MATHEMATICS OF COMPUTATION
VOLUME 45, NUMBER 172
OCTOBER 1985, PAGES 569–584

On Computing the Discriminant of an Algebraic Number Field

By Theresa P. Vaughan

Abstract. Let $f(x)$ be a monic irreducible polynomial in $\mathbf{Z}[x]$, and $\mathbf{r}$ a root of $f(x)$ in $\mathbf{C}$. Let K be the field $\mathbf{Q}(\mathbf{r})$ and $\mathscr{R}$ the ring of integers in K. Then for some $k \in \mathbf{Z}$, $\operatorname{disc}\mathbf{r} = k^2 \operatorname{disc}\mathscr{R}$. In this paper we give constructive methods for (a) deciding if a prime p divides k, and (b) if $p \mid k$, finding a polynomial $g(x) \in \mathbf{Z}[x]$ so that $g(x) \not\equiv 0 \pmod{p}$ but $g(\mathbf{r})/p \in \mathscr{R}$.

1. Introduction. Let $f(x)$ be a monic irreducible polynomial with integral coefficients, and $\mathbf{r}$ a root of $f(x)$ in $\mathbf{C}$. Let K be the field $\mathbf{Q}(\mathbf{r})$ and $\mathscr{R}$ the ring of algebraic integers in K. It is well-known that

$$N(f'(\mathbf{r})) = \operatorname{disc}\mathbf{r} = k^2 \operatorname{disc}\mathscr{R}$$

for some integer k. While $N(f'(\mathbf{r}))$ can be found by straightforward (if tedious) computation, the value of k is quite another story. According to [2, p. 77] for example, to determine k, one would have to test a finite number (which may be very large) of elements of K to see if they are integral.

In this paper, we reduce some of the difficulties involved in finding k to more manageable size; our methods do not require a search process. Consider the following problems:

(I) Given a prime p with $p^2 \mid \operatorname{disc}\mathbf{r}$, how can one tell whether or not $p \mid k$?

(II) Suppose $p \mid k$. Then, it is known there exists an element $\beta \in K$, $\beta = g(\mathbf{r})$ for $g(x) \in \mathbf{Z}[x]$ and $g(x) \not\equiv 0 \pmod{p}$ such that $\beta/p \in \mathscr{R}$. Construct such an element β.

A reasonable solution to these problems is furnished by Theorems 5.4, 5.7 and 5.9; it may be summarized as follows:

Suppose that $f(x)$ has degree n, and that $p^2 \mid \operatorname{disc}\mathbf{r}$. Factor $f(x) \bmod p$:

$$f(x) = \prod_{i=1}^{r} f_i(x)^{e_i} \qquad (f_i(x) \text{ irreducible}).$$

If all the $e_i = 1$, then $p \nmid k$. If any $e_i > 1$, then let C be the companion matrix of $f(x)$ and compute $f_i(C) \pmod{p^2}$. This matrix represents a homogeneous system of linear equations mod p^2; if this system has a nontrivial solution mod p^2, then $p \mid k$ and Theorem 5.9 enables the construction of a β as in (II) above. If for each $e_i > 1$, the system of equations has no nontrivial solution, then $p \nmid k$. The actual labor involved, amounts to the computation and row-reduction (mod p^2) of no more than $r\,n \times n$ matrices.

Received July 20, 1984.
1980 *Mathematics Subject Classification*. Primary 12A50.

In Section 2, we give the notation we need from number theory. In Section 3, we give our matrix-theoretic notation, and define one of our basic tools: an Abelian group, associated with an $n \times n$ integral matrix A:

$$R(A) = \{\bar{v} \in \mathbf{Q}^n : 0 \leqslant v_i < 1, A\bar{v} \in \mathbf{Z}^n\}$$

whose operation is addition modulo 1 on the coordinates of vectors.

In Section 4, we consider the companion matrix C of $f(x)$ as a linear transformation of $\mathbf{Z}_p^n$ (p prime); we then find that the p-component of a group $R(g(C))$ for $g(x) \in \mathbf{Z}[x]$, has a particularly nice sort of basis, which then determines (among other things) the arrangement of powers of p on the diagonal of the Smith form for $g(C)$.

In Section 5, we use the machinery developed in Section 4 to answer (I) and (II) above; in Section 6 we give some results which are helpful in computations, and two examples.

Finally, in Section 7, we list some unanswered questions and conjectures.

2. Notation (number-theoretic). Let $f(x) \in \mathbf{Z}[x]$ be monic and irreducible of degree n:

$$f(x) = x^n - a_{n-1}x^{n-1} - a_{n-2}x^{n-2} - \cdots - a_1x - a_0 \qquad (a_i \in \mathbf{Z})$$

and let $\mathbf{r}$ be a root of $f(x)$ in $\mathbf{C}$. Let K be the field $\mathbf{Q}(\mathbf{r})$ and $\mathscr{R}$ the ring of algebraic integers in K. Let $\sigma_1, \sigma_2, \ldots, \sigma_n$ be the embeddings of K in $\mathbf{C}$. If $\alpha \in K$, then the norm and trace of α are defined by $\mathscr{N}(\alpha) = \prod_{i=1}^n \sigma_i(\alpha)$; $\mathrm{Tr}(\alpha) = \sum_{i=1}^n \sigma_i(\alpha)$.

If $\{\alpha_1, \alpha_2, \ldots, \alpha_n\} \subset \mathscr{R}$ and if $\mathscr{R} = \{\sum_{i=1}^n n_i\alpha_i \mid n_i \in \mathbf{Z}\}$ then $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ is said to be an *integral basis* for $\mathscr{R}$. Then the discriminant of $\mathscr{R}$ is given by the square of the determinant of $(\sigma_i(\alpha_j))$;

$$\operatorname{disc} \mathscr{R} = \left|\sigma_i(\alpha_j)\right|^2 \in \mathbf{Z}.$$

Since $\mathbf{r}$ is a root of $f(x)$, the set $\{1, \mathbf{r}, \mathbf{r}^2, \ldots, \mathbf{r}^{n-1}\}$ is a basis for K over $\mathbf{Q}$, but not necessarily an integral basis for $\mathscr{R}$, that is, it is possible that $\mathbf{Z}[\mathbf{r}] \neq \mathscr{R}$ (of course $\mathbf{Z}[\mathbf{r}] \subseteq \mathscr{R}$ always).

Define the discriminant of $\mathbf{r}$ as

$$\operatorname{disc} \mathbf{r} = \left|\sigma_i(\mathbf{r}^{j-1})\right|^2;$$

one has $\operatorname{disc} \mathbf{r} = k^2 \operatorname{disc} \mathscr{R}$ for some $k \in \mathbf{Z}$; also $\operatorname{disc} \mathbf{r} = \mathscr{N}(f'(\mathbf{r}))$.

It is well-known that if $p \nmid k$, then the factorization of the ideal (p) in $\mathscr{R}$, into prime ideals, may be determined as follows: Write

$$f(x) \equiv \prod_{i=1}^{r} f_i(x)^{e_i} \pmod{p} \tag{1}$$

and let P_i be the ideal $(f_i(\mathbf{r}), p)$. Then $(p) = \prod_{i=1}^r P_i^{e_i}$ is the prime factorization of (p).

On the other hand, if $p \mid k$, then there exists $\beta \in \mathscr{R}$ of the form

$$\beta = \sum_{i=0}^{n-1} n_i\mathbf{r}^i \qquad (n_i \in \mathbf{Z}),$$

where not all the $n_i \equiv 0 \pmod{p}$ and $(1/p)\beta \in \mathscr{R}$; the factorization (1) does *not* yield the prime factorization of (p).

As we shall see, however, the factorization (1) may still yield partial information. Our methods rely heavily on matrix representations, described in the next section.

3. Notation (matrix-theoretic). Let C be the companion matrix for $f(x)$:

$$C = \begin{bmatrix} 0 & 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & \cdots & 0 & a_2 \\ 0 & \cdot & \cdots & 1 & a_{n-1} \end{bmatrix}.$$

The minimum and characteristic polynomials of C are both equal to $f(x)$, and C represents the $\mathbf{Q}$-linear map on K "multiplication by $\mathbf{r}$" in the basis $\{1, \mathbf{r}, \mathbf{r}^2, \ldots, \mathbf{r}^{n-1}\}$ for K over $\mathbf{Q}$. That is, if $\alpha, \beta \in K$ and

$$\alpha = \sum_{i=0}^{n-1} c_i \mathbf{r}^i; \qquad \beta = \mathbf{r}\alpha = \sum_{i=0}^{n-1} b_i \mathbf{r}^i \qquad (c_i, b_i \in \mathbf{Q}),$$

then $C \cdot \operatorname{col}(c_0, \ldots, c_{n-1}) = \operatorname{col}(b_0, \ldots, b_{n-1})$.

