the restriction (as in the case of the BERNSTEIN theorem) that the number of edges going to each vertex was finite. Still, in addition to the proof of the theorem for *any* graph of degree 2 given in 1914, here he could only present a theorem that assured the extension of this result to (any) graphs whose degree was a *power of* 2. This was given in the following theorem K:

> [K] *If every bipartite graph of degree m and every bipartite graph of degree n have a factor of degree* 1, *then a bipartite graph of degree mn also has such a factor.*

In order to prove this, KÖNIG first obtained from a graph of degree mn a graph of degree n by substituting each of its vertices by m vertices. Second, he again obtained a graph of degree mn by joining the initial m vertices to a new vertex P. The graph of degree n was still bipartite because each of its cycles corresponded to a cycle of G with the same edges. Therefore, according to the premises of the theorem, the graph had a factor of degree 1 that became of degree m when the graph became of degree mn because, during the rearrangements, m edges remained attached to every vertex. This factor of degree m had in its turn a factor of degree 1 because it was bipartite (each of its cycles was the same as in the original bipartite graph), and hence KÖNIG could apply theorem A. As the factor of degree m was a part of the graph of degree mn, its factor of degree 1 was also a factor of degree 1 of the graph of degree mn.

In practice, the factor of degree 1 of the graph of degree mn was formed by obtaining the graph of degree n from the graph of degree mn and then by returning to it. In this way KÖNIG individuated a factor of degree m (inside the graph of degree mn) that had a factor of degree 1 which was also a factor of degree 1 of the graph of degree mn.

Finally, KÖNIG introduced a theorem I, equivalent to theorem A, which was a refinement of the BERNSTEIN theorem. It affirmed that:

> [I] *If two sets have an invertible* $(1-n)$ *correspondence to each other, then there is also an invertible* $(1-1)$ *correspondence between them that only joins elements of the sets with each other that were already joined to each other by the* $(1-n)$ *correspondence.*"

As KÖNIG noticed, this theorem was equivalent to theorem A because both theorems expressed the same situation, the former in set theory, the latter in graph theory. Namely, the premises of theorem A could be represented as an invertible $(1-n)$ correspondence between two sets M and N, and the factor of degree 1 given in the theorem A connected vertices already connected in the initial graph, hence the factor of degree 1 could be considered as represented by an invertible $(1-1)$ correspondence between M and N. On the other hand, if theorem I held, its sets could be represented as a whole graph and the elements connected by the last correspondence could be considered as a factor of degree 1[12].

KÖNIG also observed that theorem I was a refinement of BERNSTEIN's theorem, since, in addition to what was expressed by the simple BERNSTEIN theorem, it required the final correspondence to join elements already joined by the initial correspondence.

5. The final solution

Finally, KÖNIG managed to prove the generalized-refined BERNSTEIN theorem helped by STEPHEN (ISTVÁN) VALKÓ [*1926*]. KÖNIG and VALKÓ used two results

---

[12] In a footnote [*1916*: *464*], KÖNIG remarked that the axiom of choice was necessary in order to prove this theorem if the graph split into infinite parts. In this paper, he also considered it impossible to prove the BERNSTEIN theorem in the form: "If $2M = 2N$, then $M = N$" without the axiom of choice. Later [*1926b*: *133* and *215*], he admitted that he had been wrong since M. SIERPINSKI [*1922*] had managed to do this.

that KÖNIG had already obtained before: theorem A for finite graphs and theorem G. In this way they had "only" to prove the theorem for denumerable graphs.

They defined the concept of the "S-system" and then, thanks to the infinity lemma, they could build a sequence $S_1, S_2, \ldots$ of S-systems that they proved to be a factor of degree 1 of a graph G.

An S-system was defined as a finite subset S of the edges of the graph G that fulfilled the following requirements:

1) two edges of S never had a common endpoint;

2) at least one of the two endpoints of any edge of S belonged to the following $2n$ points:

$$A_1, B_1, A_2, B_2, \ldots, A_n, B_n;$$

3) each of these $2n$ points was an endpoint of an edge of S.

In other words, an S-system was a chain that joined $A_1, B_1, A_2, B_2, \ldots, A_n, B_n$ with each other or also with other vertices of G by *only one edge*. The number of its edges was at least n (in the case where $A_1, B_1, A_2, B_2, \ldots, A_n, B_n$ were joined exclusively to each other) and at most $2n$ (in the case where every $A_1, B_1, A_2, B_2, \ldots, A_n, B_n$ was joined with a vertex of G which did not belong to this sequence). In order to distinguish the S-systems from each other, KÖNIG and VALKÓ used the last index of the related sequence $A_1, B_1, A_2, B_2, \ldots, A_n, B_n$: that is why they called these systems "systems S_n" or, equivalently, "S-systems belonging to n".

The main lemma (that is to say, the infinity lemma) for the proof of BERNSTEIN's theorem affirmed that:

[Infinity Lemma] *If a system S_n is contained in infinitely many S-systems, then there is a system S_{n+1} such that*: a) *S_n is contained as a part in S_{n+1}*; b) *S_{n+1} is contained as a part in infinitely many S-systems.*

In order to prove this, KÖNIG and VALKÓ introduced these other lemmas whose proofs were trivial:

i) "For each n there is an S-system S_n."

ii) "There are infinitely many S-systems."

iii) "Given a point A_i (or B_i) of a graph G and a set of infinitely many S-systems, there are among them infinitely many S-systems containing an edge which goes to that point."

