Annals of Mathematics
Vol. 43, No. 2, April, 1942

ON THEORIES WITH A COMBINATORIAL DEFINITION OF "EQUIVALENCE"

By M. H. A. Newman

(Received June 23, 1941)

1

The name "combinatorial theory" is often given to branches of mathematics in which the central concept is an equivalence relation defined by means of certain "allowed transformations" or "moves." A class of objects is given, and it is declared of certain pairs of them that one is obtained from the other by a "move"; and two objects are regarded as "equivalent" if, and only if, one is obtainable from the other by a series of moves. For example, in the theory of free groups the objects are words made from an alphabet $a, b, \cdots, a^{-1}, b^{-1}, \cdots$, and a move is the insertion or removal of a consecutive pair of letters xx^{-1} or $x^{-1}x$. In combinatorial topology the objects are complexes, and the allowed moves are "breaking an edge" by the insertion of a new vertex, or the reverse of this process.[1] In Church's "conversion calculus"[2] the rules II and III are "moves" of this kind.

In many such theories the moves fall naturally into two classes, which may be called "positive" and "negative." Thus in the free group the cancelling of a pair of letters may be called a positive move, the insertion negative; in topology the breaking of an edge, in the conversion calculus the application of Rule II (elimination of a λ), may be taken as the positive moves. In theories that have this dichotomy it is always important to discover whether there is what may be called a "theorem of confluence," namely, whether if A and B are "equivalent" it follows that there exists a third object, C, derivable both from A and from B by positive moves only. A closely connected problem is the search for "end-forms," or "normal forms," i.e. objects which admit no positive move. It is obvious that in a theory in which the confluence theorem holds no equivalence class can contain more than one end-form, but there remains the question whether in such a class any random series of positive moves must terminate at the end-form, or whether infinite series of moves may also exist.

The purpose of this paper is to make a start on a general theory of "sets of moves" by obtaining some conditions under which the answers to both the above questions are favorable. The results are essentially about "partially-ordered" systems, i.e. sets in which there is a transitive relation $>$, and sufficient conditions are given for every two elements to have a lower bound (i.e. for the set to be "directed") if it is known that every two "sufficiently near" elements have a lower bound. What further conditions are required for the existence of a *greatest* lower bound is not relevant to the present purpose, and is reserved for a later discussion.

---

[1] See Alexander [1] and Newman [1].

[2] See Church [1] and references there given.

As an application the normal form theorem of Church and Rosser [1] in the conversion calculus is derived.

2

We are concerned with two kinds of entities, "objects" and the "moves" performed on them, and each move is associated with two objects, "initial" and "final." We are therefore dealing essentially with *indexed 1-complexes* (in which, therefore, a positive sense is assigned in each 1-cell), the vertices being the "objects," and the positive 1-cells the "moves." It will be convenient to make use of this topological terminology.[3] The incidence relations are in no way restricted: there may be many cells with the same vertices, and the initial and final vertices of a cell may coincide. In diagrams the positive 1-cells slope down the paper, and some of the terms used are chosen accordingly.

Vertices are denoted by italic letters, cells (the single word is used from now on for "positive 1-cell") by the letters ξ, η, ζ, ω with various suffixes. "$x\mu y$" means "there is a cell with initial vertex x and final vertex y." An ordered set of cells $\xi_1, \xi_2, \cdots, \xi_k$, form a *path* π if there are vertices $x_0, x_1, \cdots, x_k$ such that x_{i-1} and x_i are the vertices of ξ_i for $1 \leqq i \leqq k$. The cell ξ_i is *direct* or *reversed* in π according as it runs from x_{i-1} to x_i or from x_i to x_{i-1}, and the path is denoted by $e_1\xi_1 + e_2\xi_2 + \cdots + e_k\xi_k$, where e_i is ± 1 as ξ_i is direct or reversed. If there are no reversed cells, π is a *descending* path. It is convenient to regard a single vertex, x, as a "null path" with x as initial and final vertex. A vertex which is not the initial vertex of any cell is a *minimal vertex*, or *end*.

If there is at least one non-null descending path from x to y we write $x > y$. z is a *lower* (*upper*) *bound* of x and y if $x \geqq z$ and $y \geqq z$ (if $z \geqq x$ and $z \geqq y$).

3

Expressed in this terminology the confluence property is

(A) *If x_1 and x_2 are connected by a path in the indexed complex Σ they have a lower bound.*

By a simple induction on the number of cells in a path from x_1 to x_2 this property can be deduced from the following special case of it:

(B) *If x_1 and x_2 have an upper bound they have also a lower bound.*

This in its turn is easily deduced from the still more special form (C):

(C) *If $a\mu x_1$ and $a > x_2$, x_1 and x_2 have a lower bound.*

The transition from (B) to (C) is a step towards localizing the property, and the theorems that will be proved in this paper give conditions in which the localization may be completed, i.e. in which (A) may be inferred from the following condition (holding for all a, x_1 and x_2):

(D) *If $a\mu x_1$ and $a\mu x_2$, x_1 and x_2 have a lower bound.*

NOTE. The cell and vertex terminology, although the most convenient for

---

[3] The notions that arise are closely related to those of the theory of partially ordered sets, but usually not identical. Except in the case of identity the terms of that theory are therefore avoided.

our purpose, may suggest that "$x\mu y$" implies that y is a "next" vertex below x. Actually the force of μ is that y is an element satisfying $y < x$, and lying in a certain neighborhood of x. For example, all the conditions (A) to (D) are satisfied if the vertices are taken to be the points of a vertical plane, and the positive 1-cells the downward sloping directed segments of length less than 1.

4

The simplest way of strengthening (D) so that it implies (A), is to require that paths descending from x_1 and x_2 to their lower bound shall each contain one cell; or, in terms of moves, that if two moves are possible on an object X, they can also be performed one after the other, and give the same result in either order.

THEOREM 1. *Let Σ be such that if $a\mu x$ and $a\mu y$, and $x \neq y$, there exists b such that $x\mu b$ and $y\mu b$. Then property* (A) *holds.*

Let "$x\nu y$" denote "$x\mu y$ or $x = y$." We prove that if $a\mu x$ and $a\mu y_1\mu y_2\mu \cdots \mu y_k$, there is a b_k such that $x\nu b_1\nu b_2\nu \cdots \nu b_k$ and $y_k\nu b_k$,—a stronger form of (C). Suppose this proved for $k - 1$ (the case $k = 1$ following immediately from the datum), and let $x\nu b_1\nu b_2\nu \cdots \nu b_{k-1}$, and $y_{k-1}\nu b_{k-1}$. If $y_{k-1} = b_{k-1}$ take b_k to be y_k. If $y_{k-1}\mu b_{k-1}$, since also $y_{k-1}\mu y_k$ there exists a b_k such that $b_{k-1}\nu b_k$ and $y_k\nu b_k$; and this completes the induction.

COROLLARY 1.1. *The theorem remains true if "$x\nu b$ and $y\nu b$" is substituted for "$x\mu b$ and $y\mu b$" in the enunciation.* (No change is needed in the proof.)

This almost trivial result is sufficient to settle many of the more familiar theorems of the kind that we are considering. In the "word groups" already referred to, a move is to be regarded as completely determined by the initial and final words, (so that e.g. $xx^{-1}x \to x$ is regarded as the same move whether the first or last two letters are cancelled). Hence two pairs $\mathbf{xx}^{-1}$ and $\mathbf{yy}^{-1}$ (where $\mathbf{x}$ and $\mathbf{y}$ may be of the form u^{-1}) in the same word $\mathbf{W}$, that give rise to different possible moves on $\mathbf{W}$, have no common letter and give the same result if cancelled in either order. Since every series of positive moves (cancellations) terminates it follows that all such series starting from a given word $\mathbf{W}$ lead to a common end-form.

Theorems of the Jordan-Hölder type also belong to this category. The kernel of these theorems is a theorem on modular lattices (say with the partial ordering $\succ$ and the operations $\vee$ and $\wedge$). If X, Y, Z are consecutive elements in a descending chain, $\mathfrak{S}$, in such a lattice let the chain $\mathfrak{S}'$ obtained by substituting Y' for Y be said to be *directly related* to $\mathfrak{S}$ ($\mathfrak{S}'$ dr $\mathfrak{S}$) if $X = X \vee Y'$ and $Z = Y \wedge Y'$; and $\mathfrak{S}'$ shall be *related* to $\mathfrak{S}$ if it is obtainable from $\mathfrak{S}$ by a succession of such steps. The theorem in question is then that *from any two finite descending chains, $\mathfrak{S}$ and $\mathfrak{S}'$, from A to B, a pair of related chains $\mathfrak{S}_1$ and $\mathfrak{S}_1'$, can be obtained by the insertion of a finite number of additional terms in $\mathfrak{S}$ and $\mathfrak{S}'$ respectively.* This is evidently a "confluence" theorem. To apply 1 we take as a typical vertex of Σ the class $[\mathfrak{S}]$ of all chains related to a chain $\mathfrak{S}$, and as a positive 1-cell the ordered pairs of classes $[\mathfrak{S}_1]$, $[\mathfrak{S}_2]$, where $\mathfrak{S}_2$ is obtained from $\mathfrak{S}_1$

by the insertion of one additional term,—say P—between X and Y. Then if $\mathfrak{S}_1'$ dr $\mathfrak{S}_1$, the insertion of a suitable term in $\mathfrak{S}_2$ gives a chain $\mathfrak{S}_2'$ related to $\mathfrak{S}_1'$;[4] and hence more generally any member of $[\mathfrak{S}_1]$ can be made into a chain related to $\mathfrak{S}_2$, by the insertion of one suitable term. Two successive "positive moves" on $[\mathfrak{S}_1]$ can therefore be represented by two successive insertions of new elements in the *same* chain $\mathfrak{S}$, and evidently the order in which they are inserted does not affect the result. The system therefore fulfils the conditions of Theorem 1. But any two chains descending from A to B have an "upper bound" in Σ, namely the class $[AB]$. Therefore they have a "lower bound," and this is the required result.

