Commutative Algebra with a focus on geometric and homological aspects.
Proceedings of Sevilla 2003.
Marcel Dekker's Lecture Notes in Pure and Applied Mathematics.

On free integral extensions generated by one element

IRENA SWANSON, Department of Mathematical sciences, New Mexico State University, Las Cruces, USA. *E-mail*: iswanson@nmsu.edu

ORLANDO VILLAMAYOR, Departamento de Matemáticas, Facultad de Ciencias, C-XV, Universidad Autónoma de Madrid, ctra. de Colmenar Viejo, Km. 15, 28049 Madrid, Spain. *E-mail*: villamayor@uam.es

Let R be a commutative integral domain with unity, and θ an element of an extension domain satisfying the relation

$$\theta^d = a_1\theta^{d-1} + a_2\theta^{d-2} + \cdots + a_{d-1}\theta + a_d,$$

with $a_i \in R$. We assume throughout that $R[\theta] \cong R[X]/(X^d - \sum_{i=1}^d a_i X^{d-i})$, where X is an indeterminate over R.

Suppose that R is a normal domain with quotient field K, and $K \subset L$ an algebraic extension. Let $\overline{R}$ be the integral closure of R in L, and fix $\theta \in \overline{R}$. There is information on the element θ encoded in the coefficients a_i. The first example arises when characterizing if θ belongs to the integral closure of the extended ideal $I\overline{R}$, for some ideal I in R. The objective of this paper is to study more precisely what information about θ is encoded in the coefficients a_i.

In a first approach, in Section 2, we show that for an ideal I in R, $a_i \in I^i$ for all i implies that $\theta^n R[\theta] \cap R \in I^n$ for all n, but that the converse fails. Thus contractions of powers of $\theta^n R[\theta]$ to R contain some information, but not enough.

We turn to a different approach in Sections 3 and 4, where we replace contractions by the trace functions (the image of $\theta^n R[\theta]$ in R by the trace function), and it turns out that if θ is separable over K, then the Trace codes more information.

The main results in this paper are:

a) Propositions 3.6 and 3.8 with conditions that assert that θ belongs to the integral closure of an extended ideal, and

b) Propositions 3.12 and 3.14 with conditions that assert that θ belongs to the tight closure of an extended ideal.

In all these Propositions we fix an ideal $I \subset R$ and consider the extended ideal $I.R[\theta]$. It should be pointed out that normally the condition for θ to belong to the integral closure of $I.R[\theta]$, is expressed in terms of a polynomial with coefficients in the ring $R[\theta]$; whereas we will express the same fact but in terms of a polynomial with coefficients in R; furthermore, in terms of the minimal polynomial of θ over R in case R is normal.

We also point out that we start with an ideal I in R, and an element θ in $\overline{R}$, and we study if θ belongs to integral or tight closure of the extended ideal, but only for the extension $R \subset R[\theta]$. This situation is however quite general, at least if I is a parameter ideal. In fact, given a complete local reduced ring (B, M) of dimension d containing a field, and with residue field k, and given a system of parameters $\{x_1, \ldots, x_d\}$, then B is finite over the subring $R = k[[x_1, \ldots, x_d]]$. Furthermore an element $\theta \in B$ is in the integral closure (in the tight closure) of the parameter ideal $< x_1, \ldots, x_d > B$, if and only if it is so in $< x_1, \ldots, x_d > R[\theta]$.

Throughout the previous argumentation there is a difference between characteristic zero and positive characteristic. The point is that our arguments will rely on properties of the subring of symmetric polynomials in a polynomial ring.

The relation of symmetric polynomials with our problem will arise and be discussed in the paper. We will show that the properties of θ that we are considering can be expressed in terms of symmetric functions on the roots of the minimal polynomial of θ, and hence as functions on the coefficients a_i of the minimal polynomial.

If k is a field of characteristic zero and S is a polynomial ring over k, the subring of symmetric polynomials of S can be generated in terms of the trace; however this is not so if k is of positive characteristic. In Section 4 we address the pathological behaviour in positive characteristic, and we give an example in which R is a k-algebra, k a field of positive characteristic, and the k-subalgebra generated by all the $Tr(\theta^n)$, as n varies, is not finitely generated.

We try to develop our results in maximal generality, in order to distinguish properties that hold under particular conditions (e.g. on the characteristic of R, separability of θ over K, etc.).

Our arguments rely on a precise expression of the powers θ^n of θ in terms of the natural basis $\{1, \theta, \theta^2, \ldots, \theta^{d-1}\}$ of $R[\theta]$ over R. This is done in Section 1 by using compositions, that is, ordered tuples of positive integers.

Similarly, we also develop a product formula for elements of $R[\theta]$ in terms of the natural basis.

1 Power and product formula

Every element of $R[\theta]$ can be written uniquely as an R-linear combination of $1, \theta, \theta^2, \ldots, \theta^{d-1}$. In this section we develop formulas for the R-linear combinations for all powers of θ, and for linear combinations of products.

Definition 1.1 Let e be a positive integer. A **composition** of e is an ordered tuple $(e_1, \ldots, e_k)$ of positive integers such that $\sum e_i = e$. Let $\mathcal{E}_e$ denote the set of all compositions of e.

For example, $\mathcal{E}_1 = \{(1)\}$, $\mathcal{E}_2 = \{(2), (1,1)\}$, $\mathcal{E}_3 = \{(3), (2,1), (1,2), (1,1,1)\}$.

We will express θ^n in terms of these compositions. Without loss of generality we may use the following notation:

Notation 1.2 *For $i > d$, set $a_i = 0$.*

Definition 1.3 Set $\mathcal{C}_0 = 1$, and for all positive integers e set

$$\mathcal{C}_e = \sum_{(e_1,\ldots,e_k)\in\mathcal{E}_e} a_{e_1} a_{e_2} \cdots a_{e_k}.$$

Remark 1.4 It is easy to see that for all $e > 0$, $\mathcal{C}_e = \mathcal{C}_0 a_e + \mathcal{C}_1 a_{e-1} + \cdots + \mathcal{C}_{e-1} a_1$.

