Finding generalized near-repdigit squares

1 Introduction

A *repdigit* to base b is a natural number which consists of a single repeated digit when expressed in base b. A k-digit base b repdigit has the general form

$$\underbrace{aaa\dots aaa}_{k}{}_b = a\frac{b^k-1}{b-1}$$

with the repeated digit $0 \le a < b$.

Obláth's problem is to determine all decimal repdigits which are also perfect powers. Trivially, the single-digit perfect powers 4, 8 and 9 satisfy this criteria. Obláth [1] proved that any other such number must be a *repunit*, that is, having every digit equal to 1. Bugeaud and Mignotte [2] completed his problem by showing a repunit in base 10 cannot be a perfect power.

Similarly, a *near-repdigit* to base b is a number having all digits but one equal in base b. Generalizing this, we define an *n-near-repdigit* to be a number with all equal digits except for an n digit streak of a different digit. A k-digit n-near-repdigit to base b has the general form

$$\underbrace{aa\dots aa}_{k-n-m}\underbrace{cc\dots cc}_{n}\underbrace{aa\dots aa}_{m}{}_b = a\frac{b^k-1}{b-1} + (c-a)b^m\frac{b^n-1}{b-1} \qquad (1)$$

with $k \ge n+m$, $a \ne c$ and $0 \le a, c < b$.

Gica and Panaitopol [3] determined all square near-repdigits in base 10. Our objective was to automate the process of finding all n-near-repdigits in any base b, for given n and b. A Maple program was developed which enumerates the possible cases and then attempts to find an upper bound on the n-near-repdigit digit length. If successful, the output consists of every satisfying number written in base b and an abbreviated justification why each case admits only these solutions. Any problem cases are also noted.

2 Enumerating possible cases

First, we consider all possibilities for the last d digits (base b) of any square number. Every natural number can be written in the form $b^dk + r$, for some k and $0 \le r < b^d$. Since $(b^dk+r)^2 \equiv r^2 \pmod{b^d}$, the final d digits of $(b^dk+r)^2$

and r^2 are the same. Thus by checking if the last d digits of r^2 (for $0 \le r < b^d$) are also final digits of some n-near-repdigit, we can determine the possible forms for any square n-near-repdigit.

Rather than checking r^2 for all $0 \le r < b^d$, the search can be simplified if we know the possible r's which yield acceptable $d-1$ final digits. We "extend" such an r by one digit by forming $b^{d-1}k + r$ for each $0 \le k < b$. Since every natural number can be written as $b^d k' + b^{d-1}k + r \equiv b^{d-1}k + r \pmod{b^d}$, the possible last d digits for square n-near-repdigits of this form are the last d digits of $(b^{d-1}k + r)^2$. We simply compute this for $0 \le k < b$ and record for which k (if any) yield digits of some n-near-repdigit. By starting from a small number of digits and applying this repeatedly, every possible case for the last d digits may be found without searching over every $0 \le r < b^d$.

Once the possibilities for the last d digits have been found, it is often possible to immediately identify the repeating digit (denoted a), the differing digit (denoted c), and the position at which the streak of different digits begins (denoted m). Sometimes it is not possible to distinugish the repeating digit from the differing digit[1], but this problem can be solved by choosing a larger value for d. A more serious problem occurs when the digits in a possible case are all the same; if $d > n$ then this digit cannot be c (it must be a), but it is unknown what c and (more importantly) m are (though $m \ge d$).

3 Solving cases

3.1 Trivial cases

The case where the repeated digit is 0 is easily solved, since then an n-near-repdigit has the form $cb^m \frac{b^n-1}{b-1}$, which is a perfect square when m is even and $c\frac{b^n-1}{b-1}$ is a perfect square, or when m is odd and $cb\frac{b^n-1}{b-1}$ is a perfect square. Both cases are checked for each $0 < c < b$. When successful, they yield the family of solutions $cb^{2i}\frac{b^n-1}{b-1}$ or $cb^{2i-1}\frac{b^n-1}{b-1}$ for $i > 0$.

