Friable values of binary forms

joint work with

Antal Balog, Valentin Blomer
and Gérald Tenenbaum

An y-friable integer is an integer whose all prime factors are $\leqslant y$.

Canonical decomposition $N = ab$ with a y-friable, b "sieved", *i.e* without prime factor $\leqslant y$.

For $n \in \mathbb{N}^*$, we denote by $P^+(n)$ resp. $P^-(n)$ the largest (resp. the smallest) prime factor of n, with the convention $P^+(1) = 1$, $P^-(1) = \infty$.

Let $F \in \mathbb{Z}[X]$, $2 \leqslant y \leqslant x$. We define

$$\Psi_F(x, y) = |\{n \leqslant x : P^+(F(n)) \leqslant y\}|.$$

For $F(X) = X$, we simply write $\Psi(x, y)$.
Let $u = \log x / \log y$. We have:

$$\Psi(x, y) \sim x\varrho(u)$$

in the range

$$(\log x)^{\log\log x} \leqslant y \leqslant x. \tag{1}$$

ϱ is the Dickman's function

$$\varrho(u) = u^{-u+o(u)} \quad (u \to \infty).$$

If $\{u_n\}_{n=1}^{\infty}$ is a sequence of random and independant variables uniformly distributed on $[0, 1]$ then the series $Y = u_1 + u_1u_2 + u_1u_2u_3 + \cdots$ converge with probability 1 to a random variable of density $\mathrm{e}^{-\gamma}\varrho$.

The Riemann hypothesis is equivallent to replace in (1) the exponant $\log\log x$ by $2 + \varepsilon$.

Conjecture : If F is an irreducible polynomial then $\{F(n)\}_{n\in\mathbb{N}}$ behaves like a random sequence of random integers.
Consequence : $P^+(F(n)) < n^{\varepsilon}$ with probability > 0.

Schinzel (1967) : if F has degree $g \geqslant 2$, there exists $\delta(g) \in]0,1[$ and infinitely many n such that $P^+(F(n)) \leqslant n^{g-1-\delta(g)}$, $(\delta(g) \sim 2/g$ when $g \to \infty)$.
For $F(n) = an^g + b$, for all $\varepsilon > 0$, there exists infinitely many integers n such that $P^+(an^g + b) \leqslant n^{\varepsilon}$.

Balog and Wooley (1998) : same result for

$$F(n) = \prod_{i=1}^{t} (a_i n^{g_i} + b_i).$$

Question : let $F \in \mathbb{Z}[X]$ with degree $\geqslant 2$; for which $\alpha > 0$ can we obtain an estimation of type

$$\Psi_F(x, x^\alpha) \asymp x,$$

where the implicites constants depend only on F and α?

Such estimations have been obtained for the following cases:

- $F(X) = X(aX + b)$, $a \in \mathbb{N}^*$, $b \in \mathbb{N}$, $\alpha > 0$, Balog and Ruzsa (1997)
- $F(X) = \prod_{1 \leqslant j \leqslant k}(X + j)$, $k \geqslant 2$, $\alpha > e^{-1/(k-1)}$, Hildebrand (1989)
- $F(X) = X^2 + 1$, $\alpha \geqslant 149/179$, C. D. (1996).

General result: G. Martin, G. Tenenbaum and C. D. (2001)

$$F(X) = F_1(X)_1^{\alpha} \cdots F_r(X)_r^{\alpha} \quad (F_1, \ldots, F_r \text{ irreducibles})$$

$g = \max_{1 \leqslant i \leqslant r} (\deg F_i)$,
k=number of factors F_i with degree g.
For x large enough and $y \geqslant x^{g+\varepsilon-1/k}$ we have:

(2) $$\Psi_F(x, y) \asymp x.$$

Friable values of binary forms

Motivation: the number field sieve (ANR CADO)

Let N be an integer to factorize. Let $f \in \mathbb{Z}[X]$ of degree d and m such that $N = f(m)$. Let F be the correspondant binary form: $F(a,b) = b^d f(a/b)$.

