FUNCTIONAL IDENTITIES: A SURVEY

MATEJ BREŠAR

Abstract. The paper surveys the results and applications of the theory of functional identities and generalized functional identities in rings.

1. Introduction

Our aim in writing the paper is to present a new theory, the theory of (generalized) functional identities, and its applications, to a wider audience. Therefore, we shall avoid stating the results in their most general forms; also, the main results shall not be proved, we will only try to illuminate some ideas of the proofs.

The paper is primarily addressed to algebraists whose research is connected with maps of rings (algebras) having some additional properties (e.g. Lie and Jordan maps, derivations and automorphisms, linear preservers etc.). Also, the paper is addressed to ring theorists dealing with polynomial identities and their generalizations, especially generalized polynomial identities. Functional identities have turned out to be applicable to certain problems in some other mathematical areas (in particular, in operator theory and functional analysis), and perhaps one might find some further connections elsewhere (we remark that at least at the level of basic definitions the theory of functional identities admits some parallels with that of algebraic functions). Therefore, some parts of the paper may be of some interest not only to ring theorists.

A functional identity (FI) on a ring $R$ is, roughly speaking, an identity holding for all elements in $R$ (or more generally, all elements from a certain subset of $R$) which involves maps on $R$. If the identity besides maps also includes some fixed elements in the ring we will speak about a generalized functional identity (GFI). The usual goal in the study of (G)FI’s is to find the form of the maps involved, or, when this is not possible, to determine the structure of the ring.

Over the last few decades a lot of work has been done on identities satisfied by derivations, automorphisms and some other special additive maps (see [BeMM2] for a vast literature). But this is not what we are going to discuss here. When dealing with (G)FI’s we consider either completely arbitrary maps or sometimes we assume that they are (multi–)additive; however, we do not assume in advance how the maps act on the product of elements.

Introducing the concept of (G)FI’s in such a loose manner, it might be helpful for the reader to give some concrete examples. In Section 2 we will consider in detail several simple (G)FI’s which all are just very particular cases of the identities appearing in the main results. We hope that these simple examples shall illustrate the general theory. Just glancing through them one can see what the investigation

1991 Mathematics Subject Classification. Primary 16R50. Secondary 16N60, 16W10, 17B40.

Partially supported by a grant from the Ministry of Science of Slovenia.
of (G)FI's looks like. One has to find all the "trivial" or "obvious" solutions of (G)FI's, that is, the solutions which do not depend on some structural properties of the ring but are merely consequences of a formal calculation. We call them \textit{standard solutions}. The eventual existence of a nonstandard solution implies that the ring has a very special structure. A reader familiar with the theories of rings with polynomial identities (PI's) and generalized polynomial identities (GPI's) shall note that these examples, especially 2.2 and 2.4, indicate that the concept of FI's is a generalization of the concept of PI's, and that the concept of GFI's is a generalization of the concept of GPI's. In the later sections this will become more apparent. As a matter of fact, results on FI's give definitive conclusions in non–PI rings (or better, in rings that do not satisfy a PI of some low degree). Therefore, it is perhaps more accurate to say, especially from the point of view of possible applications, that the theory of FI's is a sort of a complement to the theory of PI's (rather than its extension), and similarly, we can regard the theory of GFI's as a complement to that of GPI's.

A connection with (G)PI's is certainly one reason for studying (G)FI's. Another, perhaps the most important reason is that the results on (G)FI's make it possible to solve different problems in ring theory (and elsewhere) for which other techniques do not seem to work that efficiently. In particular, long–standing Herstein’s conjectures \cite{H} concerning Lie maps of associative rings were settled using FI's.

We shall restrict our attention to prime rings, that is, rings in which the product of two nonzero ideals is always nonzero. Most of the results can (or could) be extended to some more general rings (such as semiprime), and moreover, we have realized (but not yet published) that basically all the theory can be done in some completely different classes of rings. There are also some results on FI's in some algebras appearing in functional analysis. Nevertheless, for simplicity we assume throughout the paper that \( R \) is a prime ring. One way of looking at the (G)FI theory is that it deals with solving the equations with maps as unknowns. As it often happens in mathematics, the solution of the equation does not necessarily lie in the original set. When considering (G)FI's, rings of quotients naturally get in our way. By \( C, RC, Q_s, Q_l \) and \( Q_{ml} \) we denote the extended centroid, the central closure, the symmetric Martindale ring of quotients, the left Martindale ring of quotients and the maximal left ring of quotients of \( R \), respectively. For explanations of these, as well as many other notions appearing in this paper, we refer the reader to the book \cite{BeMM2} of Beidar, Martindale and Mikhalev. On the other hand, for a superficial understanding of the paper it is probably enough to know that \( RC, Q_s, Q_l \) and \( Q_{ml} \) are certain rings containg \( R \) (more precisely, \( R \subseteq RC \subseteq Q_s \subseteq Q_l \subseteq Q_{ml} \)) and that \( C \) is a field containing the center of \( R \) and is the center of all \( Q_s, Q_l \) and \( Q_{ml} \) (as well as of \( RC \) provided that \( R \) has the identity).

In Sections 3 and 4 we will consider FI's and GFI's, respectively, just from the point of view that they are interesting in their own right. Applications will be discussed in Section 5.

The study of (G)FI's was initiated in the beginning of the 90’s by the author in a series of papers. We remark that the terminology in these papers does not completely coincide with the one which we are using now. After that, several mathematicians have given their contributions to the theory. However, we have
to point out the important role of Beidar in developing the theory of FI's, and a similar role of Chebotar concerning GFI’s.

2. Simple examples of (G)FI’s

In all the examples below, \( f_1, f_2 : R \to R \) are assumed to be just arbitrary maps.

