ABSTRACT QUADRATIC FORMS

IRVING KAPLANSKY AND RICHARD J. SHAKER*

1. Introduction. We shall be studying the following structure, which we shall call a V-form ("vector-valued form"). Let G and W be additive abelian groups with every element of order 2 (i.e. vector spaces over the field GF(2) of two elements). Let there be given a symmetric bilinear map from $G \times G$ to W; we shall write it simply as a product ab. We define an equivalence relation on unordered n-ples of G. For $n = 2$: $(a, b) \sim (c, d)$ if $a + b = c + d$ and $ab = cd$. For $n > 2$ we define equivalence "piecewise": there is to be a chain from $(a_1, \ldots, a_n)$ to $(b_1, \ldots, b_n)$ where at each step only two elements are changed in accordance with the equivalence just defined for $n = 2$.

There are two fairly obvious invariants: $\sum a_i$ (we call this the *discriminant*), and $\sum a_i a_j$ summed over $i < j$ (we call this the *Witt invariant*).

Question. Do these invariants suffice? If not, what else is needed?

The reason for studying V-forms is that they are a generalization of ordinary quadratic form theory. Let K be any field of characteristic not 2, let K^* be its multiplicative group of non-zero elements, and write $G = K^*/(K^*)^2$. Although the natural notation for G is multiplicative, we prefer to write G additively. Let W be the subgroup of the Brauer group of K generated by all quaternion algebras; we also write W additively. Sending the pair $a, b \in K^*$ into the quaternion algebra they determine ($i^2 = a$, $j^2 = b$, $ij = -ji$) induces a map $G \times G \to W$ which is known to be symmetric and bilinear. Thus we have a V-form. Now we come to the crucial fact: for this V-form, the equivalence classes of n-ples described above are in one-to-one correspondence with equivalence classes of non-singular n-dimensional quadratic forms over K. To see this, think of the quadratic forms as diagonal matrices, and note that the diagonal elements can be thought of as lying in G, since multiplication of a diagonal element by a non-zero square is harmless; then make use of the characterization of two-dimensional quadratic forms by quaternion algebras (**5**, *Satz* 11), and Witt's theorem on piecewise equivalence (**5**, *Satz* **7**).

We devote this paper partly to exploring the theory of V-forms for its own sake, and partly to noting the implications for "concrete" quadratic forms. In § 2 we discuss completely the case where W is one-dimensional, which corresponds to local fields (the real numbers and the p-adic numbers). We are

Received May 13, 1968.

*We gratefully acknowledge partial support from the National Science Foundation, and the hospitality of the faculty of Queen Mary College of the University of London, where the ideas were initiated.

led to a local-global point of view which apes the familiar one of class field theory. In § 3 we describe the V-forms arising from the field of rational numbers and from the field of rational functions in one variable over a finite field. The interesting thing here is that the well-known analogies of class field theory can become actual identity from the V-form point of view. In § 4 the topic is algebraic function fields in one variable over a real closed field; the V-form then has additional structure which makes it a Boolean ring. In § 5 we turn to symmetric bilinear forms over fields of characteristic 2. This theory is not describable as a V-form. However, the failure of cancellation suggests that we switch to the stable point of view. When we do so, we find that this theory is an instance of V-forms; in particular, there is a Witt invariant, a fact which appears to be new. Because of the apparent lack of an appropriate analogue of quaternion algebras, the V-form has to be constructed in a devious way.

We learned, through H. Bass,† of a similar investigation due to Scharlau (**3**). He too discards the underlying field. He replaces it with a profinite group acting on a module (motivating example: the Galois group of the separable algebraic closure acting on it). Under cup-product, H^1 and H^2 give rise to a V-form. Where comparison is possible, the situation (crudely) is: Scharlau assumes more and gets more.

2. The local theorem. Let a V-form: $G \times G \to W$ be given, as defined in § 1. We use the symbol $\sim$ for the equivalence relation introduced on n-ples.

We define the *kernel* N of the V-form to be the subspace of G consisting of all a satisfying $aG = 0$. We say that the V-form is *non-singular* if $N = 0$. Let π be the natural mapping from G onto G/N. By declaring $\pi(x)\pi(y) = xy$ we define a non-singular V-form from $G/N \times G/N$ to W. In passing from the V-form on G to that on G/N we lose track of the discriminant, but no other information is lost. In detail, we have the following result.

LEMMA 1. *$(a_1, \ldots, a_n) \sim (b_1, \ldots, b_n)$ if and only if $\sum a_i = \sum b_i$ and $(\pi(a_1), \ldots, \pi(a_n)) \sim (\pi(b_1), \ldots, \pi(b_n))$.*

We leave the proof to the reader. Because of Lemma 1 we can usually restrict our deliberations to non-singular V-forms. We also leave to the reader the proofs of Lemma 2, a lemma which allows us to change our abstract quadratic forms by homothety, and Lemma 3, which asserts that equality of discriminants and Witt invariants enjoys a cancellation property.

LEMMA 2. *$(a_1, \ldots, a_n) \sim (b_1, \ldots, b_n)$ if and only if $(a_1 + c, \ldots, a_n + c) \sim (b_1 + c, \ldots, b_n + c)$.*

LEMMA 3. *Suppose that $(a, b_1, \ldots, b_n)$ and $(a, c_1, \ldots, c_n)$ have the same discriminant and the same Witt invariant. Then $(b_1, \ldots, b_n)$ and $(c_1, \ldots, c_n)$ likewise have the same discriminant and the same Witt invariant.*

†Written communication.

https://doi.org/10.4153/CJM-1969-134-5 Published online by Cambridge University Press

Consider a non-singular V-form with G one-dimensional. Then W might as well be one-dimensional. This is the V-form that we obtain from any ordered field where every positive element is a square, and there is a generalized theorem of inertia which simply states that equivalent n-ples are identical (up to a permutation, of course). The trivial proof is left to the reader.

If W is one-dimensional and G more than one-dimensional, we have a case that corresponds to p-adic fields and requires a modest investigation.

