The Theory of Irrationalities of the Third Degree

by B. N. DELONE
and D. K. FADDEEV

Volume 10

TRANSLATIONS OF
MATHEMATICAL
MONOGRAPHS

10.1090/mmono/010

Translations of Mathematical Monographs *Volume 10*

The Theory of

IRRATIONALITIES OF THE THIRD DEGREE

by

B. N. Delone and D. K. Faddeev

AMERICAN MATHEMATICAL SOCIETY
PROVIDENCE, RHODE ISLAND
1964

ТЕОРИЯ ИРРАЦИОНАЛЬНОСТЕЙ ТРЕТЬЕЙ СТЕПЕНИ

Б. Н. ДЕЛОНЕ и Д. К. ФАДДЕЕВ

Труды Математического Института
имени В. А. Стеклова XI.

Издательство Академии Наук СССР

Москва 1940 Ленинград

Translated from the Russian
by Emma Lehmer and Sue Ann Walker

Publication aided by a grant from the
NATIONAL SCIENCE FOUNDATION

Text composed on Photon, Partly subsidized by NSF Grant G21913

International Standard Book Number 0-8218-1560-1
Library of Congress Card Number 63-21548

AUTHOR'S FOREWORD TO THE TRANSLATION

The initiative for writing this book came from Edmund Landau, who suggested it to me in Göttingen in 1928. The book was begun in collaboration with the Norwegian mathematician Nagell, but we soon realized that it would be much more convenient for me to have a coauthor living in Leningrad, where I resided at that time, in the person of my pupil, a very gifted algebraist, Dimitriĭ Konstantinovič Faddeev. The whole book was conceived by me, in a sense, as a framework for my investigations into indeterminate equations of the third degree, as developed in Chapter VI.

After the works of Lagrange and Gauss in the theory of indeterminate equations of the second degree in two unknowns, which used to occupy an important part of any course in the theory of numbers, there appeared on the one hand the works of Dirichlet, and on the other those of Hermite. Dirichlet generalized the work of Lagrange and Gauss to equations of the form $\Phi = m$, where Φ is a form of the nth degree in n variables, decomposable into irrational linear factors, and m is a given number, while Hermite considered the equation $\Psi = m$, where Ψ is a quadratic form in n variables. An interesting result concerning such equations was also obtained by Dickson. As to indeterminate equations of the third and higher degree in two unknowns, only the remarkable method of Thue was applicable. This method was brought to its ultimate conclusion by Roth without introducing anything new in principle. However, for the effective solution of such equations one must apparently look for new and deeper methods.

Around 1920 there appeared two series of papers on the indeterminate equations of the third degree in two unknowns: my papers, which are given here in Chapter VI, having to do with integer solutions, and those of Mordell on a finite basis for rational solutions, which form the culmination of an idea of Poincaré.

At the present time the following two problems in the theory of indeterminate equations are unquestionably the ones to which attention should now be turned: 1) To test the effectiveness of the method for finding integer solutions of binary equations, possibly by developing further my algorithm of ascent and at the same time probably solving the problem of incomplete quotients. 2) To obtain effectively the Poincaré basis for rational solutions. The second problem is being

studied by some well-known modern mathematicians, such as Serre in Paris and Šafarevič in Moscow, but nobody has given any serious attention to the development of my method since I abandoned it in the 1930's. It should be pointed out that there is no algebraic irrational of degree higher than the second for which we know whether the partial quotients in its expansion in a continued fraction are bounded or not. However, an effective solution even of the equation $ax^3 + y^3 = \sigma$, with given integers a and σ, would apparently lead to the solution of this problem for $\sqrt[3]{a}$.

I hope that the publication of this book in the United States will serve as a stimulus to others to develop further the methods presented in Chapter VI.

B. Delone
Moscow,
July 30, 1962

Translators' note. The translators are deeply grateful to Professor Delone for his valuable assistance with certain details of the text and for his suggestions regarding the inclusion of recent or supplementary material.

TABLE OF CONTENTS

This page intentionally left blank

INTRODUCTION

A large part of the modern theory of algebraic numbers is concerned with problems whose simplest nontrivial examples can be found in the theory of quadratic irrationalities given by Gauss in his *Disquisitiones arithmeticae*. To this belong the theory of units, theory of ideals, laws of reciprocity and therefore, to a certain extent, class field theory.

A similar study of the theory of cubic irrationalities is interesting not only as the next case in complexity after the quadratic case, in which it is still possible to give solutions in terms of convenient algorithms, but primarily because it poses further problems that were so trivial in the quadratic case as to escape notice. To this belong in the first place the problem of classification of cubic irrationalities, the so-called inverse problem of the Galois theory for these irrationalities, and the problem of approximation of irrational numbers of higher degree by rationals, which is as yet not completely solved and which is closely connected with the problem of representing numbers by incomplete reducible forms (i.e., forms in which the number of variables is less than the degree). These two fundamental problems first appeared in a nontrivial way in the theory of cubic irrationalities, but they exist for irrationalities of any degree.

Until now there has appeared no monograph in the mathematical literature on the theory of cubic irrationalities. Our book fills this gap.

It is natural that this book should be published by the Academy of Sciences (U.S.S.R.), since many of the investigations in the theory of cubic irrationalities are due to mathematicians who are in one way or another connected with the Academy. They are: E. Zolotarev, A. Markov, G. Voronoĭ, myself, V. A. Tartakovskiĭ, D. K. Faddeev, E. A. Venkov, and O. K. Žitomirskiĭ. The most important contributions of foreign mathematicians to this subject are due to Eisenstein, Thue, Mordell, Nagell, A. Weil and C. L. Siegel, as well as to Dedekind and Hasse. The investigations of the last two mathematicians are not included in this monograph, since their methods are more in the nature of application of general class field theory to the special case of cubic fields.

One may hope that considerations similar to the ones discussed in Chapters I and III may lead to the construction of a theory, close to class field theory, but

which will give solutions to problems now solved by class field theory without the use of the analytical theory of numbers.

D. K. Faddeev and I are equal coauthors of this book, and about half of the material contained in it belongs to D. K. Faddeev. Usually, the plan for each section was discussed by us jointly in advance, and subsequently each of us looked over the sections written by the other. Thus §§7–9, 12, 19, 22–25, 34, 35, 42–59, 64, 70, 72–74, 79–82 were written by D. K. Faddeev, while §§1–6, 10, 11, 13–18, 20, 21, 26–33, 36–41, 60–63, 65–68, 71, 75–78 were written by me. We are indebted to V. A. Tartakovskiĭ for §69.

The plan and the idea of the book are mine, but as a result of the invaluable cooperation of D. K. Faddeev, who gave all his enthusiasm to this work, it became possible to realize a much wider program than was first envisaged, when I began to write this book with Nagell. Faddeev and I have developed especially for this book many of the results in the theory of cubic irrationalities which were not available among the known results. This is particularly true of much of the material in Chapters I and III.

I shall give a brief account of the contents of the various chapters.

Chapter I contains the most complete step-by-step geometric development of the theory of algebraic irrationalities of any degree, considered, on my suggestion, as a theory of multiplicative lattices in n-dimensional complex space K_n. It serves as an introduction to the whole book. Such lattices are somewhat more general than algebraic fields and are connected with their direct sums. They are needed in Chapter III for the solution of the inverse problem of Galois theory for fields of the third and fourth degree. The geometrical character of the exposition in Chapter I was adopted because it was necessary in Chapter III and, even more so in Chapter IV. In the beginning of Chapter I (§2) will be found my proposed proof of the theorem about the existence of infinitely many independent irreducible algebraic irrationalities of a given dimension and signature. The idea of considering an affine dilation with coefficients $r, r^2, \cdots, r^n$ along the axes in calculating the volume $Q^*(r)$ is due to a student of the Moscow State University, E. Vegeman. Further, in §3, we give the geometry of Galois theory which was developed by me [19]. [1] §4 contains a purely geometrical interpretation of Dirichlet

1) Numbers in square brackets after an author's name refer to the bibliography. If there is no such reference, then the result appears for the first time in this book.

units. §5 contains the researches of Minkowski from *Diophantische Approximationen* of the geometry of the theory of ideals (this is the only mention of the proposed geometrical theory of algebraic numbers which has so far appeared in the literature). Theorem I of §5 is due to Faddeev. §6 is devoted to the development of the theory of n-dimensional auxiliary lattices proposed by Klein, which is somewhat deeper than ideal theory. The special case $n = 2$ was considered by Klein [27] in his famous lectures on the theory of numbers, while the case $n = 3$ was the subject of a Ph. D. dissertation by Furtwängler [63]. Both the theory of units and the theory of ideals are developed in Chapter I for a most general n-dimensional maximal lattice, which may be reducible. §§7, 8, and 9 contain the theory of various forms connected with the lattices in K_n. The suggestion of considering the generalized Bézoutiants arose in connection with my plan to tabulate, in common with I. Sominskiĭ and K. Billevič, the fields of the fourth degree [18] (see §40) by making use of the projection of a field parallel to a subfield. D. K. Faddeev [60] suggested the consideration of the lattice inverse to a given lattice and correspondingly the form polar to a given reducible form. This form presents a very useful algorithmic tool, as is evident in §64.

Chapter I may be useful to anyone wishing to study the theory of algebraic numbers, as it contains a sufficiently complete step-by-step exposition of the fundamental facts of the theory.

Chapter II contains the elements of algebraic fields of the third degree. Its exposition, in contrast to Chapter I, is purely algebraic and it can be read independently of Chapter I. In Chapter II we give everywhere the most useful arithmetical algorithms which we know for actually carrying out the calculations involved and sometimes we even give numerical examples. In §11 we give a formula for raising a cubic number to any power. The method of extracting roots was suggested by Faddeev. It is useful for checking whether a given unit is fundamental or not and is used in §49 to solve the problem inverse to the Tschirnhausen problem for two equations of the fourth degree. In §13 I give my solution [15] of the problem for two equations of the third degree. §15 contains the theory developed by F. Levi [28] and by me [15]. §16 contains my method [15] for solving the problem of equivalence of two binary cubic forms without the reduction theory of forms. §17 contains an exposition of the well known method of Voronoĭ [8] for calculating a basis of a cubic field, a method which was the main result of his Master's dissertation. §18 contains the algorithm for the decomposition of a prime into prime ideals in a field of the nth degree, and in particular for cubic fields

according to Zolotarev [26].

Chapter III. §§26–30 and §§37–41 give an independent tabulation of multiplicative lattices and therefore also of fields of the third and fourth degree for all signatures. These sections close by giving tables of these lattices. The tabulation of rings of the third degree with positive discriminants was first achieved by Arndt [1–4] in 1852, using Eisenstein's idea [21], as a tabulation of clases of binary cubic forms. An analogous tabulation for negative discriminants was made by Mathews and Berwick [30, 31] and in a different manner by me [15]. The tabulation of rings of the fourth degree with signature $r = 0$ (number of pairs of complex roots) was made by I. Sominskiĭ and K. Billevič [18] and myself, while for $r = 1$ the table was calculated by Č. Poplavskiĭ. §§32–35 contain the geometry of binary cubic forms. The reduction theory was developed by Mathews [30, 31] and myself. The consideration of binary cubic forms as norms is due to Faddeev. The theorem in §36 was proved by Tartakovskiĭ in 1919 as a result of our conjecture, which arose from the study of a large table of discriminants of cubic units calculated for me in 1918 by students of the University of Kiev, using desk computers. This theorem remained unpublished up to now. As to the classification of cubic regions in terms of quadratic regions and of quartic regions in terms of cubic regions the following remarks are in order. Eisenstein [21] gave, in 1841, an interesting classification of binary cubic forms in terms of their quadratic covariants, which was later perfected by Arndt [1–4]. In my seminars at Leningrad University I often pointed out that Eisenstein's theory can be considered in the first place as a classification of cubic rings in terms of quadratic regions, in the second place it can be geometrized, and in the third place, it can be generalized to regions of higher degree. Subsequently B. A. Venkov [6] translated Eisenstein's classification into the language of algebraic number theory, while O. K. Žitomirskiĭ [24] completed its geometrization by showing how to select the axes in the projection space. Later I was able to discover the generalization of this theory to regions of the fourth degree. The detailed work on this generalization to the fourth degree was done by D. K. Faddeev [59]. At the present time Faddeev and I are constructing this theory, in [62], for fields of any degree. If by the direct problem of Galois theory we mean the problem of finding all the algebraic properties of a given field in terms of its Galois group and if by the inverse problem we mean that of finding all fields having a certain Galois group, then the theory developed in §§42–53 could be thought of as the solution of the inverse problem of Galois theory for fields of the third and fourth degree. We here give this theory (in a very careful

and detailed account of Faddeev) also for fields of the fourth degree, since their classification is based on the consideration of the fields of the third degree and also, curiously enough, on the consideration of general three-dimensional multiplicative lattices (i.e., also reducible lattices) and their auxiliary lattices.

Chapter IV is devoted to Voronoĭ's algorithm for the calculation of multiplicative automorphisms of fields of the third degree. At first we considered giving all the algorithms existing for this purpose, such as those of Zolotarev [25], Minkowski [33], Charve [67], Voronoĭ [9], Berwick [5] and Uspensky [55], but finally we decided to give only Voronoĭ's algorithm, as being the most convenient one. The case $D > 0$ was written up by D. K. Faddeev, while the case $D < 0$ was done by me (see also my note [16]). In §64 we give a refinement of Voronoĭ's algorithm for $D > 0$ which was suggested by me at the Kharkov congress and perfected by D. K. Faddeev in such a way as to reduce the calculations to rational integers. I must add that Faddeev has very elegantly perfected my calculations by noting that it is best to transform in parallel the given reducible ternary cubic form and its polar form. He also introduced the symbolic triangular notation for reducible ternary cubic forms.

Chapter V contains the exposition of Thue's theorem. The main ideas in this exposition, given in §§65, 66, 68, are due to V. A. Tartakovskiĭ [17], who is responsible for the term "boundary series." The result given in §69 is also due to V. A. Tartakovskiĭ. This result, which materially supplements Thue's result, remained unpublished up to now.

In §70 Siegel's result [46], which was obtained by him from considerations similar to those of Thue, is given in a more geometrical and therefore more elementary form, developed by Faddeev without the use of hypergeometric expansions and their estimates. A more careful estimate enabled us to give a stronger result, namely fifteen solutions instead of eighteen. This result is a generalization of my theorem in §75 to the case of positive discriminant. One must suppose that Siegel's limit of eighteen as well as Faddeev's limit of fifteen solutions is not exact (my limit of five for the case of negative discriminant is exact).

Chapter VI contains in its first part, §§71, 75, and 76, my investigations [11–14] concerning the representation of numbers by binary cubic forms of negative discriminant, and secondly (at the end of §75) the addition of Nagell [42] to my paper [12]. In §§72, 73, 74 we give the continuation of my investigation [11] by D. K. Faddeev [57, 61]. Nagell's theorem [40] is contained as a special case of these investigations. The second part of Chapter VI contains a proof of the

fundamental theorem of Mordell given by André Weil [7] and the investigations of D. K. Faddeev on the equation $x^3 + y^3 = Az^3$.

By the word "field" we mean throughout a finite algebraic extension of the field of rationals. From the point of view of multiplicative lattices considered in Chapter I the field consists of the totality of the coordinates of all the points of an irreducible multiplicative lattice, together with all the quotients of these coordinates obtained by dividing one point by another. The analogous totality of coordinates for a reducible multiplicative lattice will be called a "region."

B. Delone,
Moscow,
1940.

SUPPLEMENT I

INTRODUCTION TO DIRICHLET'S LECTURES ON THE THEORY OF NUMBERS: THE GEOMETRY OF BINARY QUADRATIC FORMS [1)]

B. N. DELONE

§1. DEFINITIONS AND SOME GENERAL THEOREMS ABOUT LATTICES

1. DEFINITION. A uniform sequence of points on a straight line in which the distance between two neighboring points is equal to a will be called a *sequence of points* with parameter a or a *one-dimensional lattice*, and will be denoted by E_1.

Let $\overrightarrow{OP}$ and $\overrightarrow{OQ}$ be two vectors of a given length forming a given angle POQ. The figure OPQ will be called a *vector-pair*, the point P being called the *end of the first vector*, and the point Q the *end of the second vector*. The collection of points of the plane OPQ whose coordinates are rational integers *with respect to the vector-pair* OPQ, i.e., with respect to the oblique-angled coordinate system with origin at the point O, with axes OP and OQ, and with scalar units on these axes equal to OP and OQ, will be called a *two-dimensional lattice* and will be denoted by E_2. The vector-pair OPQ will be said to be a *basic vector-pair* and the parallelogram constructed on it will be called a *basic parallelogram* of the lattice E_2.

In the same way, three given vectors $\overrightarrow{OP}$, $\overrightarrow{OQ}$, and $\overrightarrow{OR}$ forming a given angle in space will also be called a *vector-triple*. The collection of points of space whose coordinates are integers *with respect to the vector-triple* $OPQR$, i.e., with respect to the oblique-angled coordinate system with origin at the point O, with axes OP, OQ, and OR, and with scalar units on these axes equal to OP, OQ, and OR, will be called a *three-dimensional lattice*, and will be denoted by E_3.

---

1) This section is taken from the Introduction by B. N. Delone to the Russian translation of P. G. L. Dirichlet, *Vorlesungen über Zahlentheorie*, herausgegeben von R. Dedekind (4 Auflage, Braunschweig, Vieweg, 1894).

The vector-triple $OPQR$ will be said to be a *basic vector-triple* and the parallelepiped constructed on it will be called a *basic parallelepiped* of the lattice E_3.

If we are given an arbitrary system of points, then by a *parallelogram of the system* and by a *parallelepiped of the system* we will mean a parallelogram or parallelepiped all of whose vertices are points of the system. A parallelogram or parallelepiped of the system will be said to be *empty* if it contains no points of the system other than its vertices.

These definitions may be extended to lattices of any dimension.

2. THEOREM I. *A lattice contains no pair of points situated closer together than some given distance r. If D, E, and F are three arbitrary points of a lattice, and the segment FG is equal and parallel to the segment DE, then G is also a point of the lattice.*

The last property is called the *property of parallel translation* and follows easily from the definitions.

We will call two figures or two systems of points *homologous* with respect to a given lattice if one of them may be obtained from the other by a parallel translation of the lattice.

3. THEOREM II. *A parallelogram of some lattice E_2 or a parallelepiped of some lattice E_3 is a basic parallelogram of E_2 or a basic parallelepiped of E_3 if and only if it is empty.*

PROOF. That a basic parallelogram of a lattice E_2 is empty follows directly from its definition. If, conversely, a parallelogram of the lattice E_2 is empty, then in view of the property of parallel translation in E_2 the lattice E_2 contains the lattice E_2' constructed on this parallelogram. But E_2 cannot contain any other points; for otherwise, again by the property of parallel translation, the given lattice would have to contain points of the lattice E_2 other than its vertices. But this is impossible since the parallelogram was assumed to be empty. Hence the lattice E_2' is identical with E_2 and the given empty parallelogram is a basic parallelogram of the lattice E_2.

