INTEGER DYNAMICS

DINO LORENZINI, MENTZELOS MELISTAS, ARVIND SURESH, MAKOTO SUWAMA, AND HAIYANG WANG

3 October 2020

ABSTRACT. Let $b \geq 2$ be an integer, and write the base b expansion of any non-negative integer n as $n = x_0 + x_1 b + \cdots + x_d b^d$ with $x_d > 0$ and $0 \leq x_i < b$ for $i = 0, \ldots, d$. Let $\phi(x)$ denote an integer polynomial such that $\phi(n) > 0$ for all $n > 0$. Consider the map $S_{\phi,b} : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$, with $S_{\phi,b}(n) := \phi(x_0) + \cdots + \phi(x_d)$. It is known that the orbit set $\{n, S_{\phi,b}(n), S_{\phi,b}(S_{\phi,b}(n)), \ldots\}$ is finite for all $n > 0$. Each orbit contains a finite cycle, and for a given b, the union of such cycles over all orbit sets is finite.

Fix now an integer $\ell \geq 1$ and let $\phi(x) = x^2$. We show that the set of bases $b \geq 2$ which have at least one cycle of length ℓ always contains an arithmetic progression and thus has positive lower density. We also show that a 1978 conjecture of Hasse and Prichett on the set of bases with exactly two cycles needs to be modified, raising the possibility that this set might not be finite.

1. INTRODUCTION

Fix an integer $b \geq 2$. Any non-negative integer n can be written uniquely in base b as $n = x_0 + x_1 b + \cdots + x_d b^d$ with $x_d > 0$ and $0 \leq x_i < b$ for $i = 0, \ldots, d$. We let $n = [x_0, \ldots, x_d]_b$ denote the base b expansion of n. Fix now a function $\phi : \{0, 1, \ldots, b-1\} \to \mathbb{Z}_{\geq 0}$, and consider the map $S_{\phi,b} : \mathbb{Z}_{\geq 0} \to \mathbb{Z}_{\geq 0}$, with

$$S_{\phi,b}(n) := \phi(x_0) + \cdots + \phi(x_d).$$

For instance, when $b = 10$ and $\phi(x) = x^2$, then $S_{\phi,b}(345) = 3^2 + 4^2 + 5^2$.

The ordered sequence $[n, S_{\phi,b}(n), S_{\phi,b}(S_{\phi,b}(n)), \ldots]$ is called the *orbit* of n under $S_{\phi,b}$. We say that n has *finite orbit* under $S_{\phi,b}$ if the set $\{n, S_{\phi,b}(n), S_{\phi,b}(S_{\phi,b}(n)), \ldots\}$ is finite. Any finite orbit contains a finite *cycle*, a non-empty sequence of integers $\mathrm{cyc}(n_1, \ldots, n_\ell)$ such that $S_{\phi,b}(n_\ell) = n_1$ and when $\ell > 1$, $S_{\phi,b}(n_i) = n_{i+1}$ for $i = 1, \ldots, \ell - 1$. This cycle of *length* ℓ is unique up to cyclic permutation of its terms.

Example 1.1. Let $\phi(x) = x^2$ and $b = 30$. Take the integer $c := 315181514012209182119$, which is obtained from *coronavirus*[1] by substituting each letter with its position in the alphabet. The repeated use of the function $S_{\phi,b}$ quickly brings c to a standstill: the orbit of c is $[c, 4189, 738, 900, 1, 1, \ldots]$, and the cycle associated to this orbit is $[1]$. On the other hand, when $b = 2020$, the orbit of c has length 953 with a cycle of length 801. The outlook for 2021 is no better, as both the length of the orbit and the cycle length increase.

2000 *Mathematics Subject Classification.* 11A67, 11B25, 11D45.

Key words and phrases. Orbit; cycle; integral point; b-ary expansion; square digit sum; happy number.

[1]This project was completed in the early months of the Covid-19 pandemic. Meetings were held using Zoom.

B. M. Stewart [Ste, Theorem 1] proved that *there exists a constant γ, depending on ϕ, such that if $n > \gamma$, then $n > S_{\phi,b}(n)$. In particular every positive integer n has finite orbit under $S_{\phi,b}$.* It follows from Stewart's Theorem that the orbits of $S_{\phi,b}$ produce only finitely many distinct cycles. We will call the set of distinct cycles associated with the orbits of $S_{\phi,b}$ *the cycles associated with $S_{\phi,b}$*. The complete determination of the cycles of a given $S_{\phi,b}$ is computationally quite expensive for large b's. When $\phi(x) = x^m$, one can proceed as follows.

1.2. *Let $\phi(x) = x^m$. For each integer $n \leq (m-1)b^m - 1$, compute the cycle of $S_{\phi,b}(n)$. Then the union of all these cycles is the complete set of cycles associated with $S_{\phi,b}$.*

This statement follows from [Ste, Theorem 7], and the key to Stewart's proof is that if $n > (m-1)b^m - 1$, then $n > S_{\phi,b}(n)$. Hence, *every cycle for $S_{\phi,b}$ contains a positive integer at most equal to $(m-1)b^m - 1$. In particular, every cycle for $S_{\phi,b}$ contains a positive integer whose base b expansion has at most $m+1$ digits.*

When $\phi(x) = x^2$, the number of 1-cycles of a given $S_{x^2,b}$ is explicitly determined by the following theorem of P. Subramanian [Sub, Theorem 1.2] (see also [H-P, Section 3], and Proposition 4.1). Recall that a divisor d of a positive integer n is called *proper* if $1 \leq d < n$.

1.3. *The number of 1-cycles of $S_{x^2,b}$ is equal to the number of proper divisors of $b^2 + 1$.*

For convenience, let us call [1] the *trivial* cycle of $S_{x^m,b}$. Let $\ell \geq 1$ be any integer. Let $B(\ell)$ denote the set of bases b such that $S_{x^2,b}$ has at least one non-trivial cycle of length ℓ. Theorem 1.3 implies that the natural density of $B(1)$ is 1 (see 2.12). It is not hard to show that $B(\ell)$ is infinite for all ℓ (see 2.1), and it is natural to wonder whether $B(\ell)$ has a positive natural density. Let $S \subset \mathbb{N}$ is any subset. Let $S(n) := \{1, 2, \ldots, n\} \cap S$ and $s(n) := |S(n)|$. Recall that the *lower density* $\underline{d}(S)$ of S is defined as $\underline{d}(S) := \liminf_{n\to\infty} \frac{s(n)}{n}$. In this article, we show:

Theorem (see 2.9 and 2.13). *Let $\ell \geq 1$. Then $B(\ell)$ has a positive lower density. More precisely, $B(\ell)$ always contains an explicit arithmetic progression. The sets $B(2)$, $B(3)$, and $B(4)$ have lower density bounded below by 0.57, 0.21, and 0.11, respectively.*

1.4. The key ingredient in the proof of Theorem 2.9 is the existence of special ℓ-cycles that we now define. Let $c = \mathrm{cyc}(n_1, \ldots, n_\ell)$ denote a cycle of length ℓ for $S_{x^2,b}$. We say that c is a *propagating cycle* if (i) every integer n_i, $i = 1, \ldots, \ell$, has at most two digits when written in base b and (ii) b does not divide n_i, for all $i = 1 \ldots, \ell$. The name 'propagating' is justified by our next theorem. It is easy to check with 1.2 that all 1-cycles of $S_{x^2,b}$ are propagating.

Theorem (see 2.8). *Let $b_0 \geq 2$. Assume that S_{x^2,b_0} has s distinct propagating cycles, of lengths $\ell_1, \ldots, \ell_s$, respectively (repetitions are allowed). Let $t \geq 0$ be any integer, and let $b := b_0 + t(b_0^2 + 1)$. Then $S_{x^2,b}$ has (at least) s distinct propagating cycles, of lengths $\ell_1, \ldots, \ell_s$ respectively.*

Propagating ℓ-cycles can naturally be seen as corresponding to integer points on an algebraic variety $V_\ell/\mathbb{Q}$. It turns out that V_ℓ has the property that, through every integer point on it corresponding to an ℓ-cycle, there passes at least one integer line given by explicit equations. This arithmetico-geometrical fact underlies the proof of Theorem 2.8. When $\ell = 1$ or 2, there is in addition a second integer line passing through each point which also propagates cycles. We exploit the existence of this second line when $\ell = 2$ in 3.4 and 3.6.