One further definition: we say that $(a_1, \ldots, a_n)$ represents b, written $(a_1, \ldots, a_n) \sim b$, if there exist elements $b_2, \ldots, b_n \in G$ such that $(a_1, \ldots, a_n) \sim (b, b_2, \ldots, b_n)$.

THEOREM. *Suppose that we are given a non-singular* V*-form* $G \times G \to W$ *with* W *one-dimensional and* G *more than one-dimensional. Then the discriminant and Witt invariant are a complete set of invariants for equivalence classes of* n*-ples. Furthermore, all* 4*-ples represent* 0.

Remark. There is a possibility of confusion between our statement that 4-ples represent 0 and the traditional one that 5-forms represent 0. Because of our switch from the multiplicative to the additive notation, the corresponding classical statement is that non-singular 4-forms represent 1.

Proof. We begin by discussing triples. Suppose then that (a, b, c) and (d, e, f) have the same discriminant and Witt invariant; we are to prove them equivalent. We suppose the contrary. By Lemma 2 we can assume that the common discriminant $a + b + c = d + e + f$ is 0. It follows from Lemma 3 that (a, b, c) and (d, e, f) cannot have in common an element of G which they both represent. In particular, (a, b) cannot represent d, i.e. (a, b) is not equivalent to $(d, a + b + d)$, namely $ab \neq d(a + b + d) = d(c + d)$. Hence

$$ab = cd + d^2 + 1. \tag{1}$$

(We are writing 0, 1 for the elements of W.) By symmetry, (1) also holds after any permutation of a, b, c, and after d is replaced by e or f. Replace d by e:

$$ab = ce + e^2 + 1. \tag{2}$$

Add (1) and (2) and recall that $d + e = f$:

$$cf + f^2 = 0. \tag{3}$$

In (3) replace c by b and add; the result is $af = 0$. Thus the product of any of a, b, c by any of d, e, f is 0. Then a further use of (3) yields $f^2 = 0$. Hence also $d^2 = e^2 = 0$ and then (1) yields $ab = 1$. By symmetry between the two triples, we also have $de = 1$.

Now $ab = (a + d)(b + d)$ since ad, bd, and d^2 are all 0. Hence $(a, b) \sim (a + d, b + d)$ and $(a, b, c) \sim (a + d, b + d, c)$. Everything shown above for (a, b, c) holds for this new triple $(a + d, b + d, c)$. In particular,

$(a + d)e = 0$, whereas it is 1 since $ae = 0$ and $de = 1$. This contradiction completes the discussion of triples.

We next show that any quadruple (a, b, c, d) represents 0, and again the proof is indirect. We must have $ab = 1$, for otherwise $(a, b) \sim (0, a + b)$. Thus $ab, ac, \ldots, cd$ are all 1. We have $(a + c)d = 0$; it follows that (a, b) cannot be equivalent to $(a + c, b + c)$, for if it were we would have

$$(a, b, d) \sim (a + c, b + c, d) \sim (b + c, 0, a + c + d).$$

Hence $ab \neq (a + c)(b + c)$, i.e. $c^2 = 1$. Of course also $a^2 = b^2 = d^2 = 1$.

Next assume that $a \neq b$. There exists an element $t \in G$ with $(a + b)t = 1$. By replacing t by $t + c$ if necessary we can arrange $t^2 = 1$. Then

$$(a, b) \sim (a + t, b + t).$$

What was shown above applies to $(a + t, b + t, c, d)$. In particular, $(a + t)^2 = 1$, whereas we know it is 0. Thus we are left with the case $a = b = c = d$. We now invoke for the first time the hypothesis that the dimension of G is at least 2; we use it to assert that G contains a non-zero element u with $u^2 = 0$. Then

$$(a, a, a, a) \sim (a + u, a + u, a, a),$$

and this transfers us out of the case where all four entries are equal.

We can now swiftly conclude the proof of the theorem. Suppose that $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ have the same discriminant and Witt invariant. We are to prove them equivalent. This is true by definition for $n = 2$, and we have handled the case $n = 3$. Suppose that $n \geqq 4$. Then

$$(a_1, \ldots, a_n) \sim (0, c_2, \ldots, c_n) \quad \text{and} \quad (b_1, \ldots, b_n) \sim (0, d_2, \ldots, d_n).$$

By Lemma 3, $(c_2, \ldots, c_n)$ and $(d_2, \ldots, d_n)$ have the same discriminant and Witt invariant, and by induction they are equivalent.

Remark. We leave it to the reader to investigate the question of the existence of n-ples with prescribed discriminant and Witt invariant. The answer is entirely analogous to what holds in local fields: no condition for $n \geqq 3$ and an obvious necessary condition for $n = 2$.

We return to the case of a general V-form: $G \times G \to W$, with W having any dimension. Let M be the dual space of W, i.e. the set of all functionals from W to GF(2). For any $\psi \in M$, the V-form $x, y \to \psi(xy)$ is one to which the above discussion applies. Note that even if we assume the given V-form to be non-singular, we may have re-introduced a kernel, and we write it $N\psi$. When $\dim G/N\psi \geqq 2$, the local V-form is determined by a discriminant and Witt invariant; of course these merely paste together to form the global discriminant and Witt invariant. When $\dim(G/N\psi) = 1$ we obtain a true additional invariant: the local index. This local-global point of view is entirely analogous to the one occurring in the Minkowski-Hasse theorem (see, e.g., **5**, pp. 41–42).

https://doi.org/10.4153/CJM-1969-134-5 Published online by Cambridge University Press

In (**4**) there are a number of results asserting that under suitable hypotheses, local equivalence implies global equivalence. We mention:

(1) W two-dimensional, the V-form alternate (i.e. $a^2 = 0$ for all $a \in G$),

(2) W two-dimensional, dimension of G/N sufficiently large,

(3) W three-dimensional, the V-form alternate, dimension of G/N sufficiently large,

(4) dimension of $G/N\psi$ sufficiently large for each $\psi \in M$ (the estimate depending on the dimension of W).

Results (1)–(3) are best possible with respect to the dimension of W.