An important step of the number field sieve is to find sufficiently many (a,b) such that $F(a,b)(a-bm)$ is friable.

Let $F \in \mathbb{Z}[X,Y]$ be a binary form. The function $\Psi_F(x,y)$ is now:

$$\Psi_F(x,y) := |\{1 \leqslant a,b \leqslant x : P^+(F(a,b)) < y\}|.$$

G. Hanrot, G. Tenenbaum and J. Wu (2007): $F(a,b) = a^2 + b^2$. For all $\beta \in]0,3/5[$ they obtain an asymptotic formula for $\Psi_F(x,y)$ in the range

$$x \geqslant 3, \quad \exp\{(\log x)^{1-\beta}\} \leqslant y \leqslant x.$$

Theorem 1 (A. Balog, V. Blomer, C. D., G. Tenenbaum). *Let $F = F(X,Y)$ be a binary form with integer coefficients and degree $t \geqslant 2$. Let g be the largest degree of an irreducible factor of F and let k (resp. ℓ) denote the number of distinct irreducible factors of F having degree g—resp. $g-1$. Given any positive real number ε, the estimate*

$$\Psi_F(x,y) \asymp x^2 \tag{3}$$

holds for all large x provided $y \geqslant x^{\alpha_F+\varepsilon}$, where the exponent α_F is defined by

$$\alpha_F := \begin{cases} g - 2/k & \text{if } k \geqslant 2, \\ g - 1 - 1/(\ell+1) & \text{if } k = 1 \text{ and } (g,t) \neq (2,3) \\ \frac{1}{\sqrt{e}} = 0.6065\ldots & \text{if } t = 3 \text{ and } F \text{ is irreducible} \\ 0 & \text{if } t = 3 \text{ and } F \text{ is reducible} \\ 0 & \text{if } t = 1 \text{ or } 2. \end{cases} \tag{4}$$

Starting idea of the proof.
Suppose that F is irreducible.
If there exists a large $d|F(a,b)$ then

$$P^+(F(a,b)) \leqslant \max(d, F(a,b)/d).$$

Thus the problem is to obtain for some $\delta > 0$ the lower bound

$$|\{1 \leqslant a, b \leqslant x : \exists d|F(a,b),\ x^\alpha < d < x^{\alpha+\delta}\}| \gg x^2.$$

When F is reducible we study a vectorial version of this quantity.

Level of distribution of the sequence $\{F(a,b)\}_{a,b\in\mathbb{N}}$.
Given a binary form $F(X,Y)$, a real number $x \geqslant 1$ and a positive integer d, we define

$$A_d(x;F) := \operatorname{card}\{1 \leqslant a,b \leqslant x : F(a,b) \equiv 0 \,(\operatorname{mod} d)\}.$$

We consider the approximation

$$A_d(x;F) = x^2 \frac{\varrho_F(d)}{d^2} + r_d(x),$$

where $\varrho_F(d) := |\{0 \leqslant u,v < d : F(u,v) \equiv 0 \,(\operatorname{mod} d)\}|$, and $r_d(x)$ is an error term.
Greaves (1970) proved that if $F(X,Y)$ is irreducible and not linear the error term is small on average over d:

$$\sum_{d\leqslant z} |r_d(x)| \ll (z+x)z^{\varepsilon} \qquad (x,z \geqslant 1).$$

A Variant of Greave's result.

Let $\boldsymbol{F} = (F_1, \dots, F_m)$ be a multiset of m distinct irreducible binary forms. For all $\boldsymbol{d} = (d_1, \dots, d_m) \in \mathbb{N}^m$ and $x \geqslant 1$ we write $\boldsymbol{F}(a,b) \equiv 0 \,(\mathrm{mod}\, \boldsymbol{d}) \Leftrightarrow F_j(a,b) \equiv 0 \,(\mathrm{mod}\, d_j)$ for $1 \leqslant j \leqslant m$, $A(x; \boldsymbol{F}, \boldsymbol{d}) := |\{1 \leqslant a, b \leqslant x : \boldsymbol{F}(a,b) \equiv 0 \,(\mathrm{mod}\, \boldsymbol{d})\}|$.