5

In these examples it is obvious that if an end-form exists it is reached by random descent. This is necessarily so in all systems with non-interference of moves:

THEOREM 2. *Under the conditions of Theorem 1, if there is a descending path of k cells from a to an end e, no descending path from a contains more than k cells.*

If $k = 1$, Σ cannot contain a cell ay with $y \neq e$, since if it does b exists such that $y\mu b$ and $e\mu b$, and e is not an end. In the general case let π be a descending path $\xi_1 + \xi_2 + \cdots + \xi_k$ joining a to e, and let $\eta_1 + \eta_2 + \cdots + \eta_j$ be any descending path from a. Let ξ_1 and η_1 be cells ax and ay. If $x = y$ it follows immediately from an induction that $j \leqq k$. If not, let the cells ζ and ω descend from x and y to the common vertex w. By Theorem 1 there is a descending path σ from w to a vertex $\leqq e$, i.e., since e is an end, to e itself. Since $\xi_2 + \cdots + \xi_k$ has $k - 1$ cells, $\zeta + \sigma$ has, by an inductive hypothesis, at most $k - 1$ cells; therefore $\omega + \sigma$, and finally also $\eta_2 + \cdots + \eta_j$, have at most $k - 1$ cells,—i.e. $j \leqq k$.

COROLLARY 2.1. *Every descending path from a is part of a descending path of k cells from a to e (i.e. there is "random descent" to e).*

That Theorem 2 and Corollary 2.1 fail if the condition is weakened as in Corollary 1.1 is shown by the example in Fig. 1, (positive cells slope downward).

The main criteria for "confluence" are established in Theorems 3, 4, 5, and 9, all of which are independent. It is Theorems 5 and 9 that are used in the application to the conversion calculus.

THEOREM 3. *In an indexed complex in which all descending paths are finite,* (D) *implies* (A).

(Note that in such a complex "$>$" is a proper ordering, since if $x > x$ an infinite descending path is obtained by going round and round the re-entrant path from x to x.)

---

[4] Namely, if X and Y are in $\mathfrak{S}_1'$, insert P itself; if XYZ and $XY'Z$ are consecutive terms of $\mathfrak{S}_1$ and $\mathfrak{S}_1'$ respectively, insert $P' = Y' \wedge P$ in $\mathfrak{S}_1'$; if UXY and $UX'Y$, insert $P'' = X' \vee P$. It is easily shewn that in the second case $XPYZ$ is related to $XY'P'Z$, in the third $UXPY$ to $UP''X'Y$. Cf. Birkhoff [1] p. 37.

The symbol $[\xi]_k$ is used as an abbreviation for $\xi_1 + \xi_2 + \cdots + \xi_k$. It is convenient to allow the value $k = 0$, $[\xi]_0$ being a null path.

A *peak* of a path is the common vertex of a successive pair of cells $-\xi + \eta$, ("up" before "down").

Let $[\xi]_j$ and $[\eta]_k$ be paths descending from a vertex a to vertices b and c respectively. Let π_1 be the path $-[\xi]_j + [\eta]_k$, and let it be assumed that paths $\pi_2, \pi_3, \cdots, \pi_r$, each leading from a to b, have been defined. Let X_r be the (finite) indexed subcomplex of Σ formed by all the cells occurring in the paths $\pi_1, \cdots, \pi_r$. The *depth in* X_r of a vertex x is defined to be the maximum possible number of cells in a descending path from a to x in X_r, (or 0 if there is no such path). Thus the depth of any vertex in X_{i+1} is not less than its depth in X_i.

If π_r contains no peak, π_{r+1} is not defined. If it contains at least one peak, choose one, say y, of minimum depth in X_r among peaks of π_r. Let the vertices

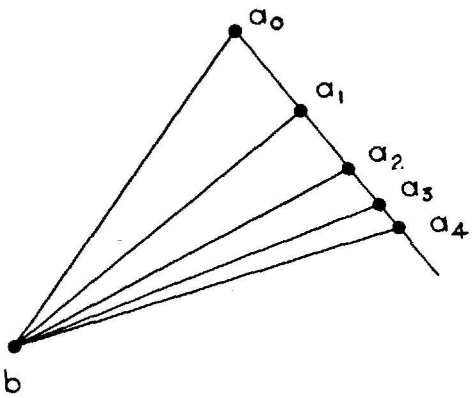

FIG. 1

immediately preceding and succeeding y on π_r be u and v, the (positive) cells yu and yv being ω and ω' respectively. There exist, by (D), paths σ and τ, (either or both of which may be null), descending from u and v to a common vertex w, and π_{r+1} is formed from π_r by substituting $\sigma - \tau$ for $-\omega + \omega'$. The effect is to replace the peak at y by at most two new ones, of depths in X_{r+1} at least 1 greater than that of y in X_r (or zero). By a simple induction it follows that π_r has at most r peaks; and if we make the inductive hypothesis that the peaks of π_{2^n} are of depth at least n in X_{2^n} it follows that, if $r \leqq 2^n$, at most $2^n - r$ peaks of π_{2^n+r} are of depth n or less in X_{2^n+r}. Thus the induction is complete and it is proved that *if* $r \geqq 2^n$ *all peaks of* π_r *are of depth at least* n.

If $[\zeta]_n$ *is a descending path of maximum length in* X_r *from* a *to a given vertex* z, *where* $r \geqq 2^m$, ζ_i *belongs to* X_{2^i} *for* $i = 1, 2, \cdots, m$. Suppose that, for a certain i, X_j is the first of the X's to contain ζ_i, where $j > 2^i$. Then $[\zeta]_i$ is a descending path in X_j of maximum length to its final vertex, z_i, since any longer one could be used as part of a longer descending path to z in X_r. Thus the depth of z_i in X_j is i. Since ζ_i belongs to X_j but to no earlier X, it is a cell of one of the descending paths that eliminate a peak, y, in the formation of π_j from π_{j-1}. By the result of the preceding paragraph, if the depth of y in X_{j-1} is $p, j - 1 < 2^{p-1}$. But the depth of z_i in X_j exceeds that of y in X_{j-1} by at least 2, $i \geqq p + 2$. Therefore $j - 1 < 2^{i-1}, j < 1 + 2^{i-1} \leqq 2^i$, contrary to the hypothesis.

The series of paths $\pi_1, \pi_2, \cdots$, *terminates.* If not choose, for each n, a maximal descending path, σ_n, in X_{2^n} from a to a peak of π_{2^n}. Since the first cell of each of these paths is in the finite complex X_2 there is at least one cell, ω_1, which is the first cell of σ_n for an infinity of n. Since the second cell of each of this infinite subsequence is in X_4 there is at least one cell, ω_2, such that $\omega_1 + \omega_2$ is the beginning of an infinity of the σ_n. Continuing in this way we obtain an infinite descending path $\omega_1 + \omega_2 + \cdots$ in Σ,—contrary to its given property.

Thus the series of paths π_r from b to c terminates in a path π_q, which, since it has no peak, must descend or ascend directly from b to c, or else descend from b to a vertex w and then rise to c.

The finiteness condition imposed on descending paths in Theorem 3 cannot be replaced by the corresponding "completeness" condition, that every descending chain of vertices has a lower bound in Σ. This is shown by the complex in Fig. 3, in which the vertices c and d are lower bounds of all sets of vertices not containing either of them; but c and d have themselves no lower bound.

6. Topology of Σ

The complex Σ can be made into a 2-complex, Σ^2, by adding a 2-cell bounded by each of the 1-cycles $\omega + \sigma - \tau - \omega'$ occurring in the proof of theorem 3 (one for each π_r). *Every component of* Σ^2 *is simply connected.* Any two paths, π and π', connecting vertices a_1 and b_1 are deformable, by the method of Theorem 3, into paths $\sigma_1 - \tau_1$ and $\sigma_1' - \tau_1'$ respectively, where σ_1 and τ_1 descend to a vertex a_2, σ_1' and τ_1' to b_2; and if $\rightsquigarrow$ stands for "is deformable into," $-\tau_1 + \tau_1' \rightsquigarrow \sigma_2 - \tau_2$, and $-\sigma_1 + \sigma_1' \rightsquigarrow \sigma_2' - \tau_2'$, where $\sigma_2, \tau_2, \sigma_2', \tau_2'$ are descending paths, the first two to a_3, the second two to b_3. In this way paths σ_n and τ_n descending to a_{n+1}, and σ_n' and τ_n' to b_{n+1}, are defined for every n. If an infinity of different paths descending from a_1 could be made from the $\sigma_i, \tau_i, \sigma_i'$ and τ_i', an infinity of them would necessarily contain one or other of σ_1, σ_1',—say σ_1; and of these an infinity would contain one or other of σ_2, σ_2',—say σ_2'; and so on. The descending path $\sigma_1 + \sigma_2' + \cdots$ so constructed would have an infinity of different paths as subsets, and would therefore be infinite, contrary to the postulated property of Σ. The number of different paths must therefore be finite.