Lemma i) served to prove lemma ii), which served to prove lemma iii), and this last was to prove the main lemma.

Proof of the main lemma: KÖNIG and VALKÓ explained that, among the infinitely many S-systems containing S_n, there were, according to lemma iii), infinitely many with an edge going to A_{n+1}. The edges that could go to A_{n+1} were totally n (that is to say, a finite number). Therefore, at least one of them — call it $A_{n+1}B_\beta$ — was contained in infinitely many of these S_n systems (if there were more than one, the vertex with the smallest β was *chosen*[13]). Again using lemma iii), KÖNIG and VALKÓ concluded that, among the infinitely many S_n systems containing $A_{n+1}B_\beta$, infinitely many also contained $A_\alpha B_{n+1}$ (if there were more than one, the vertex with the smallest α was *chosen*). Hence, if both edges $A_{n+1}B_\beta$ and $A_\alpha B_{n+1}$ were added to a system S_n, a system S_{n+1} was obtained with the properties required by the theorem (it was also possible that there was no need for adding two or even one edge: this was the case where $\alpha = n + 1$ or $\beta = n + 1$ or $\alpha = \beta = n + 1$; it is clear that in the last case $S_{n+1} = S_n$).

The final proof of BERNSTEIN's theorem: according to the main lemma, KÖNIG and VALKÓ could build the sequence:

$$S_1,\ S_2, \ldots, S_n, \ldots$$

by starting with the empty set S_0 that could be considered as a subset of any S-system; therefore, they could apply the main lemma and obtain an S_1 system,... and so on, where S_m contained S_n if $m \geqq n$.

Then they let G_1 be the set of the edges contained in any S_i. This was the required factor of degree 1 of the graph G. Namely, any vertex A_i or B_i of G had an edge of G_1 going to it (by definition of G_1 and hence of the S_1, S_2, $S_3, \ldots$). And vice versa: any vertex A_i or B_i could be the vertex of only one edge of G, because, if there had been two such edges, say A_iB_j and A_iB_k ($j \neq k$), the first one would have been contained in S_α and the other in S_β. If $\alpha \leqq \beta$, then both edges would have been contained in S_β, against the definition of the S-system itself.

The winning strategy had been the opportunity of individuating a path (the factor of degree 1) afforded by the main lemma: this possibility is exactly the meaning of the infinity lemma explicitly stated by KÖNIG in the same year [*1926*: *120*] in a set-theoretical form and one year later in a graph-theoretical form [*1927*: *121*]. As KÖNIG himself acknowledged [*1926*: *132*], the idea of the procedure had come from S. VALKÓ.

6. The infinity lemma

A second presentation of the proof of the "refined" BERNSTEIN theorem appeared in the same year [*1926*], written by KÖNIG alone. He stated (and proved) a theorem which was even more general than the refined BERNSTEIN theorem and contained a specific case equivalent to this latter: we will call it the "main theorem". He used set-theoretical, no longer graph-theoretical terminology. At the end of the article, he explained his choice by affirming [*1926*: *132*] that he had avoided the use of graphs in order to assure that his results required no special geometrical intuition (which would be automatically

[13] It should be noted that in this article DÉNES KÖNIG did not mention the axiom of choice.

supposed in a graph-theoretical treatment). Later [*1936*: *212*], he added further grounds: he had wanted to adhere to the style of *Fundamenta Mathematicae*, the journal to which he submitted his work.

In order to introduce the main theorem, KÖNIG built the set $P = MN$ as the set of non-ordered pairs $(a, b)_i$ $(i = 1, 2, \ldots)$ of elements of the sets M and N (where M and N were disjoint), and then introduced a (not specified) set G such that each of its elements corresponded to only one element of P. By definition he proposed that, if an element g of G corresponded to a (a, b) of P, then it corresponded to the elements a and b of $M + N$. He also added that this correspondence was supposed to be finite and that, by treating elements of a subset of G, the same correspondence to the elements of $M + N$ or of P was supposed for them as was given for these elements considered as elements of G.

The main theorem affirmed that:

> [Main Theorem] *If, for a natural n there are at most n elements of a set G corresponding to every element of $M + N$, then G can be split into n G_i $(i = 1, 2, \ldots)$:*
>
> $$G = G_1 + G_2 + \cdots + G_n$$
>
> *such that in every G_i there is at most one element corresponding to each element of*
>
> $M + N$ *(some of them can be empty)*[14].

The main theorem encompassed[15] the case in which *exactly n* elements of G corresponded to the same element of $M + N$. In this case, the conclusion was that it was possible to split G as above so that in G_i $(i = 1, 2, \ldots)$ *exactly one* element corresponded to the same element of $M + N$. In other words:

> *If exactly n elements of G correspond to every element of $M + N$, then G has a subset G_1 such that one and only one element in G_1 corresponds to every element of $M + N$.*

The main theorem had the refined BERNSTEIN theorem as a consequence. As KÖNIG himself explained, given the invertible $(1 - n)$ correspondence between M and N and taking the set of pairs $(a, b)_\alpha$ of P as G such that, for every element of $M + N$ (that is to say, both for every a and for every b), G contained $(a, b)_\alpha$ with the number $\alpha = 1, 2, \ldots, n$ times equal to the multiplicity of the

---

[14] Where it is not explicity affirmed, here and in the rest of the paper, "splitting" requires that the parts are disjoint pairwise.