We do not know of any fields to which these theorems are applicable.

3. Global fields. Let Q be the field of rational numbers. We wish to describe the V-form attached to Q. A natural basis for $G = Q^*/(Q^*)^2$ consists of elements x_p, p ranging over the primes, and x_{-1}. Furthermore, by the Minkowski-Hasse theorem, the same elements serve in a natural way as a basis of W. By the Minkowski-Hasse theorem, quadratic reciprocity, and the quadratic character of -1 and 2, the multiplication table is given by the following statements:

(1) $x_{-1}^2 = x_{-1} + x_2$,

(2) $x_2^2 = x_{-1}x_2 = 0$,

(3) $x_2x_p = 0$ for $p \equiv \pm 1 \pmod 8$,

(4) $x_2x_p = x_2 + x_p$ for $p \equiv \pm 3 \pmod 8$,

(5) $x_{-1}x_p = x_p^2 = 0$ for $p \equiv 1 \pmod 4$,

(6) $x_{-1}x_p = x_p^2 = x_2 + x_p$ for $p \equiv 3 \pmod 4$,

(7) If p and q are odd primes with at least one congruent to 1 (mod 4), then $x_px_q = 0$ or $x_px_q = x_p + x_q$; the product is 0 if and only if p is a quadratic residue of q,

(8) If p and q are both congruent to 3 (mod 4) and p is a quadratic residue of q, then $x_px_q = x_2 + x_p$.

Note that this complicated object is an algebra over GF(2). The elements x_p with $p \equiv 1 \pmod 4$ span a much simpler subalgebra. Let us abstract its properties.

We consider an algebra H over GF(2), with a basis $\{u_i\}$. We assume the following:

(a) $u_i^2 = 0$,

(b) $u_iu_j = 0$ or $u_iu_j = u_i + u_j$,

(c) Given a finite number of basis elements $u_1, \ldots, u_n$, there exists a basis element u different from them such that, for each i, $uu_i = 0$ or $uu_i = u + u_i$ as prescribed in advance.

We note that in the V-form for Q, the x_p's with $p \equiv 1 \pmod 4$ span an algebra of this type, the property (c) being a consequence of Dirichlet's theorem on primes in an arithmetic progression.

If we assume countable dimension, a typical stepwise argument shows that the properties (a), (b), and (c) determine H uniquely. Furthermore: regardless

of countability, for the V-form determined by H the discriminant and Witt invariant suffice. Details are given in (**4**); in addition, there is a similar treatment of the full V-form of Q, and of three more subalgebras: those spanned by the odd numbers, the positive numbers, and the odd positive numbers.

Now let k be a finite field of characteristic not 2 and $K = k(x)$ the rational function field in one indeterminate over k. A natural basis for $K^*/(K^*)^2$ consists of the monic irreducible polynomials together with an element in k but not in k^2, and W can be given the same basis. Suppose that -1 is a square in k. Then *the* V*-form of* K *is exactly the above algebra* H. More exactly, this is true provided the polynomial ring $k[x]$ enjoys a strengthened form of Dirichlet's theorem asserting that non-trivial arithmetic progressions contain an infinite number of irreducible polynomials with degrees of a prescribed parity. (We have not found this statement in the literature and have not tried to prove it.)

If -1 is not a square in k, then the V-form for K closely resembles the subalgebra spanned by the positive elements in the rational number V-form. Again the details appear in (**4**).

4. Real function fields. Let k be a real closed field. Let K be an algebraic function field in one variable over k. In this section we shall exhibit a natural ring structure on the V-form of K (or more precisely, on its V-form reduced modulo the kernel) that makes it a Boolean ring.

We assume that K is formally real. Otherwise, quadratic form theory over K is trivial. Let $L = K(i)$, $i^2 = -1$. It is known that L admits no division algebras; its Brauer group is 0. In particular, every quaternion algebra over K is split by L.

Let us, as usual, write $G = K^*/(K^*)^2$. In this section we shall write Q for the V-form: $G \times G \to W$. We write N for its kernel, and $G_0 = G/N$. The fact that every quaternion algebra is split by L has the following consequence: N is the image in G of the set of sums of squares in K^* (and it is furthermore true that every sum of squares in K is a sum of two squares). Take any quaternion algebra over K. Since it is split by L, it can be written as $Q(-1, a)$. Here a may be meaningfully taken in G_0, and it is a unique element of G_0. In this way we have defined a map $W \to G_0$. This may be compounded with $G_0 \times G_0 \to W$ to yield a ring structure on G_0. It is immediate that every element is idempotent. There is a unit element, given by the image of -1 in G_0. Less obvious is the associative law, which we now prove. For a_0, b_0, c_0 in G_0 we have to prove that $(a_0b_0)c_0 = a_0(b_0c_0)$. We take representatives a, b, c and interpret the multiplication. Thus:

(4) $Q(a, b) \sim Q(-1, d)$,
(5) $Q(d, c) \sim Q(-1, e)$,
(6) $Q(b, c) \sim Q(-1, f)$,
(7) $Q(a, f) \sim Q(-1, g)$.

We have to prove that $e_0 = g_0$, i.e. eg is a sum of squares. If this is not true,

then there is an ordering of K making eg negative, say $e < 0$, $g > 0$. Then by (5), d, $c < 0$, by (4), a, $b < 0$, by (6), $f < 0$, and by (7), $g < 0$, a contradiction.

We have thus shown that the (reduced) V-form of K carries a natural Boolean ring structure. Let us briefly consider more generally the V-form arising from any associative (and commutative) ring. What we notice at once is that the discriminant and Witt invariant are the first two elementary symmetric functions and we promptly introduce all of them, noting that they are all invariants.

When G is a Boolean ring, they constitute a complete set of invariants.

