A Diophantine representation of Wolstenholme's pseudoprimality*

Luca Vallata[1] and Eugenio G. Omodeo[2]

[1] Graduated from the University of Trieste,
email: luca.vallata@gmail.com
[2] Dipartimento di Matematica e Geoscienze / DMI, Università di Trieste,
Via Valerio 12/1, I-34127 – Trieste, Italy
email: eomodeo@units.it

Abstract. As a by-product of the negative solution of Hilbert's 10^{th} problem, various prime-generating polynomials were found. The best known *upper* bound for the number of variables in such a polynomial, to wit 10, was found by Yuri V. Matiyasevich in 1977.
We show that this bound could be lowered to 8 if the converse of Wolstenholme's theorem (1862) holds, as conjectured by James P. Jones. This potential improvement is achieved through a Diophantine representation of the set of all integers $p \geqslant 5$ that satisfy the congruence $\binom{2p}{p} \equiv 2 \mod p^3$. Our specification, in its turn, relies upon a terse polynomial representation of exponentiation due to Matiyasevich and Julia Robinson (1975), as further manipulated by Maxim Vsemirnov (1997).
We briefly address the issue of also determining a *lower* bound for the number of variables in a prime-representing polynomial, and discuss the autonomous significance of our result about Wostenholme's pseudoprimality, independently of Jones's conjecture.

Keywords. Diophantine representations, Hilbert's 10^{th} problem, DPRM theorem, Wolstenholme's theorem, Siegel's theorem on integral points.

Introduction

At the beginning of the 1960's, one decade after Martin Davis had set forth the 'daring hypothesis $\cdots$ that *every semidecidable set is Diophantine*' [Mat93, p. 99], it became clear that finding a proof of that conjecture would have entailed the possibility to construct a polynomial with integer coefficients whose positive values, as the variables run through all nonnegative integers, form the set of prime numbers.[3] The existence of such a prime-generating polynomial seemed, at the time, rather unlikely; in fact, Davis's conjecture was received with understandable skepticism.

* Work partially supported by the project FRA-UniTS (2014) *"Learning specifications and robustness in signal analysis (with a case study related to health care)."*

[3] Cf. [DMR76, Sec. 1]: "This corollary was deduced by Putnam in 1960 from the then conjectured Main Theorem and it was considered by some to be an argument against its plausibility."

With [Mat70], Yuri V. Matiyasevich positively settled Davis's conjecture and so provided a negative answer to Hilbert's 10^{th} problem [Hil00, p. 276]. Soon afterward, the same scholar obtained two polynomials representing primes and only primes, one in 24 and one in 21 variables [Mat71]; in [MR75], Matiyasevich and Julia Robinson brought the number of variables down to 14; then other researchers succeeded in bringing it further down, to 12 (cf. [JSWW76]). The record number, 10 as of today, was achieved by Matiyasevich in 1977: in fact, [Mat81] produces a prime-generating polynomial in 10 variables, of degree 15905 (reducible to 13201 (13983?) or to 11281 [Mat81, p. 44], or even to 10001 [Vse97, p. 3204]).

Although methods have significantly evolved over time, the rigmarole for getting prime-representing polynomials usually results from the combination of ideas already present in [Rob52] (see Fig. 1) with a Diophantine polynomial specification of exponentiation, such as the masterpiece proposed in [MR75] (see Fig. 2), which Maxim A. Vsemirnov refined somewhat in [Vse97].

$$a = \binom{r}{j} \quad \leftrightarrow \quad a = \left\lfloor \frac{(u+1)^r}{u^j} \right\rfloor \% \, u \;\; \& \;\; u = 2^r + 1$$

$$j! = \left\lfloor \frac{r^j}{\binom{r}{j}} \right\rfloor \quad \text{for any } r > (2\,j)^{j+1}$$

$$\begin{gathered} \neg \exists \, x \, , \, y \, \big(p = (x+2)\,(y+2) \vee p = 0 \vee p = 1\big) \\ \leftrightarrow \; \exists \, q \, , \, u \, , \, v \, \big(p = q + 2 \;\&\; p\,u - (q+1)!\,v = 1\big) \end{gathered}$$

Fig. 1. Binomial coefficient, factorial, and "p is a prime" are existentially definable by means of exponential Diophantine equations, cf. [Rob52, pp. 446–447]. Throughout, '%' designates the integer remainder operation.

Ameliorations along this pipeline are possible: e.g., Wilson's theorem enables one to state that p is a prime number through the formula $\exists \, q \, , \, u \, \big(p = q + 2 \;\&\; p\,u - (q+1)! = 1\big)$; and an improved exponential Diophantine representation of the binomial coefficient can be obtained through the theorem

$$\binom{r}{j} = \left\lfloor \frac{(u+1)^r}{u^j} \right\rfloor \% \, u \;\; \text{for } r > 0 \, , \, j > 0 \, , \text{and } u > r^j \; ,$$

as remarked in [MR75, pp. 544–545]. However, a more decisive enhancement in the formulation of a prime-generating polynomial would ensue if one could remove factorial from the pipeline and could avoid exploiting the binomial coefficient in its full strength.

