Example. Let $E = R^q$ be the free module of dimension q. Then:

$$F_k(E) = \begin{cases} 0 & \text{if } 0 \leqq k < q \\ R & \text{if } k \geqq q. \end{cases}$$

This is immediate from the definitions and the fact that the only relation of a basis for E is the trivial one.

The Fitting ideal $F_0(E)$ is called the **zero-th** or **initial Fitting ideal**. In some applications it is the only one which comes up, in which case it is called "**the**" **Fitting ideal** $F(E)$ of E. It is the ideal generated by all $q \times q$ determinants in the matrix of relations of q generators of the module.

For any module E we let $\operatorname{ann}_R(E)$ be the annihilator of E in R, that is the set of elements $a \in R$ such that $aE = 0$.

Proposition 2.5. *Suppose that E can be generated by q elements. Then*

$$(\operatorname{ann}_R(E))^q \subset F(E) \subset \operatorname{ann}_R(E).$$

In particular, if E can be generated by one element, then

$$F(E) = \operatorname{ann}_R(E).$$

Proof. Let $x_1, \ldots, x_q$ be generators of E. Let $a_1, \ldots, a_q$ be elements of R annihilating E. Then the diagonal matrix whose diagonal components are $a_1, \ldots, a_q$ is a matrix of relations, so the definition of the Fitting ideal shows that the determinant of this matrix, which is the product $a_1 \cdots a_q$ lies in $I_q(E) \subset F_0(E)$. This proves the inclusion

$$\operatorname{ann}_R(E)^q \subset F(E).$$

Conversely, let A be a $q \times q$ matrix of relations between $x_1, \ldots, x_q$. Then $\det(A)x_i = 0$ for all i so $\det(A) \in \operatorname{ann}_R(E)$. Since $F(E)$ is generated by such determinants, we get the reverse inclusion which proves the proposition.

Corollary 2.6. *Let $E = R/\mathfrak{a}$ for some ideal $\mathfrak{a}$. Then $F(E) = \mathfrak{a}$.*

Proof. The module $R/\mathfrak{a}$ can be generated by one element so the corollary is an immediate consequence of the proposition.

Proposition 2.7. *Let*

$$0 \to E' \to E \to E'' \to 0$$

be an exact sequence of finite R-modules. For integers $m, n \geqq 0$ we have

$$F_m(E')F_n(E'') \subset F_{m+n}(E).$$

In particular for $F = F_0$.

$$F(E')F(E'') \subset F(E).$$

Proof. We may assume E' is a submodule of E. We pick generators $x_1, \ldots, x_p$ of E' and elements $y_1, \ldots, y_q$ in E such that their images $y_1'', \ldots, y_q''$ in E'' generate E''. Then (x, y) is a family of generators for E. Suppose first that $m \leqq p$ and $n \leqq q$. Let A be a matrix of relations among $y_1'', \ldots, y_q''$ with q columns. If $(a_1, \ldots, a_q)$ is such a relation, then

$$a_1 y_1 + \cdots + a_q y_q \in E'$$

so there exist elements $b_1, \ldots, b_p \in R$ such that

$$\sum a_i y_i + \sum b_j x_j = 0.$$

Thus we can find a matrix B with p columns and the same number of rows as A such that (B, A) is a matrix of relations of (x, y). Let C be a matrix of relations of $(x_1, \ldots, x_p)$. Then

$$\begin{pmatrix} B & A \\ C & 0 \end{pmatrix}$$

is a matrix of relations of (x, y). If D'' is a $(q - n) \times (q - n)$ subdeterminant of A and D' is a $(p - m) \times (p - m)$ subdeterminant of C then $D''D'$ is a

$$(p + q - m - n) \times (p + q - m - n)$$

subdeterminant of the matrix

$$\begin{pmatrix} B & A \\ C & 0 \end{pmatrix}$$

and $D''D' \in F_{m+n}(E)$. Since $F_m(E')$ is generated by determinants like D' and $F_n(E'')$ is generated by determinants like D'', this proves the proposition in the present case.

If $m > p$ and $n > q$ then $F_{m+n}(E) = F_m(E') = F_n(E'') = R$ so the proposition is trivial in this case.

Say $m \leqq p$ and $n > q$. Then $F_n(E'') = R = F_q(E'')$ and hence

$$F_m(E')F_n(E'') = F_q(E'')F_m(E') \subset F_{p+n}(E) \subset F_{m+n}(E)$$

where the inclusion follows from the first case. A similar argument proves the remaining case with $m > p$ and $n \leqq q$. This concludes the proof.

Proposition 2.8. *Let E', E'' be finite R-modules. For any integer $n \geqq 0$ we have*

$$F_n(E' \oplus E'') = \sum_{r+s=n} F_r(E')F_s(E'').$$

Proof. Let $x_1, \ldots, x_p$ generate E' and $y_1, \ldots, y_q$ generate E''. Then (x, y) generate $E' \oplus E''$. By Proposition 2.6 we know the inclusion

$$\sum F_r(E')F_s(E'') \subset F_n(E' \oplus E''),$$

so we have to prove the converse. If $n \geqq p + q$ then we can take $r \geqq p$ and $s \geqq q$ in which case

$$F_r(E') = F_s(E'') = F_n(E) = R$$

and we are done. So we assume $n < p + q$. A relation between (x, y) in the direct sum splits into a relation for (x) and a relation for (y). The matrix of relations for (x, y) is therefore of the form

$$C = \begin{pmatrix} A' & 0 \\ 0 & A'' \end{pmatrix}$$

where A' is the matrix of relations for (x) and A'' the matrix of relations for (y). Thus

$$F_n(E' \oplus E'') = \sum_C I_{p+q-n}(C)$$

where the sum is taken over all matrices C as above. Let D be a

$$(p + q - n) \times (p + q - n)$$

subdeterminant. Then D has the form

$$D = \begin{vmatrix} B' & 0 \\ 0 & B'' \end{vmatrix}$$

where B' is a $k' \times (p - r)$ matrix, and B'' is a $k'' \times (q - s)$ matrix with some positive integers k', k'', r, s satisfying

$$k' + k'' = p + q - n \quad \text{and} \quad r + s = n.$$

Then $D = 0$ unless $k' = p - r$ and $k'' = q - s$. In that case

$$D = \det(B')\det(B'') \in F_r(E')F_s(E''),$$

which proves the reverse inclusion and concludes the proof of the proposition.

Corollary 2.9. *Let*

$$E = \bigoplus_{i=1}^{s} R/\mathfrak{a}_i$$

where $\mathfrak{a}_i$ is an ideal. Then $F(E) = \mathfrak{a}_1 \cdots \mathfrak{a}_s$.

Proof. This is really a corollary of Proposition 2.8 and Corollary 2.6.

§3. UNIVERSAL DERIVATIONS AND THE DE RHAM COMPLEX

In this section, all rings R, A, etc. are assumed commutative.

Let A be an R-algebra and M an A-module. By a **derivation** $D: A \to M$ (over R) we mean an R-linear map satisfying the usual rules

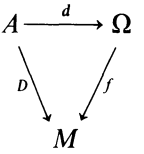
$$D(ab) = aDb + bDa.$$

Note that $D(1) = 2D(1)$ so $D(1) = 0$, whence $D(R) = 0$. Such derivations form an A-module $\mathrm{Der}_R(A, M)$ in a natural way, where aD is defined by $(aD)(b) = aDb$.

By a **universal derivation** for A over R, we mean an A-module Ω, and a derivation

$$d: A \to \Omega$$

such that, given a derivation $D: A \to M$ there exists a unique A-homomorphism $f: \Omega \to M$ making the following diagram commutative:



It is immediate from the definition that a universal derivation (d, Ω) is uniquely determined up to a unique isomorphism. By definition, we have a functorial isomorphism

$$\boxed{\mathrm{Der}_R(A, M) \approx \mathrm{Hom}_A(\Omega, M).}$$

We shall now prove the existence of a universal derivation.

The following general remark will be useful. Let

$$f_1, f_2: A \to B$$

be two homomorphisms of R-algebras, and let J be an ideal in B such that $J^2 = 0$. Assume that $f_1 \equiv f_2 \bmod J$; this means that $f_1(x) \equiv f_2(x) \bmod J$ for all x in A. Then

$$D = f_2 - f_1$$

is a derivation. This fact is immediately verified as follows:

$$\begin{aligned} f_2(ab) = f_2(a)f_2(b) &= [f_1(a) + D(a)][f_1(b) + D(b)] \\ &= f_1(ab) + f_1(b)D(a) + f_1(a)D(b). \end{aligned}$$

But the A-module structure of J is given via f_1 or f_2 (which amount to the same thing in light of our assumptions on f_1, f_2), so the fact is proved.

Let the tensor product be taken over R.

Let $\mathbf{m}_A : A \otimes A \to A$ be the multiplication homomorphism, such that $\mathbf{m}_A(a \otimes b) = ab$. Let $J = \operatorname{Ker} \mathbf{m}_A$. We define the module of **differentials**

$$\Omega_{A/R} = J/J^2,$$

as an ideal in $(A \otimes A)/J^2$. The A-module structure will always be given via the embedding on the first factor:

$$A \to A \otimes A \quad \text{by} \quad a \mapsto a \otimes 1.$$

Note that we have a direct sum decomposition of A-modules

$$A \otimes A = (A \otimes 1) \oplus J,$$

and therefore

$$(A \otimes A)/J^2 = (A \otimes 1) \oplus J/J^2.$$

Let

$$d : A \to J/J^2 \text{ be the } R\text{-linear map } a \mapsto 1 \otimes a - a \otimes 1 \bmod J^2.$$

Taking $f_1 : a \mapsto a \otimes 1$ and $f_2 : a \mapsto 1 \otimes a$, we see that $d = f_2 - f_1$. Hence d is a derivation when viewed as a map into J/J^2.

We note that J is generated by elements of the form

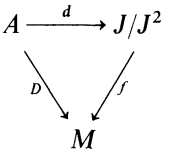
$$\sum x_i \, dy_i.$$

Indeed, if $\sum x_i \otimes y_i \in J$, then by definition $\sum x_i y_i = 0$, and hence

$$\sum x_i \otimes y_i = \sum x_i(1 \otimes y_i - y_i \otimes 1),$$

according to the A-module structure we have put on $A \otimes A$ (operation of A on the left factor.)

Theorem 3.1. *The pair $(J/J^2, d)$ is universal for derivations of A. This means: Given a derivation $D : A \to M$ there exists a unique A-linear map $f : J/J^2 \to M$ making the following diagram commutative.*



Proof. There is a unique R-bilinear map

$$f: A \otimes A \to M \quad \text{given by} \quad x \otimes y \mapsto xDy,$$

which is A-linear by our definition of the A-module structure on $A \otimes A$. Then by definition, the diagram is commutative on elements of A, when we take f restricted to J, because

$$f(1 \otimes y - y \otimes 1) = Dy.$$

Since J/J^2 is generated by elements of the form $x\,dy$, the uniqueness of the map in the diagram of the theorem is clear. This proves the desired universal property.

We may write the result expressed in the theorem as a formula

$$\boxed{\mathrm{Der}_R(A, M) \approx \mathrm{Hom}_A(J/J^2, M).}$$

The reader will find exercises on derivations which give an alternative way of constructing the universal derivation, especially useful when dealing with finitely generated algebras, which are factors of polynomial rings.

I insert here without proofs some further fundamental constructions, important in differential and algebraic geometry. The proofs are easy, and provide nice exercises.

Let $R \to A$ be an R-algebra of commutative rings. For $i \geqq 0$ define

$$\Omega^i_{A/R} = \bigwedge^i \Omega^1_{A/R},$$

where $\Omega^0_{A/R} = A$.

Theorem 3.2. *There exists a unique sequence of R-homomorphisms*

$$d_i: \Omega^i_{A/R} \to \Omega^{i+1}_{A/R}$$

such that for $\omega \in \Omega^i$ and $\eta \in \Omega^j$ we have

$$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^i \omega \wedge d\eta.$$

Furthermore $d \circ d = 0$.

The proof will be left as an exercise.

Recall that a **complex** of modules is a sequence of homomorphisms

$$\cdots \to E^{i-1} \xrightarrow{d^{i-1}} E^i \xrightarrow{d^i} E^{i+1} \to$$

such that $d^i \circ d^{i-1} = 0$. One usually omits the superscript on the maps d. With this terminology, we see that the $\Omega^i_{A/R}$ form a complex, called the **De Rham complex**.

Theorem 3.3. *Let k be a field of characteristic 0, and let $A = k[X_1, \ldots, X_n]$ be the polynomial ring in n variables. Then the De Rham complex*

$$0 \to k \to A \to \Omega^1_{A/k} \to \cdots \to \Omega^n_{A/k} \to 0$$

is exact.

Again the proof will be left as an exercise. *Hint*: Use induction and integrate formally.

Other results concerning connections will be found in the exercises below.

§4. THE CLIFFORD ALGEBRA

Let k be a field. By an **algebra** throughout this section, we mean a k-algebra given by a ring homomorphism $k \to A$ such that the image of k is in the center of A.

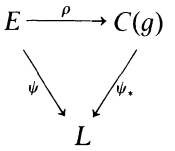
Let E be a finite dimensional vector space over the field k, and let g be a symmetric form on E. We would like to find a universal algebra over k, in which we can embed E, and such that the square in the algebra corresponds to the value of the quadratic form in E. More precisely, by a **Clifford algebra** for g, we shall mean a k-algebra $C(g)$, also denoted by $C_g(E)$, and a linear map $\rho: E \to C(g)$ having the following property: If $\psi : E \to L$ is a linear map of E into a k-algebra L such that

$$\psi(x)^2 = g(x, x) \cdot 1 \qquad (1 = \text{unit element of } L)$$

for all $x \in E$, then there exists a unique algebra-homomorphism

$$C(\psi) = \psi_* : C(g) \to L$$

such that the following diagram is commutative:



By abstract nonsense, a Clifford algebra for g is uniquely determined, up to a unique isomorphism. Furthermore, it is clear that if $(C(g), \rho)$ exists, then $C(g)$ is generated by the image of ρ, i.e. by $\rho(E)$, *as an algebra over k.*

We shall write $\rho = \rho_g$ if it is necessary to specify the reference to g explicitly.

We have trivially

$$\rho(x)^2 = g(x, x) \cdot 1$$

for all $x \in E$, and

$$\rho(x)\rho(y) + \rho(y)\rho(x) = 2g(x, y) \cdot 1$$

as one sees by replacing x by $x + y$ in the preceding relation.

Theorem 4.1. *Let g be a symmetric bilinear form on a finite dimensional vector space E over k. Then the Clifford algebra $(C(g), \rho)$ exists. The map ρ in injective, and $C(g)$ has dimension 2^n over k, if $n = \dim E$.*

Proof. Let $T(E)$ be the tensor algebra as in Chapter XVI, §7. In that algebra, we let I_g be the two-sided ideal generated by all elements

$$x \otimes x - g(x, x) \cdot 1 \text{ for } x \in E.$$

We define $C_g(E) = T(E)/I_g$. Observe that E is naturally embedded in $T(E)$ since

$$T(E) = k \oplus E \oplus (E \otimes E) \oplus \cdots.$$

Then the natural embedding of E in TE followed by the canonical homomorphisms of $T(E)$ onto $C_g(E)$ defines our k-linear map $\rho : E \to C_g(E)$. It is immediate from the universal property of the tensor product that $C_g(E)$ as just defined satisfies the universal property of a Clifford algebra, which therefore exists. The only problem is to prove that it has the stated dimension over k.

We first prove that the dimension is $\leqq 2^n$. We give a proof only when the characteristic of k is $\neq 2$ and leave characteristic 2 to the reader. Let $\{v_1, \ldots, v_n\}$ be an orthogonal basis of E as given by Theorem 3.1 of Chapter XV. Let $e_i = \psi(v_i)$, where $\psi : E \to L$ is given as in the beginning of the section. Let $c_i = g(v_i, v_i)$. Then we have the relations

$$e_i^2 = c_i, \qquad e_i e_j = -e_j e_i \text{ for all } i \neq j.$$

This immediately implies that the subalgebra of L generated by $\psi(E)$ over k is generated as a vector space over k by all elements

$$e_1^{\nu_1} \cdots e_n^{\nu_n} \text{ with } \nu_i = 0 \text{ or } 1 \text{ for } i = 1, \ldots, n.$$

Hence the dimension of this subalgebra is $\leqq 2^n$. In particular, $\dim C_g(E) \leqq 2^n$ as desired.

There remains to show that there exists at least one $\psi: E \to L$ such that L is generated by $\psi(E)$ as an algebra over k, and has dimension 2^n; for in that case, the homomorphism $\psi_* : C_g(E) \to L$ being surjective, it follows that $\dim C_g(E) \geqq 2^n$ and the theorem will be proved. We construct L in the following way. We first need some general notions.

Let M be a module over a commutative ring. Let $i, j \in \mathbf{Z}/2\mathbf{Z}$. Suppose M is a direct sum $M = M_0 \oplus M_1$ where 0, 1 are viewed as the elements of $\mathbf{Z}/2\mathbf{Z}$. We then say that M is **$\mathbf{Z}/2\mathbf{Z}$-graded**. If M is an algebra over the ring, we say

it is a $\mathbf{Z}/2\mathbf{Z}$-**graded algebra** if $M_i M_j \subset M_{i+j}$ for all $i, j \in \mathbf{Z}/2\mathbf{Z}$. We simply say **graded**, omitting the $\mathbf{Z}/2\mathbf{Z}$ prefix when the reference to $\mathbf{Z}/2\mathbf{Z}$ is fixed throughout a discussion, which will be the case in the rest of this section.

Let A, B be graded modules as above, with $A = A_0 \oplus A_1$ and $B = B_0 \oplus B_1$. Then the tensor product $A \otimes B$ has a direct sum decomposition

$$A \otimes B = \bigoplus_{i,j} A_i \otimes B_j.$$

We define a grading on $A \otimes B$ by letting $(A \otimes B)_0$ consist of the sum over indices i, j such that $i + j = 0$ (in $\mathbf{Z}/2\mathbf{Z}$), and $(A \otimes B)_1$ consist of the sum over the indices i, j such that $i + j = 1$.

Suppose that A, B are graded algebras over the given commutative ring. There is a unique bilinear map of $A \otimes B$ into itself such that

$$(a \otimes b)(a' \otimes b') = (-1)^{ij} aa' \otimes bb'$$

if $a' \in A_i$ and $b \in B_j$. Just as in Chapter XVI, §6, one verifies associativity and the fact that this product gives rise to a graded algebra, whose product is called the **super tensor product**, or **super product**. As a matter of notation, when we take the super tensor product of A and B, we shall denote the resulting algebra by

$$A \otimes_{su} B$$

to distinguish it from the ordinary algebra $A \otimes B$ of Chapter XVI, §6.

Next suppose that E has dimension 1 over k. Then the factor polynomial ring $k[X]/(x^2 - c_1)$ is immediately verified to be the Clifford algebra in this case. We let t_1 be the image of X in the factor ring, so $C_g(E) = k[t_1]$ with $t_1^2 = c_1$. The vector space E is imbedded as kt_1 in the direct sum $k \oplus kt_1$.

In general we now take the super tensor product inductively:

$$C_g(E) = k[t_1] \otimes_{su} k[t_2] \otimes_{su} \cdots \otimes_{su} k[t_n], \text{ with } k[t_i] = k[X]/(X^2 - c_i).$$

Its dimension is 2^n. Then E is embedded in $C_g(E)$ by the map

$$a_1 v_1 + \cdots + a_n v_n \mapsto a_1 t_1 \oplus \cdots \oplus a_n t_n.$$

The desired commutation rules among t_i, t_j are immediately verified from the definition of the super product, thus concluding the proof of the dimension of the Clifford algebra.

Note that the proof gives an explicit representation of the relations of the algebra, which also makes it easy to compute in the algebra. Note further that the alternating algebra of a free module is a special case, taking $c_i = 0$ for all i. Taking the c_i to be algebraically independent shows that the alternating algebra is a specialization of the generic Clifford algebra, or that Clifford algebras are what one calls perturbations of the alternating algebra. Just as for the alternating algebra, we have immediately from the construction:

Theorem 4.2. *Let g, g' by symmetric forms on E, E' respectively. Then we*

have an algebra isomorphism

$$C(g \oplus g') \approx C(g) \otimes_{su} C(g').$$

Examples. Clifford algebras have had increasingly wide applications in physics, differential geometry, topology, group representations (finite groups and Lie groups), and number theory. First, in topology I refer to Adams [Ad 62] and [ABS 64] giving applications of the Clifford algebra to various problems in topology, notably a description of the way Clifford algebras over the reals are related to the existence of vector fields on spheres. The multiplication in the Clifford algebra gives rise to a multiplication on the sphere, whence to vector fields. [ABS 64] also gives a number of computations related to the Clifford algebra and its applications to topology and physics. For instance, let $E = \mathbf{R}^n$ and let g be the negative of the standard dot product. Or more invariantly, take for E an n-dimensional vector space over $\mathbf{R}$, and let g be a *negative definite* symmetric form on E. Let $C_n = C(g)$.

The operation

$$v_1 \otimes \cdots \otimes v_r \mapsto v_r \otimes \cdots \otimes v_1 = (v_1 \otimes \cdots \otimes v_r)^* \text{ for } v_i \in E$$

induces an endomorphism of $T^r(E)$ for $r \geqq 0$. Since $v \otimes v - g(v, v) \cdot 1$ (for $v \in E$) is invariant under this operation, there is an induced endomorphism $* : C_n \to C_n$, which is actually an involution, that is $x^{**} = x$ and $(xy)^* = y^*x^*$ for $x \in C_n$. We let Spin(n) be the subgroup of units in C_n generated by the unit sphere in E (i.e. the set of elements such that $g(v, v) = -1$), and lying in the even part of C_n. Equivalently, Spin(n) is the group of elements x such that $xx^* = 1$. The name dates back to Dirac who used this group in his study of electron spin. Topologists and others view that group as being the universal covering group of the special orthogonal group $SO(n) = SU_n(\mathbf{R})$.

An account of some of the results of [Ad 62] and [ABS 64] will also be found in [Hu 75], Chapter 11. Second I refer to two works encompassing two decades, concerning the heat kernel, Dirac operator, index theorem, and number theory, ranging from Atiyah, Bott and Patodi [ABP 73] to Faltings [Fa 91], see especially §4, entitled "The local index theorem for Dirac operators". The vector space to which the general theory is applied is mostly the cotangent space at a point on a manifold. I recommend the book [BGV 92], Chapter 3.

Finally, I refer to Bröcker and Tom Dieck for applications of the Clifford algebra to representation theory, starting with their Chapter I, §6, [BtD 85].

Bibliography

[Ad 62] F. ADAMS, Vector Fields on Spheres, *Ann. Math.* **75** (1962) pp. 603–632

[ABP 73] M. ATIYAH, R. BOTT, V. PATODI, On the heat equation and the index theorem, *Invent. Math.* **19** (1973) pp. 270–330; erratum **38** (1975) pp. 277–280

[ABS 64] M. ATIYAH, R. BOTT, A. SHAPIRO, Clifford Modules, *Topology* **Vol. 3, Supp. 1** (1964) pp. 3–38

[BGV 92] N. BERLINE, E. GETZLER, and M. VERGNE, *Heat Kernels and Dirac Operators*, Springer Verlag, 1992

[BtD 85] T. BRÖCKER and T. TOM DIECK, *Representations of Compact Lie Groups*, Springer Verlag 1985

[Fa 91] G. FALTINGS, *Lectures on the arithmetic Riemann-Roch theorem*, Annals of Math. Studies 1991

[Hu 75] D. HUSEMOLLER, *Fibre Bundles*, Springer Verlag, Second Edition, 1975

EXERCISES

1. Let E be a finite dimensional vector space over a field k. Let $x_1, \ldots, x_p$ be elements of E such that $x_1 \wedge \cdots \wedge x_p \neq 0$, and similarly $y_1 \wedge \cdots \wedge y_p \neq 0$. If $c \in k$ and

$$x_1 \wedge \cdots \wedge x_p = cy_1 \wedge \cdots \wedge y_p$$

show that $x_1, \ldots, x_p$ and $y_1, \ldots, y_p$ generate the same subspace. Thus non-zero decomposable vectors in $\bigwedge^p E$ up to non-zero scalar multiples correspond to p-dimensional subspaces of E.

2. Let E be a free module of dimension n over the commutative ring R. Let $f: E \to E$ be a linear map. Let $\alpha_r(f) = \operatorname{tr} \bigwedge^r(f)$, where $\bigwedge^r(f)$ is the endomorphism of $\bigwedge^r(E)$ into itself induced by f. We have

$$\alpha_0(f) = 1, \qquad \alpha_1(f) = \operatorname{tr}(f), \qquad \alpha_n(f) = \det f,$$

and $\alpha_r(f) = 0$ if $r > n$. Show that

$$\det(1 + f) = \sum_{r \geqq 0} \alpha_r(f).$$

[*Hint*: As usual, prove the statement when f is represented by a matrix with variable coefficients over the integers.] Interpret the $\alpha_r(f)$ in terms of the coefficients of the characteristic polynomial of f.

3. Let E be a finite dimensional free module over the commutative ring R. Let $E^\vee$ be its dual module. For each integer $r \geqq 1$ show that $\bigwedge^r E$ and $\bigwedge^r E^\vee$ are dual modules to each other, under the bilinear map such that

$$(v_1 \wedge \cdots \wedge v_r, v_1' \wedge \cdots \wedge v_r') \mapsto \det(\langle v_i, v_j' \rangle)$$

where $\langle v_i, v_j' \rangle$ is the value of v_j' on v_i, as usual, for $v_i \in E$ and $v_j' \in E^\vee$.

4. Notation being as in the preceding exercise, let F be another R-module which is free, finite dimensional. Let $f: E \to F$ be a linear map. Relative to the bilinear map of the preceding exercise, show that the transpose of $\bigwedge^r f$ is $\bigwedge^r({}^t f)$, i.e. is equal to the r-th alternating product of the transpose of f.

5. Let R be a commutative ring. If E is an R-module, denote by $L_a^r(E)$ the module of

r-multilinear alternating maps of E into R itself (i.e. the r-multilinear alternating forms on E). Let $L_a^0(E) = R$, and let

$$\Omega(E) = \bigoplus_{r=0}^{\infty} L_a^r(E).$$

Show that $\Omega(E)$ is a graded R-algebra, the multiplication being defined as follows. If $\omega \in L_a^r(E)$ and $\psi \in L_a^s(E)$, and $v_1, \ldots, v_{r+s}$ are elements of E, then

$$(\omega \wedge \psi)(v_1, \ldots, v_{r+s}) = \sum \epsilon(\sigma)\omega(v_{\sigma 1}, \ldots, v_{\sigma r})\psi(v_{\sigma(r+1)}, \ldots, v_{\sigma s}),$$

the sum being taken over all permutations σ of $(1, \ldots, r+s)$ such that $\sigma 1 < \cdots < \sigma r$ and $\sigma(r+1) < \cdots < \sigma s$.

Derivations

In the following exercises on derivations, all rings are assumed commutative. Among other things, the exercises give another proof of the existence of universal derivations.

Let $R \to A$ be a R-algebra (of commutative rings, according to our convention). We denote the module of universal derivations of A over R by $(d_{A/R}, \Omega^1_{A/R})$, but we do not assume that it necessarily exists. Sometimes we write d instead of $d_{A/R}$ for simplicity if the reference to A/R is clear.

6. Let $A = R[X_\alpha]$ be a polynomial ring in variables X_α, where α ranges over some indexing set, possibly infinite. Let Ω be the free A-module on the symbols dX_α, and let

$$d : A \to \Omega$$

be the mapping defined by

$$df(X) = \sum_\alpha \frac{\partial f}{\partial X_\alpha} dX_\alpha.$$

Show that the pair (d, Ω) is a universal derivation $(d_{A/R}, \Omega^1_{A/R})$.

7. Let $A \to B$ be a homomorphism of R-algebras. Assume that the universal derivations for A/R, B/R, and B/A exist. Show that one has a natural exact sequence:

$$B \otimes_A \Omega^1_{A/R} \to \Omega^1_{B/R} \to \Omega^1_{B/A} \to 0.$$

[*Hint*: Consider the sequence

$$0 \to \mathrm{Der}_A(B, M) \to \mathrm{Der}_R(B, M) \to \mathrm{Der}_R(A, M)$$

which you prove is exact. Use the fact that a sequence of B-modules

$$N' \to N \to N'' \to 0$$

is exact if and only if its Hom into M is exact for every B-module M. Apply this to the sequence of derivations.]

8. Let $R \to A$ be an R-algebra, and let I be an ideal of A. Let $B = A/I$. Suppose that the universal derivation of A over R exists. Show that the universal derivation of B over R

also exists, and that there is a natural exact sequence

$$I/I^2 \xrightarrow{d_{A/R}} B \otimes_A \Omega^1_{A/R} \to \Omega^1_{B/R} \to 0.$$

[*Hint*: Let M be a B-module. Show that the sequence

$$0 \to \mathrm{Der}_R(B, M) \to \mathrm{Der}_R(A, M) \to \mathrm{Hom}_B(I/I^2, M)$$

is exact.]

9. Let $R \to B$ be an R-algebra. Show that the universal derivation of B over R exists as follows. Represent B as a quotient of a polynomial ring, possibly in infinitely many variables. Apply Exercises 6 and 7.

10. Let $R \to A$ be an R-algebra. Let S_0 be a multiplicative subset of R, and S a multiplicative subset of A such that S_0 maps into S. Show that the universal derivation of $S^{-1}A$ over $S_0^{-1}R$ is $(d, S^{-1}\Omega^1_{A/R})$, where

$$d(a/s) = (sd_{A/R}(a) - ad_{A/R}(s))/s^2.$$

11. Let B be an R-algebra and M a B-module. On $B \oplus M$ define a product

$$(b, x)(b', y) = (bb', by + b'x).$$

Show that $B \oplus M$ is a B-algebra, if we identify an element $b \in B$ with $(b, 0)$. For any R-algebra A, show that the algebra homomorphisms $\mathrm{Hom}_{\mathrm{Alg}/R}(A, B \oplus M)$ consist of pairs (φ, D), where $\varphi : A \to B$ is an algebra homomorphism, and $D : A \to M$ is a derivation for the A-module structure on M induced by φ.

12. Let A be an R-algebra. Let $\varepsilon : A \to R$ be an algebra homomorphism, which we call an **augmentation**. Let M be an R-module. Define an A-module structure on M via ε, by

$$a \cdot x = \varepsilon(a)x \qquad \text{for} \qquad a \in A \quad \text{and} \quad x \in M.$$

Write M_ε to denote M with this new module structure. Let:

$$\mathrm{Der}_\varepsilon(A, M) = A\text{-module of derivations for the } \varepsilon\text{-module structure on } M$$

$$I = \mathrm{Ker}\, \varepsilon.$$

Then $\mathrm{Der}_\varepsilon(A, M)$ is an A/I-module. Note that there is an R-module direct sum decomposition $A = R \oplus I$. Show that there is a natural A-module isomorphism

$$\Omega_{A/R}/I\Omega_{A/R} \approx I/I^2$$

and an R-module isomorphism

$$\mathrm{Der}_\varepsilon(A, M) \approx \mathrm{Hom}_R(I/I^2, M).$$

In particular, let $\eta : A \to I/I^2$ be the projection of A on I/I^2 relative to the direct sum decomposition $A = R \oplus I$. Then η is the universal ε-derivation.

Derivations and connections

13. Let $R \to A$ be a homomorphism of commutative rings, so we view A as an R-algebra.

Let E be an A-module. A **connection** on E is a homomorphism of abelian groups

$$\nabla : E \to \Omega^1_{A/R} \otimes_A E$$

such that for $a \in A$ and $x \in E$ we have

$$\nabla(ax) = a\nabla(x) + da \otimes x,$$

where the tensor product is taken over A unless otherwise specified. The kernel of ∇, denoted by E_∇, is called the **submodule of horizontal elements**, or the **horizontal submodule** of (E, ∇).

(a) For any integer $i \geqq 1$, define

$$\Omega^i_{A/R} = \bigwedge\nolimits^i \Omega^1_{A/R}.$$

Show that ∇ can be extended to a homomorphism of R-modules

$$\nabla_i : \Omega^i_{A/R} \otimes E \to \Omega^{i+1}_{A/R} \otimes E$$

by

$$\nabla_i(\omega \otimes x) = d\omega \otimes x + (-1)^i \omega \wedge \nabla(x).$$

(b) Define the **curvature** of the connection to be the map

$$K = \nabla_1 \circ \nabla : E \to \Omega^2_{A/R} \otimes_A E.$$

Show that K is an A-homomorphism. Show that

$$\nabla_{i+1} \circ \nabla_i(\omega \otimes x) = \omega \wedge K(x)$$

for $\omega \in \Omega^i_{A/R}$ and $x \in E$.

(c) Let $\mathrm{Der}(A/R)$ denote the A-module of derivations of A into itself, over R. Let ∇ be a connection on E. Show that ∇ induces a unique A-linear map

$$\nabla : \mathrm{Der}(A/R) \to \mathrm{End}_R(E)$$

such that

$$\nabla(D)(ax) = D(a)x + a\nabla(D)(x).$$

(d) Prove the formula

$$[\nabla(D_1), \nabla(D_2)] - \nabla([D_1, D_2]) = (D_1 \wedge D_2)(K).$$

In this formula, the bracket is defined by $[f, g] = f \circ g - g \circ f$ for two endomorphisms f, g of E. Furthermore, the right-hand side is the composed mapping

$$E \xrightarrow{K} \Omega^2_{A/R} \otimes E \xrightarrow{D_1 \wedge D_2} A \otimes E \approx E.$$

14. (a) For any derivation D of a ring A into itself, prove **Leibniz's rule**:

$$D^n(xy) = \sum_{i=0}^{n} \binom{n}{i} D^i(x) D^{n-i}(y).$$

(b) Suppose A has characteristic p. Show that D^p is a derivation.

15. Let A/R be an algebra, and let E be an A-module with a connection ∇. Assume that R has characteristic p. Define

$$\psi : \mathrm{Der}(A/R) \to \mathrm{End}_R(E)$$

by

$$\psi(D) = (\nabla(D))^p - \nabla(D^p).$$

Prove that $\psi(D)$ is A-linear. [*Hint*: Use Leibniz's formula and the definition of a connection.] Thus the image of ψ is actually in $\mathrm{End}_A(E)$.

Some Clifford exercises

16. Let $C_g(E)$ be the Clifford algebra as defined in §4. Define $F_i(C_g) = (k + E)^i$, viewing E as embedded in C_g. Define the similar object $F_i(\bigwedge E)$ in the alternating algebra. Then $F_{i+1} \supset F_i$ in both cases, and we define the i-th graded module $\mathrm{gr}_i = F_i/F_{i-1}$. Show that there is a natural (functorial) isomorphism

$$\mathrm{gr}_i(C_g(E)) \xrightarrow{\approx} \mathrm{gr}_i(\bigwedge E).$$

17. Suppose that $k = \mathbf{R}$, so E is a real vector space, which we now assume of even dimension $2m$. We also assume that g is non-degenerate. We omit the index g since the symmetric form is now fixed, and we write C^+, C^- for the spaces of degree 0 and 1 respectively in the $\mathbf{Z}/2\mathbf{Z}$-grading. For elements x, y in C^+ or C^-, define their **supercommutator** to be

$$\{x, y\} = xy - (-1)^{(\deg x)(\deg y)} yx.$$

Show that F_{2m-1} is generated by supercommutators.

18. Still assuming g non-degenerate, let J be an automorphism of (E, g) (i.e. $g(Jx, Jy) = g(x, y)$ for all $x, y \in E$) such that $J^2 = -\mathrm{id}$. Let $E_{\mathbf{C}} = \mathbf{C} \otimes_{\mathbf{R}} E$ be the extension of scalars from $\mathbf{R}$ to $\mathbf{C}$. Then $E_{\mathbf{C}}$ has a direct sum decomposition

$$E_{\mathbf{C}} = E_{\mathbf{C}}^+ \oplus E_{\mathbf{C}}^-$$

into the eigenspaces of J, with eigenvalues 1 and -1 respectively. (Proof?) There is a representation of $E_{\mathbf{C}}$ on $\bigwedge E_{\mathbf{C}}^+$, i.e. a homomorphism $E_{\mathbf{C}} \to \mathrm{End}_{\mathbf{C}}(E_{\mathbf{C}}^+)$ whereby an element of $E_{\mathbf{C}}^+$ operates by exterior multiplication, and an element of $E_{\mathbf{C}}^-$ operates by inner multiplication, defined as follows.

For $x' \in E_{\mathbf{C}}^-$ there is a unique $\mathbf{C}$-linear map having the effect

$$x'(x_1 \wedge \cdots \wedge x_r) = -2 \sum_{i=1}^{r} (-1)^{i-1} \langle x', x_i \rangle \, x_1 \wedge \cdots \wedge \hat{x}_i \wedge \cdots \wedge x_r.$$

Prove that under this operation, you get an isomorphism

$$C_g(E)_{\mathbf{C}} \to \mathrm{End}_{\mathbf{C}}(\bigwedge E_{\mathbf{C}}^{+}).$$

[*Hint*: Count dimensions.]

19. Consider the Clifford algebra over $\mathbf{R}$. The standard notation is C_n if $E = \mathbf{R}^n$ with the negative definite form, and C'_n if $E = \mathbf{R}^n$ with the positive definite form. Thus $\dim C_n = \dim C'_n = 2^n$.
    (a) Show that

$$C_1 \approx \mathbf{C} \qquad C_2 \approx \mathbf{H} \text{ (the division ring of quaternions)}$$
$$C'_1 \approx \mathbf{R} \times \mathbf{R} \qquad C'_2 \approx M_2(\mathbf{R}) \text{ (2 } \times \text{ 2 matrices over } \mathbf{R})$$

20. Establish isomorphisms:

$$\mathbf{C} \otimes_{\mathbf{R}} \mathbf{C} \approx \mathbf{C} \times \mathbf{C}; \quad \mathbf{C} \otimes_{\mathbf{R}} \mathbf{H} \approx M_2(\mathbf{C}); \quad \mathbf{H} \otimes_{\mathbf{R}} \mathbf{H} \approx M_4(\mathbf{R})$$

where $M_d(F) = d \times d$ matrices over F. For the third one, with $\mathbf{H} \otimes \mathbf{H}$, define an isomorphism

$$f : \mathbf{H} \otimes_{\mathbf{R}} \mathbf{H} \to \mathrm{Hom}_{\mathbf{R}}(\mathbf{H}, \mathbf{H}) \approx M_4(\mathbf{R})$$

by $f(x \otimes y)(z) = xz\bar{y}$, where if $y = y_0 + y_1 i + y_2 j + y_3 k$ then

$$\bar{y} = y_0 - y_1 i - y_2 j - y_3 k.$$

21. (a) Establish isomorphisms

$$C_{n+2} \approx C'_n \otimes C_2 \qquad \text{and} \qquad C'_{n+2} \approx C_n \otimes C'_2.$$

[*Hint*: Let $\{e_1, \ldots, e_{n+2}\}$ be the orthonormalized basis with $e_i^2 = -1$. Then for the first isomorphism map $e_i \mapsto e'_i \otimes e_1 e_2$ for $i = 1, \ldots, n$ and map e_{n+1}, e_{n+2} on $1 \otimes e_1$ and $1 \otimes e_2$ respectively.]
    (b) Prove that $C_{n+8} \approx C_n \otimes M_{16}(\mathbf{R})$ (which is called the **periodicity property**).
    (c) Conclude that C_n is a semi-simple algebra over $\mathbf{R}$ for all n.

From (c) one can tabulate the simple modules over C_n. See [ABS 64], reproduced in Husemoller [Hu 75], Chapter 11, §6.

Part Four

HOMOLOGICAL ALGEBRA

In the forties and fifties (mostly in the works of Cartan, Eilenberg, MacLane, and Steenrod, see [CaE 57]), it was realized that there was a systematic way of developing certain relations of linear algebra, depending only on fairly general constructions which were mostly arrow-theoretic, and were affectionately called **abstract nonsense** by Steenrod. (For a more recent text, see [Ro 79].) The results formed a body of algebra, some of it involving homological algebra, which had arisen in topology, algebra, partial differential equations, and algebraic geometry. In topology, some of these constructions had been used in part to get homology and cohomology groups of topological spaces as in Eilenberg-Steenrod [ES 52]. In algebra, factor sets and 1-cocycles had arisen in the theory of group extensions, and, for instance, Hilbert's Theorem 90. More recently, homological algebra has entered in the cohomology of groups and the representation theory of groups. See for example Curtis-Reiner [CuR 81], and any book on the cohomology of groups, e.g. [La 96], [Se 64], and [Sh 72]. Note that [La 96] was written to provide background for class field theory in [ArT 68].

From an entirely different direction, Leray developed a theory of sheaves and spectral sequences motivated by partial differential equations. The basic theory of sheaves was treated in Godement's book on the subject [Go 58]. Fundamental insights were also given by Grothendieck in homological algebra [Gro 57], to be applied by Grothendieck in the theory of sheaves over schemes in the fifties and sixties. In Chapter XX, I have included whatever is necessary of homological algebra for Hartshorne's use in [Ha 77]. Both Chapters XX and XXI give an appropriate background for the homological algebra used in Griffiths-Harris [GrH 78], Chapter 5 (especially §3 and §4), and Gunning [Gu 90]. Chapter XX carries out the general theory of derived functors. The exercises and Chapter XXI may be viewed as providing examples and computations in specific concrete instances of more specialized interest.

The commutative algebra of Chapter X and the two chapters on homological algebra in this fourth part also provide an appropriate background for certain topics in algebraic geometry such as Serre's study of intersection theory [Se 65], Grothendieck duality, and Grothendieck's Riemann-Roch theorem in algebraic geometry. See for instance [SGA 6].

Finally I want to draw attention to the use of homological algebra in certain areas of partial differential equations, as in the papers of Atiyah-Bott-Patodi and Atiyah-Singer on complexes of elliptic operators. Readers can trace some of the literature from the bibliography given in [ABP 73].

The choice of material in this part was to a large extent motivated by all the above applications.

For this chapter, considering the number of references and cross-references given, the bibliography for the entire chapter is placed at the end of the chapter.

CHAPTER XX

General Homology Theory

To a large extent the present chapter is arrow-theoretic. There is a substantial body of linear algebra which can be formalized very systematically, and constitutes what Steenrod called abstract nonsense, but which provides a well-oiled machinery applicable to many domains. References will be given along the way.

Most of what we shall do applies to abelian categories, which were mentioned in Chapter III, end of §3. However, in first reading, I recommend that readers disregard any allusions to general abelian categories and assume that we are dealing with an abelian category of modules over a ring, or other specific abelian categories such as complexes of modules over a ring.

§1. COMPLEXES

Let A be a ring. By an **open complex** of A-modules, one means a sequence of modules and homomorphisms $\{(E^i, d^i)\}$,

$$\to E^{i-1} \xrightarrow{d^{i-1}} E^i \xrightarrow{d^i} E^{i+1} \to$$

where i ranges over all integers and d_i maps E^i into E^{i+1}, and such that

$$d^i \circ d^{i-1} = 0$$

for all i.