Theorem 2. *Let $\varepsilon > 0$ and $0 \leqslant s \leqslant m$. Assume that $F_1, \dots, F_s$ are linear factors and that $F_{s+1}, \dots, F_m$ are factors of degree $\geqslant 2$. Then, uniformly for $D_1 \geqslant 1, \dots, D_m \geqslant 1$, $x \geqslant 1$, we have*

(5)
$$\sum_{\substack{d_1 \leqslant D_1, \dots, d_m \leqslant D_m \\ \mu(d_1 \cdots d_m)^2 = 1}} \left| A(x; \boldsymbol{F}, \boldsymbol{d}) - x^2 \prod_{1 \leqslant j \leqslant m} \frac{\varrho_{F_j}(d_j)}{d_j^2} \right| \ll \left\{ x \check{D}_s + \check{D}_s^2 \widehat{D}_s + D \right\} D^{\varepsilon},$$

with $\check{D}_s := 1 + \sum_{1 \leqslant j \leqslant s} D_j$, $\widehat{D}_s := \prod_{s < j \leqslant m} D_j$, $D := \prod_{1 \leqslant j \leqslant m} D_j$.

Estimate of a summatory function linked to polynomial congruences

Let $F \in \mathbb{Z}[X]$ irreducible. Let $\mathbb{K}$ be the number field generated over $\mathbb{Q}$ by a root of F and $\zeta_{\mathbb{K}}$ its Dedekind zeta function.

Theorem 3. *Let f be a multiplicative function such that its Dirichlet series $\mathcal{F}(s) := \sum_{n\geqslant 1} f(n)/n^s$ is absolutely convergent for $\Re e\, s > 1$. We suppose furthermore that in the half-plane of convergence $\Re e\, s > 1$, we have*

$$\mathcal{F}(s) = \zeta_{\mathbb{K}}(s)\mathcal{G}(s)$$

where $\mathcal{G}$ is a Dirichlet series representable as an Eulerian product, which does not vanish and is absolutely convergent in a suitable half-plane $\Re e\, s \geqslant 1-\delta$ with $\delta > 0$.

There exists an absolute constant $c > 0$ such that, for any given $\varepsilon > 0$ and uniformly in the domain H_ε, we have

$$\sum_{\substack{n\leqslant x\\ P^-(n)>y}} f(n) = \big(x\omega(u)-y\big)\frac{\mathrm{e}^\gamma}{\zeta(1,y)}+O\bigg(\frac{x\varrho(u)}{(\log y)^2}\Big\{H(u)^{-c}+Y_\varepsilon^{-1}\Big\}\bigg)$$

where ϱ and ω denote respectively the Dickman function and the Buchstab function, $\zeta(s,y)$ is the partial Eulerian product

$$\zeta(s,y) := \prod_{p\leqslant y}\big(1-1/p^s\big)^{-1} \qquad (\Re e\, s > 0)$$

the domain (H_ε) is defined by

$$(H_\varepsilon) \qquad x \geqslant 3, \quad \exp\big\{(\log_2 x)^{5/3+\varepsilon}\big\} \leqslant y \leqslant x.$$

and for $x \geqslant y \geqslant 2$, $u := (\log x)/\log y$,

$$H(u) := \exp\{u/(1+\log u)^2\}, \qquad Y_\varepsilon := \exp\big\{(\log y)^{3/2-\varepsilon}\big\}.$$

The reducible cubic case.

