Survey on counting special types of polynomials

Joachim von zur Gathen & Konstantin Ziegler
B-IT, Universität Bonn
D-53113 Bonn, Germany
{gathen,zieglerk}@bit.uni-bonn.de
http://cosec.bit.uni-bonn.de/

July 14, 2014

Abstract

Most integers are composite and most univariate polynomials over a finite field are reducible. The Prime Number Theorem and a classical result of Gauß count the remaining ones, approximately and exactly.

For polynomials in two or more variables, the situation changes dramatically. Most multivariate polynomials are irreducible. This survey presents counting results for some special classes of multivariate polynomials over a finite field, namely the the reducible ones, the s-powerful ones (divisible by the s-th power of a nonconstant polynomial), the relatively irreducible ones (irreducible but reducible over an extension field), the decomposable ones, and also for reducible space curves. These come as exact formulas and as approximations with relative errors that essentially decrease exponentially in the input size.

Furthermore, a univariate polynomial f is decomposable if $f = g \circ h$ for some nonlinear polynomials g and h. It is intuitively clear that the decomposable polynomials form a small minority among all polynomials. The tame case, where the characteristic p of $\mathbb{F}_q$ does not divide $n = \deg f$, is fairly well-understood, and we obtain closely matching upper and lower bounds on the number of decomposable polynomials. In the wild case, where p does divide n, the bounds are less satisfactory, in particular when p is the smallest prime divisor of n and divides n exactly twice. The crux of the matter is to count the number of collisions, where essentially different (g, h) yield the same f. We present a classification of all collisions at degree $n = p^2$ which yields an exact count of those decomposable polynomials.

Keywords. counting special polynomials, finite fields, combinatorics on polynomials, generating functions, analytic combinatorics, asymptotic behavior, multivariate polynomials, polynomial decomposition, Ritt's Second Theorem

2010 Mathematics Subject Classification. 00B25, 11T06, 12Y05

1 Introduction

Most integers are composite and most univariate polynomials over a finite field are reducible. The classical results of the Prime Number Theorem and a theorem of Gauß present approximations saying that randomly chosen integers up to x or polynomials of degree up to n are prime or irreducible with probability about $1/\ln x$ or $1/n$, respectively.

Concerning special classes of univariate polynomials over a finite field, Zsigmondy (1894) counts those with a given number of distinct roots or without irreducible factors of a given degree. In the same situation, Artin (1924) counts the irreducible ones in an arithmetic progression and Hayes (1965) generalizes these results. Cohen (1969) and Car (1987) count polynomials with certain factorization patterns and Williams (1969) those with irreducible factors of given degree. Polynomials that occur as a norm in field extensions are studied by Gogia & Luthar (1981).

In two or more variables, the situation changes dramatically. Most multivariate polynomials are irreducible. Carlitz (1963) provides the first count of irreducible multivariate polynomials. In Carlitz (1965), he goes on to study the fraction of irreducibles when bounds on the degrees in each variable are prescribed; see also Cohen (1968). In this survey, we opt for bounding the total degree because it has the charm of being invariant under invertible linear transformations. Gao & Lauder (2002) consider the counting problem in yet another model, namely where one variable occurs with maximal degree. The natural generating function (or zeta function) for the irreducible polynomials in two or more variables does not converge anywhere outside of the origin. Wan (1992) notes that this explains the lack of a simple combinatorial formula for the number of irreducible polynomials. But he gives a p-adic formula, and also a (somewhat complicated) combinatorial formula. For further references, see Mullen & Panario (2013, Section 3.6).

In the bivariate case, von zur Gathen (2008) proves precise approximations with an exponentially decreasing relative error. Von zur Gathen, Viola & Ziegler (2013) extend those results to multivariate polynomials and give further information such as exact formulas and generating functions. Bodin (2008) gives a recursive formula for the number of irreducible bivariate polynomials and remarks on a generalization for more than two variables; he follows up with Bodin (2010).

We present exact formulas for the numbers of reducible (Sections 2.1-2.3), s-powerful (Section 2.4), and relatively irreducible polynomials (Section 2.5). The formulas also yield simple, yet precise, approximations to these numbers, with rapidly decaying relative errors.

Geometrically, a single polynomial corresponds to a hypersurface, that is, to a cycle in affine or projective space of codimension 1. This correspondence preserves the respective notions of reducibility. Thus, Sections 2.1-2.3 can also be viewed as counting reducible hypersurfaces, in particular, planar curves, and Section 2.4 those with an s-fold component. From a geometric perspective, these results say that almost all hypersurfaces are irreducible. Can we say something similar for other types of varieties? Cesaratto, von zur Gathen & Matera (2013) give an affirmative answer for curves in $\mathbb{P}^r$ for arbitrary r. A first question is how to parametrize the curves. Moduli spaces only include irreducible curves, and systems of defining equations do not work except for complete intersections. The natural parametrization is by the Chow variety $C_{r,n}$ of curves of degree n in $\mathbb{P}^r$, for some fixed r and n. The foundation of this approach is a result by Eisenbud & Harris (1992), who identified the irreducible components of $C_{r,n}$ of maximal dimension. We present the counting results in Section 2.6.

It is intuitively clear that the decomposable polynomials form a small minority among all multivariate polynomials over a field. Von zur Gathen (2011) gives a quantitative version of this intuition (see Section 2.7). The number of multivariate decomposable polynomials is also studied by Bodin, Dèbes & Najib (2009).

This concludes the first half (Section 2) of our survey, dealing with multivariate polynomials. The second half (Section 3) is devoted to counting univariate decomposable polynomials.

Some of the results in this survey are from joint work with Raoul Blankertz, Eda Cesaratto, Mark Giesbrecht, Guillermo Matera, and Alfredo Viola.

A version of this paper is to appear in Gutierrez, Schicho & Weimann (2014). The final publication will be available at Springer after publication.

2 Counting multivariate polynomials

We work in the polynomial ring $F[x_1, \ldots, x_r]$ in $r \geq 1$ variables over a field F and consider polynomials with total degree equal to some nonnegative integer n:

$$P_{r,n}^{\text{all}}(F) = \{f \in F[x_1, \ldots, x_r] \colon \deg f = n\}.$$

The polynomials of degree at most n form an F-vector space of dimension $\binom{r+n}{r}$.

The property of a certain polynomial to be reducible, squareful, relatively irreducible, or decomposable is shared with all polynomials associated to the given one. For counting them, it is sufficient to take one representative. We choose an arbitrary monomial order, say, the degree-lexicographic one, so that the monic polynomials are those with leading coefficient 1, and write

$$P_{r,n}(F) = \{f \in P^{\text{all}}_{r,n}(F) \colon f \text{ is monic}\}.$$

We use two different methodologies to obtain such bounds: generating functions and combinatorial counting. The usual approach, see Flajolet & Sedgewick (2009), of analytic combinatorics on series with integer coefficients leads, in our case, to power series that diverge everywhere (except at 0). We have not found a way to make this work. Instead, we use power series with symbolic coefficients, namely rational functions in a variable representing the field size. Several useful relations from standard analytic combinatorics carry over to this new scenario. In a first step, this yields in a straightforward manner an exact formula for the number under consideration (Theorem 2.5). This formula is, however, not very transparent. Even the leading term is not immediately visible.

In a second step, coefficient comparisons yield easy-to-use approximations to our number (Theorem 2.7). The relative error is exponentially decreasing in the bit size of the data. Thus, Theorem 2.7 gives a "third order" approximation for the number of reducible polynomials, and thus a "fourth order" approximation for the irreducible ones. The error term is in the big-Oh form and thus contains an unspecified constant.

In a third step, a different method, namely some combinatorial counting, yields "second order" approximations with explicit constants in the error term (Theorem 2.9).

The results of Sections 2.1-2.5 are from von zur Gathen, Viola & Ziegler (2013) unless otherwise attributed, those of Section 2.6 are from Cesaratto, von zur Gathen & Matera (2013), and those of Section 2.7 are from von zur Gathen (2011).

2.1 Exact formula for reducible polynomials

To study reducible polynomials, we consider the following subsets of $P_{r,n}(F)$:

$$\begin{aligned} I_{r,n}(F) &= \{f \in P_{r,n}(F) \colon f \text{ is irreducible}\}, \\ R_{r,n}(F) &= P_{r,n}(F) \setminus I_{r,n}(F). \end{aligned}$$

In the usual notions, the polynomial 1 is neither reducible nor irreducible. In our context, it is natural to have $R_{r,0}(F) = \{1\}$ and $I_{r,0}(F) = \varnothing$.

The sets of polynomials

$$
\begin{aligned}
\mathcal{P}_r &= \bigcup_{n\geq 0} P_{r,n}(\mathbb{F}_q), \\
\mathcal{I}_r &= \bigcup_{n\geq 0} I_{r,n}(\mathbb{F}_q), \\
\mathcal{R}_r &= \mathcal{P}_r \setminus \mathcal{I}_r,
\end{aligned}
$$

are combinatorial classes with the total degree as size functions and we denote the corresponding generating functions by $\mathrm{P}_r, \mathrm{I}_r, \mathrm{R}_r \in \mathbb{Z}_{\geq 0} [\![z]\!]$, respectively. Their coefficients are

$$
\mathrm{P}_{r,n} = \# P_{r,n}(\mathbb{F}_q) = q^{\binom{r+n}{r}-1} \frac{1 - q^{-\binom{r+n-1}{r-1}}}{1 - q^{-1}}, \tag{2.1}
$$

$$
\begin{aligned}
\mathrm{R}_{r,n} &= \# R_{r,n}(\mathbb{F}_q), \\
\mathrm{I}_{r,n} &= \# I_{r,n}(\mathbb{F}_q),
\end{aligned} \tag{2.2}
$$

respectively, dropping the finite field $\mathbb{F}_q$ with q elements from the notation. By definition, $\mathcal{P}_r$ equals the disjoint union of $\mathcal{R}_r$ and $\mathcal{I}_r$, and therefore

$$
\mathrm{R}_r = \mathrm{P}_r - \mathrm{I}_r.
$$

By unique factorization, every element in $\mathcal{P}_r$ corresponds to an unordered finite sequence of elements in $\mathcal{I}_r$, where repetition is allowed, and therefore

$$
\mathrm{I}_r = \sum_{k\geq 1} \frac{\mu(k)}{k} \log \mathrm{P}_r(z^k) \tag{2.3}
$$

by Flajolet & Sedgewick (2009, Theorem I.5), where μ is the number-theoretic Möbius-function. A resulting algorithm is easy to program and returns exact results with lightning speed.

This approach quickly leads to explicit formulas. A *composition* of a positive integer n is a sequence $j = (j_1, j_2, \ldots, j_{|j|})$ of positive integers $j_1, j_2, \ldots, j_{|j|}$ with $j_1 + j_2 + \cdots + j_{|j|} = n$, where $|j|$ denotes the length of the sequence. We define the set

$$
M_n = \{\text{compositions of } n\}. \tag{2.4}
$$

This standard combinatorial notion is not to be confused with the composition of polynomials, which we discuss in Sections 2.7 and 3.

