ALGEBRAIC COMPLEXES APPLIED TO CHEMISTRY

By Peter H. Sellers

JOHNSON RESEARCH FOUNDATION, UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA

Communicated by Britton Chance and Mark Kac, February 24, 1966

1. *Introduction.*—A network of chemical reactions can be described by an algebraic complex, in which the boundary operator assigns a chemical equation to each reaction in the network. Roth (1955) in these Proceedings has given a comparable description for an electrical network; he uses a 1-dimensional simplicial complex, in which the boundary operator implements Kirchoff's laws. As observed by B. Chance (1959), a chemical network is essentially different from an electrical network, and this manifests itself in that we cannot in general use a simplicial complex to describe a chemical network.

The following is a list of some of the algebraic complexes which are treated in our detailed paper [Sellers (1966)]:

(*i*) The most fundamental example is the complex introduced by Eilenberg and Mac Lane (1950) in these Proceedings as the *generic complex* $K(F_a^*, 1)$; F_a is a free abelian group, whose free generators we shall call *primes*.

Application: We shall show that this complex represents a chemical network in which each chemical compound is characterized by a finite set of primes (i.e., molecular structure = finite set).

(*ii*) The *partition complex* is the complex obtained from the generic complex by identifying some of its primes.

Application: Chemically this is the most useful complex; each chemical compound is characterized by its empirical formula (molecular structure = finite set in which each element has a positive integral multiplicity).

(*iii*) The *complex of cubical singular homology* defined by Serre (1951) is shown in our detailed paper to be a subcomplex of the "chemical complex" defined below.

Application: This complex represents a chemical network in which each chemical compound is characterized by a sequence of empirical formulas, and each reaction is an isomerization in which some terms in a sequence are merged to form a single term or split up to form several terms (i.e., molecular structure = chain of empirical formulas).

(*iv*) The *truncated octohedral complex* is a subcomplex of the complex in (*iii*) in which the 1-cells are sequences of primes.

Application: This complex represents a chemical network in which the chemical compounds are characterized by sequences of primes, and the reactions are isomerizations which permute the primes (i.e., molecular structure = sequence).

(*v*) In § 3 we shall define the *chemical complex*, having all the above complexes as special cases or subcomplexes.

Application: This complex, as suggested by the name we have given it here, has been constructed to represent arbitrary chemical networks. It is intended to fulfill the need, stated by Aris (1965), for an algebraic treatment of chemical networks involving "structural chemical formulas." Such a network is one in which empirical formulas do not suffice to characterize its chemical compounds.

In this paper we shall consider only (*i*), the most fundamental example, and (*v*), the most general example, because of their theoretical importance.

2. *The Generic Complex.*—The generic complex $K(F_a^*, 1)$ was defined by Eilenberg and Mac Lane (1950) in these PROCEEDINGS.

F_a is a free abelian group, whose operation is called *product* and whose free generators are called *primes;* square-free products of primes are called *generic elements;* and an ordered π-tuple of generic elements whose product is generic is called a *π-cell.* A *π-chain* is an element of the free abelian group generated by all the π-cells. The boundary ∂ is the homomorphism from the group of π-chains to the group of (π-1)-chains such that for any π-cell $[g_1, g_2, \ldots, g_\pi]$ we have

$$\partial[g_1, g_2, \ldots, g_\pi] = [g_2, \ldots, g_\pi] + \sum_{\theta=1}^{\pi-1}(-1)^\theta[\ldots, g_\theta g_{\theta+1}, \ldots] + (-1)^\pi[g_1, \ldots, g_{\pi-1}].$$

To determine the chemical network described by this complex, all we have to do is to interpret the 1-cells $[g]$, $[h]$, $[gh]$, etc., as *chemical compounds*, the 2-cells $[g, h]$, etc. as *chemical reactions*, and the boundary of $[g, h]$ as the unique *chemical change* which it produces; that is,

$$\partial[g, h] = [h] - [gh] + [g] \rightleftharpoons 0;$$

this denotes a chemical change in which one molecule of $[g]$ and one of $[h]$ are used up, and one of $[gh]$ is produced. (The coefficients could be regarded as real numbers of moles per liter per second instead of whole numbers of molecules; then the groups of π-chains would have to be extended to real vector spaces. We shall not make that digression here.)