One has $\mathbf{Q}(\mathbf{r}) \cong \mathbf{Q}[C]$ and $\mathbf{Z}[\mathbf{r}] \cong \mathbf{Z}[C]$ via the usual correspondence $g(\mathbf{r}) \approx g(C)$. The eigenvalues of $g(C)$ are $\{\sigma_i(g(\mathbf{r}))\}$, $|g(C)| = \mathscr{N}(g(\mathbf{r}))$ and $\operatorname{Tr}(g(C)) = \operatorname{Tr}(g(\mathbf{r}))$. Since $f(x)$ is irreducible over $\mathbf{Q}$, any matrix $X = g(C)$ $(g(x) \in \mathbf{Q}[x])$ is singular if and only if it is 0; also if $XC = CX$ then $X = g(C)$ and conversely.

We shall need the Smith form $S(X)$ of an integral matrix $X \in \mathbf{Z}_{n\times n}$ with $|X| \neq 0$. For each such X, there exist matrices P, Q in $\mathbf{Z}_{n\times n}$ with $|P| = \pm 1$ and $|Q| = \pm 1$ and positive integers d_i with $d_i \mid d_{i+1}$ for $i = 1, 2, \ldots, n-1$ such that

$$PXQ = \begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{bmatrix} = S(X).$$

The d_i are called the *invariant factors* of X. If $p^{r_i} \| d_i$ (that is, $p^{r_i} \mid d_i$ and $p^{r_i+1} \nmid d_i$) for a prime p, then we put

$$S_p(X) = \operatorname{diag}(p^{r_1}, p^{r_2}, \ldots, p^{r_n}),$$

and call this the p-Smith form of X.

Finally, we shall require the following group associated with $X \in \mathbf{Z}_{n\times n}$, $|X| \neq 0$ (see [1] for a detailed discussion).

$$R(X) = \{\bar{v} = \operatorname{col}(v_1, v_2, \ldots, v_n) \in \mathbf{Q}^n : 0 \leqslant v_i < 1 \text{ and } X\bar{v} \in \mathbf{Z}^n\}.$$

The operation is addition modulo 1 on the coordinates of vectors. It is proved in [1] that $R(X)$ is an Abelian group with invariant factors $d_1, \ldots, d_n$; that is, $R(X) \cong C(d_1) \oplus \cdots \oplus C(d_n)$. The order of $R(X)$ is $|X|$. The p-component of $R(X)$ (elements whose order is a power of p) for a prime p, is the set of all $\bar{v} \in R(X)$ such that

$$\bar{v} = \frac{1}{p^t}\bar{a} \quad \text{where } t > 0 \text{ and } \bar{a} \in \mathbf{Z}^n.$$

If $\bar{a} \not\equiv 0 \pmod{p}$, then p^t is the order of $\bar{v}$ in $R(X)$, and we say that $\bar{v}$ is a *p^t-point* for X.

4. Preliminary Results. This section is mostly concerned with the properties of the p-component $R_p(X)$ of the group $R(X)$, for $X = g(C) \in \mathbf{Z}[C]$ (where p is a prime dividing $|X|$). From now on, unless otherwise indicated, all polynomials are in $\mathbf{Z}[x]$, all matrices are in $\mathbf{Z}[C]$, p is a prime, and we shall freely employ the following abuse of notation: If $\bar{a} \in \mathbf{Z}^n$, we say also $\bar{a} \in \mathbf{Z}_p^n$, meaning the reduced form of $\bar{a}$ modulo p; if $\bar{a} \in \mathbf{Q}^n$ we say $\bar{a} \in R(X)$, meaning the reduced form of $\bar{a}$ modulo 1.

Any integral matrix $X \in \mathbf{Z}_{n\times n}$ may be regarded as a linear transformation of the vector space $\mathbf{Z}_p^n$, and we use the notation $\ker_p X$, $\operatorname{Im}_p X$, p-rank, p-nullity, etc., in this setting.

Let p be a fixed prime, and $g(x) \in \mathbf{Z}[x]$, $X = g(C)$. For convenience, we state the following well-known facts and observations as a theorem. (For details see, e.g., [3].)

4.1. THEOREM. (a) $|X| \equiv 0 \pmod p$ *if and only if* X *is singular on* $\mathbf{Z}_p^n$; (b) $|X| \equiv 0 \pmod p$ *if and only if* $\gcd(f(x), g(x)) \not\equiv 1 \pmod p$; (c) $\dim(\ker_p X) = p$*-nullity of* $X =$ *number of invariant factors of* X *divisible by* p; (d) $\dim(\operatorname{Im}_p X) = p$*-rank* $X = n - (p$*-nullity*$)$; (e) *Every* C*-invariant subspace* W *of* $\mathbf{Z}_p^n$ *has a cyclic vector* $\bar{v}$ *for* C, *that is,* W *has a basis of the form* $\{\bar{v}, C\bar{v}, C^2\bar{v}, \ldots, C^{k-1}\bar{v}\}$ *where* $k = \dim W$; (f) *If* W *is a* C*-invariant subspace of* $\mathbf{Z}_p^n$, *then* $W = \ker_p k(C)$, *where* $k(x) \in \mathbf{Z}[x]$ *and* $k(x)h(x) = f(x) \pmod p$ *for some* $h(x)$; $\dim W =$ degree $k(x)$; *the minimum polynomial for the restriction* $C \mid W$ *is* $k(x)$; *and* $W = \operatorname{Im}_p h(C)$; (g) *Let* $f(x) = \prod_{i=1}^r f_i(x)^{e_i} \pmod p$, *where* $f_i(x)$ *is irreducible over* $\mathbf{Z}_p$ *of degree* k_i; *and put* $W_i^t = \ker_p f_i(C)^t$. *If* $e_i > 1$, *then*

$$W_i^1 \subsetneq W_i^2 \subsetneq \cdots \subsetneq W_i^{e_i} = W_i^{e_i} = \cdots.$$

If $X = g(C)$, *where*

$$\gcd(f(x), g(x)) = \prod_{i=1}^r f_i(x)^{t_i},$$

then $\ker_p X$ *is the direct sum of subspaces*:

$$\ker_p X = W_1^{t_1} \oplus W_2^{t_2} \oplus \cdots \oplus W_r^{t_r}.$$

Proof. Parts (a)–(d) are obvious; parts (e)–(g) follow from the fact that the minimum and characteristic polynomials of C are equal. □

4.2. *Definition.* Let $X = g(C)$ and let $V \neq \{0\}$ be a subset of $\ker_p X$. The *V-component* of $R_p(X)$ is the set

$$R_V(X) = \left\{\bar{v} \in R_p(X) : \bar{v} = (1/p^t)\bar{a}, \bar{a} \in \mathbf{Z}^n, \bar{a} \not\equiv 0 \pmod p \text{ and } \bar{a} \in V\right\}.$$

Before the main theorem, we need a few lemmas.

4.3. LEMMA. *Let* $X = g(C)$, $g(x) \in \mathbf{Z}[x]$. *If there is a vector* $\bar{a} \not\equiv 0$, $\bar{a} \in \ker_p X$, *then for some* $t > 0$, $(1/p^t)\bar{a} \in R_p(X)$. *Conversely, if* $(1/p^t)\bar{a} \in R_p(X)$, *where* $\bar{a} \not\equiv 0$ *and* $t > 0$, *then* $\bar{a} \in \ker_p X$.

Proof. $X\bar{a} \equiv 0 \pmod p$ if and only if $X\bar{a} = p^t\bar{b}$ for some $t > 0$, $\bar{b} \in \mathbf{Z}^n$; hence if and only if $X(1/p^t)\bar{a} = \bar{b} \in \mathbf{Z}^n$, that is, $(1/p^t)\bar{a} \in R_p(X)$. □

4.4. LEMMA. *The following are equivalent*:

(a) $\ker_p X \cap W_i \neq \{0\}$; (b) $W_i \subseteq \ker_p X$;

(c) $(1/p)W_i \subseteq R_p(X)$; (d) $f_i(x) \mid g(x) \pmod p$.

