Results. Math. 64 (2013), 59–65

1422-6383/13/010059-7
published online November 1, 2012
DOI 10.1007/s00025-012-0296-0

Results in Mathematics

Polynomials That Force a Unital Ring to be Commutative

S. M. Buckley and D. MacHale

Abstract. We characterize polynomials f with integer coefficients such that a ring with unity R is necessarily commutative if $f(R) = 0$, in the sense that $f(x) = 0$ for all $x \in R$. Such a polynomial must be primitive, and for primitive polynomials the condition $f(R) = 0$ forces R to have nonzero characteristic. The task is then reduced to considering rings of prime power characteristic and the main step towards the full characterization is a characterization of polynomials f such that R is necessarily commutative if $f(R) = 0$ and R is a unital ring of characteristic some power of a fixed prime p.

Mathematics Subject Classification (2000). 16R50.

Keywords. Unital ring, Polynomial identity, Commutativity, Monoid ring.

1. Introduction

In this paper, we discuss the sets of polynomials $f(X) = \sum_{i=0}^{n} a_i X^i \in \mathbb{Z}[X]$ that force certain classes $\mathcal{F}$ of rings to be commutative, i.e. those polynomials $f(X)$ such that if $f(x) = 0$ for all $x \in R$, where $R \in \mathcal{F}$, then R is necessarily commutative. We allow a_0 to be non-zero only when $\mathcal{F}$ is a class of rings with unity.

The equation $f(R) = 0$ means that $f(x) = 0$ for all $x \in R$. Denote by $C(\mathcal{F})$ the set of polynomials f such that R is commutative whenever $R \in \mathcal{F}$ satisfies $f(R) = 0$. Let $\mathcal{R}$ and $\widetilde{\mathcal{R}}$ be the classes of all rings, and all rings with unity, respectively. For each prime number p, define $\mathcal{R}_p$ to be the class of rings R such that $p^k R = 0$ for some $k \geq 0$, and define $\widetilde{\mathcal{R}}_p$ to be $\widetilde{\mathcal{R}} \cap \mathcal{R}_p$.

A well-known result of Jacobson [2, Theorem 11] implies that for $n > 1, X^n - X \in C(\mathcal{R})$. More generally, Herstein [1] showed that if $a_1 = \pm 1$ then $f \in C(\mathcal{R})$, and we call such polynomials *Herstein polynomials* below; see also [5]. For an overview of work on ring commutativity, see [4].

Using Herstein's result, the second author and Laffey [3] characterized $C(\mathcal{R})$: they showed that $f \in C(\mathcal{R})$ if and only if f is either a Herstein polynomial, or f satisfies the following three conditions: $a_1 = \pm 2, a_2$ is odd, and $\sum_{i=2}^{n} a_i$ is odd.

In view of this characterization of $C(\mathcal{R})$, we are led to investigate $C(\widetilde{\mathcal{R}})$, which we will see is strictly larger than $C(\mathcal{R})$. In order to state our main theorem, which characterizes $C(\widetilde{\mathcal{R}})$, we first need to define the following family of conditions on f indexed by a prime number p:

There is at least one non-multiple of p among the numbers

$$S_p := \{a_0, a_1\} \cup \{b_j, c_j \mid 0 \le j < p-1\},$$

where

$$\left.\begin{aligned} b_j &= \sum_{\substack{0\le i\le n \\ i \equiv j \pmod{p-1}}} i a_i \\ c_j &= \sum_{\substack{0\le i\le n \\ i \equiv j \pmod{p-1}}} a_i \end{aligned}\right\}, \quad 0 \le j < p-1$$

Whenever the above condition holds, we say that f *satisfies the S_p condition.* We also say that f is *primitive* if the greatest common divisor of the numbers $\{a_i\}_{i=0}^{n}$ is 1.

Theorem 1. *Suppose $f(X) = \sum_{i=0}^{n} a_i X^i \in \mathbb{Z}[X]$. Then $f \in C(\widetilde{\mathcal{R}})$ if and only if it is primitive and it satisfies the S_p condition for all primes $p \le n/2$ that divide a_1.*

After some preliminaries in Sect. 2, we prove the above theorem and related results in Sect. 3.