Example 2.1. Consider the FI

\[ f_1(x)y + f_2(y)x = 0 \]

for all \( x, y \in R \). A trivial possibility when this holds is that both \( f_1 \) and \( f_2 \) are zero. Are there any nontrivial (nonstandard) possibilities? If \( R \) is commutative, then there certainly are (just take, for example, \( f_1(x) = -f_2(x) = x \)). In general, (1) implies

\[(f_1(x)yz)w = -(f_2(yz)xw) = (f_1(xw)y)z = -(f_2(y)xwz) = f_1(x)yzw \]

for all \( x, y, z, w \in R \). That is, \( f_1(R)R[R, R] = 0 \) and so, since \( R \) is prime, either it is commutative or \( f_1 = 0 \). Clearly, \( f_1 = 0 \) yields \( f_2 = 0 \). Thus, this FI has nonstandard solutions if and only if \( R \) is commutative.

Example 2.2. Now consider a slightly more general situation when the expression \( f_1(x)y + f_2(y)x \) is not necessarily zero but is always central, that is, we consider the FI

\[ [f_1(x)y + f_2(y)x, z] = 0 \]

for all \( x, y, z \in R \). First we give an example showing when this FI can occur in a nontrivial way. Let \( R = M_2(F) \) be the ring of 2 by 2 matrices over a field \( F \) and let \( f_1(x) = f_2(x) = x - \text{tr}(x)1 \) where \( \text{tr}(x) \) denotes the trace of \( x \) and 1 denotes the identity matrix. Then (2) holds indeed for all \( x, y, z \in R \). Thus, this FI has a nonstandard solution (i.e., \( f_1 \neq 0 \) or \( f_2 \neq 0 \)) not only in commutative rings but also in the ring \( M_2(F) \), and, clearly, in some of its subrings. Let us now show that that is all, that is, that there are no other prime rings admitting nonstandard solutions. Assume, therefore, that \( R \) is arbitrary and that \( f_1 \neq 0 \). Given \( x, y, z \in R \), it follows from (2) that \([f_1(xz)y, z] = -[f_2(y)xz, z] = -[f_2(y)x, z]z = [f_1(x)y, z]z \), and hence \( f_1(x)yz^2 = (z^2f_1(x) + f_1(xz))yz + zf_1(xz)y = 0 \). Recall a well-known result of Martindale stating that \( a_1y_1 + \ldots + a_ny_n = 0 \) for all \( y \in R \), where \( a_i, b_i \in RC \) and \( a_1 \neq 0 \), implies that the \( b_i \)'s are linearly dependent over \( C \) [M5, Theorem 2] (see also [BeMM2, Corollary 6.1.3]). Therefore, fixing \( x \in R \) such that \( f_1(x) \neq 0 \) it follows that the elements \( 1, z, z^2 \) are \( C \)-dependent for any \( z \in R \). In other words, every element in \( R \) is algebraic of degree at most 2 over \( C \). It is known that such a ring \( R \) is either commutative or embeds in \( M_2(F) \) for some field \( F \) (equivalently, \( R \) satisfies \( S_1 \), the standard polynomial identity of degree 4). This simple example somehow indicates that FI’s can have a strong effect on the structure of the ring.

Example 2.3. The following FI

\[ f_1(x)y + f_2(y)x = yf_1(x) + xf_2(y) \]

for all \( x, y \in R \) is somewhat more entangled. Which maps can satisfy this identity? Note that one natural possibility is that \( f_1 \) and \( f_2 \) are of the form

\[ f_1(x) = \lambda x + \mu(x), \quad f_2(x) = \lambda x + \nu(x), \]

where \( \lambda, \mu, \nu \in R \) and \( \lambda \neq 0 \).
where \( \lambda \) is a central element and \( \mu, \nu \) are maps with range in the center. Let us show that this is essentially the only possibility, that is, this FI has only the standard solution.

Define \( B : R \times R \to R \) by \( B(x, y) = [f_1(x), y] \). Clearly, \( B \) is an inner derivation in the second argument (i.e., \( B(x, yz) = B(x, y)z + yB(x, z) \) for all \( x, y, z \in R \)). On the other hand, by (3) \( B \) can be represented as \( B(x, y) = [x, f_2(y)] \) so that \( B \) is a derivation in the first argument as well. Now let us compute \( B(xu, yv) \) in two different ways. First using the fact that \( B \) is a derivation in the first argument we get

\[
B(xu, yv) = B(x, yv)u + xB(u, yv).
\]

Since \( B \) is also a derivation in the second argument this yields

\[
B(xu, yv) = B(x, yv)u + yB(u, yv)u + xB(u, yv).
\]

On the other hand, first using the derivation law in the second and after that in the first argument we get

\[
B(xu, yv) = B(xu, y)v + yB(xu, v) = B(x, y)v + yB(x, v)u + yxB(u, v).
\]

Comparing both relations we obtain \( B(x, y)[u, v] = [x, y]B(u, v) \) for all \( x, y, u, v \in R \). Replacing \( v \) by \( rv \) and using \( [u, rv] = [u, r]v + r[u, v] \), \( B(u, rv) = B(u, r)v + rB(u, v) \) we obtain

\[
B(x, y)r[u, v] = [x, y]rB(u, v)
\]

for all \( x, y, r, u, v \in R \). Assume that \( R \) is noncommutative. Then, by Martindale’s result mentioned above, \( B(u, v) \) and \([u, v] \) are always \( C \)-dependent. Picking \( u, v \) so that \([u, v] \neq 0 \), we thus have \( B(u, v) = \lambda[u, v] \) for some \( \lambda \in C \). But then the last identity implies that \( (B(x, y) - \lambda[x, y])R[u, v] = 0 \) for all \( x, y \in R \), which further yields \( B(x, y) = \lambda[x, y] \) for all \( x, y \in R \). According to the definition of \( B \) we thus have \([f_1(x) - \lambda x, y] = 0 \) and so \( f_1(x) - \lambda x \in C \), \( x \in R \). Similarly, \( f_2(x) - \lambda x \in C \), \( x \in R \). Therefore, \( f_1, f_2 \) are of the form \( f_1(x) = \lambda x + \mu(x) \), \( f_2(x) = \lambda x + \nu(x) \), where \( \lambda \in C \) and \( \mu, \nu : R \to C \). If \( R \) is commutative then \( f_1, f_2 \) trivially take that form (say, for \( \lambda = 0 \)), so that the conclusion holds for any prime ring \( R \).