It follows that for some m, $\sigma_m = \tau_m = \sigma_m' = \tau_m' = 0$; i.e.

$$\pi - \pi' \rightsquigarrow \sigma_1 - \tau_1 - \tau_1' - \sigma_1' \rightsquigarrow \sigma_m - \tau_m + \tau_m' - \sigma_m' = 0.$$

7

To establish our second criterion, we suppose that Σ is the sum of two subcomplexes,[5] L and R, and shall use the terms "L-cell," "R-path," etc., in an obvious sense. "$x\lambda y$" and "$x\rho y$" shall mean that there is a cell xy in L or R respectively, and xLy and xRy that $x > y$ in L or R. (In diagrams the positive L- and R-cells will slope down towards the left and right respectively.) We denote by Q the following property of Σ.

(Q) *If $x\lambda y$ and $x\rho z$ there exists a vertex w such that zLw, and either $y = w$ or yRw.* (We require zLw, which is not necessarily implied by $z = w$. The possibility that $y = z$ is not excluded.)

THEOREM 4. *If, in a complex with the property* Q, *all L-paths are finite, then*

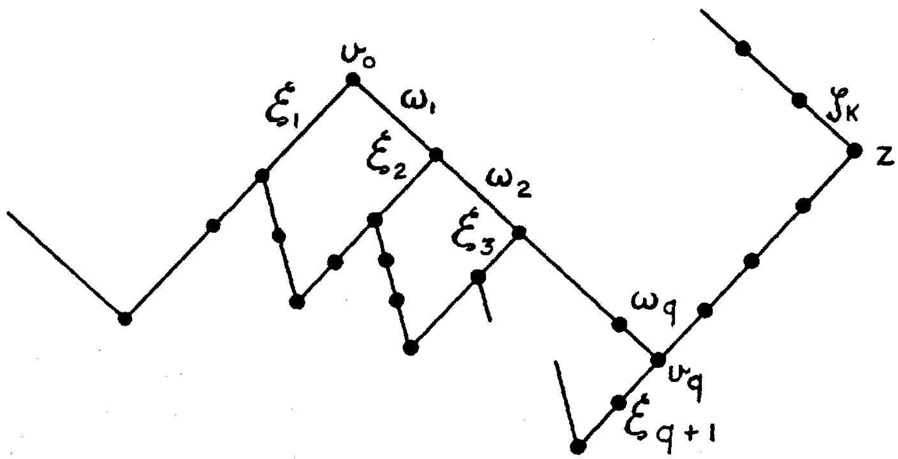

FIG. 2

if xLy and xRz there exists a vertex w such that zLw and either $y = w$ or yRw. If all L-paths are finite, then, in property (Q), $z \neq w$.

It is sufficient to prove the theorem when $x\lambda y$, the general case then following by induction. Let η be an L-cell from x to y, and $[\zeta]_k$ an R-path from x to z. Let π_1 be $-\eta + [\zeta]_k$, and suppose, inductively, that a path π_r from y to z has already been defined.

If π_r has no peak π_{r+1} is not defined; otherwise let v_0 be the last peak on π_r, from y towards z. We assume, inductively, that in proceeding from y towards z the direct ("downward") cells of π_r are in R and the reversed cells in L,—an assumption evidently satisfied for $r = 1$. The part $v_0 \cdots z$ of π_r is of the form $[\omega]_q - \sigma$ where $[\omega]_q$ is a descending R-path and σ a descending L-path. σ may be null, but $q \neq 0$ since v_0 is a peak. Let ξ_1 be the predecessor of ω_1 in π_r,

[5] This always means "indexed subcomplex," the positive direction in each 1-cell agreeing with that in Σ.

(and therefore an L-cell). Assuming inductively that ξ_m is defined, for some $m \leqq q$, as an L-cell with the same initial vertex as ω_m, let σ_m and σ_m' be the R- and L-paths which, by (Q), descend from the final vertices of ξ_m and ω_m to a common vertex. Then σ_m' is not null, and we define ξ_{m+1} to be its first cell: say $\sigma_m' = \xi_{m+1} + \tau_m$. The path π_{r+1} is now defined to be the result of substituting $\sigma_1 - \tau_1 + \sigma_2 - \tau_2 + \cdots + \sigma_q - \sigma_q'$ for $-\xi_1 + [\omega]_q$. It evidently has the property that reversed cells are in L and direct cells in R, and the inductive definition of π_r is therefore completed.

If v_q is the final vertex of ω_q, and $-\sigma_r^*$ is the portion $v_q \cdots z$ of π_r, σ_r^* is a descending L-path from z. The corresponding portion of $-\pi_{r+1}$ is $\sigma_r^* + \sigma_q'$, with at least one more cell. Since all L-paths are finite it follows that the process of constructing paths π_r terminates after a certain number, k, of steps, i.e. π_k has no peaks and is therefore a descending (possibly null) R-path from y to a

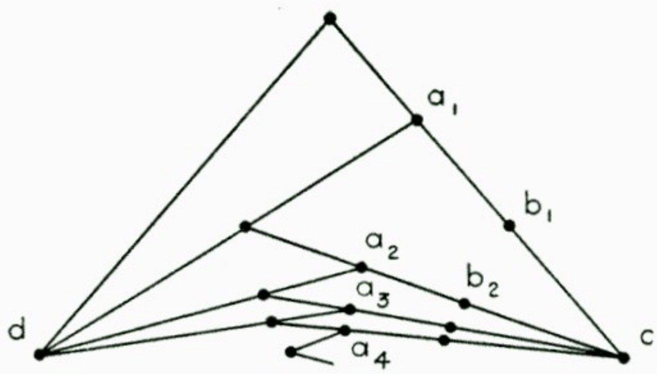

FIG. 3

vertex w, followed by an ascending (non-null) L-path from w to z. Thus yRw or $y = w$, and zLw.

COROLLARY 4.1. *If* (Q) *is strengthened by excluding the possibility* $y = w$, *Theorem* 4 *may be strengthened in the same way.* (Obvious from the method of proof.)

COROLLARY 4.2. *A descending l-path and a descending R-path have at most one common vertex.* If the two paths have their initial and final vertices, a and b in common, i.e. if aLb and aRb, there is a vertex c such that bLc and bRc (the alternative $b = c$ being impossible in this case); and a vertex d such that cLd and cRd; and so on. The path $a \cdots b \cdots c \cdots d \cdots$ is an infinite descending L-path.

In particular a cell cannot be both an L- and an R-cell. This does not mean that the condition (Q) could be weakened in Theorem 4 by adding "if $y \neq z$" at the beginning. That this would make the theorem untrue is shown by the example in Fig. 3, where segments sloping down towards the left and right belong to L and R respectively, and the cells marked $b_i c$ are in both L and R.

The condition (Q) is satisfied for pairs with $y \neq z$, and no descending L-path has more than two cells; but c and d have no lower bound.

Theorem 4 also fails if the alternative "$z = w$" is allowed in (Q). This is seen by omitting the vertices b_i in Fig. 3 so that each $a_i c$ becomes a single R-cell (but not now an L-cell).

8

We return to the consideration of indexed 1-complexes in general. A set of (positive) cells, E_x, is *at* x if x is the initial vertex of every member of E_x. (The null-set is at every vertex.) We suppose that if ξ is a cell xz, each cell η at x has a finite set of cells at z, called $\eta \mid \xi$, assigned to it as its ξ-*derivate*.[6] The ξ-derivate, $E_x \mid \xi$, of a set E_x is the logical sum of the ξ-derivates of members of E_x. (An appearance of the symbol $E_x \mid \xi$ implies that ξ is at x.) If π is a descending path from x, $E_x \mid \pi$, the derivate of E_x by *continuation* along π, is defined inductively by the equation

$$E_x \mid (\pi + \xi) = (E_x \mid \pi) \mid \xi;$$

and if π is null, $E_x \mid \pi = E_x$. We usually write $E \mid (\pi + \xi)$ without brackets: $E \mid \pi + \xi$. The path $[\xi]_m$ is a *development* of E_x if, for $1 \leqq i \leqq m$, $\xi_i \in E_x \mid [\xi]_{i-1}$. The development is *complete* if $E_x \mid [\xi]_m = 0$, *partial* if not.

The letters CD are used as an abbreviation for "complete development." We postulate the following conditions on the derivates:

(Δ_1) $\eta \mid \xi$ *is null if, and only if,* $\eta = \xi$;

(Δ_2) *if* $\eta \neq \zeta$, $(\eta \mid \xi) \cap (\zeta \mid \xi) = 0$;

(Δ_3) *if* η *and* ζ *are distinct cells at* x, *there exist developments* κ_η *and* κ_ζ *of* $\eta \mid \zeta$ *and* $\zeta \mid \eta$ *respectively, with a common final vertex* w.

(Δ_4) *with the notation of* (Δ_3), $\xi \mid (\eta + \kappa_\zeta) = \xi \mid (\zeta + \kappa_\eta)$, *for any* ξ *at* x.

It follows from (Δ_4), by summation, that the derivates of any set E_x by continuation along $\eta + \kappa_\zeta$ and $\zeta + \kappa_\eta$ are the same. A further consequence is that κ_η and κ_ζ are *complete* developments of $\eta \mid \zeta$ and $\zeta \mid \eta$ respectively. For

$$(\eta \mid \zeta) \mid \kappa_\eta = \eta \mid \zeta + \kappa_\eta$$
$$= \eta \mid \eta + \kappa_\zeta = 0.$$

From (Δ_2) it follows by induction on the length of π that if $E_x^1 \cap E_x^2$ is null, $(E_x^1 \mid \pi) \cap (E_x^2 \mid \pi)$ is also null.

