THE DEDEKIND-MERTENS LEMMA
AND THE CONTENTS OF POLYNOMIALS

DAVID E. RUSH

(Communicated by Wolmer V. Vasconcelos)

Abstract. Let \( R \) be a commutative ring, let \( X \) be an indeterminate, and let \( g \in R[X] \). There has been much recent work concerned with determining the Dedekind-Mertens number \( \mu_R(g) = \min \{ k \in \mathbb{N} | c_R(f)^{k-1}c_R(fg) = c_R(f)^kc_R(g) \text{ for all } f \in R[X] \} \), especially on determining when \( \mu_R(g) = 1 \). In this note we introduce a universal Dedekind-Mertens number \( u_R(g) \), which takes into account the fact that \( S(g) \leq \deg(g) + 1 \) for any ring \( S \) containing \( R \) as a subring, and show that \( u_R(g) \) behaves more predictably than \( \mu_R(g) \).

Introduction

Many papers (\cite{1}, \cite{2}, \cite{3}, \cite{4}, \cite{5}, \cite{7}, \cite{8}, \cite{9}, \cite{10}, \cite{13}) have recently considered questions concerning the following well-known result which is usually called the Dedekind-Mertens Lemma:

Lemma 0.1. If \( g \in R[X] \) and \( \deg(g) = n \), then
\[
  c_R(f)^n c_R(fg) = c_R(f)^{n+1} c_R(g) \text{ for all } f \in R[X].
\]

Much of this work has been on determining the smallest \( n \) for which (0.1) holds. To obtain a refinement of Lemma 0.1 in \cite{10} the authors defined the Dedekind-Mertens number \( \mu_R(g) \) of \( g \in R[X] \) to be the smallest positive integer \( k \) such that \( c_R(f)^{k-1}c_R(fg) = c_R(f)^kc_R(g) \) for all \( f \in R[X] \). Thus Lemma 0.1 states that \( \mu_R(g) \leq \deg(g) + 1 \). It follows that \( \mu_R(g) = \sup \{ \mu_{R_M}(g) | M \text{ is a maximal ideal of } R \} \). Thus in considering \( \mu_R(g) \) we may as well assume \( R \) is quasilocal. In this case, the main result of \cite{10} improves Lemma 0.1 to
\[
  \mu_R(g) \leq \mu_R(c_R(g))
\]
where \( \mu_R(M) \) denotes the minimal number of generators of the \( R \)-module \( M \). In \cite{10}, \cite{4}, the question of the opposite inequality to (0.2) was also considered. The special case of whether \( \mu_R(g) = 1 \) implies \( c_R(g) \) is principal was considered as early as \cite{14}, and several further results have recently been obtained on this case (\cite{2}, \cite{3}, \cite{4}, \cite{5}, \cite{7}, \cite{8}, \cite{9}, \cite{10}).

An important further property of the exponent \( n \) in Lemma 0.1 is that it is universal in the sense that the formula (0.1) continues to hold if \( f \) is chosen to
have coefficients in any ring $S$ containing $R$ as a subring, whereas we may have $\mu_R(g) < \mu_S(g)$ [Remark 1.7], or $\mu_R(c_R(g)) > \mu_S(c_S(g))$. Some of the history of the Dedekind-Mertens Lemma is discussed in [4] where the importance of this independence of the base ring is stressed. The object of this note is to introduce the universal Dedekind-Mertens number, and to point out that if one switches from the Dedekind-Mertens number as defined above, to the universal Dedekind-Mertens number, then the counterparts to the questions considered in [10], [4] become much simpler.

1. Strong Dedekind-Mertens Lemma

The original Dedekind-Mertens Lemma, as given for example in [11], [12], [13] and [6, p. 3] is stronger than Lemma 0.1. To explain this we extend the definition of content. If $\phi : R \to S$ is a homomorphism of rings and $f \in S[X]$, we define $c_R(f)$ to be the $R$-submodule of $S$ generated by the coefficients of $f$. If $A$, $B$ are $R$-submodules of $S$, we may define $AB$ to be the $R$-submodule of $S$ generated by $\{ab \mid a \in A, b \in B\}$. Then $c_R(f)^n c_R(fg)$ and $c_R(f)^{n+1} c_R(g)$ make sense. In particular, if $\phi$ is an inclusion of a subring $R$ into $S$, it is clear that the smaller that one chooses $R$ the stronger the condition $c_R(f)^n c_R(fg) = c_R(f)^{n+1} c_R(g)$ becomes. The Dedekind-Mertens Lemma as given for example in [13, 1] states:

