Mathematics. — On the Theory of Deduction, Part II. The Definitions of Classical and Intuitionist Negation. By K. R. POPPER. (Communicated by Prof. L. E. J. BROUWER.)

(Communicated at the meeting of November 29, 1947.)

IV

We now turn to the discussion of negation — more especially, of the intuitionist negation of BROUWER and HEYTING, and its relation to other negations. In the present section, we shall discuss the simplest characterizing rules for, and definitions of, the intuitionist negation of a, denoted by " a^{i} "; the classical negation of a, denoted by 9) " a^{k} "; and a third negation of a, denoted by " a^{m} i.e., the "minimum definable (non-modal) negation of a", or the "weakest definable (non-modal) negation of a".

The definitions and characterizing rules of a^i and a^k given in my earlier papers are adequate, but a little complicated and intuitively not as obvious as, for example, 3.71 to 3.82. A characterizing rule for a^i which in its simplicity is comparable to 3.71 etc. is this ¹⁰):

$$(4.1) a \vdash b^i \leftrightarrow a, b \vdash.$$

This may be put into the words: "The intuitionist negation of b is the weakest of those statements which are strong enough to contradict b." The corresponding explicit definition is:

$$(D 4. 1) a//b^i \leftrightarrow (c) (c \vdash a \leftrightarrow c, b \vdash)$$

Thus intutionist negation is, as it were, characterized by contradictoriness alone. One might be tempted to think that classical negation is similarly related to complementarity; but this is not the case. The dual of 4.1 leads to a new kind of negation which is weaker than classical negation. We call this negation " a^m " (the "minimum definable negation of a"); its characterizing rule may be written:

$$(4.2) am \vdash b \leftrightarrow \vdash a, b;$$

⁹) In P_1 p. 220, note 1, I have used instead of "aⁱ" and "a^k" two more complicated symbols. In P_3 , sections III ff., I have used the symbols "aⁱ" and "a^c". I have now replaced "a^c" by "a^k" because I found that the "c" in "a^c" was misleading, occurring as it does together with the statement-variable "c".

¹⁰) Rule 4.1 and all other rules in this section are purely derivational (in the sense of P_1 , p. 230), i.e. they can all be non-vacuously satisfied by non-logical (or factual) statements. Were we to write " $a \land b \vdash$ " instead of " $a, b \vdash$ ", or " $\vdash a \lor b$ " instead of " $\vdash a, b$ ", our rules would no longer be purely derivational, since these rules would need logical (demonstrable or refutable) statements to satisfy them. The reason why my earlier definitions are more complicated then the present ones is that I did not then use " \vdash " and " \nearrow " with more than one argument; this makes it impossible to obtain rules as simple as our present ones if we wish to remain within derivational logic, which I consider highly desirable for non-modal logic. (The definitions of the modalities, of course, cannot be purely derivational; cp. P_2 , end of note 20.)

that is to say, a^m can be characterized as the strongest of those statements which are weak enough to be complements of a. We can¹¹) transform (4.2) into the corresponding definition

$$(D 4. 2) a//b^m \leftarrow \rightarrow (c) (a \vdash c \leftarrow \rightarrow \vdash b, c).$$

There exist two simple rules for classical negation (see 4.311, f., below) which are analogous to 4.1 and 4.2, but less striking. The simplest and most striking characterization I have been able to find is the following explicit definition 1^2)

$$(D 4. 3) a//b^k \leftrightarrow a, b \vdash \mathcal{E} \vdash a, b.$$

That is to say, the classical negation of b can be defined (as Aristotle might have defined it) as that statement which is at once contradictory and complementary to b.

Classical negation, according to this definition, will exist in a language L if, and only if, there exists in L to every statement a a statement a^k which is both contradictory and complementary to a. It is fairly clear, from this definition, that intuitionist negation a^i or its dual a^m , or perhaps both, may exist in a language in which classical negation does not exist; and we shall prove all this by an example (in section V).