Proof. Since $f_i(x)$ is irreducible mod p, $W_i = \ker_p f_i(C)$ has no proper nontrivial C-invariant subspaces by Theorem 4.1(f); hence (a) $\leftrightarrow$ (b). The implications (b) $\leftrightarrow$ (d) also follow from Theorem 4.1(f), and (b) $\leftrightarrow$ (c) from Lemma 4.3. □

4.5. LEMMA. *Let $A, B, X, Y, Z \in \mathbf{Z}[C]$ and suppose that V is a subset of $\ker_p Z$. If $R_V(Z) \subseteq R_p(X)$ and $R_V(Z) \subseteq R_p(Y)$, then $R_V(Z) \subseteq R_p(AX + BY)$.*

Proof. Let $\bar{v} \in R_V(Z)$. By assumption, both $X\bar{v}$ and $Y\bar{v}$ are in $\mathbf{Z}^n$ and hence so is $(AX + BY)\bar{v}$; then $\bar{v} \in R_p(AX + BY)$. □

For the next theorem, we use the notation of Theorem 4.1(g).

4.6. THEOREM. *Let $X \in \mathbf{Z}[C]$ be such that $\ker_p X = W_i^{t_i}$ for some i, $1 \leqslant i \leqslant r$. Put $W = W_i$, $t = t_i$, $k = k_i$.*

(a) *There exist positive integers $s_1 \leqslant s_2 \leqslant \cdots \leqslant s_t$ such that the p-Smith form of X has the form:*

$$S_p(X) = \operatorname{diag}\Big(1, 1, \ldots, 1, \underbrace{p^{s_1}, \ldots, p^{s_1}}_{k}, \ldots, \underbrace{p^{s_t} \cdots p^{s_t}}_{k}\Big).$$

(b) *For each $i = 1, 2, \ldots, t$,*

$$s_{t-i+1} = \max\big\{s\colon (1/p^s)\bar{a} \in R_p(X), \bar{a} \in W^i - W^{i-1}\big\},$$

and if $V = W^i - W^{i-1}$, then $R_V(X)$ contains k independent $p^{s_{t-i+1}}$-points.

(c) *$R_p(X)$ does not contain any elements of order higher than p^{s_t}.*

(d) *$R_W(X)$ contains k independent p^{s_t}-points.*

Proof. (a) There exist unimodular integral matrices P, Q (not necessarily in $\mathbf{Z}[C]$) so that $PXQ = S(X)$, the Smith form of X. Then X and $S(X)$ have the same p-nullity $tk = m$, so the p-Smith form is

$$S_p(X) = \operatorname{diag}(1, 1, \ldots, 1, p^{r_1}, p^{r_2}, \ldots, p^{r_m}),$$

where the r_i are positive integers, $0 < r_1 \leqslant r_2 \leqslant \cdots \leqslant r_m$.

As is shown in [1], a basis for $R(X)$ consists of the columns of the matrix $Q \cdot S(X)^{-1}$ (reduced modulo 1, of course); hence, a basis for $R_p(X)$ is given by the last m columns of $Q \cdot S_p(X)^{-1}$, namely

$$(1/p^{r_1})Q_{n-m+1}, \ldots, (1/p^{r_m})Q_n$$

(where Q_i is the ith column of Q). By Lemma 4.3, $\{Q_i;\ i = n - m + 1, \ldots, n\} \subseteq \ker_p X$ and since Q is unimodular, these vectors are independent in $\mathbf{Z}_p^n$. Thus, this set is a basis for $\ker_p X$.

Suppose now that $Q_n \in W^j - W^{j-1}$ $(1 \leqslant j \leqslant t;\ W^0 = \{0\})$. Then Q_n is a cyclic vector for W^j, that is, the set

$$\big\{C^i(Q_n)\colon i = 0, \ldots, jk - 1\big\}$$

is a basis for W^j, and, in particular, for $0 \leqslant i < jk$, $C^i(Q_n) \not\equiv 0 \pmod{p}$. Since $X \in \mathbf{Z}[C]$, and $(1/p^{r_m})Q_n \in R_p(X)$, we have

$$X\big((1/p^{r_m})C^i(Q_n)\big) = C^i X\big((1/p^{r_m})Q_n\big) \in \mathbf{Z}_n,$$

so that all these vectors are p^{r_m}-points for X.

Since the set $\{(1/p^{r_i})Q_{n-m+i};\ i = 1,2,\ldots,m\}$ is a basis for $R_p(X)$, we have the equality (in $\mathbf{Q}^n$):

$$(1/p^{r_m})C^i(Q_n) = \sum_{u=1}^{m} n_u(1/p^{r_u})Q_{n-m+u} + \overline{m} \qquad (n_u \in \mathbf{Z} \text{ and } \overline{m} \in \mathbf{Z}^n).$$

From this

$$C^i(Q_n) = \sum_{u=1}^{m} n_u(p^{r_m - r_u})Q_{n-m+u} + p^{r_m}\overline{m}.$$

We have jk vectors $C^i(Q_n)$, independent mod p, written as linear combinations (mod p) of those Q_{n-m+u} such that $r_m - r_u = 0$; hence, the number of these must be at least jk. Since the r_i are in increasing order, we have the last $jk \geqslant k$ of the $r_i = r_m$.

The same argument now shows that if any of $Q_{n-jk+1},\ldots,Q_n$ were in some $W^u - W^{u-1}$, then we would have the last uk of the $r_i = r_m$. Thus, suppose that all of $Q_{n-jk+1},\ldots,Q_n$ are in W^j. We have jk independent vectors in a space of dimension jk, so they are a basis for W^j. It follows from the independence of the set $\{Q_i\}$ that none of $Q_{n-m+1},\ldots,Q_{n-jk}$ are in W^j (we are assuming now that $j < t$; of course, if $j = t$, we are done).

Suppose that $Q_{n-jk} \in W^a - W^{a-1}$, $a > j$. Put $r = r_{m-jk}$. The set $\{C^u(Q_{n-jk})$: $u = 0,1,\ldots,(a-j)k-1\}$ is an independent set in $W^a - W^j$, and as before, we have a set of p^r-points:

$$(1/p^r)C^u(Q_{n-jk}) \in R_p(X); \qquad u = 0,\ldots,(a-j)k-1.$$

Then we can write

$$(1/p^r)C^u(Q_{n-jk}) = \sum_{s=1}^{m-jk} n_s(1/p^{r_s})Q_{n-m+s} + \sum_{s=m-jk+1}^{m} n_s(1/p^{r_m})Q_{n-m+s} + \overline{m}$$

($n_s \in \mathbf{Z}$ and $\overline{m} \in \mathbf{Z}^n$), and from this

$$(*) \quad C^u(Q_{n-jk}) = \sum_{s=1}^{m-jk} n_s(p^{r-r_s})Q_{n-m+s} + \sum_{s=m-jk+1}^{m} n_s(1/p^{r_m-r})Q_{n-m+s} + p^r\overline{m}.$$

The middle term on the right must be integral, since all the other terms are. But the Q_i are columns of a unimodular matrix, so this implies that all coefficients of the middle term are integral: p^{r_m-r} divides n_s, $s = m - jk + 1,\ldots,m$.

Next we have (in $\mathbf{Z}_p^n$) $f_i(C)^j(W^j) = \{0\}$ while $f_i(C)^j(W^a)$ is a space of dimension $(a-j)k$, namely W^{a-j}. Applying $f_i(C)^j$ to (*), we have

$$f_i(C)^j\big(C^u(Q_{n-jk})\big) \equiv \sum_{s=1}^{m-jk} n_s p^{r-r_s} Q_{n-m+s} \pmod{p},$$

since $Q_s \in W^j$ for $s > n - jk$. Now, as before, we must have at least $(a-j)k$ of the values $r - r_s = 0$, since the set

$$\big\{f_i(C)^j\big(C^u(Q_{n-jk})\big)\colon u = 0,1,\ldots,(a-j)k-1\big\}$$

is independent mod p. By the ordering of the r_i, the last $(a-j)k$ of the integers $r_1,\ldots,r_{m-jk}$ are equal to $r = r_{m-jk}$. Continuing this process, we eventually arrive at (a), as required.