One frequently considers a finite sequence of homomorphisms, say

$$E^1 \to \cdots \to E^r$$

such that the composite of two successive ones is 0, and one can make this sequence into a complex by inserting 0 at each end:

$$\to 0 \to 0 \to E^1 \to \cdots \to E^r \to 0 \to 0 \to$$

Such a complex is called a **finite** or **bounded** complex.

Remark. Complexes can be indexed with a descending sequence of integers, namely,

$$\to E_{i+1} \xrightarrow{d_{i+1}} E_i \xrightarrow{d_i} E_{i-1} \to$$

When that notation is used systematically, then one uses upper indices for complexes which are indexed with an ascending sequence of integers:

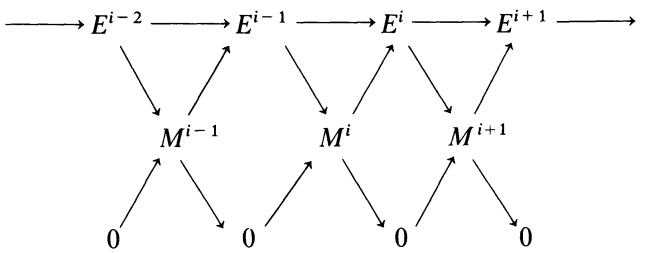
$$\to E^{i-1} \xrightarrow{d^{i-1}} E^i \xrightarrow{d^i} E^{i+1} \to$$

In this book, I shall deal mostly with ascending indices.

As stated in the introduction of this chapter, instead of modules over a ring, we could have taken objects in an arbitrary abelian category.

The homomorphisms d^i are often called **differentials**, because some of the first complexes which arose in practice were in analysis, with differential operators and differential forms. *Cf.* the examples below.

We denote a complex as above by (E, d). If the complex is exact, it is often useful to insert the kernels and cokernels of the differentials in a diagram as follows, letting $M_i = \operatorname{Ker} d^i = \operatorname{Im} d^{i-1}$.



Thus by definition, we obtain a family of short exact sequences

$$0 \to M^i \to E^i \to M^{i+1} \to 0.$$

If the complex is not exact, then of course we have to insert both the image of d^{i-1} and the kernel of d^i. The factor

$$(\operatorname{Ker} d^i)/(\operatorname{Im} d^{i-1})$$

will be studied in the next section. It is called the **homology of the complex**, and measures the deviation from exactness.

Let M be a module. By a **resolution** of M we mean an exact sequence

$$\to E_n \to E_{n-1} \to \cdots \to E_0 \to M \to 0.$$

Thus a resolution is an exact complex whose furthest term on the right before 0 is M. The resolution is indexed as shown. We usually write E_M for the part of complex formed only of the E_i's, thus:

$$E_M \text{ is:} \quad \to E_n \to E_{n-1} \to \cdots \to E_0,$$

stopping at E_0. We then write E for the complex obtained by sticking 0 on the right:

$$E \text{ is:} \quad \to E_n \to E_{n-1} \to \cdots \to E_0 \to 0.$$

If the objects E_i of the resolution are taken in some family, then the resolution is qualified in the same way as the family. For instance, if E_i is free for all $i \geqq 0$ then we say that the **resolution** is a **free resolution**. If E_i is projective for all $i \geqq 0$ then we say that the **resolution** is **projective**. And so forth. The same terminology is applied to the right, with a resolution

$$0 \to M \to E^0 \to E^1 \to \cdots \to E^{n-1} \to E^n \to,$$

also written

$$0 \to M \to E_M.$$

We then write E for the complex

$$0 \to E^0 \to E^1 \to E^2 \to \cdots.$$

See §5 for injective resolutions.

A resolution is said to be **finite** if E_i (or E^i) $= 0$ for all but a finite number of indices i.

Example. Every module admits a free resolution (on the left). This is a simple application of the notion of free module. Indeed, let M be a module, and let $\{x_j\}$ be a family of generators, with j in some indexing set J. For each j let Re_j be a free module over R with a basis consisting of one element e_j. Let

$$F = \bigoplus_{j \in J} Re_j$$

be their direct sum. There is a unique epimorphism

$$F \to M \to 0$$

sending e_j on x_j, Now we let M_1 be the kernel, and again represent M_1 as the quotient of a free module. Inductively, we can construct the desired free resolution.

Example. The Standard Complex. Let S be a set. For $i = 0, 1, 2, \ldots$ let E_i be the free module over $\mathbf{Z}$ generated by $(i+1)$-tuples $(x_0, \ldots, x_i)$ with $x_0, \ldots, x_i \in S$. Thus such $(i+1)$-tuples form a basis of E_i over $\mathbf{Z}$. There is a unique homomorphism

$$d_{i+1} : E_{i+1} \to E_i$$

such that

$$d_{i+1}(x_0, \ldots, x_{i+1}) = \sum_{j=0}^{i+1} (-1)^j (x_0, \ldots, \hat{x}_j, \ldots, x_{i+1}),$$

where the symbol $\hat{x}_j$ means that this term is to be omitted. For $i = 0$, we define $d_0 : E_0 \to \mathbf{Z}$ to be the unique homomorphism such that $d_0(x_0) = 1$. The map d_0 is sometimes called the augmentation, and is also denoted by ε. Then we obtain a resolution of $\mathbf{Z}$ by the complex

$$\to E_{i+1} \to E_i \to \cdots \to E_0 \xrightarrow{\varepsilon} \mathbf{Z} \to 0.$$

The formalism of the above maps d_i is pervasive in mathematics. See Exercise 2 for the use of the standard complex in the cohomology theory of groups. For still another example of this same formalism, compare with the Koszul complex in Chapter XXI, §4.

Given a module M, one may form $\operatorname{Hom}(E_i, M)$ for each i, in which case one gets coboundary maps

$$\delta^i : \operatorname{Hom}(E_i, M) \to \operatorname{Hom}(E_{i+1}, M), \quad \delta(f) = f \circ d^{i+1},$$

obtained by composition of mappings. This procedure will be used to obtain derived functors in §6. In Exercises 2 through 6, you will see how this procedure is used to develop the cohomology theory of groups.

Instead of using homomorphisms, one may use a topological version with simplices, and continuous maps, in which case the standard complex gives rise to the singular homology theory of topological spaces. See [GreH 81], Chapter 9.

Examples. Finite free resolutions. In Chapter XXI, you will find other examples of complexes, especially finite free, constructed in various ways with different tools. This subsequent entire chapter may be viewed as providing examples for the current chapter.

Examples with differential forms. In Chapter XIX, §3, we gave the example of the de Rham complex in an algebraic setting. In the theory of differential manifolds, the de Rham complex has differential maps

$$d^i : \Omega^i \to \Omega^{i+1},$$

sending differential forms of degree i to those of degree $i+1$, and allows for the computation of the homology of the manifold.

A similar situation occurs in complex differential geometry, when the maps d^i are given by the **Dolbeault** $\bar{\partial}$-operators

$$\bar{\partial}^i : \Omega^{p,i} \to \Omega^{p,i+1}$$

operating on forms of type (p, i). Interested readers can look up for instance Gunning's book [Gu 90] mentioned in the introduction to Part IV, Volume I, E. The associated homology of this complex is called the **Dolbeault** or **$\bar{\partial}$-cohomology** of the complex manifold.

Let us return to the general algebraic aspects of complexes and resolutions.

It is an interesting problem to discuss which modules admit finite resoutions, and variations on this theme. Some conditions are discussed later in this chapter and in Chapter XXI. If a resolution

$$0 \to E_n \to E_{n-1} \to \cdots \to E_0 \to M \to 0$$

is such that $E_m = 0$ for $m > n$, then we say that the resolution has **length** $\leqq n$ (sometimes we say it has **length** n by abuse of language).

A **closed complex** of A-modules is a sequence of modules and homomorphisms $\{(E^i, d^i)\}$ where i ranges over the set of integers mod n for some $n \geqq 2$ and otherwise satisfying the same properties as above. Thus a closed complex looks like this:

$$E^1 \to E^2 \to \cdots \to E^n \to E^1$$

We call n the **length** of the closed complex.

Without fear of confusion, one can omit the index i on d^i and write just d. We also write (E, d) for the complex $\{(E^i, d^i)\}$, or even more briefly, we write simply E.

Let (E, d) and (E', d') be complexes (both open or both closed). Let r be an integer. A **morphism** or **homomorphism** (of complexes)

$$f : (E', d') \to (E, d)$$

of **degree** r is a sequence

$$f_i : E'^i \to E^{i+r}$$

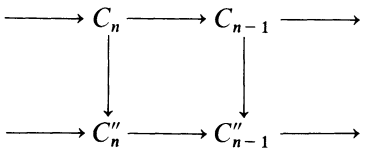
of homomorphisms such that for all i the following diagram is commutative:

$$\begin{array}{ccc} E'^{(i-1)} & \xrightarrow{f_{i-1}} & E^{i-1+r} \\ {\scriptstyle d'}\downarrow & & \downarrow{\scriptstyle d} \\ E'^i & \xrightarrow[f_i]{} & E^{i+r} \end{array}$$

Just as we write d instead of d^i, we shall also write f instead of f_i. If the complexes are closed, we define a morphism from one into the other only if they have the same length.

It is clear that complexes form a category. In fact they form an abelian category. Indeed, say we deal with complexes indexed by $\mathbf{Z}$ for simplicity, and morphisms of degree 0. Say we have a morphism of complexes $f : C \to C''$ or

putting the indices:



We let $C'_n = \mathrm{Ker}(C_n \to C''_n)$. Then the family (C'_n) forms a complex, which we define to be the kernel of f. We let the reader check the details that this and a similar definition for cokernel and finite direct sums make complexes of modules into an abelian category. At this point, readers should refer to Chapter III, §9, where kernels and cokernels are discussed in this context. The snake lemma of that chapter will now become central to the next section.

It will be useful to have another notion to deal with objects indexed by a monoid. Let G be a monoid, which we assume commutative and additive to fit the applications we have in mind here. Let $\{M_i\}_{i \in G}$ be a family of modules indexed by G. The direct sum

$$M = \bigoplus_{i \in G} M_i$$

will be called the **G-graded module associated with the family** $\{M_i\}_{i \in G}$. Let $\{M_i\}_{i \in G}$ and $\{M'_i\}_{i \in G}$ be families indexed by G, and let M, M' be their associated G-graded modules. Let $r \in G$. By a **G-graded morphism** $f : M' \to M$ of degree r we shall mean a homomorphism such that f maps M'_i into M_{i+r} for each $i \in G$ (identifying M_i with the corresponding submodule of the direct sum on the i-th component). Thus f is nothing else than a family of homomorphisms $f_i : M'_i \to M_{i+r}$.

If (E, d) is a complex we may view E as a G-graded module (taking the direct sum of the components of the complex), and we may view d as a G-graded morphism of degree 1, letting G be $\mathbf{Z}$ or $\mathbf{Z}/n\mathbf{Z}$. The most common case we encounter is when $G = \mathbf{Z}$. Then we write the complex as

$$E = \bigoplus E_i, \quad \text{and} \quad d : E \to E$$

maps E into itself. The differential d is defined as d_i on each direct summand E_i, and has degree 1.

Conversely, if G is $\mathbf{Z}$ or $\mathbf{Z}/n\mathbf{Z}$, one may view a G-graded module as a complex, by defining d to be the zero map.

For simplicity, we shall often omit the prefix "G-graded" in front of the word "morphism", when dealing with G-graded morphisms.

§2. HOMOLOGY SEQUENCE

Let (E, d) be a complex. We let

$$Z^i(E) = \operatorname{Ker} d^i$$

and call $Z^i(E)$ the module of ***i*-cycles**. We let

$$B^i(E) = \operatorname{Im} d^{i-1}$$

and call $B^i(E)$ the module of ***i*-boundaries**. We frequently write Z^i and B^i instead of $Z^i(E)$ and $B^i(E)$, respectively. We let

$$H^i(E) = Z^i/B^i = \operatorname{Ker} d^i/\operatorname{Im} d^{i-1},$$

and call $H^i(E)$ the i-th **homology group** of the complex. The graded module associated with the family $\{H^i\}$ will be denoted by $H(E)$, and will be called the **homology** of E. One sometimes writes $H^*(E)$ instead of $H(E)$.

If $f: E' \to E$ is a morphism of complexes, say of degree 0, then we get an **induced canonical homomorphism**

$$H^i(f) : H^i(E') \to H^i(E)$$

on each homology group. Indeed, from the commutative diagram defining a morphism of complexes, one sees at once that f maps $Z^i(E')$ into $Z^i(E)$ and $B^i(E')$ into $B^i(E)$, whence the induced homomorphism $H^i(f)$. Compare with the beginning remarks of Chapter III, §9. One often writes this induced homomorphism as f_{i*} rather than $H_i(f)$, and if $H(E)$ denotes the graded module of homology as above, then we write

$$H(f) = f_* : H(E') \to H(E).$$

We call $H(f)$ the map **induced** by f on homology. If $H^i(f)$ is an isomorphism for all i, then we say that f is a **homology isomorphism**.

Note that if $f: E' \to E$ and $g: E \to E''$ are morphisms of complexes, then it is immediately verified that

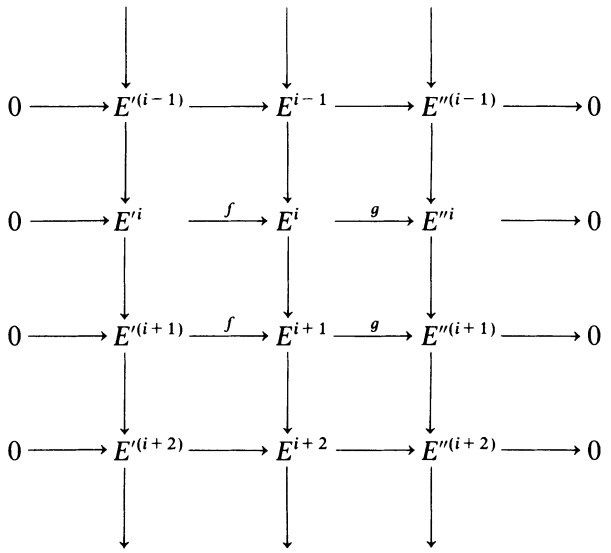
$$H(g) \circ H(f) = H(g \circ f) \qquad \text{and} \qquad H(\mathrm{id}) = \mathrm{id}.$$

Thus H is a functor from the category of complexes to the category of graded modules.

We shall consider short exact sequences of complexes with morphisms of degree 0:

$$0 \to E' \xrightarrow{f} E \xrightarrow{g} E'' \to 0,$$

which written out in full look like this:



One can define a morphism

$$\delta : H(E'') \to H(E')$$

of degree 1, *in other words, a family of homomorphisms*

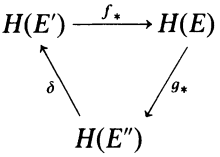
$$\delta^i : H''^i \to H'^{(i+1)}$$

by the snake lemma.

Theorem 2.1. *Let*

$$0 \to E' \xrightarrow{f} E \xrightarrow{g} E'' \to 0$$

be an exact sequence of complexes with f, g of degree 0. *Then the sequence*



is exact.

This theorem is merely a special application of the snake lemma.

If one writes out in full the homology sequence in the theorem, then it looks like this:

$$\xrightarrow{\delta} H'^i \to H^i \to H''^i \xrightarrow{\delta} H'^{(i+1)} \to H^{i+1} \to H''^{(i+1)} \xrightarrow{\delta}$$

It is clear that our map δ is functorial (in an obvious sense), and hence that our whole structure (H, δ) is a functor from the category of short exact sequences of complexes into the category of complexes.

§3. EULER CHARACTERISTIC AND THE GROTHENDIECK GROUP

This section may be viewed as a continuation of Chapter III, §8, on Euler-Poincaré maps. Consider complexes of A-modules, for simplicity.

Let E be a complex such that almost all homology groups H^i are equal to 0. Assume that E is an open complex. As in Chapter III, §8, let φ be an Euler-Poincaré mapping on the category of modules (i.e. A-modules). We define the **Euler-Poincaré characteristic** $\chi_\varphi(E)$ (or more briefly the **Euler characteristic**) with respect to φ, to be

$$\chi_\varphi(E) = \sum (-1)^i \varphi(H^i)$$

provided $\varphi(H^i)$ is defined for all H^i, in which case we say that χ_φ is **defined** for the complex E.

If E is a closed complex, we select a definite order $(E^1, \ldots, E^n)$ for the integers mod n and define the Euler characteristic by the formula

$$\chi_\varphi(E) = \sum_{i=1}^{n} (-1)^i \varphi(H^i)$$

provided again all $\varphi(H^i)$ are defined.

For an example, the reader may refer to Exercise 28 of Chapter I.

One may view H as a complex, defining d to be the zero map. In that case, we see that $\chi_\varphi(H)$ is the alternating sum given above. More generally:

Theorem 3.1. *Let F be a complex, which is of even length if it is closed. Assume that $\varphi(F^i)$ is defined for all i, $\varphi(F^i) = 0$ for almost all i, and $H^i(F) = 0$ for almost all i. Then $\chi_\varphi(F)$ is defined, and*

$$\chi_\varphi(F) = \sum_i (-1)^i \varphi(F^i).$$

Proof. Let Z^i and B^i be the groups of i-cycles and i-boundaries in F^i respectively. We have an exact sequence

$$0 \to Z^i \to F^i \to B^{i+1} \to 0.$$

Hence $\chi_\varphi(F)$ is defined, and

$$\varphi(F^i) = \varphi(Z^i) + \varphi(B^{i+1}).$$

Taking the alternating sum, our conclusion follows at once.

A complex whose homology is trivial is called **acyclic**.

Corollary 3.2. *Let F be an acyclic complex, such that $\varphi(F^i)$ is defined for all i, and equal to 0 for almost all i. If F is closed, we assume that F has even length. Then*

$$\chi_\varphi(F) = 0.$$

In many applications, an open complex F is such that $F^i = 0$ for almost all i, and one can then treat this complex as a closed complex by defining an additional map going from a zero on the far right to a zero on the far left. Thus in this case, the study of such an open complex is reduced to the study of a closed complex.

Theorem 3.3. *Let*

$$0 \to E' \to E \to E'' \to 0$$

be an exact sequence of complexes, with morphisms of degree 0. If the complexes are closed, assume that their length is even. Let φ be an Euler-Poincaré mapping on the category of modules. If χ_φ is defined for two of the above three complexes, then it is defined for the third, and we have

$$\chi_\varphi(E) = \chi_\varphi(E') + \chi_\varphi(E'').$$

Proof. We have an exact homology sequence

$$\to H''^{(i-1)} \to H'^i \to H^i \to H''^i \to H'^{(i+1)} \to$$

This homology sequence is nothing but a complex whose homology is trivial. Furthermore, each homology group belonging say to E is between homology groups of E' and E''. Hence if χ_φ is defined for E' and E'' it is defined for E. Similarly for the other two possibilities. If our complexes are closed of even length n, then this homology sequence has even length $3n$. We can therefore apply the corollary of Theorem 3.1 to get what we want.

For certain applications, it is convenient to construct a universal Euler mapping. Let $\mathfrak{A}$ be the set of isomorphism classes of certain modules. If E is a module, let $[E]$ denote its isomorphism class. We require that $\mathfrak{A}$ satisfy the **Euler-Poincaré condition**, i.e. if we have an exact sequence

$$0 \to E' \to E \to E'' \to 0,$$

then $[E]$ is in $\mathfrak{A}$ if and only if $[E']$ and $[E'']$ are in $\mathfrak{A}$. Furthermore, the zero module is in $\mathfrak{A}$.

Theorem 3.4. *Assume that* $\mathfrak{A}$ *satisfies the Euler-Poincaré condition. Then there is a map*

$$\gamma : \mathfrak{A} \to \mathbf{K}(\mathfrak{A})$$

of $\mathfrak{A}$ *into an abelian group* $\mathbf{K}(\mathfrak{A})$ *having the universal property with respect to Euler-Poincaré maps defined on* $\mathfrak{A}$.

To construct this, let $F_{ab}(\mathfrak{A})$ be the free abelian group generated by the set of such $[E]$. Let B be the subgroup generated by all elements of type

$$[E] - [E'] - [E''],$$

where

$$0 \to E' \to E \to E'' \to 0$$

is an exact sequence whose members are in $\mathfrak{A}$. We let $\mathbf{K}(\mathfrak{A})$ be the factor group $F_{ab}(\mathfrak{A})/B$, and let $\gamma : \mathfrak{A} \to \mathbf{K}(\mathfrak{A})$ be the natural map. It is clear that γ has the universal property.

We observe the similarity of construction with the Grothendieck group of a monoid. In fact, the present group is known as the **Euler-Grothendieck group** of $\mathfrak{A}$, with Euler usually left out.

The reader should observe that the above arguments are valid in abelian categories, although we still used the word **module**. Just as with the elementary isomorphism theorems for groups, we have the analogue of the Jordan-Hölder theorem for modules. Of course in the case of modules, we don't have to worry about the normality of submodules.

We now go a little deeper into **K**-theory. Let $\mathfrak{A}$ be an abelian category. In first reading, one may wish to limit attention to an abelian category of modules over a ring. Let $\mathfrak{C}$ be a family of objects in $\mathfrak{A}$. We shall say that $\mathfrak{C}$ is a **K-family** if it satisfies the following conditions.

K 1. $\mathfrak{C}$ is closed under taking finite direct sums, and 0 is in $\mathfrak{C}$.

K 2. Given an object E in $\mathfrak{A}$ there exists an epimorphism

$$L \to E \to 0$$

with L in $\mathfrak{C}$.

K 3. Let E be an object admitting a finite resolution of length n

$$0 \to L_n \to \cdots \to L_0 \to E \to 0$$

with $L_i \in \mathfrak{C}$ for all i. If

$$0 \to N \to F_{n-1} \to \cdots \to F_0 \to E \to 0$$

is a resolution with N in $\mathfrak{A}$ and $F_0, \ldots, F_{n-1}$ in $\mathfrak{C}$, then N is also in $\mathfrak{C}$.

We note that it follows from these axioms that if F is in $\mathcal{C}$ and F' is isomorphic to F, then F' is also in $\mathcal{C}$, as one sees by looking at the resolution

$$0 \to F' \to F \to 0 \to 0$$

and applying **K 3**. Furthermore, given an exact sequence

$$0 \to F' \to F \to F'' \to 0$$

with F and F'' in $\mathcal{C}$, then F' is in $\mathcal{C}$, again by applying **K 3**.

Example. One may take for $\mathcal{A}$ the category of modules over a commutative ring, and for $\mathcal{C}$ the family of projective modules. Later we shall also consider Noetherian rings, in which case one may take finite modules, and finite projective modules instead. Condition **K 2** will be discussed in §8.

From now on we assume that $\mathcal{C}$ is a **K**-family. For each object E in $\mathcal{A}$, we let $[E]$ denote its isomorphism class. An object E of $\mathcal{A}$ will be said to have **finite $\mathcal{C}$-dimension** if it admits a finite resolution with elements of $\mathcal{C}$. We let $\mathcal{A}(\mathcal{C})$ be the family of objects in $\mathcal{A}$ which are of finite $\mathcal{C}$-dimension. We may then form the

$$\mathbf{K}(\mathcal{A}(\mathcal{C})) = \mathbf{Z}[\mathcal{A}(\mathcal{C})]/R(\mathcal{A}(\mathcal{C}))$$

where $R(\mathcal{A}(\mathcal{C}))$ is the group generated by all elements $[E] - [E'] - [E'']$ arising from an exact sequence

$$0 \to E' \to E \to E'' \to 0$$

in $\mathcal{A}(\mathcal{C})$. Similarly we define

$$\mathbf{K}(\mathcal{C}) = \mathbf{Z}[(\mathcal{C})]/R(\mathcal{C}),$$

where $R(\mathcal{C})$ is the group of relations generated as above, but taking E', E, E'' in $\mathcal{C}$ itself.

There are natural maps

$$\gamma_{\mathcal{A}(\mathcal{C})} : \mathcal{A}(\mathcal{C}) \to \mathbf{K}(\mathcal{A}(\mathcal{C})) \quad \text{and} \quad \gamma_{\mathcal{C}} : \mathcal{C} \to \mathbf{K}(\mathcal{C}),$$

which to each object associate its class in the corresponding Grothendieck group. There is also a natural homomorphism

$$\epsilon : \mathbf{K}(\mathcal{C}) \to \mathbf{K}(\mathcal{A}(\mathcal{C}))$$

since an exact sequence of objects of $\mathcal{C}$ can also be viewed as an exact sequence of objects of $\mathcal{A}(\mathcal{C})$.

Theorem 3.5. *Let $M \in \mathfrak{A}(\mathcal{C})$ and suppose we have two resolutions*

$$L_M \to M \to 0 \quad \text{and} \quad L'_M \to M \to 0,$$

by finite complexes L_M and L'_M in $\mathcal{C}$. Then

$$\sum (-1)^i \gamma_e(L_i) = \sum (-1)^i \gamma_e(L'_i).$$

Proof. Take first the special case when there is an epimorphism $L'_M \to L_M$, with kernel E illustrated on the following commutative and exact diagram.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & E & \longrightarrow & L'_M & \longrightarrow & L_M & \longrightarrow & 0 \\ & & & & \downarrow & & \downarrow & & \\ & & & & M & \xrightarrow[\text{id}]{} & M & & \\ & & & & \downarrow & & \downarrow & & \\ & & & & 0 & & 0 & & \end{array}$$

The kernel is a complex

$$0 \to E_n \to E_{n-1} \to \cdots \to E_0 \to 0$$

which is exact because we have the homology sequence

$$H_p(E) \to H_p(L') \to H_p(L) \to H_{p-1}(E)$$

For $p \geqq 1$ we have $H_p(L) = H_p(L') = 0$ by definition, so $H_p(E) = 0$ for $p \geqq 1$. And for $p = 0$ we consider the exact sequence

$$H_1(L) \to H_0(E) \to H_0(L') \to H_0(L)$$

Now we have $H_1(L) = 0$, and $H_0(L') \to H_0(L)$ corresponds to the identity morphisms on M so is an isomorphism. It follows that $H_0(E) = 0$ also.

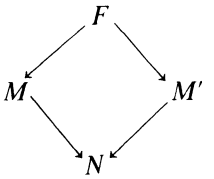
By definition of **K**-family, the objects E_p are in $\mathcal{C}$. Then taking the Euler characteristic in $\mathbf{K}(\mathcal{C})$ we find

$$\chi(L') - \chi(L) = \chi(E) = 0$$

which proves our assertion in the special case.

The general case follows by showing that given two resolutions of M in $\mathcal{C}$ we can always find a third one which tops both of them. The pattern of our construction will be given by a lemma.

Lemma 3.6. *Given two epimorphisms $u: M \to N$ and $v: M' \to N$ in $\mathfrak{A}$, there exist epimorphisms $F \to M$ and $F \to M'$ with F in $\mathfrak{C}$ making the following diagram commutative.*

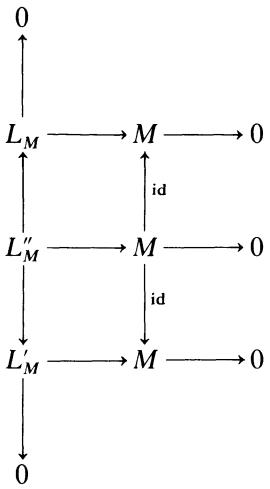


Proof. Let $E = M \times_N M'$, that is E is the kernel of the morphism

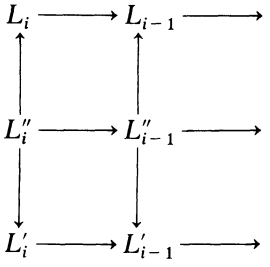
$$M \times M' \to N$$

given by $(x, y) \mapsto ux - vy$. (Elements are not really used here, and we could write formally $u - v$ instead.) There is some F in $\mathfrak{C}$ and an epimorphism $F \to E \to 0$. The composition of this epimorphism with the natural projections of E on each factor gives us what we want.

We construct a complex L''_M giving a resolution of M with a commutative and exact diagram:



The construction is done inductively, so we put indices:



Suppose that we have constructed up to L''_{i-1} with the desired epimorphisms on L_{i-1} and L'_{i-1}. We want to construct L''_i. Let $B_i = \operatorname{Ker}(L_{i-1} \to L_{i-2})$ and similarly for B'_i and B''_i. We obtain the commutative diagram:

$$\begin{array}{ccccccc}
L_i & \longrightarrow & B_i & \longrightarrow & L_{i-1} & \longrightarrow & L_{i-2} \\
 & & \uparrow & & \uparrow & & \uparrow \\
 & & B''_i & \longrightarrow & L''_{i-1} & \longrightarrow & L''_{i-2} \\
 & & \downarrow & & \downarrow & & \downarrow \\
L'_i & \longrightarrow & B'_i & \longrightarrow & L'_{i-1} & \longrightarrow & L'_{i-2}
\end{array}$$

If $B''_i \to B_i$ or $B''_i \to B'_i$ are not epimorphisms, then we replace L''_{i-1} by

$$L''_{i-1} \oplus L_i \oplus L'_i.$$

We let the boundary map to L''_{i-2} be 0 on the new summands, and similarly define the maps to L_{i-1} and L'_{i-1} to be 0 on L'_i and L_{i-1} respectively.

Without loss of generality we may now assume that

$$B''_i \to B_i \quad \text{and} \quad B''_i \to B'_i$$

are epimorphisms. We then use the construction of the preceding lemma. We let

$$E_i = L_i \oplus_{B_i} B''_i \quad \text{and} \quad E'_i = B''_i \oplus_{B'_i} L'_i.$$

Then both E_i and E'_i have natural epimorphisms on B''_i. Then we let

$$N_i = E_i \oplus_{B''_i} E'_i$$

and we find an object L''_i in $\mathcal{C}$ with an epimorphism $L''_i \to N_i$. This gives us the inductive construction of L'' up to the very end. To stop the process, we use **K 3** and take the kernel of the last constructed L''_i to conclude the proof.

Theorem 3.7. *The natural map*

$$\epsilon : \mathbf{K}(\mathcal{C}) \to \mathbf{K}(\mathfrak{A}(\mathcal{C}))$$

is an isomorphism.

Proof. The map is surjective because given a resolution

$$0 \to F_n \to \cdots \to F_0 \to M \to 0$$

with $F_i \in \mathcal{C}$ for all i, the element

$$\sum (-1)^i \gamma_{\mathcal{C}}(F_i)$$

maps on $\gamma_{\mathfrak{a}(\mathcal{C})}(M)$ under ϵ. Conversely, Theorem 3.5 shows that the association

$$M \mapsto \sum (-1)^i \gamma_{\mathcal{C}}(F_i)$$

is a well-defined mapping. Since for any $L \in \mathcal{C}$ we have a short exact sequence $0 \to L \to L \to 0$, it follows that this mapping following ϵ is the identity on $\mathbf{K}(\mathcal{C})$, so ϵ is a monomorphism. Hence ϵ is an isomorphism, as was to be shown.

It may be helpful to the reader actually to see the next lemma which makes the additivity of the inverse more explicit.

Lemma 3.8. *Given an exact sequence in* $\mathfrak{A}(\mathcal{C})$

$$0 \to M' \to M \to M'' \to 0$$

there exists a commutative and exact diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & L_{M'} & \longrightarrow & L_M & \longrightarrow & L_{M''} & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & M' & \longrightarrow & M & \longrightarrow & M'' & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

with finite resolutions $L_{M'}$, L_M, $L_{M''}$ *in* $\mathcal{C}$.

Proof. We first show that we can find L', L, L'' in $\mathcal{C}$ to fit an exact and commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & L' & \longrightarrow & L & \longrightarrow & L'' & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & M' & \longrightarrow & M & \longrightarrow & M'' & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & & 0 & &
\end{array}$$

We first select an epimorphism $L'' \to M''$ with L'' in $\mathcal{C}$. By Lemma 3.6 there exists $L_1 \in \mathcal{C}$ and epimorphisms $L_1 \to M$, $L_1 \to L''$ making the diagram commutative. Then let $L_2 \to M'$ be an epimorphism with $L_2 \in \mathcal{C}$, and finally define $L = L_1 \oplus L_2$. Then we get morphisms $L \to M$ and $L \to L''$ in the obvious way. Let L' be the kernel of $L \to L''$. Then $L_2 \subset L'$ so we get an epimorphism $L' \to M'$.

This now allows us to construct resolutions inductively until we hit the n-th step, where n is some integer such that M, M'' admit resolutions of length n in $\mathcal{C}$. The last horizontal exact sequence that we obtain is

$$0 \to L'_n \to L_n \to L''_n \to 0$$

and L''_n can be chosen to be the kernel of $L''_{n-1} \to L''_{n-2}$. By **K 3** we know that L''_n lies in $\mathcal{C}$, and the sequence

$$0 \to L''_n \to L''_{n-1}$$

is exact. This implies that in the next inductive step, we can take $L''_{n+1} = 0$. Then

$$0 \to L'_{n+1} \to L_{n+1} \to 0 \to 0$$

is exact, and at the next step we just take the kernels of the vertical arrows to complete the desired finite resolutions in $\mathcal{C}$. This concludes the proof of the lemma.

Remark. The argument in the proof of Lemma 3.8 in fact shows:

If

$$0 \to M' \to M \to M'' \to 0$$

is an exact sequence in $\mathfrak{A}$, and if M, M'' have finite $\mathcal{C}$-dimension, then so does M'.

In the category of modules, one has a more precise statement:

Theorem 3.9. *Let $\mathfrak{A}$ be the category of modules over a ring. Let $\mathcal{P}$ be the family of projective modules. Given an exact sequence of modules*

$$0 \to E' \to E \to E'' \to 0$$

if any two of E', E, E'' admit finite resolutions in $\mathcal{P}$ then the third does also.

Proofs in a more subtle case will be given in Chapter XXI, Theorem 2.7.

Next we shall use the tensor product to investigate a ring structure on the Grothendieck group. We suppose for simplicity that we deal with an abelian category of modules over a commutative ring, denoted by $\mathfrak{A}$, together with a **K**-family $\mathcal{C}$ as above, but we now assume that $\mathfrak{A}$ is closed under the tensor product. The only properties we shall actually use for the next results are the following ones, denoted by **TG** (for "tensor" and "Grothendieck" respectively):

TG 1. There is a bifunctorial isomorphism giving commutativity

$$M \otimes N \approx N \otimes M$$

for all M, N in $\mathfrak{A}$; and similarly for distributivity over direct sums, and associativity.

TG 2. For all L in $\mathcal{C}$ the functor $M \mapsto L \otimes M$ is exact.

TG 3. If L, L' are in $\mathcal{C}$ then $L \otimes L'$ is in $\mathcal{C}$.

Then we may give $\mathbf{K}(\mathcal{C})$ the structure of an algebra by defining

$$\mathrm{cl}_{\mathcal{C}}(L)\,\mathrm{cl}_{\mathcal{C}}(L') = \mathrm{cl}_{\mathcal{C}}(L \otimes L').$$

Condition **TG 1** implies that this algebra is commutative, and we call it the **Grothendieck algebra**. In practice, there is a unit element, but if we want one in the present axiomatization, we have to make it an explicit assumption:

TG 4. There is an object R in $\mathcal{C}$ such that $R \otimes M \approx M$ for all M in $\mathcal{A}$.

Then $\mathrm{cl}_{\mathcal{C}}(R)$ is the unit element.

Similarly, condition **TG 2** shows that we can define a module structure on $\mathbf{K}(\mathcal{A})$ over $\mathbf{K}(\mathcal{C})$ by the same formula

$$\mathrm{cl}_{\mathcal{C}}(L)\,\mathrm{cl}_{\mathcal{A}}(M) = \mathrm{cl}_{\mathcal{A}}(L \otimes M),$$

and similarly $\mathbf{K}(\mathcal{A}(\mathcal{C}))$ is a module over $\mathbf{K}(\mathcal{C})$, where we recall that $\mathcal{A}(\mathcal{C})$ is the family of objects in $\mathcal{A}$ which admit finite resolutions by objects in $\mathcal{C}$.

Since we know from Theorem 3.7 that $\mathbf{K}(\mathcal{C}) \approx \mathbf{K}(\mathcal{A}(\mathcal{C}))$, we also have a ring structure on $\mathbf{K}(\mathcal{A}(\mathcal{C}))$ via this isomorphism. We then can make the product more explicit as follows.

Proposition 3.10. *Let $M \in \mathcal{A}(\mathcal{C})$ and let $N \in \mathcal{A}$. Let*

$$0 \to L_n \to \cdots \to L_0 \to M \to 0$$

be a finite resolution of M by objects in $\mathcal{C}$. Then

$$\mathrm{cl}_{\mathcal{C}}(M)\,\mathrm{cl}_{\mathcal{A}}(N) = \sum (-1)^i\,\mathrm{cl}_{\mathcal{A}}(L_i \otimes N).$$

$$= \sum (-1)^i\,\mathrm{cl}_{\mathcal{A}}(H_i(K))$$

where K is the complex

$$0 \to L_n \otimes N \to \cdots \to L_0 \otimes N \to M \otimes N \to 0$$

and $H_i(K)$ is the i-th homology of this complex.

Proof. The formulas are immediate consequences of the definitions, and of Theorem 3.1.

Example. Let $\mathcal{A}$ be the abelian category of modules over a commutative ring. Let $\mathcal{C}$ be the family of projective modules. From §6 on derived functors the reader will know that the homology of the complex K in Proposition 3.10 is just $\mathrm{Tor}(M, N)$. Therefore the formula in that proposition can also be written

$$\mathrm{cl}_{\mathcal{C}}(M)\,\mathrm{cl}_{\mathcal{A}}(N) = \sum (-1)^i\,\mathrm{cl}_{\mathcal{A}}(\mathrm{Tor}_i(M, N)).$$

Example. Let k be a field. Let G be a group. By a **(G, k)-module**, we shall mean a pair (E, ρ), consisting of a k-space E and a homomorphism

$$\rho: G \to \mathrm{Aut}_k(E).$$

Such a homomorphism is also called a **representation** of G in E. By abuse of language, we also say that the k-space E is a G-module. The group G operates on E, and we write σx instead of $\rho(\sigma)x$. The field k will be kept fixed in what follows.

Let $\mathrm{Mod}_k(G)$ denote the category whose objects are (G, k)-modules. A morphism in $\mathrm{Mod}_k(G)$ is what we call a **G-homomorphism**, that is a k-linear map $f: E \to F$ such that $f(\sigma x) = \sigma f(x)$ for all $\sigma \in G$. The group of morphisms in $\mathrm{Mod}_k(G)$ is denoted by Hom_G.

If E is a G-module, and $\sigma \in G$, then we have by definition a k-automorphism $\sigma: E \to E$. Since T^r is a functor, we have an induced automorphism

$$T^r(\sigma): T^r(E) \to T^r(E)$$

for each r, and thus $T^r(E)$ is also a G-module. Taking the direct sum, we see that $T(E)$ is a G-module, and hence that T is a functor from the category of G-modules to the category of graded G-modules. Similarly for $\bigwedge^r$, S^r, and $\bigwedge$, S.

It is clear that the kernel of a G-homomorphism is a G-submodule, and that the factor module of a G-module by a G-submodule is again a G-module so the category of G-modules is an abelian category.

We can now apply the general considerations on the Grothendieck group which we write

$$\mathbf{K}(G) = \mathbf{K}(\mathrm{Mod}_k(G))$$

for simplicity in the present case. We have the canonical map

$$\mathrm{cl}: \mathrm{Mod}_k(G) \to \mathbf{K}(G).$$

which to each G-module associates its class in $\mathbf{K}(G)$.

If E, F are G-modules, then their tensor product over k, $E \otimes F$, is also a G-module. Here again, the operation of G on $E \otimes F$ is given functorially. If $\sigma \in G$, there exists a unique k-linear map $E \otimes F \to E \otimes F$ such that for $x \in E$, $y \in F$ we have $x \otimes y \mapsto (\sigma x) \otimes (\sigma y)$. The tensor product induces a law of composition on $\mathrm{Mod}_k(G)$ because the tensor products of G-isomorphic modules are G-isomorphic.

Furthermore all the conditions **TG 1** through **TG 4** are satisfied. Since k is a field, we find also that tensoring an exact sequence of G-modules over k with any G-module over k preserves the exactness, so **TG 2** is satisfied for all (G, k)-modules. Thus the Grothendieck group $\mathbf{K}(G)$ is in fact the Grothendieck ring, or the Grothendieck algebra over k.

By Proposition 2.1 and Theorem 2.3 of Chapter XVIII, we also see:

The Grothendieck ring of a finite group G consisting of isomorphism classes of finite dimensional (G, k)-spaces over a field k of characteristic 0 is naturally isomorphic to the character ring $X_{\mathbf{Z}}(G)$.

We can axiomatize this a little more. We consider an abelian category of modules over a commutative ring R, which we denote by $\mathfrak{A}$ for simplicity. For two modules M, N in $\mathfrak{A}$ we let $\mathrm{Mor}(M, N)$ as usual be the morphisms in $\mathfrak{A}$, but $\mathrm{Mor}(M, N)$ is an abelian subgroup of $\mathrm{Hom}_R(M, N)$. For example, we could take $\mathfrak{A}$ to be the category of (G, k)-modules as in the example we have just discussed, in which case $\mathrm{Mor}(M, N) = \mathrm{Hom}_G(M, N)$.

We let $\mathfrak{C}$ be the family of finite free modules in $\mathfrak{A}$. *We assume that $\mathfrak{C}$ satisfies* **TG 1**, **TG 2**, **TG 3**, **TG 4**, *and also that $\mathfrak{C}$ is closed under taking alternating products, tensor products and symmetric products.* We let $\mathbf{K} = \mathbf{K}(\mathfrak{C})$. As we have seen, $\mathbf{K}$ is itself a commutative ring. We abbreviate $\mathrm{cl}_{\mathfrak{C}} = \mathrm{cl}$.

We shall define non-linear maps

$$\lambda^i : \mathbf{K} \to \mathbf{K}$$

using the alternating product. If E is finite free, we let

$$\lambda^i(E) = \mathrm{cl}(\textstyle\bigwedge^i E).$$

Proposition 1.1 of Chapter XIX can now be formulated for the $\mathbf{K}$-ring as follows.

Proposition 3.11. *Let*

$$0 \to E' \to E \to E'' \to 0$$

be an exact sequence of finite free modules in $\mathfrak{A}$. Then for every integer $n \geqq 0$ we have

$$\lambda^n(E) = \sum_{i=0}^{n} \lambda^i(E')\lambda^{n-i}(E'').$$

As a result of the proposition, we can define a map

$$\lambda_t : \mathbf{K} \to 1 + t\mathbf{K}[[t]]$$

of $\mathbf{K}$ into the multiplicative group of formal power series with coefficients in $\mathbf{K}$, and with constant term 1, by letting

$$\lambda_t(x) = \sum_{i=0}^{\infty} \lambda^i(x)t^i.$$

Proposition 1.4 of Chapter XIX can be formulated by saying that:

The map

$$\lambda_t : \mathbf{K} \to 1 + t\mathbf{K}[[t]]$$

is a homomorphism.

We note that if L is free of rank 1, then

$$\lambda^0(L) = \text{ground ring};$$

$$\lambda^1(L) = \text{cl}(L);$$

$$\lambda^i(L) = 0 \quad \text{for} \quad i > 1.$$

This can be summarized by writing

$$\lambda_t(L) = 1 + \text{cl}(L)t.$$

Next we can do a similar construction with the symmetric product instead of the alternating product. If E is a finite free module in $\mathcal{C}$ we let as usual:

$S(E)$ = symmetric algebra of E;

$S^i(E)$ = homogeneous component of degree i in $S(E)$.