Proposition 1.5 *For all $e \geq 0$,*

$$\theta^{d+e} = \sum_{i=0}^{d-1} \left(\mathcal{C}_0 a_{d+e-i} + \mathcal{C}_1 a_{d+e-i-1} + \mathcal{C}_2 a_{d+e-i-2} + \cdots + \mathcal{C}_e a_{d-i}\right) \theta^i.$$

Proof: The proof follows by induction on e. When $e = 0$, the coefficient of θ^i in the expression on the left above is $\mathcal{C}_0 a_{d-i} = a_{d-i}$, so the proposition holds for the base case by definition.

Now let $e > 0$. Then

$$\begin{aligned}\theta^{d+e} &= \theta^{d+e-1}\theta \\ &= \sum_{i=0}^{d-1} \left(\mathcal{C}_0 a_{d+e-i-1} + \mathcal{C}_1 a_{d+e-i-2} + \mathcal{C}_2 a_{d+e-i-3} + \cdots + \mathcal{C}_{e-1} a_{d-i}\right) \theta^{i+1}\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=0}^{d-2} \left(\mathcal{C}_0 a_{d+e-i-1} + \mathcal{C}_1 a_{d+e-i-2} + \mathcal{C}_2 a_{d+e-i-3} + \cdots + \mathcal{C}_{e-1} a_{d-i}\right) \theta^{i+1} \\
&\quad + \left(\mathcal{C}_0 a_e + \mathcal{C}_1 a_{e-1} + \mathcal{C}_2 a_{e-2} + \cdots + \mathcal{C}_{e-1} a_1\right) \theta^d \\
&= \sum_{i=1}^{d-1} \left(\mathcal{C}_0 a_{d+e-i} + \mathcal{C}_1 a_{d+e-i-1} + \mathcal{C}_2 a_{d+e-i-2} + \cdots + \mathcal{C}_{e-1} a_{d-i+1}\right) \theta^i \\
&\quad + \mathcal{C}_e \sum_{i=0}^{d-1} a_{d-i}\theta^i = \sum_{i=0}^{d-1} \sum_{j=0}^{e} \mathcal{C}_j a_{d+e-i-j}\theta^i. \qquad \square
\end{aligned}
$$

Recall that $a_i = 0$ if $i > d$. Thus in the expression for θ^{d+e} in the proposition above, many of the terms $\mathcal{C}_j a_{d+e-i-j}$ are trivially zero.

We similarly determine the product formula:

Let $f = \sum_{i=0}^{d-1} f_i\theta^i$, $g = \sum_{i=0}^{d-1} g_i\theta^i$ be two elements in $R[\theta]$. Write fg as an R-linear combination of $1, \theta, \ldots, \theta^{d-1}$. (Here, $f_i = g_i = 0$ if $i < 0$ or $i \geq d$.)

$$
\begin{aligned}
fg &= \sum_{i=0}^{2d-2} \sum_{k=0}^{d-1} f_k g_{i-k}\theta^i \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{d-1} f_k g_{i-k}\theta^i + \sum_{i=d}^{2d-2} \sum_{k=0}^{d-1} f_k g_{i-k}\theta^i \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{d-1} f_k g_{i-k}\theta^i + \sum_{e=0}^{d-2} \sum_{k=0}^{d-1} f_k g_{d+e-k}\theta^{d+e} \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{d-1} f_k g_{i-k}\theta^i + \sum_{e=0}^{d-2} \sum_{k=0}^{d-1} f_k g_{d+e-k} \sum_{i=0}^{d-1} \sum_{j=0}^{e} C_j a_{d+e-i-j}\theta^i \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{d-1} \left(f_k g_{i-k} + \sum_{e=0}^{d-2} f_k g_{d+e-k} \sum_{j=0}^{e} C_j a_{d+e-i-j} \right) \theta^i \\
&= \sum_{i=0}^{d-1} \sum_{k=0}^{d-1} f_k \left(g_{i-k} + \sum_{e=0}^{d-2} g_{d+e-k} \sum_{j=0}^{e} C_j a_{d+e-i-j} \right) \theta^i.
\end{aligned}
$$

We will use this expression mainly for the cases when $fg \in R$. Then the coefficients of θ^i in the expression above, for $i > 0$, are 0, and the constant coefficient is

$$\begin{aligned}
&\sum_{k=0}^{d-1} f_k \left(g_{-k} + \sum_{e=0}^{k-1} g_{d+e-k}(\mathcal{C}_0 a_{d+e} + \mathcal{C}_1 a_{d+e-1} + \cdots + \mathcal{C}_e a_d) \right) \\
&= f_0 g_0 + \sum_{k=0}^{d-1} f_k \sum_{e=0}^{k-1} g_{d+e-k} \mathcal{C}_e a_d.
\end{aligned}$$

2 Contractions

In this section we examine implications between $a_i \in I^i$ for all i, and $\theta^n R[\theta] \cap R \in I^n$ for all n, where I is an ideal of R. In case R is an $\mathbb{N}$-graded ring with $R = R_0[R_1]$ and $I = R_1 R$, then $a_i \in I^i$ is equivalent to saying that $\deg(a_i) \geq i$. (The two statements are not equivalent in general.)

We examine how under some $\mathbb{N}$-gradings on R, the degrees of the a_i affect and are affected by the degrees of the elements of $\theta^n R[\theta] \cap R$.

PROPOSITION 2.1 *With set-up on R, $a_1, \ldots, a_d$ and θ as in the introduction, if I is any ideal of R and $a_i \in I^i$ for all i, then $\theta^n R[\theta] \cap R \in I^n$ for all n.*

Similarly, if R is an $\mathbb{N}$-graded regular ring with a_i an element of R of degree at least i, then for all $n \geq 0$, $\theta^n R[\theta] \cap R$ is an ideal all of whose elements lie in degrees at least n.

Proof: First let $n < d$. Let $g = \sum_{i=0}^{d-1} g_i \theta^i$ be such that $\theta^n g \in R$. By the product formula from the previous section, the constant coefficient of $\theta^n g$ is

$$\delta_{n0} g_0 + \sum_{k=0}^{d-1} \delta_{kn} \sum_{e=0}^{k-1} g_{d+e-k} \mathcal{C}_e a_d,$$

where δ_{ij} is the Kronecker delta function. If $n = 0$, the proposition follows trivially, and if $n > 0$, $\theta^n g$ is a multiple of a_d, so it is in $I^d \subseteq I^n$.