Joseph Wolstenholme proved the congruence $\binom{2\,p-1}{p-1} \equiv 1 \mod p^3$ for all prime numbers $p > 3$ in 1862 [Wol62]; and it was conjectured by James P. Jones (cf. [Rib04, p. 23] and [McI95, p. 381]) that, conversely, every integer $p > 3$ satisfying the said congruence is prime. If true, this conjecture would ease our

$$Q = \Box \;\leftrightarrow_{\mathrm{Def}}\; Q = h^2 \quad \text{for some } h \in \mathbb{N}\,,$$

$$X \mid Y \;\leftrightarrow_{\mathrm{Def}}\; Y = \pm h\, X \text{ for some } h \in \mathbb{N}\,.$$

A1	$D\,F\,I = \Box\,,\quad F \mid H - C\,,\quad B \leqslant C$	**E1**	$(M^2-1)\,L^2+1=\Box$
A2	$D \risingdotseq (A^2-1)\,C^2+1$	**E2**	$L^2 - 4\,(C - L\,y)^2\,x\,y\,n > 0$
A3	$E \risingdotseq 2\,(i+1)\,C^2\,D$	**E3**	$M \risingdotseq 4\,n\,(y+1)+x+2$
A4	$F \risingdotseq (A^2-1)\,E^2+1$	**E4**	$L \risingdotseq n+1+\ell\,(M-1)$
A5	$G \risingdotseq (F-A)\,F+A$	**E5**	$A \risingdotseq M\,x$
A6	$H \risingdotseq B+2\,j\,C$	**E6**	$B \risingdotseq n+1$
A7	$I \risingdotseq (G^2-1)\,H^2+1$	**E7**	$C \risingdotseq k+B$

Fig. 2. Polynomial specification of the triadic relation $x^n = y$. Besides the parameters x, y, n, this involves four existential variables (also ranging over $\mathbb{N}$): i, j, k, ℓ; a fifth unknown is implicit in the constraint **A1** stating that the product DFI must be a perfect square with F dividing $H - C$. The notation '$V \risingdotseq P$' defines V to be an alias for the (integer-valued) polynomial P; hence all uppercase letters can be eliminated, e.g. in the order: M, B; A, C, L; H, D; E; F; G; I. By themselves, **A1**–**A7** form a polynomial specification of the relation $\psi_A(B) = C$ defined by the recurrence $\psi_A(0) = 0$, $\psi_A(1) = 1$, and $\psi_A(h+2) = 2A\psi_A(h+1) - \psi_A(h)$, if one takes A, B, C as parameters subject to the preconditions $A > 1$, $B > 0$, $C > 0$.

present task, enabling us to express primality without factorial and in terms of the *central* binomial coefficient $\binom{2\,p}{p}$.

After recalling, in Sec. 1 the basic definitions and techniques we need, in Sec. 2 we produce a Diophantine polynomial generator in 8 variables for the numbers meeting the just mentioned 'Wolstenholme's pseudoprimality' criterion. In Sec. 3, we give clues about the proof that the proposed polynomial operates properly. In the conclusions, we briefly discuss the autonomous significance of our specification independently of Jones's conjecture, and address the issue of determining a lower bound for the number of variables in a polynomial representation of primality.

1 Main definitions and presupposed notions

Let us recall here the notion of *Diophantine representation* of a relation $\mathcal{R}$, which historically played an essential role in the study of Hilbert's 10^{th} problem:

Definition 1. *A relation $\mathcal{R}$ among n natural numbers is said to be* Diophantine *if one can precisely characterize which are the n-tuples $\langle \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \rangle$ constituting $\mathcal{R}$ through a bi-implication of the form*

$$\mathcal{R}(a_1,\ldots,a_n) \leftrightarrow \exists\, x_1 \cdots \exists\, x_m \left(D(\overbrace{\underbrace{a_1,\ldots,a_n}_{\text{parameters}},\ \underbrace{x_1,\ldots,x_m}_{\text{unknowns}}}^{\text{variables}}) = 0 \right)$$

which musto be true under the replacement $a_1 \mapsto \boldsymbol{a}_1, \ldots, a_n \mapsto \boldsymbol{a}_n$, where D is a polynomial with coefficients in $\mathbb{Z}$ whose variables are seen as ranging over $\mathbb{N}$.

In the common case when $n = 1$ one calls such an $\mathcal{R}$ a *Diophantine set*, and one readily gets from the defining D the polynomial $(x_0+1)\big(1-D^2(x_0,\dots,x_m)\big)-1$, whose non-negative values (under replacement of the variables $x_0,\dots,x_m$ by natural numbers $\boldsymbol{x}_0,\dots,\boldsymbol{x}_m$) are precisely the elements of $\mathcal{R}$.