THEOREM. *Let G be a Boolean ring, and consider the* V*-form it defines. Let $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ be n-ples having the same elementary symmetric functions. Then $(a_1, \ldots, a_n) \sim (b_1, \ldots, b_n)$.*

Proof. For any a and b in G we have $ab = ab(a + b + ab)$. Hence $(a, b) \sim (ab, a + b + ab)$, i.e. (a, b) represents the product ab. If we iterate this remark we see that (a, b, c) represents abc, etc. Thus $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ both represent $a_1 \ldots a_n = b_1 \ldots b_n$. The complementary $(n - 1)$-ples again have equal elementary symmetric functions (by a simple computation valid in any commutative ring). Induction on n completes the proof.

Remarks. (1) It is furthermore true that we can omit the elementary symmetric functions of odd degree, starting at degree 3.

(2) Our description of quadratic forms over real function fields can be seen (after some technical manœuvres) to be equivalent to that of Witt (**5**, pp. 143–144) and also to that of Knight (**2**).

5. Symmetric bilinear forms over fields of characteristic 2. Over a field F of characteristic 2 there exists a well-developed theory of quadratic forms. That is not our topic in this final section. Rather, we shall discuss the entirely distinct subject of symmetric bilinear forms. More exactly, we shall discuss the non-alternate ones; alternate forms have an entirely trivial theory and we say no more about them. Any non-alternate symmetric bilinear form can be diagonalized (**1**, Theorem 6), and we are instantly ready to discuss them in the same way we did quadratic forms over fields of characteristic not 2. Immediately, we have a first question: can their theory be given by a V-form? The answer is negative. The argument is brief. We use the notation $Q(a, b)$ for the hypothesized V-form, ordinary parentheses for diagonalized symmetric bilinear forms, $\sim$ for equivalence of the latter. Let a, b, $a + b$ be non-zero elements of K. Since $(a, a + b) \sim (b, ab(a + b))$ and

$$(a, ab(a + b)) \sim (b, a + b),$$

we have: $Q(a, a + b) = Q(b, ab(a + b))$, $Q(a, ab(a + b)) = Q(b, a + b)$, thus by the bilinearity $Q(a, ab) = Q(b, ab)$, hence $Q(a, a) = Q(b, b)$. This

would imply, however, that $(a, a) \sim (b, b)$, which in fact is true only if ab is a square.

All this suggests, however, that we consider a strengthening of the relation of equivalence so as to allow (a, a) and (b, b) to be equivalent. We define a new equivalence relation (denoted by $\sim_N$) so that $(a_1, \ldots, a_n) \sim_N (b_1, \ldots, b_n)$ if and only if there exists a finite chain beginning with $(a_1, \ldots, a_n)$ and ending with $(b_1, \ldots, b_n)$ such that adjacent forms are either equivalent or of the form $(c, c, c_3, \ldots, c_n)$, $(d, d, c_3, \ldots, c_n)$. We introduce the concept this way because it best fits our needs. However, $\sim_N$ is in fact the same as "stable equivalence"; that is, $(a_1, \ldots, a_n) \sim_N (b_1, \ldots, b_n)$ if and only if there exist $c_1, \ldots, c_r$ with

$$(c_1, \ldots, c_r, a_1, \ldots, a_n) \sim (c_1, \ldots, c_r, b_1, \ldots, b_n).$$

(We could also express this in the language of the appropriate Grothendieck group.) We omit the rather long proof, which is given in (**4**). However, we need and shall prove the lowest-dimensional case (Lemma 7).

Let us note the following obvious fact: the vector space $\langle a_1, \ldots, a_n \rangle$ over F^2 spanned by $a_1, \ldots, a_n$ is an invariant of $(a_1, \ldots, a_n)$ under $\sim$. One might regard this as a second invariant to be considered right after the discriminant. We mention, without proof, that if the dimension of $\langle a_1, \ldots, a_n \rangle$ does not exceed 3, then $\langle a_1, \ldots, a_n \rangle$ and the discriminant are a complete set of invariants for $\sim$. This does not work starting at dimension 4, there being in fact a counter-example in 4×4 matrices.

Plainly, $\langle a_1, \ldots, a_n \rangle$ is not an invariant under new equivalence. (However, equality of this space plus new equivalence implies equivalence.) But in the extreme case, dim $\langle a_1, \ldots, a_n \rangle = n$, it is an invariant, for the extra changes allowed in new equivalence never get a chance to come into play. We state this formally as a lemma.

LEMMA 4. *Suppose that* $a_1, \ldots, a_n$ *are linearly independent* mod F^2, *and*

$$(a_1, \ldots, a_n) \sim_N (b_1, \ldots, b_n).$$

Then $(a_1, \ldots, a_n) \sim (b_1, \ldots, b_n)$.

In the case of 3-forms of discriminant 1 we note what happens in the dependent case.

LEMMA 5. *If* a, b, *and* ab *are linearly dependent* mod F^2, *then*
(1) *Either* a *is a square or* $(b, ab) \sim (1, a)$,
(2) $(a, b, ab) \sim_N (1, 1, 1)$.

Proof. We have, say, $x^2a + y^2b + z^2ab = 0$. If $y^2b + z^2ab \neq 0$, then (b, ab) represents a, i.e. $(b, ab) \sim (1, a)$. If $y^2b + z^2ab = 0$, we note that y and z must be non-zero and conclude that a is a square. This proves part (1). Part (2) follows at once from part (1).

Before proving Lemma 7 we treat an extreme case of cancellation which is easy since the relevant vector is unique.

LEMMA 6. *Suppose that* $a, a_2, \ldots, a_n$ *are linearly independent* mod F^2 *and* $(a, a_2, \ldots, a_n) \sim (a, b_2, \ldots, b_n)$. *Then* $(a_2, \ldots, a_n) \sim (b_2, \ldots, b_n)$.

Proof. Let us think of the underlying vector space V equipped with a basis $\epsilon_1, \ldots, \epsilon_n$ and an inner product ϕ satisfying $\phi(\epsilon_1, \epsilon_1) = a$, $\phi(\epsilon_i, \epsilon_i) = a_i$ for $i = 2, \ldots, n$, $\phi(\epsilon_i, \epsilon_j) = 0$ for $i \neq j$. The linear independence mod F^2 of $a, a_2, \ldots, a_n$ tells us that ϵ_1 is the only vector in V satisfying $\phi(\epsilon_1, \epsilon_1) = a$. Thus if $\eta_1, \ldots, \eta_n$ is the second basis, we must have $\eta_1 = \epsilon_1$. The lemma now follows from the fact that $\epsilon_2, \ldots, \epsilon_n$ and $\eta_2, \ldots, \eta_n$ both span the orthogonal complement of ϵ_1.