LEMMA 1. *If* π *is a development of* E_x^1, *and* $E_x^2 \mid \pi \subseteq E_x^1 \mid \pi$, *then* $E_x^2 \subseteq E_x^1$.

Let π be $[\xi]_m$. Let j be the least integer such that $E_x^2 \mid [\xi]_j \subseteq E_x^1 \mid [\xi]_j$. If the lemma is false $j \geqq 1$, and $E_x^2 \mid [\xi]_{j-1}$ contains a cell ζ not in $E_x^1 \mid [\xi]_{j-1}$. Hence $\zeta \neq \xi_j$, and $\zeta \mid \xi_j$ is a non-null subset of $E_x^2 \mid [\xi]_j$ not contained ln $E_x^1 \mid [\xi]_j$, contrary to the hypothesis.

In particular if $E_x^2 \mid \pi = 0$, $E_x^2 \subseteq E_x^1$.

It is assumed, further, that a relation J holds between certain of the pairs of

[6] For an illustrative example of derivates see §13.

cells at a vertex, and a set E_a is defined to be a *J-set* if $\xi J\eta$ and $\eta J\xi$ for every distinct pair ξ, η in E_a. (Thus all sets with less than two members are J-sets.)

(J_1) *If $\xi J\eta$, $\xi \mid \eta$ has precisely one member.*

(J_2) *If $\eta_1 \epsilon \xi_1 \mid \zeta$ and $\eta_2 \epsilon \xi_2 \mid \zeta$, and if $\xi_1 J\xi_2$ or $\xi_1 = \xi_2$, then $\eta_1 J\eta_2$ or $\eta_1 = \eta_2$.*

It follows from J_2 that if E is a J-set, $E \mid \zeta$, and more generally $E \mid \pi$, is a J-set. From J_1 it follows that for no ξ does $\xi J\xi$.

It is now agreed that a set denoted by E, E_a, etc., shall be finite. (A CD of any set is finite by definition.)

LEMMA 2. *If the J-set E has k members, all CD's of E have k cells and the same final vertex, and all partial developments are parts of CD's.*

If η and ξ are in E, $\eta \mid \xi$ has one member if $\eta \neq \xi$, and none if $\eta = \xi$. Thus $E \mid \xi$ is a J-set with $k - 1$ members, and the development comes to an end after k steps.

Let $\eta + \sigma$ and $\zeta + \tau$ be any two CD's of E, ending at y and z respectively; and let $\eta \neq \xi$. By J_1 and Δ_3 the sets $\eta \mid \zeta$ and $\zeta \mid \eta$ are single cells, η' and ζ', with a common final vertex, w. By Δ_4, $E \mid \eta + \zeta' = E \mid \zeta + \eta'$, a set with $k - 2$ members. Let π be a CD of this set, ending at u. By an inductive hypothesis $\zeta' + \pi$ and σ, being CD's of $E \mid \eta$, a J-set with $k - 1$ members, have the same final vertex: $u = y$. Similarly $u = z$, and so $y = z$.

LEMMA 3. *If E_x is a J-set, and E_x^1 any set at the same vertex x, all derivates of E_x^1 by continuation along CD's of E_x are identical.*

With the notation of the previous lemma,

$$\begin{aligned} E_x^1 \mid \eta + \sigma &= E_x^1 \mid \eta + \zeta' + \pi, && \text{(inductive hypothesis)}, \\ &= E_x^1 \mid \zeta + \eta' + \pi, && (\Delta_4), \\ &= E_x^1 \mid \zeta + \tau, && \text{(inductive hypothesis)}. \end{aligned}$$

9

If E_x^1 is a J-set, $E_x \mid E_x^1$ denotes the continuation of E_x along a CD of E_x^1. By Lemma 3 it is independent of the CD chosen. $E_x \mid E_x^1 + E_x^2 + \cdots + E_x^k$ is defined inductively to be $(E_x \mid E_x^1 + \cdots + E_x^{k-1}) \mid (E_x^k \mid E_x^1 + \cdots + E_x^{k-1})$. Thus if $[\xi]_k$ is a CD of E_x^1, $E_x \mid E_x^1$ and $E_x \mid [\xi]_k$ are alternative notations for the same set.

We now come to the main results of the paper. All the conditions Δ and J are purely local, and involve only a fixed number of given cells.

THEOREM 5. *Let π_1 and π_2 be paths in a 1-complex with the properties J_{1-3} and Δ_1–Δ_4 descending from a to vertices b and c. Then there exist paths π_3 and π_4 descending from b and c to a common vertex d, such that if E_a is a set at a,*

$$E_a \mid \pi_1 + \pi_3 = E_a \mid \pi_2 + \pi_4 .$$

We first prove the following special case.

LEMMA 5.1. *If E_a^1 and E_a^2 are J-sets, the CD's of $E_a^1 \mid E_a^2$ and $E_a^2 \mid E_a^1$ have the same final vertex, and if E_a is any set at a, $E_a \mid E_a^1 + E_a^2 = E_a \mid E_a^2 + E_a^1$.*

(Since all sets called "E", "E^1", etc., in the following proof are at a, the suffix a will be omitted.

[7]CASE 1. $E^1 \cap E^2 = 0$. Let $n(E)$ denote the number of elements in E. We proceed by induction on $m, = n(E^1 \mid E^2) + n(E^2 \mid E^1)$. Excluding the trivial case where one set E^i is null, the minimum possible value of m is 2. This minimum is only attained if $n(E^1) = n(E^2) = 1$, and Lemma 5.1 then follows from Δ_3 and Δ_4. We may therefore assume that $m > 2$, and also that $n(E^1) > 1$ or $n(E^2) > 1$, —say $E^1 = E^3 \cup \eta$, where $E^3 \neq 0$.

The proof depends on the fact that if ξ is not in E, $n(E \mid \xi) \geqq n(E)$; and hence if E^p and E^q are J-sets satisfying $E^p \cap E^q = 0$, and $E^r \subseteq E^q$, then $n(E^p \mid E^r) \leqq n(E^p \mid E^q)$. Thus $n(E^2 \mid E^3) \leqq n(E^2 \mid E^1)$ and $n(E^3 \mid E^2) < n(E^1 \mid E^2)$. Therefore, by the inductive hypothesis, CD's of $E^2 \mid E^3$ and $E^3 \mid E^2$ have the same final vertex, z. Since E^1 is a J-set, $\eta \mid E^3$ is a single cell, ξ, and

$$\xi \mid (E^2 \mid E^3) = \eta \mid E^3 + E^2 = \eta \mid E^2 + E^3 \qquad \text{(by the inductive hypothesis)}$$
$$= E^1 \mid E^2 + E^3,$$

since $E^3 \mid E^2 + E^3 = E^3 \mid E^3 + E^2 = 0$. Thus there is a CD of $E^1 \mid E^2$ consisting of a CD of $E^3 \mid E^2$ followed by a CD of $\xi \mid (E^2 \mid E^3)$. Since $E^3 \mid E^2$ is not null it follows that $n(\xi \mid (E^2 \mid E^3)) < n(E^1 \mid E^2)$, and since also $(E^2 \mid E^3) \mid \xi = E^2 \mid E^1$ the inductive hypothesis may be applied to the sets ξ and $E^2 \mid E^3$ at z. The final vertices of CD's of $(E^2 \mid E^3) \mid \xi$, i.e. $E^2 \mid E^1$, and of $\xi \mid (E^2 \mid E^3)$ are therefore identical, and the latter set has been seen to be the end portion of a CD of $E^2 \mid E^1$. The first part of the induction is therefore complete. If E is any set at a,

$$\begin{aligned} E \mid E^2 + E^1 &= E \mid E^2 + E^3 + \eta, \\ &= E \mid E^3 + E^2 + \eta, \quad \text{(by the inductive hypothesis applied to } E^2 \text{ and } E^3) \\ &= E \mid E^3 + \eta + E^2, \quad \text{(by the inductive hypothesis applied to } \xi \text{ and } E^2 \mid E^3) \\ &= E \mid E^1 + E^2. \end{aligned}$$

CASE 2. As CASE 1 save that $E^1 \cap E^2 \neq 0$. Let $E^i = E^0 \cup E^{i+2}$, for $i = 1, 2$, where $E^3 \cap E^4 = 0$. By CASE 1, applied to $E^4 \mid E^0$ and $E^3 \mid E^0$, $E^4 \mid E^0 + E^3$ and $E^3 \mid E^0 + E^4$ have the same final vertex, and since $E^0 \mid E^0 = 0$ these two sets are $E^2 \mid E^0 + E^3$ and $E^1 \mid E^0 + E^4$, i.e. $E^2 \mid E^1$ and $E^1 \mid E^2$.

In the *general case*, to which we now turn, the result may be stated more explicitly as follows, taking π_1 and π_2 to be $[\eta]_j$ and $[\zeta]_k$.

LEMMA 5.2. *If $[\eta]_j$ and $[\zeta]_k$ are any paths descending from a, to b and c there exist paths σ_{rs} and τ_{rs}, (possibly null, $r = 1, \cdots, j+1, s = 1, \cdots, k+1$) such that*

(1) $\eta_s = \sigma_{1s}$, $\zeta_r = \tau_{r1}$,

(2) $\sigma_{r+1,s}$ *and* $\tau_{r,s+1}$ *have the same final vertex,*

[7] This proof of Case 1 was suggested by Dr. J. H. C. Whitehead, in place of one based on Theorem 4.