[15] It is interesting to specify the difference between this theorem and the refined BERNSTEIN theorem when expressed in graph-theoretical terms. The main theorem does not require the graphs to be regular (it requires that *at most* — and not exactly — n elements of G correspond to the same element of $M + N$), while the latter theorem does. Therefore, the main theorem is more general than the refined BERNSTEIN theorem.

correspondence between a and b (if there were no correspondence between them, no $(a, b)_\alpha$ would be put into G). In this case, the main theorem affirmed that in G there was a subset G_1 that contained only one $(a,b)_\alpha$ for every element of $M + N$. Since $(a, b)_\alpha$ was an element of P, the theorem affirmed that, given an invertible $(1 - n)$ correspondence between M and N, it was possible to find an invertible $(1 - 1)$ correspondence between the same elements of M and N.

First KÖNIG proved the main theorem for the case of G *finite*. The core of the proof was theorem D:

> [D] *Let $(a, b)_\alpha$ be an element of the set G and let G' be the set formed by all the other elements of G. Let the number of elements of G corresponding to the same element of the set $M + N$ be at most equal to n. If it is possible to split G' into a sum $G' = \sum_{i=1}^{n} G'_i$, (where some G'_i can be empty) so that in every G'_i $(i = 1, 2, \ldots)$ there is at most one element corresponding to the same element of $M + N$, then such a splitting also exists for G with the same number n.*

Theorem D was a lemma for the main theorem. As KÖNIG noticed, for any natural number n, it was apparent that the hypotheses of the main theorem were satisfied for a set containing a number of elements $< n$, because it was easy to obtain the splitting in such a case where each part was either an empty set or a singleton. At this point, theorem D proved the step of induction, that is to say, that if the main theorem held for G containing $n - 1$ elements, then the main theorem held for G with n elements.

Hence, KÖNIG had to prove theorem D. For this purpose, he had to consider two possibilities:

1) One of the G'_1 — say G'_k — was such that none of its elements corresponded to either an a or a b.
2) Every element of G corresponds to an (a,b).

1) The first case was very simple. He added the element $(a,b)_\alpha$ to G'_k. The splitting required was:

$$G = G_1 + G_2 + \cdots + G_n$$

where $G_k = G'_k + \{(a, b)_\alpha\}$ and $G_i = G'_i$ $(i \neq k)$.
2) On the contrary, the case in which every element of G corresponded to an (a,b) was more complicated. When $(a,b)_\alpha$ was left out of G, at most $n - 1$ elements of G' corresponded to the element of $M + N$ to which $(a, b)_\alpha$ corresponded (otherwise $n + 1$ elements of G would correspond to it — that is, the n's of G and $(a, b)_\alpha$). Hence, there was a part G'_k of G' whose elements did not correspond to any a and a part of G'_l of G' whose elements did not correspond to any b. (G'_l surely contained an element corresponding to a, otherwise it would be the first case, hence $k \neq l$).

KÖNIG supposed $a \in M$ and $b \in N$. The trick consisted in reducing this case to the first one.

He wrote the sequence S:

$$(S) \qquad (a,b_1),\ (b_1,a_1),\ (a_1,b_2),\ (b_2,a_2),\ldots$$

where (a,b_1) was the element of G'_l corresponding to a; (b_1,a_1) the element of G'_k corresponding to b_1 (if one existed); (a_1,b_2) the element of G'_l corresponding to a_1; and so on[16]. As the elements of S belonged alternately to G'_l and G'_k, he could put

$$S_l = \{(a,b_1),\ (a_1,b_2),\ldots\}, \qquad S_k = \{(b_1,a_1),\ (b_2,a_2),\ldots\}$$

so that

$$S = S_l + S_k;$$

S_k and S_l were subsets of G'_k and G'_l respectively. He indicated the complementary sets of S_k in G'_k and of S_l in G'_l respectively as T_k and T_l, that is:

$$G'_k = S_k + T_k$$

$$G'_l = S_l + T_l.$$

Then KÖNIG defined:

$$G''_k = S_l + T_k$$

$$G''_l = S_k + T_l$$

so that:

$$G'_k + G'_l = G''_k + G''_l.$$

Hence, he was able to substitute, in the sum constituting G':

$$G' = G'_1 + \cdots + G'_k + \cdots + G'_l + \cdots + G'_n$$

G'_k and G'_l by G''_k and G''_l respectively, obtaining[17]:

$$G' = G'_1 + \cdots + G''_k + \cdots + G''_l + \cdots + G'_n.$$

At this point, he had to prove that the new sum:
2.1) satisfied the hypotheses of the theorem and
2.2) really belonged to the first case (that is to say, that G''_l did not have any element corresponding to a or b).

As for 2.1), KÖNIG had to prove that neither G''_k nor G''_l contained two elements corresponding to the same element x of $M + N$.

[16] KÖNIG [*1926b*: *117*] specified that in this case he omitted the indices α, β, ... and wrote (a,b_1), (b_1,a_1)... instead of $(a,b_1)_\alpha$, $(b_1,a_1)_\beta$... because in this proof *two* elements of G never occurred corresponding to the same element $(a_\rho,\ b_\sigma)$ of P.

[17] At the end of the paper [*1926b*: *128*], KÖNIG acknowledged that the source for arriving at the idea for this last composition had been KEMPE's work [*1879*]. In effect, KEMPE [*1879*: *198*] in his proof — that was then revealed to be wrong — used a procedure of splitting and re-composing in a new manner.