Theorem 2.5 (Exact counting). *Let $r \geq 1$, $q \geq 2$, $\mathrm{P}_{r,n}$ as in (2.1), and $\mathrm{I}_{r,n}$ the number of irreducible monic r-variate polynomials of degree n over $\mathbb{F}_q$. Then we have*

$$\begin{aligned}
\mathrm{I}_{r,0} &= 0, \\
\mathrm{I}_{r,n} &= -\sum_{k \mid n} \frac{\mu(k)}{k} \sum_{j \in M_{n/k}} \frac{(-1)^{|j|}}{|j|} \mathrm{P}_{r,j_1} \mathrm{P}_{r,j_2} \cdots \mathrm{P}_{r,j_{|j|}},
\end{aligned}$$

for $n \geq 1$, and therefore for the number $\mathrm{R}_{r,n}$ of reducible monic r-variate polynomials of degree n over $\mathbb{F}_q$

$$\begin{aligned}
\mathrm{R}_{r,0} &= 1, \\
\mathrm{R}_{r,n} &= \mathrm{P}_{r,n} + \sum_{k \mid n} \frac{\mu(k)}{k} \sum_{j \in M_{n/k}} \frac{(-1)^{|j|}}{|j|} \mathrm{P}_{r,j_1} \mathrm{P}_{r,j_2} \cdots \mathrm{P}_{r,j_{|j|}},
\end{aligned}$$

for $n \geq 1$.

The formula of Theorem 2.5 is exact but somewhat cumbersome. The following two sections provide simple yet precise approximations, with rapidly decaying error terms.

2.2 Symbolic approximation for reducible polynomials

For $r \geq 2$, the power series P_r, I_r, and R_r do not converge anywhere except at 0, and the standard asymptotic arguments of analytic combinatorics are inapplicable. We now deviate from this approach and move from power series in $\mathbb{Q}[\![z]\!]$ to power series in $\mathbb{Q}(\mathbf{q})[\![z]\!]$, where $\mathbf{q}$ is a symbolic variable representing the field size. For $r \geq 2$ and $n \geq 0$ we let

$$\mathsf{P}_{r,n}(\mathbf{q}) = \mathbf{q}^{\binom{r+n}{r}-1} \frac{1 - \mathbf{q}^{-\binom{r+n-1}{r-1}}}{1 - \mathbf{q}^{-1}} \in \mathbb{Z}[\mathbf{q}]$$

in analogy to (2.1). We define the power series $\mathsf{P}_r, \mathsf{I}_r, \mathsf{R}_r \in \mathbb{Q}(\mathbf{q})[\![z]\!]$ by

$$\begin{aligned}
\mathsf{P}_r(\mathbf{q}, z) &= \sum_{n \geq 0} \mathsf{P}_{r,n}(\mathbf{q}) z^n, \qquad (2.6) \\
\mathsf{I}_r(\mathbf{q}, z) &= \sum_{k \geq 1} \frac{\mu(k)}{k} \log \mathsf{P}_r(\mathbf{q}, z^k), \\
\mathsf{R}_r(\mathbf{q}, z) &= \mathsf{P}_r(\mathbf{q}, z) - \mathsf{I}_r(\mathbf{q}, z).
\end{aligned}$$

Then $\mathsf{R}_{r,n}(\mathbf{q})$ denotes the coefficient of z^n in R_r and counts symbolically the reducible monic r-variate polynomials of degree n.

For nonzero $f \in \mathbb{Q}(\mathbf{q})$, $\deg_{\mathbf{q}} f$ is the degree of f, that is, the numerator degree minus the denominator degree. The appearance of $O(\mathbf{q}^{-m})$ with a positive integer m in an equation means the existence of some f with degree at most $-m$ that makes the equation valid. If a term $O(\mathbf{q}^{-m})$ appears, then we may conclude a numerical asymptotic result for growing prime powers q.

Theorem 2.7 (Symbolic approximation)**.** *Let $r \geq 2$ and*

$$\rho_{r,n}(\mathbf{q}) = \mathbf{q}^{\binom{r+n-1}{r}+r-1} \frac{1-\mathbf{q}^{-r}}{(1-\mathbf{q}^{-1})^2} \in \mathbb{Q}(\mathbf{q}).$$

Then the symbolic formula $\mathsf{R}_{r,n}(\mathbf{q})$ for the number of reducible monic r-variate polynomials of degree n over $\mathbb{F}_q$ satisfies

$$\mathsf{R}_{r,0}(\mathbf{q}) = 1, \quad \mathsf{R}_{r,1}(\mathbf{q}) = 0, \quad \mathsf{R}_{r,2}(\mathbf{q}) = \frac{\rho_{r,2}(\mathbf{q})}{2} \cdot (1-\mathbf{q}^{-r-1}),$$

$$\mathsf{R}_{r,3}(\mathbf{q}) = \rho_{r,3}(\mathbf{q})\Big(1-\mathbf{q}^{-r(r+1)/2} + \mathbf{q}^{-r(r-1)/2} \frac{1-2\mathbf{q}^{-r}+2\mathbf{q}^{-2r-1}-\mathbf{q}^{-2r-2}}{3(1-\mathbf{q}^{-1})}\Big),$$

$$\mathsf{R}_{r,4}(\mathbf{q}) = \rho_{r,4}(\mathbf{q}) \cdot \Big(1+\mathbf{q}^{-\binom{r+1}{3}} \cdot \frac{1+O(\mathbf{q}^{-r(r-1)/2})}{2(1-\mathbf{q}^{-r})}\Big),$$

and for $n \geq 5$

$$\mathsf{R}_{r,n}(\mathbf{q}) = \rho_{r,n}(\mathbf{q}) \cdot \Big(1+\mathbf{q}^{-\binom{r+n-2}{r-1}+r(r+1)/2} \cdot \frac{1+O(\mathbf{q}^{-r(r-1)/2})}{1-\mathbf{q}^{-r}}\Big). \tag{2.8}$$

Alekseyev (2006) lists $(\#I_{r,n}(\mathbb{F}_q))_{n\geq 0}$ as A115457–A115472 in The On-Line Encyclopedia of Integer Sequences, for $2 \leq r \leq 6$ and prime $q \leq 7$. Bodin (2008, Theorem 7) states (in our notation)

$$1-\frac{\#I_{r,n}}{\#P_{r,n}} \sim q^{-\binom{n+r-1}{r-1}-r} \frac{1-q^{-r}}{1-q^{-1}}.$$

Hou & Mullen (2009) provide results for $\#I_{r,n}(\mathbb{F}_q)$. These do not yield error bounds for the approximation of $\#R_{r,n}(\mathbb{F}_q)$. Bodin (2010) also uses (2.3) to claim a result similar to (2.8).

2.3 Explicit bounds for reducible polynomials

The third approach by "combinatorial counting" is somewhat more involved. The payoff of this additional effort is an explicit relative error bound. However, the calculations are sufficiently complicated for us to stop at the first error term. Thus we replace the asymptotic $1+O(\mathbf{q}^{-r(r-1)/2})$ in (2.8) by $1/(1-q^{-1})$.

Theorem 2.9 (Explicit approximation)**.** *Let $r, q \geq 2$, and $\rho_{r,n}$ as in Theorem 2.7. For the number $\#R_{r,n}(\mathbb{F}_q)$ of reducible monic r-variate polynomials of degree n over $\mathbb{F}_q$ we have*

$$\#R_{r,0}(\mathbb{F}_q) = 1, \quad \#R_{r,1}(\mathbb{F}_q) = 0, \quad \#R_{r,2}(\mathbb{F}_q) = \frac{\rho_{r,2}(q)}{2} \cdot (1 - q^{-r-1}),$$
$$|\#R_{r,3}(\mathbb{F}_q) - \rho_{r,3}(q)| = \rho_{r,3}(q) \cdot q^{-r(r-1)/2} \frac{1 - 2q^{-r} + 2q^{-2r-1} - q^{-2r-2}}{3(1-q^{-1})}$$
$$\leq \rho_{r,3}(q) \cdot q^{-r(r-1)/2},$$

and for $n \geq 4$

$$|\#R_{r,n}(\mathbb{F}_q) - \rho_{r,n}(q)| \leq \rho_{r,n}(q) \cdot \frac{q^{-\binom{r+n-2}{r-1}+r(r+1)/2}}{(1-q^{-1})(1-q^{-r})}$$
$$\leq \rho_{r,n}(q) \cdot 3q^{-\binom{r+n-2}{r-1}+r(r+1)/2}.$$

Remark 2.10. How close is our relative error estimate to being exponentially decaying in the input size? The usual dense representation of a polynomial in r variables and of degree n requires $b_{r,n} = \binom{r+n}{r}$ monomials, each of them equipped with a coefficient from $\mathbb{F}_q$, using about $\log_2 q$ bits. Thus the total input size is about $\log_2 q \cdot b_{r,n}$ bits. This differs from $\log_2 q \cdot (b_{r-1,n-1} - b_{r-1,2})$ by a factor of

$$\frac{b_{r,n}}{b_{r-1,n-1} - b_{r-1,2}} < \frac{b_{r,n}}{\frac{1}{2} b_{r-1,n-1}} = \frac{2(n+r)(n+r-1)}{nr}.$$

Up to this polynomial difference (in the exponent), the relative error is exponentially decaying in the bit size of the input, that is, $(\log q)$ times the number of coefficients in the usual dense representation. In particular, it is exponentially decaying in any of the parameters r, n, and $\log_2 q$, when the other two are fixed.

2.4 Powerful polynomials

For an integer $s \geq 2$, a polynomial is called *s-powerful* if it is divisible by the sth power of some nonconstant polynomial, and *s-powerfree* otherwise; it is *squarefree* if $s = 2$. Let

$$Q_{r,n,s}(F) = \{f \in P_{r,n}(F) \colon f \text{ is } s\text{-powerful}\},$$
$$S_{r,n,s}(F) = P_{r,n}(F) \setminus Q_{r,n,s}(F).$$

As in the previous section, we restrict our attention to a finite field $F = \mathbb{F}_q$, which we omit from the notation.

For the approach by generating functions, we consider the combinatorial classes $\mathcal{Q}_{r,s} = \bigcup_{n \geq 0} Q_{r,n,s}$ and $\mathcal{S}_{r,s} = \mathcal{P}_r \setminus \mathcal{Q}_{r,s}$. Any monic polynomial f factors uniquely as $f = g \cdot h^s$ where g is a monic s-powerfree polynomial and h an arbitrary monic polynomial, hence

$$\mathrm{P}_r = \mathrm{S}_{r,s} \cdot \mathrm{P}_r(z^s) \tag{2.11}$$

and by definition $\mathrm{Q}_{r,s} = \mathrm{P}_r - \mathrm{S}_{r,s}$ for the generating functions of $\mathcal{S}_{r,s}$ and $\mathcal{Q}_{r,s}$, respectively. For univariate polynomials, Carlitz (1932) derives (2.11) directly from generating functions to prove the counting formula (2.13) for $r = 1$. Flajolet, Gourdon & Panario (2001, Section 1.1) use (2.11) for $s = 2$ to count univariate squarefree polynomials, see also Flajolet & Sedgewick (2009, Note I.66).

As in Theorem 2.5, this approach quickly leads to explicit formulas.