3.2 Remaining cases

We start knowing b and n, and from Section 2 we found all possible combinations of a, c and m. Thus, we now need to find values of k such that (1) is a perfect square.

3.2.1 Simple test

Let A_k represent the k-digit n-near-repdigit of the case we are testing. Writing A_k in terms of A_{k-1} we find the recurrence relation $A_k = bA_{k-1} + a - (c-a)b^m(b^n - 1)$. Thus for any h the sequence $\{A_k \bmod h\}_{k=n+m}^{\infty}$ must be periodic with period at most h. By computing this sequence for various primes

[1] For example, if the last 4 digits of a 3-near-repdigit are 1222, it could be of the form $11\cdots11222$ or $22\cdots22111222$.

(or prime powers) h, it is sometimes possible to find a sequence whose period is entirely composed of quadratic nonresidues (mod h).

3.2.2 Main test setup

Multiplying (1) by $b-1$, we find that we want to solve the Diophantine equation

$$(b-1)x^2 = ab^k - a + (c-a)b^m(b^n - 1)$$

for x and k. We split this into two cases, for even and odd k (say $k = 2M$ and $k = 2M+1$). To ensure that the coefficient on x^2 is a perfect square we multiply by $b-1$ again (this step can be skipped if $\sqrt{b-1} \in \mathbb{Z}$), and then make the following substitutions:

$$\begin{aligned} y &= (b-1)x \\ z &= b^M \\ N &= (b-1)\big((c-a)b^m(b^n-1) - a\big) \end{aligned}$$

Then we want to find the solutions (y, z) for the pair of equations

$$y^2 - a(b-1)z^2 = N \qquad \text{and} \qquad y^2 - a(b-1)bz^2 = N,$$

thus reducing the problem to solving equations of the form

$$x^2 - Dy^2 = N. \tag{2}$$

3.2.3 Solving $x^2 - Dy^2 = N$

When D is a perfect square then (2) is a difference of squares and factorizes $\big(x + \sqrt{D}y\big)\big(x - \sqrt{D}y\big) = N$, thus all solutions can be found by examining the divisors of N. If $N = d_1 d_2$ then we have $x = (d_1 + d_2)/2$ and $y = (d_1 - d_2)/2$, although this will only be an integer solution when $d_1 \equiv d_2 \pmod 2$.

When D is not a perfect square then (2) is a generalized form of Pell's equation

$$x^2 - Dy^2 = 1, \tag{3}$$

which plays an important role when solving (2). Note that if (x, y) is a solution of (2) and (u, v) is a solution of (3), then $(ux + vyD, uy + vx)$ is also a solution of (2). Using this fact, solutions of (2) can be partitioned into equivalence classes: we say that $(x_1, y_1) \equiv (x_2, y_2)$ if there exists some solution (u, v) of (3) such that $(ux_1 + vy_1D, uy_1 + vx_1) = (x_2, y_2)$.

We define the *primitive solution* to be the smallest positive solution of (3); it may be computed by examining convergents to $\sqrt{D}$ as described in [4] and [5]. Also, for each class of solutions of (2) we define the *fundamental solution* to be the solution (x, y) with the smallest $y \geq 0$ in the class. The following theorem demonstrates how all solutions to (2) may be determined once all fundamental solutions are known.

Theorem 1 *Let (p_x, p_y) be the primitive solution of $x^2 - Dy^2 = 1$, and (f_x, f_y) be a fundamental solution of $x^2 - Dy^2 = N$. Define the pair of linear recurrence relations:*

$$\begin{aligned} X_i &= 2p_x\, X_{i-1} - X_{i-2} \\ Y_i &= 2p_x\, Y_{i-1} - Y_{i-2} \end{aligned} \tag{4}$$

with initial conditions

$$\begin{aligned} X_0 &= f_x \quad & X_1 &= p_x f_x + p_y f_y D \\ Y_0 &= f_y \quad & Y_1 &= p_x f_y + p_y f_x. \end{aligned}$$

Then all solutions to $x^2 - Dy^2 = N$ in the class of (f_x, f_y) are given by $\pm(X_i, Y_i)$ for $i \in \mathbb{Z}$.