LEMMA 7. *If* $(a, b, ab) \sim_N (c, abc, ab)$, *then* $(a, b) \sim_N (c, abc)$.

Proof. If a, b, and ab are linearly independent mod F^2, then the result follows from Lemmas 4 and 6. If they are dependent, we have $(a, b) \sim_N (1, ab)$ by Lemma 5, and similarly, $(c, abc) \sim_N (1, ab)$.

By X (or X_F if necessary) we mean the set of $\sim_N$ classes of 3-forms of discriminant 1. (*Motivation.* In characteristic not 2, the isomorphism classes of quaternion algebras are in one-to-one correspondence with $\sim$ classes of 3-forms of discriminant -1 via the mapping $Q(a, b) \leftrightarrow (a, b, -ab)$.) If the elements a, b, and ab are linearly independent over F^2, the $\sim$ and $\sim_N$ classes of (a, b, ab) coincide (Lemma 4); in the dependent case, $(a, b, ab) \sim_N (1, 1, 1)$ (Lemma 6). By abuse of notation we will usually write (a, b, ab) for either the form or its $\sim_N$ class.

We introduce a partially defined addition $\oplus$ on X by decreeing that $(a, b, ab) \oplus (a, c, ac) = (a, bc, abc)$. (*Motivation.* In the characteristic not 2 case, the "sum" $(a, bc, -abc)$ of $(a, b, -ab)$ and $(a, c, -ac)$ is defined if and only if the tensor product of the corresponding quaternion algebras splits at least partially, and the two sums then correspond. This follows from an unpublished theorem of Albert which asserts that the tensor product of two quaternion algebras fails to be a division algebra if and only if the two quaternion algebras have a common quadratic subfield.)

The first urgent task is to show that $\oplus$ is well-defined. We state this as an explicit lemma.

LEMMA 8. *Suppose that*

$$(a, b, ab) \sim_N (d, e, de) \tag{4}$$

and

$$(a, c, ac) \sim_N (d, f, df). \tag{5}$$

Then

$$(a, bc, abc) \sim_N (d, ef, def). \tag{6}$$

Proof. We break the proof into three cases.

Case I. The forms in (4) $\sim_N (1, 1, 1)$. We shall prove that

$$(a, c, ac) \sim_N (a, bc, abc). \tag{7}$$

Since, similarly, the forms on the right of (5) and (6) are $\sim_N$, this will suffice. We apply Lemma 5 to (a, b, ab). If a is a square, then both forms in (7) are $\sim_N (1, 1, 1)$. Otherwise, we have $(b, ab) \sim (1, a)$, which can be multiplied by c to yield $(bc, abc) \sim (c, ac)$, and (7) follows.

We may henceforth assume that none of the forms in (4) and (5) are $\sim_N (1, 1, 1)$. This means (Lemma 5) that a, b, and ab are linearly independent mod F^2, and similarly for the three other forms in (4) and (5). Also, by Lemma 4, $\sim_N$ may be replaced by $\sim$ in (4) and (5).

Case II. $a = d$. By Lemma 6 we deduce that

$$(b, ab) \sim (e, de) \tag{7}$$

from (4) and

$$(c, ac) \sim (f, df) \tag{8}$$

from (5). We multiply (7) through by c and (8) by e. The result is

$$(bc, abc) \sim (ef, aef),$$

which verifies (6).

Case III. No assumption.

Since (a, b, ab) represents d, we have $x^2a + y^2b + z^2ab = d$. Set $b^* = y^2b + z^2ab$. We can suppose that $b^* \neq 0$, for if $b^* = 0$, then d is a square times a, covered by Case II. Note, further, that (a, b^*) represents d. Similarly, we have $u^2a + b^2c + w^2ac = d$, and we set $c^* = v^2c + w^2ac$. This gives us two chains of equivalences:

$$(a, b, ab) \sim (a, b^*, ab^*) \sim (d, ab^*d, ab^*) \sim (d, e, de), \tag{9}$$

$$(a, c, ac) \sim (a, c^*, ac^*) \sim (d, ac^*d, ac^*) \sim (d, f, df). \tag{10}$$

We can apply Case II to the initial pairs of forms in (9) and (10), obtaining the following relation:

$$(a, bc, abc) \sim_N (a, b^*c^*, ab^*c^*). \tag{11}$$

Similarly, from the final pairs in (9) and (10) we obtain

$$(d, ef, def) \sim_N (d, b^*c^*, db^*c^*). \tag{12}$$

Thus to prove (6) it will suffice to prove the new equivalence of the right-hand forms in (11) and (12). This we do as follows. We have $(a, b^*) \sim (d, dab^*)$, hence $(ad, b^*d) \sim (1, ab^*)$, therefore

$$(ad, b^*d, ab^*) \sim_N (1, 1, 1). \tag{13}$$

https://doi.org/10.4153/CJM-1969-134-5 Published online by Cambridge University Press

Similarly,

$$(ad, c^*d, ac^*) \sim_N (1, 1, 1). \tag{14}$$

To (13) and (14) we may apply the information obtained in Case I. The result is

$$(ad, b^*c^*, adb^*c^*) \sim_N (1, 1, 1). \tag{15}$$

We also have

$$(1, b^*c^*, b^*c^*) \sim_N (1, 1, 1). \tag{16}$$

Apply Lemma 7 to (15) and (16), and we see that

$$(ad, adb^*c^*) \sim_N (1, b^*c^*). \tag{17}$$

If we multiply (17) by a we obtain: $(d, db^*c^*) \sim_N (a, ab^*c^*)$, which is just what is needed to verify new equivalence of the right-hand forms in (11) and (12). We have completed the proof of Lemma 8.