Lemma 1.1. Let $g \in R[X]$ with $\deg(g) = n$. Then

$$cz(f)^n cz(fg) = cz(f)^{n+1} cz(g)$$

for all $f \in R[X]$. In particular, unlike the inequality (0.2) this condition is universal in the sense that it continues to hold if $R$ is replaced with any ring containing the coefficients of $f$ and $g$. Thus we may as well take the coefficients of $f$ to be independent indeterminates. Because of the importance of this universality in Kronecker’s use of the content to develop his theory of divisors [3], and other reasons it is of interest to have a universal version of the inequality (0.2) and other results as well. For example if $A$ is a subring of $B$, $f \in A[X]$ with $c_A(f) = A$ and $g \in B[X]$ with $fg \in A[X]$, it is immediate from Lemma 1.1 but not from Lemma 0.1 that $g \in A[X]$. If $M$ is an $R$-module and $g \in M[X]$, let $c_{RM}(g)$ denote the $R$-submodule of $M$ generated by the coefficients of $g$. Observe that if $M$ is a submodule of an $R$-module $N$, then $c_{RM}(g) = c_{RN}(g)$. Thus we may just write $c_R(g)$. We will let $R$ be a fixed quasilocal ring throughout. In considering the relationship between $\mu_R(g)$ and $\mu_R(c_R(g))$ for $g \in R[X]$ in [4], the authors defined the polarized Dedekind-Mertens number $\mu_R(g)$ of $g$ with respect to $R$ to be the smallest positive integer $k$ such that

$$\sum_{i=1}^{k} c_R(f_i) c_R(f_1) c_R(f_2) \cdots c_R(f_i) \cdots c_R(f_k) c_R(f_k) c_R(g)$$

for all $f_1, \ldots, f_k \in R$. We also define the universal polarized Dedekind-Mertens number, and show that it is the same as the universal Dedekind-Mertens number.

Definition 1.2. Let $M$ be an $R$-module and let $g \in M[X]$. Let $T = \{t_i \mid i \in \mathbb{N}\}$ be a new set of indeterminates. The universal Dedekind-Mertens number $u \mu_R(g)$ of $g$ with respect to $R$ is the smallest positive integer $k$ such that, for each $f \in R[T][X]$, it holds that

$$c_R(f)^{k-1} c_R(fg) = c_R(f)^k c_R(g)$$

as submodules of $M[T] = M \otimes_R R[T]$. 
The universal polarized Dedekind-Mertens number $\tilde{u}\mu_R(g)$ of $g$ with respect to $R$ is the smallest positive integer $k$ such that for all $f_1, \ldots, f_k \in R[T][X]$ we have
\[
\sum_{i=1}^{k} c_R(f_i) c_R(f_1) c_R(f_2) \cdots \tilde{c}_R(f_i) \cdots c_R(f_k) = c_R(f_1) c_R(f_2) \cdots c_R(f_k) c_R(g)
\]
as submodules of $M[T] = M \otimes_R R[T]$.

It is clear that $u\mu_R(g) \leq u\tilde{\mu}_R(g)$ and the proof of the Dedekind-Mertens Lemma given in [13] actually shows that $u\tilde{\mu}_R(g) \leq \deg(g) + 1$. Also, if $R$ is a subring of $S$, it follows that $u\mu_R(g) \geq u\mu_S(g)$ and $u\tilde{\mu}_R(g) \geq u\tilde{\mu}_S(g)$.

2. Results

While the focus of much of [2], [3], [4], [5], [7], [9], [11] has been on determining the Dedekind-Mertins number $\mu_R(g)$, with many interesting partial results, if we switch to the universal Dedekind-Mertins number $u\mu_R(g)$, we have the following result.

**Theorem 2.1.** Let $(R, m)$ be a quasi-local ring, let $M$ be an $R$-module and let $g \in M[X]$. Then $\mu_R(\mu_R(g)) = u\mu_R(g) = u\mu_R(g)$.

**Proof.** As noted before we have $u\mu_R(g) \leq u\tilde{\mu}_R(g)$. The proof of the opposite inequality is similar to that of [4] Lemma 2.4. Let $u\mu_R(g) = k$ and let $f_1, \ldots, f_k \in R[T][X]$. Let $N_i > \max\{\deg_X(f_1), \deg_X(f_1 g)\}$, and if $N_1, \ldots, N_{i-1}$ have been defined, let $N_i > \max\{\deg_X(f_i), \deg_X(f_i g) + N_{i-1}\}$. Let $t_1, \ldots, t_{k-1}$ be members of $T$ which do not appear in any of the $f_i$, and let $T' = T - \{t_1, \ldots, t_{k-1}\}$.