(b) Let $j_1,\dots,j_s$ be the indices i so that $s_i \neq s_{i+1}$ $(1 \leqslant s \leqslant t)$; $j_s = t$. Then the p-Smith form is divided into s segments of like powers of p, where the ith segment has length $k(j_i - j_{i-1})$. Group the last m columns of Q correspondingly, into s sets $A_1,\dots,A_s$; e.g., A_s consists of the last $k(j_s - j_{s-1})$ columns of Q. It is seen in the proof of (a) that $s_i < s_{i+1}$ is only possible in case $A_{i+1} \cup \cdots \cup A_s$ is a basis for W^j where $j = j_s - j_i$. Thus, A_s is a basis for W^{j_s}, A_{s-1} is an independent set of cardinality $k(j_{s-1})$ in $W^a - W^b$, where $a = j_{s-1} + j_s$ and $b = j_s$; and so on.

If $\bar{v} \in W^a - W^{a-1}$ for some a, then there exists some i so that $\bar{v}$ is in the span of $A_i \cup \cdots \cup A_s$, but not of $A_{i+1} \cup \cdots \cup A_s$. The span of $A_i \cup \cdots \cup A_s$ contains $W^b - W^{b-1}$, where $b = j_i + \cdots + j_s$. Then

$$f_i(C)^{b-a}(W^b - W^{b-1}) = W^a - W^{a-1}.$$

As in the proof of (a), since $R_{W^b}(X)$ has a basis of elements all of whose orders are $\geqslant p^{s_i}$, then so does $R_{W^a}(X)$. Put $V = W^a - W^{a-1}$; then $R_V(X)$ must contain k independent p^{s_i}-points.

We must now show that $R_V(X)$ contains no elements of higher order. Let p^{r_i} denote the power of p corresponding to column Q_i of Q. Suppose that $s > s_i$, that $(1/p^s)\bar{v} \in R_p(X)$ and $\bar{v}$ is in the span of $A_i \cup \cdots \cup A_s$ but not of $A_{i+1} \cup \cdots \cup A_s$. We can write

$$(1/p^s)\bar{v} = \sum_{r_i \leqslant s_i} n_i(1/p^{r_i})Q_i + \sum_{r_i > s_i} m_i(1/p^{r_i})Q_i + \bar{m}$$

($m_i, n_i \in \mathbf{Z}$ and $\bar{m} \in \mathbf{Z}^n$). Then

$$\bar{v} = \sum_{r_i \leqslant s_i} n_i(p^{s-r_i})Q_i + \sum_{r_i > s_i} m_i(1/p^{r_i - s})Q_i + p^s\bar{m}.$$

As before, we must have $p^{r_i - s}$ dividing each m_i. But $s > s_i$ implies $p^{s-r_i} \geqslant p$ for all $r_i \leqslant s_i$; thus mod p, we have $\bar{v}$ in the span of those Q_i with corresponding $r_i \geqslant s$, that is, in the span of $A_{i+1} \cup \cdots \cup A_s$, a contradiction.

Statements (c) and (d) follow from the fact that W is always in the span of A_s.

This completes the proof.

The next two results indicate how a knowledge of the group $R(X)$ may be helpful in factorization questions.

4.7. THEOREM. *Let $A, B \in \mathbf{Z}_{n\times n}$ be nonsingular. Then $R_p(A) \subseteq R_p(B)$ if and only if there is some integer k such that $(k, p) = 1$ and an integral matrix Y so that $kB = YA$.*

Proof. $R(A)$ is generated by the columns of A^{-1} reduced modulo 1. We can write $A^{-1} = (1/d_n)D$ where d_n is the largest invariant factor of A and D is integral. Suppose $p^r \| d_n$. Then $R_p(A)$ is generated by the columns of $(1/p^r)D$, reduced modulo 1. Put $d_n = kp^r$. Clearly, $R_p(A) \subseteq R_p(B)$ if and only if $B((1/p^r)D) = Y$ is integral, and the result follows. □

4.8. THEOREM. *Let $f_i(x)$ be an irreducible factor of $f(x)$ mod p, of degree $k_i = k$, with $f_i(x)^{e_i} \| f(x)$. Put $A = f_i(C)$ and $W = \ker_p(A)$. Suppose that $R_p(A) = R_W(A)$ does not contain any p^2-points. Then,*

(a) *$R_p(A) = (1/p)W$ and $|A| = p^k a$ with $(a, p) = 1$.*

(b) *$R_p(A^t) = (1/p)W^t$ for $t = 1, 2,\dots,e_i$.*

(c) *If* $B = h(C) \not\equiv 0 \pmod{p}$ *and if* $W \subseteq \ker_p B$, *then there exist integers* a *with* $(a, p) = 1$ *and* $t > 0$, *and an integral polynomial* $k(x)$ *such that*

$$W \cap \ker_p k(C) = \{0\} \quad \text{and} \quad aB = A^t k(C).$$

(d) *In* (c), *if* B *has a* p^2*-point for* W, *then* $t \geq e_i + 1$.

Proof. (a) By Lemma 4.4, $(1/p)W \subseteq R_p(A)$; by assumption, A has no p^2-points for W, hence $(1/p)W = R_W(A) = R_p(A)$. Then, by Theorem 4.6, the p-Smith form of A has $\dim W = k$ entries on its diagonal equal to p, and the rest equal 1. Then $|A| = p^k a$ for some integer a with $(a, p) = 1$.

(b) By Theorem 4.1(f) the dimension of $\ker_p A^t$ for $1 \leq t \leq e_i$ is tk (the degree of $f_i(x)^t$); so tk is the p-nullity of A^t. But $|A^t| = p^{kt}a^t$ and so the p-Smith form of A^t must have tk entries equal p and the rest equal 1. Then, A^t has no p^2-points, and (b) follows as in (a).

(c) This follows from a finite number of applications of Theorem 4.7, since whenever $W \subseteq \ker_p B$, then by (a), $R_p(A) \subseteq R_p(B)$.

(d) Let B be as in (c) above, and suppose $1 \leq t \leq e_i$. Put $B_1 = k(C)$, and suppose that there is a p^2-point for B in W, say $(1/p^2)\bar{v}$, $\bar{v} \not\equiv 0 \pmod{p}$. Since $W \cap \ker_p B_1 = \{0\}$, then $B_1\bar{v} \not\equiv 0 \pmod{p}$ and since $B_1 = k(C)$ and W is a C-invariant subspace, then $B_1\bar{v} \in W$. Then, we have

$$aB(1/p^2)\bar{v} = A^t(1/p^2)B_1\bar{v} \in \mathbf{Z}^n,$$

contradicting the fact that for $1 \leq t \leq e_i$, A^t has no p^2-points in W. Thus, $t \geq e_i + 1$ as required. □

We conclude this section with

4.9. THEOREM. *Let* $A = g(C) \in \mathbf{Z}[C]$ *with invariant factors* $d_1, d_2, \ldots, d_n$. *Suppose* $A \neq 0$, *so that* A *is nonsingular. Then* $\operatorname{adj} A = d_1 d_2 \cdots d_{n-1} B$, *where* $B = h(C) \in \mathbf{Z}[C]$ *and* $B \not\equiv 0 \pmod{p}$ *for every prime* p *dividing* d_n.

Proof. Let the characteristic polynomial for A be $a_0 + a_1 x + \cdots + x^n$. Then

$$A^{-1} = (1/|A|) \operatorname{adj} A = (-1/a_0)(A^{n-1} + a_{n-1}A^{n-2} + \cdots + a_1 I).$$

We have $a_0 = (-1)^n|A|$, $\pm|A| = d_1 d_2 \cdots d_n$, and d_n is the smallest positive integer such that $d_n A^{-1}$ is integral. Then

$$\operatorname{adj} A = \pm(A^{n-1} + \cdots + a_1 I) \equiv 0 \pmod{d_1 d_2 \cdots d_{n-1}}.$$

Now $A = g(C)$, so we can write $\operatorname{adj} A = k(C)$ for some $k(x) \in \mathbf{Z}[x]$. Since C is a companion matrix, then for $i = 0, 1, \ldots, n-1$, the first column of C^i is $\operatorname{col}(0, \ldots, 0, 1, 0, \ldots,)$ with the 1 in the $(i+1)$st place. Then, the first column of $k(C)$ consists of the coefficients of $k(x)$, and so all these coefficients are divisible by $d_1 \cdots d_{n-1}$. So $k(x) = d_1 \cdots d_{n-1} h(x)$ for $h(x) \in \mathbf{Z}[x]$, and we have $B = h(C)$. Since d_n is the smallest positive integer so that $d_n A^{-1} = B$ is integral, then $B \not\equiv 0 \pmod{p}$ for any prime $p \mid d_n$.