Some of Stewart's 1960 results [Ste] have been independently rediscovered by H. Hasse and G. Prichett in 1978 ([H-P, Theorem 4.1]). At the end of [H-P], Hasse and Prichett propose the following conjecture:

> *Let $\phi(x) = x^2$ and consider the set $L(x^2, 2)$ of all integers $b \geq 2$ such that the list of cycles associated with $S_{\phi,b}$ consists of the trivial cycle* [1] *and exactly one additional cycle. Then $L(x^2, 2) = \{6, 10, 16, 20, 26, 40\}$.*

Hasse and Prichett made this conjecture after having numerically verified it for $b \leq 500$.

Let $\phi(x)$ be any polynomial taking positive values on $\mathbb{Z}_{>0}$. Let $L(\phi, i)$ denote the set of integers $b \geq 2$ such that the list of cycles associated with $S_{\phi,b}$ consists of exactly i distinct cycles. It is natural to wonder whether the Hasse–Prichett conjecture for $\phi(x) = x^2$ and $i = 2$ is in fact only a specific instance of a much more general phenomenon, namely that all the sets $L(\phi, i)$ are finite, for all $i \geq 1$.

Curiously, Hasse and Prichett do not mention in [H-P] a similar conjecture for the set $L(x^2, 1)$. In this case, that $L(x^2, 1) = \{2, 4\}$ seems to be by now a folklore conjecture. It is stated in [OEI], A161872, that the conjecture has been verified for all $b < 500,000,000$.

Subramanian's Theorem 1.3 shows that if the set $L(x^2, 2)$ is infinite, it will indeed be very sparse, since if $b \in L(x^2, 1)$ or $L(x^2, 2)$, then $b^2 + 1$ is prime. To justify this claim, note that 1.3 implies that $b^2 + 1$ can only have at most one proper divisor bigger than 1. This can happen only when $b^2 + 1 = p^2$ for some prime p. But the factorization $1 = (p - b)(p + b)$ has no integer solutions when $b \geq 2$.

The unboundedness of the set of integers b such that $b^2 + 1$ is prime is implied by a general 1857 conjecture of Victor Bouniakowsky [Bou, page 328], that any irreducible polynomial $f(x) \in \mathbb{Z}[x]$ with positive leading coefficient takes infinitely many prime values if the values $f(1)$, $f(2)$, $f(3)$, ... have no common factor. This conjecture in the case of $f(x) = x^2 + 1$ was one of E. Landau's four problems presented at the 1912 International Congress of Mathematicians (see [H-L], pp 46-48). Note that a *negative* answer to the Hasse–Prichett Conjecture (in the strong sense where $L(x^2, 2)$ would be proved to be infinite) would provide a positive answer to Landau's problem.

The computations below were done using the cluster Sapelo2 at the Georgia Advanced Computing Resource Center.

Theorem 1.5. *Let $b \leq 1000000$. If $b \in L(x^2, 2)$, then*

$$b \in \{6, 10, 16, 20, 26, 40, 8626, 481360\}.$$

Thus the Hasse–Prichett Conjecture at the very least needs to be modified to include the bases $b = 8626$ and $b = 481360$. The existence of the large gap between these two bases might be seen as evidence against the validity of the modified conjecture.

Searching for other types of finiteness, one may wonder for instance whether, for a given integer d, the set $M(x^2, d)$ of all bases b such that $S_{x^2,b}$ only has cycles of length at most d is finite. We have $\{2, 4\} \subseteq M(x^2, 1)$ and $\{2, 3, 4, 13, 18, 92\} \subseteq M(x^2, 2)$.

Theorem 1.6. *Let $400 \leq b \leq 1100000$. If $b \in M(x^2, 10)$, then*

$$b \in \{432, 596, 687, 1068, 1932, 3918, 288504\}.$$

Moreover, $452808 \in M(x^2, 12)$. The presence of the large gap in $M(x^2, 10)$ might be seen as evidence that $M(x^2, 10)$ might be infinite.

For comparison with the case $\phi(x) = x^2$, let us note the following results for $\phi(x) = x^3$.

Theorem (see 5.1 and 5.4). *Let $b > 2$ be a square, or an integer that is not divisible by 9. Then $S_{x^3,b}$ has at least one non-trivial 1-cycle. In particular, the set of bases $b \geq 2$ such that $S_{x^3,b}$ has a non-trivial 1-cycle has lower density bounded below by $8/9$.*

Theorem (see 5.7 and 6.3). Let $k \geq 1$ be any integer.

(a) *Let $b = 3k + 1$. Then $S_{x^3,b}$ has at least five distinct cycles.*
(b) *Let $b = 9k^2 + 15k + 7$ or $9k^2 + 21k + 13$. Then $S_{x^3,b}$ admits at least one 2-cycle.*

Part (a) of the above theorem can be interpreted as saying that at least $1/3$ of the integers do not belong to $L(x^3, i)$ with $i \leq 4$. In the spirit of the Hasse–Prichett conjecture, we offer the following questions in the case where $\phi(x) = x^3$: do the equalities $L(x^3, 1) = \{2\}$, $L(x^3, 2) = \emptyset$, $L(x^3, 3) = \{3, 26\}$, and $L(x^3, 4) = \{5, 90, 188\}$ hold?

2. Propagating ℓ-cycles

Let $\phi(x)$ be any polynomial taking positive values on $\mathbb{Z}_{>0}$, and let $\ell \geq 1$ be any integer. It is natural to wonder whether there exist bases $b \geq 2$ such that $S_{\phi,b}$ admits cycles of length ℓ. We consider this question in this section mainly when $\phi(x) = x^2$. We start with some general observations for $\phi(x) = x^m$, $m \geq 2$.

Example 2.1. Let $\ell > 1$ and $m \geq 2$ be any integers. Let $b := c^{m^\ell - 1}$. Then the orbit of c^m under $S_{x^m,b}$ is a cycle of length ℓ, namely the cycle $\mathrm{cyc}(c^m, c^{m^2}, \ldots, c^{m^\ell} = bc)$. Thus we can quantify the infinitude of the set of bases b such that $S_{x^m,b}$ contains a cycle of length ℓ by noting that this set contains all integer values of the polynomial $f(t) = t^{m^\ell - 1}$ when $t > 1$. We show in Corollary 2.9 that when $m = 2$, the same statement holds with a polynomial $f(t)$ of degree 1.

Example 2.2. Take a prime $p > m$, and let ℓ denote the order of m in $(\mathbb{Z}/p\mathbb{Z})^*$. We find that the base $b = c^p$ admits the orbit of c^m under $S_{x^m,b}$ as a cycle of length ℓ. Thus, when the order of m equals $\ell := p - 1$, we find that all integer values of the polynomial $f(t) = t^{\ell+1}$ when $t > 1$ are contained in the set of bases b such that $S_{x^m,b}$ contains a cycle of length ℓ.

For instance when $\ell = 4$ and $m = 2$, we have for all bases $b = c^5$ the 4-cycle $\mathrm{cyc}(c^2, c^4, c^8, c^6)$. When $\ell = 2$, we can consider the 2-cycle $\mathrm{cyc}(c^m, c^{m^2})$ in base $b = c^{m+1}$, since in base b, $c^{m^2} = c(c^{m+1})^{m-1}$.

Example 2.3. *Existence of propagating ℓ-cycles.* None of the examples of ℓ-cycles exhibited above when $\phi(x) = x^2$ are examples of propagating cycles in the sense of 1.4, since although each integer in the cycle has at most two digits when written in base b, at least one integer in the cycle is divisible by b. Our next example is an example of a propagating cycle.

Fix $\ell \geq 2$. Choose coprime positive integers α and β such that β divides $\alpha^{2^\ell} - \alpha$, and set $\gamma := \alpha^2 + \beta^2$. For instance, one can choose $\alpha = \beta = 1$ and $\gamma = 2$. Set $b := (\gamma^{2^{\ell-1}} - \alpha)/\beta$. Since β divides $\alpha^{2^\ell} - \alpha$ and $\alpha^2 = \gamma - \beta^2$, we find that b is an integer. Then

$$\mathrm{cyc}(\gamma, \gamma^2, \gamma^4, ..., \gamma^{2^{\ell-1}} = \alpha + \beta b)$$

is an ℓ-cycle for $S_{x^2,b}$. Indeed, it is easy to verify that $\alpha, \beta < b$, so that $\gamma^{2^{\ell-1}} = [\alpha, \beta]_b$.

It is easy to check that b does not divide any of the integers γ^{2^k} for $k = 0, \ldots, \ell - 1$. Since $\gamma^{2^{\ell-1}}$ has only two digits in base b, the smaller integers γ^{2^k} have at most two digits. Hence, this cycle is a propagating cycle.