Now let $n \geq d$. Write $n = d+e$. Let $g \in R[\theta]$ such that $\theta^{d+e} g \in R$. Write $\theta^{d+e} = \sum_{i=0}^{d-1} f_i \theta^i$. By assumption each a_i is in I^i, so that each $a_{e_1} a_{e_2} \cdots a_{e_k}$ lies in I raised to the power $\sum e_i$. Thus each $\mathcal{C}_e$ is in I^e. It follows that the coefficient f_i of θ^i in the expression of θ^{d+e} above is in I^{d+e-i}. Then by the product formula the constant part of $\theta^{d+e} g$ is in I raised to the power

$$\begin{aligned}
&\min\{\deg f_0, \deg(f_k \mathcal{C}_e a_d) | k = 0, \ldots, d-1; e = 0, \ldots, k-1\} \\
&\geq \min\{d+e, d+e-k+e+d | k = 0, \ldots, d-1; e = 0, \ldots, k-1\} = d+e,
\end{aligned}$$

which equals n. This proves the proposition. $\square$

However, the converse does not hold in general:

PROPOSITION 2.2 *Let R be a regular local ring with maximal ideal m, and let $a_1, \dots, a_d$ be a regular sequence. Then for all $n \geq 0$, $\theta^n R[\theta] \cap R \subseteq m^n$ (yet the a_i need not be in progressively higher powers of m).*

Proof: Let $n \geq 0$ and f a non-zero element of $\theta^n R[\theta] \cap R$. Write $f = \theta^n(s_0 + s_1\theta + \cdots + s_{d-1}\theta^{d-1})$ for some $s_i \in R$. Let $s = s_0 + s_1\theta + \cdots + s_{d-1}\theta^{d-1}$.

For each non-negative integer n, repeatedly rewrite each occurrence of θ^d in $\theta^n \cdot s$ as $\sum_{i=1}^d a_i \theta^{d-i}$ until $\theta^n s$ is in the form $\sum_{i=0}^{d-1} \sum_{j=0}^{d-1} b_{ij} s_j \theta^i$ for some $b_{ij} \in R$. In other words, $\sum_{i=0}^{d-1} \sum_{j=0}^{d-1} b_{ij} s_j \theta^i$ is the reduction of $\theta^n \cdot s$ with respect to the polynomial $\theta^d - \sum_{i=1}^d a_i \theta^{d-i}$. Set B_n to be the $d \times d$ matrix (b_{ij}).

Note that if $\theta^n s$ reduces to $\sum_{i=0}^{d-1} \sum_{j=0}^{d-1} b_{ij} s_j \theta^i$, then $\theta^{n+1} s$ reduces to the same polynomial as $\sum_{i=0}^{d-1} \sum_{j=0}^{d-1} b_{ij} s_j \theta^{i+1}$. But this is

$$\sum_{i=0}^{d-2} \sum_{j=0}^{d-1} b_{ij} s_j \theta^{i+1} + \sum_{j=0}^{d-1} b_{d-1,j} s_j \sum_{i=1}^{d} a_i \theta^{d-i}.$$

Thus the first row of B_{n+1} is a_d times the last row of B_n, and row i of B_{n+1}, with $i > 1$, equals row $i-1$ of B_n plus a_{d-i+1} times row d of B_n.

Note that B_0 is the identity matrix. Then by induction on n one can easily prove that for all $n \geq 0$, $\det B_n = \pm a_d^n$.

Now let C_n be the submatrix of B_n obtained from B_n by removing the first row and the first column. We claim that for all $n \geq 1$, $\det C_n = \pm a_{d-1}^{n-1} + p_n$ for some $p_n \in (a_1, \dots, a_{d-2}, a_d)$.

As B_0 is the identity matrix, then C_1 is the identity matrix, and the claim holds for $n = 1$. Suppose that the claim holds for $n \geq 1$. Let R_i be the ith row of B_n after deleting the first column. Then

$$C_{n+1} = \begin{bmatrix} R_1 + a_{d-1} R_d \\ R_2 + a_{d-2} R_d \\ \vdots \\ R_{d-2} + a_2 R_d \\ R_{d-1} + a_1 R_d \end{bmatrix}.$$

Then modulo $(a_1, \dots, a_{d-2}, a_d)$, as R_1 is a multiple of a_d,

$$\det(C_{n+1}) \equiv \det \begin{bmatrix} a_{d-1} R_d \\ R_2 \\ \vdots \\ R_{d-2} \\ R_{d-1} \end{bmatrix} = \pm a_{d-1} \det \begin{bmatrix} R_2 \\ R_3 \\ \vdots \\ R_{d-1} \\ R_d \end{bmatrix} = \pm a_{d-1} \det C_n,$$

so that the claim holds by induction.

We have proved that $\det(B_n) = \pm a_0^n \neq 0$. As $B_n(s_0, s_1, \ldots, s_{d-1})^T = (f, 0, \ldots, 0)^T$, by Cramer's rule $s_0 = \pm f \det(C_n)/a_d^n$. But $\det(C_n)$ and a_d are relatively prime, so that as $s_0 \in R$, necessarily f is a multiple of a_d^n. Thus $f \in m^n$. □

3 Trace

In the previous section we showed that $a_i \in I^i$ for all i implies that $\theta^n R[\theta] \cap R \in I^n$ for all n, but that the converse fails. In this section we analyze the situation when the contraction is replaced with the trace function. Namely, we prove that the condition $a_i \in I^i$ for all i implies that $Tr(\theta^n) \in I^n$ for all n, that the converse fails in general, but holds in several cases, for example in characteristic 0, see Proposition 3.6. Other special cases of the converse assume that θ is separable over R.

