

"DUALITIES" IN GRAPH THEORY AND IN THE RELATED FIELDS  
VIEWED FROM THE METATHEORETICAL STANDPOINT

M. Iri

Department of Mathematical Engineering and Instrumentation Physics  
Faculty of Engineering, University of Tokyo  
Hongo, Bunkyo-ku, Tokyo, Japan 113

Abstract. The importance is emphasized of distinguishing clearly among different kinds of concepts usually referred to as "duality". Those different kinds of dualities concentrate in the "dual graph", wherefrom confusion is sometimes given rise to. The importance is illustrated by "new" theorems and concepts which are derived by understanding correctly the difference of the concepts.

1. Introduction

The concept of duality, or of being dual, is familiar in graph theory, network theory, mathematical programming, etc. It is as old as projective geometry. It is so familiar and appears so simple that most textbooks and papers deal with it in a very informal manner. However, actually, there are basically two different kinds of concepts usually called "duality", of which one is the duality *in* a theory and the other is the duality *of* a theory. Although the two kinds of concepts are closely related to each other, the confusion between them sometimes leads to meaningless arguments.

In other words, there are many types of "duality theorems" in the theory of graphs, the theory of mathematical programming, etc., which are *proved within* the theory, whereas, sometimes, the theory, either in part or as a whole, has a symmetric structure to be called "duality". The former is a *theoretical* concept (i.e., a concept *in the theory*), whereas the latter is a *metatheoretical* concept (i.e., a concept outside of the theory, or *in the metatheory*). The duality theorems in mathematical programming, the Alexander-type duality theorems in algebraic

topology, Pontrjagin's duality theorem for topological groups are of the former kind; the widest known duality in projective geometry (or, in general, that in lattice theory) is of the latter kind.

Thus, from the metatheoretical standpoint, it is also interesting to investigate whether a duality theorem in a theory with a duality structure (such as graph theory) is self-dual or not, because one is apt to regard a duality theorem as a self-dual theorem on no sound basis.

In the following, the concepts of duality and duality theorems connected with graphs are re-viewed from the metatheoretical standpoint. Specifically, it is remarked that the nonexistence of a dual graph does not contradict the duality of graph theory but reinforces it, that there are  $4 = 2 \times 2$  kinds of triangular inequalities for distances and capacities on networks, and that the concept of reciprocity in linear systems theory such as electric network theory may be related to the coexistence of the metatheoretical duality and the theoretical.

All the technical materials on which the arguments in the present paper are based, are found in references [1], [2], [3] and [4]. For the first-order and higher-order predicate calculi, any introductory books and papers (e.g. [5], [6], [7]) on mathematical logic may be referred to.

## 2. Duality of a Theory and Duality Theorems in a Theory

Let us define that a "theory"  $T$  is a collection  $S$  of "sentences" written in the language of the first-order or the higher-order predicate calculus. Some of those sentences are axioms, and others are theorems. Usually — especially, for an application-oriented theory — a certain universe is fixed together with an interpretation, i.e., all the sentences of the theory are true on that universe under that interpretation.

In the following, however, we shall not describe a sentence of a theory in the form of a closed well-formed formula of the predicate calculus, but, more informally, in the ordinary plain language — the same language in which the text of this paper is written —, for we hope it will be far more legible and there is no fear of confusion or misunderstanding.

We assume that there is a subset  $D$  (possibly empty) of predicates as well as an involutive bijection

$$\psi : D \rightarrow D , \quad \psi \circ \psi = \text{id}_D . \quad (1)$$

It should be noted that most of mathematical expressions in the present paper, such as those in (1), are "metatheoretical", i.e., outside of the theory in question. The  $\psi$  induces a mapping of  $S$  to  $\psi(S)$  in a natural way, where the image  $\psi(s)$  of a sentence  $s$  in  $S$  is defined to be a sentence obtained from  $s$  by replacing each occurrence in  $s$  of every predicate  $p$  in  $D$  by  $\psi(p)$  (in particular, if no predicate in  $D$  occurs in  $s$ , we put  $\psi(s) = s$ ).