**Example 2.4.** Now we extend the FI treated in Example 2.1 in another way. Let \( a \) be a fixed nonzero element in \( R \) and consider the identity

\[
f_1(x)ya + f_2(y)xa = 0
\]

for all \( x, y \in R \). This is a simple example of a GFI. If \( a = 1 \) or, more generally, \( a \) is an invertible element, then, by what we proved in Example 2.1, \( f_1 \) and \( f_2 \) must both be zero unless \( R \) is commutative. If \( a \) is not invertible, or better, if \( a \) is "far" from being invertible, this conclusion is no longer true. Indeed, suppose that \( R \) contains an element \( a \neq 0 \) such that \( aRa \subseteq Ca \), that is, for any \( x \in R \) there is \( \lambda_x \in C \) such that \( axa = \lambda_xa \). Then we have \( axaya = \lambda_xaya = ay(\lambda_xa) = ayaxa \). Therefore, nonzero maps \( f_1(x) = -f_2(x) = axa \) satisfy (4). Conversely, let us show that (4) with \( f_2 \neq 0 \) implies that \( aRa \subseteq Ca \). Given \( x, y, z \in R \) we have, on the one hand, \( f_1(x)ya za = -f_2(yaza)xa \), and on the other hand, \( f_1(x)ya za = -f_2(y)xa za \). Hence \( f_2(yaza)xa = f_2(y)xa za \) for all \( x, y, z \in R \). Fixing any \( y \in R \) such that \( f_2(y) \neq 0 \) and again using Martindale’s result mentioned above it follows that \( axa \) and \( a \) are \( C \)-dependent for any \( z \in R \), and so our claim is proved. Thus, a necessary and sufficient condition for the GFI of the type (4) to have a nonstandard solution...
is that $R$ contains a nonzero element $a$ such that $aRa \subseteq Ca$. The structure of such rings $R$ can be precisely described. These are the so–called GPI rings whose associated division ring is a field, and can be, roughly speaking, considered as rings of linear operators containing operators of finite rank (cf. [BeMM2, Sections 4.3, 6.1]).

**Example 2.5.** In our last example we will show, in particular, that rings of quotients appear naturally in the study of (G)FI’s. Let $a, b \in R$ be nonzero elements and consider the GFI

$$f_1(x)ya = bx f_2(y)$$

for all $x, y \in R$. Let us show that $f_1$ and $f_2$ must be of the form

$$f_1(x) = bxq, \quad f_2(y) = qya$$

for some $q \in Q_s$. That is, we will show that in any case, regardless of some features of $R$, this GFI has only the standard solution.

By (5) we have $by f_1(z)x a = byb z f_2(x) = f_1(yb z)x a$ and so $(f_1(yb z)−by f_1(z))xa = 0$ for all $x, y, z \in R$. The primeness of $R$ yields $f_1(yb z) = by f_1(z)$ for all $y, z \in R$. Next, (5) shows that the map $x \mapsto f_1(x)ya$ is additive for any fixed $y$, that is, $(f_1(x_1 + x_2) − f_1(x_1) − f_1(x_2))ya = 0$ which implies that $f_1$ is additive.

Introduce the ideal $I = RbR$ and define $\phi : I \rightarrow R$ by

$$\phi(\sum y_i b z_i) = \sum y_i f_1(z_i).$$

In order to show that $\phi$ is well–defined, assume that $\sum y_i b z_i = 0$ for some $y_i, z_i \in R$. Then $\sum r y_i b z_i = 0$ for every $r \in R$, and therefore, using the properties of $f_1$ observed above, we get

$$0 = f_1(\sum r y_i b z_i) = \sum f_1((r y_i) b z_i) = br(\sum y_i f_1(z_i)).$$

But then $\sum y_i f_1(z_i) = 0$, as desired.

Obviously, $\phi$ is a homomorphism of left $R$–modules and so there is $q \in Q_s$ such that $\phi(x) = xq$, $x \in I$ [BeMM2, Proposition 2.2.1]. Therefore, $ybxq = \phi(ybx) = y f_1(x)$, i.e., $y f_1(x) − bx q) = 0$ for all $x, y \in R$ which yields $f_1(x) = bx q$. But then $bx f_2(y) − qya) = f_1(x)ya − bxq y a = 0$, $x, y \in R$, which gives $f_2(y) = qya$. Finally, $qRaR = f_2(R)R \subseteq R^2 \subseteq R$ and so $q \in Q_s$.

3. Functional identities

Before stating the main result of this section, Theorem 3.4, we will consider a few other results that have initiated the FI theory. Though less general, they are still of some interest and also easier to understand. Let us first say a few words about the background of first results on FI’s.