Some characterizing rules for classical negation, if written in terms of relative demonstrability, are only slightly different from 4.1 and 4.2, and may be described as generalizations of these rules; 4.31 is equivalent to D4.3, and so is 4.32, which is the dual rule of 4.31:

$$(4.31) a \vdash b^k, c \leftarrow \rightarrow a, b \vdash c$$

 $(4. 32) a, b^k \vdash c \leftarrow \rightarrow a \vdash b, c.$

¹¹) The method here used to obtain relatively simple formulae incorporating the condition "the weakest (or strongest) statement such that ..." is capable of fairly wide application. This may be illustrated by the example of the definition of identity. We use "Idt (x, y)" as the metalinguistic name of a statement-function expressing identity between individuals represented by the individual variables x and y. (Note that, in this characterization, "x" and "y" must not be put into quotes.) We introduce the abbreviating notation:

$$a/|a_{xy} \leftrightarrow (w) (a/|a(x) \otimes a/|a(y)).$$

We can define Idt(x, y) as the weakest statement strong enough to imply what HILBERT-BERNAYS (I, p. 65) call the "second identity axiom", as follows (cp. P_1 , D 6.2, and the correction to it in P_3 , note 1):

$$\frac{a}{Idt}(x, y) \longleftrightarrow (b)(z)((b/|b_{xy}) \to a, b(x)/b(y)) & (((c)(u)(c/|c_{xy}) \to b, c(u)/c(y))) \to b/a)).$$

Adopting the method used in 4.1 of formalizing "the weakest statement such that", this may be replaced by the rule (or the corresponding definition):

$$a/Idt(x, y) \leftrightarrow (b)(z)(b/b_{xy} \rightarrow a, b(x)/b(y)).$$

¹²) The presence of 3.71 or of 3.72 is assumed. For the derivational character of D 4.2 see note 10 above. An identical definition is given in P_2 , p. 284, rule 7.7.

The close relationship to 4.1 and 4.2, respectively, is clear, but the rules are intuitively less satisfactory than D4.3 — especially in view of the fact that D4.3 is an explicit definition. Two other characterizing rules, each of them equivalent to D4.3, may be obtained by inverting 4.1 and 4.2:

$$(4. 311) a, b^k \vdash \leftarrow \rightarrow a \vdash b.$$

$$(4.312) \qquad \qquad \vdash a^k, b \leftarrow \rightarrow a \vdash b$$

The first of these may be expressed in words: "The classical negation of b is a statement that contradicts every statement a which is at least as strong as b". This is similar to 4.1, but, surely, more involved and less striking. 4.312 may be read: "The classical negation of a is a statement which is complementary to every statement b which is at most as strong as a."

Each of the rules and definitions 4.1; D4.1; D4.3 (in the presence of conjunction or disjunction); 4.31 to 4.312; can be shown to be equivalent to the corresponding definitions given in my earlier papers.

V

In the presence of classical negation, that is to say, of characterizing rules for, or of a definition of, classical negation, it is possible to prove

(5.11)
$$a^i / |a^k|$$

(5.12) $a^m/|a^k|$

and therefore also

(5.13) $a^i//a^m$

This follows simply from the fact that the characterizing rules for a^i and a^m allow us to prove equivalence for every statement which satisfies these rules. (This is, precisely, the point which makes what may be called a "fully characterizing rule" equivalent to a definition.) But a^k satisfies the characterizing rules for a^i and for a^m . Thus we obtain 5.11 to 5.13.

This result may be generalized. Whenever we have two logical functions of statements (or two formative signs) S_1 and S_2 , which have been introduced by way of two sets of primitive rules, R_1 and R_2 , such that R_2 is obtained by the omission of some rules of R_1 , then we can prove in the presence of S_1 , the equivalence of (the full expressions of) S_1 and S_2 whenever both are definable. For example, if we introduce "a > b" by rule 3.71 and another function, say " $a \supset b$ ", by the two ¹³) rules (of which the first is like 3.71)

 $(5.21) a \vdash b \supset c \longleftrightarrow a, b \vdash c;$

$$(5. 22) a, b \supset c \vdash b \leftarrow \rightarrow a \vdash b;$$

¹³) The second of these two rules is discussed in section VII below (rule 7.3^k); for another discussion of the same rule see P_1 , pp. 215 f. (rule 4.2e).

then, since 3.71 or 5.21 can be transformed into a definition, we can prove

$$(5. 23) a > b/|a \supset b,$$

in spite of the fact that 5.22 is independent of 5.21, that is to say, that its addition changes the meaning of the sign " \supset ".

The problem arises whether, in the presence of both a^i and a^m in a language L_1 , a^k is always present, so that 5.13 holds, or whether a^i and a^m can coexist in L_1 without becoming equivalent.