We start by proving the positive results. We first introduce some more notation. Throughout this section let k be a ring; in our applications it will be either the ring of integers, or a field, and R will be a k-algebra. (This imposes no condition on R if k is the ring of integers.) Let Y_i, $i = 1, \ldots, d$ and Z be variables over k. Consider the polynomial

$$(Z - Y_1) \cdots (Z - Y_d) = Z^d - s_1 \cdot Z^{d-1} + \cdots + (-1)^d s_d$$

in $k[Y_1, \ldots, Y_d, Z]$, where $s_i = s_i(Y_1, \ldots, Y_d)$ denotes the elementary symmetric polynomials. It is well known that $k[s_1, \ldots, s_d] \subset k[Y_1, \ldots, Y_d]$ is the subring of invariants by permutations, that the extension is finite, and hence that $k[s_1, \ldots, s_d]$ is also polynomial ring over k.

Since each s_i is homogeneous of degree i in the graded ring $k[Y_1, \ldots, Y_d]$, a natural weighted homogeneous structure is defined in the polynomial ring $k[s_1, \ldots, s_d]$ by setting $\deg(s_i) = i$, which makes the inclusion an homogeneous morphism of graded rings.

Remark 3.1 Set $v_i = Y_1^i + Y_2^i + \cdots + Y_d^i$, for $i \geq 0$. Then $k[v_1, v_2, \ldots] \subset k[s_1, \ldots, s_d]$, and since each v_i is homogeneous of degree i in $k[Y_1, \ldots, Y_d]$, the inclusion is homogeneous by setting $\deg(v_i) = i$. In other words, $v_i = v_i(s_1, \ldots, s_d)$ is weighted homogeneous of degree i in $k[s_1, \ldots, s_d]$. Let us finally recall that when k is a field of characteristic zero, then $k[v_1, \ldots, v_d] = k[s_1, \ldots, s_d]$.

Remark 3.2 The ring $k[s_1, \ldots, s_d][\Theta] = k[s_1, \ldots, s_d][Z]/ < Z^d - s_1 \cdot Z^{d-1} + \cdots + (-1)^d . s_d >$ is a free module of rank d over $k[s_1, \ldots, s_d]$. The trace of the endomorphism, on this finite module, defined by multiplication by Θ^i, is the weighted homogeneous polynomial $v_i \in k[s_1, \ldots, s_d]$ mentioned above.

In fact there are d different embeddings $\sigma_i : k[s_1, \dots, s_d][\Theta] \to k[Y_1, \dots, Y_d]$ of $k[s_1, \dots, s_d]$-algebras, each defined by $\sigma_i(\Theta) = Y_i$, and the trace (of the endomorphism) of any element $\Gamma \in k[s_1, \dots, s_d][\Theta]$ is $\sum \sigma_i(\Gamma)$.

Remark 3.3 Any primitive extension over a ring R, say

$$R[\theta] = R[Z]/ < Z^d - a_1 \cdot Z^{d-1} + \cdots + (-1)^d . a_d >$$

is

$$k[s_1, \dots, s_d][Z]/ < Z^d - s_1 \cdot Z^{d-1} + \cdots + (-1)^d . s_d > \otimes_{k[s_1,\dots,s_d]} R,$$

where k denotes here the ring of integers, and $\phi : k[s_1, \dots, s_d] \to R$ defined by $\phi(s_i) = a_i$. By change of base rings it follows that the trace of the endomorphism of R modules defined by $\theta^i : R[\theta] \to R[\theta]$ is $\phi(v_i(s))$. When R is a normal domain with quotient field K, and θ is an algebraic element over K with minimal polynomial $Z^d - a_1 \cdot Z^{d-1} + \cdots + (-1)^d . a_d \in R[Z]$, then the trace of the endomorphism $\theta^i : R[\theta] \to R[\theta]$ is $Tr(\theta^i)$, where Tr denotes the trace of the field extension $K \subset K[\theta]$. In what follows, for an arbitrary ring R, we abuse notation and set $Tr(\theta^i) = \phi(v_i(s))$.

Remark 3.4 Fix an ideal I in a k-algebra R. Suppose that a weighted homogeneous structure on the polynomial ring $k[T_1, \dots, T_d]$ is defined by setting $\deg(T_i) = m_i$, and let $G(T_1, \dots, T_d)$ be weighted homogeneous element of degree m. If $\phi : k[T_1, \dots, T_d] \to R$ is a morphism of k-algebras and $\phi(T_i) \in I^{m_i}$, then $\phi(G) \in I^m$.

Now we can finally prove that the analog of Proposition 2.1 holds also for the Trace function:

Proposition 3.5 *Let I be an ideal of R. Assume that for each $i = 1, \dots, d$, $a_i \in I^i$. Then $Tr(\theta^n) \in I^n$ for all positive integers n.*

Proof: The polynomial $Z^d - \sum_{i=0}^{d-1} a_i Z^i$ is the image of $Z^d - \sum_{i=0}^{d-1} (-1)^{i+1} s_i Z^i$ by the morphism $\phi : k[s_1, \dots, s_d] \to R$, $\phi(s_i) = (-1)^i a_i \in I^i$, so we may apply Remark 3.4. □

The converse holds easily when k is a field of characteristic zero:

Proposition 3.6 *If the ring R contains a field, say k, of characteristic zero then $a_i \in I^i$ for $i = 1, \dots, d$ if and only if $Tr(\theta^n) \in I^n$ for $1 \leq n \leq d$.*

Proof: The proof follows from the proof of previous Proposition and the second assertion in Remark 3.1. □

Furthermore, the converse holds in a much greater generality, see Proposition 3.8 below. We first introduce some conditions, and show some implications among them, culminating in Proposition 3.8.

Let R be an excellent normal domain, and K the quotient field of R. Normality asserts that if θ is a root of a polynomial $Z^n + b_1 \cdot Z^{n-1} + \cdots + b_n \in R[Z]$, then the minimal polynomial of θ over K is also in $R[Z]$. For an ideal I in R we study the following conditions:

Condition 1): The minimal polynomial of θ, $Z^d + a_1 \cdot Z^{d-1} + \cdots + a_d$, is such that $a_i \in I^i$.

Condition 2): The minimal polynomial of θ, $Z^d + a_1 \cdot Z^{d-1} + \cdots + a_d$, is such that $a_i \in \overline{I^i}$, the integral closure of I^i.

Condition 3): The element θ satisfies a polynomial equation $Z^n + b_1 \cdot Z^{n-1} + \cdots + b_n$, for some n, all $b_i \in I^i$.