2.4. Let $\ell \geq 1$. Consider the affine space $\mathbb{A}^{2\ell+1}$ and label its coordinates by

$$b, \text{ and } x_i, y_i, \text{ for } i = 1, \ldots, \ell.$$

Let V_ℓ denote the algebraic subvariety of $\mathbb{A}^{2\ell+1}$ defined by $x_1^2 + y_1^2 = x_1 + by_1$ when $\ell = 1$, and in general by the ℓ equations

$$\begin{array}{l} x_i^2 + y_i^2 = x_{i+1} + by_{i+1} \quad \text{for } i = 1, \ldots, \ell - 1, \text{ and} \\ x_\ell^2 + y_\ell^2 = x_1 + by_1. \end{array}$$

Let now $\mathrm{cyc}(n_1, \ldots, n_\ell)$ be a propagating ℓ-cycle for S_{x^2,b_0}. Write $n_i = [\mathbf{x}_i, \mathbf{y}_i]_{b_0}$ with $0 \leq \mathbf{x}_i, \mathbf{y}_i \leq b_0 - 1$ and $\mathbf{x}_i \neq 0$. Then the integer point $(b_0, \mathbf{x}_i, \mathbf{y}_i, i = 1, \ldots, \ell)$ satisfies the equations of the variety V_ℓ, and thus a propagating ℓ-cycle for S_{x^2,b_0} corresponds to an integer point on the variety V_ℓ which satisfies the added requirement that $0 \leq \mathbf{x}_i, \mathbf{y}_i \leq b_0 - 1$ and $\mathbf{x}_i \neq 0$.

Theorem 2.5. *Let $P := (b_0, \mathbf{x}_1, \mathbf{y}_1, \ldots, \mathbf{x}_\ell, \mathbf{y}_\ell)$ be any point on the variety V_ℓ. Then there exists a line in $\mathbb{A}^{2\ell+1}$ that passes through P and is fully contained in V_ℓ. This line can be given by the parametric equations*

$$\begin{array}{lcl} b(t) & := & b_0 + (b_0^2 + 1)t, \\ x_i(t) & := & \mathbf{x}_i + (\mathbf{x}_i b_0 - \mathbf{y}_i)t, \\ y_i(t) & := & \mathbf{y}_i + (\mathbf{y}_i b_0 + \mathbf{x}_i)t, \text{ for } i = 1, \ldots, \ell. \end{array} \tag{2.1}$$

If P is a point which corresponds to a propagating cycle of S_{x^2,b_0}, then for every non-negative integer $\mathbf{t}$, the point

$$P(\mathbf{t}) := (b(\mathbf{t}), x_1(\mathbf{t}), y_1(\mathbf{t}), \ldots, x_\ell(\mathbf{t}), y_\ell(\mathbf{t}))$$

corresponds to a propagating cycle for $S_{x^2,b(\mathbf{t})}$.

Proof. The proof of the theorem is not difficult, once the parametric equations (2.1) are available. Indeed, one finds that in $\mathbb{Q}[t]$,

$$\begin{array}{lcl} x_i(t)^2 + y_i(t)^2 - x_{i+1}(t) - b(t)x_{i+1}(t) & = & (\mathbf{x}_i^2 + \mathbf{y}_i^2 - \mathbf{x}_{i+1} - b_0\mathbf{y}_{i+1})((b_0 t + 1)^2 + t^2) \\ & = & 0, \end{array}$$

for all $i = 1, \ldots, \ell$ (where the index $\ell + 1$ is set to mean index 1). We further need to show that for all $\mathbf{t} \geq 0$, $0 \leq x_i(\mathbf{t}), y_i(\mathbf{t}) < b(\mathbf{t})$, and that $x_i(\mathbf{t}) \neq 0$. These inequalities follow from the the fact that since $\mathbf{x}_i \neq 0$ and $\mathbf{y}_i < b_0$, we have $\mathbf{x}_i b_0 - \mathbf{y}_i > 0$, $\mathbf{x}_i + b_0\mathbf{y}_i > 0$, and $(\mathbf{x}_i b_0 - \mathbf{y}_i), (\mathbf{x}_i + b_0\mathbf{y}_i) \leq b_0^2 - 1$. □

Remark 2.6. Consider an ℓ-cycle $\mathrm{cyc}(n_1, \ldots, n_\ell)$ as in Theorem 2.5, with $n_i := \mathbf{x}_i + \mathbf{y}_i b_0$. Let $g := \gcd(b_0^2 + 1, n_1, \ldots, n_\ell)$. Given the parametrization (2.1), we will call the following parametrization of the same line the *reduced integer parametrization* of the line:

$$\begin{array}{lcl} b(t) & := & b_0 + t(b_0^2 + 1)/g, \\ x_i(t) & := & \mathbf{x}_i + t(\mathbf{x}_i b_0 - \mathbf{y}_i)/g, \\ y_i(t) & := & \mathbf{y}_i + t(\mathbf{y}_i b_0 + \mathbf{x}_i)/g, \text{ for } i = 1, \ldots, \ell. \end{array}$$

Note that g divides $\mathbf{x}_i b_0 - \mathbf{y}_i$ since $\mathbf{x}_i b_0 - \mathbf{y}_i = b_0(\mathbf{y}_i b_0 + \mathbf{x}_i) - \mathbf{y}_i(b_0^2 + 1)$. Note also that $g < b_0^2 + 1$ since if $g = b_0^2 + 1$, then $b_0^2 + 1$ would divide $n_i := \mathbf{x}_i + \mathbf{y}_i b_0$, but this is not possible since $n_i \leq b_0^2 - 1$.

Remark 2.7. For each $\mathbf{t} \in \mathbb{C}$, we can define an endomorphism $\varphi_{\mathbf{t}} : V_\ell \to V_\ell$ on the affine variety V_ℓ using the ring homomorphism $\varphi_{\mathbf{t}}^*$ on functions on V_ℓ defined as:

$$\begin{aligned} \varphi_{\mathbf{t}}^*(b) &:= b + \mathbf{t}(b^2+1), \\ \varphi_{\mathbf{t}}^*(x_i) &:= x_i + \mathbf{t}(x_i b - y_i), \\ \varphi_{\mathbf{t}}^*(y_i) &:= y_i + \mathbf{t}(y_i b + x_i), \text{ for } i = 1, \ldots, \ell. \end{aligned}$$

Theorem 2.8. *Let $b_0 \geq 2$. Assume that S_{x^2,b_0} has s distinct propagating cycles, of lengths $\ell_1, \ldots, \ell_s$, respectively (repetitions are allowed). Let $\mathbf{t} \geq 0$ be any integer, and let $b := b_0 + \mathbf{t}(b_0^2+1)$. Then $S_{x^2,b}$ has (at least) s distinct propagating cycles, of lengths $\ell_1, \ldots, \ell_s$ respectively.*

Proof. When the integers $\ell_1, \ldots, \ell_s$ are distinct, the statement of the theorem follows immediately from the existence of the 'propagating' lines proved in Theorem 2.5. Suppose now that for some integer ℓ, there are exactly $j > 1$ indices i such that $\ell_i = \ell$. Since we start with j distinct propagating ℓ-cycles, we have j distinct points on V_ℓ, and Theorem 2.5 proved the existence of j distinct lines on V_ℓ. To conclude the proof of Theorem 2.8, it suffices to prove that these lines do not intersect in V_ℓ at a point where $\mathbf{t}$ is a positive integer. This can be checked directly. Assume that $\mathbf{t} \neq 0$, and that we have two ℓ-cycles $(b_0, \mathbf{x}_1, \mathbf{y}_1, \ldots)$ and $(b_0, \overline{\mathbf{x}_1}, \overline{\mathbf{y}_1}, \ldots)$ with

$$\begin{aligned} \mathbf{x}_i + (\mathbf{x}_i b_0 - \mathbf{y}_i)\mathbf{t} &= \overline{\mathbf{x}_i} + (\overline{\mathbf{x}_i} b_0 - \overline{\mathbf{y}_i})\mathbf{t}, \\ \mathbf{y}_i + (\mathbf{x}_i + \mathbf{y}_i b_0)\mathbf{t} &= \overline{\mathbf{y}_i} + (\overline{\mathbf{x}_i} + \overline{\mathbf{y}_i} b_0)\mathbf{t}, \end{aligned}$$

for $i = 1, \ldots, \ell$. Then

$$\begin{aligned} (\mathbf{x}_i - \overline{\mathbf{x}_i})(1 + b_0\mathbf{t}) &= (\mathbf{y}_i - \overline{\mathbf{y}_i})\mathbf{t}, \\ (\mathbf{y}_i - \overline{\mathbf{y}_i})(1 + b_0\mathbf{t}) &= -(\mathbf{x}_i - \overline{\mathbf{x}_i})\mathbf{t}. \end{aligned}$$

We must have $(\mathbf{x}_i - \overline{\mathbf{x}_i}) = 0$ and, hence, $(\mathbf{y}_i - \overline{\mathbf{y}_i}) = 0$, since otherwise, the above equations imply that $(b_0\mathbf{t}+1)^2 + \mathbf{t}^2 = 0$. This latter equation is not possible when both $\mathbf{t}$ and b_0 are real, which we assume. □

Denote by $\mathrm{PB}(\ell)$ the set of bases $b \geq 2$ such that $S_{x^2,b}$ has a propagating ℓ-cycle.