As already mentioned, numerous publications have been devoted to derivations and automorphisms satisfying some identities of polynomial type. Quite often the results were proved using elementary, computational methods. A more systematic and uniform approach was found by Kharchenko (and extended to anti–automorphisms by Chuang), see [BeMM2, K]. Although (G)FI’s involve more general maps, some results on derivations and automorphisms have been our original motivation for treating the first FI’s. More than 40 years ago Posner [P] proved that if a derivation $d$ of $R$ satisfies $[d(x), x] \in Z$ for every $x \in R$, where $Z$ is the center of $R$, then either $d = 0$ or $R$ is commutative (recall that $R$ is assumed to be a
prime ring throughout). A number of authors have extended this theorem in many ways (see e.g. [BeFW, BM1, BM2, La1, La2, La3, LL1, Ma, Mi5, V] where a number of further references can be found). Although somewhat out of the context, we mention as a curiosity that in order to get noncommutative extensions of the classical Singer–Wermer theorem [SW] on bounded derivations of commutative Banach algebras (and Thomas’ generalization to the unbounded case [T]), some analogues of Posner’s theorem for derivations on Banach algebras have been obtained (see e.g. [Br7, BrV, MM, MR, R]). A typical ring-theoretic generalization of Posner’s theorem for derivations on Banach algebras have been obtained (see e.g. [AM, BeFW, Br1, Br2, Br4, Br6, Br8, Br10, Br11, BrH1, BrH2, BrMM1, BrM2, BrM3, BrSV, L, LL2, Le, LeL]). Let us state, for instance, the result from [Br8].

Theorem 3.1. [Br3] If an additive map \( f : R \to R \) satisfies \( [f(x), x] = 0 \) for all \( x \in R \), then there exist \( \lambda \in C \) and an additive map \( \mu : R \to C \) such that

\[
\pi(x) = f_1(x)y + f_2(y)x + x f_3(y) + y f_4(x).
\]

(a) If \( \pi(x) \) lies in the center of \( R \) for all \( x, y \in I \) and \( \text{char}(R) \neq 2,3 \), then either \( R \) satisfies \( S_4 \) or \( \pi(x, y) = 0 \) for all \( x, y \in I \);

(b) If \( \pi(x, y) = 0 \) for all \( x, y \in I \), then there exist \( p, q \in Q_2 \) and additive maps \( \lambda, \mu : R \to C \) such that

\[
f_1(x) = xp + \mu(x), \quad f_2(x) = xq + \lambda(x),
\]

This fact was first observed in [BrMM1] (see also a generalization in [Br6]) and is of independent interest since biderivations appear in the very definition of Poisson algebras [FL].

As we saw, the proof of Theorem 3.1 is neither long nor difficult, so it comes as no surprise that it is possible to extend it in many different ways. In particular, this result served as an original inspiration for numerous results on additive maps satisfying some special FI’s on (semi)prime rings and also on some Banach algebras [AM, BeFLW, Br1, Br2, Br4, Br6, Br8, Br10, Br11, BrH1, BrH2, BrMM1, BrM2, BrM3, BrSV, L, LL2, Le, LeL]. Let us state, for instance, the result from [Br8].

Theorem 3.2. [Br8] Let \( I \) be an ideal of \( R \), let \( f_1, f_2, f_3, f_4 : I \to R \) be additive maps and set

\[
\pi(x, y) = f_1(x)y + f_2(y)x + x f_3(y) + y f_4(x).
\]

(a) If \( \pi(x, y) \) lies in the center of \( R \) for all \( x, y \in I \) and \( \text{char}(R) \neq 2,3 \), then either \( R \) satisfies \( S_4 \) or \( \pi(x, y) = 0 \) for all \( x, y \in I \);

(b) If \( \pi(x, y) = 0 \) for all \( x, y \in I \), then there exist \( p, q \in Q_2 \) and additive maps \( \lambda, \mu : R \to C \) such that

\[
f_1(x) = xp + \mu(x), \quad f_2(x) = xq + \lambda(x),
\]
The next step in the study of FI’s has been the treatment of maps of several variables. The first result that was obtained is an analogue of Theorem 3.1 for traces of biadditive maps (i.e., maps \( q \) of the form \( q(x) = B(x, x) \) where \( B \) is a biadditive map).

**Theorem 3.3.** [Br5] Let \( q : R \to R \) be the trace of a biadditive map such that \( [q(x), x] = 0 \) for all \( x \in R \). If \( \text{char}(R) \neq 2 \) and \( R \) does not satisfy \( S_4 \), then there exist \( \lambda \in C \) and maps \( \mu, \nu : R \to C \) such that \( q(x) = \lambda x^2 + \mu(x)x + \nu(x) \) for all \( x \in R \).

Theorem 3.3 is important because of its applications (see Section 5). Some its extensions can be found in [BaM, BeMM1, BrM1, LL2, LLWW].

In the papers mentioned so far there has been a sort of a lack of systematic approach to numerous different but related problems. From this point of view Beidar’s work [Be] can be considered as a break-through in the study of FI’s. In this paper Beidar proved a generalization of Theorem 3.2 for maps of several variables. We will now state a simplified version of this result which, on the other hand, already includes an improvement (namely, the maps are not assumed to be multi–additive) that was discovered somewhat later by Beidar and Martindale [BeM].

First, however, we have to introduce some notation. Let \( r \geq 2 \) be an integer, and let \( F \) be a map defined on \( R^{r-1} \). Given \( 1 \leq i \leq r \) we define a map \( F^i \) on \( R^r \) by

\[
F^i(x_1, \ldots, x_r) = F(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_r).
\]

Also, given a map \( p \) defined on \( R^{r-2} \) (in the case \( r = 2 \) this should be understood as that \( p \) is a constant) and \( 1 \leq i < j \leq r \), we define

\[
p^{ij}(x_1, \ldots, x_r) = p^{ij}(x_1, \ldots, x_r) = p(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{j-1}, x_{j+1}, \ldots, x_r).
\]

**Theorem 3.4.** [Be, BeM] Let \( E_i, F_i : R^{r-1} \to Q_{ml}, i = 1, \ldots, r \) be any maps and set

\[
\pi(x_1, \ldots, x_r) = \sum_{i=1}^{r} E_i^i(x_1, \ldots, x_r)x_i + \sum_{j=1}^{r} x_j F_j^i(x_1, \ldots, x_r).
\]

(a) If \( \pi(x_1, \ldots, x_r) \in C \) for all \( x_1, \ldots, x_r \in R \), then either

(i) every element in \( R \) is algebraic over \( C \) of degree \( \leq r \); or

(ii) \( \pi(x_1, \ldots, x_r) = 0 \) for all \( x_1, \ldots, x_r \in R \).