(3) σ_{rs} *is a* CD *of* E^1_{rs}, $= \eta_s \mid \tau_{1s} + \tau_{2s} + \cdots + \tau_{r-1,s}$, *and* τ_{rs} *of* E^2_{rs}, $= \zeta_r \mid \sigma_{r1} + \cdots + \sigma_{r,s-1}$.

(4) *for any* E_a, $E_a \mid \pi_1 + \tau_{1,k+1} + \cdots + \tau_{j,k+1} = E_a \mid \pi_2 + \sigma_{j+1,1} + \cdots + \sigma_{j+1,k}$.

Starting from the two given paths $[\eta]_j$ and $[\zeta]_k$ we add, one by one, the pairs of paths $\sigma_{r+1,s}$ and $\tau_{r,s+1}$ for the couples (r, s) in the standard "triangular" order (1, 1), (1, 2), (2, 1), (1, 3), $\cdots$. When the time comes for $\sigma_{r+1,s}$ and $\tau_{r,s+1}$ to be added, the paths σ_{rs} and τ_{rs}, descending from a vertex x_{rs}, and corresponding to the earlier couples $(r, s - 1)$ and $(r - 1, s)$, have already been constructed as CD's of E^1_{rs} and E^2_{rs}. Hence by the cases of Theorem 5 already settled, CD's $E^1_{rs} \mid \tau_{rs}$ and $E^2_{rs} \mid \sigma_{rs}$, i.e. of $E^1_{r,s+1}$ and $E^2_{r+1,s}$, meet at a common vertex. These CD's are $\tau_{r,s+1}$ and $\sigma_{r+1,s}$; the induction is complete. (In the limiting cases $r = 1$ and $s = 1$ the single cells η_r and ζ_s play the parts of E^1_a and E^2_a in the earlier cases.)

The proof just given provides a method of deforming $\pi_1 + \pi_3$ into $\pi_2 + \pi_4$ by a series of steps in each of which a path $\tau_{rs} + \sigma_{r+1,s}$ is replaced by a path $\sigma_{rs} + \tau_{r,s+1}$. By Lemma 5.1 a set E_x at the common initial vertex of σ_{rs} and τ_{rs}, when continued along either of these paths gives the same result, and therefore the continuation of E_a along the whole path is unaffected by a single step.

THEOREM 6. *Any two* CD's *of a (finite) set* E_x *have the same final vertex.*

If $[\eta]_j$ and $[\zeta]_k$, ending in b and c, are the developments then, with the notations of Lemma 5.2, since $\eta_s \in E_x \mid [\eta]_{s-1}$,

$$E^1_{rs} \subseteq E_x \mid [\eta]_{s-1} + \tau_{1s} + \cdots + \tau_{r-1,s},$$

$$= E_x \mid [\zeta]_{r-1} + \sigma_{r1} + \cdots + \sigma_{r,s-1},$$

and therefore $\sigma_{r1} + \sigma_{r2} + \cdots$ is a development of $E_x \mid [\zeta]_{r-1}$; and in particular $\sigma_{k+1,1} + \sigma_{k+1,2} + \cdots$ is a development of $E_x \mid [\zeta]_k$, $= 0$, since $[\zeta]_k$ is a CD. Thus $c = d$, and similarly $b = d$.

COROLLARY 6.1. *Continuation of a set* E_a *along any two* CD's *of a set* E^1_a *gives the same result.* This now follows from Theorem 5, π_3 and π_4 being null.

Corollary 6.1 cannot be extended to give the general monodromy property, "continuation of E_a along any two descending paths from a to b gives the same result." Consider the 1-complex in Fig. 5, in which the vertices marked x are identical. Derivates are defined by parallel displacement downward, except that the derivates of xy and xz at z and y are zw and yw respectively. All sets are J-sets. The conditions Δ and J are satisfied in this complex, but continuation of ab to x *via* b gives the null set, *via* c the cell xz.

THEOREM 7. *In a complex satisfying,* J *and* Δ, *all developments of a finite set* E_x *are finite.*

Every set is a sum of J-sets, namely its individual members. We proceed by induction on the smallest number, k, of J-sets, E^i_x, whose sum is the given set E_x. (The case $k = 1$ is Lemma 2.)

There is at least one CD of E_x, namely $[\sigma]_k$, where σ_r is a CD of the J-set $E_x^r \mid [\sigma]_{r-1}$. Suppose that $\zeta_1 + \zeta_2 + \cdots$ is an infinite development of E_x, and let σ_{rs}, τ_{rs}, E_{rs}^1 and E_{rs}^2 be as in Lemma 5.1, save that σ_r replaces η_r. Then just as in Theorem 6, $\tau_{1s} + \tau_{2s} + \cdots$ is a development of $E_x \mid [\sigma]_{s-1}$. Since, for $i < k$, E_x^i is annihilated by continuation along σ_i, $E_x \mid [\sigma]_{k-1} = E_x^k \mid [\sigma]_{k-1}$.

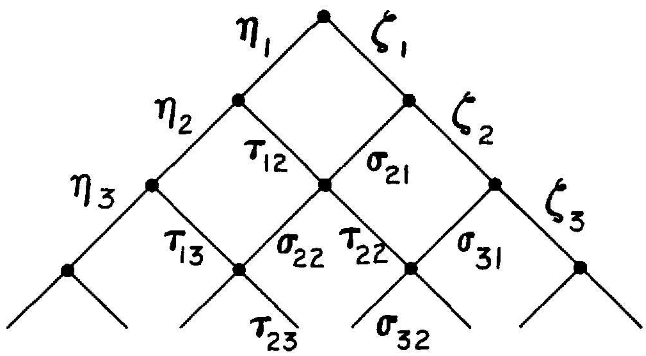

FIG. 4

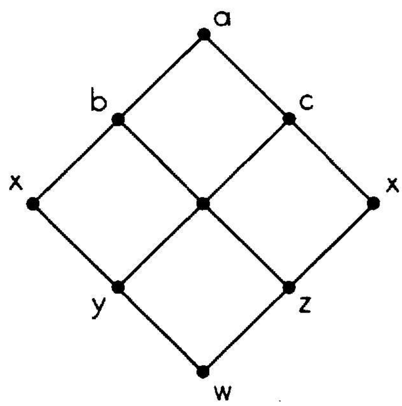

FIG. 5

Thus $\tau_{1k} + \tau_{2k} + \cdots$ is a development of a J-set, and so $\tau_{rk} = 0$ if r exceeds a certain q. Therefore if $r > q$,

$$0 = E_{rk}^2 = \zeta_r \mid \sigma_{r1} + \cdots + \sigma_{r,k-1}.$$

Now $\sigma_{r1} + \cdots + \sigma_{r,k-1}$ is a development of H_r, $= (E_x^1 \cup E_x^2 \cup \cdots \cup E_x^{k-1}) \mid [\zeta]_{r-1}$, and hence by Lemma 1, $\zeta_r \in H_r$. Therefore the infinite path $\zeta_{q+1} + \zeta_{q+2} + \cdots$ is a development of H_{q+1}, a sum of $(k - 1)$ J-sets,—contrary to the inductive hypothesis.

COROLLARY 7.1. *There are only a finite number of different developments of* E_x.

If there are an infinity, some cell ξ_1 of the finite set E_x must come first in an infinity of developments; and some cell ξ_2 of the finite set $E_x \mid \xi_1$ must be second in an infinity of these developments; and so on. The path $\xi_1 + \xi_2 + \cdots$ is an infinite development of E_x, contrary to Theorem 7.

10

Theorem 7 is connected with the problem of "random reduction". To a "normal form" or "end-form" in a system with moves there corresponds an *end* of Σ, and to a normal form *of* X an end connected by a path to a given vertex x. It follows from Theorem 5 that there is a descending path from x to the end, and that a vertex cannot be connected to two ends, i.e. that an "object" in the corresponding system cannot have two different normal forms. There remains, however, the possibility of an infinite descending path from a vertex which is also connected to an end. It will now be shown that this possibility is not realised in complexes satisfying the conditions Δ and J.

THEOREM 8. *If, in a complex satisfying the conditions* Δ *and* J, *there is a path descending from* x *to an end* e *of* Σ, *all descending paths from* x *are finite, and all maximal paths end at* e.

That all maximal descending paths from x end at e is obvious in view of Theorem 5; only the finiteness remains to be proved.

Let $[\eta]_m$ be a descending path from x to e, and (if possible) $\zeta_1 + \zeta_2 + \cdots$ an infinite descending path from x. Let the paths σ_{rs} and τ_{rs}, and the sets E^1_{rs} and E^2_{rs}, be constructed as in Lemma 5.2. Since e is an end all the $\tau_{r,m+1}$ are null. Let j be the largest number such that τ_{rj} is non-null for an infinity of values of r, and k a number such that $\tau_{r,j+1} = 0$ if $r \geqq k$. Then $E^2_{r,j+1}$, of which $\tau_{r,j+1}$ is a CD, is also null, giving $E^2_{rj} \mid \sigma_{rj} = E^2_{r,j+1} = 0$. Since σ_{rj} is a CD of E^1_{rj} it follows (Lemma 1) that, for $r \geqq k$,

$$E^2_{rj} \subseteq E^1_{rj} = E^1_{kj} \mid \tau_{kj} + \cdots + \tau_{r-1,j}.$$

Thus $\tau_{kj} + \tau_{k+1,j} + \cdots$ is a development of E^1_{kj}, and by Theorem 7 cannot be infinite,—contrary to the definition of j.