Corollary 2.9. *Let $\ell \geq 2$. Let $b_0 := 2^{2^{\ell-1}} - 1$. Then the set $\mathrm{PB}(\ell)$ contains an arithmetic progression, and has lower density bounded below by $2/(b_0^2+1)$.*

Proof. The existence of a propagating ℓ-cycle $\mathrm{cyc}(\gamma, \gamma^2, \ldots)$ for the base $b_0 := 2^{2^{\ell-1}} - 1$ is established in Example 2.3. Since $\gamma = 2$, we find that the greatest common divisor of the elements in the cycle is 2. Theorem 2.5 and Remark 2.6 show the existence of an arithmetic progression $b(\mathbf{t}) = b_0 + \mathbf{t}(b_0^2+1)/2$ such that for every integer $\mathbf{t} \geq 1$, $S_{x^2,b(\mathbf{t})}$ has a propagating ℓ-cycle. The natural density of the set of positive integers in an arithmetic progression $\{a\mathbf{t} + b \mid \mathbf{t} \geq 0\}$ is $1/a$. □

Example 2.10. Every base b whose last digit in base 10 is 3 or 8 has a propagating 2-cycle. This follows from the fact that in base $b_0 = 3$, the 2-cycle $\mathrm{cyc}(2, 4)$ is propagating.

2.11. For later use, we note here the following facts. Consider a set S of positive integers which contains a union $U := \bigcup_{i=1}^{n} \left(\bigcup_{j=1}^{r_i} \{a_i t + b_{ij} \mid t \geq 0\} \right)$ of arithmetic progressions. Then the lower density $\underline{d}(S)$ of S satisfies $\underline{d}(S) \geq d(U)$. When the a_i are pairwise coprime, we find that

$$d(U) = 1 - \prod_{i=1}^{n} \left(1 - \frac{r_i}{a_i}\right).$$

2.12. Let us show now that PB(1) has natural density 1. More generally, let $f(x) \in \mathbb{Z}[x]$ be such that $f(\mathbb{Z}_{\geq 0}) \subseteq \mathbb{Z}_{\geq 0}$, and consider $B := \{b \in \mathbb{N} \mid f(b) \text{ is not prime}\}$. Subramanian's Theorem 1.3 shows that PB(1) has the same natural density as the set B when $f(x) = x^2+1$.

Let S denote the set of primes p such that there exists $b_p \in \mathbb{N}$ with $f(b_p)$ divisible by p. The set B then contains the arithmetic progression $b_p + pt$ for each $p \in S$. Thus the lower density of B is bounded below by the product $1 - \prod_{p\in S}\left(1 - \frac{1}{p}\right)$. The product $\prod_{p\in S}\left(1 - \frac{1}{p}\right)$ converges to 0 if and only if the sum $\sum_{p\in S} \frac{1}{p}$ diverges. When $f(x) = x^2+1$, the set S consists of 2 and all primes p congruent to 1 mod 4. It follows from Dirichlet's theorem on primes in arithmetic progression that $\sum_{p\in S} \frac{1}{p}$ diverges, so that the density of B is 1 in this case.

Proposition 2.13. *For $\ell = 2, 3, 4, 5$, the lower density $\underline{d}(\mathrm{PB}(\ell))$ is bounded below as follows:*

ℓ	2	3	4	5
$\underline{d}(\mathrm{PB}(\ell)) \geq$	0.5763	0.2127	0.1144	0.0429

Proof. The steps in the computations of the four lower bounds given in the above table are the same for each ℓ. First compute a set P of propagating cycles. In our case, we used all propagating cycles with $2 \leq b \leq 500$. For each propagating cycle $\mathrm{cyc}(n_1, n_2, \dots)$ in base b, Theorem 2.5 produces integer lines, which we parametrize using their reduced integer parametrization (see 2.6). The slope of each line is $\lambda := (b^2+1)/\gcd(b^2+1, n_1, n_2, \dots)$.

Thus we now have a set of explicit arithmetic progressions of the form $b + \lambda t$ which are contained in $\mathrm{PB}(\ell)$. To compute a lower bound for the lower density, we further prune the set of lines from all the lines which do not have prime power slope. The lower density of the set of arithmetic progressions with λ a prime power can be easily bounded below since the slopes are all pairwise coprime and we can use the formula in 2.11 to obtain a lower bound for $\underline{d}(\mathrm{PB}(\ell))$. Our computations produce the following data:

ℓ	1	2	3	4	5
$\lvert P\rvert$	2444	1163	391	190	77
$\underline{d}(\mathrm{PB}(\ell)) \geq$	0.8917	0.3507	0.2127	0.1144	0.0429

The data when $\ell = 1$ is only included for information, since we know already that $\underline{d}(\mathrm{PB}(1)) = 1$. The lower bound for $\underline{d}(\mathrm{PB}(2))$ is improved to $\underline{d}(\mathrm{PB}(2)) \geq 0.5763$ in 3.3. □

Remark 2.14. Computations indicate that the first integer b such that $S_{x^2,b}$ has a propagating ℓ-cycle might be much smaller than $2^{2^{\ell-1}} - 1$ when $\ell > 2$. One may wonder whether the first such b might even be bounded by a polynomial function in ℓ. Computations show that for each $\ell \leq 20$, there exists a basis $b \leq 1230$ with an ℓ-cycle.

Remark 2.15. Among the first 40000 bases b, only 1353, or about 3.38%, do not have a 2-cycle. The lower density of the set of bases b with a 2-cycle might thus be quite larger than 0.5763, and it would be interesting to determine if it is actually equal to 1. Among the first 25000 bases b, 17155, or about 68.62%, have a 3-cycle.

3. 2-cycles for $S_{x^2,b}$

In addition to the line described in Theorem 2.5, both varieties V_1 and V_2 contain a second integer line through each integer point. We prove this fact in this section for the variety V_2 and exploit the existence of this second line to study the 2-cycles of $S_{x^2,b}$.

Proposition 3.1. *Let $P := (b_0, x_0, y_0, u_0, v_0)$ be any point with non-negative rational coefficients on the threefold V_2 (defined in 2.4) outside of the lines $(t, 0, 0, 0, 0)$ and $(t, 1, 0, 1, 0)$. Then there exist two lines in $\mathbb{A}^5$, defined by equations with coefficients in $\mathbb{Q}$, which are entirely contained in V_2 and pass through P.*

Proof. Start with ten variables $x_0, x_1, y_0, y_1, u_0, u_1, v_0, v_1$, and b_0, b_1. Evaluate the two equations for V_2 at the linear polynomials $b(t) := b_0 + tx_1$, $x(t) := x_0 + tx_1$, $y(t) := y_0 + ty_1$, $u(t) := u_0 + tu_1$, and $v(t) := v_0 + tv_1$ to obtain two quadratic polynomials in t, say $f := f_2t^2+f_1t+f_0$ and $g := g_2t^2+g_1t+g_0$. Forcing these two polynomials to vanish identically produces six equations in the ten variables. The constant terms $f_0 = x_0^2 + y_0^2 - (u_0 + b_0v_0)$ and $g_0 = u_0^2 + v_0^2 - (x_0 + b_0y_0)$ are just the two equations of V_2 evaluated at the 0-variables.