The proof is analogous for the case of a three-dimensional parallelepiped.

4. REMARK. A *triangle* of a lattice E_2 will be a *fundamental triangle* of this lattice (a triangle constructed on the basic vector-pair of the lattice E_2) if and only if it is empty, for then the parallelegram constructed on it will also be empty. However, a *tetrahedron* of the lattice E_3 may be empty, while the parallelepiped of this lattice constructed on it may turn out to be not empty, i.e., it will

not be a basic parallelepiped for the lattice.

5. THEOREM III. *If in an arbitrary system of points* 1) *there is at least one point such that the distance between the given point and any other point of the system is not less than some given value* r *and* 2) *the system possesses the property of parallel translation, then the system is a lattice.*

This fundamental theorem is, so to speak, a converse of Theorem I.

PROOF. Let O be the point of the given system that is such that no other point of the system is situated at a distance less than r from it. Then by the property of parallel translation no two points of the system can be closer than the distance r to each other. Hence, no bounded region can contain an unbounded collection of points of the system.

If the system does not consist of only the one point O (a case which, strictly speaking, does not contradict the theorem), then let $\bar{P}$ be some point of the system other than O. On the segment $O\bar{P}$ there is only a finite number of points of the system, and thus there is a point closest to O. Let this be the point P. If the segment $O\bar{P}$ contains no points of the system, then this point P will be the point $\bar{P}$. The straight line $O\bar{P}$ contains, in view of the property of parallel translation, all the points of the point sequence OP (if we take for the points D, E, and F the points O, P, and $\bar{P}$, and so on), and it contains no other points of the system; for in the contrary case, again by the property of parallel translation, such points would then belong also to the segment OP, which would contradict the assumption that P is the point closest to O on the segment $O\bar{P}$. Hence, if the given system is one-dimensional, it is identical with the sequence of points OP and is thus a one-dimensional lattice.

If the system is not one-dimensional, i.e., it possesses points that do not lie on the straight line OP, then in view of the property of parallel translation there emanates from each of these points a sequence of points equal and parallel to the sequence OP, the so-called sequence of points homologous to the sequence OP. Let $\bar{Q}$ be some point of the system that does not lie on the straight line OP, and let $\bar{Q}\bar{Q}'$ be the sequence of points passing through $\bar{Q}$ that is homologous to the sequence OP. If in the plane $OP\bar{Q}$ there is a sequence of points $\bar{\bar{Q}}\bar{\bar{Q}}'$ of the given system that is homologous to the sequence of points OP, and whose straight line $\bar{\bar{Q}}\bar{\bar{Q}}'$ passes between the lines OP and $\bar{Q}\bar{Q}'$, then the straight line $\bar{\bar{Q}}\bar{\bar{Q}}'$ has a segment in common with the parallelogram $OP\bar{Q}$ that is equal in magnitude and direction to the segment $\bar{\bar{Q}}\bar{\bar{Q}}'$. Thus, the sequence of points $\bar{\bar{Q}}\bar{\bar{Q}}'$ undoubtedly

contains a point belonging to this parallelogram (lying either within it or on its boundary). Hence, there can be only a finite number of such intervening sequences (if they exist at all). This means that there is some sequence QQ' that is closest to the sequence OP with no sequence parallel to it passing between the straight lines OP and QQ'. Then in the plane OPQ, by the property of parallel translation of the given system, there exists a two-dimensional lattice with the basic parallelogram OPQ. No points of the given system other than the points of this lattice can lie in this plane for otherwise, again because of the property of parallel translation, the parallelogram OPQ would contain points of the system other than its vertices. But this would contradict the assumption that P is the point of the straight line OP that is closest to O, while QQ' is the parallel sequence of points of the plane OPQ that is closest to OP.

Hence if the system is two-dimensional, it is identical with this two-dimensional lattice.

Finally, if the system is three-dimensional, then it possesses points that do not lie in the plane OPQ. Let $\bar{R}$ be one of these points. By the property of parallel translation, there passes through this point a lattice equal and parallel to the two-dimensional lattice OPQ, namely, the so-called lattice homologous to OPQ. Let $\bar{R}\bar{R}'\bar{R}''$ be this two-dimensional lattice. If there exists in the system a lattice $\bar{\bar{R}}\bar{\bar{R}}'\bar{\bar{R}}''$ homologous to OPQ whose plane passes between the planes OPQ and $\bar{R}\bar{R}'\bar{R}''$, then the plane $\bar{\bar{R}}\bar{\bar{R}}'\bar{\bar{R}}''$ has a parallelogram in common with the parallelepiped $OPQ\bar{R}$, and this parallelogram is equal in magnitude and direction to the parallelogram $\bar{\bar{R}}\bar{\bar{R}}'\bar{\bar{R}}''$. Thus, there is undoubtedly a point of this parallelogram in the lattice $\bar{\bar{R}}\bar{\bar{R}}'\bar{\bar{R}}''$, and hence also in the parallelepiped $OPQ\bar{R}$. In any case, there may be only a finite number of such intervening lattices (if they exist at all). This means that there is some two-dimensional lattice $RR'R''$ of our system that is homologous to the lattice OPQ and which is closest to OPQ, so that there is no two-dimensional lattice of our system parallel to it between the planes OPQ and $RR'R''$. Thus the given system of points belongs to the three-dimensional lattice with basic vector-triple $OPQR$. Moreover, there cannot be any other points in the lattice under consideration, for otherwise, again by the property of parallel translation, the parallelepiped $OPQR$ would also have to contain points of the system other than its vertices, which would contradict the assumption that no other points of the system in the plane OPQ belong to the parallelogram OPQ and that $RR'R''$ is the closest homologous lattice of our system to the lattice

OPQ. Thus the given system is identical with the three-dimensional lattice *OPQR* and the theorem is proved.

6. Theorem III shows that, conversely, a lattice may be defined as a system of points possessing the properties of *discreteness* (i.e., the distance between any two points is not less than some determined finite number r) and parallel translation. Given the property of parallel translation, we can weaken the first condition, requiring only the existence of one point of the system such that a sphere or circle with center at this point and a positive radius r contains no other point of the system. From this point of view a lattice may be characterized differently: it is a system of all the points into which a point of a plane (or of a space) may go under some discrete group of parallel translations.

A group of parallel translations of a space, if it is not merely the identity, is clearly infinite and abelian, but in general it may be either discrete or not discrete. In other words, it may or may not contain infinitely small translations. Clearly the set of all parallel translations of a lattice forms a discrete group of parallel translations. Theorem III shows, conversely, that the set of all points *homologous* to a given point of a space with respect to a given discrete group of parallel translations, i.e., obtained from this point by all the translations of this group, is a lattice. An n-dimensional lattice is a model of the most general infinite Abelian group with n independent generators, all the elements of which, other than zero, are of infinite order.

7. One may make the following remark about the freedom of choice of a basic vector-pair in the lattice E_2 or in the lattice E_3.

The concept of a *lattice* includes only a system of points, and does *not* include the straight lines on which these points lie. The same lattice E_2 or E_3 may be given by means of very different basic vector-pairs or vector-triples.

From the proof of the preceding theorem it is easy to see that necessary and sufficient conditions for the vector-pair *OPQ* to be a basic vector-pair in a given lattice E_2 are as follows: 1) O is an arbitrary point of the lattice E_2; 2) P is another point of E_2 satisfying the one condition that there is no point of the lattice E_2 within the segment OP; 3) Q is an arbitrary point of one of the two sequences of points of the lattice E_2 that are homologous to the sequence of points OP and that are closest to it. Necessary and sufficient conditions for a vector-triple *OPQR* to be a basic vector-triple of some given lattice E_3 are as follows: 1) O is an arbitrary point of E_3; 2) P is another point of E_3 satisfying

the one condition that the segment OP is empty; 3) Q is a point of E_3 belonging in the plane OPQ to one of the two sequences of points of the lattice E_3 that lie in this plane, are homologous to the sequence of points OP, and are closest to it; 4) R is an arbitrary point of one of the two two-dimensional lattices of the lattice E_3 that are homologous to the two-dimensional lattice OPQ and are closest to it.

8. THEOREM IV. *All the basic parallelograms of the same lattice E_2 have the same area and all the basic parallelepipeds of the same lattice E_3 have the same volume.*

PROOF. Since the method of proof used for the lattice E_2 is the same as for the lattice E_3, we will prove this theorem only for the lattice E_3.

Let $OPQR$ and $O\overline{P}\overline{Q}\overline{R}$ be two different basic vector-triples of the same lattice E_3, and let v and $\overline{v}$ be the volumes of basic parallelograms constructed on them. It is easy to calculate that $\overline{v} = \Delta v$, where Δ is the absolute value of the determinant $\begin{vmatrix} p & q & r \\ p' & q' & r' \\ p'' & q'' & r'' \end{vmatrix}$, whose rows (p, q, r), (p', q', r') and (p'', q'', r'') are the coordinates of the points $\overline{P}$, $\overline{Q}$ and $\overline{R}$ with respect to the vector-triple $OPQR$. But these coordinates are *integers*. Thus Δ is a nonzero integer. Since $O\overline{P}\,\overline{Q}\,\overline{R}$ is in its turn a basic vector-triple, the coordinates of the points P, Q and R with respect to the vector-triple $O\overline{P}\overline{Q}\overline{R}$ are also integers, and this means that we have analogously $v = \overline{\Delta}\overline{v}$, where $\overline{\Delta}$ is again a rational integer different from zero. Hence we have

$$v = \overline{\Delta}\overline{v} = \overline{\Delta}\Delta v,$$

from which it follows that $\overline{\Delta}\Delta = 1$, i.e., $\Delta = 1$, and thus $v = \overline{v}$.

Another proof of the same theorem. Let us take a large sphere of radius R, and let it include N points of our lattice E_3. Let A and B be two different basic parallelepipeds of the lattice E_3, i.e., such that B cannot be obtained from A by a parallel translation. To each of the N given points we assign a parallelepiped A, for instance that parallelepiped for which the given point is the lower left front vertex. All these parallelepipeds are contained in a sphere of radius R without being one inside the other. Moreover, some of them protrude outside the sphere, while a portion of the sphere near its surface is not completely filled. If the volume of such a parallelepiped is V_A, then the volume of the sphere V_R is approximately equal to NV_A, $V_R \approx NV_A$, i.e., $V_A \approx V_R/N$. Analogously, we find that $V_B \approx V_R/N$. Increasing the radius of the sphere R

unboundedly

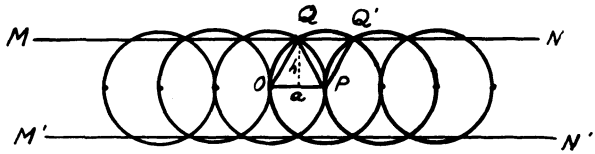

Figure 1

and taking the limit, we find that $V_A = V_B$.

REMARK. All the definitions, theorems and proofs given in this section are also true for lattices of any dimension considered in n-dimensional Euclidean space.

§2. FURTHER THEOREMS ABOUT LATTICES IN A PLANE

9. THEOREM V. *If the area of a basic parallelogram of the lattice E_2 is equal to s, then the distance between the point O of the lattice E_2 and the closest point to it of the same lattice does not exceed the value $\sqrt{2s/\sqrt{3}}$.*

PROOF. Let P be the point of the lattice E_2 closest to the point O (Figure 1). In other words, within the circle of radius $a = OP$ described about the point O, there is no other point of the lattice E_2. The same is true for similar circles described around all the remaining points of the sequence OP. If through the points of intersection of these circles we now draw the straight lines MN and $M'N'$, then within the band MN, $M'N'$ there lie no points of the lattice E_2 other than points of the sequence OP, since each interior point of this band lies within at least one of our circles. The sequence homologous and closest to the sequence OP consequently lies at least at a distance of $h = a\sqrt{3}/2$. Hence, the area s of a basic parallelogram of the lattice E_2 can not be less than $a^2\sqrt{3}/2$, from which we get that $a \leq \sqrt{2s/\sqrt{3}}$. Clearly, equality will hold if and only if the points of the homologous sequence closest to OP coincide with the points Q, Q', $\cdots$, i.e., when the fundamental triangle OPQ is equilateral. That is, the exact limit is obtained in this case only.

10. If we describe around each point of the lattice E_2 a circle of radius $a/2$, where a is the least distance between two points of the lattice E_2, then these circles will not intersect, since no two points of the lattice E_2 are within a distance less than a of each other. Thus, the lattice just found with basic vector-pair OPQ gives the *closest packing of equal circles* of diameter a, where no two circles intersect and where the centers of the circles form a lattice.

11. THEOREM VI. *In each lattice E_2 there is in general one and only one*

acute-angled fundamental triangle (if we do not count the triangle symmetric to it).

PROOF. Let P be one of the points of the lattice E_2 that are closest to O. If there are constructed through the points O and P perpendiculars to the segment OP, λ and μ (Figure 2), then either there is one point of each of the two

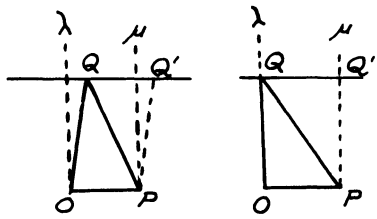

Figure 2

homologous sequences closest to the sequence OP within the band formed by these perpendiculars, or there are in each of these sequences two points lying on the perpendiculars λ and μ themselves. Let Q be one of these points. The triangle OPQ will then be empty and thus will be a fundamental triangle of the lattice E_2. This triangle has no obtuse angles. In fact, the angle Q is in any case acute, for the side OP of the triangle OPQ is the shortest segment in the lattice E_2, and thus can certainly be no longer than the other sides of this triangle. The angles O and P are not obtuse because the point Q lies either between the perpendiculars λ and μ or on one of these perpendiculars. The triangle PQQ' is also a fundamental triangle without obtuse angles, but it is symmetric to the triangle OPQ.

12. We must still decide the question of whether there are any other such nonobtuse fundamental triangles in the lattice E_2.

That part of the plane of the lattice E_2 in which each point lies as close to the point O as to any other point of the lattice is called the *Dirichlet region* of the point O in the lattice E_2. Let OPQ be an acute fundamental triangle of the lattice E_2. We consider the six triangles OPQ, OQR, ORP', $OP'Q'$, $OQ'R'$, and $OR'P$ (Figure 3). Three of these triangles are equal and the remaining three are symmetric to them. If we construct inside any one of these triangles perpendiculars at the midpoints of its sides, then they will intersect within the triangle since it is acute, and they will divide the triangle into three quadrangles. We construct such triangles for all the points of the lattice E_2, for which we need only divide into two equal parts all the basic parallelograms of the lattice E_2 homologous to the basic parallelogram $OPQR$, using diagonals homologous to the

diagonal OQ. We divide all these triangles into the small quadrangles shown in the figure. Each point of the lattice E_2 will then be surrounded by six such

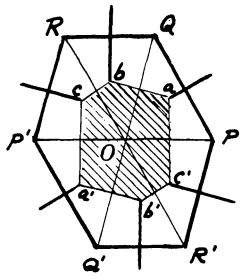

Figure 3

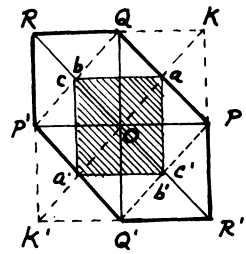

Figure 4

quadrangles, forming together a hexagon homologous to the hexagon $abca'b'c'$. These hexagons thus cover the whole plane without overlapping each other. The Dirichlet region of the point O lies in any case within the hexagon $abca'b'c'$, since each of its points is closer, for example, to O than P, i.e., lies on the "interior side" of the straight line ac', and further, is closer to O than to Q, i.e., lies on the "interior side" of the straight line ab, and so on. The Dirichlet regions of other points of the lattice E_2 have a similar position inside their corresponding hexagons. But by definition, the Dirichlet regions cover the whole plane, since for each point of the plane there is a point of the lattice E_2 that the given point is as close to as it is to any other. Hence a Dirichlet region cannot fill only a part of a hexagon; it must identically coincide with the hexagon. Thus the hexagon $abca'b'c'$ is the Dirichlet region of the point O in the lattice E_2.

By its definition the Dirichlet region is uniquely determined by the lattice E_2. As we saw, however, the acute fundamental triangles are uniquely associated with these "Dirichlet hexagons": the vertices of the triangle are the centers of those three Dirichlet regions that have a common vertex inside the triangle. Thus an acute fundamental triangle is uniquely determined by the lattice E_2.

13. In the limiting case, when the triangle OPQ is right-angled, the Dirichlet region is also degenerate: it is not a hexagon, but a quadrangle (Figure 4). Two of the sides of the hexagon, bc and $b'c'$, become equal to zero. In this case, besides the previous six nonobtuse fundamental triangles, six other triangles also meet at the point O; namely, OPQ, OKQ, OQP', $OP'K'$, $OK'Q'$, $OQ'P$.

14. THEOREM VII. *The sides of an acute fundamental triangle are the three shortest parameters of the lattice* E_2.

Parameters of a lattice are segments connecting two points of the lattice but not containing any other points of the lattice.

PROOF. Let OPQ be an acute fundamental triangle of the lattice E_2, where $PQ \leq OP \leq OQ$ and OD is a perpendicular dropped from the point O onto the straight sequence of points II (Figure 5). Then the angle PQO is greater than 45°. In fact, the angle POQ is less than or equal to the angle PQO and if the angle PQO were less than 45°, then the angle OPQ would have to be greater than 90°; but the triangle OPQ has no obtuse angles. It follows from this that the angle OQD is greater than 90°, i.e., that $OD > OQ$.

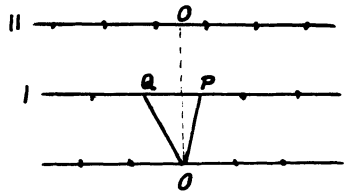

Figure 5

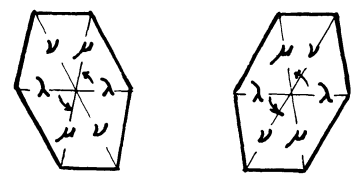

Figure 6

It is understood that OP and OQ are the two smallest parameters going from the point O to the points of the sequence I. All the parameters going to the points of the sequences II, III, and so on, are larger than OD, and thus larger than OQ. Thus PQ, OP, and OQ are the three least parameters of the lattice E_2.

15. In the future we will call a certain direction of rotation in the plane of the lattice E_2 *right* in contradistinction to rotation in the opposite direction, which will be said to be *left*. Corresponding to this, we will speak of *right* and *left vector-pairs*, meaning by the *angle of a vector-pair* that angle between its vectors which is less than 180°, always considering the direction of rotation of the angle of the vector-pair as being from its first vector to its second.

16. We will say that a vector-pair is *reduced* if 1) its first vector is the smallest and the second the next smallest parameter of the lattice and 2) the vector-pair is a right vector-pair. The three least parameters $\lambda \leq \mu \leq \nu$ may be situated in two different ways with respect to the established positive direction of rotation: either in the order λ, μ, ν or in the order λ, ν, μ. In the first case the reduced vector-pair will be acute, while in the second it will be obtuse (Figure 6).