Example. Let $n = 3$ and p be a fixed prime, $X = g(C)$, $g(x) \in \mathbf{Z}[x]$. If $S_p(X) = \operatorname{diag}(1, 1, p^t)$ (i.e., if X has p-rank 2), then $g(x)$ is divisible by a linear factor of $f(x) \pmod{p}$. If $S_p(X) = \operatorname{diag}(1, p^t, p^t)$, then $g(x)$ is either divisible by two linear factors of $f(x)$ or by an irreducible quadratic factor of $f(x) \pmod{p}$. If $S_p(X) = \operatorname{diag}(1, p^u, p^v)$ and $0 < u < v$, then $g(x)$ is divisible by two linear factors of $f(x)$

(mod p), and *not* by an irreducible quadratic factor. Of course, if $S_p(X) = \operatorname{diag}(p^u, p^v, p^t)$, then $X \equiv 0 \bmod p$ anyway, and $g(x)$ is divisible by $f(x)$ mod p.

In the next section it is shown that much stronger statements can be made from a knowledge of $S_p(X)$ and $R_p(X)$.

5. Powers. Throughout this section, p is a fixed prime, and

$$f(x) = \prod_{i=1}^{r} f_i(x)^{e_i} \pmod p; \qquad \deg f_i(x) = k_i \tag{5.1}$$

is the factorization of $f(x)$ into prime factors in $\mathbf{Z}_p[x]$. We let $W_i^t = \ker_p f_i(C)^t$; $W_i^1 = W_i$; all polynomials will be assumed to have coefficients in $\mathbf{Z}$ (or $\mathbf{Z}_p$, according to context).

5.2. *Definition.* If $X = g(C)$ and $W_i \subseteq \ker_p(X)$, and if

$$t = \max\left\{ j\colon (1/p^j)\bar{a} \in R_{W_i}(X), \bar{a} \not\equiv 0 \bmod p \right\},$$

then we say that X *has power* t *for* W_i. Now put

$$j_i = \min\left\{ t\colon t \text{ is the power for } W_i \text{ for some } X = g(C) \text{ where } R_p(X) = R_{W_i}(X) \right\}.$$

Then we say that j_i is the *least power for* W_i, or for $f_i(x)$.

Finally, we say that $f_i(x)$ is an *honest factor* (of $f(x)$, mod p) provided the integer $f_i(\mathbf{r})$ (recall $\mathbf{r}$ is a root of $f(x)$ in $\mathbf{C}$) satisfies

$$(p) = (f_i(\mathbf{r}), p)^{e_i} V$$

for some ideal V which is relatively prime to $(f_i(\mathbf{r}), p)$; that is, the p-ideal for $f_i(\mathbf{r})$ actually divides (p) to the exact exponent e_i.

We shall see that the least power for W_i determines the honesty of $f_i(x)$; that it is possible for some f_i to be honest and not others; that if all f_i are honest, then (5.1) yields the complete prime factorization for the ideal (p) in $\mathscr{R}$, and finally, that a dishonest factor may be used to construct an integer $\beta \in \mathscr{R}$ so that $\beta/p \in \mathscr{R}$ also.

The hypotheses required for the results in this section may appear rather technical and difficult of application. We remedy this situation in the next section, where it is seen that the necessary conditions may be decided in a constructive way (which is simpler than one might expect).

We shall need

5.3. LEMMA. (a) *Let* $u_i = f_i(\mathbf{r}) \in \mathscr{R}$. *Then, the ideals* (u_i, p) *and* (u_j, p) *where* $i \neq j$, *are relatively prime.*

(b) *We may always suppose without loss of generality, that* $f_i(C)$ *has the least power* j *for* $f_i(x)$, *hence* $S_p(f_i(C)) = \operatorname{diag}(1, \ldots, p^j, \ldots, p^j)$ *with determinant* p^{kj}.

Proof. (a) This follows from the fact that, for $i \neq j$, the polynomials $f_i(x)$ and $f_j(x)$ are relatively prime in $\mathbf{Z}_p[x]$.

(b) Put $W = W_i$, $k = k_i$. Suppose $g(C)$ has $\ker_p g(C) = W$ and power j for W. Then, we must have $g(x) = q(x)f_i(x) + pr(x)$, where $\gcd(q(x), f(x)) = 1 \pmod p$; hence, we can write

$$\begin{aligned} a(x)q(x) + b(x)f(x) &\equiv 1 \pmod p, \\ a(x)g(x) &\equiv f_i(x) \pmod{p, f(x)}, \\ a(C)g(C) &= f_i(C) + pk(C), \end{aligned}$$

for some $k(x) \in \mathbf{Z}[x]$. Since $\gcd(a(x), f(x)) = 1$, then

$$R_p(a(C)g(C)) = R_p(g(C)) = R_p(f_i(C) + pk(C)).$$

Then, in the factorization (5.1), we can replace the factor $f_i(x)$ by the factor $f_i(x) + pk(x)$. By Theorem 4.6 we get the p-Smith form. □

5.4. THEOREM. *If $f_i(x)$ has least power 1, then it is honest.*

Proof. Put $A = g(C)$ with $R_p(A) = R_{W_i}(A) = (1/p)W_i$. Then $\gcd(f(x), g(x)) = f_i(x) \bmod p$. By Theorem 4.8, $|A| = p^{k_i}a$ where $(a, p) = 1$. By Theorem 4.9, we have $\operatorname{adj} A = p^{k_i-1}B$, where $B = h(C)$ for some $h(x) \in \mathbf{Z}[x]$, $B \not\equiv 0 \pmod p$ and $AB = paI_n$. Since $AB \equiv 0 \pmod p$ then $f(x)/f_i(x)$ divides $h(x) \pmod p$ and since $B \not\equiv 0 \pmod p$, $f(x) \nmid h(x)$. Then, by Theorem 4.8, we have some integer b with $(b, p) = 1$, and $k(x) \in \mathbf{Z}[x]$, so that

$$bB = A^{e_i-1}k(C) \quad \text{and} \quad W_i \cap \ker_p k(C) = \{0\}.$$

Put $\alpha = g(\mathbf{r})$ and $\beta = k(\mathbf{r})$; by Lemma 5.3, (α, p) and (β, p) are relatively prime. We have $bAB = A^{e_i}k(C) = pabI_n$, and so in $\mathscr{R}$,

$$\alpha^{e_i}\beta = pab.$$

Since $(ab, p) = 1$, then $(\alpha, p)^{e_i} \| (p)$ as required. □

The next result implies that any nonrepeated factor is honest.

5.5. THEOREM. *If the least power j_i for W_i is $\geqslant 2$, then $e_i \geqslant 2$.*

Proof. Assume that $A = f_i(C)$ has the least power j and put $W = W_i$. Write $f(x) = f_i(x)h(x) + r(x)$, where either $r(x) = 0$ or $0 \leqslant \deg r(x) < \deg f_i(x)$.

If $r(x) = 0$, then $f_i(x) = f(x)$ is irreducible mod p; $|X| \equiv 0 \pmod p$ if and only if $X \equiv 0 \bmod p$ for any $X = g(C)$, and the least power $j = 1$ trivially.

If $r(x) \neq 0$, then since $f_i(x) \mid f(x) \bmod p$, we can write $r(x) = p^t k(x)$ where $t \geqslant 1$ and $k(x) \not\equiv 0 \,(p)$. Since $\deg k(x) < \deg f_i(x)$, we have $W \cap \ker_p k(C) = \{0\}$ by Lemma 4.4. Since C is the companion matrix of $f(x)$,

$$f(C) = 0 = f_i(C)h(C) + p^t k(C), \qquad f_i(C)h(C) = -p^t k(C).$$

There is a p^j-point in $R_W(A)$, say $(1/p^j)\bar{a}$, and $k(C)\bar{a} \not\equiv 0 \bmod p$. But then

$$-k(C)(1/p^{j-t})\bar{a} = h(C)f_i(C)(1/p^j)\bar{a} \in \mathbf{Z}^n$$

implies $j \leqslant t$.