Magma [BCP] can verify that (f_0, g_0) is a prime ideal in $\mathbb{Q}[x_0, y_0, u_0, v_0, b_0]$. Let F denote the field of fractions of the ring $\mathbb{Q}[x_0, y_0, u_0, v_0, b_0]/(f_0, g_0)$. Working now in the polynomial ring $F[x_1, y_1, u_1, v_1, b_1]$, consider the ideal $I := (f_1, f_2, g_1, g_2)$. Use the Magma [BCP] function `PrimaryDecomposition`(I) to produce the primary decomposition of this ideal. After about 34 hours of computing time, Magma will produce a decomposition which consists of three distinct ideals. Two of these ideals have generators that can be used to produce the parametric formulas for two different lines defined over $\mathbb{Q}$. Both lines at this point have parametric equations which are too long and complicated to be printed in this article. We succeeded in simplifying the parametrization of one of the lines, and checked that this line is the same line as the line exhibited in Theorem 2.5.

For the remainder of this section, let us call *the second line* through a point P on V_2 the line whose existence is established in the proposition and which is not equal to the line exhibited in Theorem 2.5. The Magma computation allows us to give this line in parametric form

$$\begin{aligned} b(t) := b_0 + t \quad x(t) := x_0 + tX_1/D, \quad y(t) := y_0 + tY_1/D, \\ u(t) := u_0 + tU_1/D, \quad v(t) := v_0 + tV_1/D, \end{aligned}$$

where the coefficients D, X_1, Y_1, U_1, V_1 are long formulas in the variables b_0, x_0, y_0, u_0, v_0. For instance,

$$\begin{aligned} D := u_0^2 + v_0^2 + 4y_0^2u_0^4 + 8y_0^2u_0^3 + 8y_0^2u_0^2v_0^2 + 4y_0^2u_0^2 + 8y_0^2u_0v_0^2 + 4y_0^2v_0^4 + 4y_0^2v_0^2 + 4y_0u_0^4v_0 \\ +8y_0u_0^2v_0^3 + 4y_0v_0^5 + u_0^6 + 2u_0^5 + 3u_0^4v_0^2 + 3u_0^4 + 4u_0^3v_0^2 + 2u_0^3 + 3u_0^2v_0^4 + 6u_0^2v_0^2 \\ +2u_0v_0^4 + 2u_0v_0^2 + v_0^6 + 3v_0^4, \end{aligned}$$

and we see that since D is a sum of monomials which includes $u_0^2 + v_0^2$, we must have $D > 0$ at P since the point P has non-negative coefficients and since u_0 and v_0 are not both zero by hypothesis. □

Remark 3.2. The line exhibited in Theorem 2.5 is remarkable since it allows us to 'propagate' any given propagating cycle. We believe that the second line on V_2 has the same property. In particular, starting with a propagating 2-cycle, we expect that the expressions X_1, Y_1, U_1, V_1 are all non-negative. This is immediately true for Y_1 since Magma produces a formula which is a sum of monomials, but for the other expressions, the formula involves some negative signs. The propagating property on the other hand can always be checked directly given an explicit propagating cycle, and this is what we do in order to establish our next proposition.

Assume that the point $P := (b_0, x_0, y_0, u_0, v_0)$ on V_2 has integer coefficients. Then the second lines can be reparametrized using a change of variables of the form $t := \lambda s$ with

$\lambda \in \mathbb{N}$ so that the new equations for the lines have only integer coefficients. In general there are very large cancellations in the fractions $X_1/D, Y_1/D, U_1/D$ and V_1/D and we set in this case λ to be the least common multiple of the denominators of $X_1/D, Y_1/D, U_1/D$ and V_1/D. We can use the second line through propagating 2-cycles to improve the lower bound given Proposition 2.13.

Proposition 3.3. *The set* PB(2) *of bases $b \geq 2$ such that $S_{x^2,b}$ has a propagating 2-cycle has lower density bounded below by* 0.5763.

Proof. Consider the set P_N of all propagating 2-cycles with $2 \leq b \leq N$. As in Proposition 2.13, for each propagating cycle $\mathrm{cyc}(n_1, n_2)$ in base b in P_N, Theorem 2.5 produces an integer line, which we parametrize using its reduced integer parametrization (see 2.6). The slope of the line is $\lambda_1 := (b^2+1)/\gcd(b^2+1, n_1, n_2)$. Consider the set S_1 of all the lines found this way.

Now for each propagating 2-cycle in base b in P_N, say $\mathrm{cyc}(x_0 + by_0, u_0 + bv_0)$, compute the reduced integer parametrization of the second line, with $b(t) = b + \lambda_2 t$, $x(t) = x_0 + x_2 t$, $y(t) = x_0 + y_2 t$, $u(t) = u_0 + u_2 t$, and $v(t) = v_0 + v_2 t$, and check that this line allows us to propagate the 2-cycle. To check this, we verified that $0 \leq x_2, y_2, u_2, v_2 \leq \lambda_2$. Consider the set S_2 of all the second lines found this way. We now have a set of explicit arithmetic progressions of the form $b + \lambda_1 t$ or $b + \lambda_2 t$ which are contained in PB(2).

To compute a lower bound for the lower density, we further prune the set $S_1 \cup S_2$ from all the lines whose slope is not a power of a prime. The lower density of the set of arithmetic progressions associated with the remaining lines can be easily bounded below since the lines have slopes that are all pairwise coprime and we can use the formula in 2.11 to obtain a lower bound of 0.5457 when $N = 1000$ and $|P_N| = 2885$. We can do slightly better by also considering some lines whose slope is not a power of a prime. For instance, when also considering the lines with slopes dividing $2 \cdot 17^3$, we obtain a lower bound of 0.5763 when $N = 1000$. □

Let $P = (b_0, x_0, y_0, u_0, v_0)$ be a propagating cycle on V_2, and consider the two lines passing through it and their reduced integer parametrization with $b_1(t) = b_0 + \lambda_1 t$ and $b_2(t) = b_0 + \lambda_2 t$. We have noted already that $\lambda_1 = (b_0^2+1)/\gcd(b_0^2+1, n_1, n_2) > 1$, so that for any positive integer $\mathbf{t}$, $b_1^2(\mathbf{t}) + 1$ is never prime. Thus no propagated cycle on the first line can have a base b such that b^2+1 is prime. On the other hand, quite often, the second line can produce propagated cycles that have a base b such that b^2+1 is prime. Our next proposition exploits this property.

Proposition 3.4. (a) *There exist infinitely many integers $b \geq 2$ such that $S_{x^2,b}$ has exactly two non-trivial 1-cycles, but $b \notin L(x^2, 3)$ because $S_{x^2,b}$ also has a 2-cycle.*

(b) *The Bouniakovsky Conjecture implies that there exist infinitely many integers $b \geq 2$ such that $S_{x^2,b}$ has no non-trivial 1-cycles, but $b \notin L(x^2, 2)$ because $S_{x^2,b}$ also has two distinct 2-cycles.*

Proof. Part (a) follows from the existence of the second line on V_2 passing through the point $P := (8, 2, 3, 5, 1)$, and given by

$$b = 17t + 8, x = 3t + 2, y = 5t + 3, u = 9t + 5, v = 2t + 1.$$

Indeed, this line has $b(t) = 17t + 8$, with $\gcd(17, 65) = 1$. It follows from Lemma 3.5 that the integer values of $(17x + 8)^2 + 1$ are coprime and so we can use [Iwa, Theorem, p. 172], fully

proved in [Oli, Theorem 1], applied to the polynomial $(17x+8)^2+1$ to obtain that there are infinitely many values b in the arithmetic progression $b(t) = 17t + 8$ such that $b^2 + 1$ is the product of two primes. It is clear from the equation of the line that for every positive t, $S_{x^2,b(t)}$ has a propagating 2-cycle.