The symbol $\rightleftharpoons$ denotes homological equivalence. If we transpose the above equivalence so as to have positive coefficients, we obtain the traditional chemical equation:

$$[g] + [h] \rightleftharpoons [gh].$$

We remark that this equation is balanced in the chemical sense. It suggests the possibility of a chemical change rather than the occurrence of the change itself, which we have denoted above by the 1- chain

$$[h] - [gh] + [g].$$

A 2-chain, a linear combination of reactions, is called a *reaction process.* If its boundary is zero, then it is said to be in "equilibrium"; therefore a 2-cycle is called an *equilibrium state.* If the boundary of a reaction process is a linear combination of chemical compounds which have been designated in advance as "substrates" and "products," then the process is said to be in a "stationary state"; therefore, what is called homologically a relative 2-cycle is a *stationary state* chemically.

In the cochain complex corresponding to $K(F_a^*, 1)$ the 1-cells and 2-cells are functions called *chemical fluxes* and *reaction fluxes*, respectively, with values in the domain of coefficients—integers here.

Question: Does every integral chain complex represent a chemical network in the manner described? The answer is no, if we wish to obey the law of conservation of mass by having only "balanced chemical equations."

Definition: Let x be a 1-chain in $K(F_a^*, 1)$; x is *balanced*, if and only if for any prime g of F_a

$$\xi_g(x) = 0,$$

where ξ_g is a cochain such that for any 1-cell $[m]$

$$\xi_g([m]) = \begin{cases} 1, \text{ if } g \text{ divides } m. \\ 0 \text{ otherwise.} \end{cases}$$

This definition can be extended to any dimension of cell by a termwise definition of divisibility, as is done in our detailed paper. We could describe the mathematical content of our subject as a study of integral chain complexes in which boundaries are balanced with respect to a finite set of primes.

THEOREM 1. *Let x be a 1-chain in a subcomplex of $K(F_a^*, 1)$; if $x \rightleftharpoons 0$, then x is balanced.*

Therefore we may say that if a chemical complex is represented by a subcomplex of a generic complex, then all its chemical equations are balanced. This is equivalent to saying that ξ_g is a cocycle, because for any 2-cell $[m, n]$ the definition of a coboundary gives this:

$$(\delta\xi_g)([m, n]) = \xi_g(\partial[m, n]) = \xi_g([n] - [mn] + [m]) = 0.$$

In the case of the generic complex itself, rather than a proper subcomplex, the converse of Theorem 1 holds, and we may say that every cocycle is generated by cocycles of the form ξ_g.

A 1-cocycle is a linear combination of fluxes which is zero when applied to the chemical change caused by any reaction $[m, n]$. Such linear functions are called *conservation conditions*. A 1-cocycle generated by ξ_g, ξ_h, ... may be called a *mass conservation condition*, and any others [there are none in $K(F_a^*, 1)$] may be called *stoichiometric conservation conditions*.

Having defined various chemical terms mathematically, we can easily deduce theorems about them. The following one is valid for any chemical network regardless of the complex which represents it. More refined versions of the theorem may be obtained when the complex is known.

THEOREM 2. *Let $C_2 \xrightarrow{\partial} C_1$ denote a chemical network, in which C_1 and C_2 are the free abelian groups generated by the finite sets of all chemicals and reactions, respectively, and ∂ the homomorphism which maps any chemical reaction to the chemical change it produces; let*

γ_1 = *the number of chemicals,*
γ_2 = *the number of reactions,*
η_1 = *the maximum number of linearly independent conservation conditions,*
η_2 = *the maximum number of linearly independent equilibrium states;*

then

$$\gamma_1 - \gamma_2 = \eta_1 - \eta_2.$$

This theorem may be regarded as an application of the Euler-Poincaré theorem to the complex $C_2 \xrightarrow{\partial} C_1 \rightarrow 0$.

Problem: Consider a chemical network which is characterized by a generic

complex in which F_a has μ free generators. Find the maximum number η_2 of linearly independent equilibrium states.

By a simple combinatorial calculation we get:

$$\gamma_1 = 2^\mu - 1,$$

$$\gamma_2 = 3^\mu - 2 \cdot 2^\mu + 1.$$

According to our observation that in the present case all conservation conditions are generated by the mass conservation conditions, we conclude:

$$\eta_1 = \mu.$$

By Theorem 2,

$$\eta_2 = \gamma_2 - \gamma_1 + \eta_1 = 3^\mu - 3 \cdot 2^\mu + 2 + \mu.$$

This gives us the number of basic equilibrium states, and we can get an explicit expression for each of them by working out the 2-cycles in the complex

$$C_3 \xrightarrow{\partial} C_2 \xrightarrow{\partial} C_1 \rightarrow 0,$$

where C_3 is generated by all 3-cells of the form $[l, m, n]$. We find that there are $\binom{\mu}{2}$ nonbounding 2-cycles of the form $[g, h] - [h, g]$, where g and h are prime, and that the remaining independent 2-cycles are of the form $\partial[l, m, n] = [m, n] - [lm, n] + [l, mn] - [l, m]$. The first type consists of a chemical reaction accompanied by its reverse reaction, and the second type consists of four reactions—obviously in equilibrium—whose chemical equations are as follows:

$$[m] + [n] \rightleftharpoons [mn],$$
$$[l] + [mn] \rightleftharpoons [lmn],$$
$$[lmn] \rightleftharpoons [lm] + [n],$$
$$[lm] \rightleftharpoons [l] + [m].$$

An application of the generic complex to a problem in biochemistry (hexokinase) is given in the detailed paper.