If the vector-pair OPQ is reduced, then, in general it follows directly from Theorems VI and VII that only the reflected vector-pair $OP'Q'$ will also be

reduced (Figure 8).

17. The only exceptions are the following three cases: 1) when $\lambda = \mu < \nu$; 2) when $\lambda < \mu = \nu$; 3) when $\lambda = \mu = \nu$. In these cases there are respectively 4, 4, and 12 reduced vector-pairs, which are pairwise reflections of each other (Figure 7).

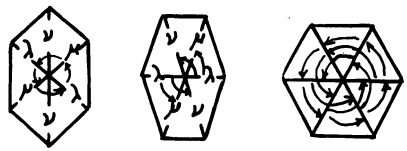

Figure 7

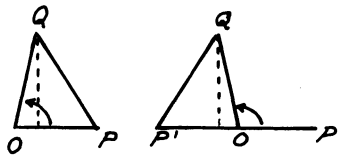

Figure 8

The case of a right-angled or square Dirichlet region is not an exception in this sense.

18. THEOREM VIII. *A basic vector-pair OPQ will be reduced if and only if* 1) *the projection of its second side OQ onto the straight line of its first side OP is in absolute value less than or equal to half of its first side*; 2) *its first side OP is less than or equal to its second side OQ*; 3) *it is a right vector-pair.*

PROOF. In fact, if these conditions are satisfied and the angle of the vector-pair is not obtuse, then in the fundamental triangle $PQ \geq OQ \geq OP$ (Figure 8). This means that its largest angle is located at the point O, i.e., the triangle is not obtuse, from which it follows by Theorem VII that the vector-pair OPQ is reduced. If the angle of the vector-pair is obtuse, then the triangle OQP' has the indicated properties, from which it is again clear that the vector-pair OPQ is reduced.

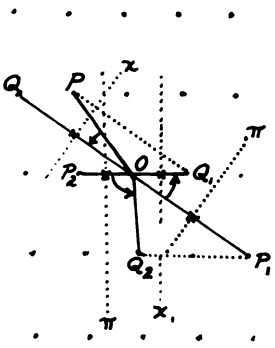

Figure 9

19. **Algorithm of reduction.** Let there be given a vector-pair OPQ. If the point Q' is symmetric to the point Q with respect to O and if R is an arbitrary point of the sequence passing through P parallel to the sequence OQ, then we will say that the vector-pair $OQ'R$ is adjacent on the right to the vector-pair OPQ. This vector-pair has the same direction of rotation as the vector-pair OPQ.

The transition from an arbitrary basic vector-pair of the lattice (the direction of rotation of which is taken to be positive) to a reduced vector-pair may be made (following Gauss) in the following manner (see Figure 9).

We construct the vector-pair OP_1Q_1 adjacent on the right to the given vector-pair OPQ so that the projection of its second vector OQ_1 onto the first vector OP_1 is less than or equal to the first vector in absolute value, i.e., so that the point Q_1 lies between the perpendiculars π and κ. There is always one such point, or two (if they lie on the perpendiculars π and κ themselves). Further, we again construct a vector-pair OP_2Q_2 adjacent on the right to the vector-pair OP_1Q_1, possessing the same property, i.e., it is such that the point Q_2 lies between the perpendiculars π_1 and κ_1, and so on.

For each such vector-pair condition 1) of Theorem VII is satisfied, and it cannot happen that condition 2) is never satisfied. In fact, it would then be true that $OQ_1 > OQ_2 > OQ_3 > \cdots$. But only a finite number of points of the lattice are situated in the circle of radius OQ_1 with center at O. Consequently, condition 2) must be satisfied after a finite number of such transformations. Then the vector-pair will be reduced.

20. **The Pell angle of a lattice.** Every lattice coincides with itself after rotation by 180° around its point O. But it may occur that a lattice will coincide with itself under rotation by a smaller angle. Since the Dirichlet region is a rectangle or a hexagon with center of symmetry at the point Q, this may occur only when the Dirichlet region is a square or a regular hexagon, and then this angle is 90° or 60°. We will call this least angle of repetition the *Pell angle* of the lattice.

21. **The vector-pair corresponding to a given parameter of a lattice.** Every parameter OM of the lattice may be taken for the first side of some basic vector-pair of the lattice. That one of these vector-pairs which has a right direction of rotation and which satisfies condition 1) of Theorem VIII will be called the *vector-pair corresponding to the parameter* OM, and the vector-pair itself will be said to be *semireduced.*

In the general case there are in all two distinct semireduced vector-pairs, while in the case when the Pell angle is 90° or 60°, there are always four or six such distinct vector-pairs respectively, but there can not be more since if there were a vector-pair equal to the given one, i.e., obtained from it by means of a rotation around the point O, then the lattice would coincide with itself under this rotation, since the given vector-pair is a basic one.

§3. THEORY OF THE DISTRIBUTION OF THE POINTS OF A LATTICE WITH RESPECT TO GIVEN ASYMPTOTES

We turn now to lattice theorems of a slightly different character, namely relating to the distribution of points of the lattice E_2 with respect to some infinite straight line.

22. DEFINITION. Let there be given in the plane of some given lattice E_2 two straight lines $O\xi$ and $O\eta$ passing through the point O of the lattice and irrational with respect to the lattice E_2, i.e., meeting no other points of the lattice. We will call these lines *axes* or *asymptotes*.

Let P be an arbitrary point of the lattice E_2. In the future a *coordinate parallelogram of a point* P will be taken to mean a parallelogram which has its center at the point O, one of its vertices at the point P, and sides parallel to the axes $O\xi$ and $O\eta$. Since the axes are irrational, no side of the coordinate parallelogram can contain two points of the lattice E_2, since otherwise there would be a point of the lattice other than the point O lying on the corresponding axis. Hence if P is a point of the lattice E_2, then there lies on the boundary of its coordinate parallelogram only one more point P' of the lattice, symmetric to P with respect to the point O, and located at the opposite vertex of the parallelogram.

A point P of the lattice E_2 whose coordinate parallelogram contains within itself no points of the system other than the point O is called a *relative minimum* of the lattice E_2 with respect to the asymptotes $O\xi$ and $O\eta$.

23. THEOREM IX. *The given lattice* E_2 *has infinitely many relative minima with respect to the given asymptotes* $O\xi$, $O\eta$.

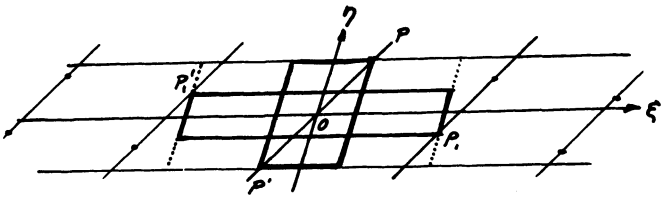

Figure 10

PROOF. It is first necessary to show that there exists at least one such relative minimum. This is clear; for example, if the point Q of the lattice E_2 is not a relative minimum, i.e., if there are other points of the lattice within its coordinate parallelogram, then it is possible to apply the argument to one of these

points, and so on. In the end we are led in this manner to a relative minimum, since there are only a finite number of points of the lattice E_2 in the coordinate parallelogram of the point Q.

24. Now let P be a relative minimum, with a positive abscissa ξ for example, (Figure 10). If we extend indefinitely the sides ot the coordinate parallelogram of the point P parallel to the axis ξ, we will obtain a band containing the axis ξ. The right side of the coordinate parallelogram of the point P parallel to the axis η is now allowed to slide inside this band, remaining constantly parallel to itself, in the direction of increasing abscissa. Eventually, it will necessarily cross some point of the lattice E_2, since every sequence of points of the lattice E_2 parallel to the sequence of points OP has two points within the band. In fact, the segment of the straight line of such a sequence of points lying within the band has length $PP' = 2OP$, and there are no points of the lattice E_2 other than O on the axis ξ.

Let P_1 be the first point through which this sliding side passes. Then only this point lies on it, since the axis η also contains no points of the lattice other than the point O. Thus P_1 is clearly a relative minimum. We may now repeat this step, starting with the point P_1, and so on. We thus obtain an infinite set of relative minima $P_1, P_2, P_3, \cdots$ lying along the positive semi-axis ξ.

25. The points $P_1', P_2', P_3', \cdots$ symmetric to these minima with respect to the point O are also relative minima, but they lie along the negative semi-axis ξ.

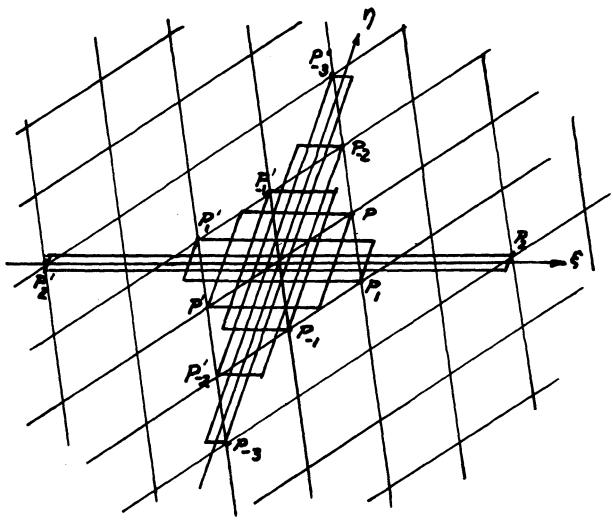

Figure 11

If, remaining in the half-plane $\xi > 0$, we apply this method, starting from the same point P and in the direction of the axis η, we obtain a chain of relative minima $P_{-1}, P_{-2}, P_{-3}, \cdots$ which approach the axis η. There are also relative minima $P'_{-1}, P'_{-2}, P'_{-3}, \cdots$ symmetric to them with respect to the point O (Figure 11).

26. THEOREM X. *The method of Theorem IX yields all of the relative minima of the lattice* E_2.

PROOF. In fact, every relative minimum that has an abscissa larger than that of P must have an ordinate larger in absolute value than the ordinate of the point P; in the contrary case P would lie in its coordinate parallelogram. Hence, each such relative minimum must lie in the band considered in Theorem IX.

But P_1 was the first point of the lattice E_2 that had an abscissa larger than that of the point P. The same is true for P_2 with respect to P_1 and so on.

Two relative minima situated with respect to each other in the same way as the points P and P_1 will be said to be *successive* and P_1 will be said to be the *first successor* with respect to the minimum P along the semi-axis $+\xi$. Thus for example, P'_{-2} and P'_{-3} are successive relative minima and P'_{-3} is the first successor with respect to the minimum P'_{-2} along the semi-axis $-\xi$.

27. THEOREM XI. *A vector-pair constructed on two successive relative minima is a basic vector-pair of the lattice* E_2.

PROOF. The triangle OPP_1 (Figure 12) lies in the parallelogram *abcd*, and this parallelogram is empty except for the point O. The triangle OPP_1 is thus empty, and hence a fundamental triangle. This means that the vector-pair OPP_1 is a basic vector-pair of the lattice E_2.

It follows from this theorem that P_1 lies in the parallel sequence of points of the lattice E_2 that is closest to the sequence OP.

28. THEOREM XII. *Successive minima lie on different sides of the corresponding axis.*

PROOF. In fact, the two points of the sequence parallel to the sequence OP that are included within the band of Theorem IX lie on different sides of the axis ξ, and both are situated to the right of the right side of the coordinate parallelogram of the point P since it is empty.

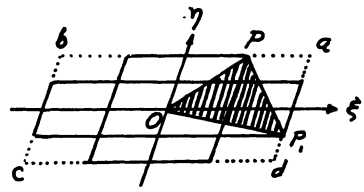

Figure 12

Under the application of the method used in the proof of Theorem IX, this side clearly first meets that one of the two points that lies on the other side of the axis ξ with respect to the point P.

29. DEFINITION. By the *angle of a vector-pair* we will again mean the angle formed by the vectors that is less than 180°.

We will say that a vector-pair *includes* a given axis or asymptote (these terms are synonymous) parallel to one of the two chosen axes if this asymptote passes within the angle of the vector-pair. It is clear that there can be only three possibilities: the vector-pair may include one, two, or no asymptotes.

Which of the ends of the vector-pair we will call *first* and which *second* is arbitrary; but in the symbol for the vector-pair we will always place the letter denoting the end of its first vector immediately after the O that designates its vertex.

A segment drawn from the end of the *first* vector of the vector-pair in a direction parallel to its second vector will be called the *beak* of the vector-pair. Thus the beak always proceeds from the first side of the vector-pair. Its length may be measured by taking the second side for unit length. If it is positive, it is directed inside the angle of the vector-pair, while if it is negative, it goes outside this angle. We will say that we *extend* the beak only when we draw it inside the frame.

A basic vector-pair of the lattice E_2 will be said to be *reduced* with respect to some asymptote if the ends of its vectors are successive minima such that the second is the successor with respect to the first along this asymptote.

THEOREM XIII. *A basic vector-pair of the lattice E_2 is reduced if and only if* 1) *it includes one and only one asymptote and* 2) *the end of its second vector lies further along this asymptote, but closer to it, than the end of its first vector.*

PROOF. The necessity of these conditions follows from the constructions of Theorem IX and Theorem XI. However, they are also sufficient.

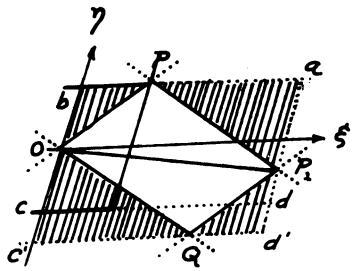

Figure 13

In fact, we continue all four of the sides of the parallelogram OPP_1Q (Figure 13). Then we obtain two empty intersecting bands, since OPP_1Q is a basic parallelogram. As is easy to see, these two bands cover the parallelogram $abc'd'$, which means that they also cover $abcd$, since the latter contains only the points

O, P, and P_1 of the lattice E_2. Thus P and P_1 are successive minima and the second is the successor with respect to the first. Thus conditions 1) and 2) are also sufficient.

31. Let OPP_1 be a reduced frame, and let $P_2, P_3, P_4, \cdots$ be a sequence of relative minima following one another after the minimum P_1 along the axis included in the vector-pair. Then all the vector-pairs OP_1P_2, OP_2P_3, $OP_3P_4, \cdots$ are also reduced basic vector-pairs. We call them *successors* with respect to the reduced vector-pair OPP_1. The collection of all these vector-pairs will be called a *chain* of the reduced vector-pair OPP_1. Starting with the first frame, this chain may also proceed in the opposite direction along the same semi-axis, where the vector-pairs $OP_{-1}P$, $OP_{-2}P_{-1}$, $OP_{-3}P_{-2}, \cdots$ are predecessors of the vector-pair OPP_1 (see Figure 11).

The fundamental problem is the following: given an arbitrary basic vector-pair of the lattice E_2, to find a method of going from it to some reduced vector-pair (it is immaterial to us along which axis it will be reduced), and further to go step by step to all the successive vector-pairs in its chain.

32. Preliminary transformation of the given vector-pair. If the given vector-pair OPP_1 contains no asymptotes, then we go from it to the vector-pair OPP_1' containing two asymptotes (Figure 14). If it is necessary at all, this preliminary transformation will be made only once, at the very beginning of the proposed algorithm. Thus we may confine ourselves in the future to the consideration of vector-pairs including at least one asymptote. For simplicity, we will say that such vector-pairs are prepared.

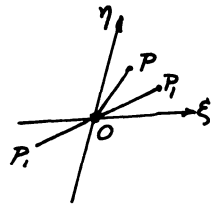

Figure 14

33. Transformation of an arbitrary prepared vector-pair into a chain of reduced vector-pairs, and progress along this chain. Assume that we are given the vector-pair OPP_1 of the lattice E_2 already prepared, i.e., such that it includes at least one asymptote. Then the reduction algorithm consists of the following.

We extend the beak of the vector-pair OPP_1 to the last point P_2 for which the vector-pair OP_1P_2 still includes at least one asymptote. We then extend the beak of the basic vector-pair OP_1P_2 thus obtained to the last point P_3 for which the vector-pair OP_2P_3 still includes at least one asymptote, and so on.

We will now show that by this process we always go from the given vector-pair

to some reduced one, and then step by step to all its successors in its chain.

34. CASE I. The prepared vector-pair is already reduced, i.e., it includes one and only one axis, and the end of its second side lies further along this axis, but closer to it, than the end of its first side (Figure 15).

Since the point P is further from the asymptote than P_1, the length of the beak to the point of its intersection with the included asymptote is greater than unity. This means that the last point P_2 on the beak which still lies on the same side as the point P, and hence forms another vector-pair OP_1P_2 which includes the asymptote, is different from P. The positive integer δ showing how many

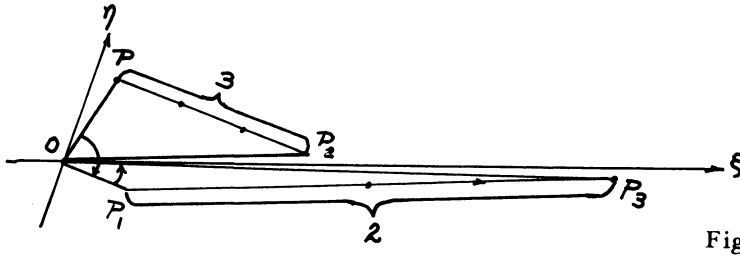

Figure 15

times it is necessary to lay off the segment OP_1 from the point P to the point P_2, i.e., the *length* of the beak PP_2, is equal to 3 in our drawing. The vector-pair OP_1P_2 is once more reduced, since again 1) it includes the asymptote $O\xi$ and 2) the end of its second side lies further along this asymptote, but closer to it, than the end of its first side.

We now consider the vector-pair OP_1P_2 and extend its beak in the same way as we did the beak of the vector-pair OPP_1. This second beak P_1P_3 has length 2 in our drawing. We continue this operation with the vector-pair OP_2P_3, and so on. Thus extending the beaks alternately, first on one side, and then on the other side of the asymptote, we will obtain, one after the other, successive vector-pairs of the chain of the vector-pair OPP_1.

35. CASE II. The prepared vector-pair, although including only one asymptote, is not reduced. Here there may be three possibilities; shown respectively in the drawings α), β) and γ) of Figure 16.

α) The end of the first side, but not of the second, lies further along the included asymptote, but closer to it. The beak, proceeding as always from the end of the first side, i.e., from the point P to the point of its intersection with the included asymptote, is shorter than the second side OP_1 of the vector-pair. The point P_2 is the point P, and $\delta = 0$. The first step in the algorithm consists

simply in going from the vector-pair OPP_1 to the vector-pair $OP_1P_2 = OP_1P$, i.e., in the permutation of successive sides of the vector-pair. The vector-pair OP_1P_2 will already be reduced.