(X_i, Y_i) is well-defined for $i < 0$ since rearranging (1) yields

$$\begin{aligned} X_i &= 2p_x\, X_{i+1} - X_{i+2} \\ Y_i &= 2p_x\, Y_{i+1} - Y_{i+2}. \end{aligned}$$

The theorem follows from Remark 6.2.1 in Mollin [4] after demonstrating that

$$X_i + Y_i\sqrt{D} = \left(p_x + p_y\sqrt{D}\right)^i \left(f_x + f_y\sqrt{D}\right), \tag{5}$$

which can be established from the closed-form expressions for X_i and Y_i.

In the standard fashion, we find that for both recurrences the characteristic polynomial is $t^2 - 2p_x t + 1 = 0$ and it has roots $p_x \pm \sqrt{p_x^2 - 1} = p_x \pm p_y\sqrt{D}$. (The latter expression following from the definition $p_x^2 - Dp_y^2 = 1$.) After solving for the proper coefficients we find the closed-form expressions:

$$X_i = \frac{f_x + f_y\sqrt{D}}{2}\left(p_x + p_y\sqrt{D}\right)^i + \frac{f_x - f_y\sqrt{D}}{2}\left(p_x - p_y\sqrt{D}\right)^i,$$

$$Y_i = \frac{f_x + f_y\sqrt{D}}{2\sqrt{D}}\left(p_x + p_y\sqrt{D}\right)^i - \frac{f_x - f_y\sqrt{D}}{2\sqrt{D}}\left(p_x - p_y\sqrt{D}\right)^i.$$

And (5) immediately follows.

Then it is necessary to find all fundamental solutions to (2). The following bounds are given by Robertson [5], very similar to those in Mollin [4].

Theorem 2 *Let (p_x, p_y) be the primitive solution of $x^2 - Dy^2 = 1$, and (f_x, f_y) be a fundamental solution of $x^2 - Dy^2 = N$. Then*

$$\begin{aligned} 0 \le f_y &\le \sqrt{\tfrac{N(p_x-1)}{2D}} \qquad \text{if } N > 0, \\ \sqrt{\tfrac{|N|}{D}} \le f_y &\le \sqrt{\tfrac{|N|(p_x+1)}{2D}} \qquad \text{if } N < 0. \end{aligned}$$

Then for all integers y within these bounds we check if $\left(\sqrt{Dy^2+N}, y\right)$ is an integer solution, in which case $\left(\pm\sqrt{Dy^2+N}, y\right)$ are both fundamental solutions (unless they belong to the same class—in which case the class is called ambiguous and we choose the fundamental solution to have $x > 0$).

3.2.4 Comparing $y^2 - Dz^2 = N$ solutions with $z = b^M$

After solving (2) we want to be able to check which solutions actually correspond with n-near-repdigits. Because of our choosen substitution, z must be some power of b. We will attempt to show that z cannot be a power of b when z is larger than some explicitly computed bound.

Let $z_1(h)$ be the sequence $\{b^i \bmod h\}_{i=0}^{\infty}$, this sequence remains the same while solving each n-near-repdigit case. Also, let $z_2(h)$ be the sequence $\{Z_i \bmod h\}_{i=0}^{\infty}$, where Z_i represents the recurrence relation in Theorem 1 applied to $y^2 - Dz^2 = N$, the relevant equation for whichever case is currently trying to be solved. It is obvious that both $z_1(h)$ and $z_2(h)$ must be periodic for all h; the following propositions are more exact.

Proposition 1 *Let $\{p_i\}$ be the prime factors of b; write $h = h' \prod p_i^{e_i}$, where* $\gcd(h, h') = 1$. *Then the pre-period of $z_1(h)$ is given by* $\max\{e_i\}$ *and the period of $z_1(h)$ is given by* $\operatorname{ord}_{h'}(b)$.