3. *The Chemical Complex.*—Now we define a complex which was constructed to represent chemical networks generally and contains the other complexes listed in section 1 as subcomplexes and special cases.

Let $\{A_1, A_2, \ldots, A_\mu\}$ be a set of finite cardinality μ. Its elements are called *primes.*

Let $\otimes_\theta$ be called a *tensor product symbol* for $\theta = 1, 2, \ldots, \omega$, and let $\bigcirc_\theta$ be called the *derived product symbol* of $\otimes_\theta$, or just a *product symbol.*

Definition: A *chemical* (or a chemical 1-cell) J is a finite sequence of primes with a product symbol between each consecutive pair; hence we write

$$J = E_1 \bigcirc_1 E_2 \bigcirc_2 \ldots \bigcirc_{\theta-1} E_\theta,$$

where $E_1, E_2, \ldots, E_\theta$ are primes.

Definition: The $\bigcirc_\theta$*-product* $J \bigcirc_\theta K$ of any pair of chemicals is a chemical of the following form:

$$J \bigcirc_\theta K = E_1 \bigcirc_1 \ldots \bigcirc_{\theta-1} E_\theta \bigcirc_\theta E_{\theta+1} \bigcirc_{\theta+1} \ldots \bigcirc_{\theta+\phi-1} E_{\theta+\phi},$$

where $J = E_1 O_1 \ldots O_{\theta-1} E_\theta$ and $K = E_{\theta+1} O_{\theta+1} \ldots O_{\theta+\phi-1} E_{\theta+\phi}$.

Let $\mathfrak{R}$ be a fixed sequence $(\rho_1, \rho_2, \ldots, \rho_\mu)$ of natural numbers, in which ρ_ν is called the *maximum power* of the prime A_ν for $1 \leq \nu \leq \mu$. $\mathfrak{R}$ is a constant of the complex defined here.

Definition: The *chemical network* of $\mathfrak{R}$, ω is denoted by

$$C_2^{\mathfrak{R},\omega} \xrightarrow{\partial} C_1^{\mathfrak{R},\omega};$$

(*i*) $C_1^{\mathfrak{R},\omega}$ is the free abelian group generated by all chemicals such that any one contains at most ρ_ν appearances of A_ν in its sequence of primes and contains any of the product symbols $O_1, O_2, \ldots, O_\omega$; (*ii*) $C_2^{\mathfrak{R},\omega}$ is the free abelian group generated by all elements of the form $J \otimes_\theta K$ $(1 \leqslant \theta \leqslant \omega)$ such that J, K, and $J O_\theta K$ are in $C_1^{\mathfrak{R},\omega}$; (*iii*) ∂ is the homomorphism determined by

$$\partial(J \otimes_\theta K) = K - J O_\theta K + J.$$

Since the groups in this definition are finitely generated, it is always possible to compute the equilibrium states, conservation conditions, stationary states, etc., for explicit small values of $\rho_1, \rho_2, \ldots, \rho_\mu$, and ω. However, to work out these things in general we need to embed the network in a longer algebraic complex, which is constructed as follows:

Definition: The *chemical complex* of $\mathfrak{R}$, ω is the sequence $C_\rho^{\mathfrak{R},\omega} \xrightarrow{\partial} C_{\rho-1}^{\mathfrak{R},\omega} \xrightarrow{\partial} \ldots \xrightarrow{\partial} C_1^{\mathfrak{R},\omega} \to 0$ such that (*i*) $C_1^{\mathfrak{R},\omega}$ is defined as above; (*ii*) for $1 < \pi \leqslant \rho_1 + \ldots + \rho_\mu = \rho$ $C_\pi^{\mathfrak{R},\omega}$ is the free abelian group generated by all π-termed sequences of the form $J_1 \otimes_1 J_2 \otimes_2 \ldots \otimes_{\pi-1} J_\pi$, called a *chemical π-cell*, such that $J_1, J_2, \ldots, J_\pi$, and $J_1 O_1 J_2 O_2 \ldots O_{\pi-1} J_\pi$ are in $C_1^{\mathfrak{R},\omega}$, and $\otimes_1, \otimes_2, \ldots, \otimes_{\pi-1}$ are any tensor product symbols; (*iii*) ∂ is the homomorphism determined by