We define

$$\sigma^i(E) = \text{cl}(S^i(E))$$

and the corresponding power series

$$\sigma_t(E) = \sum \sigma^i(E)t^i.$$

Theorem 3.12. *Let E be a finite free module in $\mathcal{C}$, of rank r. Then for all integers $n \geqq 1$ we have*

$$\sum_{i=0}^{r} (-1)^i \lambda^i(E)\sigma^{n-i}(E) = 0,$$

where by definition $\sigma^j(E) = 0$ for $j < 0$. Furthermore

$$\sigma_t(E)\lambda_{-t}(E) = 1,$$

so the power series $\sigma_t(E)$ and $\lambda_{-t}(E)$ are inverse to each other.

Proof. The first formula depends on the analogue for the symmetric product and the alternating product of the formula given in Proposition 1.1 of Chapter

XIX. It could be proved directly now, but the reader will find a proof as a special case of the theory of Koszul complexes in Chapter XXI, Corollary 4.14. The power series relation is essentially a reformulation of the first formula.

From the above formalism, it is possible to define other maps besides λ^i and σ^i.

Example. Assume that the group G is trivial, and just write $\mathbf{K}$ for the Grothendieck ring instead of $\mathbf{K}(1)$. For $x \in \mathbf{K}$ define

$$\psi_{-t}(x) = -t \frac{d}{dt} \log \lambda_t(x) = -t \lambda_t'(x)/\lambda_t(x).$$

Show that ψ_{-t} is an additive and multiplicative homomorphism. Show that

$$\psi_t(E) = 1 + \mathrm{cl}(E)t + \mathrm{cl}(E)^2 t^2 + \cdots .$$

This kind of construction with the logarithmic derivative leads to the **Adams operations** ψ^i in topology and algebraic geometry. See Exercise 22 of Chapter XVIII.

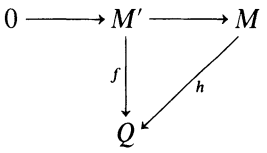
Remark. If it happens in Theorem 3.12 that E admits a decomposition into 1-dimensional free modules in the $\mathbf{K}$-group, then the proof trivializes by using the fact that $\lambda_t(L) = 1 + \mathrm{cl}(L)t$ if L is 1-dimensional. But in the example of (G, k)-spaces when k is a field, this is in general not possible, and it is also not possible in other examples arising naturally in topology and algebraic geometry. However, by "changing the base," one can sometimes achieve this simpler situation, but Theorem 3.12 is then used in establishing the basic properties. *Cf.* Grothendieck [SGA 6], mentioned in the introduction to Part IV, and other works mentioned in the bibliography at the end, namely [Ma 69], [At 61], [At 67], [Ba 68], [Bo 62]. The lectures by Atiyah and Bott emphasize the topological aspects as distinguished from the algebraic-geometric aspects. Grothendieck [Gr 68] actually shows how the formalism of Chern classes from algebraic geometry and topology also enters the theory of representations of linear groups. See also the exposition in [FuL 85], especially the formalism of Chapter I, §6. For special emphasis on applications to representation theory, see Bröcker-tom Dieck [BtD 85], especially Chapter II, §7, concerning compact Lie groups.

§4. INJECTIVE MODULES

In Chapter III, §4, we defined projective modules, which have a natural relation to free modules. By reversing the arrows, we can define a module Q to be **injective** if it satisfies any one of the following conditions which are equivalent:

I 1. Given any module M and a submodule M', and a homomorphism $f: M' \to Q$, there exists an extension of this homomorphism to M,

that is there exists $h: M \to Q$ making the following diagram commutative:



I 2. The functor $M \mapsto \mathrm{Hom}_A(M, Q)$ is exact.

I 3. Every exact sequence $0 \to Q \to M \to M'' \to 0$ splits.

We prove the equivalence. General considerations on homomorphisms as in Proposition 2.1, show that exactness of the homed sequence may fail only at one point, namely given

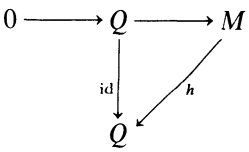
$$0 \to M' \to M \to M'' \to 0,$$

the question is whether

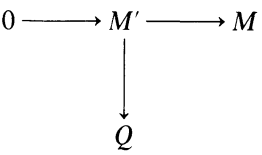
$$\mathrm{Hom}_A(M, Q) \to \mathrm{Hom}_A(M', Q) \to 0$$

is exact. But this is precisely the hypothesis as formulated in **I 1**, so **I 1** implies **I 2** is essentially a matter of linguistic reformulation, and in fact **I 1** is equivalent to **I 2**.

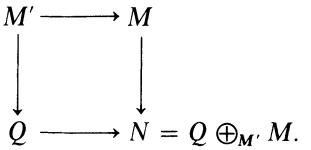
Assume **I 2** or **I 1**, which we know are equivalent. To get **I 3** is immediate, by applying **I 1** to the diagram:



To prove the converse, we need the notion of push-out (cf. Exercise 52 of Chapter I). Given an exact diagram



we form the push-out:



Since $M' \to M$ is a monomorphism, it is immediately verified from the construction of the push-out that $Q \to N$ is also a monomorphism. By **I 3**, the sequence

$$0 \to Q \to N$$

splits, and we can now compose the splitting map $N \to Q$ with the push-out map $M \to N$ to get the desired $h: M \to Q$, thus proving **I 1**.

We saw easily that every module is a homomorphic image of a free module. There is no equally direct construction for the dual fact:

Theorem 4.1. *Every module is a submodule of an injective module.*

The proof will be given by dualizing the situation, with some lemmas. We first look at the situation in the category of abelian groups. If M is an abelian group, let its dual group be $M^\wedge = \mathrm{Hom}(M, \mathbf{Q}/\mathbf{Z})$. If F is a free abelian group, it is reasonable to expect, and in fact it is easily proved that its dual $F^\wedge$ is an injective module, since injectivity is the dual notion of projectivity. Furthermore, M has a natural map into the double dual $M^{\wedge\wedge}$, which is shown to be a monomorphism. Now represent $M^\wedge$ as a quotient of a free abelian group,

$$F \to M^\wedge \to 0.$$

Dualizing this sequence yields a monomorphism

$$0 \to M^{\wedge\wedge} \to F^\wedge,$$

and since M is embedded naturally as a subgroup of $M^{\wedge\wedge}$, we get the desired embedding of M as a subgroup of $F^\wedge$.

This proof also works in general, but there are details to be filled in. First we have to prove that the dual of a free module is injective, and second we have to be careful when passing from the category of abelian groups to the category of modules over an arbitrary ring. We now carry out the details.

We say that an abelian group T is **divisible** if for every integer m, the homomorphism

$$m_T : x \mapsto mx$$

is surjective.

Lemma 4.2. *If T is divisible, then T is injective in the category of abelian groups.*

Proof. Let $M' \subset M$ be a subgroup of an abelian group, and let $f: M' \to T$ be a homomorphism. Let $x \in M$. We want first to extend f to the module (M', x) generated by M' and x. If x is free over M', then we select any value $t \in T$, and it is immediately verified that f extends to (M', x) by giving the value $f(x) = t$. Suppose that x is torsion with respect to M', that is there is a positive integer m such that $mx \in M'$. Let d be the period of x mod M', so

$dx \in M'$, and d is the least positive integer such that $dx \in M'$. By hypothesis, there exists an element $u \in T$ such that $du = f(dx)$. For any integer n, and $z \in M'$ define

$$f(z + nx) = f(z) + nu.$$

By the definition of d, and the fact that $\mathbf{Z}$ is principal, one sees that this value for f is independent of the representation of an element of (M', x) in the form $z + nx$, and then it follows at once that this extended definition of f is a homomorphism. Thus we have extended f to (M', x).

The rest of the proof is merely an application of Zorn's lemma. We consider pairs (N, g) consisting of submodules of M containing M', and an extension g of f to N. We say that $(N, g) \leqq (N_1, g_1)$ if $N \subset N_1$ and the restriction of g_1 to N is g. Then such pairs are inductively ordered. Let (N, g) be a maximal element. If $N \neq M$ then there is some $x \in M$, $x \notin N$ and we can apply the first part of the proof to extend the homomorphism to (N, x), which contradicts the maximality, and concludes the proof of the lemma.

Example. The abelian groups $\mathbf{Q}/\mathbf{Z}$ and $\mathbf{R}/\mathbf{Z}$ are divisible, and hence are injective in the category of abelian groups.

We can prove Theorem 4.1 in the category of abelian groups following the pattern described above. If F is a free abelian group, then the dual $F^\wedge$ is a direct product of groups isomorphic to $\mathbf{Q}/\mathbf{Z}$, and is therefore injective in the category of abelian groups by Lemma 4.2. This concludes the proof.

Next we must make the necessary remarks to extend the system to modules. Let A be a ring and let T be an abelian group. We make $\mathrm{Hom}_{\mathbf{Z}}(A, T)$ into an A-module as follows. Let $f: A \to T$ be an abelian group homomorphism. For $a \in A$ we define the operation

$$(af)(b) = f(ba).$$

The rules for an operation are then immediately verified. Then for any A-module X we have a natural isomorphism of abelian groups:

$$\boxed{\mathrm{Hom}_{\mathbf{Z}}(X, T) \xrightarrow{\sim} \mathrm{Hom}_A(X, \mathrm{Hom}_{\mathbf{Z}}(A, T)).}$$

Indeed, let $\psi : X \to T$ be a $\mathbf{Z}$-homomorphism. We associate with ψ the homomorphism

$$f: X \to \mathrm{Hom}_{\mathbf{Z}}(A, T)$$

such that

$$f(x)(a) = \psi(ax).$$

The definition of the A-module structure on $\mathrm{Hom}_{\mathbf{Z}}(A, T)$ shows that f is an A-homomorphism, so we get an arrow from $\mathrm{Hom}_{\mathbf{Z}}(X, T)$ to

$$\mathrm{Hom}_A(X, \mathrm{Hom}_{\mathbf{Z}}(A, T)).$$

Conversely, let $f: X \to \mathrm{Hom}_{\mathbf{Z}}(A, T)$ be an A-homomorphism. We define the corresponding ψ by

$$\psi(x) = f(x)(1).$$

It is then immediately verified that these maps are inverse to each other.

We shall apply this when T is any divisible group, although we think of T as being $\mathbf{Q}/\mathbf{Z}$, and we think of the homomorphisms into T as representing the dual group according to the pattern described previously.

Lemma 4.3. *If T is a divisible abelian group, then $\mathrm{Hom}_{\mathbf{Z}}(A, T)$ is injective in the category of A-modules.*

Proof. It suffices to prove that if $0 \to X \to Y$ is exact in the category of A-modules, then the dual sequence obtained by taking A-homomorphisms into $\mathrm{Hom}_{\mathbf{Z}}(A, T)$ is exact, that is the top map in the following diagram is surjective.

$$\begin{array}{ccccc}
\mathrm{Hom}_A(Y, \mathrm{Hom}_{\mathbf{Z}}(A, T)) & \longrightarrow & \mathrm{Hom}_A(X, \mathrm{Hom}_{\mathbf{Z}}(A, T)) & \xrightarrow{?} & 0 \\
\uparrow \approx & & \uparrow \approx & & \\
\mathrm{Hom}_{\mathbf{Z}}(Y, T) & \longrightarrow & \mathrm{Hom}_{\mathbf{Z}}(X, T) & \longrightarrow & 0
\end{array}$$

But we have the isomorphisms described before the lemma, given by the vertical arrows of the diagram, which is commutative. The bottom map is surjective because T is an injective module in the category of abelian groups. Therefore the top map is surjective, thus proving the lemma.

Now we prove Theorem 4.1 for A-modules. Let M be an A-module. We can embed M in a divisible abelian group T,

$$0 \to M \xrightarrow{f} T.$$

Then we get an A-homomorphism

$$M \to \mathrm{Hom}_{\mathbf{Z}}(A, T)$$

by $x \mapsto f_x$, where $f_x(a) = f(ax)$. One verifies at once that $x \mapsto f_x$ gives an embedding of M in $\mathrm{Hom}_{\mathbf{Z}}(A, T)$, which is an injective module by Lemma 4.3. This concludes the proof of Theorem 4.1.

§5. HOMOTOPIES OF MORPHISMS OF COMPLEXES

The purpose of this section is to describe a condition under which homomorphisms of complexes induce the same map on the homology and to show that this condition is satisfied in an important case, from which we derive applications in the next section.

The arguments are applicable to any abelian category. The reader may prefer to think of modules, but we use a language which applies to both, and is no more complicated than if we insisted on dealing only with modules.

Let $E = \{(E^n, d^n)\}$ and $E' = \{(E'^n, d'^n)\}$ be two complexes. Let

$$f, g : E \to E'$$

be two morphisms of complexes (of degree 0). We say that f is **homotopic to** g if there exists a sequence of homomorphisms

$$h_n : E^n \to E'^{(n-1)}$$

such that

$$f_n - g_n = d'^{(n-1)}h_n + h_{n+1}d^n.$$

Lemma 5.1. *If f, g are homotopic, then f, g induce the same homomorphism on the homology $H(E)$, that is*

$$H(f_n) = H(g_n) : H^n(E) \to H^n(E').$$

Proof. The lemma is immediate, because $f_n - g_n$ vanishes on the cycles, which are the kernel of d^n, and the homotopy condition shows that the image of $f_n - g_n$ is contained in the boundaries, that is, in the image of $d'^{(n-1)}$.

Remark. The terminology of homotopy is used because the notion and formalism first arose in the context of topology. Cf. [ES 52] and [GreH 81].

We apply Lemma 5.1 to injective objects. Note that as usual the definition of an injective module applies without change to define an injective object in any abelian category. Instead of a submodule in **I 1**, we use a subobject, or equivalently a monomorphism. The proofs of the equivalence of the three conditions defining an injective module depended only on arrow-theoretic juggling, and apply in the general case of abelian categories.

We say that an abelian category has **enough injectives** if given any object M there exists a monomorphism

$$0 \to M \to I$$

into an injective object. We proved in §4 that the category of modules over a ring has enough injectives. *We now assume that the abelian category we work with has enough injectives.*

By an **injective resolution** of an object M one means an exact sequence

$$0 \to M \to I^0 \to I^1 \to I^2 \to$$

such that each I_n ($n \geqq 0$) is injective. Given M, such a resolution exists. Indeed, the monomorphism

$$0 \to M \to I^0$$

exists by hypothesis. Let M^0 be its image. Again by assumption, there exists a monomorphism

$$0 \to I^0/M^0 \to I^1,$$

and the corresponding homomorphism $I^0 \to I^1$ has kernel M^0. So we have constructed the first step of the resolution, and the next steps proceed in the same fashion.

An injective resolution is of course not unique, but it has some uniqueness which we now formulate.

Lemma 5.2. *Consider two complexes*:

$$\begin{array}{ccccccccccc}
0 & \longrightarrow & M & \longrightarrow & E^0 & \longrightarrow & E^1 & \longrightarrow & E^2 & \longrightarrow & \cdots \\
 & & \downarrow{\scriptstyle \varphi} & & & & & & & & \\
0 & \longrightarrow & M' & \longrightarrow & I^0 & \longrightarrow & I^1 & \longrightarrow & I^2 & \longrightarrow & \cdots
\end{array}$$

Suppose that the top row is exact, and that each I^n ($n \geqq 0$) is injective. Let $\varphi : M \to M'$ be a given homomorphism. Then there exists a morphism f of complexes such that $f_{-1} = \varphi$; and any two such are homotopic.

Proof. By definition of an injective, the homomorphism $M \to I^0$ via M' extends to a homomorphism

$$f_0 : E^0 \to I^0,$$

which makes the first square commute:

$$\begin{array}{ccc}
M & \longrightarrow & E_0 \\
{\scriptstyle \varphi}\downarrow & & \downarrow{\scriptstyle f_0} \\
M' & \longrightarrow & I^0
\end{array}$$

Next we must construct f_1. We write the second square in the form

$$\begin{array}{ccc} 0 & \longrightarrow & E^0/M \longrightarrow E^1 \\ & & \downarrow f_0 \\ & & I^0 \longrightarrow I^1 \end{array}$$

with the exact top row as shown. Again because I^1 is injective, we can apply the same argument and find f_1 to make the second square commute. And so on, thus constructing the morphism of complexes f.

Suppose f, g are two such morphisms. We define $h_0 : E^0 \to M'$ to be 0. Then the condition for a homotopy is satisfied in the first instance, when

$$f_{-1} = g_{-1} = \varphi.$$

Next let $d^{-1} : M \to E^0$ be the embedding of M in E^0. Since I^0 is injective, we can extend

$$d^0 : E^0/\operatorname{Im} d^{-1} \to E_1$$

to a homomorphism $h_1 : E^1 \to I^0$. Then the homotopy condition is verified for $f_0 - g_0$. Since $h_0 = 0$ we actually have in this case

$$f_0 - g_0 = h_1 d^0,$$

but this simplification is misleading for the inductive step which follows. We assume constructed the map h_{n+1}, and we wish to show the existence of h_{n+2} satisfying

$$f_{n+1} - g_{n+1} = d'^n h_{n+1} + h_{n+2} d^{n+1}.$$

Since $\operatorname{Im} d^n = \operatorname{Ker} d^{n+1}$, we have a monomorphism $E^{n+1}/\operatorname{Im} d^n \to E^{n+2}$. By the definition of an injective object, which in this case is I^{n+1}, it suffices to prove that

$$f_{n+1} - g_{n+1} - d'^n h_{n+1} \quad \text{vanishes on the image of } d^n,$$

and to use the exact diagram:

$$\begin{array}{ccc} 0 & \longrightarrow & E^{n+1}/\operatorname{Im} d^n \longrightarrow E^{n+2} \\ & & \downarrow f_{n+1} - g_{n+1} \\ & & I^{n+1} \end{array}$$

to get the existence of $h_{n+2} : E^{n+2} \to I^{n+1}$ extending $f_{n+1} - g_{n+1}$. But we have:

$$\begin{aligned} &(f_{n+1} - g_{n+1} - d'^n h_{n+1}) d^n \\ &\quad = (f_{n+1} - g_{n+1}) d^n - d'^n h_{n+1} d^n \end{aligned}$$

$$= (f_{n+1} - g_{n+1})d^n - d'^n(f_n - g_n - d'^{(n-1)}h_n) \quad \text{by induction}$$
$$= (f_{n+1} - g_{n+1})d^n - d'^n(f_n - g_n) \quad \text{because } d'd' = 0$$
$$= 0 \quad \text{because } f, g \text{ are homomorphisms of complexes.}$$

This concludes the proof of Lemma 5.2.

Remark. Dually, let $P_{M'} \to M' \to 0$ be a complex with P^i projective for $i \geqq 0$, and let $E_M \to M \to 0$ be a resolution. Let $\varphi: M' \to M$ be a homomorphism. Then φ extends to a homomorphism of complex $P \to E$. The proof is obtained by reversing arrows in Lemma 5.2. The books on homological algebra that I know of in fact carry out the projective case, and leave the injective case to the reader. However, one of my motivations is to do here what is needed, for instance in [Ha 77], Chapter III, on derived functors, as a preliminary to the cohomology of sheaves. For an example of projective resolutions using free modules, see Exercises 2–7, concerning the cohomology of groups.

§6. DERIVED FUNCTORS

We continue to work in an abelian category. A covariant additive functor

$$F: \mathfrak{A} \to \mathfrak{B}$$

is said to be **left exact** if it transforms an exact sequence

$$0 \to M' \to M \to M''$$

into an exact sequence $0 \to F(M') \to F(M) \to F(M'')$. We remind the reader that F is called **additive** if the map

$$\mathrm{Hom}(A', A) \to \mathrm{Hom}(FA', FA)$$

is additive.

We assume throughout that F is left exact unless otherwise specified, and additive. We continue to assume that our abelian category has enough injectives.

Given an object M, let

$$0 \to M \to I^0 \to I^1 \to I^2 \to$$

be an injective resolution, which we abbreviate by

$$0 \to M \to I_M,$$

where I_M is the complex $I^0 \to I^1 \to I^2 \to$. We let I be the complex

$$0 \to I^0 \to I^1 \to I^2 \to$$

We define the **right-derived functor** R^nF by

$$R^nF(M) = H^n(F(I)),$$

in other words, the n-th homology of the complex

$$0 \to F(I^0) \to F(I^1) \to F(I^2) \to$$

Directly from the definitions and the monomorphism $M \to I_0$, we see that there is an isomorphism

$$R^0F(M) = F(M).$$

This isomorphism seems at first to depend on the injective resolution, and so do the functors $R^nF(M)$ for other n. However, from Lemmas 5.1 and 5.2 we see that given two injective resolutions of M, there is a homomorphism between them, and that any two homomorphisms are homotopic. If we apply the functor F to these homomorphisms and to the homotopy, then we see that the homology of the complex $F(I)$ is in fact determined up to a unique isomorphism. One therefore omits the resolution from the notation and from the language.

Example 1. Let R be a ring and let $\mathfrak{A} = \mathrm{Mod}(R)$ be the category of R-modules. Fix a module A. The functor $M \mapsto \mathrm{Hom}(A, M)$ is left exact, i.e. given an exact sequence $0 \to M' \to M \to M''$, the sequence

$$0 \to \mathrm{Hom}(A, M') \to \mathrm{Hom}(A, M) \to \mathrm{Hom}\ (A, M'')$$

is exact. Its right derived functors are denoted by $\mathrm{Ext}^n(A, M)$ for M variable. Similarly, for a fixed module B, the functor $X \mapsto \mathrm{Hom}\ (X, B)$ is right exact, and it gives rise to its **left derived functors**. For the explicit mirror image of the terminology, see the end of this section. In any case, we may consider A as variable. In §8 we shall go more deeply into this aspect of the formalism, by dealing with bifunctors. It will turn out that $\mathrm{Ext}^n\ (A, B)$ has a dual interpretation as a left derived functor of the first variable and right derived functor of the second variable. See Corollary 8.5.

In the exercises, you will prove that $\mathrm{Ext}^1(A, M)$ is in bijection with isomorphism classes of extensions, of M by A, that is, isomorphism classes of exact sequences

$$0 \to A \to E \to M \to 0.$$

The name Ext comes from this interpretation in dimension 1.

For the computation of Ext^i in certain important cases, see Chapter XXI, Theorems 4.6 and 4.11, which serve as examples for the general theory.

Example 2. Let R be commutative. The functor $M \mapsto A \otimes M$ is right exact, in other words, the sequence

$$A \otimes M' \to A \otimes M \to A \otimes M'' \to 0$$

is exact. Its left derived functors are denoted by $\mathrm{Tor}_n(A, M)$ for M variable.

Example 3. Let G be a group and let $R = \mathbf{Z}[G]$ be the group ring. Let $\mathfrak{A}$ be the category of G-modules, i.e. $\mathfrak{A} = \mathrm{Mod}(R)$, also denoted by $\mathrm{Mod}(G)$. For a G-module A, let A^G be the submodule (abelian group) consisting of those elements v such that $xv = v$ for all $x \in G$. Then $A \mapsto A^G$ is a left exact functor from $\mathrm{Mod}(R)$ into the category of abelian groups. Its left derived functors give rise to the cohomology of groups. Some results from this special cohomology will be carried out in the exercises, as further examples of the general theory.

Example 4. Let X be a topological space (we assume the reader knows what this is). By a **sheaf** $\mathfrak{F}$ of abelian groups on X, we mean the data:

(a) For every open set U of X there is given an abelian group $\mathfrak{F}(U)$.

(b) For every inclusion $V \subset U$ of open sets there is given a homomorphism

$$\mathrm{res}_V^U : \mathfrak{F}(U) \to \mathfrak{F}(V),$$

called the **restriction** from U to V, subject to the following conditions:

SH 1. $\mathfrak{F}(\text{empty set}) = 0$.

SH 2. res_U^U is the identity $\mathfrak{F}(U) \to \mathfrak{F}(U)$.

SH 3. If $W \subset V \subset U$ are open sets, then $\mathrm{res}_W^V \circ \mathrm{res}_V^U = \mathrm{res}_W^U$.

SH 4. Let U be an open set and $\{V_i\}$ be an open covering of U. Let $s \in \mathfrak{F}(U)$. If the restriction of s to each V_i is 0, then $s = 0$.

SH 5. Let U be an open set and let $\{V_i\}$ be an open covering of U. Suppose given $s_i \in \mathfrak{F}(V_i)$ for each i, such that given i, j the restrictions of s_i and s_j to $V_i \cap V_j$ are equal. Then there exists a unique $s \in \mathfrak{F}(U)$ whose restriction to V_i is s_i for all i.

Elements of $\mathfrak{F}(U)$ are called **sections** of $\mathfrak{F}$ over U. Elements of $\mathfrak{F}(X)$ are called **global sections**. Just as for abelian groups, it is possible to define the notion of homomorphisms of sheaves, kernels, cokernels, and exact sequences. The association $\mathfrak{F} \mapsto \mathfrak{F}(X)$ (global sections functor) is a functor from the category of sheaves of abelian groups to abelian groups, and this functor is left exact. Its right derived functors are the basis of cohomology theory in topology and algebraic geometry (among other fields of mathematics). The reader will find a self-contained brief definition of the basic properties in [Ha 77], Chapter II, §1, as well as a proof that these form an abelian category. For a more extensive treatment I recommend Gunning's [Gu 91], mentioned in the introduction to Part IV, notably Volume III, dealing with the cohomology of sheaves.

We now return to the general theory of derived functors. The general theory tells us that these derived functors do not depend on the resolution by projectives or injectives according to the variance. As we shall also see in §8, one can even use other special types of objects such as acyclic or exact (to be defined), which gives even more flexibility in the ways one has to compute homology. Through certain explicit resolutions, we obtain means of computing the derived functors

explicitly. For example, in Exercise 16, you will see that the cohomology of finite cyclic groups can be computed immediately by exhibiting a specific free resolution of $\mathbf{Z}$ adapted to such groups. Chapter XXI will contain several other examples which show how to construct explicit finite free resolutions, which allow the determination of derived functors in various contexts.

The next theorem summarizes the basic properties of derived functors.

Theorem 6.1. *Let $\mathfrak{A}$ be an abelian category with enough injectives, and let $F : \mathfrak{A} \to \mathfrak{B}$ be a covariant additive left exact functor to another abelian category $\mathfrak{B}$. Then:*

(i) *For each $n \geqq 0$, R^nF as defined above is an additive functor from $\mathfrak{A}$ to $\mathfrak{B}$. Furthermore, it is independent, up to a unique isomorphism of functors, of the choices of resolutions made.*

(ii) *There is a natural isomorphism $F \approx R^0F$.*

(iii) *For each short exact sequence*

$$0 \to M' \to M \to M'' \to 0$$

and for each $n \geqq 0$ there is a natural homomorphism

$$\delta^n : R^nF(M'') \to R^{n+1}F(M)$$

such that we obtain a long exact sequence:

$$\to R^nF(M') \to R^nF(M) \to R^nF(M'') \xrightarrow{\delta^n} R^{n+1}F(M') \to.$$

(iv) *Given a morphism of short exact sequences*

$$\begin{array}{ccccccccc} 0 & \longrightarrow & M' & \longrightarrow & M & \longrightarrow & M'' & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & N' & \longrightarrow & N & \longrightarrow & N'' & \longrightarrow & 0 \end{array}$$

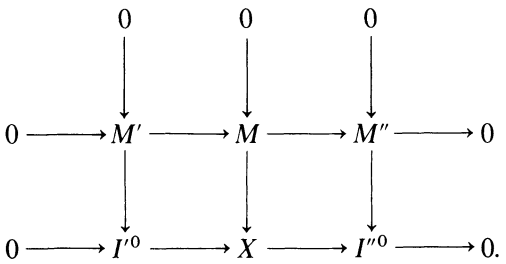
the δ's give a commutative diagram:

$$\begin{array}{ccc} R^nF(M'') & \xrightarrow{\delta^n} & R^{n+1}F(M') \\ \downarrow & & \downarrow \\ R^nF(N'') & \xrightarrow[\delta^n]{} & R^{n+1}F(N') \end{array}$$

(v) *For each injective object I of A and for each $n > 0$ we have $R^nF(I) = 0$.*

Properties (i), (ii), (iii), and (iv) essentially say that R^nF is a delta-functor in a sense which will be expanded in the next section. The last property (v) will be discussed after we deal with the delta-functor part of the theorem.

We now describe how to construct the δ-homomorphisms. Given a short exact sequence, we can find an injective resolution of M', M, M'' separately, but they don't necessarily fit in an exact sequence of complexes. So we must achieve this to apply the considerations of §1. Consider the diagram:



We give monomorphisms $M' \to I'^0$ and $M'' \to I''^0$ into injectives, and we want to find X injective with a monomorphism $M \to X$ such that the diagram is exact. We take X to be the direct sum

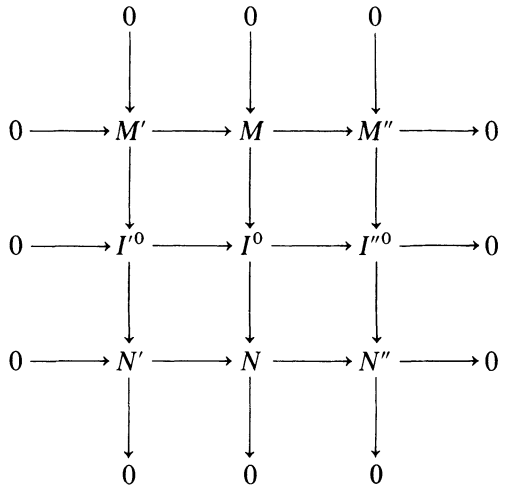
$$X = I'^0 \oplus I''^0.$$

Since I'^0 is injective, the monomorphism $M' \to I'^0$ can be extended to a homomorphism $M \to I'^0$. We take the homomorphism of M into $I'^0 \oplus I''^0$ which comes from this extension on the first factor I'^0, and is the composite map

$$M \to M'' \to I''^0$$

on the second factor. Then $M \to X$ is a monomorphism. Furthermore $I'^0 \to X$ is the monomorphism on the first factor, and $X \to I''^0$ is the projection on the second factor. So we have constructed the diagram we wanted, giving the beginning of the compatible resolutions.

Now we take the quotient homomorphism, defining the third row, to get an exact diagram:



where we let $I^0 = X$, and N', N, N'' are the cokernels of the vertical maps by definition. The exactness of the N-sequence is left as an exercise to the reader. We then repeat the construction with the N-sequence, and by induction construct injective resolutions

$$\begin{array}{ccccccccc} & & 0 & & 0 & & 0 & & \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & M' & \longrightarrow & M & \longrightarrow & M'' & \longrightarrow & 0 \\ & & \downarrow & & \downarrow & & \downarrow & & \\ 0 & \longrightarrow & I'_{M'} & \longrightarrow & I_M & \longrightarrow & I''_{M''} & \longrightarrow & 0 \end{array}$$

of the M-sequence such that the diagram of the resolutions is exact.

We now apply the functor F to this diagram. We obtain a short sequence of complexes:

$$0 \to F(I') \to F(I) \to F(I'') \to 0,$$

which is exact because $I = I' \oplus I''$ is a direct sum and F is left exact, so F commutes with direct sums. We are now in a position to apply the construction of §1 to get the coboundary operator in the homology sequence:

$$R^nF(M') \to R^nF(M) \to R^nF(M'') \xrightarrow{\delta^n} R^{n+1}F(M').$$

This is legitimate because the right derived functor is independent of the chosen resolutions.

So far, we have proved (i), (ii), and (iii). To prove (iv), that is the naturality of the delta homomorphisms, it is necessary to go through a three-dimensional commutative diagram. At this point, I feel it is best to leave this to the reader, since it is just more of the same routine.

Finally, the last property (v) is obvious, for if I is injective, then we can use the resolution

$$0 \to I \to I \to 0$$

to compute the derived functors, from which it is clear that $R^nF = 0$ for $n > 0$.

This concludes the proof of Theorem 6.1.

In applications, it is useful to determine the derived functors by means of other resolutions besides injective ones (which are useful for theoretical purposes, but not for computational ones). Let again F be a left exact additive functor. An object X is called ***F*-acyclic** if $R^nF(X) = 0$ for all $n > 0$.

Theorem 6.2. *Let*

$$0 \to M \to X^0 \to X^1 \to X^2 \to \cdots$$

be a resolution of M by F-acyclics. Let

$$0 \to M \to I^0 \to I^1 \to I^2 \to \cdots$$

be an injective resolution. Then there exists a morphism of complexes $X_M \to I_M$ extending the identity on M, and this morphism induces an isomorphism

$$H^nF(X) \approx H^nF(I) = R^nF(M) \qquad \text{for all } n \geqq 0.$$

Proof. The existence of the morphism of complexes extending the identity on M is merely Lemma 5.2. The usual proof of the theorem via spectral sequences can be formulated independently in the following manner, shown to me by David Benson. We need a lemma.

Lemma 6.3. *Let Y^i $(i \geqq 0)$ be F-acyclic, and suppose the sequence*

$$0 \to Y^0 \to Y^1 \to Y^2 \to \cdots$$

is exact. Then

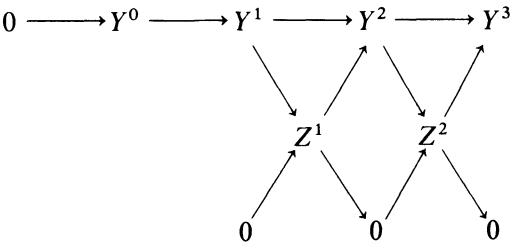
$$0 \to F(Y^0) \to F(Y^1) \to F(Y^2) \to \cdots$$

is exact.

Proof. Since F is left exact, we have an exact sequence

$$0 \to F(Y^0) \to F(Y^1) \to F(Y^2).$$

We want to show exactness at the next joint. We draw the cokernels:



So $Z_1 = \text{Coker}(Y^0 \to Y^1)$; $Z_2 = \text{Coker}(Y^1 \to Y^2)$; etc. Applying F we have an exact sequence

$$0 \to F(Y^0) \to F(Y^1) \to F(Z^1) \to R^1F(Y^0) = 0.$$

So $F(Z_1) = \operatorname{Coker}(F(Y^0) \to F(Y^1))$. We now consider the exact sequence

$$0 \to Z_1 \to Y_2 \to Y_3$$

giving the exact sequence

$$0 \to F(Z^1) \to F(Y^2) \to F(Y^3)$$

by the left-exactness of F, and proving what we wanted. But we can now continue by induction because Z_1 is F-acyclic, by the exact sequence

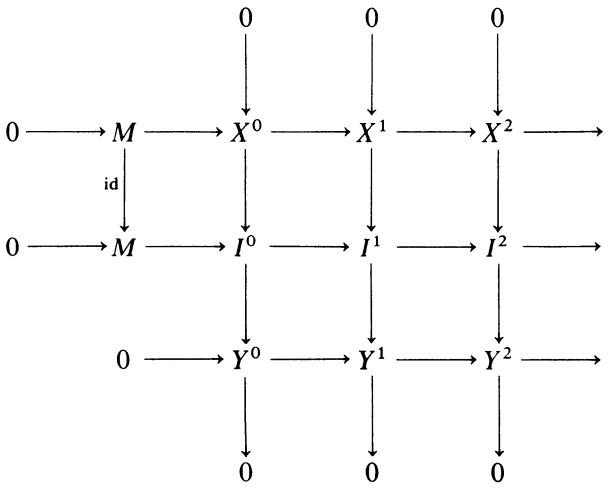
$$0 \to R^nF(Y^1) \to R^nF(Z^1) \to R^{n+1}F(Y^0) = 0.$$

This concludes the proof of Lemma 6.3.

We return to the proof of Theorem 6.2. The injective resolution

$$0 \to M \to I_M$$

can be chosen such that the homomorphisms $X_n \to I_n$ are monomorphisms for $n \geqq 0$, because the derived functor is independent of the choice of injective resolution. Thus we may assume without loss of generality that we have an exact diagram:



defining Y^n as the appropriate cokernel of the vertical map.

Since X^n and I^n are acyclic, so is Y^n from the exact sequence

$$R^kF(I^n) \to R^kF(Y^n) \to R^{k+1}F(X^n).$$

Applying F we obtain a short exact sequence of complexes

$$0 \to F(X) \to F(I) \to F(Y) \to 0.$$

whence the corresponding homology sequence

$$H^{n-1}F(Y) \to H^nF(X) \to H^nF(I) \to H^nF(Y).$$

Both extremes are 0 by Lemma 6.3, so we get an isomorphism in the middle, which by definition is the isomorphism

$$H^nF(X) \approx R^nF(M),$$

thus proving the theorem.

Left derived functors

We conclude this section by a summary of the properties of left derived functors.

We consider complexes going the other way,

$$\to X_n \to \cdots \to X_2 \to X_1 \to X_0 \to M \to 0$$

which we abbreviate by

$$X_M \to M \to 0.$$

We call such a complex a **resolution** of M if the sequence is exact. We call it a **projective resolution** if X_n is projective for all $n \geqq 0$.

Given projective resolutions X_M, $Y_{M'}$ and a homomorphism

$$\varphi : M \to M'$$

there always exists a homomorphism $X_M \to Y_{M'}$ extending φ, and any two such are homotopic.

In fact, one need only assume that X_M is a projective resolution, and that $Y_{M'}$ is a resolution, not necessarily projective, for the proof to go through.

Let T be a covariant additive functor. Fix a projective resolution of an object M,

$$P_M \to M \to 0.$$

We define the **left derived functor** L_nT by

$$L_nT(M) = H_n(T(P)),$$

where $T(P)$ is the complex

$$\to T(P_n) \to \cdots \to T(P_2) \to T(P_1) \to T(P_0) \to 0.$$

The existence of homotopies shows that $L_nT(M)$ is uniquely determined up to a unique isomorphism if one changes the projective resolution.

We define T to be **right exact** if an exact sequence

$$M' \to M \to M'' \to 0$$

yields an exact sequence

$$T(M') \to T(M) \to T(M'') \to 0.$$

If T is right exact, then we have immediately from the definitions

$$L_0 T(M) \approx M.$$

Theorems 6.1 and 6.2 then go over to this case with similar proofs. One has to replace "injectives" by "projectives" throughout, and in Theorem 6.1, the last condition states that for $n > 0$,

$$L_n T(P) = 0 \qquad \text{if } P \text{ is projective.}$$

Otherwise, it is just a question of reversing certain arrows in the proofs. For an example of such left derived functors, see Exercises 2–7 concerning the cohomology of groups.

§7. DELTA-FUNCTORS

In this section, we axiomatize the properties stated in Theorem 6.1 following Grothendieck.

Let $\mathfrak{A}$, $\mathfrak{B}$ be abelian categories. A (covariant) **δ-functor** from $\mathfrak{A}$ to $\mathfrak{B}$ is a family of additive functors $F = \{F_n\}_{n \geqq 0}$, and to each short exact sequence

$$0 \to M' \to M \to M'' \to 0$$

an associated family of morphisms

$$\delta^n : F^n(M'') \to F^{n+1}(M')$$

with $n \geqq 0$, satisfying the following conditions:

DEL 1. For each short exact sequence as above, there is a long exact sequence

$$0 \to F^0(M') \to F^0(M) \to F^0(M'') \to F^1(M') \to \cdots$$

$$\to F^n(M') \to F^n(M) \to F^n(M'') \to F^{n+1}(M') \to$$

DEL 2. For each morphism of one short exact sequence as above into another $0 \to N' \to N \to N'' \to 0$, the δ's give a commutative diagram:

$$\begin{array}{ccc} F^n(M'') & \xrightarrow{\delta} & F^{n+1}(M') \\ \downarrow & & \downarrow \\ F^n(N'') & \xrightarrow[\delta]{} & F^{n+1}(N'). \end{array}$$

Before going any further, it is useful to give another definition. Many proofs in homology theory are given by induction from one index to the next. It turns out that the only relevant data for going up by one index is given in two successive dimensions, and that the other indices are irrelevant. Therefore we generalize the notion of δ-functor as follows.

A **δ-functor defined in degrees 0, 1** is a pair of functors (F^0, F^1) and to each short exact sequence

$$0 \to A' \to A \to A'' \to 0$$

an associated morphism

$$\delta : F^0(A'') \to F^1(A'')$$

satisfying the two conditions as before, but putting $n = 0$, $n + 1 = 1$, and forgetting about all other integers n. We could also use any two consecutive positive integers to index the δ-functor, or any sequence of consecutive integers $\geqq 0$. In practice, only the case of all integers $\geqq 0$ occurs, but for proofs, it is useful to have the flexibility provided by using only two indices, say 0, 1.

The δ-functor F is said to be **universal**, if given any other δ-functor G of $\mathfrak{A}$ into $\mathfrak{B}$, and given any morphism of functors

$$f_0 : F^0 \to G^0,$$

there exists a unique sequence of morphisms

$$f_n : F^n \to G^n$$

for all $n \geqq 0$, which commute with the δ^n for each short exact sequence.

By the definition of universality, a δ-functor G such that $G^0 = F^0$ is uniquely determined up to a unique isomorphism of functors. We shall give a condition for a functor to be universal.

An additive functor F of $\mathfrak{A}$ into $\mathfrak{B}$ is called **erasable** if to each object A there exists a monomorphism $u : A \to M$ for some M such that $F(u) = 0$. In practice, it even happens that $F(M) = 0$, but we don't need it in the axiomatization.

Linguistic note. Grothendieck originally called the notion "effaceable" in French. The dictionary translation is "erasable," as I have used above. Apparently people who did not know French have used the French word in English, but there is no need for this, since the English word is equally meaningful and convenient.

We say the functor is erasable by **injectives** if in addition M can be taken to be injective.

Example. Of course, a right derived functor is erasable by injectives, and a left derived functor by projectives. However, there are many cases when one wants erasability by other types of objects. In Exercises 9 and 14, dealing with the cohomology of groups, you will see how one erases the cohomology functor with induced modules, or regular modules when G is finite. In the category of coherent sheaves in algebraic geometry, one erases the cohomology with locally free sheaves of finite rank.

Theorem 7.1. *Let $F = \{F^n\}$ be a covariant δ-functor from $\mathfrak{A}$ into $\mathfrak{B}$. If F^n is erasable for each $n > 0$, then F is universal.*

Proof. Given an object A, we erase it with a monomorphism u, and get a short exact sequence:

$$0 \to A \xrightarrow{\varphi} M \to X \to 0.$$

Let G be another δ-functor with given $f_0 : F^0 \to G^0$. We have an exact commutative diagram

$$\begin{array}{ccccccc} F^0(M) & \longrightarrow & F^0(X) & \xrightarrow{\delta'} & F^1(A) & \longrightarrow & 0 \\ \downarrow{\scriptstyle f_0} & & \downarrow{\scriptstyle f_0} & & \vdots{\scriptstyle f_1?} & & \\ G^0(M) & \longrightarrow & G^0(X) & \xrightarrow[\delta_G]{} & G^1(A) & & \end{array}$$

We get the 0 on the top right because of the erasability assumption that

$$F^1(\varphi) = 0.$$

We want to construct

$$f_1(A) : F^1(A) \to G^1(A)$$

which makes the diagram commutative, is functorial in A, and also commutes with the δ. Commutativity in the left square shows that $\operatorname{Ker} \delta_F$ is contained in the kernel of $\delta_G \circ f_0$. Hence there exists a unique homomorphism

$$f_1(A) : F^1(A) \to G^1(A)$$

which makes the right square commutative. We are going to show that $f_1(A)$ satisfies the desired conditions. The rest of the proof then proceeds by induction following the same pattern.