For example, classical results on the so-called Pell equation tell us that the equation $x^2 - d\,(y+1)^2 - 1 = 0$ in the parameter d and in the unknowns x, y makes a Diophantine representation of the set

$$\mathcal{R} = \{\,0\,\} \;\cup\; \{\, d \in \mathbb{N} \;|\; d \text{ is } \textit{not} \text{ a perfect square}\,\}\,;$$

therefore the non-negative values of the polynomial

$$(z+1)\,\big(\,1-(x^2-z\,(y+1)^2-1)^2\,\big)-1\,,$$

as x, y, z range over $\mathbb{N}$, will form this $\mathcal{R}$.

The Pell equation of the special form $x^2 = (a^2-1)\,y^2+1$ enters extensively in the ongoing; thus we find it convenient to denote its right-hand side as $\mathsf{Pell}(a, y)$. We adopt $\mathsf{Pell}(S, T)$ as an analogous syntactic abbreviation also in the case when S and T are Diophantine polynomials, as shown in Fig. 6 (top).

As is well-known (see, e.g., [Dav73]), the solutions to the said equation $x^2 = \mathsf{Pell}(\boldsymbol{a}, y)$ when $\boldsymbol{a} \geqslant 2$ form a doubly recurrent infinite sequence

$$\langle 1\,,\ 0\rangle,\ \langle \boldsymbol{a}\,,\ 1\rangle,\ \langle 2\,\boldsymbol{a}^2-1\,,\ 2\,\boldsymbol{a}\rangle,\ \langle 4\,\boldsymbol{a}^3-3\,\boldsymbol{a}\,,\ 4\,\boldsymbol{a}^2-1\rangle,\ \dots$$

of pairs whose first and second components constitute the respective increasing progressions $\chi_{\boldsymbol{a}}(0), \chi_{\boldsymbol{a}}(1),\ \chi_{\boldsymbol{a}}(2), \dots$ and $\psi_{\boldsymbol{a}}(0), \psi_{\boldsymbol{a}}(1), \dots$ shown in Fig. 3 (the latter was formerly introduced in the caption of Fig. 2). Figures 4, 5 recapitulate important properties enjoyed by these sequences.

$\chi_{\boldsymbol{a}}(0) = 1$	$\chi_{\boldsymbol{a}}(1) = \boldsymbol{a}$	$\chi_{\boldsymbol{a}}(h+2) = 2\,\boldsymbol{a}\,\chi_{\boldsymbol{a}}(h+1) - \chi_{\boldsymbol{a}}(h)$
$\psi_{\boldsymbol{a}}(0) = 0$	$\psi_{\boldsymbol{a}}(1) = 1$	$\psi_{\boldsymbol{a}}(h+2) = 2\,\boldsymbol{a}\,\psi_{\boldsymbol{a}}(h+1) - \psi_{\boldsymbol{a}}(h)$

Fig. 3. Recurrent specification of the solutions $\boldsymbol{x} = \chi_{\boldsymbol{a}}(b)$, $\boldsymbol{y} = \psi_{\boldsymbol{a}}(b)$ of Pell's equation $x^2 - (\boldsymbol{a}^2-1)\,y^2 = 1$. (These make sense even for $\boldsymbol{a} = 1$.)

2 How to represent Wolstenholme's pseudoprimality via a Diophantine polynomial

To be better aligned with [Vse97], let us now agree that the variables appearing in our Diophantine constraints must range over *positive* (instead of non-negative) integers. A refined polynomial specification of the components which occupy odd positions b in the progression $\psi_{\boldsymbol{a}}(b) = c$ discussed above is shown in Fig. 6

$$
\begin{array}{l}
n < \boldsymbol{a}^n \leqslant \chi_{\boldsymbol{a}}(n) \leqslant \dfrac{\chi_{\boldsymbol{a}}(n+1)}{\boldsymbol{a}} < \begin{cases} \chi_{\boldsymbol{a}}(n+1)\,, \\ (2\,\boldsymbol{a})^n + 1\,; \end{cases} \\[2ex]
\qquad\quad n \leqslant \psi_{\boldsymbol{a}}(n) < \dfrac{\psi_{\boldsymbol{a}}(n+1)}{\boldsymbol{a}} < \quad \psi_{\boldsymbol{a}}(n+1)\,; \\[2ex]
\qquad\qquad\quad \psi_{\boldsymbol{a}}(n) < \begin{cases} \frac{1}{2}\,\chi_{\boldsymbol{a}}(n+1)\,, \\ \frac{1}{2}\,\chi_{\boldsymbol{a}}(n) \quad \text{if } \boldsymbol{a} > 2\,; \end{cases} \\[2ex]
\qquad (2\,\boldsymbol{a}-1)^n \leqslant \psi_{\boldsymbol{a}}(n+1) \leqslant (2\,\boldsymbol{a})^n\,.
\end{array}
$$

Fig. 4. Noteworthy inequalities holding for the progressions $\chi_{\boldsymbol{a}}, \psi_{\boldsymbol{a}}$ ($\boldsymbol{a} \geqslant 2$).