$$\partial(J_1 \otimes_1 \ldots \otimes_{\pi-1} J_\pi) = (J_2 \otimes_2 \ldots \otimes_{\pi-1} J_\pi) + \sum_{\theta=1}^{\pi-1} (-1)^\theta (J_1 \otimes_1 \ldots \otimes_{\theta-1} J_\theta O_\theta J_{\theta+1} \otimes_{\theta+1} \ldots \otimes_{\pi-1} J_\pi) + (-1)^\pi (J_1 \otimes_1 \ldots \otimes_{\pi-2} J_{\pi-1}).$$

THEOREM 3. *If* $x \epsilon C_\pi^{\mathfrak{R},\omega}$, *then* $\partial\partial x = 0$.

This theorem justifies the above definition.

4. *Application of the Chemical Complex.*—In the detailed treatment [Sellers (1966)] of our subject, the chemical complex is referred to as a "sequential complex," but the present name is appropriate, because the complex was constructed specifically to provide an algebraic description of a chemical network having the following features:

(i) *Structural formulas:* It is assumed that each chemical in the network has a unique empirical formula—mathematically, a unique prime factorization. However, a chemical is not uniquely determined in general by its empirical formula, because there may be different ways of combining a given sequence of primes chemically—mathematically, two products with the same factorization may use different product symbols. If a chemical cannot be uniquely defined by an empirical formula, then we use products other than the free abelian group product and call the representation a *structural formula.*

(ii) *Balanced equations:* Let $x \rightleftharpoons y$ be a chemical equation written in the traditional manner of chemistry with positive coefficients on each side; x and y are

in $C_1^{(R,\omega}$. The equation is balanced in the chemical sense if (regard + as a product symbol) both sides have the same prime factorization. It is easily seen that this is equivalent to saying that $x - y$ satisfies the definition of balance given in § 2.

THEOREM 4. *Let x and y be 1-chains in a subcomplex of a chemical complex; if x and y are homologically equivalent, then $x \rightleftharpoons y$ is balanced. The converse holds in the nonproper subcomplex.*

When the chemical complex is applied to a problem, the product symbols will have particular properties, determining a special complex. In the detailed paper we consider the theory of a few special complexes which appear to be important chemically.

The only law of chemistry which goes into the construction of our complexes is conservation of mass. Therefore they are by no means applicable only to chemistry. It is easy to conceive of a wide class of combinatorial problems which could be put into the language of algebraic complexes, as has been done here for chemistry.

Aris, R., "Prolegomena to the rational analysis of systems of chemical reactions," *Arch. Rational Mech. Anal.*, **19**, 81 (1965).

Chance, B., in *CIBA Foundation Symposium on the Regulation of Cell Metabolism* (Boston; Little, Brown and Co., 1959), pp. 353–368.

Eilenberg, S., and S. Mac Lane, "Cohomology theory of abelian groups and homotopy theory, II," these PROCEEDINGS, **36**, 657–663 (1950).

Roth, J. P., "An application of algebraic topology to numerical analysis: On the existence of a solution to the network problem," these PROCEEDINGS, **41**, 518–521 (1955).

Sellers, P. H., "Algebraic complexes which characterize chemical networks," submitted to *J. Soc. Ind. Appl. Math.* (1966).

Serre, J.-P., "Homologie singulière des espaces fibrés," *Ann. of Math.*, **54**, 425–505 (1951).

---

*ESR STUDY OF γ-IRRADIATED POLYNUCLEOTIDES**

BY JANKO N. HERAK† AND WALTER GORDY

DEPARTMENT OF PHYSICS, DUKE UNIVERSITY

Communicated February 11, 1966

Electron spin resonance investigations of radiation-induced free radicals in DNA and RNA, their constituent bases, and their component nucleosides and nucleotides have previously been made.[1–13] This study is concerned with homopolymers and copolymers of certain of the nucleotides. Some results on three dry polynucleotides at room temperature have previously been reported by Muller.[9]

Perhaps the most important knowledge gained from these previous ESR studies of the nucleic acids and their constituents is that H atoms released from bound water, perhaps also from the sugar group, can add to certain of the base rings (apparently to all except cytosine) to produce H-addition radicals. (*See* Note added in proof.) This H-addition was proved for thymidine by analysis of the single crystal;[13] for DNA, by deuterium substitution[12] and by irradiation of DNA under H_2 pressure.[11] For guanine, adenine, and uracil, as well as for thymine, it was strikingly proved by subjection of the powdered samples to H