Part (b) follows from the existence of a second line with a similar property. Starting with the 2-cycle $(24, 16, 6, 4, 12)$, we find using the proof of Proposition 3.1 that the second line on V_2 through that point is given by

$$b = 53t + 24, x = 34t + 16, y = 13t + 6, u = 8t + 4, v = 25t + 12.$$

Again, all the coefficients of the line are positive, and it is easy to verify that for all $t > 0$, $x, y, u, v < b$. Thus for each t, the corresponding b is such that $S_{x^2,b}$ has a 2-cycle. It is easy to verify that $\gcd(53, 24^2 + 1) = 1$. Finally, we can find a point in the intersection of the arithmetic progressions $17t + 8$ and $53t + 24$; for instance when $x_0 = 400$ and $x_1 = 128$, we have $17x_0 + 8 = 53x_1 + 24 = 6808$. Thus we can consider the progression $17 \cdot 53t + 6808$, and Lemma 3.5 (b) shows that the integer values of $(17 \cdot 53x + 6808)^2 + 1$ are coprime. Hence, the Bouniakovsky Conjecture implies that there exist infinitely many integers t such that $(17 \cdot 53t + 6808)^2 + 1$ is prime and, therefore, there exist infinitely many integers b of the form $b = 17 \cdot 53t + 6808$ such that $b^2 + 1$ is prime. For each such integer, we find that $S_{x^2,b}$ has two 2-cycles by construction. □

Lemma 3.5. (a) *Let $c, d \in \mathbb{Z}$. The integer values of the polynomial $(cx+d)^2+1$ are coprime if and only if $\gcd(c, d^2 + 1) = 1$.*

(b) *Let $c_0, d_0, c_1, d_1 \in \mathbb{Z}$. Suppose that $\gcd(c_0, d_0^2 + 1) = 1$ and $\gcd(c_1, d_1^2 + 1) = 1$. Suppose that there exist integers x_0 and x_1 such that $c_0x_0 + d_0 = c_1x_1 + d_1$. Then the integer values of the polynomial $(c_0c_1x + c_0x_0 + d_0)^2 + 1$ are coprime.*

Proof. (a) If p is a prime which divides all the integer values of $(cx + d)^2 + 1$, then p divides $d^2 + 1$, $c(c + 2d)$ and $c(c - 2d)$. Hence, if p does not divide c, then p divides $c + 2d$ and $c - 2d$, and thus divides $4d$. It follows that $p = 2$. But this is a contradiction, since then c is odd, and then $c^2 + 2d$ is also odd. Thus p divides $\gcd(c, d^2 + 1)$. Reciprocally, if the values of $(cx + d)^2 + 1$ are coprime, then $d^2 + 1$ and $c(c + 2d)$ are coprime and, hence, $d^2 + 1$ and c are coprime, as desired.

(b) In view of part (a), it suffices to prove that $gcd((c_0x_0 + d_0)^2 + 1, c_0c_1) = 1$. □

Remark 3.6. The two lines used in the proof of Proposition 3.4 are far from unique, and many other such lines could have been used. In fact, we believe that there are infinitely many propagating 2-cycles P on V_2 such that the second line passing through P in reduced integer parametrization with $b_2(t) = b_0 + \lambda_2 t$ is such that $\gcd(b_0^2 + 1, \lambda_2) = 1$.

To speed up the verification of 1.5, we created in advance a set of about 150 different second lines in reduced integer parametrization with $\gcd(b_0^2 + 1, \lambda_2) = 1$, and computed all bases $b \leq 10^6$ such that $b^2 + 1$ is prime and such that $S_{x^2,b}$ has a known 2-cycle on one of our 150 such lines. To eliminate such a b from $L(x^2, 2)$ required then to produce only one cycle of length greater than 2 for $S_{x^2,b}$, which is a very quick computation.

Example 3.7. Consider $b_0 := 288504$, an unusual base discovered when verifying Theorem 1.6. In this case, S_{x^2,b_0} has exactly 104 distinct cycles, all of them of length at most 7. All non-trivial cycles are propagating, with forty-seven 1-cycles, thirty-nine 2-cycles, ten 3-cycles, six 5-cycles, and one cycle of length $4, 6$, and 7, respectively.

To illustrate propagation, note that the very first base $b > 2$ to have a 7-cycle is $b = 15$, with $c := \text{cyc}(50, 34, 20, 26, 122, 68, 80)$ (see [H-P], page 10).[2] The reduced integer parametrization of the line passing through this 7-cycle starts with $b(t) = 113t + 15$, $x_1(t) = 36t + 5$, $y_1(t) = 25t + 3$, etc. We find that when $\mathbf{t} = 2553$, the corresponding integer point on the line produces the 7-cycle for $b(\mathbf{t}) = 288504$ found in our search.

All cycles of length at least 2 are found among the orbits of n with $1 \leq n \leq 1,964,329,269$. Among the thirty-nine propagating 2-cycles, one of them, $\text{cyc}(36253850477, 38091031810)$ is such that the second line associated to it has $\gcd(b_0^2 + 1, \lambda_2) = 1$ when written in reduced integer parametrization.

4. 1-cycles of $S_{x^2,b}$

In this section, we complement Subramanian's Theorem 1.3 with Proposition 4.1, and further study the surface V_1 associated in 2.4 with 1-cycles when $\phi(x) = x^2$.

Proposition 4.1. *Let $n := x+by$ be a non-trivial 1-cycle for $S_{x^2,b}$, and let $d := \gcd(b^2+1, n)$. Then $d > 1$, and there exists another 1-cycle $N := x + b(b - y)$ for $S_{x^2,b}$ such that, letting $D := \gcd(b^2 + 1, N)$, we have $D > 1$ and $b^2 + 1 = dD$.*

Let now $g := \gcd(x, y)$, $g' := \gcd(x, b - y)$, $h := \gcd(x - 1, y)$, and $h' := \gcd(x - 1, b - y)$. Then

(a) *$x = gg'$, $y = gh$, $b - y = g'h'$, and $x - 1 = hh'$.*
(b) *$d = n/g^2$, and $D = N/g'^2$.*

Proof. We leave it to the reader to check that N is a non-trivial 1-cycle. By hypothesis, $x(x - 1) = y(b - y)$. Since x and $x - 1$ are coprime, we find that $y = gh$ and $b - y = g'h'$. Then $x(x-1) = (gg')(hh')$. By definition, again since x and $x-1$ are coprime, we find that gg' is coprime to $x-1$, and it follows that gg' divides x. The same argument shows that hh' divides $x-1$. The equality $x(x-1) = (gg')(hh')$ implies then that $gg' = x$ and $hh' = x-1$, proving Part (a).

We claim that $d > 1$. Indeed, if $d = 1$, then $n = x + by$ divides $x + by - y^2$ and, hence, divides y^2. This is a contradiction since $0 < y^2 < x + by$. Similarly, $D > 1$ since otherwise N divides $x + by - y^2$, which is also a contradiction since $N > x + by - y^2$. The reader will check directly that

$$nN = (b^2 + 1)(x + by - y^2) = (b^2 + 1)x^2.$$

Note that it follows from this equality that $n \neq N$, since otherwise $b^2 + 1$ would be a square, which is not possible since $b > 0$. Using Part (a) and this equality, we find that

$$(n/g^2)(N/g'^2) = b^2 + 1.$$

It is then clear from this latter equality that (n/g^2) divides d, and that (N/g'^2) divides D. To finish the proof of Part (b), it suffices to prove that $dD = b^2 + 1$.

For this, it suffices to show that for every prime p, we have $\text{ord}_p(dD) = \text{ord}_p(b^2+1)$. First, note that dD and b^2+1 have the same prime divisors. Indeed, it is clear from the definitions that if p divides either d or D, then it divides $b^2 + 1$. On the other hand, if p divides $b^2 + 1$, then it divides nN and, hence, it divides d or D.

Let us show now that for every prime p, $\text{ord}_p(dD) \leq \text{ord}_p(b^2+1)$. The inequality is clear if either $\alpha := \text{ord}_p(d) = 0$ or $\beta := \text{ord}_p(D) = 0$, so we may assume that $\alpha, \beta > 0$. Then $p^\alpha \cdot p^\beta$

[2]Note a typo in [H-P] in the list of cycles for $b = 15$: the cycle just before c should be the non-propagating 5-cycle $\text{cyc}(41, 125, 89, 221, 317)$.

divides nN and since $nN = (b^2+1)x^2$, we obtained the desired inequality if we show that p does not divide x. Assume by contradiction that p divides x. Since by hypothesis, p also divides $n = x + by$, we find that p divides by and, hence, p divides y because p cannot divide b since it divides b^2+1. Again by hypothesis, p divides N, so p divides $N - x + by = b^2$, which is impossible. As a result, p does not divide x.