The new equivalence class (1, 1, 1) serves as an identity element for $\oplus$, therefore we denote it by 0. In certain cases, X is a group. For example, if $[F: F^2] = 1$ or 2, then $X = 0$. We now show that X is a group if $[F: F^2] = 4$. If $\alpha = (a, b, ab)$, we write V_α for $\langle a, b, ab\rangle$. If $\alpha, \beta \in X$, then $\alpha \oplus \beta$ is defined, for either one of α, β is 0, otherwise both V_α and V_β are three-dimensional, so that $V_\alpha \cap V_\beta \neq 0$. Furthermore, the operation $\oplus$ is associative, for if $\alpha, \beta, \gamma \in X$, either one of the three is 0 (making associativity trivial), otherwise V_α, V_β, V_γ are all three-dimensional and again the intersection $V_\alpha \cap V_\beta \cap V_\gamma \neq 0$. We may write $\alpha = (a, b, ab), \beta = (a, c. ac), \gamma = (a, d, ad)$ and we find: $\alpha \oplus (\beta \oplus \gamma) = (\alpha \oplus \beta) \oplus \gamma = (a, bcd, abcd)$.

It is possible by a frontal assault to prove that the associative law for $\oplus$ holds when the relevant sums exist, but any further direct verification of the properties of $\oplus$ seems to be difficult and inconclusive. We propose instead to embed X into a group. More precisely, we establish the existence of a vector space W over GF(2) and a one-to-one mapping I from X to W preserving $\oplus$ (that is, whenever $\alpha \oplus \beta$ is defined, $I(\alpha \oplus \beta) = I(\alpha) + I(\beta)$). Let us, for brevity, call a field *favourable* if a W and an I exist for it with this property. We have seen in the preceding paragraph that F is favourable if $[F: F^2] \leqq 4$. We will, in due course, prove that all fields are favourable, but we do this first under the assumption that $[F: F^2]$ is finite. With $[F: F^2] = 2^r$, we make an induction on r and we may assume that $r \geqq 3$.

Fix an algebraic closure A of F (the purely inseparable closure $F^{2^{-\infty}}$ would actually do). Let S be the set of favourable subfields of A which contain F. We set W equal to the direct sum of W_K, K ranging over S, and we let I be the mapping from X to W such that if Π_K is the natural projection from W onto W_K, $\Pi_K I(\alpha) = I_K(\alpha)$ for all $\alpha \in X$. It is trivial that I preserves $\oplus$. With this background, we isolate the next step of the discussion as a lemma.

https://doi.org/10.4153/CJM-1969-134-5 Published online by Cambridge University Press

LEMMA 9. *If* $\alpha = (a, b, ab)$, $\beta = (c, d, cd)$, *and* $I(\alpha) = I(\beta)$, *then* $(a, b, ab) \sim_N c$, *i.e. there exists* e *with* $(a, b, ab) \sim_N (c, e, ce)$.

Proof. If $\alpha = 0$, there is nothing to prove, for a 3-form new-equivalent to 0 new-represents anything. Thus we assume that $\alpha \neq 0$.

Let $x_1, \ldots, x_r$ be a 2-basis for F, that is, the elements $x_1^{e_1} \ldots x_r^{e_r}$, $e_i = 0$ or 1, form a vector space basis of F over F^2. Let $F_1 = F(x_1^{2^{-\infty}})$ be the field obtained by adjoining to F all 2^nth roots of x_1. We leave it to the reader to show that the set B consisting of 1 and all x_1^q, where $q = m/2^n$ with m odd and less than 2^n, is a basis for F_1 over F; also, that $x_2, \ldots, x_r$ form a 2-basis for the field F_1. Note, in particular, that if a square in F_1 is written as a linear sum with respect to B, then the coefficient of 1 is of the form $y^2 + z^2x_1$, where $y, z \in F$. By induction on r, $F_1 \in S$, i.e. F_1 is favourable.

First suppose that $\alpha = 0$ when viewed as a form over F_1. Then there exist $u, v, w \in F_1$, not all zero, so that

$$u^2a + v^2b + w^2ab = 0. \tag{18}$$

We can assume, without loss of generality, that when u^2a is written as a linear sum with respect to B, the coefficient of 1 is non-zero, for if necessary we can multiply both sides of (18) by a power of x_1, which is always a square. Since the coefficient of 1 in the linear expression of $u^2a + v^2b + w^2ab$ is 0, we obtain the existence of $u_1, u_2, v_1, v_2, w_1, w_2 \in F$ with $(u_1^2 + u_2^2x_1)a + (v_1^2 + v_2^2x_1)b + (w_1^2 + w_2^2x_1)ab = 0$,

$$u_1^2a + v_1^2b + w_1^2ab = (u_2^2a + v_2^2b + w_2^2ab)x_1. \tag{19}$$

Since each side of (19) must be non-zero (otherwise $\alpha = 0$), we can solve for x_1, and find that in this case $x_1 \in \langle 1, a, b, ab\rangle$.

Next suppose that $\alpha \neq 0$ over F_1. Then a, b, and ab are linearly independent over F_1^2 and the equality of $I_{F_1}(\alpha)$ and $I_{F_1}(\beta)$ makes $(a, b, ab) \sim (c, d, cd)$ (and not just $\sim_N$). Thus there exist $u, v, w \in F_1$ such that

$$u^2a + v^2b + w^2ab = c. \tag{20}$$

Again we can write both sides of (20) as linear combinations of elements of B. Equating the coefficients of 1 we find that $u_1, u_2, v_1, v_2, w_1, w_2 \in F$ with $(u_1^2 + u_2^2x_1)a + (v_1^2 + v_2^2x_1)b + (w_1^2 + w_2^2x_1)ab = c$,

$$u_1^2a + v_1^2b + w_1^2ab + c = (u_2^2a + v_2^2b + w_2^2ab)x_1. \tag{21}$$

If both sides of (21) vanish, we have established that $(a, b, ab) \sim c$ over F, as required. Otherwise we can solve for x_1 and find:

$$x_1 \in \langle 1, a, b, ab, ac, bc, abc\rangle. \tag{22}$$

We put this together with the earlier case where we had found that $x_1 \in \langle 1, a, b, ab\rangle$, and see that if our conclusion fails then (22) must hold.