He had to consider six possibilities.
1. x was not an element of the sequence

$$a,\ b_1, a_1,\ b_2,\ a_2, \ldots$$

He proved it *by reductio ad absurdum*. Namely, if there were in G_k'' two elements g_1 and g_2 corresponding to x, then they could not both belong to S_l, hence they must belong to T_k. But this would imply that they belonged to G_k', in contrast with the definition of the sum G'. The same reasoning held if G_k'' was replaced by G_l''.
2. $x = a$. By definition, $(a,\ b_1)$ of G_l' corresponded to a, while no element of G_k' corresponded to a. Therefore, in the sum $G_k' + G_l'$ there was only one element corresponding to a. Since this sum was equal to $G_l'' + G_k''$, also in this last sum there was only one element corresponding to a. Hence, neither G_l'' nor G_k'' could contain two elements corresponding to a.
3. $x = a_\rho$ (where a_ρ was not the last element of the sequence $a,\ b_1,\ a_1,\ b_2,\ a_2, \ldots$). In this case, only two elements, that is $(b_\rho,\ a_\rho)$ — belonging to S_k and hence to G_l'' —, and $(a_\rho,\ b_{\rho-1})$ — belonging to S_l and hence to G_k'' —, corresponded to a_ρ. Therefore, in $G_l'' + G_k''$ another element corresponding to a_ρ could not exist, otherwise there would be three elements in $G_l'' + G_k''$ corresponding to a_ρ, and hence three elements in $G_k' + G_l'$ corresponding to a_ρ, while every G_k' and G_l' by definition could not contain more then one element corresponding to a_ρ.
4. $x = b_\rho$ (where b_ρ was not the last element of the sequence $a,\ b_1,\ a_1,\ b_2,\ a_2, \ldots$). For this case the same considerations held as for 3), where the two elements at issue were $(a_{\rho-1},\ b_\rho)$ and $(b_\rho,\ a_\rho)$. If $\rho = 1$, König put $a_0 = a$.
5. $x = a_\delta$ (where a_δ *was* the last element of the sequence $a,\ b_1,\ a_1,\ b_2,\ a_2, \ldots$). This case meant that S had a last element $(b_\delta,\ a_\delta)$, and no element of G_l' corresponded to a_δ, otherwise S would continue after $(b_\delta,\ a_\delta)$. Consequently, König observed, $G_k' + G_l'$ could not contain more than one element corresponding to a_δ. The same held for $G_l'' + G_k''$, which was equal to this sum. Hence, neither G_l'' nor G_k'' could contain two elements corresponding to a_δ.
6. $x = b_\delta$ (where b *was* the last element of the sequence $a,\ b_1,\ a_1,\ b_2,\ a_2, \ldots$). The reasoning was the same as in 5).

So far, König proved that the splitting

$$G' = G_1' + \cdots + G_k'' + \cdots + G_l'' + \cdots + G_n'$$

satisfied the hypothesis of theorem D.

Then, he had to prove 2.2), that is to say, that G_l'' contained no element corresponding to either a or b.

G_l'' did not contain elements corresponding to a. Namely, König has shown in case 2 above that $G_l'' + G_k''$ contained only *one* element corresponding to a, that is, $(a,\ b_1)$. This belonged to S_l, hence to G_k'', and not to G_l''.

G_l'' did not contain elements corresponding to b. Namely, König noticed that the only elements of N to which an element of S corresponded were b_1, $b_2, \ldots$. But b could not belong to this sequence, because to any b_i of this sequence an element (a_{i-1}, b_i) belonging to S_l corresponded, hence to G_l' (if $i = 1$, put $a_0 = a$), and no element of G_l' corresponded to b. Therefore, no element of S, and *a fortiori* no element of S_k, belonged to b. The same held for T_l, which was a subset of G_l', and finally for G_l'' as $G_l' = S_k + T_l$.

After this proof of theorem D, KÖNIG had to prove the main theorem for *G denumerable*. He needed theorem E:

> [E] *Let E_1, E_2, $E_3, \ldots$ be a denumerable sequence of finite, non empty sets, and R a relationship such that to every element x_{n+1} of every set E_{n+1} at least one element x_n of E_n corresponds joined with x_{n+1} by R: write x_nRx_{n+1}, where $n = 1, 2, \ldots$). Then it is possible to choose from every E_n an element a_n such that a_nRa_{n+1} holds for the infinite sequence S $a_1, a_2, a_3, \ldots$.*

This was the infinity lemma in a set-theoretical form.

In order to build the sequence S, KÖNIG introduced the notation "S_n" for a sequence S of n terms $a_1, a_2, a_3, \ldots, a_n$ where a_1 belonged to E_1, a_2 to E_2, and so on. He called its first m terms $(m < n)$ "a segment" of S_n. Then he introduced the following statement:

"Let a set I of infinitely many sequences S be such that a certain sequence S_n

$$S_n^0 = a_1, a_2, a_3, \ldots, a_n$$

is the segment of every element of I, then there is a sequence S_{n+1} — call it S_{n+1}^0 — such that: S_n^0 is a segment of S_{n+1}^0; S_{n+1}^0 is the segment of infinitely many S sequences."

Once this statement was proved, he could start with the empty set, obtain S_1^0 (it is common to all the possible S sequences) and then write a sequence

$$S_1^0, S_2^0, \ldots$$

where $S_1^0 = (a_1)$, $S_2^0 = (a_1, a_2), \ldots$.

