Can. J. Math., Vol. XXIX, No. 1, 1977, pp. 155–160

SUM OF SQUARES RINGS

BENJAMIN FINE

Introduction. One of the nicest results in elementary number theory is the following, giving the relation between quadratic residues and sums of squares.

THEOREM. *Let n be a positive integer. If -1 is a quadratic residue* mod n *then $n = u^2 + v^2$. Conversely, if $n = u^2 + v^2$ with $(u, v) = 1$, then -1 is a quadratic resdidue* mod n.

There are various proofs of this theorem (see [**4**]). In a recent paper [**2**] the above was reinterpreted and a new proof given in terms of the structure of the modular group $M = PSL_2(Z)$ – the group of linear fractional transformations $z' = (az + b)/(cz + d)$, with $ad - bc = 1$, and a, b, c, d integers. The result was shown to depend on the fact that M is group-theoretically a free product $Z_2 * Z_3$ [**6**]. It has been found that for many rings R (especially rings of integers over number fields) $PSL_2(R)$ is either a free product or a generalized free product [**1**; **3**]. The question then arises as to whether these rings satisfy sum of squares properties similar to Z.

1. Let R be a commutative ring with an identity (not a field) with -1 not a square in R. Then R is a *sum of squares ring* if it satisfies:

*SS*1. If $r \in R$ and -1 is a quadratic residue mod r then $r = \pm (u^2 + v^2)$.
*SS*2. If $r = u^2 + v^2$ with $(u, v) = 1$ then -1 is a quadratic residue mod r.

By the theorem, Z is a sum of squares ring. In this paper we will give sufficient conditions for rings to be sum of squares rings, and then give certain families of concrete examples (other than Z).

Our results depend on the following two lemmas; for simplicity, ring means commutative ring with identity.

LEMMA 1. *If R is a ring with no non-trivial idempotents and if $PSL_2(R)$ has only one conjugacy class of trace* 0 *(or more strongly of order* 2*) then R satisfies SS*1.

Proof. We must show that if -1 is a quadratic residue mod r, where $r \in R$, then $r = \pm (u^2 + v^2)$.

(1) Suppose that $\alpha^2 + 1 \equiv 0 \pmod{r}$ for some $\alpha \in R$ implies that $\alpha^2 + 1 = rs$ or, equivalently, $-\alpha^2 + rs = 1$. This implies that there is a matrix $T = \begin{bmatrix} \alpha & r \\ -s & -\alpha \end{bmatrix}$ of trace 0 and determinant $+1$. Thus $T \in SL_2(R)$.

Received March 27, 1976.

Since R has no non-trivial idempotents, $PSL_2(R) \cong SL_2(R)/(+I)$ (because $Z(SL_2(R)) = \{\pm I\}$) so the elements of $PSL_2(R)$ can be considered as matrices of determinant 1 which are identified with their negatives.

Let $\hat{T} = \pm \begin{bmatrix} \alpha & r \\ -s & -\alpha \end{bmatrix} \in PSL_2(R)$ correspond to T.

(2) Since there is only one conjugacy class of trace 0, $\hat{T}$ must be conjugate to

$$\hat{A} = \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

(3) Conjugating $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ by an arbitrary $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2(R)$ gives

$$BAB^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} d & -b \\ c & a \end{bmatrix} = \begin{bmatrix} -(ac+bd) & a^2+b^2 \\ -(c^2+d^2) & ac+bd \end{bmatrix}.$$

Now $\hat{T}$ is conjugate to $\hat{A}$, so $\hat{T}$ must be of this form with r as the upper right entry. However, since BAB^{-1} is unique only up to $+$ or $-$, it follows that $r = \pm (a^2 + b^2)$.

Note that in the general case, if there exist non-trivial idempotents then the same proof shows that if -1 is a quadratic residue mod r then $r = e(u^2 + v^2)$ for some idempotent e.

Furthermore, any element of trace 0 has order 2 in $PSL_2(R)$ [**6**], so the lemma is true under the stronger condition that $PSL_2(R)$ has only one conjugacy class of elements of order 2.

Condition $SS2$ is handled by a second lemma.

LEMMA 2. *If R is a ring such that the* g.c.d. *of any two elements is expressible as a linear combination of them, then R satisfies $SS2$.*

Proof. Suppose $r = u^2 + v^2$ with $(u, v) = 1$. Since g.c.d.'s are linearly expressible, there exist x and y in R such that $ux + vy = 1$. This implies that there is a matrix $M = \begin{bmatrix} u & v \\ -y & x \end{bmatrix}$ with determinant 1; hence $M \in SL_2(R)$.