(b) If \( \pi(x_1, \ldots, x_r) = 0 \) for all \( x_1, \ldots, x_r \in R \), then either

(iii) every element in \( R \) is algebraic over \( C \) of degree \( \leq r - 1 \); or

(iv) there exist unique maps \( p_{ij} : R^{r-2} \to Q_{ml} \) and \( \lambda_i : R \to C \) such that

\[
E_i^i(x_1, \ldots, x_r) = \sum_{1 \leq j \leq r} x_j p_{ij}^i(x_1, \ldots, x_r) + \lambda_i(x_1, \ldots, x_r),
\]

\[
F_j^i(x_1, \ldots, x_r) = \sum_{1 \leq j \leq r} -p_{ij}^i(x_1, \ldots, x_r)x_i - \lambda_j(x_1, \ldots, x_r).
\]
Moreover, if all the maps $E_i$ and $F_j$ are additive in each argument, then the same is true for the maps $p_{ij}$ and $λ_i$.

Assuming that all the maps $E_i$ and $F_j$ are "monomials" $λ_i x_{i_1} \ldots x_{i_r}$, $λ_i \in C$, the FI treated in Theorem 3.4 turns into an ordinary polynomial identity. Thus we see indeed that the concept of FIs can be regarded as a generalization of the concept of PI’s. A rough summary of Theorem 3.4 is that either the FI under consideration have only the standard solutions (i.e. the ones given in (iv)) or the ring satisfies some PI of "small" degree. Namely, it can be easily deduced from the standard PI theory that the condition that every element in $R$ is algebraic over $C$ of degree $≤ r$ is equivalent to the condition that $R$ satisfies a (standard) polynomial identity of degree $≤ 2r$ (which is further equivalent to the condition that $R$ can be embedded into the matrix ring $M_r(F)$ for some field $F$).

In the case when the expression $π(x_1, \ldots, x_r)$ in Theorem 3.4 consists of one sum only, say, $π(x_1, \ldots, x_r) = \sum_{i=1}^{r} E_i(x_1, \ldots, x_r)x_i$, the standard solution given in (iv) reduces to $E_i = 0$. This observation enables a brief explanation of why the conclusion of Theorem 3.4 is the best possible. We shall extend the counterexample given in Example 2.2 from $r = 2$ to an arbitrary $r$. Consider the ring $M_r(F)$ with $F$ a field. By the Cayley–Hamilton theorem every matrix $x ∈ M_r(F)$ satisfies $x^r - \text{tr}(x)x^{r-1} + \ldots + (-1)^{r-1}\text{det}(x)1 = 0$ (in particular, each $x$ is algebraic of degree $≤ r$). A complete linearization gives the relation of the form $\sum_{i=1}^{r} E_i(x_1, \ldots, x_r)x_i ∈ F1$ with the $E_i$’s being nonzero. That is, we arrived at a functional identity with a nonstandard solution.

Let us also mention that Beidar and Martindale [BeM] (see also a continuation [BeBrCM1]) proved a noteworthy generalization of Theorem 3.4 for rings with involution $\ast$. They treated FIs involving sums such as

$$\sum_{i=1}^{r} E_i(x_1, \ldots, x_r)x_i + \sum_{j=1}^{r} x_j F_j(x_1, \ldots, x_r)$$

$$+ \sum_{k=1}^{r} G_k(x_1, \ldots, x_r)x_k^\ast + \sum_{l=1}^{r} x_l^\ast H_l(x_1, \ldots, x_r).$$

What else can be done in the theory of FIs? Of course, one can treat FIs similar to those in Theorem 3.4 on different kinds of rings and on some special subsets of rings. A much more unpredictable project seems to be the treatment of FIs involving summands such as

$$E(x_1, \ldots, x_{i-1})x_i F(x_{i+1}, \ldots, x_r).$$

It is our impression that obtaining a very general result on FIs of such type is a very difficult problem. In [BrC] a sort of a testing case when nonzero additive maps $f, g : R → R$ satisfy $f(x)yg(x) = 0$, $x ∈ R$ was treated. It turns out that this condition is equivalent to the condition discussed in Example 2.4, i.e., that the ring is GPI and its associated division ring is a field. Another difficult problem is to treat FIs in which the maps are not separated by a variable (for instance, if the expressions such as $f(x)yg(y)$ appear in the identity). Of course, it is senseless to consider just every identity of that type (for instance, it is impossible to get any reasonable conclusion from the simplest identities such as $f(x)yg(y) = 0$, $[f(x), g(y)] = 0$ etc.). Nevertheless, some indeed quite special examples [BrH1, BrM2] suggest that there is some hope here after all.
4. Generalized functional identities

The study of GFI’s was initiated in [Br9] where the following identity was treated:

\[ \sum_{i=1}^{n} F_i(x)xa_i + \sum_{i=1}^{m} G_i(x)yb_i + \sum_{i=1}^{k} c_i yH_i(x) + \sum_{i=1}^{l} d_i xK_i(y) = 0 \]

for all \( x, y \in R \). Here, \( F_i, G_i, H_i, K_i : R \to RC \) are additive maps and \( \{a_1, \ldots, a_n\}, \{b_1, \ldots, b_m\}, \{c_1, \ldots, c_k\}, \{d_1, \ldots, d_l\} \) are \( C \)-independent subsets of \( R \). The conclusion is that either the ring is GPI or the form of the maps can be precisely described; for example, the \( F_i \)'s are of the form \( F_i(y) = \sum_{j=1}^{k} c_{ij}yp_{ji} + \sum_{j=1}^{l} \lambda_{ji}(y)d_j \), where \( p_{ji} \) are elements in the symmetric Martindale ring of quotients of \( RC \) and \( \lambda_{ji} : R \to C \) are additive maps.