By the definition of $u\mu_R(g)$ we have
\[
c_R(f_1 + \sum_{i=2}^{k} f_it_{i-1}X^{N_{i-1}})^{k-1}c_R([f_1 + \sum_{i=2}^{k} f_it_{i-1}X^{N_{i-1}}]g)
\]
\[= c_R(f_1 + \sum_{i=2}^{k} f_it_{i-1}X^{N_{i-1}})^{k}c_R(g).
\]

By the choice of the $N_i$ this is
\[
[c_R(f_1) + \sum_{i=2}^{k} c_R(f_i)t_{i-1}]^{k-1}[c_R(f_1 g) + \sum_{i=2}^{k} c_R(f_1 g)t_{i-1}]
\]
\[= [c_R(f_1) + \sum_{i=2}^{k} c_R(f_i)t_{i-1}]^{k}c_R(g).
\]

Considering these as polynomials in $t_1, \ldots, t_{k-1}$ with coefficients in $R[X, T']$, and comparing the coefficients of the monomial $t_1 t_2 \cdots t_{k-1}$, we get
\[
\sum_{i=1}^{k} c_R(f_i) c_R(f_1) c_R(f_2) \cdots \tilde{c}_R(f_i) \cdots c_R(f_k) = c_R(f_1) c_R(f_2) \cdots c_R(f_k) c_R(g).
\]

Thus $u\mu_R(g) \geq u\tilde{\mu}_R(g)$.

The proof that $u\mu_R(g) \leq \mu_R(\mu_R(g))$ is very similar to the proof given in [10] that $\mu_R(g) \leq \mu_R(\mu_R(g))$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
Lemma 2.2. Let \((R, m)\) be a quasilocal ring, let \(T\) be a countably infinite set of independent indeterminates over \(R[X]\), let \(M\) be an \(R\)-module and let \(g \in M[X]\). Let \(b \in mc_R(g)\) and \(h = g + bX^i\). Let \(A\) be a finitely generated \(R\)-module over \(R[T]\) and let \(f \in R[T][X]\). If \(A_cR(f)c_R(h) = A_cR(fh)\), then \(A_cR(f)c_R(g) = A_cR(fg)\).

Proof. It suffices to show \(A_cR(f)c_R(g) \subseteq A_cR(fg)\). Since \(b \in mc_R(g)\) and \(h = g + bX^i\), we have \(c_R(h) \subseteq c_R(g) \subseteq c_R(h) + mc_R(g)\), and thus \(c_R(g) = c_R(h)\) by Nakayama’s Lemma. Thus

\[
A_cR(f)c_R(g) = A_cR(f)c_R(h) = A_cR(fh) = A_cR(f(g + bX^i))
\]

\[
\subseteq A_cR(fg) + A_cR(f) \subseteq A_cR(fg) + mA_cR(f)c_R(g).
\]

By Nakayama’s Lemma we have \(A_cR(f)c_R(g) = A_cR(fg)\).

To show \(u\mu_R(g) \leq \mu_R(c_R(g))\), we may assume that \(c_R(g)\) is minimally generated by \(k \geq 2\) elements, and that if \(h \in M[X]\) with \(c_R(h)\) minimally generated by fewer than \(k\) elements, then for any \(f \in R[T][X]\) we have \(c_R(f)k^{-2}c_R(fh) = c_R(f)^{k-1}c_R(h)\).

Let \(g = b_mX^m + \cdots + b_1X + b_0\). By the above lemma we may assume \(b_m\) is a minimal generator of \(c_R(g)\). Then \(g = b_mh + g_1\), where \(c_R(g_1)\) is generated by fewer than \(k\) elements and \(h \in R[X]\) with \(c_R(h) = R\).

Write \(f = a_nX^n + f_1\) where \(\deg(f_1) < \deg(f) = n\). By induction on \(\deg(f)\) we may also assume \(c_R(f_1)k^{-1}c_R(fg) = c_R(f_1)^{k-1}c_R(g)\).

Claim 1. \(c_R(fg_1) \subseteq c_R(fg) + b_mC_R(f_1)\).

Indeed we have

\[
c_R(fg_1) = c_R(f(g - b_mh)) \subseteq c_R(fg) + c_R(b_mhf) = c_R(fg) + b_mC_R(f)
\]

\[
= c_R(fg) + b_mC_R(a_nX^n + f_1) \subseteq c_R(fg) + a_nb_mR + b_mC_R(f_1),
\]

and since \(a_nb_m \in c_R(fg)\), this is \(c_R(fg) + b_mC_R(f_1)\). This proves Claim 1.