Now $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \in SL_2(R)$. Conjugating A by M gives

$$MAM^{-1} = \begin{bmatrix} -\alpha & u^2+v^2 \\ -(x^2+y^2) & \alpha \end{bmatrix} \quad \text{where } \alpha = uy + vx \Rightarrow \begin{bmatrix} -\alpha & r \\ s & \alpha \end{bmatrix}.$$

Since determinants are preserved by conjugation, $-\alpha^2 - rs = 1$ or $\alpha^2 + 1 = -rs$ which implies that -1 is a quadratic residue mod r.

COROLLARY. *Euclidean rings, quotients of Euclidean rings, and polynomial rings satisfy $SS2$.*

These follow from the fact that g.c.d.'s are linearly expressible in these types of rings.

2. Using these results, we first give sufficient conditions for a Euclidean domain to be a sum of squares ring.

THEOREM 1. *A Euclidean domain D with trivial units and* char $D \neq 2$ *is a sum of squares ring if its norm function N is subadditive and $0 \neq N(b) \leqq N(a)$ implies $N(a + kb) < N(a)$ for some k in R.*

Proof. By the corollary, a Euclidean ring satisfies $SS2$ so we need only show $SS1$. We do this by showing $PSL_2(D)$ has one conjugacy class of trace 0. Since there are trivial units, this is sufficient by Lemma 1.

Let $\hat{T} \in PSL_2(D)$ with tr $\hat{T} = 0$. So $\hat{T} = \pm \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$. We will show $\hat{T}$ is conjugate to $\hat{A} = \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$.

Let $S = \{\text{conjugates of } \hat{T} \text{ in } PSL_2(D)\}$ and let $\hat{V} = \pm \begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix} \in$ with $N(\alpha)$ minimal (since N is integer valued, $\hat{V}$ exists).

If $N(\alpha) = 0$ then $\alpha = 0$, so

$$\hat{V} = \pm \begin{bmatrix} 0 & \beta \\ \gamma & 0 \end{bmatrix} = \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

since $-\beta\gamma = 1$ and D has trivial units.

Suppose $N(\alpha) > 0$. Then since $\alpha^2 + 1 = -\beta\gamma$ it follows that $N(\beta) \cdot N(\gamma) = N(\alpha^2 + 1) \leqq N(\alpha)^2 + 1$ (since N is subadditive). Therefore, $N(\beta) \leqq N(\alpha)$ or $N(\gamma) \leqq N(\alpha)$, for if $N(\beta) \geqq N(\alpha) + 1$, or $N(\gamma) \geqq N(\alpha) + 1$ we would have that $N(\beta) \cdot N(\alpha) \geqq N(\alpha)^2 + 2N(\alpha) + 1 > N(\alpha)^2 + 1$.

Assume $N(\gamma) \leqq N(\alpha)$ (the proof works equally well if $N(\beta) \leqq N(\alpha)$). Then by assumption, there exists a k in R such that $N(\alpha + k\gamma) < N(\alpha)$, unless $\gamma = 0$.

Conjugate $\hat{V} = \pm \begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix}$ by $\hat{U} = \pm \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$. Then $= \hat{W}\ \hat{U}\hat{V}\hat{U}^{-1} \in S$, but

$$\hat{W} = \pm \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & -\alpha \end{bmatrix} \begin{bmatrix} 1 & -k \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha + k\gamma & \beta - k^2\gamma - \alpha k\alpha \\ \gamma & -\alpha - k\gamma \end{bmatrix}.$$

But $N(\alpha + k\gamma) < N(\alpha)$, contradicting the minimality of $N(\alpha)$, unless $\gamma = 0$. But then $\hat{T} = \pm \begin{bmatrix} \alpha & \beta \\ 0 & -\alpha \end{bmatrix}$ and this implies that $-\alpha^2 = 1$, or $\alpha^2 = -1$, which is impossible since D has trivial units and char $D \neq 2$. It follows that $N(\alpha) = 0$ and $\hat{V} = \pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, completing the proof.

From the above theorem we can recover the fact that Z is a sum of squares ring since absolute value is subadditive and $0 < b \leqq a$ implies that $0 \leqq a - b < a$.

https://doi.org/10.4153/CJM-1977-014-3 Published online by Cambridge University Press

3. We now present some concrete examples.

THEOREM 2. *Let p be a prime. If $p \equiv 3 \pmod 4$ and $n > 1$, then Z_{p^n}, the ring of integers* mod p^n, *is a sum of squares ring.*

Proof. Z_{p^n} is a commutative ring with identity. The condition $p \equiv 3 \pmod 4$ guarantees that -1 is not a square in Z_{p^n}. Also, since $n > 1$, Z_{p^n} is not a field.