We first prove the functoriality in A.

Let $u : A \to B$ be a morphism. We form the push-out P in the diagram

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & M \\ \downarrow{\scriptstyle u} & & \downarrow \\ B & \longrightarrow & P \end{array}$$

Since φ is a monomorphism, it follows that $B \to P$ is a monomorphism also. Then we let $P \to N$ be a monomorphism which erases F_1. This yields a commutative diagram

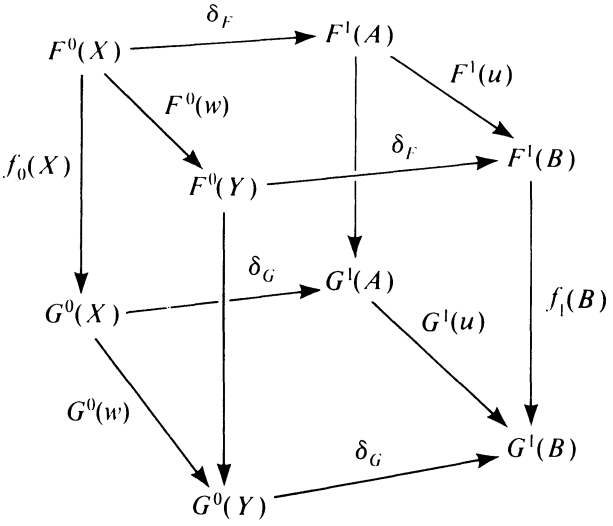
$$\begin{array}{ccccccccc} 0 & \longrightarrow & A & \longrightarrow & M & \longrightarrow & X & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle u} & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle w} & & \\ 0 & \longrightarrow & B & \longrightarrow & N & \longrightarrow & Y & \longrightarrow & 0 \end{array}$$

where $B \to N$ is the composite $B \to P \to N$, and Y is defined to be the cokernel of $B \to N$.

Functoriality in A means that the following diagram is commutative.

$$\begin{array}{ccc} F^1(A) & \xrightarrow{F^1(u)} & F^1(B) \\ \downarrow{\scriptstyle f_1(A)} & & \downarrow{\scriptstyle f_1(B)} \\ G^1(A) & \xrightarrow[F^1(u)]{} & G^1(B) \end{array}$$

This square is the right-hand side of the following cube:



All the faces of the cube are commutative except possibly the right-hand face. It is then a general fact that if the top maps here denoted by δ_F are epimorphisms,

then the right-hand side is commutative also. This can be seen as follows. We start with $f_1(B)F^1(u)\delta_F$. We then use commutativity on the top of the cube, then the front face, then the left face, then the bottom, and finally the back face. This yields

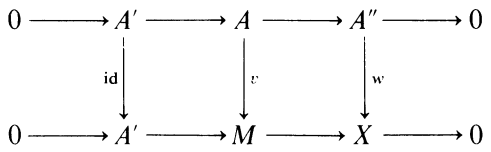
$$f_1(B)F^1(u)\delta_F = G^1(u)f_1(A)\delta_F.$$

Since δ_F is an epimorphism, we can cancel δ_F to get what we want.

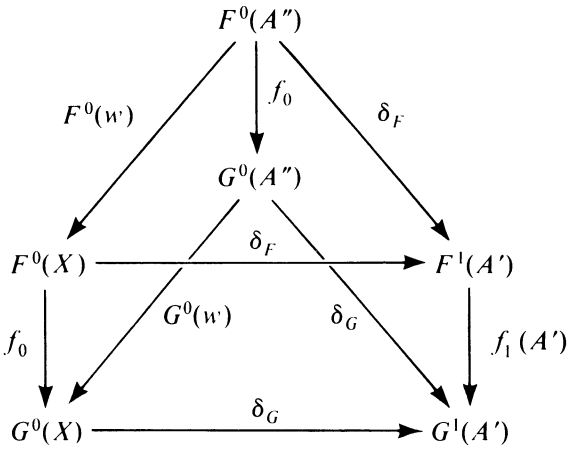
Second, we have to show that f_1 commutes with δ. Let

$$0 \to A' \to A \to A'' \to 0$$

be a short exact sequence. The same push-out argument as before shows that there exists an erasing monomorphism $0 \to A' \to M$ and morphisms v, w making the following diagram commutative:



Here X is defined as the appropriate cokernel of the bottom row. We now consider the following diagram:



Our purpose is to prove that the right-hand face is commutative. The triangles on top and bottom are commutative by the definition of a δ-functor. The

left-hand square is commutative by the hypothesis that f_0 is a morphism of functors. The front square is commutative by the definition of $f_1(A')$. Therefore we find:

$$\begin{aligned} f_1(A')\delta_F &= f_1(A')\delta_F F^0(w) && \text{(top triangle)} \\ &= \delta_F f_0 F^0(w) && \text{(front square)} \\ &= \delta_F G^0(w) f_0 && \text{(left square)} \\ &= \delta_F f_0 && \text{(bottom triangle).} \end{aligned}$$

This concludes the proof of Theorem 7.1, since instead of the pair of indices (0, 1) we could have used $(n, n+1)$.

Remark. The morphism f_1 constructed in Theorem 7.1 depends functorially on f_0 in the following sense. Suppose we have three delta functors F, G, H defined in degrees 0, 1. Suppose given morphisms

$$f_0 : F^0 \to G^0 \quad \text{and} \quad g_0 : G^0 \to H^0.$$

Suppose that the erasing monomorphisms erase both F and G. Then we can construct f_1 and g_1 by applying the theorem. On the other hand, the composite

$$g_0 f_0 = h_0 : F^0 \to H^0$$

is also a morphism of functors, and the theorem yields the existence of a morphism

$$h_1 : F^1 \to H^1$$

such that (h_0, h_1) is a δ-morphism. By uniqueness, we therefore have

$$h_1 = g_1 f_1.$$

This is what we mean by the functorial dependence as mentioned above.

Corollary 7.2. *Assume that $\mathfrak{A}$ has enough injectives. Then for any left exact functor $F : \mathfrak{A} \to \mathfrak{B}$, the derived functors R^nF with $n \geqq 0$ form a universal δ-functor with $F \approx R^0F$, which is erasable by injectives. Conversely, if $G = \{G^n\}_{n \geqq 0}$ is a universal δ-functor, then G^0 is left exact, and the G^n are isomorphic to R^nG^0 for each $n \geqq 0$.*

Proof. If F is a left exact functor, then the $\{R^nF\}_{n \geqq 0}$ form a δ-functor by Theorem 6.1. Furthermore, for any object A, let $u : A \to I$ be a monomorphism of A into an injective. Then $R^nF(I) = 0$ for $n > 0$ by Theorem 6.1(iv), so $R^nF(u) = 0$. Hence R^nF is erasable for all $n > 0$, and we can apply Theorem 7.1.

Remark. As usual, Theorem 7.1 applies to functors with different variance. Suppose $\{F^n\}$ is a family of contravariant additive functors, with n ranging over

a sequence of consecutive integers, say for simplicity $n \geqq 0$. We say that F is a **contravariant δ-functor** if given an exact sequence

$$0 \to M' \to M \to M'' \to 0$$

then there is an associated family of morphisms

$$\delta^n : F^n(M') \to F^{n+1}(M')$$

satisfying **DEL 1** and **DEL 2** with M' interchanged with M'' and N' interchanged with N''. We say that F is **coerasable** if to each object A there exists an epimorphism $u : M \to A$ such that $F(u) = 0$. We say that F is **universal** if given any other δ-functor G of $\mathfrak{A}$ into $\mathfrak{B}$ and given a morphism of functors

$$f_0 : F^0 \to G^0$$

there exists a unique sequence of morphisms

$$f_n : F^n \to G^n$$

for all $n \geqq 0$ which commute with δ for each short exact sequence.

Theorem 7.1′. *Let $F = \{F^n\}$ (n ranging over a consecutive sequence of integers $\geqq 0$) be a contravariant δ-functor from $\mathfrak{A}$ into $\mathfrak{B}$, and assume that F^n is coerasable for $n \geqq 1$. Then F is universal.*

Examples of δ-functors with the variances as in Theorems 7.1 *and* 7.1′ will be given in the next section in connection with bifunctors.

Dimension shifting

Let $F = \{F^n\}$ be a contravariant delta functor with $n \geqq 0$. Let $\mathcal{E}$ be a family of objects which erases F^n for all $n \geqq 1$, that is $F^n(E) = 0$ for $n \geqq 1$ and $E \in \mathcal{E}$. Then such a family allows us to do what is called **dimension shifting** as follows. Given an exact sequence

$$0 \to Q \to E \to M \to 0$$

with $E \in \mathcal{E}$, we get for $n \geqq 1$ an exact sequence

$$0 = F^n(E) \to F^n(Q) \to F^{n+1}(M) \to F^{n+1}(E) = 0,$$

and therefore an isomorphism

$$F^n(Q) \xrightarrow{\approx} F^{n+1}(M),$$

which exhibits a shift of dimensions by one. More generally:

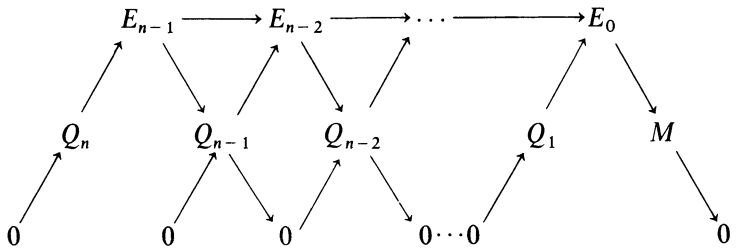
Proposition 7.3. *Let*

$$0 \to Q \to E_{n-1} \to \cdots \to E_0 \to M \to 0$$

be an exact sequence, such that $E_i \in \mathcal{E}$. *Then we have an isomorphism*

$$F^p(Q) \approx F^{p+n}(M) \qquad \text{for } p \geqq 1.$$

Proof. Let $Q = Q_n$. Also without loss of generality, take $p = 1$. We may insert kernels and cokernels at each step as follows:



Then shifting dimension with respect to each short exact sequence, we find isomorphisms

$$F^1(Q_n) \approx F^2(Q_{n-1}) \approx \cdots \approx F^{n+1}(M).$$

This concludes the proof.

One says that M has F-**dimension** $\leqq d$ if $F^n(M) = 0$ for $n \geqq d + 1$. By dimension shifting, we see that if M has F-dimension $\leqq d$, then Q has F-dimension $\leqq d - n$ in Proposition 7.3. In particular, if M has F-dimension n, then Q has F-dimension 0.

The reader should rewrite all this formalism by changing notation, using for F the standard functors arising from Hom in the first variable, on the category of modules over a ring, which has enough projectives to erase the left derived functors of

$$A \mapsto \mathrm{Hom}(A, B),$$

for B fixed. We shall study this situation, suitably axiomatized, in the next section.

§8. BIFUNCTORS

In an abelian category one often deals with Hom, which can be viewed as a functor in two variables; and also the tensor product, which is a functor in two variables, but their variance is different. In any case, these examples lead to the notion of **bifunctor**. This is an association

$$(A, B) \mapsto T(A, B)$$

where A, B are objects of abelian categories $\mathfrak{A}$ and $\mathfrak{B}$ respectively, with values in some abelian category. This means that T is functorial in each variable, with the appropriate variance (there are four possibilities, with covariance and contravariance in all possible combinations); and if, say, T is covariant in all variables, we also require that for homomorphisms $A' \to A$ and $B' \to B$ there is a commutative diagram

$$\begin{array}{ccc} T(A', B') & \longrightarrow & T(A', B) \\ \downarrow & & \downarrow \\ T(A, B') & \longrightarrow & T(A, B). \end{array}$$

If the variances are shuffled, then the arrows in the diagram are to be reversed in the appropriate manner. Finally, we require that as a functor in each variable, T is additive.

Note that Hom is a bifunctor, contravariant in the first variable and covariant in the second. The tensor product is covariant in each variable.

The Hom functor is a bifunctor T satisfying the following properties:

HOM 1. *T is contravariant and left exact in the first variable.*

HOM 2. *T is covariant and left exact in the second variable.*

HOM 3. *For any injective object J the functor*

$$A \mapsto T(A, J)$$

is exact.

They are the only properties which will enter into consideration in this section. There is a possible fourth one which might come in other times:

HOM 4. *For any projective object Q the functor*

$$B \mapsto T(Q, B)$$

is exact.

But we shall deal *non-symmetrically, and view T as a functor of the second variable, keeping the first one fixed, in order to get derived functors of the second variable.* On the other hand, we shall also obtain a δ-functor of the first variable by using the bifunctor, even though this δ-functor is not a derived functor.

If $\mathfrak{B}$ has enough injectives, then we may form the right derived functors with respect to the second variable

$$B \mapsto R^n T(A, B), \qquad \text{also denoted by } R^n T_A(B),$$

fixing A, and viewing B as variable. If $T = \mathrm{Hom}$, then this right derived functor is called **Ext**, so we have by definition

$$\mathrm{Ext}^n(A, X) = R^n \mathrm{Hom}(A, X).$$

We shall now give a criterion to compute the right derived functors in terms of the other (first) variable. We say that an object A is ***T*-exact** if the functor $B \mapsto T(A, B)$ is exact. By a ***T*-exact resolution** of an object A, we mean a resolution

$$\to M_1 \to M_0 \to A \to 0$$

where M_n is T-exact for all $n \geqq 0$.

Examples. Let $\mathfrak{A}$ and $\mathfrak{B}$ be the categories of modules over a commutative ring. Let $T = \mathrm{Hom}$. Then a T-exact object is by definition a projective module. Now let the **transpose** of T be given by

$${}^tT(A, B) = T(B, A).$$

Then a ${}^tT$-exact object is by definition an injective module.

If T is the tensor product, such that $T(A, B) = A \otimes B$, then a T-exact object is called **flat**.

Remark. In the category of modules over a ring, there are enough projectives and injectives. But there are other situations when this is not the case. Readers who want to see all this abstract nonsense in action may consult [GriH 78], [Ha 77], not to speak of [SGA 6] and Grothendieck's collected works. It may genuinely happen in practice that $\mathfrak{B}$ has enough injectives but $\mathfrak{A}$ does not have enough projectives, so the situation is not all symmetric. Thus the functor $A \mapsto R^nT(A, B)$ for fixed B is *not* a derived functor in the variable A. In the above references, we may take for $\mathfrak{A}$ the category of coherent sheaves on a variety, and for $\mathfrak{B}$ the category of all sheaves. We let $T = \mathrm{Hom}$. The locally free sheaves of finite rank are T-exact, and there are enough of them in $\mathfrak{A}$. There are enough injectives in $\mathfrak{B}$. And so it goes. The balancing act between T-exacts on one side, and injectives on the other is inherent to the situation.

Lemma 8.1. *Let T be a bifunctor satisfying* **HOM 1, HOM 2.** *Let $A \in \mathfrak{A}$, and let $M_A \to A \to 0$, that is*

$$\to M_1 \to M_0 \to A \to 0$$

be a T-exact resolution of A. Let $F^n(B) = H^n(T(M, B))$ for $B \in \mathfrak{B}$. Then F is a δ-functor and $F^0(B) = T(A, B)$. If in addition T satisfies **HOM 3,** *then $F^n(J) = 0$ for J injective and $n \geqq 1$.*

Proof. Given an exact sequence

$$0 \to B' \to B \to B'' \to 0$$

we get an exact sequence of complexes

$$0 \to T(M, B') \to T(M, B) \to T(M, B'') \to 0,$$

whence a cohomology sequence which makes F into a δ-functor. For $n = 0$ we get $F^0(B) = T(A, B)$ because $X \mapsto T(X, B)$ is contravariant and left exact for $X \in \mathfrak{A}$. If B is injective, then $F^n(B) = 0$ for $n \geqq 1$ by **HOM 3**, because $X \mapsto T(X, B)$ is exact. This proves the lemma.

Proposition 8.2. *Let T be a bifunctor satisfying* **HOM 1**, **HOM 2**, **HOM 3**. *Assume that $\mathfrak{B}$ has enough injectives. Let $A \in \mathfrak{A}$. Let*

$$M_A \to A \to 0$$

be a T-exact resolution of A. Then the two δ-functors

$$B \mapsto R^n T(A, B) \quad \text{and} \quad B \mapsto H^n(T(M, B))$$

are isomorphic as universal δ-functors vanishing on injectives, for $n \geqq 1$, and such that

$$R^0 T(A, B) = H^0(T(M), B) = T(A, B).$$

Proof. This comes merely from the universality of a δ-functor erasable by injectives.

We now look at the functoriality in A.

Lemma 8.3. *Let T satisfy* **HOM 1**, **HOM 2**, *and* **HOM 3**. *Assume that $\mathfrak{B}$ has enough injectives. Let*

$$0 \to A' \to A \to A'' \to 0$$

be a short exact sequence. Then for fixed B, we have a long exact sequence

$$0 \to T(A'', B) \to T(A, B) \to T(A', B) \to$$
$$\to R^1 T(A'', B) \to R^1 T(A, B) \to R^1 T(A', B) \to$$

such that the association

$$A \mapsto R^n T(A, B)$$

is a δ-functor.

Proof. Let $0 \to B \to I_B$ be an injective resolution of B. From the exactness of the functor $A \mapsto T(A, J)$, for J injective we get a short exact sequence of complexes

$$0 \to T(A'', I_B) \to T(A, I_B) \to T(A', I_B) \to 0.$$

Taking the associated long exact sequence of homology groups of these complexes yields the sequence of the proposition. (The functorality is left to the readers.)

If $T = \text{Hom}$, then the exact sequence looks like

$$0 \to \text{Hom}(A'', B) \to \text{Hom}(A, B) \to \text{Hom}(A', B) \to$$

$$\to \text{Ext}^1(A'', B) \to \text{Ext}^1(A, B) \to \text{Ext}^1(A', B) \to$$

and so forth.

We shall say that $\mathfrak{A}$ has **enough** T-exacts if given an object A in $\mathfrak{A}$ there is a T-exact M and an epimorphism

$$M \to A \to 0.$$

Proposition 8.4. *Let T satisfy* **HOM 1, HOM 2, HOM 3**. *Assume that $\mathfrak{B}$ has enough injectives. Fix $B \in \mathfrak{B}$. Then the association*

$$A \mapsto R^n T(A, B)$$

is a contravariant δ-functor on $\mathfrak{A}$ which vanishes on T-exacts, for $n \geqq 1$. If $\mathfrak{A}$ has enough T-exacts, then this functor is universal, coerasable by T-exacts, with value

$$R^0 T(A, B) = T(A, B).$$

Proof. By Lemma 8.3 we know that the association is a δ-functor, and it vanishes on T-exacts by Lemma 8.1. The last statement is then merely an application of the universality of erasable δ-functors.

Corollary 8.5. *Let $\mathfrak{A} = \mathfrak{B}$ be the category of modules over a ring. For fixed B, let $\text{ext}^n(A, B)$ be the left derived functor of $A \mapsto \text{Hom}(A, B)$, obtained by means of projective resolutions of A. Then*

$$\text{ext}^n(A, B) = \text{Ext}^n(A, B).$$

Proof. Immediate from Proposition 8.4.

The following proposition characterizes T-exacts cohomologically.

Proposition 8.6. *Let T be a bifunctor satisfying* **HOM 1, HOM 2, HOM 3.** *Assume that $\mathfrak{B}$ has enough injectives. Then the following conditions are equivalent:*

TE 1. *A is T-exact.*

TE 2. *For every B and every integer $n \geqq 1$, we have $R^nT(A, B) = 0$.*

TE 3. *For every B we have $R^1T(A, B) = 0$.*

Proof. Let

$$0 \to B \to I^0 \to I^1 \to$$

be an injective resolution of B. By definition, $R^nT(A, B)$ is the n-th homology of the sequence

$$0 \to T(A, I^0) \to T(A, I^1) \to T(A, I^2) \to$$

If A is T-exact, then this sequence is exact for $n \geqq 1$, so the homology is 0 and **TE 1** implies **TE 2**. Trivially, **TE 2** implies **TE 3**. Finally assume **TE 3**. Given an exact sequence

$$0 \to B' \to B \to B'' \to 0,$$

we have the homology sequence

$$0 \to T(A, B') \to T(A, B) \to T(A, B'') \to R^1T(A, B') \to.$$

If $R^1T(A, B') = 0$, then by definition A is T-exact, thus proving the proposition.

We shall say that an object A has T-**dimension** $\leqq d$ if

$$R^nT(A, B) = 0 \qquad \text{for } n > d \text{ and all } B.$$

Then the proposition states in particular that *A is T-exact if and only if A has T-dimension* 0.

Proposition 8.7. *Let T satisfy* **HOM 1, HOM 2, HOM 3.** *Assume that $\mathfrak{B}$ has enough injectives. Suppose that an object A admits a resolution*

$$0 \to E_d \to E_{d-1} \to \cdots \to E_0 \to A \to 0$$

where $E_0, \ldots, E_d$ are T-exact. Then A has T-dimension $\leqq d$. Assume this is the case. Let

$$0 \to Q \to L_{d-1} \to \cdots \to L_0 \to A \to 0$$

be a resolution where $L_0, \ldots, L_{d-1}$ are T-exact. Then Q is T-exact also.

Proof. By dimension shifting we conclude that Q has T-dimension 0, whence Q is T-exact by Proposition 8.6.

Proposition 8.7, like others, is used in the context of modules over a ring. In that case, we can take $T = \mathrm{Hom}$, and

$$R^nT(A, B) = \mathrm{Ext}^n(A, B).$$

For A to have T-dimension $\leqq d$ means that

$$\mathrm{Ext}^n(A, B) = 0 \qquad \text{for } n > d \text{ and all } B.$$

Instead of T-exact, one can then read projective in the proposition.

Let us formulate the analogous result for a bifunctor that will apply to the tensor product. Consider the following properties.

TEN 1. *T is covariant and right exact in the first variable.*

TEN 2. *T is covariant and right exact in the second variable.*

TEN 3. *For any projective object P the functor*

$$A \mapsto T(A, P)$$

is exact.

As for Hom, there is a possible fourth property which will play no role in this section:

TEN 4. *For any projective object Q the functor*

$$B \mapsto T(Q, B)$$

is exact.

Proposition 8.2′. *Let T be a bifunctor satisfying* **TEN 1**, **TEN 2**, **TEN 3**. *Assume that $\mathfrak{B}$ has enough projectives. Let $A \in \mathfrak{A}$. Let*

$$M_A \to A \to 0$$

be a T-exact resolution of A. Then the two δ-functors

$$B \mapsto L_nT(A, B) \quad \text{and} \quad B \mapsto H_n(T(M, B))$$

are isomorphic as universal δ-functors vanishing on projectives, and such that

$$L_0T(A, B) = H_0(T(M), B) = T(A, B).$$

Lemma 8.3′. *Assume that T satisfies* **TEN 1**, **TEN 2**, **TEN 3**. *Assume that $\mathfrak{B}$ has enough projectives. Let*

$$0 \to A' \to A \to A'' \to 0$$

be a short exact sequence. Then for fixed B, we have a long exact sequence:

$$\to L_1T(A', B) \to L_1T(A, B) \to L_1T(A'', B) \to$$
$$\to T(A', B) \to T(A, B) \to T(A'', B) \to 0$$

which makes the association $A \mapsto L_nT(A, B)$ *a* δ*-functor.*

Proposition 8.4′. *Let T satisfy* **TEN 1, TEN 2, TEN 3.** *Assume that* $\mathfrak{B}$ *has enough projectives. Fix* $B \in \mathfrak{B}$. *Then the association*

$$A \mapsto L_nT(A, B)$$

is a contravariant δ*-functor on* $\mathfrak{A}$ *which vanishes on T-exacts for* $n \geqq 1$. *If* $\mathfrak{A}$ *has enough T-exacts, then this functor is universal, coerasable by T-exacts, with the value*

$$L_0T(A, B) = T(A, B).$$

Corollary 8.8. *If there is a bifunctorial isomorphism* $T(A, B) \approx T(B, A)$, *and if B is T-exact, then for all A,* $L_nT(A, B) = 0$ *for* $n \geqq 1$. *In short, T-exact implies acyclic.*

Proof. Let $M_A = P_A$ be a projective resolution in Proposition 8.2′. By hypotheses, $X \mapsto T(X, B)$ is exact so $H_n(T(P, B)) = 0$ for $n \geqq 1$; so the corollary is a consequence of the proposition.

The above corollary is formulated so as to apply to the tensor product.

Proposition 8.6′. *Let T be a bifunctor satisfying* **TEN 1, TEN 2, TEN 3.** *Assume that* $\mathfrak{B}$ *has enough projectives. Then the following conditions are equivalent:*

TE 1. *A is T-exact.*

TE 2. *For every B and every integer* $n \geqq 1$ *we have* $L_nT(A, B) = 0$.

TE 3. *For every B, we have* $L_1T(A, B) = 0$.

Proof. We repeat the proof of 8.6 so the reader can see the arrows pointing in different ways.

Let

$$\to Q_1 \to Q_0 \to B \to 0$$

be a projective resolution of B. By definition, $L_nT(A, B)$ is the n-th homology of the sequence

$$\to T(A, Q_1) \to T(A, Q_0) \to 0.$$

If A is T-exact, then this sequence is exact for $n \geqq 1$, so the homology is 0, and **TE 1** implies **TE 2**. Trivially, **TE 2** implies **TE 3**. Finally, assume **TE 3**. Given an exact sequence

$$0 \to B' \to B \to B'' \to 0$$

we have the homology sequence

$$\to L_1 T(A, B'') \to T(A, B') \to T(A, B) \to T(A, B'') \to 0.$$

If $L_1 T(A, B'')$ is 0, then by definition, A is T-exact, thus proving the proposition.

§9. SPECTRAL SEQUENCES

This section is included for convenience of reference, and has two purposes: first, to draw attention to an algebraic gadget which has wide applications in topology, differential geometry, and algebraic geometry, see Griffiths-Harris, [GrH 78]; second, to show that the basic description of this gadget in the context in which it occurs most frequently can be done in just a few pages.

In the applications mentioned above, one deals with a filtered complex (which we shall define later), and a complex may be viewed as a graded object, with a differential d of degree 1. To simplify the notation at first, we shall deal with filtered objects and omit the grading index from the notation. This index is irrelevant for the construction of the spectral sequence, for which we follow Godement.

So let F be an object with a differential (i.e. endomorphism) d such that $d^2 = 0$. We assume that F is **filtered**, that is that we have a sequence

$$F = F^0 \supset F^1 \supset F^2 \supset \cdots \supset F^n \supset F^{n+1} = \{0\},$$

and that $dF^p \subset F^p$. This data is called a **filtered differential object**. (We assume that the filtration ends with 0 after a finite number of steps for convenience.)

One defines the **associated graded object**

$$\operatorname{Gr} F = \bigoplus_{p \geqq 0} \operatorname{Gr}^p F \quad \text{where} \quad \operatorname{Gr}^p F = F^p/F^{p+1}.$$

In fact, $\operatorname{Gr} F$ is a complex, with a differential of degree 0 induced by d itself, and we have the homology $H(\operatorname{Gr}^p F)$.

The filtration $\{F^p\}$ also induces a filtration on the homology $H(F, d) = H(F)$; namely we let

$$H(F)^p = \text{image of } H(F^p) \text{ in } H(F).$$

Since d maps F^p into itself, $H(F^p)$ is the homology of F^p with respect to the restriction of d to F^p, and it has a natural image in $H(F)$ which yields this filtration. In particular, we then obtain a graded object associated with the filtered homology, namely

$$\operatorname{Gr} H(F) = \bigoplus \operatorname{Gr}^p H(F).$$

A **spectral sequence** is a sequence $\{E_r, d_r\}$ $(r \geqq 0)$ of graded objects

$$E_r = \bigoplus_{p \geqq 0} E_r^p$$

together with homomorphisms (also called **differentials**) of degree r,

$$d_r : E_r^p \to E_r^{p+r}$$

satisfying $d_r^2 = 0$, and such that the homology of E_r is E_{r+1}, that is

$$H(E_r) = E_{r+1}.$$

In practice, one usually has $E_r = E_{r+1} = \cdots$ for $r \geqq r_0$. This limit object is called E_∞, and one says that the spectral sequence **abuts** to E_∞. Actually, to be perfectly strict, instead of equalities one should really be given isomorphisms, but for simplicity, we use equalities.

Proposition 9.1. *Let F be a filtered differential object. Then there exists a spectral sequence $\{E_r\}$ with:*

$$E_0^p = F^p/F^{p+1}; \qquad E_1^p = H(\operatorname{Gr}^p F); \qquad E_\infty^p = \operatorname{Gr}^p H(F).$$

Proof. Define

$$\boxed{\begin{aligned} Z_r^p &= \{x \in F^p \text{ such that } dx \in F^{p+r}\} \\ E_r^p &= Z_r^p/[dZ_{r-1}^{p-(r-1)} + Z_{r-1}^{p+1}]. \end{aligned}}$$

The definition of E_r^p makes sense, since Z_r^p is immediately verified to contain $dZ_{r-1}^{p-(r-1)} + Z_{r-1}^{p+1}$. Furthermore, d maps Z_r^p into Z_r^{p+r}, and hence includes a homomorphism

$$d_r : E_r^p \to E_r^{p+r}.$$

We shall now compute the homology and show that it is what we want.

First, for the cycles: An element $x \in Z_r^p$ represents a cycle of degree p in E_r if and only if $dx \in dZ_{r+1}^{p+1} + Z_{r-1}^{p+r+1}$, in other words

$$dx = dy + z, \qquad \text{with } y \in Z_{r-1}^{p+1} \quad \text{and} \quad z \in Z_{r-1}^{p+r+1}.$$

Write $x = y + u$, so $du = z$. Then $u \in F^p$ and $du \in F^{p+r+1}$, that is $u \in Z^p_{r+1}$. It follows that

$$p\text{-cycles of } E_r = (Z^p_{r+1} + Z^{p+1}_{r-1})/(dZ^{p-r+1}_{r-1} + Z^{p+1}_{r-1}).$$

On the other hand, the p-boundaries in E_r are represented by elements of dZ^{p-r}_r, which contains dZ^{p-r+1}_{r-1}. Hence

$$p\text{-boundaries of } E_r = (dZ^{p-r}_r + Z^{p+1}_{r-1})/(dZ^{p-r+1}_{r-1} + Z^{p+1}_{r-1}).$$

Therefore

$$\begin{aligned} H^p(E_r) &= (Z^p_{r+1} + Z^{p+1}_{r-1})/(dZ^{p-r}_r + Z^{p+1}_{r-1}) \\ &= Z^p_{r+1}/(Z^p_{r+1} \cap (dZ^{p-r}_r + Z^{p+1}_{r-1})). \end{aligned}$$

Since

$$Z^p_{r+1} \supset dZ^{p-r}_r \quad \text{and} \quad Z^p_{r+1} \cap Z^{p+1}_{r-1} = Z^{p+1}_r,$$

it follows that

$$H^p(E_r) = Z^p_{r+1}/(dZ^{p-r}_r + Z^{p+1}_r) = E^p_{r+1},$$

thus proving the property of a spectral sequence.

Remarks. It is sometimes useful in applications to note the relation

$$dZ^{p-(r-1)}_{r-1} + Z^{p+1}_{r-1} = Z^p_r \cap (dF^{p-r+1} + F^{p+1}).$$

The verification is immediate, but Griffiths-Harris use the expression on the right in defining the spectral sequence, whereas Godement uses the expression on the left as we have done above. Thus the spectral sequence may also be defined by

$$\boxed{E^p_r = Z^p_r \mod(dF^{p-r+1} + F^{p+1}).}$$

This is to be interpreted in the sense that Z mod S means

$$(Z + S)/S \quad \text{or} \quad Z/(Z \cap S).$$

The term E^p_0 is F^p/F^{p+1} immediately from the definitions, and by the general property already proved, we get $E^p_1 = H(F^p/F^{p+1})$. As to E^p_∞, for r large we have $Z^p_r = Z^p =$ cycles in F^p, and

$$\boxed{E^p_\infty = Z^p/(Z^{p+1} + (dF^0 \cap F^p))}$$

which is independent of r, and is precisely $\mathrm{Gr}^p H(F)$, namely the p-graded component of $H(F)$, thus proving the theorem.

The differential d_1 can be specified as follows.

Proposition 9.2. *The homomorphism*

$$d_1 : E_1^p \to E_1^{p+1}$$

is the coboundary operator arising from the exact sequence

$$0 \to F^{p+1}/F^{p+2} \to F^p/F^{p+2} \to F^p/F^{p+1} \to 0$$

viewing each term as a complex with differential induced by d.

Proof. Indeed, the coboundary

$$\delta : E_1^p = H(F^p/F^{p+1}) \to H(F^{p+1}/F^{p+2}) = E_1^{p+1}$$

is defined on a representative cycle z by dz, which is the same way that we defined d_1.

In most applications, the filtered differential object is itself graded, because it arises from the following situation. Let K be a complex, $K = (K^p, d)$ with $p \geqq 0$ and d of degree 1. By a **filtration** FK, also called a **filtered complex**, we mean a decreasing sequence of subcomplexes

$$K = F^0K \supset F^1K \supset F^2K \supset \cdots \supset F^nK \supset F^{n+1}K = \{0\}.$$

Observe that a short exact sequence of complexes

$$0 \to K' \to K \to K'' \to 0$$

gives rise to a filtration $K \supset K' \supset \{0\}$, viewing K' as a subcomplex.

To each filtered complex FK we associated the complex

$$\mathrm{Gr}\, FK = \mathrm{Gr}\, K = \bigoplus_{p \geqq 0} \mathrm{Gr}^p K,$$

where

$$\mathrm{Gr}^p K = F^pK/F^{p+1}K,$$

and the differential is the obvious one. The filtration F^pK on K also induces a filtration $F^pH(K)$ on the cohomology, by

$$F^pH^q(K) = F^pZ^q/F^pB^q.$$

The associated graded homology is

$$\mathrm{Gr}\, H(K) = \bigoplus_{p,q} \mathrm{Gr}^p\, H^q(K),$$

where

$$\mathrm{Gr}^p\, H^q(K) = F^p H^q(K)/F^{p+1} H^q(K).$$

A **spectral sequence** is a sequence $\{E_r, d_r\}$ $(r \geqq 0)$ of bigraded objects

$$E_r = \bigoplus_{p,q \geqq 0} E_r^{p,q}$$

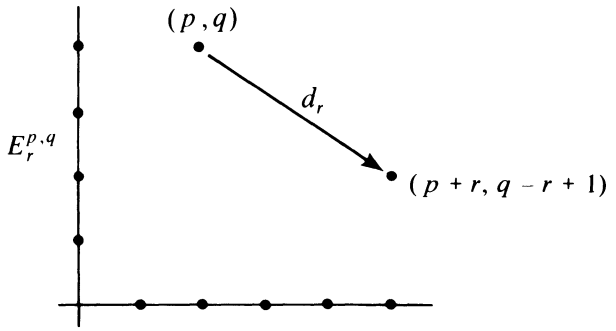
together with homomorphisms (called **differentials**)

$$d_r : E_r^{p,q} \to E_r^{p+r,q-r+1} \quad \text{satisfying} \quad d_r^2 = 0,$$

and such that the homology of E_r is E_{r+1}, that is

$$H(E_r) = E_{r+1}.$$

A spectral sequence is usually represented by the following picture:



In practice, one usually has $E_r = E_{r+1} = \cdots$ for $r \geqq r_0$. This limit object is called E_∞, and one says that the spectral sequence **abuts** to E_∞.

Proposition 9.3. *Let FK be a filtered complex. Then there exists a spectral sequence $\{E_r\}$ with:*

$$E_0^{p,q} = F^p K^{p+q}/F^{p+1} K^{p+q};$$

$$E_1^{p,q} = H^{p+q}(\mathrm{Gr}^p\, K);$$

$$E_\infty^{p,q} = \mathrm{Gr}^p\, (H^{p+q}(K)).$$

The last relation is usually written

$$E_r \Rightarrow H(K),$$

and we say that the spectral sequence **abuts** *to $H(K)$.*

The statement of Proposition 9.3 is merely a special case of Proposition 9.1, taking into account the extra graduation.

One of the main examples is the spectral sequence associated with a double complex

$$K = \bigoplus_{p,q \geqq 0} K^{p,q}$$

which is a bigraded object, together with differentials

$$d' : K^{p,q} \to K^{p+1,q} \quad \text{and} \quad d'' : K^{p,q} \to K^{p,q+1}$$

satisfying

$$d'^2 = d''^2 = 0 \quad \text{and} \quad d'd'' + d''d' = 0.$$

We denote the double complex by (K, d', d''). The associated single complex $(\mathrm{Tot}(K), D)$ (Tot for **total complex**), abbreviated K^*, is defined by

$$K^n = \bigoplus_{p+q=n} K^{p,q} \quad \text{and} \quad D = d' + d''.$$

There are two filtrations on (K^*, D) given by

$$'F^pK^n = \bigoplus_{\substack{p'+q=n \\ p' \geqq p}} K^{p',q}$$

$$''F^qK^n = \bigoplus_{\substack{p+q'=n \\ q' \geqq q}} K^{p,q''}.$$

There are two spectral sequences $\{'E_r\}$ and $\{''E_r\}$, both abutting to $H(\mathrm{Tot}(K))$. For applications, see [GrH 78], Chapter 3, §5; and also, for instance, [FuL 85], Chapter V. There are many situations when dealing with a double complex directly is a useful substitute for using spectral sequences, which are derived from double complexes anyhow.

We shall now derive the existence of a spectral sequence in one of the most important cases, the **Grothendieck spectral sequence** associated with the composite of two functors. *We assume that our abelian category has enough injectives.*

Let $C = \bigoplus C^p$ be a complex, and suppose $C^p = 0$ if $p < 0$ for simplicity. We define **injective resolution** of C to be a resolution

$$0 \to C \to I^0 \to I^1 \to I^2 \to \cdots$$

written briefly

$$0 \to C \to I_C$$

such that each I^j is a complex, $I^j = \bigoplus I^{j,p}$, with differentials

$$d^{j,p} : I^{j,p} \to I^{j,p+1}$$

and such that $I^{j,p}$ is an injective object. Then in particular, for each p we get an injective resolution of C^p, namely:

$$0 \to C^p \to I^{0,p} \to I^{1,p} \to \cdots$$

We let:

$$Z^{j,p} = \operatorname{Ker} d^{j,p} = \text{cycles in degree } p$$

$$B^{j,p} = \operatorname{Im} d^{j,p-1} = \text{boundaries in degree } p$$

$$H^{j,p} = Z^{j,p}/B^{j,p} = \text{homology in degree } p.$$

We then get complexes

$$0 \to Z^p(C) \to Z^{0,p} \to Z^{1,p} \to$$

$$0 \to B^p(C) \to B^{0,p} \to B^{1,p} \to$$

$$0 \to H^p(C) \to H^{0,p} \to H^{1,p} \to$$

We say that the resolution $0 \to C \to I_C$ is **fully injective** if these three complexes are injective resolutions of $Z^p(C)$, $B^p(C)$ and $H^p(C)$ respectively.

Lemma 9.4. *Let*

$$0 \to M' \to M \to M'' \to 0$$

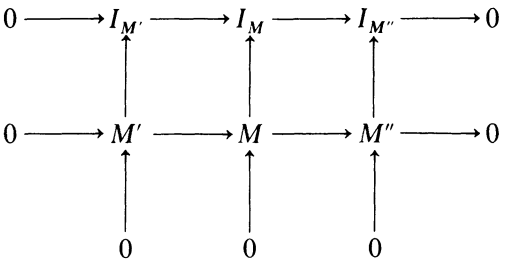
be a short exact sequence. Let

$$0 \to M' \to I_{M'} \quad \text{and} \quad 0 \to M'' \to I_{M''}$$

be injective resolutions of M' and M''. Then there exists an injective resolution

$$0 \to M \to I_M$$

of M and morphisms which make the following diagram exact and commutative:



Proof. The proof is the same as at the beginning of the proof of Theorem 6.1.

Lemma 9.5. *Given a complex C there exists a fully injective resolution of C.*

Proof. We insert the kernels and cokernels in C, giving rise to the short exact sequences with boundaries B^p and cycles Z^p:

$$0 \to B^p \to Z^p \to H^p \to 0$$

$$0 \to Z^{p-1} \to C^{p-1} \to B^p \to 0.$$

We proceed inductively. We start with an injective resolution of

$$0 \to Z^{p-1} \to C^{p-1} \to B^p \to 0$$

using Lemma 9.4. Next let

$$0 \to H^p \to I_{H^p}$$

be an injective resolution of H^p. By Lemma 9.4 there exists an injective resolution

$$0 \to Z^p \to I_{Z^p}$$

which fits in the middle of the injective resolutions we already have for B^p and H^p. This establishes the inductive step, and concludes the proof.

Given a left exact functor G on an abelian category with enough injectives, we say that an object X is **G-acyclic** if $R^pG(X) = 0$ for $p \geqq 1$. Of course,

$$R^0G(X) = G(X).$$

Theorem 9.6. (Grothendieck spectral sequence). *Let*

$$T: \mathfrak{A} \to \mathfrak{B} \quad \text{and} \quad G: \mathfrak{B} \to \mathfrak{C}$$

be covariant left exact functors such that if I is injective in $\mathfrak{A}$, then $T(I)$ is G-acyclic. Then for each A in $\mathfrak{A}$ there is a spectral sequence $\{E_r(A)\}$, such that

$$E_2^{p,q}(A) = R^pG(R^qT(A))$$

and $E_r^{p,q}$ abuts (with respect to p) to $R^{p+q}(GT)(A)$, where q is the grading index.

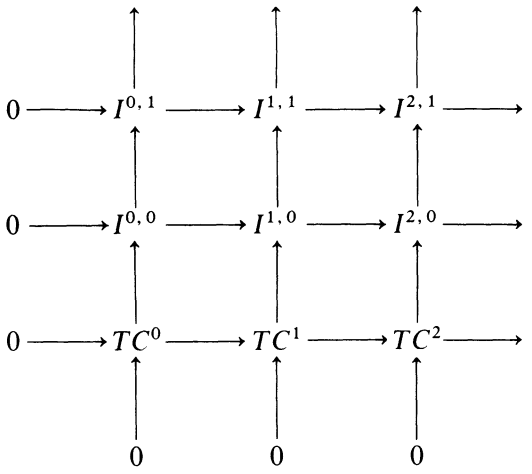
Proof. Let A be an object of $\mathfrak{A}$, and let $0 \to A \to C_A$ be an injective resolution. We apply T to get a complex

$$TC: \quad 0 \to TC^0 \to TC^1 \to TC^2 \to$$

By Lemma 9.5 there exists a fully injective resolution

$$0 \to TC \to I_{TC}$$

which has the 2-dimensional representation:



Then GI is a double complex. Let $\mathrm{Tot}(GI)$ be the associated single complex. We now consider each of the two possible spectral sequences in succession, which we denote by ${}^1E_r^{p,q}$ and ${}^2E_r^{p,q}$.