0.	$\chi_{\boldsymbol{a}}(n) - \psi_{\boldsymbol{a}}(n)\,(\boldsymbol{a}-\ell) \equiv \ell^n \mod (\,2\,\boldsymbol{a}\,\ell - \ell^2 - 1\,)\,;$
1.	$\psi_{\boldsymbol{a}}(n) \equiv n \mod (\boldsymbol{a}-1)$ and $\psi_{\boldsymbol{a}}(n) \equiv n \mod 2\,;$
2.	$p \equiv q \mod r$ implies $\begin{cases} \chi_p(n) \equiv \chi_q(n) \mod r\,, \\ \psi_p(n) \equiv \psi_q(n) \mod r\,; \end{cases}$ $r \mid (\,p-1\,)$ implies $\psi_p(n) \equiv n \mod r\,;$
3.	$\psi_{\boldsymbol{a}}(n) \mid \psi_{\boldsymbol{a}}(n\,k)\,;$
4.	$\psi_{\boldsymbol{a}}(n) \mid \psi_{\boldsymbol{a}}(\ell)$ iff $n \mid \ell\,;$
5.	$\psi_{\boldsymbol{a}}(m\,r) \equiv r\,\chi_{\boldsymbol{a}}^{r-1}(m)\,\psi_{\boldsymbol{a}}(m) \mod \psi_{\boldsymbol{a}}^3(m)\,;$
6.	$\psi_{\boldsymbol{a}}(n) \mid \ell$ if $\psi_{\boldsymbol{a}}^2(n) \mid \psi_{\boldsymbol{a}}(\ell)\,;$
7.	$\psi_{\boldsymbol{a}}^2(n) \mid \psi_{\boldsymbol{a}}(n\,\psi_{\boldsymbol{a}}(n))\,;$
8.	$\psi_{\boldsymbol{a}}(i) \equiv \psi_{\boldsymbol{a}}(j) \mod \chi_{\boldsymbol{a}}(m)$ implies $(\,i \equiv j \vee i \equiv -j\,) \mod (\,2\,m\,)\,.$

Fig. 5. Noteworthy congruences holding for the progressions $\chi_{\boldsymbol{a}}, \psi_{\boldsymbol{a}}$. Here it is assumed that $n \geqslant 0$, $k \geqslant 0$, $\ell \geqslant 0$ and that $p > 0$, $q > 0$, $r > 0$, $m > 0$.

(right) and in Fig. 7 (left); in Fig. 7 (right) we extend it into an alike specification, to be discussed next, of Wolstenholme's pseudoprimality. In addition to the 6 unknowns z, w, s, h, i, j which appear explicitly in this system of Diophantine constraints, additional unknowns enter into play due to the presence of the constructs '$\square$', '$>$', '$|$', and of a congruence. Eliminating such abbreviations seems, at first glance, to call for five extra variables; a single, 7-th unknown suffices, though, thanks to the following proposition:

Theorem 1 (Relation-combining theorem, [MR75, pp. 525–527]). *To each q in $\mathbb{N}$ there corresponds a polynomial M_q with coefficients in $\mathbb{Z}$ such that, for all integers $X_1, \ldots, X_q$, J, R, V with $J \neq 0$, the conditions*

$$X_1 = \square\,, \ldots, X_q = \square\,, J \mid R\,, V > 0$$

$$\boxed{\mathsf{Pell}(\mathrm{S},\mathrm{T}) \;=_{\mathrm{Def}}\; (\mathrm{S}^2-1)\,\mathrm{T}^2+1}$$

<table>
<tr><td colspan="2">$\mathsf{Pell}(A\,,\,C)\,F\,I = \square\,,\;\; F \mid B+2\,j\,C-C\,,\;\; B \leqslant C$
$D \leftrightharpoons \mathsf{Pell}(A\,,\,C)$</td></tr>
<tr><td>$F \leftrightharpoons \mathsf{Pell}\big(A\,,\,2\,(i+1)\,C^2\,D\big)$
$I \leftrightharpoons \mathsf{Pell}\big((F-A)\,F+A\,,\,B+2\,j\,C\big)$</td><td>$F \leftrightharpoons \mathsf{Pell}\big(A\,,\,2\,i\,C^2\,D\big)$
$I \leftrightharpoons \mathsf{Pell}\big((A+1)\,F-A\,,\,B+2\,j\,C\big)$</td></tr>
</table>