2. Preliminaries

Primitivity is a rather obvious necessary condition for $f \in C(\widetilde{\mathcal{R}})$: given any prime p, the ring $GL_2(\mathbb{F}_p)$ is noncommutative and of characteristic p, so if every coefficient of f is divisible by p then $f \notin C(\widetilde{\mathcal{R}}_p) \supset C(\widetilde{\mathcal{R}})$.

Suppose $f(X) \in \mathbb{Z}[X]$ is not the zero polynomial, and $f(R) = 0$ for some $R \in \widetilde{\mathcal{R}}$. The fact that $f(n \cdot 1) = 0$ for all $n \in \mathbb{Z}$ means that R cannot have characteristic 0, so suppose it has characteristic $m > 0$.

Suppose $m = \prod_{p|m} p^{k_p}$ is the prime factorization of the characteristic m of a ring R such that $f(R) = 0$. We write $m_p := m/p^{k_p}$ and $R_p := m_p R$. The following statements are readily verified.

- Each R_p is an ideal in R.
- R_p has characteristic p^{k_p} and $f(R_p) = 0$.
- If R has a unity, so do each of the rings R_p. In fact, since m_p is coprime to p, it has some inverse m_p^{-1} in $\mathbb{Z}_{p^{k_p}}$, and $m_p(m_p^{-1} \cdot 1)$ is a unity in R_p.

– Since the gcd of the numbers $\{m_p : p \mid n\}$ is $1, R$ is the sum of these ideals: in fact if we choose numbers n_p such that $\sum_p n_p m_p$ equals 1 mod m, then $x = \sum_p n_p(m_p x)$.
– $R_p \cap \left(\sum_{q \neq p} R_q\right) = 0$.

It follows from the above that a ring R of characteristic m is an internal direct sum of rings R_p of prime power characteristic. Since the conditions that R has a unity, $f(R) = 0$, and R is commutative each hold if and only if the same condition holds on R_p for each p, it follows that

$$C(\widetilde{\mathcal{R}}) = \bigcap_{p \text{ prime}} C(\widetilde{\mathcal{R}}_p).$$

This allows us to characterize $C(\widetilde{\mathcal{R}})$ by characterizing $C(\widetilde{\mathcal{R}}_p)$ for all primes p.

If $R \in \widetilde{\mathcal{R}}$ has characteristic p^k, and a_0 is not a multiple of p^k, then $f(0) = a_0 \cdot 1 = 0$ contradicts the fact that R has characteristic p^k. We may therefore assume that a_0 is a multiple of p^k when handling rings of characteristic p^k. Since multiples of p^k can be ignored in such rings, we may in fact assume that $a_0 = 0$. This argument works for all primes p and $k \in \mathbb{N}$, thus reducing the task at hand to characterizing when polynomials satisfying $a_0 = 0$ lie in $C(\widetilde{\mathcal{R}})$ (or $C(\widetilde{\mathcal{R}}_p)$).

3. Proof of Main Result

The main step in proving Theorem 1 is the following characterization of $C(\widetilde{\mathcal{R}}_p)$.

Theorem 2. *Suppose $f(X) = \sum_{i=0}^{n} a_i X^i \in \mathbb{Z}[X]$, and let p be a prime. Then $f \in C(\widetilde{\mathcal{R}}_p)$ if and only if f satisfies the S_p condition.*

Proof. To prove necessity of the S_p condition, we first define the monoid ring R_T to be the collection of all functions $\phi : G \to T$, where $(T, +, \cdot)$ is a commutative ring with unity, and $G = \{1, u, v\}$ is the monoid with the following multiplication table:

$\cdot$	1	u	v
1	1	u	v
u	u	u	u
v	v	v	v

Thus $(R_T, +)$ consists of the direct sum of three copies of T, and multiplication in R_T is given in coordinate form by

$$(\alpha, \beta, \gamma) \cdot (\alpha', \beta', \gamma') = (\ \alpha\alpha',\ \alpha\beta' + \beta(\alpha' + \beta' + \gamma'),\ \alpha\gamma' + \gamma(\alpha' + \beta' + \gamma')\ ). \qquad (1)$$

Examining the multiplication table for G, it is clear that $(xy)z = x(yz)$ if any of x, y, z equals 1, while if $x, y, z \in \{u, v\}$, then $(xy)z = x(yz) = x$. Thus $\cdot$ is associative in G, and so R_T is a noncommutative ring with unity.