Theorem 2.12 (Exact counting)**.** *For $r \geq 1$, $q, s \geq 2$, $\mathrm{P}_{r,n}$ as in (2.1), and M_n as in (2.4), we have for the number $\mathrm{Q}_{r,n,s} = \#Q_{r,n,s}(\mathbb{F}_q)$ of s-powerful monic r-variate polynomials of degree n over $\mathbb{F}_q$*

$$\mathrm{Q}_{r,n,s} = - \sum_{\substack{1 \leq i \leq n/s \\ j \in M_i}} (-1)^{|j|} \mathrm{P}_{r,j_1} \mathrm{P}_{r,j_2} \cdots \mathrm{P}_{r,j_{|j|}} \mathrm{P}_{r,n-is}. \tag{2.13}$$

To study the asymptotic behavior of $\mathrm{Q}_{r,n,s}$ for $r \geq 2$ we again deviate from the standard approach and move to power series in $\mathbb{Q}(\mathbf{q})[\![z]\!]$. With P_r from (2.6), we define $\mathsf{S}_{r,s}, \mathsf{Q}_{r,s} \in \mathbb{Q}(\mathbf{q})[\![z]\!]$ by

$$\begin{aligned} \mathsf{P}_r &= \mathsf{S}_{r,s} \cdot \mathsf{P}_r(z^s), \\ \mathsf{Q}_{r,s} &= \mathsf{P}_r - \mathsf{S}_{r,s}. \end{aligned}$$

The approach by generating functions now yields the following result. Its "general" case is (iv). We give exact expressions in special cases, namely for $n < 3s$ in (ii) and for $(n, s) = (6, 2)$ in (iii), which also apply when we substitute the size q of a finite field $\mathbb{F}_q$ for $\mathbf{q}$.

Theorem 2.14 (Symbolic approximation)**.** *Let $r, s \geq 2$, $n \geq 0$, and*

$$\begin{aligned} \eta_{r,n,s}(\mathbf{q}) &= \mathbf{q}^{\binom{r+n-s}{r}+r-1} \frac{(1-\mathbf{q}^{-r})(1-\mathbf{q}^{-\binom{r+n-s-1}{r-1}})}{(1-\mathbf{q}^{-1})^2} \in \mathbb{Q}(\mathbf{q}), \\ \delta &= \binom{r+n-s}{r} - \binom{r+n-2s}{r} - \frac{r(r+1)}{2}. \end{aligned}$$

Then the symbolic formula $\mathsf{Q}_{r,n,s}(\mathbf{q})$ for the number of s-powerful monic r-variate polynomials of degree n over $\mathbb{F}_q$ satisfies the following.

(i) If $n \geq 2s$, then $\delta \geq r$.

(ii)

$$\mathsf{Q}_{r,n,s}(\mathbf{q}) = \begin{cases} 0 & \text{for } n < s, \\ \eta_{r,n,s}(\mathbf{q}) & \text{for } s \leq n < 2s, \\ \eta_{r,n,s}(\mathbf{q})\Big(1 + \mathbf{q}^{-\delta} \cdot \frac{1-\mathbf{q}^{-\binom{n+r-2s-1}{r-1}}}{1-\mathbf{q}^{-\binom{n+r-s-1}{r-1}}} & \\ \quad \cdot \Big(\frac{1-\mathbf{q}^{-r(r+1)/2}}{1-\mathbf{q}^{-r}} - \mathbf{q}^{-r(r-1)/2}\frac{1-\mathbf{q}^{-r}}{1-\mathbf{q}^{-1}}\Big)\Big) & \text{for } 2s \leq n < 3s. \end{cases} \tag{2.15}$$

(iii) For $(n, s) = (6, 2)$, we have

$$\mathsf{Q}_{r,6,2}(\mathbf{q}) = \eta_{r,6,2}(\mathbf{q})\Big(1 + \mathbf{q}^{-\delta+(r-2)(r-1)(r+3)/6}(1 + O(\mathbf{q}^{-1}))\Big). \tag{2.16}$$

(iv) For $n \geq 2s$ and $(n, s) \neq (6, 2)$, we have

$$\mathsf{Q}_{r,n,s}(\mathbf{q}) = \eta_{r,n,s}(\mathbf{q})\Big(1 + \mathbf{q}^{-\delta}(1 + O(\mathbf{q}^{-1}))\Big).$$

For $r \geq 3$, we can replace $1 + O(\mathbf{q}^{-1})$ in (2.16) by $\mathbf{q}^{-1} + O(\mathbf{q}^{-2})$. The combinatorial approach replaces the asymptotic $1 + O(\mathbf{q}^{-1})$ for $n \geq 3s$ with an explicit bound. For $n < 3s$ the exact formula (2.15) of Theorem 2.14 (ii) applies.

Theorem 2.17 (Explicit approximation)**.** *Let $r, s, q \geq 2$, $\#Q_{r,n,s}(\mathbb{F}_q)$ the number of s-powerful monic r-variate polynomials of degree n over $\mathbb{F}_q$, and $\eta_{r,n,s}$ and δ as in Theorem 2.14.*

(i) For $(n, s) = (6, 2)$, we have $\delta = r(r + 1)(r^2 + 9r + 2)/24$ and

$$|\#Q_{r,6,2}(\mathbb{F}_q) - \eta_{r,6,2}(q)| \leq \eta_{r,6,2}(q) \cdot 2q^{-\delta+(r-2)(r-1)(r+3)/6}.$$

(ii) For $n \geq 3s$ and $(n, s) \neq (6, 2)$, we have

$$|\#Q_{r,n,s}(\mathbb{F}_q) - \eta_{r,n,s}(q)| \leq \eta_{r,n,s}(q) \cdot 6q^{-\delta}.$$

As noted in Remark 2.10 for reducible polynomials, the relative error term is (essentially) exponentially decreasing in the input size, and exponentially decaying in any of the parameters r, n, s, and $\log_2 q$, when the other three are fixed.

2.5 Relatively irreducible polynomials

A polynomial over F is *absolutely irreducible* if it is irreducible over an algebraic closure of F, and *relatively irreducible* (or *exceptional*) if it is irreducible over F but factors over some extension field of F. We define

$$\begin{aligned}
A_{r,n}(F) &= \{f \in P_{r,n}(F) \colon f \text{ is absolutely irreducible}\} \subseteq I_{r,n}(F),\\
E_{r,n}(F) &= I_{r,n}(F) \setminus A_{r,n}(F).
\end{aligned}$$

As before, we restrict ourselves to finite fields and recall that all our polynomials are monic. We relate the generating function $\mathrm{A}_r(\mathbb{F}_q)$ of $\#A_{r,n}(\mathbb{F}_q)$ to the generating function $\mathrm{I}_r(\mathbb{F}_q)$ of irreducible polynomials as introduced in Section 2.1 and obtain

$$\begin{aligned}
[z^n]\,\mathrm{I}_r(\mathbb{F}_q) &= \sum_{k\,|\,n} \frac{1}{k} \sum_{s\,|\,k} \mu(k/s) \cdot [z^{n/k}]\,\mathrm{A}_r(\mathbb{F}_{q^s}),\\
[z^n]\,\mathrm{A}_r(\mathbb{F}_q) &= \sum_{k\,|\,n} \frac{1}{k} \sum_{s\,|\,k} \mu(s) \cdot [z^{n/k}]\,\mathrm{I}_r(\mathbb{F}_{q^s})
\end{aligned} \tag{2.18}$$

with Möbius inversion. For an explicit formula, we combine the expression for $\mathrm{I}_{r,n}(\mathbb{F}_q)$ from Theorem 2.5 with (2.18).

Theorem 2.19 (Exact counting)**.** *For $r, n \geq 1$, $q \geq 2$, M_n as in* (2.4)*, $\mathrm{P}_{r,n}$ as in* (2.1)*, and $\mathrm{I}_{r,n}$ as in* (2.2)*, we have for the number $\mathrm{E}_{r,n}$ of relatively irreducible monic r-variate polynomials of degree n over $\mathbb{F}_q$*

$$\begin{aligned}
\mathrm{E}_{r,0}(\mathbb{F}_q) &= 0,\\
\mathrm{E}_{r,n}(\mathbb{F}_q) &= -\sum_{1<k\,|\,n} \frac{1}{k} \sum_{s\,|\,k} \mu(s)\mathrm{I}_{r,n/k}(\mathbb{F}_{q^s})\\
&= \sum_{1<k\,|\,n} \frac{1}{k} \sum_{\substack{s\,|\,k\\ m\,|\,n/k}} \frac{\mu(s)\mu(m)}{m}\\
&\qquad \cdot \sum_{j\in M_{n/(km)}} \frac{(-1)^{|j|}}{|j|} \mathrm{P}_{r,j_1}(\mathbb{F}_{q^s})\mathrm{P}_{r,j_2}(\mathbb{F}_{q^s}) \cdots \mathrm{P}_{r,j_{|j|}}(\mathbb{F}_{q^s}).
\end{aligned}$$

The approach by generating functions gives the following result.

Theorem 2.20 (Symbolic approximation)**.** *Let $r, n \geq 2$, let ℓ be the smallest prime divisor of n, and*

$$\epsilon_{r,n}(\mathbf{q}) = \frac{\mathbf{q}^{\ell(\binom{r+n/\ell}{r}-1)}}{\ell(1-\mathbf{q}^{-\ell})} \in \mathbb{Q}(\mathbf{q}),$$

$$\kappa = (\ell - 1)(\binom{r - 1 + n/\ell}{r - 1} - r) + 1.$$

Then the symbolic formula $\mathsf{E}_{r,n}(\mathbf{q})$ for the number of relatively irreducible monic r-variate polynomials of degree n over $\mathbb{F}_q$ satisfies the following.

(i) $\mathsf{E}_{r,1}(\mathbf{q}) = 0$.

(ii) If n is prime, then

$$\mathsf{E}_{r,n}(\mathbf{q}) = \epsilon_{r,n}(\mathbf{q})(1 - \mathbf{q}^{-nr})\Big(1 - \mathbf{q}^{-r(n-1)} \frac{(1 - \mathbf{q}^{-r})(1 - \mathbf{q}^{-n})}{(1 - \mathbf{q}^{-1})(1 - \mathbf{q}^{-nr})}\Big).$$

(iii) If n is composite, then $\kappa \geq 2$ and

$$\mathsf{E}_{r,n}(\mathbf{q}) = \epsilon_{r,n}(\mathbf{q})(1 + O(\mathbf{q}^{-\kappa})).$$

While (i) and (ii) yield explicit bounds, the combinatorial approach does this for (iii).

Theorem 2.21 (Explicit approximation). *Let $r, q \geq 2$, and $\epsilon_{r,n}$ and κ as in Theorem 2.20, and n be composite. Then for the number $\#E_{r,n}(\mathbb{F}_q)$ of relatively irreducible monic r-variate polynomials of degree n over $\mathbb{F}_q$ we have*

$$|\#E_{r,n}(\mathbb{F}_q) - \epsilon_{r,n}(q)| \leq \epsilon_{r,n}(q) \cdot 3q^{-\kappa}.$$

2.6 Reducible space curves

The *Chow variety* of curves of degree n in the r-dimensional projective space $\mathbb{P}^r = \mathbb{P}^r(\overline{\mathbb{F}_q})$ over an algebraic closure $\overline{\mathbb{F}_q}$ is denoted by $C_{r,n}$. Each point of the Chow variety $C_{r,n}$ actually corresponds to a unique *effective cycle* in $\mathbb{P}^r$ of dimension 1 and degree n, that is, to a formal linear combination $\sum a_i C_i$, where each C_i is an irreducible curve in $\mathbb{P}^r$, each a_i is a positive integer and $\sum a_i \deg(C_i) = n$.