$\psi(S)$  does not in general coincide with  $S$ , but, if it does, i.e., if

$$\psi(S) = S, \quad (2)$$

then the theory  $T$  is said to have the dual structure with respect to  $(\psi, D)$ , or the duality  $(\psi, D)$  holds in  $T$ . (Any theory has the trivial dual structure with respect to  $(\psi, \emptyset)$ .) Even if the entire theory has not a dual structure, it may happen that part of it, i.e., not  $S$  but a subset  $S'$  of  $S$ , has one. In such a case we say that that part of the theory has the dual structure.

For any theory and any  $(\psi, D)$ , we may find the (maximal) part  $\hat{S}$  having the dual structure with respect to  $(\psi, D)$  by means of the equation

$$\hat{S} = S \cap \psi(S). \quad (3)$$

Since those predicates which do not occur in  $\hat{S}$  are inessential, only the subset  $\hat{D}$  obtained from  $D$  by removing them, as well as the restriction  $\hat{\psi}$  of  $\psi$  to  $\hat{D}$ , is of significance. Thus, that part of the theory which consists of sentences of  $\hat{S}$  has the dual structure with respect to  $(\hat{\psi}, \hat{D})$ . It is usual that we are interested mainly in those sentences of  $\hat{S}$  in which at least one predicate from  $\hat{D}$  occurs.

Two sentences  $s_1$  and  $s_2$  of  $\hat{S}$  are said to be *the dual* of each other if  $s_1 = \psi(s_2)$  (equivalently,  $s_2 = \psi(s_1)$ ), or, more generally, if  $s_1$  and  $\psi(s_2)$  is logically equivalent:  $s_1 \equiv \psi(s_2)$ . In particular, a sentence  $s$  such that  $s \equiv \psi(s)$  is said to be *self-dual*.

In the case of graph theory, we usually put

$$D = \{\text{arc, cycle, cocycle, circuit, cocircuit, tree, cotree, ...}\} \quad (4)^{*1}$$

and

$$\psi(\text{arc}) = \text{arc}, \quad \psi(\text{cycle}) = \text{cocycle}, \quad \text{etc.} \quad (5)$$

Thus, the sentences\*<sup>2</sup>

[A ring sum of elementary cycles is the union of elementary cycles.]

and

[A ring sum of elementary cocycles is the union of elementary cocycles.]

are the dual of each other, and the sentences

[A tree is a maximal cycle-free subset of arcs.]

and

[A cotree is a maximal cocycle-free subset of arcs.]

are also the dual of each other. The sentence

[A circuit and a cocircuit are orthogonal to each other under a suitably defined product operation.]

is self-dual, and so is the sentence

[The complement of a tree is a cotree, and vice versa.]

It should be remarked that no predicate related to vertices belongs to  $D$ , nor does the most important predicate of *being a graph*.

Matroid theory may be regarded as a theory which is developed with those sentences in graph theory as axioms which belong to the part of graph theory having the dual structure.

\*1: This is an informal description. To be more formal, we should describe, e.g., "the predicate of *being an arc*", "the predicate of *being a cycle*", etc.

Among a number of different terminologies and notations in the existing literature on graph theory, we shall follow those in [8] as far as possible.

\*2: Here, as well as in the following, the informal description of a sentence of a theory is written in the brackets.

In contrast with the duality *of* a theory, the conventional characterization of duality theorems *in* a theory is somewhat vague, although the mathematical contents of the theorems themselves are clear and rigorous. However, it may be said that, in most cases, a bilinear form which can be regarded as the inner product of two vectors from dual vector spaces or modules plays a fundamental role in the theorem. Thus, Pontrjagin's duality theorem in topological group theory [9] states the relation between the topological structures of two topological groups  $G_1$  and  $G_2$  which determine each other by means of a bilinear map  $\omega : G_1 \times G_2 \rightarrow C$  to the unit circle  $C$ . The Alexander-type duality theorems in algebraic topology [10] are based on the inner product between the family of chain groups and that of cochain groups of a complex. Likewise, the duality theorem in linear programming can be regarded as a theorem stating the relation between a subspace of a vector space and a subspace of the dual space, the two subspaces being defined by mutually contragredient constraint matrices [11].

Since duality theorems are objects *in* a theory, there is little to discuss about them from the metatheoretical viewpoint.

### 3. Problems related to the Concept of Dual Graphs

The concept of dual graphs is in an interesting situation from our standpoint. Among possible different definitions of dual graphs, we shall adopt here the following.