Condition 4): θ is separable over K and $Tr_{K[\theta]/K}(\theta^i) \in I^i$.

It is clear that 1) implies both 2) and 3).

PROPOSITION 3.7 *Condition 3) implies Condition 2).*

Proof: : (Case I principal) If $I = < t >$ is a principal ideal and Condition 3) holds, it follows that θt^{-1} is an integral element over the ring R. If $Z^m + c_1 Z^{m-1} + \cdots + c_m \in R[Z]$ denotes the minimal polynomial of θt^{-1}; it is easy to check that $Z^m + tc_1 Z^{m-1} + t^2 c_2 Z^{m-2} + \cdots + t^m c_m$ is the minimal polynomial of θ over R. Hence, even Condition 1) holds in this case.

(The general case) Assume that, for some n, the element θ satisfies a polynomial equation $Z^n + b_1 Z^{n-1} + \cdots + b_n$, all $b_i \in I^i$. Let $Z^d + a_1 Z^{d-1} + \cdots + a_d$ denote the minimal polynomial of θ. We claim that $a_i \in \overline{I^i}$. Let S be the integral closure of the Rees algebra $R[It, t^{-1}]$ of I. Here t is a variable over R. As R is excellent, S is still Noetherian, excellent, normal. Its quotient field is $K(t)$. The minimal polynomial of θ over $K(t)$ is the same as the minimal polynomial of θ over K. Also, θ satisfies the polynomial equation $Z^n + b_1 Z^{n-1} + \cdots + b_n$, all $b_i \in I^i S = (It)^i t^{-i} S$, so that θ is integral over the principal ideal $t^{-1}S$. By the principal ideal case then all $a_i \in \overline{t^{-i}S} \cap R = \overline{I^n}$. □

PROPOSITION 3.8 *If θ is separable over K, and $Tr(\theta^r) \in I^r$ for all r big enough, then Condition 3) holds. In particular, Condition 2) holds.*

Proof: Let R be a normal ring with quotient field K, and set $L = K[\theta]$, where θ has minimal polynomial $f = Z^d + a_1 Z^{d-1} + \cdots + a_d$ with coefficients in R. So $\{1, \theta, \ldots, \theta^{d-1}\}$ is a basis of $R[\theta]$ over R.

For each index $j = 0, 1, \ldots, d-1$ we define $Tr(\theta^j.V)$ as a K-linear function on the variable V, say $Tr(\theta^j.V) : L \to K$. In addition $\{Tr(\theta^j.V) \mid j = 0, 1, \ldots, d-1\} \subset Hom_R(R[\theta], R)$ is a subset of the R- dual of the free module $R[\theta]$.

We will assume that the extension $K \subset L$ is separable, namely, that the discriminant Δ_f of the minimal polynomial f is non-zero in K (actually $\Delta_f \in R$), and we now argue as in [3] (Prop 11, page 40). Recall that setting $N = (n_{i,j})$ the $d \times d$ matrix where $n_{i,j} = Tr(\theta^i.\theta^j)$, then $\Delta_f = det(N)$. Since $\Delta_f \neq 0$ and $\{1, \theta, \dots, \theta^{d-1}\}$ is a basis of $L = K[\theta]$ over K, it follows that $\{Tr(\theta^j.V), j = 0, 1, \dots, d-1\}$ is a basis of $L^* = Hom_K(L, K)$.

Let T denote the free R-submodule in L^* generated by $\{Tr(\theta^j.V) \mid j = 0, 1, \dots, d-1\}$. So $T \subset Hom_R(R[\theta], R)$ is an inclusion of two free R submodules in L^*. Since the functor $Hom_R(-, R)$ reverses inclusions

$$R[\theta] = Hom_R(Hom_R(R[\theta], R), R) \subset Hom_R(T, R) \subset L.$$

Let $\{\omega_i, i = 0, 1, \dots, d-1\}$ be the dual basis of $\{Tr(\theta^j.V), j = 0, 1, \dots, d-1\}$ over the field K; it is also a basis of the R-module $Hom_R(T, R)$. Furthermore, for any element $\beta \in L$:

$$\beta = \sum_i Tr(\theta^i.\beta)\omega_i$$

is the expression of β as K-linear combination in the basis $\{\omega_i, i = 0, 1, \dots, d-1\}$. Note also that if $\beta \in R[\theta]$, all $Tr(\theta^i.\beta)$ are elements in R.

Set $R[\theta] = R^d$ by choosing basis $\{1, \theta, \dots, \theta^{d-1}\}$, and $Hom_R(T, R) = R^d$ with basis $\{\omega_i, i = 0, 1, \dots, d-1\}$, so the inclusion $R[\theta] \subset Hom_R(T, R)$ defines a short exact sequence

$$0 \to R^d \to R^d \to C \to 0$$

where C denotes the cokernel of the morphism given by the square matrix $N = (n_{i,j})$ mentioned above. Since $\Delta_f = det(N)$ it follows that $\Delta_f.Hom_R(T, R) \subset R[\theta]$; in fact $\Delta_f \in Ann(C)$.

Assume that for some ideal $I \subset R$, $Tr(\theta^r) \in I^r$ and all r big enough. In order to prove that Condition 3) holds we first note that

$$\theta^r = \sum_i Tr((\theta)^{i+r}).\omega_i \in I^r.Hom_R(T, R).$$

In fact, for r big enough:

$$J_r =< Tr(\theta^r), Tr(\theta^{r+1}), \dots, Tr(\theta^{r+d-1}) >\subset I^r$$

in R. But then,

$$\Delta_f \theta^r \in I^r \cdot \Delta_f \cdot Hom_R(T, R) \subset I^r R[\theta]$$

for all r big enough. This already shows that θ is in the integral closure of $IR[\theta]$ (integral closure in the ring $R[\theta]$). That means that θ satisfies a polynomial equation $Z^n + b_1.Z^{n-1} + \dots + b_n \in R[\theta][Z]$ with $b_i \in J^i$, $J = IR[\theta]$.