It follows that if a 2-complex Σ^2 is constructed as in §5, all its components containing ends of Σ are simply connected.

11

The theorems that have been proved indicate that complications will arise when the descending paths that join the "y" and "z" of condition (D) to "w" have either more or less than one member each, and that the difficulties are of a different kind in the two cases. In the foregoing group of theorems the second possibility, (corresponding to $\xi \mid \eta = 0$ for $\xi \neq \eta$), was excluded. The following theorem allows this possibility, but is in other ways more special than Theorem 5, and the meaning of the conditions imposed is less obvious. The theorem is used in extending the Church-Rosser Theorem to an enlarged calculus.

We suppose that derivates are defined in Σ, and satisfy Δ_2–Δ_4, but that Δ_1 holds only in the weakened form

(Δ_1^*) $\xi \mid \xi = 0$, and if $\xi \mid \eta = 0$ then $\eta A \xi$, where $\eta A \xi$ stands for "either $\eta = \xi$ or $\eta \mid \xi$ has just one member". The following additional "J-condition" is imposed, $\bar{A}$ denoting "not A":

(J_3) *if $\xi \bar{A} \eta$ and $\xi J \zeta$, then $\eta \bar{J} \zeta$.*

(Condition J_3 does not imply the second half of Δ_1^*, since $\xi \bar{J} \xi$.)

Lemmas 2 and 3 remain true under these conditions, and are proved as before. The notation $E_a \mid E_a^1 + \cdots + E_a^k$ may therefore be introduced for J-sets E_a^i.

THEOREM 9. *A complex with the properties Δ_1^*, Δ_2–Δ_4 and J_1–J_3 has the property* (A).

It is sufficient to prove the following special case, since the extension to the general case then proceeds exactly as in Theorem 5.

LEMMA 9.1. *If E_a^1 and E_a^2 are J-sets, and E_a is any set at a, the* CD's *of $E_a^1 \mid E_a^2$ and $E_a^2 \mid E_a^1$ have the same final vertex, and $E_a \mid E_a^1 + E_a^2 = E_a \mid E_a^2 + E_a^1$.*

Let $[\eta]_k$ and $[\zeta]_k$ be CD's of E_a^1 and E_a^2, η_i and ζ_i having final vertices b_i and c_i. "$E_a^1 J E_a^2$" *means* "$\eta J \zeta$ if $\eta \in E_a^1$ and $\zeta \in E_a^2$." From J_2 it follows that if $E_a^1 J E_a^2$, $(E_a^1 \mid \xi) J (E_a^2 \mid \xi)$.

CASE 1: $E_a^1 J E_a^2$. We show further that *in this case $E_a^1 \mid E_a^2$ has j cells.* First let $j = 1$. If also $k = 1$ the result follows immediately from J_1 and Δ_3. For

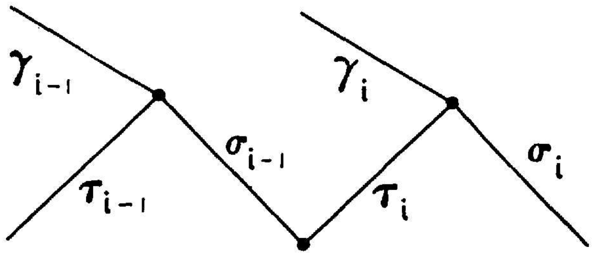

FIG. 6

general k, we have $(\eta_1 \mid \zeta_1) J (E_a^2 \mid \zeta_1)$; and since $\eta_1 \mid \zeta_1$ is a single cell, and $E_a^2 \mid \zeta_1$ has $k - 1$ cells, an inductive hypothesis shows that $\eta_1 \mid E_a^2$ is a single cell and has the same final vertex as a CD, π', of $(E_a^2 \mid \zeta_1) \mid (\eta_1 \mid \zeta_1)$, which by Δ_4 is $E_a^2 \mid \eta_1 + \zeta_1$. Hence π' is a CD of $E_a^2 \mid \eta_1$. That $E_a \mid E_a^1 + E_a^2 = E_a \mid E_a^2 + E_a^1$ is proved, as in Theorem 5, by repeated applications of Δ_4. Case 1 for general j is now completed by applying the case $j = 1$ successively to η_r and $E_a^2 \mid [\eta]_{r-1}$, for $r = 1, 2, \cdots, j$, and using the last part of the result for $r - 1$.

CASE 2: $E_a^1 \mid E_a^2 = 0$. We show further that, *in this case, $E_a^2 \mid E_a^1$ has k members or less.* If $j = k = 1$ the result is clear from J and Δ. Suppose that $j = 1$, $k > 1$. Then $\zeta_1 A \eta_1$, for if $\zeta_1 \bar{A} \eta_1$, by J_3 and J_2 $(\eta_1 \mid \zeta_1) J (E_a^2 \mid \zeta_1)$, and hence by Case 1 $\eta_1 \mid E_a^2 \neq 0$, contrary to the hypothesis. Thus $\zeta_1 \mid \eta_1$ is a single cell. By a k-induction applied to $\eta_1 \mid \zeta_1$ and $E_a^2 \mid \zeta_1$, (in place of E_a^1 and E_a^2), all CD's of $E_a^2 \mid \zeta_1 + \eta_1$ have $k - 1$ cells or less, and end at c_k. Since this set is also $E_a^2 \mid \eta_1 + \zeta_1$, the CD of $E_a^2 \mid \eta_1$ is the cell $\zeta_1 \mid \eta_1$, followed by the $k - 1$ cells, (or less), of $E_a^2 \mid \zeta_1 + \eta_1$. The final part follows by repeated applications of Δ_4.

The extension to general j is as in Case 1.

General Case. In view of Lemma 3 it may be assumed that if $E_a^1 \mid E_a^2 \neq 0$, $[\eta]_j$ is chosen so that $\eta_1 \mid E_a^2 \neq 0$. A series of "zig-zag" paths, $\pi_1, \pi_2, \cdots$, from b_j to c_k, is constructed, π_1 being $-[\eta]_j + [\zeta]_k$. Suppose π_r already constructed,

$$\pi_r = \tau_0 - \sigma_1 + \tau_1 - \cdots + \tau_{m-1} - \sigma_m,$$

where σ_i and τ_i are CD's of subsets, E_i^1 and E_i^2, of $E_a^1 \mid \gamma_i$ and $E_a^2 \mid \nu_i$ respectively. The γ_i are descending paths from a satisfying

(i) γ_m is $[\zeta]_k$

(ii) for any E_a at a, $E_a \mid \gamma_i + \sigma_i = E_a \mid \gamma_{i-1} + \tau_{i-1}$.

This whole inductive hypothesis is satisfied by π_1 if $m = 2$, $\tau_0 = \sigma_2 = \gamma_1 = 0$, $\sigma_1 = \gamma_0 = [\eta]_j$, $\tau_1 = \gamma_2 = [\zeta]_k$.

Let π_r have a peak at the join of σ_{m-1} and τ_{m-1}, and let u be the final vertex of τ_{m-1}. First suppose that $E_{m-1}^1 \mid \tau_{m-1} = 0$. By Case 2 a CD, θ, of $E_{m-1}^2 \mid \sigma_{m-1}$ ends at u, and we define π_{r+1} to be $\tau_0 - \sigma_1 + \cdots + \tau_{m-2} + \theta - \sigma_m$, the new m', τ'_{m-2} and σ'_{m-1} being $m - 1$, $\tau_{m-2} + \theta$ and σ_m. The γ'_i up to γ'_{m-2} are the corresponding γ_i, and $\gamma'_{m-1} = \gamma_m$. Since θ is a CD of

$$E_{m-1}^2 \mid \sigma_{m-1} \subseteq E_a^2 \mid \gamma_{m-1} + \sigma_{m-1}$$
$$= E_a^2 \mid \gamma_{m-2} + \tau_{m-2},$$

$\tau_{m-2} + \theta$ is a CD of a subset of $E_a^2 \mid \gamma_{m-2}$. For any E_a at a,

$$\begin{aligned} E_a \mid \gamma'_{m-1} + \sigma'_{m-1} &= E_a \mid \gamma_m + \sigma_m \\ &= E_a \mid \gamma_{m-1} + \tau_{m-1} \\ &= E_a \mid \gamma_{m-1} + \sigma_{m-1} + \theta \qquad \text{(Case 2)} \\ &= E_a \mid \gamma_{m-2} + \tau'_{m-2}. \end{aligned}$$

Secondly let $E'_{m-1} \mid \tau_{m-1} \neq 0$. By Lemmas 2 and 3, if a CD of E'_{m-1}, whose first cell, $\xi^{(1)}$, satisfies $\xi^{(1)} \mid \tau_{m-1} \neq 0$, is substituted for σ_{m-1}, all the conditions imposed on π_r remain satisfied, and it may therefore be assumed that σ_{m-1} itself is such a CD. Let τ_{m-1} be $[\omega]_p$, ($p \neq 0$ in view of the peak). We construct successively, as in Theorem 5, for $i = 1, 2, \cdots, p$, pairs of descending paths τ'_{m-2+i} and $\xi^{(i+1)} + \sigma'_{m-1+i}$, which are CD's of $\omega_i \mid \xi^{(i)}$ and $\xi^{(i)} \mid \omega_i$ respectively, and, by Δ_3, have a common final vertex. The notation "$\xi^{(i+1)} + \sigma'_{m-1+i}$" implies that $\xi^{(i)} \mid \omega_i \neq 0$, which is justified by $\xi^{(i-1)} \mid \omega_{i-1} \neq 0$, derived ultimately from $\xi^{(1)} \mid \tau_{m-1} \neq 0$. If σ_{m-1} is $\xi^{(1)} + \sigma'_{m-1}$, π_{r+1} is defined to be

$$\tau_0 - \sigma_1 + \cdots + \tau_{m-2} - \sigma'_{m-1} + \tau'_{m-1} - \cdots + \tau'_{m-2+p} - \sigma'_{m-1+p} - \xi^{(p+1)} - \sigma_m.$$