We shall prove that the second alternative holds, by constructing, as an example, a language L_1 in which a^i is equivalent to Ia, i.e. to the modal statement which asserts that the state of affairs described by a is *im*possible ¹⁴), and in which a^m is equivalent to Ua, i.e. to the modal statement asserting that the state of affairs described by a is *uncertain* (not necessary).

We have, of course, to choose a language L_1 in which at least one nonlogical statement exists, since we have otherwise Ia//Ua, which would lead to $a^i//a^m$, i.e. to the case we wish to avoid. We shall construct a language L_1 which contains one factual statement s, together with all the compounds which may be constructed from it with the help of the four functions characterized by 3.71 to 3.82 and of the definitions for a^i , a^m , Ia, and Ua. (a^k , of course, does not occur.) We shall have in L_1 (1) demonstrable statements such as s > s; $s \lor Us$; (2) factual statements such as s; $s \land Us$; $s \lor Is$, and (3) refutable statements such as $s \land Is$; $s \preceq s$.

In order to show that a^i may be here without contradiction identified with Ia, and a^m with Ua, we construct an arithmetical model of our metalanguage, interpreting our variables "a", "b", etc., as variables whose values are the three numbers 1, 2, and 3, and " $a_1, \ldots, a_n/b$ " as the statement asserting that the greatest of the numbers a_1, \ldots, a_n is at least equal to b.

On the basis of our characterizing rules for Ia; Ua; $a \land b$; $a \lor b$; a > b, and $a \not\preceq b$, this interpretation forces us to accept the following: " $a \land b$ " is to be interpreted as the greater of the two numbers a and b; " $a \lor b$ " as the smaller of them; etc. We obtain the matrices (also useful for showing that D 4.3 is independent of 4.1, etc.):

а	Ia	Ua	a∧b	123	a∨b	123	a>b	123	a¥b	123
1	3	3	1	123	1	111	1	123	1	311
2	3	1	2	223	2	122	2	113	2	332
3	1	1	3	333	3	123	3	111	3	333

¹⁴) *Ia*, *Ua*, and the other modal functions are defined in P_3 , section VIII. A simpler definition (taken from P_2 , p. 283, note 20) of *Ia* is given in section VII below, D.7. I. The dual of this definition is:

$$a||Ua \leftrightarrow (c) ((c/a V c/b) \& (a/b \rightarrow a/c)).$$

In order to show that in L_1 , $a^i//Ia$ and $a^m//Ua$, we can use the generally valid dual rules:

(5.31)
$$a^i / |a > Ia$$

or alternatively 15)

(5. 41)
$$a^i//a > (a \searrow a)$$

(5.42)
$$a^m/(a > a) \searrow a$$

Evaluating with the help of our numerical tables either 5.31 and 5.32 or 5.41 and 5.42, the equivalences $a^i //Ia$ and $a^m //Ua$ can easily be shown to be valid in L_1 .

This result is not in general true, but the fact that it holds for L_1 establishes that, without risking a contradiction, we may postulate $a^i//Ia$, provided that classical negation does not exist in the language under consideration. This justifies the well known intuitionistic identification of a^i with "a is impossible".

Our result is by no means trivial since from our definition of a^i (which is equivalent with the implicit characterization of HEYTING's calculus) we obtain

$$(5.5) a, b|a^i \to b|a^i,$$

while we cannot obtain the corresponding formula for Ia from its definition; if we could, the equivalence of Ia and a^k could be established in the presence of a^k , which is obviously not possible, in view of our definitions of Ia and a^k . The place of 5.5 is taken by the weaker rules

$$(5.51) a, b/Ia \to b/a^i.$$

$$(5.52) a/Ia \to b/Ia.$$

The fact that a rule which is like 5.5 but with "Ia" instead of "a" cannot be shown to follow from our definition of Ia might be considered, at first sight, as speaking against the intuitionist identification of a^i with Ia. But the problem is whether a^i is equivalent to Ia in the absence of a^k . That this is the problem may be seen from the fact that intuitionism admits that the law of the excluded middle holds in a wide range of cases, which means, in our way of speaking, that for certain statements a it may happen that a^i is complementary to a and thus coincident with a^k .