We now show that for every prime p, $\mathrm{ord}_p(dD) \geq \mathrm{ord}_p(b^2+1)$. Let $\gamma := \mathrm{ord}_p(b^2+1) > 0$, and assume by contradiction that $\gamma > \alpha + \beta$. Then $\gamma > \alpha$ and $\gamma > \beta$, which by definition implies that $\alpha = \mathrm{ord}_p(n)$ and $\beta = \mathrm{ord}_p(N)$. This is a contradiction, since we can conclude from $(b^2+1)x^2 = nN$ that $\gamma \leq \mathrm{ord}_p(nN) = \alpha + \beta$. □

Proposition 4.2. *Recall* (2.4) *the surface* V_1 *in* $\mathbb{A}^3$ *given by the equation* $x^2+y^2-(x+by) = 0$. *Given any point* $P := (b_0, x_0, y_0)$ *on* V_1 *outside of the lines* $(t,0,0)$ *and* $(t,1,0)$, *there exist exactly two lines in* $\mathbb{A}^3$ *that are entirely contained in* V_1 *and pass through* P, *namely the two lines given by the parametric equations*

$$\begin{array}{lll} b(t) := b_0 + t((x_0-1)^2 + y_0^2), & x(t) := x_0 + t(x_0-1)y_0, & y(t) := y_0 + ty_0^2, \\ b(t) := b_0 + t(x_0^2 + y_0^2), & x(t) := x_0 + tx_0y_0, & y(t) := y_0 + ty_0^2. \end{array}$$

The first parametric equation parametrizes the same line through P *as the line given in Theorem* 2.5. *Let* $d := \gcd(b_0^2+1, x_0 + b_0y_0)$, *and* $D := (b_0^2+1)/d$. *In reduced integer form, the first line has* $b(t) = b_0 + Dt$ *and the second line has* $b(t) = b_0 + dt$.

Proof. It is straightforward to check that the two lines described in the proposition lie on the surface V_1. The equations for these two lines were found using the same method as in the proof of Proposition 3.1, and this method shows that exactly two lines through P exist. Recall that the line in Theorem 2.5 is given by the parametric equations

$$B(t) = b_0 + (b_0^2+1)t, \quad X(t) = x_0 + (x_0b_0 - y_0)t, \quad Y(t) = y_0 + (x_0 + b_0y_0)t.$$

To see that it equals the first line of the proposition, we make the change of variable $t = (b_0^2+1)s$ in the first line, and $t = ((x_0-1)^2 + y_0^2)s$ in the other line, so that $b((b_0^2+1)s) = B(((x_0-1)^2+y_0^2)s)$. It remains to note that

$$\begin{array}{rcll} \frac{x((b_0^2+1)s)-x_0}{s} & = & (x_0-1)y_0(b_0^2+1) & \\ & = & (x_0b_0 - y_0)((x_0-1)^2+y_0^2) & = \frac{X(((x_0-1)^2+y_0^2)s)-x_0}{s}, \end{array}$$

and

$$\begin{array}{rcll} \frac{y((b_0^2+1)s)-y_0}{s} & = & y_0^2(b_0^2+1) & \\ & = & (x_0 + b_0y_0)((x_0-1)^2+y_0^2) & = \frac{Y(((x_0-1)^2+y_0^2)s)-y_0}{s}. \end{array}$$

□

5. Short cycles of $S_{x^3,b}$

In this section, $\phi(x) = x^3$. We exhibit below several parametric families of 1-cycles for $S_{x^3,b}$. After we became aware of [D-J], we noted that most of Proposition 5.1 already appears as Theorems 2-5 in [D-J]. Only parts (c) and (d) in Proposition 5.1 are in slightly stronger form than in [D-J].

Proposition 5.1. *Let* $k \geq 1$ *be a positive integer.*

(a) *Let* $b = 3k+1$. *Then* $n := [2k+1, 0, k+1]_b$, $n := [0, 2k+1, k]_b$, *and* $n := [1, 2k+1, k]_b$ *are* 1*-cycles for* $S_{x^3,b}$.

(b) *Let* $b = 3k+2$. *Then* $n := [2k+1, 0, k]_b$ *is a* 1*-cycle for* $S_{x^3,b}$.

(c) *Let* $b = 9k + 3$. *Then* $n := [6k + 2, 4k + 2, 5k + 1]_b$ *is a* 1*-cycle for* $S_{x^3,b}$.
(d) *Let* $b = 9k + 6$. *Then* $n := [6k + 4, 2k + 1, 7k + 5]_b$ *is a* 1*-cycle for* $S_{x^3,b}$.

Proof. An integer $n := [x, y, z]_b$ is a 1-cycle for $S_{x^3,b}$ if and only if the equation

$$x^3 + y^3 + z^3 = x + yb + zb^2$$

is satisfied. That this is the case can be checked directly. □

Remark 5.2. There are bases b of the form $b = 9k$, such as $b = 72$, 90, or 270, for which $S_{x^3,b}$ does not have any non-trivial 1-cycle. The bases $b = 18, 27$, and 54 have exactly one non-trivial 1-cycle, and $b = 108$ and 153 have exactly one non-trivial 1-cycle, which has 4 digits when written in base b (note it follows from 1.2 that a 1-cycle $[n]$ for $S_{x^3,b}$ is such that n has at most 4 digits in base b).

Thus Proposition 5.1 cannot immediately be generalized to include the case where $b = 9k$. But when $b = 9k^2$, Proposition 5.4 shows that $S_{x^3,b}$ has at least six non-trivial 1-cycles. When 9 divides b, we have only succeeded in producing parametric families of 1-cycles where b is a quadratic function of k, as in our next proposition.

Proposition 5.3. *Let* $b = 9(730k^2 - 1)$. *Then* $n := [27k, 3k]_b = 730(3k)^3$ *is a* 1*-cycle for* $S_{x^3,b}$.

Proof. An integer $n := [x, y]_b$ is a 1-cycle for $S_{x^3,b}$ if and only if the equation $x^3 + y^3 = x + yb$ is satisfied. Looking at this equation in the form $x(x-1)(x+1) = y(b-y^2)$, we can impose that y divide one of the factors x, $x + 1$, or $x - 1$, and solve for $b := y^2 + x(x-1)(x+1)/y$. If we want for 9 to divide b, we need to impose that $y = 3k$, and when we impose that y divide x, we can take for instance $x = 27k$, leading to the statement of the proposition. □

Proposition 5.4. *Let* $k \geq 2$ *be a positive integer.*
(a) *Suppose that* $b = k^2$. *Then* $[0, k]_b$ *and* $[1, k]_b$ *are* 1*-cycles for* $S_{x^3,b}$.
(b) *Suppose that* $b = (3k + 1)^2$. *Then* $[2k + 1, k + 1]_b$ *is a* 1*-cycle for* $S_{x^3,b}$.
(c) *Suppose that* $b = (3k + 2)^2$. *Then* $[2k + 1, k]_b$ *is a* 1*-cycle for* $S_{x^3,b}$.
(d) *Suppose that* $b = (3k)^2$. *Then* $[0, 6k^2 + k, 3k^2 + 2k]_b$, $[1, 6k^2 + k, 3k^2 + 2k]_b$, $[0, 6k^2 - k, 3k^2 - 2k]_b$ *and* $[1, 6k^2 - k, 3k^2 - 2k]_b$ *are* 1*-cycles for* $S_{x^3,b}$.

Proof. An integer $n := [x, y]_b$ is a 1-cycle for $S_{x^3,b}$ if and only if the equation $x^3 + y^3 = x + yb$ is satisfied. That this is the case can be checked directly. Similarly, for (d), an integer $n := [x, y, z]_b$ is a 1-cycle for $S_{x^3,b}$ if and only if the equation $x^3 + y^3 + z^3 = x + yb + zb^2$ is satisfied. □

Remark 5.5. When $\phi(x) = x^2$ and x^3, the set of bases b such that $S_{\phi,b}$ has a 1-cycle has positive lower density (see 1.3 and 5.1). We do not know if this remains the case when $\phi(x) = x^m$ and $m \geq 4$.

When $\phi(x) = x^m$ with $m \geq 3$, we only found the following parametric families, which show that the sets of integer values of certain polynomials $f(t)$ of degree $m-1$ are contained in the set of bases b where $S_{x^m,b}$ has a 1-cycle.

b	c^{m-1}	$2c^{m-1} - 1$	$c\frac{c^{m-1}-1}{c-1} + (c-1)^{m-1}$	m odd, $c\frac{c^{m-1}-1}{c+1} + (c+1)^{m-1}$
1-cycles	$[c^m]$, $[1 + c^m]$	$[c + bc]$	$[c + b(c-1)]$	$[c + b(c+1)]$

When $m = 3$, the parametrizations above produce 1-cycles when $b = 2c^2 - 1$, $2c^2 - c + 1$, and $2c^2 + c + 1$. Unfortunately, none of these values of b are divisible by 3.