Now put $g(x) = f_i(x) + p$, and $f(x) = g(x)h_1(x) + p^m k_1(x)$. As above, and using the minimality of j, we have $m \geqslant j$. Then,

$$f_i(x)h(x) + p^t k(x) - (f_i(x) + p)h_1(x) - p^m k_1(x) = 0,$$
$$f_i(x)[h(x) - h_1(x)] - ph_1(x) + p^t k(x) - p^m k_1(x) = 0.$$

Since $R_W(A) \subseteq R_p(p^j I)$, it follows from Lemma 4.5 that $R_W(A) \subseteq R_W(ph_1(C))$. The least power j for W is greater than 1, and all terms other than $ph_1(C)$ have a common p^j-point, so $ph_1(C)$ must have at least a p^2-point for W (Lemma 4.5). Then $h_1(C)$ must have a p-point for W, and then $W \subseteq \ker_p h_1(C)$ (Lemma 4.4). From this, $f_i(x) \mid h_1(x) \bmod p$ (Lemma 4.4); of course $h(x) \equiv h_1(x) \bmod p$, and so $f_i(x) \mid h(x) \bmod p$. Then $f_i(x)$ is a repeated factor of $f(x) \bmod p$, and this completes the proof. □

Since any nonrepeated factor is honest, it is not difficult to find examples for $f(x)$ with some factors honest and some not (see Section 6).

The remainder of this section is devoted to establishing the following:

(*) There exists an integer $\beta = g(\mathbf{r}) \in \mathbf{Z}[\mathbf{r}]$ with $g(x) \not\equiv 0 \pmod{p}$ and $\beta/p \in \mathscr{R}$ if and only if at least one of the least powers for the irreducible factors $f_i(x)$ of $f(x)$ mod p, is greater than 1.

In fact, we shall show that *each* $f_i(x)$ with least power > 1, gives rise to such a β. We begin with a necessary condition for $\beta/p \in \mathscr{R}$.

5.6. LEMMA. *Let* $X = g(C)$ *and put* $y = g(\mathbf{r}) \in \mathscr{R}$. *If* $y/p \in \mathscr{R}$, *then necessarily*

$$f_1(x)f_2(x)\cdots f_r(x) \mid g(x) \pmod{p}.$$

Proof. Suppose y/p is an integer. This is to say that the characteristic polynomial of $(1/p)X$ has coefficients in $\mathbf{Z}$. Then the characteristic polynomial of X itself must have the form:

$$|\lambda I - X| = \sum_{i=0}^{n-1} \lambda^{n-i} p^i a_{n-i} \qquad (a_n = 1, a_i \in \mathbf{Z}).$$

Then over $\mathbf{Z}_p$, X is nilpotent. But $X = g(C)$ is nilpotent over $\mathbf{Z}_p$ if and only if $f_1(x)\cdots f_r(x) \mid g(x) \bmod p$. □

We now prove half of the statement (*).

5.7. THEOREM. *If every* $f_i(x)$ $(i = 1, 2, \ldots, r)$ *has least power* $j_i = 1$, *and if* $\beta = g(\mathbf{r}) \in \mathbf{Z}[\mathbf{r}]$, *where* $\beta/p \in \mathscr{R}$, *then* $g(x) \equiv 0 \bmod p$.

Proof. Assume to the contrary, that $X = g(C)$, $g(x) \not\equiv 0 \bmod p$, and $\beta = g(\mathbf{r})$ satisfies $\beta/p \in \mathscr{R}$. We may assume that $A_i = f_i(C)$ has power 1 for W_i. Put $\alpha_i = f_i(\mathbf{r})$ and $P_i = (\alpha_i, p)$. Then,

$$A_1^{e_1}A_2^{e_2}\cdots A_r^{e_r} = pyI_n \qquad ((y, p) = 1, y \in \mathbf{Z}),$$

$$P_1^{e_1}\cdots P_r^{e_r} = (p).$$

By Lemma 5.6 we have (ignoring factors relatively prime to p)

$$X = A_1^{s_1}\cdots A_r^{s_r} \qquad (s_i > 0, i = 1, \ldots, r).$$

If all $s_i \geqslant e_i$, then $g(x) \equiv 0 \bmod p$, and if all $s_i < e_i$, then $|X| < p^n$; thus our assumptions imply some $s_i \geqslant e_i$ and some $s_i < e_i$. For simplicity, suppose $s_r < e_r$ and all other $s_i \geqslant e_i$. We have

$$(1/p)X = (1/p)A_1^{e_1}A_1^{s_1 - e_1}\cdots A_r^{s_r}$$

$$(y/p)X = (A_2^{e_2}\cdots A_r^{e_r})^{-1}A_1^{s_1-e_1}A_2^{s_2}\cdots A_r^{s_r} = A_1^{s_1-e_1}\cdots A_r^{s_r-e_r}.$$

Now put $u = y\beta/p \in \mathscr{R}$; then we have

$$u\alpha_r^{e_r - s_r} = \alpha_1^{e_1 - s_1}\cdots \alpha_{r-1}^{s_{r-1} - e_{r-1}}$$

contradicting the fact that the ideals (α_i, p) are pairwise relatively prime (Lemma 5.3). □

5.8. COROLLARY. *The conditions of Theorem* 5.7 *imply*

(a) $p \nmid (\operatorname{disc} \mathbf{r})/(\operatorname{disc} \mathscr{R})$.

(b) *The factorization* (5.1) *yields the prime factorization of the ideal* (p); *that is, with* $P_i = (f_i(\mathbf{r}), p)$, *then* P_i *is a prime ideal* $(i = 1, \ldots, r)$ *and*

$$p = \prod_{i=1}^{r} P_i^{e_i}. \quad \square$$

We now complete the proof of (*) with

5.9. THEOREM. *Suppose $f_i(x)$ has least power $j \geqslant 2$ and that $A = f_i(C)$ has power j for W_i. Define the integral matrix B by*:

$$\operatorname{adj} A = p^{j(k_i-1)}B.$$

Then $B = h(C) \in \mathbf{Z}[C]$, where $h(x) \not\equiv 0 \pmod p$ and $h(\mathbf{r})/p \in \mathscr{R}$.

Proof. We have $AB = p^j \cdot y \cdot I_n$ with $(y, p) = 1$; without loss of generality assume $y = 1$. By Theorem 4.9, $B = h(C) \in \mathbf{Z}[C]$ and $h(x) \not\equiv 0 \pmod p$. Since $j \geqslant 2$, then $e = e_i \geqslant 2$ necessarily; since $AB \equiv 0 \pmod p$, then $f(x)/f_i(x) \mid h(x)$ and since $B \not\equiv 0 \pmod p$, $f_i(x)^e \nmid h(x)$; $f_i(x)^{e-1} \| h(x) \pmod p$, and we have

$$h(x) = f_i(x)^{e-1}u(x) + pv(x), \qquad (f_i(x), u(x)) = 1.$$

Then we can write $B = A^{e-1}X + pY$, where $X, Y \in \mathbf{Z}[C]$. Let the integers in $\mathscr{R}$ corresponding to A, B, X, Y be denoted by α, β, x, y. Then by Lemma 5.3, (α, p) and (x, p) are relatively prime. Suppose that in $\mathscr{R}$ the prime factorization of (p) is

$$(p) = P_1^{t_1} \cdots P_v^{t_v} Q_1^{u_1} \cdots Q_s^{u_s},$$

where the P_i are the primes dividing (α, p) and the Q_i are the primes dividing (x, p). From $\beta = \alpha^{e-1}x + py$ and $\alpha\beta = p^j$ it is clear that $Q_k^{u_k}$ divides (x, p), $k = 1, \ldots, s$.