$$\mathbf{a} \vdash b \not\perp c, d \leftarrow \rightarrow (e) (b \vdash e, c \rightarrow a \vdash e, d).$$

We also obtain, with the modus ponens, its dual:

$$a, a > b \vdash b$$
 $a \vdash a \not\ge b, b$

¹⁵) It is remarkable that both a^i and a^m can be defined in terms of ">" and its dual " \preceq ", while a^k cannot be so defined (as proved by our example L_1). We see that every language containing ">" and " \preceq " necessarily contains a^i and a^m (possibly undistinguishable from a^k) while it need not contain a^k . — It may be mentioned that the theory of the anti-conditional " \preceq " is quite interesting. We have, for example, the dual of 7.40 (see section VII, below), i.e.:

A complete formal justification of the intuitionist identification of a^{t} with the impossibility of a could be obtained by proving the following very general

Conjecture: If a language L does not contain the classical negation of every of its statements but does contain a^i and Ia for every statement a, then the following holds: if, for some b, there does not exist a b^k , then Ib is the weakest statement contradictory to b, so that b^i is equivalent to Ib.

I have so far not been able to prove or disprove this conjecture. (The problem is a straightforward calculation in our metalinguistic calculus, and most probably not difficult to solve; cp. its formal presentation at the end of section VII.) Meanwhile, our example of a language L_1 for which this conjecture holds establishes that, for a language which does not contain classical negation, we may postulate that a^i is identical with Ia, without fear of contradiction.

It should be noted that 5.5 is the only one of the more important rules which a^i and Ia do not share. Among the more important rules which can be derived from the definition of Ia and which hold for a^i as well are:

- (5.6) *a, Ia* ⊢
- $(5.7) a \vdash Ia \rightarrow \vdash Ia$
- $(5.71) Ia \vdash a \rightarrow \vdash IIa$
- (5.8) a⊢ IIa
- $(5.9) a \vdash lb \to b \vdash la.$

Equally important is that the following classical principles (which also hold for a^m and Ua) cannot be shown to hold generally of either a^i or Ia:

- $(5.6') \qquad \qquad \vdash a, a^m$
- $(5.7') a^m \dashv a \rightarrow \mu a$
- $(5.8') a^{mm} \dashv a$
- $(5.9') a^m \vdash b \rightarrow b^m \vdash a.$

VI

We shall now extend our considerations to other kinds of negation, even to fairly remote and unusual ones. Neglecting (1) Ia and Ua, which were considered in the last section, we shall now consider (2) a^k ; (3) a^i ; (4) a^m ; (5) fa, i.e. the self-contradictory compound of a (for which fa//fbholds) definable, for example, by

$$a/(fb \leftrightarrow (c)(a/b \leftrightarrow a/c);$$

and ultimately (6) three negations, a^{j} ; a^{l} ; and a^{n} , each of which is to be considered as introduced with the help of one of the following three primitive rules, respectively:

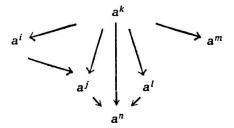
- $(6.1) a, b/c^j \to a, c/b^j.$
- $(6.2) a, b^l/c \to a, c^l/b.$
- $(6.3) a, b/c \to a, c^n/b^n.$

(These three rules will be shown *not* to be equivalent to definitions, that is to say, they are not fully characterizing rules in our sense.) 6.1 is equivalent to the axioms which introduce the negation a^{j} of JOHANSSON's so-called "Minimalkalkül". In view of 6.2, we may call a^{l} the "left-hand side negation of a" (in contradistinction to JOHANSSON's a^{j} which, in view

of 6.1, is a "right hand side negation"). a^n may be called the "neutral negation"; it is neutral with respect to right-sidedness and left-sidedness, as is a^k which, indeed, can be fully characterized by the converse of 6.3, viz. by:

$$(6.4) a, b^k/c^k \to a, c/b.$$

The following diagramme indicates the way in which the rules of the six negation a^k , a^i , a^m , a^j , a^l , and a^n are satisfied.



The arrows indicate, for example, that a^k satisfies the rules holding of all the others, or that a^i satisfies the rules holding of a^j and a^n , and that the latter ones are satisfied by all others (except a^m which is so weak ¹⁶) that it does not satisfy even 6.3).

We shall now show that the three negations listed under (6) cannot be defined.