Since Z_{p^n} is a quotient of Z, Z_{p^n} satisfies $SS2$ by the corollary to Lemma 2.

We must then show it satisfies $SS1$. Since $a^2 \equiv 1 \pmod{p^n}$ implies that $a \equiv \pm 1 \pmod{p^n}$, it must be that Z_{p^n} has no non-trivial idempotents, so $SS1$ will follow if $PSL_2(Z_{p^n})$ has one conjugacy of trace 0. But this follows from a result of D. L. McQuillan [**5**], which showed that all elements of $PSL_2(Z_{p^n})$ of trace 0 are conjugate, for all primes p and all $n \geqq 1$.

Let us now consider the case of a field, in particular a polynomial ring over a field.

THEOREM 3. *Let K be a field, in which -1 is not a square. If $PSL_2(K)$ has one conjugacy class of trace* 0, *then the polynomial ring $K[x]$ is a sum of squares ring.*

Proof. $K[x]$ satisfies $SS2$ from the corollary to Lemma 2, so we must show it satisfies $SS1$. Since $K[x]$ has no non-trivial idempotents, this will follow if $PSL_2(K[x])$ has one conjugacy class of trace 0.

Let $T = \pm \begin{bmatrix} F & G \\ H & -F \end{bmatrix} \in PSL_2(K[x])$. We will show that T is conjugate to an element of $PSL_2(K)$.

Let $S = \{\text{conjugates of } T\}$ and let $V = \pm \begin{bmatrix} f & g \\ h & -f \end{bmatrix} \in S$ with $\deg f$ minimal. If $g = 0$ or $h = 0$, then $-f^2 = 1$ or $f^2 = -1$, contradicting the fact that -1 is not a square in K (only units in $K[x]$ are in K). Hence $g \neq 0$ and $h \neq 0$. If $f \in K$, then since $f^2 + 1 = -gh$ implies that $\deg g + \deg h = 2 \deg f$, it follows that g and h are also in K so $V \in PSL_2(K)$.

Finally, assume $\deg f \geqq 1$. Since $f^2 + 1 = -gh$ implies that $\deg g + \deg h = 2 \deg f$, it follows that $\deg g \leqq \deg f$ or $\deg h \leqq \deg f$. Since K is a field, by the division algorithm we can find a $q \in K[x]$ such that $\deg (f + qg) < \deg f$. Conjugating V by $U = \begin{bmatrix} 1 & q \\ 0 & 1 \end{bmatrix}$ gives

$$W = UVU^{-1} = \pm \begin{bmatrix} 1 & q \\ 0 & 1 \end{bmatrix} \begin{bmatrix} f & g \\ h & -f \end{bmatrix} \cdot \begin{bmatrix} 1 & -q \\ 0 & 1 \end{bmatrix} = \pm \begin{bmatrix} f + qg & * \\ * & * \end{bmatrix}$$

which is in S and $\deg (f + qg) < \deg f$, contradicting the minimality of $\deg f$. Therefore, $\deg f \ngeqq 1$ so $f \in K$ implies that $g, h \in K$ (since $g, h \neq 0$) and $V = \begin{bmatrix} f & g \\ h & -f \end{bmatrix} \in PSL_2(K)$.

Since $PSL_2(K)$ has only one conjugacy class of trace 0, and every T with $\operatorname{tr} T = 0 \in PSL_2(K[x])$ is conjugate to $V \in PSL_2(K)$, then $PSL_2(K[x])$ has

only one conjugacy class of trace 0, showing $K[x]$ satisfies $SS1$ and is therefore a sum of squares ring.

THEOREM 4. *If p is a prime, $p \equiv 3 \pmod 4$, then $Zp[x]$ is a sum of squares ring (Zp = finite field of order p).*

Proof. This follows from Theorem 3 and McQuillan's results cited above. (The condition $p \equiv 3 \pmod 4$ is necessary so that -1 is not a square in Zp.)

THEOREM 5. *If F is an ordered field, and if every positive element of F has a square root, then $F[x]$ is a sum of squares ring.*

Proof. Since F is ordered, -1 is not a square. Therefore, by Theorem 3 it suffices to show $PSL_2(F)$ has only one conjugacy class of trace 0.