In this way he obtained the sequence $a_1, a_2, a_3, \ldots$ that satisfied the requirements of the infinity lemma.

Hence KÖNIG had to prove the above statement. He reasoned as follows.

The number of S_i's for which $i \leq n$ held was finite. Therefore, there was an *infinite* subset I' of I such that, if S_k belonged to I', then $k > n$. Every element of I' possessed a segment S_{n+1} having the form

$$S_{n+1} = (a_1, a_2, \ldots, a_n, x_{n+1})$$

where x_{n+1} belonged to E_{n+1}. Since E_{n+1} was finite, also the number of x_{n+1}'s and hence of S_{n+1}^0's was *finite*. Therefore, at least one of them, and only one of them, call it

$$S_{n+1}^0 = (a_1, a_2, \ldots, a_n, a_{n+1})$$

was the segment of infinitely many elements of I'.

Once theorem E had been proved, the main theorem for G denumerable could be proved.

He supposed all elements which did not correspond to any element of G, where G contained elements like $(a_\rho, b_\rho)_i$ and $i \leqq n$, to be eliminated from $M + N$.

König defined a set $G^{(k)}$ as the subset of G formed by all elements (a_ρ, b_σ) where at least one of the indices ρ and σ was $\leqq k$. Since the elements of $G^{(k)}$ were at least $2k$ and at most $2nk$, it was clear that $G^{(k)}$ itself was *finite* and the set of all $G^{(k)}$'s was *infinite* ($G^{(k)}$ could be identical to $G^{(k+1)}, G^{(k+2)}, \ldots, G^{(nk)}$ but surely not to $G^{(nk+1)}, G^{(nk+2)}, \ldots$). Hence, he could apply the main theorem of the preceding section for the finite case to $G^{(k)}$ and obtain the $G^{(k)}$ splitting:

$$G^{(k)} = \sum_{i=1}^{n} G_i^{(k)},$$

such that at most one element of each $G_i^{(k)}$ corresponded to the same element in $M + N$.

König called this splitting $\Delta^{(k)}$ and then introduced the relationship $\Delta_1 R \Delta_2$ to say that Δ_1 was a segment of Δ_2, where $G^{(k)}$ — and hence a $\Delta^{(k)}$ — was called "a segment" of $G^{(l)}$ if for every i ($i = 1, 2, \ldots$) $G_i^{(k)}$ was subset of $G_l^{(k)}$ ($i \leqq l$). Furthermore, he called E_i the set of the splittings $\Delta^{(i)}$: any E_i was non empty and contained a finite number of elements. Hence, he could apply theorem E to the infinite sequence $E_1, E_2, \ldots$ and obtain the infinite sequence of splittings[18]

$$\Delta_0^{(1)}, \Delta_0^{(2)}, \ldots$$

where every $\Delta_0^{(i)}$ was a segment of $\Delta_0^{(i+1)}$ for every i ($i = 1, 2, \ldots$). Due to the definition of a segment, $\Delta_0^{(i)}$ was also a segment of the $\Delta_0^{(i+n)}$ splitting for every natural number n. A $\Delta_0^{(i)}$ could be represented as:

$$G^{(i)} = G_1^{(i)} + G_2^{(i)} + \cdots + G_n^{(i)}.$$

At this point, König defined G_α as the set of the elements contained in at least one of the sets $G_\alpha^{(1)}, G_\alpha^{(2)}, \ldots$. Every element g of G, being contained in a set $G^{(i)}$, was contained in $G_k^{(i)}$ and consequently in G_k (where k depended only on g and not on i). This allowed him to write the following splitting:

$$G = G_1 + G_2 + \ldots.$$

This was the required splitting. Namely, it was impossible for a G_α to contain two elements — say x belonging to $G^{(i)}$ and hence to $G_\alpha^{(i)}$, and y belonging to $G^{(j)}$ and hence to $G_\alpha^{(j)}$ with $i \leqq j$ — corresponding to the same element of $M + N$. By *reductio ad absurdum*, König noticed that if they were to exist, then both of them would belong to $G_\alpha^{(j)}$, because $G_\alpha^{(i)}$ was a segment of $G_\alpha^{(j)}$. But this was an absurdity as, by definition, every $G_\alpha^{(j)}$ could only contain one element corresponding to the same element of $M + N$.

Finally, König could reconsider the set-theoretical proof of Bernstein's theorem begun in 1923 and finish it, thanks to the new theorem (and hence to

[18] This sequence is determined since M and N are already ordered sets.

the infinity lemma)[19]. He presented the definition of a *connected set*. This required the chain notion for elements of sets: if an element of the set G corresponded to every element of the set P occurring in the finite sequence

$$(a, x_1),\ (x_1,\ x_2),\ (x_2, x_3), \ldots, (x_{\rho-1},\ x_\rho),\ (x_\rho, b)$$

(where any two consecutive elements of the sequence

$$a,\ x_1,\ x_2, \ldots, x_\rho,\ b$$

alternately belonged to M and to N), then a and b of $M + N$ were joined by a chain of length ρ (every element of $M + N$ was joined with itself by a chain of length 0). A set G was connected if, a and b being any elements of $M + N$ to which elements of G corresponded, a and b were linked through a chain.

He then:

1) proved that a connected set G was either finite or denumerable;

2) showed the possibility of splitting G so that each of its parts was connected while the sum of any two of its parts was never connected;

3) applied the main theorem and obtained the final result.