[Two graphs  $G_1$  and  $G_2$  are said to be dual to each other, or  $G_i$  is a dual of  $G_j$  ( $i \neq j$ ), if they have the same set of arcs (or, more exactly, their arc sets are in one-to-one correspondence with each other), and if every cycle in one of them corresponds to a cocycle in the other, and vice versa.] (6)

This sentence (definition) is obviously self-dual. It will probably because the definition of dual graphs is self-dual that the concept of dual graphs itself is apt to be regarded as a core concept connected with the duality of graph theory. However, it should be emphasized by all means that the concept of dual graphs is primarily *within* the theory of graphs, and is not directly connected with the duality of the theory.

Sometimes, it is said that *the duality of graph theory is restricted because not every graph has its dual*, but that the theory of matroids

has the complete duality because every matroid has its dual. However, the existence of a dual is evidently not connected with the duality of a theory. Contrarily, and even paradoxically, the assertion of the nonexistence of dual graph is an evidence which supports the duality of graph theory. In fact, the sentence in graph theory

[There is a graph having no dual.]

is itself self-dual, so that it is within that part of graph theory which has the dual structure.

As has been shown, the concept of dual graphs is in a crucial position between theory and metatheory. There is another path connecting the level of metatheory and that of theory with respect to graphs. It is the so-called orthogonality of the cycle space and the cocycle space of a graph [1]. The former space is a subspace of the vector space  $V$  of the 1-dimensional chain group of the graph as a 1-dimensional topological complex, and the latter is a subspace of the vector space  $V^*$ , dual to  $V$ , of the cochain group. Thus, the sentence

[The cycle space and the cocycle space of a graph are orthogonal, or complementary, to each other in the mutually dual vector spaces defined with the arc set as the basis vectors.]

is a self-dual theorem of graph theory. Therefore, it is possible to define the concept of dual graphs in another way, i.e. by saying that

[Two graphs  $G_1$  and  $G_2$  on the same arc set are dual if and only if their cycle spaces are the orthogonal complement of each other.] (7)

So long as graphs are concerned, there is no connotative difference between the two definitions of dual graphs, i.e., one (6) based on the correspondence of the concepts of cycles and cocycles in the duality of the theory, and the other (7) based on the orthogonality of the related vector spaces.

Curiously enough, these two definitions, when extended to network theory, lead us to different definitions of dual networks. To be specific, let us consider linear electric networks. In electric network theory, the duality  $(\psi, D)$  of graph theory, exemplified in (4) and (5), is extended into  $(\tilde{\psi}, \tilde{D})$ :

$$\tilde{D} = D \cup \{\text{current, voltage, impedance, admittance, ...}\} \quad (8)$$

and

$$\begin{aligned} \tilde{\psi}|_D = \psi, & \quad \tilde{\psi} \text{ (current)} = \text{voltage,} \\ & \quad \tilde{\psi} \text{ (impedance)} = \text{admittance,} \\ & \quad \dots\dots\dots . \end{aligned} \quad (9)$$

The most fundamental laws in electric network theory, i.e., Kirchhoff's current and voltage laws, may be stated as

[The vector whose components are currents in arcs should belong to the cycle space of the graph.]

and

[The vector whose components are voltages across arcs should belong to the cocycle space of the graph.]

The two laws are obviously the dual of each other with respect to the duality  $(\tilde{\psi}, \tilde{D})$ . Thus, the duality of graph theory is naturally extended to electric network theory through the extension (8), (9) of  $\psi$  and  $D$ .

The concept of dual graphs is extended to that of dual electric networks accordingly; i.e.,

[Two electric networks  $N_1$  and  $N_2$  are said to be dual to each other, or  $N_i$  is a dual of  $N_j$  ( $i \neq j$ ), if their graphs are dual to each other, and if the physical characteristics of arcs in one of them correspond to those in the other under the correspondence  $(\tilde{\psi}, \tilde{D})$ .] (10)

According to this definition, if the current-voltage relation in  $N_1$  is expressed as

$$E_1 = Z_1 I_1 \quad (11)$$

then that in  $N_2$  is expressed as

$$I_2 = Y_2 E_2 \quad (12)$$

and

$$Y_2 = Z_1^{-1} \quad (13)$$

where  $E_i$  and  $I_i$  are, respectively, the voltage vector and the current vector in  $N_i$  ( $i = 1, 2$ ), and  $Z_1$  is the impedance matrix for arcs of  $N_1$  and  $Y_2$  is the admittance matrix for arcs of  $N_2$ .