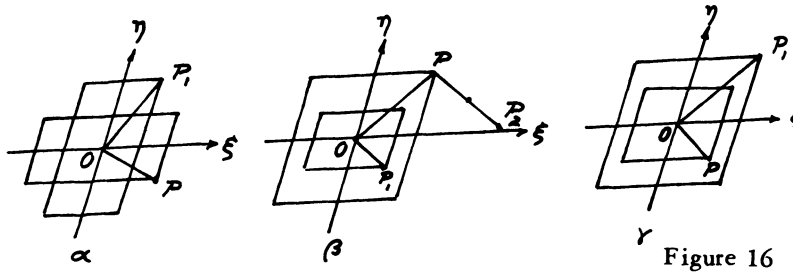

Figure 16

β) The end of the second side lies within the coordinate parallelogram of the end of the first side. The beak continued to the point of intersection from the included asymptote is longer than OP_1. Hence δ is at least equal to unity and the point P_2 is distinct from the point P. The vector-pair OP_1P_2 is already reduced.

γ) The end of the first side lies within the coordinate parallelogram of the end of the second side. The point P_2, as in case α), is the point P. The first step of the algorithm gives $\delta = 0$ and leads only to changing the places of the sides of the vector-pair. Then we obtain step β, and consequently only one more step of the algorithm is needed to give a reduced vector-pair.

Thus we see that in Case II the algorithm is applied without change, and one, or at most two, steps make the vector-pair reduced. If we continue to repeat these steps we will obtain one after another successive vector-pairs which follow this reduced vector-pair and are members of its chain, since we find ourselves already in the position of Case I.

36. CASE III. The prepared vector-pair includes two asymptotes. Here again there may be three distinct possibilities.

α) The points P and P_2 lie on opposite sides of the first asymptote (Figure 17). Of the two asymptotes covered by the vector-pair, we will take as the first one that which first intersects the positive beak, extended as always from the end of the first side of the vector-pair under consideration.

In case α) the vector-pair OP_1P_2 already includes only one asymptote. Thus after the first step of our algorithm we will obtain Case I or II.

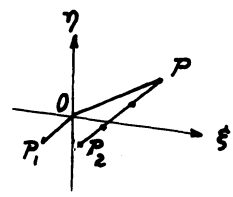

Figure 17

β) The point P_2 does not coincide with the point P although it lies in front of the first asymptote. This occurs when there are points of the lattice E_2 on the positive beak between P and the first asymptote, but there are no points of the lattice on the segment of the beak between both asymptotes (Figure 18).

In this case the point P_2 lies inside the parallelogram $OaPb$. If the following step of our algorithm yields the same result, the point P_3 lies within the parallelogram OcP_1d. If the next step of the algorithm again leads to the same situation, then the point P_4 lies inside the analogous parallelogram of the point P_2, and thus within the parallelogram $OaPb$, and so on.

But since the parallelograms $OaPb$ and OcP_1d contain only a finite number of points of the lattice E_2, it will eventually turn out that one of the successive points will lie *between* the asymptotes, and we will arrive at one of the cases that have already been considered.

γ) The point P_2 is the point P, i.e., the first point Q of the lattice E_2 on the positive beak lies already on that side of the second asymptote (Figure 19).

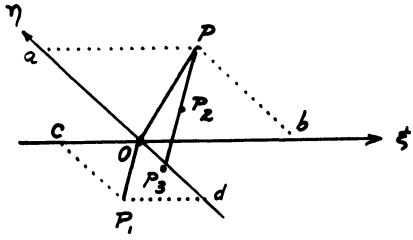

Figure 18

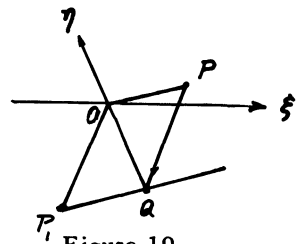

Figure 19

In our algorithm we again obtain $\delta = 0$ and the first step leads simply to an interchange of the vectors of the vector-pair, i.e., to the transition from the vector-pair OPP_1 to the vector-pair $OP_1P_2 = OP_1P$. At the next step the beak will no longer be equal to zero, i.e., $\delta_1 \neq 0$, since the first point Q on this beak which is extended now from the point P_1 still lies in front of the first asymptote. Thus after the second step we return to one of the cases that have been considered.

Thus, by means of our algorithm we can finally arrive from an arbitrary prepared vector-pair to a vector-pair covering only one asymptote. And then, repeating the algorithm without any change, we obtain a reduced vector-pair and subsequently, one after another, the successive vector-pairs of the chain following it.

37. REMARK. All the definitions and theorems concerning the properties of lattices with respect to asymptotes are *invariant* with respect to any affine trans-

formation. For example, a relative minimum remains a relative minimum, and so on.

38. Hyperbolic rotation. Let ρ be an arbitrary positive number. The transition from points (ξ, η) to points $(\rho\xi, \rho^{-1}\eta)$ is an affine transformation which takes the asymptotes into themselves. We will call this transformation a *hyperbolic rotation* and the magnitude ρ, the *parameter* of the rotation. Under such a rotation each point slides along the hyperbola (with asymptotes $O\xi$, $O\eta$) on which it lies.

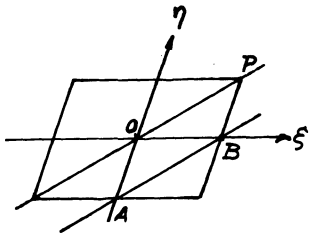

Figure 20

39. THEOREM XIV. *The area of the coordinate parallelogram of a relative minimum is less than* $4s$, *where* s *is the area of a basic parallelogram of the lattice under consideration.*

PROOF. In fact, if the point P is a relative minimum, then the parallelogram $OPAB$ must be empty (Figure 20). But this may occur only if its area, which is equal to the area of the parallelogram $OBPC$, is less than s.

40. THEOREM XV. *If all the relative minima lie on a finite number of hyperbolas, then the lattice periodically coincides with itself under a hyperbolic rotation.*

PROOF. In fact, by a suitable hyperbolic transformation, the end of the first side of the reduced vector-pair may once and for all be taken on a fixed straight line η_0 parallel to the axis η (Figure 21). But this straight line has only a finite number of points of intersection with the possible hyperbolas. Thus there are only a finite number of possible positions for the end of this side. But since s is given and does not change under hyperbolic rotation, for each such position of the end of the first vector of the reduced vector-pair, the end of the second vector must lie on some fixed straight line δ parallel to the first vector of the vector-pair. This straight line has only a finite number of points of intersection with the finite number of possible

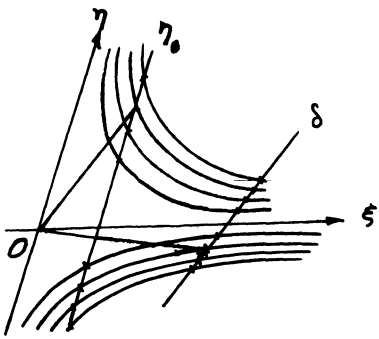

Figure 21

hyperbolas on which the end of the second vector must lie.

Thus in this case there is only a finite set of such distinct *normed* reduced basic vector-pairs. But since there is an infinite set of reduced basic vector-pairs in a lattice, there is undoubtedly an infinite set of such reduced vector-pairs that give the same normed vector-pair. In other words, there exists a hyperbolic rotation as a result of which the lattice coincides with itself. Under further rotations by the same "angle" the lattice periodically coincides with itself.

41. The Pell angle of a lattice with respect to given asymptotes. We will call the least of the indicated angles the *Pell angle*. In the same way that an arbitrary lattice coincides with itself under a usual rotation around its point O by each 180°, here the lattice coincides with itself after a hyperbolic rotation by its Pell angle.

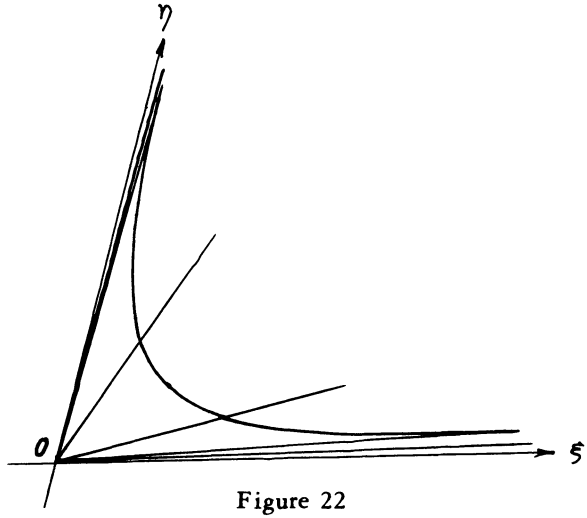

Figure 22

Each angle between asymptotes can be divided into an infinite number of such angles, where this division may begin with an arbitrary ray (Figure 22). Moreover, that part of the lattice which lies in one of these angles is identical up to a hyperbolic rotation by a corresponding multiple of this Pell angle, to the part of the lattice lying in each other such angle.

42. REMARK. It follows from the last two theorems that if a lattice does not periodically coincide with itself under hyperbolic rotation, then among the hyperbolas whose coordinate parallelograms have an area less than $4s$, there is at least one limiting hyperbola which is a position of condensation for the hyperbolas on which the points of the lattice lie.

43. THEOREM XVI. *Conversely, if a lattice periodically coincides with itself under a hyperbolic rotation, then its points, being situated under a definite hyperbola, lie on only a finite number of hyperbolas, where each hyperbola that contains at least one point contains an infinite number of points of the lattice.*

PROOF. This follows from the fact that a part of the plane lying under a given hyperbola in one of the Pell angles has a finite diameter. Consequently, there are in it only a finite number of points of the lattice, which are therefore situated on a finite number of hyperbolas. But because of the periodicity under hyperbolic rotation, the points of the lattice lie on the same hyperbolas in the other Pell angles.

Below we will see that in fact both these cases, which are so different from one another, do actually occur: there exist lattices which repeat infinitely under hyperbolic rotation with respect to a given asymptote, and there exist lattices without this property.

44. **The vector-pair corresponding to a given parameter of a lattice in the hyperbolic case.** Each parameter OM of a lattice may be taken for the first vector of some basic vector-pair. In subsection 21 we defined a "vector-pair corresponding to a given parameter OM" of a lattice, or a "semireduced vector-pair," for the usual case when the Pell angle is equal to 180°. We may further say the following: a vector-pair will be said to correspond to the parameter OM when 1) it is a right vector-pair, 2) OM is its first vector, and 3) the orthogonal projection of its second vector onto the first vector has the least possible absolute value.

We will call two directions *hyperbolically orthogonal* if they are parallel to two diameters conjugate with respect to the given asymptotes. Then we will again say that a vector-pair *corresponds to a given parameter* OM in the hyperbolic case when 1) it is a right vector-pair, 2) OM is its first vector, and 3) the orthogonal projection (in the hyperbolic sense) of its second vector onto the first vector has the least possible absolute value, i.e., the end of the second vector of the vector-pair is chosen from the one of the two closest parallel sequences to the sequence OM that gives a right vector-pair, and it is chosen in this sequence so that its projection onto the straight line OM parallel to the direction conjugate with OM lies at a minimal distance from the point O.

45. There cannot exist two distinct points M for which the corresponding vector-pairs are identical, up to a hyperbolic rotation by an angle less than the

Pell angle, for it is clear that the hyperbolic angle of two such vector-pairs would also be an angle of periodicity of the lattice (Figure 23).

§4. THE THEORY OF POSITIVE BINARY QUADRATIC FORMS

We turn now to methods that will enable us to introduce calculations into the theory of lattices.

46. The positive binary quadratic form corresponding to a given vector-pair. A particular lattice E_2 can be defined by the length of the sides of its basic vector-pair OPQ and the magnitude of the angle formed by them. However, another method is more convenient. We are given the squares A and C of the magnitude of the vectors of the basic vector-pair and the product B of these magnitudes by the cosine of the angle ϕ between the vectors. By the well-known formula for the square of the side of an oblique triangle, we find that the square of the distance from the point O to a point of the lattice E_2 that has coordinates x and y with respect to the basic vector-pair is

$$Ax^2 + 2Bxy + Cy^2.$$

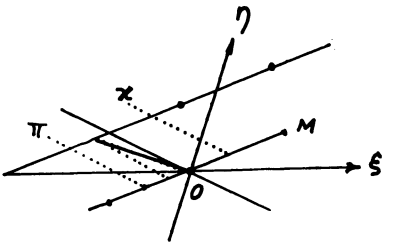

Figure 23

This connection between the quadratic form and the vector-pair OPQ may be represented in a different way. Let ξ and η be arbitrarily chosen rectangular coordinates and let the vector-pair OPQ lie in their plane so that the point O coincides with the origin of coordinates. Then if the points P and Q have the coordinates (ξ_1, η_1) and (ξ_2, η_2) clearly $\xi_1\xi_2 + \eta_1\eta_2 = \overline{OP} \cdot \overline{OQ} \cos\phi$ (i.e., is equal to the scalar product of the vectors $\overrightarrow{OP}$ and $\overrightarrow{OQ}$). Thus we have

$$(x\xi_1 + y\xi_2)^2 + (x\eta_1 + y\eta_2)^2$$
$$= (\xi_1^2 + \eta_1^2)x^2 + 2(\xi_1\xi_2 + \eta_1\eta_2)xy + (\xi_2^2 + \eta_2^2)y^2 = Ax^2 + 2Bxy + Cy^2.$$

Among other things it follows from this that the form $Ax^2 + 2Bxy + Cy^2$, which we will also abbreviate to (A, B, C), is positive.

From the point of view of vectors the form (A, B, C) is simply the scalar square of the linear vector expression $px + qy$, i.e., $Ax^2 + 2Bxy + Cy^2 = (px + qy)^2$, where p and q are the vectors $\overrightarrow{OP}$ and $\overrightarrow{OQ}$; in fact

$$A = \xi_1^2 + \eta_1^2 = p^2, \quad B = \xi_1\xi_2 + \eta_1\eta_2 = (pq), \quad C = \xi_2^2 + \eta_2^2 = q^2.$$

47. It is not difficult to see that if the form (A, B, C) undergoes a permutation $\begin{pmatrix} \alpha, & \beta \\ \gamma, & \delta \end{pmatrix}$, i.e., x and y are interchanged in the expressions $\alpha x + \beta y$ and $\gamma x + \delta y$, where α, β, γ, and δ are real numbers, we obtain the form $(\bar{A}, \bar{B}, \bar{C})$, whose vector-pair $O\bar{P}\bar{Q}$ has the ends of its vectors $\bar{P}$ and $\bar{Q}$ at the points with coordinates (α, γ) and (β, δ) with respect to the vector-pair OPQ. If α, β, γ, and δ are rational integers and $\alpha\delta - \beta\gamma = \pm 1$, then the vector-pair $O\bar{P}\bar{Q}$ is again a basic vector-pair of the lattice E_2. In fact, the points $\bar{P}$ and $\bar{Q}$ are then points of this lattice and the area of the parallelogram $O\bar{P}\bar{Q}$ is equal to the area of the parallelogram OPQ. In this case the forms (A, B, C) and $(\bar{A}, \bar{B}, \bar{C})$ are said to be *equivalent*; they are *properly equivalent* if $\alpha\delta - \beta\gamma = 1$, and *improperly equivalent* if $\alpha\delta - \beta\gamma = -1$. Properly equivalent forms correspond to basic vector-pairs with the same direction of rotation, while improperly equivalent forms correspond to vector-pairs with opposite directions of rotation.

Thus lattices E_2 and classes of properly equivalent binary positive quadratic forms correspond to each other in the sense that to each such class there corresponds a definite lattice E_2, and to each lattice E_2 there correspond in general two such classes, depending on which of the two directions of rotation is considered positive. The determinant of the form $D = B^2 - AC$ is equal to $-(\xi_1\eta_2 - \xi_2\eta_1)^2$, which is equal to $-s^2$, where s is the area of a basic parallelogram of the given lattice E_2.

48. The form (A, B, C) is said to be *reduced* (following Gauss) if its vector-pair is reduced in the sense of §2, subsection 16. The conditions indicated in Theorem VIII yield the following *conditions of reduction* expressed in terms of the quantities A, B, and C:

$$C \geq A \geq 2|B|.$$

49. The *reduction algorithm* (following Gauss) presented in §2, subsection 19 is translated into the language of calculation in the following way. The form (A', B', A'') is found, which is adjacent on the right to the given form (A, B, A'): $(A', B', A'') = (A, B, A')\begin{pmatrix} 0, & 1 \\ -1, & \delta \end{pmatrix}$, where the number δ is chosen so that $B' = -B + A'\delta$ is between $-A'/2$ and $A'/2$. Then the form (A'', B'', A''') adjacent on the right to the form (A', B', A'') is found in exactly the same way, and so on until finally one of the forms, for example the form $(A^{(n)}, B^{(n)}, A^{(n+1)})$, does not satisfy the conditions of reduction.

50. If the form (A, B, C) is reduced and $B \geq 0$, then its vector-pair determines a nonoblique triangle; but if $B < 0$, then such a triangle is given by the form $(A, B, C)\begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix} = (C, -B, A)$. If the form (A, B, C) determines an acute triangle, then the transformation for going from this vector-pair to itself and to the other vector-pairs corresponding to the five remaining acute triangles surrounding the point O are the following:

$$\begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix}, \begin{pmatrix} 0, & -1 \\ 1, & 1 \end{pmatrix}, \begin{pmatrix} -1, & -1 \\ 1, & 0 \end{pmatrix}, \begin{pmatrix} -1, & 0 \\ 0, & -1 \end{pmatrix}, \begin{pmatrix} 0, & 1 \\ -1, & -1 \end{pmatrix}, \begin{pmatrix} 1, & 1 \\ -1, & 0 \end{pmatrix}.$$

The following forms correspond to them:

$$(A, B, C), (C, -B + C, A - 2B + C), (A - 2B + C, A - B, A);$$
$$(A, B, C), (C, -B + C, A - 2B + C), (A - 2B + C, A - B, A).$$

51. **Tabulation of positive binary quadratic forms with integer coefficients.** The discussion of positive forms (A, B, C) has up to now referred to the case of completely arbitrary real coefficients A, B, C. However, one often considers forms whose coefficients A, B, C are ordinary rational integers. For forms with rational integer coefficients the following fundamental theorem holds: *The number of classes of such forms with the same discriminant D is finite.*

For the proof of this theorem we note that, from the conditions of reduction for a reduced form there follow the inequalities $A^2 \leq AC$, $4B^2 \leq A^2$, from which we have $4B^2 \leq AC$, or $3B^2 \leq AC - B^2 = |D|$, i.e., $|B| \leq \sqrt{|D|/3}$. Thus if the discriminant of a positive form with integer coefficients is equal to D, then B may take on only the values $0, \pm 1, \pm 2, \cdots, \pm \lambda$, where λ is the greatest integer $[\sqrt{|D|/3}]$, in $\sqrt{|D|/3}$. If B is now given one of these values, then $AC = B^2 + |D|$. Thus it is necessary to factor the number $B^2 + |D|$ in all possible ways into two positive factors (in the case of a positive form A and C are both always positive) and take each time for A that factor which does not exceed the other, which will be taken for C. If it turns out that $A \geq 2|B|$, then the form thus obtained is reduced and is to be recorded; in the contrary case it is omitted. By this method we necessarily obtain all the reduced forms.