As in [4] (page 348), this is equivalent to the existence of a finitely generated $R[\theta]$ submodule, say Q, in the field L, such that $\theta \cdot Q \subset J \cdot Q$. In fact Q can be chosen as the ideal $(J + \theta \cdot R[\theta])^{n-1}$ in $R[\theta]$. Finally, since Q is a finitely generated $R[\theta]$-module, it is also a finitely generated R-module. On the other hand note that $J \cdot Q = I \cdot Q$, and Condition 3) follows now from the determinant trick applied to $\theta \cdot Q \subset I \cdot Q$. □

COROLLARY 3.9 *If θ is separable over a local regular ring (R, m), then $Tr(\theta^n) \in m^n$ for all n big enough if and only if $a_i \in m^i$ for all $i = 1, \ldots, d$.* □

However, this equivalence fails in general for arbitrary rings and arbitrary ideals. The converse fails, for example, if θ is not separable over R:

EXAMPLE 3.10 Let k be a field of characteristic 2, $d = 2$, $a_1 = 0$. Then $Tr(\theta^n) = 0$ for all n, but a_2 need not be in I^2.

Another failure of the converse is if the powers of I are not integrally closed:

EXAMPLE 3.11 Let $R = k[X, Y]$ be a polynomial ring in two variables X and Y over a field k of characteristic 2. Let I be the ideal generated by $X^8, X^7Y, X^6Y^2, X^2Y^6, XY^7, Y^8$, and the minimal equation for θ being

$$\theta^2 - X^8\theta - X^{11}Y^5.$$

Note $a_1 = X^8 \in I$, $a_2 = X^{11}Y^5 \notin I^2$, but $X^{11}Y^5 \cdot I \subseteq I^3$. Hence

$$\begin{aligned} Tr(\theta) &= X^8 \in I, \\ Tr(\theta^2) &= X^8\, Tr(\theta) + Tr(X^{11}Y^5) = X^{16} \in I^2, \end{aligned}$$

and for $n \geq 3$,

$$Tr(\theta^n) = X^8\, Tr(\theta^{n-1}) + X^{11}Y^5\, Tr(\theta^{n-2}) \in I^n.$$

Set as before the ideals $J_r =< Tr(\theta^r), Tr(\theta^{r+1}), \ldots, Tr(\theta^{r+d-1}) >$ in R. Note that $\{\theta^r, \theta^{r+1}, \cdots, \theta^{r+d-1}\}$ generate the ideal $\theta^r R[\theta]$ as R module, so that J_r is the image of this ideal by the trace map.

If R is of characteristic $p > 0$, and $I =< f_1, \cdots, f_l >$, then $I^{[p^e]}$ denotes the ideal $< f_1^{p^e}, \cdots, f_l^{p^e} > \subset R$.

PROPOSITION 3.12 *Let θ be separable over a local regular ring (R, m) of characteristic p. If $J_{p^r} \subset m^{[p^r]}$ for all r big enough, then θ is in the tight closure of the parameter ideal $m.R[\theta]$.*

Proof: We apply the same argument as in the previous Proposition. Note that in this case

$$\theta^{p^r} = \sum_{0 \le i \le d-1} Tr((\theta)^{i+p^r}).\omega_i \in m^{[p^r]}.Hom_R(T, R).$$

But then,

$$\Delta_f \theta^{p^r} \in m^{[p^r]} \cdot \Delta_f \cdot Hom_R(T, R) \subset m^{[p^r]} R[\theta]$$

for r big enough. This already shows that θ is in the tight closure of $mR[\theta]$ (tight closure in the ring $R[\theta]$).

EXAMPLE 3.13 Consider $R = k[y, z]$ where k is a field of odd characteristic, and set $R[\theta]$, $\theta^2 - a_2 = 0$, where $a_2 = y^3 + z^n$, $n \ge 7$, n some integer. We will prove that $J_r \subseteq < y^{p^r}, z^{p^r} >$. Here $\{1, \theta\}$ is a basis of $R[\theta]$ over R. $Tr(1) = 2$ (invertible in k), and $Tr(\theta) = 0$. Since the trace is compatible with Frobenius, $Tr(\theta^{p^r}) = Tr(\theta)^{p^r} = 0$, so it suffices to check that $Tr(\theta^{p^r+1}) \in < y^{p^r}, z^{p^r} >$. Set $p^r + 1 = 2k$, so $(\theta)^{p^r+1} = a_2^k$, and $Tr(\theta^{p^r+1}) = 2a_2^k$. We finally refer to [1], page 14, Example 1.6.5, for a proof that $a_2^k \in < y^{p^r}, z^{p^r} >$ if $n \ge 7$ and r is sufficiently large.

PROPOSITION 3.14 *Assume that θ is separable over a local regular ring (R, m) of characteristic p, and let Δ denote the discriminant. If θ is in the tight closure of the parameter ideal $mR[\theta]$ (in a ring containing $R[\theta]$), then $\Delta.J_{p^r} \subset m^{[p^r]}$ (in R) for all r.*

Proof: Let $f(X) \in R[X]$ denote the minimal polynomial of θ. Recall that the resultant $\Delta \in < f(X), f'(X) > \cap R$ (in $R[X]$), and hence $\Delta \in < f'(\theta) >$ in $R[\theta]$. Since $f'(\theta)$ is a test element, Δ is a test element, and

$$\Delta.(\theta)^{p^r} \in m^{[p^r]} R[\theta]$$

for all r.

Note that $R[\theta] \subset Hom_R(T, R)$ (hence $m^{[p^r]} R[\theta] \subset m^{[p^r]} Hom_R(T, R)$), and that, choosing as before the basis $\{\omega_0, \omega_1, \ldots, \omega_{d-1}\}$ in $Hom_R(T, R)$:

$$\Delta\theta^{p^r} = \sum_{0 \le i \le d-1} \Delta. Tr((\theta)^{i+p^r}).\omega_i \in m^{[p^r]}.Hom_R(T, R),$$

which shows that $\Delta.J_{p^r} \subset m^{[p^r]}$ in the ring R. □

4 The subalgebra of R generated by $Tr\,\theta^n, n \ge 0$

Let R and θ be as before, so that $R[\theta] \cong R[X]/(X^d + \sum_{i=1}^{d} (-1)^i a_i X^{d-i})$. Assume now that R is an algebra over a field k. It follows from Remarks 3.1

and 3.3 that if k of characteristic zero, the k-subalgebra generated by the traces $Tr\,\theta^n$ for all n, is $k[a_1, \cdots, a_d] (\subset R)$. In particular it is finitely generated. This subalgebra need not be finitely generated over a field of positive characteristic, as we show below.