We need this construction only for $T = \mathbb{Z}_p$, where p is a prime. In this case, we write R_p in place of R_T, and we claim that for all $\alpha, \beta, \gamma \in \mathbb{Z}_p$,

$$(\alpha, \beta, \gamma)^p = \left(\alpha,\ \beta(\beta+\gamma)^{p-1},\ \gamma(\beta+\gamma)^{p-1}\right). \tag{2}$$

The equation of the first coordinates follows immediately from the fact that $\alpha^p = \alpha$ according to Fermat's Little Theorem. We get a contribution to the second coefficient from any product involving $p-k$ choices of u or v factors, as long as the leftmost of these factors involves a u, together with k choices in any positions of the factor $\alpha \cdot 1$, where $0 \leq k < p$. For each such k, we therefore get a contribution to the second coefficient of $\binom{p}{k}\alpha^k\beta(\beta+\gamma)^{p-k-1}$. Now $\binom{p}{k}$ is divisible by p for all $0 < k < p$, and we can handle the final coordinate in the same way, so the claim follows.

Suppose now that $f \in \mathbb{Z}[X]$ is such that the associated set of numbers S_p are all multiples of p. Let $x := (\alpha, \beta, \gamma)$. If $\beta + \gamma \neq 0$, then $(\beta+\gamma)^{p-1} = 1$ (in $\mathbb{Z}_p$, of course), and so it follows from (2) that $x^p = x$. It follows that $f(x) = g(x) := a_0 + \sum_{i=0}^{p-2} d_i x^i$, where $d_0 = c_0 - a_0$ and $d_i = c_i$ for all $1 \leq i < p-1$. Since $p \mid a_0$ and $p \mid c_i$ for all i, the coefficients of the polynomial g are all divisible by p. But R_p has characteristic p, so $f(x) = g(x) = 0$ for all such x.

If instead $\beta + \gamma = \alpha = 0$, then $x^2 = 0$ and so $f(x) = a_0 + a_1 x$. Since both a_1 and a_1 are divisible by p, it again follows that $f(x) = 0$.

Finally suppose $\beta + \gamma = 0$ but $\alpha \neq 0$. In this case it is readily verified that for all $i \in \mathbb{N}$, we have

$$x^i = (\alpha^i, i\alpha^{i-1}\beta, i\alpha^{i-1}\gamma). \tag{3}$$

To prove that $f(x) = 0$, it suffices to show that $\pi_k(f(x)) = 0$ for $k = 1, 2, 3$, where $\pi_k : R_p \to \mathbb{Z}_p$ is projection onto the kth coordinate.

Now $\pi_1(x^p) = \pi_1(x)$, so $\pi_1(f(x)) = a_0 \cdot 1 + \sum_{i=1}^{p-1} d_i\alpha^i$, where d_i is as before. Now a_0 and each of the d_is are divisible by p, so $\pi_1(f(x)) = 0$.

As for $\pi_2(f(x))$, it follows from (3) that

$$\pi_2(f(x)) = \left(a_0 \cdot 1 + \sum_{i=0}^{p-2} b_i \alpha^{i-1}\right)\beta.$$

Since $p \mid a_0$ and $p \mid b_i$ for all i, it follows that $\pi_2(f(x)) = 0$. Lastly $\pi_3(f(x))$ is handled like $\pi_2(f(x))$.

In summary we have shown that if S_p contains only numbers divisible by p, then $f(R_p) = 0$. Since R_p is noncommutative for all p, we have shown that $f \notin C(\widetilde{\mathcal{R}}_p)$ if all numbers in S_p are divisible by p.

We now prove sufficiency of the S_p condition. The (trivial) ring of characteristic 1 is certainly commutative, so we may assume that $R \in \widetilde{\mathcal{R}}$ has characteristic p^k for some $k \in \mathbb{N}$. When considering $f(R) = 0$ for such rings, we may treat the coefficients of f as being either elements of $\mathbb{Z}_{p^k}$, or elements

of $\mathbb{Z}$, as suits us. As discussed in Sect. 2, a_0 is necessarily a multiple of p^k, and so we may take it to be 0.