By subsection 16 of §2 we may decide whether there are equivalent forms among them. Subsections 16 and 17 of §2 show that the only cases when two nonidentical reduced forms are properly equivalent are the cases 1) $\lambda = \mu$ and 2) $\lambda < \mu = \nu$ (the case $\lambda = \mu = \nu$ is not an exception with respect to forms, for it also gives only two distinct reduced forms). In case 1) one of the reduced forms is the

form (A, B, A) and the other is the form $(A, -B, A)$ obtained from the first by the transformation $\begin{pmatrix} 0, & -1 \\ 1, & 0 \end{pmatrix}$; in case 2) the forms are $(A, \frac{1}{2}A, C)$ and $(A, -\frac{1}{2}A, C)$, where the first goes into the second under the transformation $\begin{pmatrix} 1, & -1 \\ 0, & 1 \end{pmatrix}$.

In §67 we present three examples of the determination of all nonequivalent reduced forms for a given negative determinant $D = -\Delta$.

52. **The resolution of the question of whether two given positive binary quadratic forms with integer coefficients are properly equivalent.** If the determinants are not equal the forms are not equivalent. If they are equal, we find two reduced forms of which the first is properly equivalent to the first of the given forms and the second to the second. It is clear that the given forms are properly equivalent if and only if these reduced forms are identical or if one of the exceptional cases just indicated takes place. Let S and T be the transformations by means of which the two forms are transformed into identical reduced forms. Then the first form goes into the second by means of the transformation ST^{-1}.

53. **The representation of numbers by means of positive binary forms with integer coefficients.** Let m be a given positive rational integer, and let A, B, C also be rational integers. It is required to solve the equation $Ax^2 + 2Bxy + Cy^2 = m$ in rational integers x, y. Each such solution x, y is said to be a *representation* of the number m in terms of the form (A, B, C). The determination of all these representations is the determination of all the points M of the lattice corresponding to the forms that lie on the circle of radius $\sqrt{m}$ with center at the point O.

It is sufficient to give a method for the determination of all the representations in which x and y are relatively prime. In fact, if their greatest common divisor were μ, for example, then the number m would be divisible by μ^2, and the determination of all such representations reduces to the determination of the representation of the number m/μ^2 in terms of numbers x, y that have no common divisor.

In order to find all the representations with relatively prime x and y it is only necessary to list, as in subsections 20 and 21 of §2, all the forms (m, N, L) with $N^2 - mL = D$ and $m \geq 2|N|$ (there will be as many such forms as there are solutions N of the congruence $D \equiv N^2 \pmod{m}$ for which $-m/2 \leq N \leq m/2$), since such forms correspond to all the vector-pairs with parameter OM (where

$\overline{OM}^2 = m$). Thus for each vector-pair corresponding to each listed form it is necessary to decide whether it is located in the lattice of the given form. In other words, for each of the listed forms it is necessary to decide whether it is properly equivalent to the given form (since by definition of the corresponding vector-pair we always consider it to be right, i.e., with the same direction of rotation as the given form). If

$$(A, B, C) \begin{bmatrix} x_1, & \beta \\ y_1, & \delta \end{bmatrix} = (m, N, L),$$

then x_1, y_1 is a representation of the number m.

In view of what was said in subsection 20 of §2, the representation thus obtained, together with the associated representations $-x_1, -y_1$ give in the general case all the relatively prime representations of the number m. If the Pell angle is 90°, then there are the further representations

$$-\frac{x_1B + y_1C}{\sigma}, \quad \frac{x_1A + y_1B}{\sigma}$$

and the ones associated with them, and in the case when it is equal to 60° the representations

$$\frac{1}{2}x_1 - \frac{x_1B + y_1C}{\sigma}, \frac{1}{2}y_1 + \frac{x_1A + y_1B}{\sigma};$$

$$\frac{1}{2}x_1 + \frac{x_1B + y_1C}{\sigma}, \frac{1}{2}y_1 - \frac{x_1A + y_1B}{\sigma}$$

and the representations associated with them; here σ designates the greatest common divisor of the numbers A, $2B$ and C.

We do not derive these formulas here, for they are easily obtained on the basis of what was said in subsection 21 of §2.

§5. THE THEORY OF INDEFINITE BINARY QUADRATIC FORMS

54. The indefinite binary quadratic form corresponding to a given vector-pair. A form (A, B, C) with positive determinant D is said to be indefinite. Each such form may be interpreted in the following way. Let $O\xi$, $O\eta$ be arbitrarily chosen asymptotes and let

$$Ax^2 + 2Bxy + Cy^2 = (\xi_1 x + \xi_2 y)(\eta_1 x + \eta_2 y)$$

(if $D = B^2 - AC > 0$, then ξ_1, ξ_2, η_1, and η_2 are real); then we may put in correspondence to this form the vector-pair OPQ, the ends P and Q of whose vectors have coordinates (ξ_1, η_1) and (ξ_2, η_2) with respect to the chosen asymptotes.

Thus to a given vector-pair there will correspond a unique indefinite form, and to the given form, a continuum of such vector-pairs depending on one parameter. In fact, if ρ is an arbitrary real number, then $(\rho\xi_1 x + \rho\xi_2 y)\,(\eta_1 x/\rho + \eta_2 y/\rho)$ will also be a decomposition of the same form, and there will be no other decompositions.

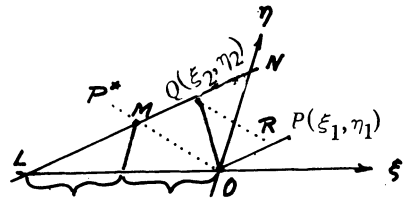

Figure 24

Such a multiplication by ρ, where ρ is positive, is a hyperbolic rotation. Under this rotation the end P of the first vector of the vector-pair OPQ of the form will slide along the hyperbola $\xi\eta = A$, and the end Q of the second vector, along the hyperbola $\xi\eta = C$, while the square of the area of the parallelogram constructed on this vector-pair will remain equal to $(\xi_1\eta_2 - \xi_2\eta_1)^2 = 4(B^2 - AC) = 4D$. But if ρ is negative, the reflected vector-pair will be obtained.

As was already mentioned, $A = \xi_1\eta_1$, $2B = \xi_1\eta_2 + \xi_2\eta_1$, $C = \xi_2\eta_2$. If the square root of the parameter of the hyperbola on which the point lies is now called the *hyperbolic distance* from this point to the point O, then A and C will be the squares of the hyperbolic lengths of the vectors OP and OQ of the vector-pair. The geometric meaning of B is the following. This B is the product of the hyperbolic length of the first vector OP by the hyperbolic length of the orthogonal projection (in the hyperbolic sense) OR of the second vector onto the first vector (Figure 24). Here again, as in subsection 44 of §3, the straight line QR is said to be hyperbolically orthogonal to the straight line OP if it is parallel to the direction of OP^* conjugate with the direction of OP with respect to the asymptotes $O\xi$, $O\eta$ (where OP^* is the diameter conjugate with OP, or, in other words, $LM = MN$, where LN is the chord parallel to OP). In fact, direct calculation gives for the coordinates of the point R the values $k\xi_1$, $k\eta_1$, where $k = \xi_1\eta_2 + \xi_2\eta_1/2\xi_1\eta_1$. From this the product of the hyperbolic lengths of OP and OR is equal to $\sqrt{\xi_1\eta_1}\ \sqrt{k\xi_1 k\eta_1} = k\xi_1\eta_1 = B$.

When translated into the language of vectors this means that the indefinite form $Ax^2 + 2Bxy + Cy^2 = (px + qy)^2$, where the expression $(px + qy)^2$ is the scalar square in the ordinary sense, the length of a vector is understood to be its hyperbolic length, and the length of an orthogonal projection of one vector onto another is understood to be the hyperbolic length of the hyperbolically orthogonal

projection.

The point with coordinates x, y with respect to the frame OPQ has coordinates $\xi_1 x + \xi_2 y, \eta_1 x + \eta_2 y$ with respect to the asymptotes $O\xi$, $O\eta$, i.e., it lies on the hyperbola $\xi\eta = Ax^2 + 2Bxy + Cy^2$ with respect to these asymptotes, or on the hyperbola whose parameter is equal to the value of the form for the pair x, y. Hence this value is equal to the square of the hyperbolic distance from the point x, y to the point O.

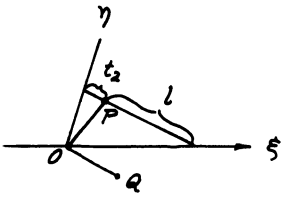

Figure 25

55. *If we want to obtain a form corresponding to the vector-pair for which the ends of the sides have the coordinates* α, γ *and* β, δ *with respect to the vector-pair* OPQ *of a given indefinite form* (A, B, C) *then in exactly the same way as for positive forms we must transform this form by the transformation* $\begin{pmatrix} \alpha, & \beta \\ \gamma, & \delta \end{pmatrix}$. If α, β, γ, and δ are integers and $\alpha\delta - \beta\gamma = \pm 1$, then the second form is also said to be *equivalent* to the first, *properly*, if $\alpha\delta - \beta\gamma = 1$, and *improperly*, if $\alpha\delta - \beta\gamma = -1$. To distinct basic vector-pairs of the same lattice E_2 with respect to the same asymptotes there correspond equivalent forms, i.e., there corresponds to the lattice E_2 with fixed asymptotes and a right direction in its plane a class of forms, whereby a class we mean a collection of properly equivalent forms.

56. Reduced indefinite binary quadratic forms. We will call such a form reduced if its vector-pair is reduced with respect to one of the asymptotes in the sense of subsection 29 of §3, i.e., by Theorem XIII of §3, if it covers one asymptote and the end of its second vector lies further along but closer to it than the end of the first vector. In order to obtain reduction conditions and methods of calculation for the case under consideration it is convenient to consider the geometrical significance of the roots of the equation $A + 2Bt + Ct^2 = 0$, which we will call the *roots of the form* (A, B, C).

If the length of the beak is t, then the coordinates of its end with respect to the vector-pair are $(1, t)$. The parameter of the hyperbola on which this point lies is equal to $A + 2Bt + Ct^2$. Thus the roots t_1 and t_2 of the form are numbers showing how many times it is necessary to lay out the second vector OQ of the form on the beak, starting from the end of its first side P, in order to reach the asymptote (Figure 25). From this we see directly that the vector-pair of the form will include two, one, or no asymptotes depending on whether the form has two,

one or no positive roots. The vector-pair includes one and only one asymptote if its roots have opposite signs, i.e., if $AC < 0$. It is easy to see that the form will be reduced if and only if the positive root $t_1 > 1$, while the negative root t_2 is in absolute value less than unity, $|t_2| < 1$. The *reduction conditions* are consequently the following:

$$0 < B + \sqrt{D} < C < -B + \sqrt{D} \quad \text{for } C > 0,$$
$$0 < -B + \sqrt{D} < |C| < B + \sqrt{D} \quad \text{for } C < 0.$$

The number C can not be equal to zero, for then the point Q would lie on the asymptote, while the asymptotes are assumed to be irrational.

57. In order to pass from an arbitrary indefinite form to the form equivalent to it (properly or improperly) and then to carry out the calculation of a chain of reduced forms following it, we need only to translate the algorithm of subsections 32 and 33 of §3 into the language of calculation.

If the form covers no asymptotes so that both its roots are negative, i.e., $AC > 0$ and $BC > 0$, then by subsection 32 of §3 it is necessary to use the transformation $\begin{pmatrix} 1, & 0 \\ 0, & -1 \end{pmatrix}$ to pass to the form $(A, -B, C)$ that already covers two asymptotes. This is the preparatory transformation. The further transformations of the algorithm of subsection 33 of §3 clearly consist in transformations of the type $\begin{pmatrix} 0, & 1 \\ 1, & \delta \end{pmatrix}$, where δ is the greatest positive integer which is less than the greatest positive root of the form. Hence, if $C > 0$, then $\delta = [(-B + \sqrt{D})/C]$, and for $C < 0$, $\delta = [(B + \sqrt{D})/|C|]$, where $[\ ]$, as always, designates the greatest integer less than or equal to the expression enclosed in the brackets.

These calculations are most simply carried out in the following way. If (A, B, A') is the prepared form and (A', B', A'') is the transformed form, then $B' = B + A'\delta$, where δ has the indicated value. Thus B' satisfies the inequalities $\sqrt{D} - A' \leq B' \leq \sqrt{D}$, if $A' > 0$, and $-\sqrt{D} \leq B' \leq -\sqrt{D} - A'$ for $A' < 0$; or assuming $\lambda = [\sqrt{D}\ ]$, we obtain for B' the inequalities

$$\left.\begin{aligned} \lambda + 1 - A' \leq B' \leq \lambda, &\quad \text{if } A' > 0, \\ -\lambda \leq B' \leq -\lambda - 1 - A', &\quad \text{if } A' < 0. \end{aligned}\right\} \qquad (*)$$

Consequently, it is necessary to look for a $B' = B + A'\delta$ satisfying the corresponding one of the inequalities (*), but this can be at once determined by a glance at the form. Clearly there is always one and only one such B'. We obtain δ in the same way. To obtain A'' we note that $B'^2 - A'A'' = B^2 - AA'$; setting

here $B' = B + A'\delta$, we find

$$A'' = A + (B + B')\delta. \qquad (**)$$

EXAMPLE. Let $(A, B, A') = \phi_0 = (3, 1, -4)$. Here $AC < 0$, i.e., the form ϕ_0 is already prepared. We have $D = B^2 - AC = 13$, $\lambda = [\sqrt{13}] = 3$. Since $A' = -4 < 0$, it is necessary to use the second of the inequalities of (*). We obtain $-3 \leq B' \leq -4 + 4 = 0$, and $B' = 1 - 4\delta$, from which it follows that $\delta = 1$, $B' = -3$; this means, in view of formula (**), that $A'' = 3 + (1 - 3)\delta = 1$. Thus the transformed form (A', B', A'') is the form $\phi_1 = (-4, -3, 1)$.

To obtain the next transformed form $(A'', B'', A''') = \phi_2$, we use the first of the inequalities of (*), since $A'' = 1 > 0$. We obtain $4 - 1 \leq B'' \leq 3$ and $B'' = -3 + 1 \cdot \delta$, from which it follows that $\delta = 6$ and $B'' = 3$. Consequently, on the basis of formula (**), $A''' = -4 + (-3 + 3)6 = -4$, and hence $\phi_2 = (1, 3, -4)$. Further calculations give the forms $\phi_3 = (-4, -1, 3)$, $\phi_4 = (3, 2, -3)$, $\phi_5 = (-3, -1, 4)$, $\phi_6 = (4, 3, -1)$, $\phi_7 = (-1, -3, 4)$, $\phi_8 = (4, 1, -3)$, $\phi_9 = (-3, -2, 3)$ and $\phi_{10} = (3, 1, -4)$, i.e., again ϕ_0.

58. The tabulation of indefinite binary quadratic forms with integer coefficients. Up to now no special assumptions have been made concerning the coefficients of indefinite forms. Now let A, B and C be rational integers, i.e., we are considering indefinite forms with integer coefficients. It turns out that the following fundamental theorem is also true for them: the number of distinct classes of forms with the same discriminant D is finite.

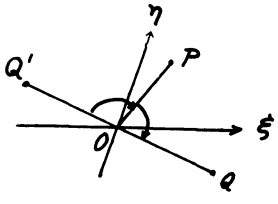

Figure 26

This quickly follows from the fact that each class contains reduced forms. In fact, if one of the reduced forms is improperly equivalent to a given one, then the form following it will be properly equivalent to the given form since two successive forms have opposite directions of rotation. If the form is reduced, then its coefficients A and C are of opposite sign, i.e., $AC < 0$. But $D = B^2 - AC > 0$, i.e., $B^2 < D$, and consequently there are only a finite number of possible values that may be taken on by the coefficient B in the reduced form with given determinant D; and for each such B there are only a finite number of values for A and C since $|A| \cdot |C|$ must be equal to $D - B^2$.

If (A, B, C) is a reduced form of a class and OPQ is its vector-pair, then the vector-pair $OQ'P$ is also reduced and has the same direction of rotation as

the vector-pair OPQ (Figure 26). This means that the form $(C, -B, A)$ corresponding to it will again be a reduced form of the same class. Thus in each class there will be reduced forms for which $B > 0$. B cannot be zero, for then we would obtain from the reduction inequalities that $\sqrt{D} = C$, and the asymptotes would not be irrational. We see from the reduction inequalities that if $B > 0$ for the reduced form, then $A > 0$ and $C < 0$. Hence for such a reduced form we have

$$A = \frac{D - B^2}{|C|} = \frac{(\sqrt{D} + B)(\sqrt{D} - B)}{|C|},$$

$$0 < \frac{-B + \sqrt{D}}{|C|} < 1 \text{ and } 1 < \frac{B + \sqrt{D}}{|C|}.$$

Consequently, for such a reduced form with $B > 0$ we have

$$\sqrt{D} - B < A < \sqrt{D} + B \text{ and } \sqrt{D} - B < |C| < \sqrt{D} + B.$$

Thus, in order to obtain all possible forms with integer coefficients for a positive determinant D and $B > 0$, for each value of B from the sequence 1, 2, 3, $\cdots$, λ, where $\lambda = [\sqrt{D}]$, it is necessary to factor $D - B^2$ in all possible ways into two positive factors lying between $\lambda - B + 1$ and $\lambda + B$ inclusive. Then one of these factors is taken for A, while the other, taken with a negative sign, is set equal to C.

EXAMPLE.

$D = 13$, $\lambda = [\sqrt{13}] = 3$,

$B =$	1,	2,	3,
$D - B^2 =$	12,	9,	4,
$A \cdot \lvert C \rvert =$	$(1 \cdot 12)$,	$(1 \cdot 9)$,	$\boxed{1 \cdot 4}$,
	$(2 \cdot 6)$	$\boxed{3 \cdot 3}$,	$\boxed{2 \cdot 2}$,
	$\boxed{3 \cdot 4}$,	$(9 \cdot 1)$,	$\boxed{4 \cdot 1}$.
	$\boxed{4 \cdot 3}$,		
	$(6 \cdot 2)$,		
	$(12 \cdot 1)$,		

The factorizations enclosed in brackets do not satisfy the inequalities for A and $|C|$. Thus we obtain for $D = 13$ six reduced forms with the coefficient $B > 0$:

$(3, 1, -4), (4, -1, 3), (3, 2, -3), (1, 3, -4), (2, 3, -2), (4, 3, -1).$

However, it is impossible to conclude from this that there are six classes of forms with integer coefficients and determinant $D = 13$, since in the case of an indefinite form some of the reduced forms with $B > 0$ may be equivalent to one another.