A very special case of such an identity was treated in Example 2.4, and even from this the very simplest case it is clear that GFI’s do not necessarily have only standard solutions in GPI rings.

The proof in [Br9] is quite long and complicated. Nevertheless, based on some ideas of this proof a young Russian mathematician Chebotar found a better approach and generalized the result of [Br9] to maps of several variables [C2], and thereby obtained an analogue of Theorem 3.4 for GFI’s (actually, Chebotar obtained this result before Theorem 3.4 was known). We will now state a somewhat improved version of Chebotar’s result which follows at once from the main theorem of [BeBrC1].

**Theorem 4.1.** [C2, BeBrC1] Let \( r \geq 2 \) be an integer and \( n_i, m_i, i = 1, \ldots, r \) be nonnegative integers. Further, let \( V \) be a finite dimensional subspace of the vector space \( Q_{ml} \) over \( C \), and let \( E_{ji}, F_{ik} : R^{r-1} \to Q_{ml} \) be maps such that

\[
\sum_{j=1}^{r} \sum_{i=1}^{n_j} E_{ji}(x_1, \ldots, x_r)x_ja_i^j + \sum_{i=1}^{m_i} b_{ik}x_iF_{ik}(x_1, \ldots, x_r) \in V
\]

for all \( x_1, \ldots, x_r \in R \), where \( \{a_{i_1}^1, \ldots, a_{i_r}^r\} \) and \( \{b_{1_i}^1, \ldots, b_{m_i}^i\} \), \( j = 1, \ldots, r \), are \( C \)-independent subsets of \( Q_{ml} \). Then either

(i) \( R \) is a GPI ring;

or

(ii) there exist unique maps \( p_{ji}^1 : R^{r-2} \to Q_{ml} \) and \( \lambda_{ik} : R^{r-1} \to C \) such that

\[
E_{ji}^1(x_1, \ldots, x_r) = \sum_{i<j} \sum_{k=1}^{m_i} b_{ik}x_ip_{ji}^1(x_1, \ldots, x_r) + \sum_{k=1}^{m_j} \lambda_{ij}^k(x_1, \ldots, x_r)b_{ik},
\]

\[
F_{ik}^1(x_1, \ldots, x_r) = -\sum_{i<j} \sum_{s=1}^{n_j} p_{ji}^1(x_1, \ldots, x_r)x_ja_s^j - \sum_{i=1}^{n_i} \lambda_{ik}^j(x_1, \ldots, x_r)a_i^j.
\]

In particular,

\[
\sum_{j=1}^{r} \sum_{i=1}^{n_j} E_{ji}^1(x_1, \ldots, x_r)x_ja_i^j + \sum_{i=1}^{m_i} b_{ik}x_iF_{ik}^1(x_1, \ldots, x_r) = 0.
\]

Moreover, if all the maps \( E_{ji} \) and \( F_{ik} \) are additive in each argument, then the same is true for the maps \( p_{ji}^1 \) and \( \lambda_{ik} \).
From this result we see that GFI’s can be viewed as a generalization of GPI’s. Indeed, assuming that all the maps $E_{ji}$ and $F_{lk}$ are "generalized monomials" $c_{j_0}x_{i_1}c_{j_1}x_{i_2}...x_{i_r}c_{j_r}$, the GFI treated in Theorem 4.1 turns into an GPI.

Theorem 4.1 tells us that either the GFI under consideration has only the standard solution or the ring $R$ is GPI and so, by a well-known theorem of Martindale [M5] (see also [BeMM2, Section 6.1]), $RC$ is a primitive ring with nonzero socle and $eRCe$ is a finite dimensional division algebra over $C$ for each primitive idempotent $e$ in $RC$. Thus, either the ring is "nice" or we have a complete "control" of the maps.

We also refer to the papers [BeBrC1, BeBrC2, C4] for various extensions of Theorem 4.1. Let us remark that [BeBrC1, BeBrC2] treat GFI’s that also involve derivations, automorphisms and anti-automorphisms, and the main results simultaneously extend Theorem 4.1 and the theory of Kharchenko and Chuang mentioned above. Finally we mention that we were informed by Chuang that he obtained some generalizations of the result of [Br9] in a somewhat different direction than discussed here.

5. Applications

We begin by a problem of determining the structure of bijective linear maps preserving commutativity, that is, maps $\theta : A \to B$ of one algebra onto another one with the property that $\theta(x)$ and $\theta(y)$ commute whenever $x$ and $y$ commute. The obvious examples are maps of the form $\theta(x) = \lambda \phi(x) + \mu(x)$ where $\lambda$ is a central element in $B$, $\mu$ is a map of $A$ into the center of $B$, and $\phi$ is either an isomorphism or an anti-isomorphism. The goal is to find reasonable conditions under which these obvious examples are basically the only possible examples. This project was started in linear algebra [W] where a result of this kind was proved for the case when $A = B = M_n(F)$ with $F$ a field and $n \geq 4$. Also, a counterexample was constructed for $n = 2$ (for $n = 3$ the result remains valid as shown later [B, PW]). Several authors have extended this result to more general algebras, in particular, to various algebras of bounded linear operators on infinite dimensional spaces [CJR, Mi6, O]. All these algebras, however, were prime and centrally closed over their centers (we say that an algebra $A$ over a field $F$ is centrally closed over $F$ if both the center and the extended centroid of $A$ are equal to $F$).