First we recall some notation. Let B_n be the matrix as in the proof of Proposition 2.2. The trace of θ^n is exactly the trace of B_n.

Remark 4.1 In the proof of Proposition 2.2 we showed that the first row of B_{n+1} is a_d times the last row of B_n, and row i of B_{n+1}, with $i > 1$, equals row $i-1$ of B_n plus a_{d-i+1} times row d of B_n.

We determine the entries of B_n more precisely:

Lemma 4.2 *For $n \leq d$,*

$$(B_n)_{ij} = \begin{cases} \delta_{i,j+n} & \text{if } j \leq d-n, \\ \sum_{k=d-n+1}^{j-1} a_{j-k}(B_n)_{ik} + a_{n-i+j} & \text{if } j > d-n. \end{cases}$$

Furthermore, for all $j > d-n$,

$$(B_n)_{ij} = (B_d)_{i,j-d+n}.$$

Proof: We proceed by induction on n. The formulation is correct for $n = 0$. Thus we assume that $n > 0$. By Remark 4.1 the formulations of the entries of B_n in the first $d-n+1$ columns are correct: in the first $d-n$ columns, the entries are $\delta_{i,j+n}$, and $(B_n)_{i,d-n+1} = a_{d-i}$.

Now let $i = 1$, $j > d-n+1$. Then

$$\begin{aligned} (B_n)_{1j} &= a_d(B_{n-1})_{dj} \\ &= a_d \left(\sum_{k=d-(n-1)+1}^{j-1} a_{j-k}(B_{n-1})_{dk} + a_{n-1-d+j} \right) \\ &= \sum_{k=d-n+2}^{j-1} a_{j-k}a_d(B_{n-1})_{dk} + a_d a_{n-1-d+j} \\ &= \sum_{k=d-n+2}^{j-1} a_{j-k}(B_n)_{1k} + (B_n)_{1,d-n+1}a_{j-(d-n+1)} \\ &= \sum_{k=d-n+1}^{j-1} a_{j-k}(B_n)_{1k} \\ &= \sum_{k=d-n+1}^{j-1} a_{j-k}(B_n)_{1k} + a_{n-1+j} \end{aligned}$$

as $n-1+j>d$ so that $a_{n-1+j}=0$. Now let $i>1$, $j>d-n+1$. Then

$$\begin{aligned}
(B_n)_{ij} &= (B_{n-1})_{i-1,j} + a_{d-i+1}(B_{n-1})_{dj} \\
&= \sum_{k=d-(n-1)+1}^{j-1} a_{j-k}(B_{n-1})_{i-1,k} + a_{(n-1)-(i-1)+j} \\
&\quad + a_{d-i+1}\left(\sum_{k=d-(n-1)+1}^{j-1} a_{j-k}(B_{n-1})_{dk} + a_{(n-1)-d+j}\right) \\
&= \sum_{k=d-n+1}^{j-1} a_{j-k}(B_{n-1})_{i-1,k} + a_{n-i+j} + a_{d-i+1}\sum_{k=d-n+1}^{j-1} a_{j-k}(B_{n-1})_{dk} \\
&\quad \text{(because for } k=d-n+1,\ (B_{n-1})_{i-1,k}=0 \text{ and } (B_{n-1})_{dk}=1) \\
&= \sum_{k=d-n+1}^{j-1} a_{j-k}(B_n)_{ik} + a_{n-i+j}.
\end{aligned}$$

Observe that the last statement is true for $j=d-n+1$. Then by induction on $j>d-n+1$,

$$\begin{aligned}
(B_n)_{ij} &= \sum_{k=d-n+1}^{j-1} a_{j-k}(B_n)_{ik} + a_{n-i+j} \\
&= \sum_{k=d-n+1}^{j-1} a_{j-k}(B_d)_{i,k-d+n} + a_{n-i+j} \\
&= \sum_{l=1}^{j-d+n-1} a_{j-l-d+n}(B_d)_{il} + a_{n-i+j} \\
&= (B_n)_{i,j-d+n}.
\end{aligned}$$

□

It then follows

COROLLARY 4.3 *Whenever* $1 \le n \le d$,

$$Tr(\theta^n) = \sum_{i=1}^{n-1} a_{n-i}\, Tr(\theta^i) + n a_n,$$

and $Tr(\theta^n)$ *is a polynomial in* $a_1, \dots, a_n$*, homogeneous of degree* n *under the weights* $\deg(a_i) = i$*.*

Proof: By definition, $Tr(\theta^n) = Tr(B_n) = \sum_{i=1}^d (B_n)_{ii}$, and by Lemma 4.2 this equals

$$Tr(\theta^n) = \sum_{i=d-n+1}^{d} (B_n)_{ii} = \sum_{i=d-n+1}^{d} (B_d)_{i,i-d+n} = \sum_{j=1}^{n} (B_d)_{d-n+j,j},$$

i.e., this is the sum of the elements of B_d on the nth diagonal, counting from the bottom leftmost corner. Hence,

$$\begin{aligned} Tr(\theta^n) &= \sum_{j=1}^{n} \left(\sum_{k=1}^{j-1} a_{j-k} (B_d)_{d-n+j,k} + a_{d-(d-n+j)+j} \right) \\ &= \sum_{j=1}^{n} \sum_{k=1}^{j-1} a_{j-k} (B_d)_{d-n+j,k} + n a_n. \end{aligned}$$

Now we change the double summation: c sums over the differences $j-k$, and k keeps the same role:

$$\begin{aligned} Tr(\theta^n) &= \sum_{c=1}^{n-1} \sum_{k=1}^{n-c} a_c (B_d)_{k+c+d-n,k} + n a_n \\ &= \sum_{c=1}^{n-1} a_c \sum_{k=1}^{n-c} (B_d)_{k+d-(n-c),k} + n a_n \\ &= \sum_{c=1}^{n-1} a_c \, Tr(\theta^{n-c}) + n a_n. \end{aligned}$$