Claim 2. \(c_R(fg) \subseteq c_R(fg) + a_nC_R(g_1)\).

Indeed we have

\[
c_R(f_1g) = c_R((f - a_nX^n)g) \subseteq c_R(fg) + a_nC_R(g) \subseteq c_R(fg) + a_nC_R(b_mh + g_1)
\]

\[
\subseteq c_R(fg) + a_nC_R(g) + a_nC_R(g_1) = c_R(fg) + a_nC_R(g_1).
\]

The last equality holds since \(a_nb_m \in c_R(fg)\). This proves Claim 2.

Now to prove \(u\mu_R(g) \leq \mu_R(c_R(g))\), it suffices to show that each term of \(c_R(f)^kC_R(g)\) of the form \(\theta = a_0^{v_0}a_1^{v_1} \cdots a_k^{v_k}b_j\), with \(\sum v_i = k\), is in \(c_R(f)^{k-1}C_R(fg)\).

Case 1. If \(v_n \neq 0\) and \(j = m\), then \(\theta = a_0^{v_0}a_1^{v_1} \cdots a_k^{v_k}b_m \in c_R(f)^{k-1}C_R(fg)\).

Case 2. If \(v_n \neq 0\) and \(j < m\), then \(b_j = b_me_j + b_1j\), where \(e_j\) is a coefficient of \(h\) and \(b_1j\) is a coefficient of \(g_1\), and \(\theta = a_0^{v_0}a_1^{v_1} \cdots a_k^{v_k}b_j = a_0^{v_0}a_1^{v_1} \cdots a_k^{v_k}(b_me_j + b_1j) \in c_R(f)^{k-1}C_R(fg) + c_R(f)^{k-1}a_nC_R(g_1)\).

Case 3. If \(v_n = 0\), then \(\theta = a_0^{v_0}a_1^{v_1} \cdots a_{n-1}^{v_{n-1}}b_j \in c_C(f_1)^kC_R(g) = c_R(f_1)^{k-1}C_R(f_1g)\) by the induction hypothesis on the degree of \(f\).

Combining the three cases we have

\[
c_R(f)^kC_R(g) \subseteq c_R(f)^{k-1}C_R(fg) + c_R(f)^{k-1}a_nC_R(g_1) + c_R(f_1)^{k-1}C_R(f_1g).
\]

Applying Claim 2 to the third term on the right, we see that this is contained in

\[
c_R(f)^{k-1}C_R(fg) + c_R(f)^{k-1}a_nC_R(g_1) + c_R(f_1)^{k-1}(c_R(fg) + a_nC_R(g_1)).
\]
Thus if \( a \)
b, and \( c \)
the product \( c_R(f_1)^{k-1}(c_R(fg) + a_n c_R(g_1)) \) is contained the other two terms. Thus
\[
(2.1) \quad c_R(f)^k c_R(g) \subseteq c_R(f)^{k-1} c_R(fg) + c_R(f)^{k-1} a_n c_R(g_1).
\]
Now since \( c_R(g_1) \) is generated by fewer than \( k \) elements, we have \( c_R(f)^{k-1} c_R(fg_1) = c_R(f)^{k-1} c_R(g_1) \) by induction on \( k \). Thus the right side of (2.1) is
\[
c_R(f)^{k-1} c_R(fg) + a_n c_R(f)^{k-2} c_R(fg_1).
\]
By Claim 1, this is contained in
\[
c_R(f)^{k-1} c_R(fg) + a_n c_R(f)^{k-2} (c_R(fg) + b_m c_R(f_1))
= c_R(f)^{k-1} c_R(fg) + a_n c_R(f)^{k-2} c_R(fg) + a_n b_m c_R(f)^{k-2} c_R(f_1).
\]
But \( a_n c_R(f)^{k-2} c_R(fg) \subseteq c_R(f)^{k-1} c_R(fg) \) and
\[
a_n b_m c_R(f)^{k-2} c_R(f_1) = (c_R(f)^{k-2} c_R(f_1))(a_n b_m) \subseteq c_R(f)^{k-1} c_R(fg).
\]
Thus \( c_R(f)^k c_R(g) = c_R(f)^{k-1} c_R(fg) \), showing that \( u \mu_R(g) \leq \mu_R(c_R(g)) \).

It remains to prove \( \mu_R(c_R(g)) \leq u \mu_R(g) \). This will follow from the next proposition.