Its final ascending part has at least one more cell than σ_m. If γ'_h is taken to be γ_h for h up to $m - 2$, $\gamma_{m-1} + [\omega]_{h-m+1} + \xi^{(h-m+2)}$ for $h = m - 1$ to $m + p - 2$, and $\gamma'_{m+p-1} = \gamma_m$, all the conditions are fulfilled, the new "m" being $m + p - 1$, the new "σ_m" $\sigma_m + \xi^{(p+1)} + \sigma'_{m+p-1}$. Condition (ii) follows for the γ'_i immediately from Δ_4 except for $i = m + p - 1$; and

$$E_a \mid \gamma_m + \sigma_m + \xi^{(p+1)} + \sigma'_{m+p-1} = E_a \mid \gamma_{m-1} + \tau_{m-1} + \xi^{(p+1)} + \sigma'_{m+p-1}$$

$$= E_a \mid \gamma_{m-1} + [\omega]_{p-1} + \xi^{(p)} + \tau'_{m+p-2} \qquad (\Delta_4)$$

$$= E_a \mid \gamma'_{m+p-2} + \tau'_{m+p-2},$$

as required.

It has thus been shown how, in all cases where π_r has a peak, a path π_{r+1} is to be constructed having either one less peak or a longer final ascending portion. Since this portion is a development of the finite J-set $E'_a \mid [\zeta]_k$, the second alternative can only occur a finite number of times. Thereafter the number of peaks decreases at each step until a path with no peaks is reached.

The extension to paths which are not CD's of J-sets now follows as in Theorem 5.

12

The theorems that follow Theorem 5, as far as Corollary 6.1, are proved under the new conditions with only minor changes in the argument. Theorem 8 fails to survive, as may be seen by considering Fig. 1: all the conditions Δ_1^*, Δ_2–Δ_4, and J_1–J_3 are satisfied by taking the derivate of a_rb at a_{r+1} to be $a_{r+1}b$, and that of a_ra_{r+1} at b to be null; and $a_rbJa_ra_{r+1}$ but not $a_ra_{r+1}Ja_rb$.

Theorem 7 still holds but its proof needs some modification.

THEOREM 10. *In a complex satisfying* Δ_1^*, Δ_2–Δ_4 and J_1–J_3 *all developments of finite sets are finite.*

We proceed as in the proof of Theorem 7. As before it follows that $\tau_{1s} + \tau_{2s} + \cdots$ is a development of $E_x \mid [\sigma]_{s-1}$ and that for some p and q the development $\tau_{p+1,q} + \tau_{p+2,q} + \cdots$ of $E_x \mid [\sigma]_{q-1} + \tau_{1q} + \cdots + \tau_{pq}$, $= E_x \mid \theta_{pq}$ say, is infinite, while $\tau_{r,q+1} = 0$ if $r > p$. Since the σ's and τ's are CD's of J-sets it follows from Case 2 of Lemma 9.1 that the number of cells in σ_{rq} is (for $r > p$) non-increasing with increasing r. If ζ_r belongs to $E_x^q \mid [\zeta]_{r-1}$, τ_{rq} is contained in $E_x^q \mid [\zeta]_{r-1} + \sigma_{r1} + \cdots + \sigma_{rq} = E_x^q \mid \theta_{r-1,q}$; by the last part of Lemma 9.1 the path σ_{rq} is a CD of this set, and may therefore be chosen in the form $\tau_{rq} + \sigma'_{rq}$, and $\sigma_{r+1,q}$ to be σ'_{rq}. Thus $\sigma_{r+1,q}$ has less cells than σ_{rq} unless $\tau_{rq} = 0$. If, for $r > r_0$, $\tau_{rq} = 0$ whenever $\zeta_r \in E_x^q \mid [\zeta]_{r-1}$, $\tau_{r_0+1,q} + \cdots$ is an infinite development of the sum of the $k - 1$ J-sets $E_x^i \mid \theta_{r_0q}$ $(i \neq q)$, contrary to the inductive hypothesis. Hence the number of cells in σ_{rq} eventually diminishes to zero, and from this point on the τ_{rq} coincide with the $\tau_{r,q+1}$ and are null,—contrary to the initial hypothesis.

COROLLARY 10.1. *Under the same conditions, the number of different developments of* E_x *is finite.* (Compare Corollary 7.1).

13. Application to the conversion calculus

The formalism first considered is that of Theorems 1 and 2 of Church and Rosser [1], but modified in two ways,—first by the adoption of the simpler bracketing of Church [1] secondly by the entire exclusion of "singular" formulae, i.e. those having "accidental" coincidences between the bound variables.[8] The WFF's are therefore rows of the symbols λ, variables, and round brackets, built up according to the following rules: (1) **x** is a WFF, (2) if **M** is a WFF containing **x** as a free variable, (λ**xM**) is a WFF, (3) if **A** and **B** are WFF's whose common variables are free in both, (**AB**) is a WFF. (A variable is *bound* in any row of symbols in which one of its occurrences immediately succeeds a λ, otherwise *free*.)

The allowed transformations that concern us are

I. To replace each specimen of a bound variable **x** in **X** by **y**, a letter not occurring in **X**.

The result, **Y**, of any series of applications of I to **X** will be called an *adjusted copy* of **X**, and **X** conv.-I **Y**.

II. To replace a part ((λ*x***M**)**N**) of **X** by the result of substituting adjusted copies of **N** for the specimens of **x** in **M**, the new bound variables being all different from each other and those of **X**.

It is agreed that a WFF denoted by one of the letters **U**, **V**, is of the form ((λ**xM**)**N**), and we accordingly speak of "the move **U** on **X**," or "the move (**X**, **U**)," if **U** is a part[9] of **X**, meaning the application of Rule II in which **U** is the part operated on. If **Y** is the WFF that thus replaces **X**, we write (**X**, **U**) →**Y** and **X** conv.-II **Y**. If a series of moves I that turns **X** into **Y** turns its part **U** into **V**, (**Y**, **V**) is an adjusted copy of (**X**, **U**).

To define the *residuals* of a part **V** of **X** after the move ((λ**xM**)**N**), suppose that each pair of brackets in **M** is provided with a numerical suffix, which is left unchanged in applying rule II, and that **V** is enclosed in the pair ${}_1(\ )$. If the move ((λ**xM**)**N**) turns **X** to **Y**, then

(a) if **V** = ((λ**xM**)**N**), **V** has no residual in **Y**;

(b) if **V** is a part of **N** its residuals are the corresponding parts of the adjusted copies of **N** that replace **x** in **M**;

(c) in all other cases the residual of **V** is the part ${}_1(\ )$ of **Y**.

The complex Σ to which our general theorems will be applied has as a typical vertex the class [**X**] of all adjusted copies of a WFF **X**. A positive 1-cell is the class [(**X**, **U**)], or briefly [**X**, **U**], consisting of (**X**, **U**) and all its adjusted copies; and its initial and final vertices are [**X**] and [**Y**], where (**X**, **U**) →**Y**. If **V** is also a part of **X**, the [**X**, **U**]-derivate of [**X**, **V**] consists of all the cells [**Y**, $\mathbf{V}_i$], where the $\mathbf{V}_i$ are the residuals of **V** in **Y**. Finally "[**X**, **U**]*J*[**X**, **V**]" means that (i) neither

[8] Cf. Newman [2] §3. After the general theoretical work the calculus may be extended, for practical convenience, by re-admitting the singular formulae and resuming the original rules I and II; and it can be shewn without difficulty that (1) every singular WFF **X** conv.-I a non-singular **X′**, and (2) if **X** conv.-II **Y**, **X** conv.-I **X′**, **Y** conv.-I **Y′** in the extended calculus, **X′** and **Y′** being non-singular, then **X′** conv.-I-II **Y′** in the restricted calculus.

[9] Defined as in Church [1].

U nor **V** is a part of the other, and (ii) the free variables of **U** and **V** are the same. Thus J is a symmetrical relation, and is independent of the WFF chosen to represent [**X**].

With these definitions the conditions J_1, J_2 *and* Δ_1–Δ_4 *are satisfied.* (J_1): no comment is necessary. (J_2): let η_1 and η_2 be determined by the parts $\mathbf{V}_1$ and $\mathbf{V}_2$ of **X**, and ξ by $\mathbf{U}, = ((\lambda \mathbf{x}\mathbf{M})\mathbf{N})$. If $\mathbf{V}_1 = \mathbf{V}_2$, distinct members of $\eta_1 \mid \xi$ are determined by different adjusted copies of $\mathbf{V}_1$, and evidently satisfy (i) and (ii). If $\mathbf{V}_1$ and $\mathbf{V}_2$ are not identical they are mutually exterior, and a residual of $\mathbf{V}_1$ could only be a part of a residual of $\mathbf{V}_2$ if $\mathbf{V}_1$ were a part of **N**, and $\mathbf{V}_2$ a part of **M** containing **x**. This contravenes the condition (ii) for $\mathbf{V}_1$ and $\mathbf{V}_2$ since **x** is free in **M** and cannot occur in **N**. A part of **X** and its residuals in **Y** have the same free variables, except that **x** is replaced at all occurrences by adjusted copies of **N**. Hence if $\mathbf{V}_1$ and $\mathbf{V}_2$ have the same free variables their residuals in **Y** have also.