Fig. 6. Polynomial specifications of the relation $\psi_A(B) = C$ (see Fig. 2). When conjoined with the constraints in the middle, the two constraints appearing on the left form an abridged formulation of the specification **A1**–**A7** recalled above from [MR75, pp. 532–533]: in this case, the unknowns i, j etc. range over $\mathbb{N}$ and the parameters A, B, C are assumed to satisfy $A > 1$, $B > 0$, $C > 0$. Likewise, the two constraints on the right must be combined with the ones in the middle to get an abridged version of the specification of [Vse97, pp. 3203–3204]: in this case, variables range over $\mathbb{N} \setminus \{0\}$ and the assumed preconditions are $A > 1$, $B > 1$, and $B \equiv 1 \mod 2$; a lower overall degree results from $(A+1)\,F - A$ having superseded $(F-A)\,F + A$.

are all met if and only if the equation $M_q(X_1, \ldots, X_q, J, R, V, m) = 0$ *admits solutions for some value* $\boldsymbol{m}$ *in* $\mathbb{N}$ *of the variable* m. ⊣

This theorem is exploitable in the case at hand, with $q = 2$, once the two divisibility conditions (one of which is hidden inside the congruence $3\,w\,C \equiv 2\,(w^2-1) \mod Q$) are combined together by resorting to the double implication

$$d_1 \mid z_1 \wedge d_2 \mid z_2 \;\leftrightarrow\; d_1\,d_2 \mid z_1 d_2 + z_2 d_1$$

which holds when d_1, d_2, z_1, z_2 are positive integers and d_1, d_2 are co-prime. All in all, we will be able to fold our constraints into a single Diophantine polynomial equation $\mathcal{W}(k, x_1, \ldots, x_7) = 0$ over $\mathbb{N}$ whose degree is 5488 (as will be assessed at the end of Sec. 3) and which admits solutions in the 7 unknowns precisely for those integer values of k which exceed 4 and which also satisfy Wolstenholme's congruence $\binom{2\,k}{k} \equiv 2 \mod k^3$.

In order to get rid of the precondition $k \geqslant 5$ (Fig. 7, right), it suffices to strenghten the inequality $K^2 - 4\,(C - K\,Y)^2 > 0$ into $(k-1)\,(k-2)\,(k-3)\,(k-4)\,\big(K^2 - 4\,(C-K\,Y)^2\big) > 0$ before resorting to Thm. 1. Accordingly, denoting by $\bar{\mathcal{W}}(k, x_1, \ldots, x_7)$ the polynomial equation that results after this preparatory retouch, our conjectured prime-generating polynomial is:

$$x_0\,\big(1 - (x_0-2)^2\,(x_0-3)^2\,\bar{\mathcal{W}}^2(x_0, x_1, \ldots, x_7)\big)\,.$$

3 Correctness of our representation of Wolstenholme's pseudoprimality

The specification of Wolstenholme's pseudoprimality which we are proposing stems from *ad hoc* modifications to [Jon82, Lemma 2.25, pp. 556–557]; hence, by bringing into our present discourse the main ingredients entering the proof thereof, we will easily get our main claim, which is:

$$\boxed{\mathsf{Pell}(\mathrm{S},\mathrm{T}) \;=_{\mathrm{Def}}\; (\mathrm{S}^2-1)\,\mathrm{T}^2+1}$$

<table>
<tr><td colspan="2">$\mathsf{Pell}(A\,,\,C)\,F\,I=\square\ \wedge\ F\mid B+2\,j\,C-C$
$D \coloneqq \mathsf{Pell}(A\,,\,C)$
$I \coloneqq \mathsf{Pell}\big((A+1)\,F-A\,,\ B+2\,j\,C\big)$
$F \coloneqq \mathsf{Pell}\big(A\,,\ 2\,i\,C^2\,D\,Q\big)$</td></tr>
<tr><td>$B \leqslant C$</td><td>$K^2-4\,(C-K\,Y)^2>0$
$\mathsf{Pell}(P\,,\ K)=\square$
$3\,w\,C \equiv 2\,(w^2-1) \mod Q$
$M \coloneqq k\,Y$
$Y \coloneqq k^3\,s+2$
$P \coloneqq 2\,M^2\,U$
$Q \coloneqq 4\,A-5$
$U \coloneqq k^3\,w$
$K \coloneqq k+1+h\,(P-1)$
$A \coloneqq M\,(U+1)$
$B \coloneqq 2\,k+1$
$C \coloneqq B+z$</td></tr>
<tr><td>Domain: $\mathbb{N}\setminus\{0\}$
Unknowns: i, j, m
Parameters: A, B, C
Precond.: $A>1\,,\ B>1\,,\ 2\nmid B\,,\ C>1$
Specifies: $\psi_A(B)=C$
Sources: [MR75], [Vse97]</td><td>Domain: $\mathbb{N}\setminus\{0\}$
Unknowns: z, w, s, h, i, j, m
Parameters: k
Precondition: $k \geqslant 5$
Specifies: $\binom{2\,k}{k} \equiv 2 \mod k^3$
Sources: [Jon82, Lemma 2.25],
L. Vallata's laurea thesis</td></tr>
</table>

Fig. 7. Polynomial specification of Wolstenholme's pseudoprimality.