However, we could just as well have used $x_2, x_3, x_1x_2, x_1x_3, x_2x_3$ or $x_1x_2x_3$

throughout the argument in place of x_1, and have reached the same conclusion. Hence

$$\langle 1, x_1, x_2, x_3, x_1x_2, x_1x_3, x_2x_3, x_1x_2x_3\rangle \subseteq \langle 1, a, b, ab, ac, bc, abc\rangle,$$

a contradiction since on the left we have a space spanned by eight linearly independent elements, and on the right a spaced spanned by seven elements.

We have proved Lemma 9, and with it at hand, we can quickly complete the proof that F is favourable. Given $\alpha = (a, b, ab)$ and $\beta = (c, d, cd)$ with $I(\alpha) = I(\beta)$, we must prove that $\alpha \sim_N \beta$. By Lemma 9 we can replace (a, b, ab) by (c, e, ce). Thus $\alpha \oplus \beta$ exists and equals (c, de, cde); call it γ. The hypothesis $I(\alpha) = I(\beta)$ now becomes $I(\gamma) = 0$. Apply Lemma 9 again, this time to γ and 0. The conclusion is $\gamma \sim_N 1$, whence $\gamma = 0$, i.e. $\alpha \oplus \beta = 0$. From this we have to conclude that $\alpha = \beta$. If $\alpha = 0$, then $0 \oplus \beta = \beta$, thus $\beta = 0$. If $\alpha, \beta \neq 0$, we have $\alpha = (e, f, ef)$, $\beta = (e, g, eg)$, $(e, fg, efg) = 0$. Then there exist $x, y, z \in F$ ($x \neq 0$ since $\beta \neq 0$) with $x^2e + y^2fg + z^2efg = 0$. Multiplying by f we obtain $ef \in \langle g, eg\rangle$, thus $(f, ef) \sim (g, eg)$, hence $\alpha = \beta$.

We proceed to prove F favourable even when $[F:F^2]$ is infinite. For this we find it advisable to introduce the "universal" W. First, it is clear what one means by a homomorphism f from X to a vector space over GF(2); f is a mapping on X and satisfies $f(\alpha \oplus \beta) = f(\alpha) + f(\beta)$ whenever $\alpha \oplus \beta$ is defined. Now let Y be the vector space over GF(2) with basis X, and note that any homomorphism defined on X extends uniquely to Y. Let Z be the subspace of Y consisting of all elements annihilated by all homomorphisms on X. Let $W = Y/Z$. When necessary, we decorate X, W, Y, Z with the subscript F. There is a natural map J_F from X_F to W_F, and the statement that F is favourable is equivalent to saying that J_F is one-to-one.

If the field G is contained in the field H, there are natural induced maps: first from X_G to X_H, then Y_G to Y_H. The latter is readily seen to carry Z_G to Z_H. There is thus finally an induced map from W_G to W_H, but we shall not make use of it.

Now given our field F with $[F:F^2] = \infty$, we introduce still another subspace Z_0 of Y_F. We let K range over the subfields of F finitely generated over the prime field GF(2). Note that every such field satisfies $[K:K^2] < \infty$. We define Z_0 to be the (set-theoretic) union of the images of Z_K in the maps from Z_K to Z_F. It is easy to see that Z_0 is a subspace of Y_F (any two K's are contained in a third one, etc.). We do not attempt to identify Z_0 and Z_F. It suffices for us to prove that the induced map I from X_F to Y_F/Z_0 is a homomorphism and one-to-one.

I is a homomorphism. Let $\alpha, \beta \in X_F$ and suppose that $\gamma = \alpha \oplus \beta$ exists. If we write u_α, u_β, and u_γ for the corresponding basis elements of Y_F, our problem is to prove that $u_\alpha + u_\beta + u_\gamma \in Z_0$. Take representatives $\alpha = (a, b, ab)$, $\beta = (a, c, ac)$ so that $\gamma = (a, bc, abc)$. We drop to the subfield K generated by a, b, and c. If we write v_α, v_β, and v_γ for the analogous basis

elements of Y_K, we have that $v_\alpha + v_\beta + v_\gamma$ lies in Z_K and is mapped into $u_\alpha + u_\beta + u_\gamma$, which therefore lies in Z_0.

I is one-to-one. Given $\alpha, \beta \in X_F$ with $I(\alpha) = I(\beta)$, we must prove that $\alpha \sim_N \beta$. The statement $I(\alpha) = I(\beta)$ means that $u_\alpha + u_\beta \in Z_0$. This can be so only if $u_\alpha + u_\beta$ comes from some $v_\alpha + v_\beta$ in the Z_K of a finitely generated subfield K. But we have proved that K is favourable. Hence $\alpha \sim_N \beta$ in K and all the more so in F.

As this long discussion nears its end, we state our main result as a formal theorem.

THEOREM. *Let F be any field of characteristic 2. On the non-singular non-alternate finite-dimensional symmetric bilinear forms over F introduce new equivalence as above. Let X be the set of equivalence classes, under new equivalence, of three-dimensional forms of discriminant 1. Introduce on X, as above, the partially defined operation $\oplus$. Then there exists an embedding I of X into a vector space W over* GF(2). *Let F^* be the multiplicative group of non-zero elements in F and $G = F^*/(F^*)^2$. We define a* V*-form Q: $G \times G \to W$ as follows. Let $a_0, b_0 \in G$, pick representatives a, b in F^*, let α be the new equivalence class of (a, b, ab), and set $Q(a_0, b_0) = I(\alpha)$. Then: equivalence with respect to this* V*-form coincides with new equivalence.*

That Q is well-defined, symmetric, and bilinear is immediate. That Q correctly describes new equivalence of 2-forms follows from Lemma 7. There is one more thing to prove: that new equivalence satisfies Witt's theorem on piecewise equivalence. We shall do better and prove the (manifestly stronger) statement that piecewise equivalence works for ordinary equivalence.