59. The periodicity of a chain of reduced forms in the case of a form with integer coefficients. This periodicity follows directly from Theorems XIV and XV (subsections 39 and 40 of §3), since a form with integer coefficients for integer x and y is itself an integer. This means that the parameters of all the hyperbolas containing the points of the lattice E_2 that correspond to the form are integers. Thus, by Theorem XIV all the relative minima lie on a finite number of hyperbolas.

This also follows from the fact just proven, that there are in general only a finite number of reduced forms with integer coefficients and the same determinant D.

If the period consists of k members and OPP_1 is a reduced vector-pair, then under the transformation $\begin{pmatrix} \alpha_1, & \beta_1 \\ \gamma_1, & \delta_1 \end{pmatrix}$ that takes the vector-pair OPP_1 into the vector-pair OP_kP_{k+1} the form (A, B, C) corresponding to the vector-pair OPP_1 goes into itself. All the remaining transformations which take the form into itself are clearly of the form $\pm \begin{pmatrix} \alpha_1, & \beta_1 \\ \gamma_1, & \delta_1 \end{pmatrix}^s$ with positive and negative integer exponents s. These tranformations are called *automorphisms* of the form.

60. The resolution of the question of whether there are two properly equivalent indefinite binary quadratic forms with integer coefficients. If the determinants of the forms under consideration are different the forms will not be equivalent. If the determinants are equal, then by means of the reduction algorithm we find a reduced form properly equivalent to the first of the given forms. It is necessary to note that all the transformations of this algorithm, including the preparatory one if it is necessary, are improper transformations. Thus if we pass to a reduced form by means of an odd number of transformations, it is necessary to calculate the next reduced form in order to obtain a reduced form properly equivalent to the given one.

Then we turn to the second form and calculate for it the whole period of reduced forms. If the reduced form (A, B, C) properly equivalent to the first form, or the reduced form $(C, -B, A)$ properly equivalent to the form (A, B, C), occurs in the indicated period and is therefore obtained from the second form by an even

number of transformations, then the two given forms are properly equivalent; in the contrary case they are improperly equivalent.

In fact, each form possesses reduced forms equivalent to it, and all reduced forms properly equivalent to some form are located either in the chain of its reduced forms or in a chain along another asymptote. But the form of the second chain can be obtained by interchanging the outermost coefficients of the form and changing the sign of the middle coefficient.

61. The representation of numbers by an indeterminate binary quadratic form with integer coefficients. Let m be a given rational integer and let A, B, and C also be rational integers. It is required to find all the representations of the number m in terms of the form (A, B, C). The determination of all such representations is exactly the same problem as the determination of all the points M of the lattice corresponding to the form (A, B, C) that lie on the hyperbola $\xi\eta = m$. As in the case of positive forms it is sufficient again to indicate the method for the determination of all representations with relatively prime values x, y.

In view of subsections 44 and 45 of §3, we will in any case find *all* the relatively prime representations x, y, the points M of which lie on one branch of the hyperbola $\xi\eta = m$, if we find all the distinct vector-pairs corresponding to the parameter OM and then rotate each of these by all the angles which are multiples of the Pell angle of the given form. Other than these representations x, y, there exist only the relatively prime representations $-x, -y$ corresponding to the points symmetric to the points M with respect to the point O and lying on the other branch of the hyperbola $\xi\eta = m$.

The middle coefficient N of the form (m, N, L) corresponding to one of the vector-pairs set into correspondence with the parameter OM satisfies, in view of its geometric interpretation and of the property of the corresponding vector-pair, the inequality $-m/2 < N < m/2$ (in general, there are a finite number of such forms with integer coefficients and with determinant D, namely, as many as the number of roots N of the congruence $D \equiv -N^2 \pmod{m}$ satisfying the condition $-m/2 < N < m/2$).

Thus in order to find *all* the representations of the number m with relatively prime values of x and y, it is necessary to write out all such forms (m, N, L), and then to decide for each of these whether or not it is properly equivalent to the fiven form (A, B, C). Those of them for which this is true give the *fundamental* solutions of the equation $(A, B, C) = m$, since if $(A, B, C)\begin{bmatrix} x_1, & \beta \\ y_1, & \delta \end{bmatrix} = (m, N, L)$,

then x_1, y_1 is a solution, namely that one which corresponds to the end of the first side of the form (m, N, L). In order to obtain all the solutions, it is necessary to find for each such solution the solutions x_s, y_s "homologous" to it in all the remaining Pell angles. For this it is necessary to transform the solution x_1, y_1 by all the automorphisms of the forms (A, B, C), i.e., by all the powers $\begin{pmatrix} \alpha_1, & \beta_1 \\ \gamma_1, & \delta_1 \end{pmatrix}$ of the basic automorphism.

If $\begin{pmatrix} \alpha_1, & \beta_1 \\ \gamma_1, & \delta_1 \end{pmatrix}^s = \begin{pmatrix} \alpha_s, & \beta_s \\ \gamma_s, & \delta_s \end{pmatrix}$, then all these representations obtained by transformation of a basic representation x_1, y_1 have the form $x_s = \alpha_s x_1 + \beta_s y_1$, $y_s = \gamma_s x_1 + \delta_s y_1$. Besides these representations homologous to the representation x_1, y_1 with respect to the periodicity of the hyperbolic rotation, there are associated with the representation x_1, y_1 the further representations $-x_s, -y_s$ symmetric with them with respect to the point O.

62. Connection with the Pell equation. All the automorphic transformations $\begin{pmatrix} \alpha_s, & \beta_s \\ \gamma_s, & \delta_s \end{pmatrix}$ and all the representations x_s, y_s homologous to the representation x_1, y_1 may be calculated much more conveniently if we turn to the parameter ρ of the Pell angle of the given form.

We consider an arbitrary nonzero point Q on the asymptote and let its coordinates with respect to the vector-pair OPP_1 be x, y. Then with respect to the vector-pair OP_kP_{k+1} (here we again assume that the period consists of k members) it has coordinates $x' = x/\rho$, $y' = y/\rho$, where ρ is the parameter of the Pell angle. From this we obtain $x/y = (\alpha_1 x + \beta_1 y)/(\gamma_1 x + \delta_1 y)$ or $\gamma_1 x^2 + (\delta_1 - \alpha_1)xy - \beta_1 y^2 = 0$. But since the point Q lies on the asymptote we have the further relation $Ax^2 + 2Bxy + Cy^2 = 0$. Hence $A : 2B : C = \gamma_1 : \delta_1 - \alpha_1 : -\beta_1$.

Among other things we here obtain the converse to the theorem about the periodicity of a lattice of a form with integer coefficients.

THEOREM XVII. *If a chain of reduced forms is periodic, then the coefficients A, B and C of the form are proportional to rational integers.*

In other words, if a lattice periodically repeats under hyperbolic rotation with respect to a given asymptote, then its form is proportional to a form with integer coefficients.

Let σ be the greatest common divisor of the numbers A, $2B$, and C. Then $\gamma_1 = Au_1/\sigma$, $\delta_1 - \alpha_1 = 2Bu_1/\sigma$, and $-\beta_1 = Cu_1/\sigma$, where u_1 is a rational integer.

If we have $\alpha_1 + \delta_1 = 2t_1/\sigma$, then we get $\alpha_1 = (t_1 - Bu_1)/\sigma$, $\beta_1 = -Cu_1/\sigma$, $\gamma_1 = Au_1/\sigma$, and $\delta_1 = (t_1 + Bu_1)/\sigma$. Or, since $\alpha_1\delta_1 - \beta_1\gamma_1 = 1$ (namely $+1$, since the two vector-pairs OPP_1 and OP_kP_{k+1} are identically oriented), we have $t_1^2 - Du_1^2 = \sigma^2$, i.e., we obtain the Pell equation. If (ξ_1, η_1) and (ξ_2, η_2) are the coordinates of the points P and P_1, while (ξ_1', η_1') and (ξ_2', η_2') are the coordinates of the points P_k and P_{k+1}, then $\xi_1' = \alpha_1\xi_1 + \gamma_1\xi_2$ and $\xi_2' = \beta_1\xi_1 + \delta_1\xi_2$. Moreover, we have $\xi_1' = \rho\xi_1$ and $\xi_2' = \rho\xi_2$; thus we obtain

$$(\alpha_1 - \rho)\xi_1 + \gamma_1\xi_2 = 0, \quad \beta_1\xi_1 + (\delta - \rho)\xi_2 = 0,$$

from which it follows that

$$\rho^2 - (\alpha_1 + \delta_1)\rho + (\alpha_1\delta_1 - \beta_1\gamma_1) = 0.$$

If we substitute here the expressions just obtained for α_1, β_1, γ_1, and δ_1 the equation will take the form $\rho^2 - 2t_1\rho/\sigma + (t_1^2 - Du_1^2)/\sigma^2 = 0$ or $\rho = (t_1 \pm \sqrt{D}\,u_1)/\sigma$.

The transition from the vector-pair OPP_1 to the vector-pair $OP_{sk}P_{sk+1}$ corresponds to the parameter ρ^s which is associated with the magnitudes of α_s, β_s, γ_s, and δ_s in the same way that the parameter ρ is associated with the magnitudes of α_1, β_1, γ_1 and δ_1. From this we obtain first, that in the relation

$$\left(\frac{t_1 + \sqrt{D}\,u_1}{\sigma}\right)^s = \frac{t_s + \sqrt{D}\,u_s}{\sigma}$$

t_s and u_s are integers, and secondly, that

$$\alpha_s = \frac{t_s - Bu_s}{\sigma}, \quad \beta_s = \frac{-Cu_s}{\sigma}, \quad \gamma_s = \frac{Au_s}{\sigma}, \quad \delta_s = \frac{t_s + Bu_s}{\sigma}.$$

Hence all the solutions homologous to the solution x_1, y_1 are obtained in the following form:

$$x_s = \frac{1}{\sigma}[x_1t_s - (x_1B + y_1C)u_s],$$

$$y_s = \frac{1}{\sigma}[y_1t_s + (x_1A + y_1B)u_s].$$

63. The case of a form with integer coefficients whose determinant is a perfect square. This case is not covered by the above theorem since the asymptotes are rational. It is possible to show that all the different classes of such forms for which $D = d^2$ have as their representatives the forms

$$(0, d, -d+1), \ (0, d, -d+2), \cdots, (0, d, 0), \cdots, (0, d, d-1), (0, d, d).$$

Since

$$Ax^2 + 2Bxy + Cy^2 = \left(\frac{-B+d}{C}x - y\right)\left(\frac{-B-d}{C}x - y\right)C,$$

the representation of the number m in terms of such a form leads simply to the solution of the determinate system of equations

$$(-B+d)x - Cy = m_1, \quad (-B-d)x - Cy = m_2$$

in integers x, y for all factorizations of the number mC into integer factors m_1, m_2.

SUPPLEMENT TWO

INVESTIGATIONS IN THE GEOMETRY OF GALOIS THEORY [1)]

B. N. DELONE and D. K. FADDEEV

§1. THE THEORY OF R-ALGEBRAS

1. The space K_n as an algebra. Let there be given a completely arbitrary (commutative) field K and an n-dimensional vector space over it. We choose in the field a basis $\mathbf{E}_n$ and besides the operations of vector addition and subtraction and multiplication by scalars (elements of K), we introduce the operation of multiplication of vectors (points). The product of two vectors will be the vector whose coordinates with respect to the chosen basis (the initial basis) are equal to the products of the corresponding coordinates of the vectors being multiplied. Under the introduction of this operation the vector space becomes a commutative algebra, which we will denote by K_n. Among the K-linear subspaces of K_n we note the *coordinate subspaces* spanned by a subset of the vectors of the initial basis, and the *bisectrices*, namely the sets of all points having equal coordinates with respect to complexes of vectors into which all the vectors of the initial basis can be decomposed. By the *initial basis of a bisectrix* we mean the collection of all vectors, each of whose coordinates is equal to 1 in one of the complexes of vectors of the initial coordinate basis $\mathbf{E}_n$ characterizing the bisectrix, and is equal to zero in the remaining axes. Correspondingly, the *initial basis of a coordinate subspace* will be the collection of vectors from the initial basis $\mathbf{E}_n$ by which the subspace is spanned. Bisectrices and coordinate subspaces are subalgebras of K_n, while multiplication of points of a bisectrix (coordinate subspace) with respect to the initial basis of the whole K_n coincides with multiplication defined for the bisectrix (coordinate subspace) with respect to its initial basis.

LEMMA 1. *Any K-subalgebra Q of the algebra K_n is either a coordinate*

1) This Supplement is a translation of the first four sections of the article in Mat. Sb. (N.S.) 15(57) (1944), 244–254.

subspace, a bisectrix of $\mathbf{E}_n$, *or a bisectrix of a coordinate subspace.*

PROOF. We choose a numbering of the axes so that points with nonzero first coordinates are located in Q. Among these points we choose a point ω, with the least number of nonzero coordinates. Then all the nonzero coordinates of the point ω are equal. In fact, if we assume that ω has a coordinate $\omega^{(2)}$ different from the first coordinate $\omega^{(1)}$ and different from zero, then the point $\omega' = \omega^2 - \omega^{(2)} \cdot \omega$, belonging to Q, would have a nonzero first coordinate and a smaller number of nonzero coordinates then ω. Together with the point ω, Q contains the point $\epsilon_1 = \omega^{(1)-1} \cdot \omega$, whose coordinates with the given numbering of axes are $(\underbrace{1, 1, \cdots, 1}_{n_1}, 0, 0, \cdots, 0)$. Moreover, the first n_1 coordinates of any point τ in Q are equal, for if there were a coordinate $\tau^{(2)}$ of the point τ, taken from among the first n_1 coordinates and not equal to the first coordinate $\tau^{(1)}$, then the point $\tau' = \tau \cdot \epsilon_1 - \tau^{(2)} \cdot \epsilon_1$ would have a nonzero first coordinate and would have fewer nonzero coordinates than ω. Applying the same argument to each coordinate axis for which there exist points of Q with the corresponding coordinate different from zero, we see that all the points of Q have coordinates

$$(\underbrace{\omega^{(1)}, \cdots, \omega^{(1)}}_{n_1}, \cdots, \underbrace{\omega^{(m)}, \cdots, \omega^{(m)}}_{n_m}, \underbrace{0, \cdots, 0}_{n'}),$$

where the points

$$\epsilon_1 = (\underbrace{1, \cdots, 1}_{n_1}, 0, \cdots, 0), \cdots, \epsilon_m = (0, \cdots, 0, \underbrace{1, \cdots, 1}_{n_m}, \underbrace{0, \cdots, 0}_{n'})$$

belong to Q, and therefore $\omega^{(1)}, \cdots, \omega^{(m)}$ are arbitrary elements of K. If $n_1 = \cdots = n_m = 1$, then Q is a coordinate subspace, and if $n' = 0$ then Q is a bisectrix of K_n; in the remaining cases Q is the bisectrix of a coordinate subspace. The lemma is proved.

2. *R-algebras of the space* K_n. *Decomposition into a direct sum.* Let there be given some field R contained in K, and let a basis $\mathfrak{E}_n$ be selected in K_n, which is, in general, not the initial one. By an R-module we will mean a collection of all points of K_n having elements of the field R as coordinates with respect to $\mathfrak{E}_n$. An R-module whose points are reproduced under multiplication will be called an R-algebra. An example of an R-algebra is an R-module constructed on the initial basis $\mathbf{E}_n$.

A linear subspace of K_n that is a K-linear envelope of some collection of vectors of an R-algebra $\mathbf{A}$ will be called an $\mathbf{A}$-*complete* subspace, and the collec-

tion of points of **A** included in it will be called its **A**-*completion.* In particular, all of K_n is **A**-complete. The dimension of an **A**-complete subspace is equal to the dimension (with respect to R) of its **A**-completion for, as it is easy to see, vectors of any R-module which are linearly independent with respect to R remain linearly independent with respect to K. Clearly, the vector sum of two **A**-complete subspaces is **A**-complete. It is also easy to see that the intersection of two **A**-complete subspaces is **A**-complete, since the dimension (with respect to K) of this intersection and the dimension of the completions contained in it are equal to the same number, namely, to the amount by which the sum of the dimensions of the given subspaces exceeds the dimension of their vector sum. A linear transformation of K_n which takes all the points of an R-algebra **A** into points of **A** will be said to be *rational* with respect to **A**. Clearly, linear transformations rational with respect to **A**, singular or nonsingular take any **A**-complete subspace of K_n (in particular all of K_n) into an **A**-complete subspace.

LEMMA 1′. *Subalgebras (over R) of an R-algebra* **A** *are* **A**-*completions of* **A**-*complete subalgebras of K_n, i.e., they are bisectrices, coordinate subspaces, or bisectrices of coordinate subspaces.*

In fact, if **B** is a subalgebra of **A**, then its K-linear envelope $\bar{\mathbf{B}}$ is a subalgebra of K_n. Any m-dimensional subalgebra of the R-algebra **A** may be considered as an R-algebra of the space K_m, put into the corresponding coordinate subspace or bisectrix, i.e., as an R-algebra of this subspace with respect to its initial basis.

A *zero divisor* will be any point that has both zero and nonzero coordinates. It is easy to see that in any R-algebra it is possible to divide by any of the points that are not zero or zero divisors. In fact, multiplication of K_n by such a point λ is a nonsingular linear transformation rational with respect to **A** which takes all of **A** into all of **A**. Hence, for each point μ in **A** there is a point μ' in **A** such that $\mu'\lambda = \mu$, i.e., $\mu' = \mu/\lambda$. It further follows from this that any R-algebra that contains at least one point that is not a zero-divisor and is not equal to zero will contain the identity $(1, 1, \cdots, 1)$.

If the R-algebra **A** can be represented in the form of a direct sum of two R-algebras, i.e., in the form of the vector sum of two algebras completing the complementary coordinate subspaces, then **A** is said to be a reducible algebra, while if such a representation is impossible, **A** is said to be irreducible. Any reducible R-algebra contains zero-divisors; for example, all the points of the algebras which are being added. The very important converse also holds.

LEMMA 2. *If an R-algebra contains a zero-divisor ω, then it is reducible and decomposes into the direct sum of two R-algebras, one of which is in the coordinate subspace K_{n_1} which includes those axes for which the coordinates of ω are different from zero, while the other R-algebra is in the complementary coordinate subspace.*

PROOF. We multiply K_n by the point ω. When this is done, K_n goes into K_{n_1}. Since multiplication by a point $\omega \in \mathbf{A}$ is a rational transformation with respect to $\mathbf{A}$, therefore K_{n_1} will be $\mathbf{A}$-complete. The completion of K_{n_1} is clearly an algebra and will be denoted by $\mathbf{A}_1$. The point ω belongs to it and is not a zero-divisor in it; hence A_1 includes a unit ϵ_1 of the space K_{n_1}. Further, the transformation $\alpha' = \alpha - \alpha\epsilon_1$ is rational with respect to A and takes K_n into the space K_{n_2} that is complementary to K_{n_1}. This subspace is also $\mathbf{A}$-complete. Its completion forms in its turn an algebra $\mathbf{A}_2$. Since for any $\alpha \in \mathbf{A}$, it is true that $\alpha = \alpha\epsilon_1 + (\alpha - \alpha\epsilon_1)$, $\alpha\epsilon_1 \in \mathbf{A}_1$, $\alpha - \alpha\epsilon_1 \in \mathbf{A}_2$ and $K_{n_1} \cap K_{n_2} = 0$, the algebra $\mathbf{A}$ is the direct sum of $\mathbf{A}_1$ and $\mathbf{A}_2$.