We follow his steps.

1) A connected set G is either finite or denumerable. In order to prove this theorem, it was sufficient for him to prove that the subset $M^* + N^*$ of $M + N$ to which the elements of G corresponded was itself either finite or denumerable, because in this case G was also such. The subset $M^* + N^*$ was the sum of the elements of $M + N$ joined to a by a chain of length 1, of the elements of $M + N$ joined to a by a chain of length 2, and so on. The number of the elements of $M + N$ joined to a by a chain of length 1 was finite; the number of the elements of $M + N$ joined to a by a chain of length 2 was the number of the elements of $M + N$ joined by a chain of length 1 to an element joined by a chain of length 1 to a, therefore, it was a finite number; and so on. Hence, their sum was either finite or denumerable.

2) At this point, KÖNIG could define a relationship R between the elements x and y of G, where x corresponded to $(a_1,\ b_1)$ and y to $(a_2,\ b_2)$, and $a_1\ a_2$ were joined by a chain. Through this relationship, he could divide G into classes of equivalence, obtaining a splitting of $G = \sum_{(\alpha)} G^{(\alpha)}$ where $\alpha = 1, 2, \ldots$. Namely, the relationship was reflexive (an element was joined to itself by a chain of length 0), symmetric and transitive: hence, it defined a partition of G. Each of these parts $G^{(\alpha)}$ was connected since a and b were two elements of $M + N$ to which x and y of $G^{(\alpha)}$ corresponded respectively. As xRy, a and b were joined by a chain. Yet, given two different $G^{(\alpha)}$ and $G^{(\beta)}$, $G^{(\alpha)} + G^{(\beta)}$ were not connected, because otherwise for two elements x and y (x belonging to $G^{(\alpha)}$ and y belonging to $G^{(\beta)}$) xRy would hold, which was impossible since they belonged to two different classes of equivalence.

[19] In [*1926b*: *133*], he observed that in this part of his paper "The axiom of choice plays an important role".

3) The statement just proved states that the $G^{(\alpha)}$'s, being connected, are either finite or denumerable. Therefore, the main theorem of the preceding section was applicable, and, for every $G^{(\alpha)}$, KÖNIG could obtain the splitting:

$$G^{(\alpha)} = G_1^{(\alpha)} + G_2^{(\alpha)} + \cdots + G_n^{(\alpha)},$$

where every $G^{(\alpha)}$ contained at most one element corresponding to the same element of $M + N$.

By putting $G_i = \sum_{(\alpha)} G_i^{(\alpha)}$ KÖNIG obtained a splitting of G:

$$G = G_1 + G_2 + \dots$$

with this same characteristic. Namely, if two elements x and y of G_i were to correspond to the same element of $M + N$, one would appear in $G_i^{(\alpha)}$ hence in $G^{(\alpha)}$, the other in $G_i^{(\beta)}$ hence in $G^{(\beta)}$, where $\alpha \neq \beta$ since $G_i^{(\alpha)}$ could not contain two elements corresponding to the same element of $M + N$. This would be in contrast with the fact that if x and y were to correspond to the same element of $M + N$, then xRy, and hence x and y, would belong to the same $G^{(\alpha)}$.

7. The infinity lemma identified

In 1927 KÖNIG presented his article "Über eine Schlussweise aus dem Endlichen ins Unendliche", which represents the *clou* of our story of the infinity lemma, since it was entirely devoted to the lemma. Here he listed some different versions of the lemma:

1) the set-theoretical version — presented in [*1926*] as "lemma E":

> [Lemma E] *Let $E_1, E_2, E_3, \dots$ be a denumerable sequence of finite, non empty sets, and R a relationship such that to each element x_{n+1} of E_{n+1} at least one element x_n of E_n linked to x_{n+1} through the relationship R corresponds (written: $x_n R x_{n+1}$ where $n = 1, 2, \dots$). Then it is possible to choose from each E_n an element a_n such that $a_n R a_{n+1} (n = 1, 2, \dots)$ holds for the infinite sequence S $a_1, a_2, a_3, \dots$*

And two graph-theoretical versions:

> 2) *If a denumerable point set of an infinite graph G splits into denumerably many sets $E_1, E_2, \dots$ which are finite and not-empty and such that every point of E_{n+1} $(n = 1, 2, \dots)$ is joined with a point of E_n by an edge, then there is, in the graph, an infinite path $a_1, a_2, \dots$ that contains from each set E_n a point a_n (also here it is not necessary that the sets E_n's be disjoint).*

> 3) *If every point of a connected infinite graph has only finitely many edges going to it, then the graph contains an infinite path.*

By giving the graph-theoretical versions, KÖNIG specified the necessity of both conditions (the finite number of edges and the connection of the graph) for the theorem to be true. Namely, if the number of edges was not finite, then the graph did not need to have an infinite path: it could be a star with infinitely many rays. As an example of the necessity of the connection of the graph, he quoted graphs consisting of non-connected polygons with two, three, four, . . . angles.

At this point, he had realized the importance and fruitfulness of the lemma and wanted to stress it. He therefore presented four applications of it in different fields:

1) a generalisation of BOREL's theorem (that is to say, DE LA VALLÉE POUSSIN's theorem);

2) a particular case of the problem of the four colours;

3) a proof of the fact that, if the hypothesis that mankind will never become extinct is assumed, then a person exists today who is the ancestor of an infinite sequence of descendants;

4) a proof that in chess there is a number N depending on the number q, so that if White plays starting from this position q, then White can win in less than N steps.