What will happen if the extension is made on the basis of the relation (7). As has already been noted, nothing new will happen as for the underlying graphs. However, the set of linear constraint equations for currents and voltages in  $N_1$  and that for currents and voltages in  $N_2$  should be related to each other in terms of orthogonality or complementarity in the dual vector spaces. More precisely, the  $Z_1$  in (11) and the  $Y_2$  in (12) should be connected with each other by the equation

$$Z_1^{-1} + Y_2 = 0 \quad \text{or} \quad Y_2 = -Z_1^{-1} \quad (14)$$

In fact, (14) is necessary and sufficient for the orthogonality:

$$E_1^{-1} \cdot E_2 + I_1^{-1} \cdot I_2 = 0 .$$

The distinction between (13) and (14) is clear. This distinction was explicitly pointed out in [3], and further generalized in [4], together with discussions on its network-theoretical implications. Comparing (14) with (13), the minus sign before  $Z_1^{-1}$  in (14) is physically not so important, whereas the transposition seems to be essential. Thus, the two definitions of dual networks coincide, up to a minus sign, with each other for "reciprocal networks" and for them only. (The reciprocity may be defined in terms of the symmetry of impedance or admittance matrix.) If they do not coincide with each other, it turns out that the two networks are "adjoint" with each other in electrical terminology.

As has been shown so far, careful consideration on duality will open new vistas to electric network theory.

#### 4. Triangular Inequalities in Network Flow Theory [2]

Triangular inequalities in network flow theory afford another example which shows the importance of precise understanding of duality concept.



As a natural extension of the duality in graph theory, network flow theory has the dual structure with respect to  $(\bar{\psi}, \bar{D})$ , where

$$\bar{D} = D \cup \{\text{flow, tension, capacity, distance, } \dots\} \quad (15)$$

and

$$\begin{aligned} \bar{\psi}|_D = \psi, & \quad \bar{\psi}(\text{flow}) = \text{tension}, \\ & \quad \bar{\psi}(\text{capacity}) = \text{distance}, \\ & \quad \dots \end{aligned} \quad (16)$$

with  $\psi$  in (5) and  $D$  in (4) (cf. [2], [12], etc.). Thus, the maximum-flow problem on a capacitated network is the dual of the shortest-path problem on a network with distances defined on arcs. (It will be worth noting that, in order to provide network flow theory with as much dual structure as possible, the concept of vertices should be expelled from the theory wherever possible; e.g., by replacing the entrance-exit pair (source-sink pair) in a two-terminal problem by a fictitious "reference arc" connecting the pair of vertices.)

The best known triangular inequality is the one for distances, which is ordinarily described as

"For any three vertices  $v_1, v_2, v_3$  on a network  $N$  with distances defined on arcs, the distance  $d_{ij}$  along the shortest path in  $N$  from  $v_i$  to  $v_j$  ( $i \neq j$ ) satisfies the inequality:  $d_{13} \leq d_{12} + d_{23}$ ." (17)

Since the maximum-flow problem was formulated, another triangular inequality has been noticed [13], which is ordinarily described as

"For any three vertices  $v_1, v_2, v_3$  on a capacitated network  $N$ , the terminal-pair capacity  $c_{ij}$ , i.e., the value of the maximum flow through  $N$ , from  $v_i$  to  $v_j$  ( $i \neq j$ ) satisfies the inequality:  $c_{13} \leq \min(c_{12}, c_{23})$ ." (18)

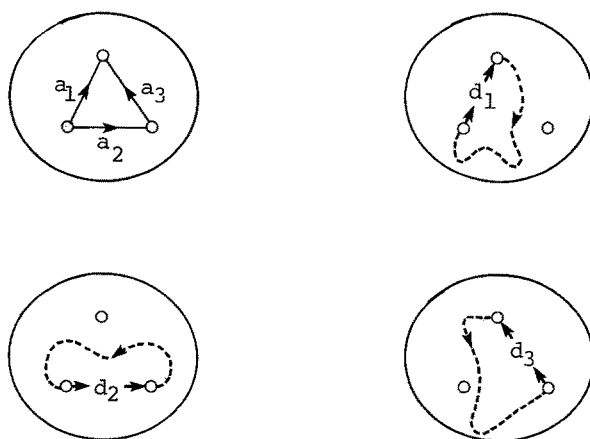
We are tempted to regard these two theorems (17) and (18) as the "dual" theorems because they are basic triangular inequalities in the dual problems. However, a little careful contemplation will make us aware that it is impossible to extend the ordinary duality of graph theory to that of network theory so as to make (17) and (18) the dual of each other.