For a subfield $F \subseteq \overline{\mathbb{F}_q}$, an effective F-cycle C is called *F-reducible* if there exist $m \geq 2$ and effective F-cycles $C_1, \ldots, C_m$ such that $C = \sum_{i=1}^{m} C_i$ holds. Let $C_{r,n}(\mathbb{F}_q)$ denote the Chow variety of effective $\mathbb{F}_q$-cycles and $R^*_{r,n}(\mathbb{F}_q)$ its closed subvariety of $\mathbb{F}_q$-reducible $\mathbb{F}_q$-cycles. Methods of algebraic geometry yield the following bounds on the probability that a random curve of degree n in $\mathbb{P}^r(\mathbb{F}_q)$ is $\mathbb{F}_q$-reducible.

Theorem 2.22. *Let $r \geq 3$ and*

$$g_{r,n} = \binom{r+n-2}{n}^2 \cdot \frac{r+n-1}{(r-1)(n+1)},$$
$$c_{r,n} = (2en)^{r(r+1)(n^2+1)+4rg_{r,n}},$$

where e denotes the basis of the natural logarithm. For the number $\#R^*_{r,n}(\mathbb{F}_q)$ of $\mathbb{F}_q$-reducible cycles of degree n we have the following.

(i) If $n \geq \min\{4r-7, 7\}$, then

$$\frac{1}{4c_{r,n}} q^{-(n-2r+3)} \leq \frac{\#R^*_{r,n}(\mathbb{F}_q)}{\#C_{r,n}(\mathbb{F}_q)} \leq c_{r,n} q^{-(n-2r+3)}.$$

(ii) If $n = 4r-8$, then

$$\frac{1}{2n!\, c_{r,n}} q^{-r+2} \leq \frac{\#R^*_{r,n}(\mathbb{F}_q)}{\#C_{r,n}(\mathbb{F}_q)} \leq c_{r,n} q^{-r+2}.$$

We call an $\overline{\mathbb{F}_q}$-reducible cycle *absolutely reducible*. An $\mathbb{F}_q$-cycle can be absolutely reducible for two reasons: either it is $\mathbb{F}_q$-reducible, as treated above, or *relatively $\mathbb{F}_q$-irreducible*, that is, is $\mathbb{F}_q$-irreducible and $\overline{\mathbb{F}_q}$-reducible. The set of relatively $\mathbb{F}_q$-irreducible (or *exceptional*) $\mathbb{F}_q$-curves of degree n in $\mathbb{P}^r$ is denoted by $E^*_{r,n}(\mathbb{F}_q)$.

Theorem 2.23. *Let $r \geq 3$, $n \geq 4r-8$, let ℓ denote the smallest prime divisor of n, and*

$$b_{r,n} = 3(r-2) + n(n+3)/2,$$
$$d_{\ell,n,r} = (en/\ell)^{r(r+1)(n^2/\ell^2+1)+4rg_{r,n/\ell}}.$$

For the number $\#E^*_{r,n}(\mathbb{F}_q)$ of relatively $\mathbb{F}_q$-irreducible cycles of degree n we have

$$q^{2n(r-1)}(1-4q^{2(1-n)(r-1)}) \leq \#E^*_{r,n}(\mathbb{F}_q) \leq 2d_{\ell,n,r} q^{2n(r-1)} \text{ for } n/\ell \leq 4r-7,$$
$$q^{\ell b_{r,n/\ell}}(1-16q^{\ell-n}) \leq \#E^*_{r,n}(\mathbb{F}_q) \leq 3d_{\ell,n,r} q^{\ell b_{r,n/\ell}} \text{ for } n/\ell \geq 4r-8.$$

2.7 Decomposable polynomials

For monic univariate $g \in F[y]$ and $h \in P_{r,n}$, we define their *composition*

$$f = g \circ h = g(h) \in P_{r,n}.$$

If $\deg g \geq 2$ and $\deg h \geq 1$, then (g, h) is a *decomposition* of f. A polynomial $f \in P_{r,n}$ is *decomposable* if there exist such g and h. There are other notions of decompositions. The present one is called uni-multivariate in von zur Gathen, Gutierrez & Rubio (2003). Another one is studied in Faugère & Perret (2008) for cryptanalytic purposes. In the context of univariate polynomials $\deg h \geq 2$ is also required, see Section 3.

It is sufficient to concentrate on polynomials with vanishing constant term, see subsection 3.1, and we denote by $D_{r,n}(F)$ the set of all decomposable polynomials $f \in P_{r,n}(F)$ with $f(0, \ldots, 0) = 0$.

Theorem 2.24. *Let $\mathbb{F}_q$ be a finite field with q elements, $r \geq 2$, and ℓ the smallest prime divisor of the composite integer $n \geq 2$. Let*

$$m = \begin{cases} n & \text{if } r = 2,\ n/\ell \text{ is prime, and } n/\ell \leq 2\ell - 5, \\ \ell & \text{otherwise,} \end{cases}$$

$$\alpha_{r,n} = q^{\binom{r+n/m}{r}+m-3} \frac{1 - q^{-\binom{r-1+n/m}{r-1}}}{1 - q^{-1}},$$

$$\beta_{r,n} = \frac{2q^{-\frac{1}{2}\binom{r-1+n/\ell}{r-1}+1}}{1 - q^{-1}}.$$

Then for the number $\#D_{r,n}(\mathbb{F}_q)$ of decomposable monic r-variate polynomials with vanishing constant term of degree n over $\mathbb{F}_q$ we have

$$|\#D_{r,n}(\mathbb{F}_q) - \alpha_{r,n}| \leq \alpha_{r,n} \cdot \beta_{r,n}.$$

3 Counting univariate decomposable polynomials

The *composition* of two univariate polynomials $g, h \in F[x]$ over a field F is denoted as $f = g \circ h = g(h)$, and then (g, h) is a *decomposition* of f, and f is *decomposable* if g and h have degree at least 2. In the 1920s, Ritt, Fatou, and Julia studied structural properties of these decompositions over $\mathbb{C}$, using analytic methods. Particularly important are two theorems by Ritt on the uniqueness, in a suitable sense, of decompositions, the first one for (many) indecomposable components and the second one for two components, as above. Engstrom (1941) and Levi (1942) proved them over arbitrary fields of characteristic zero using algebraic methods.

The theory was extended to arbitrary characteristic by Fried & MacRae (1969), Dorey & Whaples (1974), Schinzel (1982, 2000), Zannier (1993), and others. Its use in a cryptographic context was suggested by Cade (1985). In computer algebra, the decomposition method of Barton & Zippel (1985) requires exponential time. A fundamental dichotomy is between the *tame case*, where the characteristic p does not divide $\deg g$, and the *wild case*, where p divides $\deg g$, see von zur Gathen (1990a,b). (Schinzel (2000), § 1.5, uses *tame* in a different sense.) A breakthrough result of Kozen & Landau (1989) was their polynomial-time algorithm to compute tame decompositions; see also von zur Gathen, Kozen & Landau (1987); Kozen, Landau & Zippel (1996); Gutierrez & Sevilla (2006), and the survey articles of von zur Gathen (2002) and Gutierrez & Kozen (2003) with further references. Schur's conjecture, as proven by Turnwald (1995), offers a natural connection between the tame indecomposable polynomials in this section and certain absolutely irreducible bivariate polynomials, as studied in Section 2.5. More precisely, a tame polynomial f is indecomposable if $(f(x) - f(y))/(x - y)$ is absolutely irreducible. Aside from natural exceptions, the converse is also true.

In the wild case, considerably less is known, both mathematically and computationally. Zippel (1991) suggests that the block decompositions of Landau & Miller (1985) for determining subfields of algebraic number fields can be applied to decomposing rational functions even in the wild case. A version of Zippel's algorithm in Blankertz (2014) computes in polynomial time all decompositions of a polynomial that are minimal in a certain sense. Avanzi & Zannier (2003) study ambiguities in the decomposition of rational functions over $\mathbb{C}$. On a different but related topic, Zieve & Müller (2008) found interesting characterizations for Ritt's First Theorem, which deals with complete decompositions, where all components are indecomposable.

We have seen fairly precise estimates for the number of multivariate decomposable polynomials in Section 2.7. It is intuitively clear that the univariate decomposable polynomials also form only a small minority among all univariate polynomials over a field, and this second part of our survey confirms this intuition. The task is to approximate the number of decomposables over a finite field, together with a good relative error bound. One readily obtains an upper bound. The challenge then is to find an essentially matching lower bound.

A set of distinct decompositions of f is called a *collision*. The number of decomposable polynomials of degree n is thus the number of all pairs (g, h) with $\deg g \cdot \deg h = n$ reduced by the ambiguities introduced by collisions. An important tool for estimating the number of collisions is Ritt's Second Theorem. The first algebraic versions of this in positive characteristic p required $p > \deg(g \circ h)$. Zannier (1993) reduced this to the milder and more natural requirement $g' \neq 0$ for all g in the collision. His proof works over an algebraic closed field, and Schinzel's 2000 monograph adapts it to finite fields.

In Section 3.2, we provide a precise quantitative version of Ritt's Second Theorem, by determining exactly the number of such collisions in the tame case, assuming that $p \nmid n/\ell$, where n is the degree of the composition and ℓ is the smallest prime divisor of n. This is based on a unique normal form for the polynomials occurring in Ritt's Second Theorem.

Giesbrecht (1988) was the first to consider this counting problem. He showed that the decomposable polynomials form an exponentially small fraction of all univariate polynomials. General approximations to the number of univariate decomposable polynomials are shown in Section 3.3. They come with satisfactory (rapidly decreasing) relative error bounds except when p divides $n = \deg f$ exactly twice. Ziegler (2014) provides an exact count of tame univariate polynomials. In Section 3.4, we determine exactly the number of decomposable polynomials in one of the difficult wild cases, namely when $n = p^2$.

Zannier (2008) studies a different but related question, namely compositions $f = g \circ h$ in $\mathbb{C}[x]$ with a *sparse* polynomial f, having t terms. The degree is not bounded. He gives bounds, depending only on t, on the degree of g and the number of terms in h. Furthermore, he gives a parametrization of all such f, g, h in terms of varieties (for the coefficients) and lattices (for the exponents). Bodin, Dèbes & Najib (2009) also deal with counting.

Unless otherwise attributed, the results of Section 3.2 are from von zur Gathen (2014b), those of Section 3.3 from von zur Gathen (2014a), and those of Section 3.4 from Blankertz, von zur Gathen & Ziegler (2013).

3.1 Notation

A nonzero polynomial $f \in F[x]$ over a field F of characteristic $p \geq 0$ is *monic* if its leading coefficient $\operatorname{lc}(f)$ equals 1. We call f *original* if its graph contains the origin, that is, $f(0) = 0$. For $g, h \in F[x]$,

$$f = g \circ h = g(h) \in F[x] \tag{3.1}$$

is their *composition*. If $\deg g, \deg h \geq 2$, then (g, h) is a *decomposition* of f. A polynomial $f \in F[x]$ of degree at least 2 is *decomposable* if there exist such g and h, otherwise f is *indecomposable*. A decomposition (3.1) is *tame* if $p \nmid \deg g$, and f is *tame* if $p \nmid \deg f$.