**Proposition 2.3.** Let \( g \in R[X] \) have degree \( m \). If \( \mu_R(c_R(g)) > k \), and \( f \in R[T][X] \) has independent indeterminates for its coefficients, and \( \deg(f) = n > mk - k^2 \), then \( c_R(f)^{k-1} c_R(fg) \neq c_R(f)^k c_R(g) \). Thus \( u \mu_R(g) > k \).

**Proof.** Let \( \mu_R(c_R(g)) = k + j, j \geq 1 \). Then \( \mu_R(c_R(f)) = n + 1 \), and \( \mu_R(c_R(f)^k) = \binom{n+k}{n} \) = the number of monomials of degree \( k \) in \( n+1 \) variables. Also the inequality \( mk - k^2 < n \) is easily seen to be equivalent to \( \binom{n+k}{n} \) \( (m + n + 1) < \binom{n+k}{n} (k + 1) \).

Then
\[
\mu_R(c_R(f)^{k-1} c_R(fg)) \leq \binom{n+k}{n} (m + n + 1)
= \binom{n+k}{n} (k + 1) \leq \binom{n+k}{n} \mu_R(c_R(f)^k c_R(g)) = \mu_R(c_R(f)^k c_R(g)).
\]
Thus if \( \mu_R(c_R(g)) \geq k + 1 \), then \( u \mu_R(g) \geq k + 1 \).  

We state two interesting special cases of the above proposition.

(i) If \( \mu_R(c_R(g)) > 1 \) and \( f \in R[T][X] \) has independent indeterminates for its coefficients, and \( \deg(f) = n > m - 1 \), then \( c_R(fg) \neq c_R(f) c_R(g) \).

(ii) If \( \mu_R(c_R(g)) \geq m \) and \( f \in R[T][X] \) has independent indeterminates for its coefficients, and \( \deg(f) = n > m^2 - m^2 = 0 \), then \( c_R(f)^{m-1} c_R(fg) \neq c_R(f)^m c_R(g) \).

**Remark 2.4.** Let \( M \) be an \( R \)-module and let \( g \in M[X] \). The **universal Dedekind-Mertens number** \( u \mu_R(g) \) of \( g \) with respect to \( R \) is also the smallest positive integer \( k \) such that, for each commutative \( R \)-algebra \( S \) and each \( f \in S[X] \), it holds that
\[
c_R(f)^{k-1} c_R(fg) = c_R(f)^k c_R(g)
\]
as submodules of \( M \otimes_R S \).

Similarly the **universal polarized Dedekind-Mertens number** \( u \bar{\mu}_R(g) \) of \( g \) with respect to \( R \) is the smallest positive integer \( k \) such that for each commutative \( R \)-algebra \( S \).
and for all \( f_1, \ldots, f_k \in S[X] \) we have
\[
\sum_{i=1}^{k} c_R(f_i) c_R(f_1) c_R(f_2) \cdots \overline{c_R(f_i)} \cdots c_R(f_k) = c_R(f_1) c_R(f_2) \cdots c_R(f_k) c_R(g)
\]
as submodules of \( M \otimes_R S \).

**Proof.** Let \( k \) be the smallest positive integer such that for each commutative \( R \)-algebra \( S \) and each \( f \in S[X] \), it holds that
\[
c_R(f)^{k-1} c_R(fg) = c_R(f)^k c_R(g)
\]
as submodules of \( M \otimes_R S \).

By taking \( S = R[T] \) where \( T \) is a countably infinite set of indeterminates over \( R[X] \), we get \( k \geq m_R(g) \).

Also, if \( S \) is a commutative \( R \)-algebra and \( f = a_m X^m + \cdots + a_0 \in S[X] \), let \( h = t_m X^m + \cdots + t_0 \in R[T][X] \), where the \( t_i \in T \) are distinct indeterminates. Choose any \( R \)-algebra homomorphism \( \sigma : R[T][X] \to S[X] \) such that \( \sigma(X) = X \) and \( \sigma(t_i) = a_i \) for \( i = 0, \ldots, m \). Then \( \sigma \) induces a homomorphism \( M \otimes_R R[T][X] \to M \otimes_R S[X] \) which, for \( n = m_R(g) \), carries \( c_R(h)^{n-1} c_R(hg) = c_R(h)^n c_R(g) \) to \( c_R(f)^{n-1} c_R(fg) = c_R(f)^n c_R(g) \). Thus \( k \leq m_R(g) \).

\[ \square \]

**References**


Department of Mathematics, University of California, Riverside, California 92507
E-mail address: rush@math.ucr.edu