Theorem 2. *Let $\mathcal{W}(k,z,w,s,h,i,j,m)=0$ be the Diophantine polynomial equation resulting from the system in Fig. 7, right, via Thm. 1. Then the integer values $\boldsymbol{k} \geqslant 5$ for which the congruence $\binom{2\boldsymbol{k}}{\boldsymbol{k}} \equiv 2 \mod \boldsymbol{k}^3$ holds are precisely the ones for which the equation $\mathcal{W}(\boldsymbol{k},z,w,s,h,i,j,m)=0$—where $\boldsymbol{k}$ has superseded the variable k—can be solved relative to the unknowns z,w,s,h,i,j,m.* ⊣

First, we need an economical—as for the number of variables involved—representation of the triadic relation $\psi_A(B)=C$. We resort to a slight variant of the one which [Vse97, Lemma 8] proposed for an even number B, because an odd B better fits our present aims.

Lemma 1. *Let A, B, C, Q be integers with $A > 1$, $B > 1$, $C > 1$, B odd, and $Q > 0$. The relationship $\psi_A(B) = C$ holds if and only if there exist i, j such that*

$$\begin{cases} D\,F\,I = \square & \text{(P1)} \\ F \mid H - C & \text{(P2)} \\ B \leqslant C & \text{(P3)} \end{cases} \qquad \begin{array}{ll} D \leftrightharpoons \mathsf{Pell}(A\,,\ C) & \text{(P4)} \\ E \leftrightharpoons 2\,i\,C^2\,D\,Q & \text{(P5)} \\ F \leftrightharpoons \mathsf{Pell}(A\,,\ E) & \text{(P6)} \\ G \leftrightharpoons (A+1)\,F - A & \text{(P7)} \\ H \leftrightharpoons B + 2\,j\,C & \text{(P8)} \\ I \leftrightharpoons \mathsf{Pell}(G\,,\ H) & \text{(P9)} \end{array}$$

Proof: Minor modifications to the proof of [Vse97, Lemma 8, pp. 3203–3204] (see also Remark 2 therein) yield the claim of this lemma. In its turn, that proof mimicked the proof of [MR75, Theorem 4, pp. 532–533]. □

Second, we need a Diophantine representation of exponentiation:

Lemma 2. *The relationship $S^B = Y$ holds for integers S, B, Y with $S > 0$ if and only if there exist integers A, C such that*

1. $S < A$,
2. $Y^3 < A$,
3. $S^{3B} < A$,
4. $\psi_A(B) = C$,
5. $(S^2 - 1)\,Y\,C \ \equiv\ S\,(Y^2 - 1) \mod (2\,A\,S - S^2 - 1)$.

Proof: See [Jon79, Lemma 2.8, pp. 213–214], where this result is credited to Julia Robinson. A key congruence in Jones's proof just cited is

$$(\,\ell^2 - 1\,)\,\ell^n\,\psi_{\boldsymbol{a}}(n) \equiv \ell\,(\,\ell^{2\,n} - 1\,) \mod (\,2\,\boldsymbol{a}\,\ell - \ell^2 - 1\,)\,,$$

which follows easily from Fig. 5 (0), in light of the fact that $\boldsymbol{x} = \chi_{\boldsymbol{a}}(n)$, $\boldsymbol{y} = \psi_{\boldsymbol{a}}(n)$ solves the equation $x^2 = (\boldsymbol{a}^2 - 1)\,y^2 + 1$. Making use of the easy implication

$$\boldsymbol{a} \leqslant 2\,\boldsymbol{a}\,\ell - \ell^2 - 1 \ \text{ if } \ 0 < \ell < \boldsymbol{a}\,,$$

Jones gets another key ingredient for the proof:

> If $0 < \ell < \boldsymbol{a}$, $y^3 < \boldsymbol{a}$, and $z^3 < \boldsymbol{a}$ then, taken together, the congruences
>
> $$\begin{array}{l} (\,\ell^2 - 1\,)\,y\,\psi \equiv \ell\,(\,y^2 - 1\,) \mod (2\,\boldsymbol{a}\,\ell - \ell^2 - 1), \\ (\,\ell^2 - 1\,)\,z\,\psi \equiv \ell\,(\,z^2 - 1\,) \mod (2\,\boldsymbol{a}\,\ell - \ell^2 - 1) \end{array}$$
>
> imply that $y = z$, for any number ψ.

The desired conclusion follows without difficulty. □

In the light of Lemma 1 and Lemma 2, minimal clues about the proof of Theorem 2 should suffice to the reader: we will limit ourselves to indicating the modifications which the statement of the above-cited Lemma 2.25 of [Jon82] should undergo, so that its proof can then be adapted to our case without any substantial changes. Some variables of the cited lemma must be replaced by ours according to the rewritings: $B' \rightsquigarrow B$, $\phi \rightsquigarrow z$, $W \rightsquigarrow w$, $R \rightsquigarrow k$, and $N \rightsquigarrow k$ (notice that we are thus enforcing the equality $R = N$). Moreover, one should: remove condition (B11) $W = b\,w$ of the cited lemma; replace its conditions (B9) $U = N^2\,w$ and (B10) $Y = N^2\,s$ by ours, namely $U = k^3\,w$ and $Y = k^3\,s + 2$; add our condition $Q = 4\,A - 5$.