Multiplication by a unit or addition of a constant does not change decomposability, since

$$f = g \circ h \Longleftrightarrow af + b = (ag + b) \circ h$$

for all f, g, h as above and $a, b \in F$ with $a \neq 0$. In other words, the set of decomposable polynomials is invariant under this action of $F^\times \times F$ on $F[x]$.

Furthermore, any decomposition (g, h) can be normalized by this action, by taking $a = \mathrm{lc}(h)^{-1} \in F^\times$, $b = -a \cdot h(0) \in F$, $g^* = g((x-b)a^{-1}) \in F[x]$, and $h^* = ah + b$. Then $g \circ h = g^* \circ h^*$ and g^* and h^* are monic original.

It is therefore sufficient to consider compositions $f = g \circ h$ where all three polynomials are monic original. In such a tame decomposition, g and h are uniquely determined by f and $\deg g$. For $n \geq 1$ and any proper divisor e of n, we write

$$\begin{aligned} P_n(F) &= \{f \in F[x] \colon f \text{ is monic original of degree } n\}, \\ D_n(F) &= \{f \in P_n(F) \colon f \text{ is decomposable}\}, \\ D_{n,e}(F) &= \{f \in P_n(F) \colon f = g \circ h \text{ for some } (g, h) \in P_e(F) \times P_{n/e}(F)\}. \end{aligned}$$

Thus $P_n(F)$ and $D_n(F)$ are the subsets of original polynomials in the sets $P_{1,n}(F)$ and $D_{1,n}(F)$, respectively, as defined in the context of multivariate polynomials (subsection 2.7) but with right component h of degree at least 2. We sometimes leave out F from the notation when it is clear from the context and have over a finite field $\mathbb{F}_q$ with q elements

$$\begin{aligned} \#P_n &= q^{n-1}, \\ \#D_{n,e} &\leq q^{e+n/e-2}. \end{aligned}$$

The set D_n of all decomposable polynomials in P_n satisfies

$$D_n = \bigcup_{\substack{e \mid n \\ 1<e<n}} D_{n,e}.$$

In particular, $D_n = \varnothing$ if n is prime and $x \in P_1$ is neither decomposable nor indecomposable. For the resulting inclusion-exclusion formula for $\#D_n$, we have to determine the *collisions* (or nonuniqueness) of decompositions, that is, different components $(g, h) \neq (g^*, h^*)$ with equal composition $g \circ h = g^* \circ h^*$.

It is useful to single out a special case of wild compositions when $p > 0$.

Example 3.2. We call an $f \in P_n \cap F[x^p]$ a *Frobenius composition*, since then $f = g^* \circ x^p$ for some $g^* \in P_{n/p}$, and any decomposition (g, h) of $f = g \circ h$ is a *Frobenius decomposition*. We denote by $\varphi \colon F \longrightarrow F$ the Frobenius endomorphism over a field F of characteristic p, with $\varphi(a) = a^p$ for all $a \in F$, and extend it to an $\mathbb{F}_p$-linear map $\varphi \colon P_n \longrightarrow P_n$ with $\varphi(x) = x$. For $h \in P_{n/p} \setminus \{x^p\}$, this provides the collision

$$x^p \circ h = \varphi(h) \circ x^p. \tag{3.3}$$

If F is perfect – in particular if F is finite or algebraically closed – then φ is an automorphism on F and every Frobenius composition except x^{p^2} is a collision as in (3.3). Over $F = \mathbb{F}_q$, this yields $q^{p-1} - 1$ collisions in D_{p^2} and $q^{n/p-1}$ collisions in D_n for $p \mid n \neq p^2$, called *Frobenius collisions*. This example is noted in Schinzel (1982, Section I.5, page 39).

For $f \in P_n(F)$ and $a \in F$, the *original shift* of f by a is

$$f^{[a]} = (x - f(a)) \circ f \circ (x + a) \in P_n(F).$$

Original shifting defines a group action of the additive group of F on $P_n(F)$. Shifting respects decompositions in the sense that for each decomposition (g, h) of f we have a decomposition $(g^{[h(a)]}, h^{[a]})$ of $f^{[a]}$, and vice versa. We denote $(g^{[h(a)]}, h^{[a]})$ as $(g, h)^{[a]}$.

3.2 Normal form for Ritt's Second Theorem

Ritt presented two types of essential collisions:

$$\begin{aligned} x^\ell \circ x^k w(x^\ell) = x^{k\ell} w^\ell(x^\ell) = x^k w^\ell \circ x^\ell, \quad (3.4) \\ T_m(x, z^\ell) \circ T_\ell(x, z) = T_{\ell m}(x, z) = T_\ell(x, z^m) \circ T_m(x, z), \end{aligned}$$

where $w \in F[x]$, $z \in F^\times = F \setminus \{0\}$, and T_m is the mth Dickson polynomial of the first kind. And then he proved that these are all possibilities up to composition with linear polynomials. This involved four unspecified linear functions, and it is not clear whether there is a relation between the first and the second type of example.

Von zur Gathen (2014b) presents a normal form for the decompositions in Ritt's Theorem under Zannier's assumption $g'(g^*)' \neq 0$ and the standard assumption $\gcd(\ell, m) = 1$, where $m = k + \ell \deg w$ in (3.4). This normal form is unique unless $p \mid m$.

Theorem 3.5 (Ritt's Second Theorem, normal form)**.** *Let F be a field of characteristic $p \geq 0$, let $m > \ell \geq 2$ be integers with $\gcd(\ell, m) = 1$ and $n = \ell m$. Furthermore, we have monic original $f, g, h, g^*, h^* \in F[x]$ satisfying*

$$f = g \circ h = g^* \circ h^*, \quad (3.6)$$

$$f, g, h, g^*, h^* \text{ are monic original}, \quad (3.7)$$

$$\deg g = \deg h^* = m, \deg h = \deg g^* = \ell, \quad (3.8)$$

$$g'(g^*)' \neq 0, \quad (3.9)$$

where $g' = \partial g / \partial x$ is the derivative of g. Then either (i) or (ii) hold, and (iii) is also valid.

(i) (First Case) There exists a monic polynomial $w \in F[x]$ *of degree* s *and* $a \in F$ *so that*

$$f = (x^{k\ell} w^{\ell}(x^{\ell}))^{[a]},$$

where $m = s\ell + k$ *is the division with remainder of* m *by* ℓ*, with* $1 \le k < \ell$*. Furthermore, we have*

$$\begin{aligned} (g, h) &= (x^k w^{\ell}, x^{\ell})^{[a]}, \\ (g^*, h^*) &= (x^{\ell}, x^k w(x^{\ell}))^{[a]}, \end{aligned} \tag{3.10}$$

$$kw + \ell x w' \neq 0 \text{ and } p \nmid \ell. \tag{3.11}$$

Conversely, any (w, a) *as above for which* (3.11) *holds yields a collision satisfying* (3.6) *through* (3.9)*, via* (3.10)*. If* $p \nmid m$*, then* (w, a) *is uniquely determined by* f *and* ℓ*.*

(ii) (Second Case) There exist $z, a \in F$ *with* $z \neq 0$ *so that*

$$f = T_n(x, z)^{[a]}.$$

Now (z, a) *is uniquely determined by* f*. Furthermore, we have*

$$\begin{aligned} (g, h) &= (T_m(x, z^{\ell}), T_{\ell}(x, z))^{[a]}, \\ (g^*, h^*) &= (T_{\ell}(x, z^m), T_m(x, z))^{[a]}, \end{aligned} \tag{3.12}$$

$$p \nmid n. \tag{3.13}$$

Conversely, if (3.13) *holds, then any* (z, a) *as above yields a collision satisfying* (3.6) *through* (3.9)*, via* (3.12)*.*

(iii) When $\ell \ge 3$*, the First and Second Cases are mutually exclusive. For* $\ell = 2$*, the Second Case is included in the First Case.*

If $p \nmid n$, then the case where $\gcd(\ell, m) \neq 1$ is reduced to the previous one by a result of Tortrat (1988). This determines $D_{n,\ell} \cap D_{n,m}$ exactly if $p \nmid n = \ell m$.

Theorem 3.14 (Tame case)**.** *Let* $\mathbb{F}_q$ *be a finite field of characteristic* p*, let* δ *denote Kronecker's delta function, and let* $m > \ell \ge 2$ *be integers with* $p \nmid n = \ell m$*,* $i = \gcd(\ell, m)$ *and* $s = \lfloor m/\ell \rfloor$*. For the number of monic original polynomials of degree* n *over* $\mathbb{F}_q$ *with left components of degree* ℓ *and* m *we have*

$$\#(D_{n,\ell}(\mathbb{F}_q) \cap D_{n,m}(\mathbb{F}_q)) = \begin{cases} q^{2\ell+s-3} & \text{if } \ell \mid m, \\ q^{2i}(q^{s-1} + (1 - \delta_{\ell,2})(1 - q^{-1})) & \\ \qquad \le q^{2\ell+s-3} & \text{otherwise.} \end{cases}$$

In the remaining case where $p \mid n$, the Frobenius collisions are easily counted and therefore excluded. We have the following upper bounds.

Corollary 3.15 (Wild case, upper bounds). *Let $\mathbb{F}_q$ be a finite field of characteristic p and ℓ, m, $n \geq 2$ be integers with $p \mid n = \ell m$, and let $c = \#(D_{n,\ell}(\mathbb{F}_q) \cap D_{n,m}(\mathbb{F}_q) \setminus F[x^p])$ be the number of monic original polynomials of degree n over $\mathbb{F}_q$ with left components of degree ℓ and m that are not Frobenius collisions. Then the following hold.*

(i) If $p \nmid \ell$, then

$$c \leq q^{m+\lceil \ell/p \rceil -2}.$$

(ii) If $p \mid \ell$ and $\ell < m$, we set $b = \lceil (m-\ell+1)/\ell \rceil$. Then

$$c \leq q^{m+\ell-b+\lceil b/p \rceil -2}.$$

For perspective, we also note the following lower bounds on c from von zur Gathen (2013, 2014a). Unlike the exact result of Theorem 3.14, there is a substantial gap between the upper and lower bounds.

Corollary 3.16 (Wild case, lower bounds). *Let $\mathbb{F}_q$ be a finite field of characteristic p, ℓ a prime number dividing $m > \ell$, assume that $p \mid n = \ell m$, and let $c = \#(D_{n,\ell}(\mathbb{F}_q) \cap D_{n,m}(\mathbb{F}_q) \setminus F[x^p])$ be the number of monic original polynomials of degree n over $\mathbb{F}_q$ with left components of degree ℓ and m that are not Frobenius collisions. Then the following hold.*

(i) If $p = \ell \mid m$ and each nontrivial divisor of m/p is larger than p, then

$$c \geq q^{2p+m/p-3}(1-q^{-1})(1-q^{-p+1}).$$

(ii) If $p \neq \ell$ divides m exactly $d \geq 1$ times, then

$$c \geq q^{2\ell+m/\ell-3}(1-q^{-m/\ell})(1-q^{-1}(1+q^{-p+2}\frac{(1-q^{-1})^2}{1-q^{-p}}))$$

if $\ell \nmid p^d - 1$. Otherwise we set $\mu = \gcd(p^d-1, \ell)$, $r = (p^d-1)/\mu$ and have

$$c \geq q^{2\ell+m/\ell-3}\Big((1-q^{-1}(1+q^{-p+2}\frac{(1-q^{-1})^2}{1-q^{-p}}))(1-q^{-m/\ell}) \\ - q^{-m/\ell-r+2}\frac{(1-q^{-1})^2(1-q^{-r(\mu-1)})}{1-q^{-r}}(1+q^{-r(p-2)})\Big).$$

3.3 The number of decomposable univariate polynomials

The basic statement is that α_n as in (3.18) is an approximation to the number of monic original decomposable polynomials of degree n, with relative error bounds of varying quality. The following is a condensed version of the more precise bounds in von zur Gathen (2014a).