Let $(a_1, \ldots, a_n) \sim (b_1, \ldots, b_n)$. We must establish the existence of a chain of equivalences, starting at $(a_1, \ldots, a_n)$ and ending at $(b_1, \ldots, b_n)$, such that at every step only two elements get changed, and the change on those two is by equivalence of 2-forms. The proof is by induction on n.

Write V for the underlying vector space, ϕ for the form, $v_1, \ldots, v_n$ for the basis that gives us $(a_1, \ldots, a_n)$. Thus $\phi(v_i, v_i) = a_i$ and $\phi(v_i, v_j) = 0$ for $i \neq j$. Let $w_1, \ldots, w_n$ analogously be the basis that leads to $(b_1, \ldots, b_n)$. Choose notation so that $w_1 = e_1 v_1 + \ldots + e_r v_r$ with $e_i \neq 0$. We make a second induction on r. If $r = 1$, then b_1 is a multiple of a_1 by a square and we may suppose that $b_1 = a_1$. The subspaces $\langle v_2, \ldots, v_n \rangle$ and $\langle w_2, \ldots, w_n \rangle$ are identical, since each is the orthogonal complement of v_1. Thus $(a_2, \ldots, a_n) \sim (b_2, \ldots, b_n)$ and our induction on n is applicable.

We assume that $r \geqq 2$. Let $s = e_1 v_1 + e_2 v_2$. If $\phi(s, s) \neq 0$, the subspace $\langle v_1, v_2 \rangle$ can be given a new basis s, t with t orthogonal to s. After we make this change we find that the length of w_1 has decreased to $r - 1$ and our induction on r applies. Therefore, we may assume that $\phi(s, s) = 0$, which has (as a consequence): a_1 and a_2 are the same (up to a square, as always). We can make the same argument on each $e_i v_i + e_j v_j$. Thus, further activity is only

needed in the following circumstances: $a_1, \ldots, a_r$ are all equal (we can take them to be 1) and $w_1 = v_1 + \ldots + v_r$. In particular, r is odd.

We next dispose of the extreme case: $r = n$. Then the a's are all 1's, the b's might as well be all 1's, and there is nothing to prove.

Suppose that $n > r + 1$. We work within the even-dimensional space $S = \langle v_1, \ldots, v_{r+1} \rangle$. The orthogonal complement of w_1 in S cannot be alternate since its dimension is odd. Thus we may complete w_1 to an orthogonal basis of S and (by induction on n) we may pass to this basis from $v_1, \ldots, v_{r+1}$ piecewise. We complete the transition to $w_1, \ldots, w_n$ by using the overlapping vector w_1 (the case $r = 1$).

There remains the case $n = r + 1$. The a's look as follows: $1, \ldots, 1, a$, where we are writing $a_{r+1} = a$. The orthogonal complement of w_1 consists of v_{r+1} and the alternate orthogonal complement of w_1 inside $\langle v_1, \ldots, v_r \rangle$. Using (**1**, Theorem 5) we complete w_1 to an orthogonal basis with diagonal elements $(1, a, \ldots, a)$. Thus this final case simplifies to the following problem: exhibit piecewise equivalence for the forms $(1, 1, 1, a)$ and $(1, a, a, a)$. We shall do this explicitly, using the abbreviations $c = 1 + a$, $d = a + a^2$ and we have:

$$(1, 1, 1, a) \sim (1, 1, c, d) \sim (1, a, d, d) \sim (a, a, c, d) \sim (a, a, 1, a).$$

We have completed the proof of the theorem and we conclude with several remarks.

Remarks. (1) Perhaps the most interesting corollary is the existence of a Witt invariant, which is indeed invariant under new equivalence, and all the more so under equivalence.

(2) When $[F:F^2] \leqq 4$ it is immediate that 4-forms represent 0. We can then argue (just as in § 2) that the discriminant and Witt invariant characterize new equivalence. Examples of such fields are polynomials or power series in two variables over a perfect field and their algebraic extensions.

(3) If $[F:F^2] > 4$, the discriminant and Witt invariant cannot suffice. For instance, if a, b, and c are part of a 2-basis, then $(1, a, b, ab)$ and (c, ac, bc, abc) are not $\sim_N$ but have the same discriminant and Witt invariant. We can amplify this example to the 8-forms

$$(1, a, b, ab, 1 + c, a + ac, b + bc, ab + abc)$$

and

$$(c, ac, bc, abc, 1 + c, a + ac, b + bc, ab + abc)$$

which have the same discriminant, same Witt invariant, and same subspace spanned over F^2, yet are not $\sim_N$.

(4) We conclude with a note on quadratic forms over a field F of characteristic 2. Let us stick to those whose attached alternate form is non-singular. Then there is a decomposition into 2-dimensional summands (which can be taken in the form $ax^2 + xy + by^2$). With these as building blocks, piecewise equivalence can be proved. However, the setup (and the Arf and Witt

invariants that go with it) does not seem to be expressible in terms of a V-form. Nevertheless, there is a certain V-form lurking in the background: send the pair a, b into $ax^2 + xy + by^2$; this V-form is defined on the additive group of F (which can be divided by the subgroup of elements $x^2 + x$) to the Grothendieck group of quadratic forms. The abstract quadratic form theory accompanying this V-form does not seem to be pertinent for quadratic forms over F.

References

1. A. A. Albert, *Symmetric and alternate matrices in an arbitrary field*, Trans. Amer. Math. Soc. *43* (1938), 386–436.
2. J. T. Knight, *Quadratic forms over $R(t)$*, Proc. Cambridge Philos. Soc. *62* (1966), 197–205.
3. Winfried Scharlau, *Quadratische Formen und Galois-Cohomologie*, Invent. Math. *4* (1967), 238–264.
4. R. J. Shaker, *Abstract quadratic forms*, Thesis, University of Chicago, Chicago, Illinois, 1968.
5. E. Witt, *Theorie der quadratischen Formen in beliebigen Körpern*, J. Reine Angew. Math. *176* (1937), 31–44.

University of Chicago,
Chicago, Illinois