The first one is the easiest. For fixed p, we have an injective resolution

$$0 \to TC^p \to I^p_{TC}$$

where we write I^p_{TC} instead of I_{TC^p}. This is the p-th column in the diagram. By definition of derived functors, GI^p is a complex whose homology is R^qG, in other words, taking homology with respect to d'' we have

$$''H^{p,q}(GI) = H^q(GI^p) = (R^qG)(TC^p).$$

By hypothesis, C^p injective implies that $(R^qG)(TC^p) = 0$ for $q > 0$. Since G is left exact, we have $R^0G(TC^p) = TC^p$. Hence we get

$$''H^{p,q}(GI) = \begin{cases} GT(C^p) & \text{if } q = 0 \\ 0 & \text{if } q > 0. \end{cases}$$

Hence the non-zero terms are on the p-axis, which looks like

$$0 \to GT(C^0) \to GT(C^1) \to GT(C^2) \to$$

Taking $'H^p$ we get

$${}^1E_2^{p,q}(A) = \begin{cases} R^p(GT)(A) & \text{if } q = 0 \\ 0 & \text{if } q > 0. \end{cases}$$

This yields

$$H^n(\mathrm{Tot}(GI)) \approx R^n(GT)(A).$$

The second one will use the full strength of Lemma 9.5, which had not been used in the first part of the proof, so it is now important that the resolution I_{TC} is fully injective. We therefore have injective resolutions

$$0 \to Z^p(TC) \to {}^1Z^{0,p} \to {}^1Z^{1,p} \to {}^1Z^{2,p} \to$$
$$0 \to B^p(TC) \to {}^1B^{0,p} \to {}^1B^{1,p} \to {}^1B^{2,p} \to$$
$$0 \to H^p(TC) \to {}^1H^{0,p} \to {}^1H^{1,p} \to {}^1H^{2,p} \to$$

and the exact sequences

$$0 \to {}^1Z^{q,p} \to I^{q,p} \to {}^1B^{q+1,p} \to 0$$
$$0 \to {}^1B^{q,p} \to {}^1Z^{q,p} \to {}^1H^{q,p} \to 0$$

split because of the injectivity of the terms. We denote by $I^{(p)}$ the p-th row of the double complex $I = \{I^{q,p}\}$. Then we find:

$$\begin{aligned} {}'H^{q,p}(GI) = H^q(GI^{(p)}) &= G^1Z^{q,p}/G^1B^{q,p} && \text{by the first split sequence} \\ &= G'H^{q,p}(I) && \text{by the second split sequence} \end{aligned}$$

because applying the functor G to a split exact sequence yields a split exact sequence.

Then

$${}^2E_2^{p,q} = {}''H^p({}'H^{q,p}(GI) = H^p(G^1H^{q,p}(I)).$$

By the full injectivity of the resolutions, the complex ${}'H^{q,p}(I)$ with $p \geqq 0$ is an injective resolution of

$$H^q(TC) = (R^qT)(A).$$

Furthermore, we have

$$H^p(G'H^{q,p}) = R^pG(R^qT(A)),$$

since a derived functor is the homology of an injective resolution. This proves that $(R^pG)R^qT(A))$ abuts to $R^n(GT)(A)$, and concludes the proof of the theorem.

Just to see the spectral sequence at work, we give one application relating it to the Euler characteristic discussed in §3.

Let $\mathfrak{A}$ have enough injectives, and let

$$T : \mathfrak{A} \to \mathfrak{B}$$

be a covariant left exact functor. Let $\mathfrak{F}_\mathfrak{a}$ be a family of objects in $\mathfrak{A}$ giving rise to a **K**-group. More precisely, in a short exact sequence in $\mathfrak{A}$, if two of the objects lie in $\mathfrak{F}_\mathfrak{a}$, then so does the third. We also assume that the objects of $\mathfrak{F}_\mathfrak{a}$ have **finite *RT*-dimension**, which means by definition that if $A \in \mathfrak{F}_\mathfrak{a}$ then $R^iT(A) = 0$

for all i sufficiently large. We could take $\mathfrak{F}_{\mathfrak{A}}$ in fact to be the family of all objects in $\mathfrak{A}$ which have finite RT-dimension.

We define the **Euler characteristic associated with** T on $\mathbf{K}(\mathfrak{F}_{\mathfrak{A}})$ to be

$$\chi_T(A) = \sum_{i=0}^{\infty} (-1)^i \operatorname{cl}(R^iT(A)).$$

The cl denotes the class in the **K**-group $K(\mathfrak{F}_{\mathfrak{B}})$ associated with some family $\mathfrak{F}_{\mathfrak{B}}$ of objects in $\mathfrak{B}$, and such that $R^iT(A) \in \mathfrak{F}_{\mathfrak{B}}$ for all $A \in \mathfrak{F}_{\mathfrak{A}}$. This is the minimum required for the formula to make sense.

Lemma 9.7. *The map χ_T extends to a homomorphism*

$$\mathbf{K}(\mathfrak{F}_{\mathfrak{A}}) \to \mathbf{K}(\mathfrak{F}_{\mathfrak{B}}).$$

Proof. Let

$$0 \to A' \to A \to A'' \to 0$$

be an exact sequence in $\mathfrak{F}$. Then we have the cohomology sequence

$$\to R^iT(A') \to R^iT(A) \to R^iT(A'') \to R^{i+1}T(A') \to$$

in which all but a finite number of terms are 0. Taking the alternating sum in the **K**-group shows that χ_T is an Euler–Poincaré map, and concludes the proof.

Note that we have merely repeated something from §3, in a jazzed up context. In the next theorem, we have another functor

$$G : \mathfrak{B} \to \mathfrak{C},$$

and we also have a family $\mathfrak{F}_{\mathfrak{C}}$ giving rise to a **K**-group $\mathbf{K}(\mathfrak{F}_{\mathfrak{C}})$. We suppose that we can perform the above procedure at each step, and also need some condition so that we can apply the spectral sequence. So, precisely, we assume:

CHAR 1. For all i, R^iT maps $\mathfrak{F}_{\mathfrak{A}}$ into $\mathfrak{F}_{\mathfrak{B}}$, R^iG maps $\mathfrak{F}_{\mathfrak{B}}$ into $\mathfrak{F}_{\mathfrak{C}}$, and $R^i(GT)$ maps $\mathfrak{F}_{\mathfrak{A}}$ into $\mathfrak{F}_{\mathfrak{C}}$.

CHAR 2. Each subobject of an element of $\mathfrak{F}_{\mathfrak{A}}$ lies in $\mathfrak{F}_{\mathfrak{A}}$ and has finite RT- and $R(GT)$-dimension; each subobject of an element of $\mathfrak{F}_{\mathfrak{B}}$ lies in $\mathfrak{F}_{\mathfrak{B}}$ and has finite RG-dimension.

Theorem 9.8. *Assume that $T : \mathfrak{A} \to \mathfrak{B}$ and $G : \mathfrak{B} \to \mathfrak{C}$ satisfy the conditions* **CHAR 1** *and* **CHAR 2**. *Also assume that T maps injectives to G-acyclics. Then*

$$\chi_G \circ \chi_T = \chi_{GT}.$$

Proof. By Theorem 9.6, the Grothendieck spectral sequence of the composite functor implies the existence of a filtration

$$\cdots \subset F^p R^n(GT)(A) \subset F^{p+1} R^n(GT)(A) \subset \cdots$$

of $R^n(GT)(A)$, such that

$$F^{p+1}/F^p \approx E_\infty^{p,n-p}.$$

Then

$$\begin{aligned}\chi_{GT}(A) &= \sum_{n=0}^{\infty} (-1)^n \operatorname{cl}(R^n(GT)(A)) \\ &= \sum_{n=0}^{\infty} (-1)^n \sum_{p=0}^{\infty} \operatorname{cl}(E_\infty^{p,n-p}) \\ &= \sum_{n=0}^{\infty} (-1)^n \operatorname{cl}(E_\infty^n).\end{aligned}$$

On the other hand,

$$\chi_T(A) = \sum_{q=0}^{\infty} (-1)^q \operatorname{cl}(R^q T(A))$$

and so

$$\begin{aligned}\chi_G \circ \chi_T(A) &= \sum_{q=0}^{\infty} (-1)^q \chi_G(R^q T(A)) \\ &= \sum_{q=0}^{\infty} (-1)^q \sum_{p=0}^{\infty} (-1)^p \operatorname{cl}(R^p G(R^q T(A)) \\ &= \sum_{n=0}^{\infty} (-1)^n \sum_{p=0}^{n} \operatorname{cl}(R^p G(R^{n-p} T(A)) \\ &= \sum_{n=0}^{\infty} (-1)^n \operatorname{cl}(E_2^n).\end{aligned}$$

Since E_{r+1} is the homology of E_r, we get

$$\sum_{n=0}^{\infty} (-1)^n \operatorname{cl}(E_2^n) = \sum_{n=0}^{\infty} (-1)^n \operatorname{cl}(E_3^n) = \cdots = \sum_{n=0}^{\infty} (-1)^n \operatorname{cl}(E_\infty^n).$$

This concludes the proof of the theorem.

EXERCISES

1. Prove that the example of the standard complex given in §1 is actually a complex, and is exact, so it gives a resolution of $\mathbf{Z}$. [*Hint:* To show that the sequence of the standard complex is exact, choose an element $z \in S$ and define $h : E^i \to E^{i+1}$ by letting

$$h(x_0, \ldots, x_i) = (z, x_0, \ldots, x_i).$$

Prove that $dh + hd = \mathrm{id}$, and that $dd = 0$. Exactness follows at once.]

Cohomology of groups

2. Let G be a group. Use G as the set S in the standard complex. Define an action of G on the standard complex E by letting

$$x(x_0, \ldots, x_i) = (xx_0, \ldots, xx_i).$$

Prove that each E_i is a free module over the group ring $\mathbf{Z}[G]$. Thus if we let $R = \mathbf{Z}[G]$ be the group ring, and consider the category $\mathrm{Mod}(G)$ of G-modules, then the standard complex gives a free resolution of $\mathbf{Z}$ in this category.

3. The standard complex E was written in homogeneous form, so the boundary maps have a certain symmetry. There is another complex which exhibits useful features as follows. Let F^i be the free $\mathbf{Z}[G]$-module having for basis i-tuples (rather than $(i+1)$-tuples) $(x_1, \ldots, x_i)$. For $i = 0$ we take $F_0 = \mathbf{Z}[G]$ itself. Define the boundary operator by the formula

$$d(x_1, \ldots, x_i) = x_1(x_2, \ldots, x_i) + \sum_{j=1}^{i-1} (-1)^j (x_1, \ldots, x_j x_{j+1}, \ldots, x_i) + (-1)^{i+1}(x_1, \ldots, x_i).$$

Show that $E \approx F$ (as complexes of G-modules) via the association

$$(x_1, \ldots, x_i) \mapsto (1, x_1, x_1x_2, \ldots, x_1x_2 \cdots x_i),$$

and that the operator d given for F corresponds to the operator d given for E under this isomorphism.

4. If A is a G-module, let A^G be the submodule consisting of all elements $v \in A$ such that $xv = v$ for all $x \in G$. Thus A^G has trivial G-action. (This notation is convenient, but is *not* the same as for the induced module of Chapter XVIII.)
   (a) Show that if $H^q(G, A)$ denotes the q-th homology of the complex $\mathrm{Hom}_G(E, A)$, then $H^0(G, A) = A^G$. Thus the left derived functors of $A \mapsto A^G$ are the homology groups of the complex $\mathrm{Hom}_G(E, A)$, or for that matter, of the complex $\mathrm{Hom}(F, A)$, where F is as in Exercise 3.
   (b) Show that the group of 1-cycles $Z^1(G, A)$ consists of those functions $f : G \to A$ satisfying

$$f(x) + xf(y) = f(xy) \text{ for all } x, y \in G.$$

Show that the subgroup of coboundaries $B^1(G, A)$ consists of those functions f for which there exists an element $a \in A$ such that $f(x) = xa - a$. The factor group is then $H^1(G, A)$. See Chapter VI, §10 for the determination of a special case.

(c) Show that the group of 2-cocycles $Z^2(G, A)$ consists of those functions $f: G \to A$ satisfying

$$xf(y, z) - f(xy, z) + f(x, yz) - f(x, y) = 0.$$

Such 2-cocycles are also called **factor sets**, and they can be used to describe isomorphism classes of group extensions, as follows.

5. **Group extensions**. Let W be a group and A a normal subgroup, written multiplicatively. Let $G = W/A$ be the factor group. Let $F: G \to W$ be a choice of coset representatives. Define

$$f(x, y) = F(x)F(y)F(xy)^{-1}.$$

(a) Prove that f is A-valued, and that $f: G \times G \to A$ is a 2-cocycle.

(b) Given a group G and an abelian group A, we view an extension W as an exact sequence

$$1 \to A \to W \to G \to 1.$$

Show that if two such extensions are isomorphic then the 2-cocycles associated to these extensions as in (a) define the same class in $H^1(G, A)$.

(c) Prove that the map which we obtained above from isomorphism classes of group extensions to $H^2(G, A)$ is a bijection.

6. **Morphisms of the cohomology functor.** Let $\lambda: G' \to G$ be a group homomorphism. Then λ gives rise to an exact functor

$$\Phi_\lambda : \mathrm{Mod}(G) \to \mathrm{Mod}(G'),$$

because every G-module can be viewed as a G'-module by defining the operation of $\sigma' \in G'$ to be $\sigma' a = \lambda(\sigma')a$. Thus we obtain a cohomology functor $H^{G'} \circ \Phi_\lambda$.

Let G' be a subgroup of G. In dimension 0, we have a morphism of functors

$$\lambda^* : H^0_G \to H^0_{G'} \circ \Phi_\lambda \text{ given by the inclusion } A^G \subset\!\to A^{G'} = \Phi_\lambda(A)^{G'}.$$

(a) Show that there is a unique morphism of δ-functors

$$\lambda^* : H_G \to H_{G'} \circ \Phi_\lambda$$

which has the above effect on H^0_G. We have the following important special cases.

Restriction. Let H be a subgroup of G. Let A be a G-module. A function from G into A restricts to a function from H into A. In this way, we get a natural homomorphism called the **restriction**

$$\mathrm{res}: H^q(G, A) \to H^q(H, A).$$

Inflation. Suppose that H is normal in G. Let A^H be the subgroup of A consisting of those elements fixed by H. Then it is immediately verified that A^H is stable under G, and so is a G/H-module. The inclusion $A^H \subset\!\to A$ induces a homomorphism

$$H^q_G(u) = u_q : H^q(G, A^H) \to H^q(A).$$

Define the **inflation**

$$\inf^H_{G/H} : H^q(G/H, A^H) \to H^q(G, A)$$

as the composite of the functorial morphism $H^q(G/H, A^H) \to H^q(G, A^H)$ followed by the induced homomorphism $u_q = H^q_G(u)$ as above.

In dimension 0, the inflation gives the identity $(A^H)^{G/H} = A^G$.

(b) Show that the inflation can be expressed on the standard cochain complex by the natural map which to a function of G/H in A^H associates a function of G into $A^H \subset A$.

(c) Prove that the following sequence is exact.

$$0 \to H^1(G/H, A^H) \xrightarrow{\inf} H^1(G, A) \xrightarrow{\text{res}} H^1(H, A).$$

(d) Describe how one gets an operation of G on the cohomology functor H_G "by conjugation" and functoriality.

(e) In (c), show that the image of restriction on the right actually lies in $H^1(H, A)^G$ (the fixed subgroup under G).

Remark. There is an analogous result for higher cohomology groups, whose proof needs a spectral sequence of Hochschild-Serre. See [La 96], Chapter VI, §2, Theorem 2. It is actually this version for H^2 which is applied to $H^2(G, K^*)$, when K is a Galois extension, and is used in class field theory [ArT 67].

7. Let G be a group, B an abelian group and $M_G(B) = M(G, B)$ the set of mappings from G into B. For $x \in G$ and $f \in M(G, B)$ define $([x]f)(y) = f(yx)$.

(a) Show that $B \mapsto M_G(B)$ is a covariant, additive, exact functor from Mod($\mathbf{Z}$) (category of abelian groups) into Mod(G).

(b) Let G' be a subgroup of G and $G = \bigcup x_j G'$ a coset decomposition. For $f \in M(G, B)$ let f_j be the function in $M(G', B)$ such that $f_j(y) = f(x_j y)$. Show that the map

$$f \mapsto \prod_j f_j$$

is a G'-isomorphism from $M(G, B)$ to $\prod_j M(G', B)$.

8. For each G-module $A \in \text{Mod}(G)$, define $\varepsilon_A : A \to M(G, A)$ by the condition $\varepsilon_A(a) =$ the function f_a such that $f_z(\sigma) = \sigma a$ for $\sigma \in G$. Show that $a \mapsto f_a$ is a G-module embedding, and that the exact sequence

$$0 \to A \xrightarrow{\varepsilon_A} M(G, A) \to X_A = \text{coker } \varepsilon_A \to 0$$

splits over $\mathbf{Z}$. (In fact, the map $f \mapsto f(e)$ splits the left side arrow.)

9. Let $B \in \text{Mod}(\mathbf{Z})$. Let H^q be the left derived functor of $A \mapsto A^G$.

(a) Show that $H^q(G, M_G(B)) = 0$ for all $q > 0$. [*Hint*: use a contracting homotopy

$$s : C^r(G, M_G(B)) \to C^{r-1}(G, M_G(B)) \qquad \text{by} \qquad (sf)_{x_2,\ldots,x_r}(x) = f_{x,x_2,\ldots,x_r}(1).$$

Show that $f = sdf + dsf$.] Thus M_G erases the cohomology functor.

(b) Also show that for all subgroups G' of G one has $H^q(G', M_G(B)) = 0$ for $q > 0$.

10. Let G be a group and S a subgroup. Show that the bifunctors

$$(A, B) \mapsto \text{Hom}_G(A, M^S_G(B)) \text{ and } (A, B) \mapsto \text{Hom}_S(A, B)$$

on Mod(G) $\times$ Mod(S) with value in Mod($\mathbf{Z}$) are isomorphic. The isomorphism is given by the maps

$$\varphi \mapsto (a \mapsto g_a), \text{ for } \varphi \in \text{Hom}_S(A, B), \text{ where } g_a(\sigma) = \varphi(\sigma a),\ g_a \in M^S_G(B).$$

The inverse mapping is given by

$$f \mapsto f(1) \text{ with } f \in \mathrm{Hom}_G(A, M_G^S(B)).$$

Recall that $M_G^S(B)$ was defined in Chapter XVIII, §7 for the induced representation. Basically you should already know the above isomorphism.

11. Let G be a group and S a subgroup. Show that the map

$$H^q(G, M_G^S(B)) \to H^q(S, B) \text{ for } B \in \mathrm{Mod}(S),$$

obtained by composing the restriction res_S^G with the S-homomorphism $f \mapsto f(1)$, is an isomorphism for $q > 0$. [*Hint*: Use the uniqueness theorem for cohomology functors.]

12. Let G be a group. Let $\varepsilon : \mathbf{Z}[G] \to \mathbf{Z}$ be the homomorphism such that $\varepsilon\left(\sum n(x)x\right) = \sum n(x)$. Let I_G be its kernel. Prove that I_G is an ideal of $\mathbf{Z}[G]$ and that there is an isomorphism of functors (on the category of groups)

$$G/G^c \approx I_G/I_G^2, \qquad \text{by} \qquad xG^c \mapsto (x - 1) + I_G^2.$$

13. Let $A \in \mathrm{Mod}(G)$ and $\alpha \in H^1(G, A)$. Let $\{a(x)\}_{x \in G}$ be a standard 1-cocycle representing α. Show that there exists a G-homomorphism $f : I_G \to A$ such that $f(x - 1) = a(x)$, so $f \in (\mathrm{Hom}(I_G, A))^G$. Show that the sequence

$$0 \to A = \mathrm{Hom}(\mathbf{Z}, A) \to \mathrm{Hom}(\mathbf{Z}[G], A) \to \mathrm{Hom}(I_G, A) \to 0$$

is exact, and that if δ is the coboundary for the cohomology sequence, then $\delta(f) = -\alpha$.

Finite groups

We now turn to the case of *finite* groups G. For such groups and a G-module A we have the **trace**

$$T_G : A \to A \qquad \text{defined by} \qquad T_G(a) = \sum_{\sigma \in G} \sigma a.$$

We define a module A to be G-**regular** if there exists a $\mathbf{Z}$-endomorphism $u : A \to A$ such that $\mathrm{id}_A = T_G(u)$. Recall that the operation of G on $\mathrm{End}(A)$ is given by

$$[\sigma]f(a) = \sigma f(\sigma^{-1}a) \text{ for } \sigma \in G.$$

14. (a) Show that a projective object in $\mathrm{Mod}(G)$ is G-regular.
    (b) Let R be a commutative ring and let A be in $\mathrm{Mod}_R(G)$ (the category of (G, R)-modules). Show that A is $R[G]$-projective if and only if A is R-projective and $R[G]$-regular, meaning that $\mathrm{id}_A = T_G(u)$ for some R-homomorphism $u : A \to A$.

15. Consider the exact sequences:

(1) $$0 \to I_G \to \mathbf{Z}[G] \xrightarrow{\varepsilon} \mathbf{Z} \to 0$$

(2) $$0 \to \mathbf{Z} \xrightarrow{\varepsilon'} \mathbf{Z}[G] \to J_G \to 0$$

where the first one defines I_G, and the second is defined by the embedding

$$\varepsilon' : \mathbf{Z} \to \mathbf{Z}[G] \text{ such that } \varepsilon'(n) = n\left(\sum \sigma\right),$$

i.e. on the "diagonal". The cokernel of ε' is J_G by definition.

(a) Prove that both sequences (1) and (2) split in $\mathrm{Mod}(G)$.

(b) Define $M'_G(A) = \mathbf{Z}[G] \otimes A$ (tensor product over $\mathbf{Z}$) for $A \in \mathrm{Mod}(G)$. Show that $M'_G(A)$ is G-regular, and that one gets exact sequences (1_A) and (2_A) by tensoring (1) and (2) with A. As a result one gets an embedding

$$\varepsilon'_A = \varepsilon' \otimes \mathrm{id} : A = \mathbf{Z} \otimes A \to \mathbf{Z}[G] \otimes A.$$

16. **Cyclic groups**. Let G be a finite cyclic group of order n. Let σ be a generator of G. Let $K^i = \mathbf{Z}[G]$ for $i > 0$. Let $\varepsilon : K^0 \to \mathbf{Z}$ be the augmentation as before. For i odd $\geqq 1$, let $d^i : K^i \to K^{i-1}$ be multiplication by $1 - \sigma$. For i even $\geqq 2$, let d^i be multiplication by $1 + \sigma + \cdots + \sigma^{n-1}$. Prove that K is a resolution of $\mathbf{Z}$. Conclude that:
For i odd: $H^i(G, A) = A^G/T_G A$ where $T_G : a \mapsto (1 + \sigma + \cdots + \sigma^{n-1})a$;
For i even $\geqq 2$: $H^i(G, A) = A_T/(1 - \sigma)A$, where A_T is the kernel of T_G in A.

17. Let G be a finite group. Show that there exists a δ-functor $\mathbf{H}$ from $\mathrm{Mod}(G)$ to $\mathrm{Mod}(\mathbf{Z})$ such that:
(1) $\mathbf{H}^0$ is (isomorphic to) the functor $A \mapsto A^G/T_G A$.
(2) $\mathbf{H}^q(A) = 0$ if A is injective and $q > 0$, and $\mathbf{H}^q(A) = 0$ if A is projective and q is arbitrary.
(3) $\mathbf{H}$ is erased by G-regular modules. In particular, $\mathbf{H}$ is erased by M_G.
The δ-functor of Exercise 17 is called the **special cohomology functor**. It differs from the other one only in dimension 0.

18. Let $\mathbf{H} = \mathbf{H}_G$ be the special cohomology functor for a finite group G. Show that:

$$\mathbf{H}^0(I_G) = 0; \mathbf{H}^0(Z) \approx \mathbf{H}^1(I) \approx \mathbf{Z}/n\mathbf{Z} \text{ where } n = \#(G);$$

$$\mathbf{H}^0(Q/Z) = \mathbf{H}^1(Z) = \mathbf{H}^2(I) = 0$$

$$\mathbf{H}^1(Q/Z) \approx \mathbf{H}^2(Z) \approx \mathbf{H}^3(I) \approx G^\wedge = \mathrm{Hom}(G, \mathbf{Q}/\mathbf{Z}) \text{ by definition.}$$

Injectives

19. (a) Show that if an abelian group T is injective in the category of abelian groups, then it is divisible.
(b) Let A be a principal entire ring. Define the notion of divisibility by elements of A for modules in a manner analogous to that for abelian groups. Show that an A-module is injective if and only if it is A-divisible. [The proof for $\mathbf{Z}$ should work in exactly the same way.]

20. Let S be a multiplicative subset of the commutative Noetherian ring A. If I is an injective A-module, show that $S^{-1}I$ is an injective $S^{-1}A$-module.

21. (a) Show that a direct sum of projective modules is projective.
(b) Show that a direct product of injective modules is injective.

22. Show that a factor module, direct summand, direct product, and direct sum of divisible modules are divisible.

23. Let Q be a module over a commutative ring A. Assume that for every left ideal J of A, every homomorphism $\varphi : J \to Q$ can be extended to a homomorphism of A into Q. Show that Q is injective. [*Hint*: Given $M' \subset M$ and $f : M' \to Q$, let $x_0 \in M$ and $x_0 \notin M'$. Let J be the left ideal of elements $a \in A$ such that $ax_0 \in M'$. Let $\varphi(a) = f(ax_0)$ and extend φ to A, as can be done by hypothesis. Then show that

one can extend f to M by the formula

$$f(x' + bx_0) = f(x') + \varphi(b),$$

for $x' \in M$ and $b \in A$. Then use Zorn's lemma. This is the same pattern of proof as the proof of Lemma 4.2.]

24. Let

$$0 \to I_1 \to I_2 \to I_3 \to 0$$

be an exact sequence of modules. Assume that I_1, I_2 are injective.

(a) Show that the sequence splits.

(b) Show that I_3 is injective.

(c) If I is injective and $I = M \oplus N$, show that M is injective.

25. (Do this exercise after you have read about Noetherian rings.) Let A be a Noetherian commutative ring, and let Q be an injective A-module. Let $\mathfrak{a}$ be an ideal of A, and let $Q^{(\mathfrak{a})}$ be the subset of elements $x \in Q$ such that $\mathfrak{a}^n x = 0$ for some n, depending on x. Show that $Q^{(\mathfrak{a})}$ is injective. [*Hint*: Use Exercise 23.]

26. Let A be a commutative ring. Let E be an A-module, and let $E^\wedge = \mathrm{Hom}_{\mathbf{Z}}(E, \mathbf{Q}/\mathbf{Z})$ be the dual module. Prove the following statements.

(a) A sequence

$$0 \to N \to M \to E \to 0$$

is exact if and only if the dual sequence

$$0 \to E^\wedge \to M^\wedge \to N^\wedge \to 0$$

is exact.

(b) Let F be flat and I injective in the category of A-modules. Show that $\mathrm{Hom}_A(F, I)$ is injective.

(c) E is flat if and only if $E^\wedge$ is injective.

27. **Extensions of modules**. Let M, N be modules over a ring. By an **extension** of M by N we mean an exact sequence

$$(*) \qquad 0 \to N \to E \to M \to 0.$$

We shall now define a map from such extensions to $\mathrm{Ext}^1(M, N)$. Let P be projective, with a surjective homomorphism onto M, so we get an exact sequence

$$(**) \qquad 0 \to K \xrightarrow{w} P \xrightarrow{p} M \to 0$$

where K is defined to be the kernel. Since P is projective, there exists a homomorphism $u: P \to E$, and depending on u a unique homomorphism $v: K \to N$ making the diagram commutative:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K & \longrightarrow & P & \longrightarrow & M & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle u} & & \downarrow{\scriptstyle \mathrm{id}} & & \\ 0 & \longrightarrow & N & \longrightarrow & E & \longrightarrow & M & \longrightarrow & 0 \end{array}$$

On the other hand, we have the exact sequence

$$(***) \quad 0 \to \mathrm{Hom}(M, N) \to \mathrm{Hom}(P, N) \to \mathrm{Hom}(K, N) \to \mathrm{Ext}^1(M, N) \to 0,$$

with the last term on the right being equal to 0 because $\mathrm{Ext}^1(P, N) = 0$. To the extension (*) we associate the image of v in $\mathrm{Ext}^1(M, N)$.

Prove that this association is a bijection between isomorphism classes of extensions (i.e. isomorphism classes of exact sequences as in (*)), and $\mathrm{Ext}^1(M, N)$. [*Hint*: Construct an inverse as follows. Given an element e of $\mathrm{Ext}^1(M, N)$, using an exact sequence (**), there is some element $v \in \mathrm{Hom}(K, N)$ which maps on e in (***). Let E be the push-out of v and w. In other words, let J be the submodule of $N \oplus P$ consisting of all elements $(v(x), -w(x))$ with $x \in K$, and let $E = (N \oplus P)/J$. Show that the map $y \mapsto (y, 0) \bmod J$ gives an injection of N into E. Show that the map $N \oplus P \to M$ vanishes on J, and so gives a surjective homomorphism $E \to M \to 0$. Thus we obtain an exact sequence (*); that is, an extension of M by N. Thus to each element of $\mathrm{Ext}^1(M, N)$ we have associated an isomorphism class of extensions of M by N. Show that the maps we have defined are inverse to each other between isomorphism classes of extensions and elements of $\mathrm{Ext}^1(M, N)$.]

28. Let R be a principal entire ring. Let $a \in R$. For every R-module N, prove:
    (a) $\mathrm{Ext}^1(R/aR, N) = N/aN$.
    (b) For $b \in R$ we have $\mathrm{Ext}^1(R/aR, R/bR) = R/(a, b)$, where (a, b) is the g.c.d of a and b, assuming $ab \neq 0$.

Tensor product of complexes.

29. Let $K = \bigoplus K_p$ and $L = \bigoplus L_q$ be two complexes indexed by the integers, and with boundary maps lower indices by 1. Define $K \otimes L$ to be the direct sum of the modules $(K \otimes L)_n$, where

$$(K \otimes L)_n = \bigoplus_{p+q=n} K_p \otimes L_q.$$

Show that there exist unique homomorphisms

$$d = d_n : (K \otimes L)_n \to (K \otimes L)_{n-1}$$

such that

$$d(x \otimes y) = d(x) \otimes y + (-1)^p x \otimes d(y).$$

Show that $K \otimes L$ with these homomorphisms is a complex, that is $d \circ d = 0$.

30. Let K, L be double complexes. We write $K_{i\cdot}$ and $L_{i\cdot}$ for the ordinary column complexes of K and L respectively. Let $\varphi: K \to L$ be a homomorphism of double complexes. Assume that each homomorphism

$$\varphi_i : K_{i\cdot} \to L_{i\cdot}$$

is a homology isomorphism.
    (a) Prove that $\mathrm{Tot}(\varphi): \mathrm{Tot}(K) \to \mathrm{Tot}(L)$ is a homology isomorphism. (If you want to see this worked out, cf. [FuL 85], Chapter V, Lemma 5.4.)
    (b) Prove Theorem 9.8 using (a) instead of spectral sequences.

Bibliography

[ArT 68] E. ARTIN and J. TATE, *Class Field Theory*, Benjamin, 1968; Addison-Wesley, 1991

[At 61] M. ATIYAH, Characters and cohomology of finite groups, *Pub. IHES* **9** (1961), pp. 5–26

[At 67] M. ATIYAH, *K-theory*, Benjamin, 1967; reprinted Addison-Wesley, 1991

[ABP 73] M. ATIYAH, R. BOTT, and R. PATODI, On the heat equation and the index theorem, *Invent. Math.* **19** (1973), pp. 279–330

[Ba 68] H. BASS, *Algebraic K-theory*, Benjamin, 1968

[Bo 69] R. BOTT, *Lectures on K(X)*, Benjamin, 1969

[BtD 85] T. BROCKER and T. TOM DIECK, *Representations of Compact Lie Groups*, Springer Verlag, 1985

[CaE 57] H. CARTAN and S. EILENBERG, *Homological Algebra*, Princeton University Press, 1957

[CuR 81] C. CURTIS and I. REINER, *Methods of Representation Theory*, John Wiley & Sons, 1981

[ES 52] S. EILENBERG and N. STEENROD, *Foundations of Algebraic Topology*, Princeton University Press, 1952

[FuL 85] W. FULTON and S. LANG, *Riemann-Roch algebra*, Springer Verlag, 1985

[Go 58] R. GODEMENT, *Théorie des faisceaux*, Hermann Paris, 1958

[GreH 81] M. GREENBERG and J. HARPER, *Algebraic Topology: A First Course*, Benjamin-Addison-Wesley, 1981

[GriH 78] P. GRIFFITHS and J. HARRIS, *Principles of algebraic geometry*, Wiley Interscience 1978

[Gro 57] A. GROTHENDIECK, Sur quelques points d'algèbre homologique, *Tohoku Math. J.* **9** (1957) pp. 119–221

[Gro 68] A. GROTHENDIECK, *Classes de Chern et représentations linéaires des groupes discrets*, Dix exposés sur la cohomologie étale des schémas, North-Holland, Amsterdam, 1968

[Gu 91] R. GUNNING, *Introduction to holomorphic functions of several variables*, Vol. III Wadsworth & Brooks/Cole, 1990

[Ha 77] R. HARTSHORNE, *Algebraic Geometry*, Springer Verlag, 1977

[HiS 70] P. J. HILTON and U. STAMMBACH, *A Course in Homological Algebra*, Graduate Texts in Mathematics, Springer Verlag, 1970.

[La 96] S. LANG, *Topics in cohomology of groups*, Springer Lecture Notes, 1996

[Man 69] J. MANIN, *Lectures on the K-functor in Algebraic Geometry, Russian Math Surveys* **24**(5) (1969) pp. 1–89

[Mat 70] H. MATSUMURA, *Commutative Algebra*, Second Edition, Benjamin-Cummings, 1981

[No 68] D. NORTHCOTT, *Lessons on Rings, Modules and Multiplicities*, Cambridge University Press, 1968

[No 76] D. NORTHCOTT, *Finite Free Resolutions*, Cambridge University Press, 1976

[Ro 79] J. ROTMAN, *Introduction to Homological Algebra*, Academic Press, 1979

[Se 64] J.-P. SERRE, *Cohomologie Galoisienne*, Springer Lecture Notes **5**, 1964

[Se 65] J.-P. SERRE, *Algèbre locale, multiplicités*, Springer Lecture Notes **11** (1965) Third Edition 1975

[SGA 6] P. BERTHELOT, A. GROTHENDIECK, L. ILLUSIE et al. *Théorie des intersections et théorème de Riemann-Roch*, Springer Lecture Notes 146, 1970

[Sh 72] S. SHATZ, *Profinite groups, arithmetic and geometry*, Ann. of Math Studies, Princeton University Press 1972

CHAPTER XXI

Finite Free Resolutions

This chapter puts together specific computations of complexes and homology. Partly these provide examples for the general theory of Chapter XX, and partly they provide concrete results which have occupied algebraists for a century. They have one aspect in common: the computation of homology is done by means of a finite free resolution, i.e. a finite complex whose modules are finite free.

The first section shows a general technique (the mapping cylinder) whereby the homology arising from some complex can be computed by using another complex which is finite free. One application of such complexes has already been given in Chapter X, putting together Proposition 4.5 followed by Exercises 10–15 of that chapter.

Then we go to major theorems, going from Hilbert's Syzygy theorem, from a century ago, to Serre's theorem about finite free resolutions of modules over polynomial rings, and the Quillen-Suslin theorem. We also include a discussion of certain finite free resolutions obtained from the Koszul complex. These apply, among other things, to the Grothendieck Riemann-Roch theorem of algebraic geometry.

Bibliographical references refer to the list given at the end of Chapter XX.

§1. SPECIAL COMPLEXES

As in the preceding chapter, we work with the category of modules over a ring, but the reader will notice that the arguments hold quite generally in an abelian category.

In some applications one determines homology from a complex which is not suitable for other types of construction, like changing the base ring. In this section, we give a general procedure which constructs another complex with

better properties than the first one, while giving the same homology. For an application to Noetherian modules, see Exercises 12–15 of Chapter X.

Let $f: K \to C$ be a morphism of complexes. We say that f is a **homology isomorphism** if the natural map

$$H(f): H(K) \to H(C)$$

is an isomorphism. The definition is valid in an abelian category, but the reader may think of modules over a ring, or abelian groups even. A family $\mathfrak{F}$ of objects will be called **sufficient** if given an object E there exists an element F in $\mathfrak{F}$ and an epimorphism

$$F \to E \to 0,$$

and if $\mathfrak{F}$ is closed under taking finite direct sums. For instance, we may use for $\mathfrak{F}$ the family of free modules. However, in important applications, we shall deal with finitely generated modules, in which case $\mathfrak{F}$ might be taken as the family of finite free modules. These are in fact the applications I have in mind, which resulted in having axiomatized the situation.

Proposition 1.1. *Let C be a complex such that $H^p(C) \neq 0$ only for $0 \leqq p \leqq n$. Let $\mathfrak{F}$ be a sufficient family of projectives. There exists a complex*

$$0 \to K^0 \to K^1 \to \cdots \to K^n \to 0$$

such that:

$$K^p \neq 0 \quad \text{only for} \quad 0 \leqq p \leqq n;$$

$$K^p \text{ is in } \mathfrak{F} \text{ for all } p \geqq 1;$$

and there exists a homomorphism of complexes

$$f: K \to C$$

which is a homology isomorphism.

Proof. We define f_m by descending induction on m:

$$\begin{array}{ccccccc}
\longrightarrow & K^m & \longrightarrow & K^{m+1} & \xrightarrow{\delta_K^{m+1}} & K^{m+2} & \longrightarrow \\
 & \downarrow{\scriptstyle f_m} & & \downarrow{\scriptstyle f_{m+1}} & & \downarrow{\scriptstyle f_{m+2}} & \\
\longrightarrow & C^m & \longrightarrow & C^{m+1} & \xrightarrow[\delta_C^{m+1}]{} & C^{m+2} & \longrightarrow
\end{array}$$

We suppose that we have defined a morphism of complexes with $p \geqq m+1$ such that $H^p(f)$ is an isomorphism for $p \geqq m+2$, and

$$f_{m+1}: Z^{m+1}(K) \to H^{m+1}(C)$$

is an epimorphism, where Z denotes the cycles, that is Ker δ. We wish to construct K^m and f_m, thus propagating to the left. First let $m \geqq 0$. Let B^{m+1} be the kernel of

$$\operatorname{Ker} \delta_K^{m+1} \to H^{m+1}(C).$$

Let K' be in $\mathfrak{F}$ with an epimorphism

$$\delta' : K' \to B^{m+1}.$$

Let $K'' \to H^m(C)$ be an epimorphism with K'' in $\mathfrak{F}$, and let

$$f'' : K'' \to Z^m(C)$$

be any lifting, which exists since K'' is projective. Let

$$K^m = K' \oplus K''$$

and define $\delta^m : K^m \to K^{m+1}$ to be δ' on K' and 0 on K''. Then

$$f_{m+1} \circ \delta'(K') \subset \delta_C(C_m),$$

and hence there exists $f' : K' \to C^m$ such that

$$\delta_C \circ f' = f_{m+1} \circ \delta'.$$

We now define $f_m : K^m \to C^m$ to be f' on K' and f'' on K''. Then we have defined a morphism of complexes truncated down to m as desired.

Finally, if $m = -1$, we have constructed down to K^0, δ^0, and f_0 with

$$K^0 \xrightarrow{f_0} H^0(C) \to 0$$

exact. The last square looks like this, defining $K^{-1} = 0$.

$$\begin{array}{ccccccc} 0 & \longrightarrow & K' \oplus K'' & \xrightarrow{\delta_0 = \delta'} & \delta' K' \subset K^1 & & \\ & & \searrow^{f'} \quad \swarrow_{f''} & & & \searrow^{f_1} & \\ 0 & \longrightarrow & C^0 & \longrightarrow & & & C^1 \end{array}$$

We replace K^0 by $K^0/(\operatorname{Ker} \delta^0 \cap \operatorname{Ker} f_0)$. Then $H^0(f)$ becomes an isomorphism, thus proving the proposition.

We want to say something more about K^0. For this purpose, we define a new concept. Let $\mathfrak{F}$ be a family of objects in the given abelian category (think of modules in first reading). We shall say that $\mathfrak{F}$ is **complete** if it is sufficient, and for any exact sequence

$$0 \to F' \to F \to F'' \to 0$$

with F'' and F in $\mathfrak{F}$ then F' is also in $\mathfrak{F}$.

Example. In Chapter XVI, Theorem 3.4 we proved that the family of finite flat modules in the category of finite modules over a Noetherian ring is complete. Similarly, the family of flat modules in the category of modules over a ring is complete. We cannot get away with just projectives or free modules, because in the statement of the proposition, K^0 is not necessarily free but we want to include it in the family as having especially nice properties. In practice, the family consists of the flat modules, or finite flat modules. Cf. Chaper X, Theorem 4.4, and Chapter XVI, Theorem 3.8.

Proposition 1.2. *Let $f: K \to C$ be a morphism of complexes, such that K^p, $H^p(C)$ are $\neq 0$ only for $p = 1, \ldots, n$. Let $\mathfrak{F}$ be a complete family, and assume that K^p, C^p are in $\mathfrak{F}$ for all p, except possibly for K^0. If f is a homology isomorphism, then K^0 is also in $\mathfrak{F}$.*

Before giving the proof, we define a new complex called the **mapping cylinder** of an arbitrary morphism of complexes f by letting

$$M^p = K^p \oplus C^{p-1}$$

and defining $\delta_M : M^p \to M^{p+1}$ by

$$\delta_M(x, y) = (\delta x, fx - \delta y).$$

It is trivially verified that M is then a complex, i.e. $\delta \circ \delta = 0$. If C' is the complex obtained from C by shifting degrees by one (and making a sign change in δ_C), so $C'^p = C^{p-1}$, then we get an exact sequence of complexes

$$0 \to C' \to M \to K \to 0$$

and hence the **mapping cylinder exact cohomology sequence**

$$\begin{array}{ccccccccc} H^p(K) & \longrightarrow & H^{p+1}(C') & \longrightarrow & H^{p+1}(M) & \longrightarrow & H^{p+1}(K) & \longrightarrow & H^{p+2}(C') \\ & & \| & & & & & & \| \\ & & H^p(C) & & & & & & H^{p+1}(C) \end{array}$$

and one sees from the definitions that the cohomology maps

$$H^p(K) \to H^{p+1}(C') \approx H^p(C)$$

are the ones induced by $f: K \to C$.

We now return to the assumptions of Proposition 1.2, so that these maps are isomorphisms. We conclude that $H(M) = 0$. This implies that the sequence

$$0 \to K^0 \to M^1 \to M^2 \to \cdots \to M^{n+1} \to 0$$

is exact. Now each M^p is in $\mathfrak{F}$ by assumption. Inserting the kernels and cokernels at each step and using induction together with the definition of a complete family, we conclude that K^0 is in $\mathfrak{F}$, as was to be shown.