Theorem 3.17. *Let $\mathbb{F}_q$ be a finite field with q elements and characteristic p, let ℓ be the smallest prime divisor of the composite integer $n \geq 2$, and*

$$\alpha_n = \begin{cases} 2q^{\ell+n/\ell-2} & \text{if } n \neq \ell^2, \\ q^{2\ell-2} & \text{if } n = \ell^2. \end{cases} \tag{3.18}$$

Then the following hold for the number $\#D_n(\mathbb{F}_q)$ of decomposable monic original polynomials of degree n over $\mathbb{F}_q$, where $p \parallel n$ means that p divides n exactly twice.

(i) $q^{2\sqrt{n}-2} \leq \alpha_n \leq 2q^{n/2}$.

(ii) $\alpha_n/2 \leq \#D_n(\mathbb{F}_q) \leq \alpha_n(1 + q^{-n/3\ell^2}) < 2\alpha_n \leq 4q^{n/2}$.

(iii) If $n \neq p^2$ *and* $q > 5$, *then* $\#D_n(\mathbb{F}_q) \geq (3 - 2q^{-1})\alpha_n/4 \geq q^{2\sqrt{n}-2}/2$.

(iv) Unless $p = \ell \parallel n$ *and , we have* $\#D_n(\mathbb{F}_q) \geq \alpha_n(1 - 2q^{-1})$.

(v) If $p \nmid n$, *then* $|\#D_n(\mathbb{F}_q) - \alpha_n| \leq \alpha_n \cdot q^{-n/3\ell^2}$.

The relative error in (v) is exponentially decreasing in the input size $n \log q$, in the tame case and for growing $n/3\ell^2$. In (iv), the factor is $1 + O(q^{-1})$ over $\mathbb{F}_q$. When $p = \ell \parallel n$, then we have a factor of about 2 in (ii), which is improved to about 4/3 in (iii). The case $n = p^2$ is settled in subsection 3.4.

Beyond the previous precise bounds, without asymptotics or unspecified constants, we now derive some conclusions about the asymptotic behavior. There are two parameters: the field size q and the degree n. When n is prime, then $\#D_n(\mathbb{F}_q) = 0$, and prime values of n are excepted in the following. We consider the asymptotics in one parameter, where the other one is fixed, and also the special situations where $\gcd(q, n) = 1$. Furthermore, we denote as "$q, n \longrightarrow \infty$" any infinite sequence of pairwise distinct (q, n). The cases $n = 4$ and $p^2 \parallel n \neq p^2$ for some prime p are the only ones where our methods do not show that $\#D_n(\mathbb{F}_q)/\alpha_n \longrightarrow 1$.

Theorem 3.19. *Let $\#D_n(\mathbb{F}_q)$ be the number of decomposable monic original polynomials of degree n over $\mathbb{F}_q$, α_n as in (3.18), and $\nu_{q,n} = \#D_n(\mathbb{F}_q)/\alpha_n$. We only consider composite n.*

(i) For any q, we have

$$\lim_{\substack{n\to\infty \\ \gcd(q,n)=1}} \nu_{q,n} = 1,$$

$$\limsup_{n\to\infty} \nu_{q,n} = 1,$$

$$\frac{1}{2} \leq \nu_{q,n} \text{ for any } n,$$

$$\frac{3-2q^{-1}}{4} \leq \nu_{q,n} \text{ for any } n \text{ if } q > 5.$$

(ii) Let n be a composite integer and ℓ its smallest prime divisor. Then

$$\lim_{\substack{q\to\infty \\ \gcd(q,n)=1}} \nu_{q,n} = 1,$$

$$\limsup_{q\to\infty} \nu_{q,n} = 1,$$

$$\liminf_{q\to\infty} \nu_{q,n} \begin{cases} = 2/3 & \text{if } n = 4, \\ \geq \frac{1}{4}(3 + \frac{1}{\ell+1}) \geq \frac{5}{6} & \text{if } \ell^2 \parallel n \text{ and } n \neq \ell^2, \\ = 1 & \text{otherwise.} \end{cases}$$

(iii) For any sequence $q, n \to \infty$, we have

$$\lim_{\substack{q,n\to\infty \\ \gcd(q,n)=1}} \nu_{q,n} = 1,$$

$$\frac{1}{2} \leq \liminf_{q,n\to\infty} \nu_{q,n} \leq \limsup_{q,n\to\infty} \nu_{q,n} = 1.$$

3.4 Collisions at degree p^2

The previous section gives satisfactory estimates for the number of decomposable polynomials at degree n unless $p^2 \parallel n$. The material of this section determines the number in the easiest of these open cases, namely for $n = p^2$.

First, we present two classes of explicit collisions at degree r^2, where r is a power of the characteristic $p > 0$ of the field F. The collisions of Fact 3.20 consist of additive and subadditive polynomials. A polynomial A of degree r^k is *r-additive* if it is of the form $A = \sum_{0\leq i\leq k} a_i x^{r^i}$ with all $a_i \in F$. We call a polynomial *additive* if it is p-additive. A polynomial is additive if and only if it acts additively on an algebraic closure $\overline{F}$ of F, that is $A(a+b) = A(a) + A(b)$ for all $a, b \in \overline{F}$; see Goss (1996, Corollary 1.1.6). The composition of additive polynomials is additive, see for instance Proposition 1.1.2 of the cited book.

The decomposition structure of additive polynomials was first studied by Ore (1933). Dorey & Whaples (1974, Theorem 4) show that all components of an additive polynomial are additive. Giesbrecht (1988) gives lower bounds on the number of decompositions and algorithms to determine them.

For a divisor m of $r-1$, the (r,m)-*subadditive* polynomial associated with the r-additive polynomial A is $S = x(\sum_{0\le i\le k} a_i x^{(r^i-1)/m})^m$ of degree r^k. Then A and S are related as $x^m \circ A = S \circ x^m$. Dickson (1897) notes a special case of subadditive polynomials, and Cohen (1985) is concerned with the reducibility of some related polynomials. Cohen (1990a,b) investigates their connection to exceptional polynomials and coins the term "sub-linearized"; see also Cohen & Matthews (1994). Coulter, Havas & Henderson (2004) derive the number of indecomposable subadditive polynomials and present an algorithm to decompose subadditive polynomials.

Ore (1933, Theorem 3) describes exactly the right components of degree p of an additive polynomial. Henderson & Matthews (1999) relate such additive decompositions to subadditive polynomials, and in their Theorems 3.4 and 3.8 describe the collisions of Fact 3.20 below. Theorem 3.24 shows that together with those of Theorem 3.22 and the Frobenius collisions of Example 3.2, these examples and their shifts comprise all collisions at degree p^2.

Fact 3.20. *Let r be a power of p, $u, s \in F^\times$, $\varepsilon \in \{0,1\}$, m a positive divisor of $r-1$, $\ell = (r-1)/m$, and*

$$\begin{aligned} f &= S(u,s,\varepsilon,m) = x(x^{\ell(r+1)} - \varepsilon u s^r x^\ell + u s^{r+1})^m \in P_{r^2}(F), \\ T &= \{t \in F \colon t^{r+1} - \varepsilon u t + u = 0\}. \end{aligned} \tag{3.21}$$

For each $t \in T$ and

$$\begin{aligned} g &= x(x^\ell - u s^r t^{-1})^m, \\ h &= x(x^\ell - st)^m, \end{aligned}$$

both in $P_r(F)$, we have $f = g \circ h$. Moreover, f has a $\#T$-collision.

The polynomials f in (3.21) are "simply original" in the sense that they have a simple root at 0. This motivates the designation S. The second construction of collisions goes as follows.

Theorem 3.22. *Let r be a power of p, $b \in F^\times$, $a \in F \setminus \{0, b^r\}$, $a^* = b^r - a$,*

m an integer with $1 < m < r-1$ and $p \nmid m$, $m^* = r - m$, and

$$\begin{aligned} f = M(a,b,m) = x^{mm^*}(x-b)^{mm^*} & \left(x^m + a^* b^{-r}((x-b)^m - x^m)\right)^m \\ & \cdot \left(x^{m^*} + ab^{-r}((x-b)^{m^*} - x^{m^*})\right)^{m^*}, \\ g &= x^m (x-a)^{m^*}, \\ h &= x^r + a^* b^{-r}(x^{m^*}(x-b)^m - x^r), \\ g^* &= x^{m^*}(x-a^*)^m, \\ h^* &= x^r + ab^{-r}(x^m(x-b)^{m^*} - x^r). \end{aligned} \tag{3.23}$$

Then $f = g \circ h = g^* \circ h^* \in P_{r^2}(F)$ has a 2-collision.

The polynomials f in (3.23) are "multiply original" in the sense that they have a multiple root at 0. This motivates the designation M. The notation is set up so that * acts as an involution on our data, leaving b, f, r, and x invariant.

Zieve (2011) points out that the rational functions of case (4) in Proposition 5.6 of Avanzi & Zannier (2003) can be transformed into (3.23). Zieve also mentions that this example already occurs in unpublished work of his, joint with Robert Beals.

Theorem 3.24. *Let F be a perfect field of characteristic p and $f \in P_{p^2}(F)$. Then f has a collision $\{(g,h),(g^*,h^*)\}$ if and only if exactly one of the following holds.*

(F) *The polynomial f is a Frobenius collision as in Example 3.2.*

(S) *The polynomial f is simply original and there are u, s, ε, and m as in Fact 3.20 and $w \in F$ such that*

$$f^{[w]} = S(u,s,\varepsilon,m)$$

and the collision $\{(g,h)^{[w]},(g^*,h^*)^{[w]}\}$ is contained in the collision described in Fact 3.20, with $\#T \geq 2$.

(M) *The polynomial f is multiply original and there are a, b, and m as in Theorem 3.22 and $w \in F$ such that*

$$f^{[w]} = M(a,b,m)$$

and the collision $\{(g,h)^{[w]},(g^*,h^*)^{[w]}\}$ is as in Theorem 3.22.

In particular, the collisions in case (S) and case (M) have exactly $\#T$ and 2 distinct decompositions, respectively. Inclusion-exclusion now yields the following exact formula for the number of decomposable polynomials of degree p^2 over $\mathbb{F}_q$.