Using FI’s instead of linear algebra and operator theory techniques a fairly more general ring-theoretic analogue of these results can be proved. Let us show just the main idea of this approach. If $\theta$ is a bijective linear map preserving commutativity, then it satisfies $[\theta(x), \theta(x^2)] = 0,$ for $x \in A$ (for $x$ certainly commutes with $x^2$). That is, $[y, \theta(\theta^{-1}(y)^2)] = 0$ for all $y \in B.$ Therefore we have arrived at the FI treated in Theorem 3.3. Applying this theorem one can prove the following result.

**Theorem 5.1.** [Br5] Let $A$ and $B$ be centrally closed prime algebras over a field $F$ with char($F$) $\neq 2, 3$, and let $\theta : A \to B$ be a bijective linear map satisfying $[\theta(x), \theta(x^2)] = 0$ for all $x \in R.$ If neither $A$ nor $B$ satisfies $S_4$, then

$$\theta(x) = \lambda \phi(x) + \mu(x)$$

for all $x \in A$, where $\lambda \in F$, $\lambda \neq 0$, $\mu$ is a linear map of $A$ into $F$ and $\phi$ is either an isomorphism or an anti-isomorphism of $A$ onto $B$.

Using adequate extensions of Theorem 3.3 similar results have been obtained for von Neumann algebras [BrM1] and semiprime rings [BaM]. We also mention
papers [BrS1, BrS2] where further applications of Theorem 3.3 to problems arising in operator theory can be found.

Lie isomorphisms of rings, that is, bijective additive maps \( \theta : S \to R \) satisfying \( \theta([x,y]) = [\theta(x), \theta(y)] \) for all \( x, y \in S \), of course, also preserve commutativity, and so the method outlined above also applies to these maps. Using it one can prove

**Theorem 5.2.** [Br5] Let \( S \) and \( R \) be prime rings with \( \text{char}(R) \neq 2 \), and let \( \theta : S \to R \) be a Lie isomorphism. If neither \( S \) nor \( R \) satisfies \( S_4 \), then \( \theta \) is of the form \( \theta = \phi + \tau \), where \( \phi \) is either an isomorphism or negative of an anti–isomorphism of \( S \) into \( RC \), and \( \tau \) is an additive map of \( S \) into \( C \) sending commutators to 0.

While the exclusion of algebras satisfying \( S_4 \) is necessary in Theorem 5.1, this is not the case with Lie isomorphisms. Actually, we were informed that P. Blau, a student of Martindale, removed this condition in Theorem 5.2.

Theorem 5.2 solves the problem posed by Herstein in his 1961 AMS Hour Talk [H]. Actually, Herstein raised questions about Lie isomorphisms of various Lie subrings of associative rings. Theorem 5.2 just answers the easiest question; nevertheless, it suggests the way how to approach the problem in a more complicated setting. Let us now discuss in somewhat greater detail Herstein’s questions.

Given any associative ring \( R \), it becomes a Lie ring when introducing a new product, the Lie product, by \( [x, y] = xy -yx \). Theorem 5.2 thus describes isomorphisms of \( S \) onto \( R \) considered as Lie rings. By a Lie subring of an associative ring \( R \) we mean an additive subgroup of \( R \) closed under the Lie product. Some Lie subrings appear naturally and are of special interest. One example is \( [R, R] \), the additive subgroup of \( R \) generated by all commutators in \( R \), or more generally, any Lie ideal of \( R \). In case \( R \) has an involution, the set of its skew elements \( K \) (i.e. elements \( x \) such that \( x^* = -x \)) as well as its derived Lie ring \( [K, K] \) are Lie subrings of \( R \). The problem that Herstein raised [H, pp. 528–529] is to characterize Lie isomorphisms (and Lie derivations) of \( R \), \( [R, R] \), \( [R, R]/Z \cap [R, R] \), \( K \), \( [K, K] \) and \( [K, K]/Z \cap [K, K] \) for the case when \( R \) is a simple ring. Here \( Z \) is the center of \( R \). The cases of \( [R, R]/Z \cap [R, R] \) and \( [K, K]/Z \cap [K, K] \) seem to be of special interest since these two Lie rings are, except in some very special situations, simple [H]. The solutions in the classical case when \( R = M_n(F) \), \( F \) a field, have been well–known for a long time (see e.g. [J, Chapter 10]). In 1951 Hua proved Theorem 5.2 for the case when \( S = R = M_n(\Delta) \), \( n \geq 3 \), \( \Delta \) a division ring [Hu]. Later on Martindale considered Herstein’s problems in a series of papers. In particular, he extended the treatment of the problems from simple to prime rings. It is interesting to note that the need to enlarge prime rings when studying Lie isomorphisms was Martindale’s original motivation for introducing the now classical concepts of the extended centroid and the central closure, cf. [M4, M5]. Martindale and some of his students have obtained solutions of Herstein’s problems [Ho, M1, M2, M3, M4, M6, M7, M8, Ro] provided that the rings contain some nontrivial idempotents. For instance, neglecting some differences in technical assumptions, Martindale [M4] proved Theorem 5.2 under the additional assumption that \( S \) contain an idempotent different from 0 and 1. Lie map problems have also been considered in certain operator algebras [A1, A2, AA, ARU, Ha, Mi1, Mi2, Mi3, Mi4] and the techniques there also rest heavily on the presence of idempotents.