In the next proposition, we have axiomatized the situation so that it is applicable to the tensor product, discussed later, and to the case when the family $\mathfrak{F}$ consists of flat modules, as defined in Chapter XVI. No knowledge of this chapter is needed here, however, since the axiomatization uses just the general language of functors and exactness.

Let $\mathfrak{F}$ be a complete family again, and let T be a covariant additive functor on the given category. We say that $\mathfrak{F}$ is **exact for T** if given an exact sequence

$$0 \to F' \to F \to F'' \to 0$$

in $\mathfrak{F}$, then

$$0 \to T(F') \to T(F) \to T(F'') \to 0$$

is exact.

Proposition 1.3. *Let $\mathfrak{F}$ be a complete family which is exact for T. Let $f : K \to C$ be a morphism of complexes, such that K^p and C^p are in $\mathfrak{F}$ for all p, and K^p, $H^p(C)$ are zero for all but a finite number of p. Assume that f is a homology isomorphism. Then*

$$T(f) : T(K) \to T(C)$$

is a homology isomorphism.

Proof. Construct the mapping cylinder M for f. As in the proof of Proposition 1.2, we get $H(M) = 0$ so M is exact. We then start inductively from the right with zeros. We let Z^p be the cycles in M^p and use the short exact sequences

$$0 \to Z^p \to M^p \to Z^{p+1} \to 0$$

together with the definition of a complete family to conclude that Z^p is in $\mathfrak{F}$ for all p. Hence the short sequences obtained by applying T are exact. But $T(M)$ is the mapping cylinder of the morphism

$$T(f) : T(K) \to T(C),$$

which is therefore an isomorphism, as one sees from the homology sequence of the mapping cylinder. This concludes the proof.

§2. FINITE FREE RESOLUTIONS

The first part of this section develops the notion of resolutions for a case somewhat more subtle than projective resolutions, and gives a good example for the considerations of Chapter XX. Northcott in [No 76] pointed out that minor adjustments of standard proofs also applied to the non-Noetherian rings, only occasionally slightly less tractable than the Noetherian ones.

Let A be a ring. A module E is called **stably free** if there exists a finite free module F such that $E \oplus F$ is finite free, and thus isomorphic to $A^{(n)}$ for some positive integer n. In particular, E is projective and finitely generated.

We say that a module M has a **finite free resolution** if there exists a resolution

$$0 \to E_n \to \cdots \to E_0 \to M \to 0$$

such that each E_i is finite free.

Theorem 2.1. *Let M be a projective module. Then M is stably free if and only if M admits a finite free resolution.*

Proof. If M is stably free then it is trivial that M has a finite free resolution. Conversely assume the existence of the resolution with the above notation. We prove that M is stably free by induction on n. The assertion is obvious if $n = 0$. Assume $n \geqq 1$. Insert the kernels and cokernels at each step, in the manner of dimension shifting. Say

$$M_1 = \operatorname{Ker}(E_0 \to P),$$

giving rise to the exact sequence

$$0 \to M_1 \to E_0 \to M \to 0.$$

Since M is projective, this sequence splits, and $E_0 \approx M \oplus M_1$. But M_1 has a finite free resolution of length smaller than the resolution of M, so there exists a finite free module F such that $M_1 \oplus F$ is free. Since $E_0 \oplus F$ is also free, this concludes the proof of the theorem.

A resolution

$$0 \to E_n \to \cdots \to E_0 \to M \to 0$$

is called **stably free** if all the modules E_i $(i = 0, \ldots, n)$ are stably free.

Proposition 2.2. *Let M be an A-module. Then M has a finite free resolution of length $n \geqq 1$ if and only if M has a stably free resolution of length n.*

Proof. One direction is trivial, so we suppose given a stably free resolution with the above notation. Let $0 \leqq i < n$ be some integer, and let F_i, F_{i+1} be finite free such that $E_i \oplus F_i$ and $E_{i+1} \oplus F_{i+1}$ are free. Let $F = F_i \oplus F_{i+1}$. Then we can form an exact sequence

$$0 \to E_n \to \cdots \to E_{i+1} \oplus F \to E_i \oplus F \to \cdots \to E_0 \to M \to 0$$

in the obvious manner. In this way, we have changed two consecutive modules in the resolution to make them free. Proceeding by induction, we can then make E_0, E_1 free, then E_1, E_2 free, and so on to conclude the proof of the proposition.

The next lemma is designed to facilitate dimension shifting.

We say that two modules M_1, M_2 are **stably isomorphic** if there exist finite free modules F_1, F_2 such that $M_1 \oplus F_1 \approx M_2 \oplus F_2$.

Lemma 2.3. *Let M_1 be stably isomorphic to M_2. Let*

$$0 \to N_1 \to E_1 \to M_1 \to 0$$

$$0 \to N_2 \to E_2 \to M_2 \to 0$$

be exact sequences, where M_1 is stably isomorphic to M_2, and E_1, E_2 are stably free. Then N_1 is stably isomorphic to N_2.

Proof. By definition, there is an isomorphism $M_1 \oplus F_1 \approx M_2 \oplus F_2$. We have exact sequences

$$0 \to N_1 \to E_1 \oplus F_1 \to M_1 \oplus F_1 \to 0$$

$$0 \to N_2 \to E_2 \oplus F_2 \to M_2 \oplus F_2 \to 0$$

By Schanuel's lemma (see below) we conclude that

$$N_1 \oplus E_2 \oplus F_2 \approx N_2 \oplus E_1 \oplus F_1.$$

Since E_1, E_2, F_1, F_2 are stably free, we can add finite free modules to each side so that the summands of N_1 and N_2 become free, and by adding 1-dimensional free modules if necessary, we can preserve the isomorphism, which proves that N_1 is stably isomorphic to N_2.

We still have to take care of **Schanuel's lemma**:

Lemma 2.4. *Let*

$$0 \to K \to P \to M \to 0$$

$$0 \to K' \to P' \to M \to 0$$

be exact sequences where P, P' are projective. Then there is an isomorphism

$$K \oplus P' \approx K' \oplus P.$$

Proof. Since P is projective, there exists a homomorphism $P \to P'$ making the right square in the following diagram commute.

$$\begin{array}{ccccccccc} 0 & \longrightarrow & K & \xrightarrow{i} & P & \longrightarrow & M & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle u} & & \downarrow{\scriptstyle w} & & \downarrow{\scriptstyle \mathrm{id}} & & \\ 0 & \longrightarrow & K' & \xrightarrow[j]{} & P' & \longrightarrow & M & \longrightarrow & 0 \end{array}$$

Then one can find a homomorphism $K \to K'$ which makes the left square commute. Then we get an exact sequence

$$0 \to K \to P \oplus K' \to P' \to 0$$

by $x \mapsto (ix, ux)$ for $x \in K$ and $(y, z) \mapsto wy - jz$. We leave the verification of exactness to the reader. Since P' is projective, the sequence splits thus proving Schanuel's lemma. This also concludes the proof of Lemma 2.3.

The minimal length of a stably free resolution of a module is called its **stably free dimension**. To construct a stably free resolution of a finite module, we proceed inductively. The preceding lemmas allow us to carry out the induction, and also to stop the construction if a module is of finite stably free dimension.

Theorem 2.5. *Let M be a module which admits a stably free resolution of length n*

$$0 \to E_n \to \cdots \to E_0 \to M \to 0.$$

Let

$$F_m \to \cdots \to F_0 \to M \to 0$$

be an exact sequence with F_i stably free for $i = 0, \ldots, m$.

(i) *If $m < n - 1$ then there exists a stably free F_{m+1} such that the exact sequence can be continued exactly to*

$$F_{m+1} \to \cdots \to F_0 \to M \to 0.$$

(ii) *If $m = n - 1$, let $F_n = \operatorname{Ker}(F_{n-1} \to F_{n-2})$. Then F_n is stably free and thus*

$$0 \to F_n \to F_{n-1} \to \cdots \to F_0 \to M \to 0$$

is a stably free resolution.

Remark. If A is Noetherian then of course (i) is trivial, and we can even pick F_{m+1} to be finite free.

Proof. Insert the kernels and cokernels in each sequence, say

$$K_m = \operatorname{Ker}(E_m \to E_{m-1}) \quad \text{if} \quad m \neq 0$$
$$K_0 = \operatorname{Ker}(E_0 \to M),$$

and define K'_m similarly. By Lemma 2.3, K_m is stably isomorphic to K'_m, say

$$K_m \oplus F \approx K'_m \oplus F'$$

with F, F' finite free.

If $m < n - 1$, then K_m is a homomorphic image of E_{m+1}; so both $K_m \oplus F$ and $K'_m \oplus F'$ are homomorphic images of $E_{m+1} \oplus F$. Therefore K'_m is a homomorphic image of $E_{m+1} \oplus F$ which is stably free. We let $F_{m+1} = E_{m+1} \oplus F$ to conclude the proof in this case.

If $m = n - 1$, then we can take $K_n = E_n$. Hence $K_m \oplus F$ is stably free, and so is $K'_m \oplus F'$ by the isomorphism in the first part of the proof. It follows trivially that K'_m is stably free, and by definition, $K'_m = F_{m+1}$ in this case. This concludes the proof of the theorem.

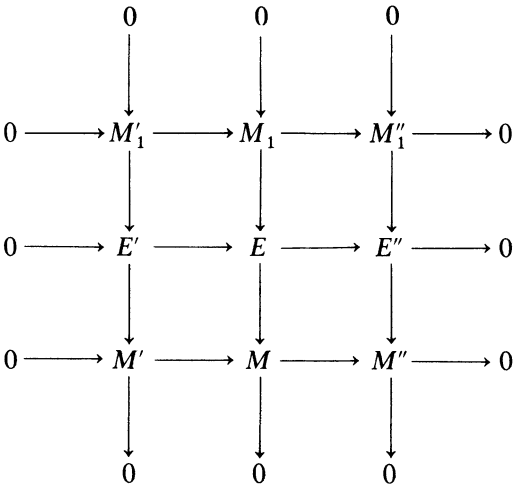
Corollary 2.6. *If* $0 \to M_1 \to E \to M \to 0$ *is exact, M has stably free dimension* $\leqq n$, *and E is stably free, then* M_1 *has stably free dimension* $\leqq n - 1$.

Theorem 2.7. *Let*

$$0 \to M' \to M \to M'' \to 0$$

be an exact sequence. If any two of these modules have a finite free resolution, then so does the third.

Proof. Assume M' and M have finite free resolutions. Since M is finite, it follows that M'' is also finite. By essentially the same construction as Chapter XX, Lemma 3.8, we can construct an exact and commutative diagram where E', E, E'' are stably free:



We then argue by induction on the stably free dimension of M. We see that M_1 has stably free dimension $\leqq n - 1$ (actually $n - 1$, but we don't care), and M'_1 has finite stably free dimension. By induction we are reduced to the case when M has stably free dimension 0, which means that M is stably free. Since by assumption there is a finite free resolution of M', it follows that M'' also has a finite free resolution, thus concluding the proof of the first assertion.

Next assume that M', M'' have finite free resolutions. Then M is finite. If both M' and M'' have stably free dimension 0, then M', M'' are projective and $M \approx M' \oplus M''$ is also stably free and we are done. We now argue by induction on the maximum of their stably free dimension n, and we assume $n \geqq 1$. We can construct an exact and commutative diagram as in the previous case with E', E, E'' finite free (we leave the details to the reader). But the maximum of the stably free dimensions of M'_1 and M''_1 is at most $n - 1$, and so by induction it follows that M_1 has finite stably free dimension. This concludes the proof of the second case.

Observe that the third statement has been proved in Chapter XX, Lemma 3.8 when A is Noetherian, taking for $\mathfrak{A}$ the abelian category of finite modules, and for $\mathscr{C}$ the family of stably free modules. Mitchell Stokes pointed out to me that the statement is valid in general without Noetherian assumption, and can be proved as follows. We assume that M, M'' have finite free resolutions. We first show that M' is finitely generated. Indeed, suppose first that M is finite free. We have two exact sequences

$$0 \to M' \to M \to M'' \to 0$$
$$0 \to K'' \to F'' \to M'' \to 0$$

where F'' is finite free, and K'' is finitely generated because of the assumption that M'' has a finite free resolution. That M' is finitely generated follows from Schanuel's lemma. If M is not free, one can reduce the finite generation of M' to the case when M is free by a pull-back, which we leave to the reader.

Now suppose that the stably free dimension of M'' is positive. We use the same exact commutative diagram as in the previous cases, with E', E, E'' finite free. The stably free dimension of M''_1 is one less than that of M'', and we are done by induction. This concludes the proof of Theorem 2.7.

This also concludes our general discussion of finite free resolutions. For more information cf. Northcott's book on the subject.

We now come to the second part of this section, which provides an application to polynomial rings.

Theorem 2.8. *Let R be a commutative Noetherian ring. Let x be a variable. If every finite R-module has a finite free resolution, then every finite $R[x]$-module has a finite free resolution.*

In other words, in the category of finite R-modules, if every object is of finite stably free dimension, then the same property applies to the category of finite $R[x]$-modules. Before proving the theorem, we state the application we have in mind.

Theorem 2.9. (Serre). *If k is a field and $x_1, \ldots, x_r$ independent variables, then every finite projective module over $k[x_1, \ldots, x_r]$ is stably free, or equivalently admits a finite free resolution.*

Proof. By induction and Theorem 2.8 we conclude that every finite module over $k[x_1, \ldots, x_r]$ is of finite stably free dimension. (We are using Theorem 2.1.) This concludes the proof.

The rest of this section is devoted to the proof of Theorem 2.8.

Let M be a finite $R[x]$-module. By Chapter X, Corollary 2.8, M has a finite filtration

$$M = M_0 \supset M_1 \supset \cdots \supset M_n = 0$$

such that each factor M_i/M_{i+1} is isomorphic to $R[x]/P_i$ for some prime P_i. In light of Theorem 2.7, it suffices to prove the theorem in case $M = R[x]/P$ where P is prime, which we now assume. In light of the exact sequence

$$0 \to P \to R[x] \to R[x]/P \to 0.$$

and Theorem 2.7, we note that M has a finite free resolution if and only if P does.

Let $\mathfrak{p} = P \cap R$. Then $\mathfrak{p}$ is prime in R. Suppose there is some $M = R[x]/P$ which does not admit a finite free resolution. Among all such M we select one for which the intersection $\mathfrak{p}$ is maximal in the family of prime ideals obtained as above. This is possible in light of one of the basic properties characterizing Noetherian rings.

Let $R_0 = R/\mathfrak{p}$ so R_0 is entire. Let $P_0 = P/\mathfrak{p}R[x]$. Then we may view M as an $R_0[x]$-module, equal to R_0/P_0. Let $f_1, \ldots, f_n$ be a finite set of generators for P_0, and let f be a polynomial of minimal degree in P_0. Let K_0 be the quotient field of R_0. By the euclidean algorithm, we can write

$$f_i = q_i f + r_i \quad \text{for} \quad i = 1, \ldots, n$$

with $q_i, r_i \in K_0[x]$ and $\deg r_i < \deg f$. Let d_0 be a common denominator for the coefficients of all q_i, r_i. Then $d_0 \neq 0$ and

$$d_0 f_i = q_i' f + r_i'$$

where $q_i' = d_0 q_i$ and $r_i' = d_0 r_i$ lie in $R_0[x]$. Since $\deg f$ is minimal in P_0 it follows that $r_i' = 0$ for all i, so

$$d_0 P_0 \subset R_0[x] f = (f).$$

Let $N_0 = P_0/(f)$, so N_0 is a module over $R_0[x]$, and we can also view N_0 as a module over $R[x]$. When so viewed, we denote N_0 by N. Let $d \in R$ be any element reducing to d_0 mod $\mathfrak{p}$. Then $d \notin \mathfrak{p}$ since $d_0 \neq 0$. The module N_0 has a finite filtration such that each factor module of the filtration is isomorphic to some $R_0[x]/Q_0$ where Q_0 is an associated prime of N_0. Let Q be the inverse image of Q_0 in $R[x]$. These prime ideals Q are precisely the associated primes of N in $R[x]$. Since d_0 kills N_0 it follows that d kills N and therefore d lies in every associated prime of N. By the maximality property in the selection of P,

it follows that every one of the factor modules in the filtration of N has a finite free resolution, and by Theorem 2.7 it follows that N itself has a finite free resolution.

Now we view $R_0[x]$ as an $R[x]$-module, via the canonical homomorphism

$$R[x] \to R_0[x] = R[x]/\mathfrak{p}R[x].$$

By assumption, $\mathfrak{p}$ has a finite free resolution as R-module, say

$$0 \to E_n \to \cdots \to E_0 \to \mathfrak{p} \to 0.$$

Then we may simply form the modules $E_i[x]$ in the obvious sense to obtain a finite free resolution of $\mathfrak{p}[x] = \mathfrak{p}R[x]$. From the exact sequence

$$0 \to \mathfrak{p}R[x] \to R[x] \to R_0[x] \to 0$$

we conclude that $R_0[x]$ has a finite free resolution as $R[x]$-module.

Since R_0 is entire, it follows that the principal ideal (f) in $R_0[x]$ is $R[x]$-isomorphic to $R_0[x]$, and therefore has a finite free resolution as $R[x]$-module. Theorem 2.7 applied to the exact sequence of $R[x]$-modules

$$0 \to (f) \to P_0 \to N \to 0$$

shows that P_0 has a finite free resolution; and further applied to the exact sequence

$$0 \to \mathfrak{p}R[x] \to P \to P_0 \to 0$$

shows that P has a finite free resolution, thereby concluding the proof of Theorem 2.8.

§3. UNIMODULAR POLYNOMIAL VECTORS

Let A be a commutative ring. Let $(f_1, \ldots, f_n)$ be elements of A generating the unit ideal. We call such elements **unimodular**. We shall say that they have the **unimodular extension property** if there exists a matrix in $GL_n(A)$ with first column ${}^t(f_1, \ldots, f_n)$. If A is a principal entire ring, then it is a trivial exercise to prove that this is always the case. Serre originally asked the question whether it is true for a polynomial ring $k[x_1, \ldots, x_r]$ over a field k. The problem was solved by Quillen and Suslin. We give here a simplification of Suslin's proof by Vaserstein, also using a previous result of Horrocks. The method is by induction on the number of variables, in some fashion.

We shall write $f = {}^t(f_1, \ldots, f_n)$ for the column vector. We first remark that f has the unimodular extension property if and only if the vector obtained by a permutation of its components has this property. Similarly, we can make

the usual row operations, adding a multiple gf_i to f_j $(j \neq i)$, and f has the unimodular extension property if and only if any one of its transforms by row operations has the unimodular extension property.

We first prove the theorem in a context which allows the induction.

Theorem 3.1. (Horrocks). *Let* $(\mathfrak{o}, \mathfrak{m})$ *be a local ring and let* $A = \mathfrak{o}[x]$ *be the polynomial ring in one variable over* $\mathfrak{o}$. *Let* f *be a unimodular vector in* $A^{(n)}$ *such that some component has leading coefficient* 1. *Then* f *has the unimodular extension property.*

Proof. (Suslin). If $n = 1$ or 2 then the theorem is obvious even without assuming that $\mathfrak{o}$ is local. So we assume $n \geqq 3$ and do an induction of the smallest degree d of a component of f with leading coefficient 1. First we note that by the Euclidean algorithm and row operations, we may assume that f_1 has leading coefficient 1, degree d, and that $\deg f_i < d$ for $j \neq 1$. Since f is unimodular, a relation $\sum g_i f_i = 1$ shows that not all coefficients of $f_2, \ldots, f_n$ can lie in the maximal ideal $\mathfrak{m}$. Without loss of generality, we may assume that some coefficient of f_2 does not lie in $\mathfrak{m}$ and so is a unit since $\mathfrak{o}$ is local. Write

$$f_1(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_0 \quad \text{with} \quad a_i \in \mathfrak{o},$$
$$f_2(x) = \qquad b_s x^s + \cdots + b_0 \quad \text{with} \quad b_i \in \mathfrak{o},\ s \leqq d - 1,$$

so that some b_i is a unit. Let $\mathfrak{a}$ be the ideal generated by all leading coefficients of polynomials $g_1 f_1 + g_2 f_2$ of degree $\leqq d - 1$. Then $\mathfrak{a}$ contains all the coefficients b_i, $i = 0, \ldots, s$. One sees this by descending induction, starting with b_s which is obvious, and then using a linear combination

$$x^{d-s} f_2(x) - b_s f_1(x).$$

Therefore $\mathfrak{a}$ is the unit ideal, and there exists a polynomial $g_1 f_1 + g_2 f_2$ of degree $\leqq d - 1$ and leading coefficient 1. By row operations, we may now get a polynomial of degree $\leqq d - 1$ and leading coefficient 1 as some component in the i-th place for some $i \neq 1, 2$. Thus ultimately, by induction, we may assume that $d = 0$ in which case the theorem is obvious. This concludes the proof.

Over any commutative ring A, for two column vectors f, g we write $f \sim g$ over A to mean that there exists $M \in GL_n(A)$ such that

$$f = Mg,$$

and we say that f is **equivalent to** g **over** A. Horrocks' theorem states that a unimodular vector f with one component having leading coefficient 1 is $\mathfrak{o}[x]$-equivalent to the first unit vector e^1. We are interested in getting a similar descent over non-local rings. We can write $f = f(x)$, and there is a natural "constant" vector $f(0)$ formed with the constant coefficients. As a corollary of Horrocks' theorem, we get:

Corollary 3.2. *Let $\mathfrak{o}$ be a local ring. Let f be a unimodular vector in $\mathfrak{o}[x]^{(n)}$ such that some component has leading coefficient 1. Then $f \sim f(0)$ over $\mathfrak{o}[x]$.*

Proof. Note that $f(0) \in \mathfrak{o}^{(n)}$ has one component which is a unit. It suffices to prove that over any commutative ring R any element $c \in R^{(n)}$ such that some component is a unit is equivalent over R to e^1, and this is obvious.

Lemma 3.3. *Let R be an entire ring, and let S be a multiplicative subset. Let x, y be independent variables. If $f(x) \sim f(0)$ over $S^{-1}R[x]$, then there exists $c \in S$ such that $f(x + cy) \sim f(x)$ over $R[x, y]$.*

Proof. Let $M \in GL_n(S^{-1}R[x])$ be such that $f(x) = M(x)f(0)$. Then $M(x)^{-1}f(x) = f(0)$ is constant, and thus invariant under translation $x \mapsto x + y$. Let

$$G(x, y) = M(x)M(x + y)^{-1}.$$

Then $G(x, y)f(x + y) = f(x)$. We have $G(x, 0) = I$ whence

$$G(x, y) = I + yH(x, y)$$

with $H(x, y) \in S^{-1}R[x, y]$. There exists $c \in S$ such that cH has coefficients in R. Then $G(x, cy)$ has coefficients in R. Since $\det M(x)$ is constant in $S^{-1}R$, it follows that $\det M(x + cy)$ is equal to this same constant and therefore that $\det G(x, cy) = 1$. This proves the lemma.

Theorem 3.4. *Let R be an entire ring, and let f be a unimodular vector in $R[x]^{(n)}$, such that one component has leading coefficient 1. Then $f(x) \sim f(0)$ over $R[x]$.*

Proof. Let J be the set of elements $c \in R$ such that $f(x + cy)$ is equivalent to $f(x)$ over $R[x, y]$. Then J is an ideal, for if $c \in J$ and $a \in R$ then replacing y by ay in the definition of equivalence shows that $f(x + cay)$ is equivalent to $f(x)$ over $R[x, ay]$, so over $R[x, y]$. Equally easily, one sees that if c, $c' \in J$ then $c + c' \in J$. Now let $\mathfrak{p}$ be a prime ideal of R. By Corollary 3.2 we know that $f(x)$ is equivalent to $f(0)$ over $R_{\mathfrak{p}}[x]$, and by Lemma 3.3 it follows that there exists $c \in R$ and $c \notin \mathfrak{p}$ such that $f(x + cy)$ is equivalent to $f(x)$ over $R[x, y]$. Hence J is not contained in $\mathfrak{p}$, and so J is unit ideal in R, so there exists an invertible matrix $M(x, y)$ over $R[x, y]$ such that

$$f(x + y) = M(x, y)f(x).$$

Since the homomorphic image of an invertible matrix is invertible, we substitute 0 for x in this last relation to conclude the proof of the theorem.

Theorem 3.5. (Quillen-Suslin). *Let k be a field and let f be a unimodular vector in $k[x_1, \ldots, x_r]^{(n)}$. Then f has the unimodular extension property.*

Proof. By induction on r. If $r = 1$ then $k[x_1]$ is a principal ring and the theorem is left to the reader. Assume the theorem for $r - 1$ variables with $r \geqq 2$, and put

$$R = k[x_1, \ldots, x_{r-1}].$$

We view f as a vector of polynomials in the last variable x_r and want to apply Theorem 3.4. We can do so if some component of f has leading coefficient 1 in the variable x_r. We reduce the theorem to this case as follows. The proof of the Noether Normalization Theorem (Chapter VIII, Theorem 2.1) shows that if we let

$$y_r = x_r$$

$$y_i = x_i - x_r^{m_i}$$

then the polynomial vector

$$f(x_1, \ldots, x_r) = g(y_1, \ldots, y_r)$$

has one component with y_r-leading coefficient equal to 1. Hence there exists a matrix $N(y) = M(x)$ invertible over $R[x_r] = R[y_r]$ such that

$$g(y_1, \ldots, y_r) = N(y_1, \ldots, y_r)g(y_1, \ldots, y_{r-1}, 0),$$

and $g(y_1, \ldots, y_{r-1}, 0)$ is unimodular in $k[y_1, \ldots, y_{r-1}]^{(n)}$. We can therefore conclude the proof by induction.

We now give other formulations of the theorem. First we recall that a module E over a commutative ring A is called **stably free** if there exists a finite free module F such that $E \oplus F$ is finite free.

We shall say that a commutative ring A has the **unimodular column extension property** if every unimodular vector $f \in A^{(n)}$ has the unimodular extension property, for all positive integers n.

Theorem 3.6. *Let A be a commutative ring which has the unimodular column extension property. Then every stably free module over A is free.*

Proof. Let E be stably free. We use induction on the rank of the free modules F such that $E \oplus F$ is free. By induction, it suffices to prove that if $E \oplus A$ is free then E is free. Let $E \oplus A = A^{(n)}$ and let

$$p : A^{(n)} \to A$$

be the projection. Let u^1 be a basis of A over itself. Viewing A as a direct summand in $E \oplus A = A^{(n)}$ we write

$$u^1 = {}^t(a_{11}, \ldots, a_{n1}) \quad \text{with} \quad a_{i1} \in A.$$

Then u^1 is unimodular, and by assumption u^1 is the first column of a matrix $M = (a_{ij})$ whose determinant is a unit in A. Let

$$u^j = Me^j \quad \text{for} \quad j = 1, \ldots, n,$$

where e^j is the j-th unit column vector of $A^{(n)}$. Note that u^1 is the first column of M. By elementary column operations, we may change M so that $u^j \in E$ for $j = 2, \ldots, n$. Indeed, if $pe^j = cu^1$ for $j \geqq 2$ we need only replace e^j by $e^j - ce^1$. Without loss of generality we may therefore assume that $u^2, \ldots, u^n$ lie in E. Since M is invertible over A, it follows that M induces an automorphism of $A^{(n)}$ as A-module with itself by

$$X \mapsto MX.$$

It follows immediately from the construction and the fact that $A^{(n)} = E \oplus A$ that M maps the free module with basis $\{e^2, \ldots, e^n\}$ onto E. This concludes the proof.

If we now feed Serre's Theorem 2.9 into the present machinery consisting of the Quillen-Suslin theorem and Theorem 3.6, we obtain the alternative version of the Quillen-Suslin theorem:

Theorem 3.7. *Let k be a field. Then every finite projective module over the polynomial ring $k[x_1, \ldots, x_r]$ is free.*

§4. THE KOSZUL COMPLEX

In this section, we describe a finite complex built out of the alternating product of a free module. This gives an application of the alternating product, and also gives a fundamental construction used in algebraic geometry, both abstract and complex, as the reader can verify by looking at Griffiths-Harris [GrH 78], Chapter V, §3; Grothendieck's [SGA 6]; Hartshorne [Ha 77], Chapter III, §7; and Fulton-Lang [FuL 85], Chapter IV, §2.

We know from Chapter XX that a free resolution of a module allows us to compute certain homology or cohomology groups of a functor. We apply this now to Hom and also to the tensor product. Thus we also get examples of explicit computations of homology, illustrating Chapter XX, by means of the Koszul complex. We shall also obtain a classical application by deriving the so-called Hilbert Syzygy theorem.

Let A be a ring (always assumed commutative) and M a module. A sequence of elements $x_1, \ldots, x_r$ in A is called M-**regular** if $M/(x_1, \ldots, x_r)M \neq 0$, if x_1

is not divisor of zero in M, and for $i \geqq 2$, x_i is not divisor of 0 in

$$M/(x_1, \ldots, x_{i-1})M.$$

It is called **regular** when $M = A$.

Proposition 4.1. *Let $I = (x_1, \ldots, x_r)$ be generated by a regular sequence in A. Then I/I^2 is free of dimension r over A/I.*

Proof. Let $\bar{x}_i$ be the class of x_i mod I^2. It suffices to prove that $\bar{x}_1, \ldots, \bar{x}_r$ are linearly independent. We do this by induction on r. For $r = 1$, if $\bar{a}\bar{x} = 0$, then $ax = bx^2$ for some $b \in A$, so $x(a - bx) = 0$. Since x is not zero divisor in A, we have $a = bx$ so $\bar{a} = 0$.

Now suppose the proposition true for the regular sequence $x_1, \ldots, x_{r-1}$. Suppose

$$\sum_{i=1}^{r} \bar{a}_i \bar{x}_i = 0 \quad \text{in} \quad I/I^2.$$

We may assume that $\sum a_i x_i = 0$ in A; otherwise $\sum a_i x_i = \sum y_i x_i$ with $y_i \in I$ and we can replace a_i by $a_i - y_i$ without changing $\bar{a}_i$.

Since x_r is not zero divisor in $A/(x_1, \ldots, x_{r-1})$ there exist $b_i \in A$ such that

$$a_r x_r + \sum_{i=1}^{r-1} a_i x_i = 0 \Rightarrow a_r = \sum_{i=1}^{r-1} b_i x_i \Rightarrow \sum_{i=1}^{r-1} (a_i + b_i x_r) x_i = 0.$$

By induction,

$$a_j + b_j x_r \in \sum_{i=1}^{r-1} A x_i \qquad (j = 1, \ldots, r-1)$$

so $a_j \in I$ for all j, so $\bar{a}_j = 0$ for all j, thus proving the proposition.

Let K, L be complexes, which we write as direct sums

$$K = \bigoplus K_p \quad \text{and} \quad L = \bigoplus L_q$$

with $p, q \in \mathbf{Z}$. Usually, $K_p = L_q = 0$ for $p, q < 0$. Then the **tensor product** $K \otimes L$ is the complex such that

$$(K \otimes L)_n = \bigoplus_{p+q=n} K_p \otimes L_q;$$

and for $u \in K_p$, $v \in L_q$ the differential is defined by

$$d(u \otimes v) = du \otimes v + (-1)^p u \otimes dv.$$

(Carry out the detailed verification, which is routine, that this gives a complex.)

Let A be a commutative ring and $x \in A$. We define the complex $K(x)$ to have $K_0(x) = A$, $K_1(x) = Ae_1$, where e_1 is a symbol, Ae_1 is the free module of rank 1 with basis $\{e_1\}$, and the boundary map is defined by $de_1 = x$, so the complex can be represented by the sequence

$$\begin{array}{ccccccc} 0 & \longrightarrow & Ae_1 & \xrightarrow{d} & A & \longrightarrow & 0 \\ & & \| & & \| & & \\ 0 & \longrightarrow & K_1(x) & \longrightarrow & K_0(x) & \longrightarrow & 0 \end{array}$$

More generally, for elements $x_1, \ldots, x_r \in A$ we define the **Koszul complex** $K(x) = K(x_1, \ldots, x_r)$ as follows. We put:

$K_0(x) = A$;

$K_1(x) =$ free module E with basis $\{e_1, \ldots, e_r\}$;

$K_p(x) =$ free module $\bigwedge^p E$ with basis $\{e_{i_1} \wedge \cdots \wedge e_{i_p}\}$, $i_1 < \cdots < i_p$;

$K_r(x) =$ free module $\bigwedge^r E$ of rank 1 with basis $e_1 \wedge \cdots \wedge e_r$.

We define the **boundary maps** by $de_i = x_i$ and in general

$$d : K_p(x) \to K_{p-1}(x)$$

by

$$d(e_{i_1} \wedge \cdots \wedge e_{i_p}) = \sum_{j=1}^{p} (-1)^{j-1} x_{i_j} e_{i_1} \wedge \cdots \wedge \widehat{e_{i_j}} \wedge \cdots \wedge e_{i_p}.$$

A direct verification shows that $d^2 = 0$, so we have a complex

$$0 \to K_r(x) \to \cdots \to K_p(x) \to \cdots \to K_1(x) \to A \to 0$$

The next lemma shows the extent to which the complex is independent of the ideal $I = (x_1, \ldots, x_r)$ generated by (x). Let

$$I = (x_1, \ldots, x_r) \supset I' = (y_1, \ldots, y_r)$$

be two ideals of A. We have a natural ring homomorphism

$$\text{can} : A/I' \to A/I.$$

Let $\{e'_1, \ldots, e'_r\}$ be a basis for $K_1(y)$, and let

$$y_i = \sum c_{ij} x_j \quad \text{with} \quad c_{ij} \in A.$$

We **define** $f_1 : K_1(y) \to K_1(x)$ by

$$f_1 e'_i = \sum c_{ij} e_j$$

and

$$f_p = f_1 \wedge \cdots \wedge f_1, \qquad \text{product taken } p \text{ times.}$$

Let $D = \det(c_{ij})$ be the determinant. Then for $p = r$ we get that

$$f_r : K_r(y) \to K_r(x) \text{ is multiplication by } D.$$

Lemma 4.2. *Notation as above, the homomorphisms f_p define a morphism of Koszul complexes:*

$$\begin{array}{ccccccccccccc}
0 & \longrightarrow & K_r(y) & \longrightarrow & \cdots & \longrightarrow & K_p(y) & \longrightarrow & \cdots & \longrightarrow & K_1(y) & \longrightarrow & A & \longrightarrow & A/I' & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle f_r = D} & & & & \downarrow{\scriptstyle f_p} & & & & \downarrow{\scriptstyle f_1} & & \downarrow{\scriptstyle \mathrm{id}} & & \downarrow{\scriptstyle \mathrm{can}} & & \\
0 & \longrightarrow & K_r(x) & \longrightarrow & \cdots & \longrightarrow & K_p(x) & \longrightarrow & \cdots & \longrightarrow & K_1(x) & \longrightarrow & A & \longrightarrow & A/I & \longrightarrow & 0
\end{array}$$

and define an isomorphism if D is a unit in A, for instance if (y) is a permutation of (x).

Proof. By definition

$$f(e'_{i_1} \wedge \cdots \wedge e'_{i_p}) = \left(\sum_{j=1}^{r} c_{i_1 j} e_j\right) \wedge \cdots \wedge \left(\sum_{j=1}^{r} c_{i_p j} e_j\right).$$

Then

$$\begin{aligned}
&fd(e'_{i_1} \wedge \cdots \wedge e'_{i_p}) \\
&\quad = f\left(\sum_k (-1)^{k-1} y_{i_k} e'_{i_1} \wedge \cdots \wedge \widehat{e'_{i_k}} \wedge \cdots \wedge e'_{i_p}\right) \\
&\quad = \sum_k (-1)^{k-1} y_{i_k} \left(\sum_{j=1}^{r} c_{i_1 j} e_j\right) \wedge \cdots \wedge \widehat{\sum_k} \wedge \cdots \wedge \left(\sum_{j=1}^{r} c_{i_p j} e_j\right) \\
&\quad = \sum (-1)^{k-1} \left(\sum_{j=1}^{r} c_{i_1 j} e_j\right) \wedge \cdots \wedge \underbrace{\left(\sum_{j=1}^{r} c_{i_k j} x_j e_j\right)}_{\text{omitted}} \wedge \cdots \wedge \left(\sum_{j=1}^{r} c_{i_p j} e_j\right) \\
&\quad = df(e'_{i_1} \wedge \cdots \wedge e'_{i_p})
\end{aligned}$$

using $y_{i_k} = \sum c_{i_k j} x_j$. This concludes the proof that the f_p define a homomorphism of complexes.

In particular, if (x) and (y) generate the same ideal, and the determinant D is a unit (i.e. the linear transformation going from (x) to (y) is invertible over the ring), then the two Koszul complexes are isomorphic.

The next lemma gives us a useful way of making inductions later.

Proposition 4.3. *There is a natural isomorphism*

$$K(x_1, \ldots, x_r) \approx K(x_1) \otimes \cdots \otimes K(x_r).$$

Proof. The proof will be left as an exercise.

Let $I = (x_1, \ldots, x_r)$ be the ideal generated by $x_1, \ldots, x_r$. Then directly from the definitions we see that the 0-th homology of the Koszul complex is simply A/IA.

More generally, let M be an A-module. Define the **Koszul complex of** M by

$$K(x; M) = K(x_1, \ldots, x_r; M) = K(x_1, \ldots, x_r) \otimes_A M$$

Then this complex looks like

$$0 \to K_r(x) \otimes M \to \cdots \to K_2(x) \otimes_A M \to M^{(r)} \to M \to 0.$$

We sometimes abbreviate $H_p(x; M)$ for $H_pK(x; M)$. The first and last homology groups are then obtained directly from the definition of boundary. We get

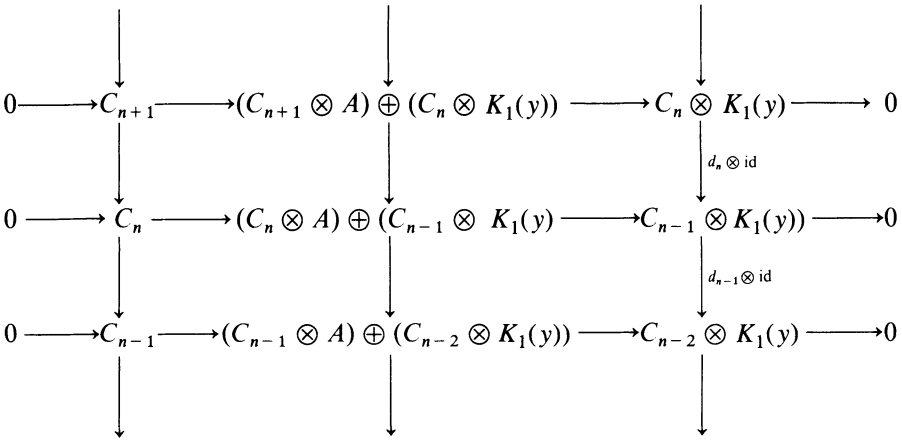
$H_0(K(x; M)) \approx M/IM$;

$H_r(K(x); M) = \{v \in M$ such that $x_iv = 0$ for all $i = 1, \ldots, r\}$.

In light of Proposition 4.3, we study generally what happens to a tensor product of any complex with $K(x)$, when x consists of a single element. Let $y \in A$ and let C be an arbitrary complex of A-modules. We have an exact sequence of complexes

$$0 \to C \to C \otimes K(y) \to (C \otimes K(y))/C \to 0 \tag{1}$$

made explicit as follows.



We note that $C \otimes K_1(y)$ is just C with a dimension shift by one unit, in other words

$$(C \otimes K_1(y))_{n+1} = C_n \otimes K_1(y). \tag{2}$$

In particular,

$$H_{n+1}(C \otimes K(y)/C) \approx H_n(C). \tag{3}$$

Associated with an exact sequence of complexes, we have the homology sequence, which in this case yields the long exact sequence

$$\longrightarrow H_{n+1}(C) \longrightarrow H_{n+1}(C \otimes K_1(y)) \longrightarrow H_{n+1}(C \otimes K(y)/C) \xrightarrow{\partial} \begin{matrix} H_n(C) \\ \wr\wr \\ H_n(C) \end{matrix}$$

which we write stacked up according to the index:

$$\begin{aligned} &\to H_{p+1}(C) \to H_{p+1}(C) \to H_{p+1}(C \otimes K(y)) \to \\ &\to H_p(C) \to H_p(C) \to H_p(C \otimes K(y)) \to \end{aligned} \tag{4}$$

ending in lowest dimension with

$$\to H_1(C) \to H_1(C \otimes K(y)) \to H_0(C) \to H_0(C). \tag{5}$$

Furthermore, a direct application of the definition of the boundary map and the tensor product of complexes yields:

The boundary map on $H_p(C)$ $(p \geqq 0)$ is induced by multiplication by $(-1)^p y$:

$$\partial = (-1)^p m(y) : H_p(C) \to H_p(C). \tag{6}$$

Indeed, write

$$(C \otimes K(y))_p = (C_p \otimes A) \oplus (C_{p-1} \otimes K_1(y)) \approx C_p \oplus C_{p-1}.$$

Let $(v, w) \in C_p \oplus C_{p-1}$ with $v \in C_p$ and $w \in C_{p-1}$. Then directly from the definitions,

$$d(v, w) = (dv + (-1)^{p-1} yw, dw). \tag{7}$$

To see (6), one merely follows up the definitions of the boundary, taking an element $w \in C_p \approx C_p \otimes K_1(y)$, lifting back to $(0, w)$, applying d, and lifting back to C_p. If we start with a cycle, i.e. $dw = 0$, then the map is well defined on the homology class, with values in the homology.

Lemma 4.4. *Let $y \in A$ and let C be a complex as above. Then $m(y)$ annihilates $H_p(C \otimes K(y))$ for all $p \geqq 0$.*

Proof. If (v, w) is a cycle, i.e. $d(v, w) = 0$, then from (7) we get at once that $(yv, yw) = d(0, (-1)^p v)$, which proves the lemma.

In the applications we have in mind, we let $y = x_r$ and

$$C = K(x_1, \ldots, x_{r-1}; M) = K(x_1, \ldots, x_{r-1}) \otimes M.$$

Then we obtain:

Theorem 4.5.(a) *There is an exact sequence with maps as above:*

$$\to H_pK(x_1, \ldots, x_{r-1}; M) \to H_pK(x_1, \ldots, x_{r-1}; M) \to H_pK(x_1, \ldots, x_r; M)$$

$$\cdots \to H_1(x_1, \ldots, x_r; M) \to H_0(x_1, \ldots, x_{r-1}; M) \xrightarrow{m(x_r)} H_0(x_1, \ldots, x_{r-1}; M).$$

(b) *Every element of* $I = (x_1, \ldots, x_r)$ *annihilates* $H_p(x; M)$ *for* $p \geqq 0$.
(c) *If* $I = A$, *then* $H_p(x; M) = 0$ *for all* $p \geqq 0$.

Proof. This is immediate from Proposition 4.3 and Lemma 4.4.