Now suppose $(\alpha) \subseteq P_1^{a_1} \cdots P_v^{a_v}$ (a_i maximal) and consider $\alpha^e x + py\alpha = p^j$. We have

$$(\alpha^e x) \subseteq P_1^{ea_1} \cdots P_v^{ea_v}; \qquad (py\alpha) \subseteq P_1^{t_1+a_1} \cdots P_v^{t_v+a_v}.$$

If $e \cdot a_i < t_i + a_i$, then $ea_i = jt_i$, and $(j-1)t_i < a_i$. Since $j > 1$, then $t_i < a_i$, and we get $P_i^{t_i} \mid (\alpha)$. If $ea_i > t_i + a_i$, then $a_i = (j-1)t_i \geqslant t_i$, and again $P_i^{t_i} \mid (\alpha)$. If $ea_i = t_i + a_i$, then $(e-1)a_i = t_i$ and $P_i^{t_i} \mid (\alpha^{e-1})$. In all cases we have

$$P_i^{t_i} \mid (\alpha^{e-1}) \quad \text{and} \quad Q_i^{u_i} \mid (x)$$

and hence, $\alpha^{e-1}x \equiv 0 \bmod p$. Then $\beta = \alpha^{e-1}x + py = pz$ for some z in $\mathscr{R}$; $\beta/p \in \mathscr{R}$ as required. □

5.10. COROLLARY. *The conditions of Theorem* 5.9 *imply that* $p \mid (\operatorname{disc} \mathbf{r})/(\operatorname{disc} \mathscr{R})$.

6. Computational Methods and Examples. Given the factorization

$$f(x) = \prod_{i=1}^{r} f_i(x)^{e_i} \pmod p,$$

we wish to know whether a factor $f_i(x)$ is honest or not. It is if $e_i = 1$, but if $e_i > 1$, the general results of Section 5 do not look very helpful. Using these results, however, we can prove some things which, while not perhaps of great theoretical interest, are yet very convenient for computational purposes. We begin with

6.1. THEOREM. *Let j be the least power for the irreducible factor $f_i(x)$ of $f(x) \bmod p$. Suppose also that $e_i > 1$. Then $j \geqslant 2$ if and only if $R_p(f_i(C))$ contains a p^2-point.*

Proof. We have $e_i > 1$, and from Theorem 4.8, if $\ker_p A = W$ precisely, for some A, and if the least power for W were equal to 1, then A could not have any p^2-points for W. Hence, if A does have such points, the least power for W must be at least 2.

On the other hand, if the least power for W is $\geqslant 2$, then every matrix $A \in \mathbf{Z}[C]$ with $\ker_p A = W$ must have p^2-points for W; in particular this applies to $A = f_i(C)$. □

Procedure. (The procedure described below is not too difficult if n and p are not "too large".) (a) Compute $f_i(C)$, working mod p^2; (b) Still working mod p^2, row-reduce $f_i(C)$, usually to at least a row-echelon form; (c) If the corresponding system of linear equations has a nontrivial solution vector mod p^2, then this gives a p^2-point for $f_i(C)$, and if there is no solution mod p^2, then $f_i(C)$ has no p^2-points.

One could also deduce this information from the p-Smith form of $f_i(C)$, or from its determinant, but these things require more work than row-reduction as above. We give some examples below. The next theorems are occasionally helpful.

6.2. THEOREM. *Let $f_i(x)$ be an irreducible factor of $f(x)$, and $j = j_i$ the least power for $f_i(x)$. Then $j \geqslant 2$ if and only if whenever*

$$f(x) = g(x)h(x) + r(x) \quad (\text{in } \mathbf{Z}[x]),$$

$$g(x) \equiv f_i(x) \pmod{p},$$

$$r(x) = 0 \quad \text{or} \quad 0 \leqslant \deg r(x) < \deg g(x),$$

we have $r(x) \equiv 0 \pmod{p^2}$.

Proof. See the first part of the proof of Theorem 5.5. □

Example. Take $p = 3, f(x) = x^3 - 19$. We can write $f(x) = (x - 1)(x^2 + x + 1) - 18$; $18 \equiv 0 \pmod 9$ is suspicious and one must investigate further. If $p = 3$, $f(x) = x^3 - 4$, we write $f(x) = (x - 1)(x^2 + x + 1) - 3$; since $3 \not\equiv 0 \pmod 9$ this factorization is honest.

This theorem has a simple corollary which is also useful.

6.3. COROLLARY. *Suppose that for some prime p, we have $p^2 \mid a_0$, $p \mid a_1$ (where $f(x) = x^n - a_{n-1}x^{n-1} - \cdots - a_1x - a_0$); put $f_1(x) = x$. Then $\pmod p$ the least power for $f_1(x)$ is $\geqslant 2$, and $p \mid (\operatorname{disc} \mathbf{r}/\operatorname{disc} \mathscr{R})$.* □

6.4. THEOREM. *With notation as above, $j \geqslant 2$ if and only if there exist matrices X and Y in $\mathbf{Z}[C]$ such that $\ker_p X = ker_p Y = W_i = W$ but*

$$R_W(X) \not\subseteq R_W(Y) \quad \text{and} \quad R_W(Y) \not\subseteq R_W(X).$$

In fact, $j \geqslant 2$ if and only if $R_W(X)$ contains even one point not in $R_W(Y)$ or vice versa.

Proof. Suppose the contrary. Let $A = f_i(C)$ have the least power j for W, and put $B = A + pI_n$. Then $\ker_p B = \ker_p A = W$ and B has power $m \geqslant j$ for W. Suppose $R_p(A) \subseteq R_p(B)$. By Theorem 4.7, there exist an integer k so that $(k, p) = 1$ and an integral matrix Y so that $kBA^{-1} = Y = k(I_n + pA^{-1})$. But since A has power j for W, we may write $A^{-1} = (1/p^j a)D$ where $(a, p) = 1$ and D is an integral matrix, $D \not\equiv 0 \pmod P$. Thus $k(I_n + pA^{-1})$ cannot be integral since $j \geqslant 2$; a contradiction. Next suppose $R_p(B) \subseteq R_p(A)$. The order of $R_p(B)$ is p^{km} (k is the degree of $f_i(x)$) and the order of $R_p(A)$ is p^{kj}; since $m \geqslant j$, we must have in this case $m = j$. Then the argument above applies, and this completes the proof. □

The following example, which is fairly simple and could doubtlessly be done in many ways, serves to illustrate the methods of this paper.

6.5. *Example.* Let $f(x) = x^4 + x^3 + 9$, with root $\mathbf{r}$. Then $f(x)$ is irreducible since $x^4 + x^3 + 1$ is irreducible over $\mathbf{Z}_2$. We have

$$C = \begin{bmatrix} 0 & 0 & 0 & -9 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix};$$

$$|f'(C)| = \begin{vmatrix} 0 & -36 & 9 & -9 \\ 0 & 0 & -36 & 0 \\ 3 & 0 & 0 & -36 \\ 4 & -1 & 1 & -1 \end{vmatrix} = 9^3 \times 11 \times 23.$$

We have $\operatorname{disc} \mathbf{r} = k^2 \operatorname{disc} \mathscr{R}$; evidently k must be either a power of 3, or 1. Since $|C| = 3^2$, we know $3 \mid k$ by Theorem 6.1; and by Theorem 5.9 $(1/3)\operatorname{adj} C$ represents an algebraic integer. Here, $\operatorname{adj} C = -C^3 - C^2$, so $(\mathbf{r}^3 + \mathbf{r}^2)/3 \in \mathscr{R}$. It is well-known that $\mathscr{R}$ has an integral basis of the form $\{1, g_1(\mathbf{r})/d_1, g_2(\mathbf{r})/d_2, g_3(\mathbf{r})/d_3\}$, where the $g_i(x)$ are monic polynomials of degree i, in $\mathbf{Z}[x]$, and $d_i \mid d_{i+1}$. We shall find such a basis for $\mathscr{R}$. We have $f(x) = x^3(x+1) \pmod 3$; by Lemma 5.6, a necessary condition for $(1/3)g(\mathbf{r}) \in \mathscr{R}$ is that $x(x+1)$ divide $g(x) \pmod 3$. Hence, if $g(x) = x + a$, then $(1/3)g(\mathbf{r}) \notin \mathscr{R}$, and so $g_1(x) = x$, $d_1 = 1$. Next, if $g(x) = x^2 + sx + t = (x^2 + x) + (s-1)x + t$, and if $s - 1$, $t \not\equiv 0 \pmod 3$, then $(1/3)g(\mathbf{r}) \notin \mathscr{R}$; thus for $g_2(\mathbf{r})$ we need only consider $(\mathbf{r}^2 + \mathbf{r})$. The characteristic polynomial of the matrix