□

For $n \geq 0$ let $\mathcal{C}_n$ be as in Definition 1.3. We adopt the notation that for $n < 0$, $\mathcal{C}_n = 0$. Then for $n \geq 0$, let P_n be the row matrix $[\mathcal{C}_n, \mathcal{C}_{n-1}, \ldots, \mathcal{C}_{n-d+1}]$, and for each $n = 1, \ldots, d$, let

$$F_n = \sum_{i=0}^{d-1} a_{d+n-1-i} \, Tr(\theta^i).$$

Let $\vec{F}$ be the vector $(F_1, \ldots, F_d)$. With this we can give another formulation of the trace of powers of θ:

LEMMA 4.4 *For each $e \geq 0$,*

$$Tr(\theta^{d+e}) = P_e \cdot \vec{F}.$$

Proof: By Proposition 1.5,

$$
\begin{aligned}
Tr(\theta^{d+e}) &= \sum_{i=0}^{d-1}\sum_{j=0}^{e}\mathcal{C}_j a_{d+e-i-j}\,Tr(\theta^i) = \sum_{j=0}^{e}\mathcal{C}_j\sum_{i=0}^{d-1}a_{d+e-i-j}\,Tr(\theta^i) \\
&= \sum_{j=e-d+1}^{e}\mathcal{C}_j\sum_{i=0}^{d-1}a_{d+e-i-j}\,Tr(\theta^i) = \sum_{j=e-d+1}^{e}\mathcal{C}_j F_{e-j+1} \\
&= P_e\cdot\vec{F}.
\end{aligned}
$$

□

Now we can give an example of a k algebra R, and θ as before, where k is a field of positive characteristic, and the subalgebra of R generated over k by $Tr(\theta^n)$ as n varies is not a finitely generated algebra (compare with Remark 3.1):

EXAMPLE 4.5 Let k be a field of positive prime characteristic p, $d=p$, and $a_1,\ldots,a_d$ indeterminates over k, $R=k[a_1,\ldots,a_d]$. Let $A=k[Tr\,\theta,Tr\,\theta^2,\ldots]$. It follows from Remark 3.3 and Remark 3.1 that $A\subseteq R$. But this A is not finitely generated over k, as we prove below.

For each $n\geq 1$, let $A_n=k[Tr\,\theta,Tr\,\theta^2,\ldots,Tr\,\theta^n]$.

Claim: For each $n\geq 0$ and $l\in\{0,\ldots,d-1\}$:

$$A_{dn+l}=k[a_i a_d^j|\text{ either }j<n\text{ or else }j=n\text{ and }i\leq l].$$

We will prove this by induction on n. It holds for $n=0$ by Corollary 4.3. Thus by the definition of the F_i and by Corollary 4.3, all F_i are in all $A_{(n+1)d+l}$. Furthermore, each F_i is linear in a_d.

By Lemma 4.4, $Tr(\theta^{(n+1)d+l})$ equals

$$\mathcal{C}_{nd+l}F_1+\cdots+\mathcal{C}_{nd+1}F_l+\mathcal{C}_{nd}F_{l+1}+\mathcal{C}_{nd-1}F_{l+1}+\cdots+\mathcal{C}_{nd-(d-i-1)}F_d.$$

By the structure of the $\mathcal{C}_i$, a_d appears in C_i with exponent at most i/d. Thus the summand $\mathcal{C}_{nd-1}F_{l+1}+\cdots+\mathcal{C}_{nd-(d-i-1)}F_d$ lies in $A_{(n+1)d+l-1}$. Also, in the expansion of the summand $\mathcal{C}_{nd+l}F_1+\cdots+\mathcal{C}_{nd+1}F_l$, in each term a_d either appears with exponent n or smaller, or else it appears with exponent exactly $n+1$ and is multiplied by one of the variables $a_1,\ldots,a_{l-1}$. Thus also this summand lies in $A_{(n+1)d+l-1}$. Thus

$$A_{(n+1)d+l}=A_{(n+1)d+l-1}[\mathcal{C}_{nd}F_{l+1}].$$

F_{l+1} is linear in a_d with leading coefficient $Tr(\theta^l)$. $\mathcal{C}_{nd}$ equals a_d^n plus terms of lower a_d-degree, so that similarly, by Corollary 4.3,

$$A_{(n+1)d+l}=A_{(n+1)d+l-1}[a_d^n a_d\,Tr(\theta^l)]=A_{(n+1)d+l-1}[a_d^{n+1}la_l].$$

This proves the claim. As $a_1,\ldots,a_d$ are variables over k, this means that A is not finitely generated over k.

As an almost immediate corollary we can give another proof of Proposition 3.8 in a special case:

PROPOSITION 4.6 *Let $d = p$, i.e., the degree of the extension is the same as the characteristic of the ring. Assume that $X^d - \sum_{i=1}^{d} a_i X^{d-i}$ is a separable polynomial. Let v be any valuation $v : R \to \mathbb{N} \cup \{\infty\}$ such that $v(x) = \infty$ if and only if $x = 0$. Then $v(Tr(\theta^n)) \geq n$ for all n if and only if $v(a_i) \geq i$ for all i.*

Proof: With notation as above, one can prove by induction on $nd + l$ that $v(a_d^n a_l) \geq nd + l$. In particular, for $l = 1, \ldots, d-1$, $v(a_l) \geq l$. Also,

$$v(a_d) = \frac{1}{n}(v(a_d^n a_l) - v(a_l)) \geq \frac{1}{n}(nd + l - v(a_l))$$

for all n and l. As at least one $v(a_l)$ is finite (by the separability assumption), it follows that $v(a_d) \geq d$. □

References

[1] C. Huneke, *Tight closure and its applications. With an appendix by Melvin Hochster.* CBMS Regional Conference Series in Mathematics, 88. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1996.

[2] J. Lipman, A. Sathaye, *Jacobian ideals and a theorem of Briancon-Skoda*, Michigan Math. J. 28 (1981), 199-222.

[3] J-P Serre, *Local Algebra*, Springer Monographs in Mathematics. Springer-Verlag Berlin Heidelberg New York.

[4] O. Zariski and P. Samuel, *Commutative Algebra II*, D. Van Nostrand, 1980.