Here we consider only the first two applications.

As for the generalization of BOREL's theorem, its statement is:

Let E be a closed subset of the interval $(0, 1)$ *and let I be a family of intervals which have the property that every point of E is contained in one of these intervals. Then there is a natural number n such that if* $(0, 1)$ *is divided into* 2^n *subintervals, those subintervals which contain a point of E are contained in an interval of I.*

The proof given by KÖNIG was by *reductio ad absurdum.* If the theorem were not true, then there would be at least one interval $(m/2^n, m + 1/2^n)$ — where $m = 0, m = 1, \dots$ or $m = 2^n - 1$ — which contained a point of E and which was not contained in any interval of I. Let E_n be the set of *these* intervals. For every n, E_n was finite and not empty. KÖNIG defined $x_n R x_{n+1}$, where x_n was an element of E_n and x_{n+1} an element of E_{n+1}, if x_{n+1} came about by bisecting x_n. The E_n sets and the relationship R satisfied the conditions of the infinity lemma. According to this latter, an infinite sequence $a_1, a_2, \dots,$ of intervals existed such that each of them had come about by bisection from the preceding one, contained a point of E and was not contained in any interval of I. Therefore, also the common point α of the intervals $a_1, a_2, \dots$ was not contained in any intervals of I. This was impossible since, because E was closed, α had to belong to E.[20]

As for the problem of the four colours, he proved the statement[21]:

If a two-sided surface is divided by an infinite graph into denumerably many regions, then it is possible to assign each region a colour among

[20] In this regard, KÖNIG underlined that, thanks to his lemma, he could avoid the use of the BOLZANO-WEIERSTRAß theorem in this proof.

[21] In a footnote [*1927*: *124*], he specified that the English mathematician L. H. THOMAS had reached the same result a short time before.

p different colours[22] *in such a way that two adjacent regions always have different colours.*

His reasoning was as follows.

Let $L_1, L_2, \ldots$ be the regions of the infinite map. The limit-edges of $L_1, L_2, \ldots$ and L_n make up a graph G_n that divides the surface into $n+1$ regions $L_1, L_2, \ldots L_n, M_n$, where M_n indicates $L_{n+1}, L_{n+2}, \ldots$ The map K_n determined by G_n is finite, hence it can be coloured by p colours in finitely many ways. The different ways of colouring $F_n, F'_n, F''_n, \ldots$ of K_n by p colours formes a finite, not empty set E_n. From every colouring F_{n+1} of K_{n+1} a colouring F_n of K_n results, according to which every $L_1, L_2, \ldots, L_n$ contains the same colour as F_{n+1}.

KÖNIG indicated this situation as $F_n R F_{n+1}$. (It was useless to consider the colours of M_n). By the infinity lemma (applicable as E_n and R satisfied its requirements) he obtained the infinite sequence of colourings $F_1, F_2, \ldots,$ that respectively coloured $K_1, K_2, \ldots,$ in such a way that every region L_i contained in all these colourings F_k the same colour. This meant that the above sequence of colourings attributed to every region $L_1, L_2, \ldots,$ a certain colour among the possible p's.

Name, definitions, applications: the infinity lemma had gained its autonomy.

We briefly recapitulate the whole history. In 1926 DÉNES KÖNIG had managed to prove the refined BERNSTEIN theorem, helped by STEPHAN VALKÓ. He himself had previously proved it for some cases only. In the same year, he isolated a part of the proof as a lemma apart. Later, he gave it different expressions, but the core was always the same. Given an infinite sequence of finite non-empty sets such that every element of one of them was joined by a certain relationship R at least with an element of the following set, KÖNIG formed a sequence containing only one element from every set, joined with the successive element by the relationship R.

Hence, to KÖNIG's great credit, he had been able to isolate the lemma and to understand its wide applicability.

As for the influence of the others, we have individuated as "sources" JULIUS KÖNIG, PETERSEN and KEMPE. The last two had a mere technical influence: PETERSEN for graph terminology and KEMPE for a procedure of splitting and composition. The relationship with his father was obviously deeper and more complex. They showed two common aspects in mathematical research. DÉNES shared his father's opinion about the axiom of choice (both when he was against it and when he accepted it). His researches into equivalence theorems began due to his father's interest into them. The starting point for DÉNES had been his father's simplification of the proof of the CANTOR–BERNSTEIN–SCHRÖDER theorem, simplification given in order to make it constructively acceptable. Nevertheless, the two KÖNIGS' difference in attitudes was also apparent in the history of the infinity lemma. As G. H. MOORE observed [*1982, 119*], JULIUS'

---

[22] KÖNIG did not specify anything about this number p.

attitude was philosophical while DÉNES' was more mathematical. We can add that, in the case of the infinity lemma, DÉNES' scant philosophical attitude was revealed by his not very deep analysis of the role of the axiom of choice in its proofs. The lemma was used inside an equivalence theorem, and DÉNES had initially applied himself to prove equivalence theorems without the well-ordering principle which is equivalent to the axiom of choice. Still, he did not seem to have noticed this linking and in his 1914 article, although he had promised at its beginning not to use the well-ordering principle[23], he then specified that the axiom of choice was necessary in his proofs. In his following articles, he only gave some hints here and there about the presence of the axiom in his proofs (reported in the footnotes) and in [*1936*: *82*], he simply affirmed that "In most applications of the infinity lemma, however, the axiom can be avoided".