We define the **augmented Koszul complex** to be

$$0 \to K_r(x; M) \to \cdots \to K_1(x; M) = M^{(r)} \to M \to M/IM \to 0.$$

Theorem 4.6. *Let M be an A-module.*
(a) *Let* $x_1, \ldots, x_r$ *be a regular sequence for M. Then* $H_pK(x; M) = 0$ *for* $p > 0$. (*Of course,* $H_0K(x; M) = M/IM$.) *In other words, the augmented Koszul complex is exact.*
(b) *Conversely, suppose A is local, and* $x_1, \ldots, x_r$ *lie in the maximal ideal of A. Suppose M is finite over A, and also assume that* $H_1K(x; M) = 0$. *Then* $(x_1, \ldots, x_r)$ *is M-regular.*

Proof. We prove (a) by induction on r. If $r = 1$ then $H_1(x; M) = 0$ directly from the definition. Suppose $r > 1$. We use the exact sequence of Theorem 4.5(a). If $p > 1$ then $H_p(x; M)$ is between two homology groups which are 0, so $H_p(x; M) = 0$. If $p = 1$, we use the very end of the exact sequence of Theorem 4.5(a), noting that $m(x_r)$ is injective, so by induction we find $H_1(x; M) = 0$ also, thus proving (a).

As to (b), by Lemma 4.4 and the hypothesis, we get an exact sequence

$$H_1(x_1, \ldots, x_{r-1}; M) \xrightarrow{m(x_r)} H_1(x_1, \ldots, x_{r-1}; M) \to H_1(x; M) = 0,$$

so $m(x_r)$ is surjective. By Nakayama's lemma, it follows that

$$H_1(x_1, \ldots, x_{r-1}; M) = 0.$$

By induction $(x_1, \ldots, x_{r-1})$ is an M-regular sequence. Looking again at the tail end of the exact sequence as in (a) shows that x_r is $M/(x_1, \ldots, x_{r-1})M$-regular, whence proving (b) and the theorem.

We note that (b), which uses only the triviality of H_1 (and not all H_p) is due to Northcott [No 68], 8.5, Theorem 8. By (a), it follows that $H_p = 0$ for $p > 0$.

An important special case of Theorem 4.6(a) is when $M = A$, in which case we restate the theorem in the form:

Let $x_1, \ldots, x_r$ be a regular sequence in A. Then $K(x_1, \ldots, x_r)$ is a free resolution of A/I:

$$0 \to K_r(x) \to \cdots \to K_1(x) \to A \to A/I \to 0.$$

In particular, A/I has Tor*-dimension $\leqq r$.*

For the Hom functor, we have:

Theorem 4.7. *Let $x_1, \ldots, x_r$ be a regular sequence in A. Then there is an isomorphism*

$$\varphi_{x,M}: H^r(\mathrm{Hom}(K(x), M)) \to M/IM$$

to be described below.

Proof. The module $K_r(x)$ is 1-dimensional, with basis $e_1 \wedge \cdots \wedge e_r$. Depending on this basis, we have an isomorphism

$$\mathrm{Hom}(K_r(x), M) \approx M,$$

whereby a homomorphism is determined by its value at the basis element in M. Then directly from the definition of the boundary map d_r in the Koszul complex, which is

$$d_r : e_1 \wedge \cdots \wedge e_r \mapsto \sum_{j=1}^{r} (-1)^{j-1} x_j e_1 \wedge \cdots \wedge \hat{e}_j \wedge \cdots \wedge e_r$$

we see that

$$\begin{aligned} H^r(\mathrm{Hom}(K_r(x), M) &\approx \mathrm{Hom}(K_r(x), M)/d^{r-1}\,\mathrm{Hom}(K_{r-1}(x), M) \\ &\approx M/IM. \end{aligned}$$

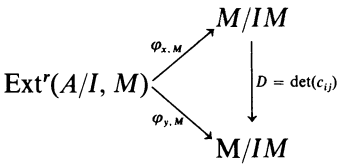
This proves the theorem.

The reader who has read Chapter XX knows that the i-th homology group of $\mathrm{Hom}(K(x), M)$ is called $\mathrm{Ext}^i(A/I, M)$, determined up to a unique isomorphism by the complex, since two resolutions of A/I differ by a morphism of complexes, and two such morphisms differ by a homotopy which induces a homology isomorphism. Thus Theorem 4.7 gives an isomorphism

$$\varphi_{x,M} : \mathrm{Ext}^r(A/I, M) \to M/IM.$$

In fact, we shall obtain morphisms of the Koszul complex from changing the sequence. We go back to the hypothesis of Lemma 4.2.

Lemma 4.8. *If $I = (x) = (y)$ where (x), (y) are two regular sequences, then we have a commutative diagram*



where all the maps are isomorphisms of A/I-modules.

The fact that we are dealing with A/I-modules is immediate since multiplication by an element of A commutes with all homomorphisms in sight, and I annihilates A/I.

By Proposition 4.1, we know that I/I^2 is a free module of rank r over A/I. Hence

$$\bigwedge^r(I/I^2)$$

is a free module of rank 1, with basis $\bar{x}_1 \wedge \cdots \wedge \bar{x}_r$ (where the bar denotes residue class mod I^2). Taking the dual of this exterior product, we see that under a change of basis, it transforms according to the inverse of the determinant mod I^2. This allows us to get a canonical isomorphism as in the next theorem.

Theorem 4.9. *Let $x_1, \ldots, x_r$ be a regular sequence in A, and let $I = (x)$. Let M be an A-module. Let*

$$\psi_{x,M} : M/IM \to (M/IM) \otimes \bigwedge^r(I/I^2)^{\mathrm{dual}}$$

be the embedding determined by the basis $(\bar{x}_1 \wedge \cdots \wedge \bar{x}_r)^{\mathrm{dual}}$ of $\bigwedge^r(I/I^2)^{\mathrm{dual}}$. Then the composite isomorphism

$$\mathrm{Ext}^r(A/I, M) \xrightarrow{\varphi_{x,M}} M/IM \xrightarrow{\psi_{x,M}} (M/IM) \otimes \bigwedge^r(I/I^2)^{\mathrm{dual}}$$

is a functorial isomorphism, independent of the choice of regular generators for I.

We also have the analogue of Theorem 4.5 in intermediate dimensions.

Theorem 4.10. *Let $x_1, \ldots, x_r$ be an M-regular sequence in A. Let $I = (x)$. Then*

$$\mathrm{Ext}^i(A/I, M) = 0 \quad \textit{for} \quad i < r.$$

Proof. For the proof, we assume that the reader is acquainted with the exact homology sequence. Assume by induction that $\mathrm{Ext}^i(A/I, M) = 0$ for

$i < r - 1$. Then we have the exact sequence

$$0 = \mathrm{Ext}^{i-1}(A/I, M/x_1M) \to \mathrm{Ext}^i(A/I, M) \xrightarrow{x_1} \mathrm{Ext}^i(A/I, M)$$

for $i < r$. But $x_1 \in I$ so multiplication by x_1 induces 0 on the homology groups, which gives $\mathrm{Ext}^i(A/I, M) = 0$ as desired.

Let $L_N \to N \to 0$ be a free resolution of a module N. By definition,

$$\mathrm{Tor}_i^A(N, M) = i\text{-th homology of the complex } L \otimes M.$$

This is independent of the choice of L_N up to a unique isomorphism. We now want to do for Tor what we have just done for Ext.

Theorem 4.11. *Let $I = (x_1, \ldots, x_r)$ be an ideal of A generated by a regular sequence of length r.*

(i) *There is a natural isomorphism*

$$\mathrm{Tor}_i^A(A/I, A/I) \approx \bigwedge_{A/I}^i(I/I^2), \quad \textit{for} \quad i \geqq 0.$$

(ii) *Let L be a free A/I-module, extended naturally to an A-module. Then*

$$\mathrm{Tor}_i^A(L, A/I) \approx L \otimes \bigwedge_{A/I}^i(I/I^2), \quad \textit{for} \quad i \geqq 0.$$

These isomorphisms will follow from the next considerations.

First we use again that the residue classes $\bar{x}_1, \ldots, \bar{x}_r$ mod I^2 form a basis of I/I^2 over A/I. Therefore we have a unique isomorphism of complexes

$$\varphi_x : K(x) \otimes A/I \to \bigwedge(I/I^2) = \bigoplus \bigwedge^i(I/I^2)$$

with zero differentials on the right-hand side, such that

$$e_{i_1} \wedge \cdots \wedge e_{i_p} \mapsto \bar{x}_{i_1} \wedge \cdots \wedge \bar{x}_{i_p}.$$

Lemma 4.12. *Let $I = (x) \supset I' = (y)$ be two ideals generated by regular sequences of length r. Let $f : K(y) \to K(x)$ be the morphism of Koszul complexes defined in Lemma* 4.2. *Then the following diagram is commutative:*

$$\begin{array}{ccc} K(y) \otimes A/I' & \xrightarrow{\varphi_y} & \bigwedge_{A/I'}(I'/I'^2) \\ \downarrow{\scriptstyle f \otimes \mathrm{can}} & & \downarrow{\scriptstyle \text{canonical hom}} \\ K(x) \otimes A/I & \xrightarrow[\varphi_x]{} & \bigwedge_{A/I}(I/I^2) \end{array}$$

Proof. We have

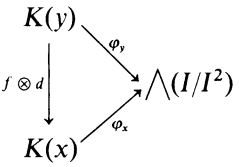
$$\varphi_x \circ (f \otimes \mathrm{can})(e'_{i_1} \wedge \cdots \wedge e'_{i_p} \otimes 1)$$

$$= \sum_{j=2}^{r} c_{i_1 j} \bar{x}_j \wedge \cdots \wedge \sum_{j=1}^{r} c_{i_p j} \bar{x}_j$$

$$= \bar{y}_{i_1} \wedge \cdots \wedge \bar{y}_{i_p} = \mathrm{can}(\varphi_y(e'_{i_1} \wedge \cdots \wedge e'_{i_p})).$$

This proves the lemma.

In particular, if $I' = I$ then we have the commutative diagram



which shows that the identification of $\mathrm{Tor}_i(A/I, A/I)$ with $\bigwedge^i(I/I^2)$ via the choices of bases is compatible under one isomorphism of the Koszul complexes, which provide a resolution of A/I. Since any other homomorphism of Koszul complexes is homotopic to this one, it follows that this identification does not depend on the choices made and proves the first part of Theorem 4.11.

The second part follows at once, because we have

$$\mathrm{Tor}_i^A(A/I, L) = H_i(K(x) \otimes L) = H_i((K(x) \otimes_A A/I) \otimes_{A/I} L$$

$$= \bigwedge^i_{A/I}(I/I^2) \otimes L.$$

This concludes the proof of Theorem 4.11.

Example. Let k be a field and let $A = k[x_1, \ldots, x_r]$ be the polynomial ring in r variables. Let $I = (x_1, \ldots, x_r)$ be the ideal generated by the variables. Then $A/I = k$, and therefore Theorem 4.11 yields for $i \geqq 0$:

$$\mathrm{Tor}_i^A(k, k) \approx \bigwedge^i_k(I/I^2)$$

$$\mathrm{Tor}_i^A(L, k) \approx L \otimes \bigwedge^i_k(I/I^2)$$

Note that in the present case, we can think of I/I^2 as the vector space over k with basis $\bar{x}_1, \ldots, \bar{x}_r$. Then A can be viewed as the symmetric algebra SE, where E is this vector space. We can give a specific example of the Koszul complex in this context as in the next theorem, given for a free module.

Theorem 4.13. *Let E be a finite free module of rank r over the ring R. For each $p = 1, \ldots, r$ there is a unique homomorphism*

$$d_p : \bigwedge^p E \otimes SE \to \bigwedge^{p-1} E \otimes SE$$

such that

$$d_i((x_1 \wedge \cdots \wedge x_p) \otimes y)$$

$$= \sum_{i=1}^{p} (-1)^{i-1}(x_1 \wedge \cdots \wedge \hat{x_i} \wedge \cdots \wedge x_p) \otimes (x_i \otimes y)$$

where $x_i \in E$ and $y \in SE$. This gives the resolution

$$0 \to \bigwedge^r E \otimes SE \to \bigwedge^{r-1} E \otimes SE \to \cdots \to \bigwedge^0 E \otimes SE \to R \to 0$$

Proof. The above definitions are merely examples of the Koszul complex for the symmetric algebra SE with respect to the regular sequence consisting of some basis of E.

Since d_p maps $\bigwedge^p E \otimes S^q E$ into $\bigwedge^{p-1} E \otimes S^{q+1} E$, we can decompose this complex into a direct sum corresponding to a given graded component, and hence:

Corollary 4.14. *For each integer $n \geqq 1$, we have an exact sequence*

$$0 \to \bigwedge^r E \otimes S^{n-r} E \to \cdots \to \bigwedge^1 E \otimes S^{n-1} E \to S^n E \to 0$$

where $S^j E = 0$ for $j < 0$.

Finally, we give an application to a classical theorem of Hilbert. The polynomial ring $A = k[x_1, \ldots, x_r]$ is naturally graded, by the degrees of the homogeneous components. *We shall consider graded modules, where the grading is in dimensions $\geqq 0$, and we assume that homomorphisms are graded of degree 0.*

So suppose M is a graded module (and thus $M_i = 0$ for $i < 0$) and M is finite over A. Then we can find a graded surjective homomorphism

$$L_0 \to M \to 0$$

where L_0 is finite free. Indeed, let $w_1, \ldots, w_n$ be homogeneous generators of M. Let $e_1, \ldots, e_n$ be basis elements for a free module L_0 over A. We give L_0 the grading such that if $a \in A$ is homogeneous of degree d then ae_i is homogeneous of degree

$$\deg ae_i = \deg a + \deg w_i.$$

Then the homomorphism of L_0 onto M sending $e_i \mapsto w_i$ is graded as desired.

The kernel M_1 is a graded submodule of L_0. Repeating the process, we can find a surjective homomorphism

$$L_1 \to M_1 \to 0.$$

We continue in this way to obtain a graded resolution of M. We want this resolution to stop, and the possibility of its stopping is given by the next theorem.

Theorem 4.15. (Hilbert Syzygy Theorem). *Let k be a field and*

$$A = k[x_1, \ldots, x_r]$$

the polynomial ring in r variables. Let M be a graded module over A, and let

$$0 \to K \to L_{r-1} \to \cdots \to L_0 \to M \to 0$$

be an exact sequence of graded homomorphisms of graded modules, such that $L_0, \ldots, L_{r-1}$ are free. Then K is free. If M is in addition finite over A and $L_0, \ldots, L_{r-1}$ are finite free, then K is finite free.

Proof. From the Koszul complex we know that $\operatorname{Tor}_i(M, k) = 0$ for $i > r$ and all M. By dimension shifting, it follows that

$$\operatorname{Tor}_i(K, k) = 0 \quad \text{for} \quad i > 0.$$

The theorem is then a consequence of the next result.

Theorem 4.16. *Let F be a graded finite module over $A = k[x_1, \ldots, x_r]$. If $\operatorname{Tor}_1(F, k) = 0$ then F is free.*

Proof. The method is essentially to do a Nakayama type argument in the case of the non-local ring A. First note that

$$F \otimes k = F/IF$$

where $I = (x_1, \ldots, x_r)$. Thus $F \otimes k$ is naturally an $A/I = k$-module. Let $v_1, \ldots, v_n$ be homogeneous elements of F whose residue classes mod IF form a basis of F/IF over k. Let L be a free module with basis $e_1, \ldots, e_n$. Let

$$L \to F$$

be the graded homomorphism sending $e_i \mapsto v_i$ for $i = 1, \ldots, n$. It suffices to prove that this is an isomorphism. Let C be the cokernel, so we have the exact sequence

$$L \to F \to C \to 0.$$

Tensoring with k yields the exact sequence

$$L \otimes k \to F \otimes k \to C \otimes k \to 0.$$

Since by construction the map $L \otimes k \to F \otimes k$ is surjective, it follows that $C \otimes k = 0$. But C is graded, so the next lemma shows that $C = 0$.

Lemma 4.17. *Let N be a graded module over $A = k[x_1, \ldots, x_r]$. Let $I = (x_1, \ldots, x_r)$. If $N/IN = 0$ then $N = 0$.*

Proof. This is immediate by using the grading, looking at elements of N of smallest degree if they exist, and using the fact that elements of I have degree > 0.

We now get an exact sequence of graded modules

$$0 \to E \to L \to F \to 0$$

and we must show that $E = 0$. But the exact homology sequence and our assumption yields

$$0 = \operatorname{Tor}_1(F, k) \to E \otimes k \to L \otimes k \to F \otimes k \to 0.$$

By construction $L \otimes k \to F \otimes k$ is an isomorphism, and hence $E \otimes k = 0$. Lemma 4.17 now shows that $E = 0$. This concludes the proof of the syzygy theorem.

Remark. The only place in the proof where we used that k is a field is in the proof of Theorem 4.16 when we picked homogeneous elements $v_1, \ldots, v_n$ in M whose residue classes mod IM form a basis of M/IM over A/IA. Hilbert's theorem can be generalized by making the appropriate hypothesis which allows us to carry out this step, as follows.

Theorem 4.18. *Let R be a commutative local ring and let $A = R[x_1, \ldots, x_r]$ be the polynomial ring in r variables. Let M be a graded finite module over A, projective over R. Let*

$$0 \to K \to L_{r-1} \to \cdots \to L_0 \to M \to 0$$

be an exact sequence of graded homomorphisms of graded modules such that $L_0, \ldots, L_{r-1}$ are finite free. Then K is finite free.

Proof. Replace k by R everywhere in the proof of the Hilbert syzygy theorem. We use the fact that a finite projective module over a local ring is free. Not a word needs to be changed in the above proof with the following exception. We note that the projectivity propagates to the kernels and cokernels in the given resolution. Thus F in the statement of Theorem 4.16 may be assumed projective, and each graded component is projective. Then F/IF is projective over $A/IA = R$, and so is each graded component. Since a finite projective module over a local ring is free, and one gets the freeness by lifting a basis from the residue class field, we may pick $v_1, \ldots, v_n$ homogeneous exactly as we did in the proof of Theorem 4.16. This concludes the proof.

EXERCISES

For exercises 1 through 4 on the Koszul complex, see [No 68], Chapter 8.

1. Let $0 \to M' \to M \to M'' \to 0$ be an exact sequence of A-modules. Show that tensoring with the Koszul complex $K(x)$ one gets an exact sequence of complexes, and therefore an exact homology sequence

$$0 \to H_r K(x; M') \to H_r K(x; M) \to H_r K(x; M'') \to \cdots$$
$$\cdots \to H_p K(x; M') \to H_p K(x; M) \to H_p K(x; M'') \to \cdots$$
$$\cdots \to H_0 K(x; M') \to H_0 K(x; M) \to H_0 K(x; M'') \to 0$$

2. (a) Show that there is a unique homomorphism of complexes

$$f : K(x; M) \to K(x_1, \ldots, x_{r-1}; M)$$

such that for $v \in M$:

$$f_p(e_{i_1} \wedge \cdots \wedge e_{i_p} \otimes v) = \begin{cases} e_{i_1} \wedge \cdots \wedge e_{i_p} \otimes x_r v & \text{if } i_p = r \\ e_{i_1} \wedge \cdots \wedge e_{i_p} \otimes v & \text{if } i_p = r. \end{cases}$$

(b) Show that f is injective if x_r is not a divisor of zero in M.

(c) For a complex C, denote by $C(-1)$ the complex shifted by one place to the left, so $C(-1)_n = C_{n-1}$ for all n. Let $\overline{M} = M/x_r M$. Show that there is a unique homomorphism of complexes

$$g: K(x_1, \ldots, x_{r-1}, 1; M) \to K(x_1, \ldots, x_{r-1}; \overline{M})(-1)$$

such that for $v \in M$:

$$g_p(e_{i_1} \wedge \cdots \wedge c_{i_p} \otimes v) = \begin{cases} e_{i_1} \wedge \cdots \wedge e_{i_{p-1}} \otimes v & \text{if } i_p = r \\ 0 & \text{if } i_p < r. \end{cases}$$

(d) If x_r is not a divisor of 0 in M, show that the following sequence is exact:

$$0 \to K(x; M) \xrightarrow{f} K(x_1, \ldots, x_{r-1}, 1; M) \xrightarrow{g} K(x_1, \ldots, x_{r-1}; \overline{M})(-1) \to 0.$$

Using Theorem 4.5(c), conclude that for all $p \geqq 0$, there is an isomorphism

$$H_p K(x; M) \xrightarrow{\simeq} H_p K(x_1, \ldots, x_{r-1}; \overline{M}).$$

3. Assume A and M Noetherian. Let I be an ideal of A. Let $a_1, \ldots, a_k$ be an M-regular sequence in I. Show that this sequence can be extended to a maximal M-regular sequence $a_1, \ldots, a_q$ in I, in other words an M-regular sequence such that there is no M-regular sequence $a_1, \ldots, a_{q+1}$ in I.

4. Again assume A and M Noetherian. Let $I = (x_1, \ldots, x_r)$ and let $a_1, \ldots, a_q$ be a maximal M-regular sequence in I. Assume $IM \neq M$. Prove that

$$H_{r-q}(x; M) \neq 0 \text{ but } H_p(x; M) = 0 \text{ for } p > r - q.$$

[See [No 68], 8.5 Theorem 6. The result is similar to the result in Exercise 5, and generalizes Theorem 4.5(a). See also [Mat 80], pp. 100-103. The result shows that

all maximal M-regular sequences in M have the same length, which is called the I-**depth** of M and is denoted by $\operatorname{depth}_I(M)$. For the proof, let s be the maximal integer such that $H_s K(x; M) \neq 0$. By assumption, $H_0(x; M) = M/IM \neq 0$, so s exists. We have to prove that $q + s = r$. First note that if $q = 0$ then $s = r$. Indeed, if $q = 0$ then every element of I is zero divisor in M, whence I is contained in the union of the associated primes of M, whence in some associated prime of M. Hence $H_r(x; M) \neq 0$.

Next assume $q > 0$ and proceed by induction. Consider the exact sequence

$$0 \to M \xrightarrow{a_1} M \to M/a_1M \to 0$$

where the first map is $m(a_1)$. Since I annihilates $H_p(x; M)$ by Theorem 4.5(c), we get an exact sequence

$$0 \to H_p(x; M) \to H_p(x; M/a_1M) \to H_{p-1}(x; M) \to 0.$$

Hence $H_{s+1}(x; M/a_1M) \neq 0$, but $H_p(x; M/a_1M) = 0$ for $p \geqq s + 2$. From the hypothesis that $a_1, \ldots, a_q$ is a maximal M-regular sequence, it follows at once that $a_2, \ldots, a_q$ is maximal M/a_1M-regular in I, so by induction, $q - 1 = r - (s + 1)$ and hence $q + s = r$, as was to be shown.]

5. The following exercise combines some notions of Chapter XX on homology, and some notions covered in this chapter and in Chapter X, §5. Let M be an A-module.

Let A be Noetherian, M finite module over A, and I an ideal of A such that $IM \neq M$. Let r be an integer $\geqq 1$. Prove that the following conditions are equivalent:

(i) $\operatorname{Ext}^i(N, M) = 0$ for all $i < r$ and all finite modules N such that $\operatorname{supp}(N) \subset \mathfrak{Z}(I)$.

(ii) $\operatorname{Ext}^i(A/I, M) = 0$ for all $i < r$.

(iii) There exists a finite module N with $\operatorname{supp}(N) = \mathfrak{Z}(I)$ such that

$$\operatorname{Ext}^i(N, M) = 0 \quad \text{for all } i < r.$$

(iv) There exists an M-regular sequence $a_1, \ldots, a_r$ in I.

[*Hint*: (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) is clear. For (iii) $\Rightarrow$ (iv), first note that

$$0 = \operatorname{Ext}^0(N, M) = \operatorname{Hom}(N, M).$$

Assume $\operatorname{supp}(N) = \mathfrak{Z}(I)$. Find an M-regular element in I. If there is no such element, then I is contained in the set of divisors of 0 of M in A, which is the union of the associated primes. Hence $I \subset P$ for some associated prime P. This yields an injection $A/P \subset M$, so

$$0 \neq \operatorname{Hom}_{A_P}(A_P/PA_P, M).$$

By hypothesis, $N_P \neq 0$ so $N_P/PN_P \neq 0$, and N_P/PN_P is a vector space over A_P/PA_P, so there exists a non-zero A_P/PA_P homomorphism

$$N_P/PN_P \to M_P,$$

so $\operatorname{Hom}_{A_P}(N_P, M_P) \neq 0$, whence $\operatorname{Hom}(N, M) \neq 0$, a contradiction. This proves the existence of one regular element a_1.

Now let $M_1 = M/a_1M$. The exact sequence

$$0 \to M \xrightarrow{a_1} M \to M/a_1M \to 0$$

yields the exact cohomology sequence

$$\to \mathrm{Ext}^i(N, M) \to \mathrm{Ext}^i(N, M/a_1M) \to \mathrm{Ext}^{i+1}(N, M) \to$$

so $\mathrm{Ext}^i(N, M/a_1M) = 0$ for $i < r - 1$. By induction there exists an M_1-regular sequence $a_2, \ldots, a_r$ and we are done.

Last, (iv) $\Rightarrow$ (i). Assume the existence of the regular sequence. By induction, $\mathrm{Ext}^i(N, a_1M) = 0$ for $i < r - 1$. We have an exact sequence for $i < r$:

$$0 \to \mathrm{Ext}^i(N, M) \xrightarrow{a_1} \mathrm{Ext}^i(N, M)$$

But $\mathrm{supp}(N) = \mathfrak{Z}(\mathrm{ann}(N)) \subset \mathfrak{Z}(I)$, so $I \subset \mathrm{rad}(\mathrm{ann}(N))$, so a_1 is nilpotent on N. Hence a_1 is nilpotent on $\mathrm{Ext}^i(N, M)$, so $\mathrm{Ext}^i(N, M) = 0$. Done.] See Matsumura's [Mat 70], p. 100, Theorem 28. The result is useful in algebraic geometry, with for instance $M = A$ itself. One thinks of A as the affine coordinate ring of some variety, and one thinks of the equations $a_i = 0$ as defining hypersurface sections of this variety, and the simultaneous equations $a_1 = \cdots = a_r = 0$ as defining a complete intersection. The theorem gives a cohomological criterion in terms of Ext for the existence of such a complete intersection.

APPENDIX 1

The Transcendence of e and π

The proof which we shall give here follows the classical method of Gelfond and Schneider, properly formulated. It is based on a theorem concerning values of functions satisfying differential equations, and it had been recognized for some time that such values are subject to severe restrictions, in various contexts. Here, we deal with the most general algebraic differential equation.

We shall assume that the reader is acquainted with elementary facts concerning functions of a complex variable. Let f be an entire function (i.e. a function which is holomorphic on the complex plane). For our purposes, we say f is of order $\leqq \rho$ if there exists a number $C > 1$ such that for all large R we have

$$|f(z)| \leqq C^{R^\rho}$$

whenever $|z| \leqq R$. A meromorphic function is said to be of order $\leqq \rho$ if it is a quotient of entire functions of order $\leqq \rho$.

Theorem. *Let K be a finite extension of the rational numbers. Let $f_1, \ldots, f_N$ be meromorphic functions of order $\leqq \rho$. Assume that the field $K(f_1, \ldots, f_N)$ has transcendence degree $\geqq 2$ over K, and that the derivative $D = d/dz$ maps the ring $K[f_1, \ldots, f_N]$ into itself. Let $w_1, \ldots, w_m$ be distinct complex numbers not lying among the poles of the f_i, such that*

$$f_i(w_\nu) \in K$$

for all $i = 1, \ldots, N$ and $\nu = 1, \ldots, m$. Then $m \leqq 10\rho[K : \mathbf{Q}]$.

Corollary 1. (Hermite-Lindemann). *If α is algebraic (over $\mathbf{Q}$) and $\neq 0$, then e^α is transcendental. Hence π is transcendental.*

Proof. Suppose that α and e^{α} are algebraic. Let $K = \mathbf{Q}(\alpha, e^{\alpha})$. The two functions z and e^{z} are algebraically independent over K (trivial), and the ring $K[z, e^{z}]$ is obviously mapped into itself by the derivative. Our functions take on algebraic values in K at $\alpha, 2\alpha, \ldots, m\alpha$ for any m, contradiction. Since $e^{2\pi i} = 1$, it follows that $2\pi i$ is transcendental.

Corollary 2. (Gelfond-Schneider). *If α is algebraic $\neq 0, 1$ and if β is algebraic irrational, then $\alpha^{\beta} = e^{\beta \log \alpha}$ is transcendental.*

Proof. We proceed as in Corollary 1, considering the functions $e^{\beta t}$ and e^{t} which are algebraically independent because β is assumed irrational. We look at the numbers $\log \alpha, 2 \log \alpha, \ldots, m \log \alpha$ to get a contradiction as in Corollary 1.

Before giving the main arguments proving the theorem, we state some lemmas. The first two, due to Siegel, have to do with integral solutions of linear homogeneous equations.

Lemma 1. *Let*

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\cdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n &= 0 \end{aligned}$$

be a system of linear equations with integer coefficients a_{ij}, and $n > r$. Let A be a number such that $|a_{ij}| \leqq A$ for all i, j. Then there exists an integral, non-trivial solution with

$$|x_j| \leqq 2(2nA)^{r/(n-r)}.$$

Proof. We view our system of linear equations as a linear equation $L(X) = 0$, where L is a linear map, $L : \mathbf{Z}^{(n)} \to \mathbf{Z}^{(r)}$, determined by the matrix of coefficients. If B is a positive number, we denote by $\mathbf{Z}^{(n)}(B)$ the set of vectors X in $\mathbf{Z}^{(n)}$ such that $|X| \leqq B$ (where $|X|$ is the maximum of the absolute values of the coefficients of X). Then L maps $\mathbf{Z}^{(n)}(B)$ into $\mathbf{Z}^{(r)}(nBA)$. The number of elements in $\mathbf{Z}^{(n)}(B)$ is $\geqq B^n$ and $\leqq (2B + 1)^n$. We seek a value of B such that there will be two distinct elements X, Y in $\mathbf{Z}^{(n)}(B)$ having the same image, $L(X) = L(Y)$. For this, it will suffice that $B^n > (2nBA)^r$, and thus it will suffice that

$$B = (2nA)^{r/(n-r)}.$$

We take $X - Y$ as the solution of our problem.

Let K be a finite extension of $\mathbf{Q}$, and let I_K be the integral closure of $\mathbf{Z}$ in K. From Exercise 5 of Chapter IX, we know that I_K is a free module over $\mathbf{Z}$, of dimension $[K : \mathbf{Q}]$. We view K as contained in the complex numbers. If

$\alpha \in K$, a conjugate of α will be taken to be an element $\sigma\alpha$, where σ is an embedding of K in $\mathbf{C}$. By the **size** of a set of elements of K we shall mean the maximum of the absolute values of all conjugates of these elements.

By the size of a vector $X = (x_1, \ldots, x_n)$ we shall mean the size of the set of its coordinates.

Let $\omega_1, \ldots, \omega_M$ be a basis of I_K over $\mathbf{Z}$. Let $\alpha \in I_K$, and write

$$\alpha = a_1\omega_1 + \cdots + a_M\omega_M.$$

Let $\omega'_1, \ldots, \omega'_M$ be the dual basis of $\omega_1, \ldots, \omega_M$ with respect to the trace. Then we can express the (Fourier) coefficients a_j of α as a trace,

$$a_j = \mathrm{Tr}(\alpha\omega'_j).$$

The trace is a sum over the conjugates. Hence the size of these coefficients is bounded by the size of α, times a fixed constant, depending on the size of the elements ω'_j.

Lemma 2. *Let K be a finite extension of $\mathbf{Q}$. Let*

$$\begin{aligned} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n &= 0 \\ &\cdots \\ \alpha_{r1}x_1 + \cdots + \alpha_{rn}x_n &= 0 \end{aligned}$$

be a system of linear equations with coefficients in I_K, and $n > r$. Let A be a number such that $\mathrm{size}(\alpha_{ij}) \leqq A$, for all i, j. Then there exists a non-trivial solution X in I_K such that

$$\mathrm{size}(X) \leqq C_1(C_2nA)^{r/(n-r)},$$

where C_1, C_2 are constants depending only on K.

Proof. Let $\omega_1, \ldots, \omega_M$ be a basis of I_K over $\mathbf{Z}$. Each x_j can be written

$$x_j = \xi_{j1}\omega_1 + \cdots + \xi_{jM}\omega_M$$

with unknowns $\xi_{j\lambda}$. Each α_{ij} can be written

$$\alpha_{ij} = a_{ij1}\omega_1 + \cdots + a_{ijM}\omega_M$$

with integers $a_{ij\lambda} \in \mathbf{Z}$. If we multiply out the $\alpha_{ij}x_j$, we find that our linear equations with coefficients in I_K are equivalent to a system of rM linear equations in the nM unknowns $\xi_{j\lambda}$, with coefficients in $\mathbf{Z}$, whose size is bounded by CA, where C is a number depending only on M and the size of the elements ω_λ, together with the products $\omega_\lambda\omega_\mu$, in other words where C depends only on K. Applying Lemma 1, we obtain a solution in terms of the $\xi_{j\lambda}$, and hence a solution X in I_K, whose size satisfies the desired bound.

The next lemma has to do with estimates of derivatives. By the size of a polynomial with coefficients in K, we shall mean the size of its set of coefficients. A **denominator** for a set of elements of K will be any positive rational integer whose product with every element of the set is an algebraic integer. We define in a similar way a denominator for a polynomial with coefficients in K. We abbreviate "denominator" by den.

Let

$$P(T_1, \ldots, T_N) = \sum \alpha_{(\nu)} M_{(\nu)}(T)$$

be a polynomial with complex coefficients, and let

$$Q(T_1, \ldots, T_N) = \sum \beta_{(\nu)} M_{(\nu)}(T)$$

be a polynomial with real coefficients $\geqq 0$. We say that Q **dominates** P if $|\alpha_{(\nu)}| \leq \beta_{(\nu)}$ for all (ν). It is then immediately verified that the relation of dominance is preserved under addition, multiplication, and taking partial derivatives with respect to the variables $T_1, \ldots, T_N$.

Lemma 3. *Let K be of finite degree over $\mathbf{Q}$. Let $f_1, \ldots, f_N$ be functions, holomorphic on a neighborhood of a point $w \in \mathbf{C}$, and assume that $D = d/dz$ maps the ring $K[f_1, \ldots, f_N]$ into itself. Assume that $f_i(w) \in K$ for all i. Then there exists a number C_1 having the following property. Let $P(T_1, \ldots, T_N)$ be a polynomial with coefficients in K, of degree $\leqq r$. If we set $f = P(f_1, \ldots, f_N)$, then we have, for all positive integers k,*

$$\text{size}(D^k f(w)) \leqq \text{size}(P) r^k k! C_1^{k+r}$$

Furthermore, there is a denominator for $D^k f(w)$ bounded by $\text{den}(P) C_1^{k+r}$.

Proof. There exist polynomials $P_i(T_1, \ldots, T_N)$ with coefficients in K such that

$$Df_i = P_i(f_1, \ldots, f_N).$$

Let h be the maximum of their degrees. There exists a unique derivation $\bar{D}$ on $K[T_1, \ldots, T_N]$ such that $\bar{D}T_i = P_i(T_1, \ldots, T_N)$. For any polynomial P we have

$$\bar{D}(P(T_1, \ldots, T_N)) = \sum_{i=1}^{N} (D_i P)(T_1, \ldots, T_N) \cdot P_i(T_1, \ldots, T_N),$$

where $D_1, \ldots, D_N$ are the partial derivatives. The polynomial P is dominated by

$$\text{size}(P)(1 + T_1 + \cdots + T_N)^r,$$

and each P_i is dominated by $\text{size}(P_i)(1 + T_1 + \cdots + T_N)^h$. Thus $\bar{D}P$ is dominated by

$$\text{size}(P) C_2 r (1 + T_1 + \cdots + T_N)^{r+h}.$$

Proceeding inductively, one sees that $\bar{D}^k P$ is dominated by

$$\text{size}(P)C_3^k r^k k!(1 + T_1 \cdots + T_N)^{r+kh}.$$

Substituting values $f_i(w)$ for T_i, we obtain the desired bound on $D^k f(w)$. The second assertion concerning denominators is proved also by a trivial induction.

We now come to the main part of the proof of our theorem. Let f, g be two functions among $f_1, \ldots, f_N$ which are algebraically independent over K. Let r be a positive integer divisible by $2m$. We shall let r tend to infinity at the end of the proof.

Let

$$F = \sum_{i,j=1}^{r} b_{ij} f^i g^j$$

have coefficients b_{ij} in K. Let $n = r^2/2m$. We can select the b_{ij} not all equal to 0, and such that

$$D^k F(w_v) = 0$$

for $0 \leqq k < n$ and $v = 1, \ldots, m$. Indeed, we have to solve a system of mn linear equations in $r^2 = 2mn$ unknowns. Note that

$$\frac{mn}{2mn - mn} = 1.$$

We multiply these equations by a denominator for the coefficients. Using the estimate of Lemma 3, and Lemma 2, we can in fact take the b_{ij} to be algebraic integers, whose size is bounded by

$$O(r^n n! C_1^{n+r}) \leqq O(n^{2n})$$

for $n \to \infty$.

Since f, g are algebraically independent over K, our function F is not identically zero. We let s be the smallest integer such that all derivatives of F up to order $s - 1$ vanish at all points $w_1, \ldots, w_m$, but such that $D^s F$ does not vanish at one of the w, say w_1. Then $s \geqq n$. We let

$$\gamma = D^s F(w_1) \neq 0.$$

Then γ is an element of K, and by Lemma 3, it has a denominator which is bounded by $O(C_1^s)$ for $s \to \infty$. Let c be this denominator. The norm of $c\gamma$ from K to $\mathbf{Q}$ is then a non-zero rational integer. Each conjugate of $c\gamma$ is bounded by $O(s^{5s})$. Consequently, we get

$$1 \leqq |N_{\mathbf{Q}}^K(c\gamma)| \leqq O(s^{5s})^{[K:\mathbf{Q}]-1} |\gamma|, \tag{1}$$

where $|\gamma|$ is the fixed absolute value of γ, which will now be estimated very well by global arguments.

Let θ be an entire function of order $\leqq \rho$, such that θf and θg are entire, and $\theta(w_1) \neq 0$. Then $\theta^{2r}F$ is entire. We consider the entire function

$$H(z) = \frac{\theta(z)^{2r}F(z)}{\prod_{v=1}^{m}(z - w_v)^s}.$$

Then $H(w_1)$ differs from $D^sF(w_1)$ by obvious factors, bounded by $C_4^s s!$. By the maximum modulus principle, its absolute value is bounded by the maximum of H on a large circle of radius R. If we take R large, then $z - w_v$ has approximately the same absolute value as R, and consequently, on the circle of radius R, $H(z)$ is bounded in absolute value by an expression of type

$$\frac{s^{3s}C_5^{2rR^\rho}}{R^{ms}}.$$

We select $R = s^{1/2\rho}$. We then get the estimate

$$|\gamma| \leqq \frac{s^{4s}C_6^s}{s^{ms/2\rho}}.$$

We now let r tend to infinity. Then both n and s tend to infinity. Combining this last inequality with inequality (1), we obtain the desired bound on m. This concludes the proof.

Of course, we made no effort to be especially careful in the powers of s occurring in the estimates, and the number 10 can obviously be decreased by exercising a little more care in the estimates.

The theorem we proved is only the simplest in an extensive theory dealing with problems of transcendence degree. In some sense, the theorem is best possible without additional hypotheses. For instance, if $P(t)$ is a polynomial with integer coefficients, then $e^{P(t)}$ will take the value 1 at all roots of P, these being algebraic. Furthermore, the functions

$$t, e^t, e^{t^2}, \ldots, e^{t^n}$$

are algebraically independent, but take on values in $\mathbf{Q}(e)$ for all integral values of t.

However, one expects rather strong results of algebraic independence to hold. Lindemann proved that *if* $\alpha_1, \ldots, \alpha_n$ *are algebraic numbers, linearly independent over* $\mathbf{Q}$, *then*

$$e^{\alpha_1}, \ldots, e^{\alpha_n}$$

are algebraically independent.

More generally, Schanuel has made the following conjecture: If $\alpha_1, \ldots, \alpha_n$ are complex numbers, linearly independent over $\mathbf{Q}$, then the transcendence degree of

$$\alpha_1, \ldots, \alpha_n, e^{\alpha_1}, \ldots, e^{\alpha_n}$$

should be $\geqq n$.

From this one would deduce at once the algebraic independence of e and π (looking at $1, 2\pi i, e, e^{2\pi i}$), and all other independence statements concerning the ordinary exponential function and logarithm which one feels to be true, for instance, the statement that π cannot lie in the field obtained by starting with the algebraic numbers, adjoining values of the exponential function, taking algebraic closure, and iterating these two operations. Such statements have to do with values of the exponential function lying in certain fields of transcendence degree $< n$, and one hopes that by a suitable deepening of Theorem 1, one will reach the desired results.

APPENDIX 2

Some Set Theory

§1. DENUMERABLE SETS

Let n be a positive integer. Let J_n be the set consisting of all integers k, $1 \leqq k \leqq n$. If S is a set, we say that S has n elements if there is a bijection between S and J_n. Such a bijection associates with each integer k as above an element of S, say $k \mapsto a_k$. Thus we may use J_n to "count" S. Part of what we assume about the basic facts concerning positive integers is that if S has n elements, then the integer n is uniquely determined by S.

One also agrees to say that a set has 0 elements if the set is empty.

We shall say that a set S is **denumerable** if there exists a bijection of S with the set of positive integers $\mathbf{Z}^+$. Such a bijection is then said to **enumerate** the set S. It is a mapping

$$n \mapsto a_n$$

which to each positive integer n associates an element of S, the mapping being injective and surjective.

If D is a denumerable set, and $f : S \to D$ is a bijection of some set S with D, then S is also denumerable. Indeed, there is a bijection $g : D \to \mathbf{Z}^+$, and hence $g \circ f$ is a bijection of S with $\mathbf{Z}^+$.

Let T be a set. A **sequence** of elements of T is simply a mapping of $\mathbf{Z}^+$ into T. If the map is given by the association $n \mapsto x_n$, we also write the sequence as $\{x_n\}_{n \geqq 1}$, or also $\{x_1, x_2, \ldots\}$. For simplicity, we also write $\{x_n\}$ for the sequence. Thus we think of the sequence as prescribing a first, second, ..., n-th element of T. We use the same braces for sequences as for sets, but the context will always make our meaning clear.

Examples. The even positive integers may be viewed as a sequence $\{x_n\}$ if we put $x_n = 2n$ for $n = 1, 2, \ldots$. The odd positive integers may also be viewed as a sequence $\{y_n\}$ if we put $y_n = 2n - 1$ for $n = 1, 2, \ldots$. In each case, the sequence gives an enumeration of the given set.

We also use the word *sequence* for mappings of the natural numbers into a set, thus allowing our sequences to start from 0 instead of 1. If we need to specify whether a sequence starts with the 0-th term or the first term, we write

$$\{x_n\}_{n \geqq 0} \quad \text{or} \quad \{x_n\}_{n \geqq 1}$$

according to the desired case. Unless otherwise specified, however, we always assume that a sequence will start with the first term. Note that from a sequence $\{x_n\}_{n \geqq 0}$ we can define a new sequence by letting $y_n = x_{n-1}$ for $n \geqq 1$. Then $y_1 = x_0, y_2 = x_1, \ldots$. Thus there is no essential difference between the two kinds of sequences.