In order to look for the dual of (17) (and that of (18)), if there is any, we have first to rewrite the statement of the triangular inequality into a dualizable form. It is in fact possible. (17) can be rewritten as

[On a network  $N$  with distances defined on arcs, we choose three arcs  $a_1, a_2, a_3$  such that  $\{-a_1, a_2, a_3\}$  is a circuit. Let  $d_i$  denote the value of the solution of the shortest-path problem with  $a_i$  as the reference arc on the network which is obtained from  $N$  by deleting (opening) the other two arcs  $a_j, a_k$  ( $j, k \neq i$ ). Then, we have  $d_1 \leq d_2 + d_3$ .] (19)

Once (17) is rewritten in this form, it is straightforward to describe the dual "theorem" for a capacitated network as follows (see Figs. 1 and 2):

[On a capacitated network  $N$ , we choose three arcs  $a_1, a_2, a_3$  such that  $\{-a_1, a_2, a_3\}$  is a cocircuit. Let  $c_i$  denote the value of the solution of the maximum-flow problem with  $a_i$  as the reference arc on the network which is obtained from  $N$  by contracting (shortening) the other two arcs  $a_j, a_k$  ( $j, k \neq i$ ). Then we have  $c_1 \leq c_2 + c_3$ .] (20)



$$d_1 \leq d_2 + d_3$$

Fig. 1. Triangular inequality for the shortest-path problem.

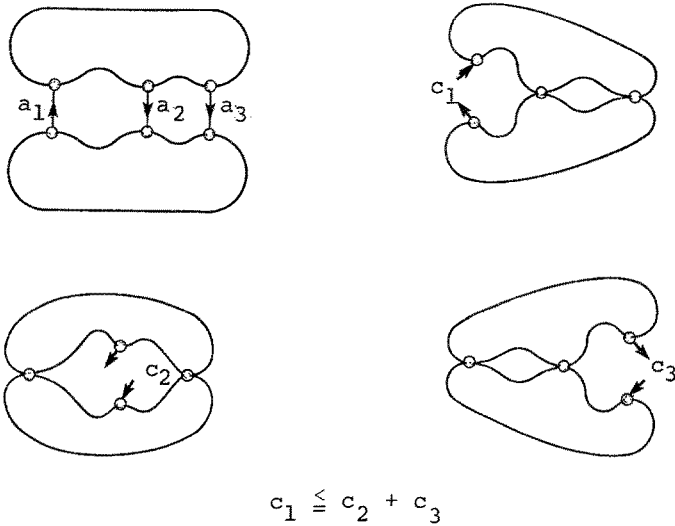


Fig. 2. Another triangular inequality for the maximum-flow problem.

Similarly, we have the dualizable form of (18):

[On a capacitated network  $N$ , we choose three arcs  $a_1, a_2, a_3$  such that  $\{-a_1, a_2, a_3\}$  is a circuit. Let  $c_i$  denote the value of the solution of the maximum-flow problem with  $a_i$  as the reference arc on the network obtained from  $N$  by deleting (opening) the other two arcs  $a_j, a_k$  ( $j, k \neq i$ ). Then we have  $c_1 \geq \min(c_2, c_3)$ .] (21)

and its dual (see Figs. 3 and 4):

[On a network  $N$  with distances defined on arcs, we choose three arcs  $a_1, a_2, a_3$  such that  $\{-a_1, a_2, a_3\}$  is a cocircuit. Let  $d_i$  denote the value of the solution of the shortest-path problem with  $a_i$  as the reference arc on the network obtained from  $N$  by contracting (shortening) the other two arcs  $a_j$  and  $a_k$  ( $j, k \neq i$ ). Then we have  $d_1 \geq \min(d_2, d_3)$ .] (22)