Theorem 3.25. *Let $\mathbb{F}_q$ be a finite field of characteristic p, δ Kronecker's delta function, and τ the number of positive divisors of $p-1$. Then for the number $\#D_{p^2}(\mathbb{F}_q)$ of decomposable monic original polynomials of degree p^2 over $\mathbb{F}_q$ we have*

$$\begin{aligned}\#D_{p^2}(\mathbb{F}_q) = q^{2p-2} - q^{p-1} + 1 - \frac{(\tau q - q + 1)(q-1)(qp-p-2)}{2(p+1)} \\ - (1-\delta_{p,2})\frac{q(q-1)(q-2)(p-3)}{4}.\end{aligned}$$

In particular, we have

$$\begin{aligned}\#D_4(\mathbb{F}_q) &= q^2 \cdot \frac{2+q^{-2}}{3} && \text{for } p=2,\\ \#D_9(\mathbb{F}_q) &= q^4\Big(1 - \frac{3}{8}(q^{-1}+q^{-2}-q^{-3}-q^{-4})\Big) && \text{for } p=3,\\ \#D_{p^2}(\mathbb{F}_q) &= q^{2p-2}\Big(1 - q^{-p+1} + O(q^{-2p+5+1/d})\Big) && \text{for } q=p^d \text{ and } p \geq 5.\end{aligned}$$

We have two independent parameters p and d, and $q = p^d$. For two eventually positive functions $f, g\colon \mathbb{N}^2 \to \mathbb{R}$, here $g \in O(f)$ means that there are constants b and c so that $g(p,d) \leq c \cdot f(p,d)$ for all p and d with $p+d \geq b$. We have the following asymptotics.

Corollary 3.26. *Let $p \geq 5$, $d \geq 1$, $q = p^d$, and $k \geq 1$. Then the number c_k of decomposable monic original polynomials of degree p^2 over $\mathbb{F}_q$ with exactly k decompositions is as follows*

$$\begin{aligned}c_1 &= q^{2p-2}(1 - 2q^{-p+1} + O(q^{-2p+5+1/d})),\\ c_2 &= q^{p-1}(1 + O(q^{-p+4+1/d})),\\ c_{p+1} &= (\tau-1)q^{3-3/d}\Big(1 + O(q^{-\max\{2/d, 1-1/d\}})\Big)\\ &\subseteq O\Big(q^{3-3/d+1/(d \log\log p)}\Big),\\ c_k &= 0 \text{ if } k \notin \{1, 2, p+1\}.\end{aligned}$$

Theorem 3.25 leads to $\lim_{q\to\infty} \nu_{q,\ell^2} = 1$ for any prime $\ell > 2$ in Theorem 3.19 (ii). For $n = 4$, the sequence has no limit, but oscillates between values close to $\liminf_{q\to\infty} \nu_{q,4} = 2/3$ and to $\limsup_{q\to\infty} \nu_{q,4} = 1$, and these are the only two accumulation points of the sequence $\nu_{q,4}$.

4 Open problems

Further types of multivariate polynomials that are examined from a counting perspective include singular bivariate ones (von zur Gathen, 2008) and pairs of coprime polynomials (Hou & Mullen, 2009). It remains open to extend the methods of Section 2 to singular multivariate ones and achieve exponentially decreasing error bounds for coprime multivariate polynomials.

For univariate decomposable polynomials, the question of good asymptotics for $\nu_{q,n}$ when q is fixed and $n \to \infty$ is still open. More work is needed to understand the case where the characteristic p is the smallest prime divisor of the degree n, divides n exactly twice, and $n \neq p^2$. Ritt's Second Theorem covers distinct-degree collisions, even in the wild case, see Zannier (1993); it would be interesting to see a parametrization even for $p \mid m$ and obtain a similar classification for general equal-degree collisions.

Finally, this survey deals with polynomials only and the study of rational functions with the same methods remains open.

5 Acknowledgments

This work was funded by the B-IT Foundation and the Land Nordrhein-Westfalen.

References

MAX ALEKSEYEV (2006). A115457–A115472. In *The On-Line Encyclopedia of Integer Sequences*. OEIS Foundation Inc. URL `http://oeis.org`. Last download 4 December 2012. (Cited on page 7.)

E. ARTIN (1924). Quadratische Körper im Gebiete der höheren Kongruenzen. II. (Analytischer Teil.). *Mathematische Zeitschrift* **19**(1), 207–246. URL `http://dx.doi.org/10.1007/BF01181075`. (Cited on page 2.)

ROBERTO M. AVANZI & UMBERTO M. ZANNIER (2003). The equation $f(X) = f(Y)$ in rational functions $X = X(t)$, $Y = Y(t)$. *Compositio Math.* **139**(3), 263–295. URL `http://dx.doi.org/10.1023/B:COMP.0000018136.23898.65`. (Cited on pages 15 and 24.)

DAVID R. BARTON & RICHARD ZIPPEL (1985). Polynomial Decomposition Algorithms. *Journal of Symbolic Computation* **1**, 159–168. (Cited on page 15.)

Raoul Blankertz (2014). A polynomial time algorithm for computing all minimal decompositions of a polynomial. *ACM Communications in Computer Algebra* **48**(1), 13–23. Issue 187. (Cited on page 15.)

Raoul Blankertz, Joachim von zur Gathen & Konstantin Ziegler (2013). Compositions and collisions at degree p^2. *Journal of Symbolic Computation* **59**, 113–145. ISSN 0747-7171. URL `http://dx.doi.org/10.1016/j.jsc.2013.06.001`. Also available at `http://arxiv.org/abs/1202.5810`. Extended abstract in *Proceedings of the 2012 International Symposium on Symbolic and Algebraic Computation ISSAC '12, Grenoble, France* (2012), 91–98. (Cited on page 16.)

Arnaud Bodin (2008). Number of irreducible polynomials in several variables over finite fields. *American Mathematical Monthly* **115**(7), 653–660. ISSN 0002-9890. (Cited on pages 2 and 7.)

Arnaud Bodin (2010). Generating series for irreducible polynomials over finite fields. *Finite Fields and Their Applications* **16**(2), 116–125. URL `http://dx.doi.org/10.1016/j.ffa.2009.11.002`. (Cited on pages 3 and 7.)

Arnaud Bodin, Pierre Dèbes & Salah Najib (2009). Indecomposable polynomials and their spectrum. *Acta Arithmetica* **139(1)**, 79–100. (Cited on pages 3 and 16.)

John J. Cade (1985). A New Public-key Cipher Which Allows Signatures. In *Proceedings of the 2nd SIAM Conference on Applied Linear Algebra*. SIAM, Raleigh NC. (Cited on page 15.)

M. Car (1987). Théorèmes de densité dans $\mathbb{F}_q[X]$. *Acta Arithmetica* **48**, 145–165. (Cited on page 2.)

Leonard Carlitz (1932). The arithmetic of polynomials in a Galois field. *American Journal of Mathematics* **54**, 39–50. (Cited on page 9.)

Leonard Carlitz (1963). The distribution of irreducible polynomials in several indeterminates. *Illinois Journal of Mathematics* **7**, 371–375. (Cited on page 2.)

Leonard Carlitz (1965). The distribution of irreducible polynomials in several indeterminates II. *Canadian Journal of Mathematics* **17**, 261–266. (Cited on page 2.)

Eda Cesaratto, Joachim von zur Gathen & Guillermo Matera (2013). The number of reducible space curves over a finite field. *Journal* *of Number Theory* **133**, 1409–1434. URL `http://dx.doi.org/10.1016/j.jnt.2012.08.027`. (Cited on pages 3 and 4.)

Stephen Cohen (1968). The distribution of irreducible polynomials in several indeterminates over a finite field. *Proceedings of the Edinburgh Mathematical Society* **16**, 1–17. (Cited on page 2.)

Stephen Cohen (1969). Some arithmetical functions in finite fields. *Glasgow Mathematical Society* **11**, 21–36. (Cited on page 2.)

Stephen D. Cohen (1985). Reducibility of sub-linear polynomials over a finite field. *Bulletin of the Korean Mathematical Society* **22**, 53–56. (Cited on page 23.)

Stephen D. Cohen (1990a). Exceptional polynomials and the reducibility of substitution polynomials. *Enseign. Math. (2)* **36**(1-2), 53–65. ISSN 0013-8584. (Cited on page 23.)

Stephen D. Cohen (1990b). The Factorable Core of Polynomials Over Finite Fields. *Journal of the Australian Mathematical Society, Series A* **49**(02), 309–318. URL `http://dx.doi.org/10.1017/S1446788700030585`. (Cited on page 23.)

Stephen D. Cohen & Rex W. Matthews (1994). A class of exceptional polynomials. *Transactions of the American Mathematical Society* **345**(2), 897–909. ISSN 0002-9947. URL `http://www.jstor.org/stable/2155005`. (Cited on page 23.)

Robert S. Coulter, George Havas & Marie Henderson (2004). On decomposition of sub-linearised polynomials. *Journal of the Australian Mathematical Society* **76**(3), 317–328. ISSN 1446-7887. URL `http://dx.doi.org/10.1017/S1446788700009885`. (Cited on page 23.)

L. E. Dickson (1897). The analytic representation of substitutions on a power of a prime number of letters with a discussion of the linear group. *Annals of Mathematics* **11**, 65–120, 161–183. (Cited on page 23.)

F. Dorey & G. Whaples (1974). Prime and Composite Polynomials. *Journal of Algebra* **28**, 88–101. URL `http://dx.doi.org/10.1016/0021-8693(74)90023-4`. (Cited on pages 14 and 23.)

David Eisenbud & Joe Harris (1992). The dimension of the Chow variety of curves. *Compositio Mathematica* **83**(3), 291–310. (Cited on page 3.)

H. T. Engstrom (1941). Polynomial Substitutions. *American Journal of Mathematics* **63**, 249–255. URL `http://www.jstor.org/stable/pdfplus/2371520.pdf`. (Cited on page 14.)

Jean-Charles Faugère & Ludovic Perret (2008). High Order Derivatives and Decomposition of Multivariate Polynomials. In *Extended Abstracts of the Second Workshop on Mathematical Cryptology WMC 08*, edited by Álvar Ibeas & Jaime Gutiérrez, 90–93. URL `http://grupos.unican.es/amac/wmc-2008/`. (Cited on page 14.)

P. Flajolet, X. Gourdon & D. Panario (2001). The Complete Analysis of a Polynomial Factorization Algorithm over Finite Fields. *Journal of Algorithms* **40**(1), 37–81. Extended Abstract in *Proceedings of the 23rd International Colloquium on Automata, Languages and Programming ICALP 1996*, Paderborn, Germany, ed. F. Meyer auf der Heide and B. Monien, Lecture Notes in Computer Science **1099**, Springer-Verlag, 1996, 232–243. (Cited on page 9.)

Philippe Flajolet & Robert Sedgewick (2009). *Analytic Combinatorics*. Cambridge University Press. ISBN 0521898064, 824 pages. (Cited on pages 4, 5, and 9.)

Michael D. Fried & R. E. MacRae (1969). On the invariance of chains of Fields. *Illinois Journal of Mathematics* **13**, 165–171. (Cited on page 14.)

Shuhong Gao & Alan G. B. Lauder (2002). Hensel Lifting and Bivariate Polynomial Factorisation over Finite Fields. *Mathematics of Computation* **71**(240), 1663–1676. (Cited on page 2.)