Given a sequence $\{x_n\}$, we call x_n the n-th term of the sequence. A sequence may very well be such that all its terms are equal. For instance, if we let $x_n = 1$ for all $n \geqq 1$, we obtain the sequence $\{1, 1, 1, \ldots\}$. Thus there is a difference between a sequence of elements in a set T, and a subset of T. In the example just given, the set of all terms of the sequence consists of one element, namely the single number 1.

Let $\{x_1, x_2, \ldots\}$ be a sequence in a set S. By a **subsequence** we shall mean a sequence $\{x_{n_1}, x_{n_2}, \ldots\}$ such that $n_1 < n_2 < \cdots$. For instance, if $\{x_n\}$ is the sequence of positive integers, $x_n = n$, the sequence of even positive integers $\{x_{2n}\}$ is a subsequence.

An enumeration of a set S is of course a sequence in S.

A set is **finite** if the set is empty, or if the set has n elements for some positive integer n. If a set is not finite, it is called **infinite**.

Occasionally, a map of J_n into a set T will be called a **finite sequence** in T. A finite sequence is written as usual,

$$\{x_1, \ldots, x_n\} \quad \text{or} \quad \{x_i\}_{i=1,\ldots,n}.$$

When we need to specify the distinction between finite sequences and maps of $\mathbf{Z}^+$ into T, we call the latter infinite sequences. Unless otherwise specified, we shall use the word sequence to mean infinite sequence.

Proposition 1.1. *Let D be an infinite subset of $\mathbf{Z}^+$. Then D is denumerable, and in fact there is a unique enumeration of D, say $\{k_1, k_2, \ldots\}$ such that*

$$k_1 < k_2 < \cdots < k_n < k_{n+1} < \cdots.$$

Proof. We let k_1 be the smallest element of D. Suppose inductively that we have defined $k_1 < \cdots < k_n$, in such a way that any element k in D which is not equal to $k_1, \ldots, k_n$ is $> k_n$. We define k_{n+1} to be the smallest element of D which is $> k_n$. Then the map $n \mapsto k_n$ is the desired enumeration of D.

Corollary 1.2. *Let S be a denumerable set and D an infinite subset of S. Then D is denumerable.*

Proof. Given an enumeration of S, the subset D corresponds to a subset of $\mathbf{Z}^+$ in this enumeration. Using Proposition 1.1, we conclude that we can enumerate D.

Proposition 1.3. *Every infinite set contains a denumerable subset.*

Proof. Let S be an infinite set. For every non-empty subset T of S, we select a definite element a_T in T. We then proceed by induction. We let x_1 be the chosen element a_S. Suppose that we have chosen $x_1, \ldots, x_n$ having the property that for each $k = 2, \ldots, n$ the element x_k is the selected element in the subset which is the complement of $\{x_1, \ldots, x_{k-1}\}$. We let x_{n+1} be the selected element in the complement of the set $\{x_1, \ldots, x_n\}$. By induction, we thus obtain an association $n \mapsto x_n$ for all positive integers n, and since $x_n \neq x_k$ for all $k < n$ it follows that our association is injective, i.e. gives an enumeration of a subset of S.

Proposition 1.4. *Let D be a denumerable set, and $f : D \to S$ a surjective mapping. Then S is denumerable or finite.*

Proof. For each $y \in S$, there exists an element $x_y \in D$ such that $f(x_y) = y$ because f is surjective. The association $y \mapsto x_y$ is an injective mapping of S into D, because if

$$y, z \in S \quad \text{and} \quad x_y = x_z$$

then

$$y = f(x_y) = f(x_z) = z.$$

Let $g(y) = x_y$. The image of g is a subset of D and D is denumerable. Since g is a bijection between S and its image, it follows that S is denumerable or finite.

Proposition 1.5. *Let D be a denumerable set. Then $D \times D$ (the set of all pairs (x, y) with $x, y \in D$) is denumerable.*

Proof. There is a bijection between $D \times D$ and $\mathbf{Z}^+ \times \mathbf{Z}^+$, so it will suffice to prove that $\mathbf{Z}^+ \times \mathbf{Z}^+$ is denumerable. Consider the mapping of $\mathbf{Z}^+ \times \mathbf{Z}^+ \to \mathbf{Z}^+$ given by

$$(m, n) \mapsto 2^n 3^m.$$

It is injective, and by Proposition 1.1, our result follows.

Proposition 1.6. *Let $\{D_1, D_2, \ldots\}$ be a sequence of denumerable sets. Let S be the union of all sets D_i $(i = 1, 2, \ldots)$. Then S is denumerable.*

Proof. For each $i = 1, 2, \ldots$ we enumerate the elements of D_i, as indicated in the following notation:

$$\begin{aligned} D_1&: \{x_{11}, x_{12}, x_{13}, \ldots\} \\ D_2&: \{x_{21}, x_{22}, x_{23}, \ldots\} \\ &\cdots \\ D_i&: \{x_{i1}, x_{i2}, x_{i3}, \ldots\} \\ &\cdots \end{aligned}$$

The map $f: \mathbf{Z}^+ \times \mathbf{Z}^+ \to D$ given by

$$f(i, j) = x_{ij}$$

is then a surjective map of $\mathbf{Z}^+ \times \mathbf{Z}^+$ onto S. By Proposition 1.4, it follows that S is denumerable.

Corollary 1.7. *Let F be a non-empty finite set and D a denumerable set. Then $F \times D$ is denumerable. If $S_1, S_2, \ldots$ are a sequence of sets, each of which is finite or denumerable, then the union $S_1 \cup S_2 \cup \cdots$ is denumerable or finite.*

Proof. There is an injection of F into $\mathbf{Z}^+$ and a bijection of D with $\mathbf{Z}^+$. Hence there is an injection of $F \times \mathbf{Z}^+$ into $\mathbf{Z}^+ \times \mathbf{Z}^+$ and we can apply Corollary 1.2 and Proposition 1.6 to prove the first statement. One could also define a surjective map of $\mathbf{Z}^+ \times \mathbf{Z}^+$ onto $F \times D$. (Cf. Exercises 1 and 4.) As for the second statement, each finite set is contained in some denumerable set, so that the second statement follows from Proposition 1.1 and 1.6.

For convenience, we shall say that a set is **countable** if it is either finite or denumerable.

§2. ZORN'S LEMMA

In order to deal efficiently with infinitely many sets simultaneously, one needs a special property. To state it, we need some more terminology.

Let S be a set. An **ordering** (also called partial ordering) of S is a relation, written $x \leqq y$, among some pairs of elements of S, having the following properties.

ORD 1. *We have $x \leqq x$.*

ORD 2. *If $x \leqq y$ and $y \leqq z$ then $x \leqq z$.*

ORD 3. *If $x \leqq y$ and $y \leqq x$ then $x = y$.*

We sometimes write $y \geqq x$ for $x \leqq y$. Note that we don't require that the relation $x \leqq y$ or $y \leqq x$ hold for every pair of elements (x, y) of S. Some pairs may not be comparable. If the ordering satisfies this additional property, then we say that it is a **total ordering**.

Example 1. Let G be a group. Let S be the set of subgroups. If H, H' are subgroups of G, we define

$$H \leqq H'$$

if H is a subgroup of H'. One verifies immediately that this relation defines an ordering on S. Given two subgroups H, H' of G, we do not necessarily have $H \leqq H'$ or $H' \leqq H$.

Example 2. Let R be a ring, and let S be the set of left ideals of R. We define an ordering in S in a way similar to the above, namely if L, L' are left ideals of R, we define

$$L \leqq L'$$

if $L \subset L'$.

Example 3. Let X be a set, and S the set of subsets of X. If Y, Z are subsets of X, we define $Y \leqq Z$ if Y is a subset of Z. This defines an ordering on S.

In all these examples, the relation of ordering is said to be that of inclusion.

In an ordered set, if $x \leqq y$ and $x \neq y$ we then write $x < y$.

Let A be an ordered set, and B a subset. Then we can define an ordering on B by defining $x \leqq y$ for $x, y \in B$ to hold if and only if $x \leqq y$ in A. We shall say that R_0 is the ordering on B **induced** by R, or is the **restriction** to B of the partial ordering of A.

Let S be an ordered set. By a **least** element of S (or a **smallest** element) one means an element $a \in S$ such that $a \leqq x$ for all $x \in S$. Similarly, by a **greatest element** one means an element b such that $x \leqq b$ for all $x \in S$.

By a **maximal element** m of S one means an element such that if $x \in S$ and $x \geqq m$, then $x = m$. Note that a maximal element need not be a greatest element. There may be many maximal elements in S, whereas if a greatest element exists, then it is unique (proof ?).

Let S be an ordered set. We shall say that S is **totally ordered** if given $x, y \in S$ we have necessarily $x \leqq y$ or $y \leqq x$.

Example 4. The integers $\mathbf{Z}$ are totally ordered by the usual ordering. So are the real numbers.

Let S be an ordered set, and T a subset. An **upper bound** of T (in S) is an element $b \in S$ such that $x \leqq b$ for all $x \in T$. A **least upper bound** of T in S is an upper bound b such that, if c is another upper bound, then $b \leqq c$. We shall say

that S is **inductively ordered** if every non-empty totally ordered subset has an upper bound.

We shall say that S is **strictly** inductively ordered if every non-empty totally ordered subset has a least upper bound.

In Examples 1, 2, 3, in each case, the set is strictly inductively ordered. To prove this, let us take Example 1. Let T be a non-empty totally ordered subset of the set of subgroups of G. This means that if $H, H' \in T$, then $H \subset H'$ or $H' \subset H$. Let U be the union of all sets in T. Then:

1. U is a subgroup. *Proof*: If $x, y \in U$, there exist subgroups $H, H' \in T$ such that $x \in H$ and $y \in H'$. If, say, $H \subset H'$, then both $x, y \in H'$ and hence $xy \in H'$. Hence $xy \in U$. Also, $x^{-1} \in H'$, so $x^{-1} \in U$. Hence U is a subgroup.
2. U is an upper bound for each element of T. *Proof*: Every $H \in T$ is contained in U, so $H \leqq U$ for all $H \in T$.
3. U is a least upper bound for T. *Proof*: Any subgroup of G which contains all the subgroups $H \in T$ must then contain their union U.

The proof that the sets in Examples 2, 3 are strictly inductively ordered is entirely similar.

We can now state the property mentioned at the beginning of the section.

Zorn's Lemma. *Let S be a non-empty inductively ordered set. Then there exists a maximal element in S.*

As an example of Zorn's lemma, we shall now prove the infinite version of a theorem given in Chapters 1, §7, and XIV, §2, namely:

Let R be an entire, principal ring and let E be a free module over R. Let F be a submodule. Then F is free. In fact, if $\{v_i\}_{i \in I}$ is a basis for E, and $F \neq \{0\}$, then there exists a basis for F indexed by a subset of I.

Proof. For each subset J of I we let E_J be the free submodule of E generated by all $v_j, j \in J$, and we let $F_J = E_J \cap F$. We let S be the set of all pairs (F_J, w) where J is a subset of I, and $w : J' \to F_J$ is a basis of F_J indexed by a subset J' of J. We write w_j instead of $w(j)$ for $j \in J'$. If (F_J, w) and (F_K, u) are such pairs, we define $(F_J, w) \leqq (F_K, u)$ if $J \subset K$, if $J' \subset K'$, and if the restriction of u to J' is equal to w. (In other words, the basis u for F_K is an extension of the basis w for F_J.) This defines an ordering on S, and it is immediately verified that S is in fact inductively ordered, and non-empty (say by the finite case of the result). We can therefore apply Zorn's lemma. Let (F_J, w) be a maximal element. We contend that $J = I$ (this will prove our result). Suppose $J \neq I$ and let $k \in I$ but $k \notin J$. Let $K = J \cup \{k\}$. If

$$E_{J \cup \{k\}} \cap F = F_J,$$

then (F_K, w) is a bigger pair than (F_J, w) contradicting the maximality assumption. Qtherwise there exist elements of F_K which can be written in the form

$$cv_k + y$$

with some $y \in E_J$ and $c \in R$, $c \neq 0$. The set of all elements $c \in R$ such that there exists $y \in E_J$ for which $cv_k + y \in F$ is an ideal. Let a be a generator of this ideal, and let

$$w_k = av_k + y$$

be an element of F, with $y \in E_J$. If $z \in F_K$ then there exists $b \in R$ such that $z - bw_k \in E_J$. But $z - bw_k \in F$, whence $z - bw_k \in F_J$. It follows at once that the family consisting of w_j $(j \in J)$ and w_k is a basis for F_K, thus contradicting the maximality again. This proves what we wanted.

Zorn's lemma could be just taken as an axiom of set theory. However, it is not psychologically completely satisfactory as an axiom, because its statement is too involved, and one does not visualize easily the existence of the maximal element asserted in that statement. We show how one can prove Zorn's lemma from other properties of sets which everyone would immediately grant as acceptable psychologically.

From now on to the end of the proof of Theorem 2.1, we let A be a non-empty partially ordered and strictly inductively ordered set. We recall that **strictly inductively ordered** means that every nonempty totally ordered subset has a least upper bound. We assume given a map $f : A \to A$ such that for all $x \in A$ we have $x \leqq f(x)$. We could call such a map an **increasing** map.

Let $a \in A$. Let B be a subset of A. We shall say that B is **admissible** if:

1. B contains a.
2. We have $f(B) \subset B$.
3. Whenever T is a non-empty totally ordered subset of B, the least upper bound of T in A lies in B.

Then B is also strictly inductively ordered, by the induced ordering of A. We shall prove:

Theorem 2.1. (Bourbaki). *Let A be a non-empty partially ordered and strictly inductively ordered set. Let $f : A \to A$ be an increasing mapping. Then there exists an element $x_0 \in A$ such that $f(x_0) = x_0$.*

Proof. Suppose that A were totally ordered. By assumption, it would have a least upper bound $b \in A$, and then

$$b \leqq f(b) \leqq b,$$

so that in this case, our theorem is clear. The whole problem is to reduce the theorem to that case. In other words, what we need to find is a totally ordered admissible subset of A.

If we throw out of A all elements $x \in A$ such that x is not $\geqq a$, then what remains is obviously an admissible subset. Thus without loss of generality, we may assume that A has a least element a, that is $a \leqq x$ for all $x \in A$.

Let M be the intersection of all admissible subsets of A. Note that A itself is an admissible subset, and that all admissible subsets of A contain a, so that M is not empty. Furthermore, M is itself an admissible subset of A. To see this, let $x \in M$. Then x is in every admissible subset, so $f(x)$ is also in every admissible subset, and hence $f(x) \in M$. Hence $f(M) \subset M$. If T is a totally ordered non-empty subset of M, and b is the least upper bound of T in A, then b lies in every admissible subset of A, and hence lies in M. It follows that M is the smallest admissible subset of A, and that any admissible subset of A contained in M is equal to M.

We shall prove that M is totally ordered, and thereby prove Theorem 2.1.

[First we make some remarks which don't belong to the proof, but will help in the understanding of the subsequent lemmas. Since $a \in M$, we see that $f(a) \in M$, $f \circ f(a) \in M$, and in general $f^n(a) \in M$. Furthermore,

$$a \leqq f(a) \leqq f^2(a) \leqq \cdots.$$

If we had an equality somewhere, we would be finished, so we may assume that the inequalities hold. Let D_0 be the totally ordered set $\{f^n(a)\}_{n \geqq 0}$. Then D_0 looks like this:

$$a < f(a) < f^2(a) < \cdots < f^n(a) < \cdots.$$

Let a_1 be the least upper bound of D_0. Then we can form

$$a_1 < f(a_1) < f^2(a_1) < \cdots$$

in the same way to obtain D_1, and we can continue this process, to obtain

$$D_1, D_2, \ldots.$$

It is clear that $D_1, D_2, \ldots$ are contained in M. If we had a precise way of expressing the fact that we can establish a never-ending string of such denumerable sets, then we would obtain what we want. The point is that we are now trying to prove Zorn's lemma, which is the natural tool for guaranteeing the existence of such a string. However, given such a string, we observe that its elements have two properties: If c is an element of such a string and $x < c$, then $f(x) \leqq c$. Furthermore, there is no element between c and $f(c)$, that is if x is an element of the string, then $x \leqq c$ or $f(c) \leqq x$. We shall now prove two lemmas which show that elements of M have these properties.]

Let $c \in M$. We shall say that c is an **extreme point** of M if whenever $x \in M$ and $x < c$, then $f(x) \leqq c$. For each extreme point $c \in M$ we let

$$M_c = \text{set of } x \in M \quad \text{such that} \quad x \leqq c \quad \text{or} \quad f(c) \leqq x.$$

Note that M_c is not empty because a is in it.

Lemma 2.2. *We have $M_c = M$ for every extreme point c of M.*

Proof. It will suffice to prove that M_c is an admissible subset. Let $x \in M_c$. If $x < c$ then $f(x) \leqq c$ so $f(x) \in M_c$. If $x = c$ then $f(x) = f(c)$ is again in M_c. If $f(c) \leqq x$, then $f(c) \leqq x \leqq f(x)$, so once more $f(x) \in M_c$. Thus we have proved that $f(M_c) \subset M_c$.

Let T be a totally ordered subset of M_c and let b be the least upper bound of T in M. If all elements $x \in T$ are $\leqq c$, then $b \leqq c$ and $b \in M_c$. If some $x \in T$ is such that $f(c) \leqq x$, then $f(c) \leqq x \leqq b$, and so b is in M_c. This proves our lemma.

Lemma 2.3. *Every element of M is an extreme point.*

Proof. Let E be the set of extreme points of M. Then E is not empty because $a \in E$. It will suffice to prove that E is an admissible subset. We first prove that f maps E into itself. Let $c \in E$. Let $x \in M$ and suppose $x < f(c)$. We must prove that $f(x) \leqq f(c)$. By Lemma 2.2, $M = M_c$, and hence we have $x < c$, or $x = c$, or $f(c) \leqq x$. This last possibility cannot occur because $x < f(c)$. If $x < c$ then

$$f(x) \leqq c \leqq f(c).$$

If $x = c$ then $f(x) = f(c)$, and hence $f(E) \subset E$.

Next let T be a totally ordered subset of E. Let b be the least upper bound of T in M. We must prove that $b \in E$. Let $x \in M$ and $x < b$. If for all $c \in T$ we have $f(c) \leqq x$, then $c \leqq f(c) \leqq x$ implies that x is an upper bound for T, whence $b \leqq x$, which is impossible. Since $M_c = M$ for all $c \in E$, we must therefore have $x \leqq c$ for some $c \in T$. If $x < c$, then $f(x) \leqq c \leqq b$, and if $x = c$, then

$$c = x < b.$$

Since c is an extreme point and $M_c = M$, we get $f(x) \leqq b$. This proves that $b \in E$, that E is admissible, and thus proves Lemma 2.3.

We now see trivially that M is totally ordered. For let $x, y \in M$. Then x is an extreme point of M by Lemma 2, and $y \in M_x$ so $y \leqq x$ or

$$x \leqq f(x) \leqq y,$$

thereby proving that M is totally ordered. As remarked previously, this concludes the proof of Theorem 2.1.

We shall obtain Zorn's lemma essentially as a corollary of Theorem 2.1. We first obtain Zorn's lemma in a slightly weaker form.

Corollary 2.4. *Let A be a non-empty strictly inductively ordered set. Then A has a maximal element.*

Proof. Suppose that A does not have a maximal element. Then for each $x \in A$ there exists an element $y_x \in A$ such that $x < y_x$. Let $f: A \to A$ be the map such that $f(x) = y_x$ for all $x \in A$. Then A, f satisfy the hypotheses of Theorem 2.1 and applying Theorem 2.1 yields a contradiction.

The only difference between Corollary 2.4 and Zorn's lemma is that in Corollary 2.4, we assume that a non-empty totally ordered subset has a *least* upper bound, rather than an upper bound. It is, however, a simple matter to reduce Zorn's lemma to the seemingly weaker form of Corollary 2.4. We do this in the second corollary.

Corollary 2.5. (Zorn's lemma). *Let S be a non-empty inductively ordered set. Then S has a maximal element.*

Proof. Let A be the set of non-empty totally ordered subsets of S. Then A is not empty since any subset of S with one element belongs to A. If $X, Y \in A$, we define $X \leqq Y$ to mean $X \subset Y$. Then A is partially ordered, and is in fact strictly inductively ordered. For let $T = \{X_i\}_{i \in I}$ be a totally ordered subset of A. Let

$$Z = \bigcup_{i \in I} X_i.$$

Then Z is totally ordered. To see this, let $x, y \in Z$. Then $x \in X_i$ and $y \in X_j$ for some $i, j \in I$. Since T is totally ordered, say $X_i \subset X_j$. Then $x, y \in X_j$ and since X_j is totally ordered, $x \leqq y$ or $y \leqq x$. Thus Z is totally ordered, and is obviously a least upper bound for T in A. By Corollary 2.4, we conclude that A has a maximal element X_0. This means that X_0 is a maximal totally ordered subset of S (non-empty). Let m be an upper bound for X_0 in S. Then m is the desired maximal element of S. For if $x \in S$ and $m \leqq x$ then $X_0 \cup \{x\}$ is totally ordered, whence equal to X_0 by the maximality of X_0. Thus $x \in X_0$ and $x \leqq m$. Hence $x = m$, as was to be shown.

§3. CARDINAL NUMBERS

Let A, B be sets. We shall say that the **cardinality** *of A is the same as the cardinality of B*, and write

$$\text{card}(A) = \text{card}(B)$$

if there exists a bijection of A onto B.

We say $\text{card}(A) \leqq \text{card}(B)$ if there exists an injective mapping (injection) $f: A \to B$. We also write $\text{card}(B) \geqq \text{card}(A)$ in this case. It is clear that if $\text{card}(A) \leqq \text{card}(B)$ and $\text{card}(B) \leqq \text{card}(C)$, then $\text{card}(A) \leqq \text{card}(C)$.

This amounts to saying that a composite of injective mappings is injective. Similarly, if $\text{card}(A) = \text{card}(B)$ and $\text{card}(B) = \text{card}(C)$ then $\text{card}(A) = \text{card}(C)$.

This amounts to saying that a composite of bijective mappings is bijective. We clearly have $\text{card}(A) = \text{card}(A)$. Using Zorn's lemma, it is easy to show (see Exercise 14) that

$$\text{card}(A) \leqq \text{card}(B) \quad \text{or} \quad \text{card}(B) \leqq \text{card}(A).$$

Let $f: A \to B$ be a surjective map of a set A onto a set B. Then

$$\text{card}(B) \leqq \text{card}(A).$$

This is easily seen, because for each $y \in B$ there exists an element $x \in A$, denoted by x_y, such that $f(x_y) = y$. Then the association $y \mapsto x_y$ is an injective mapping of B into A, whence by definition, $\text{card}(B) \leqq \text{card}(A)$.

Given two nonempty sets A, B we have $\text{card}(A) \leqq \text{card}(B)$ *or* $\text{card}(B) \leqq \text{card}(A)$.

This is a simple application of Zorn's lemma. We consider the family of pairs (S, f) where S is a subset of A and $f: S \to B$ is an injective mapping. From the existence of a maximal element, the assertion follows at once.

Theorem 3.1. (Schroeder-Bernstein). *Let A, B be sets, and suppose that* $\text{card}(A) \leqq \text{card}(B)$, *and* $\text{card}(B) \leqq \text{card}(A)$. *Then*

$$\text{card}(A) = \text{card}(B).$$

Proof. Let

$$f: A \to B \quad \text{and} \quad g: B \to A$$

be injections. We separate A into two disjoint sets A_1 and A_2. We let A_1 consist of all $x \in A$ such that, when we lift back x by a succession of inverse maps,

$$x, \ g^{-1}(x), \quad f^{-1} \circ g^{-1}(x), \quad g^{-1} \circ f^{-1} \circ g^{-1}(x), \ldots$$

then at some stage we reach an element of A which cannot be lifted back to B by g. We let A_2 be the complement of A_1, in other words, the set of $x \in A$ which can be lifted back indefinitely, or such that we get stopped in B (i.e. reach an element of B which has no inverse image in A by f). Then $A = A_1 \cup A_2$. We shall define a bijection h of A onto B.

If $x \in A_1$, we define $h(x) = f(x)$.

If $x \in A_2$, we define $h(x) = g^{-1}(x) =$ unique element $y \in B$ such that $g(y) = x$.

Then trivially, h is injective. We must prove that h is surjective. Let $b \in B$. If, when we try to lift back b by a succession of maps

$$\cdots \circ f^{-1} \circ g^{-1} \circ f^{-1} \circ g^{-1} \circ f^{-1}(b)$$

we can lift back indefinitely, or if we get stopped in B, then $g(b)$ belongs to A_2 and consequently $b = h(g(b))$, so b lies in the image of h. On the other hand, if we cannot lift back b indefinitely, and get stopped in A, then $f^{-1}(b)$ is defined (i.e., b is in the image of f), and $f^{-1}(b)$ lies in A_1. In this case, $b = h(f^{-1}(b))$ is also in the image of h, as was to be shown.

Next we consider theorems concerning sums and products of cardinalities.

We shall reduce the study of cardinalities of products of arbitrary sets to the denumerable case, using Zorn's lemma. Note first that an infinite set A always contains a denumerable set. Indeed, since A is infinite, we can first select an element $a_1 \in A$, and the complement of $\{a_1\}$ is infinite. Inductively, if we have selected distinct elements $a_1, \ldots, a_n$ in A, the complement of $\{a_1, \ldots, a_n\}$ is infinite, and we can select a_{n+1} in this complement. In this way, we obtain a sequence of distinct elements of A, giving rise to a denumerable subset of A.

Let A be a set. By a **covering** of A one means a set Γ of subsets of A such that the union

$$\bigcup_{C \in \Gamma} C$$

of all the elements of Γ is equal to A. We shall say that Γ is a **disjoint covering** if whenever $C, C' \in \Gamma$, and $C \neq C'$, then the intersection of C and C' is empty.

Lemma 3.2. *Let A be an infinite set. Then there exists a disjoint covering of A by denumerable sets.*

Proof. Let S be the set whose elements are pairs (B, Γ) consisting of a subset B of A, and a disjoint covering of B by denumerable sets. Then S is not empty. Indeed, since A is infinite, A contains a denumerable set D, and the pair $(D, \{D\})$ is in S. If (B, Γ) and (B', Γ') are elements of S, we define

$$(B, \Gamma) \leqq (B', \Gamma')$$

to mean that $B \subset B'$, and $\Gamma \subset \Gamma'$. Let T be a totally ordered non-empty subset of S. We may write $T = \{(B_i, \Gamma_i)\}_{i \in I}$ for some indexing set I. Let

$$B = \bigcup_{i \in I} B_i \quad \text{and} \quad \Gamma = \bigcup_{i \in I} \Gamma_i.$$

If $C, C' \in \Gamma$, $C \neq C'$, then there exists some indices i, j such that $C \in \Gamma_i$ and $C' \in \Gamma_j$. Since T is totally ordered, we have, say,

$$(B_i, \Gamma_i) \leqq (B_j, \Gamma_j).$$

Hence in fact, C, C' are both elements of Γ_j, and hence C, C' have an empty intersection. On the other hand, if $x \in B$, then $x \in B_i$ for some i, and hence there is some $C \in \Gamma_i$ such that $x \in C$. Hence Γ is a disjoint covering of B. Since the

elements of each Γ_i are denumerable subsets of A, it follows that Γ is a disjoint covering of B by denumerable sets, so (B, Γ) is in S, and is obviously an upper bound for T. Therefore S is inductively ordered.

Let (M, Δ) be a maximal element of S, by Zorn's lemma. Suppose that $M \neq A$. If the complement of M in A is infinite, then there exists a denumerable set D contained in this complement. Then

$$(M \cup D, \Delta \cup \{D\})$$

is a bigger pair than (M, Δ), contradicting the maximality of (M, Δ). Hence the complement of M in A is a finite set F. Let D_0 be an element of Δ. Let

$$D_1 = D_0 \cup F.$$

Then D_1 is denumerable. Let Δ_1 be the set consisting of all elements of Δ, except D_0, together with D_1. Then Δ_1 is a disjoint covering of A by denumerable sets, as was to be shown.

Theorem 3.3. *Let A be an infinite set, and let D be a denumerable set. Then*

$$\text{card}(A \times D) = \text{card}(A).$$

Proof. By the lemma, we can write

$$A = \bigcup_{i \in I} D_i$$

as a disjoint union of denumerable sets. Then

$$A \times D = \bigcup_{i \in I} (D_i \times D).$$

For each $i \in I$, there is a bijection of $D_i \times D$ on D_i by Proposition 1.5. Since the sets $D_i \times D$ are disjoint, we get in this way a bijection of $A \times D$ on A, as desired.

Corollary 3.4. *If F is a finite non-empty set, then*

$$\text{card}(A \times F) = \text{card}(A).$$

Proof. We have

$$\text{card}(A) \leqq \text{card}(A \times F) \leqq \text{card}(A \times D) = \text{card}(A).$$

We can then use Theorem 3.1 to get what we want.

Corollary 3.5. *Let A, B be non-empty sets, A infinite, and suppose*

$$\text{card}(B) \leqq \text{card}(A).$$

Then

$$\text{card}(A \cup B) = \text{card}(A).$$

Proof. We can write $A \cup B = A \cup C$ for some subset C of B, such that C and A are disjoint. (We let C be the set of all elements of B which are not elements of A.) Then $\text{card}(C) \leqq \text{card}(A)$. We can then construct an injection of $A \cup C$ into the product

$$A \times \{1, 2\}$$

of A with a set consisting of 2 elements. Namely, we have a bijection of A with $A \times \{1\}$ in the obvious way, and also an injection of C into $A \times \{2\}$. Thus

$$\text{card}(A \cup C) \leqq \text{card}(A \times \{1, 2\}).$$

We conclude the proof by Corollary 3.4 and Theorem 3.1.

Theorem 3.6. *Let A be an infinite set. Then*

$$\text{card}(A \times A) = \text{card}(A).$$

Proof. Let S be the set consisting of pairs (B, f) where B is an infinite subset of A, and f is a bijection of B onto $B \times B$. Then S is not empty because if D is a denumerable subset of A, we can always find a bijection of D on $D \times D$. If (B, f) and (B', f') are in S, we define $(B, f) \leqq (B', f')$ to mean $B \subset B'$, and the restriction of f' to B is equal to f. Then S is partially ordered, and we contend that S is inductively ordered. Let T be a non-empty totally ordered subset of S, and say T consists of the pairs (B_i, f_i) for i in some indexing set I. Let

$$M = \bigcup_{i \in I} B_i.$$

We shall define a bijection $g : M \to M \times M$. If $x \in M$, then x lies in some B_i. We define $g(x) = f_i(x)$. This value $f_i(x)$ is independent of the choice of B_i in which x lies. Indeed, if $x \in B_j$ for some $j \in I$, then say

$$(B_i, f_i) \leqq (B_j, f_j).$$

By assumption, $B_i \subset B_j$, and $f_j(x) = f_i(x)$, so g is well defined. To show g is surjective, let $x, y \in M$ and $(x, y) \in M \times M$. Then $x \in B_i$ for some $i \in I$ and $y \in B_j$ for some $j \in I$. Again since T is totally ordered, say $(B_i, f_i) \leqq (B_j, f_j)$. Thus $B_i \subset B_j$, and $x, y \in B_j$. There exists an element $b \in B_j$ such that

$$f_j(b) = (x, y) \in B_j \times B_j.$$

By definition, $g(b) = (x, y)$, so g is surjective. We leave the proof that g is injective to the reader to conclude the proof that g is a bijection. We then see

that (M, g) is an upper bound for T in S, and therefore that S is inductively ordered.

Let (M, g) be a maximal element of S, and let C be the complement of M in A. If $\text{card}(C) \leqq \text{card}(M)$, then

$$\text{card}(A) = \text{card}(M \cup C) = \text{card}(M)$$

by Corollary 3.5, and hence $\text{card}(M) = \text{card}(A)$. Since $\text{card}(M) = \text{card}(M \times M)$, we are done with the proof in this case. If

$$\text{card}(M) \leqq \text{card}(C),$$

then there exists a subset M_1 of C having the same cardinality as M. We consider

$$\begin{aligned}(M \cup M_1) \times (M \cup M_1) \\ = (M \times M) \cup (M_1 \times M) \cup (M \times M_1) \cup (M_1 \times M_1).\end{aligned}$$

By the assumption on M and Corollary 3.5, the last three sets in parentheses on the right of this equation have the same cardinality as M. Thus

$$(M \cup M_1) \times (M \cup M_1) = (M \times M) \cup M_2$$

where M_2 is disjoint from $M \times M$, and has the same cardinality as M. We now define a bijection

$$g_1 : M \cup M_1 \to (M \cup M_1) \times (M \cup M_1).$$

We let $g_1(x) = g(x)$ if $x \in M$, and we let g_1 on M_1 be any bijection of M_1 on M_2. In this way we have extended g to $M \cup M_1$, and the pair $(M \cup M_1, g_1)$ is in S, contradicting the maximality of (M, g). The case $\text{card}(M) \leqq \text{card}(C)$ therefore cannot occur, and our theorem is proved (using Exercise 14 below).

Corollary 3.7. *If A is an infinite set, and $A^{(n)} = A \times \cdots \times A$ is the product taken n times, then*

$$\text{card}(A^{(n)}) = \text{card}(A).$$

Proof. Induction.

Corollary 3.8. *If $A_1, \ldots, A_n$ are non-empty sets with A_n infinite, and*

$$\text{card}(A_i) \leqq \text{card}(A_n)$$

for $i = 1, \ldots, n$, then

$$\text{card}(A_1 \times \cdots \times A_n) = \text{card}(A_n).$$

Proof. We have

$$\text{card}(A_n) \leqq \text{card}(A_1 \times \cdots \times A_n) \leqq \text{card}(A_n \times \cdots \times A_n)$$

and we use Corollary 3.7 and the Schroeder-Bernstein theorem to conclude the proof.

Corollary 3.9. *Let A be an infinite set, and let Φ be the set of finite subsets of A. Then*

$$\text{card}(\Phi) = \text{card}(A).$$

Proof. Let Φ_n be the set of subsets of A having exactly n elements, for each integer $n = 1, 2, \ldots$. We first show that $\text{card}(\Phi_n) \leqq \text{card}(A)$. If F is an element of Φ_n, we order the elements of F in any way, say

$$F = \{x_1, \ldots, x_n\}.$$

and we associate with F the element $(x_1, \ldots, x_n) \in A^{(n)}$,

$$F \mapsto (x_1, \ldots, x_n).$$

If G is another subset of A having n elements, say $G = \{y_1, \ldots, y_n\}$, and $G \neq F$, then

$$(x_1, \ldots, x_n) \neq (y_1, \ldots, y_n).$$

Hence our map

$$F \mapsto (x_1, \ldots, x_n)$$

of Φ_n into $A^{(n)}$ is injective. By Corollary 3.7, we conclude that

$$\text{card}(\Phi_n) \leqq \text{card}(A).$$

Now Φ is the disjoint union of the Φ_n for $n = 1, 2, \ldots$ and it is an exercise to show that $\text{card}(\Phi) \leqq \text{card}(A)$ (cf. Exercise 1). Since

$$\text{card}(A) \leqq \text{card}(\Phi),$$

because in particular, $\text{card}(\Phi_1) = \text{card}(A)$, we see that our corollary is proved.

In the next theorem, we shall see that given a set, there always exists another set whose cardinality is bigger.

Theorem 3.10. *Let A be an infinite set, and T the set consisting of two elements $\{0, 1\}$. Let M be the set of all maps of A into T. Then*

$$\text{card}(A) \leqq \text{card}(M) \quad \textit{and} \quad \text{card}(A) \neq \text{card}(M).$$

Proof. For each $x \in A$ we let

$$f_x : A \to \{0, 1\}$$

be the map such that $f_x(x) = 1$ and $f_x(y) = 0$ if $y \neq x$. Then $x \mapsto f_x$ is obviously an injection of A into M, so that $\text{card}(A) \leqq \text{card}(M)$. Suppose that

$$\text{card}(A) = \text{card}(M).$$

Let

$$x \mapsto g_x$$

be a bijection between A and M. We define a map $h : A \to \{0, 1\}$ by the rule

$$h(x) = 0 \quad \text{if} \quad g_x(x) = 1,$$

$$h(x) = 1 \quad \text{if} \quad g_x(x) = 0.$$

Then certainly $h \neq g_x$ for any x, and this contradicts the assumption that $x \mapsto g_x$ is a bijection, thereby proving Theorem 3.10.

Corollary 3.11. *Let A be an infinite set, and let S be the set of all subsets of A. Then* $\text{card}(A) \leqq \text{card}(S)$ *and* $\text{card}(A) \neq \text{card}(S)$.

Proof. We leave it as an exercise. [*Hint*: If B is a non-empty subset of A, use the characteristic function φ_B such that

$$\varphi_B(x) = 1 \quad \text{if} \quad x \in B,$$

$$\varphi_B(x) = 0 \quad \text{if} \quad x \notin B.$$

What can you say about the association $B \mapsto \varphi_B$?]

§4. WELL-ORDERING

An ordered set A is said to be **well-ordered** if it is totally ordered, and if every non-empty subset B has a least element, that is, an element $a \in B$ such that $a \leqq x$ for all $x \in B$.

Example 1. The set of positive integers $\mathbf{Z}^+$ is well-ordered. Any finite set can be well-ordered, and a denumerable set D can be well-ordered: Any bijection of D with $\mathbf{Z}^+$ will give rise to a well-ordering of D.

Example 2. Let S be a well-ordered set and let b be an element of some set, $b \notin S$. Let $A = S \cup \{b\}$. We define $x \leqq b$ for all $x \in S$. Then A is totally ordered, and is in fact well-ordered.

Proof. Let B be a non-empty subset of A. If B consists of b alone, then b is a least element of B. Otherwise, B contains some element $a \in A$. Then $B \cap A$ is not empty, and hence has a least element, which is obviously also a least element for B.

Theorem 4.1. *Every non-empty set can be well-ordered.*

Proof. Let A be a non-empty set. Let S be the set of all pairs (X, ω), where X is a subset of A and ω is a well-ordering of X. Note that S is not empty because any single element of A gives rise to such a pair. If (X, ω) and (X', ω') are such pairs, we define $(X, \omega) \leqq (X', \omega')$ if $X \subset X'$, if the ordering induced on X by ω' is equal to ω, and if X is an initial segment of X'. It is obvious that this defines an ordering on S, and we contend that S is inductively ordered. Let $\{(X_i, \omega_i)\}$ be a totally ordered non-empty subset of S. Let $X = \bigcup X_i$. If $a, b \in X$, then a, b lie in some X_i, and we define $a \leqq b$ in X if $a \leqq b$ with respect to the ordering ω_i. This is independent of the choice of i (immediate from the assumption of total ordering). In fact, X is well ordered, for if Y is a non-empty subset of X, then there is some element $y \in Y$ which lies in some X_j. Let c be a least element of $X_j \cap Y$. One verifies at once that c is a least element of Y. We can therefore apply Zorn's lemma. Let (X, ω) be a maximal element in S. If $X \neq A$, then, using Example 2, we can define a well-ordering on a bigger subset than X, contradicting the maximality assumption. This proves Theorem 4.1.

Note. Theorem 4.1 is an immediate and straightforward consequence of Zorn's lemma. Usually in mathematics, Zorn's lemma is the most efficient tool when dealing with infinite processes.

EXERCISES

1. Prove the statement made in the proof of Corollary 3.9.
2. If A is an infinite set, and Φ_n is the set of subsets of A having exactly n elements, show that

$$\text{card}(A) \leqq \text{card}(\Phi_n)$$

for $n \geqq 1$.

3. Let A_i be infinite sets for $i = 1, 2, \ldots$ and assume that

$$\text{card}(A_i) \leqq \text{card}(A)$$

for some set A, and all i. Show that

$$\text{card}\left(\bigcup_{i=1}^{\infty} A_i\right) \leqq \text{card}(A).$$

4. Let K be a subfield of the complex numbers. Show that for each integer $n \geqq 1$, the cardinality of the set of extensions of K of degree n in $\mathbf{C}$ is $\leqq \text{card}(K)$.

5. Let K be an infinite field, and E an algebraic extension of K. Show that

$$\text{card}(E) = \text{card}(K).$$

6. Finish the proof of the Corollary 3.11.

7. If A, B are sets, denote by $M(A, B)$ the set of all maps of A into B. If B, B' are sets with the same cardinality, show that $M(A, B)$ and $M(A, B')$ have the same cardinality. If A, A' have the same cardinality, show that $M(A, B)$ and $M(A', B)$ have the same cardinality.

8. Let A be an infinite set and abbreviate $\text{card}(A)$ by α. If B is an infinite set, abbreviate $\text{card}(B)$ by β. Define $\alpha\beta$ to be $\text{card}(A \times B)$. Let B' be a set disjoint from A such that $\text{card}(B) = \text{card}(B')$. Define $\alpha + \beta$ to be $\text{card}(A \cup B')$. Denote by B^A the set of all maps of A into B, and denote $\text{card}(B^A)$ by β^α. Let C be an infinite set and abbreviate $\text{card}(C)$ by γ. Prove the following statements:
   (a) $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.
   (b) $\alpha\beta = \beta\alpha$.
   (c) $\alpha^{\beta+\gamma} = \alpha^\beta \alpha^\gamma$.

9. Let K be an infinite field. Prove that there exists an algebraically closed field K^a containing K as a subfield, and algebraic over K. [*Hint*: Let Ω be a set of cardinality strictly greater than the cardinality of K, and containing K. Consider the set S of all pairs (E, φ) where E is a subset of Ω such that $K \subset E$, and φ denotes a law of addition and multiplication on E which makes E into a field such that K is a subfield, and E is algebraic over K. Define a partial ordering on S in an obvious way; show that S is inductively ordered, and that a maximal element is algebraic over K and algebraically closed. You will need Exercise 5 in the last step.]

10. Let K be an infinite field. Show that the field of rational functions $K(t)$ has the same cardinality as K.

11. Let J_n be the set of integers $\{1, \ldots, n\}$. Let $\mathbf{Z}^+$ be the set of positive integers. Show that the following sets have the same cardinality:
   (a) The set of all maps $M(\mathbf{Z}^+, J_n)$.
   (b) The set of all maps $M(\mathbf{Z}^+, J_2)$.
   (c) The set of all real numbers x such that $0 \leqq x < 1$.
   (d) The set of all real numbers.

12. Show that $M(\mathbf{Z}^+, \mathbf{Z}^+)$ has the same cardinality as the real numbers.

13. Let S be a non-empty set. Let S' denote the product S with itself taken denumerably many times. Prove that $(S')'$ has the same cardinality as S'. [Given a set S whose cardinality is strictly greater than the cardinality of $\mathbf{R}$, I do not know whether it is always true that $\text{card } S = \text{card } S'$.] Added 1994: The grapevine communicates to me that according to Solovay, the answer is "no."

14. Let A, B be non–empty sets. Prove that

$$\text{card}(A) \leqq \text{card}(B) \quad \text{or} \quad \text{card}(B) \leqq \text{card}(A).$$

[Hint: consider the family of pairs (C, f) where C is a subset of A and $f: C \to B$ is an injective map. By Zorn's lemma there is a maximal element. Now finish the proof].