If $p \nmid a_1$, then a_1 is a unit mod p^k, so $g(X) := a_1^{-1} f(X) \in \mathbb{Z}_{p^k}[X]$ has the form $X + \sum_{i=2}^{n} d_i X^i$, and so it is a Herstein polynomial when we view its coefficients as being integers. In particular the condition $g(R) = 0$ forces characteristic p^k rings $R \in \widetilde{\mathcal{R}}$ to be commutative. We may therefore assume that $p \mid a_1$.

Suppose next that there exists $i, 0 \leq i < p-1$, such that $p \nmid b_i$. We treat $f(X)$ as a polynomial in $\mathbb{Z}_{p^k}[X]$, but let us also write $f_p(X)$ for $f(X)$ when instead viewed as an element of $\mathbb{Z}_p[X]$. Expanding $f_p(X+t)$ for $t \in \mathbb{Z}_p$, we see that the coefficient of X is $s_p(t) := \sum_{i=1}^{n} i a_i t^i$. Let $S_p(X) := \sum_{i=0}^{p-1} b_i X^i \in \mathbb{Z}_p[X]$. By Fermat's Little Theorem, $s_p(t) = S_p(t)$ for all $t \in \mathbb{Z}_p$. The fact that $p \nmid b_i$ for some i means that S_p is not the zero polynomial, and so it has at most $p-1$ roots. Thus there exists $t \in \mathbb{Z}_p$ such that $s_p(t) \neq 0$. It follows that the coefficient of X in the expansion of $f(X + t \cdot 1)$ is coprime to p for some $t \in \mathbb{Z}_{p^k}$. Fixing this value of t and picking $k \in \mathbb{Z}_{p^k}$ which is equivalent to t mod p, we get a polynomial $g(X) := f(X+k) \in \mathbb{Z}_{p^k}[X]$ such that $g(R) = 0$ and such that the coefficient of X in g is a unit mod p^k. This implies the commutativity of R as before.

Lastly, suppose that for all $0 \leq i < p-1, p \mid b_i$, but there exists some such i for which $p \nmid c_i$. Since $p \mid b_i$ for all i, it follows from the analysis in the previous paragraph that $s_p(t) = 0$ for all $t \in \mathbb{Z}_p$. If we expand $(X+t) f_p(X+t) \in \mathbb{Z}_p[X]$, we get a polynomial whose coefficient of X is $s'_p(t) = \sum_{i=1}^{n} (i+1) a_i t^i$. Writing $d_p(X) = \sum_{i=0}^{p-1} c_i X^i \in \mathbb{Z}_p[X]$, we see that

$$s'_p(t) = s'_p(t) - s_p(t) = \sum_{i=0}^{n} a_i t^i = d_p(t), \quad t \in \mathbb{Z}_p.$$

Since $p \nmid c_i$ for some i, it follows that $d_p(t) \neq 0$ for some $t \in \mathbb{Z}_p$. As before we get a polynomial $g(X) := (X+k) f(X+k) \in \mathbb{Z}_{p^k}$ such that $g(R) = 0$ and such that the coefficient of X in g is a unit mod p^k, and it again follows that R is commutative. □

Since

$$C(\widetilde{\mathcal{R}}) = \bigcap_{p \text{ prime}} C(\widetilde{\mathcal{R}}_p),$$

it follows that the polynomials in $C(\widetilde{\mathcal{R}})$ are precisely those that satisfy the S_p condition for all p. Now the S_p condition is clearly satisfied whenever $p \nmid a_1$, which helps to cut down on the number of conditions that need to be verified. However, it at first appears that we might still need to check an infinite number of conditions if $a_1 = 0$ before being able to conclude that $f \in C(\widetilde{\mathcal{R}})$.

Fortunately, for any given polynomial it is easy to discard all but a finite number of the S_p conditions—regardless of the value of a_1—and to deduce Theorem 1.

Proof of Theorem 1. If f is not primitive, then all coefficients a_i are divisible by some prime p, and certainly f does not satisfy the S_p condition. Thus by Theorem 2, $f \notin C(\widetilde{\mathcal{R}})$.

For the converse direction, we may assume that $p \mid a_0$ and $p \mid a_1$, since otherwise S_p trivially holds. It suffices to show that the S_p condition holds for all primes $p > n/2$ as long f is primitive. If $p > n+1$, then this is easy: all the sums in the S_p condition involve at most one term so, since the gcd of the coefficients is 1, there exists i such that $p \nmid a_i = c_i$.