Joachim von zur Gathen (1990a). Functional Decomposition of Polynomials: the Tame Case. *Journal of Symbolic Computation* **9**, 281–299. URL `http://dx.doi.org/10.1016/S0747-7171(08)80014-4`. Extended abstract in *Proceedings of the 28th Annual IEEE Symposium on Foundations of Computer Science, Los Angeles CA* (1987). (Cited on page 15.)

Joachim von zur Gathen (1990b). Functional Decomposition of Polynomials: the Wild Case. *Journal of Symbolic Computation* **10**, 437–452. URL `http://dx.doi.org/10.1016/S0747-7171(08)80054-5`. (Cited on page 15.)

Joachim von zur Gathen (2002). Factorization and Decomposition of Polynomials. In *The Concise Handbook of Algebra*, edited by Alexander V. Mikhalev & Günter F. Pilz, 159–161. Kluwer Academic Publishers. ISBN 0-7923-7072-4. (Cited on page 15.)

Joachim von zur Gathen (2008). Counting reducible and singular bivariate polynomials. *Finite Fields and Their Applications* **14**(4), 944–978. URL http://dx.doi.org/10.1016/j.ffa.2008.05.005. Extended abstract in *Proceedings of the 2007 International Symposium on Symbolic and Algebraic Computation ISSAC '07, Waterloo, Ontario, Canada* (2007), 369-376. (Cited on pages 2 and 26.)

Joachim von zur Gathen (2011). Counting decomposable multivariate polynomials. *Applicable Algebra in Engineering, Communication and Computing* **22**(3), 165–185. URL http://dx.doi.org/10.1007/s00200-011-0141-9. Abstract in *Abstracts of the Ninth International Conference on Finite Fields and their Applications*, pages 21–22, Dublin, July 2009, Claude Shannon Institute, http://www.shannoninstitute.ie/fq9/AllFq9Abstracts.pdf. (Cited on pages 3 and 4.)

Joachim von zur Gathen (2013). Lower bounds for decomposable univariate wild polynomials. *Journal of Symbolic Computation* **50**, 409–430. URL http://dx.doi.org/10.1016/j.jsc.2011.01.008. Extended abstract in *Proceedings of the 2009 International Symposium on Symbolic and Algebraic Computation ISSAC '09, Seoul, Korea* (2009). (Cited on page 20.)

Joachim von zur Gathen (2014a). Counting decomposable univariate polynomials. *To appear in Combinatorics, Probability and Computing, Special Issue.* Extended abstract in *Proceedings of the 2009 International Symposium on Symbolic and Algebraic Computation ISSAC '09, Seoul, Korea* (2009). Preprint (2008) available at http://arxiv.org/abs/0901.0054. (Cited on pages 16, 20, and 21.)

Joachim von zur Gathen (2014b). Normal form for Ritt's Second Theorem. *Finite Fields and Their Applications* **27**, 41–71. ISSN 1071-5797. URL http://dx.doi.org/10.1016/j.ffa.2013.12.004. Also available at http://arxiv.org/abs/1308.1135. (Cited on pages 16 and 18.)

Joachim von zur Gathen, Jaime Gutierrez & Rosario Rubio (2003). Multivariate polynomial decomposition. *Applicable Algebra in Engineering, Communication and Computing* **14**(1), 11–31. URL http://www.springerlink.com/content/0a5c0vvp82xbx5je/. Extended abstract in *Proceedings of the Second Workshop on Computer Algebra in Scientific Computing, CASC '99, München, Germany* (1999). (Cited on page 14.)

Joachim von zur Gathen, Dexter Kozen & Susan Landau (1987). Functional Decomposition of Polynomials. In *Proceedings of the 28th Annual*

IEEE Symposium on Foundations of Computer Science, Los Angeles CA, 127–131. IEEE Computer Society Press, Washington DC. URL http://dx.doi.org/10.1109/SFCS.1987.29. (Cited on page 15.)

Joachim von zur Gathen, Alfredo Viola & Konstantin Ziegler (2013). Counting reducible, powerful, and relatively irreducible multivariate polynomials over finite fields. *SIAM Journal on Discrete Mathematics* **27**(2), 855–891. URL http://dx.doi.org/10.1137/110854680. Also available at http://arxiv.org/abs/0912.3312. Extended abstract in *Proceedings of LATIN 2010, Oaxaca, Mexico* (2010). (Cited on pages 2 and 4.)

Mark William Giesbrecht (1988). *Some Results on the Functional Decomposition of Polynomials*. Master's thesis, Department of Computer Science, University of Toronto. Technical Report 209/88. Available as http://arxiv.org/abs/1004.5433. (Cited on pages 16 and 23.)

Sudesh K. Gogia & Indar S. Luthar (1981). Norms from certain extensions of $F_q(T)$. *Acta Arithmetica* **38**(4), 325–340. ISSN 0065-1036. (Cited on page 2.)

David Goss (1996). *Basic Structures of Function Field Arithmetic*. Springer-Verlag. ISBN 3-540-61087-1. (Cited on page 22.)

Johannes Grabmeier, Erich Kaltofen & Volker Weispfenning (editors) (2003). *Computer Algebra Handbook – Foundations, Applications, Systems*. Springer-Verlag, Berlin, Heidelberg, New York. ISBN 3-540-65466-6. URL http://www.springer.com/978-3-540-65466-7. (Cited on page 31.)

Jaime Gutierrez & Dexter Kozen (2003). Polynomial Decomposition. In Grabmeier, Kaltofen & Weispfenning (2003), section 2.2.4 (pages 26–28). URL http://www.springer.com/978-3-540-65466-7. (Cited on page 15.)

Jaime Gutierrez, Josef Schicho & Martin Weimann (editors) (2014). *Computer Algebra and Polynomials*, Lecture Notes in Computer Science. (Cited on page 3.)

Jaime Gutierrez & David Sevilla (2006). On Ritt's decomposition theorem in the case of finite fields. *Finite Fields and Their Applications* **12**(3), 403–412. URL http://dx.doi.org/10.1016/j.ffa.2005.08.004. Also available at http://arxiv.org/abs/0803.3976. (Cited on page 15.)

David. R. Hayes (1965). The Distribution of Irreducibles in GF$[q, x]$. *Transactions of the American Mathematical Society* **117**, 101–127. URL http://dx.doi.org/10.2307/1994199. (Cited on page 2.)

Marie Henderson & Rex Matthews (1999). Composition behaviour of sub-linearised polynomials over a finite field. In *Finite Fields: Theory, Applications, and Algorithms (Waterloo, ON, 1997)*, volume 225 of *Contemp. Math.*, 67–75. Amer. Math. Soc., Providence, RI. (Cited on page 23.)

Xiang-dong Hou & Gary L. Mullen (2009). Number of Irreducible Polynomials and Pairs of Relatively Prime Polynomials in Several Variables over Finite Fields. *Finite Fields and Their Applications* **15**(3), 304–331. URL http://dx.doi.org/10.1016/j.ffa.2008.12.004. (Cited on pages 7 and 26.)

Dexter Kozen & Susan Landau (1989). Polynomial Decomposition Algorithms. *Journal of Symbolic Computation* **7**, 445–456. URL http://dx.doi.org/10.1016/S0747-7171(89)80027-6. An earlier version was published as Technical Report 86-773, Cornell University, Department of Computer Science, Ithaca, New York, 1986. (Cited on page 15.)

Dexter Kozen, Susan Landau & Richard Zippel (1996). Decomposition of Algebraic Functions. *Journal of Symbolic Computation* **22**, 235–246. (Cited on page 15.)

S. Landau & G. L. Miller (1985). Solvability by Radicals is in Polynomial Time. *Journal of Computer and System Sciences* **30**, 179–208. (Cited on page 15.)

Howard Levi (1942). Composite polynomials with coefficients in an arbitrary field of characteristic zero. *American Journal of Mathematics* **64**, 389–400. (Cited on page 14.)

Gary L. Mullen & Daniel Panario (2013). *Handbook of Finite Fields*. Discrete Mathematics and Its Applications. CRC Press. ISBN 978-1-4398-7378-6 (Hardback). URL http://www.crcpress.com/product/isbn/9781439873786. (Cited on page 2.)

O. Ore (1933). On a Special Class of Polynomials. *Transactions of the American Mathematical Society* **35**, 559–584. (Cited on page 23.)

Andrzej Schinzel (1982). *Selected Topics on Polynomials*. Ann Arbor; The University of Michigan Press. ISBN 0-472-08026-1. (Cited on pages 14 and 18.)

ANDRZEJ SCHINZEL (2000). *Polynomials with special regard to reducibility.* Cambridge University Press, Cambridge, UK. ISBN 0521662257. (Cited on pages 14 and 15.)

PIERRE TORTRAT (1988). Sur la composition des polynômes. *Colloquium Mathematicum* **55**(2), 329–353. (Cited on page 19.)

GERHARD TURNWALD (1995). On Schur's Conjecture. *Journal of the Australian Mathematical Society, Series A* **58**, 312–357. URL http://anziamj.austms.org.au/JAMSA/V58/Part3/Turnwald.html. (Cited on page 15.)

DAQING WAN (1992). Zeta Functions of Algebraic Cycles over Finite Fields. *Manuscripta Mathematica* **74**, 413–444. (Cited on page 2.)

KENNETH S. WILLIAMS (1969). Polynomials with irreducible factors of specified degree. *Canadian Mathematical Bulletin* **12**, 221–223. ISSN 0008-4395. (Cited on page 2.)

U. ZANNIER (1993). Ritt's Second Theorem in arbitrary characteristic. *Journal für die reine und angewandte Mathematik* **445**, 175–203. (Cited on pages 14, 15, and 26.)

UMBERTO ZANNIER (2008). On composite lacunary polynomials and the proof of a conjecture of Schinzel. *Inventiones mathematicae* **174**, 127–138. ISSN 0020-9910 (Print) 1432-1297 (Online). URL http://dx.doi.org/10.1007/s00222-008-0136-8. Also available at http://arxiv.org/abs/0705.0911v1. (Cited on page 16.)

KONSTANTIN ZIEGLER (2014). Tame decompositions and collisions. *Preprint,* 35 pages. URL http://arxiv.org/abs/1402.5945. Extended abstract to appear in *Proceedings of the 2014 International Symposium on Symbolic and Algebraic Computation ISSAC '14, Kobe, Japan.* (Cited on page 16.)

MICHAEL ZIEVE (2011). Personal communication. (Cited on page 24.)

MICHAEL E. ZIEVE & PETER MÜLLER (2008). On Ritt's Polynomial Decomposition Theorems. *Submitted,* URL http://arxiv.org/abs/0807.3578. (Cited on page 15.)

RICHARD ZIPPEL (1991). Rational Function Decomposition. In *Proceedings of the 1991 International Symposium on Symbolic and Algebraic Computation ISSAC '91, Bonn, Germany*, edited by STEPHEN M. WATT, 1–6. ACM Press, Bonn, Germany. ISBN 0-89791-437-6. (Cited on page 15.)

K. Zsigmondy (1894). Über die Anzahl derjenigen ganzzahligen Functionen n-ten Grades von x, welche in Bezug auf einen gegebenen Primzahlmodul eine vorgeschriebene Anzahl von Wurzeln besitzen. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Abteilung II* **103**, 135–144. (Cited on page 2.)