It is now easy to prove the following important theorem:

THEOREM 1. *Each R-algebra $\mathbf{A}$ is either irreducible or uniquely decomposable into the direct sum of irreducible algebras.*

PROOF. If $\mathbf{A}$ is reducible we decompose it into the direct sum of two algebras; if one or both of these is reducible we continue the decomposition, and so on. The process of decomposition must terminate, since the number of direct summands of $\mathbf{A}$ cannot exceed its dimension. Thus $\mathbf{A}$ is decomposed into the direct sum of irreducible R-algebras $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_k$. Since an irreducible algebra can not contain a zero-divisor, the complexes of vectors of the initial basis pertaining to the subspaces containing $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_k$ may be characterized by the fact that the coordinates of all of the points of $\mathbf{A}$ that correspond to each separate complex either vanish or do not vanish simultaneously. This holds for every complex of coordinate vectors of each subspace containing an irreducible summand of the algebra $\mathbf{A}$, and hence such a summand must coincide with one of the $\mathbf{A}_1$, $\mathbf{A}_2, \cdots, \mathbf{A}_k$. The theorem is proved.

We note that it follows from this theorem that any R-algebra contains a unit of K_n. In fact, each irreducible R-algebra contains a unit because of the possibility of division by any one of its points other than zero, while the sum of units of all the subspaces containing irreducible summands of the algebra is clearly a unit of the whole K_n.

3. The direct product of algebras. Let there be given points α and β

in the spaces K_m and K_n with coordinates $(\alpha^{(1)}, \alpha^{(2)}, \cdots, \alpha^{(m)})$ and $(\beta^{(1)}, \beta^{(2)}, \cdots, \beta^{(n)})$. We associate with them a point of the space K_{mn} with coordinates $(\cdots, \alpha^{(i)} \cdot \beta^{(j)}, \cdots)$, $i = 1, \cdots, m$; $j = 1, \cdots, n$. The point constructed in such a manner will be called the *composite point* of α and β and will be denoted by $\alpha * \beta$. It is easy to convince oneself of the validity of the following rules for operations:

(1^1) $\alpha * (\beta_1 + \beta_2) = (\alpha * \beta_1) + (\alpha * \beta_2)$,

(1^2) $(\alpha_1 + \alpha_2) * \beta = (\alpha_1 * \beta) + (\alpha_2 * \beta)$,

(1^3) $\alpha * (a\beta) = (a\alpha) * \beta = a(\alpha * \beta)$,

where a denotes any scalar;

(2) $\alpha_1\alpha_2 * \beta_1\beta_2 = (\alpha_1 * \beta_1)(\alpha_2 * \beta_2)$.

The *composite of two point collections* **A** and **B** will be the collection of the composites of all the points of **A** with the points of **B**. Finally, the *direct product* **A** * **B** of the collections **A** and **B** (which are assumed to be additive and subtractive) will be the collection of all points obtained from the composites of **A** and **B** by addition and subtraction.

THEOREM 2. *The direct product of two R-algebras is an algebra.*

PROOF. From the definition of the direct product and from properties (1^1), (1^2), and (1^3) for composites of points, it follows that the direct product of linear envelopes (with respect to the field K or to one of its subfields) of two collections is the linear envelope of their composites. In particular, $K_m * K_n = K_{mn}$, for the composite of the initial basis of K_m and K_n is the initial basis of K_{mn}. Further, the composite of any coordinate systems $\mathfrak{E}_m$ and $\mathfrak{E}_n$ of the spaces K_m and K_n is always the basis of K_{mn}, since its K-linear envelope is equal to $K_m * K_n = K_{mn}$ and it consists of exactly mn vectors. From this it follows that the direct product of R-algebras is an R-module based on the composites of its bases. This R-module, in view of rule (2), is reproduced by multiplication, i.e., is an R-algebra, which is what we wanted to show.

§2. THE GALOIS GROUP OF AN R-ALGEBRA

By *axial-superpositions* we will mean linear transformations of K_n by which the vectors of the initial coordinate basis are only permuted among themselves. An *axial-superposition into itself* of an R-algebra will be an axial-superposition that takes the R-algebra into itself. Axial-superpositions of an R-algebra into itself are automorphisms of the R-algebra and clearly form a group. There exist

R-algebras admitting all $n!$ possible axial-superpositions; an example is an R-algebra constructed on the initial basis. The other extreme case is also possible, when the R-algebra does not have any axial-superpositions into itself other than the identity. For an irreducible R-algebra the number of axial-superpositions into itself cannot exceed its dimension; namely, the following assertion is true.

LEMMA 1. *Among the axial-superpositions of an irreducible R-algebra into itself there exists not more than one taking some vector of the initial basis into a particular other vector of the initial basis, i.e., there are no more than n axial-superpositions.*

PROOF. Let us assume the contrary, namely, that two distinct axial-superpositions into itself σ_1 and σ_2 of an R-algebra $\mathbf{A}$ take a vector $\mathbf{e}_1$ of the initial basis into the same vector $\mathbf{e}_2$. Then the axial-superposition $\sigma_3 = \sigma_2\sigma_1^{-1} \neq 1$ takes $\mathbf{e}_1$ into itself. Here there exists a point α of $\mathbf{A}$ such that $\alpha^{\sigma_3} = \beta \neq \alpha$, since $\sigma_3 \neq 1$. The first coordinates of the points α and β are equal, since $\mathbf{e}_1^{\sigma_3} = \mathbf{e}_1$ and hence the point $\beta - \alpha$ of $\mathbf{A}$ turns out to be a zero-divisor, which is impossible in view of the irreducibility of $\mathbf{A}$. The lemma is proved.

We will say that an R-algebra of the space K_n is *normal* if it is irreducible and has n axial-superpositions into itself.

THEOREM 1. *Every irreducible R-algebra of $\mathbf{A}$ is a subalgebra of some normal algebra.*

PROOF. We take n copies of the algebra $\mathbf{A}$, denoting them by $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$ and we form their direct product $\mathbf{D}$. It is an R-algebra in a space of dimension n^n. We will number the vectors of the initial coordinate system of this space with sets of indices $(j_1, j_2, \cdots, j_n)$, where we lay out on the axes $\mathbf{e}_{j_1 j_2 \cdots j_n}$ the products of the j_1th coordinates of the points of $\mathbf{A}_1$, the j_2th coordinates of the points of $\mathbf{A}_2$, and so on to the j_nth coordinates of the points of $\mathbf{A}_n$ and the sums of such products. We consider now the $n!$ axes whose "numbers" do not include identical indices, and we show that $\mathbf{D}$ has a direct summand $\mathbf{C}$ which is contained in the coordinate subspace $K_{n!}$ based on these axes. We introduce into the algebra $\mathbf{A}$ the basis $\omega_1, \cdots, \omega_n$; we denote by $\omega_{1,i}, \cdots, \omega_{n,i}$ the corresponding bases of the algebras $\mathbf{A}_i$, and we construct the point β, represented symbolically in the form of the determinant

$$\beta = \begin{vmatrix} \omega_{11} & \omega_{12} & \cdots & \omega_{1n} \\ \omega_{21} & \omega_{22} & \cdots & \omega_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \omega_{n1} & \omega_{n2} & \cdots & \omega_{nn} \end{vmatrix}.$$

In the calculation of this determinant it is necessary to consider its individual elements as composites of points of the algebras $\mathbf{A}_1, \cdots, \mathbf{A}_n$, while addition must be understood as addition in the algebra $\mathbf{D}$. Clearly, the coordinate of the point β that corresponds to the axis $\mathbf{e}_{j_1 j_2 \cdots j_n}$ is equal to the determinant whose columns are composed respectively of the $j_1, j_2, \cdots, j_n$th coordinates of the points $\omega_1, \omega_2, \cdots, \omega_n$. Hence all the coordinates of the point β that correspond to axes whose "numbers" contain equal indices are equal to zero, while the coordinates corresponding to axes without equal indices are equal, up to the sign, to the determinant consisting of all the coordinates of the points $\omega_1, \omega_2, \cdots, \omega_n$, and are thus different from zero. Hence, in view of Lemma 2, the algebra $\mathbf{D}$ actually has a direct summand $\mathbf{C}$ that is contained in the coordinate subspace $K_{n!}$. We now rearrange in some manner $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$. Then their direct products will not change while the axis with the number $(j_1, j_2, \cdots, j_n)$ will go into the axis with the number in which the indices $j_1, j_2, \cdots, j_n$ undergo the corresponding permutation. The algebra $\mathbf{C}$ is transformed into itself and any axis of its initial basis may be taken into any other axis of the same basis by means of the appropriate choice of the permutation of $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$. The algebra $\mathbf{C}$ may turn out to be reducible; clearly, in this case it is the direct sum of identical normal algebras. In fact, let σ be an axial-superposition into itself of the algebra $\mathbf{C}$ (induced by some permutation of $\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_n$), which takes some axis of the subspace containing an irreducible summand of $\mathbf{B}$ into another axis of the same subspace. This axial-superposition takes $\mathbf{B}$ into some irreducible summand of the algebra $\mathbf{D}$ which must coincide with $\mathbf{B}$, for coordinate subspaces containing distinct irreducible summands of algebras intersect only in zero. Hence, such an axial-superposition of $\mathbf{C}$ induces an axial-superposition of $\mathbf{B}$ into itself, where any axis of the initial basis of the algebra $\mathbf{B}$ may be taken into any other axis of its initial basis, i.e., $\mathbf{B}$ is normal. Having considered an axial-superposition taking some axis of an irreducible summand $\mathbf{B}$ into an axis of another irreducible summand $\mathbf{B}'$, we see that it takes $\mathbf{B}$ into $\mathbf{B}'$. Thus we see that all the direct summands of $\mathbf{C}$ are normal and may be obtained from one of the $\mathbf{B}$ by axial-superpositions, i.e., they are equal to each other.

Further, the algebra $\mathbf{D}$ contains the composite $\mathbf{A}_1 * 1 * \cdots * 1 = \widetilde{\mathbf{A}}_1$, which is nothing other than the algebra $\mathbf{A}$ based on some bisectrix; namely, the first coordinates of points of $\mathbf{A}$ are laid out on the complex of axes, the first index of which is equal to 1, the second on the complex of axes, the first index of which is equal to 2, and so on. Let e be a unit of the space $K_{n!}$. Then clearly $e \in \mathbf{C}$ and $\widetilde{\mathbf{A}}_i e \subset \mathbf{C}$ is again an algebra $\mathbf{A}$ put into the corresponding bisectrix of $K_{n!}$. Thus, $\mathbf{A}$ is a subalgebra of $\mathbf{C}$, based on the bisectrix of all of $K_{n!}$. But, in any case, $\mathbf{A}$ is also a subalgebra for $\mathbf{B}$, which follows from the following lemma:

LEMMA 2. *If an irreducible R-algebra* $\mathbf{A}$ *is a subalgebra of an R-algebra* $\mathbf{C}$ *and is located on the bisectrix of all of its space, then* $\mathbf{A}$ *is a subalgebra of each irreducible summand of* $\mathbf{C}$.

PROOF. Let $\mathbf{C} = \mathbf{B}_1 + \cdots + \mathbf{B}_k$ be the decomposition into irreducible summands, let $K_{n_1}, \cdots, K_{n_k}$ be the subspaces containing $\mathbf{B}_1, \cdots, \mathbf{B}_k$, let $K_{n'_1}, \cdots$ $\cdots, K_{n'_k}$ be the subspaces complementary to them, and let $\epsilon_1, \cdots, \epsilon_k$ be units of the algebras $\mathbf{B}_1, \cdots, \mathbf{B}_k$. Then clearly the collection $\epsilon_j \mathbf{A}$ is a subalgebra of $\mathbf{B}_j$, the coordinate points of which coincide with the coordinates of the corresponding points of $\mathbf{A}$, where all the coordinates will be represented unless multiplication by ϵ_j annihilates some vector of the initial coordinate system of the bisectrix containing $\mathbf{A}$. But this situation would be possible only if the bisectrix containing $\mathbf{A}$ had an intersection with the space $K_{n'_j}$ other than zero, which is impossible. In fact, if there were a bisectrix containing $\mathbf{A}$ with a nonzero intersection with $K_{n'_j}$, then this intersection, being the intersection of $\mathbf{C}$-complete subspaces, would be $\mathbf{C}$-complete, and points of this intersection would be zero-divisors and would be included in $\mathbf{A}$, which is impossible, in view of the irreducibility of $\mathbf{A}$. Thus each algebra $\epsilon_j \mathbf{A}$ is equal to $\mathbf{A}$ based on some bisectrix K_{n_j} in a corresponding way, i.e., $\mathbf{A}$ is a subalgebra of each $\mathbf{B}_j$. Lemma 2 and Theorem 1 are now proved in full.

We note that, starting with the reducible algebra $\mathbf{A}$, by the same construction we could have constructed a normal algebra $\mathbf{B}$. Having generalized Lemma 4 in a corresponding manner, it is easy to see that the normal algebra so constructed contains all the irreducible summands of $\mathbf{A}$ as subalgebras. Further, it is possible to show that the normal algebra $\mathbf{B}$ thus constructed is minimal among all the normal algebras that contain all the irreducible summands.

§3. BASIC THEOREMS OF GALOIS FOR R-ALGEBRAS

Let $\mathfrak{G}$ be the group of axial-superpositions of a normal algebra $\mathbf{A}$ and let $\mathfrak{H}$ be one of its subgroups. We decompose all the axes of the initial coordinate system into complexes, putting into one complex all the axes that can be obtained from one another by means of axial-superpositions from the group $\mathfrak{H}$. It is easy to see that such complexes do not have common axes, that the number of axes in each complex is equal to the order of $\mathfrak{H}$, and hence that the number of complexes is equal to the index of $\mathfrak{H}$. The bisectrix of the space K_n determined by such a decomposition of the initial coordinate system will be called the bisectrix *belonging to the subgroup* $\mathfrak{H}$. Its points may be characterized by the fact that they and only they remain fixed under all the axial-superpositions from $\mathfrak{G}$.

THEOREM 1. *Any R-subalgebra of a normal R-algebra* $\mathbf{A}$ *is the completion of a bisectrix belonging to some subgroups of the group of axial-superpositions* $\mathfrak{G}$ *of the algebra* $\mathbf{A}$ *into itself. Conversely, a bisectrix belonging to some subgroups of the group* $\mathfrak{G}$ *is* $\mathbf{A}$*-complete and its completion is a subalgebra of* $\mathbf{A}$.

PROOF. Let there be given a subalgebra $\mathbf{B}$ of a normal algebra $\mathbf{A}$. It completes some bisectrix $\overline{\mathbf{B}}$ of the whole K_n. Let this bisectrix be characterized by the equality of coordinates in the complexes of axes

$$(\mathbf{e}_1, \cdots, \mathbf{e}_m), \quad (\mathbf{e}_{m+1}, \cdots), \cdots.$$

We consider the collection $\mathfrak{H}_1$ of all axial-superpositions taking the axis $\mathbf{e}_1$ into the axes of the first complex. These axial-superpositions do not change all the points of $\mathbf{B}$, for if one of these axial-superpositions $\omega \in \mathbf{B}$ went into a different point ω', then, in view of the normality of $\omega' \in \mathbf{A}$, $\mathbf{A}$ would contain the zero divisor $\omega - \omega'$. Each axial-superposition taking $\mathbf{e}_1$ into the axis $\mathbf{e}_k$ of another complex changes at least one point of the subalgebra $\mathbf{B}$, since there exist in it points that have unequal coordinates in the axes $\mathbf{e}_1$ and $\mathbf{e}_k$. Thus the collection $\mathfrak{H}_1$ coincides with the collection $\mathfrak{H}$ of all the axial-superpositions that do not change all the points of the subalgebra $\mathbf{B}$. From these considerations we conclude that the collection of all axial-superpositions taking any axis into all the axes of the complex containing it coincides with the collection $\mathfrak{H}$. Clearly $\mathfrak{H}$ is a group, and the bisectrix containing $\mathbf{B}$ belongs to this group.

Let there now be given some subgroup $\mathfrak{H}$ of the group $\mathfrak{G}$ of axial-superpositions into itself of the normal algebra $\mathbf{A}$. We associate with each point $\alpha \in K_n$ the sum of all the points obtained from α by axial-superpositions from the group $\mathfrak{H}$. This determines a linear transformation of K_n which is rational with respect

to **A** because of its normality, and which takes all of K_n into a bisectrix belonging to $\mathfrak{H}$. Hence this bisectrix is **A**-complete and its completion is clearly a subalgebra of the algebra **A**. The theorem is proved in full.

THEOREM 2. *If $\mathfrak{H}$ is a normal divisor of the group $\mathfrak{G}$ of axial-superpositions of a normal R-algebra* **A**, *then the subalgebra* **B** *belonging to $\mathfrak{H}$ is normal and its group of axial-superpositions is isomorphic to $\mathfrak{G}/\mathfrak{H}$.*

PROOF. Let $\mathfrak{G} = \mathfrak{H} + \sigma_2\mathfrak{H} + \cdots + \sigma_k\mathfrak{H}$ and let $\mathbf{e}_1$ be one of the vectors of the initial coordinate system of K_n. Then the vectors of the initial coordinate system belonging to $\mathfrak{H}$ are $e_k = \sum_{\tau \in \sigma_k \mathfrak{H}} \mathbf{e}^{\tau}$. It is easy to see that if $\mathfrak{H}$ is a normal divisor of $\mathfrak{G}$, then each axial-superposition of $\mathfrak{G}$ takes into itself the initial basis of the bisectrix belonging to $\mathfrak{H}$ and consequently the whole bisectrix. In fact

$$e_k^{\sigma} = \sum_{\tau \in \sigma_k \mathfrak{H}} \mathbf{e}_1^{\tau\sigma} = \sum_{\tau' \in \sigma_k \mathfrak{H}\sigma} \mathbf{e}_1^{\tau'} = \sum_{\tau' \in \sigma_k \sigma \mathfrak{H}} \mathbf{e}_1^{\tau'}.$$

Axial-superpositions induced in the bisectrix form a group homomorphic to $\mathfrak{G}$ with the kernel of the homomorphism being $\mathfrak{H}$. Hence this group is isomorphic to $\mathfrak{G}/\mathfrak{H}$. Its order is equal to the dimension of the bisectrix. Under all these superpositions **B** goes into itself, and hence, being irreducible since it is a subalgebra of an irreducible algebra **A**, it is normal. The theorem is proved.