Finally, we can find a trace of his mathematical rather than philosophical attitude in his appreciation of evidence. This appreciation was shared by his father, but, while the latter wrote a foundational book based upon it, DÉNES regarded it more as a matter of taste, the taste for intuitive, visual tools in proofs, that led him to devote himself to graph theory.

Bibliography

BERNSTEIN, F. *1905* "Untersuchungen aus der Mengenlehre". *Mathematische Annalen* **61**, 117–160.

CANTOR, G. *1895* "Beiträge zur Begründung der transfiniten Mengenlehre". *Mathematische Annalen* **46**, 481–512; also in

CANTOR, G. *1962, Gesammelte Abhandlungen.* Olms, Hildesheim, 282–356.

DUMMETT, M. *1977 Elements of Intuitionism.* Clarendon Press, Oxford.

GALLAI, T. *1978* "The Life and Scientific Work of Dénes König (1984–1944)". *Linear Algebra and its Applications* **21**, 189–205.

KEMPE, A. B. *1879* "On the Geographical Problem of the Four Colours". *American Journal of Mathematics* **2**, 193–200.

KÖNIG, D. *1899* "Elementary Discussion of Two Maximum-Minimum Problems". (In Hungarian) *Mathematikai ès Fizikai Lapok* **8**, 271–274.

KÖNIG, D. *1902 Mathematical Recreations I*, (In Hungarian) Budapest.

KÖNIG, D. *1905 Mathematical Recreations II*, (In Hungarian) Budapest.

KÖNIG, D. *1907* "Elementary Discussion of Rotations and Finite Rotation Groups of a Space of Many Dimensions". (In Hungarian) *Mathematikai és Fizikai Lapok* **16**, 313–335, 373–390.

KÖNIG, D. *1908* "Zur Theorie der Mächtigkeiten". *Rendiconti del Circolo Matematico di Palermo* **26**, 339–342.

---

[23] He affirmed that "Direct proofs are interesting not only from the viewpoint of principles, but also because these direct proofs where the notion of the well-ordered set is not regarded let us see the nucleus of things better" [*1914*: *443*].

KÖNIG, D. *1911a* "Graphs on Two-sided Surfaces". (In Hungarian) *Mathematikai és Természettudományi Értesítö* **29**, 112–117.

KÖNIG, D. *1911b* "The Genus of Graphs". (In Hungarian) *Mathematikai és Természettudományi Értesítö* **29**, 345–350.

KÖNIG, D. *1913a* "On the One- and Two-sidedness of Manifolds of Higher Dimensions". (In Hungarian) *Mathematikai és Fizikai Lapok* **22**, 40–42.

KÖNIG, D. *1913b* "Zur Analysis Situs der Doppelmannigfaltigkeiten und der projektiven Räume", in *Proceedings of the Fifth Congress of Mathematics* II, 129–133.

KÖNIG, D. *1915* "Graphs and Determinants". (In Hungarian) *Mathematikai és Természettudományi Értesítö* **33**, 221–229.

KÖNIG, D. *1916* "Über Graphen und ihre Anwendung auf Determinantentheorie und Mengenlehre". *Mathematische Annalen* **77**, 453–465.

KÖNIG, D. *1918 Elements of Analysis Situs.* Budapest

KÖNIG, D. *1923* "Sur un problème de la théorie générale des ensembles et la théorie des graphes" (Communication faite le 7 avril 1914 au Congrès de Philosophie Mathématique à Paris). *Revue de Métaphysique et de Morale* **30**, 443–449.

KÖNIG, D. *1926* "Sur les correspondances multivoques des ensembles". *Fundamenta Mathematicae* **8**, 114–134.

KÖNIG, D. *1927* "Über eine Schlußweise aus dem Endlichen ins Unendliche". *Acta Litterarum ac Scientiarum*, Szeged **3**, 121–130.

KÖNIG, D. *1936 Theorie der endlichen und unendlichen Graphen.* Akademische Verlagsgesellschaft, Leipzig; Engl. transl. *Theory of Finite and Infinite Graphs.* Boston: Birkhäuser, 1990.

KÖNIG, D. and VALKÓ S. *1926* "Über mehrdeutige Abbildungen von Mengen". *Mathematische Annalen* **95**, 135–138.

KÖNIG, J. *1906* "Sur la théorie des esnsembles". *Comptes rendus de l'Académie de France*, 110–112.

KÖNIG, J. *1914 Neue Grundlagen der Logik, Arithmetik und Mengenlehre.* Von Veit, Leipzig.

MOORE, G. H. *1982 Zermelo's Axiom of Choice: its Origins, Development and Influence.* Springer, New York.

PETERSEN J. *1879* "Die Theorie der regulären Graphs". *Acta Mathematica* **15**, 193–220.

POINCARÉ, H. *1906* "Les mathématiques et la logique. II". *Revue de Métaphysique et de Morale* **15**, 17–34.

SIERPINSKI, M. *1922* "Sur l'égalité $2m = 2n$ pour les nombres cardinaux". *Fundamenta Mathematicae* **3**, 1.

TUTTE, W. T. *1990* "Commentary", in KÖNIG, D. *Theory of Finite and Infinite Graphs.* Boston: Birkhäuser, 1990, 1–43.

Dipartimento di Filosofia
Università degli Studi
V. Festa del Perdono 7
I-20122 Milano
Italy

(Received February 8, 1996)