When $(n+2)/2 < p \leq n+1$, all sums in the S_p condition involve at most two terms. Suppose $p \nmid a_i$ for some i. If $c_i = a_i$, then we are done. Otherwise $c_i = a_i + a_j$, where $j \geq 0$ satisfies $|j-i| = p-1$. But then $jc_i - b_i = \pm(p-1)a_i$ is not divisible by p, and so either c_i or b_i is not divisible by p. Finally if $n/2 < p \leq (n+2)/2$, then the analysis is similar except that b_j and c_j may involve three terms when $j = 0$ or $j = 1$. But in both of these sums, we can discard one of the terms since $p \mid a_0$ and $p \mid a_1$. □

Since there are no primes $p \leq 3/2$, the characterization for polynomials of degree at most 3 is particularly simple.

Corollary 1. *Suppose $f \in \mathbb{Z}[X]$ has degree at most 3. Then $f \in C(\widetilde{\mathcal{R}})$ if and only if f is primitive.*

Note that primitivity alone is not sufficient for polynomials of degree larger than 3, since $f(X) = X^4 + X^2$ fails the S_2 condition, and so $f(R_2) = 0$, where R_2 is the noncommutative monoid ring in the proof of Theorem 2.

According to [3], a polynomial f (with $a_0 = 0$) lies in $C(\mathcal{R})$ if and only if it is either a Herstein polynomial or $a_1 = \pm 2, a_2$ is odd, and the sum of the coefficients is odd. Comparing this with Corollary 1 or Theorem 1, it is easy to give examples of polynomials in $C(\widetilde{\mathcal{R}}) \setminus C(\mathcal{R})$, for instance $3X + X^2, 2X + X^2 + X^3$, or $2X^2 + X^3$.

In view of the characterization of $C(\mathcal{R})$, it is perhaps worth noting for comparison purposes the characterization of polynomials with $a_1 = \pm 2$ that lie in $C(\widetilde{\mathcal{R}})$.

Corollary 2. *Suppose $f(X) = \pm 2X + \sum_{i=2}^{n} a_i X^i \in \mathbb{Z}[X]$. Then $f \in C(\widetilde{\mathcal{R}})$ if and only if either $\sum_{0<i<n/2} a_{2i+1}$ or $\sum_{0<i\leq n/2} a_{2i}$ is odd.*

Proof. By Theorem 1, $f \in C(\widetilde{\mathcal{R}})$ if and only if either b_0 or c_0 is odd, where b_0, c_0 are the sums in the S_2 condition. But, mod $2, b_0$ is the sum of the coefficients over all odd indices, and $b_0 + c_0$ is the sum over all even indices. □

Lastly we note that the examples that prove necessity of the conditions in Theorems 1 and 2 involve only finite rings of prime characteristic.

Thus if $\mathcal{F}$ is the set of all finite rings with unity, then $C(\mathcal{F}) = C(\widetilde{\mathcal{R}})$, while if $\mathcal{F}$ consists of finite rings with unity of characteristic p, then $C(\mathcal{F}) = C(\widetilde{\mathcal{R}}_p)$. This is parallel to the fact that if $\mathcal{F}$ consists of finite rings (without the assumption of unity), then $C(\mathcal{F}) = C(\mathcal{R})$ (since the proof in [3] uses only finite rings to prove necessity).

References

[1] Herstein, I.N.: The structure of a certain class of rings. Am. J. Math. **75**, 864–871 (1953)

[2] Jacobson, N.: Structure theory for algebraic algebras of bounded degree. Ann. Math. **46**, 695–707 (1945)

[3] Laffey, T.J., MacHale, D.: Polynomials that force a ring to be commutative. Proc. R. Ir. Acad. Sect. A **92**, 277–280 (1992)

[4] Pinter-Lucke, J.: Commutativity conditions for rings: 1950–2005. Expo. Math. **25**, 165–174 (2007)

[5] Tan, K.T.: On a commutativity theorem of Jacobson. Tamkang J. Math. **4**, 53–55 (1973)

S. M. Buckley
Department of Mathematics and Statistics
National University of Ireland Maynooth
Maynooth, Co., Kildare, Ireland
e-mail: stephen.buckley@maths.nuim.ie

D. MacHale
School of Mathematical Sciences
University College Cork
Cork, Ireland
e-mail: d.machale@ucc.ie

Received: November 9, 2011.
Accepted: October 18, 2012.