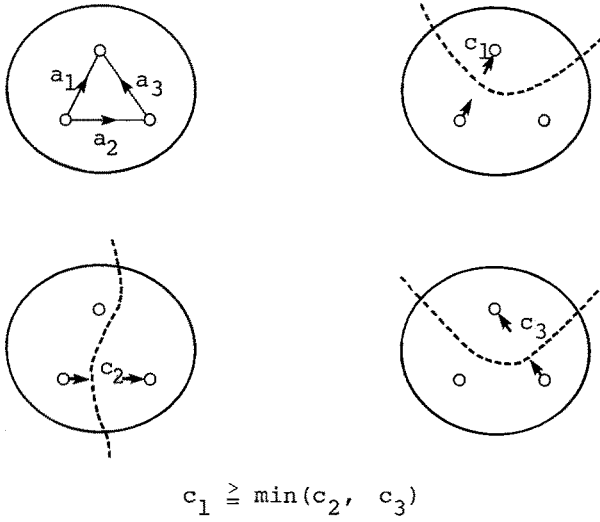


Fig. 3. Triangular inequality for the maximum-flow problem.

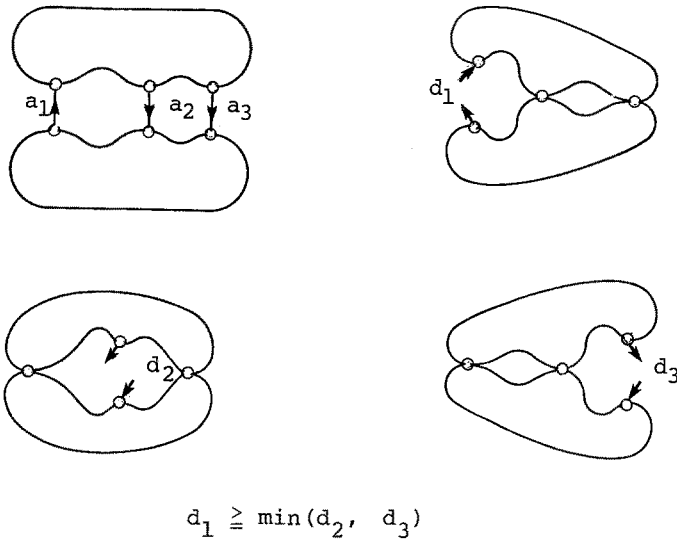


Fig. 4. Another triangular inequality for the shortest-path problem.

## References

- [1] Iri, M.: Metatheoretical considerations on duality. *RAAG Research Notes*, Third Series, No. 124, February 1968.
- [2] Iri, M.: *Network Flow, Transportation and Scheduling — Theory and Algorithms*. Academic Press, New York, 1969.
- [3] Iri, M.; Recski, A.: Reflection on the concepts of dual, inverse and adjoint networks (in Japanese). *Papers of the Technical Group on Circuits and Systems*, Institute of Electronics and Communication Engineers of Japan, CAS 79-78 (September 1979). (English translation available)
- [4] Iri, M.; Recski, A.: Reflection on the concepts of dual, inverse and adjoint networks, II — Towards a qualitative theory (in Japanese). *Papers of the Technical Group on Circuits and Systems*, Institute of Electronics and Communication Engineers of Japan, CAS 79-133 (January 1980). (English translation available)
- [5] Kemeny, J. G.: A new approach to semantics — Part I; Part II. *Journal of Symbolic Logic*, Vol. 21, No. 1 (March 1956), pp. 1-27; No. 2 (June 1956), pp. 149-161.
- [6] Robinson, A.: *On the Metamathematics of Algebra*. North-Holland Publishing Co., Amsterdam, 1951.
- [7] Smullyan, R. M.: *First-Order Logic*. Springer-Verlag, Berlin, 1968.
- [8] Berge, C.: *Graphes et Hypergraphes*. Dunod, Paris, 1970.
- [9] Pontrjagin, L.: *Topological Groups*. Princeton University Press, Princeton, 1946.
- [10] Lefschetz, S.: *Algebraic Topology*. American Mathematical Society Colloquium Publications, Vol. 27, New York, 1942.
- [11] Iri, M.: *Linear Programming* (in Japanese). Hakujitsu-sha, Tokyo, 1973.
- [12] Berge, C.; Ghouila-Houri, A.: *Programmes, Jeux et Réseaux de Transport*. Dunod, Paris, 1962.
- [13] Ford, L. R., Jr.; Fulkerson, D. R.: *Flows in Networks*, Princeton University Press, Princeton, 1962.