$$C^2 + C = \begin{bmatrix} 0 & 0 & -9 & 0 \\ 1 & 0 & 0 & -9 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

is $x^4 + c_3x^3 + c_2x^2 + c_1x + c_0$, where c_{4-i} is the sum of the principal $i \times i$ minors (except for sign). We find $c_1 = -9$, not divisible by 27, so $(1/3)(\mathbf{r}^2 + \mathbf{r}) \notin \mathscr{R}$; then $g_2(\mathbf{r}) = \mathbf{r}^2$ and $d_2 = 1$. We know that d_3 is at least 3; we inquire whether it may be 9 or more. Consider

$$g(x) = x^3 + ux^2 + sx + t = (x^2 + x)(x + u - 1) + (s - u + 1)x + t.$$

We require $s - u + 1 \equiv 0 \pmod 3$ and $t \equiv 0 \pmod 3$. Since $(\mathbf{r}^3 + \mathbf{r}^2)/3 \in \mathscr{R}$, then $g(\mathbf{r})/3 \in \mathscr{R}$ implies $(u-1)(\mathbf{r}^2 + \mathbf{r})/3 \in \mathscr{R}$ and from the preceding case we must have $u - 1 \equiv 0 \pmod 3$. Supposing $g(\mathbf{r})/3 \in \mathscr{R}$, we may write

$$g(\mathbf{r}) = (\mathbf{r}^2 + \mathbf{r})(\mathbf{r} + 3a) + 3(b\mathbf{r} + c) \qquad (a, b, c \in \mathbf{Z}),$$

$$\mathbf{r} \times g(\mathbf{r}) = (\mathbf{r}^4 + \mathbf{r}^3) + 3a(\mathbf{r}^3 + \mathbf{r}^2) + 3(b\mathbf{r}^2 + c\mathbf{r}),$$

$$9 + \mathbf{r} \times g(\mathbf{r}) = 3a(\mathbf{r}^3 + \mathbf{r}^2) + 3(b\mathbf{r}^2 + c\mathbf{r}).$$

If $(1/9)g(\mathbf{r}) \in \mathscr{R}$, then so is

$$(*) \qquad a(\mathbf{r}^3 + \mathbf{r}^2)/3 + (b\mathbf{r}^2 + c\mathbf{r})/3 \in \mathscr{R}.$$

Since $(\mathbf{r}^3 + \mathbf{r}^2)/3 \in \mathscr{R}$, then from $(*)$ $(b\mathbf{r}^2 + c\mathbf{r})/3 \in \mathscr{R}$. But then $b \equiv c \equiv 0 \pmod 3$. Reducing coefficients mod 9, we get

$$g(\mathbf{r}) = (\mathbf{r}^2 + \mathbf{r})(\mathbf{r} + 3a),$$

where a is one of 0, 1, 2. Then for the norm of $g(\mathbf{r})$, we find $|C^2 + C| = 81$ and $|C + 3aI| = 9(9a^4 - 3a^3 + 1)$, so $9^4 \nmid |g(C)|$. Then $(1/9)g(\mathbf{r}) \notin \mathscr{R}$, and we have $d_3 = 3$, $g_3(\mathbf{r}) = (\mathbf{r}^3 + \mathbf{r}^2)$. Finally, we have $\operatorname{disc} \mathscr{R} = 9^2 \times 11 \times 23$.

6.6. *Example.* Take $n = 7, p = 5$,

$$f(x) = x^7 + 48x^6 + 27x^5 + 48x^4 - 3x^3 - 3x + 48$$
$$= (x^3 + x + 1)^2(x - 2) + 25(2x^6 + x^5 + 2x^4 + 2).$$

Then $f(x)$ is irreducible by Eisenstein's criterion, and the factorization mod 5 is suspect. We know that in any case, the factor $x - 2 \pmod 5$ is honest, for it is not repeated. Put $g(x) = x^3 + x + 1$.

In order to use Theorem 6.1, it suffices to work with matrices reduced mod 25 in order to find 25-points for the originals. Below on the left is $A = C^3 + C + I_7$ (C is the companion matrix of $f(x)$), and on the right is the reduced row echelon form of A; both reduced mod 25.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2 & 4 & 6 \\ 1 & 1 & 0 & 0 & 3 & 8 & 13 \\ 0 & 1 & 1 & 0 & 0 & 3 & 8 \\ 1 & 0 & 1 & 1 & 3 & 6 & 12 \\ 0 & 1 & 0 & 1 & 3 & 7 & 12 \\ 0 & 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 1 & 2 & 3 & 5 \end{bmatrix} \qquad \left[\begin{array}{c|ccc} & 2 & 4 & 6 \\ I_4 & 1 & 4 & 7 \\ & -1 & -1 & 1 \\ & 2 & 3 & 5 \\ \hline 0 & & 0 & \end{array}\right]$$

Any vector in the solution set of the right-hand matrix yields a 25-point for A, for instance, we find $\bar{v} = (1/25)\ \mathrm{col}(-2, -1, 1, 2, 1, 0, 0)$ and $g(x)$ is dishonest. For purposes of comparison, consider $A + 5I_7$; this row-reduces to

$$\left[\begin{array}{c|ccc} & -8 & -16 & 1 \\ I_4 & 6 & 4 & 2 \\ & -1 & 4 & 1 \\ & 2 & 3 & 10 \\ \hline & -5 & -5 & -5 \\ 0 & 0 & 0 & 0 \\ & 0 & 0 & 0 \end{array}\right]$$

and this solution set gives a 25-point for $A + 5I_7$, namely

$$\bar{w} = (1/25)\ \mathrm{col}(8, 2, -5, -1, -1, 1, 0).$$

It suffices to use the (mod 25) versions to find that $A\bar{w} \in (1/5)\mathbf{Z}^7$ and not in $\mathbf{Z}^7$.

From the original factorization of $f(x)$ we know $A = g(C)$ has power 2, which must then be the least power; $\dim \ker_5 A = 3$, and $S_5(A) = \mathrm{diag}(1, 1, 1, 1, 5^2, 5^2, 5^2)$. Then $\mathrm{adj}\, A = 5^4 B$ where B is integral and $B \not\equiv 0 \pmod 5$; $(1/5)B$ represents an algebraic integer, and 5 divides $\mathrm{disc}(\mathbf{r})/\mathrm{disc}(\mathscr{R})$.

7. Open Problems.

7.1. By Lemma 5.6, a necessary condition for $(1/p)h(C)$ to represent an algebraic integer is that $h(C)$ should be nilpotent (mod p). Given $n \geqslant 3$, and supposing the factorization of $f(x)$ is dishonest, is there a $j = j(n)$ so that for any $h(C)$ with index of nilpotency $j \pmod p$, $(1/p)h(C)$ represents an algebraic integer? Does $j = 2$ always work?

7.2. If in the factorization 5.1 some $f_i(x)$ is honest (but not all), is $(f_i(\mathbf{r}), p)$ a prime ideal? If not, find a counterexample.

7.3. Suppose some of the $f_i(x)$ are honest and some not. The honest factors correspond to relatively prime ideal factors of (p) and each dishonest factor gives rise to a B with $(1/p)B$ representing an algebraic integer. Is this a new ideal factor of (p) (at least, relatively prime to the honest ones)?

7.4. What is the relation (if any) between the least powers for the $f_i(x)$ and the power of p dividing $\text{disc}(\mathbf{r})/\text{disc}(\mathscr{R})$?

7.5. If two dishonest factors have different least powers, do they give rise to different ideal factors of (p)? (Via Theorem 5.9, that is.)

Department of Mathematics
University of North Carolina at Greensboro
Greensboro, North Carolina 27412

1. KEN BYRD & THERESA P. VAUGHAN, "A group of integral points in a matrix parallelepiped," *Linear Algebra Appl.*, v. 30, 1980, pp. 155–166.

2. HARVEY COHN, *A Classical Invitation to Algebraic Numbers and Class Fields*, Springer-Verlag, Berlin and New York, 1978.

3. KENNETH HOFFMAN & RAY KUNZE, *Linear Algebra*, Prentice-Hall, Englewood Cliffs, N. J., 1961.

4. MORRIS NEWMAN, *Integral Matrices*, Academic Press, New York, 1972.