Proposition 2 *There is no pre-period of $z_2(h)$ for any h, and the period of $z_2(h)$ is no larger than $2h$.*

Proof. Let s and t be the pre-period and period, respectively, of $z_1(h)$. Then $s \geq 0$ and $t > 0$ are the smallest integers such that $b^{s+t} \equiv b^s \pmod{h}$. This is equivalent to the system of simultaneous congruences[2]:

$$\begin{aligned} b^{s+t} &\equiv b^s \pmod{h'} \\ b^{s+t} &\equiv b^s \pmod{p_i^{e_i}} \end{aligned}$$

Dividing by coprime factors, this reduces to:

$$\begin{aligned} b^{t} &\equiv 1 \pmod{h'} \\ p_i^{s+t} &\equiv p_i^s \pmod{p_i^{e_i}} \end{aligned}$$

Since $t > 1$ the last congruence implies $p_i^s \equiv 0 \pmod{p_i^{e_i}}$. The smallest s such that this is true for all $\{p_i\}$ will be $\max\{e_i\}$. And the smallest t such that $b^t \equiv 1 \pmod{h'}$ is $\operatorname{ord}_{h'}(b)$ by definition.

For the pre-period of $z_2(h)$, as already pointed out, the recurrence Z_i is bi-directional so that you can determine a term in $z_2(h)$ based on terms that succeed it in the sequence. Thus, the periodicity that occurs must also extend backwards.

Let $k(h)$ be the period of $z_2(h)$. If we write $h = \prod q_i^{f_i}$ where $\{q_i\}$ are the prime factors of h, then $k(h) = \operatorname{lcm}\{k(q_i^{f_i})\}$. The proof is similar to some found in [6].

Once we have found the pre-period and period length for both $z_1(h)$ and $z_2(h)$, all that remains is to compare the approapriate terms in the two sequences and find some h such that the terms found in the repeating portion of $z_1(h)$

[2]For conciseness, only the congruence for the general case p_i is shown, but it should be understood that there will be a single congruence for each prime factor.

and those in the repeating portion of $z_2(h)$ have no elements in common. The entire pre-period of $z_1(h)$ also must be checked for solutions. Although finding a suitable h is not difficult for most equations, unfortunately no h was found in some cases, for example in $(3x)^2 - 3(4^M)^2 = 33$.

4 Example Results

All solutions to the near-repdigit square problem:

BASE 2:
only the trivial family 10^{2i} for $i > 0$
BASE 6:
13, 24, 41, 121, 144, 244, 441, 4424
and the trivial families 10^{2i}, $4 \cdot 10^{2i}$ for $i > 0$
BASE 10:
16, 25, 36, 49, 64, 81, 121, 144, 225, 441, 484, 676, 1444, 44944
and the trivial families 10^{2i}, $4 \cdot 10^{2i}$, $9 \cdot 10^{2i}$ for $i > 0$

All solutions to the 2-near-repdigit square problem:

BASE 2:
100, 1001, 11001, 1111001
BASE 6:
100, 144, 244, 400, 441, 3344, 11441
BASE 10:
100, 144, 225, 400, 441, 900, 7744, 11881, 55225

All solutions to the 3-near-repdigit square problem:

BASE 2:
110001
BASE 6:
none
BASE 10:
1444

References

[1] R. Obláth, Une proprieté des puissances parfaites, *Mathesis* **65** (1956), 356–364.

[2] Y. Bugeaud and M. Mignotte, On integers with identical digits, *Mathematika,* **46** (1999), 411–417.

[3] A. Gica and L. Panaitopol, On Obláth's problem, *J. Integer Seq.* **6** (2003), Paper 03.3.5.

[4] R. Mollin, *Fundamental Number Theory with Applications*. C.R.C. Press, Boca Raton, 1998, pp. 298–301.

[5] J. Robertson, Solving the generalized Pell equation $x^2 - Dy^2 = N$, `http://hometown.aol.com/jpr2718/pell.pdf` (2004).

[6] D. Vella and A. Vella, Cycles in the Generalized Fibonacci Sequence Mod p, *The Mathematical Gazette* (2006).