Let $T = \pm \begin{bmatrix} \alpha & r \\ \hat{s} & -\alpha \end{bmatrix}$. We will show that T is conjugate to $\pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Since $\alpha^2 + 1 = -r\hat{s}$, we have $-r\hat{s} > 0$. Assume $r > 0$, so $\hat{s} < 0$. Let $s = -\hat{s} > 0$.

Conjugating $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ by arbitrary $\begin{bmatrix} u & v \\ m & n \end{bmatrix} \in SL_2(F)$ gives

$$\begin{bmatrix} u & v \\ m & n \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} n & -v \\ -m & u \end{bmatrix} = \begin{bmatrix} -(vn + um) & u^2 + v^2 \\ -(n^2 + m^2) & vn + um \end{bmatrix}.$$

Therefore, to show T is conjugate to $\pm\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ we must find u, v, m, n such that

(1) $u^2 + v^2 = r$,
(2) $m^2 + n^2 = S$,
(3) $un - vm = 1$, and
(4) $-(vn + um) = 2$.

Now solving (1), (2), (3) is sufficient for if u, v, m, n satisfy (1), (2), (3) then

$$\begin{bmatrix} u & v \\ m & n \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} n & -v \\ -m & u \end{bmatrix} = \begin{bmatrix} \beta & r \\ \hat{s} & -\beta \end{bmatrix}$$

implies $\beta^2 = \alpha^2$ or $\beta = \pm\alpha = \pm(vn + um)$.

If $\alpha = -vn - um$ we are done. If $\alpha = vn + um$, then $-u$, $-n$, v, m also satisfy (1), (2), (3) and give the right value for α.

Now we show that (1), (2), (3) can be solved. First choose u and v with $u^2 + v^2 = r$ and $u \neq 0$. This can be done since $r > 0$, and every positive element has a square root. We now solve for n, m. From (3), $n = (1 + vm)/u$ and from (2), $m^2 + n^2 = m^2 + (1 + vm)^2/u^2 = s$. It follows that $u^2m^2 + v^2m^2 + 2vm + 1 - su^2 = 0$ or $(u^2 + v^2)m^2 + 2vm + 1 - su^2 = 0$. Since every positive element has a square root, this is solvable for m if the discriminant is non-negative, that is, if

$$4v^2 - 4(u^2 + v^2)(1 - su^2) \geqq 0$$

$$\text{(equivalently, } v^2 - (u^2 + v^2)(1 - su^2) \geqq 0)$$

or if

$$su^4 + su^2v^2 - u^2 \geqq 0 \quad \text{(equivalently, } u^2(su^2 + sv^2 - 1) \geqq 0\text{)}.$$

But since $u \neq 0$ it must be that $su^2 + sv^2 - 1 \geqq 0$, (that is, $s(u^2 + v^2) - 1 \geqq 0$) or $sr - 1 \geqq 0$.

But $sr = 1 + \alpha^2$ implies that $sr - 1 = \alpha^2 \geqq 0$, so it is solvable for m. Therefore T is conjugate to $\pm \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $PSL_2(F)$ has one conjugacy class of trace 0, so $F[x]$ is a sum of squares ring.

As an immediate consequence we get the following:

COROLLARY. *If R denotes the real field, and A is the field of real algebraic numbers over Q, then $R[x]$, and $A[x]$ are sum of squares rings.*

4. In closing, two questions are presented which arise in the light of the examples.

(1) If R is both an integral domain and a sum of squares ring, must it be a unique factorization domain? In the examples, Z and $K[x]$ were both, and these are *UFD*'s.

(2) If R is a sum of squares ring, is $R[x]$ also a sum of squares ring? In particular, is $Z[x]$ a sum of squares ring?

REFERENCES

1. Benjamin Fine, *The HNN and generalized free product structure of certain linear groups*, Bull. A.M.S. *81* (1975) 413–416.
2. ——— *A note on quadratic residues and sums of squares*, submitted, Proceedings of AMS. 1976.
3. ——— *The structure of $PSL_2(R)$*, Annals of Mathematics Study *79*, Discontinuous Groups and Riemann Surfaces (1974), 145–169.
4. G. H. Hardy and E. M. Wright, *The theory of numbers* (Clarendon Press, 1960).
5. D. L. McQuillan, *Classification of normal congruence subgroups of the modular group*, Amer. J. Math. *87* (1965), 285–296.
6. Morris Newman, *Integral matrices* (Academic Press, New York, 1972).

Fairfield University,
Fairfield, Connecticut