The great advantage of the results on FI’s is that they do not depend on local properties of the ring. In particular, the (non)existence of idempotents is irrelevant.
in this setting. Using FI’s Herstein’s Lie map problems can be solved, modulo some low dimensional cases, in full generality.

How to get an FI when treating Lie isomorphisms of the skew elements? A square of a skew element is hardly ever skew again (so the same trick as above does not work), but the cube is. Therefore, every Lie isomorphism \( \theta \) of skew elements satisfies \( [\theta(k), \theta(k^3)] = 0 \) for every skew element \( k \). But this is an FI on \( K' = \theta(K) \). Based on this observation, Beidar, Martindale and Mikhalev [BeMM1] (see also [BeMM2, Chapter 9]) solved Herstein’s Lie isomorphism problem for skew elements in the case when the involution is of the first kind. We now state a slightly improved version of their theorem obtained by Chebotar [C3] who also found a shorter proof based on results from [BeM].

**Theorem 5.3.** [BeMM1, C3] Let \( R \) and \( R' \) be prime rings with involutions of the first kind and of characteristic \( \neq 2 \). Let \( K \) and \( K' \) denote respectively the skew elements of \( R \) and \( R' \). Assume that \( \dim_{\mathbb{C}}(RC) \neq 1, 4, 9, 16, 25, 64 \). Then any Lie isomorphism \( \theta \) of \( K \) onto \( K' \) can be extended uniquely to an associative isomorphism of \( \langle K \rangle \) onto \( \langle K' \rangle \), the associative subrings generated by \( K \) and \( K' \) respectively.

The reasons for excluding some low dimensional cases are explained in [M7].

The key part of the proofs of all applications of FI theory is to show that the initial conditions imply that a ring satisfies a certain FI. While it is quite obvious how to get appropriate FI’s when proving Theorems 5.1, 5.2 and 5.3, this is not the case when considering Lie isomorphisms of \( [R,R], [R,R]/Z \cap [R,R], [K,K] \) and \( [K,K]/Z \cap [K,K] \). In particular, these Lie rings are not closed under any powers, and so there is no such intimate relation between them and the associative structure of \( R \). Nevertheless, one can find tricks how FI’s can be produced in these cases too, and all Herstein’s Lie isomorphism conjectures were finally settled by Beidar and Chebotar [BeC4] and Beidar, Chebotar, Martindale and the author [BeBrCM2]. The proofs heavily rest on the very useful concept of \( d \)-free sets introduced by Beidar and Chebotar [BeC1, BeC2]. Roughly speaking, a subset \( S \) of a ring \( R \) is \( d \)-free if functional identities on \( S \), such as treated in Theorem 3.4, involving at most \( d \) variables have only standard solutions. Actually, the main results in [BeC4, BeBrCM2] give considerably more than just solutions of Herstein’s problems. They are stated in terms of \( d \)-free sets which, in particular, allows for a unified approach to a variety of Lie map problems involving different subsets of rings.

For other applications of (G)FI’s to Lie isomorphism and some related problems we refer to [BaM, BeBrCM3, BeC3, BeC6, BeC7, BV1, BV2, BV3, Bl, Br5, BrM1, C1, S, SB].

We close this paper with a beautiful result of Beidar and Chebotar [BeC5] on maps of a ring onto a prime ring that preserve any polynomial in noncommuting indeterminates. We need some further notation to state it. Let \( Z \) be a commutative ring with 1 and \( Z(X) \) be the free \( Z \)-algebra on a set \( X = \{x_1, x_2, \ldots \} \). A polynomial \( f(x_1, \ldots, x_r) \in Z(X) \) is said to be proper if at least one of its coefficients is equal to \( \pm 1 \). A simplified version of a result of Beidar and Chebotar (in particular, we just ignore the involvement Lie ideals in the statement) reads as

**Theorem 5.4.** [BeC5] Let \( S \) and \( R \) be \( Z \)-algebras with \( R \) prime and \( \text{char}(R) \neq 2 \). Further, let \( f(x_1, \ldots, x_r) \in Z(X) \), \( r \geq 2 \), be a proper multilinear polynomial and
\[ \theta : S \to R \text{ be a surjective additive map satisfying} \]
\[ \theta(f(x_1, \ldots, x_r)) = f(\theta(x_1), \ldots, \theta(x_r)) \]

for all \( x_1, \ldots, x_r \in S \). Then either every element in \( R \) is algebraic over \( C \) of degree \( \leq \max\{2r, 7\} \) or

\[ \theta(x) = \lambda \phi(x) + \mu(x) \]

for all \( x \in S \), where \( \lambda \in C \), \( \lambda \neq 0 \), \( \mu : S \to C \) is an additive map, and \( \phi : S \to RC \) is either a homomorphism or an anti–homomorphism.

The case when \( \theta \) is bijective was treated somewhat earlier by Beidar and Fong [BeF]. Besides the classical cases of Lie (the polynomial \( xy - yx \)) and Jordan (the polynomial \( xy + yx \)) homomorphisms, Theorem 5.4 also generalizes a recent result concerning \( n \)--Jordan maps [BrMM2] (the polynomial \( x^n \), or better, its linearized form) which was also proved using FI's. The crucial part of the proof of Theorem 5.4 is based on computing the expression \( \theta([f(x_1, \ldots, x_r), f(y_1, \ldots, y_r)]) \) in two different ways, which results in a certain FI.

The results that were explicitly stated in this section are applications of FI’s rather than GFI’s. Examples of how GFI’s can be used can be found in [Br10, BrC, BrH2, BrM3, C1, Hv].

**References**


DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARIBOR, PF, KOROŠKA 160, 2000 MARIBOR, SLOVENIA

E-mail address: bresar@uni-mb.si