In considering the conditions Δ, let ξ, η, ζ be determined by the moves $\mathbf{U}_i$ on **X**, $(i = 1, 2, 3)$, where $\mathbf{U}_i = {}_i((\lambda \mathbf{x}_i \mathbf{M}_i)\mathbf{N}_i)$, and $(\mathbf{X}, \mathbf{U}_i) \to \mathbf{Y}_i$. Thus $\mathbf{U}_i$ is the part ${}_i(\ )$ of **X**.

Δ_1 : no comment is necessary. $\Delta 2$: suppose that $\mathbf{U}_2 \neq \mathbf{U}_3$. The residuals are clearly distinct if they are determined by the original brackets, or one by old and the other by new brackets. The remaining possibility is that a residual of $\mathbf{U}_2$ is a part, $\mathbf{U}_2'$ of an adjusted copy of $\mathbf{N}_1$, and a residual of $\mathbf{U}_3$ is either a different part of the same copy, or part of a different copy,—in any case different from $\mathbf{U}_2'$. Δ_3 : the condition is obviously satisfied unless one of $\mathbf{U}_2$, $\mathbf{U}_3$ is part of the other,—say $\mathbf{U}_2$ of $\mathbf{U}_3$. If $\mathbf{U}_2$ is in $\mathbf{M}_3$ the residual of $\mathbf{U}_2$ in $\mathbf{Y}_3$, and of $\mathbf{U}_3$ in $\mathbf{Y}_2$, are determined by their original brackets, and since $\mathbf{N}_3$ contains no copy of $\mathbf{x}_2$ (a bound variable of $\mathbf{M}_3$), the order of performance of ${}_2(\ )$ and ${}_3(\ )$ is indifferent. If $\mathbf{U}_2$ is in $\mathbf{N}_3$ the final effect is the same whether ${}_2(\ )$ is performed first on $\mathbf{N}_3$, followed by ${}_3(\ )$, or the residuals of ${}_2(\ )$ on the adjusted copies of $\mathbf{N}_3$ in $\mathbf{Y}_3$. Δ_4 : Let **W** be the final result of either series of moves on **x**: it has been shown to be unique to within I-adjustment, and therefore determines a unique vertex, w, of Σ. If $\mathbf{U}_1$ (or $\mathbf{U}_2$) is not part of either of the other $\mathbf{U}_i$'s, its performance, before or after $\mathbf{U}_2$ (or $\mathbf{U}_1$), does not affect the residual of $\mathbf{U}_3$. We may therefore suppose that one of the $\mathbf{U}_i$'s contains the other two. If $\mathbf{U}_3$ is not part of either $\mathbf{N}_1$ or $\mathbf{N}_2$ the residual of $\mathbf{U}_3$ in **W** by either route is ${}_3(\ )$. We therefore assume that $\mathbf{U}_1$ contains both $\mathbf{U}_2$ and $\mathbf{U}_3$, and that $\mathbf{U}_3$ is part of either $\mathbf{N}_1$ or $\mathbf{N}_2$. Finally, if both $\mathbf{U}_2$ and $\mathbf{U}_3$ are in $\mathbf{N}_1$, the same residuals of $\mathbf{U}_3$ are evidently obtained whether $\mathbf{U}_2$ is performed on $\mathbf{N}_1$ before $\mathbf{U}_1$, or the corresponding moves on the copies of $\mathbf{N}_1$ after $\mathbf{U}_1$. There remains only the case where $\mathbf{U}_2$ is part of $\mathbf{M}_1$, and $\mathbf{U}_3$ of either, (α), $\mathbf{N}_1$, or, (β), $\mathbf{N}_2$. (α): the residuals of $\mathbf{U}_3$ by either route are the corresponding parts of the adjusted copies of $\mathbf{N}_1$ that replace $\mathbf{x}_1$ in the move ${}_1(\ )$ on $\mathbf{Y}_2$. (β): if the residuals of $\mathbf{U}_3$ in $\mathbf{Y}_2$ are the parts enclosed in the brackets ${}_{31}(\ )$, ${}_{32}(\ )$, $\cdots$, the residuals in **W** by either route are the parts enclosed in the same brackets.

The conditions for all the Theorems 5 to 8 are therefore satisfied, and we obtain the following results.

Corollary 11.1. *If* **X** conv.-I–II **Y** *and* **X** *conv.* I–II **Z**, *there is a* WFF **W** *su that* **Y** conv. I–II **W** *and* **Z** conv. I–II **W**.

Corollary 11.2. *There are only a finite number of different developments of given set of moves* II *on a* WFF **X**. *All of them are finite, and all end in adjust copies of the same* WFF.

Corollary 11.3. *A* WFF *has (apart from* I*-adjustments) at most one norm form, and if one exists all series of moves* II *terminate in this normal form or a adjusted copy.*

14

Two generalizations of these theorems were given by Church and Rosser in their paper. The first, to the formalism extended so as to include the δ-symbol, is of no interest in the present connection: it is easily shown that the original conditions Δ and J are still satisfied, and hence that the Corollaries 11 hold. The second generalization (of which the proof was not given by Church and Rosser) is to the formalism in which (λ**xM**) is counted a WFF even if **x** does not occur in **M**. The rules of procedure, and the definitions of derivates need no modification, and the conditions Δ_2–Δ_4 are proved to hold, just as before. The second part ("only if") of condition Δ_1 now fails, but the condition Δ_1^* is satisfied. The conditions J_1–J_3 are also satisfied if a different, more complicated, interpretation is given to J.

Let "**U** *S* **V**" stand for "a free variable of **U** is bound in **V**." It implies that **U** is a proper part of **V**, and if **U**′ and **V**′ are, for any **W**, **W**-residuals of **U** and **V**, **U**′ *S* **V**′ implies **U** *S* **V**. Let "**U** *Ex* **V**" stand for "neither **U** nor **V** is a part of the other." Then, with the same notation, **U**′ *Ex* **V**′ implies **U** *Ex* **V**. We now take [**X**, **U**] *J* [**X**, **V**], for any significant **U** and **V**, to mean

"(i) **U** is not a part of **V**, and (ii) there is no part **W** of **X** such that **V** *S* **W** and **U** *Ex* **W**."

J_1 is clearly satisfied.

J_2. Let the notations be those of the previous discussion of J_2, and let $\mathbf{V}_1'$ and $\mathbf{V}_2'$ be distinct residuals of $\mathbf{V}_1$ and $\mathbf{V}_2$. If $\mathbf{V}_1'$ *S* $\mathbf{V}_2'$ and $\mathbf{V}_2'$ *Ex* **W**′, then $\mathbf{V}_1$ *S* **W** and $\mathbf{V}_2$ *Ex* **W**, which is incompatible with $\eta_1 = \eta_2$ or $\eta_1 \, J \, \eta_2$. If $\mathbf{V}_1 = \mathbf{V}_2$, $\mathbf{V}_1'$ cannot be part of $\mathbf{V}_2'$. If $\mathbf{V}_1 \neq \mathbf{V}_2$, the only possibility that $\mathbf{V}_1'$ be part of $\mathbf{V}_2'$ is that $\mathbf{V}_2$ be part of **M**, with **x** as a free variable, and $\mathbf{V}_1$ be in **N**,—which in view of $\eta_1 \, J \, \eta_2$ contravenes (ii).

J_3. Let η, ξ and ζ be determined by **U**, $\mathbf{V}_1$, and $\mathbf{V}_2$, where **U** is ((λ**xM**)**N**). Suppose that $\xi \, \bar{A} \, \eta$ and $\eta \, \bar{J} \, \zeta$. Then $\mathbf{V}_1$ is part of **N** and either **U** is part of $\mathbf{V}_2$ or, for some **W**, $\mathbf{V}_2$ *S* **W** and **U** *Ex* **W**. The first alternative gives $\mathbf{V}_1$ part of $\mathbf{V}_2$; the second $\mathbf{V}_2$ *S* **W** and $\mathbf{V}_1$ *Ex* **W**; and both contradict $\xi \, J \, \zeta$. Hence

Theorem 12. *Corollaries* 11.1, 11.2 *and the first part of Corollary* 11.3 *hold in the extended calculus.*

It is easily seen that the second part of Corollary 11.3 fails to survive.

Cambridge, England

References

J. W. Alexander, [1] *Combinatorial theory of complexes*, Annals of Math., 31 (1930), pp. 292–320.

G. Birkhoff, [1] *Lattice theory*, New York, 1940.

A. Church, [1] *A formulation of the simple theory of types*, J. of Symbolic Logic, 5 (1940), pp. 56–68.

A. Church and B. Rosser, [1] *Some properties of conversion*, Trans. Amer. Math. Soc., 39 (1936), pp. 472–482.

M. H. A. Newman, [1] *A theorem in combinatory topology*, J. London Math. Soc., 6 (1931), pp. 186–192.

[2] *Stratified systems of logic*. (Forthcoming.)