§4. CONNECTION WITH THE PRESENT-DAY PRESENTATION OF GALOIS THEORY

1. On generic points. A generic point of the space K_n is a point whose coordinates are distinct and different from zero.

LEMMA 1. *Any irreducible R-algebra* **A** *contains a generic point.*

PROOF. Let us consider separately the cases when R consists of a finite and of an infinite number of elements.

Let the field R be finite. Then any irreducible R-algebra **A** will also be a finite field. As is well known, all the nonzero elements of a finite field form a *cyclic* group with respect to multiplication. Let α be a generator of this group. Then α is a generic point, for if two of its coordinates were equal, i.e., if α lay in some bisectrix, then all the elements of **A** would lie in the same bisectrix and, except for zero, would all be powers of α, which is impossible.

Now let R be infinite. We find first of all in the R-algebra **A** points $\alpha_2, \alpha_3, \cdots, \alpha_n$ such that the ith coordinate $\alpha_i^{(i)}$ of the point α_i is different from its first coordinate $\alpha_i^{(1)}$. Such points may be found, for otherwise **A** would

be in some bisectrix. We now construct successively the points $\beta_2, \beta_3, \cdots, \beta_n$ such that all the coordinates from the second to the ith of the points β_i are different from its first coordinate. As β_2 we take α_2, constructing the successive points inductively. Let β_i be already constructed. We look for β_{i+1} in the form $\beta_{i+1} = \beta_i + x\alpha_{i+1}$, where $x \in R$. For our purpose we must choose x so that none of the following equalities are fulfilled:

$$\begin{aligned}
\beta_i^{(2)} &+ x\alpha_{i+1}^{(2)} = \beta_i^{(1)} + x\alpha_{i+1}^{(1)},\\
&\cdots\cdots\cdots\cdots\cdots\\
\beta_i^{(i)} &+ x\alpha_{i+1}^{(i)} = \beta_i^{(1)} + x\alpha_{i+1}^{(1)},\\
\beta_i^{(i+1)} &+ x\alpha_{i+1}^{(i+1)} = \beta_i^{(1)} + x\alpha_{i+1}^{(1)}.
\end{aligned}$$

None of these equalities is an identity (the last in view of the choice α_{i+1}, the remaining in view of the construction of the point β_i). Hence each of these equalities can be satisfied for not more than one value of x. Taking x different from this finite number of excluded values, we construct the point $\beta_{i+1} = \beta_i + x\alpha_{i+1}$, satisfying the given requirement. The last point β_n, which we denote by γ_1, has the property that its first coordinate is different from all the remaining coordinates.

In the same way we construct points $\gamma_2, \cdots, \gamma_n$ so that the ith coordinate of the point γ_i is different from all the remaining coordinates of that point. Then we pass to the inductive construction of the points $\delta_2, \delta_3, \cdots, \delta_n$ such that the first i coordinates of the point δ_i are distinct. As δ_2 we take γ_2. Let δ_i be already constructed. We look for the point δ_{i+1} in the form $\delta_i + x\gamma_{i+1}$ for $x \in R$. We must choose x so that none of the following equalities is satisfied:

$$\begin{aligned}
\delta_i^{(j)} + x\gamma_{i+1}^{(j)} &= \delta_i^{(k)} + x\gamma_{i+1}^{(k)}, && j,\ k \le i,\\
\delta_i^{(j)} + x\gamma_{i+1}^{(j)} &= \delta_i^{(i+1)} + x\gamma_{i+1}^{(i+1)}, && j = 1, 2, \cdots, i.
\end{aligned}$$

The equalities in the first set are not satisfied identically because of the choice of δ_i, while the last are not satisfied identically because of the choice of γ_{i+1}. Hence we may again take for x any member of R other than a finite number of values.

The last of the points δ_n thus constructed will not have equal coordinates, and in view of the irreducibility of $\mathbf{A}$ all of its coordinates will be different from 0, i.e., δ_n is a generic point. The lemma is proved.

REMARK. If R is infinite the requirement of the irreducibility of $\mathbf{A}$ is not essential. It is used only to check that no coordinate of the point δ_n is equal to zero, which may be avoided by adding, when necessary, a properly chosen scalar multiple of the unit. If R is finite the requirement of irreducibility is essential, for in this case it is easy to construct an example of a reducible algebra that does not have any generic points.

2. Content of the theory thus constructed. We have the following theorem:

THEOREM. *Any irreducible R-algebra is a separable finite algebraic extension of the field R. Conversely, any separable finite algebraic extension of degree n of the field R may be represented in the form of an irreducible R-algebra in the field K_n for a suitably chosen field K.*

PROOF. Let α be a generic point of an irreducible R-algebra $\mathbf{A}$ of the space K_n. Then the points $1, \alpha, \cdots, \alpha^{n-1}$ are linearly independent with respect to K, i.e., they form a coordinate system for the space K_n. In fact, if they were linearly dependent, that is if we had $c_1\alpha^{n-1} + \cdots + c_n = 0$ with coefficients from K, then the polynomial $\phi(x) = c_1x^{n-1} + \cdots + c_n$ would have n distinct roots, namely the coordinates of the point α, which is impossible. The basis $1, \alpha, \cdots, \alpha^{n-1}$ belongs to $\mathbf{A}$, and hence all the points of $\mathbf{A}$ are representable in terms of the basis with coordinates in R. In particular, there exist $a_1, a_2, \cdots, a_n \in R$ such that $\alpha^n = a_1\alpha^{n-1} + \cdots + a_n$. The roots of the polynomial $f(x) = x^n - a_1x^{n-1} - \cdots - a_n$ are the coordinates of the point α; they are all distinct. The polynomial $f(x)$ is irreducible in R. In fact, if we had $f(x) = \phi_1(x) \cdot \phi_2(x)$, where ϕ_1, ϕ_2 are nonconstant polynomials with coefficients from R, then the point $\phi_1(\alpha)$, being different from zero, would have zero coordinates corresponding to those coordinates of α which are roots of $\phi_1(x)$. Thus $\mathbf{A} = R(\alpha)$, where α is a root of a polynomial that is irreducible in R and which does not have multiple roots, i.e., it is a separable finite algebraic extension of R.

Conversely, let there be given a field $\tilde{\mathbf{A}}$, which is a finite separable algebraic extension of R, let $\tilde{\alpha}$ be a primitive element of it, and let $f(x)$ be an irreducible polynomial determining $\tilde{\alpha}$. We take for K a field in which $f(x)$ may be decomposed into linear factors, for example, an algebraically closed field containing R. In K let $f(x) = (x - \alpha^{(1)})(x - \alpha^{(2)}) \cdots (x - \alpha^{(n)})$. In view of the separability of $\tilde{\mathbf{A}}$, all the $\alpha^{(1)}, \alpha^{(2)}, \cdots, \alpha^{(n)}$ will be distinct. We associate with the element $\tilde{\alpha}$ the point α of the space K_n with coordinates $\alpha^{(1)}, \alpha^{(2)}, \cdots, \alpha^{(n)}$. Since α is a generic point in K_n, the points $1, \alpha, \cdots, \alpha^{n-1}$ form a

coordinate system in K_n. Clearly, an R-module based on this coordinate system is an R-algebra $\mathbf{A}$, isomorphic with the field $\tilde{\mathbf{A}}$. It will be irreducible, since, in view of the isomorphism with the field $\tilde{\mathbf{A}}$, it will contain no zero-divisors. The theorem is proved.

This page intentionally left blank

BIBLIOGRAPHY

[1] F. Arndt, *Versuch einer Theorie der homogenen Funktionen des dritten Grades mit zwei Variabeln*, Arch. Math. Phys. 17 (1851), 1–85.

[2] ———, *Untersuchungen über die Anzahl der kubischen Klassen, welche zu einer determinirenden quadratischen Klasse gehören*, Arch. Math. Phys. 19 (1852), 408–418.

[3] ———, *Tabellarische Berechnung der reducirten binären kubischen Formen und Klassification derselben für alle succesiven negativen Determinanten* $(-D)$ *von* $D = 3$ *bis* $D = 2000$, Arch. Math. Phys. 31 (1858), 335–448.

[4] ———, *Zur Theorie der binären kubischen Formen*, J. Reine Angew. Math. 53 (1857), 309–321.

[5] W. E. H. Berwick, *The classification of ideal numbers that depend on a cubic irrationality*, Proc. London Math. Soc. (2) 12 (1913), 393–429.

[6] B. A. Venkov, *Classification of cubic regions by quadratic regions*, Proc. 2nd all-Union Math. Congress (Leningrad, 1934). (Russian)

[7] A. Weil, *Sur un théorème de Mordell*, Bull. Sci. Math. (2) 54 (1930), 182–191.

[8] G. Voronoĭ, *Concerning algebraic integers derivable from a root of an equation of the third degree*, Master's Thesis, St. Petersburg, 1894. (Russian)

[9] ———, *On a generalization of the algorithm for continued fractions*, Doctoral Thesis, Warsaw, 1896. (Russian)

[10] R. Dedekind, *Ueber die Anzahl der Idealklassen in reinen kubischen Zahlkörpern*, J. Reine Angew. Math. 121 (1900), 40–123.

[11] B. N. Delone, *Solution of the indeterminate equation* $X^3q + Y^3 = 1$, Izv. Akad. Nauk SSSR (6) 16 (1922), 273–280. (Russian)

[12] ———, *On the number of representations of a number by a cubic binary form with negative determinant*, Izv. Akad. Nauk SSSR (6) 16 (1922), 253–272. (Russian)

[13] ———, *Vollständige Lösung der unbestimmten Gleichung* $X^3q + Y^3 = 1$ *in ganzen Zahlen*, Math. Z. 28 (1928), 1–9; *Über die Darstellung der Zahlen durch die binären kubischen Formen von negativer Diskriminante*, Math. Z. 31 (1930), 1–26. (German translations of [11] and [12])

[14] ———, *Ueber den Algorithmus der Erhöhung*, Ž. Leningrad. Fiz.-Mat. Obšč. 1 (1927), no. 2, 257–267.

[15] ———, *Solution of the problem of equivalence and tabulation of cubic binary forms with negative determinant*, Ž. Leningrad. Fiz.-Mat. Obšč. 1 (1926), no. 1, 40–55. (Russian)

[16] ———, *Interprétation géometrique de la généralisation de l'algorithme des fractions continues donné par Voronoĭ*, C. R. Acad. Sci. Paris (1923).

[17] ———, *On indeterminate equations*, Proc. all-Russian Math. Congress (Moscow, 1927), pp. 148–161. (Russian)

[18] ———, *A table of purely real domains of the fourth order* (in collaboration with I. Sominski and K. Billevič), Izv. Akad. Nauk SSSR Otd. Mat. i Estest. Nauk (1935), 1267–1310. (Russian. French summary)

[19] ———, *On the geometry of Galois theory*, Memorial volume dedicated to D. A. Grave, pp. 52–62, Moscow, 1940. (Russian)

[20] G. Eisenstein, *Théorèmes sur les formes cubiques et solution d'une équation du quatrième degré à quatre indeterminées*, J. Reine Angew. Math. 27 (1844), 75–79.

[21] ———, *Untersuchungen über die cubischen Formen mit zwei Variabeln*, J. Reine Angew. Math. 27 (1844), 89–104.

[22] ———, *Eigenschaften und Beziehungen der Ausdrücke, welche bei der Auflösung der allgemeinen cubischen Gleichungen erscheinen*, J. Reine Angew. Math. 27 (1844), 319–329.

[23] ———, *Allgemeine Untersuchungen über die Formen dritten Grades mit drei Variabeln, welche der Kreistheilung ihre Enstehung verdanken*, J. Reine Angew. Math. 28 (1844), 289–374.

[24] O. Žitomirskiĭ, *Sur la classification des formes cubiques*, Izv. Akad. Nauk SSSR Otd. Mat. i Estest. Nauk (1935), 1299–1312.

[25] E. I. Zolotarev, *On an indeterminate equation of third degree*, St. Petersburg, 1869. (Russian)

[26] ———, *Theory of integral complex numbers with applications to the integral calculus*, Doctoral Dissertation, St. Petersburg, 1874. (Russian)

[27] F. Klein, *Ausgewählte Kapiteln der Zahlentheorie*, Göttingen, 1896–97.

[28] F. Levi, *Kubische Zahlkörper und binäre kubische Formenklassen*, Ber. Sachs. Akad. Wiss. Leipzig Mat.-Nat. Kl. 66 (1914).

[29] A. Markov, *Sur les nombres entiers dépendants d'une racine cubique d'un nombre entier ordinaire,* Mem. l'Acad. Imp. Sci. St. Pétersbourg (7) 38 (1892), no. 9, 1–37.

[30] G. B. Mathews and W. E. H. Berwick, *On the reduction of arithmetical binary cubics which have a negative discriminant,* Proc. London Math. Soc. (2) 10 (1911–1912), 48–53.

[31] G. B. Mathews, *On the reduction and classification of binary cubics which have a negative discriminant,* Proc. London Math. Soc. (2) 10 (1911–1912), 128–138.

[32] H. Minkowski, *Diophantische Approximationen,* Leipzig, 1907.

[33] ——, *Généralisation de la théorie des fractions continues,* Ann. Sci. École Norm. Sup. (3) 13 (1896), 41–60.

[34] L. J. Mordell, *Note on the integer solutions of the equation* $Ey^2 = Ax^3 + Bx^2 + Cx + D$, Messenger of Math. 51 (1922), 169–171.

[35] ——, *On the integer solutions of the equation* $ey^2 = ax^3 + bx^2 + cx + d$, Proc. London Math. Soc. (2) 21 (1922–23), 415–419.

[36] ——, *On the rational solutions of the indeterminate equations of the third and fourth degrees,* Proc. Cambridge Philos. Soc. 21 (1922), 179–192.

[37] ——, *Indeterminate equations of the third degree,* Science Progress (1923).

[38] T. Nagell, *Vollständige Lösung einiger unbestimmten Gleichungen dritten Grades,* Vid.-Selsk. Skr. I. Mat. Nat. Kl. 1922, no. 14.

[39] ——, *Über die Einheiten in reinen kubischen Zahlkörpern,* Vid.-Selsk. Skr. I. Mat. Nat. Kl. 1923, no. 11.

[40] ——, *Solution complète de quelques équations cubiques à deux indéterminées,* J. Math. Pures Appl. (9) 4 (1925), 209–270.

[41] ——, *Über einige kubische Gleichungen mit zwei Unbestimmten,* Math. Z. 24 (1925), 422–447.

[42] ——, *Darstellung ganzer Zahlen durch binäre kubische Formen mit negativer Diskriminante,* Math. Z. 28 (1928), 10–29.

[43] ——, *Zur Theorie der kubischen Irrationalitäten,* Acta Math. 55 (1930), 33–65.

[44] ——, *L'analyse indéterminée de degré supérieur,* Mémor. Sci. Math. Vol. 39, Paris, 1929.

[45] L. W. Reid, *Tafel der Klassenanzahlen für kubische Zahlkörper,* Inaugural Dissertation, Göttingen, 1899.

[46] C. L. Siegel, *Über einige Anwendungen diophantischer Approximationen,* Sitz. Preuss. Akad. Wiss. Phys.-Math. Kl. 1929, 508.

[47] ———, *Die Gleichung* $ax^n - by^n = c$, Math. Ann. 114 (1937), 57–68.

[48] V. Tartakovskiĭ, *Auflösung der Gleichung* $x^4 - \rho y^4 = 1$, Izv. Akad. Nauk SSSR (6) 20 (1926), 301–324.

[49] A. Thue, *Bemerkungen über gewisse Näherungsbrüche algebraischer Zahlen,* Vid.-Selsk. Skr. I. Mat. Nat. Kl. 1908.

[50] ———, *Ueber rationale Annäherungswerte der reellen Wurzeln der ganzer Functionen dritten Grades* $x^3 - ax - b$, Vid.-Selsk. Skr. I. Mat. Nat. Kl. 1908.

[51] ———, *On an equation which is in general insoluble,* Vid.-Selsk. Skr. I. Mat. Nat. Kl. 1908. (Norwegian)

[52] ———, *Über Annäherungswerte algebraischer Zahlen,* J. Reine Angew. Math. 135 (1910), 284–305.

[53] ———, *Eine Lösung der Gleichung* $P(x) - Q(x) = (x - \rho)^n \cdot Pr(x)$ *in ganzen Functionen* P, Q *und* R *für jede beliebige ganze Zahl, wenn* ρ *eine Wurzel einer beliebigen ganzen Function bedeutet,* Vid.-Selsk. Skr. I. Mat. Nat. Kl. 1909.

[54] ———, *Ein Fundamentaltheorem zur Bestimmung von Annäherungswerten aller Wurzeln gewisser ganzen Funktionen,* J. Reine Angew. Math. 138 (1910), 96–108.

[55] J. V. Uspensky, *A method for finding units in cubic orders of a negative discriminant,* Trans. Amer. Math. Soc. 33 (1931), 1–22.

[56] D. K. Faddeev, *Tabulation of Galois domains and rings of the third order,* Trudy Fiz.-Mat. Inst. Steklov. V (1934)), 19–24. (Russian)

[57] ———, *On the equation* $x^4 - Ay^4 = \pm 1$, Trudy Fiz.-Mat. Inst. Steklov. V (1934), 41–52. (Russian)

[58] ———, *On the equation* $x^3 + y^3 = Az^3$, Trudy Fiz.-Mat. Inst. Steklov. V (1934), 25–40. (Russian)

[59] ———, *Classification of algebraic domains by their cubic resolvents,* Proc. 2nd all-Union Math. Congress (Leningrad, 1934), Vol. 2, pp. 32–35. (Russian)

[60] ———, *On a property of the group of classes of ideals for a domain of third degree,* Proc. 2nd all-Union Math. Congress (Leningrad, 1934), Vol. 2, pp. 42–44. (Russian)

[61] ———, *On a class of indeterminate equations of the third degree,* Proc. 2nd all-Union Math. Congress (Leningrad, 1934), Vol. 2, pp. 36–41. (Russian)

[62] ———, *The structure of algebraic domains whose Galois group is the quaternion group,* Leningrad. Gos, Univ. Učen. Zap. (1937), no. 17, 17–24. (Russian)

[63] P. Furtwängler, *Kubische Zahlkörper und Zahlengitter,* Dissertation, Göttingen.

[64] H. Hasse, *Arithmetische Theorie der kubischen Zahlkörper auf Klassenkörpertheoretischer Grundlage,* Math. Z. 31 (1930), 565–582.

[65] N. G. Čebotarev, *Foundations of Galois theory.* I, II, GITTL, Moscow, 1934, 1936. (Russian)

[66] ———, *The problem inverse to the Tschirnhausen problem,* Vestnik Čist. i Prikl. Znan. 1 (1922), no. 2, 1–8. (Russian)

[67] L. Charve, *De la réduction des formes quadratiques ternaires positives et de son application aux irrationnelles du troisième degré,* Ann. Sci. École Norm. Sup. (2) 9 (1880), 3–156. (Supplement)