Degree of the polynomial through which we have represented Wolstenholme's pseudoprimality

To end, let us calculate the degree of the polynomial $\mathcal{W}(k, z, w, s, h, i, j, m)$ discussed above. To more easily get the degrees of the polynomials involved in the right-hand specification of Fig. 7, we add a few more abbreviations to it: $H \leftrightharpoons B + 2\,j\,C$, $E \leftrightharpoons 2\,i\,C^2\,D\,Q$, and $G \leftrightharpoons (A+1)\,F - A$; then we get the degree map:

$$\begin{array}{l} B/1,\;\; U/4,\, Y/4;\;\; C/1,\;\; M/5;\; H/2,\;\; A/9,\;\; P/14; \\ D/20,\, Q/9,\, K/15;\, E/32;\, F/82;\, G/91;\, I/186. \end{array}$$

To complete the assessment of the degree of $\mathcal{W}$, we need to make the polynomial M_q of Thm. 1 rather explicit: according to [MR75],

$$\begin{array}{l} M_q(\,X_1\,,\ldots,X_q\,,\;J\,,\;R\,,\;V\,,\;m\,) \quad =_{\text{Def}} \quad \prod_{\sigma\in\{0,1\}^{\{1,\ldots,q\}}} \Big(J^2\,m + \\ \quad R^2 - J^2\,(\,2\,V - 1\,)\,\Big(R^2 + W^q + \sum_{j=1}^{q}(-1)^{\sigma(j)}\sqrt{X_j}\,W^{j-1}\Big)\Big)\,, \end{array}$$

where

$$W \leftrightharpoons 1 \;+\; \textstyle\sum_{i=1}^{q} X_i^2\,.$$

In the case at hand,

$$\mathcal{W}(k, z, w, s, h, i, j, m) \;=_{\text{Def}}\; M_2(\,X_1\,,\;X_2\,,\;J\,,\;R\,,\;V\,,\;m),$$

where $X_1 \leftrightharpoons DFI$ and $X_2 \leftrightharpoons \mathsf{Pell}(P\,,\;K)$; hence $q = 2$ and $W \leftrightharpoons 1 + (D\,F\,I)^2 + \big((P^2-1)\,K^2+1\big)^2$. The polynomial V which we using in a statement $V > 0$ is $V \leftrightharpoons K^2 - 4(C - KY)^2$. The polynomials J, R of which we are stating that $J \mid R$, result from combination of the two conditions $F \mid H - C$ and $3\,w\,C \equiv 2\,(w^2-1) \bmod Q$: hence $J \leftrightharpoons F\,Q$ and $R \leftrightharpoons (H-C)\,Q + \big(2\,(w^2-1) - 3\,w\,C\big)\,F$. The polynomials just introduced have degrees:

$$W/576,\; V/38,\; J/91,\; R/84$$

and, consequently, $\mathcal{W}$ has the degree

$$\deg M_2 = 4\;\deg\big(\,J^2\,(\,2\,V - 1\,)\,W^2\,\big) = 4 \cdot 1372 = 5488\,.$$

Conclusions and future work

After explaining what it means for a relation $\varrho(x_1, \ldots, x_n)$ to be *Diophantine in a set* $\mathcal{S}$, Julia Robinson proved in [Rob69] that every recursively enumerable set is Diophantine in *any* infinite set of primes. We do not know whether Jones's conjectured converse of Wolstenholme's theorem will be proved, hence we cannot refer Robinson's result just recalled to the set $\mathbb{W}$ of all integers $k \geqslant 5$ such that $\binom{2k}{k} \equiv 2 \mod k^3$, and we feel that it would add to the autonomous significance of our polynomial representation of $\mathbb{W}$ if we succeeded in showing that every recursively enumerable set is Diophantine in $\mathbb{W}$.

Albeit subject to Jones's conjecture, the result presented in this paper suggests a new estimate for the *rank* (= least possible number of unknowns in a Diophantine representation) of the set of primes, shifting it down from 9 to 7. Although this was to be expected (cf. [Mat93, p. 56]), we could not find this result published anywhere.

We would also like to determine a non-trivial *lower* bound for the rank of primality. Pietro Corvaja gave us clues that the lower bound 2 can be obtained through direct application of Siegel's theorem on integral points (see [Sie29][4]).

It is a bit deceiving that we could not benefit from the celebrated [AKS04] for the aims of this paper; an explanation might be that the complexity of prime-number recognition has to do with bounds that one can place on the *sizes* of the unknowns in a Diophantine representation of primality rather than on the number of those unknowns.