Remark 5.6. Let $W_1/\mathbb{Q}$ denote the algebraic surface defined by the equation $x^3 + y^3 - (x + by) = 0$ in the affine space $\mathbb{A}^3$. General results on singular cubic surfaces in $\mathbb{A}^3$ predict that W_1 can contain at most 15 lines of $\mathbb{A}^3$ (use [B-W], Lemma 3 (c) and [B-W, page 255]). Unfortunately, none of these lines produces non-trivial 1-cycles for $S_{x^3,b}$.

Let $W_1'/\mathbb{Q}$ denote the algebraic surface defined by the equation $x^3 + y^3 - (x + b^2y) = 0$ in the affine space $\mathbb{A}^3$. The associated projective cubic surface in $\mathbb{P}^3$ is non-singular, and thus contains 27 lines of $\mathbb{P}^3$ (over $\mathbb{C}$). Some of these lines produce the parametrizations in Proposition 5.4 (b), (c), and (d).

Let us now consider 2-cycles of $S_{x^3,b}$. As noted already in 2.2, we have the following parametric family: when $b = c^4$, then $\mathrm{cyc}(c^3, c^9)$ is a 2-cycle for $S_{x^3,b}$, and in this example, one of the integer in the cycle is a 3-digit number in base b, since $c^9 = [0, 0, c]_b$. Using this example, we find that every value of the polynomial $f(t) = t^4$ is among the bases b such that $S_{x^3,b}$ has a 2-cycle. The following proposition allows us to prove the same statement with a quadratic polynomial $f(t)$.

Proposition 5.7. *Let W denote the algebraic variety in $\mathbb{A}^5$ defined by the equations $x^3+y^3 = u+bv$ and $u^3+v^3 = x+by$. The variety W contains the following two rational curves given by the parametrizations*

$$\begin{array}{llllll} b(t) := 9t^2 + 15t + 7, & x(t) := 2t + 2, & y(t) := t, & u(t) := t, & v(t) := t + 1, \\ b(t) := 9t^2 + 21t + 13, & x(t) := 2t + 3, & y(t) := t + 1, & u(t) := t + 1, & v(t) := t + 2. \end{array}$$

For every integer $\mathbf{t} \geq 0$, $\mathrm{cyc}(n(\mathbf{t}), m(\mathbf{t}))$ is a 2-cycle for $S_{x^3,b(\mathbf{t})}$, where $n(\mathbf{t}) := x(\mathbf{t})+y(\mathbf{t})b(\mathbf{t})$ and $m(\mathbf{t}) := u(\mathbf{t}) + v(\mathbf{t})b(\mathbf{t})$.

Proof. It is straightforward to verify that $(b(t), x(t), y(t), u(t), v(t))$ verifies the equations of W in both cases. It is also clear that $0 \leq x(\mathbf{t}), y(\mathbf{t}), u(\mathbf{t}), v(\mathbf{t}) < b(\mathbf{t})$. □

Note that $b(t)-1$ factorizes for both parametrizations, as $(3t+2)(3t+3)$ and $(3t+3)(3t+4)$, respectively. Thus for any integer $\mathbf{t}$, $b(\mathbf{t})$ is of the form $n(n + 1) + 1$ with either n or $n + 1$ divisible by 3.

6. A lower bound on the number of distinct cycles of $S_{x^m,b}$

In this section, we slightly generalize Theorem 12 of H. Grundman and E. Teeple in [G-T] from the case $\phi(x) = x^3$ to $\phi(x) = x^m$ for all $m \geq 3$. Given positive integers m and b, define

$$N = N(m, b) := \prod_{\substack{p \text{ prime} \\ p-1 \mid (m-1) \\ p \leqslant b-1}} p \quad \cdot \prod_{\substack{p \text{ prime} \\ p^{r-1}(p-1) \mid (m-1) \\ p > b-1 \\ \mathrm{ord}_p(m-1) = r-1}} p^r.$$

Proposition 6.1. *Let $\phi(x) = x^m$ with $m \geq 2$. Let $b \geq 2$. Then $S_{x^m,b}$ has at least $\gcd(b - 1, N)$ distinct cycles. In particular, when $m \geq 5$ is prime and $b = mk + 1$, then $S_{x^m,b}$ has at least m distinct cycles.*

To prove Proposition 6.1, we use the following slightly more general set-up.

Proposition 6.2. *Let $b \geq 2$ and set $B := \{0, 1, \ldots, b - 1\}$. Let $\phi : B \to \mathbb{Z}_{\geq 0}$. Suppose that there exists a positive integer ℓ such that $\ell \mid b - 1$ and such that $\phi(n) \equiv n \pmod{\ell}$ for all $n \in B$. Then*

$$S_{\phi,b}(n) \equiv n \pmod{\ell} \text{ for all } n \in \mathbb{Z}_{\geq 0}.$$

In particular, the cycles associated to the orbits of $n \in \{1, \dots, \ell\}$ under $S_{\phi,b}$ are all pairwise distinct, so that $S_{\phi,b}$ has at least ℓ distinct cycles.

Proof. Write $n = \sum_{i=0}^{d} n_i b^i$ in base b. Then

$$S_{\phi,b}(n) = \sum_{i=0}^{d} \phi(n_i) \equiv \sum_{i=0}^{d} n_i \equiv \sum_{i=0}^{d} n_i b^i = n \pmod{\ell}.$$

It follows that the cycles associated to the orbits of $n \in \{1, \dots, \ell\}$ under $S_{\phi,b}$ are all pairwise distinct. □

Proof of Proposition 6.1. Suppose that p is a prime such that $p-1$ divides $m-1$. Then for all $n \in \mathbb{Z}$, $n^m \equiv n \pmod{p}$. Suppose now that $p > b-1$ and that $\varphi(p^r) = p^{r-1}(p-1)$ divides $m-1$. Then the class of every integer $n \leq p-1$ is a unit in $\mathbb{Z}/p^r\mathbb{Z}$, and so by Euler's Theorem, $n^m \equiv n \pmod{p^r}$. It follows that $N(m,b)$ divides $n^m - n$ for all integers in B, and we can apply 6.2 with $\ell = N(m,b)$. □

Corollary 6.3. *Let $\phi(x) = x^3$. Let k be any positive integer and set $b = 3k+1$. Then $S_{x^3,b}$ has at least 5 distinct cycles.*

Proof. We know that [1] is a cycle. Using 6.2, we obtain that the orbit of $n = 3$ produces a cycle consisting entirely of integers congruent to 0 modulo 3. Proposition 5.1 (a) exhibits three non-trivial cycles consisting of integers congruent to 1 or 2 modulo 3. □

References

[BCP] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), 235-265. http://magma.maths.usyd.edu.au/magma/ 8

[Bou] V. Bouniakowsky, *Sur les diviseurs numériques invariables des fonctions rationnelles entières*, Mém. Acad. Sc. St. Pétersbourg **6** (1857), 305–329. 3

[B-W] J. W. Bruce and C. T. C. Wall, *On the classification of cubic surfaces*, J. London Math. Soc. (2) **19** (1979), no. 2, 245–256. 14

[D-J] L. E. Deimel, Jr. and M. T. Jones, *Finding pluperfect digital invariants: techniques, results and observations*, J. Recreational Math. **14** (1981/82), no. 2, 87–108. 12

[G-T] H. G. Grundman and E. A. Teeple, *Generalized happy numbers*, Fibonacci Quart. **39** (2001) 462–466. 14

[H-L] G. H. Hardy and J. E. Littlewood, *Some problems of partitio numerorum; III: On the expression of a number as a sum of primes*, Acta Math., **44** (1923), 1–70. 3

[H-P] H. Hasse and G. Prichett, *A conjecture on digital cycles*, J. reine angew. Math. **298** (1978), 8–15. 2, 3, 11

[Iwa] H. Iwaniec, *Almost-primes represented by quadratic polynomials*, Invent. math. **47** (1978), no. 2, 171–188. 9

[Oli] R. Lemke Oliver, *Almost-primes represented by quadratic polynomials*, Acta Arith. **151** (2012), no. 3, 241–261. 10

[OEI] OEIS Foundation Inc. (2020), The On-Line Encyclopedia of Integer Sequences, http://oeis.org 3

[Ste] B. M. Stewart, *Sums of functions of digits*, Canadian J. Math. **12** (1960), 374–389. 2, 3

[Sub] P. K. Subramanian, *On bases and cycles*, Math. Mag. **41** (1968), 117–123. 2

Department of Mathematics, University of Georgia, Athens, GA 30602, USA

Email address: lorenzin@uga.edu, mentzmel@uga.edu, arvind.suresh25@uga.edu, makoto.suwama@uga.edu, hywangn@uga.edu