Case $F = F_1F_2F_3$ is a product of 3 distinct linear forms (linearly independant). Writing $r = F_1(a,b)$, $s = F_2(a,b)$, we have for some constants c_1, c_2, c_3, c_4, $\alpha, \beta \in \mathbb{Z}$

$$\Psi_F(x,y) \gg |\{(r,s) \in [c_1x, c_2x] \times [c_3x, c_4x] : P^+(rs(\alpha r + \beta s)) \leqslant y\}|.$$

Let $\mathcal{A}, \mathcal{B} \subset [1,x] \cap \mathbb{N}$. R. de la Bretèche (1999) obtained an asymptotic formula for the quantity

$$\sum_{\substack{(a,b)\in\mathcal{A}\times\mathcal{B}\\ P^+(a+b)\leqslant y}} 1$$

in the range $x \geqslant 3$, $\exp((\log x)^{2/3+\varepsilon}) \leqslant y$.

We apply this result with $\mathcal{A} = \{\alpha r : r \in [c_1x, c_2x] \text{ and } P^+(r) \leqslant y\}$, $\mathcal{B} = \{\beta s : s \in [c_3x, c_4x] \text{ and } P^+(s) \leqslant y\}$.

Case $F = F_1F_2$ where F_1 is linear and F_2 is quadratic (and irreducible). Writing $r = F_1(a, b)$, $s = b$ we have

$$\Psi_F(x, y) \gg \{(r, s) \in [c_1x, c_2x] \times [c_3x, c_4x] : P^+(rF_2'(r, s)) \leqslant y\}$$

for suitable constants c_1, c_2, c_3, c_4 and a binary quadratic form F_2'.

A large sieve inequality. For $h \in \mathbb{Z}$ we define

$$\gamma_{h,r}(d) := \sum_{\substack{1\leqslant s\leqslant d\\ d|F_2'(r,s)}} \exp(2i\pi\frac{hs}{d}).$$

Theorem 4. *For any sequence $\xi_{h,r}$ of complex numbers, D, H, R, and any $\delta > 0$ we have*

$$\sum_{d\leqslant D}\left|\sum_{h\leqslant H}\sum_{r\leqslant R}\xi_{h,r}\gamma_{h,r}(d)\right| \ll D^{\frac12}(D+HR)^{\frac12}\Big(\sum_{h,r}|\xi_{h,r}|^2\Big)^{\frac12}(DHR)^{\delta}.$$

Fouvry and Iwaniec (1997) : $F(a,b) = a^2 + b^2$ (asymptotic formula for the number of the prime $p \leqslant x$ of type $p = \ell^2 + m^2$ where ℓ is a prime number.)

Idea of the proof of Theorem 4

Take for example $F_2(a,b) = a^2 + b^2$. There exists a correspondance between the roots of the congruences $F_2(v,1) = v^2 + 1 \equiv 0 \,(\mathrm{mod}\, d)$ and the representations $d = r^2 + s^2$ with $(r,s) = 1$ and $-s < r \leqslant s$. Moreover we have

$$\frac{v}{d} \equiv \frac{r}{sd} - \frac{\bar{r}}{s} \,(\mathrm{mod}\, 1) \quad \text{where } r\bar{r} \equiv 1 \,(\mathrm{mod}\, d).$$

The spacement between two fractions $\frac{v}{d}$, $\frac{v_1}{d_1}$ with $D < d, d_1 \leqslant 2D$ is ($\|\vartheta\|$ denotes the distance of ϑ to $\mathbb{Z}$)

$$\left\| \frac{v}{d} - \frac{v_1}{d_1} \right\| > \frac{1}{36D}.$$

Generalization to the other binary quadratic forms.
There exists also a correspondance between the roots v, $1 \leqslant v \leqslant d$, $F_2(v,1) \equiv 0 \,(\mathrm{mod}\, d)$ and some representation of d of type $d = Q(r,s)$ where Q belongs to a finite set of binary quadratic form of discriminant the discrimant of F_2. This correspondance provides a parametrisation of the roots v in terms of r, s and we can prove an inequality of type (writing $ru - st = 1$)

$$\frac{v}{d} = \frac{r}{\alpha_2' t} + O\Big(\frac{1}{d}\Big).$$

Thus the fractions of the set

$$\{v/d : F_2(v,1) \equiv 0 \,(\mathrm{mod}\, d), 1 \leqslant v \leqslant d, D < d \leqslant 2D\}$$

are well-spaced.