Acknowledgements

As hinted at above, we had pleasant and profitable exchanges of ideas with prof. Pietro Corvaja (University of Udine).

References

N.B. Yuri V. Matiyasevich's name was transliterated variously in his publications in English; in this bibliography, the authors have preferred conformity with the spellings found in the originals to uniformity of writing.

AKS04. Manindra Agrawal, Neeraj Kayal, and Nitin Saxena. Primes is in P. *Annals of Mathematics*, 160(2):781–793, June 2004.

CZ02. Pietro Corvaja and Umberto Zannier. A subspace theorem approach to integral points on curves. *C. R. Acad. Sci. Paris, Ser. I*, 334(4):267–271, 2002.

Dav58. Martin Davis. *Computability and Unsolvability.* McGraw-Hill, New York, 1958. Reprinted with an additional appendix, Dover 1983.

Dav73. Martin Davis. Hilbert's tenth problem is unsolvable. *The American Mathematical Monthly*, 80(3):233–269, 1973. Reprinted with corrections in the Dover edition of *Computability and Unsolvability* [Dav58].

[4] A proof of Siegel's theorem along new lines can be found in [CZ02].

DMR76. Martin Davis, Yuri Matijasevič, and Julia Robinson. Hilbert's tenth problem. Diophantine equations: positive aspects of a negative solution. In *Mathematical Developments Arising From Hilbert Problems*, volume 28 of *Proceedings of Symposia in Pure Mathematics*, pages 323–378, Providence, RI, 1976. American Mathematical Society. Reprinted in [Rob96].

Hil00. David Hilbert. Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Kongreß zu Paris 1900. *Nachrichten von der Königliche Gesellschaft der Wissenschaften zu Göttingen*, pages 253–297, 1900.

Jon79. James P. Jones. Diophantine representation of Mersenne and Fermat primes. *Acta Arithmetica*, XXXV(3):209–221, 1979.

Jon82. James P. Jones. Universal Diophantine equation. *The Journal of Symbolic Logic*, 47(3):549–571, 1982.

JSWW76. James P. Jones, Daihachiro Sato, Hideo Wada, and Douglas Wiens. Diophantine representation of the set of prime numbers. *American Mathematical Monthly*, 83(6):449–464, 1976.

Mat70. Ju. V. Matijasevič. Diofantovost' perechislimykh mnozhestv. *Doklady Akademii Nauk SSSR*, 191(2):279–282, 1970. (Russian). (Translated as Ju. V. Matijasevič. Enumerable sets are Diophantine. *Soviet Mathematics. Doklady*, 11(2):354-358, 1970.).

Mat71. Ju. V. Matijasevič. Diophantine representation of the set of prime numbers. *Soviet Mathematics. Doklady*, 12(1):249–254, 1971.

Mat81. Yu. V. Matijasevič. Primes are nonnegative values of a polynomial in 10 variables. *Journal of Soviet Mathematics*, 15(1):33–44, 1981.

Mat93. Yuri Vladimirovich Matiyasevich. *Hilbert's tenth problem*. The MIT Press, Cambridge (MA) and London, 1993.

McI95. Richard J. McIntosh. On the converse of Wolstenholme's theorem. *Acta Arithmetica*, LXXI(4):381–389, 1995.

MR75. Yuri Matijasevič and Julia Robinson. Reduction of an arbitrary diophantine equation to one in 13 unknowns. *Acta Arithmetica*, XXVII:521–553, 1975.

Rib04. Paulo Ribenboim. *The little book of bigger primes*. Springer, 2nd edition, 2004.

Rob52. Julia Robinson. Existential definability in arithmetic. *Transactions of the American Mathematical Society*, 72(3):437–449, 1952.

Rob69. Julia Robinson. Unsolvable Diophantine problems. *Proc. Amer. Math. Soc.*, 22(2):534–538, 1969.

Rob96. Julia Robinson. *The collected works of Julia Robinson*. Number 6 in Collected Works. American Mathematical Society, Providence, RI, 1996. ISBN 0-8218-0575-4. With an introduction by Constance Reid. Edited and with a foreword by Solomon Feferman. xliv+338 pp.

Sie29. Karl Ludwig Siegel. *Über einige Anwendungen diophantischer Approximationen*. Abhandlungen der Preussischer Akademie der Wissenschaften, 1. 1929. An English translation by Clemens Fuchs is available in [Zan14].

Vse97. Maxim Aleksandrovich Vsemirnov. Infinite sets of primes, admitting Diophantine representations in eight variables. *Journal of Mathematical Sciences*, 87(1):3200–3208, 1997.

Wol62. Joseph Wolstenholme. On certain properties of prime numbers. *The Quarterly Journal of Pure and Applied Mathematics*, 5:35–39, 1862.

Zan14. Umberto Zannier, editor. *On some applications of Diophantine approximations*. Edizioni della Normale. Scuola Normale Superiore, 2014.