In order to show this for a^{j} , we introduce the following auxiliary definition of "tb":

$$a||tb \leftrightarrow (c) (b|a \leftrightarrow c/a)$$

"tb" may be called the "self-complementary compound of b". We obtain ta//tb.

It can now be shown that tb satisfies 6.1. On the other hand, a^k also satisfies 6.1. Thus, if a^j were definable, we would obtain $ta//a^j$ and $a^k//a^j$ and therefore $ta//a^k$. But this is possible in contradictory languages only; since we have $(a > a)^k \vdash and \vdash t(a > a)$, it is clear that $ta//a^k$ would lead to a//b. Thus there cannot be a definition equivalent to 6.1, and a^j cannot be defined 1^7).

Similarly, 6.2 is satisfied by fa and a^k ; thus a^l cannot be defined.

¹⁶) There are, of course, dual rules of 6.1, 6.2, and 6.3, two of which are satisfied by a^m , just as 6.1 and 6.3 are satisfied by a^i . Note that our diagramme does not contain the duals of a^j , a^l , and a^k , and that it is therefore not fully symmetrical.

¹⁷) For JOHANSSON's calculus, see HILBERT-BERNAYS II, 449 f. The fact that his negation cannot be defined, and that this impossibility can be proved, is mentioned in P_2 , p. 286.

6.3 is satisfied by ta, fa and a^k ; thus a^n cannot be defined.

The fact that JOHANSSON's negation a^{j} cannot be formally distinguished from ta (which is not, by any stretch of imagination, to be called a negation) speaks strongly against its adoption. a^{l} seems slightly preferable since fa has something in common with Ia; for fa is equivalent to Ia whenever a is either factual or demonstrable. a^{n} is, perhaps, the best of the three negations under (6).

But the fact that none of the three is definable speaks very strongly against all of them; indeed, I suggested in P_3 that the term "formative sign" should be applied only to signs whose meaning is definable by definitions in terms of deducibility. Should this suggestion be accepted, then we would have to say that those signs of a language L which represent JOHANSSON's negation, or the others under (6), are not formative.

On the other hand, the very fact that these three signs cannot be defined makes it possible to combine them in the same language L with classical negation without running the risk of destroying their distinguishableness; which is not possible for the signs under (3) and (4).

VII

We now turn to some concluding remarks about the existential assumptions connected with intuitionist and classical negation. They will provide, at the same time, examples of the kind of existential problems which are far from trivial and which arise, and can in principle be solved, within the general theory of derivation.

It is well known that intuitionism does not assert that there does not exist a classical negation of any statement; on the contrary, BROUWER has not only asserted that the law of the excluded middle is valid for certain entities, but has even given a proof ¹⁸) of its validity for a certain range of entities. This means, from our point of view, that there exist some statements a to which statements b exist which are both contradictory and complementary to a. Where intuitionism deviates from classical logic is in its assertion that such statements b do not exist to every statement a.

In other words, intuitionism asserts, for the language which it considers (the language in which mathematicians deal with infinite sets):

$$(7.1i) (a) (Eb) (c) (c \vdash b \leftarrow \rightarrow a, c \vdash).$$

The corresponding assertion or postulate of classical logic is

(7.1^k) (a) (Eb) (c) (($c \vdash b \leftrightarrow a, c \vdash) \mathcal{E}(\vdash a, c \leftrightarrow b \vdash c)$). which, of course, implies 7.1ⁱ.

Now intuitionism does not only assert 7. 1^i , but it denies 7. 1^k ; that is to say, it asserts besides 7. 1^i the following principle which is a negation (an intuitionist one) of 7. 1^k :

$$(7.2^i) (Ea) (b) (Ec) (d) (e) (((c \vdash b \leftarrow \rightarrow a, c \vdash) & (\vdash a, c \leftrightarrow b \vdash c)) \rightarrow d \vdash e).$$

¹⁸) Cp. L. E. J. BROUWER, Intuitionistische Betrachtungen über den Formalismus (Sitzungsberichte Preuss. Akad., 1928, V, esp. pp. 51 f.).

In other words, the fundamental assertion of intuitionism — the one by which it is distinguished from classical logic — is an existential assertion; it asserts the existence of a statement *a* for which the application of the classical principle 7.1^k leads to a contradiction. (This is the force of the clause " $d \vdash e$ " which allows us to deduce a contradiction from a tautology.)

Intuitionism does not consider an existential assertion such as 7. 2^i as legitimate if it cannot be supported by an actual construction of an example; it is therefore a crucial task for intuitionism to give an example of a statement which, if treated classically, leads to a contradiction. In other words, the proof ¹⁹) of the contradictoriness of classical mathematics must be of crucial importance for intuitionism. It is an agreement with this result that BROUWER considers the proof of the contradictoriness of classical mathematics as one of the most central problems of intuitionist mathematics.

We now proceed to state some principles equivalent to 7.2^{i} .

One such principle is obtained by asserting the existence of an intuitionist negation which is not classical. This may be written, considering 7.1^{i} and 7.1^{k} :

(7. 21^{*i*}) (*Ea*)(*Eb*)(*c*)(*d*)(*e*)(($c \vdash b \leftarrow \rightarrow c, a \vdash$) & (($b \vdash c \leftrightarrow \vdash c, a$) $\rightarrow d \vdash e$)) In the presence of 7. 1^{*i*}, 7. 21^{*i*} is equivalent to 7. 2^{*i*}.

A less obvious equivalence can be obtained if we remember 20) that, in the presence of intuitionist negation and the conditional, classical negation can be derived from the following principle 7. 3^{k} (cp. also 5. 22):

$$(7. 3k) a, b > c/b \to a/b.$$

In this form, the existential character of 7.3^k is not very obvious, but it becomes obvious if we eliminate here the sign ">", with the help of (7.40) $a, b > c/d \leftrightarrow (e) (b, e/c \rightarrow a, e/d)$

which an alternative characterizing rule 21) for ">", obtainable from 3.81 (or from the definition of "b > c"). Applying 7.40, we can transform 7.3^k into 7.4^k:

$$(7. 4k) \qquad (a) (b) (c) (Ed) ((b, d|c \rightarrow a, d|b) \rightarrow a|b)$$

which shows its existential character. Negating 7.4^{*k*}, we obtain a principle 7.4^{*i*} which (in the presence of 7.1^{*i*} and of a formula asserting the existence of the conditional) is equivalent to 7.2^{i} :

$$(7.4^{i}) \qquad (Ea) (Eb) (Ec) (d) (e) (f) (((b, d/c \rightarrow a, d/b) \rightarrow a/b) \rightarrow e/f).$$

In this case, the equivalence is by no means obvious, but demands a fairly complicated proof, based, of course, upon BI.1 and BI.2. It shows that the consequences derivable from this basis are not all trivial.

As an example of an interesting and perhaps more difficult problem,

¹⁹) This proof must, of course, be intuitionistically valid; and therefore, a fortiori. classically valid also.

²⁰) Cp. note 13, and rule 5.22, above, and P_1 , pp. 215 f: (rule 4.2e).

²¹) For the dual of rule 7.40, see note 15 above.

I may refer to the conjecture, formulated in section V, that $a^i//la$ whenever a^i and la exist but not a^k . This conjecture can be proved or disproved by proving or disproving a metalinguistic conditional with the conjunction of the closures of BI. 1, BI. 2 and the definition of " $\vdash a$, b" and "a, $b \vdash$ " as antecedent and a consequent which asserts: "If a is (not equivalent to b^m and therefore) not the classical negation of b then, if a is the intuitionist negation of b, then a//lb." In order to formalize this assertion, we shall make use of the following comparatively simple definition ²² of lb:

$$(D7. I) a||Ib \leftrightarrow (c) ((a|c V b|c) & (b|a \rightarrow c|a)).$$

With the help of this, we can write the assertion in question as follows:

$$(a) (b) (c) (d) (Ee) (Ef) (Eg) (((a/d \leftrightarrow \rightarrow \vdash b, d) \rightarrow e/f) \rightarrow ((g/a \leftrightarrow \rightarrow b, g \vdash) \rightarrow ((a/c \ V \ b/c) \& (b/a \rightarrow c/a))).$$

If the metalinguistic conditional here described is refutable, and a counter example can be constructed, 23) i.e. a language for which it is not valid, then the problem arises of formulating the necessary and sufficient conditions under which it holds; for that it holds for some languages has been established by our example in section V.

²²) Cp. note 14, above.

²³) Note added in the proofs. I have now been able to construct various simple counter examples which refute the conjecture.